



Concept Manual

Self-Serve Data Preparation

(SSDP)

Version 5.1

Document Information	
Document ID	Smarten-SSDP-Concept-Manual
Document Version	2.0
Product Version	5.1
Date	5-December-2020
Recipient	NA
Author	EMTPL

© Copyright Elegant MicroWeb Technologies Pvt. Ltd. 2020. All Rights Reserved.

Statement of Confidentiality, Disclaimer and Copyright

This document contains information that is proprietary and confidential to EMTPL, which shall not be disclosed, transmitted, or duplicated, used in whole or in part for any purpose other than its intended purpose. Any use or disclosure in whole or in part of this information without the express written permission of EMTPL is prohibited.

Any other company and product names mentioned are used for identification purpose only, may be trademarks of their respective owners and are duly acknowledged.

Disclaimer

This document is intended to support administrators, technology managers or developers using and implementing Smarten. The business needs of each organization will vary and this document is expected to provide guidelines and not rules for making any decisions related to Smarten. The overall performance of Smarten depends on many factors, including but not limited to hardware configuration and network throughput.

Contents

1 About this document	7
1.1 Scope and Organization of Topic Areas.....	7
1.2 Conventions Used	7
2 Introducing ElegantJ BI - Smarten.....	8
3 Introducing Self-Serve Data Preparation (SSDP).....	8
4 Data Source.....	10
4.1 Creating Data Source.....	10
4.2 Example	10
4.2.1 Creating Data Source for File	10
4.3 List of Data Sources.....	11
5 Dataset	12
5.1 Types of Dataset.....	12
5.1.1 Cache Dataset	12
5.1.2 Real-Time Dataset	12
5.2 Creating a Dataset.....	12
5.2.1 Connect to Data Source	13
5.2.2 Fetch Data for Dataset from Data Source	13
5.2.2.1 Column Selection	13
5.2.3 Example	14
5.2.3.1 Creating a Dataset from a Database Data Source.....	14
6 Explore Data.....	15
6.1 Highlight	15
6.1.1 Highlight Missing Values	15
6.1.2 Highlight Spaces	16
6.1.3 Highlight Inconsistent Values.....	17
6.1.4 Highlight Duplicate columns with this column	18
6.1.5 Highlight Duplicate rows with this row	19
6.1.6 Highlight Duplicate column values.....	20
6.1.7 Highlight All duplicate rows	21
6.1.8 Highlight Rows with all null.....	21
6.1.9 Highlight Rows with all zeros	22
6.1.10 Highlight Columns with all null.....	23
6.1.11 Highlight Columns with all zeros	24
6.2 Unique Values	25
6.3 Filter	25
6.3.1 Filter Rows with this column value	25
6.3.2 Filter Duplicate rows with this row	26

6.3.3 Filter All duplicate rows.....	27
6.3.4 Filter Rows with all null	28
6.3.5 Filter Rows with all zeros	28
6.3.6 Custom	29
6.3.6.1 Filter	29
6.3.6.2 Advanced Filter.....	30
6.4 Statistics	31
7 Clean Data.....	32
7.1 Cluster & Edit.....	32
7.2 Find & replace	34
7.3 Remove.....	35
7.3.1 Remove This column	35
7.3.2 Remove This row	36
7.3.3 Remove Row with this column value	37
7.3.4 Remove Duplicate columns with this column.....	38
7.3.5 Remove Duplicate rows with this row	39
7.3.6 Remove All duplicate rows.....	40
7.3.7 Remove Rows with all null	41
7.3.8 Remove Rows with all zeros.....	42
7.3.9 Remove Columns with all null.....	42
7.3.10 Remove Columns with all zeros	43
7.3.11 Remove Rows in search result	44
7.4 Outliers	45
7.4.1 Remove Outliers.....	47
7.4.2 Replace Outliers	48
8 Transform Data	48
8.1 For text columns.....	48
8.1.1 Transform Upper case.....	48
8.1.2 Transform Lower case	49
8.1.3 Transform Capitalise	50
8.1.4 Transform Advanced functions	51
8.2 For numerical columns	54
8.2.1 Transform Advanced functions	54
8.3 For datetime columns	57
8.3.1 Transform—Make timepart zero	57
8.3.2 Transform First date of month.....	58
8.3.3 Transform Last date of month	59
8.3.4 Transform Advanced functions	60
8.4 Change data type	64
8.4.1 Timestamp.....	64

8.4.2 Date	64
8.4.3 Integer	65
8.4.4 Double	65
8.4.5 String	65
8.5 Aggregate	66
8.5.1 Sum	66
8.5.2 Min	67
8.5.3 Max	67
8.5.4 Count	68
8.5.5 Average	68
9 Shape Data	69
9.1 Split	69
9.1.1 Split to column	69
9.1.2 Split to row	73
9.2 Merge columns	75
9.3 Sort	76
9.3.1 Sort Ascending	76
9.3.2 Sort Descending	77
9.3.3 Sort None	79
9.4 Copy	80
9.4.1 Copy Column	80
9.4.2 Copy Row	81
9.5 Edit Row	82
9.6 Add column	83
9.6.1 Row number	83
9.6.2 For Datetime columns	84
9.6.3 Custom	85
9.7 Mark as	100
9.7.1 Time dimension	100
9.7.2 GeoMap dimension	101
9.7.3 Dimension	103
9.8 Fill data	104
10 Blend data	105
10.1 Join	105
10.1.1 Inner	105
10.1.2 Left	106
10.1.3 Right	107
10.1.4 Outer	108
10.2 Append	109
11 Sampling	110

11.1 Sampling—Simple random sampling	110
11.2 Sampling—Stratified sampling	111
12 Publish and Refresh Dataset.....	112
12.1 Publish Dataset.....	112
12.2 Refresh Dataset	113
12.2.1 Refresh methods	114
12.2.1.1 From scratch.....	114
12.2.1.2 Incremental	114
13 Dataset Management	115
13.1 Manage Access Rights for Dataset	115
13.2 Edit Dataset	115
13.3 Delete Dataset.....	115
13.4 Copy Dataset	115
13.5 IT Certification for Dataset.....	116
14 Action editor	116
15 Lineage diagram	117
16 Auto-suggestions and recommendations	117
17 Product and Support Information.....	119

1 About this document

This manual explains the concepts required to use Self-Serve Data Preparation (SSDP) features in Smarten Augmented Analytics Suite.

1.1 Scope and Organization of Topic Areas

Chapter 2	Introducing ElegantJ BI - Smarten
Chapter 3	Introducing Self-Serve Data Preparation (SSDP)
Chapter 4	Data Source
Chapter 5	Dataset
Chapter 6	Explore Data
Chapter 7	Clean Data
Chapter 8	Transform Data
Chapter 9	Shape Data
Chapter 10	Blend Data
Chapter 11	Sampling
Chapter 12	Publish & Refresh Dataset
Chapter 13	Dataset Management
Chapter 14	Action Editor
Chapter 15	Lineage Diagram
Chapter 16	Autosuggestions & Recommendations
Chapter 17	Product and Support Information

1.2 Conventions Used

This manual uses typographical conventions in the text to help you distinguish between the names of files, instructions, and other important notes that are relevant during installation. For example:

- Important notes are indicated in a different font colour as shown in the example below.

Note:

These are the front-end operations that highlight records in the front-end interface and do not remove any records from the Dataset.

- References to documents are highlighted as below:

Reference: Self-Serve Data Preparation (SSDP) - Concept Manual > Shape Data > Add Column > Custom

2 Introducing ElegantJ BI - Smarten

ElegantJ BI is a full-stack Business Intelligence tool that employs the “Smarten” approach to Augmented Analytics. The solution comprises a comprehensive set of tools, including Self-Serve Data Preparation, Smart Visualisation, and Plug n’ Play Predictive Analytics. These tools are designed to democratize advanced analytics and transform business users into citizen data scientists.

Self-Serve Data Preparation

Self-Serve Data Preparation allows business users to perform data preparation on their own without the assistance of IT staff or data analysts. Users are not restricted by complex tools or forced to wait for IT to deliver crucial data. Guided by smart suggestions and auto recommendations, business users can prepare, blend, and transform data and create analysis-ready data quickly and accurately without assistance rather than waiting for central metadata prepared by IT.

Smart Data Visualization

Intuitive Smart Data Visualization tools suggest the best options for visualizing and plotting a particular set or type of data based on the nature, dimensions, and trend of data so that business users can easily select the appropriate method to clearly and quickly visualize data in a way that is meaningful to the task.

Plug n’ Play Predictive Analysis

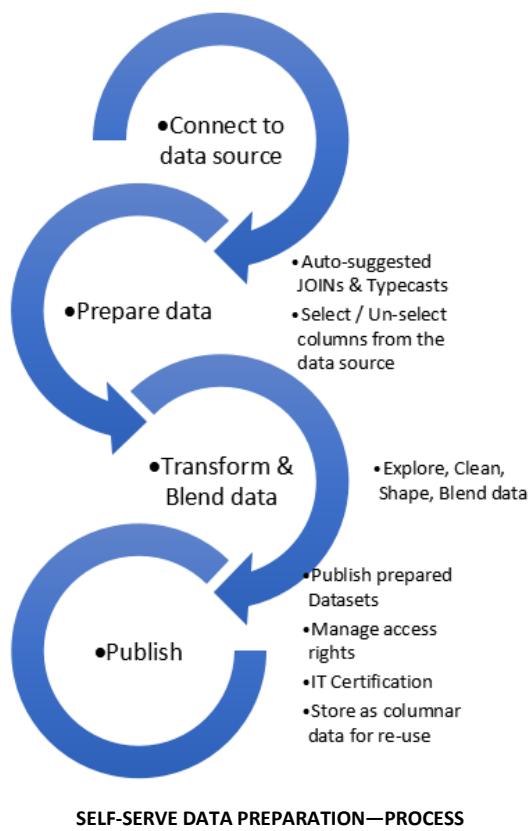
Assisted Predictive Modelling allows business users to explore predictive algorithms and models without the skill or knowledge of a data scientist or statistician. The solution considers the underlying data and use case, suggests the best-fit algorithm, and then displays output results and data visualization accompanied by an interpretation using simple human language.

3 Introducing Self-Serve Data Preparation (SSDP)

The Self-Serve Data Preparation component of the Smarten Augmented Analytics solution allows business users to prepare and analyze data with clear results without the assistance of technology staff or analysts.

The Smarten SSDP solution requires minimal training and gives business users the freedom to use simple, intuitive tools to perform sophisticated tasks, including data preparation utilizing machine learning, auto-detection, and auto-suggestion features. The comprehensive toolset allows users to gather, prepare, and analyze with only a basic working knowledge of Excel.

Users can process and work on raw data and convert and transform information into reusable analysis-ready data. The IT staff can certify the data quality so that all users understand the origin and veracity of the data as that data is published and shared with other users.



SELF-SERVE DATA PREPARATION—PROCESS

Here is a snapshot of the Self-Serve Data Preparation (SSDP) process:

Create Data Source profile:

The process of SSDP begins by identifying Data Source(s) and then creating a Data Source Profile. The Data Source contains the data the user wishes to extract, manipulate, and analyze. The Smarten Self-Serve Data Preparation component allows users to extract data from a variety of Data Source types, including Files, Databases, SAP®, R Script, and Google Analytics.

Create Dataset:

The user creates a Dataset by connecting to the Data Source to fetch the desired data. During the data fetching process, the user has the opportunity to select columns for the Dataset. Users can create many Datasets from one Data Source.

Smarten supports both Cache Datasets and Real-Time Dataset architecture. When using Real-time Datasets, the latest data is extracted from the Data Sources as and when required, and all data-related actions performed on the Dataset are performed in real time on the latest source data. For Cache Datasets, the data is not extracted in real time. Rather, it is cached and stored in a columnar data structure. Cache Datasets are updated periodically from the Data Sources with the help of a scheduler.

Process Data:

While creating a Dataset, the user can work with and explore data, and clean, shape, and blend data, employing a wide variety of functions. Users can also combine (JOIN) many Datasets with the help of auto-suggestion to identify possible JOINs and their relative value and strength. Business users can employ these tools with no required specialized skills or scripting or advanced knowledge. The intuitive interface is combined with machine learning capability and auto-detection and auto-suggest features to create analysis-ready data quickly, easily, and clearly.

Manage Data:

Users can manage Datasets, provide access rights and permissions to other users, and IT can certify data quality to help users identify quality Datasets.

Publish Data:

When a user publishes a Dataset, it is made available to other users to create Objects, such as reports, dashboards, visualization, and predictive models. During the publication process, the user can specify a Dataset as a Cache Dataset or a Real-time Dataset. When publishing a Cache Dataset, a Scheduler is created to allow for automatic update of the Dataset from Data Source(s) with a defined frequency.

The Self-Serve Data Preparation component of the Smarten solution will benefit the average organization with tools that are accessible to and suitable for a business user with average Excel skills and will provide clear, swift results that can be shared throughout the organization.

4 Data Source

Data Sources contain the raw data that has to be extracted and processed to build analysis-ready data. ElegantJ BI - Smarten allows extracting data from various types of Data Sources, such as Files, Databases, SAP®, R Script, and Google Analytics.

4.1 Creating Data Source

A Data Source Profile contains important information related to the Data Sources from which raw data has to be extracted. This information is used to establish a connection with the respective Data Sources to create a Dataset.

A Data Source Profile is created by entering required configuration parameters that are used to establish connection with the Data Source. These parameters vary with respect to the type of Data Source that is being configured. For example, while creating a Data Source Profile for a Database, users can enter configuration parameters, such as Host Name, Port Number, User Name, and Password. While creating a Data Source Profile for a File, users can upload the required files from their locations and specify the column and row delimiters.

4.2 Example

Explained below is the process of creating a Data Source Profile for a File (txt, csv, tsv) type of Data Source.

4.2.1 Creating Data Source for File

Typical steps involved in creating a Data Source Profile for File (txt, csv, tsv) are as follows:

- Select the location of Data Source, such as desktop, server, or S3
- Select source files from the desktop, server, or S3
- Specify the Row separator
- Specify the Column separator
- Specify the Text qualifier
- Save the Data Source Profile

New datasource profile

Select data source type

FILE	XLS / XLSX	XML
Text (e.g. txt, csv, tsv)		
JSON		

DATABASE

SQL server	Oracle	IBM DB2
MySQL	Postgre SQL	Sybase
Amazon redshift	SAP HANA	

Please contact your administrator to add other databases

OTHER

Google Analytics	SAP	R script
MDX		

NEXT **CANCEL**

CREATE DATASOURCE PROFILE—DATA SOURCE TYPE

New datasource profile

Text - select file(s)

Name
Datasource-1

Description
Datasource-1

Select file(s) from

My desktop Shared folder on server S3 HDFS*

Upload file(s)

age-purchase.csv (113.4kB) Success

First row contains column names

Encoding

NEXT **CANCEL** **BACK**

Success : 1 File(s) / 0 Folder(s) Failed : 0 File(s) / 0 Folder(s)

CREATE DATA SOURCE PROFILE FOR FILE

Please refer to the Smarten SSDP - User Manual for information on creating Data Sources Profiles for other types of Data Sources.

4.3 List of Data Sources

Smarten supports the following types of Data Sources:

Data Source Type	Data Source
File	
	Text (txt, csv, tsv)
	XML
	JSON
	S3
	XLS, XLSX
Web services	
	XML
	JSON
Database	
	SQL server
	MySQL

	Amazon redshift
	Oracle
	PostgreSQL
	Cassandra
	IBM DB2
	Sybase
	SAP HANA
	Amazon Athena
Other	
	Google Analytics
	SAP
	R script
	SmartenInsight

5 Dataset

Datasets are a source of analysis-ready data in the form of a columnar data structure.

The process of creating a Dataset includes extracting raw data by connecting to the Data Source and selecting the columns for it. The extracted data is displayed in a format similar to an Excel file with rows of data for selected columns. Processed with the help of a wide range of easy-to-use functions, this raw data is cleaned, shaped, and transformed into more useful analysis-ready data and stored in a columnar data structure that can be published for collaboration and reuse. Users can manage their Datasets, and IT departments can certify Datasets for validation of quality.

Users have access to a full suite of Smart Data Visualization, Plug n' Play Predictive Analysis, Dashboards, and Reports to analyze, present, and share results from the published Datasets.

5.1 Types of Dataset

Smarten supports both Cache Dataset and Real-Time Dataset architecture. Users can publish a Dataset as a Cache Dataset or a Real-Time Dataset depending upon the business need.

5.1.1 Cache Dataset

Datasets published as Cache Datasets store data in a columnar data structure.

Cache Datasets require scheduled updates to reflect the new data from the Data Sources. The automatic scheduler can be configured to update Cache Datasets automatically as per predefined frequency.

5.1.2 Real-Time Dataset

Real-time analytics is required in various use cases, such as the stock market, telecommunications, IT infrastructure management, and IoT, where users need to access data in real time.

The Real-Time Dataset does not store or cache any data. It extracts the data from Data Sources as and when required and always retrieves the latest data from Data Sources.

5.2 Creating a Dataset

A Dataset can be created from a Data Source or another Dataset.

5.2.1 Connect to Data Source

Datasets using a Data Source are created by connecting to a Data Source using a Data Source Profile and fetching the raw data from a Data Source.

5.2.2 Fetch Data for Dataset from Data Source

Raw data is fetched once connection is established with a Data Source. Due to a variety of available Data Sources, the steps for fetching data from a Data Source are unique to its Type. This may sound complex, but the data extraction process, guided by a highly intuitive interface of Smarten, empowered by features of auto suggestions and auto-detection, helps users to extract data from a variety of different Data Sources with minimal effort and expertise.

5.2.2.1 Column Selection

By allowing users to select columns from the Data Sources, Smarten SSDP provides them with the flexibility to load only the required data into the Datasets.

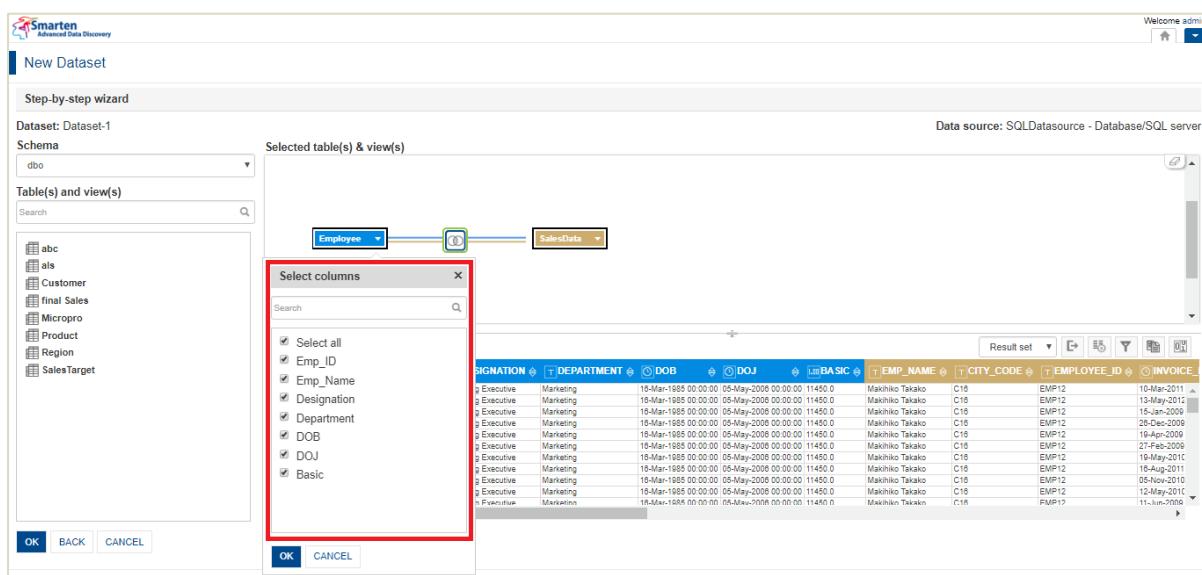
While fetching data, the option to select columns may appear at different stages for different Data Source Types.

For example, while creating Datasets from a File type of Data Source, the user is presented with a list of all columns belonging to the Data Source as soon as the connection is established.

#	YEAR	QUARTER	MONTH	DAY_OF_MONTH	DAY_OF_WEEK	FL_DATE	UNIQUE_CARRIER	AIRLINE_ID	CARRIER	TAIL_NUM	FL_NUM	ORIGIN_AIRPORT_ID	ORIGIN_AIRPORT_SEQ_ID	ORIGIN_CITY_MARKET_ID
1	2018	2	6	2	4	02-Jun-2018	AA	19805	AA	N5F5AA	1	10721	1	1
2	2018	2	6	3	5	03-Jun-2018	AA	19805	AA	N784AA	1	10721	1	1
3	2018	2	6	4	6	04-Jun-2018	AA	19805	AA	N796AA	1	10721	1	1
4	2018	2	6	5	1	05-Jun-2018	AA	19805	AA	N797AA	1	10721	1	1
5	2018	2	6	7	2	07-Jun-2018	AA	19805	AA	N798AA	1	10721	1	1
6	2018	2	6	8	3	08-Jun-2018	AA	19805	AA	N783AA	1	10721	1	1
7	2018	2	6	9	4	09-Jun-2018	AA	19805	AA	N794AA	1	10721	1	1
8	2018	2	6	10	5	10-Jun-2018	AA	19805	AA	N791AA	1	10721	1	1
9	2018	2	6	11	6	11-Jun-2018	AA	19805	AA	N788AA	1	10721	1	1
10	2018	2	6	13	1	13-Jun-2018	AA	19805	AA	N797AA	1	10721	1	1
11	2018	2	6	14	2	14-Jun-2018	AA	19805	AA	N785AA	1	10721	1	1
12	2018	2	6	15	3	15-Jun-2018	AA	19805	AA	N786AA	1	10721	1	1
13	2018	2	6	16	4	16-Jun-2018	AA	19805	AA	N784AA	1	10721	1	1
14	2018	2	6	17	5	17-Jun-2018	AA	19805	AA	N783AA	1	10721	1	1
15	2018	2	6	18	6	18-Jun-2018	AA	19805	AA	N792AA	1	10721	1	1
16	2018	2	6	20	1	20-Jun-2018	AA	19805	AA	N785AA	1	10721	1	1
17	2018	2	6	21	2	21-Jun-2018	AA	19805	AA	N784AA	1	10721	1	1
18	2018	2	6	22	3	22-Jun-2018	AA	19805	AA	N788AA	1	10721	1	1
19	2018	2	6	23	4	23-Jun-2018	AA	19805	AA	N796AA	1	10721	1	1
20	2018	2	6	24	5	24-Jun-2018	AA	19805	AA	N788AA	1	10721	1	1
21	2018	2	6	25	6	25-Jun-2018	AA	19805	AA	N795AA	1	10721	1	1
22	2018	2	6	27	1	27-Jun-2018	AA	19805	AA	N783AA	1	10721	1	1
23	2018	2	6	28	2	28-Jun-2018	AA	19805	AA	N788AA	1	10721	1	1
24	2018	2	6	29	3	29-Jun-2018	AA	19805	AA	N792AA	1	10721	1	1
25	2018	2	6	30	4	30-Jun-2018	AA	19805	AA	N797AA	1	10721	1	1
26	2018	2	6	1	3	01-Jun-2018	AA	19805	AA	N794AA	1	12478	1	1
27	2018	2	6	2	4	02-Jun-2018	AA	19805	AA	N796AA	1	12478	1	1
28	2018	2	6	3	5	03-Jun-2018	AA	19805	AA	N792AA	1	12478	1	1

COLUMN SELECTION—FILE AS DATA SOURCE

For Datasets created from a Database type of Data Source, the user has an option to select columns of each Table selected from the Database or from the result set.



COLUMN SELECTION—DATABASE AS DATA SOURCE

Please refer to the Smarten SSDP - User Manual for information on column selection from other types of Data Sources.

5.2.3 Example

Explained below is the process of creating a Dataset from a Database Data Source.

5.2.3.1 Creating a Dataset from a Database Data Source

Typical steps involved in creating a Dataset from a Database Data Source are as follows:

Step 1: Create Data Source Profile for Database

Step 2: Connect to the Data Source Profile to fetch data

Step 3: Select the data to be extracted through Graphical query builder or by just pasting the ready-to-use query

The graphical query builder provides an easy-to-use graphical interface to select Tables from a particular Database schema and establish a relationship between them. All that a user has to do is select the required Tables from the schema, and the rest is taken care by the system. The auto-detection feature of Smarten SSDP automatically detects the common columns between the selected Tables and suggests a connection between appropriate Tables. This allows users to quickly create a complex structure of interrelated Tables with ease.

As this structure of interrelated Tables is represented through a graphical chart, it is very easy for users to make changes to the structure. Users are able to perform many actions, such as adding or removing Tables in the structure, selecting columns for a particular Table, or even changing the JOINS between the Tables. Changes made to the structure are reflected instantly in the chart and in the result set.

For example, in the image shown below, “Customer” table is connected to tables “CustomerAddress,” “SalesOrderHeader,” “SalesPerson,” and “SpecialOffer,” and “SalesOrderHeader” is further connected with the table of “SalesOrderDetail.”

The screenshot shows the Smarten Advanced Data Discovery interface. On the left, there's a schema tree under 'Schema' (Sales) and a list of tables under 'Table(s) and view(s)' (e.g., Individual, SalesOrderHeaderSalesReason, SalesPersonQuotaHistory, SalesReason, SalesTaxRate, SalesTerritory, SalesTerritoryHistory, ShoppingCartItem, SpecialOfferProduct, Store, StoreContact, VIndividualCustomer, VIndividualDemographics). In the center, a 'Selected table(s) & view(s)' diagram shows entities like Customer, CustomerAddress, SalesOrderHeader, SalesOrderDetail, and SpecialOffer connected by arrows. Below this is a 'Result set' table with the following data:

#	CUSTOMERID	TERRITORYID	ACCOUNTNUMBER	CUSTOMERTYPE	ROWGUID	MODIFIEDDATE	CUSTOMERID
1	576	2	AW00000578	S	566E9FAD-066F-4FF-9B9D-D848EEA5B7AC	13-Oct-2004 11:15:07	576
2	12026	4	AW000012026	I	55AD71A8-7FB4-457B-B9B-E3C708BB505BB	13-Oct-2004 11:15:07	12026
3	20954	7	AW000020954	I	903A190C-A3E7-4211-9B75-42D2ADE882BD	13-Oct-2004 11:15:07	20954
4	11247	10	AW000011247	I	EA39D9E-FB24-445B-B181-7E6B6B4CBA05	13-Oct-2004 11:15:07	11247
5	17559	4	AW000017559	I	524D4A55-F032-4494-8843-0D2FEE33P95E9	13-Oct-2004 11:15:07	17559
6	28347	1	AW000028347	I	B6BE8EB5-A1FF-451A-944D-F50327F61643	13-Oct-2004 11:15:07	28347

Step 4: Extract the data and create a Dataset

Raw data is extracted from the Data Sources into the Dataset allowing users to clean, shape, and blend the data for further analysis.

Smarten SSDP supports many Data Sources, such as RDBMS, File (e.g., CSV, Text), Web Services, SAP, and Google Analytics. Please refer to the Smarten SSDP - User Manual for information on creating Datasets from various types of Data Sources.

6 Explore Data

Smarten SSDP allows users to explore the raw data and refine it into more useful analysis-ready data by using a wide variety of functions.

6.1 Highlight

Users can identify the data by highlighting it for missing and inconsistent values, spaces, and duplication of rows and columns and accordingly take action for data correction.

Note:

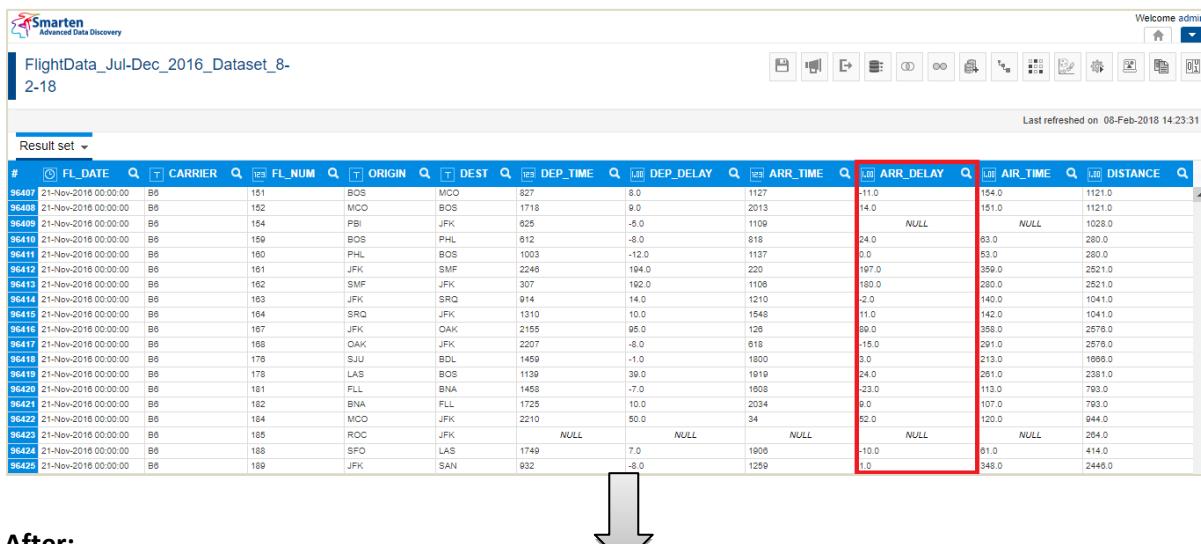
These are the front-end operations that highlight records in the front-end interface and do not remove any records from the Dataset.

6.1.1 Highlight Missing Values

This function allows users to find data containing null or blank values. It highlights all the rows of a target column that has null or blank values in the data.

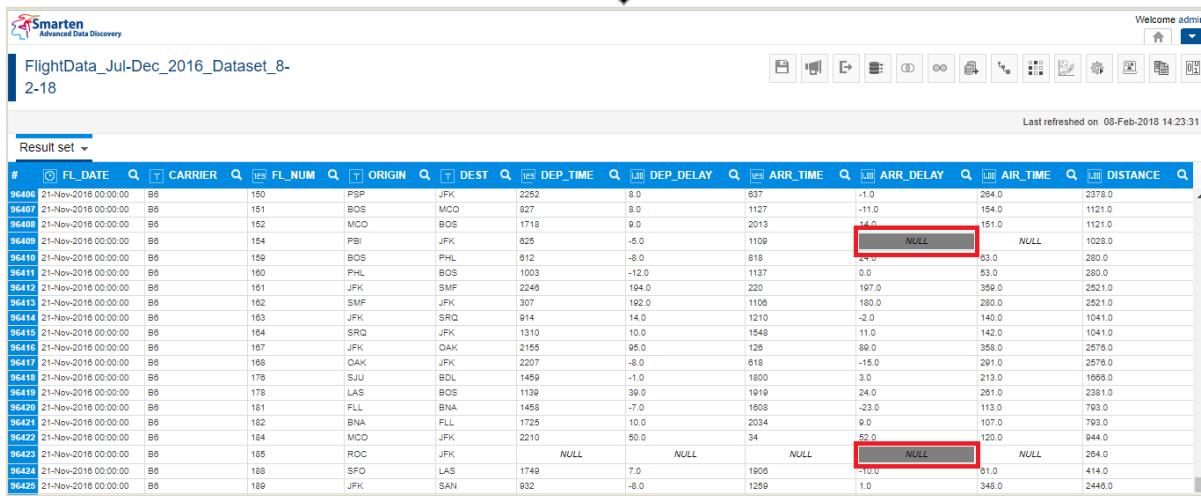
Shown below is the before and after scenario of “Highlight Missing Values” in column “ARR_DELAY”:

Before:



#	FL_DATE	CARRIER	FL_NUM	ORIGIN	DEST	DEP_TIME	DEP_DELAY	ARR_TIME	ARR_DELAY	AIR_TIME	DISTANCE
96407	21-Nov-2016 00:00:00	B6	151	BOS	MCO	827	8.0	1127	-11.0	154.0	1121.0
96408	21-Nov-2016 00:00:00	B6	152	MCO	BOS	1718	9.0	2013	-14.0	151.0	1121.0
96409	21-Nov-2016 00:00:00	B6	154	PBI	JFK	625	-5.0	1109	NULL	1028.0	
96410	21-Nov-2016 00:00:00	B6	159	BOS	PHL	612	-8.0	818	24.0	63.0	280.0
96411	21-Nov-2016 00:00:00	B6	160	PHL	BOS	1003	-12.0	1137	0.0	63.0	280.0
96412	21-Nov-2016 00:00:00	B6	161	JFK	SMF	2248	104.0	220	197.0	359.0	2521.0
96413	21-Nov-2016 00:00:00	B6	162	SMF	JFK	307	192.0	1108	180.0	280.0	2521.0
96414	21-Nov-2016 00:00:00	B6	163	JFK	SRQ	914	14.0	1210	-2.0	140.0	1041.0
96415	21-Nov-2016 00:00:00	B6	164	SRQ	JFK	1310	10.0	1548	11.0	142.0	1041.0
96416	21-Nov-2016 00:00:00	B6	167	JFK	OAK	2155	95.0	128	89.0	358.0	2578.0
96417	21-Nov-2016 00:00:00	B6	168	OAK	JFK	2207	-8.0	618	-15.0	291.0	2578.0
96418	21-Nov-2016 00:00:00	B6	178	SJU	BDL	1459	-1.0	1800	2.0	213.0	1668.0
96419	21-Nov-2016 00:00:00	B6	178	LAS	BOS	1139	30.0	1919	24.0	261.0	2381.0
96420	21-Nov-2016 00:00:00	B6	181	FLL	BNA	1458	-7.0	1608	-23.0	113.0	763.0
96421	21-Nov-2016 00:00:00	B6	182	BNA	FLL	1725	10.0	2034	9.0	107.0	763.0
96422	21-Nov-2016 00:00:00	B6	184	MCO	JFK	2210	50.0	34	52.0	120.0	944.0
96423	21-Nov-2016 00:00:00	B6	185	ROC	JFK	NULL	NULL	NULL	NULL	NULL	264.0
96424	21-Nov-2016 00:00:00	B6	188	SFO	LAS	1749	7.0	1908	-10.0	61.0	414.0
96425	21-Nov-2016 00:00:00	B6	189	JFK	SAN	932	-8.0	1259	1.0	348.0	2448.0

After:



#	FL_DATE	CARRIER	FL_NUM	ORIGIN	DEST	DEP_TIME	DEP_DELAY	ARR_TIME	ARR_DELAY	AIR_TIME	DISTANCE
96406	21-Nov-2016 00:00:00	B6	150	PSP	JFK	2252	8.0	637	-1.0	264.0	2378.0
96407	21-Nov-2016 00:00:00	B6	151	BOS	MCO	827	8.0	1127	-11.0	154.0	1121.0
96408	21-Nov-2016 00:00:00	B6	152	MCO	BOS	1718	9.0	2013	-14.0	151.0	1121.0
96409	21-Nov-2016 00:00:00	B6	154	PBI	JFK	625	-5.0	1109	NULL	1028.0	
96410	21-Nov-2016 00:00:00	B6	159	BOS	PHL	612	-8.0	818	24.0	63.0	280.0
96411	21-Nov-2016 00:00:00	B6	160	PHL	BOS	1003	-12.0	1137	0.0	63.0	280.0
96412	21-Nov-2016 00:00:00	B6	161	JFK	SMF	2248	104.0	220	197.0	359.0	2521.0
96413	21-Nov-2016 00:00:00	B6	162	SMF	JFK	307	192.0	1108	180.0	280.0	2521.0
96414	21-Nov-2016 00:00:00	B6	163	JFK	SRQ	914	14.0	1210	-2.0	140.0	1041.0
96415	21-Nov-2016 00:00:00	B6	164	SRQ	JFK	1310	10.0	1548	11.0	142.0	1041.0
96416	21-Nov-2016 00:00:00	B6	167	JFK	OAK	2155	95.0	128	89.0	358.0	2578.0
96417	21-Nov-2016 00:00:00	B6	168	OAK	JFK	2207	-8.0	618	-15.0	291.0	2578.0
96418	21-Nov-2016 00:00:00	B6	178	SJU	BDL	1459	-1.0	1800	2.0	213.0	1668.0
96419	21-Nov-2016 00:00:00	B6	178	LAS	BOS	1139	30.0	1919	24.0	261.0	2381.0
96420	21-Nov-2016 00:00:00	B6	181	FLL	BNA	1458	-7.0	1608	-23.0	113.0	763.0
96421	21-Nov-2016 00:00:00	B6	182	BNA	FLL	1725	10.0	2034	9.0	107.0	763.0
96422	21-Nov-2016 00:00:00	B6	184	MCO	JFK	2210	50.0	34	52.0	120.0	944.0
96423	21-Nov-2016 00:00:00	B6	185	ROC	JFK	NULL	NULL	NULL	NULL	NULL	264.0
96424	21-Nov-2016 00:00:00	B6	188	SFO	LAS	1749	7.0	1908	-10.0	61.0	414.0
96425	21-Nov-2016 00:00:00	B6	189	JFK	SAN	932	-8.0	1259	1.0	348.0	2448.0

HIGHLIGHT MISSING VALUES

6.1.2 Highlight Spaces

This function helps users identify data containing spaces within the string type of columns. It highlights all the rows of a target column that has spaces in the data.

Note:

This function is applicable for string type-data only.

Shown below is the before and after scenario of “Highlight Spaces” in column “FL_DATE”:

Before:

#	FL_DATE	CARRIER	FL_NUM	ORIGIN	DEST	DEP_TIME	DEP_DELAY	ARR_TIME	ARR_DELAY	AIR_TIME	DISTANCE
1	01-Nov-2016 00:00:00	AA	1	JFK	LAX	824	-6.0	1137	-4.0	330.0	2475.0
2	02-Nov-2016 00:00:00	AA	1	JFK	LAX	824	-6.0	1128	-13.0	318.0	2475.0
3	03-Nov-2016 00:00:00	AA	1	JFK	LAX	831	1.0	1138	-3.0	331.0	2475.0
4	04-Nov-2016 00:00:00	AA	1	JFK	LAX	755	-5.0	1045	-57.0	312.0	2475.0
5	05-Nov-2016 00:00:00	AA	1	JFK	LAX	753	-7.0	1043	-59.0	310.0	2475.0
6	07-Nov-2016 00:00:00	AA	1	JFK	LAX	759	-2.0	1045	-57.0	300.0	2475.0
7	08-Nov-2016 00:00:00	AA	1	JFK	LAX	755	-6.0	1033	-66.0	297.0	2475.0
8	09-Nov-2016 00:00:00	AA	1	JFK	LAX	757	-3.0	1024	-78.0	283.0	2475.0
9	10-Nov-2016 00:00:00	AA	1	JFK	LAX	758	-2.0	1044	-58.0	285.0	2475.0
10	11-Nov-2016 00:00:00	AA	1	JFK	LAX	810	16.0	1056	-48.0	312.0	2475.0
11	13-Nov-2016 00:00:00	AA	1	JFK	LAX	845	45.0	1112	-30.0	292.0	2475.0
12	14-Nov-2016 00:00:00	AA	1	JFK	LAX	841	41.0	1202	20.0	306.0	2475.0
13	15-Nov-2016 00:00:00	AA	1	JFK	LAX	821	21.0	1125	-17.0	316.0	2475.0
14	16-Nov-2016 00:00:00	AA	1	JFK	LAX	751	-6.0	1052	-50.0	340.0	2475.0

After:

#	FL_DATE	CARRIER	FL_NUM	ORIGIN	DEST	DEP_TIME	DEP_DELAY	ARR_TIME	ARR_DELAY	AIR_TIME	DISTANCE
1	01-Nov-2016 00:00:00	AA	1	JFK	LAX	824	-6.0	1137	-4.0	330.0	2475.0
2	02-Nov-2016 00:00:00	AA	1	JFK	LAX	824	-6.0	1128	-13.0	318.0	2475.0
3	03-Nov-2016 00:00:00	AA	1	JFK	LAX	831	1.0	1138	-3.0	331.0	2475.0
4	04-Nov-2016 00:00:00	AA	1	JFK	LAX	755	-5.0	1045	-57.0	312.0	2475.0
5	05-Nov-2016 00:00:00	AA	1	JFK	LAX	753	-7.0	1043	-59.0	310.0	2475.0
6	07-Nov-2016 00:00:00	AA	1	JFK	LAX	758	-2.0	1045	-57.0	300.0	2475.0
7	08-Nov-2016 00:00:00	AA	1	JFK	LAX	755	-6.0	1033	-66.0	297.0	2475.0
8	09-Nov-2016 00:00:00	AA	1	JFK	LAX	757	-3.0	1024	-78.0	283.0	2475.0
9	10-Nov-2016 00:00:00	AA	1	JFK	LAX	759	-2.0	1044	-58.0	285.0	2475.0
10	11-Nov-2016 00:00:00	AA	1	JFK	LAX	810	16.0	1056	-48.0	312.0	2475.0
11	13-Nov-2016 00:00:00	AA	1	JFK	LAX	845	45.0	1112	-30.0	292.0	2475.0
12	14-Nov-2016 00:00:00	AA	1	JFK	LAX	841	41.0	1202	20.0	306.0	2475.0
13	15-Nov-2016 00:00:00	AA	1	JFK	LAX	821	21.0	1125	-17.0	316.0	2475.0
14	16-Nov-2016 00:00:00	AA	1	JFK	LAX	751	-6.0	1052	-50.0	340.0	2475.0

HIGHLIGHT SPACES

6.1.3 Highlight Inconsistent Values

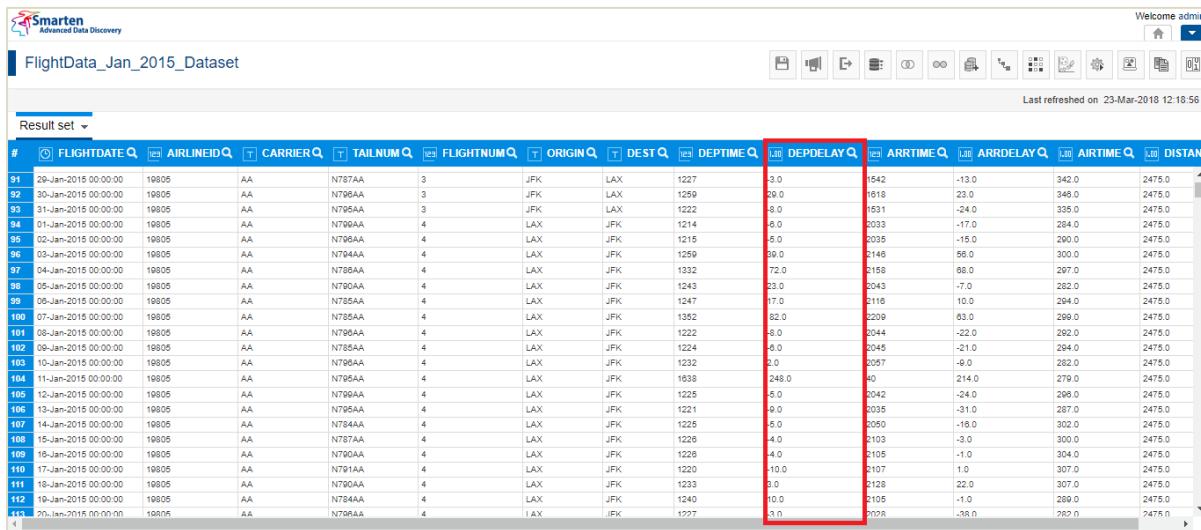
This function helps users to find discrepancies within a column, such as a string type of column having integer values and vice versa, or capitalization not staying the same throughout, and many more. It highlights all the inconsistent values of the target column.

Note:

This function is applicable for string-type data only.

Shown below is the before and after scenario of “Highlight Inconsistent values” in column “DEPDELAY.” The function highlights all those values that have a leading space:

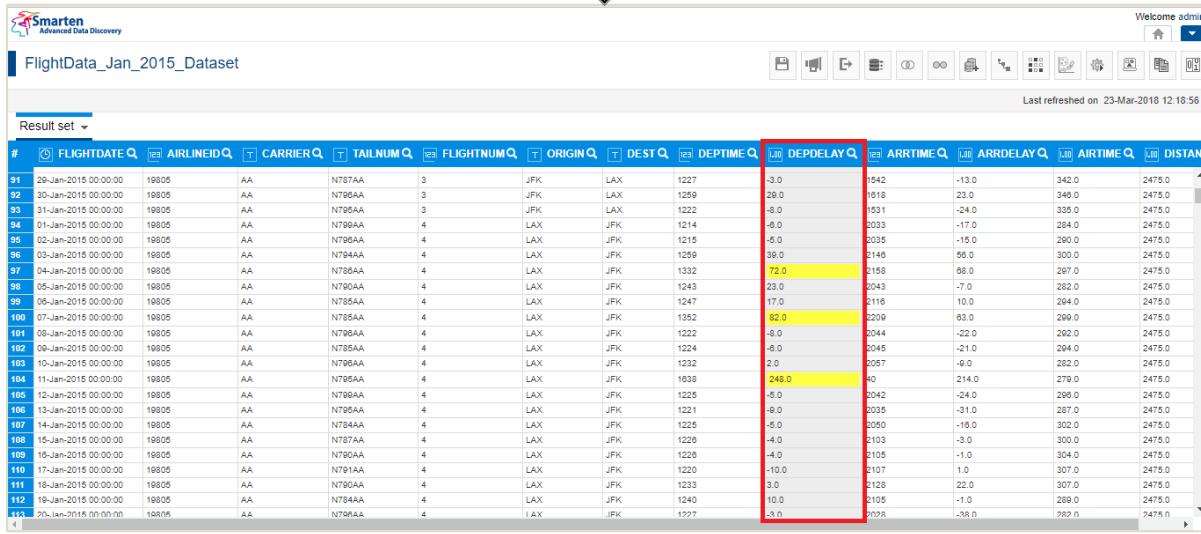
Before:



The screenshot shows a data grid with 111 rows of flight data. The columns include: #, FLIGHTDATE_Q, AIRLINEID_Q, CARRIER_Q, TAILNUM_Q, FLIGHTNUM_Q, ORIGIN_Q, DEST_Q, DEPTIME_Q, DEPDELAY_Q, ARRTIME_Q, ARRDELAY_Q, AIRTIME_Q, and DISTANCE. The 'DEPDELAY_Q' column is highlighted with a red border.



After:



The screenshot shows the same data grid as before, but the 'DEPDELAY_Q' column now contains inconsistent values. The first two rows show -3.0, while subsequent rows show values like 72.0, 82.0, and 248.0. The rest of the columns remain the same.

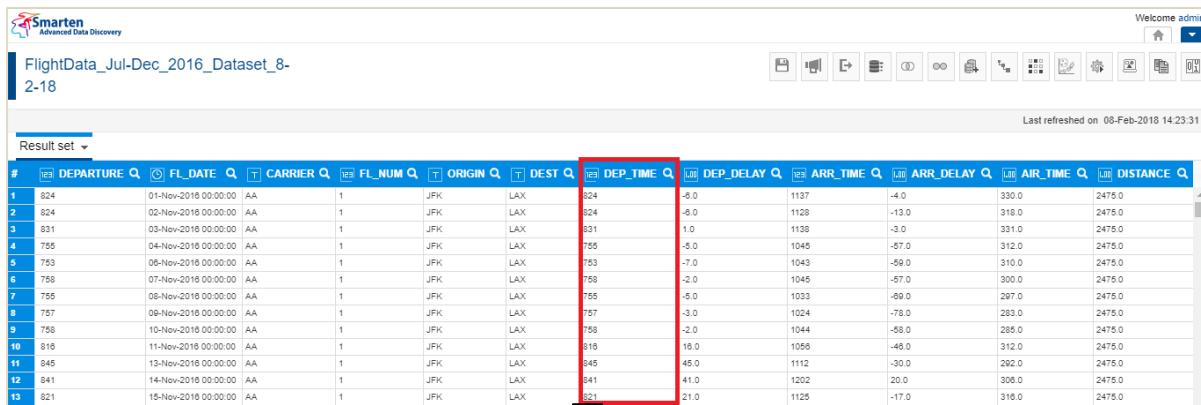
HIGHLIGHT INCONSISTENT VALUES

6.1.4 Highlight Duplicate columns with this column

This function helps users find duplicates of a column. It highlights duplicate columns of the target column. A duplicate column is a column that contains exactly the same data as that of the target column.

Shown below is the before and after scenario of “Highlight Duplicate columns with this column” for column “DEP_TIME”:

Before:



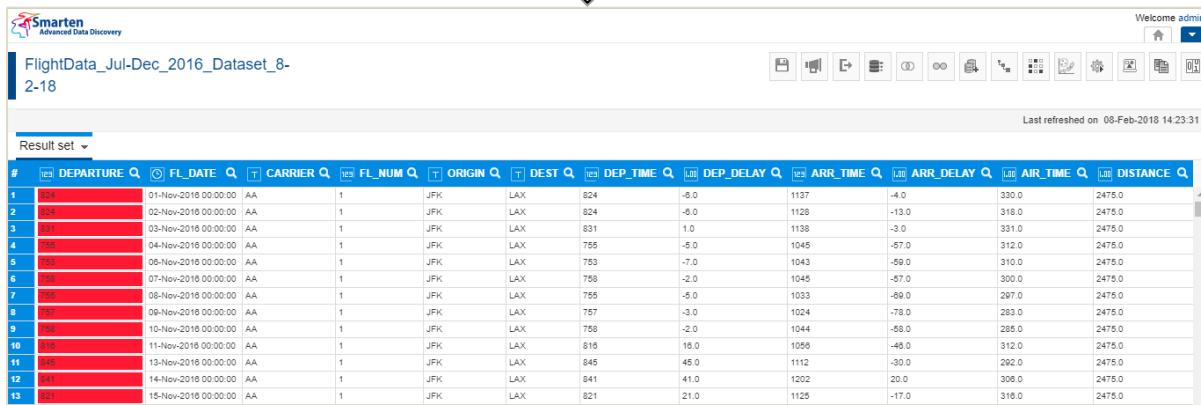
FlightData_Jul-Dec_2016_Dataset_8-
2-18

Last refreshed on 08-Feb-2018 14:23:31

Result set ▾

#	DEPARTURE Q	FL_DATE Q	CARRIER Q	FL_NUM Q	ORIGIN Q	DEST Q	DEP_TIME Q	DEP_DELAY Q	ARR_TIME Q	ARR_DELAY Q	AIR_TIME Q	DISTANCE Q
1	824	01-Nov-2016 00:00:00	AA	1	JFK	LAX	824	-8.0	1137	-4.0	330.0	2475.0
2	824	02-Nov-2016 00:00:00	AA	1	JFK	LAX	824	-8.0	1128	-13.0	318.0	2475.0
3	831	03-Nov-2016 00:00:00	AA	1	JFK	LAX	831	1.0	1138	-3.0	331.0	2475.0
4	755	04-Nov-2016 00:00:00	AA	1	JFK	LAX	755	-5.0	1045	-57.0	312.0	2475.0
5	753	06-Nov-2016 00:00:00	AA	1	JFK	LAX	753	-7.0	1043	-59.0	310.0	2475.0
6	758	07-Nov-2016 00:00:00	AA	1	JFK	LAX	758	-2.0	1045	-57.0	300.0	2475.0
7	755	08-Nov-2016 00:00:00	AA	1	JFK	LAX	755	-5.0	1033	-69.0	297.0	2475.0
8	757	09-Nov-2016 00:00:00	AA	1	JFK	LAX	757	-3.0	1024	-78.0	283.0	2475.0
9	758	10-Nov-2016 00:00:00	AA	1	JFK	LAX	758	-2.0	1044	-58.0	285.0	2475.0
10	816	11-Nov-2016 00:00:00	AA	1	JFK	LAX	816	19.0	1055	-46.0	312.0	2475.0
11	845	13-Nov-2016 00:00:00	AA	1	JFK	LAX	845	45.0	1112	-30.0	292.0	2475.0
12	841	14-Nov-2016 00:00:00	AA	1	JFK	LAX	841	41.0	1202	20.0	308.0	2475.0
13	821	15-Nov-2016 00:00:00	AA	1	JFK	LAX	821	21.0	1125	-17.0	316.0	2475.0

After:



FlightData_Jul-Dec_2016_Dataset_8-
2-18

Last refreshed on 08-Feb-2018 14:23:31

Result set ▾

#	DEPARTURE Q	FL_DATE Q	CARRIER Q	FL_NUM Q	ORIGIN Q	DEST Q	DEP_TIME Q	DEP_DELAY Q	ARR_TIME Q	ARR_DELAY Q	AIR_TIME Q	DISTANCE Q
1	824	01-Nov-2016 00:00:00	AA	1	JFK	LAX	824	-8.0	1137	-4.0	330.0	2475.0
2	824	02-Nov-2016 00:00:00	AA	1	JFK	LAX	824	-8.0	1128	-13.0	318.0	2475.0
3	831	03-Nov-2016 00:00:00	AA	1	JFK	LAX	831	1.0	1138	-3.0	331.0	2475.0
4	755	04-Nov-2016 00:00:00	AA	1	JFK	LAX	755	-5.0	1045	-57.0	312.0	2475.0
5	753	06-Nov-2016 00:00:00	AA	1	JFK	LAX	753	-7.0	1043	-59.0	310.0	2475.0
6	758	07-Nov-2016 00:00:00	AA	1	JFK	LAX	758	-2.0	1045	-57.0	300.0	2475.0
7	755	08-Nov-2016 00:00:00	AA	1	JFK	LAX	755	-5.0	1033	-69.0	297.0	2475.0
8	757	09-Nov-2016 00:00:00	AA	1	JFK	LAX	757	-3.0	1024	-78.0	283.0	2475.0
9	758	10-Nov-2016 00:00:00	AA	1	JFK	LAX	758	-2.0	1044	-58.0	285.0	2475.0
10	816	11-Nov-2016 00:00:00	AA	1	JFK	LAX	816	19.0	1055	-46.0	312.0	2475.0
11	845	13-Nov-2016 00:00:00	AA	1	JFK	LAX	845	45.0	1112	-30.0	292.0	2475.0
12	841	14-Nov-2016 00:00:00	AA	1	JFK	LAX	841	41.0	1202	20.0	308.0	2475.0
13	821	15-Nov-2016 00:00:00	AA	1	JFK	LAX	821	21.0	1125	-17.0	316.0	2475.0

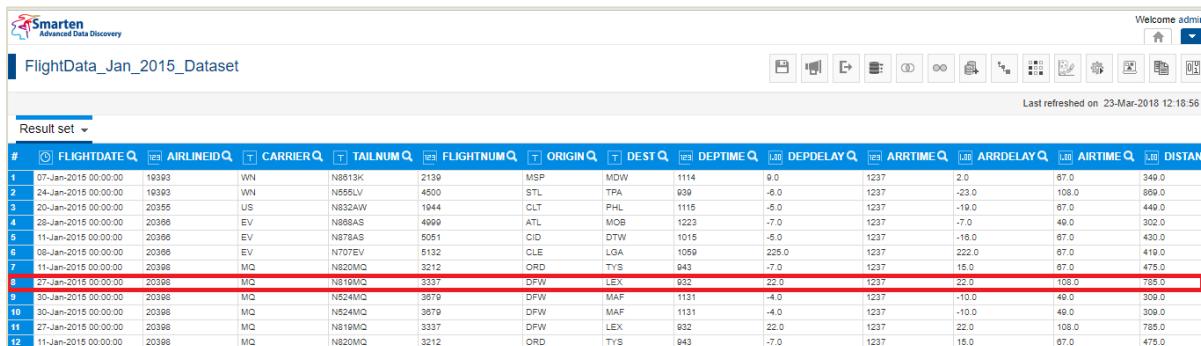
HIGHLIGHT DUPLICATE COLUMNS WITH THIS COLUMN

6.1.5 Highlight Duplicate rows with this row

This function helps users find duplicates of a row. It highlights duplicate rows of the target row. A duplicate row is a row that contains exactly the same data as that of the target row.

