

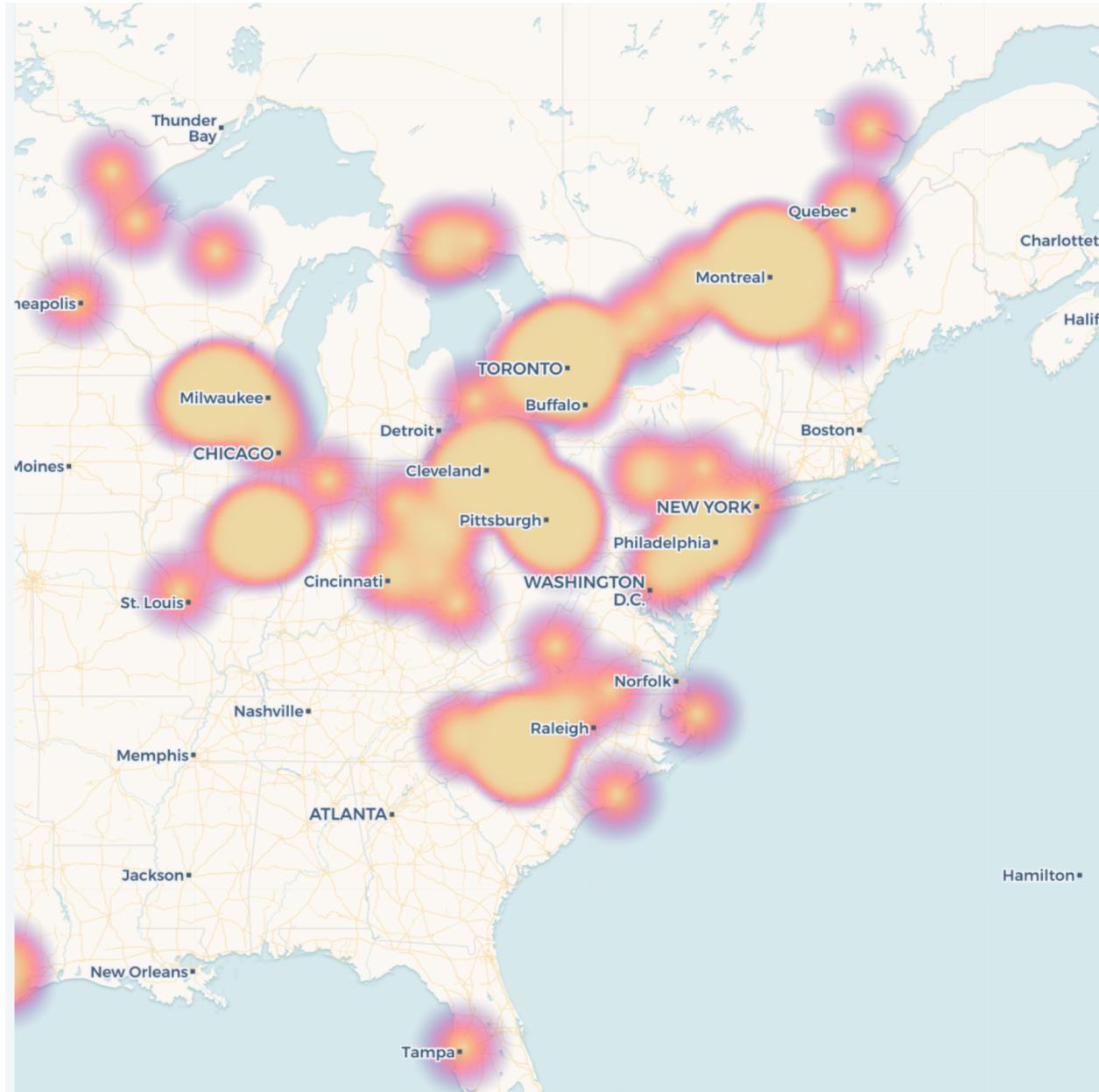
YELP US!

