United States Breweries Opened By Year

Ryan Shaver

Abstract

Dataset:
Alcohol and Tobacco, Tax and Trade Bureau from the U.S. Department of the Treasury Dataset of Opened Breweries from 1984 - December 31, 2018

Question:
Has there been a national trend in brewery openings since 1984? If so, what is it?

Method Used:
Line graph, representing breweries opened against years using Plot.ly tracers

Findings:
There has been a positive trend in brewery opening, with notable increases nation wide around 1995 and 2010.

Motivation

While sitting around with family and friends, enjoying a craft beer, speculation may occur. This speculation which I've personally experienced, is that the brewery industry in the United States has been expanding. There is never any data to supplement these speculations, and often inferences are supplied in the form of, "Look at how many craft beer options there are now in the grocery store compared to when we first started drinking." There can be many reasons why stores don't stock all the available beers, and assuredly they still do not stalk all the beer which is currently available, so this isn't a reliable metric.

But this did lend itself to be the motivating catalyst for this project, "Is there data to support there are more breweries opened now then there were back when we started drinking?"

Dataset(s)

The data set utilized for this project was acquired through Data.gov. Clicking through the link, you will see a data set which accounts for the breweries opened from 1984 to March 31, 2018. But by doing a bit more digging, I was able to locate a more up to date data set from the Alcohol and Tobaco, Tax and Trade Bureau from the U.S. Department of the Treasury

For those interested in exploring the data set themselves, the link is supplied on the TTB website. They provide an Excel format which was not readily formatted for analysis. In the project's Github repository, I will provide the file of my cleaned and transposed CSV.

Data Preparation and Cleaning

As I’m sure many students have found, sometimes it is quite difficult to find reliable data or any data at all on the subject for which you wish to analyze. The data retrieval process, thus, was quite arduous as it took quite a long time to find the most relevant dataset.

Unfortunately, the dataset has many cells empty which needed to be replaced with zeros and probably the largest change I had to make to the dataset to make it usable was a complete transpose of the entire matrix. I was surprised that the government had the dataset in the opposite orientation, as it would be more cumbersome to analyze by year, which I would assume would be the aim of an annual dataset.

Research Question

The research question aimed at being answered by this project is whether there has been a national trend here in the United States in brewery openings since 1984? If so, what is that trend?

Methods

The methods used to analyze the data were two fold:

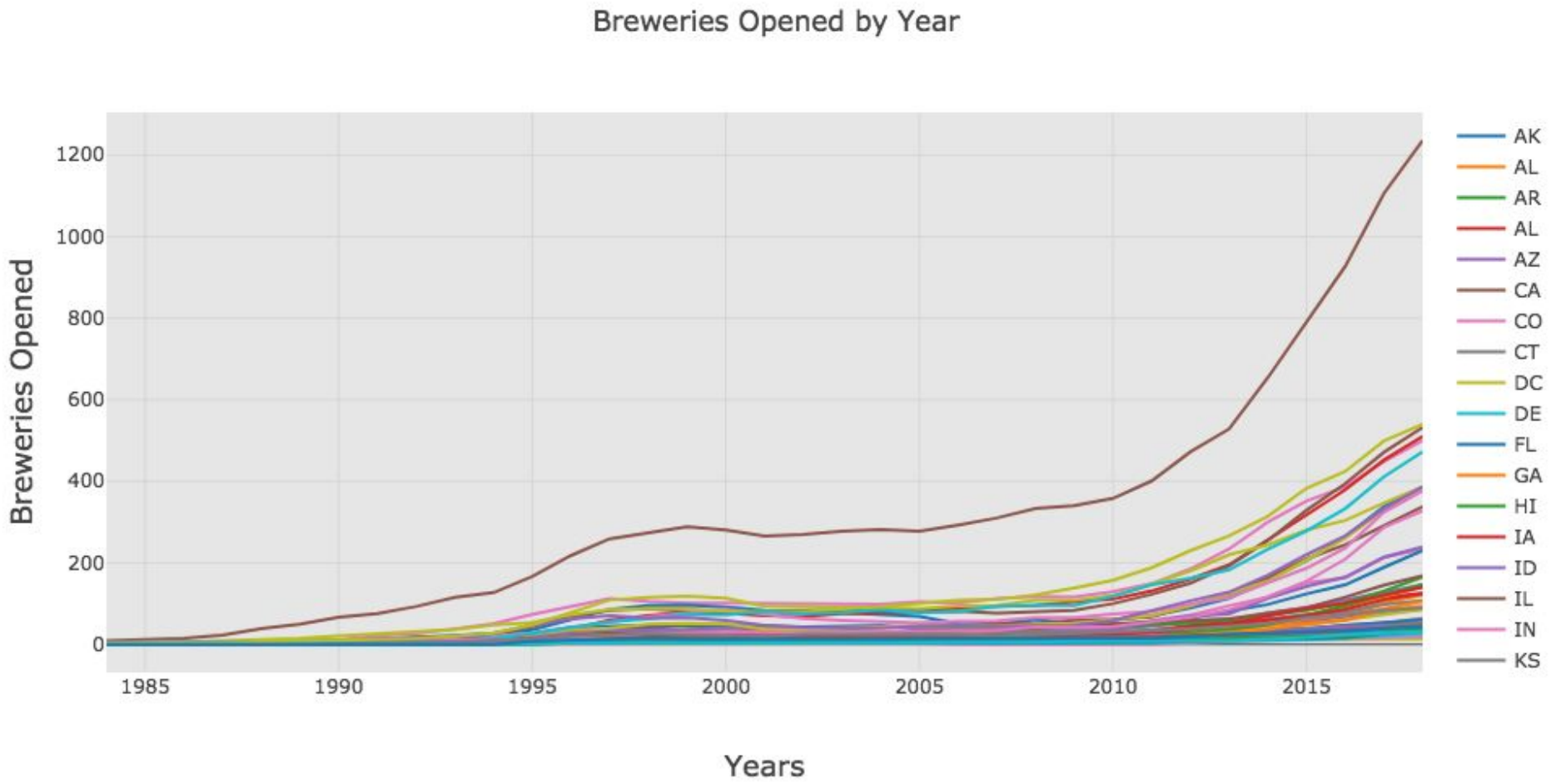
First I used a line graph to represent the breweries opened per year for each state. This visual representation allows us to easily see any correlations over time, as we can see in 1995 and 2010.