Shown below is the before and after scenario of “Highlight Duplicate rows with this row”:

Before:



FlightData_Jan_2015_Dataset

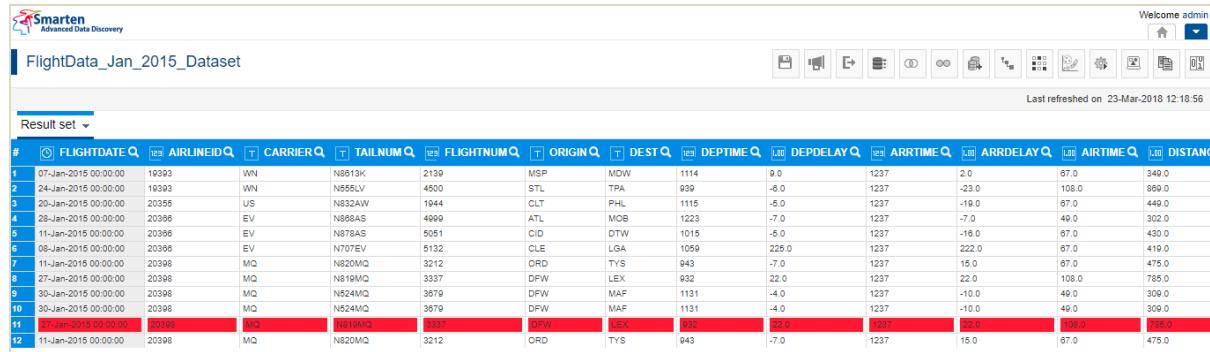
Last refreshed on 23-Mar-2018 12:18:56

Result set ▾

#	FLIGHTDATE Q	AIRLINEID Q	CARRIER Q	TAILNUM Q	FLIGHTNUM Q	ORIGIN Q	DEST Q	DEPTIME Q	DEPDDELAY Q	ARRTIME Q	ARRDELAY Q	AIRTIME Q	DISTANCE Q
1	07-Jan-2015 00:00:00	19393	WN	N8613K	2139	MSP	MDW	1114	9.0	1237	2.0	87.0	349.0
2	24-Jan-2015 00:00:00	19393	WN	N855LV	4500	STL	TPA	939	-8.0	1237	-23.0	108.0	869.0
3	20-Jan-2015 00:00:00	20355	US	N832AW	1044	CLT	PHL	1115	-8.0	1237	-19.0	87.0	449.0
4	28-Jan-2015 00:00:00	20388	EV	N868AS	4000	ATL	MDB	1223	-7.0	1237	-7.0	49.0	302.0
5	11-Jan-2015 00:00:00	20388	EV	N878AS	5051	CID	DTW	1015	-8.0	1237	-16.0	67.0	430.0
6	08-Jan-2015 00:00:00	20388	EV	N707EV	5132	CLE	LGA	1059	225.0	1237	22.0	67.0	419.0
7	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	943	-7.0	1237	15.0	87.0	475.0
8	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	932	22.0	1237	22.0	108.0	765.0
9	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3079	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
10	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3079	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
11	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	932	22.0	1237	22.0	108.0	765.0
12	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	943	-7.0	1237	15.0	87.0	475.0



After:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	07-Jan-2015 00:00:00	19393	VN	N8613K	2139	MSP	MWD	1114	-9.0	1237	-2.0	67.0	349.0
2	24-Jan-2015 00:00:00	19393	VN	N555LV	4500	STL	TPA	039	-8.0	1237	-23.0	108.0	869.0
3	20-Jan-2015 00:00:00	20355	US	N832AW	1944	CLT	FHL	1115	-5.0	1237	-19.0	67.0	449.0
4	28-Jan-2015 00:00:00	20366	EV	N868AS	4999	ATL	MOB	1223	-7.0	1237	-7.0	49.0	302.0
5	11-Jan-2015 00:00:00	20366	EV	N878AS	5051	CID	DTW	1015	-5.0	1237	-18.0	67.0	430.0
6	08-Jan-2015 00:00:00	20366	EV	N707EV	5132	CLE	LGA	1059	225.0	1237	222.0	67.0	419.0
7	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	043	-7.0	1237	15.0	67.0	475.0
8	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	032	22.0	1237	22.0	108.0	785.0
9	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3579	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
10	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3579	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
11	01-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	030	-2.0	1237	15.0	67.0	475.0
12	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	043	-7.0	1237	15.0	67.0	475.0

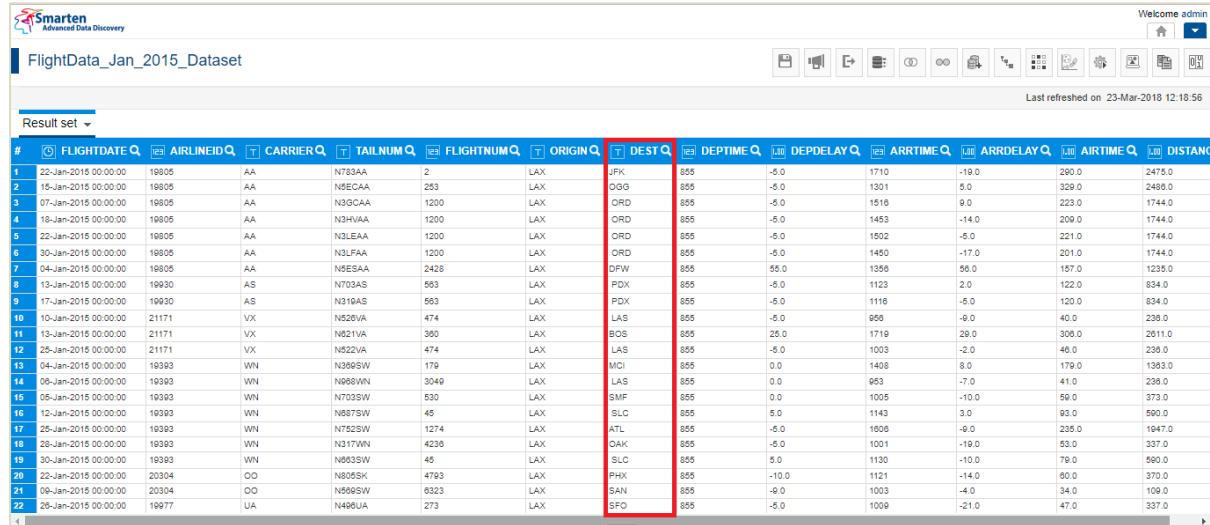
HIGHLIGHT DUPLICATE ROWS WITH THIS ROW

6.1.6 Highlight Duplicate column values

This function helps users find duplicate values in the same column. All duplicate values in the target column are highlighted.

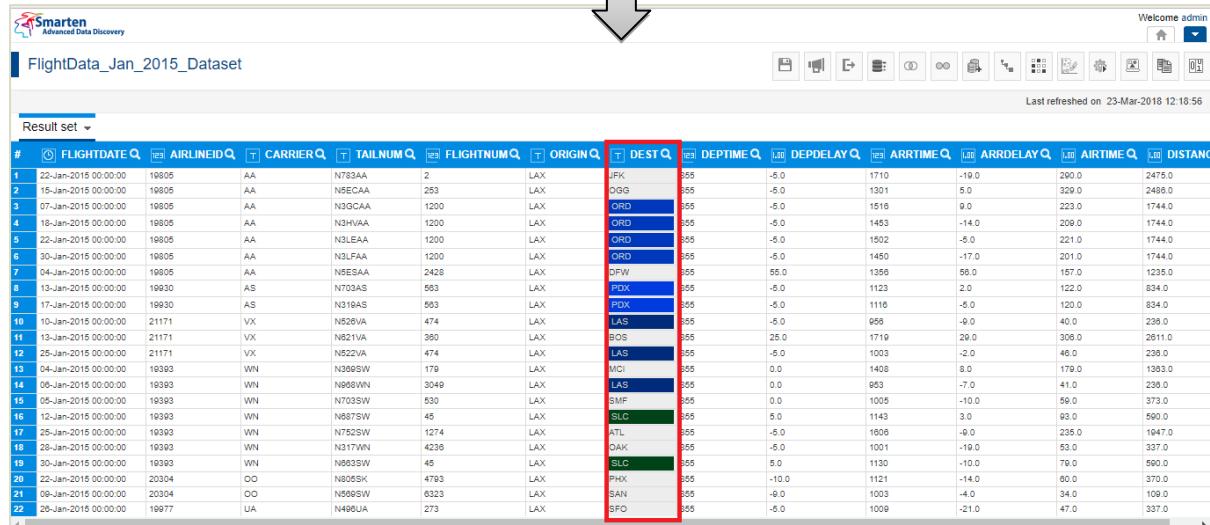
Shown below is the before and after scenario of “Highlight Duplicate column values” in column “DEST”:

Before:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	22-Jan-2015 00:00:00	19805	AA	N783AA	2	LAX	JFK	855	-5.0	1710	-19.0	200.0	2475.0
2	15-Jan-2015 00:00:00	19805	AA	N5ECAA	253	LAX	OOG	855	-5.0	1301	5.0	329.0	2486.0
3	07-Jan-2015 00:00:00	19805	AA	N3GCAA	1200	LAX	ORD	855	-5.0	1516	9.0	223.0	1744.0
4	18-Jan-2015 00:00:00	19805	AA	N3HVAA	1200	LAX	ORD	855	-5.0	1453	-14.0	209.0	1744.0
5	22-Jan-2015 00:00:00	19805	AA	N3LEAA	1200	LAX	ORD	855	-5.0	1502	-5.0	221.0	1744.0
6	30-Jan-2015 00:00:00	19805	AA	N3LFAA	1200	LAX	ORD	855	-5.0	1450	-17.0	201.0	1744.0
7	04-Jan-2015 00:00:00	19805	AA	N5ESAA	2428	LAX	DFW	855	55.0	1359	58.0	157.0	1235.0
8	13-Jan-2015 00:00:00	19930	AS	N703AS	563	LAX	PDX	855	-5.0	1123	2.0	122.0	834.0
9	17-Jan-2015 00:00:00	19930	AS	N310AS	563	LAX	PDX	855	-5.0	1118	-5.0	120.0	834.0
10	10-Jan-2015 00:00:00	21171	VX	N520VA	474	LAX	LAS	855	-5.0	058	-9.0	40.0	238.0
11	13-Jan-2015 00:00:00	21171	VX	N521VA	300	LAX	BOS	855	25.0	1719	29.0	306.0	2611.0
12	25-Jan-2015 00:00:00	21171	VX	N522VA	474	LAX	LAS	855	-5.0	1003	-2.0	46.0	238.0
13	04-Jan-2015 00:00:00	19933	WN	N369SW	179	LAX	MCI	855	0.0	1408	8.0	179.0	1383.0
14	05-Jan-2015 00:00:00	19933	WN	N968WN	3049	LAX	LAS	855	0.0	053	-7.0	41.0	238.0
15	05-Jan-2015 00:00:00	19933	WN	N703SW	530	LAX	SF	855	0.0	1005	-10.0	59.0	373.0
16	12-Jan-2015 00:00:00	19933	WN	N875SW	45	LAX	SLC	855	5.0	1143	3.0	93.0	560.0
17	25-Jan-2015 00:00:00	19933	WN	N752SW	1274	LAX	ATL	855	-5.0	1806	-9.0	235.0	1947.0
18	28-Jan-2015 00:00:00	19933	WN	N317WN	4238	LAX	OAK	855	-5.0	1001	-19.0	53.0	337.0
19	30-Jan-2015 00:00:00	19933	WN	N983SW	45	LAX	SLC	855	5.0	1130	-10.0	79.0	560.0
20	22-Jan-2015 00:00:00	20304	OO	N805SK	4793	LAX	PHX	855	-10.0	1121	-14.0	60.0	370.0
21	09-Jan-2015 00:00:00	20304	OO	N568SW	6323	LAX	SAN	855	-9.0	1003	-4.0	34.0	109.0
22	26-Jan-2015 00:00:00	19977	UA	N490UA	273	LAX	SFO	855	-5.0	1009	-21.0	47.0	337.0

After:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	22-Jan-2015 00:00:00	19805	AA	N783AA	2	LAX	JFK	855	-5.0	1710	-19.0	200.0	2475.0
2	15-Jan-2015 00:00:00	19805	AA	N5ECAA	253	LAX	OOG	855	-5.0	1301	5.0	329.0	2486.0
3	07-Jan-2015 00:00:00	19805	AA	N3GCAA	1200	LAX	ORD	855	-5.0	1516	9.0	223.0	1744.0
4	18-Jan-2015 00:00:00	19805	AA	N3HVAA	1200	LAX	ORD	855	-5.0	1453	-14.0	209.0	1744.0
5	22-Jan-2015 00:00:00	19805	AA	N3LEAA	1200	LAX	ORD	855	-5.0	1502	-5.0	221.0	1744.0
6	30-Jan-2015 00:00:00	19805	AA	N3LFAA	1200	LAX	ORD	855	-5.0	1450	-17.0	201.0	1744.0
7	04-Jan-2015 00:00:00	19805	AA	N5BSAA	2428	LAX	DFW	855	55.0	1356	58.0	157.0	1235.0
8	13-Jan-2015 00:00:00	19930	AS	N703AS	563	LAX	PDX	855	-5.0	1123	2.0	122.0	834.0
9	17-Jan-2015 00:00:00	19930	AS	N310AS	563	LAX	PDX	855	-5.0	1118	-5.0	120.0	834.0
10	10-Jan-2015 00:00:00	21171	VX	N526VA	474	LAX	LAS	855	-5.0	058	-9.0	40.0	238.0
11	13-Jan-2015 00:00:00	21171	VX	N821VA	380	LAX	BOS	855	25.0	1719	29.0	308.0	2611.0
12	25-Jan-2015 00:00:00	21171	VX	N522VA	474	LAX	LAS	855	-5.0	1003	-2.0	46.0	238.0
13	04-Jan-2015 00:00:00	19933	WN	N369SW	179	LAX	MCI	855	0.0	1408	8.0	179.0	1383.0
14	05-Jan-2015 00:00:00	19933	WN	N968WN	3049	LAX	LAS	855	0.0	053	-7.0	41.0	238.0
15	05-Jan-2015 00:00:00	19933	WN	N703SW	530	LAX	SF	855	0.0	1005	-10.0	59.0	373.0
16	12-Jan-2015 00:00:00	19933	WN	N875SW	45	LAX	SLC	855	5.0	1143	3.0	93.0	560.0
17	25-Jan-2015 00:00:00	19933	WN	N752SW	1274	LAX	ATL	855	-5.0	1806	-9.0	235.0	1947.0
18	28-Jan-2015 00:00:00	19933	WN	N317WN	4238	LAX	OAK	855	-5.0	1001	-19.0	53.0	337.0
19	30-Jan-2015 00:00:00	19933	WN	N983SW	45	LAX	SLC	855	5.0	1130	-10.0	79.0	560.0
20	22-Jan-2015 00:00:00	20304	OO	N805SK	4793	LAX	PHX	855	-10.0	1121	-14.0	60.0	370.0
21	09-Jan-2015 00:00:00	20304	OO	N568SW	6323	LAX	SAN	855	-9.0	1003	-4.0	34.0	109.0
22	26-Jan-2015 00:00:00	19977	UA	N490UA	273	LAX	SFO	855	-5.0	1009	-21.0	47.0	337.0

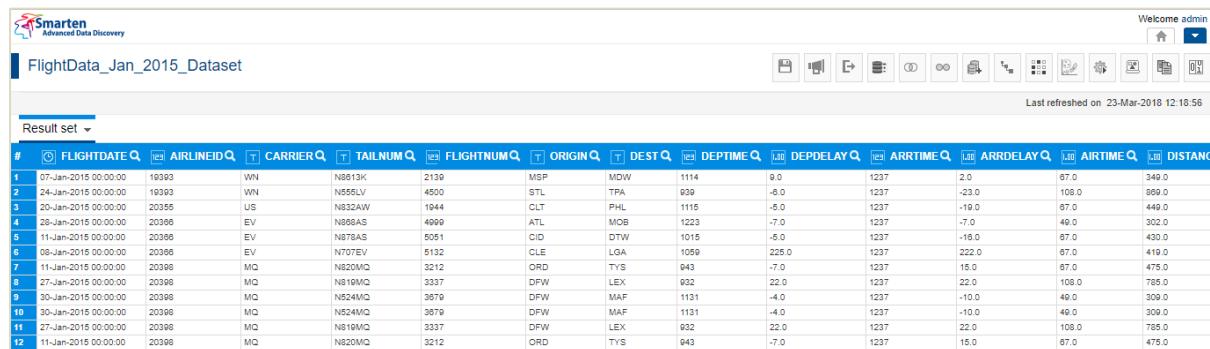
HIGHLIGHT DUPLICATE COLUMN VALUES

6.1.7 Highlight All duplicate rows

This function helps users find all duplicate rows within the Dataset. It highlights sets of rows that contain exactly the same data by highlighting them in the same color. Each set of duplicate rows is highlighted by a different color to make it easier for users to identify.

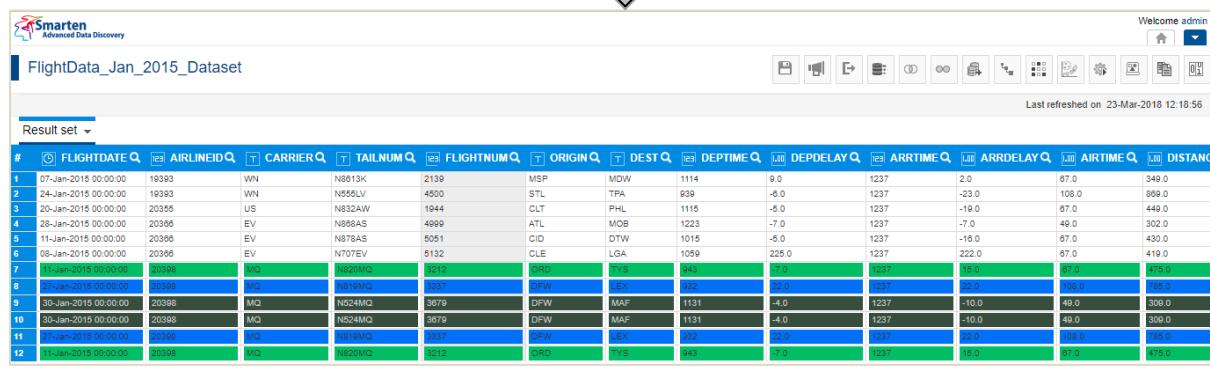
Shown below is the before and after scenario of “Highlight All duplicate rows”:

Before:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	07-Jan-2015 00:00:00	19393	WN	N8613K	2139	MSP	MDW	1114	9.0	1237	2.0	87.0	349.0
2	24-Jan-2015 00:00:00	19393	WN	N556LV	4500	STL	TPA	039	-8.0	1237	-23.0	108.0	869.0
3	20-Jan-2015 00:00:00	20355	US	N832AW	1944	CLT	PHL	1115	-5.0	1237	-19.0	87.0	449.0
4	28-Jan-2015 00:00:00	20366	EV	N868AS	4999	ATL	MOB	1223	-7.0	1237	-7.0	49.0	302.0
5	11-Jan-2015 00:00:00	20366	EV	N878AS	5051	CID	DTW	1015	-5.0	1237	-16.0	87.0	430.0
6	08-Jan-2015 00:00:00	20366	EV	N707EV	5132	CLE	LGA	1059	225.0	1237	222.0	87.0	419.0
7	11-Jan-2015 00:00:00	20368	MQ	N820MQ	3212	ORD	TYS	043	-7.0	1237	15.0	87.0	475.0
8	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	032	22.0	1237	22.0	108.0	785.0
9	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3879	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
10	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3879	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
11	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	032	22.0	1237	22.0	108.0	785.0
12	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	043	-7.0	1237	15.0	87.0	475.0

After:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	07-Jan-2015 00:00:00	19393	WN	N8613K	2139	MSP	MDW	1114	9.0	1237	2.0	87.0	349.0
2	24-Jan-2015 00:00:00	19393	WN	N556LV	4500	STL	TPA	039	-8.0	1237	-23.0	108.0	869.0
3	20-Jan-2015 00:00:00	20355	US	N832AW	1944	CLT	PHL	1115	-5.0	1237	-19.0	87.0	449.0
4	28-Jan-2015 00:00:00	20366	EV	N868AS	4999	ATL	MOB	1223	-7.0	1237	-7.0	49.0	302.0
5	11-Jan-2015 00:00:00	20366	EV	N878AS	5051	CID	DTW	1015	-5.0	1237	-16.0	87.0	430.0
6	08-Jan-2015 00:00:00	20366	EV	N707EV	5132	CLE	LGA	1059	225.0	1237	222.0	87.0	419.0
7	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	043	-7.0	1237	15.0	87.0	475.0
8	20-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	032	22.0	1237	22.0	108.0	785.0
9	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3879	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
10	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3879	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
11	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	032	22.0	1237	22.0	108.0	785.0
12	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	043	-7.0	1237	15.0	87.0	475.0

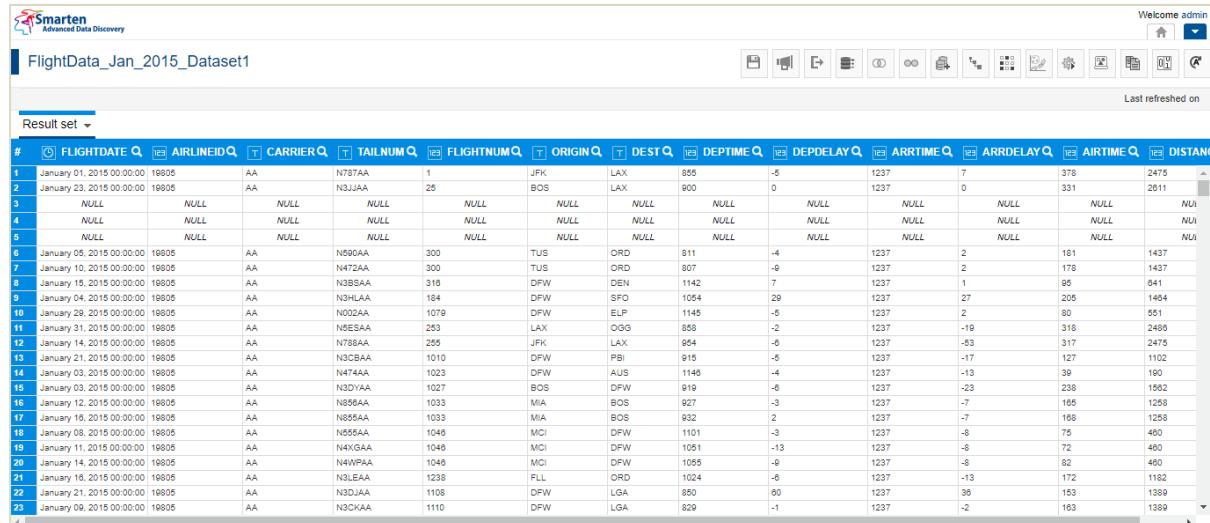
HIGHLIGHT ALL DUPLICATE ROWS

6.1.8 Highlight Rows with all null

This function helps users find rows that have null in all the columns. It highlights all such rows that have null in all the columns.

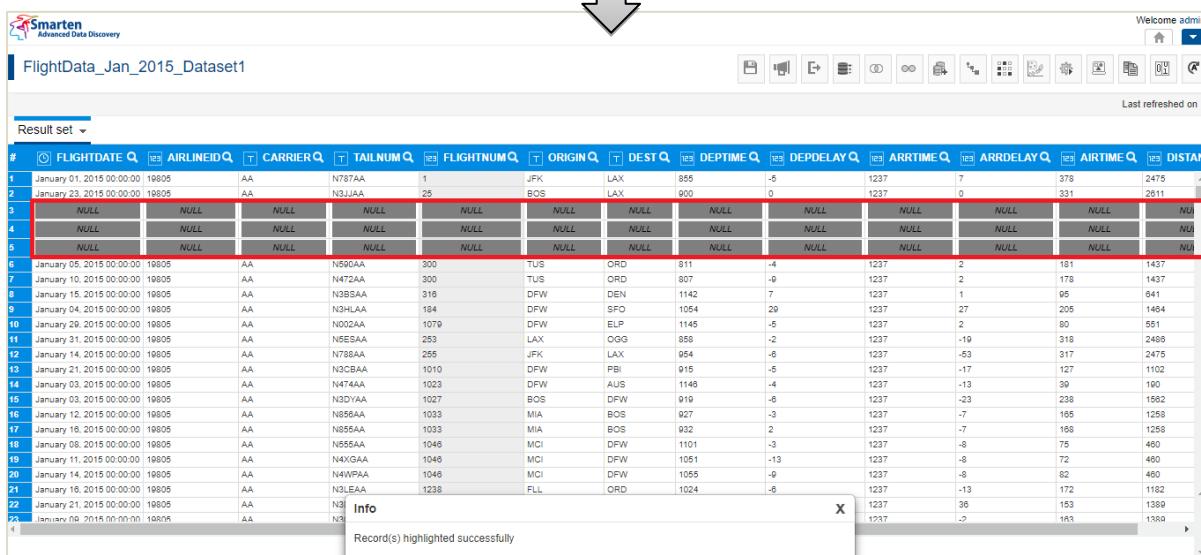
Below is the before and after scenario of “Highlight Rows with all null”:

Before:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	January 01, 2015 00:00:00	19805	AA	N878AA	1	JFK	LAX	855	-5	1237	7	378	2475
2	January 23, 2015 00:00:00	19805	AA	N3JUAA	25	BOS	LAX	900	0	1237	0	331	2611
3	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NU1
4	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NU1
5	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NU1
6	January 31, 2015 00:00:00	19805	AA	N908AA	300	TUS	ORD	811	-4	1237	2	181	1437
7	January 10, 2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9	1237	2	178	1437
8	January 15, 2015 00:00:00	19805	AA	N9B8AA	318	DFW	DEN	1142	7	1237	1	95	641
9	January 04, 2015 00:00:00	19805	AA	N9HLLAA	184	DFW	SFO	1054	29	1237	27	205	1404
10	January 28, 2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145	-5	1237	2	80	551
11	January 31, 2015 00:00:00	19805	AA	N9E5AA	253	LAX	OOG	858	-2	1237	-19	318	2485
12	January 14, 2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-6	1237	-53	317	2475
13	January 21, 2015 00:00:00	19805	AA	N3CBAAA	1010	DFW	PBI	915	-5	1237	-17	127	1102
14	January 03, 2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1145	-4	1237	-13	39	190
15	January 03, 2015 00:00:00	19805	AA	N2D9AA	1027	BOS	DFW	919	-6	1237	-23	238	1562
16	January 12, 2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3	1237	-7	185	1258
17	January 16, 2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2	1237	-7	188	1258
18	January 08, 2015 00:00:00	19805	AA	N855AA	1045	MCI	DFW	1101	-3	1237	-8	75	460
19	January 11, 2015 00:00:00	19805	AA	N4XGAA	1045	MCI	DFW	1051	-13	1237	-8	72	460
20	January 14, 2015 00:00:00	19805	AA	N4WPAA	1045	MCI	DFW	1055	-9	1237	-8	82	460
21	January 16, 2015 00:00:00	19805	AA	N2LEAA	1238	FLL	ORD	1024	-6	1237	-13	172	1182
22	January 21, 2015 00:00:00	19805	AA	N3DJA	1108	DFW	LGA	850	60	1237	38	153	1389
23	January 09, 2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1	1237	-2	163	1389

After:



FlightData_Jan_2015_Dataset1

Last refreshed on

Result set

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	January 01, 2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475
2	January 23, 2015 00:00:00	19805	AA	N3JUAA	25	BOS	LAX	900	0	1237	0	331	2611
3	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
4	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
6	January 05, 2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4	1237	2	181	1437
7	January 10, 2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9	1237	2	178	1437
8	January 15, 2015 00:00:00	19805	AA	N2BSAA	316	DFW	DEN	1142	7	1237	1	95	641
9	January 04, 2015 00:00:00	19805	AA	N2HLAA	184	DFW	SFO	1054	29	1237	27	205	1484
10	January 29, 2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145	-5	1237	2	80	551
11	January 31, 2015 00:00:00	19805	AA	N5E5AA	253	LAX	OOG	858	-2	1237	-19	318	2488
12	January 14, 2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-6	1237	-53	317	2475
13	January 21, 2015 00:00:00	19805	AA	N2C8AA	1010	DFW	PBI	915	-5	1237	-17	127	1102
14	January 03, 2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1148	-4	1237	-13	39	190
15	January 03, 2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919	-6	1237	-23	238	1562
16	January 12, 2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3	1237	-7	165	1258
17	January 16, 2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2	1237	-7	188	1258
18	January 08, 2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3	1237	-8	75	460
19	January 11, 2015 00:00:00	19805	AA	N4XGAA	1046	MCI	DFW	1051	-13	1237	-8	72	460
20	January 14, 2015 00:00:00	19805	AA	N4WPAA	1046	MCI	DFW	1055	-9	1237	-9	82	460
21	January 16, 2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6	1237	-13	172	1182
22	January 21, 2015 00:00:00	19805	AA	N3	Info					1237	36	153	1389
23	January 08, 2015 00:00:00	19805	AA	N3						1237	-7	183	1389

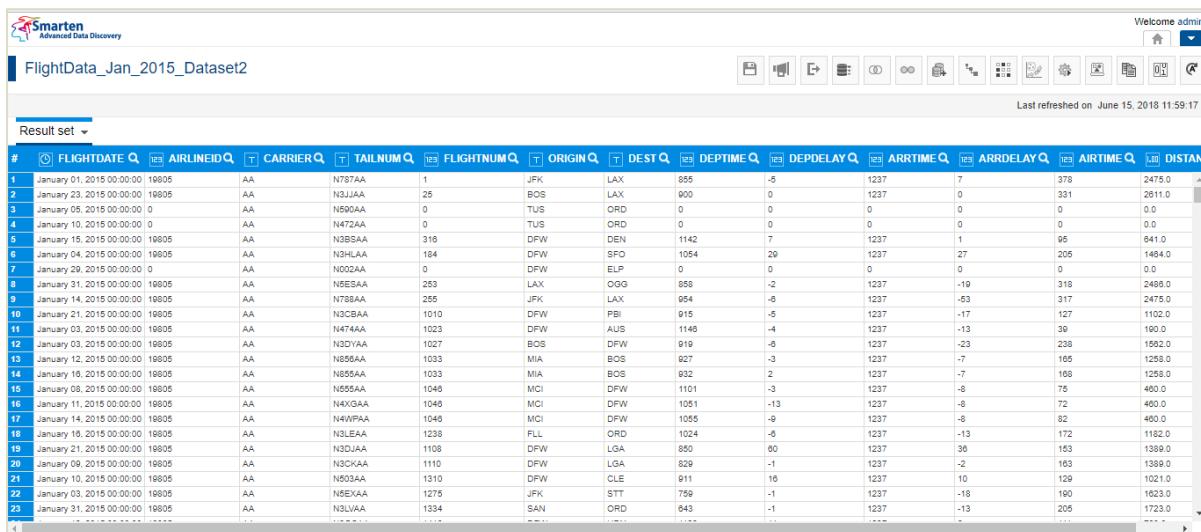
HIGHLIGHT ROWS WITH ALL NULL

6.1.9 Highlight Rows with all zeros

This function helps users find rows that have zeros in all numeric columns. It highlights all such rows that have zeros in all numeric columns.

Shown below is the before and after scenario of “Highlight Rows with all zeros”:

Before:



FlightData_Jan_2015_Dataset2

Last refreshed on June 15, 2018 11:59:17

Result set

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	January 01, 2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475
2	January 23, 2015 00:00:00	19805	AA	N3JUAA	25	BOS	LAX	900	0	1237	0	331	2611.0
3	January 05, 2015 00:00:00	0	AA	N590AA	0	TUS	ORD	0	0	0	0	0	0.0
4	January 10, 2015 00:00:00	0	AA	N472AA	0	TUS	ORD	0	0	0	0	0	0.0
5	January 15, 2015 00:00:00	19805	AA	N2BSAA	316	DFW	DEN	1142	7	1237	1	95	641.0
6	January 04, 2015 00:00:00	19805	AA	N2HLAA	184	DFW	SFO	1054	29	1237	27	205	1484.0
7	January 29, 2015 00:00:00	0	AA	N002AA	0	DFW	ELP	0	0	0	0	0	0.0
8	January 31, 2015 00:00:00	19805	AA	N5E5AA	253	LAX	OOG	858	-2	1237	-19	318	2488.0
9	January 14, 2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-6	1237	-53	317	2475.0
10	January 21, 2015 00:00:00	19805	AA	N2C8AA	1010	DFW	PBI	915	-5	1237	-17	127	1102.0
11	January 03, 2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1148	-4	1237	-13	39	190.0
12	January 03, 2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919	-6	1237	-23	238	1562.0
13	January 12, 2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	927	-3	1237	-7	165	1258.0
14	January 18, 2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2	1237	-7	168	1256.0
15	January 08, 2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3	1237	-8	75	460.0
16	January 11, 2015 00:00:00	19805	AA	N4XGAA	1046	MCI	DFW	1051	-13	1237	-8	72	460.0
17	January 14, 2015 00:00:00	19805	AA	N4WPAA	1046	MCI	DFW	1055	-9	1237	-8	82	460.0
18	January 19, 2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6	1237	-13	172	1182.0
19	January 21, 2015 00:00:00	19805	AA	N3DAAA	1108	DFW	LGA	850	60	1237	36	153	1389.0
20	January 08, 2015 00:00:00	19805	AA	N2CKAA	1110	DFW	LGA	829	-1	1237	-2	163	1389.0
21	January 10, 2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16	1237	10	129	1021.0
22	January 03, 2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759	-1	1237	-18	190	1623.0
23	January 31, 2015 00:00:00	19805	AA	N2LVA	1334	SAN	ORD	643	-1	1237	-13	205	1723.0



After:

FlightData_Jan_2015_Dataset2

Last refreshed on June 15, 2018 11:59:17

Result set

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTAN
1	January 01, 2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475.0
2	January 23, 2015 00:00:00	19805	AA	N3JJAA	25	BOS	LAX	900	0	1237	0	331	2811.0
3	January 05, 2015 00:00:00	0	AA	N500AA	0	TUS	ORD	0	0	0	0	0	0.0
4	January 10, 2015 00:00:00	0	AA	N472AA	0	TUS	ORD	0	0	0	0	0	0.0
5	January 15, 2015 00:00:00	19805	AA	N35SAAA	318	DFW	DEN	1142	7	1237	1	95	641.0
6	January 04, 2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	1054	29	1237	27	205	1484.0
7	January 16, 2015 00:00:00	19805	AA	N002AA	0	DFW	ELP	0	0	0	0	0	0.0
8	January 31, 2015 00:00:00	19805	AA	N855AA	253	LAX	Ogg	858	-2	1237	-19	318	2486.0
9	January 14, 2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-5	1237	-53	317	2475.0
10	January 21, 2015 00:00:00	19805	AA	N3CBAAA	1010	DFW	PBI	915	-5	1237	-17	127	1102.0
11	January 03, 2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1148	-4	1237	-13	39	190.0
12	January 03, 2015 00:00:00	19805	AA	N3DYYAA	1027	BOS	DFW	919	-6	1237	-23	238	1582.0
13	January 12, 2015 00:00:00	19805	AA	N858AA	1033	MIA	BOS	927	-3	1237	-7	165	1258.0
14	January 18, 2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2	1237	-7	168	1258.0
15	January 08, 2015 00:00:00	19805	AA	N555AA	1048	MCI	DFW	1101	-3	1237	-8	75	460.0
16	January 11, 2015 00:00:00	19805	AA	N4XGAA	1048	MCI	DFW	1051	-13	1237	-8	72	460.0
17	January 14, 2015 00:00:00	19805	AA	N4WPAA	1048	MCI	DFW	1055	-9	1237	-8	82	460.0
18	January 19, 2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6	1237	-13	172	1182.0
19	January 21, 2015 00:00:00	19805	AA	N2DJAA	1108	DFW	LGA	850	80	1237	38	153	1389.0
20	January 08, 2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1	1237	-2	183	1389.0
21	January 10, 2015 00:00:00	19805	AA	N003AA	1310	DFW	CLE	911	16	1237	10	129	1021.0
22	January 03, 2015 00:00:00	19805	AA	N	Info					1237	-18	180	1623.0
23	January 31, 2015 00:00:00	19805	AA	N						1237	-13	205	1723.0

Record(s) highlighted successfully

HIGHLIGHT ROWS WITH ALL ZEROS

6.1.10 Highlight Columns with all null

This function helps users find columns that have null in all rows. It highlights all such columns that have null in all rows.

Shown below is the before and after scenario of “Highlight Columns with all null”:

Before:

FlightData_Jan_2015_Dataset3

Last refreshed on May 11, 2018 12:43:59

Result set

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTAN
1	January 01, 2015 00:00:00	NULL	AA	N787AA	1	JFK	LAX	NULL	-5	1237	NULL	378	2475
2	January 23, 2015 00:00:00	NULL	AA	N3JJAA	25	BOS	LAX	NULL	0	1237	NULL	331	2811
3	January 05, 2015 00:00:00	NULL	AA	N500AA	300	TUS	ORD	NULL	-4	1237	NULL	181	1437
4	January 10, 2015 00:00:00	NULL	AA	N472AA	300	TUS	ORD	NULL	-9	1237	NULL	178	1437
5	January 15, 2015 00:00:00	NULL	AA	N2BSAA	318	DFW	DEN	NULL	7	1237	1	95	641
6	January 04, 2015 00:00:00	NULL	AA	N3HLAA	184	DFW	SFO	NULL	29	1237	27	205	1484
7	January 29, 2015 00:00:00	NULL	AA	N002AA	1079	DFW	ELP	NULL	-5	1237	2	80	551
8	January 31, 2015 00:00:00	NULL	AA	N855AA	253	LAX	Ogg	NULL	-2	1237	-19	318	2486
9	January 14, 2015 00:00:00	NULL	AA	N788AA	255	JFK	LAX	NULL	-6	1237	-53	317	2475
10	January 21, 2015 00:00:00	NULL	AA	N3CBAAA	1010	DFW	PBI	NULL	-5	1237	-17	127	1102
11	January 03, 2015 00:00:00	NULL	AA	N474AA	1023	DFW	AUS	NULL	-4	1237	-13	39	190
12	January 03, 2015 00:00:00	NULL	AA	N3DYYAA	1027	BOS	DFW	NULL	-6	1237	-23	238	1582
13	January 12, 2015 00:00:00	NULL	AA	N856AA	1033	MIA	BOS	NULL	-3	1237	-7	165	1258
14	January 18, 2015 00:00:00	NULL	AA	N855AA	1033	MIA	BOS	NULL	2	1237	-7	168	1258
15	January 08, 2015 00:00:00	NULL	AA	N555AA	1048	MCI	DFW	NULL	-3	1237	-8	75	460
16	January 11, 2015 00:00:00	NULL	AA	N4XGAA	1048	MCI	DFW	NULL	-13	1237	-8	72	460
17	January 14, 2015 00:00:00	NULL	AA	N4WPAA	1048	MCI	DFW	NULL	-9	1237	-8	82	460
18	January 19, 2015 00:00:00	NULL	AA	N3LEAA	1238	FLL	ORD	NULL	-6	1237	-18	172	1182
19	January 21, 2015 00:00:00	NULL	AA	N2DJAA	1108	DFW	LGA	NULL	80	1237	38	153	1389
20	January 08, 2015 00:00:00	NULL	AA	N3CKAA	1110	DFW	LGA	NULL	-1	1237	-2	183	1389
21	January 10, 2015 00:00:00	NULL	AA	N003AA	1310	DFW	CLE	NULL	16	1237	10	129	1021



After:

FlightData_Jan_2015_Dataset3

Last refreshed on May 11, 2018 12:43:59

Result set

#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	January 01, 2015 00:00:00	NULL	AA	N787AA	1	JFK	LAX	NULL	-5	1237	NULL	378	2475
2	January 23, 2015 00:00:00	NULL	AA	N3JUAA	25	BOS	LAX	NULL	0	1237	NULL	331	2011
3	January 05, 2015 00:00:00	NULL	AA	N590AA	300	TUS	ORD	NULL	-4	1237	NULL	181	1437
4	January 10, 2015 00:00:00	NULL	AA	N472AA	300	TUS	ORD	NULL	-9	1237	NULL	178	1437
5	January 15, 2015 00:00:00	NULL	AA	N2BSAA	318	DFW	DEN	NULL	7	1237	1	95	641
6	January 04, 2015 00:00:00	NULL	AA	N3HLAA	184	DFW	SFO	NULL	29	1237	27	205	1494
7	January 29, 2015 00:00:00	NULL	AA	N002AA	1079	DFW	ELP	NULL	-5	1237	2	80	551
8	January 31, 2015 00:00:00	NULL	AA	N5E5AA	253	LAX	OGG	NULL	-2	1237	-19	318	2488
9	January 14, 2015 00:00:00	NULL	AA	N788AA	255	JFK	LAX	NULL	-8	1237	-53	317	2475
10	January 21, 2015 00:00:00	NULL	AA	N3CBA	1010	DFW	PBI	NULL	-5	1237	-17	127	1102
11	January 03, 2015 00:00:00	NULL	AA	N474AA	1023	DFW	AUS	NULL	-4	1237	-13	39	190
12	January 03, 2015 00:00:00	NULL	AA	N3DYAA	1027	BOS	DFW	NULL	-6	1237	-23	238	1562
13	January 12, 2015 00:00:00	NULL	AA	N586AA	1033	MIA	BOS	NULL	-3	1237	-7	165	1258
14	January 19, 2015 00:00:00	NULL	AA	N855AA	1033	MIA	BOS	NULL	2	1237	-7	168	1258
15	January 08, 2015 00:00:00	NULL	AA	N555AA	1048	MCI	DFW	NULL	-3	1237	-8	75	460
16	January 11, 2015 00:00:00	NULL	AA	N4XQAA	1048	MCI	DFW	NULL	-13	1237	-8	72	460
17	January 14, 2015 00:00:00	NULL	AA	N4WPAA	1048	MCI	DFW	NULL	-9	1237	-8	82	460
18	January 16, 2015 00:00:00	NULL	AA	N3LAAA	1050	DFW	DFW	1237	-13	172	1182		
19	January 21, 2015 00:00:00	NULL	AA	N3LAAA	1050	DFW	DFW	1237	38	153	1389		

HIGHLIGHT COLUMNS WITH ALL NULL

6.1.11 Highlight Columns with all zeros

This function helps users find numeric columns that have zeros in all rows. It highlights all such numeric columns that have zero in all rows.

Shown below is the before and after scenario of “Highlight Columns with all zeros”:

Before:

FlightData_Jan_2015_Dataset4

Last refreshed on June 15, 2018 12:32:06

Result set

#	FLIGHTDATE	AIRLINEID	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	January 01, 2015 00:00:00	0	AA	1	N787AA	JFK	LAX	0	-5	1237	7	378	2475
2	January 23, 2015 00:00:00	0	AA	25	N3JUAA	BOS	LAX	0	0	1237	0	331	2011
3	January 05, 2015 00:00:00	0	AA	300	N590AA	TUS	ORD	0	-4	1237	2	181	1437
4	January 10, 2015 00:00:00	0	AA	300	N472AA	TUS	ORD	0	-9	1237	2	178	1437
5	January 15, 2015 00:00:00	0	AA	318	N2BSAA	DFW	DEN	0	7	1237	1	95	641
6	January 04, 2015 00:00:00	0	AA	184	N3HLAA	DFW	SFO	0	29	1237	27	205	1494
7	January 29, 2015 00:00:00	0	AA	1079	N002AA	DFW	ELP	0	-5	1237	2	80	551
8	January 31, 2015 00:00:00	0	AA	253	N5E5AA	LAX	OGG	0	-2	1237	-19	318	2488
9	January 14, 2015 00:00:00	0	AA	255	N788AA	JFK	LAX	0	-6	1237	-53	317	2475
10	January 21, 2015 00:00:00	0	AA	1010	N3CBA	DFW	PBI	0	-5	1237	-17	127	1102
11	January 03, 2015 00:00:00	0	AA	1023	N474AA	DFW	AUS	0	-4	1237	-13	39	190
12	January 03, 2015 00:00:00	0	AA	1027	N3DYAA	BOS	DFW	0	-6	1237	-23	238	1562
13	January 12, 2015 00:00:00	0	AA	1033	N855AA	MIA	BOS	0	-3	1237	-7	165	1258
14	January 18, 2015 00:00:00	0	AA	1033	N855AA	MIA	BOS	0	2	1237	-7	168	1258
15	January 08, 2015 00:00:00	0	AA	1048	N555AA	MCI	DFW	0	-3	1237	-8	75	460
16	January 11, 2015 00:00:00	0	AA	1048	N4XQAA	MCI	DFW	0	-13	1237	-8	72	460
17	January 14, 2015 00:00:00	0	AA	1048	N4WPAA	MCI	DFW	0	-9	1237	-8	82	460
18	January 16, 2015 00:00:00	0	AA	1238	N3LEAA	FLL	ORD	0	-6	1237	-13	172	1182
19	January 21, 2015 00:00:00	0	AA	1108	N2DQAA	DFW	LGA	0	80	1237	38	153	1389
20	January 08, 2015 00:00:00	0	AA	1110	N3CKAA	DFW	LGA	0	-1	1237	-2	163	1389
21	January 10, 2015 00:00:00	0	AA	1310	N503AA	DFW	CLE	0	18	1237	10	129	1021
22	January 03, 2015 00:00:00	0	AA	1275	N5EXAA	JFK	STT	0	-1	1237	-18	190	1623
23	January 31, 2015 00:00:00	0	AA	1334	N3LVAA	SAN	ORD	0	-1	1237	-13	205	1723



After:

The screenshot shows the Smarten Advanced Data Discovery interface. The title bar says "FlightData_Jan_2015_Dataset4". The main area displays a table of flight data with 21 rows and 17 columns. The columns are labeled: #, FLIGHTDATE, AIRLINEID, CARRIER, FLIGHTNUM, TAILNUM, ORIGIN, DEST, DEPETIME, DEPDELAY, ARRTIME, ARRDELAY, AIRTIME, DISTANCE. Rows 1 through 20 show various flight details, while row 21 is a summary row. A tooltip "Column(s) highlighted successfully" appears over the last column. The interface includes a toolbar at the top with various icons and a message bar at the bottom right.

HIGHLIGHT COLUMNS WITH ALL ZEROS

6.2 Unique Values

This function helps users to know the number of times a value is repeated in a particular column. It also allows them to delete and edit the value. Deleting a value deletes all rows containing that particular value, and editing allows replacing the value with some other value.

The screenshot shows the Smarten interface with the title "FlightData_Jul-Dec_2016_Dataset_8-218". A modal window titled "Unique values - ORIGIN" is open over the main table. The modal contains a search bar, a page navigation bar (Page 1 of 7), and a list of unique origin cities with their counts. The list includes: ATL (194014), ORD (124715), DEN (116190), LAX (110065), DFW (97648), SFO (88691), PHX (78605), LAS (76027), SEA (69060), IAH (67797), MSP (66380), DTW (62377), BOS (62037), MCO (61005), EWR (60092), CLT (55632), SLC (55328), RMI (49717). At the bottom of the modal, it says "311 Unique values from 2840195 rows". The main table below the modal shows flight data with 24 rows and 17 columns, including columns for FL_DATE, CARRIER, FL_NUM, ORIGIN, DEST, DEP_TIME, DEP_DELAY, and others.

UNIQUE VALUES

6.3 Filter

Users can filter data for a particular column value, duplicate rows, or rows with all null and zero values.

6.3.1 Filter Rows with this column value

This function allows users find all rows containing a particular value in a column. It returns only those rows that have the current column value in the target column.

Note:

This is a back-end data filter operation.

Shown below is the before and after scenario of “Filter Rows with this column value” in column “ORIGIN”:

Before:

#	FLIGHDATE	AIRLINEID	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTAN
1	January 01, 2015 00:00:00	19805	AA	1	N787AA	JFK	LAX	-5.00	1237	7.00	378.00	2475.00	
2	January 23, 2015 00:00:00	19805	AA	25	N3JAA	BOS	Highlight	0.00	1237	0.00	331.00	2811.00	
3	January 05, 2015 00:00:00	19805	AA	300	N590AA	TUS	Unique values	-4.00	1237	2.00	181.00	1437.00	
4	January 10, 2015 00:00:00	19805	AA	300	N472AA	TUS	Cluster & edit	-6.00	1237	2.00	178.00	1437.00	
5	January 15, 2015 00:00:00	19805	AA	315	N383AA	DFW	Find & replace	2.00	1237	1.00	95.00	641.00	
6	January 04, 2015 00:00:00	19805	AA	184	N3H1AA	DFW	Remove	-5.00	1237	27.00	205.00	1464.00	
7	January 28, 2015 00:00:00	19805	AA	1079	N002AA	DFW	Mark as	-2.00	1237	2.00	80.00	551.00	
8	January 31, 2015 00:00:00	19805	AA	253	N56SA	LAX	Copy	-6.00	1237	-19.00	318.00	2484.00	
9	January 14, 2015 00:00:00	19805	AA	255	N783AA	JFK	Sort	-5.00	1237	-53.00	317.00	2474.00	
10	January 21, 2015 00:00:00	19805	AA	1010	N3CBA	DFW	Transform	-4.00	1237	-17.00	127.00	1102.00	
11	January 03, 2015 00:00:00	19805	AA	1023	N474AA	DPW	Add column	-8.00	1237	-13.00	39.00	190.00	
12	January 03, 2015 00:00:00	19805	AA	1027	N3D1AA	BOS	Split	-3.00	1237	-23.00	238.00	1562.00	
13	January 12, 2015 00:00:00	19805	AA	1033	N896AA	MIA	Merge columns	2.00	1237	-7.00	165.00	1258.00	
14	January 18, 2015 00:00:00	19805	AA	1033	N855AA	MIA	Filter	-3.00	1237	-7.00	168.00	1258.00	
15	January 08, 2015 00:00:00	19805	AA	1048	N565AA	MCI	Duplicate rows with this row	-8.00	1237	-8.00	75.00	460.00	
16	January 11, 2015 00:00:00	19805	AA	1048	N4XGAA	MCI	All duplicate rows	-8.00	1237	-8.00	72.00	460.00	
17	January 14, 2015 00:00:00	19805	AA	1048	N4VPA	MCI	Rows with all null	-13.00	1237	-8.00	82.00	460.00	
18	January 18, 2015 00:00:00	19805	AA	1238	N3LEA	FLL	Rows with all zeros	-13.00	1237	-13.00	172.00	1182.00	
19	January 21, 2015 00:00:00	19805	AA	1108	N3D2AA	DFW	Rows with all zeros	-2.00	1237	36.00	153.00	1388.00	
20	January 08, 2015 00:00:00	19805	AA	1110	N3C2AA	DFW	Custom	-2.00	1237	10.00	129.00	1021.00	
21	January 10, 2015 00:00:00	19805	AA	1310	N503AA	DFW	Rows with all zeros	-18.00	1237	-18.00	190.00	1823.00	
22	January 03, 2015 00:00:00	19805	AA	1275	N5EXAA	JFK	Rows with all zeros	-13.00	1237	-13.00	205.00	1723.00	
23	January 31, 2015 00:00:00	19805	AA	1334	N3LVA	SAN	Rows with all zeros	-13.00	1237	-13.00	205.00	1723.00	



After:

#	FLIGHDATE	AIRLINEID	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTAN
1	January 01, 2015 00:00:00	19805	AA	1	N787AA	JFK	LAX	855	-5.00	1237	7.00	378.00	2475.00
2	January 14, 2015 00:00:00	19805	AA	255	N788AA	JFK	LAX	954	-6.00	1237	-53.00	317.00	2475.00
3	January 03, 2015 00:00:00	19805	AA	1275	N5EXAA	JFK	STT	759	-1.00	1237	-18.00	190.00	1823.00
4	January 17, 2015 00:00:00	20409	B6	223	N4R2JB	JFK	LAX	931	-8.00	1237	-22.00	338.00	2475.00
5	January 24, 2015 00:00:00	20409	B6	411	N834JB	JFK	LAS	953	-4.00	1237	-11.00	263.00	2246.00
6	January 28, 2015 00:00:00	21171	VX	11	N849VA	JFK	SFO	907	97.00	1237	82.00	334.00	2560.00
7	January 28, 2015 00:00:00	19877	UA	541	N568UA	JFK	SFO	930	0.00	1237	-24.00	347.00	2598.00
8	January 17, 2015 00:00:00	20355	US	887	N157UW	JFK	CUL	1044	29.00	1237	18.00	84.00	541.00
9	January 24, 2015 00:00:00	19790	DL	2348	N822DL	MCO	819	4.00	1237	77.00	107.00	644.00	

FILTER ROWS WITH THIS COLUMN VALUE

6.3.2 Filter Duplicate rows with this row

This function helps users find duplicates of a row. It returns only duplicate rows of the selected row, including the selected row. A duplicate row is a row that contains exactly the same data for all columns as that of the selected row.

Note:

This is a back-end data filter operation.

Shown below is the before and after scenario of “Filter Duplicate rows with this row”:

Before:

FlightData_Jan_2015_Dataset

Last refreshed on

#	FLIGHDATE	AIRLINEID	ROW_NUMBER	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY
1	January 01, 2015 00:00:00	19805	0	AA	1	N787AA	JFK	LAX	855	-5:00	1237	7:00
2	January 23, 2015 00:00:00	19805	1	AA	25	N3JAA	BOS	LAX	900	0:00	1237	0:00
3	January 05, 2015 00:00:00	19805	2	AA	300	N90AA	TUS	ORD	811	-4:00	1237	2:00
4	January 10, 2015 00:00:00	19805	3	AA	300	N47AA	TUS	ORD	807	-9:00	1237	2:00
5	January 15, 2015 00:00:00	19805	4	AA	318	N3BSAA	DFW	DEN	1142	7:00	1237	1:00
6	January 04, 2015 00:00:00	19805	5	AA	184	N3HAA	DFW	SFO	1054	29:00	1237	27:00
7	January 28, 2015 00:00:00	19805	6	AA	1079	N002AA	DFW	ELP	1145	-5:00	1237	2:00
8	January 31, 2015 00:00:00	19805	7	AA	253	N3EAA	LAX	OGG	858	-2:00	1237	-19:00
9	January 14, 2015 00:00:00	19805	8	AA	255	N788AA	JFK	LAX	054	-6:00	1237	-53:00
10	January 21, 2015 00:00:00	19805	9	AA	1010	N3CBA	DFW	PBI	915	-5:00	1237	-17:00
11	January 03, 2015 00:00:00	19805	10	AA	1023	N47AA	DFW	AUS	1148	-4:00	1237	-13:00
12	January 03, 2015 00:00:00	19805	11	AA	1027	N3DAA	BOS	DFW	919	-6:00	1237	-23:00
13	January 12, 2015 00:00:00	19805	12	AA	1033	N856AA	MIA	BOS	927	-3:00	1237	-7:00
14	January 18, 2015 00:00:00	19805	13	AA	1033	N855AA	MIA	BOS	932	2:00	1237	-7:00
15	January 08, 2015 00:00:00	19805	14	AA	1048	N556AA	MCI	DFW	1101	-3:00	1237	-8:00
16	January 11, 2015 00:00:00	19805	15	AA	1048	N4KGA	MCI	DFW	1051	-13:00	1237	-8:00
17	January 14, 2015 00:00:00	19805	16	AA	1048	N4WPAA	MCI	DFW	1055	-9:00	1237	-8:00
18	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-8:00	1237	-13:00
19	January 21, 2015 00:00:00	19805	18	AA	1108	N3DAA	DFW	LGA	850	60:00	1237	38:00
20	January 08, 2015 00:00:00	19805	19	AA	1110	N3CKAA	DFW	LGA	829	-1:00	1237	-2:00
21	January 10, 2015 00:00:00	19805	20	AA	1310	N503AA	DFW	CLE	911	16:00	1237	10:00
22	January 03, 2015 00:00:00	19805	21	AA	1275	N5EXAA	JFK	STT	759	-1:00	1237	-18:00
23	January 31, 2015 00:00:00	19805	22	AA	1334	N3LVA	SAN	ORD	643	-1:00	1237	-13:00

After:

FlightData_Jan_2015_Dataset

Last refreshed on

#	FLIGHDATE	AIRLINEID	ROW_NUMBER	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY
1	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-8:00	1237	-13:00
2	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-8:00	1237	-13:00
3	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-8:00	1237	-13:00
4	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-8:00	1237	-13:00

FILTER DUPLICATE ROWS WITH THIS ROW

6.3.3 Filter All duplicate rows

This function helps users find all duplicate rows within the Dataset. It returns sets of rows that contain exactly the same data in all columns in the Dataset.

