



FOR THE ON-THE-GO BEER CONNOISSEUR!

# Go Beers

gocalbeers.com

Alexina Chong  
Calvin Nutley  
Jonathan Manalus

## Why Big Beer Data?

- Amidst a declining nationwide beer market, the craft brewing industry has flourished, growing 18% in 2013 alone.
- Our recommendation system aims to expedite this process by suggesting beers tailored to a user's flavor preferences and that brewery's specific menu offerings.

## Competitors & Tools

- Craftbeers.com offers a similar app, but wasn't tailored to a user's location.
- We used OpenRefine, Pandas, and Python to clean and analyze our data.

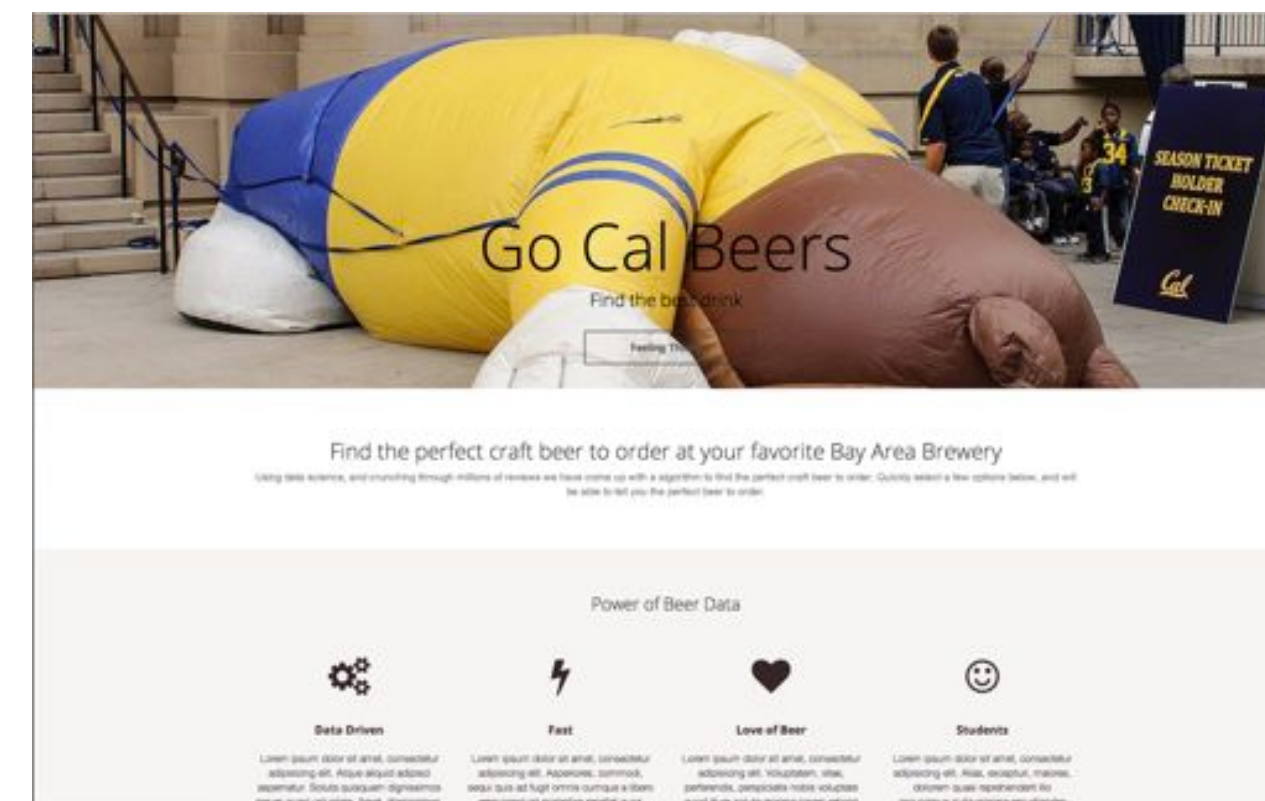
## Beer Advocate Data Sets

- A 1.5GB text file with over 1.5 million reviews from Jan 1998 until Nov 2011.
- Data found through Stanford SNAP.
- Schema for the dataset [beer name, beer id, brewer id, ABV, beer style, appearance score, aroma score, palate score, taste score, profile name, review text]
- Second dataset was manually collected from Beer Advocate to find all 39 Bay Area breweries.

## Data Cleaning & Integration

- Created our own script to parse the data text file into a proper CSV file.
- Developed a parser to handle the large text file and to help us better analyze the data.
- Our parser helped us clean the data and discover punctuation in string values of the data that interfered with iPython.
- OpenRefine helped us better understand our data.

