NuVista
Federal Credit Union

- VIJAYA RANA
- UB:50419571
- Email- vrana3@buffalo.edu
- Linkedin-linkedin.com/in/ranavijaya
- Github-github.com/ranavijaya

Topics Covered

- Dataset Modification
- Methodology used
- Top 10 predictors
- Visualizing Important features
- Key Brand Insights- 1
- Key Brand Insights - 2

Dataset Modification

- Few columns are dropped for the analysis shown below

```
#dropping record+status columns
df.drop(df.iloc[:,0:5], inplace=True, axis=1)
```

```
#dropping (use MRK_Age)
df.drop(df.iloc[:,5:11], inplace=True, axis=1)
```

```
#dropping S4 columns
df.drop(df.iloc[:,14:65], inplace=True, axis=1)
```

```
#drooping state divison
df.drop(df.iloc[:,18:27], inplace=True, axis=1)
```

```
#dropping Q5
df.drop(df.iloc[:,68:70], inplace=True, axis=1)
```

Methodology used

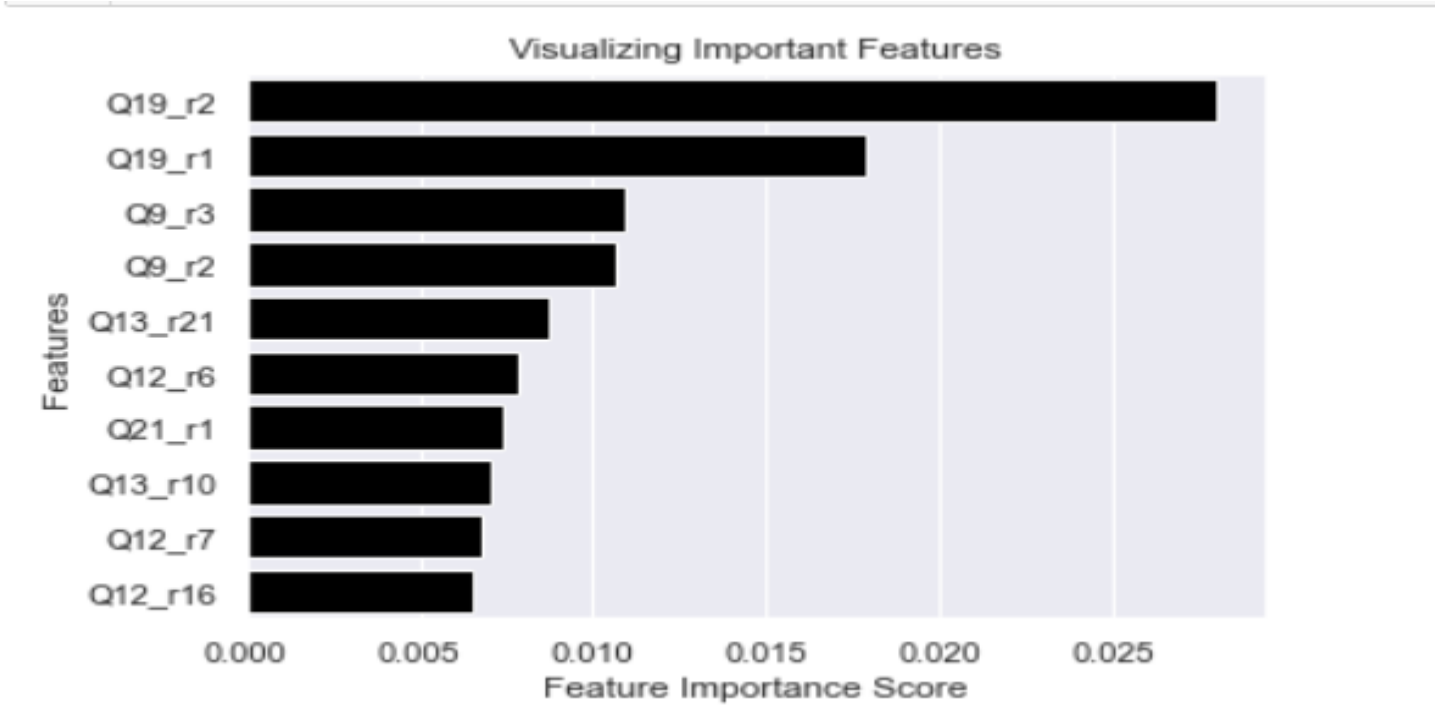
- Used RandomForestClassifier for finding the top features model
- Found feature importance for Top 10 predictors
- Used matplotlib and sns for visualization of the predictors

Q19_r2	0.027995
Q19_r1	0.017832
Q9_r3	0.010904
Q9_r2	0.010698
Q13_r21	0.008698
Q12_r6	0.007865
Q21_r1	0.007402
Q13_r10	0.007042
Q12_r7	0.006779
Q12_r16	0.006524

