

```
In [1]: import pandas as pd
import numpy as np

In [2]: #Adding Column Names
column_names = ['DCNumber', 'LastName', 'FirstName', 'MiddleName', 'NameSuffix', 'Race', 'Sex', 'BirthDate', 'PrisonReleaseDate', 'ReceiptDate', 'releasedateflag_descr', 'race_descr', 'custody_description', 'FACILITY_description', 'Age']

In [3]: #Importing Data
data1 = pd.read_csv('1.csv', sep=',', header=None, names=column_names)

In [4]: data1

Out[4]:
   DCNumber  LastName  FirstName  MiddleName  NameSuffix  Race  Sex  BirthDate  PrisonReleaseDate  ReceiptDate  releasedateflag_descr  race_descr  custody_description  FACILITY_description  Age
0      000084      ODDM      EDDIE      NaN      NaN      W      M      1940-01-01 00:00:00.00000000      NaN      1976-10-08 00:00:00.00000000      life sentence      WHITE      CLOSE      UNION C.I.      84
1      000130  SANTIAGO      DANIEL      A      NaN      H      M      1988-01-28 00:00:00.00000000      NaN      2018-11-14 00:00:00.00000000      life sentence      HISPANIC      CLOSE      WAKULLA C.I.      36
2      000132      TELFAIR      MICHAEL      NaN      NaN      B      M      1967-12-04 00:00:00.00000000      10/04/2033      2018-10-08 00:00:00.00000000      valid release date      BLACK      MINIMUM      MADISON C.I.      56
3      000150      KAAR      NICHOLAS      J      NaN      W      M      2001-01-25 00:00:00.00000000      07/22/2024      2023-01-26 00:00:00.00000000      valid release date      WHITE      MEDIUM      LAWTEY C.I.      23
4      000256      COON      DOUGLAS      D      NaN      B      M      1944-06-28 00:00:00.00000000      NaN      1971-10-12 00:00:00.00000000      life sentence      BLACK      CLOSE      EVERGLADES C.I.      79
...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...      ...
86959  Y92115      SARTIN      NICHOLAS      C      NaN      W      M      1994-03-11 00:00:00.00000000      10/13/2028      2023-11-16 00:00:00.00000000      valid release date      WHITE      MEDIUM      OKEECHOBEE C.I.      29
86960  Y92120  DISRUDE      BRANDON      J      NaN      W      M      1987-07-31 00:00:00.00000000      07/25/2025      2023-10-26 00:00:00.00000000      valid release date      WHITE      NaN      CFRC-MAIN      36
86961  Y92171      HEARD      JALEEL      A      NaN      B      M      1997-04-17 00:00:00.00000000      06/22/2028      2023-12-18 00:00:00.00000000      valid release date      BLACK      MEDIUM      NWFRC ANNEX.      26
86962  Y92210      JONES      GEORGE      F      JR      B      M      1968-10-09 00:00:00.00000000      02/29/2024      2024-01-29 00:00:00.00000000      valid release date      BLACK      NaN      S.F.R.C.      55
86963  Y92215      STACEY      CHRISTOPHER      J      NaN      W      M      2000-01-04 00:00:00.00000000      08/04/2059      2024-01-31 00:00:00.00000000      valid release date      WHITE      NaN      NWFRC ANNEX.      24

86964 rows x 15 columns

In [5]: #Adding Column Names
column_names2=['DCNumber', 'Sequence', 'OffenseDate', 'DateAdjudicated', 'County', 'CaseNumber', 'prisonterm', 'adjudicationcharge_descr']

In [6]: #Importing Data
data2 = pd.read_csv('2.csv', low_memory=False, sep=',', header=None, names=column_names2)

In [7]: data2

Out[7]:
   DCNumber  Sequence  OffenseDate  DateAdjudicated  County  CaseNumber  prisonterm  adjudicationcharge_descr
0      000084      1      1976-01-27      1976-10-08      DUVAL      7600296.0      9999998      1ST DG MUR/PREMED. OR ATT.
