



BEER REVIEW ANALYSIS

Laura Liu

lauraliu0824@gmail.com

50,000 observations from 1999-2012

The goal of this project is to determine the most important characteristics of beer that lead a good review.



Data demographic

Overview of the main variables of 50,000 observations, especially about distributions and trends.



Text mining

Analyzing customer experience feedback using text mining and exploring the trends for flavor customers use when giving positive comments.



Modeling

Discussion about different models adapted to predict overall ratings and investigation about importance of different explanatory variables



Collected beer reviews from BeerAdvocate.com



Harpoon Sweet Spot

Harpoon Brewery & Beer Hall
American Blonde Ale / 4.80% ABV

3.5/5 rDev -2% | Score: 3.57

look: 3.5 | smell: 3.5 | taste: 3.5 | feel: 3.5 | overall: 3.5

L: brilliant gold in the glass supporting a 1" white foam cap with good retention and filled with streams of rapidly rising bubbles

S: Initial moderate levels of biscuit aromas yield to a pleasant hop note and then settle down to onion, garlic, some wood and fruit.

T: Fruit flavors on the palate suppress the onion and garlic but there is a trace of over ripe fruit that gives way to a herbal hop note on the finish. The malt biscuit flavor is medium and the hop flavors are medium high.

F: medium body and carbonation

O: full flavored beer has low alcohol so it is sessionable but not all of the flavors - in particular the fruit vs. the vegetable - fit well together

675 characters

pat61, 17 minutes ago



Madagascar

4 Hands Brewing Co.
Milk / Sweet Stout / 9.30% ABV

4.5/5 rDev +1.1% | Score: 4.45

look: 4.5 | smell: 4.5 | taste: 4.5 | feel: 4.5 | overall: 4.5

Yeah, this works for me! Lots of vanilla on the nose and taste. Just the right amount of barrel and sweetness. Feel is medium and slick. Overall a delicious stout that i would gladly drink anytime.

197 characters

BEER88, 18 minutes ago



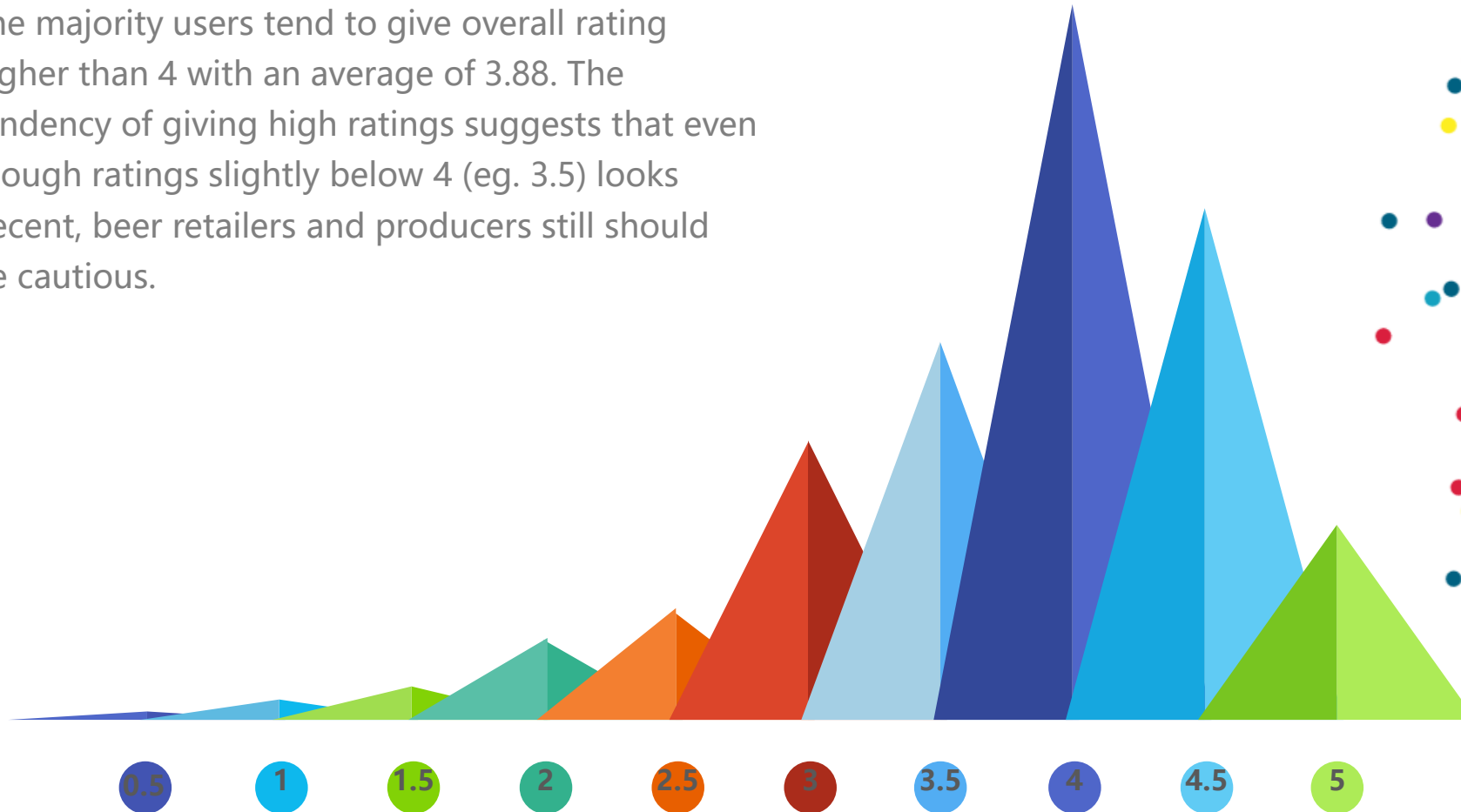
- The original reviews, before they were collected into a JSON file looked like the picture above.

