

Leonardo Ignatius

UKRAINIAN COFFEE SHOPS

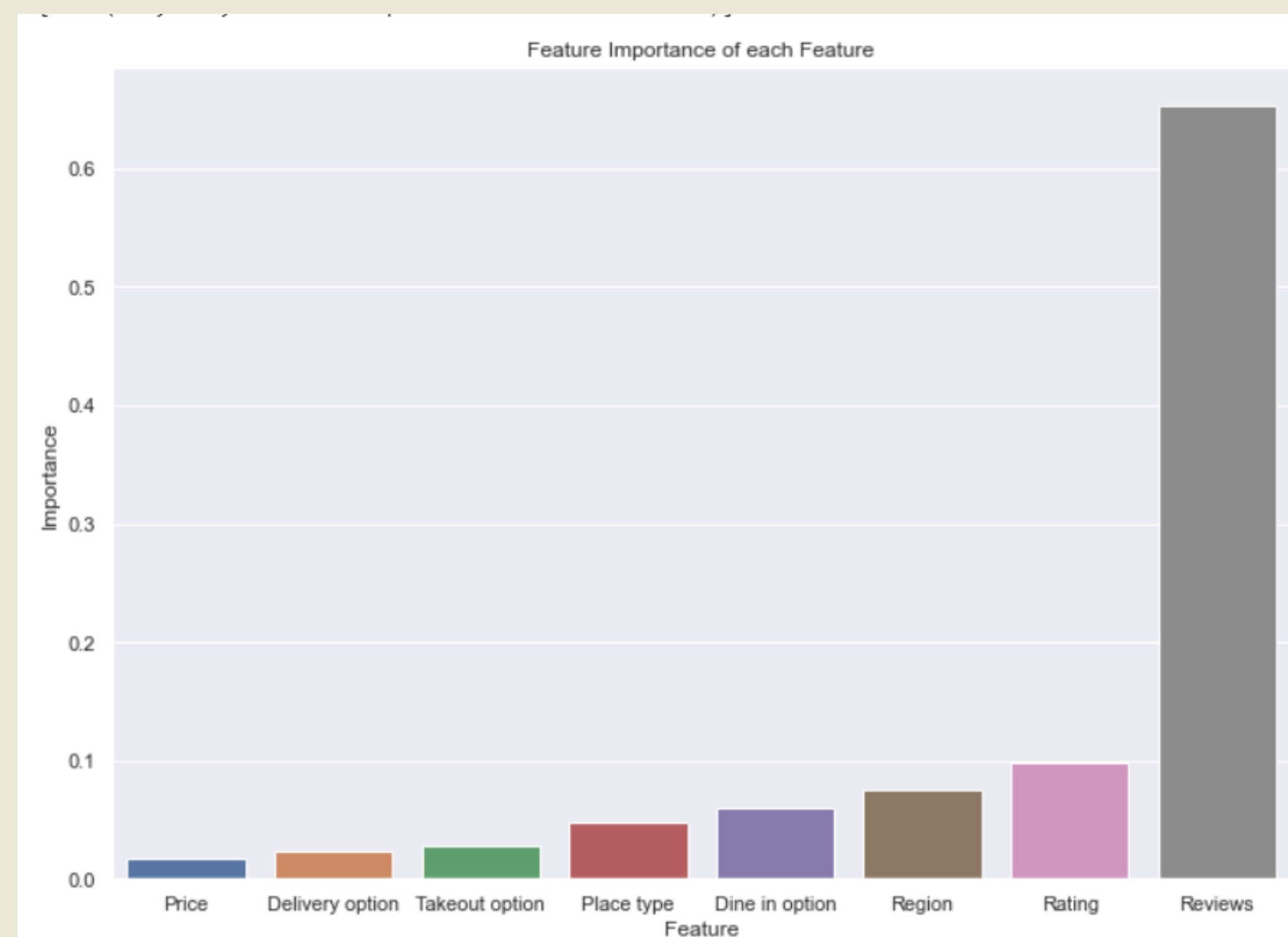
