

Explorations of Beer Data



Case Study 1
Andrew Leppla
Anish Patel

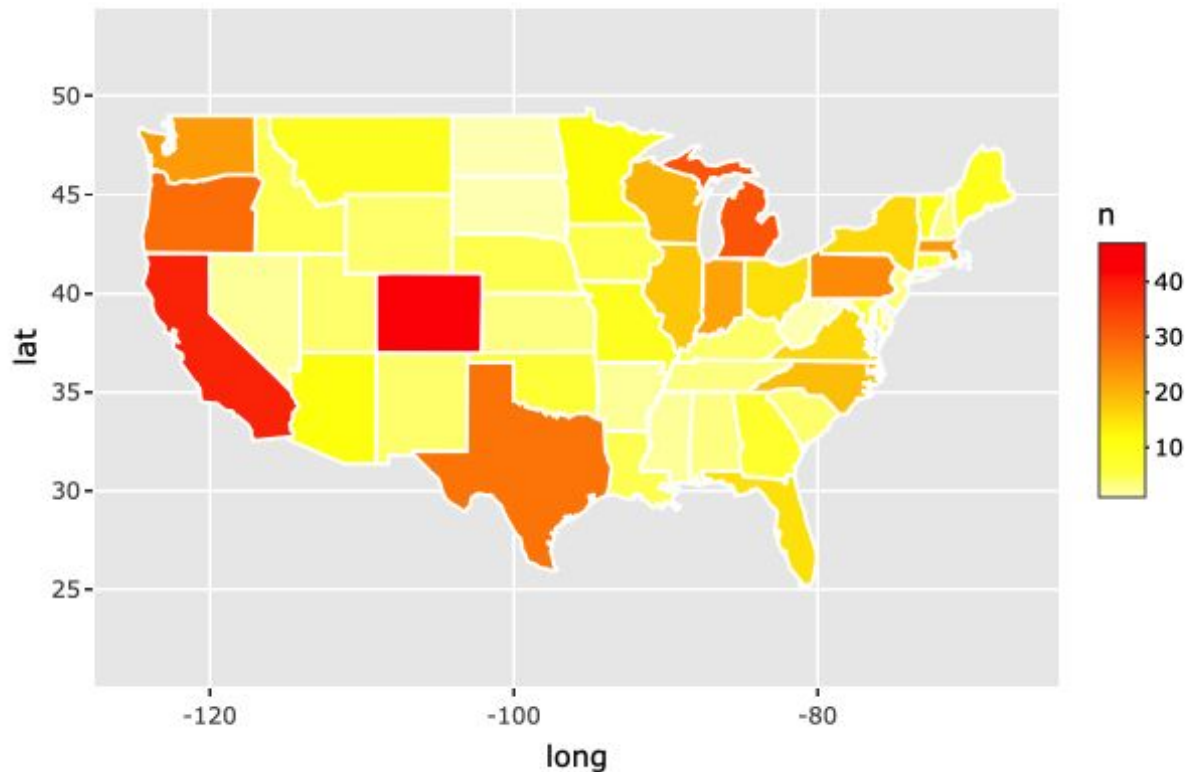
How Many Breweries are in Each State?

Each US Region has Dominant States

Top 10 States:

1. **Colorado (47)**
2. **California (39)**
3. **Michigan (32)**
4. **Oregon (29)**
5. **Texas (28)**
6. **Pennsylvania (25)**
7. **Massachusetts (23)**
8. **Washington (23)**
9. **Indiana (22)**
10. **Wisconsin (20)**

Count of Breweries by State (Continental US)