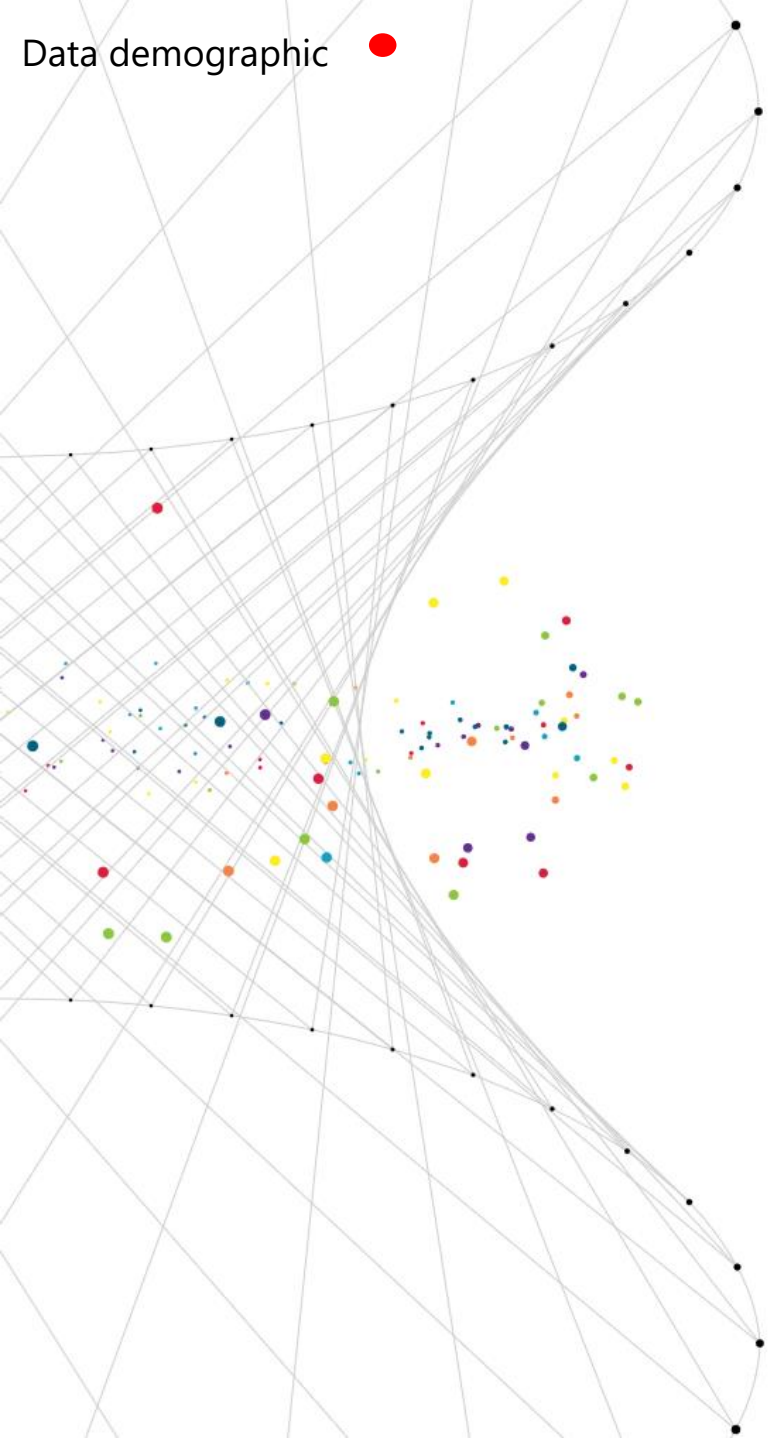


Data demographic

Distribution of overall rating

The majority users tend to give overall rating higher than 4 with an average of 3.88. The tendency of giving high ratings suggests that even though ratings slightly below 4 (eg. 3.5) looks decent, beer retailers and producers still should be cautious.

