Data Rapid

Data Rapid

Table of Contents

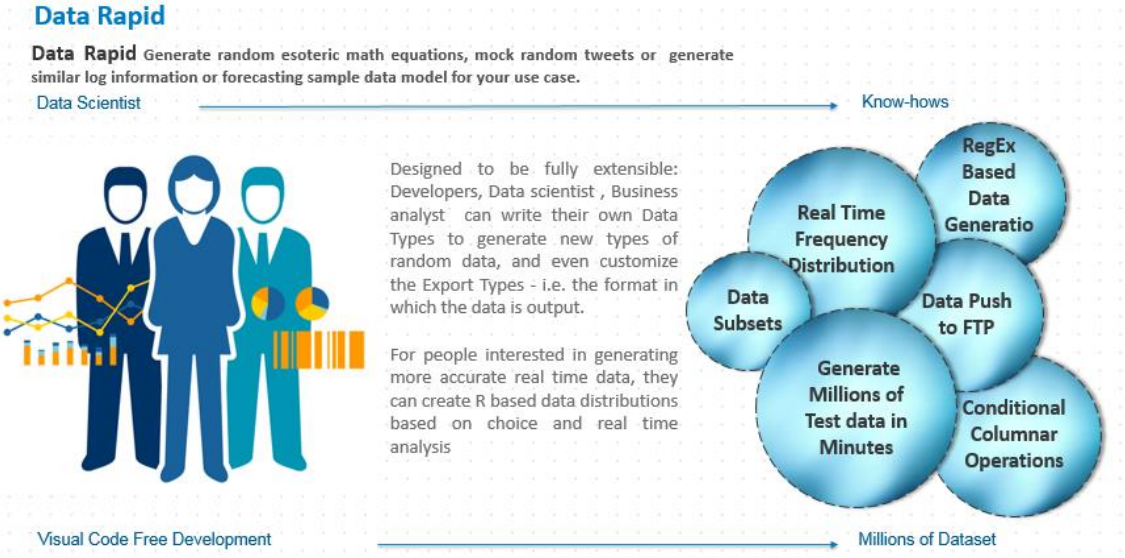
Data Rapid

1. Introduction

Data Rapid is a web based test data generation tool which uses regex pattern and predefined static content to generate the realistic data independent of any domain. Data Rapid eliminates the need of coding to generate the data. Data Rapid adds the capabilities of constant data generation and distribution data generation. With rapidly changing dynamics in market place, extracting insights really fast to support real-time business processes have become a must-have competitive advantage. Data Rapid enables enterprises to build meaningful constant and real data there by building the correlations across diverse data sets.

Generate realistic huge data-sets within seconds

Why Data Rapid? – We all have been working on different POC's (Proof of concept) and for every scenario we require huge data sets. In many cases we do not get the data from the customer and then we have to generate the data ourselves. In such a case it becomes difficult to create data manually. In such situations the data rapid comes in handy which enables us to generate huge datasets in a very short span of time.