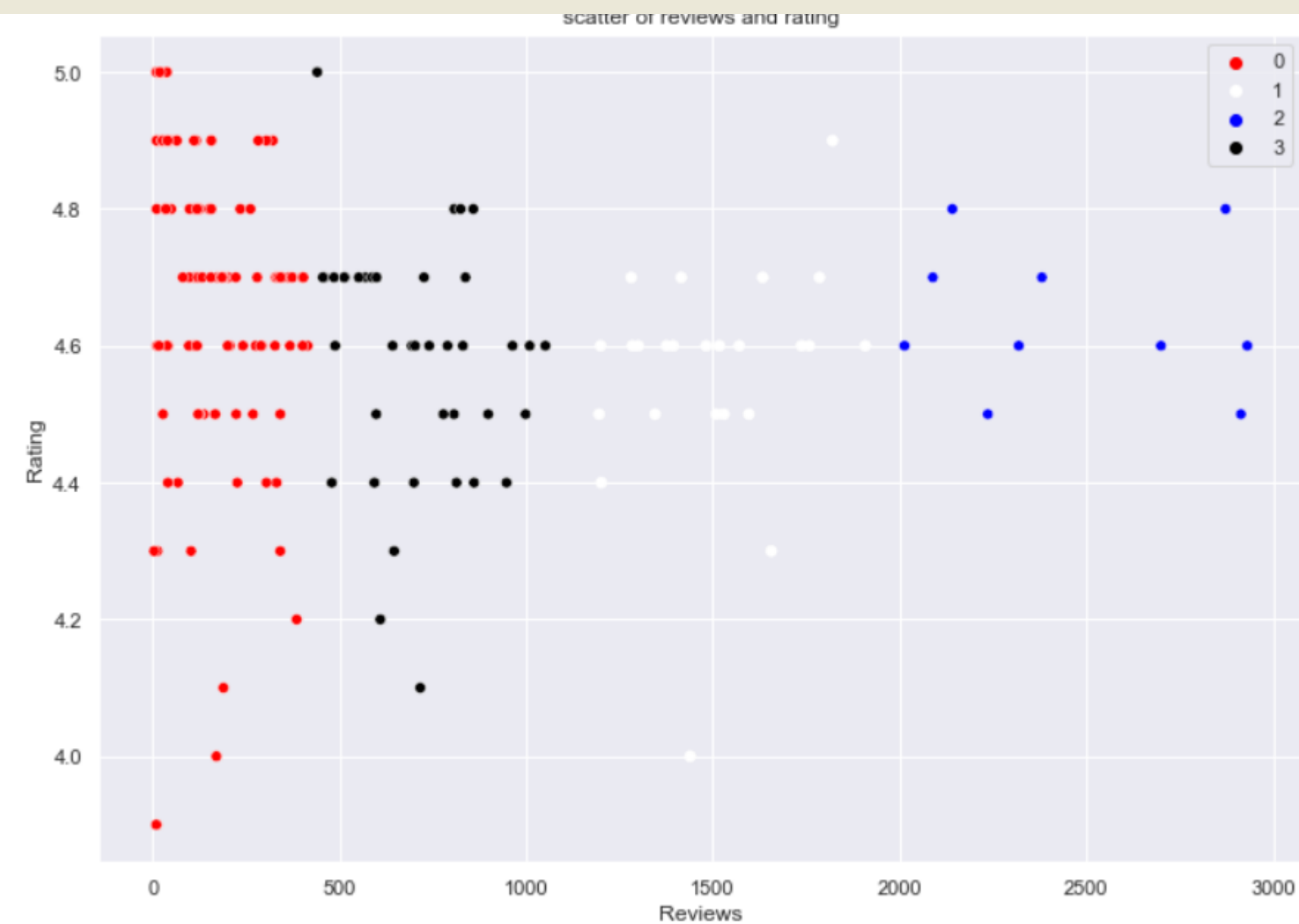
GOALS