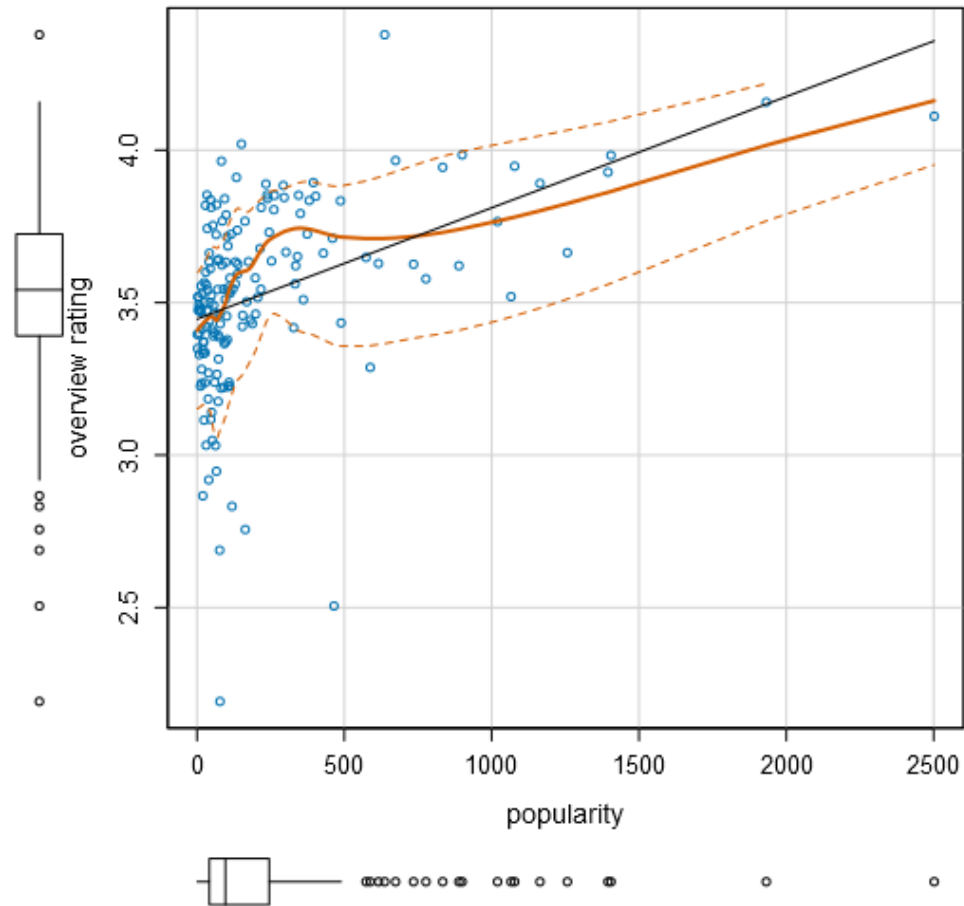




Overall rating for beers with different popularity

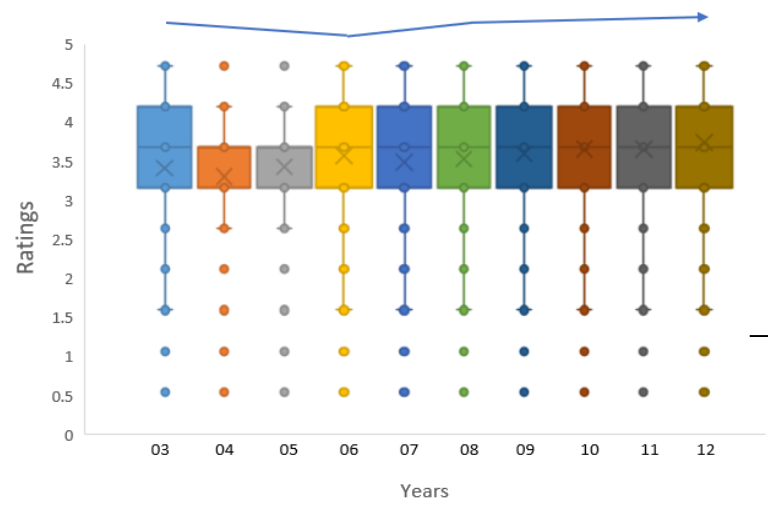
Popularity(number of reviews posted from tasters) and overall rating has significant correlation above 700 reviews. The beers with great popularity always have good ratings in customer feedback, or we could say with more and more good feedback, the beers tend to popular.