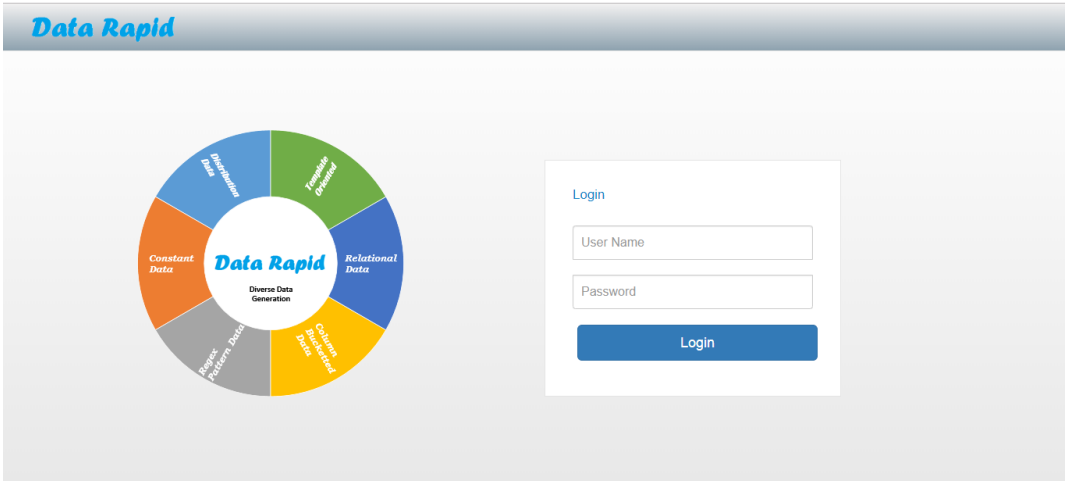


2. Data Rapid – Detailed Walkthrough

2.1. Pre Requisites

2.1.1. Data Rapid Access

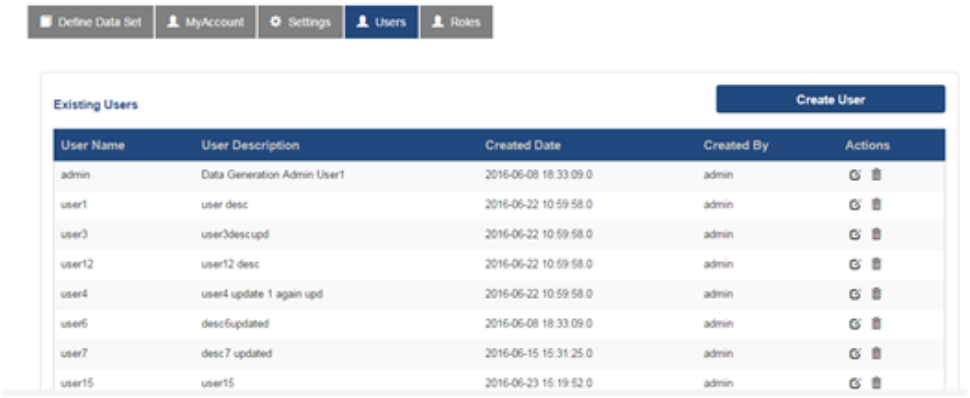
The users those who needs to generate data using the Data Rapid has to get the user and credential from the Admin team. The default credential is **admin/datagenadmin**



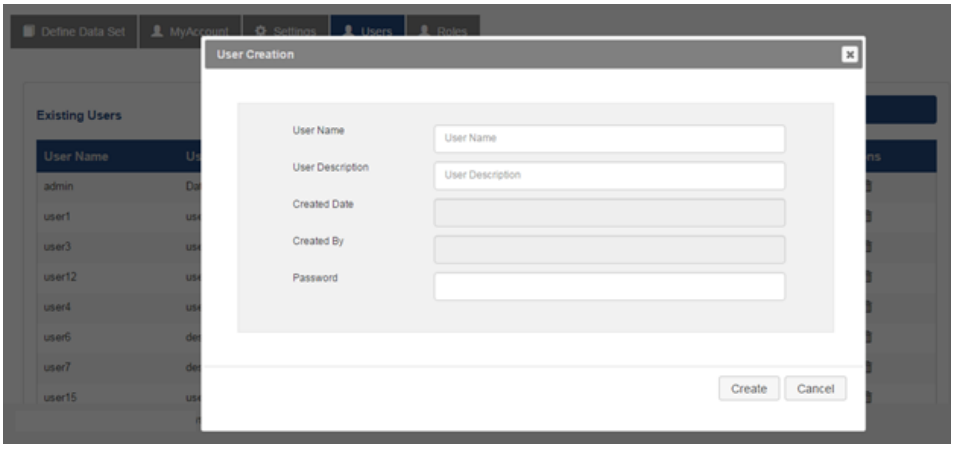
After the successful login data can be created from the define data set tab.

2.1.2. Data Rapid User Management

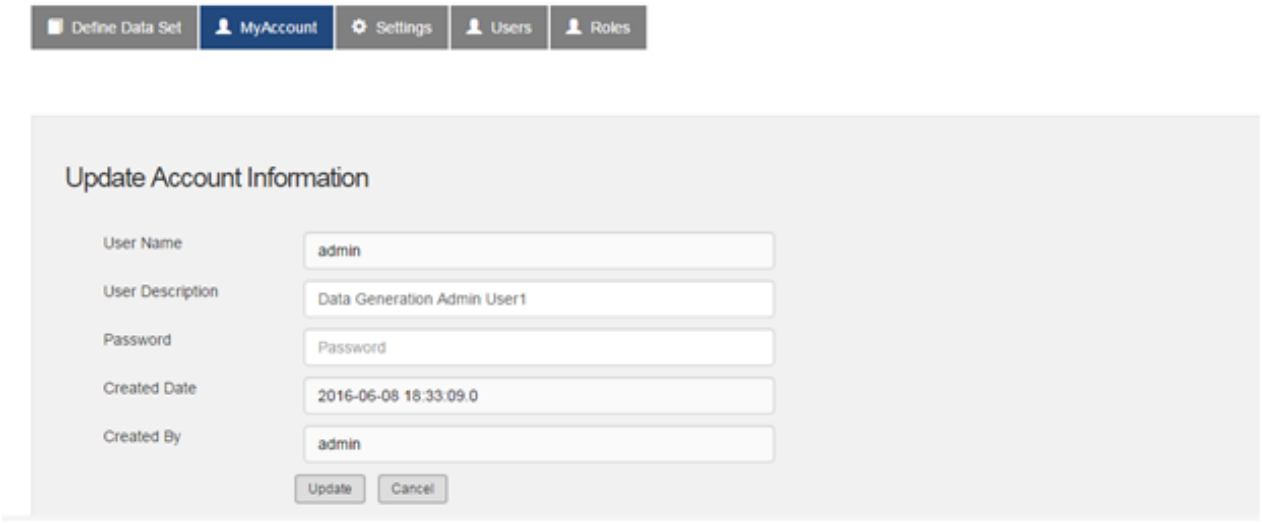
Admin user has the permission to create/update/delete user using the admin credential.



Using the create button admin can create a new user with user name, description and password.

