Wireframe Documentation

Analysing Swiggy: Bangalore delivery outlet data

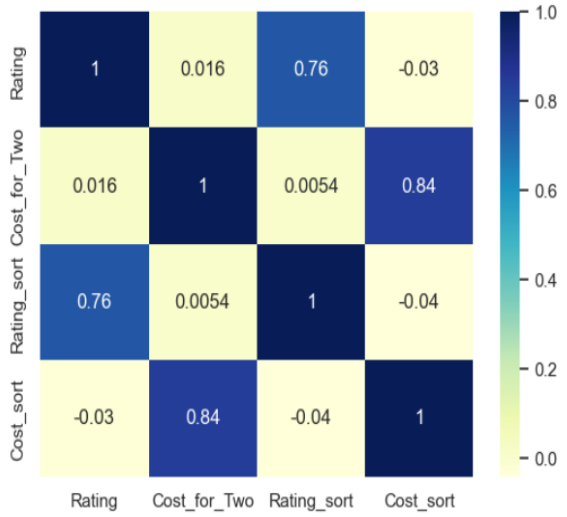
By
Mamta M. Natu

2. Wireframe Documentation

Wireframe for Exploratory data analysis

1) Heat map: To find correlation between different attributes in data.

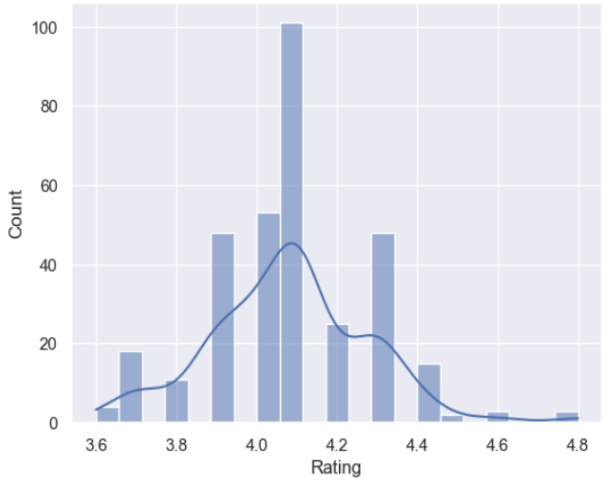
```
In [25]: #Using heatmap from seaborn package we can check correlation between different features.
dataplot = sns.heatmap(swiggy_df.corr(), cmap="YlGnBu", annot=True)

# Display the Pharma Sector Heatmap
plt.show()
```



From above heat map we can easily find the correlation between different attributes. As 'cost_sort' and 'cost of two' are directly related it is showing positiv corelation between them obviouslu.

2) Histplot: Histplot shows the distribution of 'ratings' count.

```
In [26]: sns.histplot(data = swiggy_df['Rating'], kde = True)
Out[26]: <AxesSubplot:xlabel='Rating', ylabel='Count'>
```

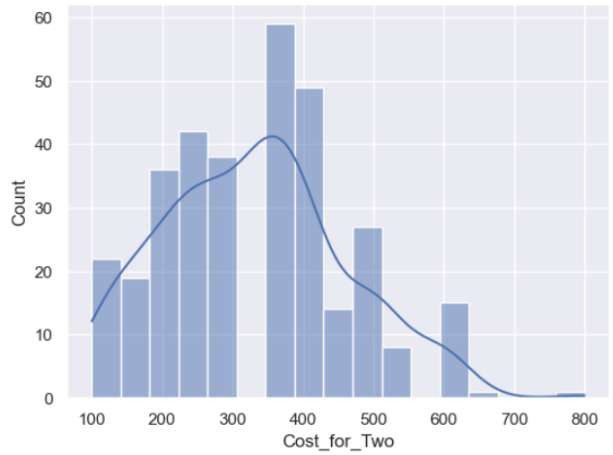


From above visualisation we can say that, Mean Rating is 4.1 approximately and count is 100 is maximum for it.

