

Notebook

May 4, 2019

Contents

0.1 Crime dataset exploration	4
-------------------------------------	---

```
Collecting package metadata: done
Solving environment: -
The environment is inconsistent, please check the package plan
carefully
The following packages are causing the inconsistency:

- conda-forge/linux-64::opencv==3.4.4=py37_blas_openblash85ad109_1203
- conda-forge/linux-64::matplotlib==3.0.3=py37_1
done

## Package Plan ##

environment location: /opt/conda

added / updated specs:
- geopy

The following packages will be downloaded:
```

package	build	
-----	-----	
blas-2.8	openblas	6 KB
conda-forge		
geographiclib-1.49	py_0	32 KB
conda-forge		
geopy-1.19.0	py_0	53 KB
conda-forge		
libblas-3.8.0	8_openblas	6 KB
conda-forge		
libcbblas-3.8.0	8_openblas	6 KB
conda-forge		
liblapack-3.8.0	8_openblas	6 KB
conda-forge		
liblapacke-3.8.0	8_openblas	6 KB
conda-forge		
numpy-1.15.4	py37h8b7e671_1002	4.2 MB
conda-forge		
openblas-0.3.6	h6e990d7_1	15.8 MB

```

conda-forge
  scikit-learn-0.20.3          |    py37ha8026db_1          5.7 MB  conda
-forge
  scipy-1.2.1                  |    py37h09a28d5_1          18.1 MB
conda-forge
-----
Total:                          43.9 MB

```

The following NEW packages will be INSTALLED:

```

geographiclib    conda-forge/noarch::geographiclib-1.49-py_0
geopy            conda-forge/noarch::geopy-1.19.0-py_0
libblas          conda-forge/linux-64::libblas-3.8.0-8_openblas
libcblas        conda-forge/linux-64::libcblas-3.8.0-8_openblas
liblapack        conda-forge/linux-64::liblapack-3.8.0-8_openblas
liblapacke       conda-forge/linux-64::liblapacke-3.8.0-8_openblas

```

The following packages will be UPDATED:

```

blas              1.1-openblas --> 2.8-
openblas
numpy             1.15.4-py37_blas_openblash1522bff_1000 --> 1.15.4-
py37h8b7e671_1002
openblas         0.3.3-h9ac9557_1001 --> 0.3.6-
h6e990d7_1
scikit-learn     0.20.3-py37_blas_openblashebf5e3_0 --> 0.20.3-
py37ha8026db_1
scipy            1.2.1-py37_blas_openblash1522bff_0 --> 1.2.1-
py37h09a28d5_1

```

Downloading and Extracting Packages

```

blas-2.8          | 6 KB      |
#####           | 100%
scipy-1.2.1       | 18.1 MB   |
#####           | 100%
geographiclib-1.49 | 32 KB     |
#####           | 100%
openblas-0.3.6    | 15.8 MB   |
#####           | 100%
numpy-1.15.4      | 4.2 MB    |
#####           | 100%
liblapacke-3.8.0  | 6 KB      |
#####           | 100%
liblapack-3.8.0   | 6 KB      |
#####           | 100%
scikit-learn-0.20.3 | 5.7 MB   |
#####           | 100%
libblas-3.8.0     | 6 KB      |
#####           | 100%
geopy-1.19.0      | 53 KB     |
#####           | 100%
libcblas-3.8.0    | 6 KB      |
#####           | 100%

```

Preparing transaction: done

Verifying transaction: done

```

Executing transaction: done
Collecting package metadata: done
Solving environment: \
The environment is inconsistent, please check the package plan
carefully
The following packages are causing the inconsistency:

- conda-forge/linux-64::opencv==3.4.4=py37_blas_openblash85ad109_1203
- conda-forge/linux-64::matplotlib==3.0.3=py37_1
done

## Package Plan ##

environment location: /opt/conda

added / updated specs:
- folium=0.5.0

The following packages will be downloaded:

package | build
-----|-----
altair-3.0.1 | py37_0 727 KB
conda-forge
branca-0.3.1 | py_0 25 KB
conda-forge
folium-0.5.0 | py_0 45 KB
conda-forge
-----
Total: 797 KB

The following NEW packages will be INSTALLED:

altair conda-forge/linux-64::altair-3.0.1-py37_0
branca conda-forge/noarch::branca-0.3.1-py_0
folium conda-forge/noarch::folium-0.5.0-py_0

Downloading and Extracting Packages
altair-3.0.1 | 727 KB |
##### | 100%
folium-0.5.0 | 45 KB |
##### | 100%
branca-0.3.1 | 25 KB |
##### | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
Folium installed
Libraries imported.

