DAT-A-VENGERS

1.Total credits give to every participants for whole game **75,000**.
2.On every click the credits for that feature will automatically deductued.So be patient after a click.

So lets get to the work !!

Options you are getting:-
Null values
1.Show columns that has null (CREDITS COST:-**1500**)
-It will give you a list of total no. of columns with True(if there is a null element in the column) OR False(if the column is full).

2.Number Of Null in cols (CREDITS COST:-**300**)
-You have to choose the column in which you want to check the total no. Of null elements

3.Number Of columns Having Null (CREDITS COST:-**800**)
-This will give the total number of columns which have atleast one empty data.

NORMALIZATION
Here you have to select a column name and normalistion way.

$$S = \sqrt{\frac{\Sigma(X - \overline{X})^2}{N}}$$

where S = the standard deviation of a sample,
Σ means "sum of,"
X = each value in the data set,
$\overline{X}$ = mean of all values in the data set,
N = number of values in the data set.

1ST Way (CREDITS COST:-400)
The every element(x) will be replaced by **(x - Mean) / Deviation.**

Where Mean is the mean of the data of whole column.
Deviation is the standard deviation of the data.

2ND Way (CREDITS COST:-400)

The every element(x) will be replaced by **(x - Mean)**
Where Mean is the mean of the data of whole column.

3RD Way (CREDITS COST:-400)

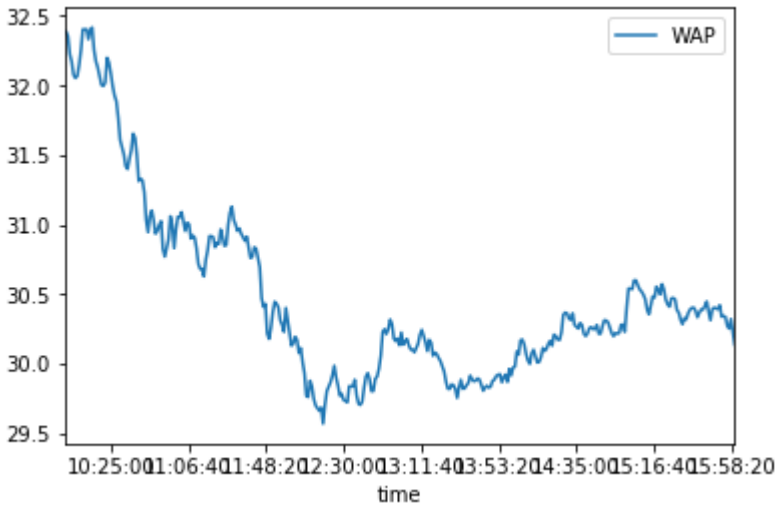
The every element(x) will be replaced by **x / Deviation.**
Where Deviation is the standard deviation of the data.

DATA VISUALIZATION

Here you have to select a graph type and column name

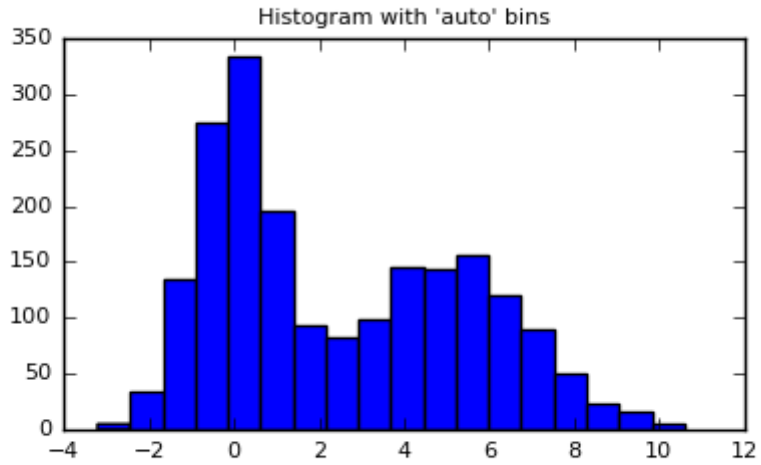
LINE (CREDITS COST:-300)

-you will get the graph of values of column vs entry number.