The second method used was the Plot.ly tracers, which I was able to write a personalized function for, allowing us to take each year and create the aforementioned line graph for each state.

Findings

As you can observe on the following page and on the projects notebook, there was an overall national trend in the United States in the opening of breweries since 1984.

That trend is overwhelming positive, positive in the direction of the slope. Not only can we identify that there have been more and more breweries opening over time, we were able to identify that in 1995 and 2010 there seemed to be something which percipitated an increase in brewery openings.


Breweries Opened by Year
Breweries Opened
1200
1000
800
600
400
200
0
1985
1990
1995
2000
2005
2010
2015
Years
AK
AL
AR
AL
AZ
CA
CO
CT
DC
DE
FL
GA
HI
IA
ID
IL
IN
KS

Limitations

One of the limitations of this dataset and conclusion, is that it only indicates overall breweries opened at the time of investigation, and not the longevity of individual breweries.

The longevity of individual breweries in important, because it could indicate more evidence about the popularity of breweries. For example, if the rate at which breweries opened and then closed was also high, it would tell us more about the state of breweries in the United States.

Unfortunately, our dataset does not break down the individual brewery longevity.

Conclusions

As can be observed from the above plot, there seemed to be a slight bump after 1995 and then a rather dramatic increase starting around 2009 to 2010. This information in itself is quite useful as it indicates a nation wide trend in brewery openings.

This information can lead economists and industry brewers to investigate what may have been the proximal causes for these two industry increases.

Conclusions continued…

Further information I found worth noting are the states with the most breweries as of 2018. The usual suspects of California, Washington, and Colorado were among the top five, but I was surprised to see New York and Michigan in the top five as well. These top five were followed by Pennsylvania, Texas, North Carolina, Oregon, and Florida. Again where California, Colorado, Oregon, and Texas are well known in the beer community as having outstanding breweries, it was enlightining to see Pennslvania, North Carolina, Florida, and Michigan in the mix.

For those planning brewery excurssions, this plot can certainly steer brewery goers in novel and unexcpected directions.

Acknowledgements

I’d like to thank my friend Sean for helping optomize my code. As I am fairly new to coding, I was unsure of how to make the tracer function work well. I had given up at some point, and just wrote out all fifty-one tracers by hand, which I knew was not efficient.

Sean took my unsuccesful tracer function and showed me the points I could change to get it to work. I owe his a great deal of thanks and acknowledgment, as it not only optomized the code, but left the notebook looking more professional and elegant.

References

Plot.ly

- https://plot.ly/python/ipython-notebook-tutorial/
- https://plot.ly/python/choropleth-maps/
- https://www.youtube.com/watch?v=hA39KSTb3dY
- https://plot.ly/python/#animations
- https://plot.ly/python/gapminder-example/
- https://plot.ly/python/bubble-maps/

Folium / Leaflet

- https://www.youtube.com/watch?v=4RnU5qKTfYY
- https://www.youtube.com/watch?v=xN2N-p33V1k

References Continued...

Widgets

- https://www.youtube.com/watch?v=1ndo6C1KWjI
- https://www.youtube.com/watch?v=i40d8-Hu4vM
- ipywidegs (core UI controls / sliders)
- bqplot (2d plotting)
- pythreejs, ipyvolume (3d plotting)
- ipyleaflet (maps)

CSV Analysis / ML

- https://www.youtube.com/watch?v=-0NwrcZOKhQ
- https://www.youtube.com/watch?v=Q73ADVZCqSU
- https://www.youtube.com/watch?v=OBPjFnyxoCc
- https://www.youtube.com/watch?v=zJ4RK6jtYCU&list=PLbD3QT5__Llz88nB-B-Kp5s118DOkAHr1
-

References Continued...

Resources

- https://plot.ly/python/choropleth-maps/
- https://www.kaggle.com/rdoume/beerreviews
- https://www.kaggle.com/ehallmar/beers-breweries-and-beer-reviews
- http://beer.tany.kim/
- https://untappd.com/api/docs
- https://www.reddit.com/r/Untappd/comments/41i45t/mass_data_export/
- https://www.kaggle.com/nickhould/craft-cans
- https://data.world/datafiniti/breweries-brew-pubs-in-the-usa
- https://www.reddit.com/r/datasets/comments/6i0v3g/craft_beer_dataset/
- https://github.com/nickhould/craft-beers-dataset
- https://catalog.data.gov/dataset?tags=beer
- https://catalog.data.gov/dataset/yearly-statistical-beer-data-by-state-2007-2016
- https://catalog.data.gov/dataset/beer-production-and-operations-reports
- https://catalog.data.gov/dataset/brewery-count-by-state-1984-march-31-2017
- https://catalog.data.gov/dataset/brewery-count-by-state-1984-march-31-2018
- https://www.ttb.gov/foia/frl.shtml