San Francisco Business Locations

https://github.com/aravindnatarajan/SFAlpine

https://data.sfgov.org/Economy-and-Community/
Open-Business-Locations-San-Francisco/g8m3-pdis

dbaName	city	state	zip	bizStartDate	locStartDate	lat	lon
3101 LAGUNA APTS	San Francisco	CA	94123	1993-09-30	1993-09-30	43.1613	-77.6193
IDEAL NOVAK CORP	San Francisco	CA	94116	1968-10-01	1968-10-01	37.7998	-122.431
TOURNAHU ARMS	San Francisco	CA	94123	1968-10-01	1968-10-01	37.7489	-122.465
3301 BRODERICK APARTMENTS	San Francisco	CA	94123	1968-10-01	1988-05-01	37.8048	-122.443
3301 BRODERICK APARTMENTS	San Francisco	CA	94123	1968-10-01	1988-05-01	37.8009	-122.445
1840-42 JEFFERSON ST	San Francisco	CA	94109	1968-10-01	1991-10-07	37.8009	-122.445
1601 GRANT APTS	San Francisco	CA	94133	1968-10-01	1968-10-01	37.7915	-122.42
1601 GRANT PARKING	San Francisco	CA	94133	1968-10-01	2013-01-01	37.8017	-122.408
ABBETT ELEC CORP	San Francisco	CA	94110	1968-10-01	1968-10-01	37.8017	-122.408
ABBEY CARPET	San Francisco	CA	94118	1968-10-01	2012-07-27	37.7637	-122.41
ABBEY CARPET OF SAN FRANCISCO	San Francisco	CA	94118	1968-10-01	2013-10-01	37.7822	-122.45
ABC AUTO PARTS	San Francisco	CA	94124	1968-10-01	1968-10-01	37.7822	-122.45
ABC INSURANCE	San Francisco	CA	94124	1968-10-01	2002-10-24	37.7462	-122.392
J & E CONSTRUCTION	San Francisco	CA	94124	1968-10-01	1999-08-27	37.7361	-122.393
ABC WINDOW CONTRACTORS INC	San Francisco	CA	94124	1968-10-01	1997-09-29	37.7361	-122.393
ABLE BUILDING MAINTENANCE CO	San Francisco	CA	94107	1968-10-01	1968-10-01	37.7503	-122.386

http://www.city-data.com/zipmaps/San-Francisco-California.html#top

zip	population	col	popDens	house	income	unemp	area
94102	25881	176.9	38579	574100	22517	9.3	0.7
94103	23907	155.5	17630	628000	43364	9.8	1.4
94104	211	166.3	2719	1000001	48750	10.3	0.1
94105	5553	140.6	15033	740900	163949	2.8	0.4
94107	27856	144.1	15546	719700	114439	7.5	1.8
94108	13709	175.1	50764	856900	35427	9.3	0.3
94109	55519	152.5	46737	723200	60722	6.7	1.2
94110	72128	147.8	31018	768200	82111	7.9	2.3
94111	3496	151.1	10160	882000	89722	12.3	0.3
94112	82726	145.6	24592	602400	71625	10.9	3.4
94114	31397	147.9	22028	1000001	114689	6	1.4
94115	33782	154.5	30278	965800	76952	6.2	1.1
94116	45157	147.5	17465	734400	83407	8.2	2.6
94117	41568	149.6	24633	911500	94479	6	1.7
94118	40292	153.8	20675	1000001	82857	7	1.9
94121	42104	152.5	13662	846400	72756	7.7	3.1
94122	57369	148.9	24269	774500	81768	7.1	2.4
94123	24500	148.5	24005	1000001	112650	7	1
94124	35309	145.3	7167	470400	50146	16.2	4.9