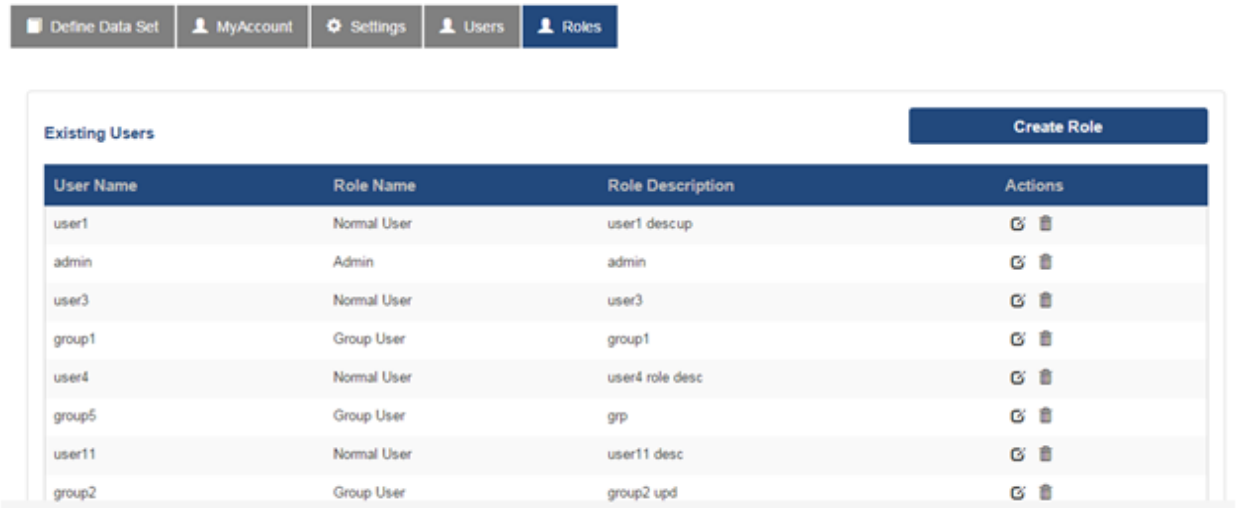


User Specific information can be viewed from the MyAccount tab. Password and description can be edited from this tab.