HISTOGRAM (CREDITS:-250)

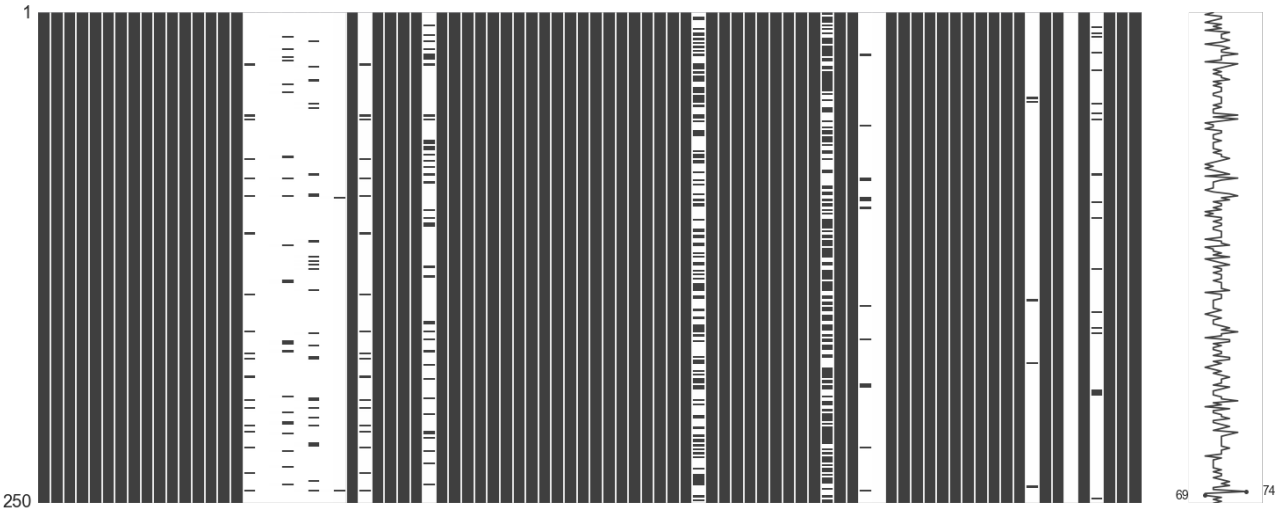
-it will give you the the graph between column entry and frequency.
It would look like..



MISSING NUMBER VISUALISATION

Matrix (CREDITS:-600)

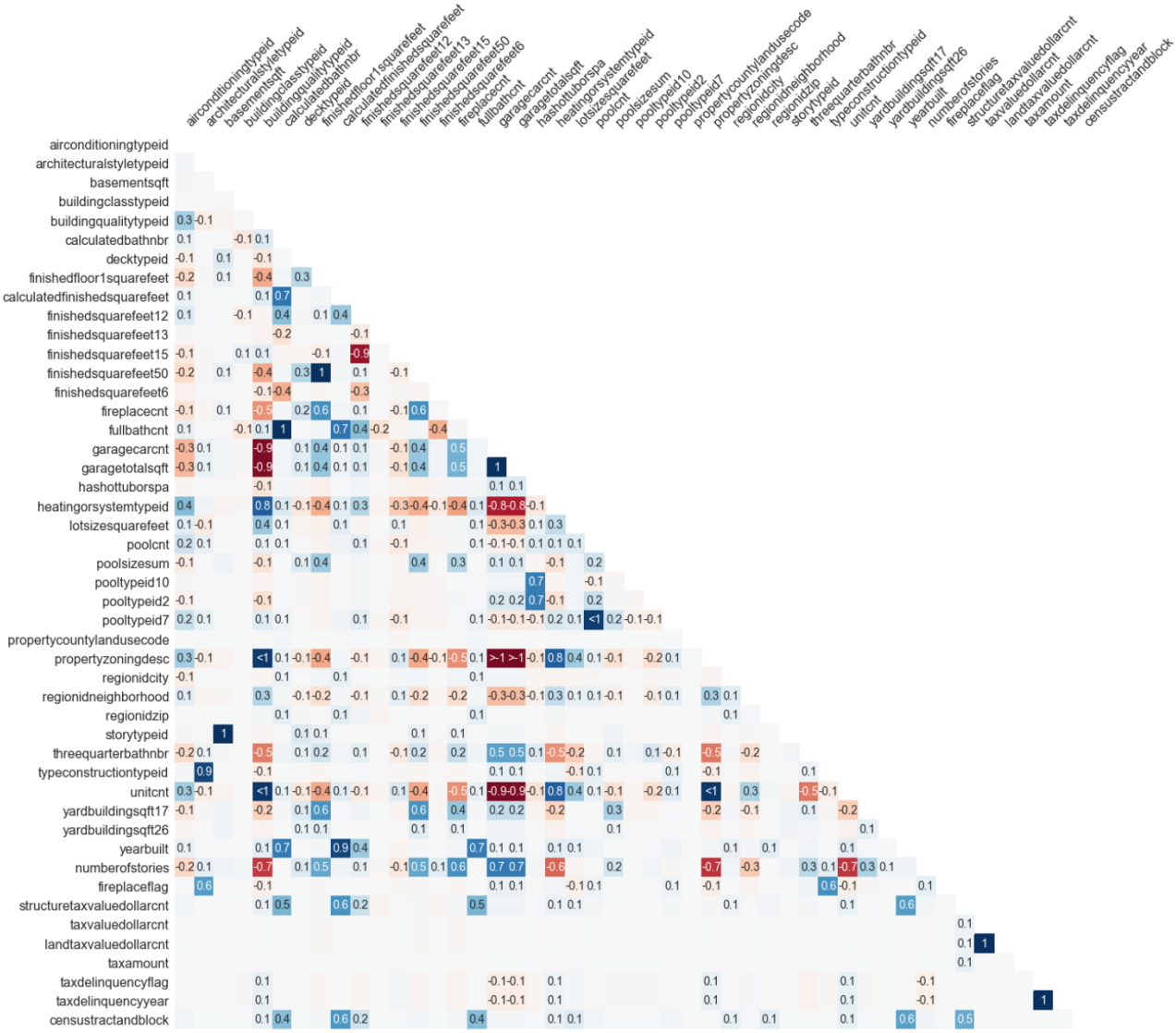
- The nullity matrix gives you a data-dense display which lets you quickly visually pick out the missing data patterns in the dataset.