```

```

Your credentials:
CLIENT_ID: R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0
CLIENT_SECRET: RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UHOMV32HP

```

0.1 Crime dataset exploration

Incident Datetime Incident Date Incident Time Incident Year						
\						
0	2018/12/02	12:45:00	AM	2018/12/02	00:45	2018
1	2018/12/01	08:30:00	PM	2018/12/01	20:30	2018
2	2019/03/18	02:01:00	PM	2019/03/18	14:01	2019
3	2019/03/20	08:00:00	AM	2019/03/20	08:00	2019
4	2019/03/12	01:30:00	PM	2019/03/12	13:30	2019
Incident Day of Week			Report Datetime		Row ID	Incident
ID \						
0	Sunday	2018/12/02	01:56:00	AM	74374327130	
743743						
1	Saturday	2018/12/01	09:18:00	PM	74370071000	
743700						
2	Monday	2019/03/18	02:21:00	PM	78164004134	
781640						
3	Wednesday	2019/03/20	02:06:00	PM	78169706244	
781697						
4	Tuesday	2019/03/15	06:02:00	PM	78154706372	
781547						
Incident Number		CAD Number	...	Longitude	\	
0	180908554	183360210.0	...	-122.404795		
1	180908112	183353564.0	...	-122.408036		
2	190194129	190772267.0	...	-122.406699		
3	190199583	190792201.0	...	-122.404865		
4	196055103	NaN	...	NaN		
point SF Find Neighborhoods						\
0	(37.78490829943, -122.40479506276)					32.0
1	(37.786409612811, -122.408036237445)					19.0
2	(37.756833733806, -122.406699002688)					53.0
3	(37.78400661242, -122.404864795177)					32.0
4	NaN					NaN
Current Police Districts			Current Supervisor Districts		\	
0	5.0		10.0			
1	6.0		3.0			
2	3.0		2.0			
3	1.0		10.0			
4	NaN		NaN			
Analysis Neighborhoods HSOC Zones as of 2018-06-05 OWED Public Spaces						
\						
0	8.0		NaN			
NaN						
1	36.0		NaN			
NaN						
2	20.0		3.0			
NaN						
3	34.0		NaN			
NaN						
4	NaN		NaN			

NaN

```
Central Market/Tenderloin Boundary Polygon - Updated \
0 NaN
1 1.0
2 NaN
3 NaN
4 NaN
```

```
Parks Alliance CPSI (27+TL sites)
0 NaN
1 NaN
2 NaN
3 NaN
4 NaN
```

[5 rows x 34 columns]

	Police District	Latitude	Longitude	Incident Category
0	Southern	37.784908	-122.404795	Other Miscellaneous
1	Central	37.786410	-122.408036	Lost Property
2	Mission	37.756834	-122.406699	Assault
3	Southern	37.784007	-122.404865	Larceny Theft
4	Central	NaN	NaN	Larceny Theft