Dealing with Missing Values

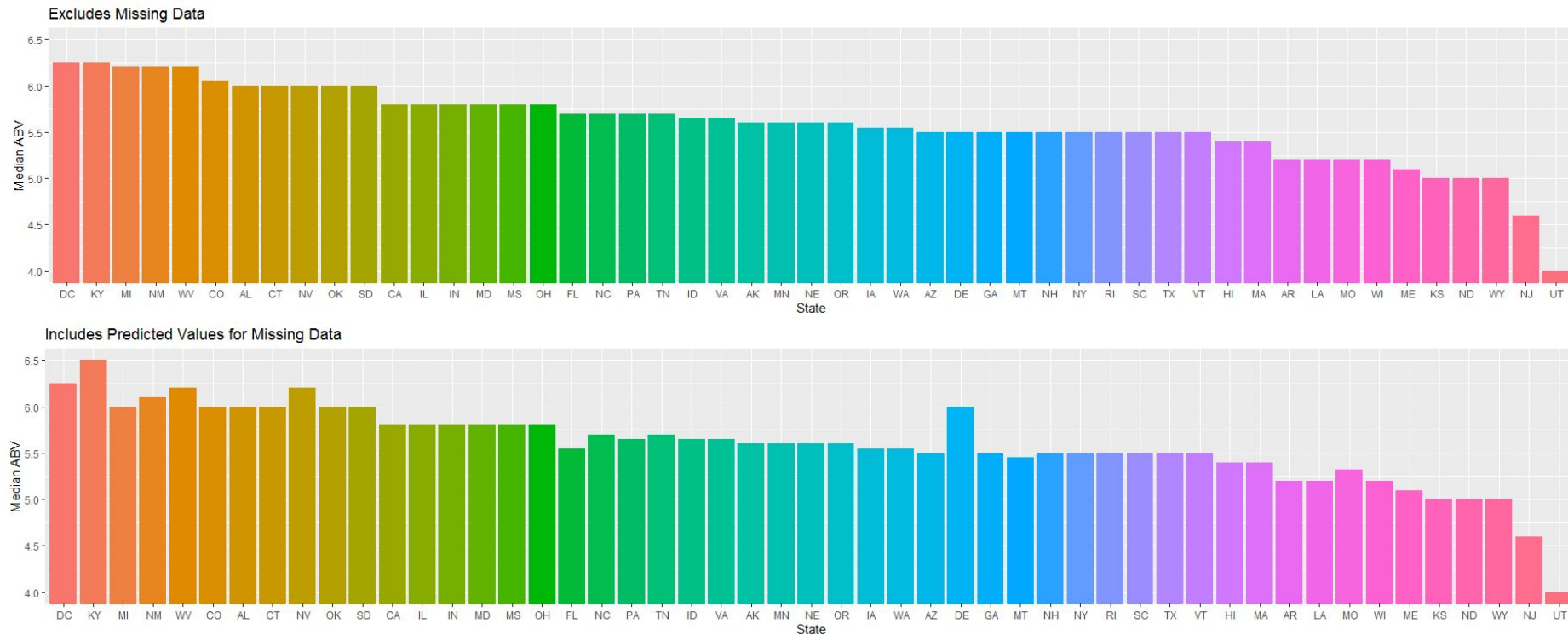
Two methods addressed Missing Values:

1. Exclude missing values
 - 5 Beer Styles, 62 ABVs, 1005 IBUs (2410 beers total)
2. Replace missing values based on Beer Style
 - Lagers have moderate ABV, low IBU
 - Double / Imperial IPAs have high ABV, high IBU
 - Ciders and Meads typically don't add hops (IBU=0)
 - A few exceptions