Taylor Gunter
Matthew Coker
Devin Arnold

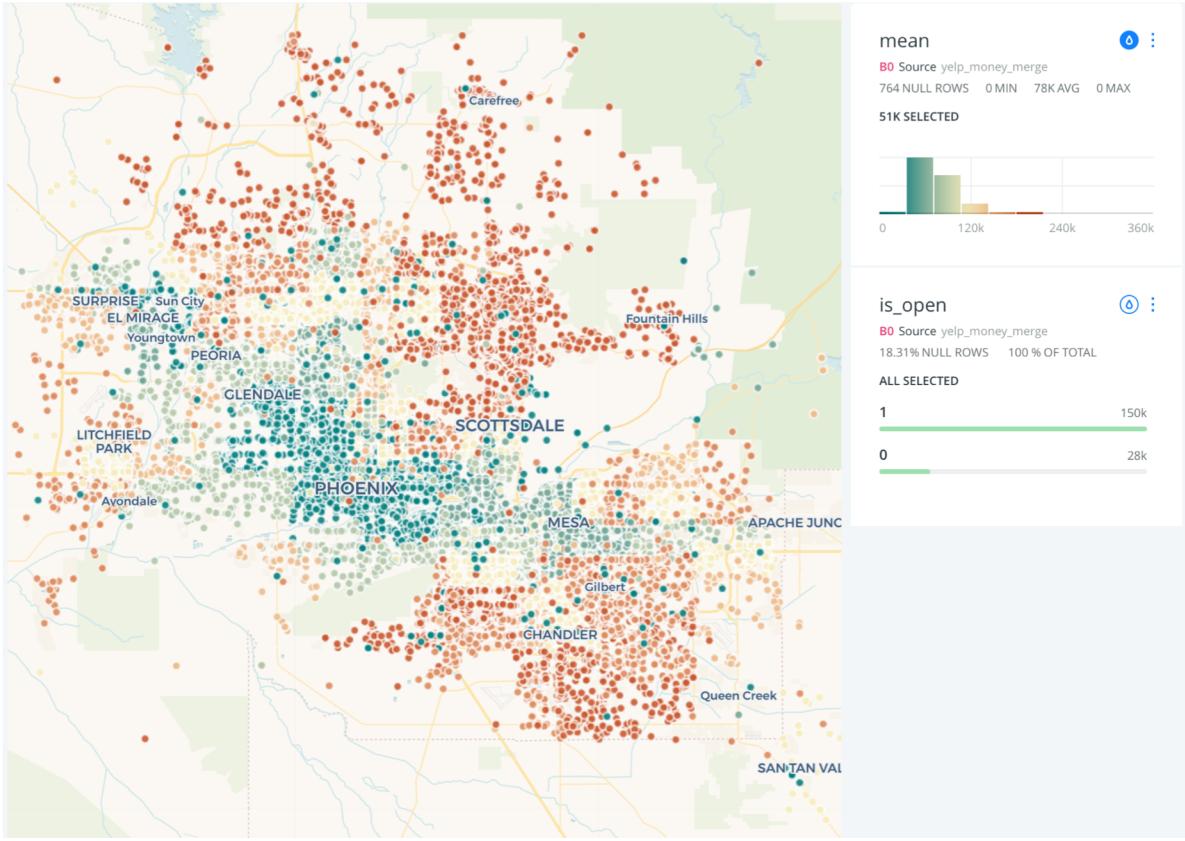


Questions to Answer

- Prove successful restaurants have more in common than just great food
- What do successful restaurants have in common?
- What failed successful restaurants have in common?
- How does mean and median household income of the area a restaurant is located in affect restaurant success?

