CCAA_Python_Training_Part_2

September 27, 2021

0.0.1 CCAA Python Training Part 2

Date Created: September 8th 2021

Created By: Analyst 134

Data Sources and Examples Located in NFCAPythonTraining.ipynb

```
   year   INCIDENT_ID    OFFENSE_ID  OFFENSE_CODE  OFFENSE_CODE_EXTENSION  \
0  2020  2.020132e+09  2.020000e+15          5707                       0
1  2020  2.020121e+09  2.020000e+15          5441                       0
2  2020  2.020298e+08  2.020000e+14          2305                       0
3  2020  2.020431e+08  2.020000e+14          5212                       1
4  2020  2.020466e+08  2.020000e+14          2303                       0

                OFFENSE_TYPE_ID       OFFENSE_CATEGORY_ID  \
0           criminal-trespassing          all-other-crimes
1              traffic-accident          traffic-accident
2      theft-items-from-vehicle  theft-from-motor-vehicle
3  weapon-by-prev-offender-powpo          all-other-crimes
4                theft-shoplift                   larceny

  FIRST_OCCURRENCE_DATE LAST_OCCURRENCE_DATE    REPORTED_DATE  …  \
0       2/29/2020 20:15                  NaN  2/29/2020 21:16  …
1       2/25/2020 11:00                  NaN  2/25/2020 11:02  …
2       1/13/2020 23:00       1/14/2020 5:00  1/14/2020 11:07  …
3        1/20/2020 2:02                  NaN   1/20/2020 3:36  …
4       1/21/2020 15:43                  NaN  1/21/2020 17:13  …

      GEO_LON    GEO_LAT  DISTRICT_ID  PRECINCT_ID            NEIGHBORHOOD_ID  \
0 -104.988366  39.754698          6.0        611.0                five-points
1 -104.975266  39.730449          6.0        623.0              capitol-hill
2 -105.010399  39.679879          4.0        422.0  college-view-south-platte
3 -104.988962  39.752955          6.0        611.0                five-points
4 -104.891880  39.784132          5.0        511.0                  stapleton

   IS_CRIME  IS_TRAFFIC    New_Date  Months  monthName
0         1           0  2020-02-29       2   February
1         0           1  2020-02-25       2   February
```

2	1	0	2020-01-14	1	January
3	1	0	2020-01-20	1	January
4	1	0	2020-01-21	1	January

[5 rows x 23 columns]

0.1 Step 3: Create Frequency Tables 2020 Denver Data Frequency Tables

How to Create Frequency Tables

Frequency Tables in Python are useful for looking at one variable. If you need to compare two columns, look at pivot tables. Note: piviot tables cant compare columns that have duplicates

How to Create Pivot Tables

0.1.1 Crime Types

col_0	count	percent
OFFENSE_CATEGORY_ID		
traffic-accident	14603	17.72%
all-other-crimes	12012	14.57%
theft-from-motor-vehicle	10179	12.35%
larceny	9851	11.95%
public-disorder	9762	11.84%
auto-theft	8273	10.04%
burglary	5141	6.24%
other-crimes-against-persons	3911	4.74%
aggravated-assault	2892	3.51%
drug-alcohol	2538	3.08%
robbery	1177	1.43%
white-collar-crime	1146	1.39%
sexual-assault	705	0.86%
arson	154	0.19%
murder	85	0.10%

0.1.2 Neighborhoods Crime Rate

col_0	count	percent
NEIGHBORHOOD_ID		
five-points	4431	5.38%
capitol-hill	3363	4.08%
montbello	2653	3.22%
cbd	2576	3.13%
gateway-green-valley-ranch	2283	2.77%
stapleton	2228	2.70%
union-station	2129	2.58%
lincoln-park	2044	2.48%
baker	2035	2.47%