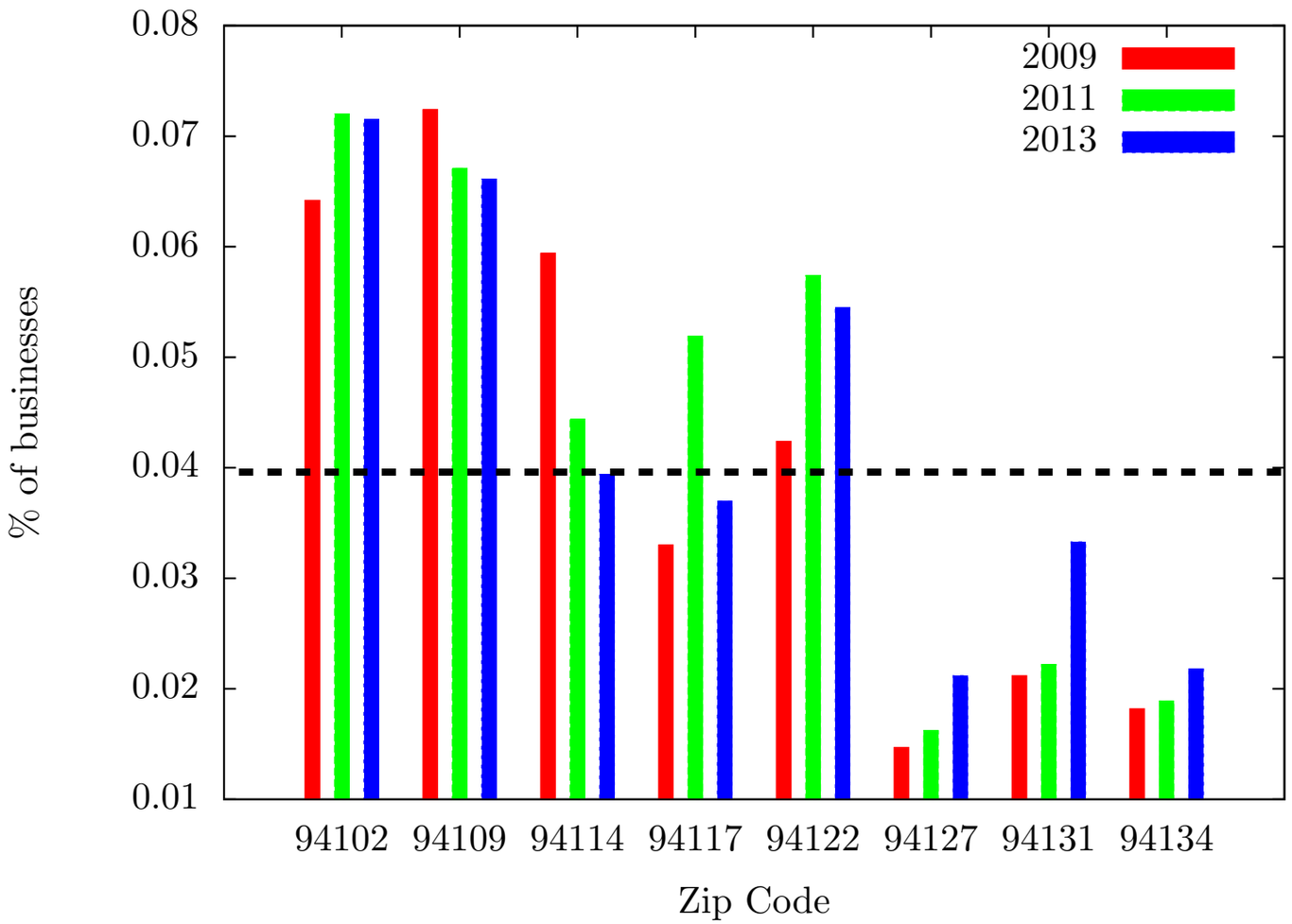
Troubles with the data:

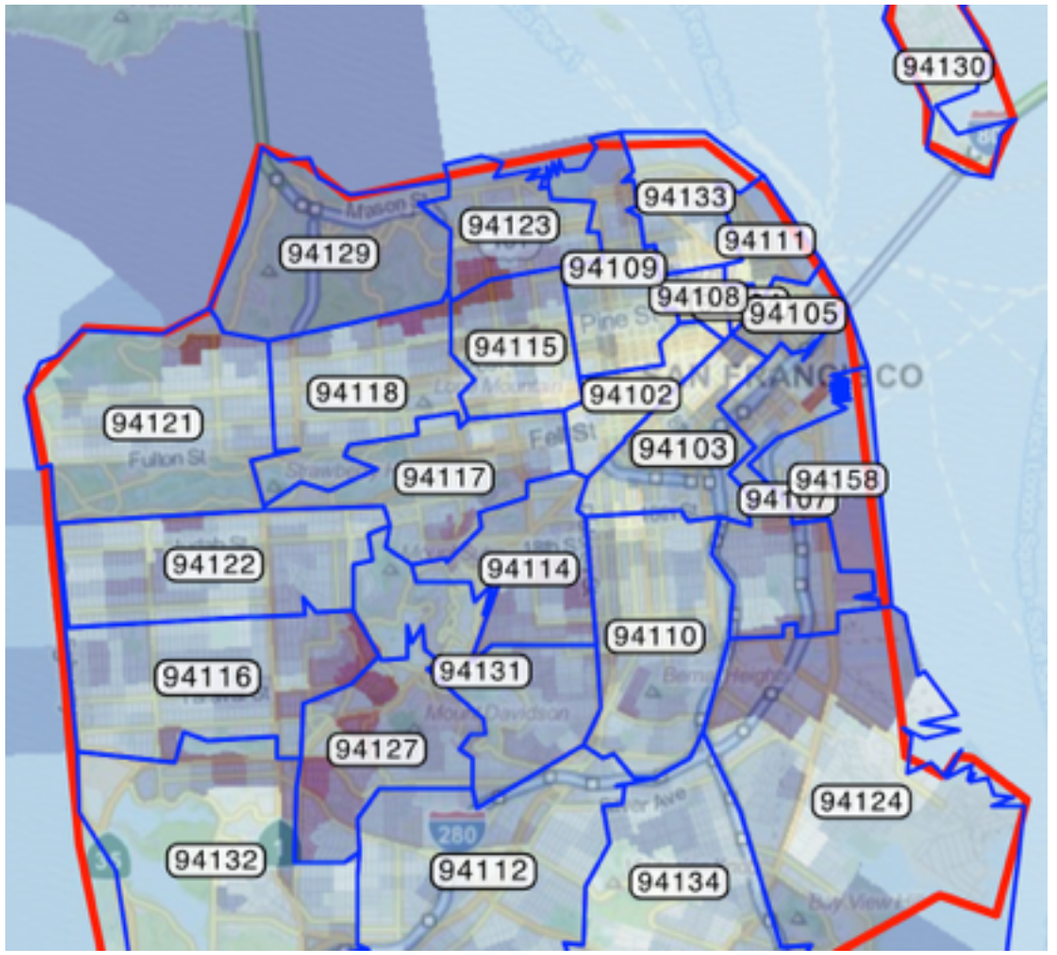
1. csv files shouldn't have commas in fields!
   e.g. "Mary, John, and Paul place"

2. Lat/Lon info difficult to separate.

Data put into a MySQL database.
Data retrieved through Python/MySQLdb
and analysed with Python/numpy, Python/matplotlib.