- Find out how many segment of coffee shop in the new market in Ukraine
- Find out which key features effect the segmentation of the coffee shops

The Data Used

- Region
- Place Name
- Place Type
- Ratings
- Number of Reviews
- Takeout Options
- Price Category
- Delivery Options
- Dine in Options

Result

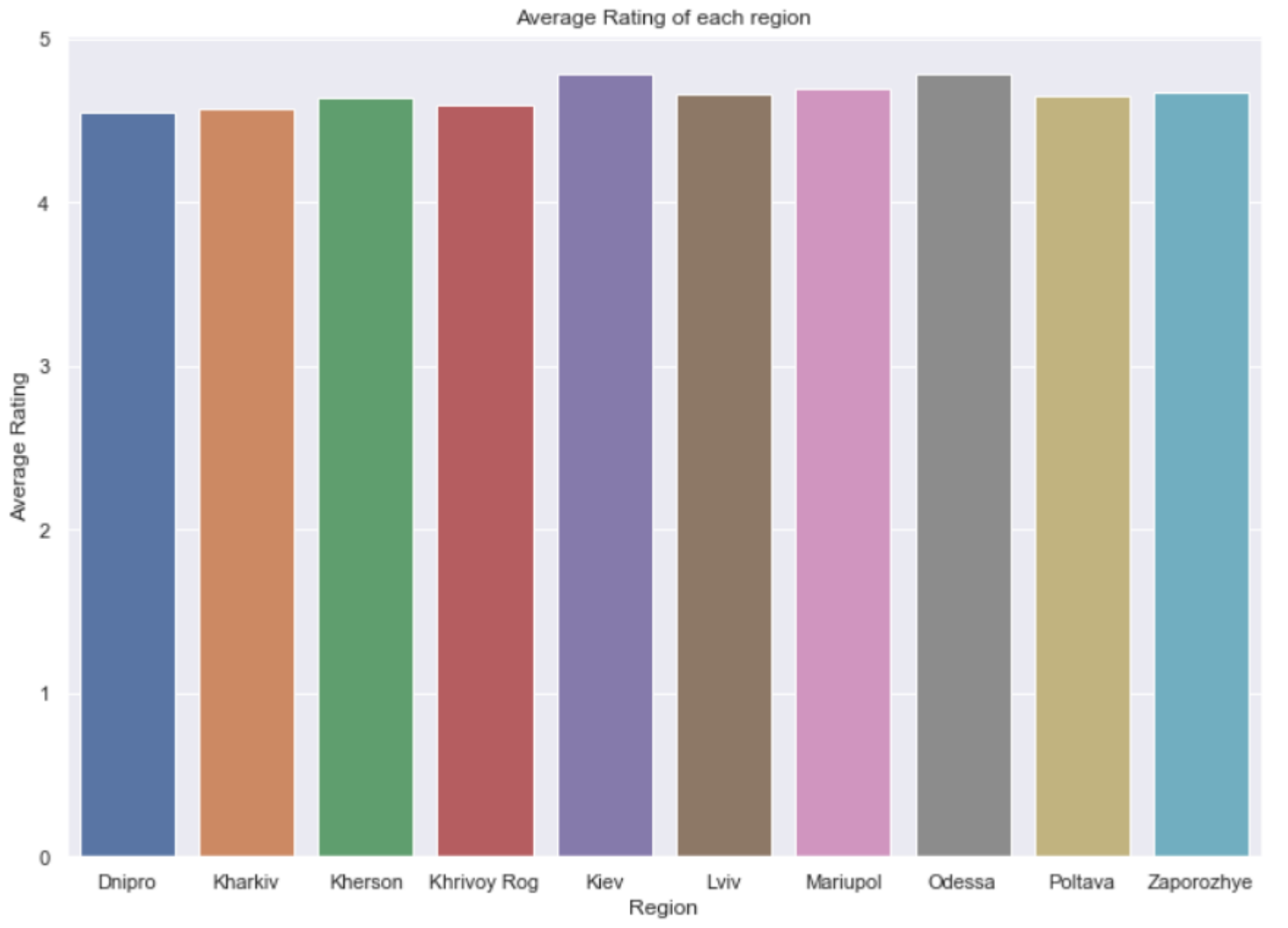


By Using Machine Learning Technique based on the data, we found out that the segmentation is diverse into four segment

