

Final Project - Breweries Opened By Year (UCSD DSE 200x)

February 23, 2019

1 United States Breweries Opened By Year

Final Project for UCSD DSE200x Python for Data Science Ryan Shaver

1.1 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.

1.1.1 The Data Set

The data set utilized for this project was aquired 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 Tobacco, 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 web-site](#). 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.

```
In [8]: # Import
import pandas as pd
import numpy as np
import plotly.plotly as py
import plotly.graph_objs as go
```

```
In [9]: # Read in and display the DataFrame
df = pd.read_csv("brewery_count_transposed.csv", header=0)
df
```

```
Out[9]:
```

	Year	AK	AL	AR	AZ	CA	CO	CT	DC	DE	FL	GA	HI	IA	ID	\
0	1984	0	0	0	0	9	0	0	0	0	0	0	0	0	0	
1	1985	0	0	0	0	12	0	0	0	0	0	0	0	0	0	
2	1986	0	0	0	0	15	0	0	0	0	0	0	0	0	0	

3	1987	0	0	0	0	23	0	0	0	0	0	0	0	0	3
4	1988	0	0	0	0	39	3	0	0	0	3	0	0	0	3
5	1989	0	0	0	6	50	7	0	0	0	6	0	0	0	3
6	1990	0	0	0	8	67	11	0	0	0	7	0	0	0	3
7	1991	3	0	0	8	76	22	0	0	0	13	0	0	4	5
8	1992	4	0	0	7	93	27	0	0	0	17	0	0	5	8
9	1993	3	0	0	7	115	38	0	0	0	22	0	0	6	10
10	1994	3	0	3	11	127	51	0	0	0	27	3	0	6	11
11	1995	5	5	3	11	167	74	0	0	4	36	6	4	6	15
12	1996	11	8	3	17	218	92	12	3	5	62	18	6	9	24
13	1997	16	8	4	34	259	112	22	5	6	86	24	8	12	28
14	1998	14	9	5	40	273	105	25	5	8	96	26	13	13	29
15	1999	19	9	4	48	289	102	26	4	9	97	30	16	14	25
16	2000	17	7	5	45	281	102	27	4	8	92	29	15	13	24
17	2001	17	4	5	45	266	102	22	5	8	84	30	14	16	20
18	2002	19	3	4	39	270	102	17	5	8	83	25	13	15	19
19	2003	15	3	4	36	278	99	17	4	9	78	29	13	17	20
20	2004	13	5	4	31	282	99	20	5	9	74	29	13	21	19
21	2005	15	5	4	31	278	104	19	4	8	69	25	12	18	20
22	2006	14	4	4	33	293	102	18	4	9	51	22	12	19	20
23	2007	15	5	4	33	310	112	19	3	10	50	23	12	20	20
24	2008	16	5	4	33	333	117	21	3	9	61	20	11	21	24
25	2009	18	5	5	34	340	116	19	3	11	52	22	11	28	25
26	2010	20	7	5	37	358	129	21	4	11	60	26	11	27	25
27	2011	23	9	7	45	401	149	20	6	12	71	25	12	35	33
28	2012	25	15	13	55	472	185	24	5	13	89	31	15	46	39
29	2013	25	23	18	59	528	234	39	11	13	113	37	17	54	42
30	2014	28	25	23	77	654	300	50	12	15	158	48	20	60	51
31	2015	35	30	29	91	788	352	59	13	21	205	54	26	71	57
32	2016	36	37	34	110	927	386	76	13	25	264	69	23	94	67
33	2017	45	52	44	130	1,106	448	103	13	33	338	102	28	115	76
34	2018	51	55	53	146	1,236	500	124	13	38	386	121	38	125	87

