Analysis of Domestic Taste Trends

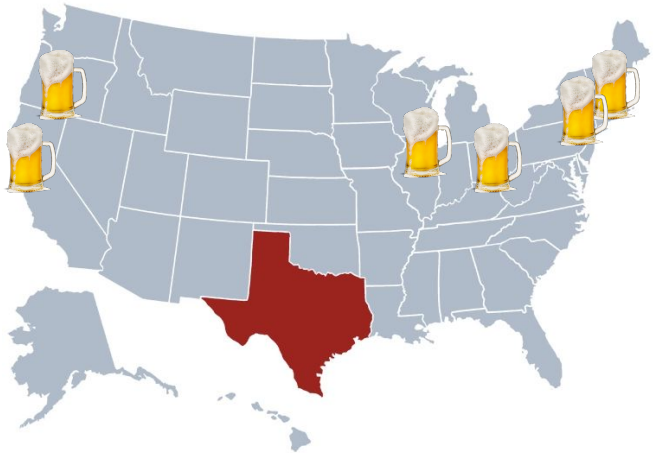
Prepared by TART

Context

June 2018 - Good Beers (GB) is looking to expand its existing chain of local breweries and gain additional market share of the domestic Bar and Nightclub industry. To date, GB has maintained a stronghold in Texas by operating 12 breweries across Dallas, Austin, Houston, and San Antonio, and it has become a main distributor of its classic beer, which contains a relatively high ABV and IBU percentages in comparison to their peers.

Business Strategy

GB has engaged TART to perform an analysis of raw data collected over a period of 6 months to determine whether the ABV and IBU scores of its most popular selling beer will be palatable to domestic targets identified for its initial expansion. **This includes Portland, San Francisco, Chicago, Cleveland, Boston, and New York.**



The following slides will provide the results of TART's analysis.

Breweries in US

Target areas selected by GB seem aggressive based on the existing presence of regional breweries. A more conservative approach may be to lean into growing markets with less competition.

##	State	Total Breweries
## 1	AK	7
## 2	AL	3
## 3	AR	2
## 4	AZ	11
## 5	**CA**	**39**
## 6	CO	47
## 7	CT	8
## 8	DC	1
## 9	DE	2
## 10	FL	15
## 11	GA	7
## 12	HI	4
## 13	IA	5
## 14	ID	5
## 15	**IL**	**18**
## 16	IN	22
## 17	KS	3

##	State	Total Breweries
## 18	KY	4
## 19	LA	5
## 20	**MA**	**23**
## 21	MD	7
## 22	ME	9
## 23	MI	32
## 24	MN	12
## 25	MO	9
## 26	MS	2
## 27	MT	9
## 28	NC	19
## 29	ND	1
## 30	NE	5
## 31	NH	3
## 32	NJ	3
## 33	NM	4
## 34	NV	2

##	State	Total Breweries
## 35	**NY**	**16**
## 36	**OH**	**15**
## 37	OK	6
## 38	**OR**	**29**
## 39	PA	25
## 40	RI	5
## 41	SC	4
## 42	SD	1
## 43	TN	3
## 44	TX	28
## 45	UT	4
## 46	VA	16
## 47	VT	10
## 48	WA	23
## 49	WI	20
## 50	WV	1
## 51	WY	4

```
library(plyr)
breweries_state <-
count(breweries$State)
colnames(breweries_state) <-
c("State", "Number of
Breweries")
breweries_state
```

Code uses the **PLYR library in R** to leverage the count function to determine the number of breweries in each state.

Data Integrity

Upon merging the two (2) datasets provided, TART identified instances where several values were missing - this gap was in the ABV and IBU scores.

… more specifically, 62 observations did not contain an ABV score and 1005 were missing IBU

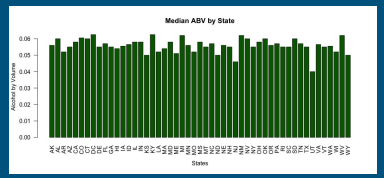
Tahir Ahmad, TART

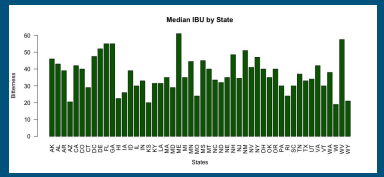
```
colnames(beers)[5] <- "Brew_ID"
beers2 <- merge.data.frame(beers, breweries, by =
"Brew_ID")
colnames(beers2)[2] <- "Beers Name"
colnames(beers2)[8] <- "Breweries Name"
#Print the first 6 observation
head(beers2,6)
tail(beers2,6)
colSums(is.na(beers2))
```

Column names were modified to create a common Brew_ID **before invoking the merge function** and to help delineate between beers names and breweries.