Note:

This is a back-end data filter operation.

Shown below is the before and after scenario of “FilterAll duplicate rows”:

Before:

FlightData_Jan_2015_Dataset

Last refreshed on

#	FLIGHDATE	AIRLINEID	ROW_NUMBER	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY
1	January 01, 2015 00:00:00	19805	0	AA	1	N787AA	JFK	LAX	855	-5:00	1237	7:00
2	January 23, 2015 00:00:00	19805	1	AA	25	N3JAA	BOS	LAX	900	0:00	1237	0:00
3	January 05, 2015 00:00:00	19805	2	AA	300	N90AA	TUS	ORD	811	-4:00	1237	2:00
4	January 10, 2015 00:00:00	19805	3	AA	300	N47AA	TUS	ORD	807	-9:00	1237	2:00
5	January 15, 2015 00:00:00	19805	4	AA	318	N3BSAA	DFW	DEN	1142	7:00	1237	1:00
6	January 04, 2015 00:00:00	19805	5	AA	184	N3HAA	DFW	SFO	1054	29:00	1237	27:00
7	January 28, 2015 00:00:00	19805	6	AA	1079	N002AA	DFW	ELP	1145	-5:00	1237	2:00
8	January 31, 2015 00:00:00	19805	7	AA	253	N3EAA	LAX	OGG	858	-2:00	1237	-19:00
9	January 14, 2015 00:00:00	19805	8	AA	255	N788AA	JFK	LAX	054	-6:00	1237	-53:00
10	January 21, 2015 00:00:00	19805	9	AA	1010	N3CBA	DFW	PBI	915	-5:00	1237	-17:00
11	January 03, 2015 00:00:00	19805	10	AA	1023	N47AA	DFW	AUS	1148	-4:00	1237	-13:00
12	January 03, 2015 00:00:00	19805	11	AA	1027	N3DAA	BOS	DFW	919	-5:00	1237	-23:00
13	January 12, 2015 00:00:00	19805	12	AA	1033	N856AA	MIA	BOS	927	-3:00	1237	-7:00
14	January 18, 2015 00:00:00	19805	13	AA	1033	N855AA	MIA	BOS	932	2:00	1237	-7:00
15	January 08, 2015 00:00:00	19805	14	AA	1048	N4KGA	MCI	DFW	1101	-3:00	1237	-8:00
16	January 11, 2015 00:00:00	19805	15	AA	1048	N4WPAA	MCI	DFW	1051	-13:00	1237	-8:00
17	January 14, 2015 00:00:00	19805	16	AA	1048	N3LEAA	FLL	ORD	1024	-8:00	1237	-13:00
18	January 18, 2015 00:00:00	19805	17	AA	1238	N3DAA	DFW	LGA	850	60:00	1237	38:00
19	January 21, 2015 00:00:00	19805	18	AA	1108	N3CKAA	DFW	LGA	829	-1:00	1237	-2:00
20	January 08, 2015 00:00:00	19805	19	AA	1110	N503AA	DFW	CLE	911	16:00	1237	10:00
21	January 10, 2015 00:00:00	19805	20	AA	1310	N5EXAA	JFK	STT	759	-1:00	1237	-18:00
22	January 03, 2015 00:00:00	19805	21	AA	1275	N3LVA	SAN	ORD	643	-1:00	1237	-13:00
23	January 31, 2015 00:00:00	19805	22	AA	1334



After:

#	FLIGHTDATE	AIRLINEID	ROW_NUMBER	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	January 23, 2015 00:00:00	19805	1	AA	25	N3JJA	BOS	LAX	900	0.0	1237	0.0	331.0	
2	January 23, 2015 00:00:00	19805	1	AA	25	N3JJA	BOS	LAX	900	0.0	1237	0.0	331.0	
3	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-8.0	1237	-13.0	172.0	
4	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-8.0	1237	-13.0	172.0	
5	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-8.0	1237	-13.0	172.0	
6	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-8.0	1237	-13.0	172.0	
7	January 09, 2015 00:00:00	20398	450	MQ	3318	N876MQ	DFW	LFT	1121	16.0	1237	18.0	57.0	
8	January 09, 2015 00:00:00	20398	450	MQ	3318	N876MQ	DFW	LFT	1121	16.0	1237	18.0	57.0	

FILTER ALL DUPLICATE ROWS

6.3.4 Filter Rows with all null

This function helps users find rows that have null in all columns. It returns all such rows that have null in all columns.

Note:

This is a back-end data filter operation.

Shown below is the before and after scenario of “Filter Rows with all null”:

Before:

#	FLIGHTDATE	AIRLINEID	ROW_NUMBER	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	January 01, 2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475	
2	January 23, 2015 00:00:00	19805	AA	N3JJA	25	BOS	LAX	900	0	1237	0	331.0	2611	
3	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
4	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
6	January 05, 2015 00:00:00	19805	AA	N980AA	300	TUS	ORD	811	-4	1237	2	181	1437	
7	January 10, 2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9	1237	2	178	1437	
8	January 15, 2015 00:00:00	19805	AA	N2B5AA	316	DFW	DEN	1142	7	1237	1	95	841	
9	January 04, 2015 00:00:00	19805	AA	N3H2AA	184	DFW	SFO	1054	29	1237	27	205	1464	
10	January 28, 2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145	-5	1237	2	80	551	
11	January 31, 2015 00:00:00	19805	AA	N8E5AA	253	LAX	OIG	858	-2	1237	-19	318	2485	
12	January 14, 2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-6	1237	-53	317	2475	
13	January 21, 2015 00:00:00	19805	AA	N3C8AA	1010	DFW	PBI	915	-5	1237	-17	127	1102	
14	January 03, 2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4	1237	-13	39	190	
15	January 03, 2015 00:00:00	19805	AA	N3D9AA	1027	BOS	DFW	919	-6	1237	-23	238	1562	
16	January 12, 2015 00:00:00	19805	AA	N565AA	1033	MIA	BOS	927	-3	1237	-7	165	1258	
17	January 19, 2015 00:00:00	19805	AA	N555AA	1033	MIA	BOS	932	2	1237	-7	168	1258	
18	January 08, 2015 00:00:00	19805	AA	N555AA	1048	MCI	DFW	1101	-3	1237	-8	75	460	
19	January 11, 2015 00:00:00	19805	AA	N4X9AA	1048	MCI	DFW	1051	-13	1237	-8	72	460	
20	January 14, 2015 00:00:00	19805	AA	N4WPAA	1048	MCI	DFW	1055	-9	1237	-8	82	460	
21	January 18, 2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6	1237	-13	172	1182	
22	January 21, 2015 00:00:00	19805	AA	N3DJAA	1108	DFW	LGA	850	60	1237	38	153	1389	
23	January 09, 2015 00:00:00	19805	AA	N2CKAA	1110	DFW	LGA	829	-1	1237	-2	163	1389	



After:

#	FLIGHTDATE	AIRLINEID	ROW_NUMBER	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
2	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	

FILTER ROWS WITH ALL NULL

6.3.5 Filter Rows with all zeros

This function helps users find rows that have zeros in all numeric columns. It returns all such rows that have zeros in all numeric columns.

Note:

This is a back-end data filter operation.

Showed below is the before and after scenario of “Filter Rows with all zeros”:

Before:

#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRELAY	AIRTIME	DISTANCE
1	January 01, 2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475.0
2	January 23, 2015 00:00:00	19805	AA	N8JAA	25	BOS	LAX	900	0	1237	0	331	2611.0
3	January 08, 2015 00:00:00	0	AA	N590AA	0	TUS	ORD	0	0	0	0	0	0.0
4	January 10, 2015 00:00:00	0	AA	N472AA	0	TUS	ORD	0	0	0	0	0	0.0
5	January 15, 2015 00:00:00	19805	AA	N2BSAA	318	DFW	DEN	1142	7	1237	1	95	641.0
6	January 04, 2015 00:00:00	19805	AA	N2HLAA	184	DFW	SFO	1054	29	1237	27	205	1484.0
7	January 29, 2015 00:00:00	0	AA	N02AA	0	DFW	ELP	0	0	0	0	0	0.0
8	January 31, 2015 00:00:00	19805	AA	N5E3AA	253	LAX	OOG	858	-2	1237	-19	318	2488.0
9	January 14, 2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-6	1237	-53	317	2475.0
10	January 21, 2015 00:00:00	19805	AA	N2CBAA	1010	DFW	PBI	915	-5	1237	-17	127	1102.0
11	January 03, 2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1148	-4	1237	-13	39	190.0
12	January 03, 2015 00:00:00	19805	AA	N2DYAA	1027	BOS	DFW	919	-6	1237	-23	238	1582.0
13	January 12, 2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3	1237	-7	165	1259.0
14	January 18, 2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2	1237	-7	168	1259.0
15	January 08, 2015 00:00:00	19805	AA	N555AA	1045	MCI	DFW	1101	-3	1237	-8	75	400.0
16	January 11, 2015 00:00:00	19805	AA	N4X9AA	1045	MCI	DFW	1051	-13	1237	-8	72	400.0
17	January 14, 2015 00:00:00	19805	AA	N4WPAA	1045	MCI	DFW	1055	-9	1237	-8	82	400.0
18	January 18, 2015 00:00:00	19805	AA	N2LEAA	1238	FLL	ORD	1024	-6	1237	-13	172	1182.0
19	January 21, 2015 00:00:00	19805	AA	N2DJA	1108	DFW	LGA	850	60	1237	38	153	1388.0
20	January 08, 2015 00:00:00	19805	AA	N2CKAA	1110	DFW	LGA	829	-1	1237	-2	163	1388.0
21	January 10, 2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16	1237	10	129	1021.0
22	January 03, 2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759	-1	1237	-18	190	1023.0
23	January 31, 2015 00:00:00	19805	AA	N2UAA	1334	SAN	ORD	643	-1	1237	-13	205	1723.0



After:

#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRELAY	AIRTIME	DISTANCE
1	January 05, 2015 00:00:00	0	AA	N590AA	0	TUS	ORD	0	0	0	0	0	0.0
2	January 10, 2015 00:00:00	0	AA	N472AA	0	TUS	ORD	0	0	0	0	0	0.0
3	January 29, 2015 00:00:00	0	AA	N002AA	0	DFW	ELP	0	0	0	0	0	0.0

FILTER ROWS WITH ALL ZEROS

6.3.6 Custom

6.3.6.1 Filter

Data can be filtered through different expressions created using one or more columns of the Dataset. Users can create many different conditions using “AND” and “OR” operators.

Note:

This is a back-end data filter operation.

For example, a user may want to filter data for all flights departing from “JFK” (New York) and landing at “LAX” (Los Angeles).

The screenshot shows the Smarten Advanced Data Discovery interface. On the left, there is a table titled "FlightData_Jan_2015_Dataset" with columns: #, FLIGHDATE, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEPTEIME. The table contains 23 rows of flight data from January 2015. On the right, an "Add Filter" dialog box is open. It has two tabs: "Filter" (selected) and "Advanced Filter". Under "Filter", there is a dropdown for "Column name" set to "Dest" with the value "LAX". Below it are "ADD" and "O₊" buttons. Under "Advanced Filter", there are sections for "Column", "Operator", and "Value". Two filters are applied: "Origin = JFK" and "Dest = LAX", both with the operator "AND". An expression box at the bottom shows "(Origin = JFK AND Dest = LAX)". At the bottom of the dialog are "APPLY" and "CANCEL" buttons.

ADD FILTER

The resultant data will be displayed as follows:

The screenshot shows the Smarten Advanced Data Discovery interface after applying the filter. The table now only contains three rows of data where the destination is LAX. The columns include: #, FLIGHDATE, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEPTEIME, DEPDELAY, ARRTIME, ARRDELAY, AIRTIME, DISTANCE. The rows are: 1. 01-Jan-2015 00:00:00, 19805, AA, N78AA, 1, JFK, LAX, 855, -5.0, 1237, 7.0, 378.0, 2475.0; 2. 14-Jan-2015 00:00:00, 19805, AA, N78AA, 255, JFK, LAX, 954, -6.0, 1237, -53.0, 317.0, 2475.0; 3. 17-Jan-2015 00:00:00, 20408, B6, N64JB, 223, JFK, LAX, 931, -8.0, 1237, -22.0, 338.0, 2475.0. At the bottom of the dialog is a "FILTER" button.

6.3.6.2 Advanced Filter

The advanced filter is a type of filter that can be applied to the columns of a Dataset. Users can create filters based on various string, arithmetic, date, or miscellaneous statements using various arithmetic operators (such as +, -, /) or comparison operators (such as =, >, <).

Note:

This is a back-end data filter operation.

For example, a user may want to filter data for all flights scheduled on the 26th of any month. The expression in this case can be as follows:

`day(FlighDate) == 26`

where `day()` is the Date function and `FlightDate` is the column name.

The screenshot shows the Smarten Advanced Data Discovery interface with an "Add Filter" dialog box. The "Advanced Filter" tab is selected. The "Expression" field contains `day(FlightDate) == 26`. The "Dimension values" section lists columns: FlightDate, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEPTEIME. The "Functions" section includes Date, DateAdd, DateDiff, DatePart, DateTime, Day, DayName, and AirTime. The "Operators" section includes standard arithmetic and comparison operators. At the bottom of the dialog are "APPLY", "VERIFY EXPRESSION", and "CANCEL" buttons.

ADVANCED FILTER

The resultant data will be displayed as follows:

This screenshot shows the Smarten Advanced Data Discovery interface. At the top, there's a navigation bar with icons for Home, Logout, and Help. Below it is a toolbar with various data manipulation and visualization tools. The main area is titled "FlightData_Jan_2015_Dataset". A "Result set" dropdown is open, showing a preview of the data. The data table has 20 columns: #, FLIGHDATE, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEPETIME, DEPDELAY, ARRTIME, ARRDELAY, AIRTIME, and DISTANCE. The preview shows approximately 20 rows of flight information from January 28, 2015. A "FILTER" button is located at the bottom of the preview area.

Note:

Please refer to the following document for more information on different expressions.

Reference: Self-Serve Data Preparation (SSDP) - Concept Manual > Shape Data > Add Column > Custom

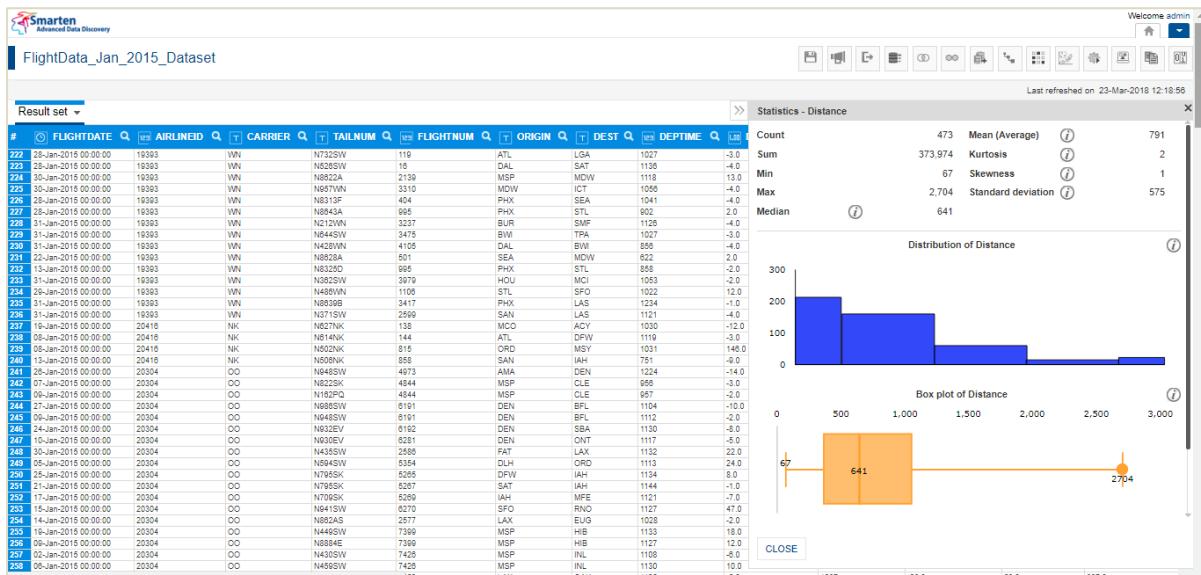
6.4 Statistics

This function allows users to explore data with the help of various statistics, including skewness and kurtosis.

Note:

Statistics is applicable for a Numeric type of data only.

Shown below are the statistics for the column “Distance” of the Dataset:



STATISTICS

Statistics are provided as follows:

Heading	Description
Count	Displays the count of the column
Sum	Displays the sum of the column
Min	Displays the minimum value of the column
Max	Displays the maximum value of the column
Median	Displays the value in the middle when the data items are arranged in ascending order
Mean	Displays the average of all data values of the column
Kurtosis	Displays the measure of the peakedness of the Dataset
Skewness	Displays the measure of symmetry. A Dataset is symmetric if it looks the same to the left and right of the center point
Standard deviation	Displays the measure of how spread out the Dataset is
Distribution	A graphical display where the data is grouped into buckets and then plotted as bars
Box Plot	A standardized way of displaying the distribution of data based on the five-number summary: minimum, first quartile, median, third quartile, and maximum

7 Clean Data

Smarten SSDP allows users to clean the raw data to derive more useful analysis-ready data by using a wide variety of functions.

7.1 Cluster & Edit

This function allows users to edit values in bulk. It creates groups of similar data values and displays them in clusters, allowing users to edit them in bunches.

Note:

This function is applicable for string-type data only.

Shown below is the before and after scenario of “Cluster & edit” for column “ORIGIN”:

Before:

The screenshot shows the Smarten interface with a 'FlightData_Jan_2015_Dataset' view. A 'Cluster & edit - Origin' dialog is open on the right. The dialog has a 'SIZE' column, a 'VALUES' column, and a 'NAME' column. It lists three clusters: 'SEA' (size 4, values: SEA, SEA, SJC, SJU, SGU), 'MCO' (size 4, values: MCO, MCO, MKE, MCI, MSY), and 'BUR' (size 3, values: BUR, BUR). The 'MCO' cluster has three rows highlighted with a red box. At the bottom of the dialog are 'APPLY' and 'CANCEL' buttons.

SIZE	VALUES	NAME
4	SEA SEA SJC SJU SGU	SEA
4	MCO MCO MKE MCI MSY	MCO
3	BUR BUR	BUR

CLUSTER & EDIT

As you can see in the above image, there are 7 rows for flights originating from "MKE" (General Mitchell Airport) and 4 rows for flights originating from "MCI" (Kansas City Airport).

Now, let us assume we want to replace all flights originating from "MKE" with "MCI" as shown below:

The screenshot shows the Smarten Advanced Data Discovery interface. On the left, a table titled "FlightData_Jan_2015_Dataset" displays flight information. On the right, a "Cluster & edit - Origin" dialog is open, showing a list of current origin codes and their counts. A red box highlights the row for "MKE" (count 7), which is being replaced by "MCI".

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTH
1	01-Jan-2015 00:00:00	19805	AA	N187AA	1	JFK	LAX	855
2	23-Jan-2015 00:00:00	19805	AA	N32JAA	25	BOS	LAX	900
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807
5	15-Jan-2015 00:00:00	19805	AA	N2BSAA	318	DFW	DEN	1142
6	04-Jan-2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	1054
7	20-Jan-2015 00:00:00	19805	AA	N02AA	1079	DFW	ELP	1145
8	31-Jan-2015 00:00:00	19805	AA	N5E5AA	253	LAX	OOG	858
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954
10	21-Jan-2015 00:00:00	19805	AA	N3CBAA	1010	DFW	PBI	915
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1145
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101
16	11-Jan-2015 00:00:00	19805	AA	N4XGAA	1046	MCI	DFW	1051
17	14-Jan-2015 00:00:00	19805	AA	N4WPAA	1046	MCI	DFW	1055
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024
19	21-Jan-2015 00:00:00	19805	AA	N3DAAA	1108	DFW	LGA	850
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911
22	03-Jan-2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759
23	31-Jan-2015 00:00:00	19805	AA	N3LVAA	1334	SAN	ORD	643

Cluster & edit - Origin

SIZE	VALUES	NAME
4	SEA	19 SEA
	SEA	11
	SJC	6
	SJU	1
	SQU	1
<input checked="" type="checkbox"/>	4 MCO	23 MCI
	MKE	7

APPLY CANCEL

After:

As a result, all rows having Origin as "MKE" will be replaced with "MCI," and now there will be 11 rows for flights originating from "MCI" as shown below:

The screenshot shows the Smarten interface after the replacement. The "Cluster & edit - Origin" dialog now shows the updated counts for each origin code. A red box highlights the row for "MKE" (count 11), which has been successfully replaced by "MCI".

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTH
1	01-Jan-2015 00:00:00	19805	AA	N187AA	1	JFK	LAX	855
2	23-Jan-2015 00:00:00	19805	AA	N32JAA	25	BOS	LAX	900
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807
5	15-Jan-2015 00:00:00	19805	AA	N2BSAA	318	DFW	DEN	1142
6	04-Jan-2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	1054
7	20-Jan-2015 00:00:00	19805	AA	N02AA	1079	DFW	ELP	1145
8	31-Jan-2015 00:00:00	19805	AA	N5E5AA	253	LAX	OOG	858
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954
10	21-Jan-2015 00:00:00	19805	AA	N3CBAA	1010	DFW	PBI	915
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1145
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101
16	11-Jan-2015 00:00:00	19805	AA	N4XGAA	1046	MCI	DFW	1051
17	14-Jan-2015 00:00:00	19805	AA	N4WPAA	1046	MCI	DFW	1055
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024
19	21-Jan-2015 00:00:00	19805	AA	N3DAAA	1108	DFW	LGA	850
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911
22	03-Jan-2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759
23	31-Jan-2015 00:00:00	19805	AA	N3LVAA	1334	SAN	ORD	643

Cluster & edit - Origin

SIZE	VALUES	NAME
3	LAX	35 LAX
	LAX	15
	LGA	12
	LAS	8
<input checked="" type="checkbox"/>	3 MCI	23 MCI
	MCI	11
	MCO	10
	MSY	2
<input checked="" type="checkbox"/>	2 RIC	3 RIC
	RIC	2
	RSW	1

APPLY CANCEL

7.2 Find & replace

This function allows users to find a particular value in a column and replace it with some other value.

The following options are provided while finding and replacing a value:

Find		
Option	Description	Example (TAILNUM)
Equals	Finds values that have exactly the same value as the entered one	Find="N002AA" Considers all rows that have TAILNUM = "N002AA"
Starts with	Finds values that start with the entered value	Find="N0" Considers all rows that have TAILNUM starting with "N0." For example, "N002AA" and "NODYAA"
Ends with	Finds values that end with the entered value	Find="AA" Considers all rows that have TAILNUM ending with "AA." For example, "N3JJAA" and "N788AA"
Contains	Finds values that contain the entered value	Find="00" Considers all rows in which TAILNUM contains "00" at any place. For example, "N002AA" and "N27400"

Ignore case	
Option	Description
When Selected	Finds the entered value irrespective of its case. For example, while finding "ABCD," data containing "abcd" or "ABcd" is also considered
When Not selected	Finds values that contain the same case as the entered value. For example, while finding "ABCD," data containing "ABCD" will only be considered, and "abcd" or "ABcd" will not be considered

Replace with	
Option	Description
Entire cell value	Replaces the entire value with the new value. For example, "N002AA" will be replaced with "99" when finding values containing "00" and replacing them with "99"
Only matched value	Replaces only the matched value with the new value. For example, "N002AA" will be replaced with "N992AA" when finding values containing "00" and replacing them with "99"

For example, in the image below, all rows having Origin as "MCI" are replaced with "MKE" irrespective of the case.

The screenshot shows the Smarten Advanced Data Discovery interface. On the left, there is a table titled "FlightData_Jan_2015_Dataset" with columns: #, FLIGHDATE, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEPETIME, DEPDELAY, ARRTIME, ARRDELAY, AIRTIME, DISTANCE. The data consists of 23 rows of flight information. On the right, a "Find & replace - Origin" dialog box is open, showing a search field with "MCI", a replace field with "MKE", and options for "Entire cell value" or "Only matched value". Below the dialog are "APPLY" and "CANCEL" buttons.

7.3 Remove

Users can clean data by removing unnecessary and duplicate rows and columns, or rows having null or zero values or having a particular value in a column.

7.3.1 Remove This column

This function allows users to remove a particular column from the Dataset.

Shown below is the before and after scenario of “Remove This column” for column “CARRIER”:

Before:

The screenshot shows the Smarten Advanced Data Discovery interface with the "FlightData_Jan_2015_Dataset" table. The "CARRIER" column is highlighted with a red border. A large gray arrow points downwards towards the bottom of the table, indicating the removal of the "CARRIER" column.

After:

FlightData_Jan_2015_Dataset

Last refreshed on 23-Mar-2018 12:18:56

Result set ▾

#	FLIGHDATE	AIRLINEID	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 00:00:00	19805	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	378.0	2475.0
2	23-Jan-2015 00:00:00	19805	N3JJAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2611.0
3	05-Jan-2015 00:00:00	19805	N900AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.0
4	10-Jan-2015 00:00:00	19805	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.0
5	15-Jan-2015 00:00:00	19805	N3BSAA	318	DFW	DEN	1142	7.0	1237	1.0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	N3HJAA	184	DFW	SFO	1054	29.0	1237	27.0	205.0	1464.0
7	29-Jan-2015 00:00:00	19805	N002AA	1079	DFW	ELP	1145	-5.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	N5E3AA	253	LAX	OGG	858	-2.0	1237	-19.0	318.0	2468.0
9	14-Jan-2015 00:00:00	19805	N788AA	255	JFK	LAX	954	-6.0	1237	-53.0	317.0	2475.0
10	21-Jan-2015 00:00:00	19805	N3CBAAA	1010	DFW	PBI	915	-5.0	1237	-17.0	127.0	1102.0
11	03-Jan-2015 00:00:00	19805	N474AA	1023	DFW	AUS	1146	-4.0	1237	-13.0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	N3DVAA	1027	BOS	DFW	919	-6.0	1237	-23.0	238.0	1562.0
13	12-Jan-2015 00:00:00	19805	N856AA	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1258.0
14	16-Jan-2015 00:00:00	19805	N855AA	1033	MIA	BOS	932	2.0	1237	-7.0	168.0	1258.0
15	08-Jan-2015 00:00:00	19805	N555AA	1048	MCI	DFW	1101	-3.0	1237	-8.0	75.0	460.0
16	11-Jan-2015 00:00:00	19805	N4XGAA	1048	MCI	DFW	1051	-13.0	1237	-8.0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	N4WPAA	1048	MCI	DFW	1055	-6.0	1237	-8.0	82.0	460.0
18	10-Jan-2015 00:00:00	19805	N3LEAA	1238	FLL	ORD	1024	-6.0	1237	-13.0	172.0	1162.0
19	21-Jan-2015 00:00:00	19805	N3DAAA	1108	DFW	LGA	850	60.0	1237	38.0	153.0	1369.0
20	09-Jan-2015 00:00:00	19805	N3CKAA	1110	DFW	LGA	829	-1.0	1237	-2.0	163.0	1369.0
21	10-Jan-2015 00:00:00	19805	N503AA	1310	DFW	CLE	911	16.0	1237	10.0	129.0	1021.0
22	03-Jan-2015 00:00:00	19805	N5EXAA	1275	JFK	STT	759	-1.0	1237	-18.0	190.0	1623.0
23	31-Jan-2015 00:00:00	19805	N3LVAA	1334	SAN	ORD	643	-1.0	1237	-13.0	205.0	1723.0
24	22-Jan-2015 00:00:00	19805	N5FAAA	-----	-----	-----	-----	-----	-----	-----	-----	-----

REMOVE THIS COLUMN

7.3.2 Remove This row

This function allows users to remove a particular row from the Dataset.

Shown below is the before and after scenario of “Remove This row”:

Before:

FlightData_Jan_2015_Dataset

Last refreshed on 23-Mar-2018 12:18:56

Result set ▾

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	378.0	2475.0
2	23-Jan-2015 00:00:00	19805	AA	N3JJAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2611.0
3	05-Jan-2015 00:00:00	19805	AA	N900AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.0
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.0
5	15-Jan-2015 00:00:00	19805	AA	N3BSAA	318	DFW	DEN	1142	7.0	1237	1.0	95.0	541.0
6	04-Jan-2015 00:00:00	19805	AA	N3HJAA	184	DFW	SFO	1054	29.0	1237	2.0	205.0	1464.0
7	20-Jan-2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145	-6.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	N5E3AA	253	LAX	OGG	858	-2.0	1237	-19.0	318.0	2468.0
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-6.0	1237	-53.0	317.0	2475.0
10	21-Jan-2015 00:00:00	19805	AA	N3CBAAA	1010	DFW	PBI	915	-5.0	1237	-17.0	127.0	1102.0
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4.0	1237	-13.0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	AA	N3DVAA	1027	BOS	DFW	919	-6.0	1237	-23.0	238.0	1562.0
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1258.0
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2.0	1237	-7.0	168.0	1258.0
15	08-Jan-2015 00:00:00	19805	AA	N4XGAA	1048	MCI	DFW	1101	-3.0	1237	-8.0	75.0	460.0
16	11-Jan-2015 00:00:00	19805	AA	N4WPAA	1048	MCI	DFW	1051	-13.0	1237	-8.0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6.0	1237	-13.0	172.0	1162.0
18	20-Jan-2015 00:00:00	19805	AA	N3DAAA	1108	DFW	LGA	850	60.0	1237	38.0	153.0	1369.0
19	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1.0	1237	-2.0	163.0	1369.0
20	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16.0	1237	10.0	129.0	1021.0
21	03-Jan-2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759	-1.0	1237	-18.0	190.0	1623.0
22	31-Jan-2015 00:00:00	19805	AA	N3LVAA	1334	SAN	ORD	643	-1.0	1237	-13.0	205.0	1723.0
23	22-Jan-2015 00:00:00	19805	AA	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



After:

#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	378.0	2475.0
2	23-Jan-2015 00:00:00	19805	AA	N3JUAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2011.0
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.0
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.0
5	15-Jan-2015 00:00:00	19805	AA	N3BSAA	316	DFW	DEN	1142	7.0	1237	1.0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	1054	28.0	1237	27.0	205.0	1484.0
7	29-Jan-2015 00:00:00	19805	AA	N002AA	1078	DFW	ELP	1145	-5.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	N5E5AA	253	LAX	OOG	856	-2.0	1237	-18.0	318.0	2486.0
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-5.0	1237	-53.0	317.0	2475.0
10	21-Jan-2015 00:00:00	19805	AA	N3CBAA	1010	DFW	PBI	915	-5.0	1237	-17.0	127.0	1102.0
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1148	-4.0	1237	-13.0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919	-5.0	1237	-23.0	238.0	1562.0
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1258.0
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2.0	1237	-7.0	168.0	1258.0
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1045	MCI	DFW	1101	-3.0	1237	-8.0	75.0	460.0
16	11-Jan-2015 00:00:00	19805	AA	N4XGAA	1045	MCI	DFW	1051	-13.0	1237	-8.0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	AA	N4WPAA	1045	MCI	DFW	1055	-9.0	1237	-8.0	82.0	460.0
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6.0	1237	-13.0	172.0	1182.0
19	21-Jan-2015 00:00:00	19805	AA	N3DAA	1108	DFW	LGA	850	60.0	1237	38.0	153.0	1388.0
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1.0	1237	-2.0	163.0	1389.0
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16.0	1237	10.0	129.0	1021.0
22	03-Jan-2015 00:00:00	19805	AA	N5F1AA	1274	DTT	SFO	1111	-1.0	1237	-18.0	160.0	1623.0
23	31-Jan-2015 00:00:00	19805	AA	N3LInfo	1237	-13.0	205.0	1723.0

7.3.3 Remove Row with this column value

This function allows users to remove all rows that have currently selected value in a particular column.

For example:

Column	Value	Result
TAILNUM	N002AA	Removes all rows where TAILNUM = "N002AA"
FLIGHTNUM	1046	Removes all rows where FLIGHTNUM = "1046"
FLIGHTDATE	03-Jan-2015 00:00:00	Removes all rows where FLIGHTDATE = "03-Jan-2015 00:00:00"

Shown below is the before and after scenario of “Remove Row with this column value” for column “Carrier” that has value = “AA”:

Before:

#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	378.0	2475.0
2	23-Jan-2015 00:00:00	19805	AA	N3JUAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2011.0
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.0
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.0
5	15-Jan-2015 00:00:00	19805	AA	N3BSAA	316	DFW	DEN	1142	7.0	1237	1.0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	1054	28.0	1237	27.0	205.0	1484.0
7	29-Jan-2015 00:00:00	19805	AA	N002AA	1078	DFW	ELP	1145	-5.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	N5E5AA	253	LAX	OOG	856	-2.0	1237	-18.0	318.0	2486.0
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-5.0	1237	-53.0	317.0	2475.0
10	21-Jan-2015 00:00:00	19805	AA	N3CBAA	1010	DFW	PBI	915	-5.0	1237	-17.0	127.0	1102.0
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1148	-4.0	1237	-13.0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919	-5.0	1237	-23.0	238.0	1562.0
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1258.0
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2.0	1237	-7.0	168.0	1258.0
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1045	MCI	DFW	1101	-3.0	1237	-8.0	75.0	460.0
16	11-Jan-2015 00:00:00	19805	AA	N4XGAA	1045	MCI	DFW	1051	-13.0	1237	-8.0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	AA	N4WPAA	1045	MCI	DFW	1055	-9.0	1237	-8.0	82.0	460.0
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6.0	1237	-13.0	172.0	1182.0
19	21-Jan-2015 00:00:00	19805	AA	N3DAA	1108	DFW	LGA	850	60.0	1237	38.0	153.0	1388.0
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1.0	1237	-2.0	163.0	1389.0
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16.0	1237	10.0	129.0	1021.0
22	03-Jan-2015 00:00:00	19805	AA	N5F1AA	1275	DTT	SFO	759	-1.0	1237	-18.0	160.0	1623.0
23	31-Jan-2015 00:00:00	19805	AA	N3LInfo	1334	SAN	ORD	943	-1.0	1237	-13.0	205.0	1723.0

After:

The screenshot shows the Smarten Advanced Data Discovery interface. The title bar says "FlightData_Jan_2015_Dataset". The top right corner shows "Welcome admin". Below the title bar are various icons for file operations like save, open, and search. A message at the top right says "Last refreshed on 23-Mar-2018 12:18:56". The main area is a table titled "Result set" with 23 columns and 23 rows. The columns are labeled: #, FLIGHDATE, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEPETIME, DEPDELAY, ARRTIME, ARRDELAY, AIRTIME, DISTANCE. The first few rows of data are as follows:

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 00:00:00	19930	AS	N520AS	358	SEA	SF	1100	0.0	1237	-8.0	79.0	605.0
2	06-Jan-2015 00:00:00	19930	AS	N562AS	55	BRW	FAI	1119	1.0	1237	1.0	69.0	503.0
3	08-Jan-2015 00:00:00	19930	AS	N797AS	75	SEA	JNU	1119	9.0	1237	-15.0	120.0	909.0
4	08-Jan-2015 00:00:00	19930	AS	N506AS	481	SAN	SEA	058	-9.0	1237	-23.0	139.0	1050.0
5	09-Jan-2015 00:00:00	19930	AS	N516AS	62	SIT	KTN	1148	-10.0	1237	-13.0	37.0	183.0
6	14-Jan-2015 00:00:00	19930	AS	N513AS	662	SEA	DFW	648	-4.0	1237	11.0	205.0	1860.0
7	14-Jan-2015 00:00:00	19930	AS	N564AS	879	SEA	LH	825	-5.0	1237	-38.0	354.0	2701.0
8	13-Jan-2015 00:00:00	19930	AS	N786AS	62	SIT	KTN	1152	-8.0	1237	-11.0	33.0	183.0
9	14-Jan-2015 00:00:00	19930	AS	N792AS	65	PSG	JNU	1149	-17.0	1237	-8.0	33.0	123.0
10	11-Jan-2015 00:00:00	19930	AS	N791AS	422	SEA	LAX	063	-7.0	1237	-7.0	131.0	954.0
11	12-Jan-2015 00:00:00	19930	AS	N782AS	61	YAK	CDV	1149	-6.0	1237	-6.0	37.0	213.0
12	15-Jan-2015 00:00:00	19930	AS	N793AS	229	SJC	SEA	1027	27.0	1237	33.0	104.0	897.0
13	24-Jan-2015 00:00:00	19930	AS	N560AS	422	SEA	LAX	062	-6.0	1237	-7.0	125.0	954.0
14	24-Jan-2015 00:00:00	19930	AS	N410AS	627	LAS	PDX	1030	-10.0	1237	-16.0	112.0	763.0
15	26-Jan-2015 00:00:00	19930	AS	N593AS	879	SEA	LH	827	-3.0	1237	-38.0	340.0	2701.0
16	26-Jan-2015 00:00:00	19930	AS	N764AS	43	ANC	BET	1116	8.0	1237	-2.0	60.0	399.0
17	15-Jan-2015 00:00:00	20409	B6	N775JB	5	EWR	FLL	031	-14.0	1237	-15.0	144.0	1055.0
18	23-Jan-2015 00:00:00	19930	AS	N565AS	662	SEA	DFW	700	10.0	1237	11.0	108.0	1860.0
19	05-Jan-2015 00:00:00	20409	B6	N807JB	133	BOS	SFO	827	-3.0	1237	18.0	399.0	2704.0
20	20-Jan-2015 00:00:00	19930	AS	N560AS	835	SJC	KOA	021	-14.0	1237	-35.0	291.0	2384.0
21	21-Jan-2015 00:00:00	19930	AS	N613AS	501	SNA	SEA	050	-10.0	1237	-3.0	151.0	978.0
22	07-Jan-2015 00:00:00	20409	B6	N507JB	308	FLL	EWR	658	0.0	1237	-8.0	143.0	1055.0
23	22-Jan-2015 00:00:00	20409	B6	N388JB	315	SYR	JFK	1125	-8.0	1237	-4.0	42.0	209.0

REMOVE ROW WITH THIS COLUMN VALUE

7.3.4 Remove Duplicate columns with this column

This function allows users to remove all duplicate columns of a target column. A duplicate column is a row that contains exactly the same data as the target row.

Shown below is the before and after scenario of “Remove Duplicate columns with this column” for column “DEST”:

Before:

The screenshot shows the Smarten Advanced Data Discovery interface. The title bar says "FlightData_Jan_2015_Dataset". The top right corner shows "Welcome admin". Below the title bar are various icons for file operations like save, open, and search. A message at the top right says "Last refreshed on 23-Mar-2018 12:18:56". The main area is a table titled "Result set" with 14 columns and 23 rows. The columns are labeled: #, FLIGHDATE, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEST_1, DEPETIME, DEPDELAY, ARRTIME, ARRDELAY, AIRTIME. The first few rows of data are as follows:

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEST_1	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	LAX	855	-5.0	1237	7.0	378.0
2	23-Jan-2015 00:00:00	19805	AA	N3JUA	25	BOS	LAX	LAX	900	0.0	1237	0.0	331.0
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	ORD	811	-4.0	1237	2.0	181.0
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	ORD	807	-6.0	1237	2.0	178.0
5	15-Jan-2015 00:00:00	19805	AA	N3BSAA	316	DFW	DEN	DEN	1142	7.0	1237	1.0	95.0
6	04-Jan-2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	SFO	1054	29.0	1237	27.0	205.0
7	29-Jan-2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	ELP	1145	-5.0	1237	2.0	80.0
8	31-Jan-2015 00:00:00	19805	AA	N5E5AA	253	LAX	OOG	OOG	858	-2.0	1237	-18.0	318.0
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	LAX	954	-6.0	1237	-53.0	317.0
10	21-Jan-2015 00:00:00	19805	AA	N3CBAA	1010	DFW	PBI	PBI	915	-5.0	1237	-17.0	127.0
11	03-Jan-2015 00:00:00	19805	AA	N4T4AA	1023	DFW	AUS	AUS	1146	-4.0	1237	-13.0	38.0
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	DFW	919	-6.0	1237	-23.0	238.0
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	BOS	927	-3.0	1237	-7.0	165.0
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	BOS	932	2.0	1237	-7.0	168.0
15	08-Jan-2015 00:00:00	19805	AA	N556AA	1046	MCI	DFW	DFW	1101	-3.0	1237	-8.0	75.0
16	11-Jan-2015 00:00:00	19805	AA	N4XGAA	1046	MCI	DFW	DFW	1051	-13.0	1237	-8.0	72.0
17	14-Jan-2015 00:00:00	19805	AA	N4WPAA	1046	MCI	DFW	DFW	1055	-9.0	1237	-8.0	82.0
18	18-Jan-2015 00:00:00	19805	AA	N3L6AA	1238	FLL	ORD	ORD	1024	-6.0	1237	-13.0	172.0
19	21-Jan-2015 00:00:00	19805	AA	N3DAAA	1108	DFW	LGA	LGA	850	80.0	1237	38.0	153.0
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	LGA	829	-1.0	1237	-2.0	163.0
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	CLE	911	16.0	1237	10.0	129.0
22	03-Jan-2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	STT	759	-1.0	1237	-18.0	190.0
23	31-Jan-2015 00:00:00	19805	AA	N3LVAA	1334	SAN	ORD	ORD	843	-1.0	1237	-13.0	205.0



After:

REMOVE DUPLICATE COLUMNS WITH THIS COLUMN

7.3.5 Remove Duplicate rows with this row

This function allows users to remove all duplicate rows of a target row. A duplicate row is a row that contains exactly the same data as the target row.

Shown below is the before and after scenario of “Remove Duplicate rows with this row”:

Before:

After:

The screenshot shows a table with 474 rows of flight data. A specific row at index 444 is highlighted with a red border. After performing an action, a message box appears stating "All the duplicate rows for selected row removed successfully".

REMOVE DUPLICATE ROWS WITH THIS ROW

7.3.6 Remove All duplicate rows

This function allows users to remove all duplicate rows in the Dataset.

Shown below is the before and after scenario of “Remove All duplicate rows”:

Before:

The image below highlights all duplicate rows in the Dataset.

The screenshot shows a table with 13 rows of flight data. Most rows are highlighted with different colors (green, blue, black) to identify them as duplicates. An arrow points downwards from this screenshot to the 'After' screenshot.

After:

The screenshot shows the same dataset after duplicates have been removed. The table now contains only 10 rows. A message box at the bottom states "All the duplicate rows removed successfully".

REMOVE ALL DUPLICATE ROWS

7.3.7 Remove Rows with all null

This function allows users to remove all rows that have null in all their string and integer type of columns.

Shown below is the before and after scenario of “Remove Rows with all null”:

Before:

#	FLIGHTDATE	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPODELAY	ARRTIME	ARREDELAY	AIRTIME	DISTANCE
1	January 14, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
2	January 14, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
3	January 14, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
4	January 14, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	January 14, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
6	January 14, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
7	January 14, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
8	January 14, 2015 00:00:00	N201AA	1084	TUS	DFW	041	-5	1237	-18	99	813
9	January 21, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
10	January 31, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
11	January 31, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
12	January 31, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
13	January 31, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
14	January 31, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
15	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
16	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
17	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
18	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
19	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
20	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
21	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
22	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
23	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
24	January 07, 2015 00:00:00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
25	January 15, 2015 00:00:00	N472AA	0	TUS	ORD	0	0	0	0	0	0
26	January 15, 2015 00:00:00	N775JB	5	EWR	FLL	931	-14	1237	-15	144	1055
27	January 14, 2015 00:00:00	N792AS	65	PSG	JNU	1149	-17	1237	-13	33	123
28	January 23, 2015 00:00:00	N201AA	2428	LAX	DFW	749	-6	1237	-23	208	1592
29	January 15, 2015 00:00:00	N201AA	1027	BOS	DFW	016	48	1237	-23	146	1235
30	January 12, 2015 00:00:00	N503AA	1310	DFW	CLE	911	15	1237	10	129	1021
31	January 21, 2015 00:00:00	N813AS	501	SNA	SEA	950	-10	1237	-3	151	978
32	January 08, 2015 00:00:00	N797AS	75	SEA	JNU	1119	9	1237	-15	120	909
33	January 03, 2015 00:00:00	N841AB	722	PBI	BOS	949	-8	1237	-14	151	1197
34	January 13, 2015 00:00:00	N793AS	22	GTE	KTN	1152	-45	1237	-11	33	163
35	January 13, 2015 00:00:00	N3GCAA	1418	DFW	HGN	1125	11	1237	2	111	769
36	January 12, 2015 00:00:00	N856AA	1033	MIA	BOS	927	-3	1237	-7	185	1288
37	January 23, 2015 00:00:00	N298JB	1334	SJU	MCO	1025	-15	1237	-11	176	1189
38	January 15, 2015 00:00:00	N019AA	1513	ATL	DFW	1054	7	1237	17	118	731
39	January 01, 2015 00:00:00	N3DEAA	1554	DFW	SNA	1130	-5	1237	-13	105	1205
40	January 23, 2015 00:00:00	N201AA	1437	MIA	DFW	952	39	1237	62	166	1121
41	January 23, 2015 00:00:00	N015AA	1182	ATL	MIA	1050	63	1237	51	74	594
42	January 09, 2015 00:00:00	N516AS	62	SIT	KTN	1148	-10	1237	-13	37	183
43	January 24, 2015 00:00:00	N419AS	627	LAS	FOX	1030	-10	1237	-16	112	763
44	January 25, 2015 00:00:00	N865JB	271	LGA	FLL	044	14	1237	-4	155	1079
45	January 09, 2015 00:00:00	N863AS	4681	SAN	SEA	955	-40	1237	-23	159	1050
46	January 08, 2015 00:00:00	N863AS	55	BRW	FLL	1119	1	1237	1	89	503
47	January 23, 2015 00:00:00	N020AA	0	DFW	ELP	0	0	0	0	0	0
48	January 22, 2015 00:00:00	N3FAAA	1418	DFW	HGN	1135	20	1237	2	106	769
49	January 23, 2015 00:00:00	N865AS	662	SEA	DFW	700	10	1237	11	198	1660
50	January 11, 2015 00:00:00	N781AS	422	DFW	SFO	1054	29	1237	27	205	1660
51	January 24, 2015 00:00:00	N834JB	411						-7	131	654
52	January 17, 2015 00:00:00	N489AA	1584						-11	283	2248
53	January 30, 2015 00:00:00	N865JB	354						-18	101	813
54									32	138	1028



After:

#	FLIGHTDATE	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPODELAY	ARRTIME	ARREDELAY	AIRTIME	DISTANCE
1	January 21, 2015 00:00:00	N201AA	1084	TUS	DFW	041	-5	1237	-18	99	813
2	January 10, 2015 00:00:00	N472AA	0	TUS	ORD	0	0	0	0	0	0
3	January 15, 2015 00:00:00	N775JB	5	EWR	FLL	931	-14	1237	-15	144	1055
4	January 14, 2015 00:00:00	N792AS	65	PSG	JNU	1149	-17	1237	-8	33	123
5	January 19, 2015 00:00:00	N813AS	2428	LAX	DFW	749	-6	1237	-13	146	1235
6	January 15, 2015 00:00:00	N201AA	1027	BOS	DFW	016	48	1237	-23	208	1592
7	January 12, 2015 00:00:00	N503AA	1310	DFW	CLE	911	15	1237	10	129	1021
8	January 21, 2015 00:00:00	N813AS	501	SNA	SEA	950	-10	1237	-3	151	978
9	January 08, 2015 00:00:00	N797AS	75	SEA	JNU	1119	9	1237	-15	120	909
10	January 03, 2015 00:00:00	N841AB	722	PBI	BOS	949	-8	1237	-14	151	1197
11	January 13, 2015 00:00:00	N793AS	22	GTE	KTN	1152	-45	1237	-11	33	163
12	January 13, 2015 00:00:00	N3GCAA	1418	DFW	HGN	1125	11	1237	2	111	769
13	January 12, 2015 00:00:00	N856AA	1033	MIA	BOS	927	-3	1237	-7	185	1288
14	January 23, 2015 00:00:00	N298JB	1334	SJU	MCO	1025	-15	1237	-11	176	1189
15	January 15, 2015 00:00:00	N019AA	1513	ATL	DFW	1054	7	1237	17	118	731
16	January 01, 2015 00:00:00	N3DEAA	1554	DFW	SNA	1130	-5	1237	-13	105	1205
17	January 23, 2015 00:00:00	N201AA	1437	MIA	DFW	952	39	1237	62	166	1121
18	January 23, 2015 00:00:00	N015AA	1182	ATL	MIA	1050	63	1237	51	74	594
19	January 09, 2015 00:00:00	N516AS	62	SIT	KTN	1148	-10	1237	-13	37	183
20	January 24, 2015 00:00:00	N419AS	627	LAS	FOX	1030	-10	1237	-16	112	763
21	January 25, 2015 00:00:00	N865JB	271	LGA	FLL	044	14	1237	-4	155	1079
22	January 09, 2015 00:00:00	N863AS	4681	SAN	SEA	955	-40	1237	-23	159	1050
23	January 08, 2015 00:00:00	N863AS	55	BRW	FLL	1119	1	1237	1	89	503
24	January 23, 2015 00:00:00	N020AA	0	DFW	ELP	0	0	0	0	0	0
25	January 22, 2015 00:00:00	N3FAAA	1418	DFW	HGN	1135	20	1237	2	106	769
26	January 23, 2015 00:00:00	N865AS	662	SEA	DFW	700	10	1237	11	198	1660
27	January 11, 2015 00:00:00	N781AS	422	DFW	SFO	1054	29	1237	27	205	1660
28	January 24, 2015 00:00:00	N834JB	411						-7	131	654
29	January 17, 2015 00:00:00	N489AA	1584						-11	283	2248
30	January 30, 2015 00:00:00	N865JB	354						-18	101	813
31									32	138	1028

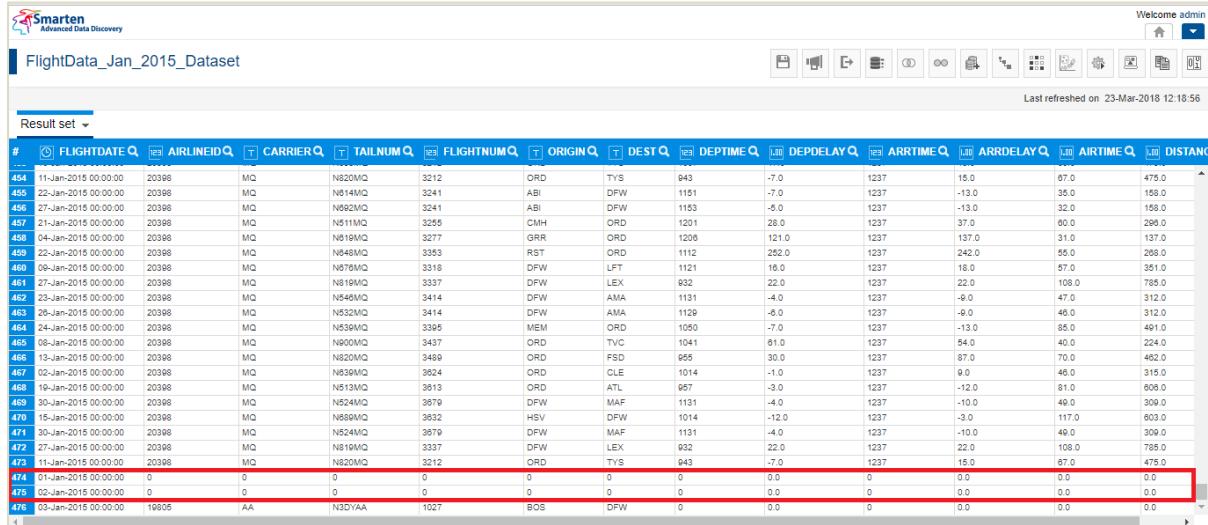
REMOVE ROWS WITH ALL NULL

7.3.8 Remove Rows with all zeros

This function allows users to remove all rows that have zero in all columns.

Shown below is the before and after scenario of “Remove Rows with all zeros”:

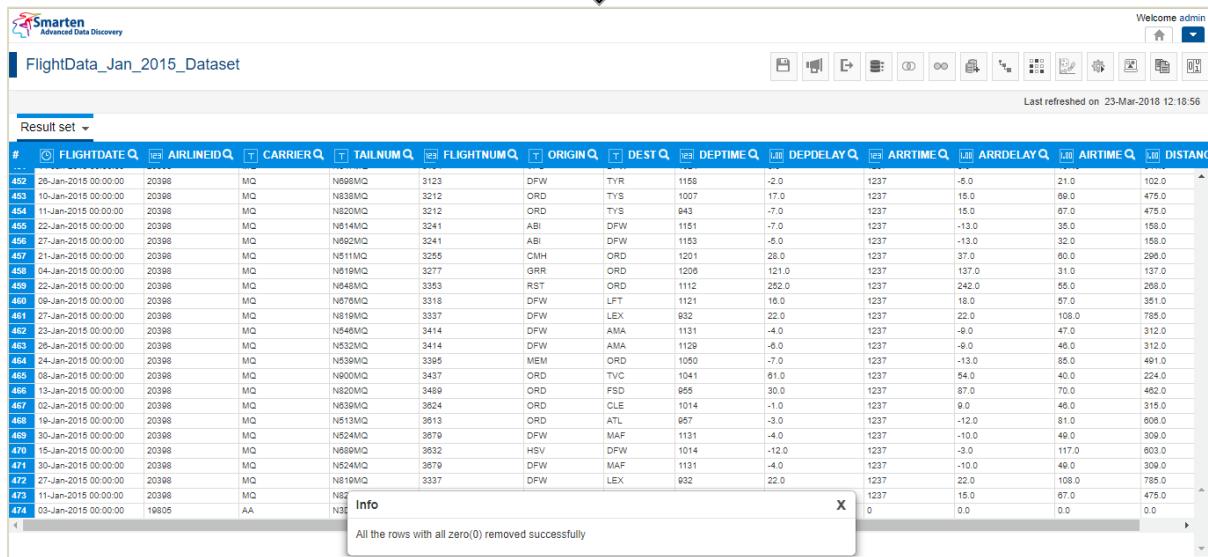
Before:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTHI	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
454	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	943	-7.0	1237	15.0	07.0	475.0
455	22-Jan-2015 00:00:00	20398	MQ	N814MQ	3241	ABI	DFW	1151	-7.0	1237	-13.0	35.0	158.0
456	27-Jan-2015 00:00:00	20398	MQ	N829MQ	3241	ABI	DFW	1153	-5.0	1237	-13.0	32.0	158.0
457	21-Jan-2015 00:00:00	20398	MQ	N511MQ	3255	CMH	ORD	1201	28.0	1237	37.0	60.0	298.0
458	04-Jan-2015 00:00:00	20398	MQ	N819MQ	3277	GRR	ORD	1203	121.0	1237	137.0	31.0	137.0
459	22-Jan-2015 00:00:00	20398	MQ	N849MQ	3353	RST	ORD	1112	252.0	1237	242.0	55.0	288.0
460	09-Jan-2015 00:00:00	20398	MQ	N876MQ	3316	DFW	LFT	1121	16.0	1237	18.0	57.0	351.0
461	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	932	22.0	1237	22.0	108.0	758.0
462	23-Jan-2015 00:00:00	20398	MQ	N846MQ	3414	DFW	AMA	1131	-4.0	1237	-9.0	47.0	312.0
463	26-Jan-2015 00:00:00	20398	MQ	N832MQ	3414	DFW	AMA	1129	-6.0	1237	-9.0	46.0	312.0
464	24-Jan-2015 00:00:00	20398	MQ	N539MQ	3395	MEM	ORD	1050	-7.0	1237	-13.0	85.0	491.0
465	08-Jan-2015 00:00:00	20398	MQ	N900MQ	3437	ORD	TVC	1041	61.0	1237	54.0	40.0	224.0
466	13-Jan-2015 00:00:00	20398	MQ	N820MQ	3489	ORD	FSD	955	30.0	1237	87.0	70.0	462.0
467	02-Jan-2015 00:00:00	20398	MQ	N839MQ	3624	ORD	CLE	1014	-1.0	1237	9.0	46.0	315.0
468	19-Jan-2015 00:00:00	20398	MQ	N813MQ	3913	ORD	ATL	957	-3.0	1237	-12.0	81.0	608.0
469	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3879	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
470	15-Jan-2015 00:00:00	20398	MQ	N889MQ	3832	HSV	DFW	1014	-12.0	1237	-3.0	117.0	603.0
471	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3879	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
472	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	932	22.0	1237	22.0	108.0	758.0
473	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	943	-7.0	1237	15.0	87.0	475.0
474	01-Jan-2015 00:00:00	0	0	0	0	0	0	0.0	0.0	0	0.0	0.0	0.0
475	02-Jan-2015 00:00:00	0	0	0	0	0	0	0.0	0.0	0	0.0	0.0	0.0
476	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	0	0.0	0	0.0	0.0	0.0



After:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTHI	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
452	26-Jan-2015 00:00:00	20398	MQ	N868MQ	3123	DFW	TYR	1158	-2.0	1237	-5.0	21.0	102.0
453	10-Jan-2015 00:00:00	20398	MQ	N838MQ	3212	ORD	TYS	1007	17.0	1237	15.0	69.0	475.0
454	11-Jan-2015 00:00:00	20398	MQ	N820MQ	3212	ORD	TYS	943	-7.0	1237	15.0	87.0	475.0
455	22-Jan-2015 00:00:00	20398	MQ	N814MQ	3241	ABI	DFW	1151	-7.0	1237	-13.0	35.0	158.0
456	27-Jan-2015 00:00:00	20398	MQ	N829MQ	3241	ABI	DFW	1153	-5.0	1237	-13.0	32.0	158.0
457	21-Jan-2015 00:00:00	20398	MQ	N511MQ	3255	CMH	ORD	1201	28.0	1237	37.0	60.0	298.0
458	04-Jan-2015 00:00:00	20398	MQ	N819MQ	3277	GRR	ORD	1206	121.0	1237	137.0	31.0	137.0
459	22-Jan-2015 00:00:00	20398	MQ	N848MQ	3353	RST	ORD	1112	252.0	1237	242.0	55.0	288.0
460	09-Jan-2015 00:00:00	20398	MQ	N876MQ	3316	DFW	LFT	1121	16.0	1237	18.0	57.0	351.0
461	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	932	22.0	1237	22.0	108.0	758.0
462	23-Jan-2015 00:00:00	20398	MQ	N546MQ	3414	DFW	AMA	1131	-4.0	1237	-9.0	47.0	312.0
463	20-Jan-2015 00:00:00	20398	MQ	N532MQ	3414	DFW	AMA	1129	-6.0	1237	-9.0	46.0	312.0
464	24-Jan-2015 00:00:00	20398	MQ	N539MQ	3395	MEM	ORD	1050	-7.0	1237	-13.0	85.0	491.0
465	08-Jan-2015 00:00:00	20398	MQ	N900MQ	3437	ORD	TVC	1041	61.0	1237	54.0	40.0	224.0
466	13-Jan-2015 00:00:00	20398	MQ	N820MQ	3489	ORD	FSD	955	30.0	1237	87.0	70.0	462.0
467	02-Jan-2015 00:00:00	20398	MQ	N839MQ	3624	ORD	CLE	1014	-1.0	1237	9.0	46.0	315.0
468	19-Jan-2015 00:00:00	20398	MQ	N813MQ	3913	ORD	ATL	957	-3.0	1237	-12.0	81.0	608.0
469	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3879	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
470	15-Jan-2015 00:00:00	20398	MQ	N889MQ	3832	HSV	DFW	1014	-12.0	1237	-3.0	117.0	603.0
471	30-Jan-2015 00:00:00	20398	MQ	N524MQ	3879	DFW	MAF	1131	-4.0	1237	-10.0	49.0	309.0
472	27-Jan-2015 00:00:00	20398	MQ	N819MQ	3337	DFW	LEX	932	22.0	1237	22.0	108.0	758.0
473	11-Jan-2015 00:00:00	20398	MQ	N820MQ	Info					1237	15.0	87.0	475.0
474	03-Jan-2015 00:00:00	19805	AA	N3DYAA	0				0.0	0	0.0	0.0	0.0

All the rows with all zero(0) removed successfully

X REMOVE ROWS WITH ALL ZEROS

7.3.9 Remove Columns with all null

This function allows users to remove one or more columns that have null value in all the rows.