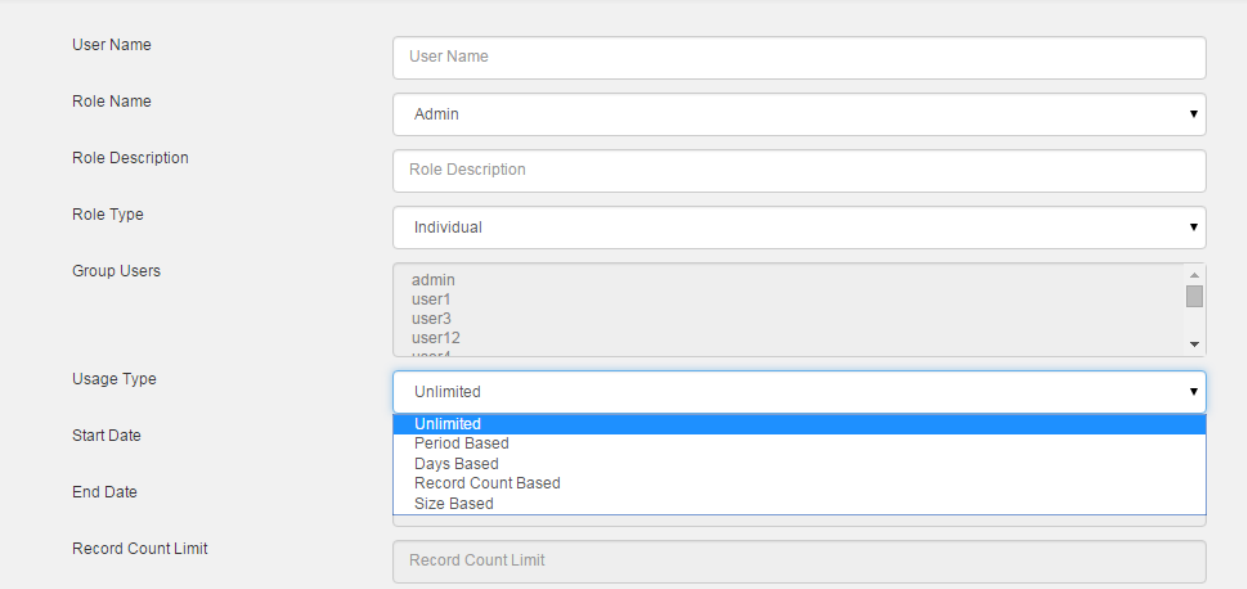


2.1.3. Data Rapid Role/License Management

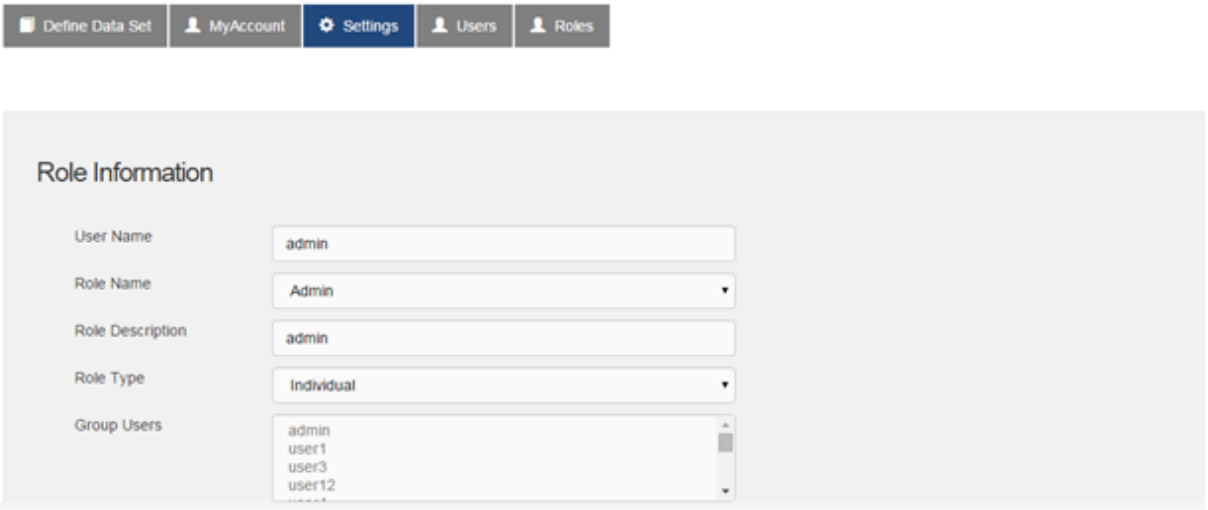
Admin user has the permission to create/update/delete user role or license using the admin credential.



Using the create button admin can create a new user role/ license with user name, role name, role description, role type, group users, usage type of the user, start date, end date, record count limit, usage limit, days limit. Any of the available license types of usages like unlimited, period based, days based, record count based, size based can be added for a user based on the licensing model.



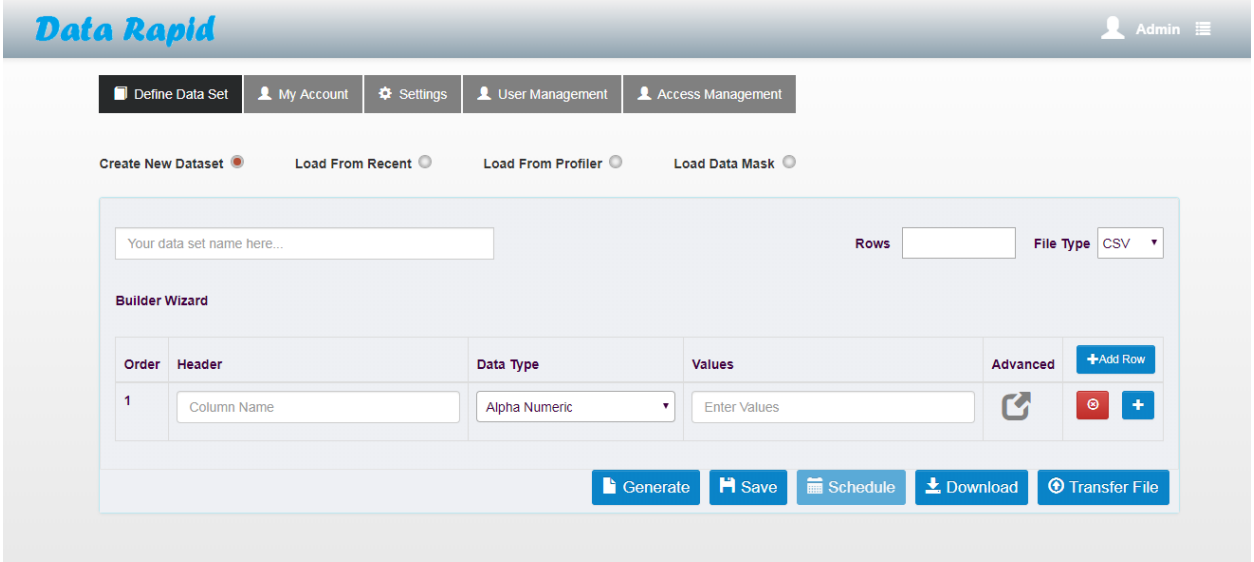
Role Specific details can be viewed from the setting tab



2.2. Data Rapid Builder Details

The Data Rapid can be used to generate data for any of the simple as well as complex constant and real distribution data.

Data Rapid



Some of the features of this tool are as follows

- ✓ A template based data generation which has a visual code free development
- ✓ Create large volumes of data within a couple of clicks.
- ✓ Generate meaningful test data at row level.
- ✓ Extremely fast data generation (Generates more than 10 Million records in less than 3 minutes).
- ✓ Improve operational efficiency.
- ✓ Generates the statistical distribution data as well as constant data.
- ✓ Facility to group the column values into factors/segments using column bucketing functionality.
- ✓ Generate user defined data using the regex pattern.
- ✓ Generate huge log files within seconds.
- ✓ Facility to download the file directly to your system.
- ✓ Reuse the configuration after saving the dataset.
- ✓ Sftp/ftp transfer facility for the files.

2.2.1. Builder Data Types

The Data Builder feature allows us to build constant data such as country names, zip-codes, phone numbers, and alpha-numeric codes and so on. The user can generate any specific data which he wants using the simple java regex pattern. If there are use cases wherein he wants to specify certain repetitive values, he can generate that data also.