	IL	IN	KS	KY	LA	MA	MD	ME	MI	MN	MO	MS	MT	NC	ND	NE	\
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
2	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	
3	0	0	0	0	0	3	0	0	0	4	0	0	0	0	0	0	
4	4	0	0	0	0	3	0	0	3	5	0	0	0	0	0	0	
5	5	4	0	0	0	5	3	0	3	7	0	0	0	6	0	0	
6	6	4	0	0	0	5	3	0	3	7	0	0	3	8	0	0	
7	9	5	0	0	3	6	4	3	3	8	3	0	5	11	0	0	
8	12	6	0	0	3	9	5	5	4	8	4	0	7	12	0	5	
9	16	7	5	3	4	9	7	8	6	9	5	0	10	14	0	6	
10	22	9	7	3	3	17	10	12	12	13	12	0	12	20	0	6	
11	26	10	13	7	4	27	12	21	24	18	13	0	13	25	3	8	
12	35	21	14	11	8	42	22	27	36	26	14	0	16	36	3	8	
13	39	26	17	14	11	44	27	28	60	37	22	0	20	40	4	9	

14	44	25	17	15	14	48	27	33	71	35	26	0	24	50	6	8
15	47	24	15	12	11	50	30	37	79	36	30	3	24	52	3	9
16	45	26	10	15	12	45	25	42	77	41	32	4	25	51	3	9
17	41	23	8	10	9	46	26	39	71	37	26	4	26	38	3	11
18	43	23	8	10	7	43	26	38	69	37	28	4	28	38	3	10
19	43	23	7	10	6	44	27	37	76	31	28	3	24	41	3	12
20	45	23	10	10	9	46	29	43	78	30	28	3	22	43	4	12
21	43	27	11	10	10	38	25	43	84	31	33	3	21	45	4	16
22	45	25	17	10	10	40	26	43	91	35	34	0	22	46	3	16
23	52	30	18	12	10	43	24	43	94	34	38	0	26	45	3	17
24	51	30	17	12	10	45	24	43	96	36	38	0	31	50	3	16
25	60	39	17	15	11	44	26	42	104	39	47	0	31	52	3	17
26	62	43	17	14	12	48	24	44	111	42	51	3	31	61	3	16
27	68	59	21	21	12	52	26	49	131	56	55	0	35	71	3	18
28	95	72	21	25	12	65	40	53	160	70	60	4	41	94	7	20
29	123	95	25	25	15	82	43	60	195	75	63	7	49	125	8	25
30	164	116	27	38	17	98	55	71	256	113	77	10	62	155	10	35
31	210	151	37	48	25	124	73	84	316	142	90	14	74	207	11	39
32	244	163	47	60	34	146	88	102	379	165	116	14	79	260	15	47
33	291	213	53	73	43	189	116	131	452	214	145	16	98	330	22	53
34	338	234	64	86	47	230	141	165	510	239	168	19	108	387	26	60

	NH	NJ	NM	NV	NY	OH	OK	OR	PA	RI	SC	SD	TN	TX	UT	VA \
0	0	0	0	0	0	0	0	0	4	0	0	0	0	4	0	0
1	0	0	0	0	0	0	0	3	4	0	0	0	0	4	0	0
2	0	0	0	0	3	0	0	5	5	0	0	0	0	4	0	0
3	0	0	0	0	4	0	0	8	6	0	0	0	0	4	0	0
4	0	0	0	0	5	3	0	13	6	0	0	0	0	5	0	0
5	0	0	0	0	6	5	0	15	8	0	0	0	0	7	0	0
6	0	0	0	0	8	6	0	21	11	0	0	0	0	8	0	4
7	0	0	3	0	13	9	0	26	12	0	0	0	0	7	3	5
8	3	0	5	0	16	12	0	30	12	0	0	0	3	8	4	7
9	4	0	8	3	20	15	4	37	12	0	0	0	5	11	4	12
10	9	5	10	3	27	20	6	48	15	3	0	0	9	27	8	9
11	14	7	12	5	44	28	8	54	27	6	11	0	13	48	13	12
12	18	20	20	10	70	40	9	73	42	5	19	5	15	63	15	17
13	19	23	26	18	84	55	9	86	56	6	21	5	21	71	17	23
14	19	27	27	20	88	68	10	88	63	5	23	5	25	65	17	27
15	19	28	26	24	81	77	10	88	74	6	22	7	23	66	20	29
16	19	24	26	22	78	87	9	82	73	8	19	7	23	59	20	31
17	17	22	23	16	83	74	6	82	81	7	19	6	23	47	16	29
18	17	20	23	17	84	64	6	81	76	7	19	7	22	42	17	30
19	16	22	20	16	86	59	6	87	78	6	21	7	22	41	15	30
20	16	21	22	17	80	55	7	92	86	6	21	6	24	40	16	31
21	14	21	22	16	82	53	8	88	78	5	19	6	19	45	15	34
22	14	21	23	16	82	57	7	95	81	5	16	7	19	47	17	35
23	15	21	21	16	77	58	6	96	94	5	17	8	22	45	15	35
24	16	20	20	17	81	66	9	108	97	5	16	7	23	47	15	43