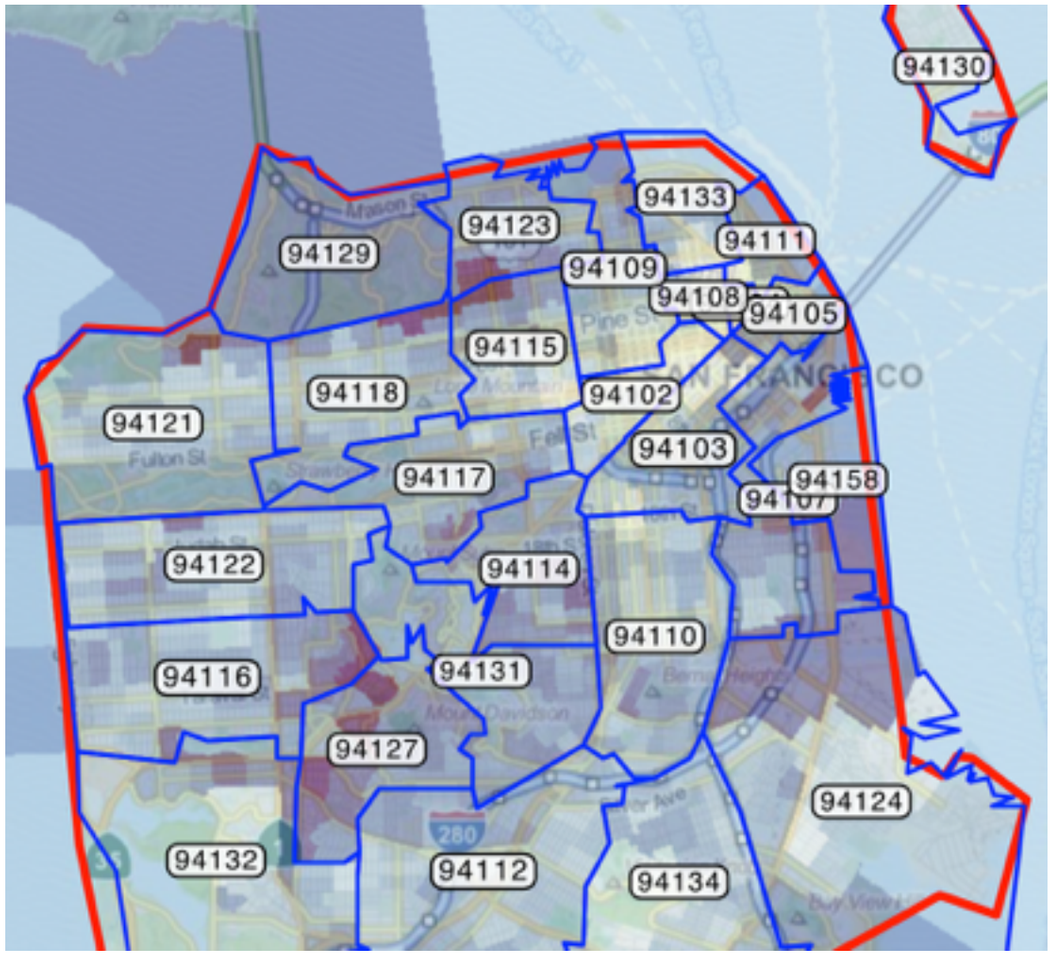
How many businesses are started each year?





Business Details are not known!

But we can make some assumptions:
e.g. businesses with the word
“Apartment” or “Apt”
are likely, apartments :)



Why are there so many new apartments in 2007?

```
select year(biz.locStartDate),
count(biz.dbaName) from biz
where (biz.dbaName like "%Apartment%"
or biz.dbaName like "%Apt%")
group by year(biz.locStartDate)
```

Query Favorites Query History

ar(biz.locStartDate)	count(biz.dbaName)
2003	74
2004	62
2005	81
2006	74
2007	1029
2008	59
2009	30
2010	33
2011	47
2012	30
2013	42