Scatterplot of popularity versus overview rating



Data demographic

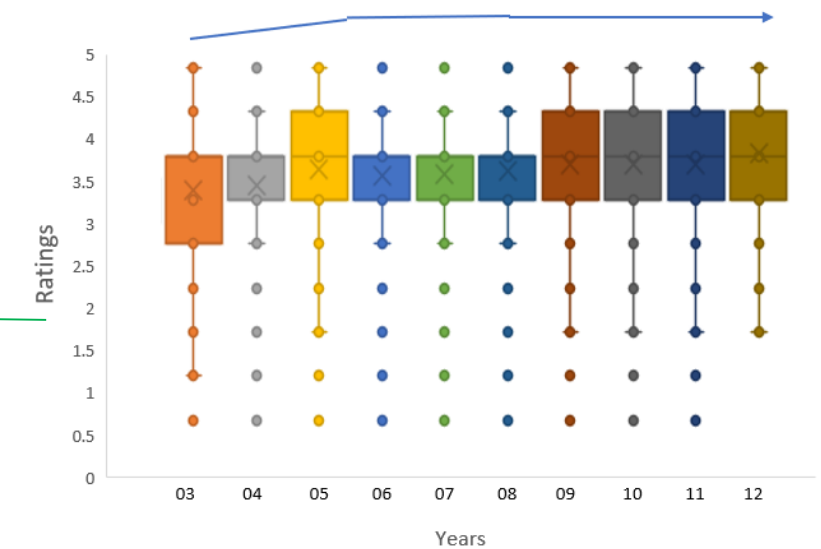
Taste 2003-2012



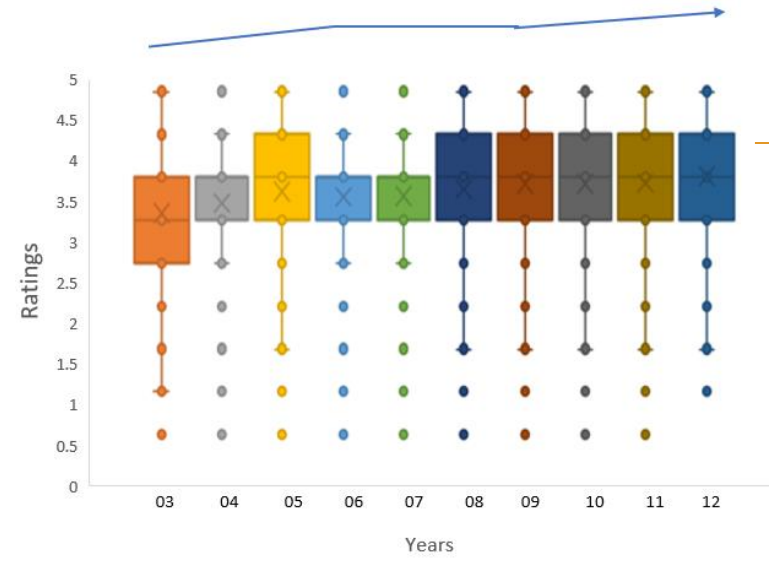
Taste

Feel

Feel 2003-2012



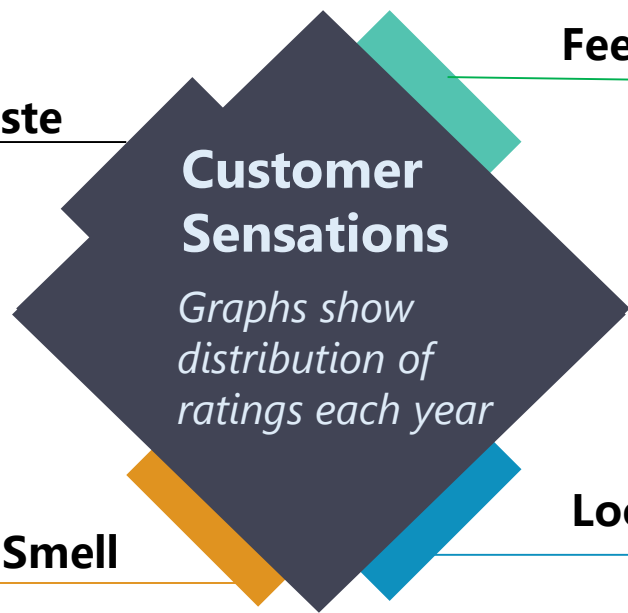
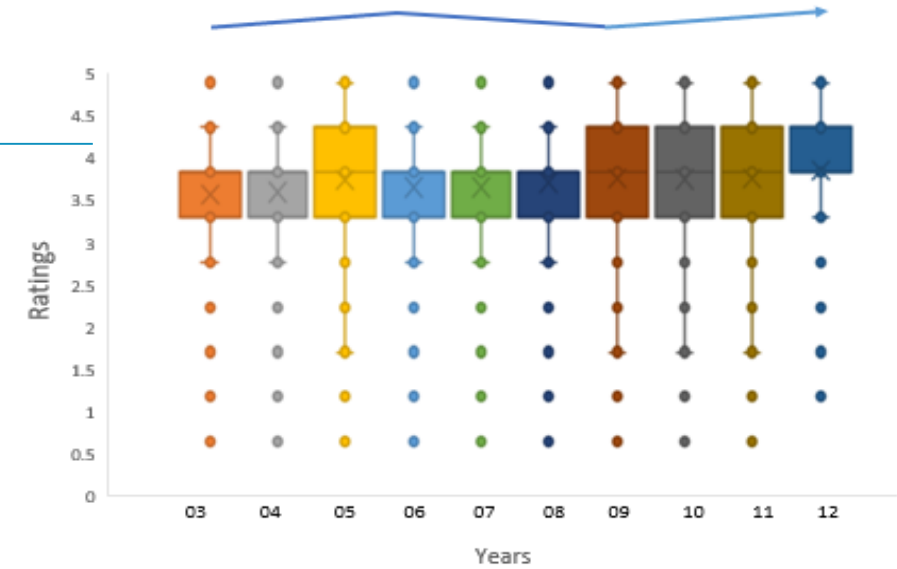
Smell 2003-2012