These are the importance of each features regarding the segmentation

Data Analysis

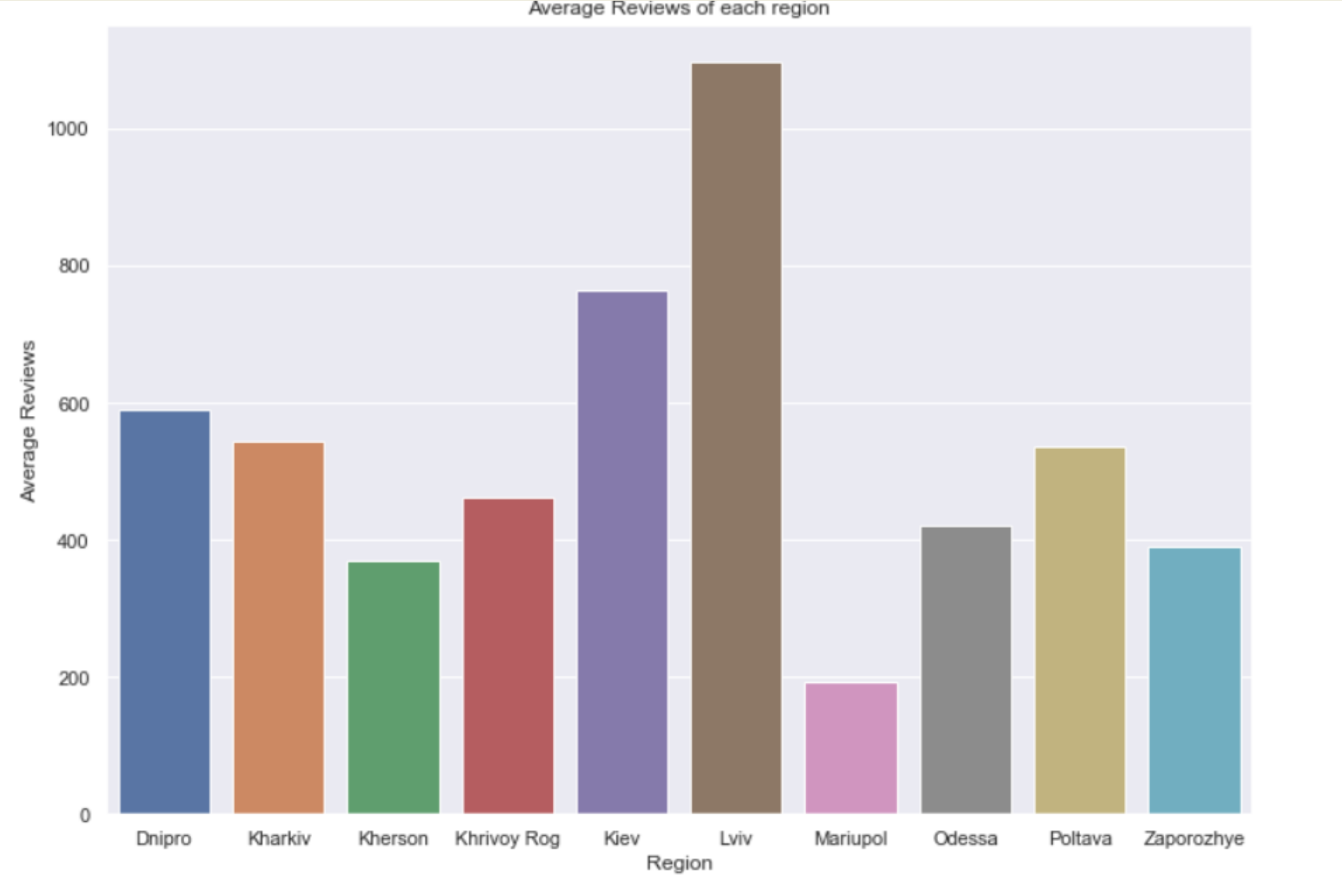
Average Rating Each Region



Region	
Kiev	4.785000
Odessa	4.785000
Mariupol	4.700000
Zaporozhye	4.670000
Lviv	4.668421
Poltava	4.655000
Kherson	4.642105
Khrivoy Rog	4.600000
Kharkiv	4.575000
Dnipro	4.547368
Name: Rating, dtype: float64	