25	17	22	21	18	84	66	10	108	96	5	15	7	23	47	18	42
26	19	24	26	20	100	75	10	120	118	5	16	8	26	59	18	44
27	21	31	30	20	123	81	10	148	146	6	17	9	31	84	18	54
28	25	29	38	23	150	99	13	180	162	9	23	12	45	107	22	67
29	35	37	48	26	193	114	17	220	183	12	30	12	51	128	24	85
30	46	49	60	34	255	151	18	244	233	14	37	15	69	170	27	117
31	63	71	71	39	329	187	21	281	278	15	51	19	88	220	29	155
32	73	96	86	44	394	236	26	304	333	17	59	21	101	266	34	209
33	88	123	110	46	471	324	43	347	411	27	84	28	120	333	39	287
34	106	146	124	50	532	377	55	386	472	33	94	35	140	387	46	328

	VT	WA	WI	WV	WY
0	0	4	5	0	0
1	0	4	6	0	0
2	0	4	7	0	0
3	0	4	10	0	0
4	0	7	12	0	0
5	3	8	14	0	0
6	4	10	17	0	0
7	6	13	18	0	0
8	6	15	19	0	0
9	7	21	23	0	0
10	10	28	29	3	0
11	12	48	35	0	8
12	20	79	50	4	10
13	24	109	57	4	14
14	27	115	65	5	15
15	24	118	77	5	13
16	23	113	91	6	13
17	19	94	82	4	11
18	25	94	90	5	11
19	25	91	90	5	11
20	26	93	97	4	13
21	24	100	98	4	14
22	24	108	103	6	14
23	22	110	105	6	14
24	24	121	110	6	14
25	27	138	108	8	13
26	32	157	119	8	16
27	34	188	126	8	15
28	39	230	146	8	19
29	45	266	146	9	25
30	51	314	168	13	28
31	66	383	189	17	32
32	73	424	217	24	33
33	84	499	261	27	39
34	90	540	303	32	44

```

In [10]: # Construct Tracers for Plot.ly to Display
states = ['AK', 'AL', 'AR', 'AZ', 'CA', 'CO', 'CT', 'DC', 'DE', 'FL', 'GA', 'HI', 'IA', 'ID']

def createTraces(states):
    traces = []
    for abbreviation in states:
        traces.append(go.Scatter(x=df['Year'], y=df[abbreviation], mode='lines', name=
    return traces

layout = go.Layout(title='Breweries Opened by Year',
                    plot_bgcolor='rgb(230, 230,230)',
                    xaxis=dict(
                        title='Years',
                        titlefont=dict(
                            #family='Courier New, monospace',
                            size=18,
                            #color='#7f7f7f'
                        )
                    ),
                    yaxis=dict(
                        title='Breweries Opened',
                        titlefont=dict(
                            #family='Courier New, monospace',
                            size=18,
                            #color='#000000'
                        )))

data = createTraces(states)
fig = go.Figure(data, layout=layout)

# Plot data in the notebook
py.iplot(fig, filename='simple-plot-from-csv')

```

```
Out[10]: <plotly.tools.PlotlyDisplay object>
```

1.2 Conclusion

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.

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 enlightening to see Pennsylvania, North Carolina, Florida, and Michigan in the mix.

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

1.3 References

1.3.1 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/>

1.3.2 Folium / Leaflet

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

1.3.3 Widgets

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

1.3.4 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

1.3.5 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/nickould/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/fri.shtml>