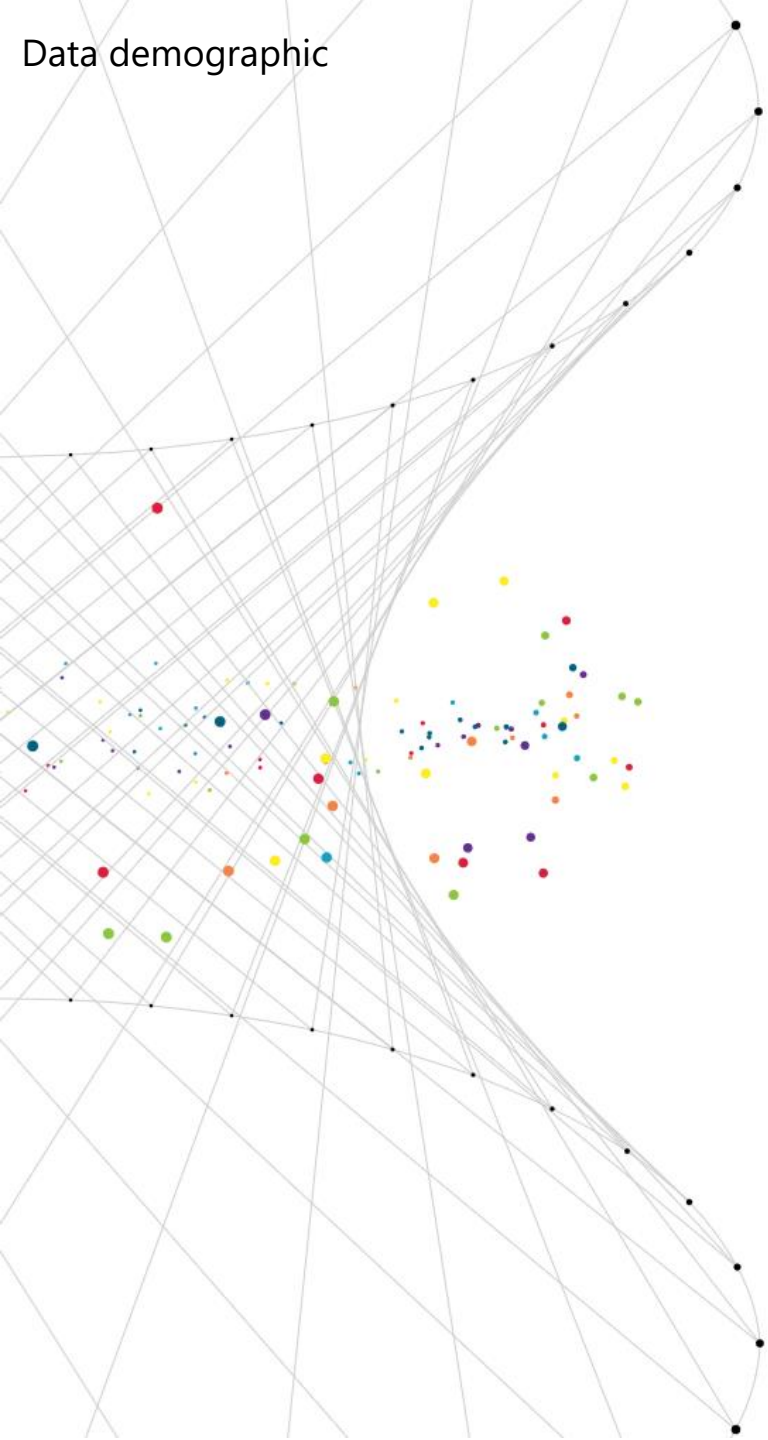


Smell

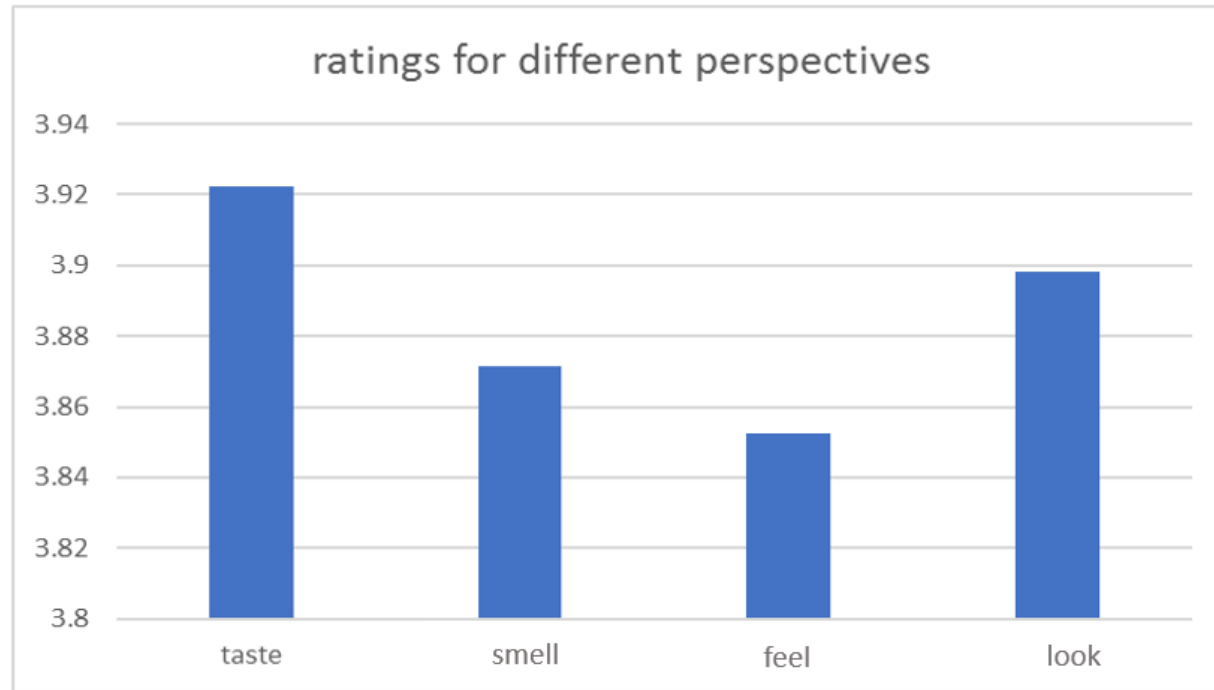
Look

Look 2003-2012

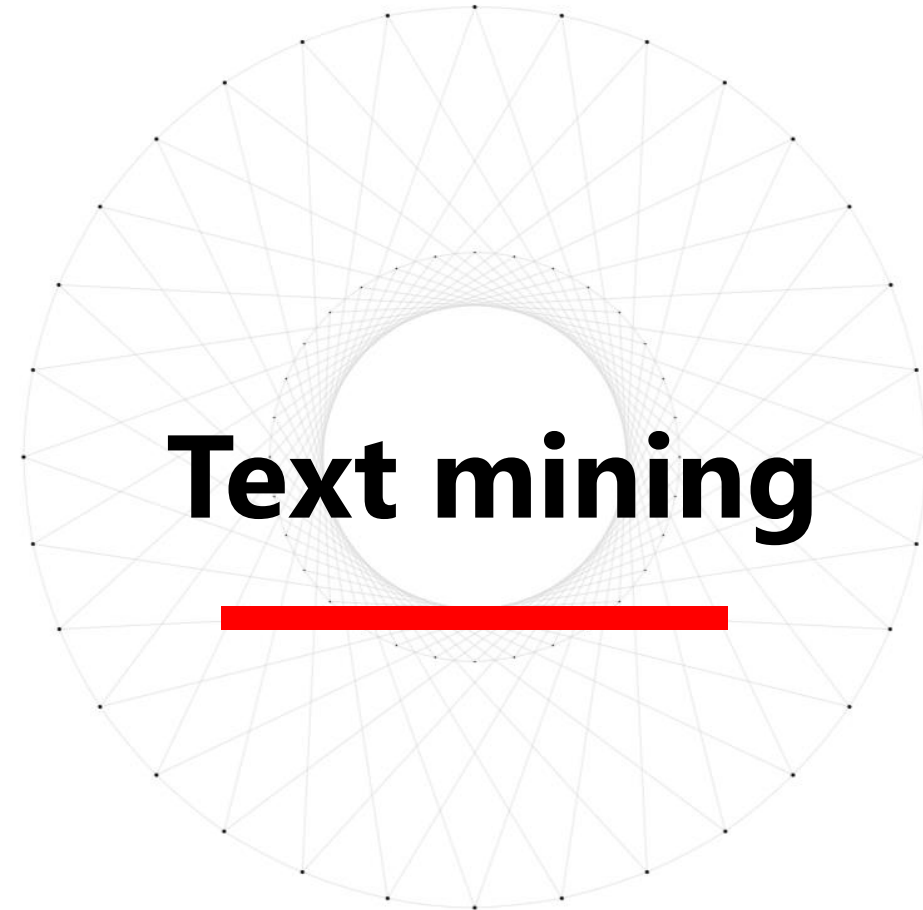




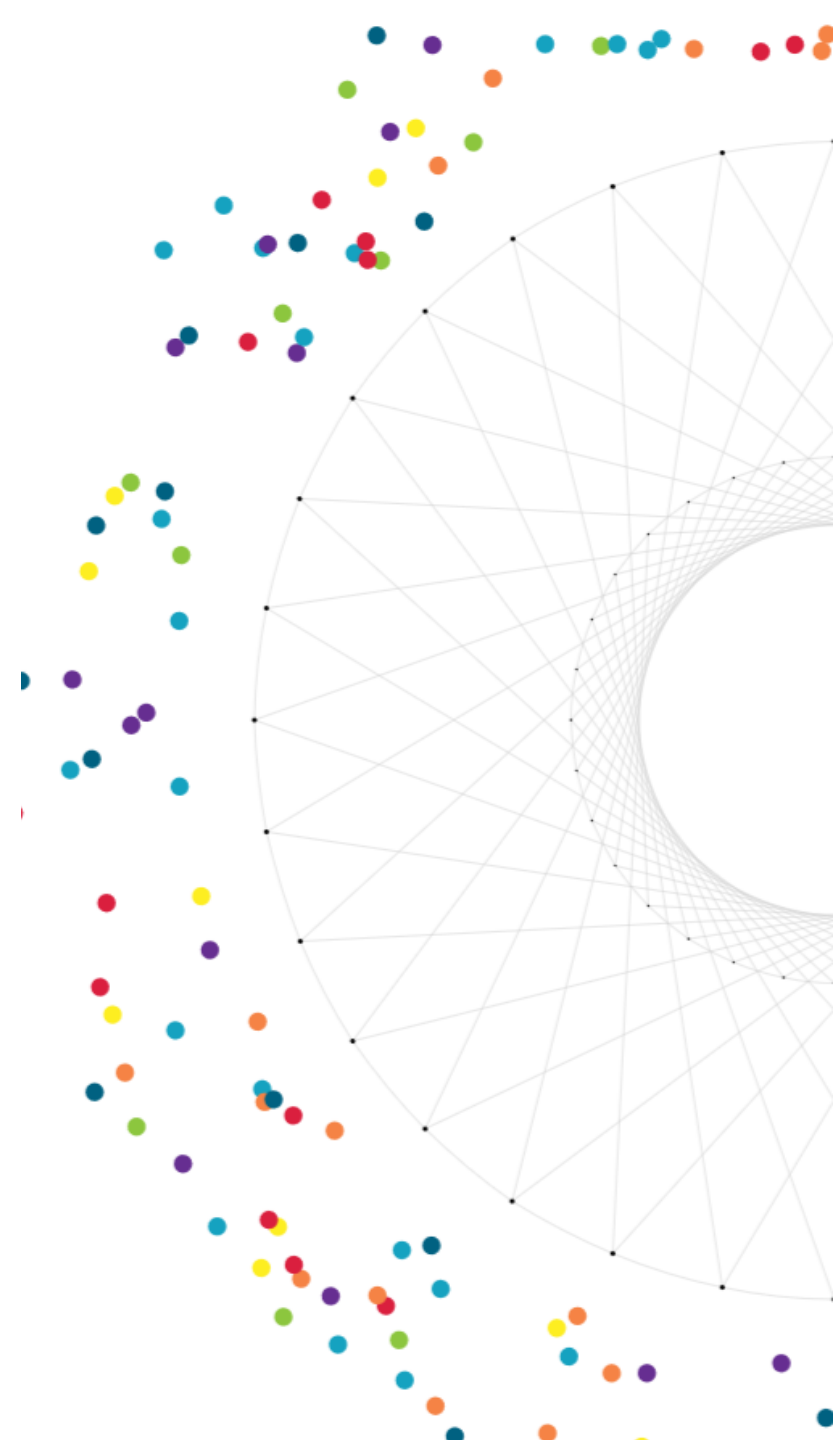
Average ratings for customer sensations