The supported data types are AlphaNumeric, AmericanCard, City, Country, CVV, Date, Default Set, Digit Format, EmailAddress, FirstName, FloatRange, Guid, HexadecimalCode, HexColors, IntegerRange, IPAddress, LastName, Location, MACAddress, MastercardNumber, NumberFormat, Discovery Card number, Password, PhoneNumber,

Data Rapid

PhoneNumberWithExt, RCommand, SSN, State, Timestamp, UniqueValues, UserDefiendRegexPattern, VisaCreditCardNumber, ZipCode, IMEI number, DiscoveryCreditCard and IncrementalUniqueValues.

2.2.1.1. Available Data Types

The following are the data builder type configuration details.

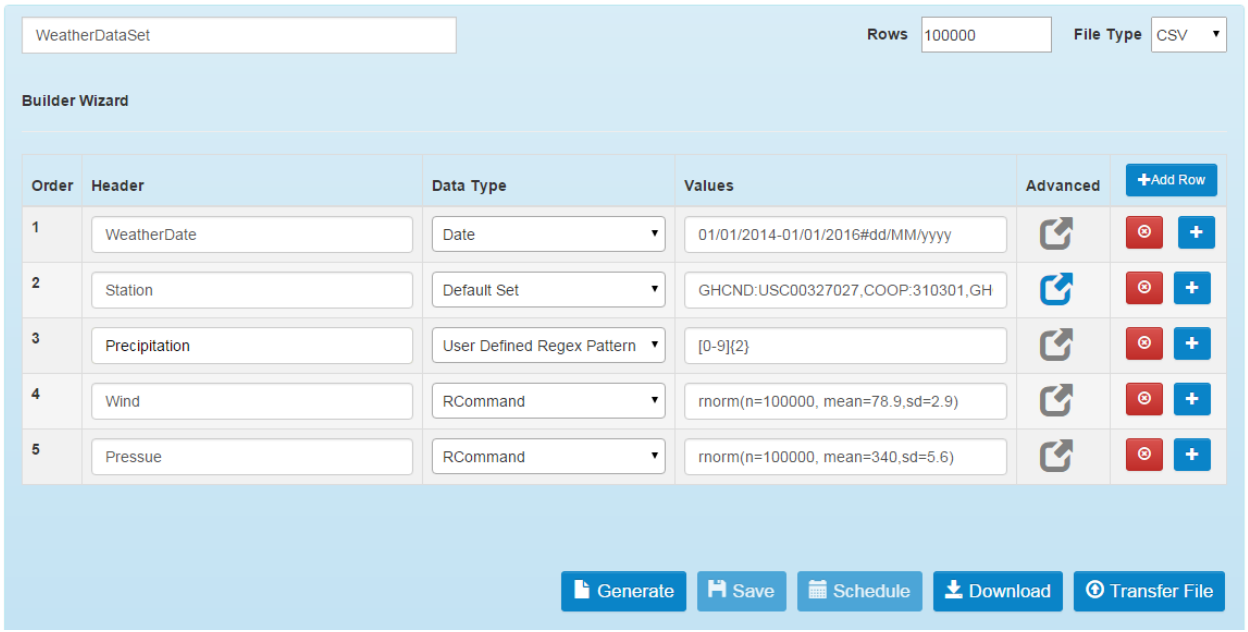
Data Rapid Data Types			
Sr No.	**Data Types**	**Comments**	**Values field Samples**
1	Alpha Numeric	Generate Alphanumeric codes based on the parameters in values field. Usage : Select drop down Alpha Numeric and set the values field in the format <numberOfDigits,numberOfAlphabets>	Eg : 3,2 Will generates an alpha numeric code like 459db
2	American Card	Generate Americancard numbers automatically	No values allowed
3	City	Generate City automatically	No values allowed
4	Country	Generate Country automatically	No values allowed
5	CVV	Generate CVV numbers automatically	No values allowed
6	Date	Generate random dates between start date and the end date based on the parameters in values field. Usage : Select drop down Date and set the values field in the format <start date in dd/MM/yyyy>-<end date in dd/MM/yyyy>#<output date format>	Eg : 01/01/2014-01/01/2016#dd/MMM/yyyy will generate dates between 01/01/2014 and 01/01/2016 in dd/MMM/yyyy format
7	Default Set	Generate user defined values from the set provided in the values field separated by comma. Usage : Select drop down Default Set and set the values field in the format <attributes separated by comma>	Eg: Machine1,Machine2,Machine3 Will generate values from any of the Machine1 or Machine2 or Machine3
8	Digit format	Generate random digits based on the parameters in values field. Usage : Select drop down Digit Format and set the values field in the format <n digit number>	Eg: 5 Will generate numbers like 45689
9	Discovery Credit Card	Generate DiscoveryCreditCard numbers automatically	No values allowed
10	Email Address	Generate emailaddress automatically	No values allowed
11	First Name	Generate firstname automatically	No values allowed
13	Float Range	Generate random float values between the given ranges. Usage : <lower range> to <upper range>	Eg : 1 to 10000 Will generate float range between 1 and 10000
14	Guid	Generate guid automatically	No values allowed