3) Histplot: By plotting histplot of ‘cost_for_two’ on x axis and count on y axis gives distribution of its count.

```
In [27]: sns.histplot(data = swiggy_df['Cost_for_Two'], kde = True)
```

```
Out[27]: <AxesSubplot:xlabel='Cost_for_Two', ylabel='Count'>
```

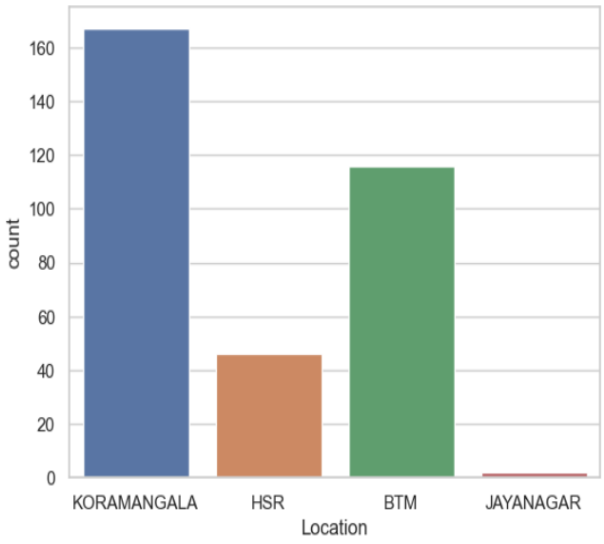


Mean cost of two persons is about 350 rs.

4) Count plot: Location on x axis and count on y axis gives bins showing count of location.

```
In [28]: ##Viewing survived data using countplot
         sns.set_style('whitegrid')
         sns.countplot(x='Location',data=swiggy_df)
```

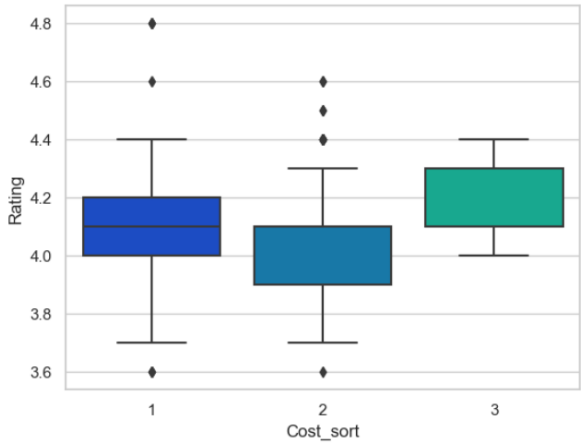
```
Out[28]: <AxesSubplot:xlabel='Location', ylabel='count'>
```



5) Boxplot: Boxplot gives outliers in cost_for_two.

Divide the "cost for two" in different categories
Cost_for_two>=200 ➔ Affordable=1,
200<Cost_for_two<500 ➔ Reasonable=2,
Cost_for_two>=500➔ Premium=3

```
In [29]: sns.boxplot(x='Cost_sort',y='Rating',data=swiggy_df,palette='winter')
Out[29]: <AxesSubplot:xlabel='Cost_sort', ylabel='Rating'>
```

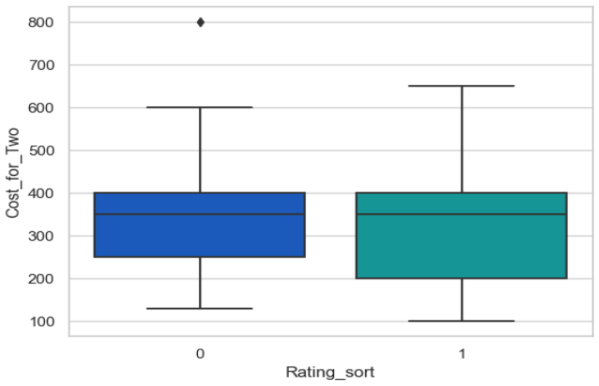


In above box plot shows For all cost sort type wether it is affordable, reasonable or premium maximul rating lies between 3.9 to 4.4 It show cost and rating are not that much correlated.