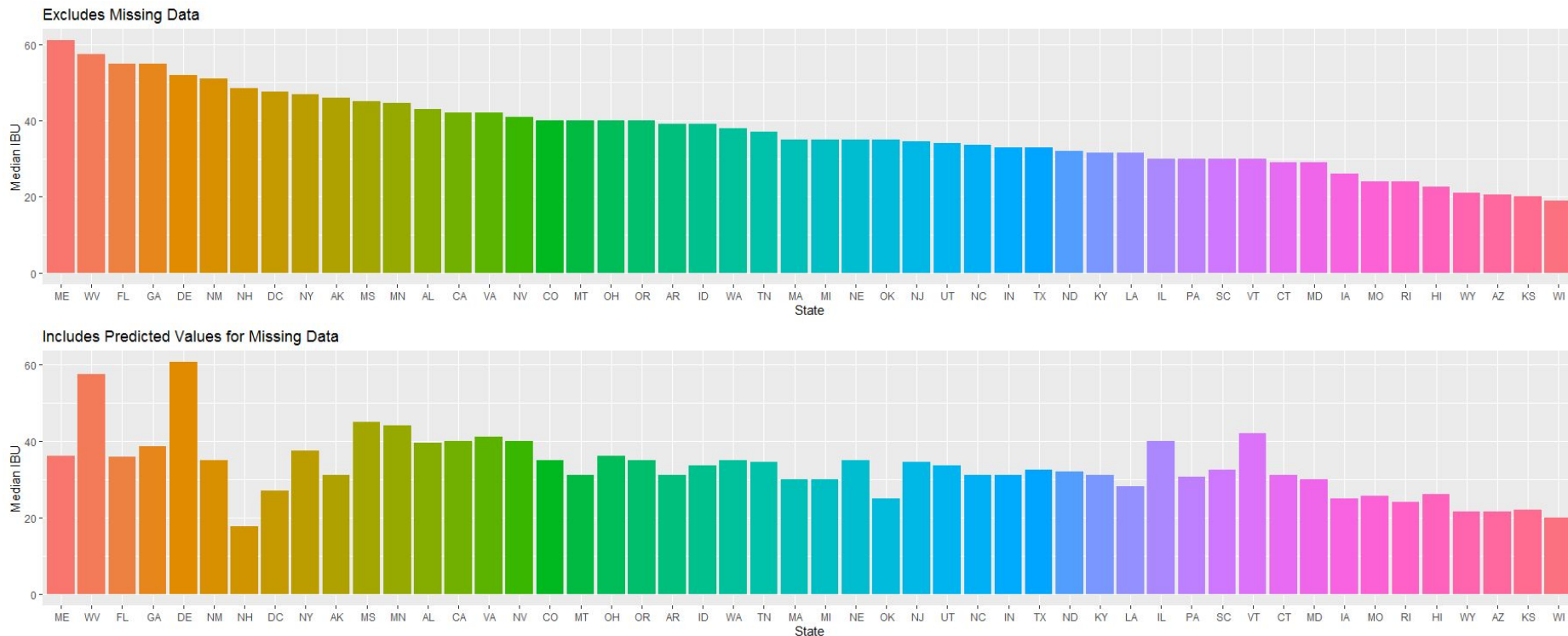
Median ABV by State - Utah is Lowest

2-3% Missing ABV Values have little effect



Median IBU by State

40% Missing IBU Values are Important! Need Data for Trends in IPAs.



States pushing the limits of ABV and IBU

Colorado has the max. alcoholic
(ABV) beer at 12.8%

"Lee Hill Series Vol. 5 - Belgian Style Quadrupel
Ale" by Upslope Brewing



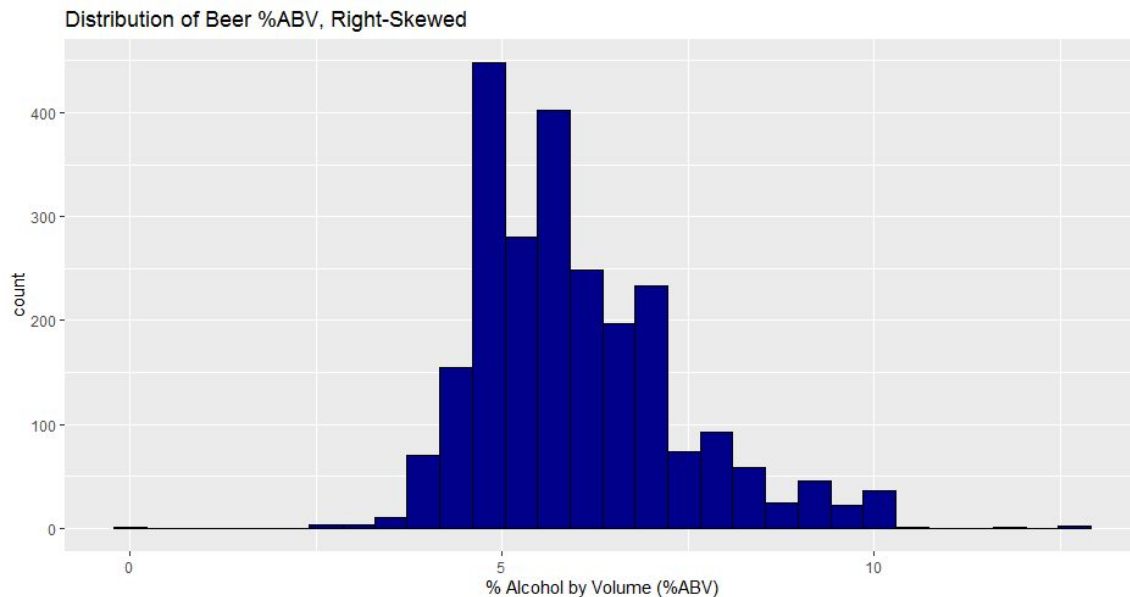
Oregon has the most bitter
(highest IBU) beer at 138 IBU