	(Globally Unique Identifier)		
15	Hexa Decimal Code	Generate hexadecimalcode automatically	No values allowed
16	Hex Colors	Generate hexcolors automatically	No values allowed
17	IncrementalUniq ue Values	Generate IncrementalUniqueValues between the range. Specify the start and end value to generate the unique key and the value by which you want to increment the values. Usage : <lower range>to<Upper Range>,<increment sequence>	Eg : 10 to 100,2 Will generate the values between 10 and 100 increment by 2
18	Integer Range	Generate random integer values between the given ranges. Usage : <lower range>to<Upper Range>	Eg : -100 to 100 Will generate numbers between -100 to 100
19	IMEI Number (International Mobile Station Equipment Identity)	Generate imei number automatically	No values allowed
20	IP Address	Generate ip-address automatically	No values allowed
21	Last Name	Generate last name automatically	No values allowed
22	Location	Generate location automatically	No values allowed
23	MAC Address	Generate macaddress automatically	No values allowed
24	MastercardNumb er	Generate mastercardnumber automatically	No values allowed
25	Number Format	Generate numbers n,m with length m and precision of m.	Eg: 4,2 Will generate numbers like 45.23
26	Password	Generate password automatically	No values allowed
27	Phone Number	Generate phonenumber automatically	No values allowed
28	Phone Number WithExt	Generate phonenumberwithext automatically	No values allowed
29	RCommand	Can be used to provide the distribution R command. Execute the rcommand specified. The value of n should be equal to the total number of rows. R Enabled distribution data generated will reverse the seed generation with respect to the expected frequency distribution like rnorm(4, mean=50, sd=10), floor(runif(100, min=0, max=101)) etc.	Eg: rnorm(n=1000,mean=100,sd=5.2) Will generate distribution values with mean 100 and standard deviation 5.2

30	SSN	Generate ssn automatically	No values allowed
31	State	Generate state automatically	No values allowed
32	Timestamp	Generate the timestamp.Enter the fromDate-ToDate to generate standard Timestamp.	Eg: 01/01/2014-01/01/2016 Will generate timestamp values between 01/01/2014 and 01/01/2016
33	Unique Values	Generate uniquevalues. Specify the start value and end value to generate the unique key.	Eg : 1 to 1000 Will generate unique values between 1 and 1000
34	User DefinedRegex Pattern	Generate user defined regex pattern.	Eg : [0-9]{2} Will generate a number with the regex pattern
35	VisaCreditCard Number	Generate visacreditcardnumber automatically	No values allowed
36	ZipCode	Generate zipcode automatically	No values allowed

2.2.1.2. Sample Screenshots for weather data set

The Weather data set configuration with weather date, station, precipitation, wind, pressure.