Median ABV and IBU by State

The median ABV and IBU scores across the US are 5.6% and 35 respectively. Medians for ABV ranged between .05 and .06, with the exception of UT and NJ. IBU medians fluctuated more; however, targeted states* scored in the upper quadrant and ranged between 40 and 50.





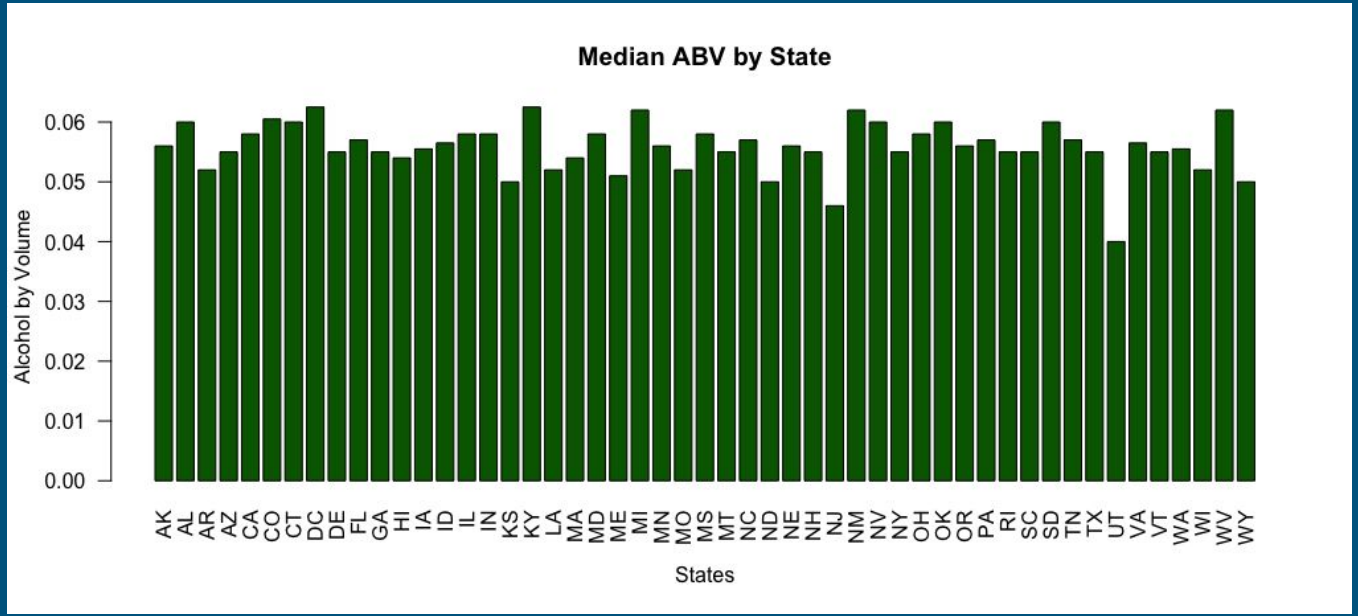

```
median_ABV <- aggregate(ABV~State,
beers2, FUN = median)
barplot(median_ABV$ABV,
        space = .25,
        xlab = "States",
        ylab = "Alcohol by Volume",
        main = "Median ABV by State",
        names.arg = median_ABV$State,
        las=2, col = "darkgreen")
```

```
median_IBU <-
aggregate(IBU~State, beers2, FUN =
median)
barplot(median_IBU$IBU,
        space = .25,
        xlab = "States",
        ylab = "Bitterness",
        main = "Median IBU by State",
        names.arg =
median_IBU$State,
        las=2, col = "darkgreen")
```