Shown below is the before and after scenario of “Remove Columns with all null”:

Before:

Result set ▾

#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRELAY	AIRTIME	DISTANCE
1	January 01, 2015 00:00:00	NULL	AA	N787AA	1	JFK	LAX	NULL	-5	1237	NULL	378	2475
2	January 23, 2015 00:00:00	NULL	AA	N3JAA	25	BOS	LAX	NULL	0	1237	NULL	331	2811
3	January 05, 2015 00:00:00	NULL	AA	N590AA	300	TUS	ORD	NULL	-4	1237	NULL	181	1437
4	January 10, 2015 00:00:00	NULL	AA	N472AA	300	TUS	ORD	NULL	-9	1237	NULL	178	1437
5	January 15, 2015 00:00:00	NULL	AA	N295AA	318	DFW	DEN	NULL	7	1237	1	95	641
6	January 04, 2015 00:00:00	NULL	AA	N2HLAA	184	DFW	SFO	NULL	29	1237	27	205	1464
7	January 29, 2015 00:00:00	NULL	AA	N002AA	1079	DFW	ELP	NULL	-5	1237	2	80	551
8	January 31, 2015 00:00:00	NULL	AA	N5E5AA	253	LAX	OOG	NULL	-2	1237	-19	318	2488
9	January 14, 2015 00:00:00	NULL	AA	N788AA	255	JFK	LAX	NULL	-5	1237	-53	317	2475
10	January 21, 2015 00:00:00	NULL	AA	N3CBA	1010	DFW	PBI	NULL	-5	1237	-17	127	1102
11	January 03, 2015 00:00:00	NULL	AA	N474AA	1023	DFW	AUS	NULL	-4	1237	-13	39	190
12	January 03, 2015 00:00:00	NULL	AA	N3DYAA	1027	BOS	DFW	NULL	-6	1237	-23	238	1562
13	January 12, 2015 00:00:00	NULL	AA	N565AA	1033	MIA	BOS	NULL	-3	1237	-7	165	1258
14	January 15, 2015 00:00:00	NULL	AA	N855AA	1033	MIA	BOS	NULL	2	1237	-7	168	1258
15	January 03, 2015 00:00:00	NULL	AA	N555AA	1046	MCI	DFW	NULL	-3	1237	-8	75	460
16	January 11, 2015 00:00:00	NULL	AA	N4X0AA	1045	MCI	DFW	NULL	-13	1237	-8	72	460
17	January 14, 2015 00:00:00	NULL	AA	N4WPAA	1045	MCI	DFW	NULL	-9	1237	-8	82	460
18	January 16, 2015 00:00:00	NULL	AA	N3LEAA	1238	FLL	ORD	NULL	-6	1237	-13	172	1182
19	January 21, 2015 00:00:00	NULL	AA	N3DJA	1108	DFW	LGA	NULL	60	1237	38	153	1389
20	January 08, 2015 00:00:00	NULL	AA	N3CKAA	1110	DFW	LGA	NULL	-1	1237	-2	163	1389
21	January 10, 2015 00:00:00	NULL	AA	N503AA	1310	DFW	CLE	NULL	16	1237	10	129	1021

After:

Result set ▾

#	FLIGHTDATE	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPDELAY	ARRTIME	ARRELAY	AIRTIME	DISTANCE
1	January 01, 2015 00:00:00	AA	N787AA	1	JFK	LAX	-5	1237	NULL	378	2475
2	January 23, 2015 00:00:00	AA	N3JAA	25	BOS	LAX	0	1237	NULL	331	2811
3	January 05, 2015 00:00:00	AA	N590AA	300	TUS	ORD	-4	1237	NULL	181	1437
4	January 10, 2015 00:00:00	AA	N472AA	300	TUS	ORD	-9	1237	NULL	178	1437
5	January 15, 2015 00:00:00	AA	N3BSAA	318	DFW	DEN	7	1237	1	95	641
6	January 04, 2015 00:00:00	AA	N3HLLA	184	DFW	SFO	29	1237	27	205	1464
7	January 29, 2015 00:00:00	AA	N002AA	1079	DFW	ELP	-5	1237	2	80	551
8	January 31, 2015 00:00:00	AA	N5E5AA	253	LAX	OOG	-2	1237	-19	318	2488
9	January 14, 2015 00:00:00	AA	N788AA	255	JFK	LAX	-6	1237	-53	317	2475
10	January 21, 2015 00:00:00	AA	N3CBA	1010	DFW	PBI	-5	1237	-17	127	1102
11	January 03, 2015 00:00:00	AA	N474AA	1023	DFW	AUS	-4	1237	-13	39	190
12	January 12, 2015 00:00:00	AA	N3DYAA	1027	BOS	DFW	-6	1237	-23	238	1562
13	January 15, 2015 00:00:00	AA	N855AA	1033	MIA	BOS	-3	1237	-7	165	1258
14	January 16, 2015 00:00:00	AA	N565AA	1033	MIA	BOS	2	1237	-7	168	1258
15	January 08, 2015 00:00:00	AA	N555AA	1046	MCI	DFW	-3	1237	-8	75	460
16	January 11, 2015 00:00:00	AA	N4X0AA	1045	MCI	DFW	-13	1237	-8	72	460
17	January 14, 2015 00:00:00	AA	N4WPAA	1045	MCI	DFW	-9	1237	-8	82	460
18	January 16, 2015 00:00:00	AA	N3LEAA	1238	FLL	ORD	-6	1237	-13	172	1182
19	January 21, 2015 00:00:00	AA	N3DJA	1108	DFW	LGA	60	1237	38	153	1389
20	January 09, 2015 00:00:00	AA	N3CKAA	1110	DFW	LGA	-1	1237	-2	163	1389
21	January 10, 2015 00:00:00	AA	N503AA	1310	DFW	CLE	16	1237	10	129	1021

Info
 All the columns with all null removed successfully

REMOVE COLUMNS WITH ALL NULL

7.3.10 Remove Columns with all zeros

This function allows users to remove one or more columns that have zero value in all rows.

Shown below is the before and after scenario of “Remove Columns with all zeros”:

Before:

FlightData_Jan_2015_Dataset4

Last refreshed on May 11, 2018 13:36:07

Result set

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM_Q	ORIGIN	DEST_Q	DEPTIME_Q	DEPDELAY_Q	ARRTIME_Q	ARRDELAY_Q	AIRTIME_Q	DISTANCE_Q
1	January 01, 2015 00:00:00	19805	AA	N787AA	0	JFK	0	855	-5	1237	7	378	2475
2	January 23, 2015 00:00:00	19805	AA	N3JUAA	0	BOS	0	900	0	1237	0	331	2611
3	January 05, 2015 00:00:00	19805	AA	N590AA	0	TUS	0	811	-4	1237	2	181	1437
4	January 10, 2015 00:00:00	19805	AA	N472AA	0	TUS	0	807	-9	1237	2	178	1437
5	January 15, 2015 00:00:00	19805	AA	N2BSAA	0	DFW	0	1142	7	1237	1	95	641
6	January 04, 2015 00:00:00	19805	AA	N3HLAA	0	DFW	0	1054	29	1237	27	205	1464
7	January 29, 2015 00:00:00	19805	AA	N002AA	0	DFW	0	1145	-5	1237	2	80	551
8	January 31, 2015 00:00:00	19805	AA	N5E5AA	0	LAX	0	858	-2	1237	-19	318	2486
9	January 14, 2015 00:00:00	19805	AA	N788AA	0	JFK	0	954	-6	1237	-53	317	2475
10	January 21, 2015 00:00:00	19805	AA	N3CBAA	0	DFW	0	915	-5	1237	-17	127	1102
11	January 03, 2015 00:00:00	19805	AA	N474AA	0	DFW	0	1146	-4	1237	-13	39	190
12	January 03, 2015 00:00:00	19805	AA	N2DYAA	0	BOS	0	919	-6	1237	-23	238	1562
13	January 12, 2015 00:00:00	19805	AA	N565AA	0	MIA	0	927	-3	1237	-7	165	1255
14	January 18, 2015 00:00:00	19805	AA	N555AA	0	MIA	0	932	2	1237	-7	168	1256
15	January 08, 2015 00:00:00	19805	AA	N555AA	0	MCI	0	1101	-3	1237	-8	75	460
16	January 11, 2015 00:00:00	19805	AA	N4XGAA	0	MCI	0	1051	-13	1237	-8	72	460
17	January 14, 2015 00:00:00	19805	AA	N4WPAA	0	MCI	0	1055	-9	1237	-8	82	460
18	January 18, 2015 00:00:00	19805	AA	N3LEAA	0	FLL	0	1024	-6	1237	-13	172	1182
19	January 21, 2015 00:00:00	19805	AA	N3DJA	0	DFW	0	850	60	1237	38	153	1389
20	January 09, 2015 00:00:00	19805	AA	N3CKAA	0	DFW	0	829	-1	1237	-2	163	1389
21	January 10, 2015 00:00:00	19805	AA	N503AA	0	DFW	0	911	16	1237	10	129	1021
22	January 03, 2015 00:00:00	19805	AA	N5EXAA	0	JFK	0	759	-1	1237	-18	190	1623
23	January 31, 2015 00:00:00	19805	AA	N3LVAA	0	SAN	0	643	-1	1237	-13	205	1723

After:

FlightData_Jan_2015_Dataset4

Last refreshed on May 11, 2018 13:36:07

Result set

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM_Q	ORIGIN	DEST_Q	DEPTIME_Q	DEPDELAY_Q	ARRTIME_Q	ARRDELAY_Q	AIRTIME_Q	DISTANCE_Q
1	January 01, 2015 00:00:00	19805	AA	N787AA	JFK	855	-5	1237	7	378	2475		
2	January 23, 2015 00:00:00	19805	AA	N3JUAA	BOS	900	0	1237	0	331	2611		
3	January 05, 2015 00:00:00	19805	AA	N590AA	TUS	811	-4	1237	2	181	1437		
4	January 10, 2015 00:00:00	19805	AA	N472AA	TUS	807	-9	1237	2	178	1437		
5	January 15, 2015 00:00:00	19805	AA	N3BSAA	DFW	1142	7	1237	1	95	641		
6	January 04, 2015 00:00:00	19805	AA	N3HLAA	DFW	1054	29	1237	27	205	1464		
7	January 29, 2015 00:00:00	19805	AA	N002AA	DFW	1145	-5	1237	2	80	551		
8	January 31, 2015 00:00:00	19805	AA	N5E5AA	LAX	858	-2	1237	-19	318	2486		
9	January 14, 2015 00:00:00	19805	AA	N788AA	JFK	954	-6	1237	-53	317	2475		
10	January 21, 2015 00:00:00	19805	AA	N3CBAA	DFW	915	-5	1237	-17	127	1102		
11	January 03, 2015 00:00:00	19805	AA	N474AA	DFW	1146	-4	1237	-13	39	190		
12	January 03, 2015 00:00:00	19805	AA	N2DYAA	BOS	919	-6	1237	-23	238	1562		
13	January 12, 2015 00:00:00	19805	AA	N565AA	MIA	927	-3	1237	-7	165	1255		
14	January 18, 2015 00:00:00	19805	AA	N555AA	MIA	932	2	1237	-7	168	1256		
15	January 08, 2015 00:00:00	19805	AA	N555AA	MCI	1101	-3	1237	-8	75	460		
16	January 11, 2015 00:00:00	19805	AA	N4XGAA	MCI	1051	-13	1237	-8	72	460		
17	January 14, 2015 00:00:00	19805	AA	N4WPAA	MCI	1055	-9	1237	-8	82	460		
18	January 18, 2015 00:00:00	19805	AA	N3LEAA	FLL	1024	-6	1237	-13	172	1182		
19	January 21, 2015 00:00:00	19805	AA	N3DJA	DFW	850	60	1237	38	153	1389		
20	January 09, 2015 00:00:00	19805	AA	N3CKAA	DFW	829	-1	1237	-2	163	1389		
21	January 10, 2015 00:00:00	19805	AA	N503AA	DFW	911	16	1237	10	129	1021		
22	January 03, 2015 00:00:00	19805	AA	N5EXAA	JFK	759	-1	1237	-18	190	1623		
23	January 31, 2015 00:00:00	19805	AA	N3LVAA					-13	205	1723		
24	January 13, 2015 00:00:00	19805	AA						2	111	769		
25	January 22, 2015 00:00:00	19805	AA						2	108	769		

REMOVE COLUMNS WITH ALL ZEROS

7.3.11 Remove Rows in search result

This function allows users to remove one or more rows that are filtered through search criteria.

These rows are removed from the dataset data.

For example, a user has searched records for the unique_carrier column with “WN” criteria. The user can use this feature to delete all searched records from the dataset. The system will delete searched records from the dataset data and reset search criteria.

Shown below is the before and after scenario of “Remove Rows in search result”:

Before:

Last refreshed on 14 September, 2018 06:50 PM

#	DATE	UNIQUE_CARRIER	FL_NUM	ORIGIN	DEST	DEP_TIME	DEP_DELAY	ARR_TIME	ARR_DELAY	AIR_TIME	DISTANCE
1	27 January, 2016 12:00 AM	WN	2513	GRR	DEN	620	-5.0	710	-20.0	131.0	1015.0
2	14 January, 2016 12:00 AM	WN	44	PHX	LAX	1925	1.0	1941	-4.0	66.0	370.0
3	26 January, 2016 12:00 AM	WN	2114	OAK	AUS	1040	0.0	1559	-6.0	189.0	1497.0
4	24 January, 2016 12:00 AM	WN	4158	HOU	DAL	1654	-8.0	1749	-11.0	42.0	239.0
5	19 January, 2016 12:00 AM	WN	358	MSP	MDW	835	-5.0	758	-12.0	65.0	349.0
6	20 January, 2016 12:00 AM	WN	1380	PHX	PHL	1005	0.0	1821	-14.0	235.0	2075.0
7	14 January, 2016 12:00 AM	WN	2733	DEN	BUR	840	0.0	1027	-13.0	79.0	602.0
8	1 January, 2016 12:00 AM	WN	253	HOU	DAL	824	-6.0	920	-10.0	40.0	239.0
9	8 January, 2016 12:00 AM	WN	1195	AUS	STL	919	-1.0	1011	-4.0	37.0	169.0
10	21 January, 2016 12:00 AM	WN	2529	RSW	SJC	1641	1.0	1812	-8.0	135.0	979.0
11	29 January, 2016 12:00 AM	WN	2704	SJC	QNT	1840	5.0	1937	-8.0	45.0	333.0
12	15 January, 2016 12:00 AM	WN	583	BWI	SJU	1210	-5.0	1711	-9.0	223.0	1585.0
13	16 January, 2016 12:00 AM	WN	1834	PHX	SAN	1944	-1.0	1942	-8.0	47.0	304.0
14	26 January, 2016 12:00 AM	WN	854	ATL	1815	1622	-3.0	2118	-2.0	102.0	680.0
15	17 January, 2016 12:00 AM	WN	611	PBI	ATL	1815	35.0	2003	28.0	84.0	545.0
16	28 January, 2016 12:00 AM	WN	2346	BUR	OAK	1819	-1.0	1727	-3.0	50.0	325.0
17	27 January, 2016 12:00 AM	WN	1261	PHX	SJC	1139	-6.0	1226	-19.0	90.0	621.0
18	21 January, 2016 12:00 AM	WN	3085	OAK	LAX	NULL	NULL	NULL	NULL	NULL	337.0
19	9 January, 2016 12:00 AM	WN	564	MCO	611	11.0	+	1214	NULL	NULL	849.0
20	24 January, 2016 12:00 AM	WN	2974	RDU	RDU	NULL	NULL	NULL	NULL	NULL	1436.0
21	19 January, 2016 12:00 AM	WN	503	BNA	HOU	1120	0.0	1331	-4.0	113.0	670.0
22	13 January, 2016 12:00 AM	WN	6147	LAX	LAS	1925	1.0	2022	-3.0	38.0	238.0
23	2 January, 2016 12:00 AM	WN	2931	CMI	OAK	621	1.0	827	-13.0	275.0	2110.0
24	11 January, 2016 12:00 AM	WN	3915	ATL	PBI	1958	33.0	2145	30.0	89.0	545.0
25	7 January, 2016 12:00 AM	DL	629	ATL	DTW	922	7.0	1106	-4.0	85.0	594.0
26	11 January, 2016 12:00 AM	DL	921	MSP	MKE	707	-3.0	819	-10.0	43.0	297.0
27	3 January, 2016 12:00 AM	DL	2541	PHX	SLC	658	-2.0	831	-12.0	73.0	507.0
28	3 January, 2016 12:00 AM	EV	5358	DSM	ATL	1659	4.0	2002	6.0	97.0	743.0
29	21 January, 2016 12:00 AM	DL	808	MSN	DTW	1648	3.0	1903	-2.0	54.0	311.0
30	1 January, 2016 12:00 AM	EV	3915	IAH	OKC	NULL	NULL	NULL	NULL	NULL	395.0
31	9 January, 2016 12:00 AM	EV	5945	BIF	ORD	653	-2.0	742	-16.0	68.0	474.0
32	30 January, 2016 12:00 AM	EV	4276	SYR	ORD	1655	-13.0	1807	-21.0	107.0	807.0
33	10 January, 2016 12:00 AM	EV	2745	GRK	DFW	635	1.0	738	5.0	31.0	134.0
34	1 January, 2016 12:00 AM	AS	469	SAN	SEA	629	-1.0	904	-29.0	135.0	1050.0
35	26 January, 2016 12:00 AM	DL	1122	MSP	ATL	729	0.0	1116	12.0	122.0	607.0
36	18 January, 2016 12:00 AM	EV	5224	FAY	ATL	1650	-1.0	1815	-6.0	65.0	331.0
37	13 January, 2016 12:00 AM	EV	5476	LGA	STL	1418	4.0	1559	-15.0	138.0	888.0
38	9 January, 2016 12:00 AM	DL	738	SAN	DTW	657	-3.0	1439	20.0	238.0	1956.0
39	29 January, 2016 12:00 AM	EV	2741	HOU	DFW	1914	-6.0	2030	-4.0	48.0	247.0
40	10 January, 2016 12:00 AM	EV	4543	STL	EWR	1528	-1.0	1938	46.0	113.0	872.0
41	25 January, 2016 12:00 AM	DL	1390	BIL	SLC	616	-4.0	729	-18.0	56.0	387.0
42	4 January, 2016 12:00 AM	UA	500	IND	SFO	611	9.0	757	-20.0	250.0	1943.0
43	16 January, 2016 12:00 AM	EV	4314	EVR	CLT	1205	-6.0	1357	-20.0	90.0	529.0
44	15 January, 2016 12:00 AM	EV	4728	CLL	IAH	1544	-12.0	1631	-10.0	22.0	74.0
45	10 January, 2016 12:00 AM	EV	4202	EWR	STL	1828	-1.0	2015	-9.0	134.0	872.0
46	22 January, 2016 12:00 AM	DL	2444	SLC	DTW	1525	0.0	2057	7.0	188.0	1481.0
47	1719	Info	All the rows in search result removed successfully	X							

After:

Last refreshed on 14 September, 2018 06:50 PM

#	DATE	UNIQUE_CARRIER	FL_NUM	ORIGIN	DEST	DEP_TIME	DEP_DELAY	ARR_TIME	ARR_DELAY	AIR_TIME	DISTANCE
1	7 January, 2016 12:00 AM	DL	629	ATL	DTW	922	7.0	1106	-4.0	85.0	594.0
2	11 January, 2016 12:00 AM	DL	921	MSP	MKE	707	-3.0	819	-10.0	43.0	297.0
3	3 January, 2016 12:00 AM	DL	2541	PHX	SLC	658	-2.0	831	-12.0	73.0	507.0
4	3 January, 2016 12:00 AM	EV	5358	DSM	ATL	1659	4.0	2002	6.0	97.0	743.0
5	21 January, 2016 12:00 AM	DL	808	MSN	DTW	1648	3.0	1903	-2.0	54.0	311.0
6	1 January, 2016 12:00 AM	EV	3915	IAH	OKC	NULL	NULL	NULL	NULL	NULL	395.0
7	9 January, 2016 12:00 AM	EV	5945	BIF	ORD	653	-2.0	742	-16.0	68.0	474.0
8	30 January, 2016 12:00 AM	EV	4276	SYR	ORD	1655	-13.0	1807	-21.0	107.0	807.0
9	10 January, 2016 12:00 AM	EV	2745	GRK	DFW	635	1.0	738	5.0	31.0	134.0
10	1 January, 2016 12:00 AM	AS	469	SAN	SEA	629	-1.0	904	-29.0	135.0	1050.0
11	26 January, 2016 12:00 AM	DL	1122	MSP	ATL	729	0.0	1116	12.0	122.0	607.0
12	18 January, 2016 12:00 AM	EV	5224	FAY	ATL	1650	-1.0	1815	-6.0	65.0	331.0
13	13 January, 2016 12:00 AM	EV	5476	LGA	STL	1418	4.0	1559	-15.0	138.0	888.0
14	9 January, 2016 12:00 AM	DL	738	SAN	DTW	657	-3.0	1439	20.0	238.0	1956.0
15	29 January, 2016 12:00 AM	EV	2741	HOU	DFW	1914	-6.0	2030	-4.0	48.0	247.0
16	10 January, 2016 12:00 AM	EV	4543	STL	EWR	1528	-1.0	1938	46.0	113.0	872.0
17	25 January, 2016 12:00 AM	DL	1390	BIL	SLC	616	-4.0	729	-18.0	56.0	387.0
18	4 January, 2016 12:00 AM	UA	500	IND	SFO	611	9.0	757	-20.0	250.0	1943.0
19	16 January, 2016 12:00 AM	EV	4314	EVR	CLT	1205	-6.0	1357	-20.0	90.0	529.0
20	15 January, 2016 12:00 AM	EV	4728	CLL	IAH	1544	-12.0	1631	-10.0	22.0	74.0
21	10 January, 2016 12:00 AM	EV	4202	EWR	STL	1828	-1.0	2015	-9.0	134.0	872.0
22	22 January, 2016 12:00 AM	DL	2444	SLC	DTW	1525	0.0	2057	7.0	188.0	1481.0
23	1719	Info	All the rows in search result removed successfully	X							
24	6 January, 2016 12:00 AM	EV	4655								
25	20 January, 2016 12:00 AM	DL	1319								
26	7 January, 2016 12:00 AM	DL	9311								

REMOVE ROWS IN SEARCH RESULT

7.4 Outliers

Outliers are the observations lying outside the overall pattern of distribution.

This function allows users to identify outlier values in the data and replace or remove them from the Dataset. Users can also download records detected as outliers in the form of a text file.

Note:

Outliers are applicable for numeric-type data only.

As shown in the image below, the columns of “AirTime” and “Distance” are selected to process Outliers.

FlightData_Jan_2015_Dataset

Last refreshed on 23-Mar-2018 12:18:56

Result set ▾

#	FLIGHTDATE	AIRLINED	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEP
1	19690	HA	N489HA	182	HNL	ITO	LIH	1144	-5.0
2	19690	HA	N478HA	182	HNL	ITO	LIH	1146	-3.0
3	19690	HA	N486HA	185	OIG	HNL	LIH	1139	-6.0
4	19690	HA	N487HA	313	HNL	LIH	LIH	1153	-8.0
5	19690	HA	N477HA	313	HNL	LIH	LIH	1153	-8.0
6	20398	MQ	N845MQ	2939	DFW	FWA	925	30.0	
7	20398	MQ	N923MQ	2939	DFW	FWA	846	-4.0	
8	20398	MQ	N528MQ	3031	LGA	ATL	946	-13.0	
9	20398	MQ	N658MQ	3049	IND	ORD	1156	82.0	
10	20398	MQ	N620MQ	3062	ORD	OIC	1022	92.0	
11	20398	MQ	N692MQ	3007	DFW	GRR	900	25.0	
12	20398	MQ	N614MQ	3011	DFW	JAN	1102	-3.0	
13	20398	MQ	N620MQ	3075	ORD	GRB	1145	75.0	
14	20398	MQ	N653MQ	2981	RIC	ORD	1101	-5.0	
15	20398	MQ	N647MQ	3154	VPS	DFW	1021	6.0	
16	20398	MQ	N698MQ	3123	DFW	TYR	1168	-2.0	
17	20398	MQ	N838MQ	3212	ORD	TYS	1007	17.0	
18	20398	MQ	N820MQ	3212	ORD	TYS	943	-7.0	
19	20398	MQ	N614MQ	3241	ABI	DFW	1151	-7.0	
20	20398	MQ	N692MQ	3241	ABI	DFW	1153	-5.0	
21	20398	MQ	N511MQ	3256	CMH	ORD	1201	28.0	
22	20398	MQ	N619MQ	3277	GRR	ORD	1206	121.0	
23	20398	MQ	N648MQ	3353	RST	ORD	1112	252.0	
24	20398	MQ	N676MQ	3318	DFW	LFT	1121	16.0	
25	20398	MQ	N819MQ	3337	DFW	LEX	932	22.0	
26	20398	MQ	N546MQ	3414	DFW	AMA	1131	-4.0	
27	20398	MQ	N532MQ	3414	DFW	AMA	1129	-6.0	
28	20398	MQ	N530MQ	3395	MEM	ORD	1050	-7.0	
29	20398	MQ	N900MQ	3437	ORD	JVC	1041	81.0	
30	20398	MQ	N620MQ	3489	ORD	FSD	955	30.0	
31	20398	MQ	N659MQ	3624	ORD	CLE	1014	-1.0	
32	20398	MQ	N513MQ	3613	ORD	ATL	957	-3.0	
33	20398	MQ	N524MQ	3679	DFW	MAF	1131	-4.0	
34	20398	MQ	N659MQ	3632	HSV	DFW	1014	-12.0	
35	20398	MQ	N524MQ	3679	DFW	MAF	1131	-4.0	
36	20398	MQ	N819MQ	3237	DFW	LEX	932	22.0	
37	20398	MQ	N820MQ	3212	ORD	TYS	943	-7.0	
38	19805	AA	N3DYAA	1027	BOS	DFW	0	0.0	

Outliers

Select columns to process outliers

Available columns	Selected columns
AirlineID	+ \$ AirTime
FlightNum	+ \$ Distance
DepTime	+ \$
DepDelay	+ \$
ArrTime	+ \$
ArrDelay	+ \$

APPLY CANCEL

10.0.0.22:8080/smarten/datasets/162519e1267.ds?isFromRepository=true#

OUTLIERS—COLUMN SELECTION

FlightData_Jan_2015_Dataset

Last refreshed on 23-Mar-2018 12:18:56

Result set ▾

#	FLIGHTDATE	AIRLINED	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTH
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855
2	23-Jan-2015 00:00:00	19805	AA	N3JUAA	25	BOS	LAX	900
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807
5	15-Jan-2015 00:00:00	19805	AA	N3BSAA	316	DFW	DEN	1142
6	04-Jan-2015 00:00:00	19805	AA	N3HLLA	184	DFW	SFO	1054
7	28-Jan-2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145
8	31-Jan-2015 00:00:00	19805	AA	N6E5AA	253	LAX	OIG	858
9	14-Jan-2015 00:00:00	19805	AA	N783AA	255	JFK	LAX	954
10	21-Jan-2015 00:00:00	19805	AA	N3CBAA	1010	DFW	PBI	915
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1146
12	03-Jan-2015 00:00:00	19805	AA	N3DAAA	1027	BOS	DFW	919
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932
15	08-Jan-2015 00:00:00	19805	AA	N655AA	1046	MCI	DFW	1101
16	11-Jan-2015 00:00:00	19805	AA	N4X6AA	1046	MCI	DFW	1051
17	14-Jan-2015 00:00:00	19805	AA	N4WPAA	1046	MCI	DFW	1055
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024
19	21-Jan-2015 00:00:00	19805	AA	N3DAAA	1108	DFW	LGA	850
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911
22	03-Jan-2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759
23	31-Jan-2015 00:00:00	19805	AA	N3LVA	1334	SAN	ORD	643
24	13-Jan-2015 00:00:00	19805	AA	N3C5AA	1419	DFW	HND	1126
25	22-Jan-2015 00:00:00	19805	AA	N3FAAA	1419	DFW	HND	1135
26	01-Jan-2015 00:00:00	19805	AA	N3DUAA	1482	SFO	DFW	703
27	29-Jan-2015 00:00:00	19805	AA	N3HDA	1482	SFO	DFW	711
28	30-Jan-2015 00:00:00	19805	AA	N4Q4AA	1083	DFW	AUS	1145
29	28-Jan-2015 00:00:00	19805	AA	N016AA	1162	ATL	MIA	1050
30	01-Jan-2015 00:00:00	19805	AA	N481AA	1174	RNO	DFW	712
31	09-Jan-2015 00:00:00	19805	AA	N3EMAA	1175	MIA	IAH	1046
32	01-Jan-2015 00:00:00	19805	AA	N3DEAA	1554	DFW	SNA	1130
33	26-Jan-2015 00:00:00	19805	AA	N503AA	1627	DFW	FAT	1129
34	17-Jan-2015 00:00:00	19805	AA	N489AA	1584	TUS	DFW	940
35	21-Jan-2015 00:00:00	19805	AA	N201AA	1584	TUS	DFW	941
36	27-Jan-2015 00:00:00	19805	AA	N496AA	1584	TUS	DFW	937
37	03-Jan-2015 00:00:00	19805	AA	N3LTA	1694	MIA	LGA	953
38	09-Jan-2015 00:00:00	19805	AA	N3HEAA	1507	SEA	DFW	653

Outliers

Total 1 records (0% of total records) detected as outliers with 100% accuracy.
[Click here to download.](#)

COLUMN NAME	#
AirTime	#1

Distance

0 200 400

0 2,000 4,000

REMOVE OUTLIERS REPLACE OUTLIERS CANCEL BACK

www.ElegantBI.com

OUTLIERS

FlightData_Jan_2015_Dataset

Last refreshed on 23-Mar-2018 12:18:56

Result set

#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTH
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855
2	23-Jan-2015 00:00:00	19805	AA	N3JUA	25	BOS	LAX	900
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807
5	15-Jan-2015 00:00:00	19805	AA	N38SA	316	DFW	DEN	1142
6	04-Jan-2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	1054
7	29-Jan-2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145
8	31-Jan-2015 00:00:00	19805	AA	N5ESA	253	LAX	DGG	858
9	14-Jan-2015 00:00:00	19805	AA	N789AA	255	JFK	LAX	954
10	21-Jan-2015 00:00:00	19805	AA	N3CBA	1010	DFW	PBI	915
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1146
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919
13	12-Jan-2015 00:00:00	19805	AA	N595AA	1033	MIA	BOS	927
14	16-Jan-2015 00:00:00	19805	AA	N655AA	1033	MIA	BOS	932
15	09-Jan-2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101
16	11-Jan-2015 00:00:00	19805	AA	N4X6AA	1046	MCI	DFW	1051
17	14-Jan-2015 00:00:00	19805	AA	N4WIPAA	1046	MCI	DFW	1055
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024
19	21-Jan-2015 00:00:00	19805	AA	N3DJA	1108	DFW	LGA	850
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829
21	10-Jan-2015 00:00:00	19805	AA	N630AA	1310	DFW	CLE	911
22	03-Jan-2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759
23	31-Jan-2015 00:00:00	19805	AA	N3LVA	1334	SAN	ORD	843
24	13-Jan-2015 00:00:00	19805	AA	N3GCAA	1419	DFW	HND	1126
25	22-Jan-2015 00:00:00	19805	AA	N3FAAA	1418	DFW	HND	1135
26	01-Jan-2015 00:00:00	19805	AA	N3DUAA	1482	SFO	DFW	703
27	28-Jan-2015 00:00:00	19805	AA	N3HDA	1482	SFO	DFW	711
28	30-Jan-2015 00:00:00	19805	AA	N499AA	1083	DFW	AUS	1145
29	28-Jan-2015 00:00:00	19805	AA	N016AA	1162	ATL	MIA	1050
30	01-Jan-2015 00:00:00	19805	AA	N481AA	1174	RNO	DFW	712
31	09-Jan-2015 00:00:00	19805	AA	N3EMAA	1175	MIA	IAH	1046
32	01-Jan-2015 00:00:00	19805	AA	N3DEAA	1554	DFW	SNA	1130
33	26-Jan-2015 00:00:00	19805	AA	N503AA	1627	DFW	FAT	1129
34	17-Jan-2015 00:00:00	19805	AA	N489AA	1584	TUS	DFW	940
35	21-Jan-2015 00:00:00	19805	AA	N201AA	1594	TUS	DFW	941
36	22-Jan-2015 00:00:00	19805	AA	N3DJA	1604	TUS	DFW	947

Outliers

Total 1 records (0% of total records) detected as outliers with 100% accuracy.
[Click here to download.](#)

COLUMN NAME #

AirTime #1

Upper bound: 392
Upper quartile: 143
Median: 93
Lower quartile: 60
Lower bound: 0

Distance #0

Upper bound: 2704
Upper quartile: 641
Median: 0
Lower quartile: 0
Lower bound: 0

OUTLIER DETAILS

OUTLIERS—REMOVE OUTLIERS

REMOVE OUTLIERS REPLACE OUTLIERS CANCEL BACK

7.4.1 Remove Outliers

This option allows users to remove all rows of selected columns that contain outlier values.

FlightData_Jan_2015_Dataset

Last refreshed on 23-Mar-2018 12:18:56

Result set

#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTH
1	015 00:00:00	19690	HA	N489HA	182	HNL	ITO	1144
2	015 00:00:00	19690	HA	N478HA	182	HNL	ITO	1146
3	015 00:00:00	19690	HA	N485HA	185	OIG	HNL	1139
4	015 00:00:00	19690	HA	N487HA	313	HNL	LIH	1153
5	015 00:00:00	19690	HA	N477HA	313	HNL	LIH	1153
6	015 00:00:00	20988	MQ	N846MQ	2939	DFW	FIA	925
7	015 00:00:00	20988	MQ	N923MQ	2939	DFW	FIA	930
8	015 00:00:00	20988	MQ	N629MQ	3031	LGA	ATL	946
9	015 00:00:00	20988	MQ	N658MQ	3049	IND	ORD	1166
10	015 00:00:00	20988	MQ	N620MQ	3062	ORD	OIC	1022
11	015 00:00:00	20988	MQ	N692MQ	3007	DFW	GRR	900
12	015 00:00:00	20988	MQ	N614MQ	3011	DFW	JAN	1102
13	015 00:00:00	20988	MQ	N620MQ	3075	ORD	GRB	1146
14	015 00:00:00	20988	MQ	N853MQ	2981	RIC	ORD	1101
15	015 00:00:00	20988	MQ	N647MQ	3164	VPS	DFW	1021
16	015 00:00:00	20988	MQ	N698MQ	3123	DFW	TYR	1168
17	015 00:00:00	20988	MQ	N838MQ	3212	ORD	TYS	1007
18	015 00:00:00	20988	MQ	N820MQ	3212	ORD	TYS	943
19	015 00:00:00	20988	MQ	N614MQ	3241	ABI	DFW	1151
20	015 00:00:00	20988	MQ	N692MQ	3241	ABI	DFW	1153
21	015 00:00:00	20988	MQ	N511MQ	3255	CMH	ORD	1201
22	015 00:00:00	20988	MQ	N619MQ	3277	GRB	ORD	1206
23	015 00:00:00	20988	MQ	N648MQ	3353	RST	ORD	1112
24	015 00:00:00	20988	MQ	N676MQ	3318	DFW	LFT	1121
25	015 00:00:00	20988	MQ	N819MQ	3337	DFW	LEX	932
26	015 00:00:00	20988	MQ	N648MQ	3414	DFW	AMA	1131
27	015 00:00:00	20988	MQ	N532MQ	3414	DFW	AMA	1129
28	015 00:00:00	20988	MQ	N539MQ	3395	MEM	ORD	1050
29	015 00:00:00	20988	MQ	N900MQ	3437	ORD	TVC	1041
30	015 00:00:00	20988	MQ	N620MQ	3489	ORD	FSD	995
31	015 00:00:00	20988	MQ	N639MQ	3624	ORD	CLE	1014
32	015 00:00:00	20988	MQ	N513MQ	3613	ORD	ATL	957
33	015 00:00:00	20988	MQ	N524MQ	3679	DFW	MAF	1131
34	015 00:00:00	20988	MQ	N689MQ	3632	HSV	DFW	1014
35	015 00:00:00	20988	MQ	N524MQ	3679	DFW	MAF	1131
36	015 00:00:00	20988	MQ	N819MQ	3337	DFW	LEX	932
37	015 00:00:00	20988	MQ	N620MQ	3212	ORD	TYS	943
38	015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	0

Outliers

Total 1 records (0% of total records) detected as outliers with 100% accuracy.
[Click here to download.](#)

COLUMN NAME #

AirTime #1

Upper bound: 392
Upper quartile: 143
Median: 93
Lower quartile: 60
Lower bound: 0

Distance #0

Upper bound: 2704
Upper quartile: 641
Median: 0
Lower quartile: 0
Lower bound: 0

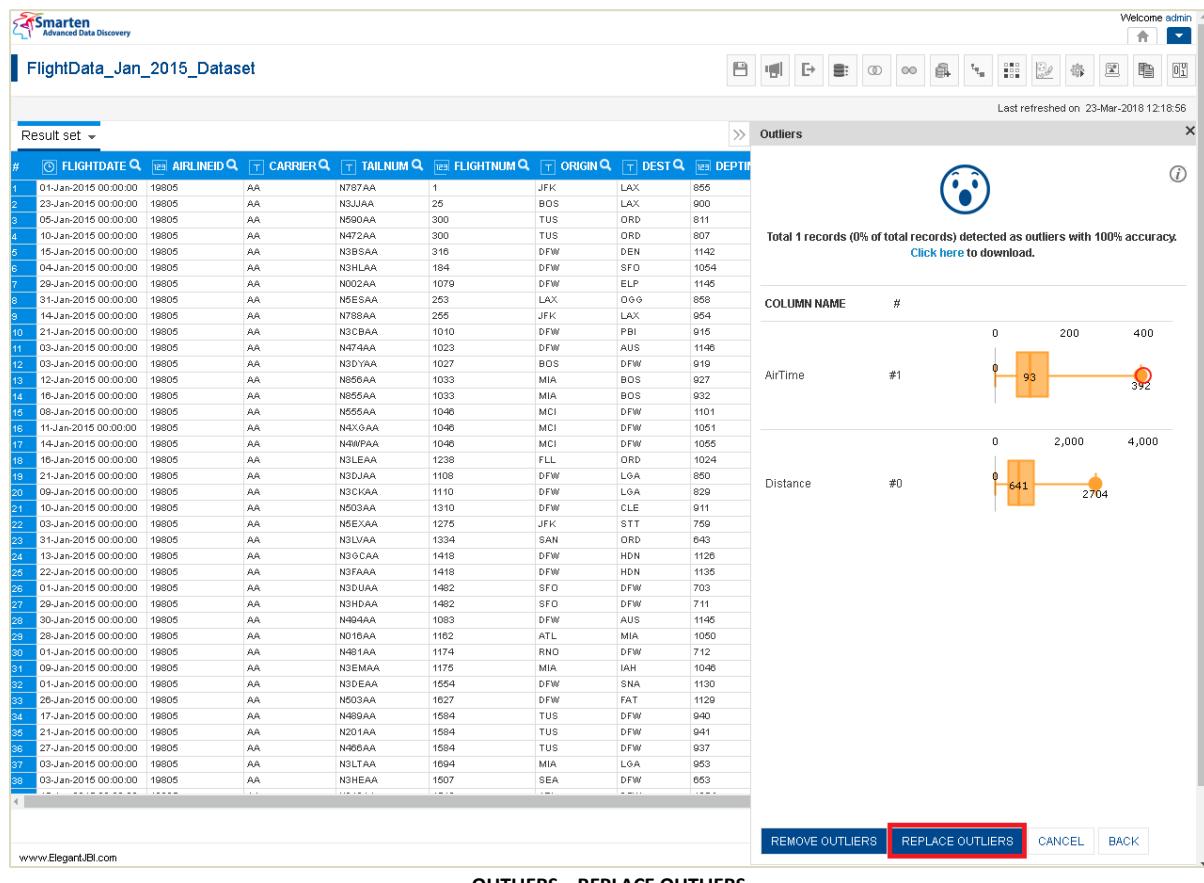
OUTLIER DETAILS

OUTLIERS—REMOVE OUTLIERS

REMOVE OUTLIERS REPLACE OUTLIERS CANCEL BACK

7.4.2 Replace Outliers

This option allows users to replace all outlier values of selected columns with the median value of the respective column.



8 Transform Data

Smarten SSDP allows users to transform data with the help of easy-to-use functions that change the data type and format and perform various operations on them.

8.1 For text columns

Shown below are the functions that can be used to transform string-type data.

8.1.1 Transform Upper case

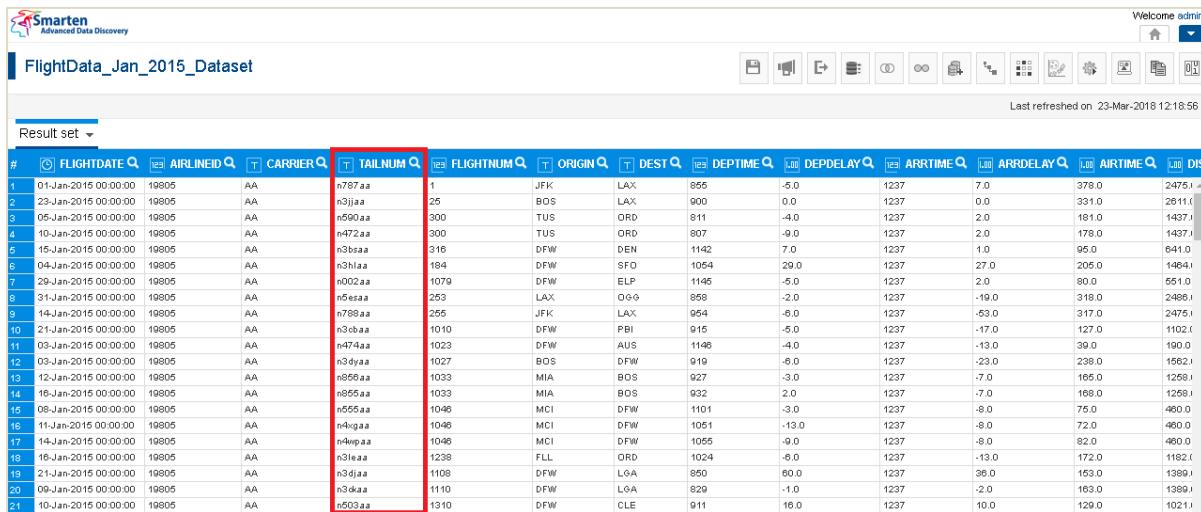
This function allows users to convert data from a target column containing lowercase characters into uppercase.

For example:

Original value	Transformed value
n3jjaa	N3JJAA
N3jjaa	N3JJAA
N3JJAA	N3JJAA

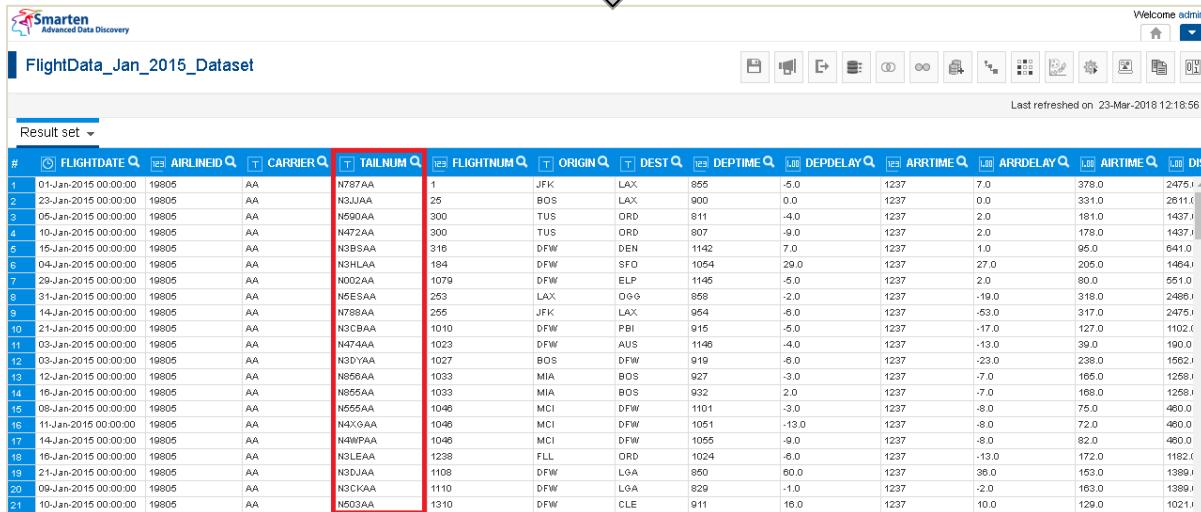
Shown below is the before and after scenario of “Upper case” for column “TAILNUM”:

Before:



#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DIS
1	01-Jan-2015 00:00:00	19805	AA	n787aa	1	JFK	LAX	855	-5.0	1237	7.0	379.0	24751.▲
2	23-Jan-2015 00:00:00	19805	AA	N3JJAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2611.▲
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.▲
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.▲
5	15-Jan-2015 00:00:00	19805	AA	N3B3AA	316	DFW	DEN	1142	7.0	1237	1.0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	AA	n3h1aa	184	DFW	SFO	1054	29.0	1237	27.0	205.0	1464.▲
7	20-Jan-2015 00:00:00	19805	AA	n002aa	1079	DFW	ELP	1145	-5.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	n5e5aa	253	LAX	OGG	858	-2.0	1237	-19.0	318.0	2486.▲
9	14-Jan-2015 00:00:00	19805	AA	n789aa	255	JFK	LAX	954	-6.0	1237	-63.0	317.0	2475.▲
10	21-Jan-2015 00:00:00	19805	AA	n3ebaa	1010	DFW	PBI	915	-5.0	1237	-17.0	127.0	1102.▲
11	03-Jan-2015 00:00:00	19805	AA	n474aa	1023	DFW	AUS	1148	-4.0	1237	-13.0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	AA	n3dyaa	1027	BOS	DFW	919	-6.0	1237	-23.0	238.0	1562.▲
13	12-Jan-2015 00:00:00	19805	AA	n059aa	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1269.▲
14	16-Jan-2015 00:00:00	19805	AA	n055aa	1033	MIA	BOS	932	2.0	1237	-7.0	168.0	1269.▲
15	08-Jan-2015 00:00:00	19805	AA	n555aa	1046	MCI	DFW	1101	-3.0	1237	-8.0	75.0	460.0
16	11-Jan-2015 00:00:00	19805	AA	n4kgaa	1046	MCI	DFW	1051	-13.0	1237	-8.0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	AA	n4wipaa	1048	MCI	DFW	1055	-9.0	1237	-8.0	82.0	460.0
18	16-Jan-2015 00:00:00	19805	AA	n3teaa	1238	FLL	ORD	1024	-6.0	1237	-13.0	172.0	1182.▲
19	21-Jan-2015 00:00:00	19805	AA	n3dja	1108	DFW	LGA	850	60.0	1237	36.0	153.0	1369.▲
20	09-Jan-2015 00:00:00	19805	AA	n3okaa	1110	DFW	LGA	829	-1.0	1237	-2.0	163.0	1369.▲
21	10-Jan-2015 00:00:00	19805	AA	n503aa	1310	DFW	CLE	911	16.0	1237	10.0	129.0	1021.▲

After:



#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DIS
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	379.0	24751.▲
2	23-Jan-2015 00:00:00	19805	AA	N3JJAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2611.▲
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.▲
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.▲
5	15-Jan-2015 00:00:00	19805	AA	N3B3AA	316	DFW	DEN	1142	7.0	1237	1.0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	AA	N3H1AA	184	DFW	SFO	1054	29.0	1237	27.0	205.0	1464.▲
7	20-Jan-2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145	-5.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	N5E5AA	253	LAX	OGG	858	-2.0	1237	-19.0	318.0	2486.▲
9	14-Jan-2015 00:00:00	19805	AA	N789AA	255	JFK	LAX	954	-6.0	1237	-63.0	317.0	2475.▲
10	21-Jan-2015 00:00:00	19805	AA	N3CBAA	1010	DFW	PBI	915	-5.0	1237	-17.0	127.0	1102.▲
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1148	-4.0	1237	-13.0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919	-6.0	1237	-23.0	238.0	1562.▲
13	12-Jan-2015 00:00:00	19805	AA	N059AA	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1269.▲
14	16-Jan-2015 00:00:00	19805	AA	N055AA	1033	MIA	BOS	932	2.0	1237	-7.0	168.0	1269.▲
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3.0	1237	-8.0	75.0	460.0
16	11-Jan-2015 00:00:00	19805	AA	N4KGAA	1046	MCI	DFW	1051	-13.0	1237	-8.0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	AA	N4WIPEA	1048	MCI	DFW	1055	-9.0	1237	-8.0	82.0	460.0
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6.0	1237	-13.0	172.0	1182.▲
19	21-Jan-2015 00:00:00	19805	AA	N3DJA	1108	DFW	LGA	850	60.0	1237	36.0	153.0	1369.▲
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1.0	1237	-2.0	163.0	1369.▲
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16.0	1237	10.0	129.0	1021.▲

TRANSFORM UPPER CASE

8.1.2 Transform Lower case

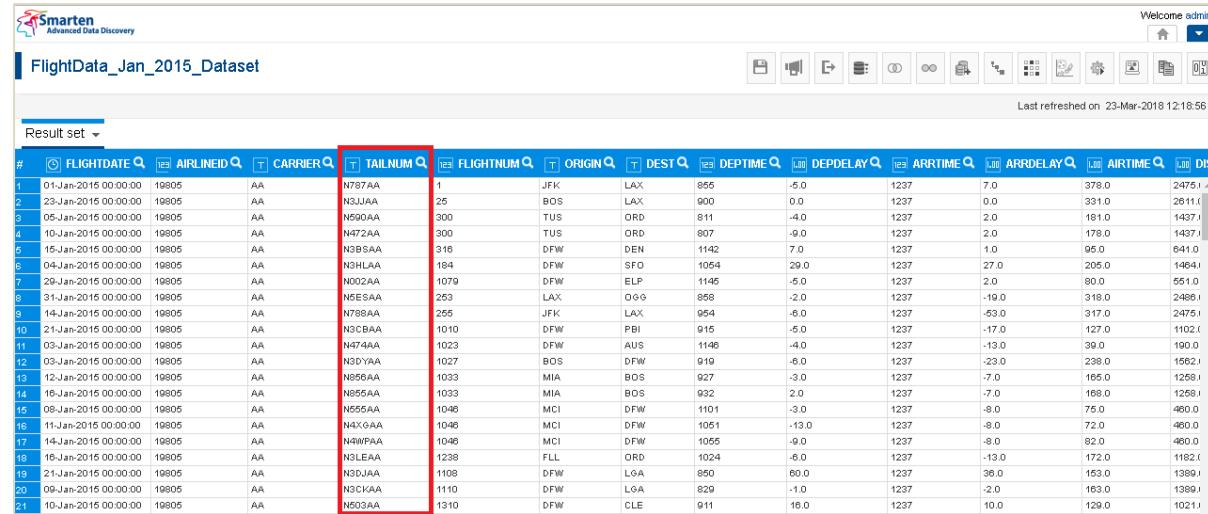
This function allows users to convert data from a target column containing uppercase characters into lowercase.