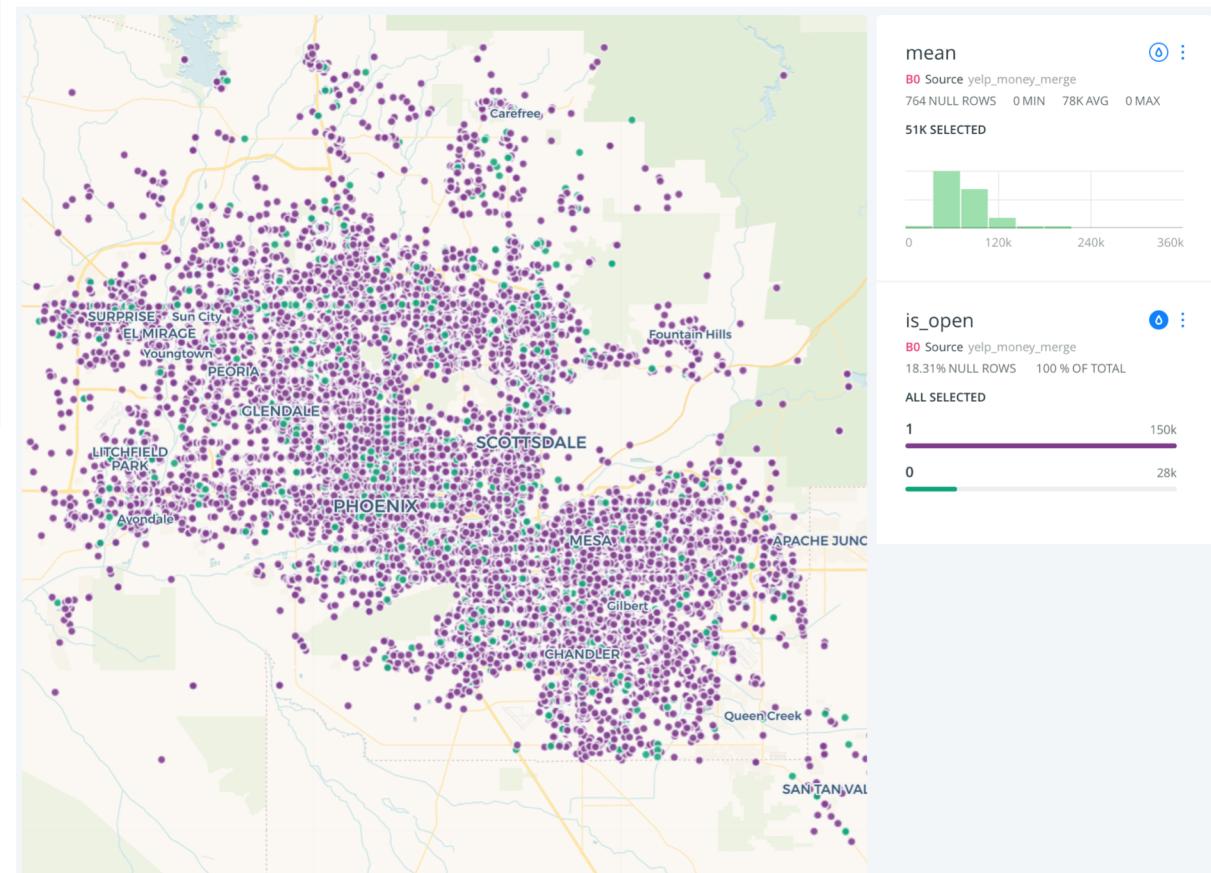


- ▶ Heatmap of the East Coast of the United States
- ▶ Multiple clusters in major cities
- ▶ Change focus



Mean Income

Open vs. Closed

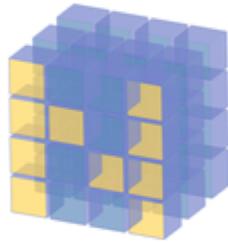


Data Preparation

- Done in Jupyter Notebook using Pandas
- Discarded attributes we deemed unimportant
- Cleaned empty attribute values
- Changed certain attribute values(such as alcohol, smoking, restaurant pricing, etc.) to integers from 0-4. Places them on a scale
- Simplified some attributes(ages allowed, BYOB, etc.) to have less unique values
- Discarding sparse tuples and non business tuples

Tools

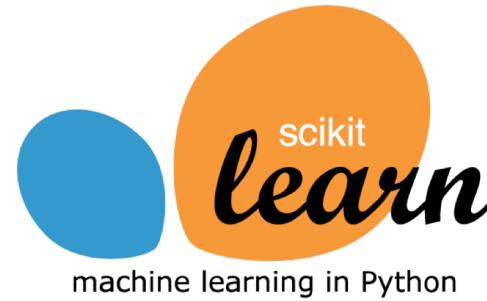
- Python
- Numpy
- Pandas
- Scikit-learn
- Jupyter Notebook
- Mlxtend
- Matplotlib
- Bokeh