Kiev and Odessa Have the highest rating average, and Kharkiv and Dnipro have the lowest rating average

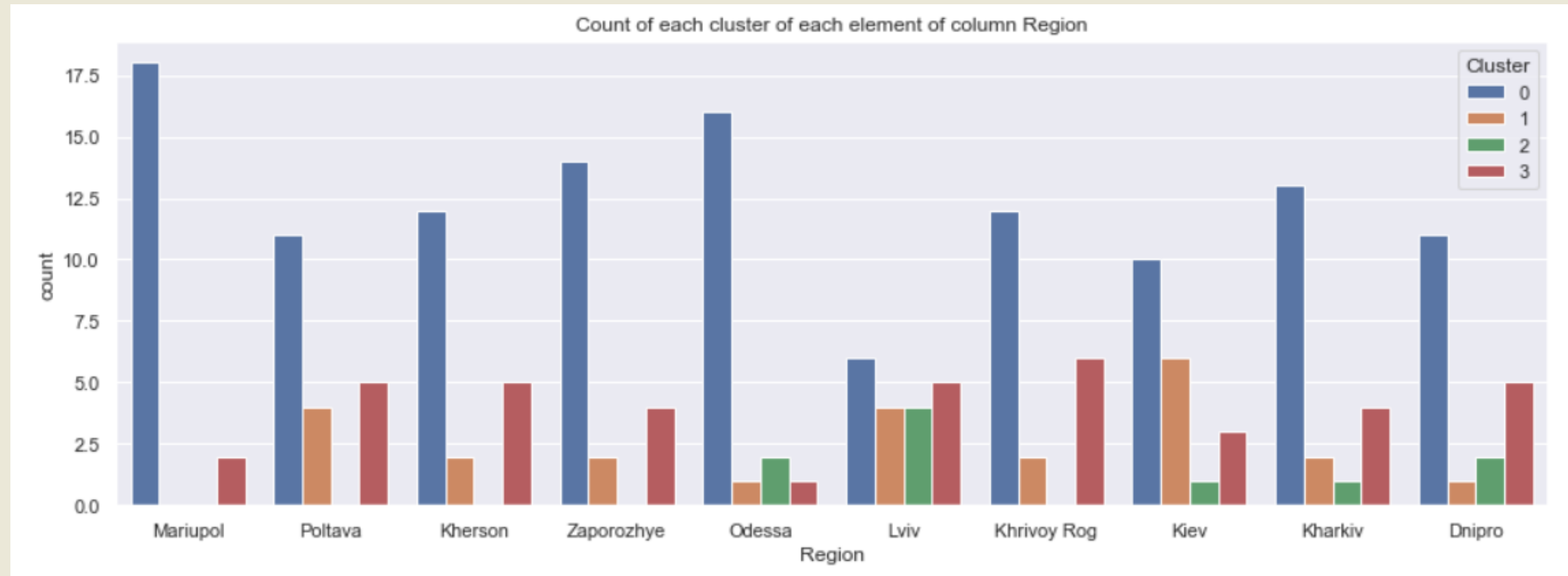
Average Reviews Each Region



Region	
Lviv	1096.631579
Kiev	764.600000
Dnipro	589.947368
Kharkiv	544.450000
Poltava	535.250000
Khrivoy Rog	461.900000
Odessa	422.000000
Zaporozhye	389.950000
Kherson	370.157895
Mariupol	193.800000

Lviv Have the highest review average,
Mariupol have the lowest review average

Number Clusters from Each Region



Lviv has the equally distributed Segment
Mariupol segment is imbalanced

LVIV

Cluster 0

	Region	Place type	Rating	Reviews	Price	Delivery option	Dine in option	Takeout option	Cluster
107	Lviv	Coffee shop	4.3	102.0	NaN	NaN	True	True	0
111	Lviv	Coffee shop	5.0	24.0	NaN	True	True	True	0
112	Lviv	Coffee shop	4.7	363.0	ss	False	True	True	0
116	Lviv	Cafe	5.0	10.0	NaN	NaN	False	True	0
118	Lviv	Cafe	4.9	157.0	ss	False	True	True	0
119	Lviv	Coffee shop	4.9	321.0	ss	False	True	True	0