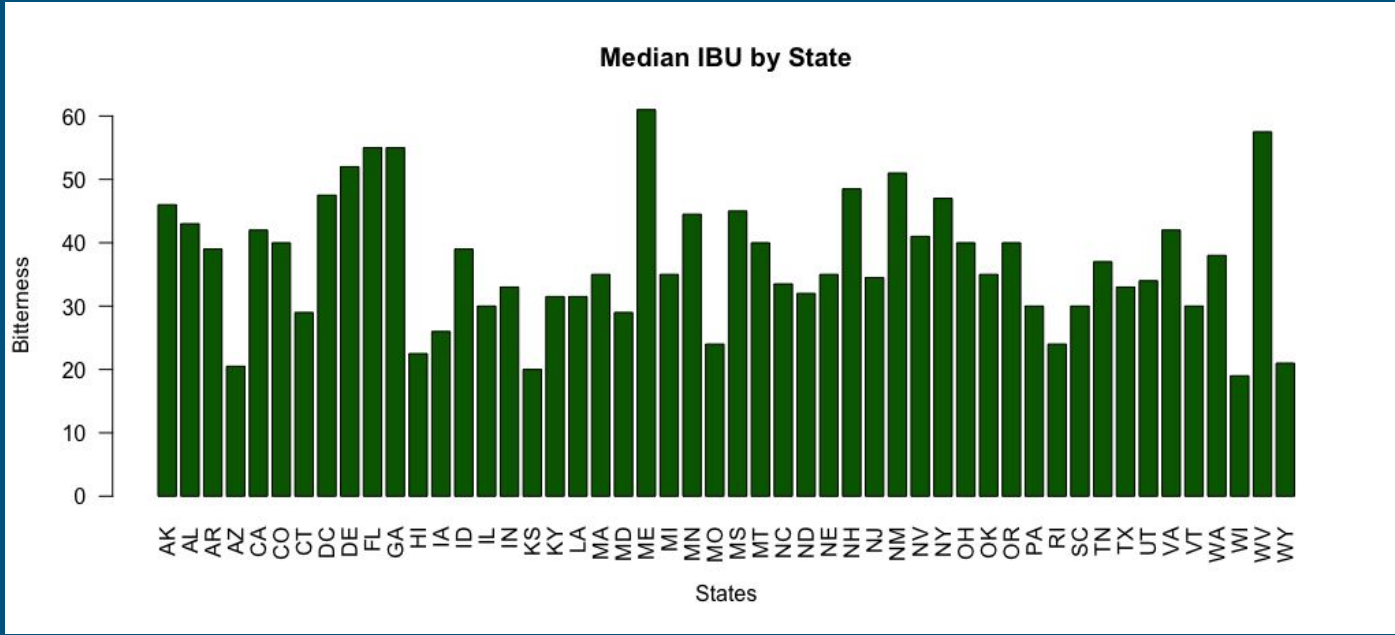
Using the **aggregate function**, median ABV and IBU data were computed for each state. Subsequently, dataset was plotted with the **barplot function**.

* Only exception was IL, which was identified to hold a median IBU of ~30

Median ABV by State



Median IBU by State



States with maximum ABV and IBU

The state with the maximum ABV and IBU score was Colorado and Oregon respectively. This should provide comfort to GB given the bitter taste of its best selling beer.

In reviewing the Five-Number Summary, the IQR for ABV is between 5% and 6.7%, with a median of 5.6%, and a mean of 5.98%.

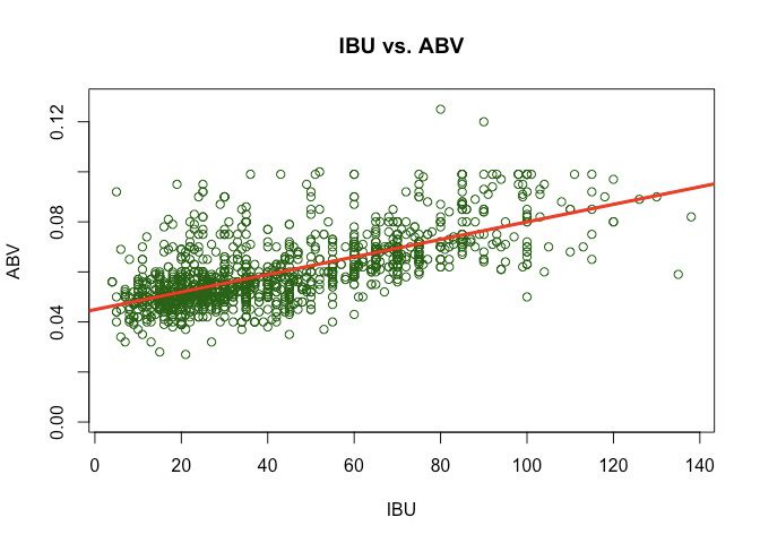
TQ Senkungu , TART

```
beers2$State[which.max(beers2$ABV)]
beers2$State[which.max(beers2$IBU)]
summary(beers2$ABV)
```