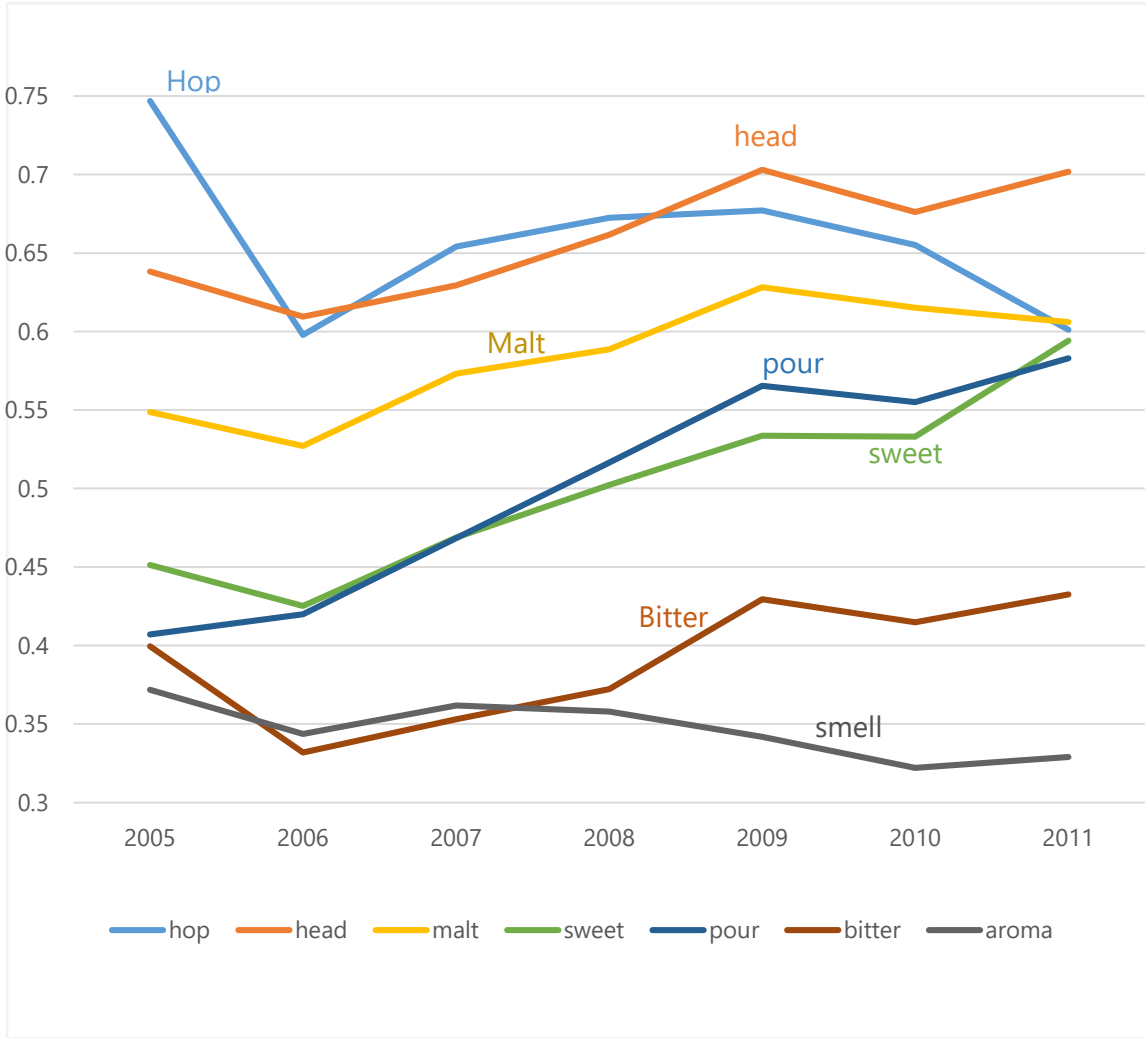
- Customers tend to give higher rating for taste and lower rating for feel.
- Customers tend to feel positive about the look over time.
- Feel is the primary factor to improve, since the customers opinion about