For example:

Original value	Transformed value
N3JJAA	n3jjaa
n3JJaa	n3jjaa
n3jjaa	n3jjaa

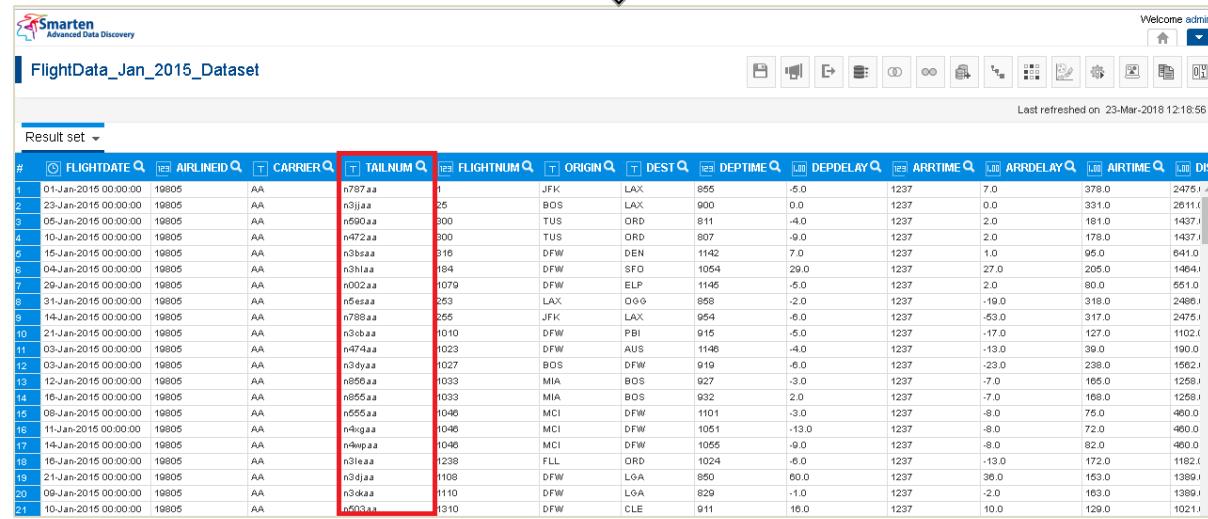
Shown below is the before and after scenario of “Lower case” for column “TAILNUM”:

Before:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DIS
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5:0	1237	7:0	379:0	24751:+
2	23-Jan-2015 00:00:00	19805	AA	N3JJA	25	BOS	LAX	900	0:0	1237	0:0	331:0	2611:+
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4:0	1237	2:0	181:0	1437:1
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9:0	1237	2:0	179:0	1437:1
5	15-Jan-2015 00:00:00	19805	AA	N3BAA	316	DFW	DEN	1142	7:0	1237	1:0	95:0	641:0
6	04-Jan-2015 00:00:00	19805	AA	N3HAA	184	DFW	SFO	1054	29:0	1237	27:0	205:0	1464:1
7	29-Jan-2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145	-6:0	1237	2:0	80:0	551:0
8	31-Jan-2015 00:00:00	19805	AA	N2EAA	253	LAX	Ogg	858	-2:0	1237	-19:0	318:0	2486:1
9	14-Jan-2015 00:00:00	19805	AA	N785AA	255	JFK	LAX	954	-6:0	1237	-53:0	317:0	2475:1
10	21-Jan-2015 00:00:00	19805	AA	N3CBA	1010	DFW	PBI	915	-5:0	1237	-17:0	127:0	1102:1
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4:0	1237	-13:0	39:0	190:0
12	03-Jan-2015 00:00:00	19805	AA	N3DAA	1027	BOS	DFW	919	-6:0	1237	-23:0	239:0	1562:1
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3:0	1237	-7:0	165:0	1269:1
14	16-Jan-2015 00:00:00	19805	AA	N955AA	1033	MIA	BOS	932	2:0	1237	-7:0	168:0	1269:1
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3:0	1237	-8:0	75:0	460:0
16	11-Jan-2015 00:00:00	19805	AA	N4XGAA	1046	MCI	DFW	1051	-13:0	1237	-8:0	72:0	460:0
17	14-Jan-2015 00:00:00	19805	AA	N4WPA	1046	MCI	DFW	1055	-9:0	1237	-8:0	82:0	460:0
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6:0	1237	-13:0	172:0	1182:1
19	21-Jan-2015 00:00:00	19805	AA	N3DJA	1108	DFW	LGA	850	60:0	1237	36:0	163:0	1389:1
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1:0	1237	-2:0	163:0	1389:1
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16:0	1237	10:0	129:0	1021:1

After:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DIS
1	01-Jan-2015 00:00:00	19805	AA	n787aa	1	JFK	LAX	855	-5:0	1237	7:0	379:0	24751:+
2	23-Jan-2015 00:00:00	19805	AA	n3jja	25	BOS	LAX	900	0:0	1237	0:0	331:0	2611:+
3	05-Jan-2015 00:00:00	19805	AA	n590aa	300	TUS	ORD	811	-4:0	1237	2:0	181:0	1437:1
4	10-Jan-2015 00:00:00	19805	AA	n472aa	300	TUS	ORD	807	-9:0	1237	2:0	179:0	1437:1
5	15-Jan-2015 00:00:00	19805	AA	n3baa	316	DFW	DEN	1142	7:0	1237	1:0	95:0	641:0
6	04-Jan-2015 00:00:00	19805	AA	n3haa	184	DFW	SFO	1054	29:0	1237	27:0	205:0	1464:1
7	29-Jan-2015 00:00:00	19805	AA	n002aa	1079	DFW	ELP	1145	-6:0	1237	2:0	80:0	551:0
8	31-Jan-2015 00:00:00	19805	AA	n2eaa	253	LAX	Ogg	858	-2:0	1237	-19:0	318:0	2486:1
9	14-Jan-2015 00:00:00	19805	AA	n785aa	255	JFK	LAX	954	-6:0	1237	-53:0	317:0	2475:1
10	21-Jan-2015 00:00:00	19805	AA	n3cba	1010	DFW	PBI	915	-5:0	1237	-17:0	127:0	1102:1
11	03-Jan-2015 00:00:00	19805	AA	n474aa	1023	DFW	AUS	1146	-4:0	1237	-13:0	39:0	190:0
12	03-Jan-2015 00:00:00	19805	AA	n3daa	1027	BOS	DFW	919	-6:0	1237	-23:0	239:0	1562:1
13	12-Jan-2015 00:00:00	19805	AA	n856aa	1033	MIA	BOS	927	-3:0	1237	-7:0	165:0	1269:1
14	16-Jan-2015 00:00:00	19805	AA	n955aa	1033	MIA	BOS	932	2:0	1237	-7:0	168:0	1269:1
15	08-Jan-2015 00:00:00	19805	AA	n555aa	1046	MCI	DFW	1101	-3:0	1237	-8:0	75:0	460:0
16	11-Jan-2015 00:00:00	19805	AA	n4xgaa	1046	MCI	DFW	1051	-13:0	1237	-8:0	72:0	460:0
17	14-Jan-2015 00:00:00	19805	AA	n4wpa	1046	MCI	DFW	1055	-9:0	1237	-8:0	82:0	460:0
18	16-Jan-2015 00:00:00	19805	AA	n3leaa	1238	FLL	ORD	1024	-6:0	1237	-13:0	172:0	1182:1
19	21-Jan-2015 00:00:00	19805	AA	n3djaa	1108	DFW	LGA	850	60:0	1237	36:0	163:0	1389:1
20	09-Jan-2015 00:00:00	19805	AA	n3ckaa	1110	DFW	LGA	829	-1:0	1237	-2:0	163:0	1389:1
21	10-Jan-2015 00:00:00	19805	AA	n503aa	1310	DFW	CLE	911	16:0	1237	10:0	129:0	1021:1

TRANSFORM LOWER CASE

8.1.3 Transform Capitalise

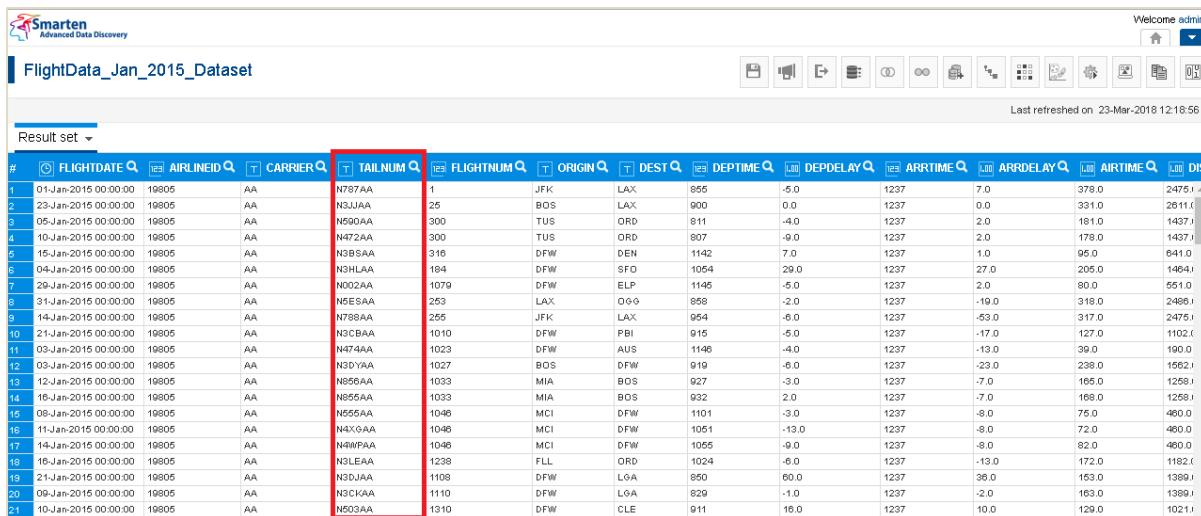
This function allows users to capitalize the data of a target column.

For example:

Original value	Transformed value
N3JJAA	N3jjaa
n3JJaa	N3jjaa
N3JJaa	N3jjaa

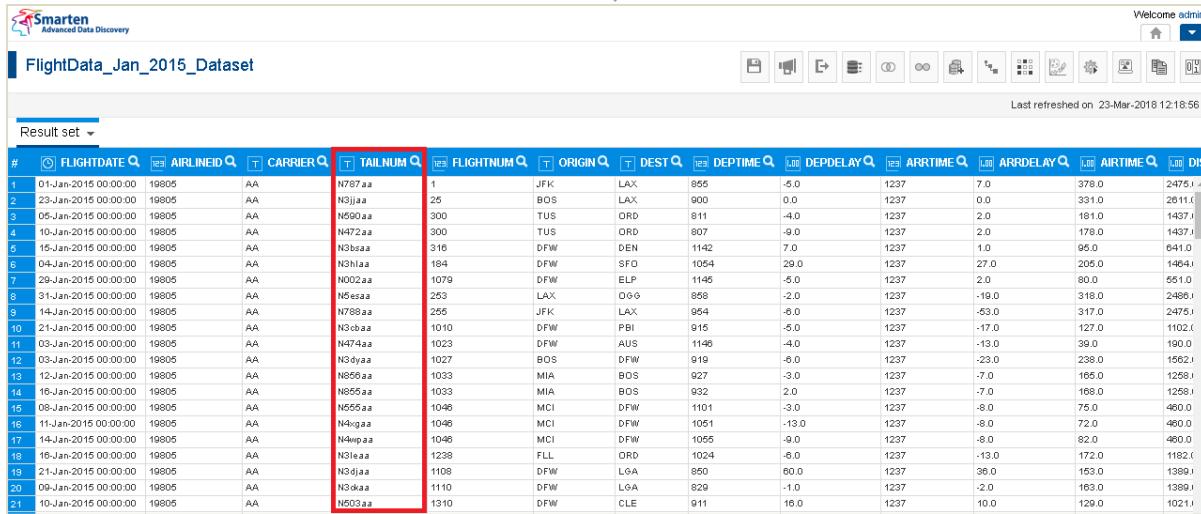
Shown below is the before and after scenario of “Capitalise” for column “TAILNUM”:

Before:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARTRIME	ARRDELAY	AIRTIME	DIS
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	378.0	2475.1
2	23-Jan-2015 00:00:00	19805	AA	N93JAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2611.1
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.1
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.1
5	15-Jan-2015 00:00:00	19805	AA	N3BSAA	316	DFW	DEN	1142	7.0	1237	1.0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	AA	N3HLLA	184	DFW	SFO	1054	29.0	1237	27.0	205.0	1464.1
7	29-Jan-2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145	-5.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	N56SA	263	LAX	OOG	858	-2.0	1237	-19.0	318.0	2486.1
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-6.0	1237	-53.0	317.0	2475.1
10	21-Jan-2015 00:00:00	19805	AA	N5CBA	1010	DFW	PBI	915	-8.0	1237	-17.0	127.0	1102.1
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4.0	1237	-13.0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	AA	N9DYAA	1027	BOS	DFW	919	-6.0	1237	-23.0	238.0	1562.1
13	12-Jan-2015 00:00:00	19805	AA	N956AA	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1259.1
14	16-Jan-2015 00:00:00	19805	AA	N955AA	1033	MIA	BOS	932	2.0	1237	-7.0	165.0	1259.1
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3.0	1237	-8.0	75.0	460.0
16	11-Jan-2015 00:00:00	19805	AA	N4XGAA	1048	MCI	DFW	1051	-13.0	1237	-8.0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	AA	N9WMPAA	1046	MCI	DFW	1055	-9.0	1237	-8.0	82.0	460.0
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6.0	1237	-13.0	172.0	1162.1
19	21-Jan-2015 00:00:00	19805	AA	N9DJAA	1108	DFW	LGA	850	60.0	1237	36.0	153.0	1389.1
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1.0	1237	-2.0	163.0	1389.1
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16.0	1237	10.0	129.0	1021.1

After:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARTRIME	ARRDELAY	AIRTIME	DIS
1	01-Jan-2015 00:00:00	19805	AA	N787aa	1	JFK	LAX	855	-5.0	1237	7.0	378.0	2475.1
2	23-Jan-2015 00:00:00	19805	AA	N93Jaa	25	BOS	LAX	900	0.0	1237	0.0	331.0	2611.1
3	05-Jan-2015 00:00:00	19805	AA	N590aa	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.1
4	10-Jan-2015 00:00:00	19805	AA	N472aa	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.1
5	15-Jan-2015 00:00:00	19805	AA	N3bsaa	316	DFW	DEN	1142	7.0	1237	1.0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	AA	N3hlla	184	DFW	SFO	1054	29.0	1237	27.0	205.0	1464.1
7	29-Jan-2015 00:00:00	19805	AA	N002aa	1079	DFW	elp	1145	-5.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	N56sa	263	LAX	OOG	858	-2.0	1237	-19.0	318.0	2486.1
9	14-Jan-2015 00:00:00	19805	AA	N788aa	255	JFK	LAX	954	-6.0	1237	-53.0	317.0	2475.1
10	21-Jan-2015 00:00:00	19805	AA	N3oba	1010	DFW	pbi	915	-8.0	1237	-17.0	127.0	1102.1
11	03-Jan-2015 00:00:00	19805	AA	N474aa	1023	DFW	AUS	1146	-4.0	1237	-13.0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	AA	N3dyaa	1027	BOS	DFW	919	-6.0	1237	-23.0	238.0	1562.1
13	12-Jan-2015 00:00:00	19805	AA	N859aa	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1259.1
14	16-Jan-2015 00:00:00	19805	AA	N855aa	1033	MIA	BOS	932	2.0	1237	-7.0	165.0	1259.1
15	08-Jan-2015 00:00:00	19805	AA	N655aa	1046	MCI	DFW	1101	-3.0	1237	-8.0	75.0	460.0
16	11-Jan-2015 00:00:00	19805	AA	N4qgaa	1046	MCI	DFW	1051	-13.0	1237	-8.0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	AA	N4eqaaa	1046	MCI	DFW	1055	-9.0	1237	-8.0	82.0	460.0
18	16-Jan-2015 00:00:00	19805	AA	N3leaa	1238	FLL	ORD	1024	-6.0	1237	-13.0	172.0	1162.1
19	21-Jan-2015 00:00:00	19805	AA	N3djaa	1108	DFW	LGA	850	60.0	1237	36.0	153.0	1389.1
20	09-Jan-2015 00:00:00	19805	AA	N3kcaa	1110	DFW	LGA	829	-1.0	1237	-2.0	163.0	1389.1
21	10-Jan-2015 00:00:00	19805	AA	N503aa	1310	DFW	CLE	911	16.0	1237	10.0	129.0	1021.1

TRANSFORM CAPITALISE

8.1.4 Transform Advanced functions

Users can perform various operations to transform string-type data by using an extensive collection of advanced functions.

Shown below is the list of all Data Operations:

Operation	Description	Example
concat("string", "string")	Returns a resulting string after concatenating specified strings Argument 1: The text that has to be concatenate with argument 2 Argument 2: The text that has to be concatenate with argument 1 Returns: A string	Argument 1 = "N787" Argument 2 = "AA" Returns "N787AA"

isNull(object)	Determines if the argument is NULL	Argument 1 = "N787AA" Returns "false"
	Argument 1: The object that is to be checked	Argument 1 = "NULL" Returns "true"
	Returns: A boolean	
isNumber("string")	Determines if the specified string contains a number	Argument 1 = "N787AA" Returns "false"
	Argument 1: The string that is to be checked	Argument 1 = "787" Returns "true"
	Returns: A boolean	
left("string", i)	Returns a specified number of characters from a string starting with the first character	Argument 1 = "N787AA" Argument 2 = 2 Returns "N7"
	Argument 1: The text from which the partial words are to be returned	Argument 1 = "N787AA"
	Argument 2: The number of characters to be extracted from the beginning of the text	Argument 2 = 8 Returns "N787AA"
leftTrim("string")	Returns a copy of a specified string with leading blanks removed	Argument 1 = " 87AA" Returns "87AA"
	Argument 1: The text for which blank spaces are to be removed from left	Argument 1 = "87AA " Returns "87AA "
	Returns: A string	Argument 1 = "87AA " Returns "87AA "
lpad("string", i, "string")	Returns a string with left-pad of specified length	Argument 1 = "N787AA" Argument 2 = 10 Argument 3 = "X" Returns "XXXXN787AA"
	Argument 1: The text in which left-pad is to be added	Argument 1 = "N787AA"
	Argument 2: The length of string to be returned	Argument 2 = 3 Argument 3 = "X" Returns "N78"
match("string", "string")	Argument 3: The text that is to be padded in specified string	Argument 1 = "N787AA" Argument 2 = 6 Argument 3 = "X" Returns "N787AA"
	Returns: A string	
match("string", "string")	Returns a determination of whether or not a string contains a particular pattern of characters	Argument 1 = "AA" Argument 2 = "N787AA" Returns 1
	Argument 1: The text that has to be searched in argument 2	Argument 1 = "aa" Argument 2 = "N787AA" Returns 0
	Argument 2: The text in which argument 1 has to be searched	Argument 1 = "AB" Argument 2 = "N787AA" Returns 0

	Returns: A number	
reverse("string")	<p>Reverses the order of characters in a string</p> <p>Argument 1: The text that needs to be reversed</p> <p>Returns: A string</p>	<p>Argument 1 = "N208WN"</p> <p>Returns "NW802N"</p>
right("string", i)	<p>Returns the specified number of characters from the end of a specified string</p> <p>Argument 1: The text from which the specified number of characters should be returned from the end</p> <p>Argument 2: The number of characters to be returned from the string</p> <p>Returns: A string</p>	<p>Argument 1 = "N208WN"</p> <p>Argument 2 = 3</p> <p>Returns "8WN"</p> <hr/> <p>Argument 1 = "N208WN"</p> <p>Argument 2 = 8</p> <p>Returns "N208WN"</p>
rightTrim("string")	<p>Returns a copy of the specified string with trailing blanks removed</p> <p>Argument 1: The text from which extra spaces have to be removed from the right</p> <p>Returns: A string</p>	<p>Argument 1 = "N208 "</p> <p>Returns "N208"</p> <hr/> <p>Argument 1 = " 08WN"</p> <p>Returns " 08WN"</p> <hr/> <p>Argument 1 = " 208W "</p> <p>Returns " 208W"</p>
rpad("string", i, "string")	<p>Returns a string with right-pad of specified length</p> <p>Argument 1: The text in which left-pad is to be added</p> <p>Argument 2: The length of string to be returned</p> <p>Argument 3: The text that is to be padded in specified string</p> <p>Returns: A string</p>	<p>Argument 1 = "N208WN"</p> <p>Argument 2 = 9</p> <p>Argument 3 = "XY"</p> <p>Returns "N208WNXYX"</p> <hr/> <p>Argument 1 = "N208WN"</p> <p>Argument 2 = 6</p> <p>Argument 3 = "XY"</p> <p>Returns "N208WN"</p> <hr/> <p>Argument 1 = "N208WN"</p> <p>Argument 2 = 4</p> <p>Argument 3 = "XY"</p> <p>Returns "N208"</p>
substring("string", i, i)	<p>Returns a string containing a character copied (starting at a specified position and ending at a specified position) from a specified string</p> <p>Argument 1: The text from which the characters have to be copied</p> <p>Argument 2: Starting position from which the characters have to be copied considering the position of the first character at 0</p> <p>Argument 3: Ending position up to which the</p>	<p>Argument 1 = "N208WN"</p> <p>Argument 2 = 2</p> <p>Argument 3 = 4</p> <p>Returns "08"</p> <hr/> <p>Argument 1 = "N208WN"</p> <p>Argument 2 = 2</p> <p>Argument 3 = 6</p> <p>Returns "08WN"</p>

	characters in the text are to be copied Returns: A string	
trim("string")	Returns a string with leading and trailing blanks removed Argument 1: The text from which the extra spaces are to be removed Returns: A string	Argument 1 = " 08WN" Returns "08WN" Argument 1 = "N208 " Returns "N208" Argument 1 = "208W " Returns "208W"

8.2 For numerical columns

Shown below are the functions that can be used to transform numeric data.

8.2.1 Transform Advanced functions

Users can perform various operations to transform numeric data by using an extensive collection of advanced functions.

Shown below is the list of all Data Operations:

Operation	Description	Example
abs(number)	Return absolute value of a number, a number without its sign Argument 1: The number for which absolute value is required Returns: A number	Argument 1 = 32 Returns 32 Argument 1 = 67.98 Returns 67.98 Argument 1 = -23 Returns 23
ceil(d)	Returns the smallest whole number that is greater than or equal to a specified number Argument 1: The number that has to be rounded up Returns: A number	Argument 1 = 26 Returns 26 Argument 1 = 26.7 Returns 27 Argument 1 = -26.7 Returns -26
divide(number, number)	Returns the quotient of two numbers Argument 1: Dividend number that is to be divided by the divisor Argument 1: Divisor number Returns: A number	Argument 1 = 551 Argument 2 = 2 Returns 275.50 Argument 1 = -450 Argument 2 = 3 Returns -150.00 Argument 1 = 551 Argument 2 = 0 Returns NULL Argument 1 = 0 Argument 2 = 551 Returns 0.00
exp(d)	Returns the exponential value of a number	Argument 1 = 1145

	Argument 1: The exponent applied to base e Returns: A number	Returns "Infinity" Argument 1 = 12 Returns 162754.79 Argument 1 = -25 Returns 0.00
fact(i)	Returns the factorial of a number Argument 1: The number for which factorial is to be calculated Returns: A number	Argument 1 = 7 Returns 5040 Argument 1 = -5 Returns NULL
floor(d)	Returns the largest whole number that is smaller than or equal to a specified number Argument 1: The number to be rounded down Returns: A number	Argument 1 = 26 Returns 26 Argument 1 = 26.7 Returns 26 Argument 1 = -26.7 Returns -27
log(d)	Returns natural logarithm (base e) of a number Argument 1: A value greater than 0 for which logarithm is to be calculated Returns: A number	Argument 1 = 551 Returns 6.31 Argument 1 = -551 Returns NULL Argument 1 = 551.45 Returns 6.31
logTen(d)	Returns decimal logarithm (base 10) of a number Argument 1: The value greater than 0 for which logarithm is to be calculated Returns: A number	Argument 1 = 551 Returns 2.74 Argument 1 = -551 Returns NULL Argument 1 = 551.45 Returns 2.74
max(number, number)	Returns larger of two numbers Argument 1: First number to find out if it is larger than the second number Argument 2: Second number to find out if it is larger than the first number Returns: A number	Argument 1 = 198 Argument 2 = 1660 Returns 1660.00 Argument 1 = 198 Argument 2 = -1660 Returns 198.00
min(number, number)	Returns smaller of two numbers Argument 1: First number to find out if it is smaller than the second number Argument 2: Second number to find out if it is smaller than the first number Returns: A number	Argument 1 = 198 Argument 2 = 1660 Returns 198.00 Argument 1 = 198 Argument 2 = -1660 Returns -1660.00
minus(number, number)	Returns the subtraction of two numbers	Argument 1 = -5 Argument 2 = 1237

	<p>Argument 1: A base number Argument 2: A number that is to be subtracted from the base number Returns: A number</p>	Returns -1242.00 ----- Argument 1 = 1237 Argument 2 = 29 Returns 1208.00
mod(number, number)	<p>Returns modulus of two numbers Argument 1: Dividend: The number to be divided Argument 2: Divisor: The number by which the dividend has to be divided Returns: A number</p>	Argument 1 = 460 Argument 2 = 72 Returns 28.00 ----- Argument 1 = -460 Argument 2 = 72 Returns 44.00 ----- Argument 1 = 460 Argument 2 = -72 Returns 28.00 ----- Argument 1 = -460 Argument 2 = -72 Returns -28.00
multiply(number, number)	<p>Returns the product of two numbers Argument 1: A base number Argument 2: A number that is to be multiplied by the base number Returns: A number</p>	Argument 1 = 460 Argument 2 = 72 Returns 33120.00 ----- Argument 1 = -460 Argument 2 = 72 Returns -33120.00 ----- Argument 1 = -460 Argument 2 = -72 Returns 33120.00
pi(d)	<p>Returns pi times a number Argument 1: The number Returns: A number</p>	Argument 1 = 641 Returns 2013.76 ----- Argument 1 = -3 Returns -9.42
plus(number, number)	<p>Returns the sum of two numbers Argument 1: A base number Argument 2: A number that is to be added to the base number Returns: A number</p>	Argument 1 = 460 Argument 2 = 72 Returns 532.00 ----- Argument 1 = 460 Argument 2 = -72 Returns 388.00 ----- Argument 1 = -460 Argument 2 = -72 Returns -532.00
round(d, i)	<p>Returns the number rounded to a specified number of decimal places Argument 1: The number to be rounded Argument 2: The number of places to which the number is to be rounded</p>	Argument 1 = 12.356 Argument 2 = 1 Returns 12.40 ----- Argument 1 = -12.356 Argument 2 = 1 Returns -12.40

	Returns: A number	Argument 1 = 12.356 Argument 2 = 2 Returns 12.36
		Argument 1 = 12.356 Argument 2 = 3 Returns 12.35
sign(d)	Returns a number (-1, 0, or 1) indicating the sign of a number Argument 1: The number for which the algebraic sign is to be determined Returns: A number	Argument 1 = -5 Returns -1 Argument 1 = 0 Returns 0 Argument 1 = 29 Returns 1
sqrt(d)	Returns the square root of a number Argument 1: A positive value for which square root is to be calculated Returns: A number	Argument 1 = 100 Returns 10.00 Argument 1 = 588 Returns 24.24 Argument 1 = -588 Returns NaN (Not a number)

8.3 For datetime columns

Shown below are the functions that can be used to transform datetime data.

8.3.1 Transform—Make timepart zero

This function converts the timepart of the datetime type of column into zero.

For example:

Original value	Transformed value
01-Jan-2015 08:55:00	01-Jan-2015 00:00:00
23-Feb-2015 10:45:55	23-Feb-2015 00:00:00
30-Apr-2015 15:30:20	30-Apr-2015 00:00:00

Shown below is the before and after scenario of “Make timepart zero” for column “FLIGHTDATE”:

Before:

Flight_Data_Jan_2015_Dataset

Last refreshed on May 09, 2018 14:32:56

#	FLIGHTDATE Q	AIRLINEID Q	CARRIER Q	TAILNUM Q	FLIGHTNUM Q	ORIGIN Q	DEST Q	DEPETIME Q	DEPDELAY Q	ARRTIME Q	ARRDELAY Q	AIRTIME Q	DISTANCE
1	01 - Jan - 2015 06:55:00	19805	AA	N787AA		JFK	LAX	855	-5	1237	7	378	2475
2	23 - Jan - 2015 09:00:00	19805	AA	N3JAA		BOS	LAX	900	0	1237	0	331	2611
3	05 - Feb - 2015 08:11:00	19805	AA	N590AA		TUS	ORD	811	-4	1237	2	181	1437
4	10 - Jan - 2015 06:07:00	19805	AA	N472AA		TUS	ORD	807	-9	1237	2	178	1437
5	15 - Jan - 2015 11:42:00	19805	AA	N3BSAA		DFW	DEN	1142	7	1237	1	95	641
6	04 - Feb - 2015 10:54:00	19805	AA	N3HLAA		DFW	SFO	1054	29	1237	27	205	1464
7	29 - Jan - 2015 11:45:00	19805	AA	N002AA		DFW	ELP	1145	-5	1237	2	80	551
8	31 - Jan - 2015 05:58:00	19805	AA	N5E5AA		LAX	OGG	858	-2	1237	-19	318	2486
9	14 - Mar - 2015 09:54:00	19805	AA	N788AA		JFK	LAX	954	-6	1237	-53	317	2475
10	21 - Mar - 2015 09:15:00	19805	AA	N3CBAA		DFW	PBI	915	-5	1237	-17	127	1102
11	03 - Jan - 2015 11:46:00	19805	AA	N474AA		DFW	AUS	1146	-4	1237	-13	39	190
12	03 - Mar - 2015 09:19:00	19805	AA	N3DYAA		BOS	DFW	919	-6	1237	-23	238	1562
13	12 - Jan - 2015 09:27:00	19805	AA	N565AA		MIA	BOS	927	-3	1237	-7	165	1258
14	16 - Jan - 2015 09:32:00	19805	AA	N855AA		MIA	BOS	932	2	1237	-7	168	1258
15	08 - Aug - 2015 11:01:00	19805	AA	N555AA		MCI	DFW	1101	-3	1237	-8	75	460
16	11 - Jan - 2015 01:51:00	19805	AA	N4XGAA		MCI	DFW	1051	-13	1237	-8	72	460
17	14 - Jan - 2015 10:59:00	19805	AA	N4WPAA		MCI	DFW	1055	-9	1237	-8	82	460
18	16 - Aug - 2015 10:40:00	19805	AA	N3LEAA		FLL	ORD	1024	-6	1237	-13	172	1182
19	21 - Jan - 2015 06:50:00	19805	AA	N3DJAA		DFW	LGA	850	60	1237	38	153	1389
20	09 - Jan - 2015 06:29:00	19805	AA	N3CKAA		DFW	LGA	829	-1	1237	-2	163	1389
21	10 - Dec - 2015 09:11:00	19805	AA	N503AA		DFW	CLE	911	16	1237	10	129	1021
22	03 - Dec - 2015 07:59:00	19805	AA	N5EXAA		JFK	STT	759	-1	1237	-18	190	1623
23	31 - Jan - 2015 08:43:00	19805	AA	N3LVAA		SAN	ORD	843	-1	1237	-13	205	1723

After:

Flight_Data_Jan_2015_Dataset

Last refreshed on May 09, 2018 14:32:56

Result Set

#	FLIGHTDATE Q	AIRLINEID Q	CARRIER Q	TAILNUM Q	FLIGHTNUM Q	ORIGIN Q	DEST Q	DEPETIME Q	DEPDELAY Q	ARRTIME Q	ARRDELAY Q	AIRTIME Q	DISTANCE
1	01 - Jan - 2015 00:00:00	19805	AA	N787AA		JFK	LAX	855	-5	1237	7	378	2475
2	23 - Jan - 2015 00:00:00	19805	AA	N3JAA		BOS	LAX	900	0	1237	0	331	2611
3	05 - Feb - 2015 00:00:00	19805	AA	N590AA		TUS	ORD	811	-4	1237	2	181	1437
4	10 - Jan - 2015 00:00:00	19805	AA	N472AA		TUS	ORD	807	-9	1237	2	178	1437
5	15 - Jan - 2015 00:00:00	19805	AA	N3BSAA		DFW	DEN	1142	7	1237	1	95	641
6	04 - Feb - 2015 00:00:00	19805	AA	N3HLAA		DFW	SFO	1054	29	1237	27	205	1464
7	29 - Jan - 2015 00:00:00	19805	AA	N002AA		DFW	ELP	1145	-5	1237	2	80	551
8	31 - Jan - 2015 00:00:00	19805	AA	N5E5AA		LAX	OGG	858	-2	1237	-19	318	2486
9	14 - Mar - 2015 00:00:00	19805	AA	N788AA		JFK	LAX	954	-6	1237	-53	317	2475
10	21 - Mar - 2015 00:00:00	19805	AA	N3CBAA		DFW	PBI	915	5	1237	-17	127	1102
11	03 - Jan - 2015 00:00:00	19805	AA	N474AA		DFW	AUS	1146	-4	1237	-13	39	190
12	03 - Mar - 2015 00:00:00	19805	AA	N3DYAA		BOS	DFW	919	-6	1237	-23	238	1562
13	12 - Jan - 2015 00:00:00	19805	AA	N855AA		MIA	BOS	927	-3	1237	-7	165	1258
14	16 - Jan - 2015 00:00:00	19805	AA	N555AA		MIA	BOS	932	2	1237	-7	168	1258
15	08 - Aug - 2015 00:00:00	19805	AA	N555AA		MCI	DFW	1101	-3	1237	-8	75	460
16	11 - Jan - 2015 00:00:00	19805	AA	N4XGAA		MCI	DFW	1051	-13	1237	-8	72	460
17	14 - Jan - 2015 00:00:00	19805	AA	N4WPAA		MCI	DFW	1055	-9	1237	-8	82	460
18	16 - Aug - 2015 00:00:00	19805	AA	N3LEAA		FLL	ORD	1024	-6	1237	-13	172	1182
19	21 - Jan - 2015 00:00:00	19805	AA	N3DJAA		DFW	LGA	850	60	1237	38	153	1389
20	09 - Jan - 2015 00:00:00	19805	AA	N3CKAA		DFW	LGA	829	-1	1237	-2	163	1389
21	10 - Dec - 2015 00:00:00	19805	AA	N503AA		DFW	CLE	911	16	1237	10	129	1021
22	03 - Dec - 2015 00:00:00	19805	AA	N5EXAA		JFK	STT	760	-1	1237	-18	190	1623
23	31 - Jan - 2015 00:00:00	19805	AA	N3LVAA						1237	-13	205	1723

TRANSFORM—MAKE TIMEPART ZERO

8.3.2 Transform First date of month

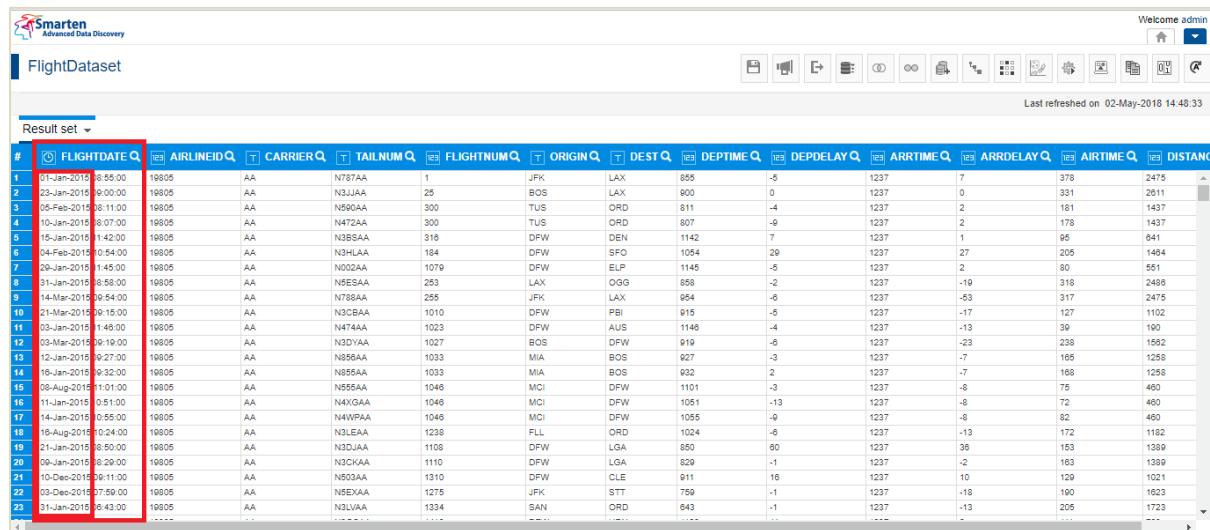
This function converts all dates into the first date of their respective month.

For example:

Original value	Transformed value
01-Jan-2015	01-Jan-2015
23-Feb-2015	01-Feb-2015
30-Apr-2015	01-Apr-2015

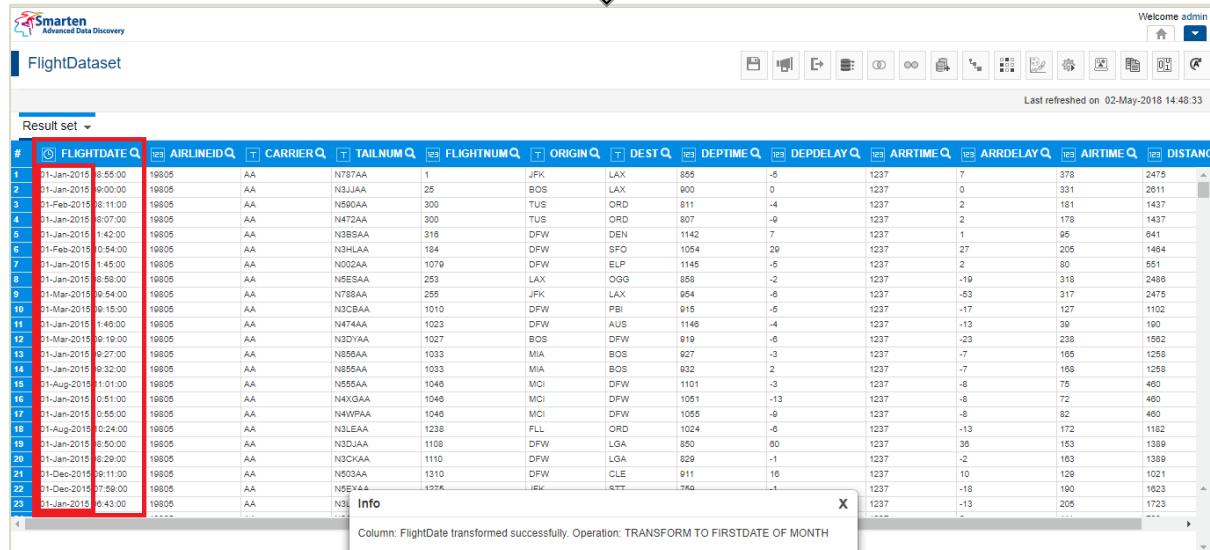
Shown below is the before and after scenario of “First date of month” for column “FLIGHTDATE”:

Before:



#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 08:55:00	19805	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475
2	23-Jan-2015 09:00:00	19805	AA	N3JJAA	25	BOS	LAX	900	0	1237	0	331	2611
3	05-Feb-2015 08:11:00	19805	AA	N590AA	300	TUS	ORD	811	-4	1237	2	181	1437
4	10-Jan-2015 08:07:00	19805	AA	N472AA	300	TUS	ORD	807	-9	1237	2	178	1437
5	15-Jan-2015 11:42:00	19805	AA	N3BSAA	316	DFW	DEN	1142	7	1237	1	95	641
6	04-Feb-2015 05:44:00	19805	AA	N3HLAA	184	DFW	SFO	1054	29	1237	27	205	1464
7	20-Jan-2015 11:45:00	19805	AA	N02AA	1079	DFW	ELP	1145	-5	1237	2	80	551
8	31-Jan-2015 08:58:00	19805	AA	N5ESAA	253	LAX	OOG	858	-2	1237	-19	318	2488
9	14-Mar-2015 09:54:00	19805	AA	N788AA	255	JFK	LAX	954	-8	1237	-53	317	2475
10	21-Mar-2015 09:15:00	19805	AA	N3CBA	1010	DFW	PBI	915	-5	1237	-17	127	1102
11	03-Jan-2015 11:49:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4	1237	-13	39	190
12	03-Mar-2015 09:19:00	19805	AA	N3DVA	1027	BOS	DFW	919	-8	1237	-23	238	1582
13	12-Jan-2015 09:27:00	19805	AA	N56AA	1033	MIA	BOS	927	-3	1237	-7	165	1258
14	16-Jan-2015 02:32:00	19805	AA	N856AA	1033	MIA	BOS	932	2	1237	-7	168	1258
15	08-Aug-2015 11:01:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3	1237	-8	75	460
16	11-Jan-2015 09:10:00	19805	AA	N4XGAA	1046	MCI	DFW	1051	-13	1237	-8	72	460
17	14-Jan-2015 05:55:00	19805	AA	N4WPAA	1046	MCI	DFW	1055	-9	1237	-8	82	460
18	16-Aug-2015 02:40:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-8	1237	-13	172	1182
19	21-Jan-2015 08:50:00	19805	AA	N3DJAA	1108	DFW	LGA	850	60	1237	38	153	1389
20	09-Jan-2015 08:29:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1	1237	-2	163	1389
21	10-Dec-2015 09:11:00	19805	AA	N503AA	1310	DFW	CLE	911	16	1237	10	129	1021
22	03-Dec-2015 07:59:00	19805	AA	N5EXAA	1275	JFK	STT	759	-1	1237	-18	190	1623
23	31-Jan-2015 08:43:00	19805	AA	N3LVAA	1334	SAN	ORD	643	-1	1237	-13	205	1723

After:



#	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 08:55:00	19805	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475
2	01-Jan-2015 09:00:00	19805	AA	N3JJAA	25	BOS	LAX	900	0	1237	0	331	2611
3	01-Feb-2015 08:11:00	19805	AA	N590AA	300	TUS	ORD	811	-4	1237	2	181	1437
4	01-Jan-2015 08:07:00	19805	AA	N472AA	300	TUS	ORD	807	-9	1237	2	178	1437
5	01-Jan-2015 11:42:00	19805	AA	N3BSAA	316	DFW	DEN	1142	7	1237	1	95	641
6	01-Feb-2015 05:44:00	19805	AA	N3HLAA	184	DFW	SFO	1054	29	1237	27	205	1464
7	01-Jan-2015 11:45:00	19805	AA	N02AA	1079	DFW	ELP	1145	-5	1237	2	80	551
8	01-Jan-2015 08:58:00	19805	AA	N5ESAA	253	LAX	OOG	858	-2	1237	-19	318	2488
9	01-Mar-2015 09:44:00	19805	AA	N788AA	255	JFK	LAX	954	-8	1237	-53	317	2475
10	01-Mar-2015 09:15:00	19805	AA	N3CBA	1010	DFW	PBI	915	-5	1237	-17	127	1102
11	01-Jan-2015 11:49:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4	1237	-13	39	190
12	01-Mar-2015 09:19:00	19805	AA	N3DJAA	1027	BOS	DFW	919	-8	1237	-23	238	1582
13	01-Jan-2015 09:27:00	19805	AA	N856AA	1033	MIA	BOS	927	-3	1237	-7	165	1258
14	01-Jan-2015 09:32:00	19805	AA	N856AA	1033	MIA	BOS	932	2	1237	-7	168	1258
15	01-Aug-2015 11:01:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3	1237	-8	75	460
16	01-Jan-2015 05:10:00	19805	AA	N4XGAA	1046	MCI	DFW	1051	-13	1237	-8	72	460
17	01-Jan-2015 05:55:00	19805	AA	N4WPAA	1046	MCI	DFW	1055	-9	1237	-8	82	460
18	01-Aug-2015 02:40:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-8	1237	-13	172	1182
19	01-Jan-2015 08:50:00	19805	AA	N3DJAA	1108	DFW	LGA	850	60	1237	38	153	1389
20	01-Jan-2015 08:29:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1	1237	-2	163	1389
21	01-Dec-2015 09:11:00	19805	AA	N503AA	1310	DFW	CLE	911	16	1237	10	129	1021
22	01-Dec-2015 07:59:00	19805	AA	N5EXAA	1275	JFK	STT	740	-1	1237	-18	190	1623
23	01-Jan-2015 08:43:00	19805	AA	N3LVAA	1334	SAN	ORD	643	-1	1237	-13	205	1723

TRANSFORM—FIRST DATE OF MONTH

8.3.3 Transform Last date of month

This function converts all dates into the last date of their respective month.

For example:

Original value	Transformed value
01-Jan-2015	31-Jan-2015
23-Feb-2015	28-Feb-2015
30-Apr-2015	30-Apr-2015

Shown below is the before and after scenario of “Last date of month” for column “FLIGHTDATE”:

Before:

#	FlightDate	AirlineID	Carrier	TailNum	FlightNum	Origin	Dest	DepartureTime	Delay	ArrivalTime	ArrDelay	Airtime	Distance
1	01-Jan-2015 08:55:00	19805	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475
2	23-Jan-2015 09:00:00	19805	AA	N3JAA	25	BOS	LAX	900	0	1237	0	331	2611
3	05-Feb-2015 08:11:00	19805	AA	N590AA	300	TUS	ORD	811	-4	1237	2	181	1437
4	10-Jan-2015 08:07:00	19805	AA	N472AA	300	TUS	ORD	807	-9	1237	2	178	1437
5	15-Jan-2015 11:42:00	19805	AA	N3BSAA	316	DFW	DEN	1142	7	1237	1	95	641
6	04-Feb-2015 09:54:00	19805	AA	N3HLAA	184	DFW	SFO	1054	29	1237	27	205	1464
7	20-Jan-2015 11:45:00	19805	AA	N02AA	1079	DFW	ELP	1145	-5	1237	2	80	551
8	31-Jan-2015 08:58:00	19805	AA	N5E5AA	253	LAX	OOG	858	-2	1237	-19	318	2486
9	14-Mar-2015 09:54:00	19805	AA	N788AA	255	JFK	LAX	954	-8	1237	-53	317	2475
10	21-Mar-2015 09:15:00	19805	AA	N3CBA	1010	DFW	PBI	915	-5	1237	-17	127	1102
11	03-Jan-2015 11:49:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4	1237	-13	39	190
12	03-Mar-2015 09:19:00	19805	AA	N3DVA	1027	BOS	DFW	919	-8	1237	-23	238	1582
13	12-Jan-2015 09:27:00	19805	AA	N565AA	1033	MIA	BOS	927	-3	1237	-7	165	1258
14	16-Jan-2015 09:32:00	19805	AA	N565AA	1033	MIA	BOS	932	2	1237	-7	168	1258
15	08-Aug-2015 01:01:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3	1237	-8	75	460
16	11-Jan-2015 09:11:00	19805	AA	N4XGAA	1046	MCI	DFW	1051	-13	1237	-8	72	460
17	14-Jan-2015 08:55:00	19805	AA	N4WPAA	1046	MCI	DFW	1055	-9	1237	-8	82	460
18	16-Aug-2015 02:40:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-8	1237	-13	172	1182
19	21-Jan-2015 08:50:00	19805	AA	N3DJAA	1108	DFW	LGA	850	60	1237	38	153	1389
20	09-Jan-2015 08:29:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1	1237	-2	163	1389
21	10-Dec-2015 09:11:00	19805	AA	N503AA	1310	DFW	CLE	911	16	1237	10	129	1021
22	03-Dec-2015 07:59:00	19805	AA	N5EXAA	1275	JFK	STT	759	-1	1237	-18	190	1623
23	31-Jan-2015 08:43:00	19805	AA	N3LVAA	1334	SAN	ORD	843	-1	1237	-13	205	1723



After:

#	FlightDate	AirlineID	Carrier	TailNum	FlightNum	Origin	Dest	DepartureTime	Delay	ArrivalTime	ArrDelay	Airtime	Distance
1	31-Jan-2015 08:55:00	19805	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475
2	31-Jan-2015 09:00:00	19805	AA	N3JAA	25	BOS	LAX	900	0	1237	0	331	2611
3	28-Feb-2015 08:11:00	19805	AA	N590AA	300	TUS	ORD	811	-4	1237	2	181	1437
4	31-Jan-2015 08:07:00	19805	AA	N472AA	300	TUS	ORD	807	-9	1237	2	178	1437
5	31-Jan-2015 11:42:00	19805	AA	N3BSAA	316	DFW	DEN	1142	7	1237	1	95	641
6	28-Feb-2015 09:54:00	19805	AA	N3HLAA	184	DFW	SFO	1054	29	1237	27	1464	
7	31-Jan-2015 11:45:00	19805	AA	N02AA	1079	DFW	ELP	1145	-5	1237	2	80	551
8	31-Jan-2015 08:58:00	19805	AA	N5E5AA	253	LAX	OOG	858	-2	1237	-19	318	2486
9	31-Jan-2015 09:54:00	19805	AA	N788AA	255	JFK	LAX	954	-8	1237	-53	317	2475
10	31-Mar-2015 09:15:00	19805	AA	N3CBA	1010	DFW	PBI	915	-5	1237	-17	127	1102
11	31-Jan-2015 11:49:00	19805	AA	N474AA	1023	DFW	AUS	1148	-4	1237	-13	39	190
12	31-Mar-2015 08:19:00	19805	AA	N3DVA	1027	BOS	DFW	919	-6	1237	-23	238	1582
13	31-Jan-2015 09:27:00	19805	AA	N565AA	1033	MIA	BOS	927	-3	1237	-7	165	1258
14	31-Jan-2015 09:32:00	19805	AA	N565AA	1033	MIA	BOS	932	2	1237	-7	168	1258
15	31-Aug-2015 11:01:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3	1237	-8	75	460
16	31-Jan-2015 09:11:00	19805	AA	N4XGAA	1046	MCI	DFW	1051	-13	1237	-8	72	460
17	31-Jan-2015 08:55:00	19805	AA	N4WPAA	1046	MCI	DFW	1055	-9	1237	-8	82	460
18	31-Aug-2015 02:40:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-8	1237	-13	172	1182
19	31-Jan-2015 08:50:00	19805	AA	N3DJAA	1108	DFW	LGA	850	60	1237	38	153	1389
20	31-Jan-2015 08:29:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1	1237	-2	163	1389
21	31-Dec-2015 09:11:00	19805	AA	N503AA	1310	DFW	CLE	911	16	1237	10	129	1021
22	31-Dec-2015 07:59:00	19805	AA	N5EXAA	1275	JFK	STT	759	-1	1237	-18	190	1623
23	31-Jan-2015 08:43:00	19805	AA	N3LVAA	1334	SAN	ORD	843	-1	1237	-13	205	1723

TRANSFORM—LAST DATE OF MONTH

8.3.4 Transform Advanced functions

Users can perform various operations to transform datetime data by using an extensive collection of advanced functions.

Shown below is the list of all Data Operations:

Operation	Description	Example
date(Timestamp)	Returns the date part of a Timestamp Argument 1: The timestamp for which the date has to be returned Returns: A date	Argument 1 = 2018-02-16 20:38:40 Returns 2018-02-16

dateAdd("string", i , date)	<p>Adds a certain date or time interval to a date</p> <p>Argument 1: The type of interval to be added (where the type of interval can be: Year / Month / Day / Hour / Minute / Second)</p> <p>Argument 2: The interval to be added</p> <p>Argument 3: The date or timestamp to which the specified interval has to be added or subtracted</p> <p>Returns: A timestamp</p>	<p>Argument 1 = "Year" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2020-02-16 20:38:40</p> <p>Argument 1 = "Month" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2018-04-16 20:38:40</p> <p>Argument 1 = "Day" Argument 2 = 10 Argument 3 = 2018-02-16 20:38:40 Returns 2018-02-26 20:38:40</p> <p>Argument 1 = "Hour" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2018-02-16 22:38:40</p> <p>Argument 1 = "Minute" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2018-02-16 22:40:40</p> <p>Argument 1 = "Second" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2018-02-16 22:38:42</p>
dateDiff("string", date, date)	<p>Returns the number of intervals between two dates or times</p> <p>Argument 1: The type of interval to be calculated (where the type of interval to be calculated can be: Year / Month / Day / Hour / Minute / Second)</p> <p>Argument 2: The start date or time</p> <p>Argument 3: The end date or time</p> <p>Returns: A number</p>	<p>Argument 1 = "Year" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2016-02-16 20:30:20 Returns 2</p> <p>Argument 1 = "Month" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2018-05-16 20:38:40 Returns -3</p> <p>Argument 1 = "Day" Argument 2 = 2018-02-20 20:38:40 Argument 3 = 2018-02-16 20:38:40 Returns 4</p> <p>Argument 1 = "Hour" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2018-02-16 10:38:40 Returns 10</p> <p>Argument 1 = "Minute" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2018-02-16 10:18:40 Returns 10</p> <p>Argument 1 = "Second" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2018-02-16 10:38:10 Returns 30</p>
day(date)	<p>Returns the day of a date represented by a number (an integer between 1 and 31)</p> <p>Argument 1: The date or timestamp argument whose day of month is to</p>	<p>Argument 1 = 2018-02-16 20:38:40 Returns 16</p>

	<p>be returned</p> <p>Returns: A number</p>	
dayName(date)	<p>Returns the name of the day of the week</p> <p>Argument 1: The date or timestamp for which day of the week is to be returned</p> <p>Returns: A string</p>	<p>Argument 1 = 2018-02-16 20:38:40 Returns Friday</p>
dayOfWeek(date)	<p>Returns a number (between 1 and 7) representing the day of the week(for example, Monday is 1, Wednesday is 3, and Sunday is 7)</p> <p>Argument 1: The date or timestamp for which day of the week is to be returned</p> <p>Returns: A number</p>	<p>Argument 1 = 2018-02-16 20:38:40 Returns 5</p>
dayOfYear(date)	<p>Returns a number representing the day of the year (an integer between 1 and 366)</p> <p>Argument 1: The date or timestamp for which day of the year is to be returned</p> <p>Returns: A number</p>	<p>Argument 1 = 2018-02-16 20:38:40 Returns 47</p>
daysAfter(date, date)	<p>Returns the count of number of days after specified date</p> <p>Argument 1: The start date Argument 2: The end date</p> <p>Returns: A number</p>	<p>Argument 1 = 2018-02-16 20:38:40 Argument 2 = 2018-02-10 20:38:40 Returns 6</p>
formatDate(date, "string")	<p>Returns the date or timestamp in specified format as a string datatype</p> <p>Argument 1: The date or timestamp Argument 2: The format in which specified date or timestamp will be transformed (where the format can be user defined such as "dd-mm-yy hh:mm:ss")</p> <p>Returns: A string</p>	<p>Argument 1 = 2018-02-16 Argument 2 = "yy/mm/dd" Returns 18/02/16</p> <p>Argument 1 = 2018-02-16 20:38:40 Argument 2 = "MM/dd/yyyy" Returns 02/16/2018</p>
hour(date)	<p>Returns the hour of a time value (an integer ranging between 0 [12:00 AM] to 23 [11:00 PM])</p>	<p>Argument 1 = 2018-02-16 20:38:40 Returns 20</p>

	<p>Argument 1: The timestamp for which hours are to be returned</p> <p>Returns: A number</p>	
minute(date)	<p>Returns the minutes of a time value (an integer ranging from 0 to 59)</p> <p>Argument 1: The timestamp for which minutes are to be returned</p> <p>Returns: A number</p>	Argument 1 = 2018-02-16 20:38:40 Returns 38
month(date)	<p>Returns the month (an integer between 1 and 12)</p> <p>Argument 1: The date or timestamp for which month is to be returned</p> <p>Returns: A number</p>	Argument 1 = 2018-02-16 20:38:40 Returns 2
monthName (date)	<p>Returns the month name for a given date or timestamp</p> <p>Argument 1: The date or timestamp for which month name is to be returned</p> <p>Returns: A string</p>	Argument 1 = 2018-02-16 20:38:40 Returns February
quarter(date)	<p>Returns the quarter corresponding to a date (an integer between 1 and 4)</p> <p>Argument 1: The date or timestamp for which quarter is to be returned</p> <p>Returns: A number</p>	Argument 1 = 2018-02-16 20:38:40 Returns 1
second(timestamp)	<p>Returns the seconds of a time value (an integer in the range 0 to 59)</p> <p>Argument 1: The timestamp for which seconds are to be returned</p> <p>Returns: A number</p>	Argument 1 = 2018-02-16 20:38:40 Returns 40
time(timestamp)	<p>Returns the time part from a given timestamp as a string datatype</p> <p>Argument 1: The timestamp for which time part is to be returned</p> <p>Returns: A string</p>	Argument 1 = 2018-02-16 20:38:40 Returns "20:38:40"
weekOfMonth (date)	<p>Returns a number (between 1 and 5) representing the week of the month (for example, 1 for the first week and 3 for the third week)</p> <p>Argument 1: The date or timestamp for which week of month is to be returned</p>	Argument 1 = 2018-02-16 20:38:40 Returns 3

	Returns: A number	
weekOfYear (date)	<p>Returns a number (between 1 and 52) representing the week of the year (for example, 8 for the eighth week, 19 for the nineteenth week)</p> <p>Argument 1: The date or timestamp for which week of the year is to be returned</p> <p>Returns: A number</p>	Argument 1 = 2018-02-16 20:38:40 Returns 7
year(date)	<p>Returns the year part of the Date or Timestamp (for example, 2001, 2018, 3000)</p> <p>Argument 1: The date or timestamp for which year part is to be returned</p> <p>Returns: A number</p>	Argument 1 = 2018-02-16 20:38:40 Returns 2018

8.4 Change data type

This function allows the flexibility to change the current data type of a column into another. Users can preview a sample of transformed values before actually changing the data type of a column. Data types can be changed as follows:

8.4.1 Timestamp

The Timestamp data type can be converted into following new data types:

New data type	Example		
	Original value	Format	Transformed value
String (user can specify the format of the Timestamp target variable)	2015-01-03 11:46:00	dd/mm/yyyy hh:mm:ss	"03/01/2015 11:46:00"
		dd MMMM, yyyy hh:mm	"03 January, 2015 11:46"
		dd - M - yyyy hh:mm a	"03 - 1 - 2015 11:46 AM"
		d MMMM, yyyy	"3 January, 2015"
Date	2015-01-03 11:46:00	Not applicable	2015-01-03

8.4.2 Date

The Date can be converted into the following new data types:

New data type	Example		
	Original value	Format	Transformed value
String	2015-01-03	dd MMM, YYYY	"03 Jan, 2015"
		d/M/yyyy EEEE	"3/1/2015 Saturday"

(user can specify the format of the Date target variable)		d/M/yy	"3/1/15"
		d.M.yyyy	"3.1.2015"
Timestamp	2015-01-03	Not applicable	2015-01-03 00:00:00

8.4.3 Integer

The Integer type of data can be converted into the following new data types:

New data type	Example	
	Original value	Transformed value
String	799	"799"
Double	799	799.00
BigInteger	799	799

8.4.4 Double

The Double type of data can be converted into the following new data types:

New data type	Example	
	Original value	Transformed value
String	61.0	"61.0"
Integer	61.0	61
BigInteger	61.0	61

8.4.5 String

The String type of data can be converted into the following new data types:

New data type	Example		
	Original value	Format	Transformed value
Date (user needs to specify the format of the Date source variable)	N016AA	Not applicable	NULL
	1027	Not applicable	NULL
	21/03/2015	dd/mm/yyyy	2015-03-21
	21/3/15	d/M/yy	2015-03-21
	21.3.2015	d.M.yyyy	2015-03-21
Timestamp (user needs to specify the format of the Timestamp source variable)	N016AA	Not applicable	NULL
	1027	Not applicable	NULL
	21/03/2015 09:15:00	dd/mm/yyyy hh:mm:ss	2015-03-21 09:15:00
	21 Mar. 09:15:00 AM	d MMM. hh:mm:ss a	2015-03-21 09:15:00
	21-3-2015 09:15:00	d-M-yyyy hh:mm:ss	2015-03-21 09:15:00
Double	N016AA	Not applicable	NULL
	21/03/2015 09:15:00	Not applicable	NULL
	1027	Not applicable	1027.00

Boolean	N016AA	Not applicable	NULL
	21/03/2015 09:15:00	Not applicable	NULL
	1027	Not applicable	NULL
Integer	N016AA	Not applicable	NULL
	21/03/2015 09:15:00	Not applicable	NULL
	1027	Not applicable	1027
BigInteger	N016AA	Not applicable	NULL
	21/03/2015 09:15:00	Not applicable	NULL
	1027	Not applicable	1027

8.5 Aggregate

Aggregate allows users to aggregate and reduce the Dataset by using various data operations on Measure columns.