HEATMAP (CREDITS:-1000)

- This map describes the degree of nullity relationship between the different features. The range of this nullity correlation is from -1 to 1 ($-1 \leq R \leq 1$). Features with no missing value are excluded in the *heatmap*. If the nullity correlation is very close to zero (-

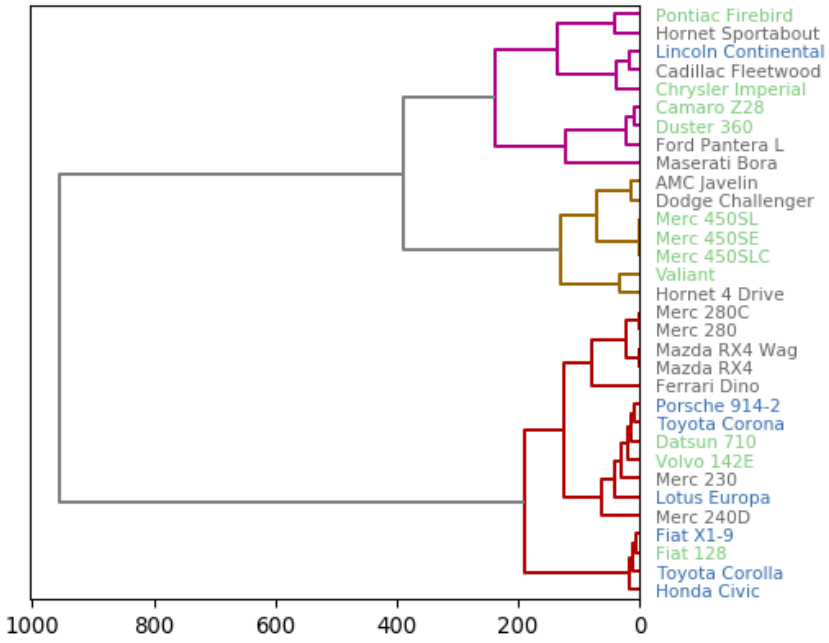
0.05 < R < 0.05), no value will be displayed. Also, a perfect positive nullity correlation (R=1) indicates when the first feature and the second feature both have corresponding missing values while a perfect negative nullity correlation (R=-1) means that one of the features is missing and the second is not missing.