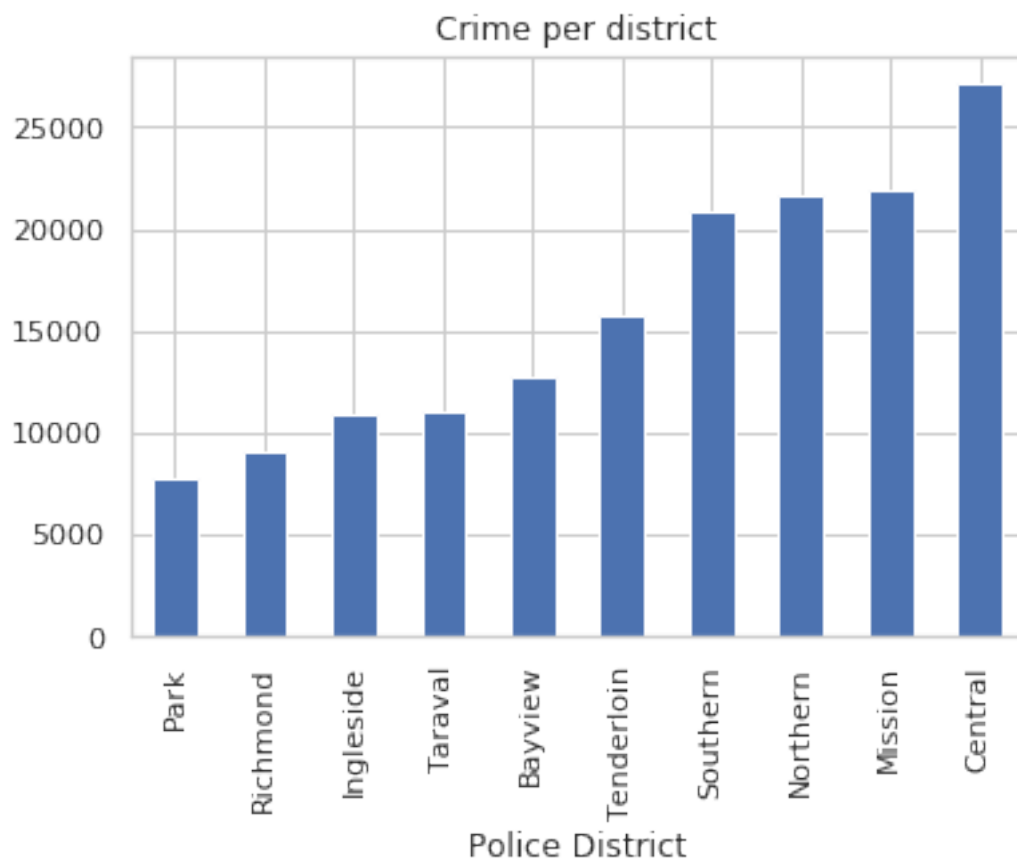
	Police District	Latitude	Longitude	Incident Category
0	60407 Bayview	37.728051	-122.382126	Assault
1	192597 Central	37.791716	-122.398264	Larceny Theft
2	109345 Ingleside	37.709540	-122.457670	Larceny Theft
3	62767 Mission	37.765606	-122.410487	Disorderly Conduct
4	70315 Northern	37.785373	-122.431366	Larceny Theft
5	122448 Park	37.767462	-122.433408	Lost Property
6	47072 Richmond	37.768157	-122.500460	Fire Report
7	173837 Southern	37.789369	-122.391858	Non-Criminal
8	194798 Taraval	37.763943	-122.468451	Assault
9	29790 Tenderloin	37.786520	-122.414807	Fraud

	Incident Category	Count
0	Human Trafficking, Commercial Sex Acts	5
1	Weapons Offence	6
2	Gambling	17
3	Homicide	30
4	Suspicious	32
5	Motor Vehicle Theft?	37
6	Liquor Laws	42
7	Rape	56
8	Drug Violation	61
9	Suicide	81
10	Vehicle Misplaced	82
11	Human Trafficking (A), Commercial Sex Acts	85
12	Vehicle Impounded	130
13	Courtesy Report	131
14	Fire Report	189
15	Vandalism	235