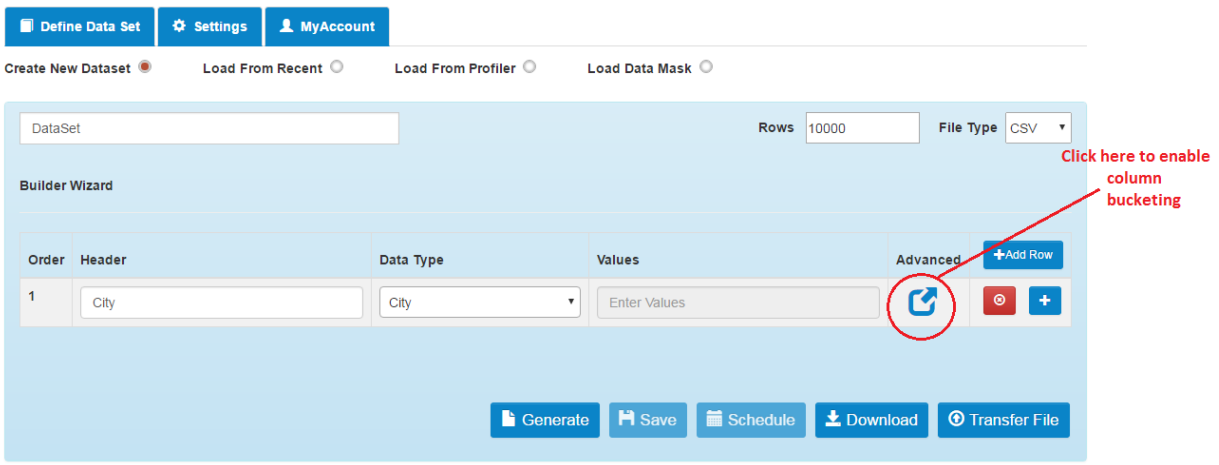


The Weather data output after execution

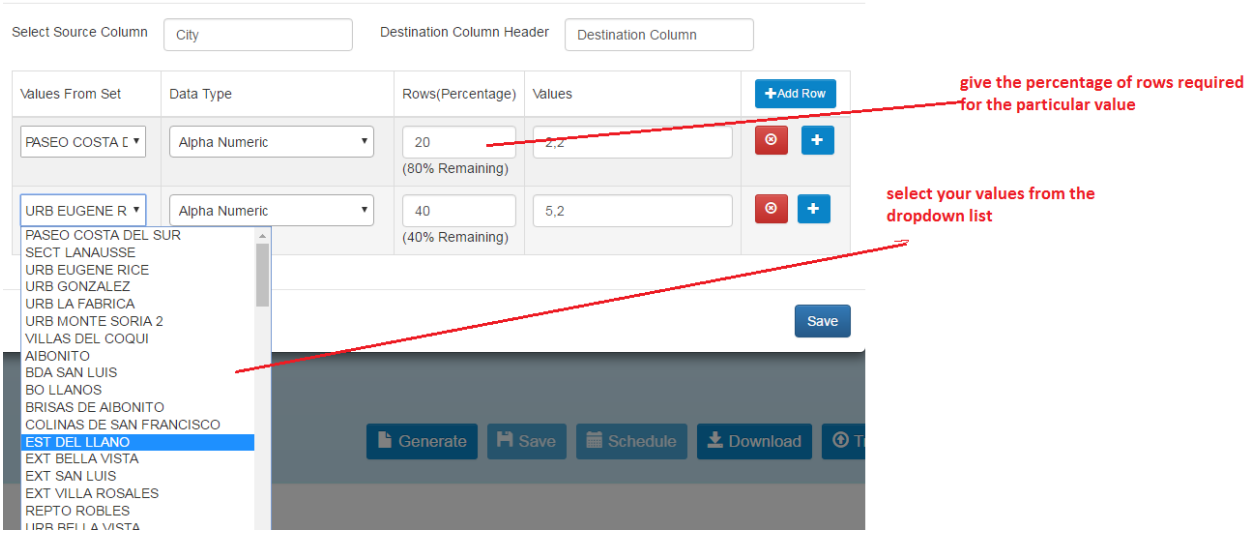
WeatherDate	Station	Precipitation	Wind	Pressue
23/07/2015	GHCND:US1MDAA0054	6	76.41587665	341.3249663
23/01/2014	GHCND:USC00185111	55	74.65827527	339.0613792
24/02/2015	COOP:310301	79	82.34218173	329.2484299
14/12/2014	GHCND:US1MDAA0054	73	81.1000724	336.612273
16/09/2014	GHCND:USC00327027	11	77.4949381	322.2757338
9/1/2014	GHCND:USC00327027	72	77.41014961	340.3022644
3/10/2015	GHCND:USC00327027	66	79.20894858	343.8944301
19/01/2014	COOP:310301	49	82.38824712	343.4221984
17/10/2015	COOP:310301	57	81.97793064	338.8873696
27/08/2014	GHCND:USC00327027	14	81.58238286	330.5650885
30/12/2015	COOP:310301	26	78.71868435	344.0371409
12/7/2014	GHCND:USC00327027	4	82.47183492	341.5226284
24/12/2014	COOP:310301	63	79.05143914	345.9040872
24/01/2014	COOP:310301	24	77.62761962	343.0703837
30/03/2014	GHCND:US1MDAA0054	25	75.81395306	342.3171025
28/02/2014	GHCND:US1MDAA0054	83	78.20148884	344.3103316
7/9/2015	GHCND:USC00327027	49	86.18755539	335.4723183
21/07/2014	GHCND:USC00185111	20	84.04989291	335.8929582
6/3/2015	COOP:310301	35	74.25520203	332.8416709
26/09/2014	GHCND:US1MDAA0054	55	77.57357246	346.1753735
12/9/2014	GHCND:USC00327027	35	73.40233912	336.8254029
30/03/2015	GHCND:USC00185111	93	79.2158134	338.8555504

2.2.2. Column Bucketing

The Column Bucketing functionality helps us to group the column values into factors/segments and to map the relations between different columns. Using this feature, we can map the values from one column to the destination column. This functionality is present for the following data types: - City, Country, Default Set, EmailAddress, FirstName, LastName, Location, State and ZipCode. When you select any of the following data types the column bucketing functionality is enabled.



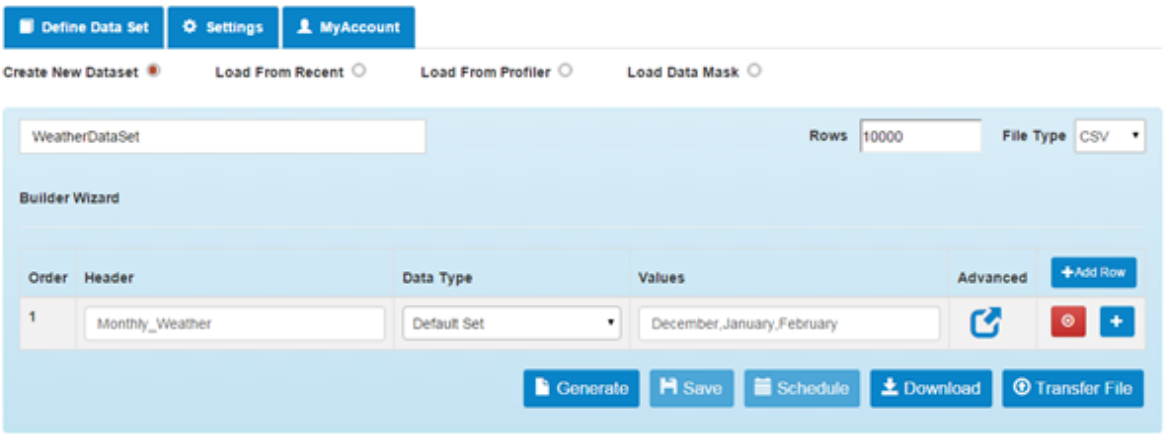
After clicking on the advanced field you are directed towards the bucketing screen wherein you can select your values and their percentage required in the destination column.