## Application



Beer Survey

Please Note: This is a proof of concept, and that seasonal beers are constantly changing, and may not be available.

Where brewery in San Francisco are you currently located at?

21st Amendment   Anchor Brewing   Beach   Magnolia Gastrobrewery

Pizza Organics   Social Kitchen & Brewery   Southpaw BROS   Southern Pacific Brewing

Speakeasy Ales & Lagers   Thricebrewing

What style of beer would you prefer?

Blonde & Pale Ale   Brown Ale   Dark Lager   Fruit/Veggie/Novelty

IPA   Pils Lager & Pilsner   Stout   Strong Ale & Sour Wine   Wheat

Winter Warmer

What characteristics would you like in your beer?

Malty   Hoppy   Caramel   Sweet   Floral   Bitter

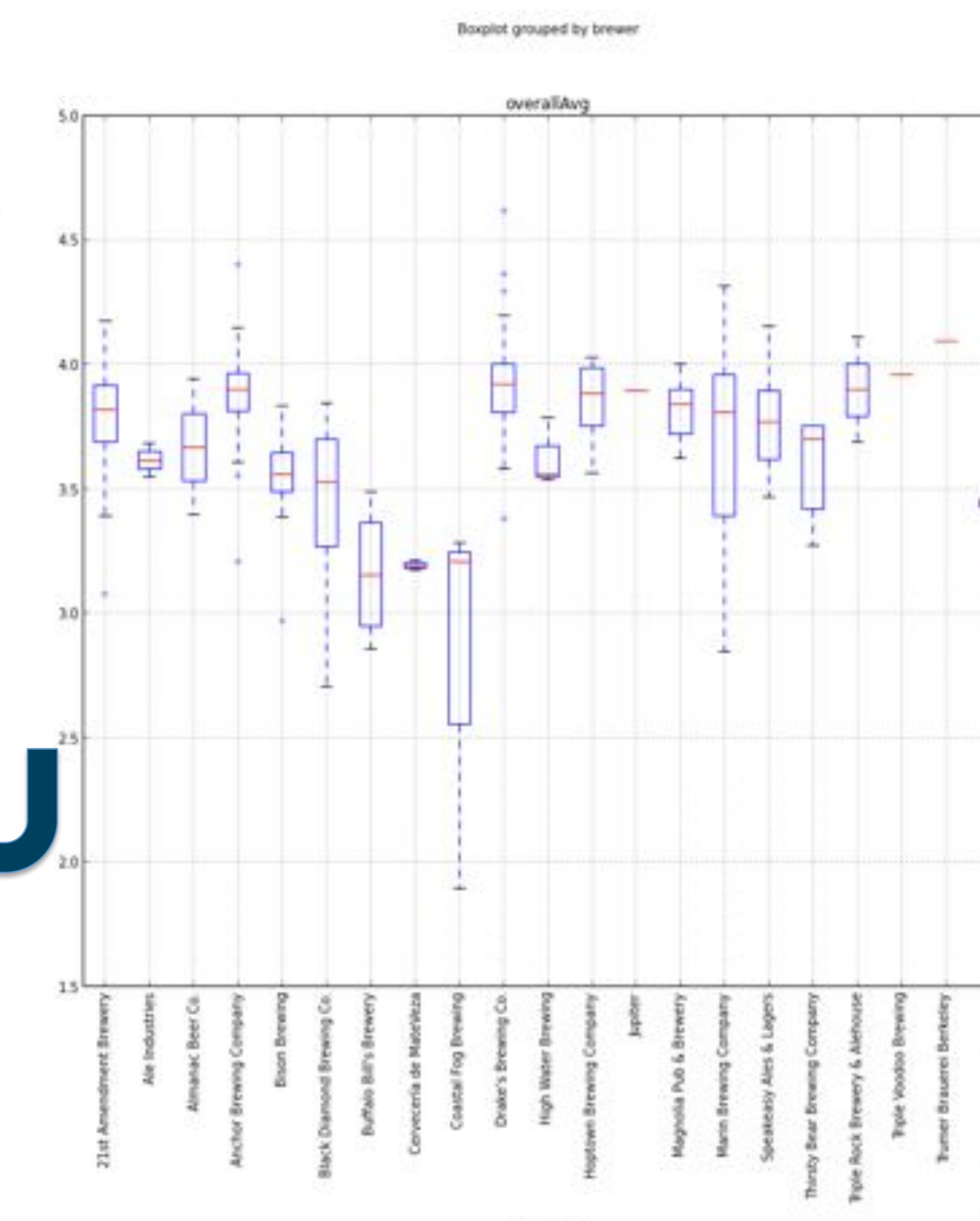
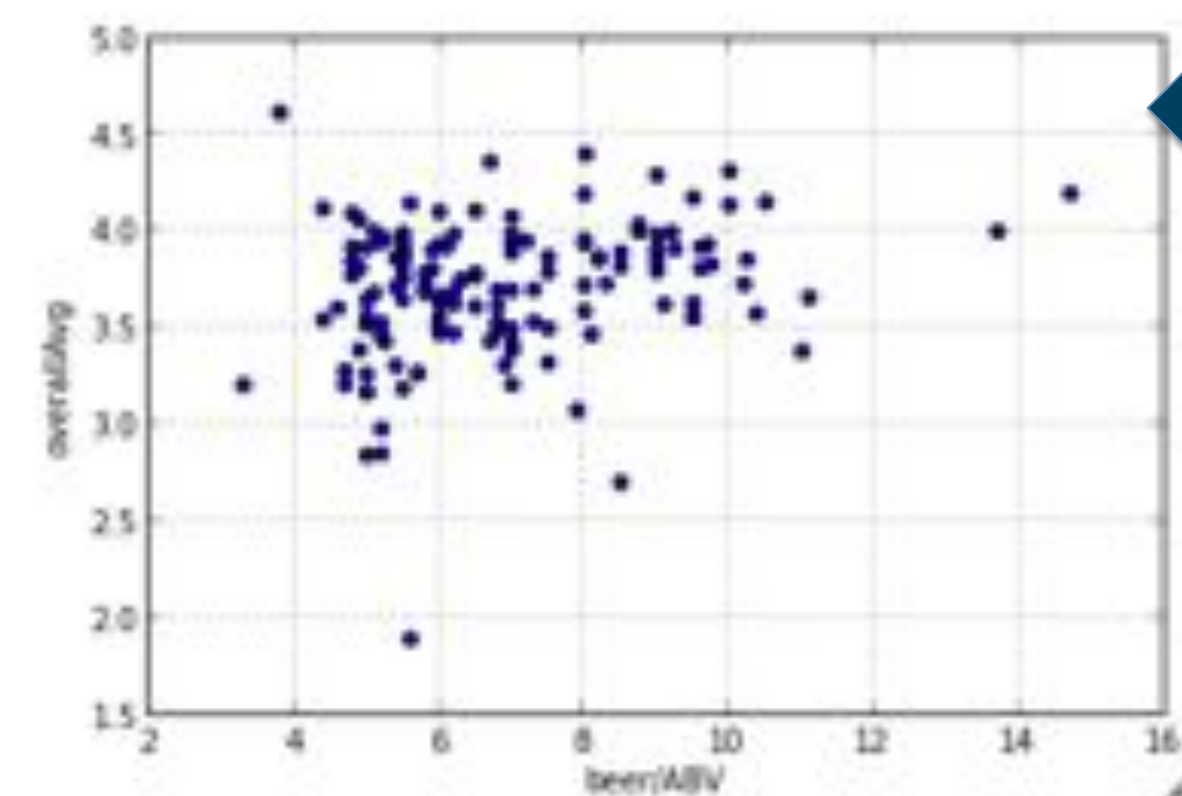
Nutty   Creamy   Light   Fruity   Earthy   Sticky

Chocolate

- Go Beers allows you to find the perfect craft beer based on your preferences at SF breweries, using a simple survey to find out what you enjoy in a beer.
- Responsive design works on any device.

## Data Graphs

- Distribution of beer ratings by brewery.
- ABV vs overall average rating: shows that beers with a lower ABV have a greater range of ratings.



- 50 most frequently used words in beer reviews for Bay Area breweries.



## Challenges & Lessons

- Our large dataset is a 1.2GB text file that kept crashing our VMs.
- Our dataset is outdated which caused problems since new beers are constantly being released, while old ones are discontinued.
- While we all consider ourselves to be decently familiar with broad beer style categories, we soon realized that the level of detail present in Beer Advocate's classifications were far beyond our understanding. Thus, we had to conduct extensive online research about beer styles so that we could accurately reclassify them into broader, more approachable categories.

## Open Issues & Extensions

- With outdated reviews, we have no way of knowing if our beers are current without checking the breweries' websites.
- We would love to create an API that would allow breweries to instantly update our app with any changes to their menus.
- We would like to extend our app by integrating map and location services to further streamline the recommendation system, and by expanding to other markets.
- In addition, we would like to use machine learning techniques to refine our results based on a user's prior selections and avoid recommending beers that the user has tried and disliked.