civic-center	2016	2.45%
east-colfax	1788	2.17%
west-colfax	1687	2.05%
hampden	1680	2.04%
northeast-park-hill	1679	2.04%
north-capitol-hill	1655	2.01%
westwood	1625	1.97%
dia	1607	1.95%
central-park	1585	1.92%
highland	1555	1.89%
hampden-south	1520	1.84%
speer	1420	1.72%
elyria-swansea	1265	1.53%
globeville	1233	1.50%
washington-virginia-vale	1164	1.41%
cheesman-park	1114	1.35%
mar-lee	1097	1.33%
villa-park	1095	1.33%
city-park-west	1084	1.32%
virginia-village	1009	1.22%
ruby-hill	1008	1.22%
athmar-park	977	1.19%
overland	963	1.17%
harvey-park	961	1.17%
college-view-south-platte	960	1.16%
sunnyside	928	1.13%
congress-park	919	1.11%
cherry-creek	878	1.07%
windsor	831	1.01%
berkeley	816	0.99%
west-highland	796	0.97%
goldsmith	747	0.91%
lowry-field	742	0.90%
washington-park-west	729	0.88%
barnum	719	0.87%
jefferson-park	698	0.85%
hale	688	0.83%
university-park	663	0.80%
montclair	654	0.79%
harvey-park-south	629	0.76%
sloan-lake	627	0.76%
marston	600	0.73%
platt-park	595	0.72%
valverde	593	0.72%
south-park-hill	576	0.70%
university-hills	571	0.69%
clayton	549	0.67%

```
university                    548   0.66%
bear-valley                   543   0.66%
north-park-hill               523   0.63%
southmoor-park                519   0.63%
cole                          509   0.62%
auraria                       500   0.61%
hilltop                       497   0.60%
whittier                      486   0.59%
sun-valley                    457   0.55%
belcaro                       454   0.55%
regis                         452   0.55%
kennedy                       443   0.54%
washington-park               434   0.53%
chaffee-park                  430   0.52%
barnum-west                   418   0.51%
city-park                     412   0.50%
fort-logan                    387   0.47%
cory-merrill                  376   0.46%
skyland                       302   0.37%
rosedale                      251   0.30%
country-club                  204   0.25%
indian-creek                  155   0.19%
wellshire                      91   0.11%

NEIGHBORHOOD_ID             NEIGHBORHOOD_ID
five-points                 five-points                 4431
capitol-hill                capitol-hill                3363
montbello                   montbello                   2653
cbd                         cbd                         2576
gateway-green-valley-ranch  gateway-green-valley-ranch  2283
Name: NEIGHBORHOOD_ID, dtype: int64
```

0.1.3 Frequency By Month

```
col_0      count percent
monthName
August      7746   9.40%
January     7666   9.30%
October     7587   9.20%
September   7559   9.17%
July        7500   9.10%
November    6939   8.42%
February    6889   8.36%
May         6699   8.13%
June        6647   8.06%
March       6049   7.34%
December    5791   7.03%
```

```
April       5357   6.50%
```