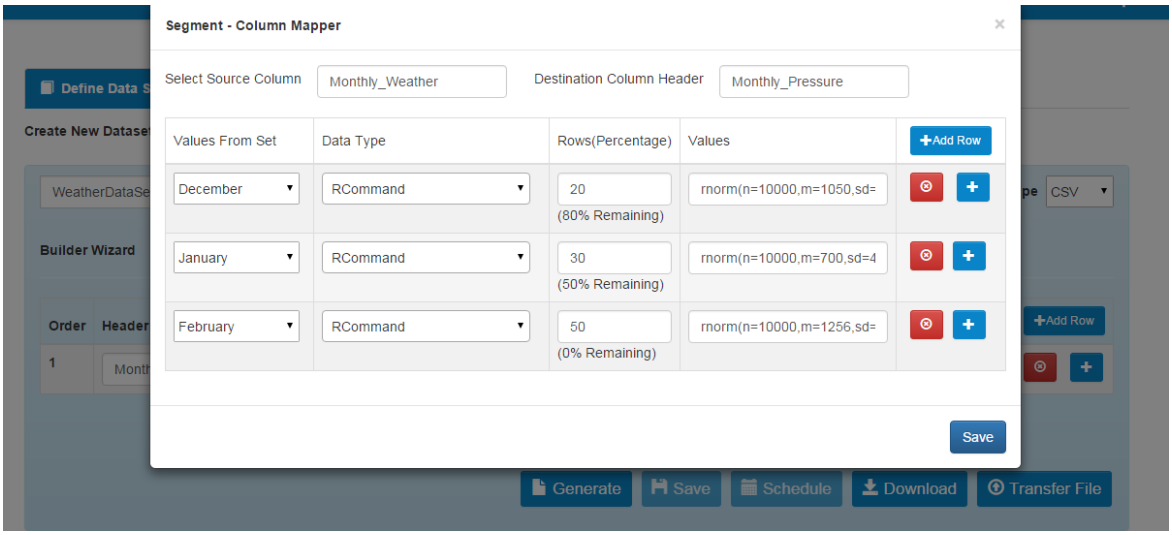
2.2.2.1. Sample Screenshots for weather data set with bucketing enabled

The sample configuration will create a monthly pressure with 20% of total rows having a distribution of defined mean and standard deviation for December, 30% of total rows having a distribution of defined mean and standard deviation for January and 50% of total rows having a distribution of defined mean and standard deviation for February.

The Weather data set configuration with weather monthly weather and pressure distribution.



The Weather data set configuration with weather monthly weather and pressure distribution.



The Weather data output after execution

Monthly_Weather	Monthly_Pressure
December	1056.478677
December	1045.643212
December	1054.050288
December	1046.276757
December	1049.45727
December	1055.257289
December	1041.921356
December	1054.399928
December	1059.581363
December	1040.192679
December	1058.962005
December	1040.91448
December	1053.716422
December	1057.899396
December	1056.29503
December	1057.16438
December	1041.725043

2.3. Sample Data Generation Details

Following are the sample data generation configurations.

2.3.1. Customer Data Set

Sr No.	Column Header	Data Type	Values	Comments
1	DateId	Unique Values	1 to 1000	Will generate the unique values from 1 till 1000
2	Date	Date	01/01/2014-01/01/2016#dd/MMM/yyyy	Will generate the dates between 1st Jan 2014 to 1st Jan 2016 and will generate the output in the user defined output format
3	CustomerId	Unique Values	1 to 1000	Will generate the unique values from 1 till 1000

4	Customer Name	First Names		Automatically Populated
5	Customer Location	Location		Automatically Populated
6	Zip Code	Zip Code		Will generate the zip codes.
7	RequestChannelId	Integer Range	1 to 4	Will generate the numbers from 1 to 4 randomly
8	RequestChannel	Default Set	Call,Email,Online,Facebook	Will generate the data from the column values that we have given.
9	TranChannelId	Integer Range	1 to 3	Will generate the numbers from 1 to 3 randomly
10	TranChannel	Default Set	Store,Mobile App,Web	Will generate the data from the column values that we have given.
11	ProblemAreaId	Integer Range	1 to 25	Will generate the numbers from 1 to 25 randomly
12	ProblemArea	Default Set	Shopping,Order,Payment,Wallet,Others,Shipping	Will generate the data from the column values that we have given.
13	ProblemType	Default Set	Ambiguios product description,Incorrect delivery,Taxes,Unable to fund,Image not available,Shipping charges	Will generate the data from the column values that we have given.
14	NumberofInteractions	Integer Range	1 to 3	Will generate the numbers from 1 to 3 randomly

2.3.2. Log File Data Set

Sr No.	Column Header	Data Type	Values	Comments
1	Date	Date	01/01/2014-01/01/2016#dd/MMM/yyyy	Will generate the dates between 1st Jan 2014 to 31 Dec 2014
2	Time	Date	01/01/2014-01/01/2016#HH:mm:ss	Will generate the time
3	cs-ip	IpAddress		Will generate the ip addresses
4	sc-bytes	Integer Range	1 to 255	Will generate the numbers from 1 to 255 randomly
5	cs-method	Default Set	GET,POST	Will generate method according to the values given
6	Phone number	PhoneNumber		Will generate the phone numbers.