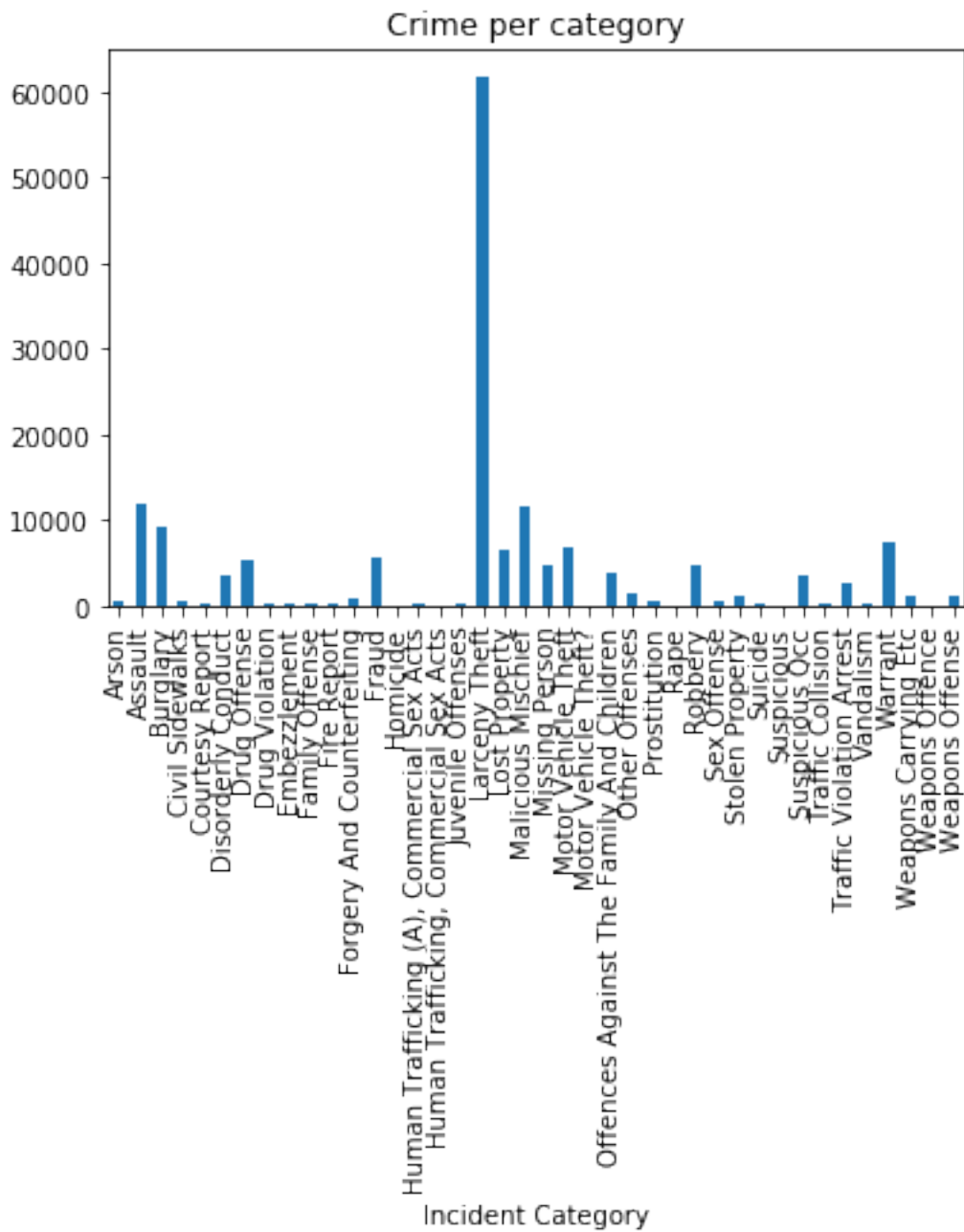
16	Embezzlement	236
17	Family Offense	255
18	Juvenile Offenses	263
19	Traffic Collision	354
20	Sex Offense	402
21	Arson	453
22	Prostitution	569
23	Civil Sidewalks	640
24	Case Closure	773
25	Forgery And Counterfeiting	882
26	Stolen Property	1037
27	Weapons Offense	1047
28	Weapons Carrying Etc	1082
29	Other	1239
30	Other Offenses	1360
31	Miscellaneous Investigation	1731
32	Traffic Violation Arrest	2721
33	Disorderly Conduct	3430
34	Suspicious Occ	3578
35	Offences Against The Family And Children	3886
36	Recovered Vehicle	4218
37	Missing Person	4648
38	Robbery	4718
39	Drug Offense	5361
40	Fraud	5628
41	Lost Property	6590
42	Motor Vehicle Theft	6820
43	Warrant	7423
44	Burglary	9077
45	Malicious Mischief	11589
46	Assault	11813
47	Non-Criminal	12347
48	Other Miscellaneous	15545
49	Larceny Theft	61902
Cleaned 36124 records, number of incidents left		is 158729
Incident Category		Count
0	Human Trafficking, Commercial Sex Acts	5
1	Weapons Offence	6
2	Homicide	30
3	Suspicious	32
4	Motor Vehicle Theft?	37
5	Rape	56
6	Drug Violation	61
7	Suicide	81
8	Human Trafficking (A), Commercial Sex Acts	85
9	Courtesy Report	131
10	Fire Report	189
11	Vandalism	235
12	Embezzlement	236
13	Family Offense	255
14	Juvenile Offenses	263
15	Traffic Collision	354
16	Sex Offense	402
17	Arson	453
18	Prostitution	569

19		Civil Sidewalks	640
20		Forgery And Counterfeiting	882
21		Stolen Property	1037
22		Weapons Offense	1047
23		Weapons Carrying Etc	1082
24		Other Offenses	1360
25		Traffic Violation Arrest	2721
26		Disorderly Conduct	3430
27		Suspicious Occ	3578
28	Offences Against The Family And Children		3886
29		Missing Person	4648
30		Robbery	4718
31		Drug Offense	5361
32		Fraud	5628
33		Lost Property	6590
34		Motor Vehicle Theft	6820
35		Warrant	7423
36		Burglary	9077
37		Malicious Mischief	11589
38		Assault	11813
39		Larceny Theft	61902