0.2 Step 4: Using Heat Map Graph To Visualize Data

Using Seaborn Heatmap Calendar

Github Notebook Example

```
            year   INCIDENT_ID    OFFENSE_ID  OFFENSE_CODE  \
Datetime
2020-02-29  2020  2.020132e+09  2.020000e+15          5707
2020-02-25  2020  2.020121e+09  2.020000e+15          5441
2020-01-14  2020  2.020298e+08  2.020000e+14          2305
2020-01-20  2020  2.020431e+08  2.020000e+14          5212
2020-01-21  2020  2.020466e+08  2.020000e+14          2303

            OFFENSE_CODE_EXTENSION               OFFENSE_TYPE_ID  \
Datetime
2020-02-29                       0          criminal-trespassing
2020-02-25                       0             traffic-accident
2020-01-14                       0      theft-items-from-vehicle
2020-01-20                       1  weapon-by-prev-offender-powpo
2020-01-21                       0                theft-shoplift

                 OFFENSE_CATEGORY_ID FIRST_OCCURRENCE_DATE  \
Datetime
2020-02-29           all-other-crimes       2/29/2020 20:15
2020-02-25           traffic-accident       2/25/2020 11:00
2020-01-14  theft-from-motor-vehicle       1/13/2020 23:00
2020-01-20           all-other-crimes        1/20/2020 2:02
2020-01-21                    larceny       1/21/2020 15:43

           LAST_OCCURRENCE_DATE    REPORTED_DATE  …     GEO_LON    GEO_LAT  \
Datetime                                          …
2020-02-29                  NaN  2/29/2020 21:16  … -104.988366  39.754698
2020-02-25                  NaN  2/25/2020 11:02  … -104.975266  39.730449
2020-01-14      1/14/2020 5:00  1/14/2020 11:07  … -105.010399  39.679879
2020-01-20                  NaN   1/20/2020 3:36  … -104.988962  39.752955
2020-01-21                  NaN  1/21/2020 17:13  … -104.891880  39.784132

            DISTRICT_ID  PRECINCT_ID            NEIGHBORHOOD_ID  IS_CRIME  \
Datetime
2020-02-29          6.0        611.0                five-points         1
2020-02-25          6.0        623.0               capitol-hill         0
2020-01-14          4.0        422.0  college-view-south-platte         1
2020-01-20          6.0        611.0                five-points         1
2020-01-21          5.0        511.0                  stapleton         1
```

```
            IS_TRAFFIC    New_Date  Months  monthName
Datetime
2020-02-29           0  2020-02-29       2   February
2020-02-25           1  2020-02-25       2   February
2020-01-14           0  2020-01-14       1    January
2020-01-20           0  2020-01-20       1    January
2020-01-21           0  2020-01-21       1    January

[5 rows x 23 columns]
```

```
True
```

```
  NEIGHBORHOOD_ID monthName  Count
0     five-points   January    554
1    capitol-hill       May    499
2     five-points    August    424
3       stapleton   January    418
4    capitol-hill      June    413
```

```
                      Count                                                 \
NEIGHBORHOOD_ID athmar-park auraria  baker barnum barnum-west bear-valley
monthName
April                  59.0    23.0   99.0   46.0        25.0         41.0
August                102.0    56.0  226.0   62.0        42.0         56.0
December               55.0    12.0  122.0   47.0        30.0         38.0
February               85.0    70.0  186.0   50.0        25.0         35.0
January               107.0    71.0  188.0   58.0        46.0         35.0
July                   77.0    32.0  209.0   61.0        39.0         50.0
June                   77.0    27.0  152.0   62.0        37.0         46.0
March                  95.0    54.0  139.0   58.0        25.0         39.0
May                    78.0    30.0  134.0   62.0        41.0         36.0
November               72.0    38.0  191.0   69.0        39.0         46.0
October                87.0    38.0  197.0   64.0        41.0         65.0
September              83.0    49.0  192.0   80.0        28.0         56.0

                                                          …                 \
NEIGHBORHOOD_ID belcaro berkeley capitol-hill    cbd  … virginia-village
monthName                                             …
April              15.0     54.0        213.0  176.0  …             64.0
August             44.0     70.0        286.0  235.0  …            103.0
December           37.0     65.0        177.0  146.0  …             98.0
February           44.0     67.0        227.0  274.0  …             78.0
January            39.0     79.0        254.0  290.0  …             75.0
July               33.0     73.0        305.0  226.0  …             85.0
June               51.0     79.0        413.0  172.0  …             76.0
March              29.0     61.0        200.0  234.0  …             72.0
May                22.0     57.0        499.0  205.0  …             79.0
```