Cluster 1

	Region	Place type	Rating	Reviews	Price	Delivery option	Dine in option	Takeout option	Cluster
104	Lviv	Coffee shop	4.6	1517.0	ss	False	True	True	1
113	Lviv	Coffee roasters	4.6	1199.0	NaN	False	True	True	1
114	Lviv	Coffee shop	4.6	1284.0	ss	False	True	True	1
115	Lviv	Cafe	4.6	1570.0	ss	True	True	True	1

Cluster 2

	Region	Place type	Rating	Reviews	Price	Delivery option	Dine in option	Takeout option	Cluster
101	Lviv	Coffee shop	4.6	2931.0	ss	NaN	True	True	2
102	Lviv	Coffee shop	4.7	2089.0	ss	False	True	True	2
110	Lviv	Coffee shop	4.8	2873.0	ss	False	True	True	2
117	Lviv	Cafe	4.5	2236.0	ss	False	True	True	2

Cluster 3

	Region	Place type	Rating	Reviews	Price	Delivery option	Dine in option	Takeout option	Cluster
103	Lviv	Coffee shop	4.4	813.0	ss	False	True	True	3
105	Lviv	Coffee shop	4.8	858.0	ss	NaN	True	True	3
106	Lviv	Coffee shop	4.6	702.0	ss	False	True	True	3
108	Lviv	Coffee shop	4.5	778.0	ss	False	True	True	3
109	Lviv	Coffee shop	4.6	1009.0	ss	False	True	True	3

Average of Each Clusters

Cluster 0

Rating	4.800000
Reviews	162.833333
Delivery option	0.250000
Dine in option	0.833333
Takeout option	1.000000
Cluster	0.000000

Cluster 1

Rating	4.60
Reviews	1392.50
Delivery option	0.25
Dine in option	1.00
Takeout option	1.00
Cluster	1.00

Cluster 2

Rating	4.65
Reviews	2532.25
Delivery option	0.00
Dine in option	1.00
Takeout option	1.00
Cluster	2.00

Cluster 3

Rating	4.58
Reviews	832.00
Delivery option	0.00
Dine in option	1.00
Takeout option	1.00
Cluster	3.00

Correlation of clusters in LVIV

Reviews	0.435301
Rating	0.416455
Place type	0.309728
Delivery option	0.292698
Dine in option	0.282913
Price	0.083826

Doing Test on a random dataset with conditions

Data Declaration & Expectation

	Region	Place type	Rating	Reviews	Price	Delivery option	Dine in option	Takeout option
0	5	1	3	1500	1	0	1	0
1	2	2	5	100	2	1	0	0
2	4	3	1	500	2	1	1	1
3	5	3	5	10000	4	1	0	1

0 = Medium Ratings , High Reviews (Expected Cluster 1)

1 = High Ratings , low Reviews (Expected Cluster 0)

2 = Low Ratings , medium Reviews (Expected Cluster 3)

3 = High Ratings , High Reviews (Expected Cluster 2)

Doing Test on a random dataset with conditions

Test Result

	Region	Place type	Rating	Reviews	Price	Delivery option	Dine in option	Takeout option	cluster
0	5	1	3	1500	1	0	1	0	1
1	2	2	5	100	2	1	0	0	0
2	4	3	1	500	2	1	1	1	3
3	5	3	5	10000	4	1	0	1	2

0 = Cluster 1 (Expected Cluster 1)

1 = Cluster 0 (Expected Cluster 0)

2 = Cluster 3 (Expected Cluster 3)

3 = Cluster 2 (Expected Cluster 2)

CONCLUSION /OUTCOME




**In This New Market There
are 4 Segmentations of
Coffee Shops in Ukraine**




**The top 3 Key Features Are
Reviews, Rating, and Region**




**The Most Common Segment
is segment 0**



**In LVIV, the segment is well
distributed evenly , and in
mariupol the cluster is imbalance**



Reviews and Ratings is the most important Feature to determine segment, so to adjust the segment the coffee shop need to add more reviews and adjust the rating



To do that we can use promotions such as giving discounts, if the customer create a new review to our shops

FUTURE WORK

THANK YOU