Index containing the max ABV was used to cross-reference the state in the original beer dataset. Further, **R's summary function was used to calculate the Min, Max, Median, and IQR results for ABV.**

Correlation between ABV and IBU

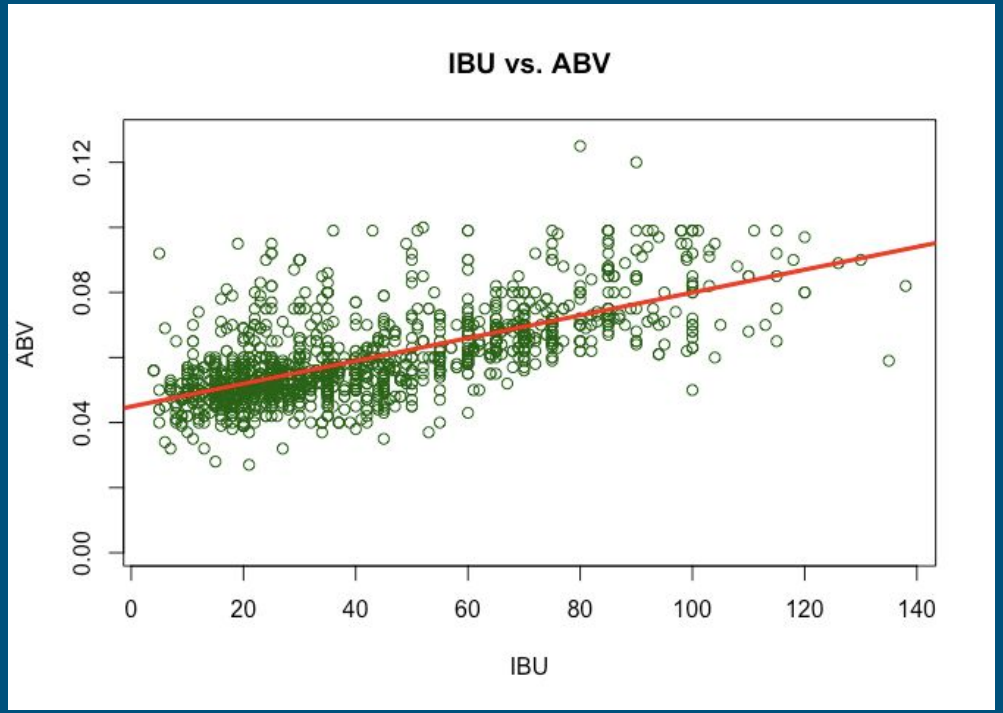
Although a preliminary view of the plot between ABV and IBU may be perceived as a cluster and correlation is difficult to discern, we believe that a large set of the data is positively correlated.




```
regline <-
lm(beers2$ABV~beers2$IBU,data=beers2)
        summary(regline)
with(beers2,plot(beers2$IBU,
beers2$ABV,xlab="IBU",ylab="ABV",main=
"IBU vs. ABV",col="green"))
    abline(regline,col="red",lwd=3)
```

Used the plot function in R to create a scatter plot of ABV and IBU values. Performing a linear regression, the r squared value was calculated at 0.4493 - a relatively weak correlation.

ABV vs IBU



Recommendations

Based on the our preliminary analysis, we believe that GB will benefit from the regional expansion as taste and alcohol content found in target areas aligns to GB's existing product. Initial beer launch should include beers with IBU rating between 15-30 and alcohol content from 4-6%. This is the sweet spot for production and maximize your expertise.