"Bitter Bitch Imperial IPA"
by Astoria Brewing Company



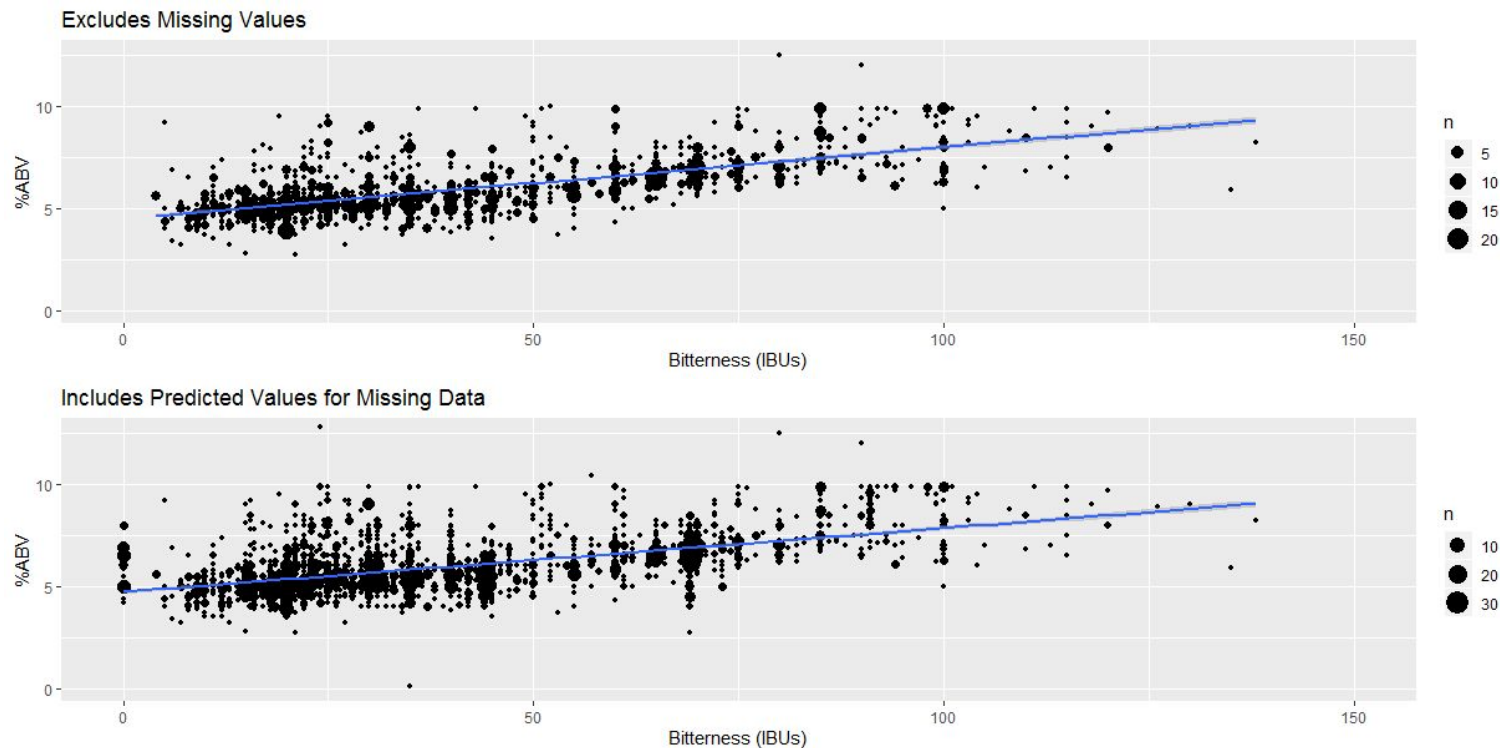
ABV Summary Stats

- Range of 0.1% to 12.8% Alcohol By Volume
- Median 5.6% < Mean 6.0%
- More “big” beers in the market are shifting mean ABV higher



Balanced Beers - Higher IBU's tend to have higher ABV

Roughly linear trend, Missing values don't really change it



