

# Final Project



Lisa McDermott, Eric Thiry, Clayton Remsik, John VanVliet

# Selected Topic

- Analyzing the distribution of craft breweries across the United States and determining optimal locations for expansion taking employment, population density and income data into account

# Reasons for Selection

- The data is readily available
  - Census Data
  - Unemployment and Income Data
  - Brewery Data
- The analysis is relevant
  - Breweries are growing rapidly in popularity
- Craft beer is a shared interest among the team

# Data Source

- Public information gathered from census and brewery data
  - US City, County and State Data
  - Unemployment and Income Data

# Questions We Hope to Answer

- What are optimal locations for new breweries?
- How do population densities, median incomes, breweries per 50,000 people, and employment statistics affect brewery locations?

# Data Exploration

- Four CSV files were uploaded to Jupyter Notebook and cleaned. Exported to PostgreSQL via string method
- Used Machine Learning to find clusters of the population with high median incomes and low unemployment to find the best possible opportunities of new brewery locations
- Created US Map and table visualizations using Tableau

# Analysis

- The machine learning model identified four clusters, each with five counties that are optimal locations to add new breweries

Summary Tab						
Cluster	Total FIPS	Avg. Median Household Income	Avg. Population 2010	Brewery Count	Avg. Brewery Count	Avg. Breweries per 50k
0	805	56,099	111,850	2,558	3.18	1.55
1	617	54,309	76,373	1,443	2.34	1.39
2	532	55,219	140,530	1,846	3.47	1.85
3	1,089	55,829	84,066	2,628	2.41	1.50

- A detailed summary of each cluster is shown in the preceding slides
- Visualizations are provided in Tableau and can be found using the link below
  - <https://public.tableau.com/app/profile/clayton.remsik/viz/BootcampFinalGroupproject-USCraftBreweries/Cluster3Selections>

# Analysis Continued - Cluster 0

## Cluster 0 Selections

County Name	State	Population	Avg Median Household Income	Breweries per 50k	Brewery Count
Arlington County	VA	207,627	\$118,986.00	0.241	1
Bergen County	NJ	905,116	\$107,971.00	0.221	4
Charles County	MD	146,551	\$102,510.00	0.341	1
Fairfax County	VA	1,081,726	\$127,898.00	0.231	5
Wilson County	TN	113,993	\$80,080.00	0.000	0



# Analysis Continued - Cluster 1

## Cluster 1 Selections

County Name	State	Population	Avg Median Household Income	Breweries per 50k	Brewery Count
Anne Arundel Co..	MD	537,656	\$100,916.00	0.465	5
Fort Bend County	TX	585,375	\$101,361.00	0.427	5
Oldham County	KY	60,316	\$101,152.00	0.000	0
Richmond County	NY	468,730	\$86,624.00	0.213	2
Stafford County	VA	128,961	\$109,090.00	0.388	1

## Analysis Continued - Cluster 2

### Cluster 2 Selections

County Name	State	Population	Avg Median Household Income	Breweries per 50k	Brewery Count
Morris County	NJ	492,276	\$116,328.00	0.609	6
Norfolk County	MA	670,850	\$106,851.00	0.745	10
Putnam County	NY	99,710	\$105,600.00	0.000	0
Santa Clara Coun..	CA	1,781,642	\$132,444.00	0.589	21
Westchester Co..	NY	949,113	\$101,741.00	0.579	11

## Analysis Continued - Cluster 3

### Cluster 3 Selections

County Name	State	Population	Avg Median Household Income	Breweries per 50k	Brewery Count
Contra Costa County	CA	1,049,025	\$106,555.00	0.667	14
Kendall County	IL	114,736	\$105,586.00	0.436	1
Montgomery County	MD	971,777	\$110,012.00	0.566	11
San Mateo County	CA	718,451	\$135,234.00	0.766	11
Williamson County	TN	183,182	\$119,637.00	0.819	3