NumPy



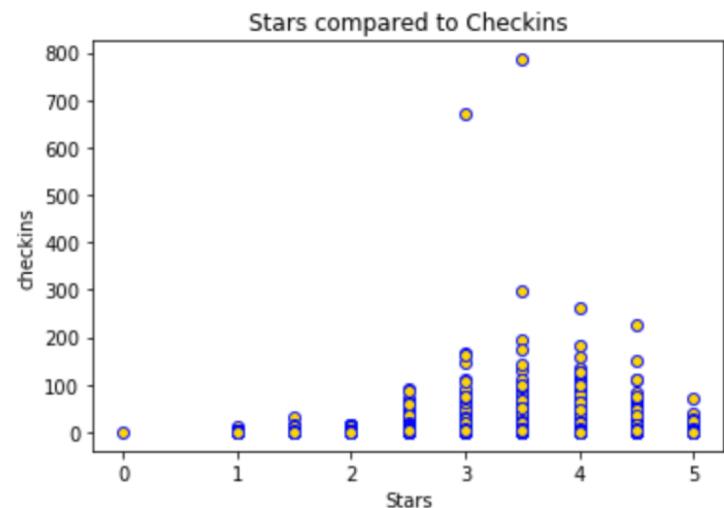
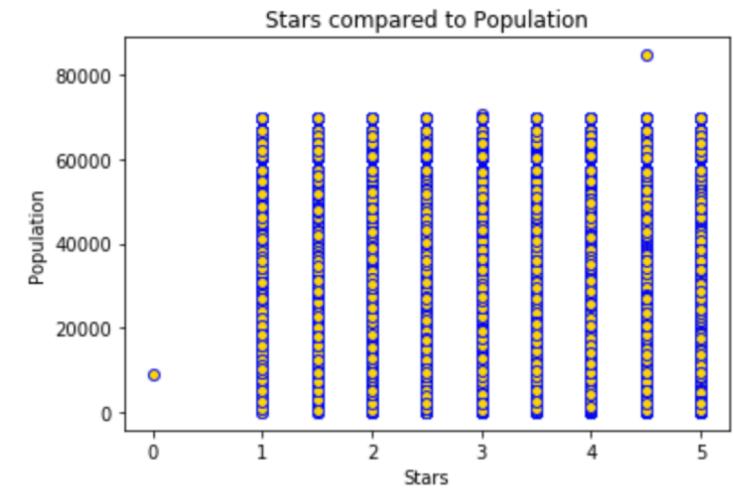
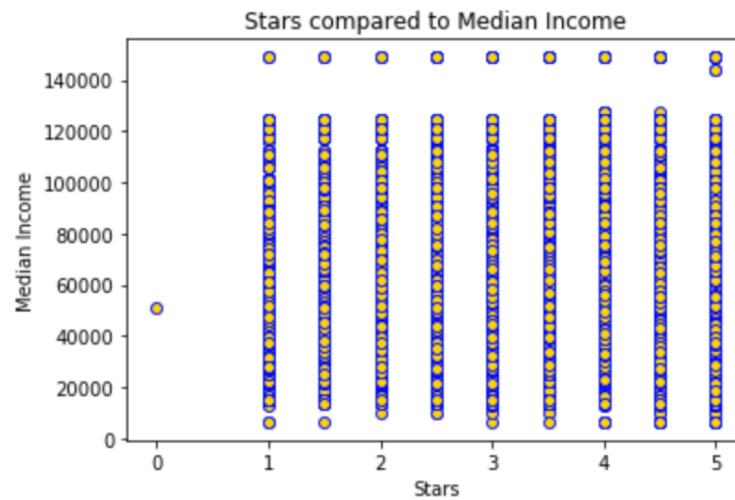
