Variable Description Document

Data Set – Beers

* Name: Name of the beer.
* Beer\_ID: Unique identifier of the beer.
* ABV: Alcohol by volume of the beer.
* IBU: International Bitterness Units of the beer.
* Brewery\_ID: Brewery id associated with the beer.
* Style: Style of the beer.
* Ounces: Ounces of beer.

Data Set – Breweries

* Brew\_ID: Unique identifier of the brewery.
* Name: Name of the brewery.
* City: City where the brewery is located.
* State: U.S. State where the brewery is located.

Data Set – BeerBrewery

* Merged From the 2 data sets above by Brewery\_ID and Bew ID
* 2 Added Columns
  + IBUFac – A factorized version of IBU that makes the NA observations into a level
  + ABVFac – Factorized version of ABV cut into 3 levels
    - Strong ( > 0.07)
    - Normal (0.04 - 0.07)
    - Weak (< 0.04)

Data Set – BeerBreweryNA.Delete

* Same as BeerBrewery but with NA IBU entries dropped

Data Set – BeerBreweryLMPredicted

* Same as BeerBrewery but with NA IBU entries imputed with predictions from linear model 2.
* Not in the codebooks due to not being used in final presentation.

Data Set – BeerBreweryKNNPredicted

* Same as BeerBrewery but with NA IBU entries imputed with a KNN 20
* Included in Code Books – data was used in final presentation
* Includes the ABV and IBU columns – but these have been scaled due to the imputation
* ABV1 and IBUPred are the unscaled data.

Data Set - Bud

* Used to create the IPA or Ale classification

Data Set - Alcohol Consumption Ranked by State

* State
* Consumption in Gallons
* Consumption in Liters
* State population