IBM Data Science Professional Certificate

Capstone Project

Researching Crime Rates and Coffee Shops in Orange County, California, United States of America

By: Kelly Anderson – April 2019

Description of the data – Collection Methods and Datasets:

The city crime rates and population data can be found on a Wikipedia page which shows a grid. This data was cited as taken from 2013, so it should be relevant given the timeframe.

https://en.wikipedia.org/wiki/Orange_County,_California

Notes: United States Department of Justice, Federal Bureau of Investigation. Crime in the United States, 2012, Table 8 (California). Retrieved November 14, 2013.
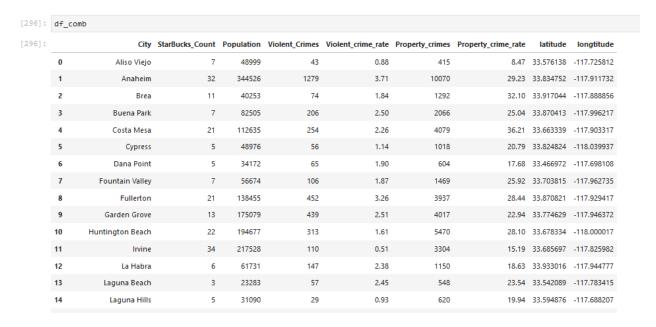
DataGrid:

Cities by population and crime rates   [edit]

Cities by population and crime rates					[hide]
City	Population[91]	Violent crimes[91]	Violent crime rate per 1,000 persons	Property crimes[91]	Property crime rate per 1,000 persons
Aliso Viejo	48,999	43	0.88	415	8.47
Anaheim	344,526	1,279	3.71	10,070	29.23
Brea	40,253	74	1.84	1,292	32.10
Buena Park	82,505	206	2.50	2,066	25.04
Costa Mesa	112,635	254	2.26	4,079	36.21
Cypress	48,976	56	1.14	1,018	20.79
Dana Point	34,172	65	1.90	604	17.68
Fountain Valley	56,674	106	1.87	1,469	25.92
Fullerton	138,455	452	3.26	3,937	28.44
Garden Grove	175,079	439	2.51	4,017	22.94
Huntington Beach	194,677	313	1.61	5,470	28.10
Irvine	217,528	110	0.51	3,304	15.19
Laguna Beach	23,283	57	2.45	548	23.54
Laguna Hills	31,090	29	0.93	620	19.94
Laguna Niguel	64,533	47	0.73	764	11.84
Laguna Woods	16,590	4	0.24	148	8.92
La Habra	61,731	147	2.38	1,150	18.63

The numbers of Starbucks and Sushi restaurants will be collected using FourSquare API calls. This will be done by looping through all cities and compiling the unique items that meet the requirements.

Example DataFrame from results from Foursquare API:

```
[269]: dfc.head()
```

```
[269]:
                            id       location.address  location.city      name
    0  4b4e6d97f964a5206bed26e3     27020 Alicia Pkwy   Laguna Niguel  Starbucks
    1  4ac8ea1cf964a52045bd20e3  23411 Aliso Viejo Pkwy   Aliso Viejo  Starbucks
    2  4bca132a511f9521d5f5aec7       27072 La Paz Rd     Aliso Viejo  Starbucks
    3  4b3bc69df964a520e97a25e3  28391 Marguerite Pkwy  Mission Viejo  Starbucks
    4  4baa360af964a5209e533ae3     25630 Alicia Pkwy    Laguna Hills  Starbucks
```

```
[150]: dfc.to_csv('Starbucks_out.csv')
```

The datasets can then be merged to create the datapoints needed for the mapping, graphing, and correlation models.

```
[296]: df_comb
```

	City	StarBucks_Count	Population	Violent_Crimes	Violent_crime_rate	Property_crimes	Property_crime_rate	latitude	longtitude
0	Aliso Viejo	7	48999	43	0.88	415	8.47	33.576138	-117.725812
1	Anaheim	32	344526	1279	3.71	10070	29.23	33.834752	-117.911732
2	Brea	11	40253	74	1.84	1292	32.10	33.917044	-117.888856
3	Buena Park	7	82505	206	2.50	2066	25.04	33.870413	-117.996217
4	Costa Mesa	21	112635	254	2.26	4079	36.21	33.663339	-117.903317
5	Cypress	5	48976	56	1.14	1018	20.79	33.824824	-118.039937
6	Dana Point	5	34172	65	1.90	604	17.68	33.466972	-117.698108
7	Fountain Valley	7	56674	106	1.87	1469	25.92	33.703815	-117.962735
8	Fullerton	21	138455	452	3.26	3937	28.44	33.870821	-117.929417
9	Garden Grove	13	175079	439	2.51	4017	22.94	33.774629	-117.946372
10	Huntington Beach	22	194677	313	1.61	5470	28.10	33.678334	-118.000017
11	Irvine	34	217528	110	0.51	3304	15.19	33.685697	-117.825982
12	La Habra	6	61731	147	2.38	1150	18.63	33.933016	-117.944777
13	Laguna Beach	3	23283	57	2.45	548	23.54	33.542089	-117.783415
14	Laguna Hills	5	31090	29	0.93	620	19.94	33.594876	-117.688207

These datasets should provide a sufficient amount of records to try to draw conclusions as to relationships, as well as classify the cities to find common groups.