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f50898420b8>
```



<matplotlib.axes._subplots.AxesSubplot at 0x7f5226914198>

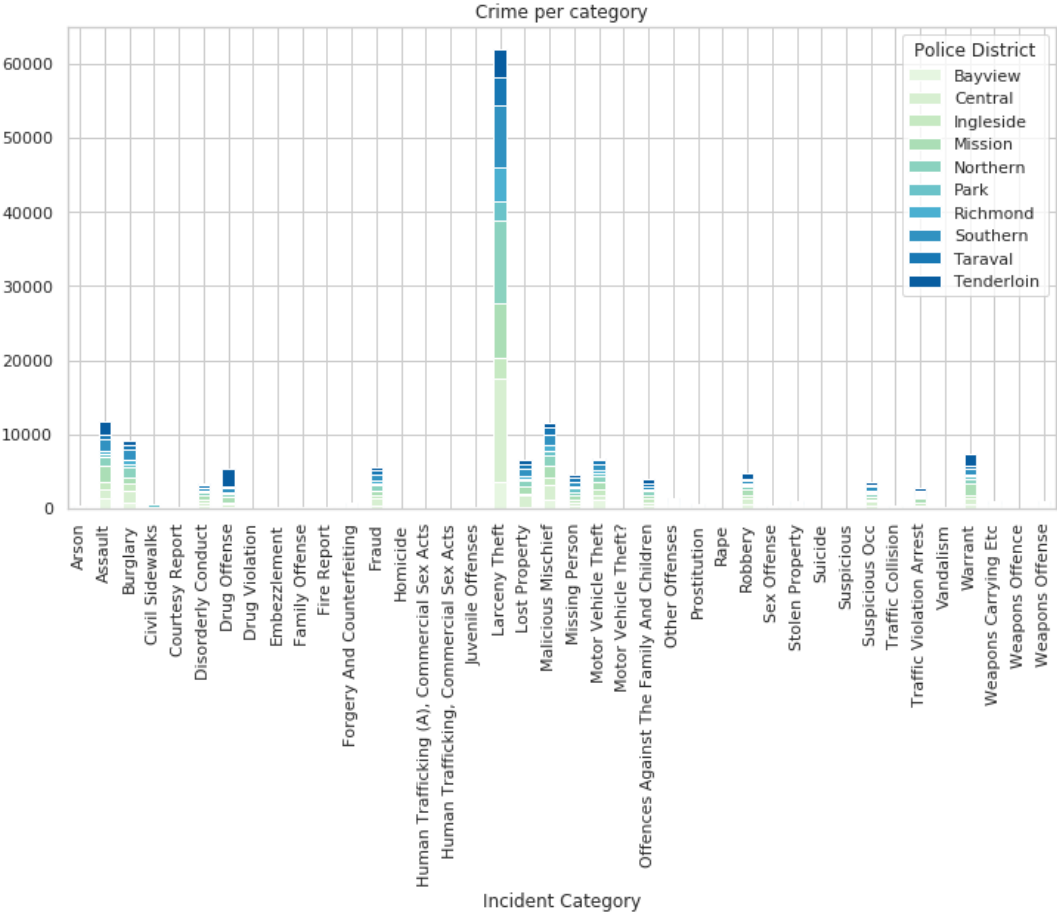


	Incident Category	Police District	count
0	Larceny Theft	Central	14091
1	Larceny Theft	Northern	11128
2	Larceny Theft	Southern	8234
3	Larceny Theft	Mission	7334
4	Larceny Theft	Richmond	4533
5	Larceny Theft	Taraval	3866
6	Larceny Theft	Tenderloin	3734
7	Larceny Theft	Bayview	3487

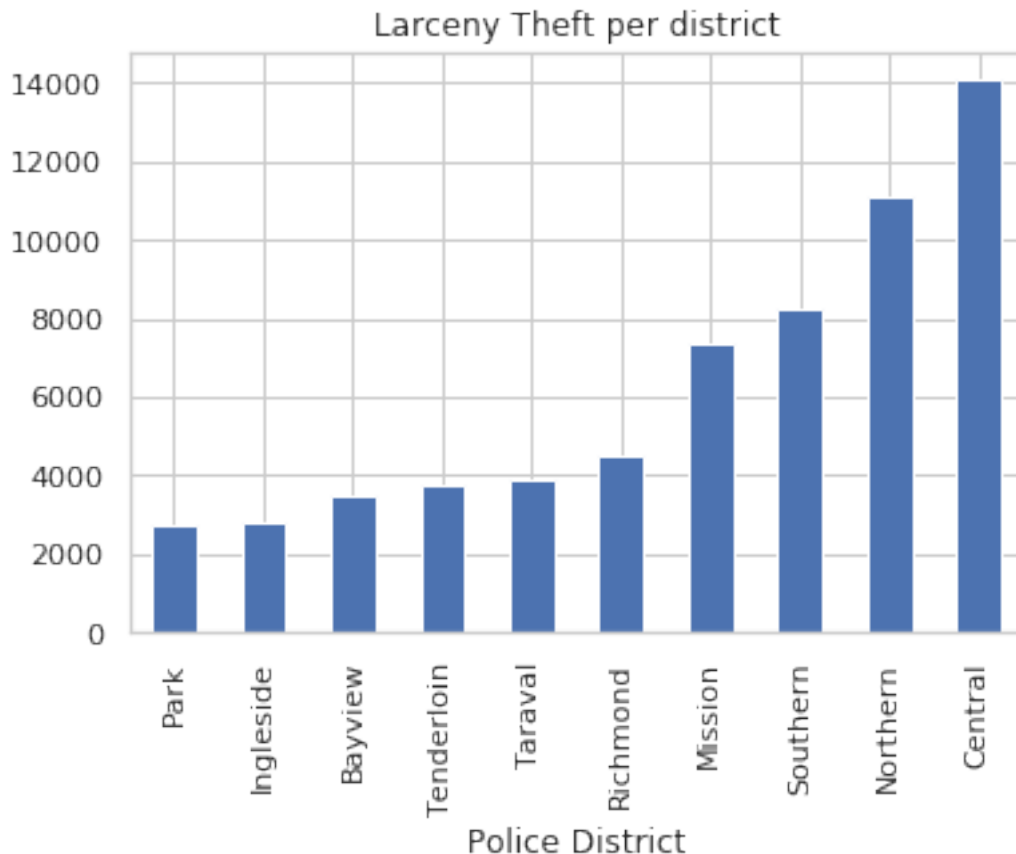
8		Larceny Theft	Ingleside	2792
9		Larceny Theft	Park	2703
10		Drug Offense	Tenderloin	2402
11		Assault	Mission	2054
12		Malicious Mischief	Central	1915
13		Assault	Tenderloin	1815
14		Warrant	Tenderloin	1729
15		Burglary	Central	1590
16		Assault	Southern	1586
17		Malicious Mischief	Southern	1570
18		Malicious Mischief	Mission	1540
19		Malicious Mischief	Northern	1512
20		Warrant	Mission	1462
21		Lost Property	Central	1461
22		Burglary	Southern	1407
23		Assault	Central	1363
24		Burglary	Northern	1358
25		Assault	Bayview	1311
26		Malicious Mischief	Bayview	1275
27		Assault	Northern	1253
28		Motor Vehicle Theft	Bayview	1118
29		Lost Property	Southern	1032