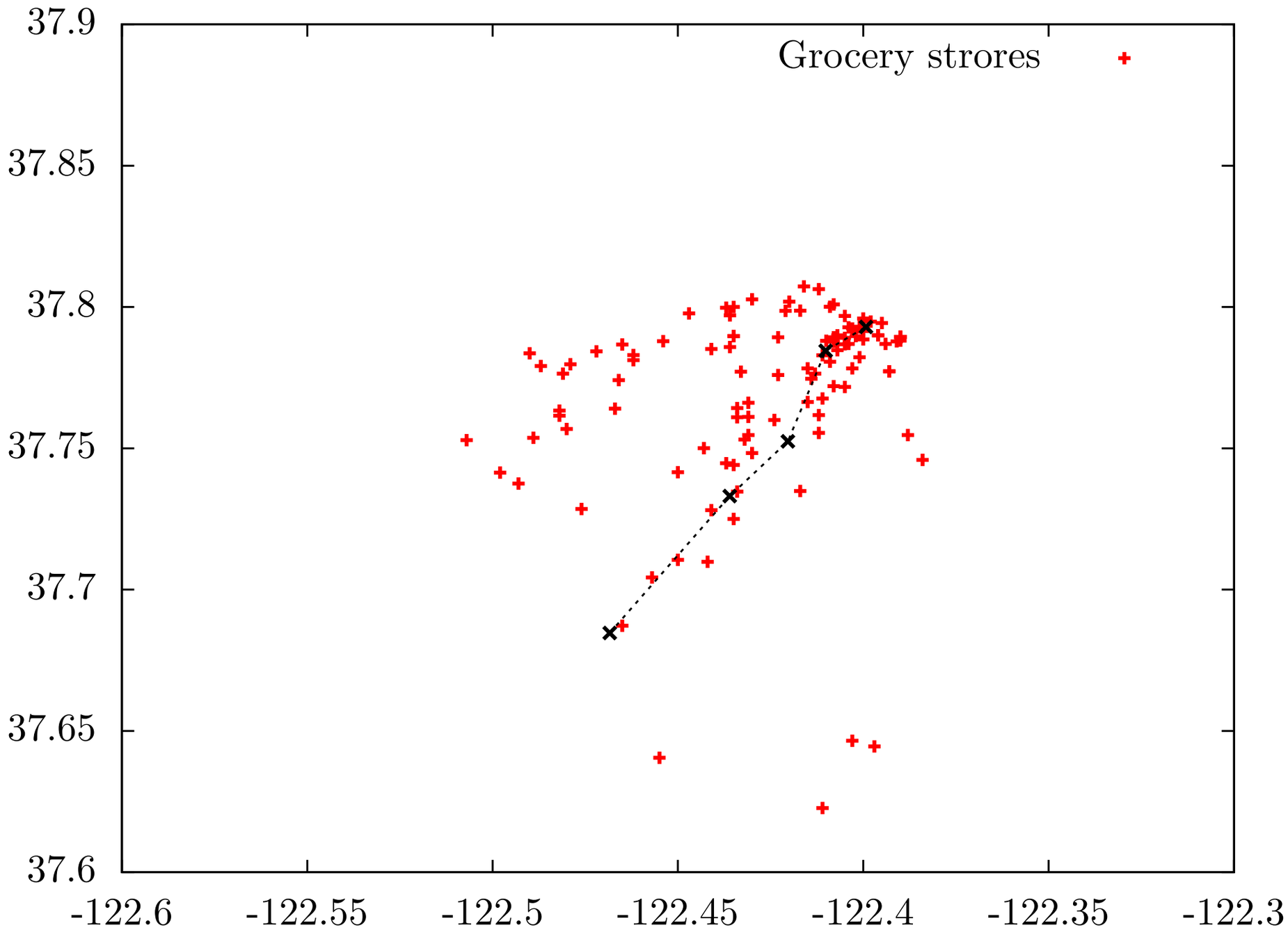
Why are there so many new apartments in 2007?