it has not changed and also it is currently the lowest average rating which allows for a lot of improvement.



Most frequent stems in positive comments



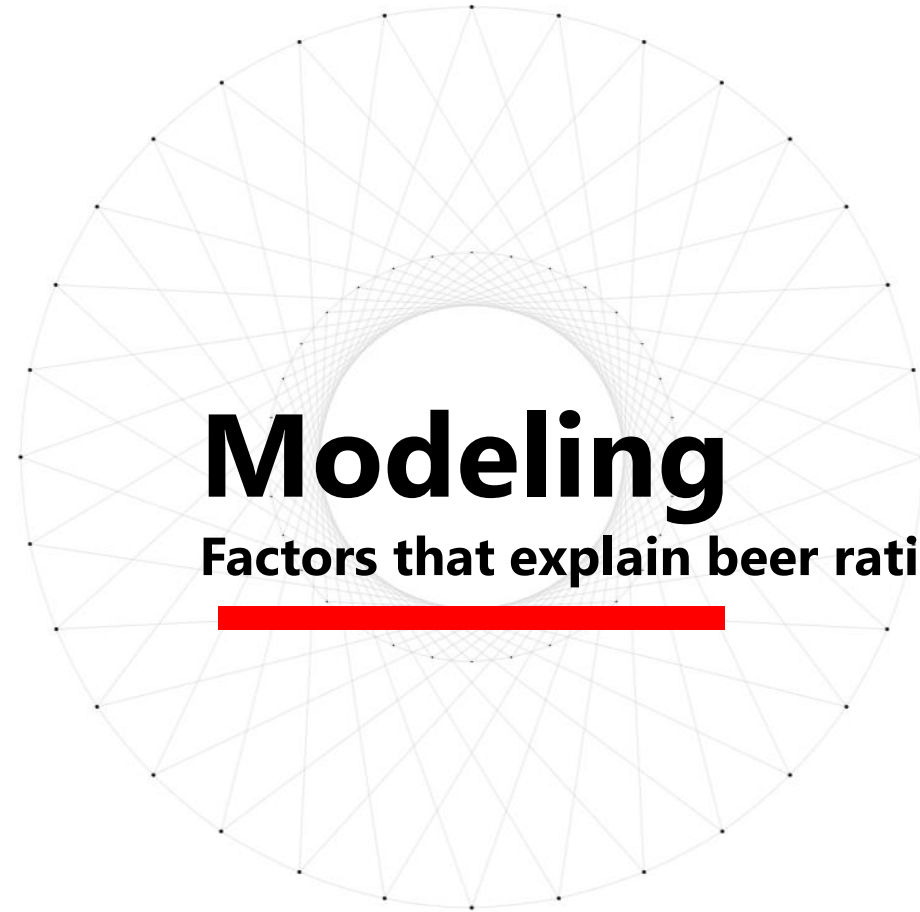
Trend of key words in positive review for beer