6) Boxplot: Boxplot gives outliers in ratings. Also gives distribution of ratings.

We divide Ratings in two halves, below and above average.
Ratings<4➔below average ratings.
Ratings>=4➔ above average ratings.

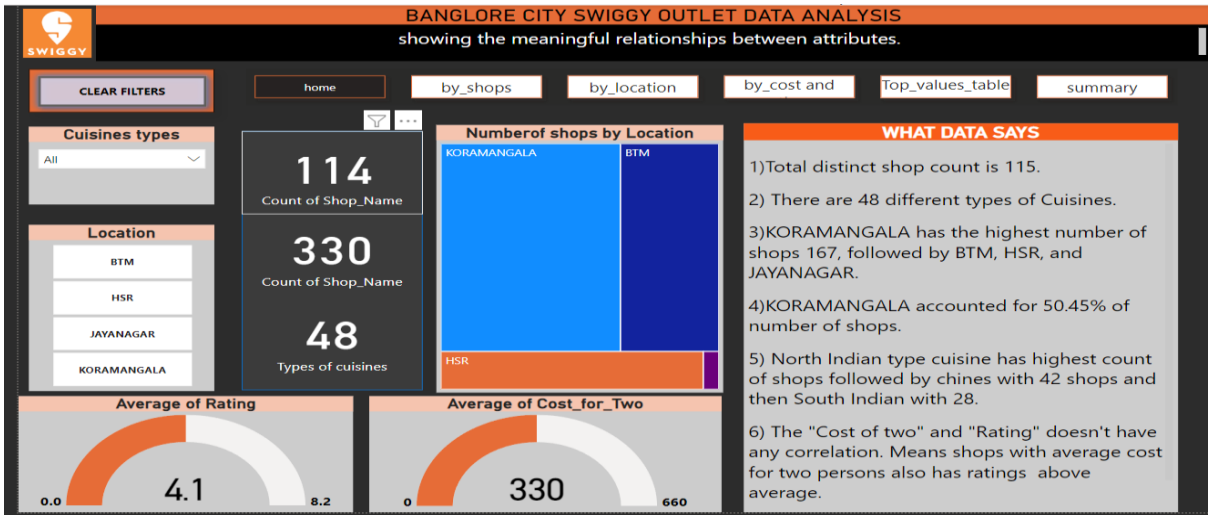
```
In [30]: sns.boxplot(x='Rating_sort',y='Cost_for_Two',data=swiggy_df,palette='winter')
Out[30]: <AxesSubplot:xlabel='Rating_sort', ylabel='Cost_for_Two'>
```



Wireframe Documentation of PowerBi Report

6) Homepage: It gives overview on numeric analysis of swiggy data

- It has slicers on Cuisine types and Location.
- Cards with different quantatitive information about data.
- Treemap showing distribution of shops over diferent Locations.
- Guages showing average ratings and average cost_for_two.
- Text box giving basic information about data.



7) Analysis of data by distinct_shop_Count:

Decomposition tree, stacked column chart, cluster bar chart, line and cluster column chart, filled map used to show distribution of distinct shop count.

6. Wireframe Documentation

8) Analysis of data by Location:

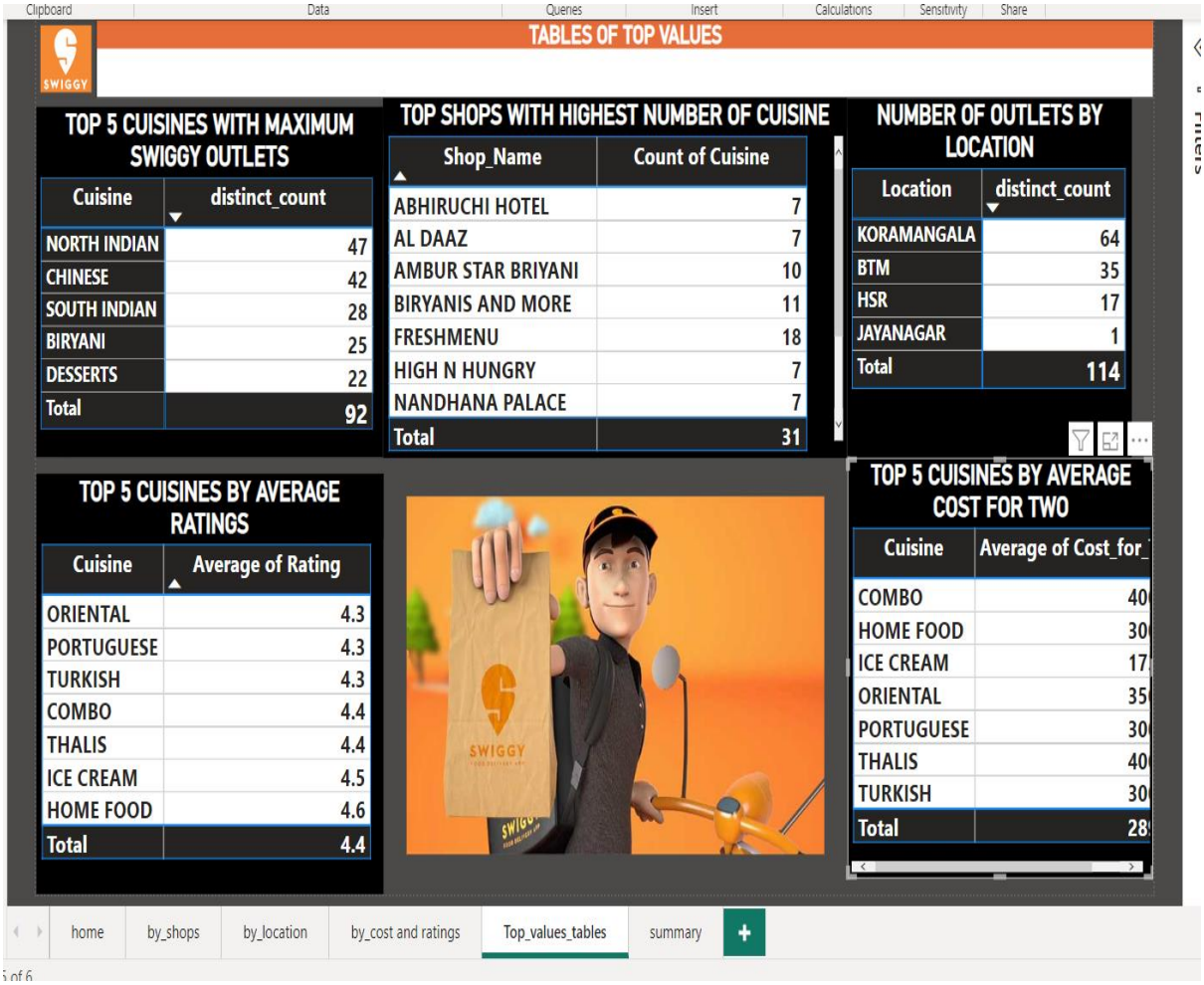
Area chart, Ribbon chart, cluster bar chart, cluster line chart, cluster column chart, line and stacked column chart used to show distribution of different attributes by location.

9) Analysis of data by cost for two persons and ratings:

Area chart, Line and stacked column chart, line , line and stacked column chart used to show distribution of cost for two and ratings.

10) Tables for top values of different attributes:

- Top 5 cuisines with Maximum outlets
- Top shops with highest number if cuisines
- Top 5 cuisines by average cost for two
- Top 5 cuisines by average ratings
- Number of outlets by location.



8. Wireframe Documentation

6)Summary page:

Heat map used to show correlation between differnet attributs in data.
Key influencer is used to find the effect of ratings on cost for two persons.
Text box showing question answer about relationship beween different attributes in data.