..	
344	Human Trafficking, Commercial Sex Acts		Southern	3
345		Civil Sidewalks	Tenderloin	3
346		Homicide	Southern	3
347		Homicide	Central	3
348		Weapons Offence	Mission	2
349		Prostitution	Ingleside	2
350		Rape	Central	2
351		Drug Violation	Northern	2
352		Weapons Offence	Central	2
353		Motor Vehicle Theft?	Tenderloin	2
354		Rape	Richmond	2
355		Homicide	Northern	2
356		Rape	Park	2
357		Suspicious	Northern	1
358	Human Trafficking, Commercial Sex Acts		Mission	1
359		Drug Violation	Park	1
360	Human Trafficking (A), Commercial Sex Acts		Ingleside	1
361		Homicide	Richmond	1
362		Motor Vehicle Theft?	Park	1
363		Motor Vehicle Theft?	Richmond	1
364	Human Trafficking, Commercial Sex Acts		Ingleside	1
365		Juvenile Offenses	Central	1
366		Suspicious	Mission	1
367		Homicide	Ingleside	1
368		Suspicious	Taraval	1
369		Suspicious	Park	1
370		Weapons Offence	Tenderloin	1
371		Weapons Offence	Taraval	1
372		Suspicious	Bayview	1
373	Human Trafficking (A), Commercial Sex Acts		Bayview	1

[374 rows x 3 columns]

<matplotlib.axes._subplots.AxesSubplot at 0x7f5089de1828>



<matplotlib.axes._subplots.AxesSubplot at 0x7f50897540b8>