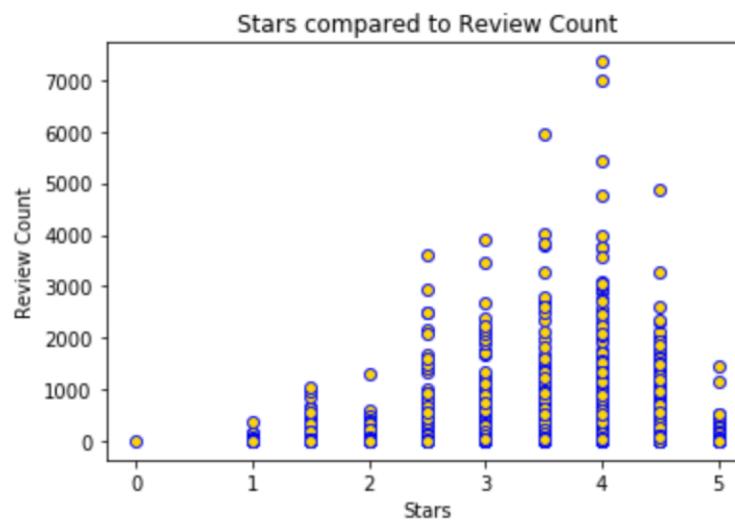
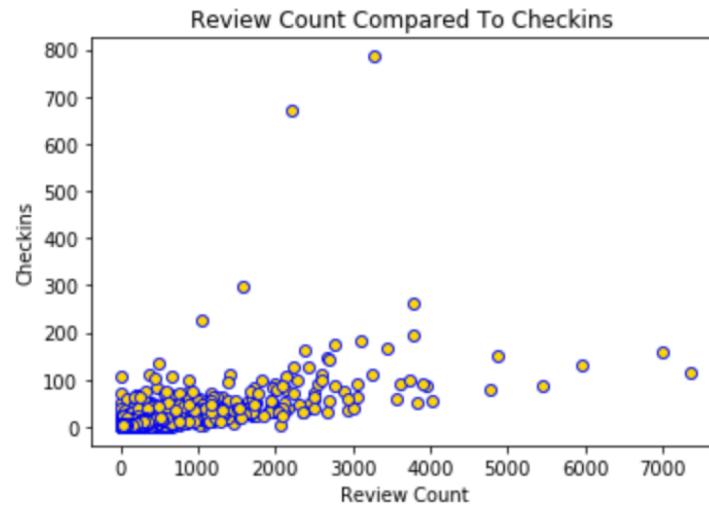
MLXTEND



