**Assignment 4 - Pandas Data Interpretation**

**Introduction**

My father owns a pub and with the COVID-19 public safety measures, he was forced to close the doors for longer than a single day for the first time in 58 years! Before the Lockdown, the pub was only closed on Christmas Day and Good Friday.

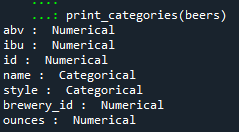
Although the pubs are closed, people in Ireland are still enjoying alcohol, albeit a shop-bought and home-consumed experience. Craft beer is one product which consumers have shown interest in in recent years (Bord Bia, 2018).

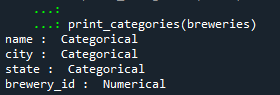
With this in mind, I examined a Craft Beers dataset from Kaggle. This dataset comprised of two CSV files on information about **canned**beers and breweries in the **United States**. This dataset was described as being a sample of the craft beers in the United States because it did not include every single beer available.

**Exploratory Data Analysis**

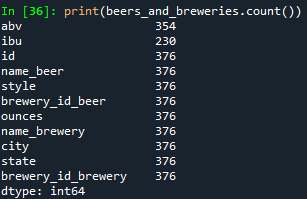
First I joined the beers and breweries datasets together. This was achieved using a Pandas Merge.

Then I identified the data types of each of the variables in the data set.

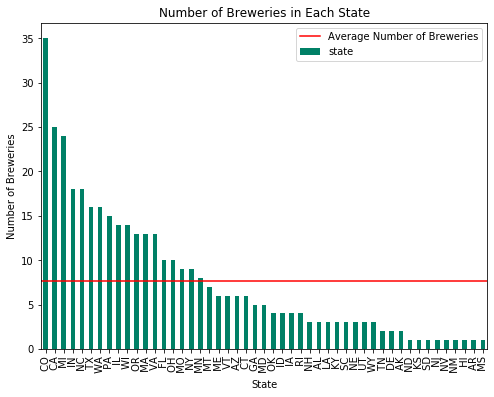
 *Image 1: Data types of the Beer dataset*

* Image 2: Data types of Breweries dataset*

Using a df.shape function, it was determined that there 376 rows and 11 variables in the dataset. Using df.count(), it was shown that there was some missing data in the ABV and IBU columns (see image 3).

 *Image 3: Identifying missing data*

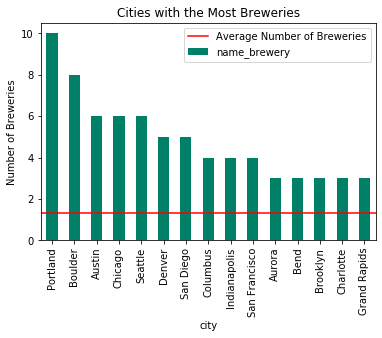
**Q.1 Which State has the most breweries?**



*Figure 1: Number of breweries per State*

As evidenced in Figure 1, Colorado had the most breweries than any other State.

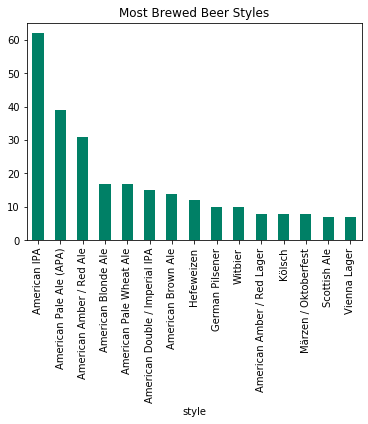
**Q.2 If you were interested in brewery tours, which is the best city to visit?**



*Figure 2: The number of breweries per U.S. city*

As evidenced in Figure 2, Portland (OR) and Boulder (CO) are the best cities to visit if you are interested in visiting craft breweries.

**Q. 3 What is the most brewed beer style in U.S. craft breweries?**

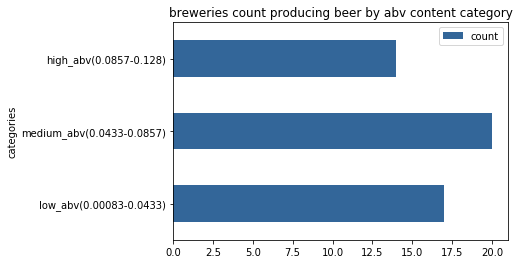
****

*Figure 3: Ranking the most produced U.S. craft beer types*

As, evidenced in Figure 3, American IPAs are the most commonly brewed beer at craft breweries in the United States.

**Q. 4 What is the average Alcohol Per Volume (ABV) metric for the beer in this dataset?**

There was a total of 376 individual beers in the dataset. 22 beers had no ABV value. The ABV level of the beer ranges from 0.1% to 12.8%. The average ABV reading is 5.97%. As evidenced in figure 4, most breweries produced medium strength beer.

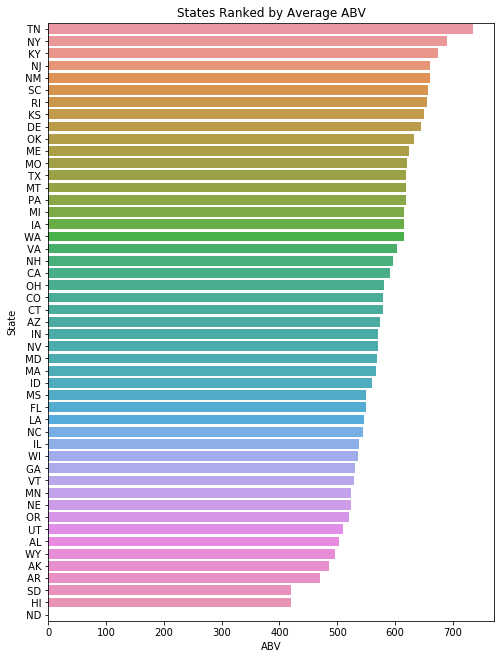
****

*Figure 4: The proportion of beer produced per ABV category*

**Q. 6 Which States brew the strongest beer?**

Considering that most breweries produced medium ABV beer, which State produced the strongest beer?

As evidenced in figure 6, Tennessee has the highest average ABV in their beers, while Hawaii has the lowest.

****

*Figure 6: State ranking per average ABV in beer produced*

**References**

Bord Bia (2018) *Craft Beer and Microbreweries in Ireland Report* [PDF], available at <http://icbi.ie/wp-content/uploads/2018/11/Craft-Beer-Report-2018.pdf>, date acccessed: 15/05/2020

Dataset used: <https://www.kaggle.com/nickhould/craft-cans/data?select=breweries.csv>