+ Hops contain an essential oil with a very bitter flavor, which relates to another important key word "bitter"

+ Beer head beats hop as the most aware key word after year 2008

+ People have increasing preference for terms sweet and pour



Modeling

Factors that explain beer ratings



Modeling

Linear regression

Generalized Linear Model

- The most important factors are **customer sensations**: taste, smell, feel, and look. **Taste** is the most important
- Beers with **higher alcohol volume** tend to have lower rating
- People mentioning **hops** in comments tend to give negative rating
- Higher **popularity** does not necessarily lead to higher average overall rating

Generalized Linear Model with Cross Validation

- Raters preferences do not change over time.

Generalized Linear Model with Interactions

- **Hop** has an increasingly positive effect on overall rating over time.
- **Hop** and **flavor** have increasingly positive effects on overall ratings with **alcohol volume**.
- **Malt** and **bitter** have increasingly negative effects on overall ratings with **alcohol volume**

Classification

Random Forest

- The most important factor is **taste**
- **User preferences** exhibit higher importance compared to popularity of beer name and beer style.

Support Vector Regression

- **Lowest MSE (0.1525)**
- **Look** has the least impact on overall rating compared with taste, feel, and smell.
- **Hop** is the most important factor among users preference in writing reviews.

Customer Sensations

According to the modeling, customer sensations are worthwhile to be explored

Look

- Describe the beer's color, clarity, head retention and lacing
- Look sets expectations and is frequently a sign of good or bad things to come
- Pour the beer into a clean glass and raise it to eye level for better viewing

Taste

- Describe any malt, hops, fermentation byproducts, balance, finish or aftertaste and other flavor characteristics
- Sip, let it warm and wander on the palate, swallow, and then breathe out. This process of exhaling is called retro-olfaction and will release retained stimulations at the mucus and mouthfeel level, but at a higher temperature

Smell

- Describe any malt, hops, yeast and other aromatics
- Breathe through your nose, then with your mouth open, and finally through your mouth only; the nose and mouth are connected in the experience. This process is called olfaction

Feel

- Describe the beer's body, carbonation, warmth, creaminess, astringency and other palate sensations
- Take another sip and let it wander, focusing on the beer's texture on your palate

BEER REVIEW ANALYSIS

Laura Liu

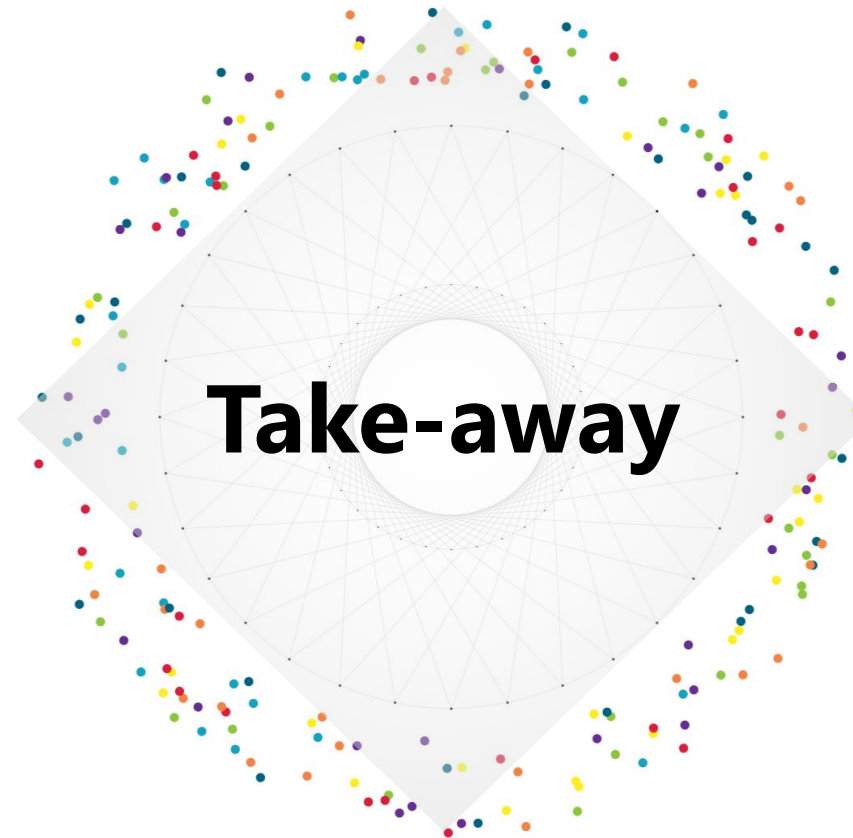
Customer sensations:

Customer ratings for four different Sensations have the greatest prediction power. More specifically, taste is the most important aspect and look is the least important. However, look has been given high ratings over years.

Feel needs to be improved and has the potential for development.

Popularity

The higher the popularity the higher correlation between popularity and overall rating. Lower popularity beer tend to have widely varying ratings.



Flavors for beer

Sweet beer tends to win more interest over time. Also, customers focus more on smell over time.

Hop is the most important factor. Its effects on overall rating increase with time and volume of alcohol. On the other hand, malt and bitter have an increasingly negative effect on overall ratings with alcohol volume.



Thank you

Laura Liu

50,000 observations from 1999-2012