Q1 How many breweries are present in each state?

Using this bar chart we show the number of breweries in each state from the most breweries to the least per state.

Q2 Merge beer data with the breweries data. Print the first 6 observations and the last six observations to check the merged file.

Using the head function we show the first six observations and the last six observations.

Q3 Address the missing values in each column

We filtered the N/A’s out only when we did statistical calculations. We did not remove or delete any data for any other reason.

Q4 Compute the median alcohol content and international bitterness unit for each state. Plot a bar chart to compare.

Using the bar chart we were able to display the median ABV and IBU per state using an overlay of the two.

Q5a Which state has the maximum ABV beer?

Colorado has the highest ABV beer with a 12.8% alcohol beer

Q5b Which state has the maximum IBU beer?

Oregon has the beer with the highest IBU at 138

Q6 Comment on the summary statistics and distribution of the ABV variable

It appears the beer sweet spot is around 5.5% ABV as that’s the average mean of all beers in the data set. There are quite a few states that prefer a higher ABV and those typically have a higher number of breweries associated.

Q7 Is there an apparent relationship between the bitterness of the beer and its alcoholic content? Draw a scatter plot. Make your best judgement of a relationship and EXPLAIN your answer.

Yes, using the scatter plot we can see the higher the IBU (the more bitter) the beer the higher the ABV. The judgement can be made that the longer the hops are brewed the more bitter the beer. The end result also brings a higher alcohol content. Alcohol is could also be considered bitter so simply the higher the ABV the more bitter the beer due to the alcohol alone.

Q8 Budweiser would also like to investigate the difference with respect to IBU and ABV between IPA’s (Indian Pale Ales) and other types of Ale (any beer with “Ale” in its name other than IPA.) You decide to use KNN clustering to investigate this relationship. Provide statistical evidence one way or the other.

Using KNN we were able to predict with 83% success rate if a beer was an IPA or not. With IPA having a much higher alcohol content and higher IBU than traditional ale style beers we could predict the style of beer simply based on ABV and IBU.

Q9 Knock their socks off! Find one other useful inference from the data that you feel Budweiser may be able to find value in. you must convince them why it’s important and back up your conviction with appropriate statistical evidence.

In conclusion we suggest you sell new conceptual products to the states that have the most breweries per capita. Also, you should continue to do a hard push in the states that have the least amount of breweries per capita. Even possibly a new product with a lower ABV as it appears there is a market for a lower ABV product.

We would like to do more research on this with actual sales data to provide even more confirmation to where they should focus on new and existing products. Possibly partner with our distribution and retail chains for their beer year over year beer sales.