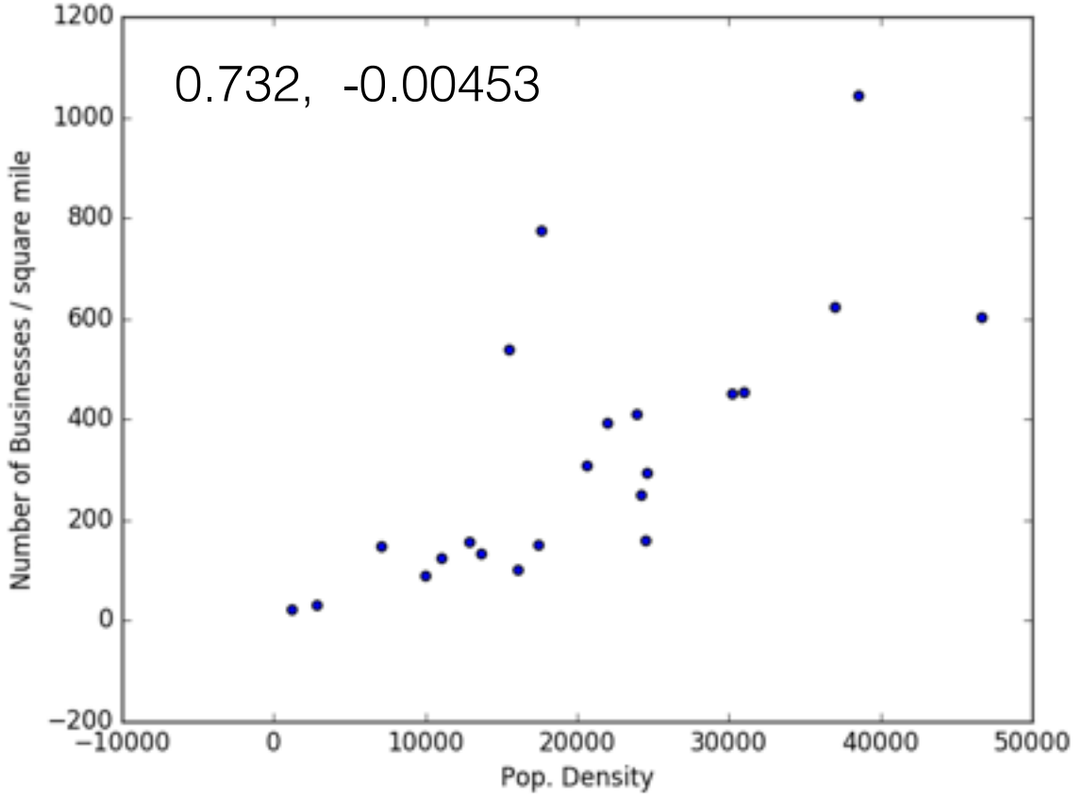
```
select month(biz.locStartDate),
count(biz.dbaName) from biz
where (biz.dbaName like "%Apartment%"
or biz.dbaName like "%Apt%")
and year(biz.locStartDate) = 2007
group by month(biz.locStartDate)
```