Top 10 predictors

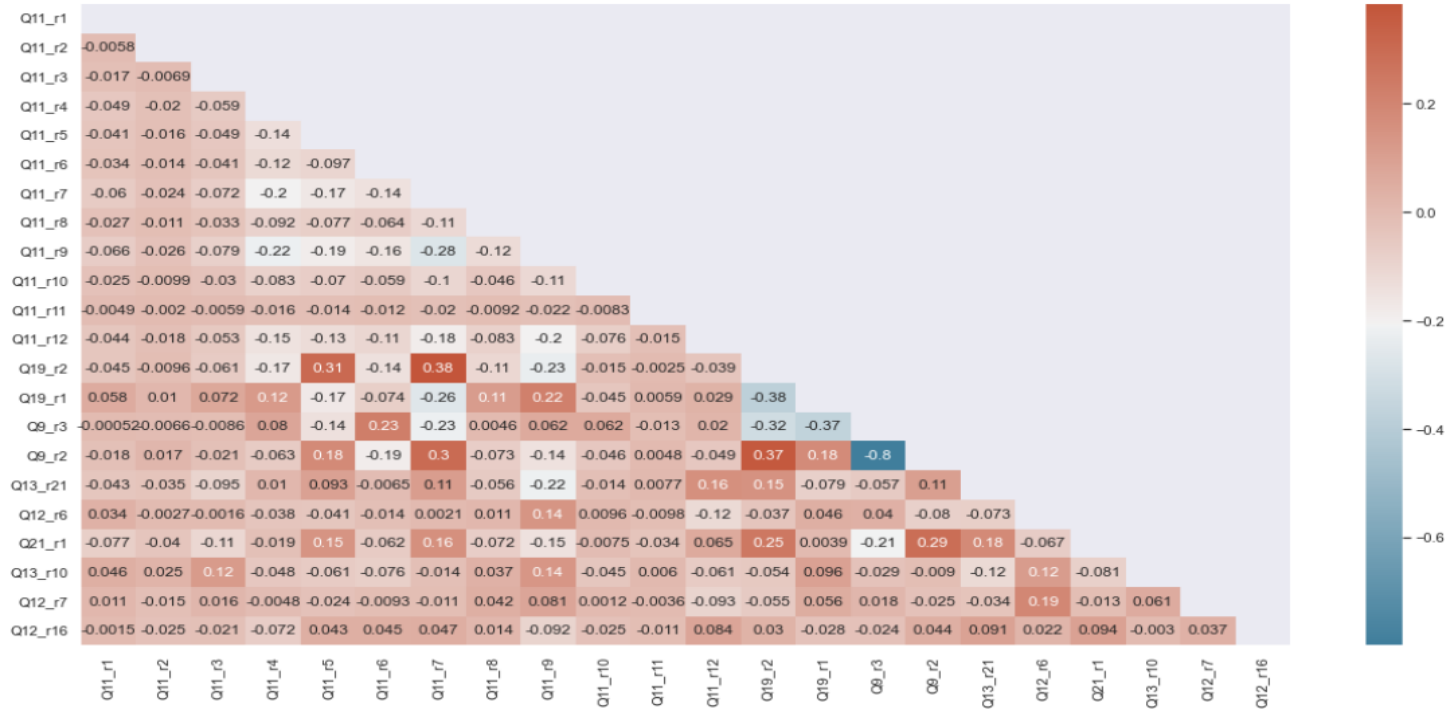
- 'Q19_r2'- Pack Type:Can
- 'Q19_r1'- Pack Type: Bag
- 'Q9_r3'- Category: Coffee Pod or Capsule
- 'Q9_r2'- Category: Ground Coffee
- 'Q13_r21'-Product Attributes: It's affordable
- 'Q12_r6'- Benefits: It's a premium brand
- 'Q21_r1'-Price Tier: Under $10
- 'Q13_r10'-Product Attributes: It's strong
- 'Q12_r7'- Benefits: It's a high-quality brand
- 'Q12_r16'-Benefits : It has good aroma

Visualizing Important Features



Key Brand Insights

- Pack Type : Can & Brand : Maxwell House have positive 0.31 correlation
- Pack Type : Can & Brand : Folgers have positive 0.38 correlation
- Category : Ground Coffee & Pack Type : Can have positive 0.37 correlation
- Category : Ground Coffee & Category: Coffee Pod or Capsule has highest negative -0.8 correlation



Summary

- Classification report was generated consisting of performance evaluation metrics.

```
#checking the classification report
from sklearn.metrics import classification_report
print(classification_report(y_test, y_predict))

```

	precision	recall	f1-score	support
0	0.00	0.00	0.00	9
1	0.00	0.00	0.00	2
2	0.00	0.00	0.00	13
3	1.00	0.01	0.02	107
4	0.00	0.00	0.00	80
5	0.00	0.00	0.00	60
6	0.52	0.16	0.24	146
7	0.00	0.00	0.00	34
8	0.71	0.05	0.10	183
9	0.00	0.00	0.00	29
10	0.00	0.00	0.00	0
11	0.00	0.00	0.00	96
12	0.00	0.00	0.00	0
micro avg	0.57	0.04	0.08	759
macro avg	0.17	0.02	0.03	759
weighted avg	0.41	0.04	0.07	759
samples avg	0.04	0.04	0.04	759

THANK YOU