```

/opt/conda/lib/python3.7/site-packages/numpy/core/_methods.py:32:
RuntimeWarning: invalid value encountered in reduce
    return umr_minimum(a, axis, None, out, keepdims, initial)
/opt/conda/lib/python3.7/site-packages/numpy/core/_methods.py:28:
RuntimeWarning: invalid value encountered in reduce
    return umr_maximum(a, axis, None, out, keepdims, initial)
/opt/conda/lib/python3.7/site-packages/matplotlib/colors.py:512:
RuntimeWarning: invalid value encountered in less
    xa[xa < 0] = -1

```

```
<pandas.io.formats.style.Styler at 0x7f50896f2940>
```

```

https://api.foursquare.com/v2/venues/explore?client_id=
R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=
RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll
=37.728051442991145,-122.38212570745505&v=20180604&radius=1000&limit
=3000
https://api.foursquare.com/v2/venues/explore?client_id=
R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=
RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll
=37.791716057660935,-122.39826444939199&v=20180604&radius=1000&limit
=3000
https://api.foursquare.com/v2/venues/explore?client_id=
R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=
RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll

```

```
=37.709539891678034,-122.45767030673143&v=20180604&radius=1000&limit=3000
https://api.foursquare.com/v2/venues/explore?client_id=R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll=37.76560583760297,-122.41048669029504&v=20180604&radius=1000&limit=3000
https://api.foursquare.com/v2/venues/explore?client_id=R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll=37.78537264618818,-122.43136639807877&v=20180604&radius=1000&limit=3000
https://api.foursquare.com/v2/venues/explore?client_id=R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll=37.76746158246703,-122.43340825880185&v=20180604&radius=1000&limit=3000
https://api.foursquare.com/v2/venues/explore?client_id=R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll=37.76815687456544,-122.50045983043844&v=20180604&radius=1000&limit=3000
https://api.foursquare.com/v2/venues/explore?client_id=R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll=37.78936926678204,-122.39185829145077&v=20180604&radius=1000&limit=3000
https://api.foursquare.com/v2/venues/explore?client_id=R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll=37.7639429851715,-122.46845112003868&v=20180604&radius=1000&limit=3000
https://api.foursquare.com/v2/venues/explore?client_id=R3IASVU5PXYSX2RCMNB3KJTEKMUSRXIOZKQX1NMBOV1LY4W0&client_secret=RHZYEXWQ5AUX3D1VFFGPR4QXKLZPOEJJZFMPZX5UH0MV32HP&ll=37.7865198190659,-122.41480664938405&v=20180604&radius=1000&limit=3000
```

Police District	reasons.count	\
0	Bayview	0
1	Bayview	0
2	Bayview	0
3	Bayview	0
4	Bayview	0

	reasons.items	\
0	[{'summary': 'This spot is popular', 'type': '...	
1	[{'summary': 'This spot is popular', 'type': '...	
2	[{'summary': 'This spot is popular', 'type': '...	
3	[{'summary': 'This spot is popular', 'type': '...	
4	[{'summary': 'This spot is popular', 'type': '...	

	referralId	\
0	e-0-523a4dfb8bbd84de289ee1ba-0	
1	e-0-4f3709abe4b0aa62a1d35520-1	
2	e-0-5703cab0498e52dae7c36bcf-2	
3	e-0-4c5a47e9aeb7b71319b373cd-3	
4	e-0-4d76aae2b483a0932ebe9305-4	

	venue.categories	venue.delivery.id
\		
0	[{'id': '50327c8591d4c4b30a586d5d', 'name': 'B...	NaN