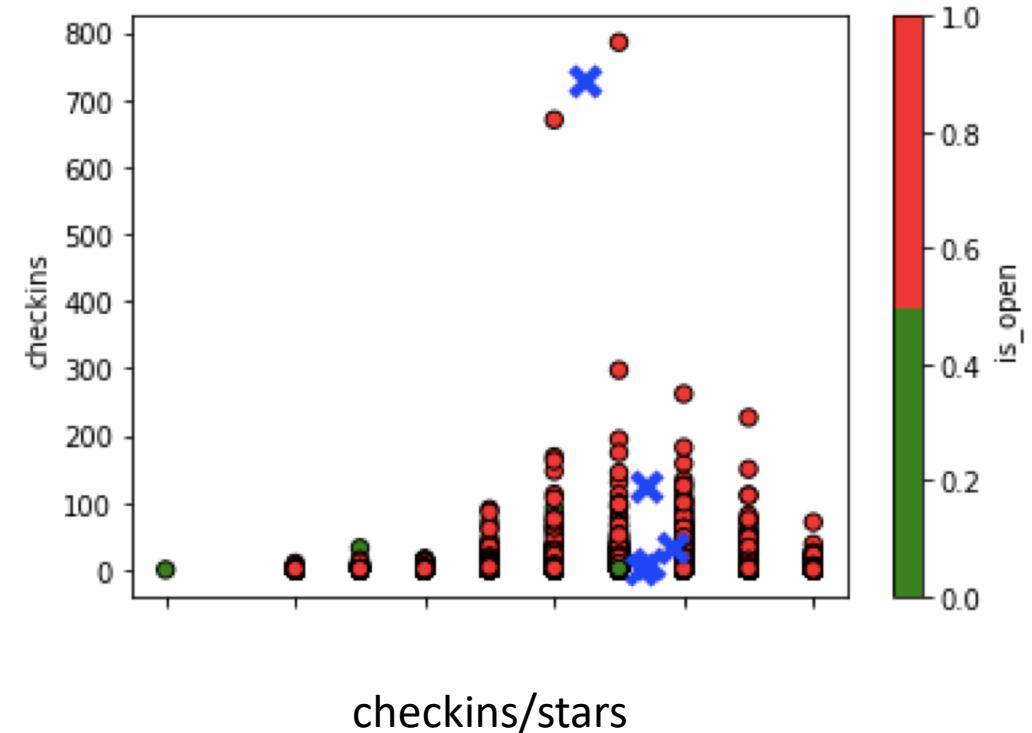
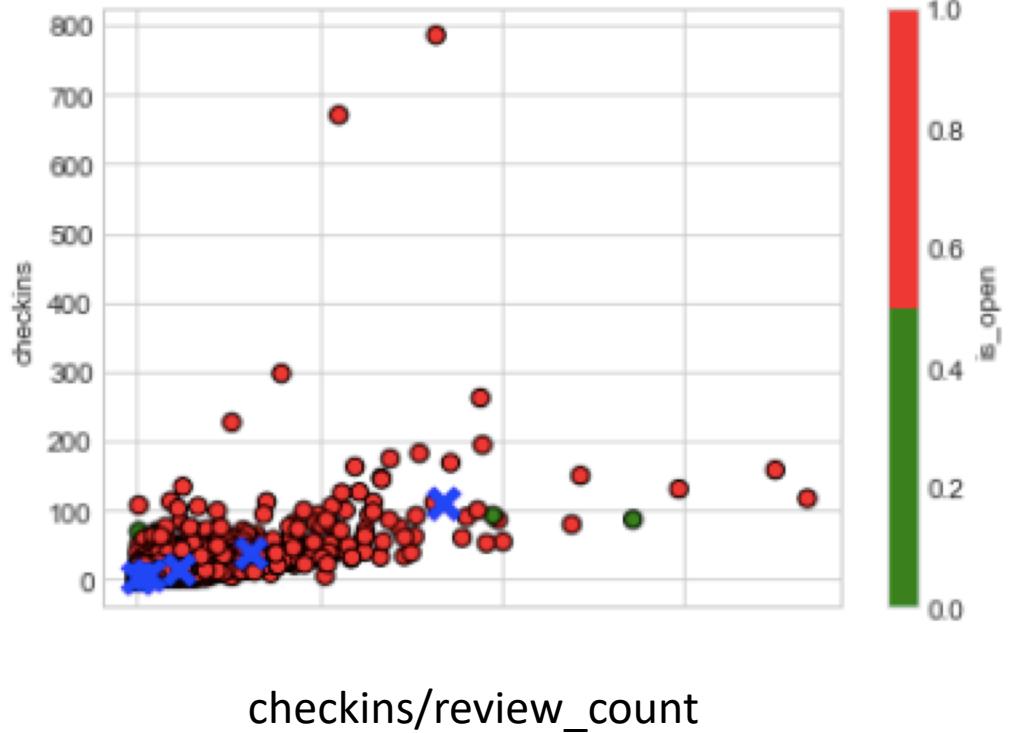
matplotlib