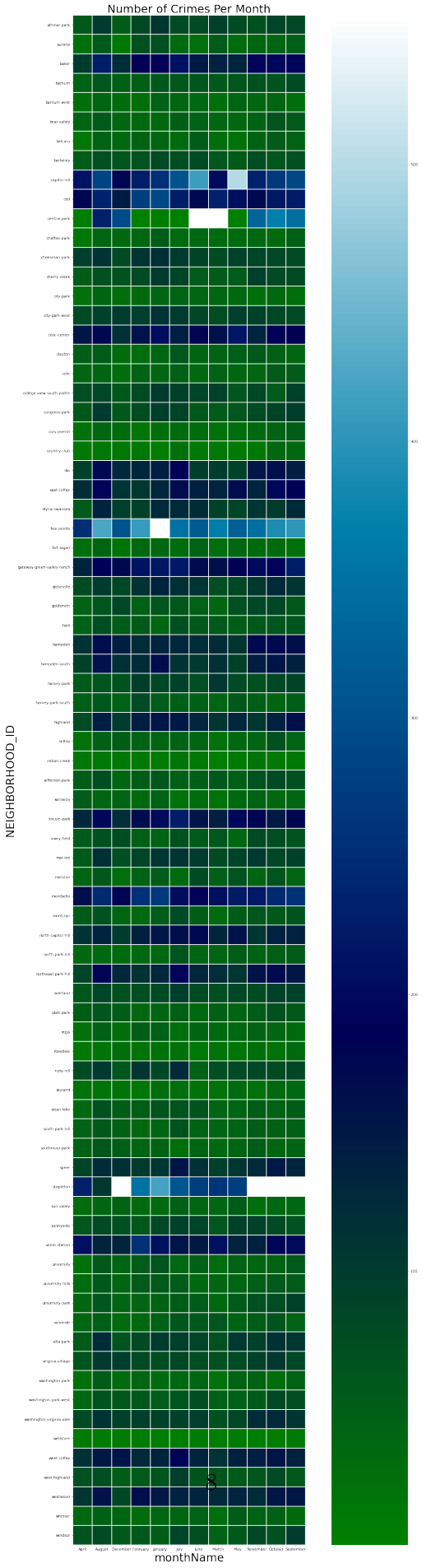
November	43.0	76.0	233.0	175.0	…	90.0
October	54.0	63.0	267.0	219.0	…	105.0
September	43.0	72.0	289.0	224.0	…	84.0

\

NEIGHBORHOOD_ID	washington-park	washington-park-west	washington-virginia-vale
monthName			
April	24.0	37.0	77.0
August	53.0	67.0	113.0
December	30.0	61.0	91.0
February	38.0	61.0	94.0
January	32.0	55.0	92.0
July	36.0	64.0	93.0
June	39.0	76.0	96.0
March	22.0	47.0	70.0
May	44.0	48.0	81.0
November	43.0	57.0	123.0
October	29.0	83.0	126.0
September	44.0	73.0	108.0

NEIGHBORHOOD_ID	wellshire	west-colfax	west-highland	westwood	whittier	windsor
monthName						
April	11.0	115.0	71.0	108.0	36.0	67.0
August	8.0	148.0	74.0	159.0	47.0	77.0
December	8.0	157.0	51.0	81.0	41.0	59.0
February	8.0	126.0	56.0	166.0	31.0	61.0
January	6.0	136.0	64.0	149.0	38.0	51.0
July	4.0	182.0	96.0	138.0	42.0	86.0
June	8.0	135.0	65.0	134.0	45.0	76.0
March	7.0	113.0	64.0	125.0	29.0	45.0
May	8.0	140.0	57.0	137.0	53.0	58.0
November	5.0	142.0	62.0	122.0	43.0	83.0
October	7.0	155.0	78.0	153.0	48.0	68.0
September	11.0	138.0	58.0	153.0	33.0	100.0

[12 rows x 79 columns]


Number of Crimes Per Month
NEIGHBORHOOD_ID
monthName
April
August
December
February
January
July
June
March
May
November
October
September
500
400
300
200
100

1 Conclusion:

Key Takeaways:

1.) Python is a useful tool for cleaning and showcasing large datasets

2.) Python Can be used in addition to Excel.

3.) Python Can be Automated to create monthly reports

4.) Python Is a good beginner programmer language

1.1 Contact Info:

Rachel Stevenson

Email: rachel.stevenson@state.co.us

CDPS Github Repository