```

1  [{'id': '4bf58dd8d48988d1ed941735', 'name': 'S...      NaN
2  [{'id': '4bf58dd8d48988d16a941735', 'name': 'B...      NaN
3  [{'id': '4bf58dd8d48988d176941735', 'name': 'G...      NaN
4  [{'id': '4bf58dd8d48988d14f941735', 'name': 'S...      807702

    venue.delivery.provider.icon.name \
0                                     NaN
1                                     NaN
2                                     NaN
3                                     NaN
4  /delivery_provider_grubhub_20180129.png

    venue.delivery.provider.icon.prefix \
0                                     NaN
1                                     NaN
2                                     NaN
3                                     NaN
4  https://fastly.4sqi.net/img/general/cap/

venue.delivery.provider.icon.sizes venue.delivery.provider.name ...
\
0                                     NaN      NaN ...
1                                     NaN      NaN ...
2                                     NaN      NaN ...
3                                     NaN      NaN ...
4                                     [40, 50]      grubhub ...

    venue.location.labeledLatLngs  venue.location.
lat \
0  [{'label': 'display', 'lat': 37.7268123517755,...
37.726812
1  [{'label': 'display', 'lat': 37.73052542272909...
37.730525
2  [{'label': 'display', 'lat': 37.72690804644706...
37.726908
3  [{'label': 'display', 'lat': 37.73185052412636...
37.731851
4  [{'label': 'display', 'lat': 37.7322674, 'lng'...
37.732267

    venue.location.lng venue.location.neighborhood venue.location.
postalCode \
0          -122.382880      NaN
94124
1          -122.372998      NaN
94124
2          -122.391536      Bayview
94124
3          -122.389733      NaN
94124
4          -122.391890      NaN
94124

    venue.location.state      venue.name venue.photos.
count \

```

```

0          CA  The San Francisco Mead Company
0
1          CA          Archimedes Banya
0
2          CA          Craftsman and Wolves Den
0
3          CA          Bayview Hunters Point YMCA
0
4          CA          JJ Fish & Chicken
0

```

```

venue.photos.groups  venue.venuePage.id
0                    []                NaN
1                    []                NaN
2                    []                NaN
3                    []                NaN
4                    []                NaN

```

```
[5 rows x 31 columns]
```

Empty DataFrame

```

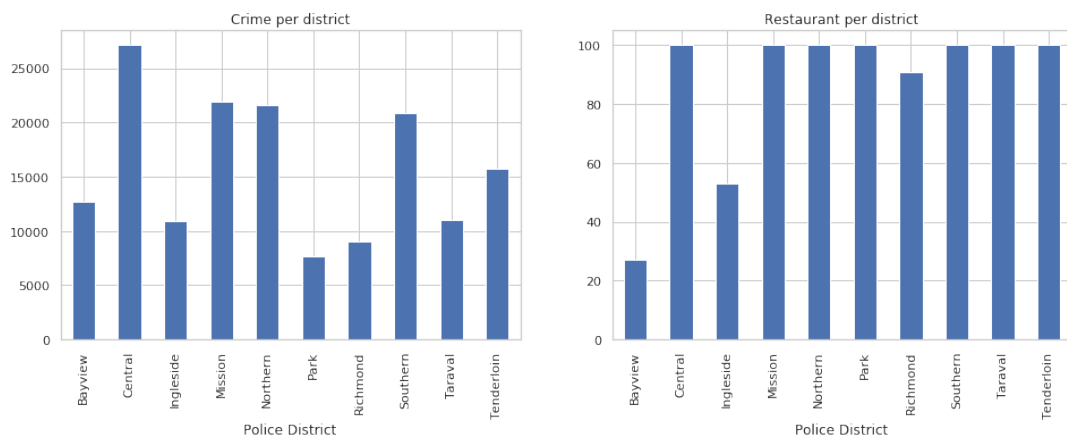
Columns: [Police District, reasons.count, reasons.items, referralId,
venue.categories, venue.delivery.id, venue.delivery.provider.icon.name,
venue.delivery.provider.icon.prefix, venue.delivery.provider.icon.sizes,
venue.delivery.provider.name, venue.delivery.url, venue.events.count,
venue.events.summary, venue.id, venue.location.address, venue.location.cc,
venue.location.city, venue.location.country, venue.location.crossStreet,
venue.location.distance, venue.location.formattedAddress, venue.location.labeledLatLngs,
venue.location.lat, venue.location.lng, venue.location.neighborhood, venue.location.postalCode,
venue.location.state, venue.name, venue.photos.count, venue.photos.groups, venue.venuePage.id]
Index: []

```

```
[0 rows x 31 columns]
```

914

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f52d6dbaeb8>
```



```
<folium.folium.Map at 0x7f50896eee10>
```

```
GeoJSON file downloaded!
```

```
<folium.folium.Map at 0x7f5261558390>
```