Investigate differences in IBUs and ABV for India Pale Ales (IPAs) vs. other Ales

Approach: Use k Nearest Neighbors (kNN) to identify beers as IPAs or Ales based on IBUs and ABV

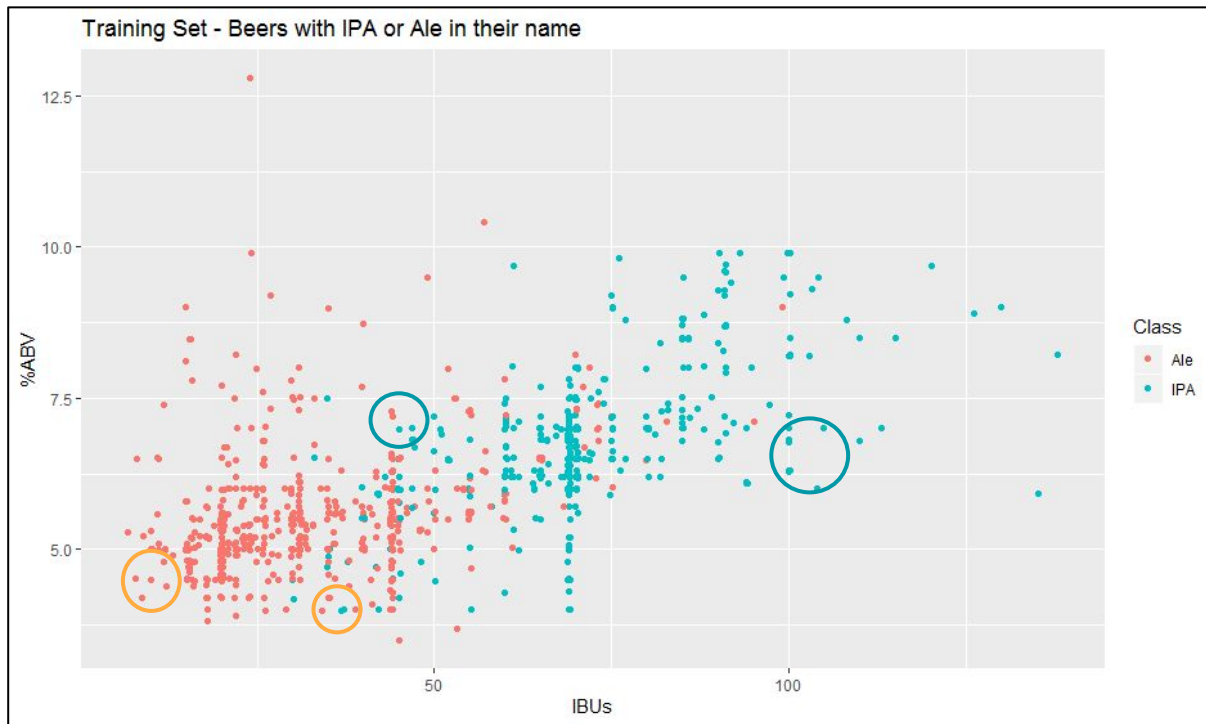
- Use beers with “IPA” or “Ale” in their name to predict if a beer without “IPA” or “Ale” in its name is an IPA or an Ale:

Train (60%): “Ranger **IPA**”, “Maine Island Trail **Ale**”



Predict (40%): “Peacemaker” = Ale

Model Visual & Prediction Results



Results

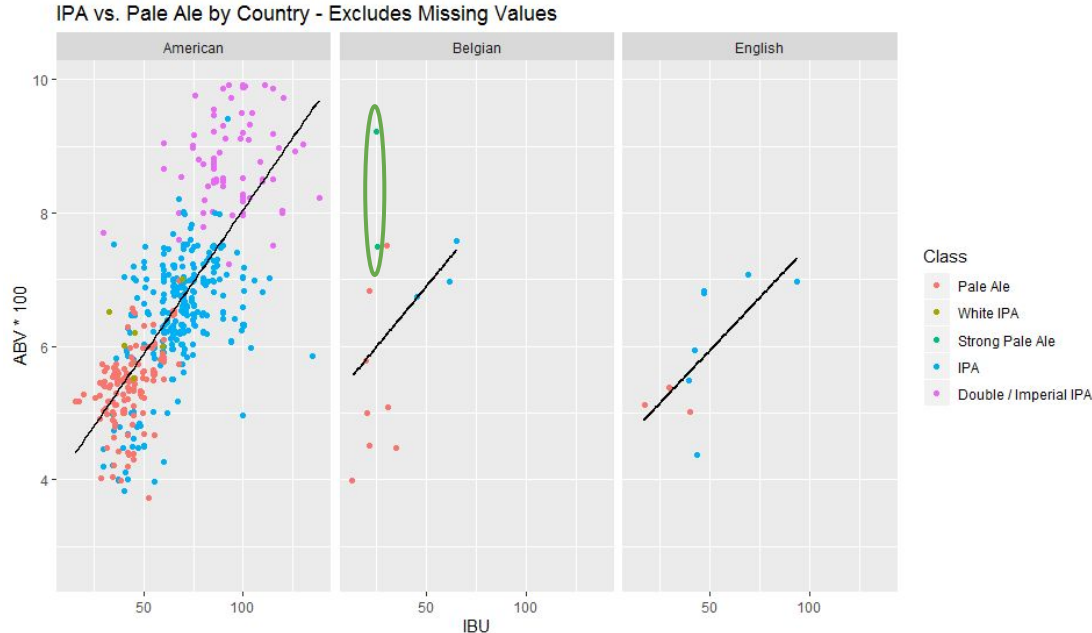
IPAs and Ales:
Accuracy = 90%

Ales:
Sensitivity = 95%

IPAs:
Specificity = 80%

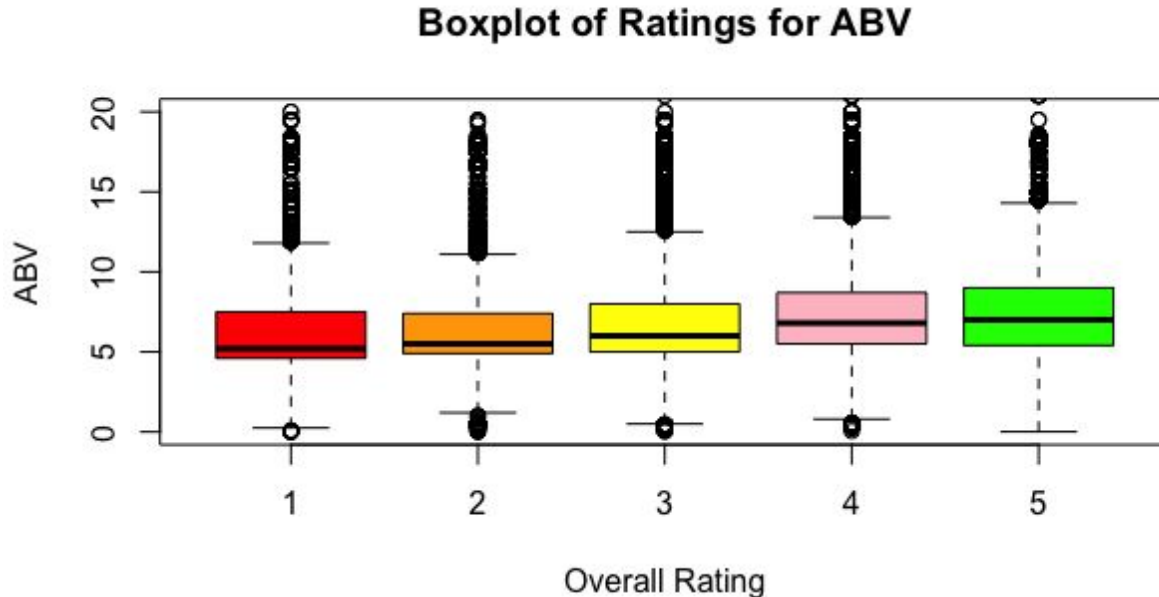
Deeper Dive - Pale Ales

- What makes an India Pale Ale (IPA) vs. a regular Pale Ale?
 - More Hops and Alcohol, Depends on the Country / Style
- American Double / Imperial IPAs are pushing beers into new territory



Deeper Dive - Beer Ratings and ABV

- What associations are seen between overall beer rating and ABV ?
 - Visually appears ABV increase, the overall beer rating increases



Links to YouTube Videos

Andrew Leppla:

<https://youtu.be/ZNBv8CjmSfY>

Anish Patel:

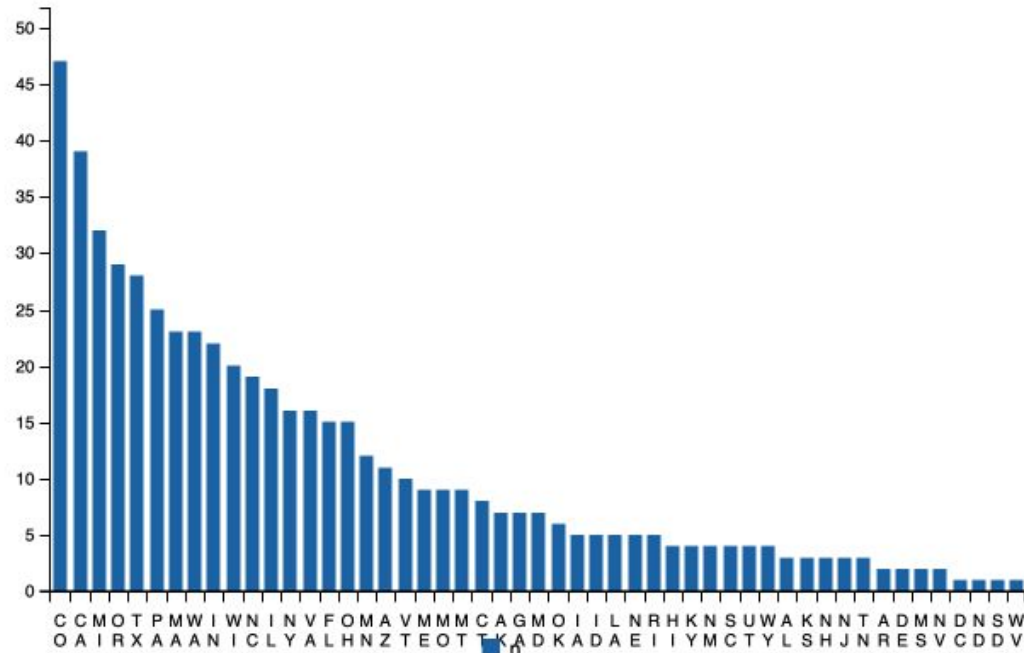
<https://youtu.be/QkKDy7jOZhA>

Appendix

States and Beer Markets

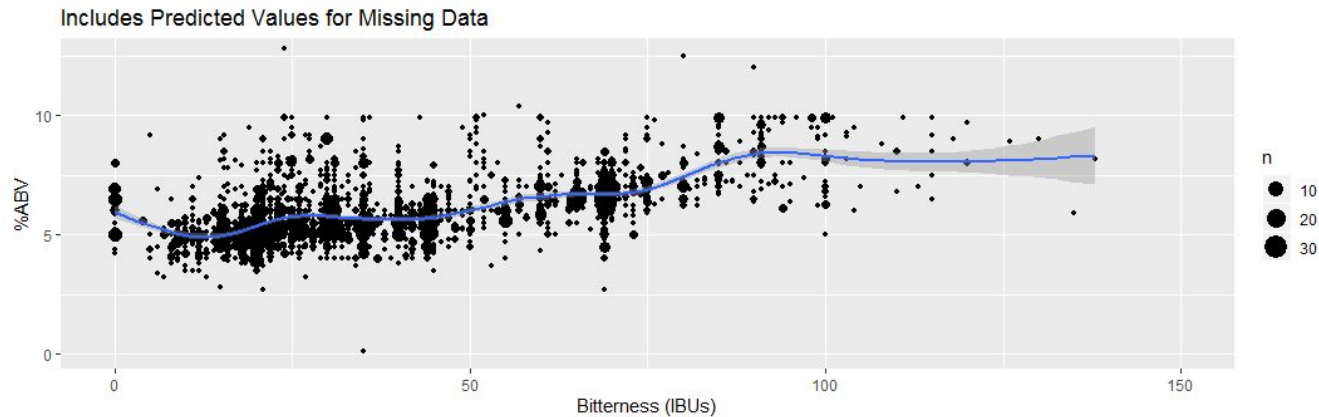
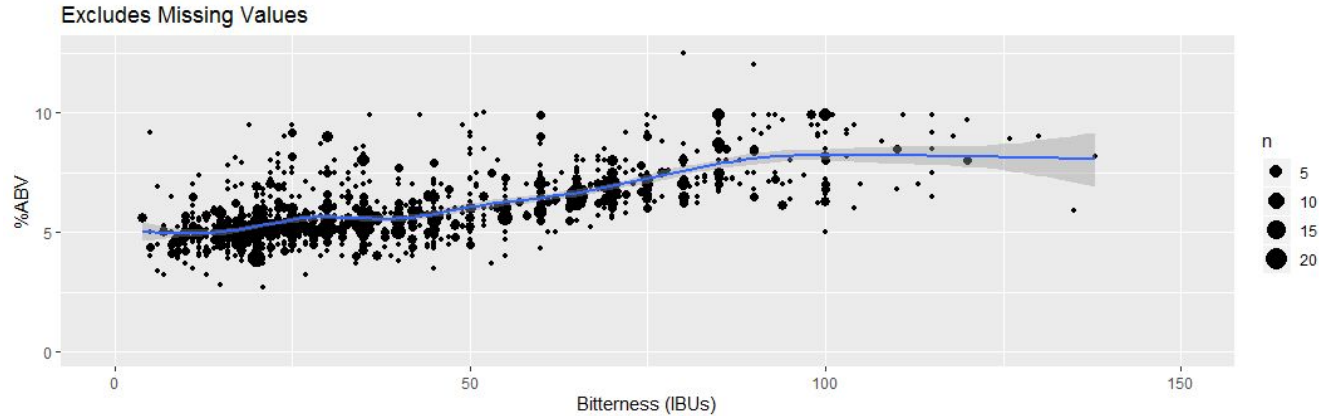
- States with many breweries are large markets but highly competitive.
- States with fewer breweries may be better for mass-market beers.

Count of Breweries by State



IBU-ABV Trend could be more complex

Some “bumps” and different linear sections



Next Steps

- Explore more current data. As of June 2019, there were 7480 active craft breweries in the USA (Active Licenses per Alcohol and Tobacco Tax and Trade Bureau) 1016 more than the prior year reported by [Forbes](#).
- Use web data scrapers to gather more data from social beer apps like Untappd