The screenshot shows the Smarten platform's data exploration interface. A modal window titled 'Aggregate' is open, allowing users to perform calculations on selected columns. The 'Measure columns' section contains four items, each with a dropdown menu: DEP_DELAY (SUM), ARR_DELAY (SUM), AIR_TIME (SUM), and DISTANCE (SUM). The 'Columns' section lists 'FL_NUM' and 'DEP_TIME'. The background shows a table of flight data with columns like FL_DATE, CARRIER, FL_NUM, ORIGIN, DEST, DEP_TIME, and DEP_DELAY. The top navigation bar includes the Smarten logo, user information ('Welcome admin'), and various toolbar icons.

Shown below is a list of operations applicable to Measure columns:

8.5.1 Sum

This function returns the sum of the selected Measure columns for the aggregated row.

For example, the sum total of delays in Departure and Arrival of flights starting from “LAX” (Los Angeles) to various destinations each day.

Fl_Date	Origin	Destination	Dep_delay	Arr_delay
01-Jan-2018	LAX	OGG	2.00	3.00
01-Jan-2018	LAX	OGG	5.00	8.00
10-Jan-2018	LAX	OGG	1.00	2.00
10-Jan-2018	LAX	OGG	5.00	3.00
01-Jan-2018	LAX	JFK	4.00	2.00
01-Jan-2018	LAX	JFK	6.00	4.00

DATASET



Fl_Date	Origin	Destination	Dep_delay	Arr_delay
01-Jan-2018	LAX	OGG	7.00	11.00
10-Jan-2018	LAX	OGG	6.00	5.00
01-Jan-2018	LAX	JFK	10.00	6.00

AGGREGATED DATASET

8.5.2 Min

This function returns the minimum value of the selected Measure column for the aggregated row.

For example, the minimum delays in Departure and Arrival of flights starting from “LAX” (Los Angeles) to various destinations each day.

Fl_Date	Origin	Destination	Dep_delay	Arr_delay
01-Jan-2018	LAX	OGG	2.00	8.00
01-Jan-2018	LAX	OGG	5.00	3.00
10-Jan-2018	LAX	OGG	1.00	3.00
10-Jan-2018	LAX	OGG	5.00	2.00
01-Jan-2018	LAX	JFK	4.00	4.00
01-Jan-2018	LAX	JFK	6.00	2.00

DATASET



Fl_Date	Origin	Destination	Dep_delay	Arr_delay
01-Jan-2018	LAX	OGG	2.00	3.00
10-Jan-2018	LAX	OGG	1.00	2.00
01-Jan-2018	LAX	JFK	4.00	2.00

AGGREGATED DATASET

8.5.3 Max

This function returns the maximum value of the selected measure column for the aggregated row.

For example, the maximum delays in Departure and Arrival of flights starting from “LAX” (Los Angeles) to various destinations each day.

Fl_Date	Origin	Destination	Dep_delay	Arr_delay
01-Jan-2018	LAX	OGG	2.00	8.00
01-Jan-2018	LAX	OGG	5.00	3.00
10-Jan-2018	LAX	OGG	1.00	3.00
10-Jan-2018	LAX	OGG	5.00	2.00
01-Jan-2018	LAX	JFK	4.00	4.00
01-Jan-2018	LAX	JFK	6.00	2.00

DATASET



Fl_Date	Origin	Destination	Dep_delay	Arr_delay
01-Jan-2018	LAX	OGG	5.00	8.00
10-Jan-2018	LAX	OGG	5.00	3.00
01-Jan-2018	LAX	JFK	6.00	4.00

AGGREGATED DATASET

8.5.4 Count

This function returns the count of the selected Measure column for the aggregated row.

For example, the count of flights with delayed Departure and Arrival, starting from “LAX” (Los Angeles) to various destinations each day.

Fl_Date	Origin	Destination	Dep_delay	Arr_delay
01-Jan-2018	LAX	OGG	2.00	3.00
01-Jan-2018	LAX	OGG	5.00	8.00
10-Jan-2018	LAX	OGG	1.00	2.00
10-Jan-2018	LAX	OGG	5.00	3.00
01-Jan-2018	LAX	JFK	4.00	2.00
01-Jan-2018	LAX	JFK	6.00	4.00
01-Jan-2018	LAX	JFK	7.00	6.00

DATASET



Fl_Date	Origin	Destination	Dep_delay	Arr_delay
01-Jan-2018	LAX	OGG	2.00	2.00
10-Jan-2018	LAX	OGG	2.00	2.00
01-Jan-2018	LAX	JFK	3.00	3.00

AGGREGATED DATASET

8.5.5 Average

This function returns the average value of the selected Measure column for the aggregated row.

For example, the average delays in Departure and Arrival of flights starting from “LAX” (Los Angeles) to various destinations each day.



9 Shape Data

Smarten SSDP allows users to shape raw data by splitting, merging, sorting, and copying it or by adding custom columns and marking it as per requirement.

9.1 Split

Users can split a column into many columns or rows.

Note:

This function is applicable for String and Numeric data only.

9.1.1 Split to column

This function allows users to split a column into two columns based on a given separator or other criteria selected by the user. The split values are displayed in two new columns that are added to the Dataset. Users can change the default names of the new columns and also preview the resultant values before actually splitting the data of the target column.

For example, let us split the values of column “FLIGHT” into two new columns called “CARRIER” and “FLIGHTNUM” using a hyphen (-) as a separator:

Original column	New columns	
FLIGHT	CARRIER	FLIGHTNUM
AA-1010	AA	1010
DL-1803	DL	1803
WN-3771	WN	3771

Shown below is the before and after scenario of “Split - Split to column” for column “FLIGHT”:

Before:

Smarten Advanced Data Discovery

FlightData_Jan_2015_Dataset

Last refreshed on

Result set

Split column - FLIGHT

Split by: Match, Regex, Length

Separator: -

Split from: Left, Right

New column name: CARRIER, FLIGHTNUM

PREVIEW

#	FLIGHTDATE	AIRLINEID	FLIGHT	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY
1	January 01, 2015 00:00:00	19805	AA-1	N787AA	JFK	LAX	865	-5.0
2	January 23, 2015 00:00:00	19805	AA-25	N3JUAA	BOS	LAX	900	0.0
3	January 05, 2015 00:00:00	19805	AA-300	N590AA	TUS	ORD	811	-4.0
4	January 10, 2015 00:00:00	19805	AA-300	N472AA	TUS	ORD	807	-9.0
5	January 15, 2015 00:00:00	19805	AA-316	N9B6AA	DFW	DEN	1142	7.0
6	January 04, 2015 00:00:00	19805	AA-184	N3HLAA	DFW	SFO	1054	29.0
7	January 29, 2015 00:00:00	19805	AA-1079	N002AA	DFW	ELP	1145	-5.0
8	January 31, 2015 00:00:00	19805	AA-253	N5E5AA	LAX	O9G	858	-2.0
9	January 14, 2015 00:00:00	19805	AA-255	N789AA	JFK	LAX	954	-6.0
10	January 21, 2015 00:00:00	19805	AA-1010	N3CBAA	DFW	PBI	915	-5.0
11	January 03, 2015 00:00:00	19805	AA-1023	N474AA	DFW	AUS	1148	-4.0
12	January 03, 2015 00:00:00	19805	AA-1027	N3D7AA	BOS	DFW	919	-6.0
13	January 12, 2015 00:00:00	19805	AA-1033	N565AA	MIA	BOS	927	-3.0
14	January 16, 2015 00:00:00	19805	AA-1033	N565AA	MIA	BOS	932	2.0
15	January 08, 2015 00:00:00	19805	AA-1046	N556AA	MCI	DFW	1101	-3.0
16	January 11, 2015 00:00:00	19805	AA-1046	N4XGAA	MCI	DFW	1051	-13.0
17	January 14, 2015 00:00:00	19805	AA-1046	N4WPPA	MCI	DFW	1055	-9.0
18	January 16, 2015 00:00:00	19805	AA-1238	N3LEAA	FLL	ORD	1024	-6.0
19	January 21, 2015 00:00:00	19805	AA-1083	N3DQAA	DFW	LGA	660	60.0
20	January 09, 2015 00:00:00	19805	AA-1110	N3C7AA	DFW	LGA	829	-1.0
21	January 10, 2015 00:00:00	19805	AA-1310	N503AA	DFW	CLE	911	16.0
22	January 03, 2015 00:00:00	19805	AA-1275	N5E7AA	JFK	STT	759	-1.0
23	January 31, 2015 00:00:00	19805	AA-1334	N3LVAA	SAN	ORD	643	-1.0
24	January 13, 2015 00:00:00	19805	AA-1418	N3GCCA	DFW	HND	1128	11.0
25	January 22, 2015 00:00:00	19805	AA-1418	N3FAAA	DFW	HND	1135	20.0
26	January 01, 2015 00:00:00	19805	AA-1482	N3DQAA	SFO	DFW	703	-7.0
27	January 29, 2015 00:00:00	19805	AA-1482	N3HDAA	SFO	DFW	711	-4.0
28	January 30, 2015 00:00:00	19805	AA-1083	N494AA	DFW	AUS	1145	-5.0
29	January 28, 2015 00:00:00	19805	AA-1165	N016AA	ATL	MIA	1050	63.0
30	January 01, 2015 00:00:00	19805	AA-1174	N481AA	RNO	DFW	712	-3.0
31	January 09, 2015 00:00:00	19805	AA-1175	N3EMAA	MIA	IAH	1046	71.0
32	January 01, 2015 00:00:00	19805	AA-1554	N3DEAA	DFW	SNA	1130	-5.0
33	January 26, 2015 00:00:00	19805	AA-1627	N502AA	DFW	FAT	1129	-6.0
34	January 17, 2015 00:00:00	19805	AA-1684	N489AA	TUS	DFW	940	-7.0
35	January 21, 2015 00:00:00	19805	AA-1584	N201AA	TUS	DFW	941	-6.0
36	January 27, 2015 00:00:00	19805	AA-1584	N465AA	TUS	DFW	937	-10.0
37	January 03, 2015 00:00:00	19805	AA-1694	N3LTAA	MIA	LGA	953	-2.0
38	January 03, 2015 00:00:00	19805	AA-1507	N3HEAA	SEA	DFW	663	-7.0
39	January 15, 2015 00:00:00	19805	AA-1513	N019AA	ATL	DFW	1054	7.0
40	January 08, 2015 00:00:00	19805	AA-2333	N3F1AA	TPA	DFW	1102	5.0

www.ElegantBI.com

APPLY CANCEL



After:

Smarten Advanced Data Discovery

FlightData_Jan_2015_Dataset

Last refreshed on

Result set

Column: FLIGHT split to column: CARRIER and column: FLIGHTNUM successfully

SPLIT - SPLIT TO COLUMN

#	FLIGHTDATE	AIRLINEID	FLIGHT	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AI
1	January 01, 2015 00:00:00	19805	AA-1	AA	1	N787AA	JFK	LAX	865	-5.0	1237	7.0	378 ▲
2	January 23, 2015 00:00:00	19805	AA-25	AA	25	N3JUAA	BOS	LAX	900	0.0	1237	0.0	331
3	January 05, 2015 00:00:00	19805	AA-300	AA	300	N590AA	TUS	ORD	811	-4.0	1237	2.0	181
4	January 10, 2015 00:00:00	19805	AA-300	AA	300	N472AA	TUS	ORD	807	-9.0	1237	2.0	178
5	January 15, 2015 00:00:00	19805	AA-316	AA	316	N3B5AA	DFW	DEN	1142	7.0	1237	1.0	95.0
6	January 04, 2015 00:00:00	19805	AA-184	AA	184	N3HLAA	DFW	SFO	1054	29.0	1237	27.0	205
7	January 29, 2015 00:00:00	19805	AA-1079	AA	1079	N002AA	DFW	ELP	1145	-5.0	1237	2.0	80.0
8	January 31, 2015 00:00:00	19805	AA-253	AA	253	N5E5AA	LAX	O9G	858	-2.0	1237	-19.0	318
9	January 14, 2015 00:00:00	19805	AA-255	AA	255	N788AA	JFK	LAX	954	-6.0	1237	-53.0	317
10	January 21, 2015 00:00:00	19805	AA-1010	AA	1010	N3CBAA	DFW	PBI	915	-5.0	1237	-17.0	127
11	January 03, 2015 00:00:00	19805	AA-1023	AA	1023	N474AA	DFW	AUS	1148	-4.0	1237	-13.0	39.0
12	January 03, 2015 00:00:00	19805	AA-1027	AA	1027	N3D7AA	BOS	DFW	919	-6.0	1237	-23.0	238
13	January 12, 2015 00:00:00	19805	AA-1033	AA	1033	N855AA	MIA	BOS	927	-3.0	1237	-7.0	165
14	January 18, 2015 00:00:00	19805	AA-1033	AA	1033	N855AA	MIA	BOS	932	2.0	1237	-7.0	168
15	January 08, 2015 00:00:00	19805	AA-1046	AA	1046	N555AA	MCI	DFW	1101	-3.0	1237	-8.0	75.0
16	January 11, 2015 00:00:00	19805	AA-1046	AA	1046	N4XGAA	MCI	DFW	1051	-13.0	1237	-8.0	72.0
17	January 14, 2015 00:00:00	19805	AA-1046	AA	1046	N4WPPA	MCI	DFW	1055	-9.0	1237	-8.0	82.0
18	January 16, 2015 00:00:00	19805	AA-1238	AA	1238	N3LEAA	FLL	ORD	1024	-6.0	1237	-13.0	172
19	January 21, 2015 00:00:00	19805	AA-1108	AA	1108	N3D9AA	DFW	LGA	850	60.0	1237	36.0	153
20	January 09, 2015 00:00:00	19805	AA-1110	AA	1110	N3CKAA	DFW	LGA	829	-1.0	1237	-2.0	163
21	January 10, 2015 00:00:00	19805	AA-1310	AA	1310	N503AA	DFW	CLE	911	16.0	1237	10.0	129
22	January 03, 2015 00:00:00	19805	AA-1275	AA	1275	N5EXAA	JFK	STT	759	-1.0	1237	-18.0	190
23	January 31, 2015 00:00:00	19805	AA-1334	AA	1334	N3LVAA	SAN	ORD	843	-1.0	1237	-13.0	205
24	January 13, 2015 00:00:00	19805	AA-1418	AA	1418	N3GCCA	DFW	HND	1128	11.0	1237	2.0	111
25	January 22, 2015 00:00:00	19805	AA-1418	AA	1418	N3FAAA	DFW	HND	1135	20.0	1237	2.0	106
26	January 01, 2015 00:00:00	19805	AA-1482	AA	1482	N3DUAA	SFO	DFW	703	-7.0	1237	-3.0	185
27	January 29, 2015 00:00:00	19805	AA-1482	AA	1482	N3HDAA	SFO	DFW	711	-4.0	1237	-3.0	175
28	January 30, 2015 00:00:00	19805	AA-1083	AA	1083	N494AA	DFW	AUS	1145	-5.0	1237	-11.0	35.0
29	January 28, 2015 00:00:00	19805	AA-1162	AA	1162	N016AA	ATL	MIA	1050	63.0	1237	51.0	74.0
30	January 01, 2015 00:00:00	19805	AA-1174	AA	1174	N481AA	RNO	DFW	712	-3.0	1237	2.0	175
31	January 09, 2015 00:00:00	19805	AA-1175	AA	1175	N3EMAA	MIA	IAH	1046	71.0	1237	55.0	148
32	January 01, 2015 00:00:00	19805	AA-1554	AA	1554	N3DEAA	DFW	SNA	1130	-5.0	1237	-13.0	166
33	January 26, 2015 00:00:00	19805	AA-1627	AA	1627	N503AA	DFW	FAT	1129	-6.0	1237	-34.0	169
34	January 17, 2015 00:00:00	19805	AA-1584	AA	1584	N489AA	TUS	DFW	940	-7.0	1237	-18.0	101
35	January 21, 2015 00:00:00	19805	AA-1584	AA	1584	N201AA	TUS	DFW	941	-6.0	1237	-18.0	99.0
36	January 27, 2015 00:00:00	19805	AA-1584	AA	1584	N466AA	TUS	DFW	937	-10.0	1237	-18.0	103
37	January 03, 2015 00:00:00	19805	AA-1694	AA	1694	N3LTAA	MIA	LGA	953	-2.0	1237	-21.0	142.0
38	January 03, 2015 00:00:00	19805	AA-1507	AA	1507	1237	-13.0	194

SPLIT - SPLIT TO COLUMN

Smarten SSDP provides the following options to split a column:

Split by: Match		
Option	Description	Example
Separator	Splits the target column value from the position where the value entered as a separator is found	FLIGHT = "AA-1027" Separator = "-" Result: FLIGHT_1= "AA" FLIGHT_2= "1027"
Split from		
Left	Splits column value starting from the left	FLIGHT = "US-698" Separator = "-" Result: FLIGHT_1= "US", FLIGHT_2= "698"
		FLIGHT = "US-698-432" Separator = "-" Result: FLIGHT_1= "US" FLIGHT_2= "698-432"
Right	Splits column value starting from the right	FLIGHT = "US-698" Separator = "-" Result: FLIGHT_1= "US", FLIGHT_2= "698"
		FLIGHT = "US-698-432" Separator = "-" Result: FLIGHT_1= "US-698" FLIGHT_2= "432"
New column name		
Column_1	Default name of the first column	FLIGHT_1
Column_2	Default name of the second column	FLIGHT_2
Include separator		
When Selected for column_1	Suffixes the separator to the column_1 value	FLIGHT = "US-698" Separator = "-" Result: FLIGHT_1= "US-", FLIGHT_2= "698"
		FLIGHT = "US-698-432" Separator = "-" Result: FLIGHT_1= "US-" FLIGHT_2= "698-432"
When Selected for column_2	Prefixes the separator to the column_2 value	FLIGHT = "US-698" Separator = "-" Result: FLIGHT_1= "US", FLIGHT_2= "-698"
		FLIGHT = "US-698-432" Separator = "-" Result: FLIGHT_1= "US" FLIGHT_2= "-698-432"
When Not selected	Excludes the separator in the column value	

Split by: Regex		
Option	Description	
Expression	Splits column value from the position where the value entered as an expression is found	
Split from		
Left	Splits column value starting from the left	FLIGHT = "AA-1027" Separator = "\-" Result: FLIGHT_1="AA" FLIGHT_2="1027"
		FLIGHT = "US-698-432" Separator = "\-" Result: FLIGHT_1="US" FLIGHT_2="698-432"
Right	Splits column value starting from the right	FLIGHT = "US-698" Separator = "\-" Result: FLIGHT_1="US", FLIGHT_2="698"
		FLIGHT = "US-698-432" Separator = "\-" Result: FLIGHT_1="US-698" FLIGHT_2="432"
New column name		
Column 1	Default name of the first column	FLIGHT_1
Column 2	Default name of the second column	FLIGHT_2

Split by: Length		
Option	Description	Example
Length	Number of characters at which the target column should be split	FLIGHT = "AA-1027" Length = 1 Result: FLIGHT_1="A" FLIGHT_2="A-1027"
Split from		
Left	Splits column value starting from the left up to the entered length	FLIGHT = "US-698" Length = 2 Result: FLIGHT_1="US", FLIGHT_2="-698"
		FLIGHT = "US-698" Length = 3 Result: FLIGHT_1="US-", FLIGHT_2="698"
		FLIGHT = "US-698" Length = 6 Result: FLIGHT_1="US-698", FLIGHT_2=" "

Right	Splits column value starting from the right up to the entered length	FLIGHT = "US-698" Length = 2 Result: FLIGHT_1="US-", FLIGHT_2="98"
		FLIGHT = "US-698" Length = 3 Result: FLIGHT_1="US-", FLIGHT_2="698"
		FLIGHT = "US-698" Length = 6 Result: FLIGHT_1=" ", FLIGHT_2="US-698"
New column name		
Column 1	Default name of the first column	FLIGHT_1
Column 2	Default name of the second column	FLIGHT_2

9.1.2 Split to row

This function splits the value of a column into one or more rows based on the given separator. The split values are displayed in many rows of a new column that is added to the Dataset. Users can change the default name of the new column.

For example, let us split the values of column “STOPOVERS” into many rows of a new column called “STOPOVER” using comma (,) as a separator. Given below is the original data and its result:

Original Dataset:

FLIGHTDATE	ORIGIN	DEST	FLIGHT	STOPOVERS
January 03, 2015	LAX	LGA	AA-1027	DFW
January 24, 2015	LGA	LAX	B6-411	ATL,DEN
January 16, 2015	LAX	LGA	DL-1651	LAS,DEN,FLL

Result:

FLIGHTDATE	ORIGIN	DEST	FLIGHT	STOPOVERS	STOPOVER
January 03, 2015	LAX	LGA	AA-1027	DFW	DFW
January 24, 2015	LGA	LAX	B6-411	ATL,DEN	ATL
January 24, 2015	LGA	LAX	B6-411	ATL,DEN	DEN
January 16, 2015	LAX	LGA	DL-1651	LAS,DEN,FLL	LAS
January 16, 2015	LAX	LGA	DL-1651	LAS,DEN,FLL	DEN
January 16, 2015	LAX	LGA	DL-1651	LAS,DEN,FLL	FLL

Shown below is the before and after scenario of “Split - Split to row” for column “FLIGHT” into a new column called “FLIGHTNUM” using a hyphen (-) as a separator:

Before:

After:

SPLIT - SPLIT TO ROW

9.2 Merge columns

This function allows users to merge data from many columns into a single column.

For example, let us merge the values of two columns of CARRIER and FLIGHTNUM into one new column called FLIGHT, using a hyphen (-) as a separator.

Original columns		New column
CARRIER	FLIGHTNUM	FLIGHT
AA	1010	AA-1010
DL	1803	DL-1803
US	698-432	US-698-432

Shown below is the before and after scenario of “Merge columns”:

Before:

The screenshot shows the Smarten Advanced Data Discovery interface. On the left, the 'FlightData_Jan_2015_Dataset' is loaded, displaying a table with columns: #, FLIGHTDATE, AIRLINEDID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEPT. The 'CARRIER' and 'FLIGHTNUM' columns are highlighted with red boxes. On the right, the 'Merge columns' dialog is open. It has a 'New column name' field containing 'FLIGHT'. Below it, under 'Available column(s)', 'Carrier' and 'FlightNum' are listed and mapped to the 'FLIGHT' column using a plus sign and a separator dash ('-'). Other columns like 'FlightDate', 'AirlineID', 'TailNum', 'Origin', 'Dest', 'DepTime', 'DepDelay', and 'ArrTime' are also listed but not selected. A 'Separator' dropdown is set to '-'.

After:

#	FLIGHDATE	AIRLINEDID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	
1	January 01, 2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	
2	January 23, 2015 00:00:00	19805	AA	N3JUAA	25	BOS	LAX	900	0.0	1237	0.0	
3	January 05, 2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	
4	January 10, 2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	
5	January 15, 2015 00:00:00	19805	AA	N3BSAA	316	DFW	DEN	1142	7.0	1237	1.0	
6	January 04, 2015 00:00:00	19805	AA	N3H1AA	184	DFW	SFO	1054	29.0	1237	27.0	
7	January 29, 2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145	-5.0	1237	2.0	
8	January 31, 2015 00:00:00	19805	AA	N6ESAA	253	LAX	O+G	858	-2.0	1237	-19.0	
9	January 14, 2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-6.0	1237	-53.0	
10	January 21, 2015 00:00:00	19805	AA	AA-1010	1010	DFW	PBI	915	-5.0	1237	-17.0	
11	January 03, 2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4.0	1237	-13.0	
12	January 03, 2015 00:00:00	19805	AA	N3D1AA	1027	BOS	DFW	919	-6.0	1237	-23.0	
13	January 12, 2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3.0	1237	-7.0	
14	January 16, 2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2.0	1237	-7.0	
15	January 08, 2015 00:00:00	19805	AA	AA-1046	1046	MCI	DFW	1101	-3.0	1237	-8.0	
16	January 11, 2015 00:00:00	19805	AA	AA-1046	1046	MCI	DFW	1051	-13.0	1237	-8.0	
17	January 14, 2015 00:00:00	19805	AA	AA-1046	1046	MCI	DFW	1055	-9.0	1237	-8.0	
18	January 16, 2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6.0	1237	-13.0	
19	January 21, 2015 00:00:00	19805	AA	AA-1108	1108	DFW	LGA	850	6.0	1237	36.0	
20	January 09, 2015 00:00:00	19805	AA	AA-1110	1110	DFW	LGA	829	-1.0	1237	-2.0	
21	January 10, 2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16.0	1237	10.0	
22	January 03, 2015 00:00:00	19805	AA	AA-1275	1275	JFK	STT	759	-1.0	1237	-18.0	
23	January 31, 2015 00:00:00	19805	AA	N2LVAA	1334	SAN	ORD	843	-1.0	1237	-13.0	
24	January 13, 2015 00:00:00	19805	AA	AA-1418	1418	DFW	HON	1126	11.0	1237	2.0	
25	January 22, 2015 00:00:00	19805	AA	AA-1418	1418	DFW	HON	1135	20.0	1237	2.0	
26	January 01, 2015 00:00:00	19805	AA	AA-1482	1482	SFO	DFW	703	-7.0	1237	-3.0	
27	January 29, 2015 00:00:00	19805	AA	AA-1482	1482	SFO	DFW	711	-4.0	1237	-3.0	
28	January 30, 2015 00:00:00	19805	AA	AA-1083	1083	DFW	AUS	1145	-5.0	1237	-11.0	
29	January 28, 2015 00:00:00	19805	AA	AA-1162	1162	ATL	MIA	1050	63.0	1237	51.0	
30	January 01, 2015 00:00:00	19805	AA	AA-1174	1174	RNO	DFW	712	-3.0	1237	2.0	
31	January 09, 2015 00:00:00	19805	AA	AA-1175	1175	MIA	IAH	1048	71.0	1237	55.0	
32	January 01, 2015 00:00:00	19805	AA	AA-1554	1554	DFW	SNA	1130	-5.0	1237	-13.0	
33	January 26, 2015 00:00:00	19805	AA	AA-1627	1627	DFW	FAT	1129	-6.0	1237	-34.0	
34	January 17, 2015 00:00:00	19805	AA	AA-1594	1594	TUS	DFW	940	-7.0	1237	-18.0	
35	January 21, 2015 00:00:00	19805	AA	AA-1594	1594	N201AA	TUS	DFW	941	-6.0	1237	-18.0
36	January 27, 2015 00:00:00	19805	AA	AA-1594	1594	N460AA	TUS	DFW	937	-10.0	1237	-18.0
37	January 03, 2015 00:00:00	19805	AA	N3LTAA	1604	MIA	IAD	963	-2.0	1237	-21.0	
38	January 03, 2015 00:00:00	19805	AA	N3LTAA	1604	MIA	IAD	974	-7.0	1237	-13.0	

9.3 Sort

Users can arrange data of a particular column in ascending or descending order. A sort operation is applied on dataset data, and dataset records will be stored in the same order as per the sort function.

9.3.1 Sort Ascending

This function arranges the data of the target column in ascending order.

For example:

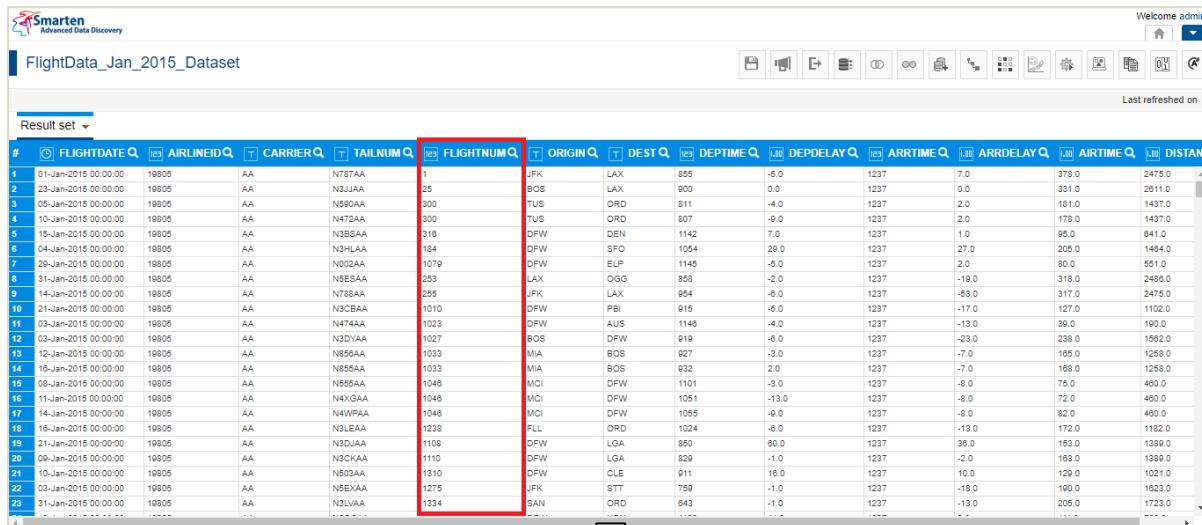
Numeric type data		String type data		Datetime type data	
Original value	Sorted value	Original value	Sorted value	Original value	Sorted value
855	25	N787AA	A5EXAA	August 08, 2015 11:01:00	January 01, 2015 09:00:00
900	698	N3CBAA	L199AA	January 23, 2015 08:55:00	January 23, 2015 08:55:00
698	855	A5EXAA	N3CBAA	January 01, 2015 09:00:00	March 14, 2015 09:54:00
1010	900	L199AA	N787AA	March 14, 2015 09:54:00	March 18, 2015 04:40:00
25	1010	P9DXAA	P9DXAA	March 18, 2015 04:40:00	

March 18, 2015 04:4:00

August 08, 2015 11:01:00

Shown below is the before and after scenario of “Sort Ascending” in column “FLIGHTNUM”:

Before:

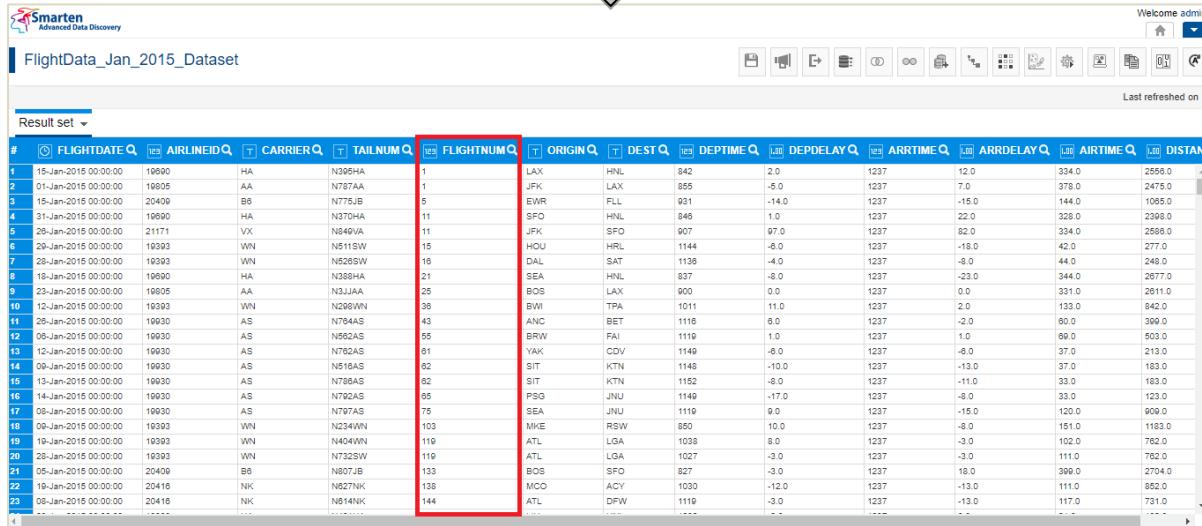


FlightData_Jan_2015_Dataset

Result set

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	378.0	2475.0
2	23-Jan-2015 00:00:00	19805	AA	N3JAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2611.0
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.0
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	173.0	1437.0
5	15-Jan-2015 00:00:00	19805	AA	N3BSAA	318	DFW	DEN	1142	7.0	1237	1.0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	1054	28.0	1237	27.0	205.0	1494.0
7	28-Jan-2015 00:00:00	19805	AA	N02AA	1079	DFW	ELP	1145	-5.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	N5E5AA	253	LAX	OOG	858	-2.0	1237	-19.0	318.0	2480.0
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-6.0	1237	-53.0	317.0	2475.0
10	21-Jan-2015 00:00:00	19805	AA	N3CBAA	1010	DFW	PBI	915	-5.0	1237	-17.0	127.0	1102.0
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4.0	1237	-13.0	39.0	160.0
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919	-6.0	1237	-23.0	238.0	1952.0
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1285.0
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2.0	1237	-7.0	168.0	1258.0
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3.0	1237	-8.0	75.0	460.0
16	11-Jan-2015 00:00:00	19805	AA	N4XGAA	1046	MCI	DFW	1051	-13.0	1237	-8.0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	AA	N4WPAA	1048	MCI	DFW	1055	-9.0	1237	-8.0	82.0	460.0
18	18-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6.0	1237	-13.0	172.0	1182.0
19	21-Jan-2015 00:00:00	19805	AA	N3DAAA	1108	DFW	LGA	850	60.0	1237	36.0	153.0	1389.0
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1.0	1237	-2.0	163.0	1389.0
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16.0	1237	10.0	129.0	1021.0
22	03-Jan-2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759	-1.0	1237	-18.0	190.0	1623.0
23	31-Jan-2015 00:00:00	19805	AA	N3LVAA	1334	SAN	ORD	843	-1.0	1237	-13.0	205.0	1723.0

After:



FlightData_Jan_2015_Dataset

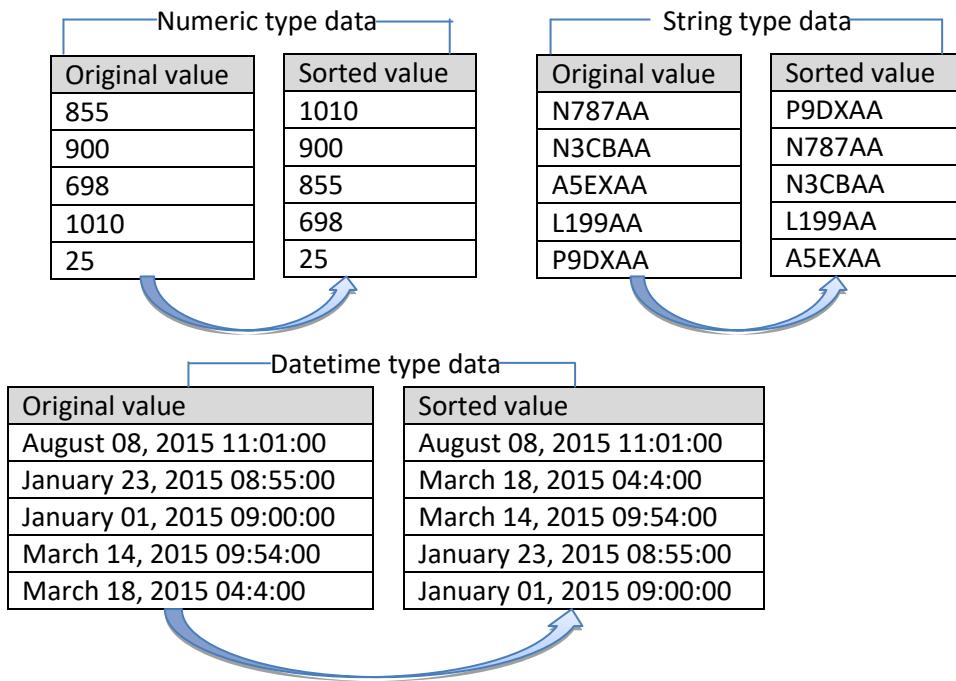
Result set

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	15-Jan-2015 00:00:00	19800	HA	N395HA	1	LAX	HNL	842	2.0	1237	12.0	334.0	2558.0
2	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	378.0	2475.0
3	15-Jan-2015 00:00:00	20409	B6	N775JB	5	EWR	FLL	931	-14.0	1237	-15.0	144.0	1085.0
4	31-Jan-2015 00:00:00	19800	HA	N370HA	11	SFO	HNL	848	1.0	1237	22.0	328.0	2398.0
5	26-Jan-2015 00:00:00	21171	VX	N849VA	11	JFK	SFO	907	97.0	1237	82.0	334.0	2568.0
6	20-Jan-2015 00:00:00	19393	WN	N511SW	15	HOU	HRL	1144	-6.0	1237	-18.0	42.0	277.0
7	28-Jan-2015 00:00:00	19393	WN	N526SW	18	DAL	SAT	1138	-4.0	1237	-8.0	44.0	248.0
8	18-Jan-2015 00:00:00	19800	HA	N388HA	21	SEA	HNL	937	-6.0	1237	-23.0	344.0	2677.0
9	23-Jan-2015 00:00:00	19805	AA	N3JAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2611.0
10	12-Jan-2015 00:00:00	19393	WN	N298WN	36	BWI	TPA	1011	11.0	1237	2.0	133.0	842.0
11	28-Jan-2015 00:00:00	19930	AS	N784AS	43	ANC	BET	1116	6.0	1237	-2.0	60.0	399.0
12	06-Jan-2015 00:00:00	19930	AS	N562AS	55	BRW	FAI	1119	1.0	1237	1.0	69.0	503.0
13	12-Jan-2015 00:00:00	19930	AS	N762AS	61	YAK	CDV	1149	-6.0	1237	-6.0	37.0	213.0
14	09-Jan-2015 00:00:00	19930	AS	N516AS	62	SIT	KTN	1148	-10.0	1237	-13.0	37.0	183.0
15	13-Jan-2015 00:00:00	19930	AS	N788AS	62	SIT	KTN	1152	-8.0	1237	-11.0	33.0	183.0
16	14-Jan-2015 00:00:00	19930	AS	N792AS	65	PSG	JNU	1149	-17.0	1237	-8.0	33.0	123.0
17	08-Jan-2015 00:00:00	19930	AS	N797AS	75	SEA	JNU	1119	9.0	1237	-15.0	120.0	909.0
18	09-Jan-2015 00:00:00	19393	WN	N234WN	103	MKE	RSW	850	10.0	1237	-9.0	151.0	1183.0
19	19-Jan-2015 00:00:00	19393	WN	N404WN	119	ATL	LGA	1038	8.0	1237	-3.0	102.0	782.0
20	28-Jan-2015 00:00:00	19393	WN	N732SW	119	ATL	LGA	1027	-3.0	1237	-3.0	111.0	782.0
21	05-Jan-2015 00:00:00	20409	B6	N807JB	133	BOS	SFO	827	-3.0	1237	10.0	398.0	2170.0
22	19-Jan-2015 00:00:00	20416	NK	N827NK	138	MCO	ACY	1030	-12.0	1237	-13.0	111.0	852.0
23	08-Jan-2015 00:00:00	20416	NK	N814NK	144	ATL	DFW	1119	-3.0	1237	-13.0	117.0	731.0

SORT ASCENDING

9.3.2 Sort Descending

This function arranges the data of the target column in descending order.



Shown below is the before and after scenario of “Sort Descending” in column “ORIGIN”:

Before:

The screenshot shows the Smarten Advanced Data Discovery interface with the following details:

- Header:** Welcome admin, FlightData_Jan_2015_Dataset, Last refreshed on [date].
- Table Headers:** #, FLIGHDATE, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEPTIME, DEPDELAY, ARRTIME, ARRDDELAY, AIRTIME, DISTANCE.
- Data:** The table contains 23 rows of flight data. The 'ORIGIN' column is highlighted with a red border.
- Toolbar:** Includes icons for search, refresh, export, and other database operations.

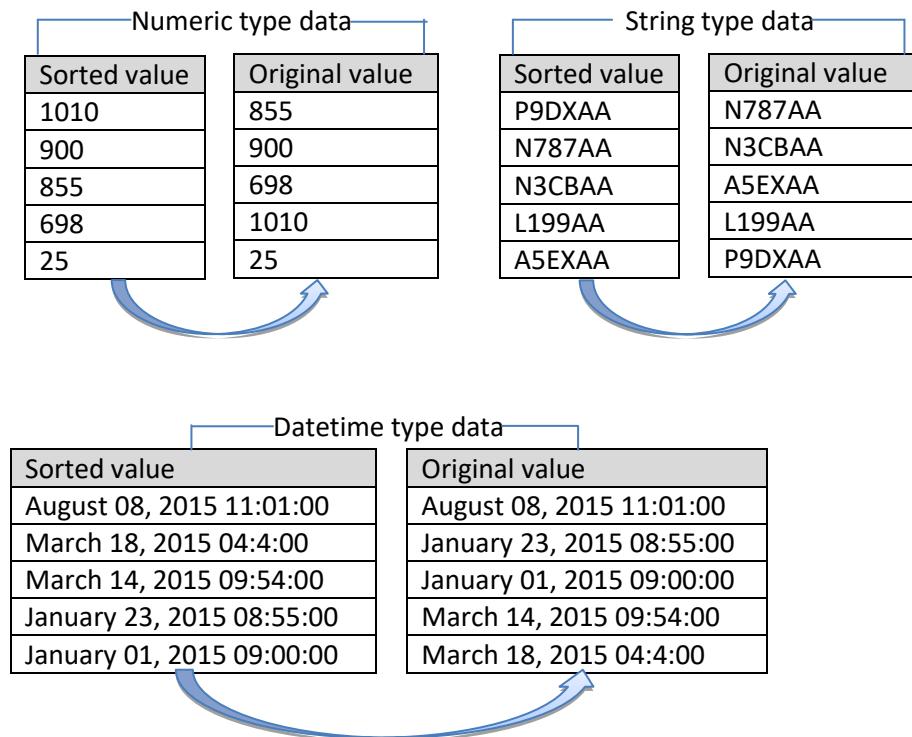
After:

SORT DESCENDING

#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPETIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	12-Jan-2015 00:00:00	19930	AS	N782AS	61	YAK	COV	1149	-8.0	1237	-8.0	37.0	213.0
2	08-Jan-2015 00:00:00	20398	EV	N12145	6070	XNA	DEN	1130	-11.0	1237	-9.0	101.0	618.0
3	15-Jan-2015 00:00:00	20398	EV	N27200	6070	XNA	DEN	1136	-5.0	1237	-9.0	97.0	618.0
4	16-Jan-2015 00:00:00	20398	EV	N11121	6070	XNA	DEN	1131	-10.0	1237	-9.0	101.0	618.0
5	09-Jan-2015 00:00:00	20304	OO	N9625W	8533	XNA	ORD	1055	17.0	1237	12.0	79.0	522.0
6	23-Jan-2015 00:00:00	19790	DL	N981AT	2145	VPS	ATL	1017	-8.0	1237	-2.0	41.0	250.0
7	14-Jan-2015 00:00:00	20398	MQ	N947MQ	3154	VPS	DFW	1021	6.0	1237	3.0	107.0	641.0
8	13-Jan-2015 00:00:00	19790	DL	N968AT	2033	TVS	ATL	1151	20.0	1237	3.0	33.0	152.0
9	21-Jan-2015 00:00:00	19805	AA	N201AA	1584	TUS	DFW	941	-8.0	1237	-18.0	99.0	813.0
10	25-Jan-2015 00:00:00	20304	OO	N494CA	4769	TUS	SLC	1047	-10.0	1237	-17.0	91.0	601.0
11	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.0
12	17-Jan-2015 00:00:00	19805	AA	N489AA	1584	TUS	DFW	940	-7.0	1237	-18.0	101.0	813.0
13	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.0
14	27-Jan-2015 00:00:00	19805	AA	N486AA	1584	TUS	DFW	937	-10.0	1237	-18.0	103.0	813.0
15	27-Jan-2015 00:00:00	19393	WN	N848SW	4955	TUL	DAL	1145	-10.0	1237	-18.0	40.0	237.0
16	19-Jan-2015 00:00:00	20398	EV	N14592	3933	TUL	ORD	1034	-3.0	1237	7.0	93.0	565.0
17	04-Jan-2015 00:00:00	19393	WN	N850SW	4722	TPA	AUS	1058	3.0	1237	2.0	143.0	928.0
18	08-Jan-2015 00:00:00	19805	AA	N3FUAA	2333	TPA	DFW	1102	5.0	1237	-8.0	130.0	929.0
19	26-Jan-2015 00:00:00	19393	WN	N238WN	403	TPA	BWI	1028	-4.0	1237	-8.0	116.0	842.0
20	21-Jan-2015 00:00:00	19805	AA	N3JUAA	2333	TPA	DFW	1053	-4.0	1237	-8.0	143.0	929.0
21	12-Jan-2015 00:00:00	19393	WN	N852SW	548	TPA	FHL	1013	38.0	1237	37.0	128.0	920.0
22	18-Jan-2015 00:00:00	20398	EV	N854AS	5328	TLH	ATL	1139	59.0	1237	48.0	48.0	223.0
23	22-Jan-2015 00:00:00	20409	BB	N368JB	315	SYR	JFK	1125	-8.0	1237	-4.0	42.0	209.0

9.3.3 Sort None

This function removes any sorting that was applied on the column and displays the data in the order of records that were written in the Dataset files.



Shown below is the before and after scenario of “Sort None” in column “ORIGIN”:

Before:

FlightData_Jan_2015_Dataset

Last refreshed on

Result set

#	FLIGHDATE Q	AIRLINEID Q	CARRIER Q	TAILNUM Q	FLIGHTNUM Q	ORIGIN Q	DEST Q	DEPTIME Q	DEPDELAY Q	ARRTIME Q	ARRDELAY Q	AIRTIME Q	DISTANCE Q
1	12-Jan-2015 00:00:00	19930	AS	N762AS	61	YAK	CDV	1149	-8.0	1237	-8.0	37.0	213.0
2	08-Jan-2015 00:00:00	20366	EV	N12145	6070	XNA	DEN	1130	-11.0	1237	-8.0	101.0	616.0
3	15-Jan-2015 00:00:00	20366	EV	N27200	6070	XNA	DEN	1136	-8.0	1237	-9.0	97.0	616.0
4	16-Jan-2015 00:00:00	20366	EV	N11121	6070	XNA	DEN	1131	-10.0	1237	-9.0	101.0	616.0
5	09-Jan-2015 00:00:00	20304	OO	N982SW	5533	XNA	ORD	1055	17.0	1237	12.0	79.0	522.0
6	23-Jan-2015 00:00:00	19790	DL	N981AT	2146	VPS	ATL	1017	-8.0	1237	-2.0	41.0	250.0
7	14-Jan-2015 00:00:00	20398	MQ	N847MQ	3154	VPS	DFW	1021	8.0	1237	3.0	107.0	641.0
8	13-Jan-2015 00:00:00	19790	DL	N989AT	2033	TYS	ATL	1151	20.0	1237	3.0	33.0	152.0
9	21-Jan-2015 00:00:00	19805	AA	N201AA	1584	TUS	DFW	041	-8.0	1237	-18.0	99.0	813.0
10	25-Jan-2015 00:00:00	20304	OO	N494CA	4799	TUS	SLC	1047	-10.0	1237	-17.0	91.0	601.0
11	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.0
12	17-Jan-2015 00:00:00	19805	AA	N489AA	1584	TUS	DFW	040	-7.0	1237	-18.0	101.0	813.0
13	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.0
14	27-Jan-2015 00:00:00	19805	AA	N466AA	1584	TUS	DFW	037	-10.0	1237	-18.0	103.0	813.0
15	27-Jan-2015 00:00:00	19392	VN	N849SW	4955	TUL	DAL	1145	-10.0	1237	-18.0	40.0	237.0
16	10-Jan-2015 00:00:00	20366	EV	N14682	3033	TUL	ORD	1034	-3.0	1237	7.0	93.0	585.0
17	04-Jan-2015 00:00:00	19393	VN	N850SW	4722	TPA	AUS	1058	3.0	1237	2.0	143.0	928.0
18	08-Jan-2015 00:00:00	19805	AA	N32UAA	2333	TPA	DFW	1102	5.0	1237	-8.0	139.0	629.0
19	20-Jan-2015 00:00:00	19393	VN	N238WN	403	TPA	BWI	1026	-4.0	1237	-8.0	116.0	842.0
20	20-Jan-2015 00:00:00	19805	AA	N32UAA	2333	TPA	DFW	1053	-4.0	1237	-8.0	143.0	629.0
21	12-Jan-2015 00:00:00	19393	VN	N832SW	548	TPA	FHL	1013	38.0	1237	37.0	128.0	920.0
22	10-Jan-2015 00:00:00	20366	EV	N854AS	5328	TLH	ATL	1139	59.0	1237	48.0	46.0	223.0
23	22-Jan-2015 00:00:00	20409	B6	N386JB	315	SYR	JFK	1125	-8.0	1237	-4.0	42.0	209.0

After:

FlightData_Jan_2015_Dataset

Last refreshed on

Result set

#	FLIGHDATE Q	AIRLINEID Q	CARRIER Q	TAILNUM Q	FLIGHTNUM Q	ORIGIN Q	DEST Q	DEPTIME Q	DEPDELAY Q	ARRTIME Q	ARRDELAY Q	AIRTIME Q	DISTANCE Q
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5.0	1237	7.0	378.0	2475.0
2	23-Jan-2015 00:00:00	19805	AA	N37JAA	25	BOS	LAX	900	0.0	1237	0.0	331.0	2611.0
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4.0	1237	2.0	181.0	1437.0
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9.0	1237	2.0	178.0	1437.0
5	15-Jan-2015 00:00:00	19805	AA	N3BSAA	318	DFW	DEN	1142	7.0	1237	1.0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	1054	29.0	1237	27.0	205.0	1484.0
7	20-Jan-2015 00:00:00	19805	AA	N020AA	1079	DFW	ELP	1145	-5.0	1237	2.0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	N5E5AA	253	LAX	OGG	858	-2.0	1237	-19.0	318.0	2488.0
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-8.0	1237	-53.0	317.0	2475.0
10	21-Jan-2015 00:00:00	19805	AA	N3C8AA	1010	DFW	PBI	915	-8.0	1237	-17.0	127.0	1102.0
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1146	-4.0	1237	-13.0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919	-8.0	1237	-23.0	238.0	1592.0
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3.0	1237	-7.0	165.0	1258.0
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2.0	1237	-7.0	168.0	1258.0
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1048	MCI	DFW	1101	-3.0	1237	-8.0	75.0	480.0
16	11-Jan-2015 00:00:00	19805	AA	N4XGA	1048	MCI	DFW	1051	-13.0	1237	-8.0	72.0	480.0
17	14-Jan-2015 00:00:00	19805	AA	N4WPAA	1048	MCI	DFW	1055	-9.0	1237	-8.0	82.0	480.0
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-8.0	1237	-13.0	172.0	1182.0
19	21-Jan-2015 00:00:00	19805	AA	N3DJAA	1108	DFW	LGA	850	60.0	1237	38.0	153.0	1389.0
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	S29	929	-1.0	1237	-2.0	163.0	1389.0
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	18.0	1237	10.0	129.0	1021.0
22	03-Jan-2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759	-1.0	1237	-18.0	190.0	1023.0
23	31-Jan-2015 00:00:00	19805	AA	N3LVA	1334	SAN	ORD	643	-1.0	1237	-13.0	205.0	1723.0

SORT NONE

9.4 Copy

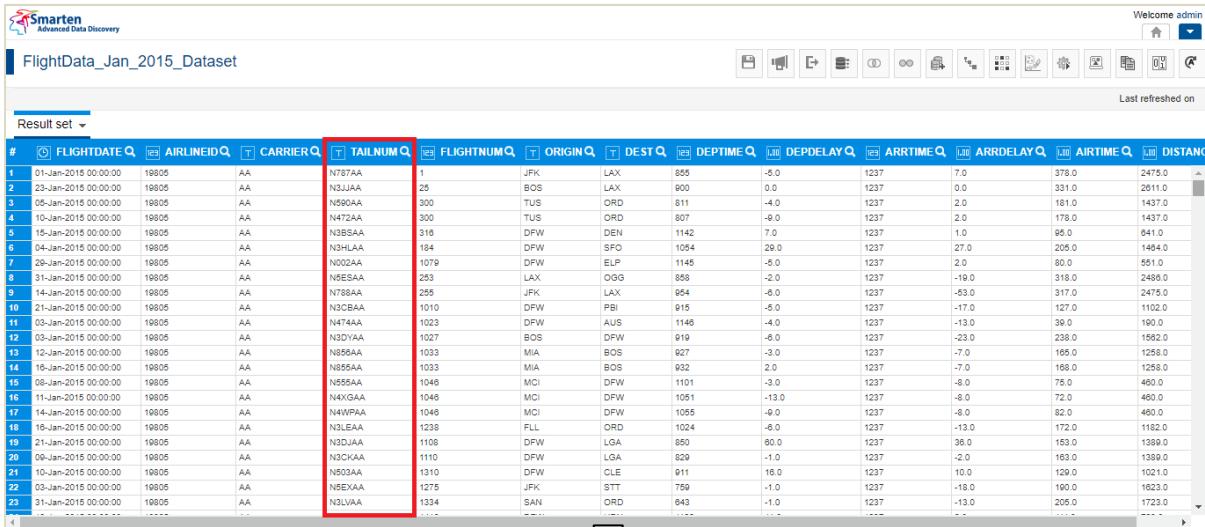
Users have the facility to copy a particular column or row. A replicate of the target column or row is added to the Dataset.