DENDROGRAM (CREDITS:-700)

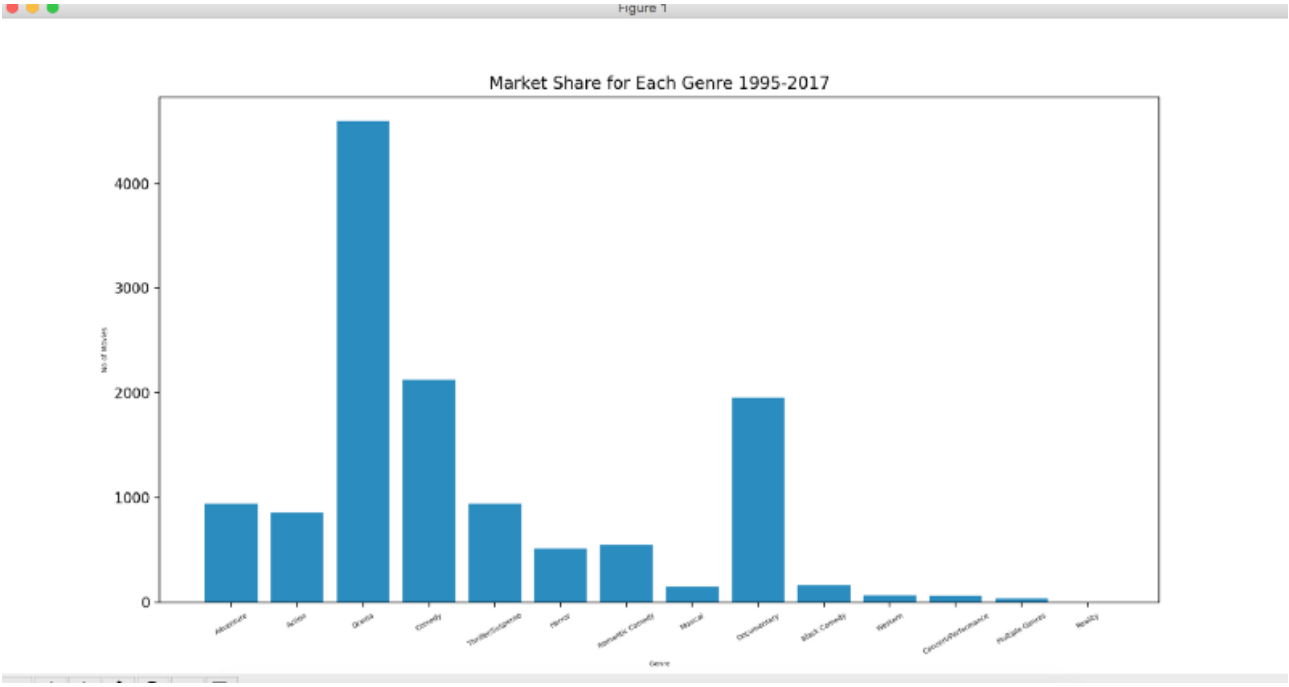
-Its a hierarchical clastering analysis used to analyse the order of dependencies of partially filled columns

The sample graph wold look like the below.



BAR (CREDITS:-900)

-Its a bar graph of all the cols vs the fraction of non empty data in that col.



FILL NULL VALUES

Here you have to select a column name and normalistion way.

Mean (CREDITS:-300)

The every element(x) will be replaced by **Mean**
Where Mean is the mean of the data of whole column.

Zero (CREDITS:-300)
The null element will be replaced by **Zero**

Standard Deviation (CREDITS:-300)
The every null element will be replaced by **Standard Deviation.**
Where Deviation is the standard deviation of the data.

LINEAR REGRESSION (CREDITS:-5 per row and after **3** linear regression **3000 per call)**
- Here your model will be trained and then everytime you have to pick the no of data you want (one row is one data).

DROP COLUMN (CREDITS:-700)
- It drops the column from the dataset. It helps to remove the useless data to get more accurate answer.

CHECKPOINT (CREDITS:-2000)
- Whenever you clicks on checkpoint it stores your current situation and deduct the credit.

REVERT (CREDITS:-500)
- It takes you back to your position of your checkpoint.

TEST ACCURACY (Only be used thrice in a event no credits)
- It gives the accuracy of your model on our test set (test data).