Pandas

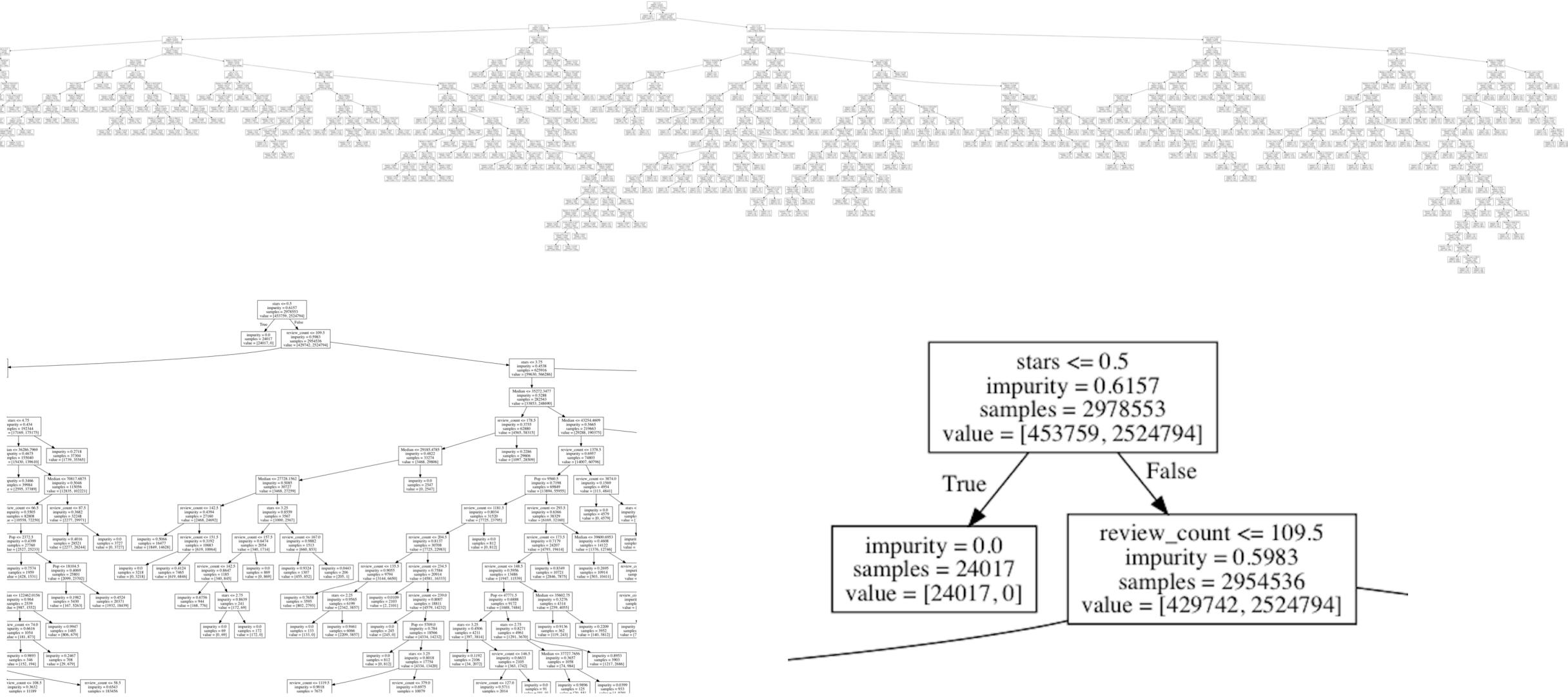




Clustering



Decision Tree



Confusion Matrix

	Predicted Failure	Predicted Success
True Failure	11331	140710
True Success	697	840114

Decision Tree

K-Nearest Neighbors

	Predicted Failure	Predicted Success
True Failure	6	82
True Success	30	655

Association Rules

	antecedants	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(GoodForMeal_dinner)	(CoatCheck)	0.007426	0.001638	0.001217	0.163862	100.055083	0.001205	1.194016
1	(CoatCheck)	(GoodForMeal_dinner)	0.001638	0.007426	0.001217	0.742972	100.055083	0.001205	3.861735
2	(GoodForMeal_brunch)	(CoatCheck)	0.001631	0.001638	0.001059	0.649194	396.401752	0.001056	2.845906
3	(CoatCheck)	(GoodForMeal_brunch)	0.001638	0.001631	0.001059	0.646586	396.401752	0.001056	2.824930
4	(BestNights_monday, BusinessParking_valet)	(GoodForMeal_dinner)	0.002092	0.007426	0.001605	0.767296	103.330726	0.001589	4.265387
5	(GoodForMeal_dinner)	(BestNights_monday, BusinessParking_valet)	0.007426	0.002092	0.001605	0.216120	103.330726	0.001589	1.273038
6	(GoodForMeal_breakfast, BusinessParking_valet)	(GoodForMeal_dinner)	0.004953	0.007426	0.003762	0.759628	102.298161	0.003725	4.129329
7	(GoodForMeal_dinner)	(GoodForMeal_breakfast, BusinessParking_valet)	0.007426	0.004953	0.003762	0.506643	102.298161	0.003725	2.016891
8	(BestNights_monday, GoodForMeal_breakfast)	(GoodForMeal_dinner)	0.001664	0.007426	0.001408	0.845850	113.909522	0.001395	6.439008
9	(GoodForMeal_dinner)	(BestNights_monday, GoodForMeal_breakfast)	0.007426	0.001664	0.001408	0.189548	113.909522	0.001395	1.231827
10	(BestNights_monday, GoodForMeal_breakfast, BusinessParking_valet)	(GoodForMeal_dinner)	0.001256	0.007426	0.001105	0.879581	118.452080	0.001096	8.242683

Knowledge Gained

- Datasets very sparsely populated
- Support and confidence levels are low for any combination of attributes
- Checkins and stars are the largest factors to determine whether a business succeeded or failed
- After, population size and median income make a difference
- Population over 33,000 and income over 29,000 is a more successful application
- Checkins are important
- High star ratings are most important

Knowledge Application

- What factors matter most for a business to be successful
- Locations where businesses can do best based on population and income data
- Finding correlations between successful businesses