9.4.1 Copy Column

This function allows copying the data of a target column into a new column.

Shown below is the before and after scenario of “Copy Column” for column “TAILNUM”:

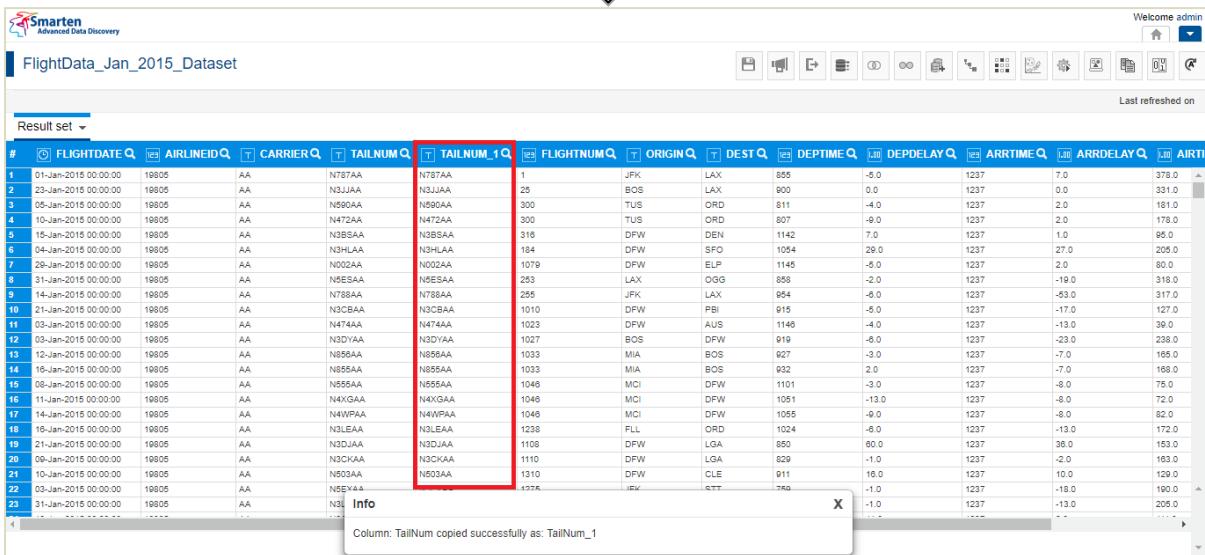
Before:



The screenshot shows a table with 23 rows of flight data. The columns include: #, FLIGHTDATE, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, DEPTIME, DEPDELAY, ARRTIME, ARRDDELAY, AIRTIME, and DISTANCE. The 'TAILNUM' column is highlighted with a red border.



After:



The screenshot shows the same table as before, but with a new column 'TAILNUM_1' added between 'TAILNUM' and 'FLIGHTNUM'. The 'TAILNUM' column is now empty. A tooltip at the bottom left says 'Column: TailNum copied successfully as: TailNum_1'.

COPY COLUMN

9.4.2 Copy Row

This function allows copying the data of a target row into a new row. While copying, the user has an option to change the value of any column as per requirement.

Shown below is the before and after scenario of “Copy Row”:

Before:

After:

9.5 Edit Row

With the help of this function, users can change the values of one or more columns for a particular row provided there is at least one column in the Dataset that has unique values.

Users cannot change the value of the column not having unique values. In such a case, users can take help of functions, such as “Add column Row number,” which creates a new column containing the unique value of each row number in the Dataset.

For example, let us change the value of column “FLIGHTNUM” from “316” to “300” for row number 4. Below is the before and after scenario of “Edit Row” for row number 4:

Before:

#	ROW_NUMBER	FLIGHTDATE	AIRLINEID	CARRIER	TAILNUM	ORIGIN	
1	0	January 01, 2015 00:00:00	19805	AA	1	N787AA	JFK
2	1	January 23, 2015 00:00:00	19805	AA	25	N3JAA	BOS
3	2	January 05, 2015 00:00:00	19805	AA	300	N690AA	TUS
4	3	January 10, 2015 00:00:00	19805	AA	300	N472AA	TUS
5	4	January 15, 2015 00:00:00	19805	AA	316	N3BSAA	DFW
6	5	January 04, 2015 00:00:00	19805	AA	184	N3HCAA	DFW
7	6	January 29, 2015 00:00:00	19805	AA	1079	N002AA	DFW
8	7	January 31, 2015 00:00:00	19805	AA	253	N5E5AA	LAX
9	8	January 14, 2015 00:00:00	19805	AA	255	N785AA	JFK
10	9	January 21, 2015 00:00:00	19805	AA	1010	N3C8AA	DFW
11	10	January 03, 2015 00:00:00	19805	AA	1023	N474AA	DFW
12	11	January 03, 2015 00:00:00	19805	AA	1027	N3D7AA	BOS
13	12	January 12, 2015 00:00:00	19805	AA	1033	N856AA	MIA
14	13	January 16, 2015 00:00:00	19805	AA	1033	N855AA	MIA
15	14	January 08, 2015 00:00:00	19805	AA	1046	N655AA	MCI
16	15	January 11, 2015 00:00:00	19805	AA	1046	N4X5AA	MCI
17	16	January 14, 2015 00:00:00	19805	AA	1046	N4W5AA	MCI
18	17	January 16, 2015 00:00:00	19805	AA	1238	N3LEAA	FL
19	18	January 21, 2015 00:00:00	19805	AA	1108	N3D5AA	DFW
20	19	January 09, 2015 00:00:00	19805	AA	1110	N3CKAA	DFW
21	20	January 10, 2015 00:00:00	19805	AA	1310	N505AA	DFW
22	21	January 03, 2015 00:00:00	19805	AA	1275	N5EXAA	JFK
23	22	January 31, 2015 00:00:00	19805	AA	1334	N3LVAA	SAN
24	23	January 13, 2015 00:00:00	19805	AA	1418	N3G5AA	DFW
25	24	January 22, 2015 00:00:00	19805	AA	1418	N3F5AA	DFW
26	25	January 01, 2015 00:00:00	19805	AA	1482	N3D5AA	SFO
27	26	January 29, 2015 00:00:00	19805	AA	1482	N3HDA	SFO
28	27	January 30, 2015 00:00:00	19805	AA	1083	N9Q4AA	DFW
29	28	January 28, 2015 00:00:00	19805	AA	1162	N916AA	ATL
30	29	January 01, 2015 00:00:00	19805	AA	1174	N4B1AA	RNO
31	30	January 09, 2015 00:00:00	19805	AA	1175	N3EMAA	MIA
32	31	January 01, 2015 00:00:00	19805	AA	1554	N3DEAA	DFW
33	32	January 26, 2015 00:00:00	19805	AA	1627	N505AA	DFW
34	33	January 17, 2015 00:00:00	19805	AA	1594	N489AA	TUS
35	34	January 21, 2015 00:00:00	19805	AA	1594	N201AA	TUS
36	35	January 27, 2015 00:00:00	19805	AA	1594	N495AA	TUS
37	36	January 03, 2015 00:00:00	19805	AA	1694	N3LTAA	MIA
38	37	January 03, 2015 00:00:00	19805	AA	1507	N3HEAA	SEA

↓

After:

#	FLIGHTDATE	AIRLINEID	ROW_NUMBER	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	January 01, 2015 09:00:00	19805	0	AA	N787AA	1	JFK	LAX	855	-5	1237	7	378	2475
2	January 23, 2015 08:55:00	19805	1	AA	N3JAA	25	BOS	LAX	900	0	1237	0	331	2811
3	February 01, 2015 09:11:00	19805	2	AA	N690AA	300	TUS	ORD	911	-4	1237	2	181	1457
4	February 14, 2015 09:42:00	19805	3	AA	N101AA	100	ORD	CDG	920	-1	1237	1	138	1427
5	February 14, 2015 09:42:00	19805	4	AA	N3BSAA	300	DFW	DEN	1142	7	1237	1	66	341
6	January 29, 2015 11:45:00	19805	5	AA	N3PLAA	154	DFW	SFO	1054	-9	1237	2	203	1454
7	January 29, 2015 11:45:00	19805	6	AA	N002AA	1079	DFW	ELP	1145	-5	1237	2	80	551
8	January 29, 2015 08:58:00	19805	7	AA	N5ESAA	253	LAX	OOG	858	-2	1237	-19	318	2486
9	January 29, 2015 08:58:00	19805	8	AA	N785AA	255	LAX	OOG	854	-2	1237	-53	317	2475
10	March 21, 2015 09:15:00	19805	9	AA	N3CBAA	1010	DFW	PBI	915	-6	1237	-17	127	1102
11	January 03, 2015 11:48:00	19805	10	AA	N474AA	1023	DFW	AUS	1148	-4	1237	-13	39	190
12	March 03, 2015 09:16:00	19805	11	AA	N3D5AA	1027	BOS	DFW	919	-6	1237	-23	238	1562
13	January 12, 2015 09:27:00	19805	12	AA	N856AA	1033	MIA	BOS	927	-3	1237	-7	165	1258
14	January 10, 2015 09:22:00	19805	13	AA	N855AA	1033	MIA	BOS	932	2	1237	-7	168	1258
15	January 10, 2015 09:22:00	19805	14	AA	N3M5AA	1048	MIA	DFW	1103	-9	1237	-8	76	440
16	January 11, 2015 10:51:00	19805	15	AA	N4K5AA	1048	MCI	DFW	1051	-13	1237	-8	72	460
17	January 14, 2015 10:55:00	19805	16	AA	N4WPAA	1048	MCI	DFW	1055	-9	1237	-8	82	480
18	August 10, 2015 10:24:00	19805	17	AA	N3LEAA	1238	FLL	ORD	1024	-6	1237	-13	172	1182
19	January 21, 2015 08:50:00	19805	18	AA	N3Q5AA	1108	DFW	LGA	860	0	1237	38	153	1389
20	January 29, 2015 08:50:00	19805	19	AA	N3Q5AA	1110	DFW	LGA	859	-1	1237	2	163	1399
21	December 10, 2015 09:11:00	19805	20	AA	N503AA	1310	DFW	CLE	911	18	1237	10	129	1021
22	December 03, 2015 07:59:00	19805	21	AA	NRE5AA	1275	JFK	STT	759	-1	1237	-18	190	1823
23	January 31, 2015 08:43:00	19805	22	AA	N3L5AA	1334	SAN	ORD	843	-1	1237	-13	205	1723
24	January 13, 2015 11:28:00	19805	23	AA	N3SGAA	1418	DFW	HDL	1125	11	1237	2	111	79
25	January 13, 2015 11:28:00	19805	24	AA	N3P5AA	1418	HDL	1135	20	1237	2	105	745	
26	January 01, 2015 08:00:00	19805	25	AA	N3D5AA	1482	SFO	DFW	703	-7	1237	-3	195	1484
27	January 29, 2015 08:00:00	19805	26	AA	N3H5AA	1482	SFO	DFW	711	-4	1237	-3	175	1494
28	January 30, 2015 08:00:00	19805	27	AA	N494AA	1083	DFW	AUS	1145	-5	1237	-11	35	190
29	January 28, 2015 00:00:00	19805	28	AA	N015AA	1162	ATL	MIA	1050	63	1237	51	74	894

EDIT ROW

9.6 Add column

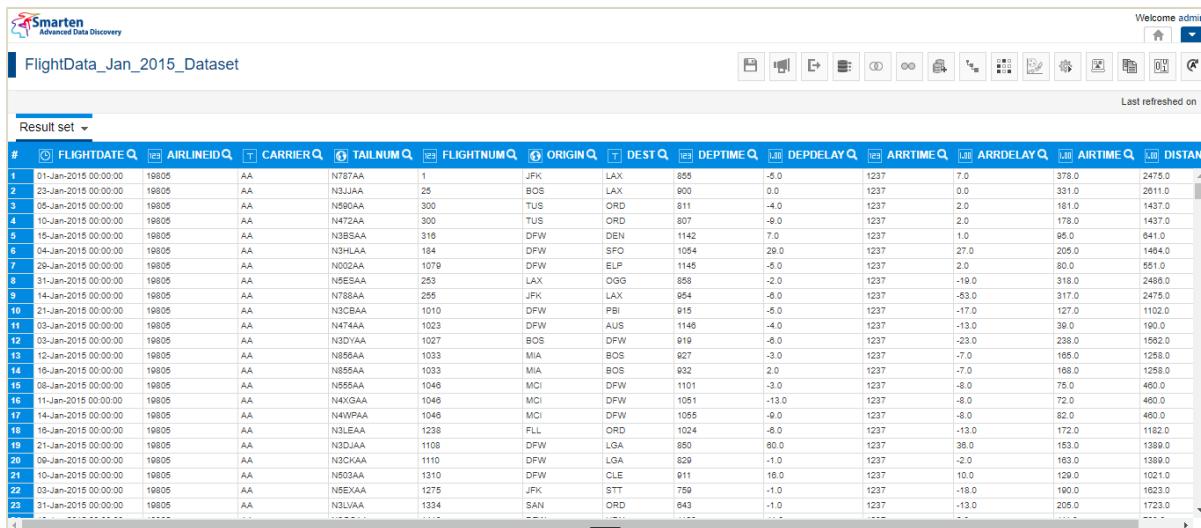
Smarten SSDP allows users to add new custom columns to the Dataset as per their requirement.

9.6.1 Row number

This function allows users to add a new column that contains the row number in ascending order, starting with zero.

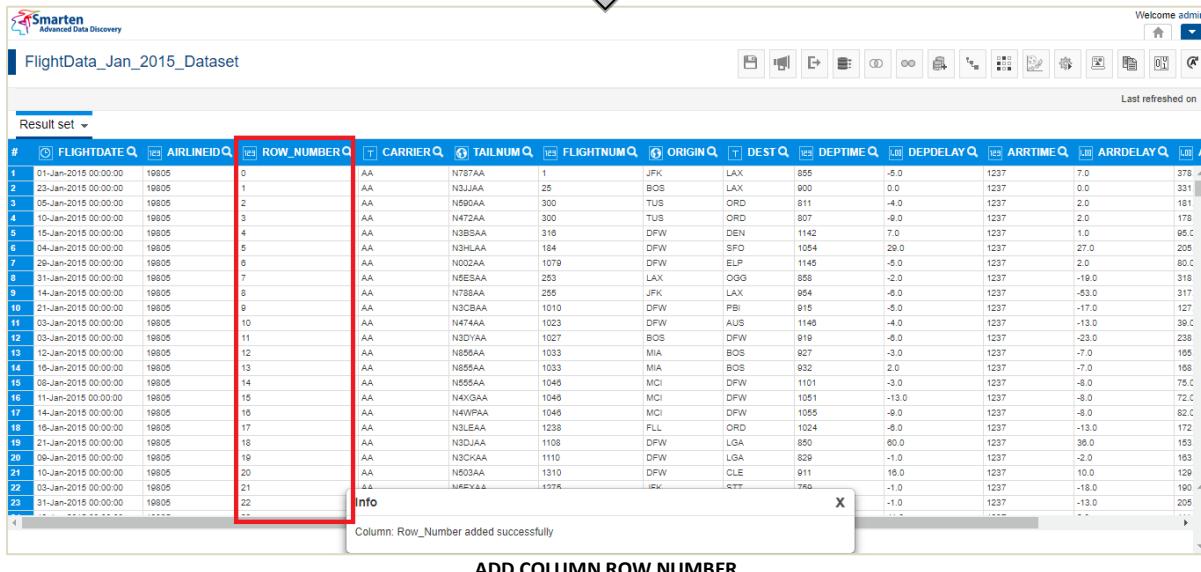
Shown below is the before and after scenario of “Add column Row Number”:

Before:



#	FLIGHDATE	AIRLINEID	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 00:00:00	19805	AA	N787AA	1	JFK	LAX	855	-5:0	1237	7:0	378.0	2475.0
2	23-Jan-2015 00:00:00	19805	AA	N3JAA	25	BOS	LAX	800	0:0	1237	0:0	331.0	2611.0
3	05-Jan-2015 00:00:00	19805	AA	N590AA	300	TUS	ORD	811	-4:0	1237	2:0	181.0	1437.0
4	10-Jan-2015 00:00:00	19805	AA	N472AA	300	TUS	ORD	807	-9:0	1237	2:0	178.0	1437.0
5	15-Jan-2015 00:00:00	19805	AA	N3BSAA	316	DFW	DEN	1142	7:0	1237	1:0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	AA	N3HLAA	184	DFW	SFO	1054	28:0	1237	27:0	205.0	1444.0
7	29-Jan-2015 00:00:00	19805	AA	N002AA	1079	DFW	ELP	1145	-5:0	1237	2:0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	AA	N5E3AA	253	LAX	OOG	858	-2:0	1237	-19:0	318.0	2486.0
9	14-Jan-2015 00:00:00	19805	AA	N788AA	255	JFK	LAX	954	-5:0	1237	-53:0	317.0	2475.0
10	21-Jan-2015 00:00:00	19805	AA	N3CBA	1010	DFW	PBI	915	-5:0	1237	-17:0	127.0	1102.0
11	03-Jan-2015 00:00:00	19805	AA	N474AA	1023	DFW	AUS	1148	-4:0	1237	-13:0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	AA	N3DYAA	1027	BOS	DFW	919	-6:0	1237	-23:0	238.0	1582.0
13	12-Jan-2015 00:00:00	19805	AA	N856AA	1033	MIA	BOS	927	-3:0	1237	-7:0	165.0	1259.0
14	16-Jan-2015 00:00:00	19805	AA	N855AA	1033	MIA	BOS	932	2:0	1237	-7:0	168.0	1259.0
15	08-Jan-2015 00:00:00	19805	AA	N555AA	1046	MCI	DFW	1101	-3:0	1237	-8:0	76.0	460.0
16	11-Jan-2015 00:00:00	19805	AA	N4XGAA	1046	MCI	DFW	1051	-13:0	1237	-8:0	72.0	460.0
17	14-Jan-2015 00:00:00	19805	AA	N4WPAA	1046	MCI	DFW	1055	-9:0	1237	-8:0	82.0	460.0
18	16-Jan-2015 00:00:00	19805	AA	N3LEAA	1238	FLL	ORD	1024	-6:0	1237	-13:0	172.0	1162.0
19	21-Jan-2015 00:00:00	19805	AA	N3DAAA	1108	DFW	LGA	850	60:0	1237	38:0	153.0	1389.0
20	09-Jan-2015 00:00:00	19805	AA	N3CKAA	1110	DFW	LGA	829	-1:0	1237	-2:0	163.0	1389.0
21	10-Jan-2015 00:00:00	19805	AA	N503AA	1310	DFW	CLE	911	16:0	1237	10:0	129.0	1021.0
22	03-Jan-2015 00:00:00	19805	AA	N5EXAA	1275	JFK	STT	759	-1:0	1237	-18:0	190.0	1823.0
23	31-Jan-2015 00:00:00	19805	AA	N3LVA	1334	SAN	ORD	643	-1:0	1237	-13:0	205.0	1723.0

After:



#	FLIGHDATE	AIRLINEID	ROW_NUMBER	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY	AIRTIME	DISTANCE
1	01-Jan-2015 00:00:00	19805	0	AA	N787AA	1	JFK	LAX	855	-5:0	1237	7:0	378.0	2475.0
2	23-Jan-2015 00:00:00	19805	1	AA	N3JAA	25	BOS	LAX	800	0:0	1237	0:0	331.0	2611.0
3	05-Jan-2015 00:00:00	19805	2	AA	N590AA	300	TUS	ORD	811	-4:0	1237	2:0	181.0	1437.0
4	10-Jan-2015 00:00:00	19805	3	AA	N472AA	300	TUS	ORD	807	-9:0	1237	2:0	178.0	1437.0
5	15-Jan-2015 00:00:00	19805	4	AA	N3BSAA	316	DFW	DEN	1142	7:0	1237	1:0	95.0	641.0
6	04-Jan-2015 00:00:00	19805	5	AA	N3HLAA	184	DFW	SFO	1054	29:0	1237	27:0	205.0	1444.0
7	29-Jan-2015 00:00:00	19805	6	AA	N002AA	1079	DFW	ELP	1145	-5:0	1237	2:0	80.0	551.0
8	31-Jan-2015 00:00:00	19805	7	AA	N5E3AA	253	LAX	OOG	858	-2:0	1237	-19:0	318.0	2486.0
9	14-Jan-2015 00:00:00	19805	8	AA	N788AA	255	JFK	LAX	954	-5:0	1237	-53:0	317.0	2475.0
10	21-Jan-2015 00:00:00	19805	9	AA	N3CBA	1010	DFW	PBI	915	-5:0	1237	-17:0	127.0	1102.0
11	03-Jan-2015 00:00:00	19805	10	AA	N474AA	1023	DFW	AUS	1148	-4:0	1237	-13:0	39.0	190.0
12	03-Jan-2015 00:00:00	19805	11	AA	N3DYAA	1027	BOS	DFW	919	-6:0	1237	-23:0	238.0	1582.0
13	12-Jan-2015 00:00:00	19805	12	AA	N856AA	1033	MIA	BOS	927	-3:0	1237	-7:0	165.0	1259.0
14	16-Jan-2015 00:00:00	19805	13	AA	N855AA	1046	MCI	DFW	1101	-3:0	1237	-8:0	75.0	460.0
15	08-Jan-2015 00:00:00	19805	14	AA	N4XGAA	1046	MCI	DFW	1051	-13:0	1237	-8:0	72.0	460.0
16	11-Jan-2015 00:00:00	19805	15	AA	N4WPAA	1046	MCI	DFW	1055	-9:0	1237	-8:0	82.0	460.0
17	14-Jan-2015 00:00:00	19805	16	AA	N3LEAA	1238	FLL	ORD	1024	-6:0	1237	-13:0	172.0	1162.0
18	16-Jan-2015 00:00:00	19805	17	AA	N3DAAA	1108	DFW	LGA	850	60:0	1237	38:0	153.0	1389.0
19	21-Jan-2015 00:00:00	19805	18	AA	N3CKAA	1110	DFW	LGA	829	-1:0	1237	-2:0	163.0	1389.0
20	09-Jan-2015 00:00:00	19805	19	AA	N503AA	1310	DFW	CLE	911	16:0	1237	10:0	129.0	1021.0
21	10-Jan-2015 00:00:00	19805	20	AA	N5EXAA	1275	JFK	STT	759	-1:0	1237	-18:0	190.0	1823.0
22	03-Jan-2015 00:00:00	19805	21	AA	N3LVA	1334	SAN	ORD	643	-1:0	1237	-13:0	205.0	1723.0
23	31-Jan-2015 00:00:00	19805	22	AA	N5E3AA	1108	DFW	LGA	850	60:0	1237	38:0	153.0	1389.0

9.6.2 For Datetime columns

This function helps users separate the various parts of a Datetime type of column and store the information in a separate column.

Note:

This function is applicable for Datetime data only.

Shown below is a list of functions for Datetime type of columns:

Function	Description	Example (23-Jan-2015 03:45:00)
Year	Returns the year part of the date	2015
Quarter	Returns the quarter corresponding to the date with year starting from January	1
Month	Returns the month number of the date with year starting from January	1
Month name	Returns the name of the month	January
Week of Year	Returns a number (between 1 and 52) representing the week of the year starting from January	4

Week of Month	Returns a number (between 1 and 5) representing the week of the month with year starting from January	4
Day of Week	Returns a number (between 1 and 7) representing the day of the week with year starting from January	5
Weekday Name	Returns the name of the weekday	Friday
Day of Year	Returns a number representing the day of the year (an integer between 1 and 366) with year starting from January	23
Day of Month	Returns the day of the month	23
Hour	Returns the hour of a time value (an integer ranging between 0 (12:00 A.M.) to 23 (11:00 P.M.))	03
Minute	Returns the minutes of a time value (an integer ranging from 0 to 59)	45
Second	Returns the seconds of a time value (an integer in the range from 0 to 59)	00

Financial		
Function	Description	Example (23-Jan-2015 03:45:00) Financial year start - April
Year	Returns the year part of the date	2014
Quarter	Returns the quarter corresponding to the date with the year starting from the user's selected month	4
Month	Returns the month number of the date with the year starting from the user's selected month	10
Week of Year	Returns a number (between 1 and 52) representing the week of the year starting from the user's selected month	43
Week of Month	Returns a number (between 1 and 5) representing the week of the month with the year starting from the user's selected month	4
Day of Month	Returns the day of the month	23
Hour	Returns the hour of a time value (an integer ranging from 0 (12:00 A.M.) to 23 (11:00 P.M.))	03
Minute	Returns the minutes of a time value (an integer ranging from 0 to 59)	45
Second	Returns the seconds of a time value (an integer ranging from 0 to 59)	00

9.6.3 Custom

Users can quickly create their own custom columns by creating expressions based on various string, arithmetic, date, or miscellaneous statements using various arithmetic operators (such as +, -, /,) or comparison operators (such as =, >, <).

Arithmetic Functions		
Function	Description	Example
abs(number)	Return absolute value of a number, a number without its sign	Argument 1 = 32 Returns 32
	Argument 1: The number for which absolute value is required	Argument 1 = 67.98 Returns 67.98
	Returns: A number	Argument 1 = -23 Returns 23
ceil(d)	Returns the smallest whole number that is greater than or equal to a specified number	Argument 1 = 26 Returns 26
	Argument 1: The number that has to be rounded up	Argument 1 = 26.7 Returns 27
	Returns: A number	Argument 1 = -26.7 Returns -26
exp(d)	Returns the exponential value of a number	Argument 1 = 1145 Returns "Infinity"
	Argument 1: The exponent applied to base e	Argument 1 = 12 Returns 162754.79
	Returns: A number	Argument 1 = -25 Returns 0.00
fact(i)	Returns the factorial of a number	Argument 1 = 7 Returns 5040
	Argument 1: The number for which factorial is to be calculated	Argument 1 = -5 Returns NULL
	Returns: A number	
floor(d)	Returns the largest whole number that is smaller than or equal to a specified number	Argument 1 = 26 Returns 26
	Argument 1: The number to be rounded down	Argument 1 = 26.7 Returns 26
	Returns: A number	Argument 1 = -26.7 Returns -27
log(d)	Returns natural logarithm (base e) of a number	Argument 1 = 551 Returns 6.31
	Argument 1: A value greater than 0 for which logarithm is to be calculated	Argument 1 = -551 Returns NULL
	Returns: A number	Argument 1 = 551.45 Returns 6.31
logTen(d)	Returns decimal logarithm (base 10) of a number	Argument 1 = 551 Returns 2.74
	Argument 1: The value greater than 0 for which logarithm is to be calculated	Argument 1 = -551 Returns NULL
	Returns: A number	Argument 1 = 551.45 Returns 2.74

max(number, number)	Returns larger of two numbers Argument 1: First number to find out if it is larger than the second number Argument 2: Second number to find out if it is larger than the first number Returns: A number	Argument 1 = 198 Argument 2 = 1660 Returns 1660.00
	Argument 1 = 198 Argument 2 = -1660 Returns 198.00	
min(number, number)	Returns smaller of two numbers Argument 1: First number to find out if it is smaller than the second number Argument 2: Second number to find out if it is smaller than the first number Returns: A number	Argument 1 = 198 Argument 2 = 1660 Returns 198.00
	Argument 1 = 198 Argument 2 = -1660 Returns -1660.00	
mod(number, number)	Returns modulus of two numbers Argument 1: Dividend: The number to be divided Argument 2: Divisor: The number by which the dividend has to be divided Returns: A number	Argument 1 = 460 Argument 2 = 72 Returns 28.00
	Argument 1 = -460 Argument 2 = 72 Returns 44.00	
	Argument 1 = 460 Argument 2 = -72 Returns 28.00	
pi(d)	Returns pi times a number Argument 1: The number Returns: A number	Argument 1 = 641 Returns 2013.76
	Argument 1 = -3 Returns -9.42	
random(number, number)	Returns a random number between two specified numbers Argument 1: The smallest integer value Argument 2: The largest integer value Returns: A number	Argument 1 = 54 Argument 2 = 55 Returns 54.45/54.51 /54.95...
	Argument 1 = 72 Argument 2 = 80 Returns 72.89/ 73.94/75.20/76.47..	
	Argument 1 = 20 Argument 2 = -10 Returns -7.68/-9.75/-2.65/5.97	
round(d, i)	Returns a number rounded to a specified number of decimal places	Argument 1 = 12.356 Argument 2 = 1 Returns 12.40

	<p>Argument 1: The number to be rounded Argument 2: The number of places to which the number is to be rounded</p> <p>Returns: A number</p>	<p>Argument 1 = -12.356 Argument 2 = 1 Returns -12.40</p> <p>Argument 1 = 12.356 Argument 2 = 2 Returns 12.36</p> <p>Argument 1 = 12.356 Argument 2 = 3 Returns 12.35</p>
sign(d)	<p>Returns a number (-1, 0, or 1) indicating the sign of a number</p> <p>Argument 1: The number for which the algebraic sign is to be determined</p> <p>Returns: A number</p>	<p>Argument 1 = -5 Returns -1</p> <p>Argument 1 = 0 Returns 0</p> <p>Argument 1 = 29 Returns 1</p>
sqrt(d)	<p>Returns the square root of a number</p> <p>Argument 1: A positive value for which the square root is to be calculated</p> <p>Returns: A number</p>	<p>Argument 1 = 100 Returns 10.00</p> <p>Argument 1 = 588 Returns 24.24</p> <p>Argument 1 = -588 Returns NaN (Not a number)</p>
truncate(d, i)	<p>Returns a number truncated to a specified number of decimal places</p> <p>Argument 1: The number to be truncated Argument 2: The scale of the truncation</p> <p>Returns: A number</p>	<p>Argument 1 = 10.54 Argument 2 = 1 Returns 10.50</p> <p>Argument 1 = 10.54 Argument 2 = 2 Returns 10.54</p> <p>Argument 1 = 10.54 Argument 2 = 0 Returns 11.00</p> <p>Argument 1 = 10.25 Argument 2 = 0 Returns 10.00</p> <p>Argument 1 = -10.54 Argument 2 = 1 Returns -10.50</p>

String Functions		
Function	Description	Examples
asc(c)	<p>Returns ASCII value of a character</p> <p>Argument 1: The character for which the ASCII value is to be returned</p> <p>Returns: A number</p>	<p>Argument 1 = A Returns 65</p> <p>Argument 1 = a Returns 97</p> <p>Argument 1 = "1" Returns 49</p>

booleanValue ("string")	Returns the content of a string as a Boolean	Argument 1 = "True" Returns true
	Argument 1: The string from which boolean is to be returned	Argument 1 = "TRUE" Returns true
	Returns: A boolean	Argument 1 = "abc" Returns false
		Argument 1 = "748" Returns false
byteValue(object)	Returns the content of a string as a byte	Argument 1 = "N787AA" Returns 0
	Argument 1: The object from which byte is to be returned	Argument 1 = "-128" Returns -128
	Returns: A byte	Argument 1 = "-129" Returns 0
		Argument 1 = "127" Returns 127
		Argument 1 = "128" Returns 0
		Argument 1 = "120.54" Returns 120
charValue(i)	Returns the content of an integer as a character	Argument 1 = 65 Returns "A"
	Argument 1: The number from which character is to be returned	Argument 1 = 97 Returns "a"
	Returns: A character	Argument 1 = 49 Returns "1"
doubleValue(object)	Returns the content of a string as double	Argument 1 = "748" Returns 748.00
	Argument 1: The object from which double is to be returned	Argument 1 = "748.52" Returns 748.52
	Returns: A double	Argument 1 = "-748.52" Returns -748.52
		Argument 1 = "abc" Returns 0.00
		Argument 1 = "ABC" Returns 0.00
fill("string", i)	Returns a string of specified length, filled with occurrences of a specified string	Argument 1 = "N787AA" Argument 2 = 2 Returns "N7"
	Argument 1: The string that has to be filled	Argument 1 = "N787AA" Argument 2 = 6 Returns "N787AA"
	Argument 2: The length of the filled string	
	Returns: A string	Argument 1 = "N787AA" Argument 2 = 9 Returns "N787AAN78"
floatValue(object)	Returns the content of a string as a float	Argument 1 = "748" Returns 748.00
	Argument 1: The object from which float is to be returned	Argument 1 = "-748.52" Returns -748.52
		Argument 1 = "abc" Returns 0.00

	Returns: A float	Argument 1 = "ABC" Returns 0.00
indexOfChar("string", c, i)	Returns the starting position of a character within a specified string Argument 1: The string from which the index is to be returned Argument 2: The character to find the index Argument 3: The starting position of the string in number Returns: A number	Argument 1 = "N787AA" Argument 2 = '7' Argument 3 = 1 Returns 1 Argument 1 = "N787AA" Argument 2 = '7' Argument 3 = 3 Returns 3 Argument 1 = "N787AA" Argument 2 = 'A' Argument 3 = 3 Returns 4 Argument 1 = "N787AA" Argument 2 = 'A' Argument 3 = 6 Returns 5 Argument 1 = "N787AA" Argument 2 = 'Y' Argument 3 = 1 Returns -1
indexOfString("string ", "string", i)	Returns the starting position of a string within a specified string Argument 1: The string from which the index is to be returned Argument 2: The string to find index Argument 3: The starting position of the string in number Returns: A number	Argument 1 = "N208WN" Argument 2 = '208' Argument 3 = 1 Returns 1 Argument 1 = "N208WN" Argument 2 = '208' Argument 3 = 4 Returns -1 Argument 1 = "N208WN" Argument 2 = 'WN' Argument 3 = 4 Returns 4
intValue(object)	Returns the contents of a string as an integer Argument 1: The object from which integer is to be returned Returns: An integer	Argument 1 = "N787AA" Returns 0 Argument 1 = "748" Returns 748 Argument 1 = "748.52" Returns 748 Argument 1 = "-748" Returns -748
isDate("string")	Determine if the specified string contains a valid date Argument 1: The string that is to be checked Returns: A boolean	Argument 1 = "2015-01-09" Returns "true" Argument 1 = "N787AA" Returns "false"
isNull(object)	Determines if the argument is NULL Argument 1: The object that is to be checked	Argument 1 = "N787AA" Returns "false" Argument 1 = NULL Returns "true"

	Returns: A boolean	
isNumber("string")	Determines if the specified string contains a number Argument 1: The string that is to be checked Returns: A boolean	Argument 1 = "N787AA" Returns "false"
		Argument 1 = "787" Returns "true"
isTime("string")	Determines if the specified string contains a valid time Argument 1: The string that is to be checked Returns: A boolean	Argument 1 = "15:30:00" Returns "true"
		Argument 1 = "N787AA" Returns "false"
left("string", i)	Returns a specified number of characters from a string starting with the first character Argument 1: The text from which the partial words are to be returned Argument 2: The number of characters to be extracted from the beginning of the text Returns: A string	Argument 1 = "N787AA" Argument 2 = 2 Returns "N7" Argument 1 = "N787AA" Argument 2 = 8 Returns "N787AA"
leftTrim("string")	Returns a copy of a specified string with leading blanks removed Argument 1: The text for which blank spaces are to be removed from left Returns: A string	Argument 1 = " 87AA" Returns "87AA"
		Argument 1 = "87AA " Returns "87AA "
length("string")	Returns the length of a string Argument 1: The string for which length is to be checked Returns: A number	Argument 1 = "N787AA" Returns 6
		Argument 1 = "748" Returns 3
longValue(object)	Returns the content of a string as long Argument 1: The object from which long is to be returned Returns: A long	Argument 1 = "748.52" Returns 6
		Argument 1 = "-748.52" Returns 7
match("string", "string")	Returns a determination whether or not a string contains a particular pattern of characters	Argument 1 = "AA" Argument 2 = "N787AA" Returns 1

	<p>Argument 1: The text that has to be searched in argument 2 Argument 2: The text in which the argument 1 has to be searched Returns: A number</p>	<p>Argument 1 = "aa" Argument 2 = "N787AA" Returns 0</p> <p>Argument 1 = "AB" Argument 2 = "N787AA" Returns 0</p>
replace("string", i, i, "string")	<p>Returns a copy of a specified string in which a specified number of characters, starting with a specified character, have been replaced with characters from another specified string</p> <p>Argument 1: The string to be processed Argument 2: Start index Argument 3: End index Argument 4: The string to be replaced</p> <p>Returns: A string</p>	<p>Argument 1 = "N208WN" Argument 2 = 1 Argument 3 = 2 Argument 4 = "3" Returns "N308WN"</p> <p>Argument 1 = "N208WN" Argument 2 = 4 Argument 3 = 6 Argument 4 = "ML" Returns "N208ML"</p> <p>Argument 1 = "N208WN" Argument 2 = 0 Argument 3 = 1 Argument 4 = "M" Returns "M208WN"</p>
reverse("string")	<p>Reverses the order of characters in a string</p> <p>Argument 1: The text that needs to be reversed</p> <p>Returns: A string</p>	<p>Argument 1 = "N208WN" Returns "NW802N"</p>
right("string", i)	<p>Returns the specified number of characters from the end of a specified string</p> <p>Argument 1: The text from which the specified number of characters should be returned from the end Argument 2: The number of characters to be returned from the string</p> <p>Returns: A string</p>	<p>Argument 1 = "N208WN" Argument 2 = 3 Returns "8WN"</p> <p>Argument 1 = "N208WN" Argument 2 = 8 Returns N208WN"</p>
rightTrim("string")	<p>Returns a copy of a specified string with trailing blanks removed</p> <p>Argument 1: The text from which extra spaces have to be removed from the right</p> <p>Returns: A string</p>	<p>Argument 1 = "N208 " Returns "N208"</p> <p>Argument 1 = " 08WN" Returns " 08WN"</p> <p>Argument 1 = " 208W " Returns " 208W"</p>
shortValue(object)	<p>Returns contents of a string as short</p> <p>Argument 1: The object from which short to be returned</p>	<p>Argument 1 = "N787AA" Returns 0</p> <p>Argument 1 = "748" Returns 748</p> <p>Argument 1 = "748.52" Returns 748</p>

	Returns: A long	Argument 1 = “-748.52” Returns -748
space(i)	Returns the string of a specified length, filled with a specified number of spaces Argument 1: Number of space Returns: A string	Argument 1 = 5 Returns “ ”
substring(“string”, i, i)	Returns a string containing a character copied (starting at a specified position and ending at a specified position) from a specified string Argument 1: The text from which the characters have to be copied Argument 2: Starting position from which the characters have to be copied Argument 3: Ending position up to which the characters in the text are to be copied Returns: A string	Argument 1 = “N208WN” Argument 2 = 2 Argument 3 = 4 Returns “08” Argument 1 = “N208WN” Argument 2 = 2 Argument 3 = 6 Returns “08WN”
toLowerCase(“string”)	Returns a copy of a specified string with all uppercase letters converted to lowercase Argument 1: The text for which the uppercase letters are to be converted into lowercase Returns: A string	Argument 1 = “N208WN” Returns “n208wn” Argument 1 = “N208wN” Returns “n208wn” Argument 1 = “n208wn” Returns “n208wn”
toString(object)	Returns a string representation of a specified object Argument 1: The object for which string is to be returned Returns: A string	Argument 1 = 748 Returns “748” Argument 1 = 748.52 Returns “748.52” Argument 1 = -748.52 Returns “-748.52” Argument 1 = 16-02-2018 20:38:40 Returns “16-02-2018 20:38:40”
toUpperCase(“string”)	Returns a copy of a specified string with all lowercase letters converted to uppercase Argument 1: The text for which the lowercase letters are to be converted into uppercase Returns: A string	Argument 1 = “n208wn” Returns “N208WN” Argument 1 = “n208Wn” Returns “N208WN” Argument 1 = “N208WN” Returns “N208WN”
trim(“string”)	Returns a string with leading and trailing blanks removed Argument 1: The text from which the extra spaces are to be removed Returns: A string	Argument 1 = “ 08WN” Returns “08WN” Argument 1 = “N208 ” Returns “N208” Argument 1 = “ 208W ” Returns “208W”

Miscellaneous Functions		
Functions	Description	Examples
ifCase(condition, truevalue, falsevalue)	<p>Returns TRUE if the condition is validated and returns FALSE if invalidated</p> <p>Argument 1: The condition Argument 2: True value Argument 3: False value</p> <p>Returns: An object</p>	<p>Argument 1 = origin=="LAX" Argument 2 = "Los Angeles" Argument 3 = "Others"</p> <p>Returns "Los Angeles" if the value of origin is "LAX" or else returns "Others"</p>
noOfDaysByDate(startDate, endDate)	<p>Returns the number of days between a given start and end date</p> <p>Argument 1: Start date Argument 2: End date</p> <p>Returns: A number</p>	<p>Argument 1 = 2014-03-10 Argument 2 = 2014-04-10 Returns 32</p>
noOfHalfYearsByDate(startDate, endDate)	<p>Returns a number of half years between a given start and end date</p> <p>Argument 1: Start date Argument 2: End date</p> <p>Returns: A number</p>	<p>Argument 1 = 2014-01-01 Argument 2 = 2014-12-31 Returns 2</p>
		<p>Argument 1 = 2014-01-01 Argument 2 = 2014-05-31 Returns 0</p>
		<p>Argument 1 = 2014-01-01 Argument 2 = 2014-08-31 Returns 1</p>
noOfMonthsByDate(startDate, endDate)	<p>Returns the number of months between a given start and end date</p> <p>Argument 1: Start date Argument 2: End date</p> <p>Returns: A number</p>	<p>Argument 1 = 2014-01-01 Argument 2 = 2014-12-31 Returns 12</p>
		<p>Argument 1 = 2014-01-01 Argument 2 = 2014-07-10 Returns 6</p>
		<p>Argument 1 = 2014-01-01 Argument 2 = 2014-05-15 Returns 4</p>
noOfQuartersByDate(startDate, endDate)	<p>Returns a number of quarters between a given start and end date</p> <p>Argument 1: Start date Argument 2: End date</p> <p>Returns: A number</p>	<p>Argument 1 = 2014-01-01 Argument 2 = 2014-12-31 Returns 4</p>
		<p>Argument 1 = 2014-01-01 Argument 2 = 2014-08-15 Returns 2</p>
noOfWeeksByDate(startDate, endDate)	<p>Returns the number of weeks between a given start and end date</p> <p>Argument 1: Start date Argument 2: End date</p>	<p>Argument 1 = 2014-01-01 Argument 2 = 2015-01-01 Returns 52</p>
		<p>Argument 1 = 2014-01-01 Argument 2 = 2014-07-01 Returns 25</p>

	Returns: A number	Argument 1 = 2014-01-01 Argument 2 = 2014-01-03 Returns 0
		Argument 1 = 2014-01-01 Argument 2 = 2014-01-12 Returns 1
whenThen(columnname, whenvalue1, thenresult1, whenvalue2, thenresult2, ..., elseresult)	Tests values of a column or expression and returns values based on the results of the test	Argument 1 = Origin Argument 2 = "LAX" Argument 3 = "Los Angeles" Argument 4 = "JFK" Argument 5 = "John F. Kennedy" Argument 6 = "Others" Returns "Los Angeles" if the value of column "origin" is "LAX" and "John F. Kennedy" if the value is "JFK." For rest of the values, returns "Others"

Date functions		
Function	Description	Example
date(Timestamp)	Returns the date part of a Timestamp Argument 1: The timestamp for which the date has to be returned Returns: A date	Argument 1 = 2018-02-16 20:38:40 Returns 2018-02-16
dateAdd ("string", i , date)	Adds a certain date or time interval to a date Argument 1: The interval of time (where the type of interval can be: Year / Month / Day / Hour / Minute / Second) Argument 2: The number of interval to be added to the time Argument 3: The date and time to add the interval to Returns: A date	Argument 1 = "y" or "Y" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2020-02-16 20:38:40 Argument 1 = "m" or "M" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2018-04-16 20:38:40 Argument 1 = "d" or "D" Argument 2 = 10 Argument 3 = 2018-02-16 20:38:40 Returns 2018-02-26 20:38:40 Argument 1 = "h" or "H" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2018-02-16 22:38:40 Argument 1 = "n" or "N" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2018-02-16 22:40:40

		Argument 1 = "s" or "S" Argument 2 = 2 Argument 3 = 2018-02-16 20:38:40 Returns 2018-02-16 22:38:42
dateDiff ("string", date, date)	Returns the number of intervals between two dates or times Argument 1: The interval of time (where the type of interval to be calculated can be: Year / Month / Day / Hour / Minute / Second) Argument 2: The first date or time Argument 3: The second date or time Returns: A number	Argument 1 = "y" or "Y" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2016-02-16 20:30:20 Returns 2 Argument 1 = "m" or "M" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2018-05-16 20:38:40 Returns -3 Argument 1 = "d" or "D" Argument 2 = 2018-02-20 20:38:40 Argument 3 = 2018-02-16 20:38:40 Returns 4 Argument 1 = "h" or "H" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2018-02-16 10:38:40 Returns 10 Argument 1 = "n" or "N" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2018-02-16 10:18:40 Returns 10 Argument 1 = "s" or "S" Argument 2 = 2018-02-16 20:38:40 Argument 3 = 2018-02-16 10:38:10 Returns 30
datePart ("string", date)	Returns the specified part of a given date Argument 1: The interval of time (where the part of the date can be: Year / Month / Day / Hour / Minute / Second) Argument 2: The date Returns: A number	Argument 1 = "y" or "Y" Argument 2 = 2018-02-16 20:38:40 Returns 2018 Argument 1 = "m" or "M" Argument 2 = 2018-02-16 20:38:40 Returns 2 Argument 1 = "d" or "D" Argument 2 = 2018-02-16 20:38:40 Returns 16 Argument 1 = "h" or "H" Argument 2 = 2018-02-16 20:38:40 Returns 20 Argument 1 = "n" or "N" Argument 2 = 2018-02-16 20:38:40 Returns 38 Argument 1 = "s" or "S" Argument 2 = 2018-02-16 20:38:40 Returns 40
dateTime("string")	Returns contents of a string as date-time Argument 1: The string for which date-time is to be returned Returns: A date-time	Argument 1 = "2018-02-16 20:38:40" Returns 2018-02-16 20:38:40
day(date)	Returns the day of a date	Argument 1 = 2018-02-16 20:38:40

	<p>represented by a number (an integer between 1 and 31)</p> <p>Argument 1: The date or timestamp for which day part is to be returned</p> <p>Returns: A number</p>	Returns 16
dayName(date)	<p>Returns the name of the day of the week</p> <p>Argument 1: The date or timestamp for which day of the week is to be returned</p> <p>Returns: A string</p>	Argument 1 = 2018-02-16 20:38:40 Returns Friday
dayOfWeek(date)	<p>Returns a number (between 1 and 7) representing the day of the week</p> <p>Argument 1: The date or timestamp for which day of the week is to be returned</p> <p>Returns: A number</p>	Argument 1 = 2018-02-16 20:38:40 Returns 5
daysAfter(date, date)	<p>Returns the count of number of days after specified date</p> <p>Argument 1: The start date Argument 2: The end date</p> <p>Returns: A number</p>	Argument 1 = 2018-02-16 20:38:40 Argument 2 = 2018-02-10 20:38:40 Returns 6
formatDate (date, "string")	<p>Returns the date format for a given pattern</p> <p>Argument 1: The target date Argument 2: The string (where the format can be user defined, such as "dd-mm-yy hh:mm:ss")</p> <p>Returns: A date</p>	Argument 1 = 2018-02-16 Argument 2 = "yy/mm/dd" Returns 18/02/16 Argument 1 = 2018-02-16 20:38:40 Argument 2 = "MM/dd/yyyy" Returns 02/16/2018
hour(date)	<p>Returns the hour of a time value (an integer ranging between 0 [12:00 AM] to 23 [11:00 PM])</p> <p>Argument 1: The timestamp for which hours are to be returned</p> <p>Returns: A number</p>	Argument 1 = 2018-02-16 20:38:40 Returns 20
minute(date)	<p>Returns the minutes of a time value (an integer ranging from 0 to 59)</p> <p>Argument 1: The timestamp for which minutes are to be returned</p>	Argument 1 = 2018-02-16 20:38:40 Returns 38

	Returns: A number	
month(date)	<p>Returns the month (an integer between 1 and 12)</p> <p>Argument 1: The date or timestamp for which month is to be returned</p> <p>Returns: A number</p>	<p>Argument 1 = 2018-02-16 20:38:40</p> <p>Returns 2</p>
monthName(i, [b], [i])	<p>Returns the month name for a given month number</p> <p>Argument 1: The number for month</p> <p>Argument 2: True if the month name is abbreviated, otherwise False (Optional to enter. Default is False)</p> <p>Argument 3: The starting month of year in number (Optional to enter. Default is 1 for January)</p> <p>Returns: A string</p>	<p>Argument 1 = 1</p> <p>Argument 2 = True</p> <p>Argument 3 = 1</p> <p>Returns Jan</p> <hr/> <p>Argument 1 = 3</p> <p>Argument 2 = True</p> <p>Argument 3 = 4</p> <p>Returns Jun</p> <hr/> <p>Argument 1 = 9</p> <p>Argument 2 = False</p> <p>Argument 3 = 1</p> <p>Returns September</p> <hr/> <p>Argument 1 = 2</p> <p>Argument 2 = False</p> <p>Argument 3 = 12</p> <p>Returns January</p> <hr/> <p>Argument 1 = 2</p> <p>Argument 2 = ""</p> <p>Argument 3 = ""</p> <p>Returns February</p>
now()	<p>Returns the current time</p> <p>Returns: A timestamp</p>	Returns 20:38:40
relativeDate(timestamp, i)	<p>Returns the date that occurs n days after a given date</p> <p>Argument 1: The date or timestamp</p> <p>Argument 2: The number of days to be added to the date-timestamp</p> <p>Returns: A timestamp</p>	<p>Argument 1 = 2018-02-16 20:38:40</p> <p>Argument 2 = 5</p> <p>Returns 2018-02-21</p>
relativeTime(timestamp, i)	<p>Returns the time that occurs n seconds after a given time</p> <p>Argument 1: The timestamp</p> <p>Argument 2: The number of seconds to be added to the timestamp</p> <p>Returns: A timestamp</p>	<p>Argument 1 = 20:38:40</p> <p>Argument 2 = 5</p> <p>Returns 20:38:45</p>
second(timestamp)	<p>Returns the seconds of a time value (an integer in the range 0 to 59)</p> <p>Argument 1: The timestamp for which seconds are to be returned</p> <p>Returns: A number</p>	<p>Argument 1 = 2018-02-16 20:38:40</p> <p>Returns 40</p>

time(timestamp)	Returns the time part from a given timestamp as a string datatype Argument 1: The timestamp for which time part is to be returned Returns: A string	Argument 1 = 2018-02-16 20:38:40 Returns "20:38:40"
today()	Returns the current system date Returns: A date	Returns 2018-02-16
weekdayName(i, [b], [i])	Returns the day name for a given day number of a week Argument 1: The number for day of week Argument 2: True if the day name is abbreviated, otherwise False (Optional to enter. Default is False) Argument 3: The first day of the week in number (Optional to enter. Default is 1 for Sunday) Returns: A string	Argument 1 = 1 Argument 2 = True Argument 3 = 1 Returns Sun
		Argument 1 = 1 Argument 2 = False Argument 3 = 1 Returns Sunday
		Argument 1 = 5 Argument 2 = False Argument 3 = 1 Returns Thursday
		Argument 1 = 1 Argument 2 = False Argument 3 = 3 Returns Tuesday
		Argument 1 = 1 Argument 2 = False Argument 3 = 5 Returns Thursday
year(date)	Returns the year corresponding to a date (an integer between 1000 and 3000) Argument 1: The date or timestamp for which year part is to be returned Returns: A number	Argument 1 = 2018-02-16 20:38:40 Returns 2018

9.7 Mark as

Users can mark a particular column of a Dataset as a Geomap dimension, Time dimension, or Dimension column depending upon the data of that column.

Marking columns as GeoMap dimension makes them available to be used in GeoMap objects. Similarly, columns marked as Time dimension makes them available to be used in objects where they are needed for time-based analysis. Marking a column as Dimension allows them to be used as Dimension columns.

Note:

This function is applicable for String and Numeric data only.

9.7.1 Time dimension

Columns can be marked as Time dimension for the following intervals of time:

- Year
- Quarter
- Month
- Week
- Day
- Hour
- Minute
- Second

For example, let us mark the column of “DEPTIME IN HRS” that contains the departure time of flights in hours as a Time dimension of Hour. Shown below is the before and after scenario of “Mark as Time dimension” for column “DEPTIME IN HRS”:

Before:

#	FLIGHDATE	AIRLINEID	ROW_NUMBER	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME IN HRS	DEPTIME IN MIN	DEPDELAY	ARRTIME
1	January 01, 2015 00:00:00	19805	0	AA	1	N178AA	JFK	LAX	8	480	-5.0	1237
2	January 23, 2015 00:00:00	19805	1	AA	25	N3JAA	BOS	LAX	9	540	0.0	1237
3	January 05, 2015 00:00:00	19805	2	AA	300	N50AA	TUS	ORD	8	480	-4.0	1237
4	January 10, 2015 00:00:00	19805	3	AA	300	N472AA	TUS	ORD	8	480	-9.0	1237
5	January 15, 2015 00:00:00	19805	4	AA	318	N3B3AA	DFW	DEN	11	660	7.0	1237
6	January 04, 2015 00:00:00	19805	5	AA	184	N3H4AA	DFW	SFO	10	600	29.0	1237
7	January 28, 2015 00:00:00	19805	6	AA	1079	N020AA	DFW	ELP	11	660	237	1237
8	January 31, 2015 00:00:00	19805	7	AA	283	N5E5AA	LAX	OGG	8	480	0.0	1237
9	January 03, 2015 00:00:00	19805	8	AA	255	N783AA	JFK	LAX	9	540	-4.0	1237
10	January 21, 2015 00:00:00	19805	9	AA	1010	N3C9AA	DFW	PBI	9	540	7.0	1237
11	January 02, 2015 00:00:00	19805	10	AA	1023	N474AA	DFW	AUS	11	720	237	1237
12	January 03, 2015 00:00:00	19805	11	AA	1027	N3D7AA	BOS	DFW	9	540	0.0	1237
13	January 12, 2015 00:00:00	19805	12	AA	1033	N856AA	MIA	BOS	9	540	-4.0	1237
14	January 18, 2015 00:00:00	19805	13	AA	1033	N855AA	MIA	BOS	9	540	7.0	1237
15	January 08, 2015 00:00:00	19805	14	AA	1045	N555AA	MCI	DFW	11	720	237	1237
16	January 11, 2015 00:00:00	19805	15	AA	1045	N4XGAA	MCI	DFW	10	720	0.0	1237
17	January 14, 2015 00:00:00	19805	16	AA	1045	N4WPAA	MCI	DFW	10	720	-4.0	1237
18	January 18, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	10	720	60.0	1237
19	January 21, 2015 00:00:00	19805	18	AA	1108	N3DJAA	DFW	LGA	8	540	-1.0	1237
20	January 09, 2015 00:00:00	19805	19	AA	1110	N3CKAA	DFW	LGA	8	540	16.0	1237
21	January 10, 2015 00:00:00	19805	20	AA	1310	N503AA	DFW	CLE	9	540	-1.0	1237
22	January 03, 2015 00:00:00	19805	21	AA	1275	N5EXAA	JFK	STT	7	420	-1.0	1237
23	January 31, 2015 00:00:00	19805	22	AA	1334	N3LVAA	SAN	ORD	6	432	-1.0	1237

After:

The screenshot shows the Smarten Advanced Data Discovery interface. A large red arrow points downwards from the 'Before' section to the 'After' section. In the 'After' section, the 'DEPTIME' column is highlighted with a red border. A tooltip window titled 'Info' appears over the 'DEPTIME' column, containing the text 'Column: DepTime in Hrs marked as time dimension - Hour'. The main table displays flight data for January 2015, with columns including FLIGHDATE, AIRLINEID, ROW_NUMBER, CARRIER, FLIGHTNUM, TAILNUM, ORIGIN, DEST, DEPTIME (marked with a red border), DEPTIME_IN_HRS, DEPTIME_IN_MIN, DEPDELAY, and ARRTIME.