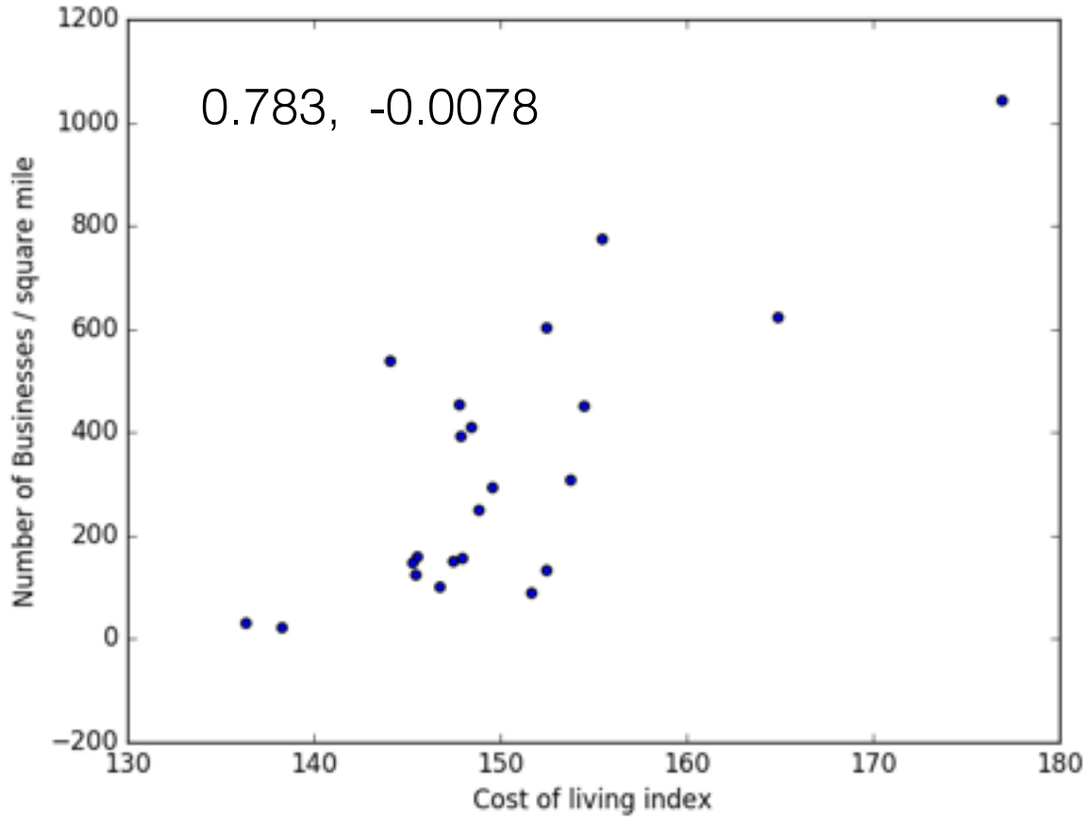
Query Favorites Query History

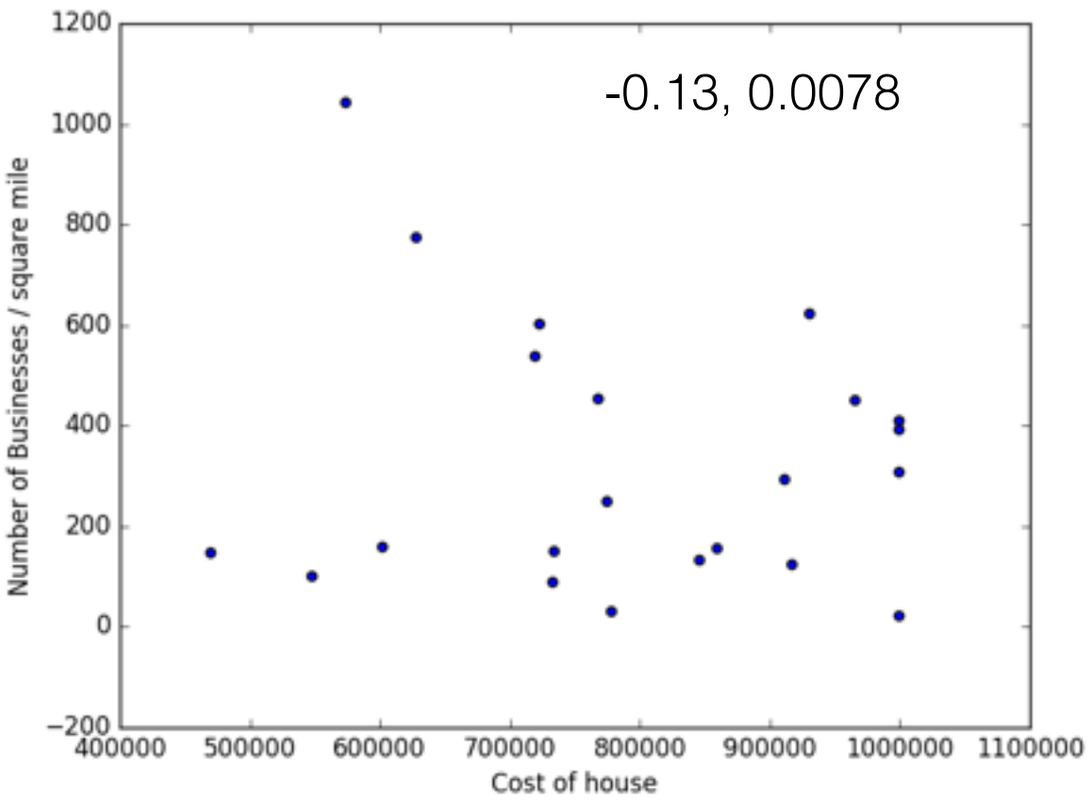
nth(biz.locStartDate)	count(biz.dbaName)
1	27
2	5
3	612
4	297
5	27
6	11
7	19
8	10
9	7
10	1
11	7
12	6

