

PROJECT TITLE: BEER PREFERENCES ACROSS AMERICA

TEAM MEMBERS:

GUANBO BIAN, JESSICA CARPENTER, THOMAS CHAVEZ

PROJECT DESCRIPTION

- In our project we would like to compare beer preferences across America by using the Kaggle data set, "Beer Reviews."
 - What is the correlation between taste and appearance?
 - What is the relationship between aroma and taste?
 - How many clusters do we have based on the appearance/taste ratings?
 - Are there outliers in the dataset?

PRIOR WORK

- <https://www.visualcapitalist.com/cp/mapped-beer-consumption-in-the-u-s/>
 - Mappings of popular mainstream beers (e.g. Budweiser, Stella) by state. This is different
- Surveys (ex: Akron Beacon Journal)
 - Found that Americans prefer lagers over pale ales, IPAs, and stouts
 - However, it does not go into more specifics
- There is an opportunity with this dataset to get more granular with what Americans prefer most about different types of beers.

DATASETS:

- List of dataset: Beer Reviews
- URL: <https://www.kaggle.com/datasets/rdoume/beerreviews>
- Where its downloaded- It has been downloaded to Tom's computer.

PROPOSED WORK:

- Data Cleaning:
 - check for null values
 - convert data types
 - remove duplicate data
- Data Preprocessing:
 - normalize values
 - Remove attributes that are not relevant to our analysis
- Data Integration:
 - Integrate with location data or beer style data if available

TOOLS

- The tools we will be using are:
 - Python
 - Pandas
 - Matplotlib
 - Numpy
 - Github
 - Kaggle Resources

EVALUATION

- To evaluate our project, we will:
 - show a trend in data between beer types and palate preferences
 - find clusters of beer preference correlations
 - identify popular styles and their related features