MARK AS TIME DIMENSION

9.7.2 GeoMap dimension

Columns can be marked as GeoMap dimensions of:

- Country
- State
- City
- Area
- ZIP code
- Latitude
- Longitude

For example, let us mark the column of “ORIGIN” containing the origin cities of flights as a GeoMap dimension of City. Shown below is the before and after scenario of “Mark as GeoMap dimension” for column “ORIGIN”:

Before:

The screenshot shows the Smarten interface with the 'Before' scenario for marking the 'ORIGIN' column as a GeoMap dimension. A context menu is open over the 'TUS' value in the 'ORIGIN' column, listing options such as 'Highlight', 'Unique values', 'Cluster & edit', 'Find & replace', 'Remove', 'Mark as', 'Copy', 'Sort', 'Transform', 'Add column', 'Split', 'Merge columns', 'Filter', and 'Edit row'. The 'Mark as' option is highlighted, and a dropdown menu shows 'GeoMap dimension+' followed by 'Country', 'Area', and 'City'. The main table displays flight data for January 2015, with columns including FLIGHDATE, AIRLINEID, ROW_NUMBER, CARRIER, FLIGHTNUM, TAILNUM, ORIGIN, DEST, DEPTIME, DEPDELAY, ARRTIME, and ARRDELAY.

After:

FlightData_Jan_2015_Dataset

Last refreshed on

Result set ▾

#	FLIGHTDATE	AIRLINEID	ROW_NUMBER	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY
1	January 01, 2015 00:00:00	19805	0	AA	1	N787AA	JFK	LAX	855	-5.0	1237	7.0
2	January 23, 2015 00:00:00	19805	1	AA	25	N3JAA	BOS	LAX	900	0.0	1237	0.0
3	January 05, 2015 00:00:00	19805	2	AA	300	N90AA	TUS	ORD	811	-4.0	1237	2.0
4	January 10, 2015 00:00:00	19805	3	AA	300	N472AA	TUS	ORD	807	-9.0	1237	2.0
5	January 15, 2015 00:00:00	19805	4	AA	316	N3BSAA	DFW	DEN	1142	7.0	1237	1.0
6	January 04, 2015 00:00:00	19805	5	AA	184	N3HAA	DFW	SFO	1054	20.0	1237	27.0
7	January 29, 2015 00:00:00	19805	6	AA	1079	N002AA	DFW	ELP	1145	-5.0	1237	2.0
8	January 31, 2015 00:00:00	19805	7	AA	253	N6ESAA	LAX	OGG	856	-2.0	1237	-19.0
9	January 14, 2015 00:00:00	19805	8	AA	256	N786AA	JFK	LAX	954	-6.0	1237	-63.0
10	January 21, 2015 00:00:00	19805	9	AA	1010	N3CBA	DFW	PBI	915	-5.0	1237	-17.0
11	January 03, 2015 00:00:00	19805	10	AA	1023	N474AA	DFW	AUS	1146	-4.0	1237	-13.0
12	January 03, 2015 00:00:00	19805	11	AA	1027	N3D'YAA	BOS	DFW	919	-6.0	1237	-23.0
13	January 12, 2015 00:00:00	19805	12	AA	1033	N856AA	MIA	BOS	927	-3.0	1237	-2.0
14	January 16, 2015 00:00:00	19805	13	AA	1033	N856AA	MIA	BOS	932	2.0	1237	10.0
15	January 08, 2015 00:00:00	19805	14	AA	1048	N555AA	MCI	DFW	1101	-3.0	1237	-8.0
16	January 11, 2015 00:00:00	19805	15	AA	1048	N4XGAA	MCI	DFW	1051	-13.0	1237	-72.0
17	January 14, 2015 00:00:00	19805	16	AA	1048	N4WPAA	MCI	DFW	1055	-9.0	1237	-8.0
18	January 19, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-6.0	1237	-13.0
19	January 21, 2015 00:00:00	19805	18	AA	1108	N3D'JAA	DFW	LGA	850	60.0	1237	38.0
20	January 09, 2015 00:00:00	19805	19	AA	1110	N3CKAA	DFW	LGA	829	-1.0	1237	-2.0
21	January 10, 2015 00:00:00	19805	20	AA	1310	N503AA	DFW	CLE	911	16.0	1237	10.0
22	January 03, 2015 00:00:00	19805	21	AA	1275	N5EXAA	JFK	STT	759	-1.0	1237	-18.0
23	January 31, 2015 00:00:00	19805	22	AA	1334	N3LVAA	SAN	ORD	643	-1.0	1237	-13.0

MARK AS GEOMAP DIMENSION

Similarly, let us mark the column of “DEPDELAY” as a GeoMap dimension of Latitude. Shown below is the before and after scenario of “Mark as GeoMap dimension” for column “DEPDELAY”:

Before:

FlightData_Jan_2015_Dataset

Last refreshed on

Result set ▾

#	FLIGHTDATE	AIRLINEID	ROW_NUMBER	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY
1	January 01, 2015 00:00:00	19805	0	AA	1	N787AA	JFK	LAX	855	-5.0	1237	7.0
2	January 23, 2015 00:00:00	19805	1	AA	25	N3JAA	BOS	LAX	900	0.0	1237	0.0
3	January 05, 2015 00:00:00	19805	2	AA	300	N90AA	TUS	ORD	811	-4.0	1237	2.0
4	January 10, 2015 00:00:00	19805	3	AA	300	N472AA	TUS	ORD	807	-9.0	1237	2.0
5	January 15, 2015 00:00:00	19805	4	AA	316	N3BSAA	DFW	DEN	1142	7.0	1237	1.0
6	January 04, 2015 00:00:00	19805	5	AA	184	N3HAA	DFW	SFO	1054	20.0	1237	27.0
7	January 29, 2015 00:00:00	19805	6	AA	1079	N002AA	DFW	ELP	1145	-5.0	1237	2.0
8	January 31, 2015 00:00:00	19805	7	AA	253	N6ESAA	LAX	OGG	856	-2.0	1237	-19.0
9	January 14, 2015 00:00:00	19805	8	AA	256	N786AA	JFK	LAX	954	-6.0	1237	-63.0
10	January 21, 2015 00:00:00	19805	9	AA	1010	N3CBAA	DFW	PBI	915	-5.0	1237	-17.0
11	January 03, 2015 00:00:00	19805	10	AA	1023	N474AA	DFW	AUS	1146	-4.0	1237	0.0
12	January 03, 2015 00:00:00	19805	11	AA	1027	N3D'YAA	BOS	DFW	919	-6.0	1237	-23.0
13	January 12, 2015 00:00:00	19805	12	AA	1033	N856AA	MIA	BOS	927	-3.0	1237	-10.0
14	January 16, 2015 00:00:00	19805	13	AA	1033	N856AA	MIA	BOS	932	2.0	1237	10.0
15	January 08, 2015 00:00:00	19805	14	AA	1048	N555AA	MCI	DFW	1101	-3.0	1237	-8.0
16	January 11, 2015 00:00:00	19805	15	AA	1048	N4XGAA	MCI	DFW	1051	-13.0	1237	-72.0
17	January 14, 2015 00:00:00	19805	16	AA	1048	N4WPAA	MCI	DFW	1055	-9.0	1237	-8.0
18	January 19, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-6.0	1237	-13.0
19	January 21, 2015 00:00:00	19805	18	AA	1108	N3D'JAA	DFW	LGA	850	60.0	1237	38.0
20	January 09, 2015 00:00:00	19805	19	AA	1110	N3CKAA	DFW	LGA	829	-1.0	1237	-2.0
21	January 10, 2015 00:00:00	19805	20	AA	1310	N503AA	DFW	CLE	911	16.0	1237	10.0
22	January 03, 2015 00:00:00	19805	21	AA	1275	N5EXAA	JFK	STT	759	-1.0	1237	-18.0
23	January 31, 2015 00:00:00	19805	22	AA	1334	N3LVAA	SAN	ORD	643	-1.0	1237	-13.0



After:

FlightData_Jan_2015_Dataset

Last refreshed on

Result set ▾

#	FLIGHTDATE	AIRLINEID	ROW_NUMBER	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY
1	January 01, 2015 00:00:00	19805	0	AA	1	N787AA	JFK	LAX	855	-5.0	1237	7.0
2	January 23, 2015 00:00:00	19805	1	AA	25	N3JAA	BOS	LAX	900	0.0	1237	0.0
3	January 05, 2015 00:00:00	19805	2	AA	300	N90AA	TUS	ORD	811	-4.0	1237	2.0
4	January 10, 2015 00:00:00	19805	3	AA	300	N472AA	TUS	ORD	807	-9.0	1237	2.0
5	January 15, 2015 00:00:00	19805	4	AA	316	N3BSAA	DFW	DEN	1142	7.0	1237	1.0
6	January 04, 2015 00:00:00	19805	5	AA	184	N3HAA	DFW	SFO	1054	20.0	1237	27.0
7	January 29, 2015 00:00:00	19805	6	AA	1079	N002AA	DFW	ELP	1145	-5.0	1237	2.0
8	January 31, 2015 00:00:00	19805	7	AA	253	N6ESAA	LAX	OGG	856	-2.0	1237	-19.0
9	January 14, 2015 00:00:00	19805	8	AA	256	N786AA	JFK	LAX	954	-6.0	1237	-63.0
10	January 21, 2015 00:00:00	19805	9	AA	1010	N3CBAA	DFW	PBI	915	-5.0	1237	-17.0
11	January 03, 2015 00:00:00	19805	10	AA	1023	N474AA	DFW	AUS	1146	-4.0	1237	0.0
12	January 03, 2015 00:00:00	19805	11	AA	1027	N3D'YAA	BOS	DFW	919	-6.0	1237	-23.0
13	January 12, 2015 00:00:00	19805	12	AA	1033	N856AA	MIA	BOS	927	-3.0	1237	-10.0
14	January 16, 2015 00:00:00	19805	13	AA	1033	N856AA	MIA	BOS	932	2.0	1237	10.0
15	January 08, 2015 00:00:00	19805	14	AA	1048	N555AA	MCI	DFW	1101	-3.0	1237	-8.0
16	January 11, 2015 00:00:00	19805	15	AA	1048	N4XGAA	MCI	DFW	1051	-13.0	1237	-72.0
17	January 14, 2015 00:00:00	19805	16	AA	1048	N4WPAA	MCI	DFW	1055	-9.0	1237	-8.0
18	January 19, 2015 00:00:00	19805	17	AA	1238	N3LEAA	FLL	ORD	1024	-6.0	1237	-13.0
19	January 21, 2015 00:00:00	19805	18	AA	1108	N3D'JAA	DFW	LGA	850	60.0	1237	38.0
20	January 09, 2015 00:00:00	19805	19	AA	1110	N3CKAA	DFW	LGA	829	-1.0	1237	-2.0
21	January 10, 2015 00:00:00	19805	20	AA	1310	N503AA	DFW	CLE	911	16.0	1237	10.0
22	January 03, 2015 00:00:00	19805	21	AA	1275	N5EXAA	JFK	STT	759	-1.0	1237	-18.0
23	January 31, 2015 00:00:00	19805	22	AA	1334	N3LVAA	SAN	ORD	643	-1.0	1237	-13.0

MARK AS GEOMAP DIMENSION

9.7.3 Dimension

Columns having numeric values can be marked as Dimension columns.

Note:

This function is applicable for Numeric data only.

For example, let us mark the column of “FLIGHTNUM” containing the flight numbers as a Dimension column. Shown below is the before and after scenario of “Mark as Dimension” for column “FLIGHTNUM”:

Before:

#	FLIGHDATE	AIRLINEID	ROW_NUMBER	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEDELAY	ARRTIME	ARRDELAY
1	January 01, 2015 00:00:00	19805	0	AA	1	N787AA	JFK	LAX	855	-5.0	1237	7.0
2	January 23, 2015 00:00:00	19805	1	AA	25	N3JJAA	BOS	LAX	900	0.0	1237	0.0
3	January 08, 2015 00:00:00	19805	2	AA	300	N590AA	TUS	ORD	811	-4.0	1237	2.0
4	January 10, 2015 00:00:00	19805	3	AA	300	N472AA	TUS	ORD	807	-9.0	1237	2.0
5	January 15, 2015 00:00:00	19805	4	AA	318	N5BSAA	DFW	DEN	1142	7.0	1237	1.0
6	January 04, 2015 00:00:00	19805	5	AA	184	N3HLLA	DFW	SFO	1054	29.0	1237	27.0
7	January 29, 2015 00:00:00	19805	6	AA	1079	N002AA	DFW	ELP	1145	-5.0	1237	2.0
8	January 31, 2015 00:00:00	19805	7	AA	253	N5E5AA	LAX	OGG	858	-2.0	1237	-19.0
9	January 14, 2015 00:00:00	19805	8	AA	255	N788AA	JFK	LAX	864	-6.0	1237	-53.0
10	January 21, 2015 00:00:00	19805	9	AA	1010	N5C8AA	DFW	PBI	915	-5.0	1237	-17.0
11	January 03, 2015 00:00:00	19805	10	AA	1023	NDFWAA	DFW	AUS	1148	-4.0	1237	-13.0
12	January 03, 2015 00:00:00	19805	11	AA	1027	N919AA	BOS	DFW	919	-6.0	1237	-23.0
13	January 12, 2015 00:00:00	19805	12	AA	1033	N855AA	MIA	BOS	927	-3.0	1237	-7.0
14	January 16, 2015 00:00:00	19805	13	AA	1046	N555AA	MIA	BOS	932	2.0	1237	-10.0
15	January 08, 2015 00:00:00	19805	14	AA	1046	N4XQAA	MCI	DFW	1101	-3.0	1237	-8.0
16	January 11, 2015 00:00:00	19805	15	AA	1046	N4WPAA	MCI	DFW	1055	-9.0	1237	-17.0
17	January 14, 2015 00:00:00	19805	16	AA	1045	N474AA	DFW	AUS	1148	-4.0	1237	-13.0
18	January 18, 2015 00:00:00	19805	17	AA	1238	N3DYYA	BOS	DFW	919	-6.0	1237	-23.0
19	January 21, 2015 00:00:00	19805	18	AA	1108	N3DJAA	DFW	LGA	850	60.0	1237	36.0
20	January 09, 2015 00:00:00	19805	19	AA	1110	N3CKAA	DFW	LGA	829	-1.0	1237	-2.0
21	January 10, 2015 00:00:00	19805	20	AA	1310	N503AA	DFW	CLE	911	16.0	1237	10.0
22	January 03, 2015 00:00:00	19805	21	AA	1275	N4KEYAA	JFK	STT	759	-1.0	1237	-18.0
23	January 31, 2015 00:00:00	19805	22	AA	1334	N3LVA	SAN	ORD	643	-1.0	1237	-13.0

After:

#	FLIGHDATE	AIRLINEID	ROW_NUMBER	CARRIER	FLIGHTNUM	TAILNUM	ORIGIN	DEST	DEPTIME	DEDELAY	ARRTIME	ARRDELAY
1	January 01, 2015 00:00:00	19805	0	AA	1	N787AA	JFK	LAX	855	-5.0	1237	7.0
2	January 23, 2015 00:00:00	19805	1	AA	25	N3JJAA	BOS	LAX	900	0.0	1237	0.0
3	January 08, 2015 00:00:00	19805	2	AA	300	N590AA	TUS	ORD	811	-4.0	1237	2.0
4	January 10, 2015 00:00:00	19805	3	AA	300	N472AA	TUS	ORD	807	-9.0	1237	2.0
5	January 15, 2015 00:00:00	19805	4	AA	318	N5BSAA	DFW	DEN	1142	7.0	1237	1.0
6	January 04, 2015 00:00:00	19805	5	AA	184	N3HLLA	DFW	SFO	1054	29.0	1237	27.0
7	January 29, 2015 00:00:00	19805	6	AA	1079	N002AA	DFW	ELP	1145	-5.0	1237	2.0
8	January 31, 2015 00:00:00	19805	7	AA	253	N5E5AA	LAX	OGG	858	-2.0	1237	-19.0
9	January 14, 2015 00:00:00	19805	8	AA	255	N788AA	JFK	LAX	864	-6.0	1237	-53.0
10	January 21, 2015 00:00:00	19805	9	AA	1010	N5C8AA	DFW	PBI	915	-5.0	1237	-17.0
11	January 03, 2015 00:00:00	19805	10	AA	1023	NDFWAA	DFW	AUS	1148	-4.0	1237	-13.0
12	January 03, 2015 00:00:00	19805	11	AA	1027	N3DYYA	BOS	DFW	919	-6.0	1237	-23.0
13	January 12, 2015 00:00:00	19805	12	AA	1033	N855AA	MIA	BOS	927	-3.0	1237	-7.0
14	January 16, 2015 00:00:00	19805	13	AA	1046	N555AA	MIA	BOS	932	2.0	1237	-10.0
15	January 08, 2015 00:00:00	19805	14	AA	1046	N4XQAA	MCI	DFW	1101	-3.0	1237	-8.0
16	January 11, 2015 00:00:00	19805	15	AA	1046	N4WPAA	MCI	DFW	1055	-9.0	1237	-17.0
17	January 14, 2015 00:00:00	19805	16	AA	1046	N474AA	DFW	AUS	1148	-4.0	1237	-13.0
18	January 18, 2015 00:00:00	19805	17	AA	1238	N3LVA	SAN	ORD	643	-1.0	1237	-13.0
19	January 21, 2015 00:00:00	19805	18	AA	1108	N3DJAA	DFW	LGA	850	60.0	1237	36.0
20	January 09, 2015 00:00:00	19805	19	AA	1110	N3CKAA	DFW	LGA	829	-1.0	1237	-2.0
21	January 10, 2015 00:00:00	19805	20	AA	1310	N503AA	DFW	CLE	911	16.0	1237	10.0
22	January 03, 2015 00:00:00	19805	21	AA	1275	N4KEYAA	JFK	STT	759	-1.0	1237	-18.0
23	January 31, 2015 00:00:00	19805	22	AA	1334	N3LVA	SAN	ORD	643	-1.0	1237	-13.0

MARK AS DIMENSION

9.8 Fill data

Users can fill null or empty values in rows for a column based on previous row values. This function can be applied to all data types. For numeric data type columns, apart from the “Previous value” option, users can fill data using the Mean, Median, Minimum, and Maximum value of the column.

For example, the dep_time column has some NULL values in rows 53 and 55 as shown below. Once the user fills the value with the “Previous value” option, the value of row no. 52 (1139) will be filled in row no. 53, and in the same way the value of row no. 54 (611) will be filled in row no. 55.

Before:

#	DATE	UNIQUE_CARRIER	FL_NUM	ORIGIN	DEST	DEP_TIME	DEP_DELAY	ARR_TIME	ARR_DELAY	AIR_TIME	DISTANCE
42	30 January, 2016 12:00 AM	DL	1434	HNL	LAX	1545	-2.0	2250	-24.0	256.0	2566.0
43	13 January, 2016 12:00 AM	EV	5903	ORD	GSO	1411	-1.0	1738	54.0	80.0	580.0
44	2 January, 2016 12:00 AM	EV	2816	DFW	JAN	1047	-3.0	1218	3.0	58.0	408.0
45	4 January, 2016 12:00 AM	DL	1942	EWR	ATL	1533	3.0	1742	-28.0	103.0	746.0
46	17 January, 2016 12:00 AM	WN	854	PBI	ATL	1815	35.0	2003	28.0	84.0	545.0
47	5 January, 2016 12:00 AM	DL	2655	DEN	SLC	840	-5.0	1018	-1.0	75.0	391.0
48	8 January, 2016 12:00 AM	EV	5099	DTW	ALB	1957	-3.0	2128	-17.0	70.0	489.0
49	26 January, 2016 12:00 AM	WN	611	BUR	OAK	1619	-1.0	1727	-3.0	50.0	325.0
50	17 January, 2016 12:00 AM	DL	1220	DTW	DCA	1957	-8.0	2121	-19.0	65.0	405.0
51	1 January, 2016 12:00 AM	DL	2053	FNT	1530	25.0	1714	9.0	90.0	845.0	
52	27 January, 2016 12:00 AM	WN	2346	PHX	SJC	1139	-6.0	1228	-19.0	90.0	621.0
53	21 January, 2016 12:00 AM	WN	1281	OAK	LAX	1139	NULL	NULL	NULL	NULL	337.0
54	9 January, 2016 12:00 AM	WN	3085	HOU	MCO	811	11.0	1214	NULL	NULL	849.0
55	24 January, 2016 12:00 AM	WN	564	DEN	RDU	811	NULL	NULL	NULL	NULL	1438.0
56	19 January, 2016 12:00 AM	WN	2974	BNA	HOU	1120	0.0	1331	-4.0	113.0	670.0
57	12 January, 2016 12:00 AM	EV	3848	IAH	GRK	2110	-5.0	2202	-12.0	32.0	166.0
58	13 January, 2016 12:00 AM	WN	503	LAX	LAS	1928	1.0	2022	-3.0	38.0	236.0
59	2 January, 2016 12:00 AM	WN	6147	CMH	OAK	621	1.0	827	-13.0	275.0	2110.0
60	11 January, 2016 12:00 AM	WN	2931	ATL	PBI	1958	33.0	2145	30.0	89.0	545.0
61	2 January, 2016 12:00 AM	WN	4946	RDU	TPA	857	-8.0	1047	-8.0	99.0	587.0
62	7 January, 2016 12:00 AM	WN	1857	BWI	PBI	2013	3.0	2224	-28.0	119.0	883.0
63	15 January, 2016 12:00 AM	WN	3642	SEA	SJC	645	-5.0	840	-15.0	90.0	997.0
64	19 January, 2016 12:00 AM	DL	158	DTW	BOS	1159	-4.0	1350	-10.0	92.0	632.0
65	22 January, 2016 12:00 AM	WN	2425	TPA	DEN	1703	8.0	1854	-8.0	207.0	1508.0
66	30 January, 2016 12:00 AM	WN	2755	DSM	MDW	1755	-5.0	1856	-24.0	51.0	306.0

After:

#	DATE	UNIQUE_CARRIER	FL_NUM	ORIGIN	DEST	DEP_TIME	DEP_DELAY	ARR_TIME	ARR_DELAY	AIR_TIME	DISTANCE
42	30 January, 2016 12:00 AM	WN	2754	BWI	SJC	1210	-5.0	1717	-6.0	225.0	1650.0
39	26 January, 2016 12:00 AM	DL	2531	BOS	LAX	731	-2.0	1058	-17.0	360.0	2611.0
40	16 January, 2016 12:00 AM	WN	1834	PHX	SAN	1644	-1.0	1942	-8.0	47.0	304.0
41	20 January, 2016 12:00 AM	WN	583	DEN	MSP	1822	-3.0	2118	-2.0	102.0	680.0
42	30 January, 2016 12:00 AM	DL	1434	HNL	LAX	1545	-2.0	2250	-24.0	256.0	2566.0
43	13 January, 2016 12:00 AM	EV	5903	ORD	GSO	1411	17.0	1738	54.0	80.0	580.0
44	2 January, 2016 12:00 AM	EV	2816	DFW	JAN	1047	-3.0	1218	3.0	58.0	408.0
45	4 January, 2016 12:00 AM	DL	1942	EWR	ATL	1533	3.0	1742	-28.0	103.0	746.0
46	17 January, 2016 12:00 AM	WN	854	PBI	ATL	1815	35.0	2003	28.0	84.0	545.0
47	5 January, 2016 12:00 AM	DL	2655	DEN	SLC	840	-5.0	1018	-1.0	75.0	391.0
48	8 January, 2016 12:00 AM	EV	5099	DTW	ALB	1957	-3.0	2128	-17.0	70.0	489.0
49	26 January, 2016 12:00 AM	WN	611	BUR	OAK	1619	-1.0	1727	-3.0	50.0	325.0
50	17 January, 2016 12:00 AM	DL	1220	DTW	DCA	1957	-8.0	2121	-19.0	65.0	405.0
51	1 January, 2016 12:00 AM	DL	2053	FNT	1530	25.0	1714	9.0	90.0	845.0	
52	27 January, 2016 12:00 AM	WN	2346	PHX	SJC	1139	-6.0	1228	-19.0	90.0	621.0
53	21 January, 2016 12:00 AM	WN	1281	OAK	LAX	1139	NULL	NULL	NULL	NULL	337.0
54	9 January, 2016 12:00 AM	WN	3085	HOU	MCO	811	11.0	1214	NULL	NULL	849.0
55	24 January, 2016 12:00 AM	WN	564	DEN	RDU	811	NULL	NULL	NULL	NULL	1438.0
56	19 January, 2016 12:00 AM	WN	2974	BNA	HOU	1120	0.0	1331	-4.0	113.0	670.0
57	12 January, 2016 12:00 AM	EV	3848	IAH	GRK	2110	-5.0	2202	-12.0	32.0	166.0
58	13 January, 2016 12:00 AM	WN	503	LAX	LAS	1928	1.0	2022	-3.0	38.0	236.0
59	2 January, 2016 12:00 AM	WN	6147	CMH	OAK	621	1.0	827	-13.0	275.0	2110.0
60	11 January, 2016 12:00 AM	WN	2931	ATL	PBI	1958	33.0	2145	30.0	89.0	545.0
61	2 January, 2016 12:00 AM	WN	4946	RDU	TPA	857	-8.0	1047	-8.0	99.0	587.0
62	7 January, 2016 12:00 AM	WN	1857	BWI	PBI	2013	3.0	2224	-28.0	119.0	883.0
63	15 January, 2016 12:00 AM	WN	3642	SEA	SJC	645	-5.0	840	-15.0	90.0	697.0

FILL DATA

10 Blend data

Users can blend data from many Datasets based on common fields between them. Assisted by the auto-suggestions and recommendations by Smarten SSDP, users are able to blend the data quickly, efficiently, and without any formal knowledge of SQL or scripting.

10.1 Join

Using this functionality, users can combine (JOIN) many Datasets with the help of auto-suggestion to identify possible JOINs and their relative value and strength. Users also have the facility to create a copy of resultset before applying the blend operation. Different types of Joins are available, such as Left, Inner, Right, and Outer.

The screenshot shows the Smarten interface with a 'BLEND—JOIN' dialog open. On the left, there's a table titled 'Flight_Data_Jan_2015_Dataset' with columns like FLIGHTDATE, AIRLINEID, CARRIER, TAILNUM, FLIGHTNUM, ORIGIN, DEST, and DEPTIME. On the right, the 'Blend - JOIN' dialog lists 'FlightData_Jul-Dec_2016_Dataset_8-2-18' as the target dataset. It shows three selected JOINs: Carrier (string) JOINED to CARRIER (string) at 77%, Origin (string) JOINED to ORIGIN (string) at 100%, and Dest (string) JOINED to DEST (string) at 100%. Buttons for 'APPLY' and 'CANCEL' are at the bottom.

10.1.1 Inner

Inner Join is also known as the Equi Join, and it finds and returns all rows from one or more tables as long as there is a match between the columns. An inner join focuses on the commonality between two tables. When using an inner join, there must be at least some matching data between two (or more) tables that are being compared.

For example, let us assume there are two tables; one contains information about the date of flight, origin, and departure time of different flights, and the other stores information about the destination and arrival time of those flights. The common field between these two tables is the Flight number.

DEPARTURE Table:

Flight_Num	Fl_Date	Origin	Dep_time
253	01-Jan-2016	LAX	08:58
927	10-Jan-2016	LAX	11:22
641	30-Jan-2016	LAX	08:55
491	18-Jan-2016	LAX	10:11
611	30-Jan-2016	LAX	08:30

ARRIVAL Table:

Fl_Num	Dest	Arr_time
253	OGG	15:00
611	HNL	14:25
927	SFO	13:52
1212	MSP	12:37
1337	SMF	12:37

```
SQL> SELECT FLIGHT_NUM, FL_DATE, ORIGIN, DEP_TIME, DEST, ARR_TIME
   FROM DEPARTURE
   INNER JOIN ARRIVAL
  ON DEPARTURE.FLIGHT_NUM = ARRIVAL.FL_NUM;
```

Result:

Flight_Num	Fl_Date	Origin	Dep_time	Dest	Arr_time
253	01-Jan-2016	LAX	08:58	OGG	15:00
611	30-Jan-2016	LAX	08:30	HNL	14:25
927	10-Jan-2016	LAX	11:22	SFO	13:52

10.1.2 Left

A Left Join returns all the values from the left table plus matched values from the right table or NULL in case of no matching join predicate.

For example, let us assume there are two tables; one contains information about the date of flight, origin, and departure time of different flights, and the other stores information about the destination and arrival time of those flights. The common field between these two tables is the Flight number.

DEPARTURE Table:

Flight_Num	Fl_Date	Origin	Dep_time
253	01-Jan-2016	LAX	08:58
927	10-Jan-2016	LAX	11:22
641	30-Jan-2016	LAX	08:55
491	18-Jan-2016	LAX	10:11
611	30-Jan-2016	LAX	08:30

ARRIVAL Table:

Fl_Num	Dest	Arr_time
253	OGG	15:00
611	HNL	14:25
927	SFO	13:52
1212	MSP	12:37
1337	SMF	12:37

```
SQL> SELECT FLIGHT_NUM, FL_DATE, ORIGIN, DEP_TIME, DEST, ARR_TIME
   FROM DEPARTURE
  LEFT JOIN ARRIVAL
```

ON DEPARTURE.FLIGHT_NUM = ARRIVAL.FL_NUM;

Result:

Flight_Num	Fl_Date	Origin	Dep_time	Dest	Arr_time
253	01-Jan-2016	LAX	08:58	OGG	15:00
927	10-Jan-2016	LAX	11:22	SFO	13:52
641	30-Jan-2016	LAX	08:30	NULL	NULL
491	18-Jan-2016	LAX	10:11	NULL	NULL
611	30-Jan-2016	LAX	08:30	HNL	14:25

10.1.3 Right

A Right Join returns all the values from the right table plus matched values from the left table or NULL in case of no matching join predicate.

For example, let us assume there are two tables; one contains information about the date of flight, origin, and departure time of different flights, and the other stores information about the destination and arrival time of those flights. The common field between these two tables is the Flight number.

DEPARTURE Table:

Flight_Num	Fl_Date	Origin	Dep_time
253	01-Jan-2016	LAX	08:58
927	10-Jan-2016	LAX	11:22
641	30-Jan-2016	LAX	08:55
491	18-Jan-2016	LAX	10:11
611	30-Jan-2016	LAX	08:30

ARRIVAL Table:

Fl_Num	Dest	Arr_time
253	OGG	15:00
611	HNL	14:25
927	SFO	13:52
1212	MSP	12:37
1337	SMF	12:37

```
SQL> SELECT FLIGHT_NUM, FL_DATE, ORIGIN, DEP_TIME, DEST, ARR_TIME
   FROM DEPARTURE
   RIGHT JOIN ARRIVAL
   ON DEPARTURE.FLIGHT_NUM = ARRIVAL.FL_NUM;
```

Result:

Flight_Num	Fl_Date	Origin	Dep_time	Dest	Arr_time
253	01-Jan-2016	LAX	08:58	OGG	15:00
611	30-Jan-2016	LAX	08:30	HNL	14:25
927	10-Jan-2016	LAX	11:22	SFO	13:52
NULL	NULL	NULL	NULL	MSP	12:37
NULL	NULL	NULL	NULL	SMF	12:37

10.1.4 Outer

An outer join returns a set of records (or rows) that include what an inner join would return but also includes other rows for which no corresponding match is found in the other table. It returns all data from two or more tables regardless of whether or not there is shared information. Think of a full join as simply duplicating all the specified information, but in one table rather than many separate tables. Where matching data is missing, nulls will be produced.

For example, let us assume there are two tables; one contains information about the date of flight, origin, and departure time of different flights, and the other stores information about the destination and arrival time of those flights. The common field between these two tables is the Flight number.

DEPARTURE Table:

Flight_Num	Fl_Date	Origin	Dep_time
253	01-Jan-2016	LAX	08:58
927	10-Jan-2016	LAX	11:22
641	30-Jan-2016	LAX	08:55
491	18-Jan-2016	LAX	10:11
611	30-Jan-2016	LAX	08:30

ARRIVAL Table:

Fl_Num	Dest	Arr_time
253	OGG	15:00
611	HNL	14:25
927	SFO	13:52
1212	MSP	12:37
1337	SMF	12:37

```
SQL> SELECT FLIGHT_NUM, FL_DATE, ORIGIN, DEP_TIME, DEST, ARR_TIME
   FROM DEPARTURE
   FULL OUTER JOIN ARRIVAL
   ON DEPARTURE.FLIGHT_NUM = ARRIVAL.FL_NUM;
```

Result:

Flight_Num	Fl_Date	Origin	Dep_time	Dest	Arr_time
253	01-Jan-2016	LAX	08:58	OGG	15:00
927	10-Jan-2016	LAX	11:22	SFO	13:52
641	30-Jan-2016	LAX	08:55	NULL	NULL
491	18-Jan-2016	LAX	10:11	NULL	NULL
611	30-Jan-2016	LAX	08:30	HNL	14:25
NULL	NULL	NULL	NULL	MSP	12:37
NULL	NULL	NULL	NULL	SMF	12:37

10.2 Append

Using this functionality, users can append Datasets with the help of auto-match column suggestions provided by Smarten SSDP. Users also have the facility to create a copy of resultset before applying the blend operation.

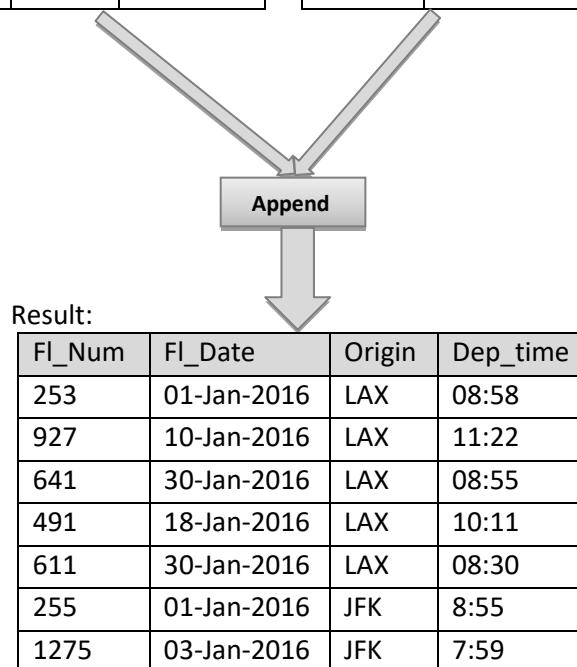
For example, let us assume there are two tables; one contains information about the flights originating from the Los Angeles airport, and the other stores information about the flights originating from the John F. Kennedy airport.

Table 1:

Fl_Num	Fl_Date	Origin	Dep_time
253	01-Jan-2016	LAX	08:58
927	10-Jan-2016	LAX	11:22
641	30-Jan-2016	LAX	08:55
491	18-Jan-2016	LAX	10:11
611	30-Jan-2016	LAX	08:30

Table 2:

Fl_Num	Fl_Date	Origin	Dep_time
255	01-Jan-2016	JFK	8:55
1275	03-Jan-2016	JFK	7:59
223	17-Jan-2016	JFK	9:30
541	29-Jan-2016	JFK	9:30
887	17-Jan-2016	JFK	10:45



223	17-Jan-2016	JFK	9:30
541	29-Jan-2016	JFK	9:30
887	17-Jan-2016	JFK	10:45

11 Sampling

Sampling is a statistical procedure that is concerned with the selection of a subset (a statistical sample) from within a statistical population. By studying the sample, we may fairly generalize our results back to the population from which they were chosen. This not only keeps the cost low but also allows analyzing the sample faster as compared with the entire population.

Smarten SSDP allows users to take a sample from a Dataset using two types of sampling methods.

Smarten SSDP recommends size of the sampling data, which can be changed by the users. Users also have a choice to enter the number of records of a Dataset to be considered as the sample size.

11.1 Sampling—Simple random sampling

Simple random sampling is a method of sampling in which the selection is based purely on a chance, and every item has an equal chance of being selected. For example, in a lottery system, each member of the population is assigned a number after which the numbers are selected at random.

Shown below is the before and after scenario of “Sampling—Simple random sampling” for a Dataset:

Before:

The screenshot shows the Smarten Advanced Data Discovery application. At the top, there's a navigation bar with the Smarten logo and a user welcome message. Below it is a toolbar with several icons. A sampling dialog box is overlaid on the main area. The dialog box has a title 'Sampling' and a sub-section 'Select sampling method' with two radio buttons: 'Simple random sampling' (selected) and 'Stratified sampling'. It also displays a note: '30% sample size has been recommended and it will result in 144 records in final sample'. Below this is a 'Change sample size' section with two radio buttons: one set to '30.0 %' and another set to '144 records'. At the bottom of the dialog are 'APPLY' and 'CANCEL' buttons. The main area of the application shows a table titled 'FlightData_Jan_2015_Dataset' with columns like '#', 'FLIGHDATE', 'AIRLINEID', 'ROW_NUMBER', 'CARRIER', 'FLIGHTNUM', 'TAILNUM', 'ORIGIN', etc. The table contains approximately 200 rows of flight data. The bottom left of the screen shows the URL 'www.ElegantJBI.com'.

After:

The screenshot shows the Smarten Advanced Data Discovery interface. At the top, there's a navigation bar with icons for home, refresh, search, and other functions. The main area is titled 'FlightData_Jan_2015_Dataset'. Below the title, it says 'Result set'. A table displays flight data with columns: #, FLIGHITDATE, AIRLINED, ROW_NUMBER, CARRIER, FLIGHTNUM, TAILNUM, ORIGIN, DEST, DEPETIME, DEPDELAY, ARRTIME, ARRDELAY, and ARRIVEDAY. The table has 23 rows of data. At the bottom of the table, a modal window titled 'Info' states 'Sampling applied on the Dataset successfully'. In the bottom right corner of the main interface, a status bar says '132 records' and 'Last refreshed on June 08, 2018 14:18:40'.

SAMPLING—SIMPLE RANDOM SAMPLING

11.2 Sampling—Stratified sampling

Stratified random sampling is a method of sampling that involves the division of a population into smaller groups known as strata. In stratified random sampling, or stratification, the strata are formed based on members' shared attributes or characteristics. For example, subgroups of customers can be formed on the basis of any of their demographics, such as income group, region, and gender. A random sample from each of these subgroups is taken in proportion to the subgroup size relative to the population size, and these samples are then added to form a final stratified random sample. So, if the original Dataset had a 1:3 ratio of males and females, then the stratified random sample based on gender will also have a 1:3 ratio of males and females.

For example, let us take “CARRIER” as the class to create a sample from a Dataset, and there are 5 subgroups within carriers, which are WN, DL, EV, AA, and UA. Smarten SSDP will take a random sample from each of these subgroups in proportion to the subgroup size relative to the Dataset size.

Original data

CARRIER	No of records
WN	101
DL	71
EV	62
AA	50
UA	32

Total number of records in the Dataset: 480

30% Sampling

Sample data

CARRIER	No of records (approx)
WN	34
DL	19
EV	17
AA	16
UA	5

Total number of records in the sample (approx): 135

Showed below is the before and after scenario of “Sampling—Stratified sampling” for a Dataset using CARRIER as the subgroup:

Before:

The screenshot shows the Smarten interface with the 'FlightData_Jan_2015_Dataset' loaded. A sampling dialog box is overlaid on the screen. The dialog title is 'Sampling'. It contains two radio button options: 'Simple random sampling' and 'Stratified sampling'. The 'Stratified sampling' option is selected. Below it, a note says 'Stratified sampling is done with respect to a particular class to come up with proportionate class counts in a sample. Please select the class' with a dropdown menu set to 'CARRIER'. At the bottom of the dialog, it says '30% sample size has been recommended and it will result in 144 records in final sample'. There are two radio buttons for 'Change sample size': '30.0 %' and '144 records'. The '144 records' button is selected. Buttons for 'APPLY' and 'CANCEL' are at the bottom right.

After:

The screenshot shows the Smarten interface with the 'FlightData_Jan_2015_Dataset' loaded. The sampling dialog box is no longer present. The top status bar shows '135 records' and 'Last refreshed on June 08, 2018 14:18:40'. A large downward arrow is centered between the 'Before' and 'After' sections.

SAMPLING—STRATIFIED SAMPLING

12 Publish and Refresh Dataset

Smarten SSDP allows users to shape raw data by splitting, merging, sorting, and copying it or by adding custom columns and marking it as per requirement.

12.1 Publish Dataset

Publishing a Dataset makes a Dataset available for collaboration and reuse by other Smarten users, who have access to a full suite of Smart Data Visualization, Plug n' Play Predictive Analysis, Dashboards, and Reports to analyze, present, and share results from the published Datasets. By default, a Dataset is unpublished, and only the user who created the Dataset can access that

Dataset. Once a Dataset is published, other users who have been given access rights by the creator of the Dataset can access the Dataset as per the permissions given to them.

Only those users who have been given the permission to “Write” the Dataset by the creator of the Dataset can modify that Dataset.

During the publishing process, the user can specify a Dataset as a Cache Dataset or a Real-time Dataset depending upon the business need. Datasets published as Cache Datasets store data in a columnar data structure and need to be updated periodically from the data sources with the help of a scheduler, whereas Real-Time Datasets do not store or cache any data. They extract the latest data from Data Sources as and when required.

12.2 Refresh Dataset

Datasets published as Cache Datasets store data in a columnar data structure.

Cache Datasets require scheduled updates to reflect the new data from the Data Sources. While publishing a Cache Dataset, users can create a Scheduler that can be configured to update Cache Datasets automatically as per predefined frequency and duration.

While updating the current dataset, users can also specify dependant datasets which need to be updated before updating current dataset. A dependent datasets are datasets that are used in blend operations -JOIN, APPEND or a parent dataset. This feature provides facility to rebuild dependant datasets first and then rebuild current dataset. So, current dataset contains latest data from dependant datasets as well. It helps avoid overlapping of different rebuild tasks which generally causes reflection of older data.

For example, Dataset AB is created from dataset A JOIN dataset B.

In such scenario, if we use Rebuild dependant dataset feature, while rebuilding dataset AB, system will first rebuild dataset A and dataset B first, then it will rebuild dataset AB, and dataset AB will always have latest data from data A and dataset B.

An e-mail notification is sent to the creator of the Dataset about the success or failure of the Dataset update process.

The screenshot shows the Smarten platform's data discovery interface. A central dialog box titled 'Publish Dataset' is open, overlaid on a dataset view. The dialog includes fields for 'Cache' (selected) and 'Real-Time', 'Scheduler settings', 'Frequency' (set to 'One time' on '11-5-2020'), 'Start time' (set to '0' on '0'), 'Email notification for Dataset publish process' (None selected), 'Data Refresh Scheduler' (Active selected), and options for 'Rebuild method' and 'Rebuild dependent dataset(s)'. At the bottom are 'PUBLISH' and 'CANCEL' buttons. In the background, the 'dataset-loan-test1' dataset is visible with columns '#', 'LOAN_AMOUNT', and 'APPL'. Another dataset, 'SHIP_STATUS', is also visible with columns 'SHIP_STATUS' and 'ANNUAL_INCOM'. The footer of the page includes the URL 'www.smarten.com', the text 'Powered by ElegantJ BI Version 5.0.2', and a 'SCHEDULER' link.

12.2.1 Refresh methods

User can refresh dataset using two methods – From scratch and Incremental

12.2.1.1 From scratch

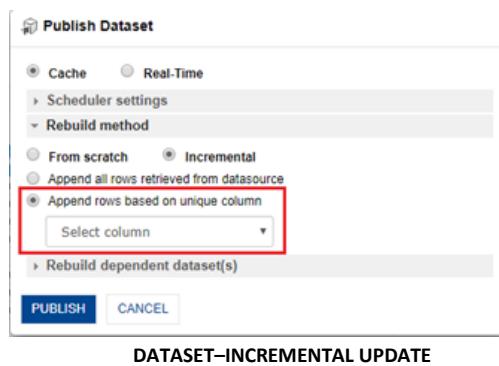
In this method, the system retrieves the latest data from the data source and overwrites the data in the dataset.

12.2.1.2 Incremental

User can update dataset with incremental option. In incremental option, system retrieves data from data source and appends only new data into the dataset. Smarten supports two options for incremental update, one is, **append all rows retrieved from data source** and another is, **append new rows identified based on unique ID column**.

For example, if you have selected the 'ID' column as a unique column from a dataset and the highest value in that column is '250' in the dataset. When you update the dataset, the system retrieves only those records that have value greaterthan '250' in the 'ID' column and appends that data to the dataset.

Same way, if you have selected the 'Date' column as a unique column from a dataset and the highest value in that column is '10-10-2020' in the dataset. When you update the dataset, the system retrieves only those records that have value greaterthan '10-10-2020' in the 'Date' column and appends that data to the dataset.



13 Dataset Management

Managing a Dataset involves providing access rights for a Dataset; editing, deleting, and copying a Dataset; changing the Data Source for a Dataset; and getting it certified by the IT department. A Dataset can be managed either by the user who created it or by the Administrator.

13.1 Manage Access Rights for Dataset

Dataset access permission is about granting or restricting access to Datasets by users. The permissions are provided to view, edit, or delete a Dataset. For example, team members who are power users require performing all actions on a Dataset and hence should be given all permissions, whereas some team members may just need to view the data, and so they should be given the permission to only view the Dataset. Access permissions can be given as per Roles or to individual users of ElegantJ BI – Smarten by the user who created the Dataset or by the Administrator. The access rights provided by the last user whether creator or Administrator are applicable.

13.2 Edit Dataset

This feature enables authorized users to edit a Dataset. Users can change the name and description of the Dataset.

Smarten SSDP also allows users to change the Data Source of the Dataset provided the columnar data structure of the new Data Source matches that of the Dataset. Users can also change the columns of the Dataset and the steps of extracting data from the new Data Source.

13.3 Delete Dataset

This feature enables authorized users to delete a Dataset. A deleted Dataset is no longer available in the system.

While deleting a Dataset, users may or may not delete Objects associated with that Dataset. Users can reuse these objects by associating them with another Dataset having the same columnar data structure as the deleted one.

13.4 Copy Dataset

This feature enables authorized users to replicate a Dataset.

It allows users to create a copy of a Dataset without going through the process of creating a Dataset from scratch.

13.5 IT Certification for Dataset

This feature enables IT staff to approve a Dataset. Marking a Dataset as IT approved certifies it for data quality and helps users in identifying quality Datasets.

NAME	CREATED	UPDATED
20170719_Tally_2	admin 19-Jul-2017 14:39:55	admin 05-Sep-2017 11:39:24
<input checked="" type="checkbox"/> ARIMA-Sales consumption forecasting-Dataset	Sahil 21-Feb-2018 15:14:38	Dhaval Oza 12-Mar-2018 13:34:53
Association Rule Mining-Product Bundling-Dataset	Jalpa 23-Feb-2018 15:35:51	Jalpa 23-Feb-2018 15:35:51
BinaryClassification-Hot lead Prediction-Dataset	Sahil 21-Feb-2018 18:46:04	Sahil 21-Feb-2018 18:46:04
BinaryClassification-Predicting Defaulters-Dataset	Sahil 21-Feb-2018 13:51:14	Sahil 21-Feb-2018 13:51:14

MARK IT APPROVED

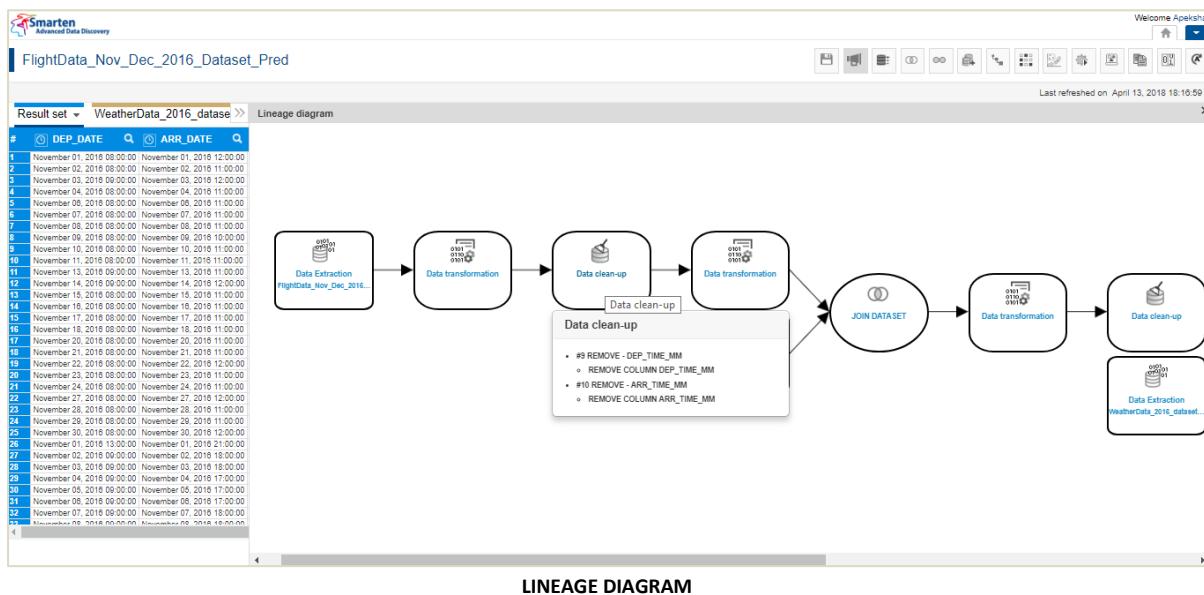
14 Action editor

Users can view all data-related actions performed on the Dataset right from the extraction process to preparing analysis-ready data. Users can roll back the effects of actions by deleting and activating or inactivating the actions. Some actions can also be edited.

#	YEAR	QUARTER	MONTH	DAY_OF_MONTH	DAY_OF_WEEK	FL_DATE	UNIQUE_CARRIER
1	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
2	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
3	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
4	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
5	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
6	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
7	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
8	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
9	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
10	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
11	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
12	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
13	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
14	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
15	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
16	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
17	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
18	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
19	2016	1	1	1	5	01-Jan-2016 00:00:00	DL
20	2016	1	1	1	5	01-Jan-2016 00:00:00	DL

15 Lineage diagram

A diagram representation of all the actions performed by a user, a Lineage diagram helps users to view the complete data flow and transformation steps in a single glance.



16 Auto-suggestions and recommendations

The intuitively guided interface of Smarten SSDP provides users with relevant auto-suggestions and recommendations to help users improve the quality of data.

Guided by smart suggestions and auto recommendations, business users can prepare, blend, and transform data and create analysis-ready data quickly and accurately without assistance or any specialized skills or scripting or advanced knowledge.

For example, Smarten SSDP will provide suggestions when there are blank values in a String column, outlier values in a Numeric column, or when all rows of a String column contain numeric values, and many other suggestions to help users create analysis-ready data quickly.

Shown below is the image of auto-suggestions and recommendations provided by Smarten SSDP for a Dataset:

Smarten Advanced Data Discovery

Welcome admin

Last refreshed on June 15, 2018 14:09:50

FlightData_Jan_2015_Dataset1

Result set ▾

#	FLIGHTDATE	AIRLINEID	ROW_NUMBER	CARRIER	TAILNUM	FLIGHTNUM	ORIGIN	DEST	DEPTIME	DEPDELAY	ARRTIME	ARRDELAY
1	January 01, 2015 00:00:00	19805	0	AA	N787AA	1	JFK	LAX	855	-5	1237	7
2	January 23, 2015 00:00:00	19805	1	AA	N3JJA	25	BOS	LAX	800	0	1237	0
3	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
4	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
6	January 04, 2015 00:00:00	19805	5	AA	N3HLA	184	DFW	SFO	1054	29	1237	27
7	January 29, 2015 00:00:00	19805	6	AA	N002AA	1079	DFW	ELP	1145	-5	1237	2
8	January 31, 2015 00:00:00	19805	7	AA	N5ESAA	253	LAX	OGG	858	-2	1237	-19
9	January 14, 2015 00:00:00	19805	8	AA	N788AA	255	JFK	LAX	854	-5	1237	-53
10	January 21, 2015 00:00:00	19805	9	AA	N3CBA	1010	DFW	PBI	915	-5	1237	-17
11	January 03, 2015 00:00:00	19805	10	AA	N474AA	1023	DFW	AUS	1148	-4	1237	-13
12	January 03, 2015 00:00:00	19805	11	AA	N3DYA	1027	BOS	DFW	919	-6	1237	-23
13	January 12, 2015 00:00:00	19805	12	AA	N858AA	1033	MIA	BOS	927	-3	1237	-7
14	January 16, 2015 00:00:00	19805	13	AA	N855AA	1033	MIA	BOS	932	2	1237	-7
15	January 08, 2015 00:00:00	19805	14	AA	N555AA	1049	MCI	DFW	1101	-3	1237	-8
16	January 11, 2015 00:00:00	19805	15	AA	N4YKA	1049	MCI	DFW	1051	-13	1237	-8
17	January 14, 2015 00:00:00	19805	16	AA	N4V1PA	1045	MCI	DFW	1055	-9	1237	-8
18	January 18, 2015 00:00:00	19805	17	AA	N3LEA	1238	FLL	ORD	1024	-6	1237	-13
19	January 21, 2015 00:00:00	19805	18	AA	N3DJAA	1108	DFW	LGA	850	60	1237	36
20	January 06, 2015 00:00:00	19805	19	AA	N3DVA	1110	DFW	LGA	850	-1	1237	-2
21	January 13, 2015 00:00:00	19805	20	AA	N3DVA	1110	DFW	LGA	850	16	1237	10
22	January 03, 2015 00:00:00	19805	21	AA	N3DVA	1110	DFW	LGA	850	-1	1237	-18
23	January 31, 2015 00:00:00	19805	22	AA	N3DVA	1110	DFW	LGA	850	-1	1237	-13

Suggestion X
NULL value(s) found in column: Row_Number. You can replace or remove NULL values using Unique values feature.

AUTO-SUGGESTIONS AND RECOMMENDATIONS—DATASET

Shown below is the image of auto-suggestions and recommendations provided by Smarten SSDP while blending Datasets using JOIN function:

Smarten Advanced Data Discovery

Welcome Apksa

Last refreshed on

Flight_Data_Jan_2015_Dataset

Result set ▾ FlightData_Jul-Dec_2016_Dataset_8-2-18

Blend - JOIN

Select dataset

Result set - Flight_Data_Jan_2015_Dataset

FlightData_Jul-Dec_2016_Dataset_8-2-18

Selected JOINS

Result set - Flight_Data_Jan_2015_Dataset

Carrier (string) 77% CARRIER (string)

Origin (string) 100% ORIGIN (string)

Dest (string) 100% DEST (string)

Create a copy of resultset before applying blend operation as: Copy_Flight_Data_Jan_2015_Dataset

APPLY CANCEL

AUTO-SUGGESTIONS AND RECOMMENDATIONS—BLEND (JOIN)

Shown below is the image of auto-suggestions and recommendations provided by Smarten SSDP while blending Datasets using APPEND function:

The screenshot shows the Smarten Advanced Data Discovery interface. On the left, there is a table titled "FlightData_Jan_2015_Dataset" with columns: #, FLIGHTDATE, AIRLINEID, ROW_NUMBER, CARRIER, FLIGHTNUM, TAILNUM, ORIGIN, DEST, and DISTANCE. The table contains 23 rows of flight data from January 1st to January 31st, 2015. On the right, a modal dialog titled "Blend - append" is open, showing a "Selected matches" section with three pairs of datasets being compared: "Origin (string)" vs "ORIGIN (string)", "Dest (string)" vs "DEST (string)", and "Distance (double)" vs "DISTANCE (double)". Below this, there is a checkbox for "Create a copy of resultset before applying blend operation as: Copy_FlightData_Jan_2015_Dataset_1", and buttons for "APPLY" and "CANCEL".

AUTO-SUGGESTIONS AND RECOMMENDATIONS—BLEND (APPEND)

17 Product and Support Information

Find more information about ElegantJ BI-Smartten and its features at www.smarten.com

Support: support@smarten.com

Sales: sales@smarten.com

Feedback & Suggestions: support@smarten.com

Support & Knowledgebase Portal: support.smarten.com