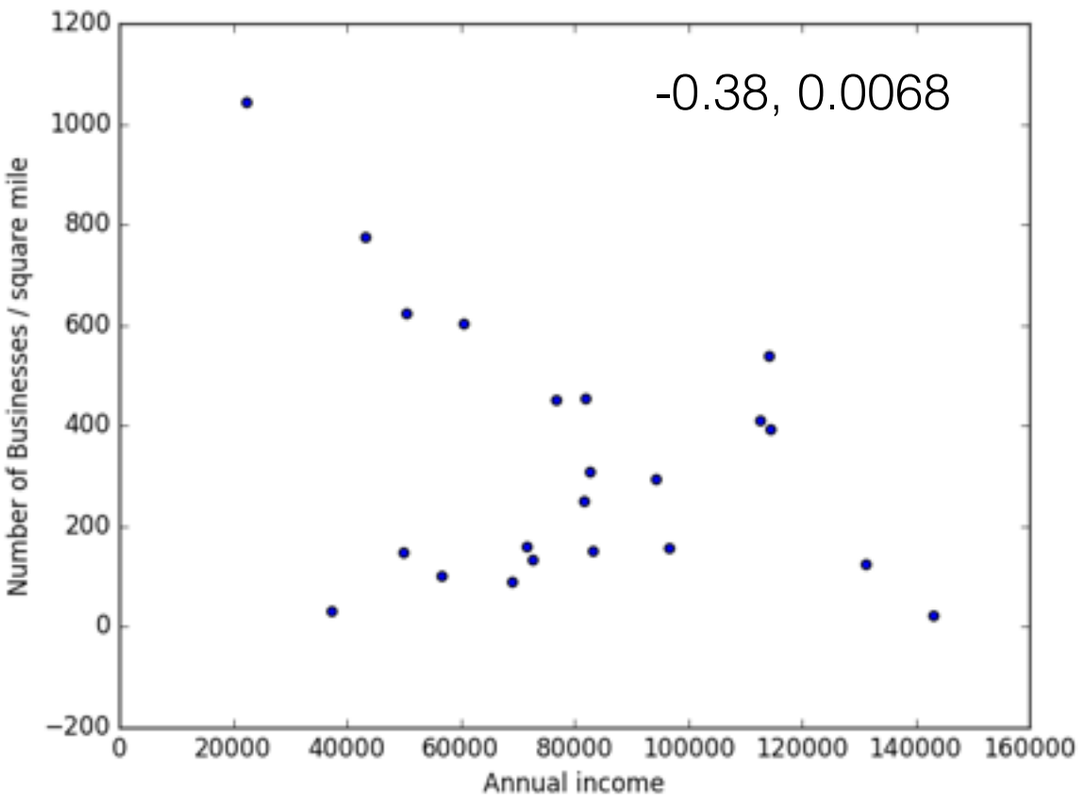
select biz.lat, biz.lon from biz
where (biz.dbaName like "%SAFEWAY%" or biz.dbaName like "%FOOD%"
or biz.dbaName like "%FARMER%" or biz.dbaName like "%GROCERY%")
and (lat > 0 and lon < 0)




0.732, -0.00453
Number of Businesses / square mile
Pop. Density
1200
1000
800
600
400
200
0
−200
−10000
0
10000
20000
30000
40000
50000


0.783, -0.0078
Number of Businesses / square mile
Cost of living index
1200
1000
800
600
400
200
0
−200
130
140
150
160
170
180


-0.13, 0.0078
Number of Businesses / square mile
Cost of house
1200
1000
800
600
400
200
0
−200
400000
500000
600000
700000
800000
900000
1000000
1100000


-0.38, 0.0068
Number of Businesses / square mile
Annual income
1200
1000
800
600
400
200
0
−200
0
20000
40000
60000
80000
100000
120000
140000
160000

Use Google Maps API for better visualization.

Use Yelp API to better understand the businesses.

1. Where should I live?
[Rent, Distance to bart, Distance to restaurants, noise, crime, etc.]

2. Where should I start my new business?
[Yelp reviews will help here]