1      000130      1      2014-09-21      2018-11-01      HILLSBOROUGH      1708585.0      9999998      SEX BAT BY ADULT/VCTM LT 12
2      000130      2      2014-09-21      2018-11-01      HILLSBOROUGH      1708585.0      9999998      SEX BAT BY ADULT/VCTM LT 12
3      000130      3      2014-09-21      2018-11-01      HILLSBOROUGH      1708585.0      9999998      SEX BAT BY ADULT/VCTM LT 12
4      000130      4      2017-06-06      2018-11-01      HILLSBOROUGH      1708585.0      50000      POSS PHOTO ETC CHILD SEX PERF
...      ...      ...      ...      ...      ...      ...      ...
355573  Y92210      4      2015-12-22      2024-01-22      MIAMI-DADE      1713900.0      51000      FRAUD USE ID $50K+/20+ PERSONS
355574  Y92210      5      2015-12-22      2024-01-22      MIAMI-DADE      1713900.0      51000      MORTGAGE FRAUD > $100K
355575  Y92210      6      2015-12-22      2024-01-22      MIAMI-DADE      1713900.0      51000      UTTER FORGED INSTRUMENT
355576  Y92215      1      2018-06-02      2024-01-12      ESCAMBIA      1803125.0      300000      1ST DG MUR/PREMED. OR ATT.
355577  Y92215      2      2018-06-02      2024-01-12      ESCAMBIA      1803125.0      100000      1ST DG MUR/PREMED. OR ATT.

355578 rows x 8 columns

In [8]: #Saving the file as csv
data1.to_csv('InMateroot.csv', index=False)

In [9]: #Saving the file as csv
data2.to_csv('InmateOffense.csv', index=False)

In [10]: import pandas as pd

#Load the data
inmate_active_root = pd.read_csv('InMateroot.csv', sep=',', dtype={'DCNumber': str})
inmate_active_offenses = pd.read_csv('InmateOffense.csv', sep=',', dtype={'DCNumber': str})

#Merge the DataFrames on 'DCNumber'
merged_df = pd.merge(inmate_active_root, inmate_active_offenses, on='DCNumber', how='inner')

#Filter for 'FLORIDA STATE PRISON' in 'FACILITY_description'
florida_state_prison = merged_df[merged_df['FACILITY_description'] == 'FLORIDA STATE PRISON']

#Count by offense
offense_counts = florida_state_prison.groupby('adjudicationcharge_descr')['DCNumber'].count().reset_index(name='Count')

#Sort the counts in descending order
offense_counts_sorted = offense_counts.sort_values(by='Count', ascending=False)

print(offense_counts_sorted)

   adjudicationcharge_descr  Count
210      ROBB. GUN OR DEADLY WPN      491
98      FEL/DELI W/GUN/CONC WPN/AMMO      319
48      BURGUNKCSTRUC/CLV OR ATT.      252
4      2ND DEG MURD/DANGEROUS ACT      192
13      AGG ASSLT-W/MPN NO INTENT TO K      162
..      ..      ..
184      PETITT-THEFT-MISD      1
181      OWN/RENT STRUC/CONV TRAF/MAN      1
179      ORGANIZES THEFT PROPERTY ECT.      1
178      ORG-FRAUD-$50K OR MORE      1
67      CONSTRUCTIVE POSSESSION      1

[299 rows x 2 columns]

In [11]: #Number of Inmates by Facility
inmates_by_facility = merged_df['FACILITY_description'].value_counts().reset_index()
inmates_by_facility.columns = ['Facility', 'Number of Inmates']

In [12]: #Number of Offenses by Type
offenses_by_type = merged_df['adjudicationcharge_descr'].value_counts().reset_index()
offenses_by_type.columns = ['Offense Type', 'Count']

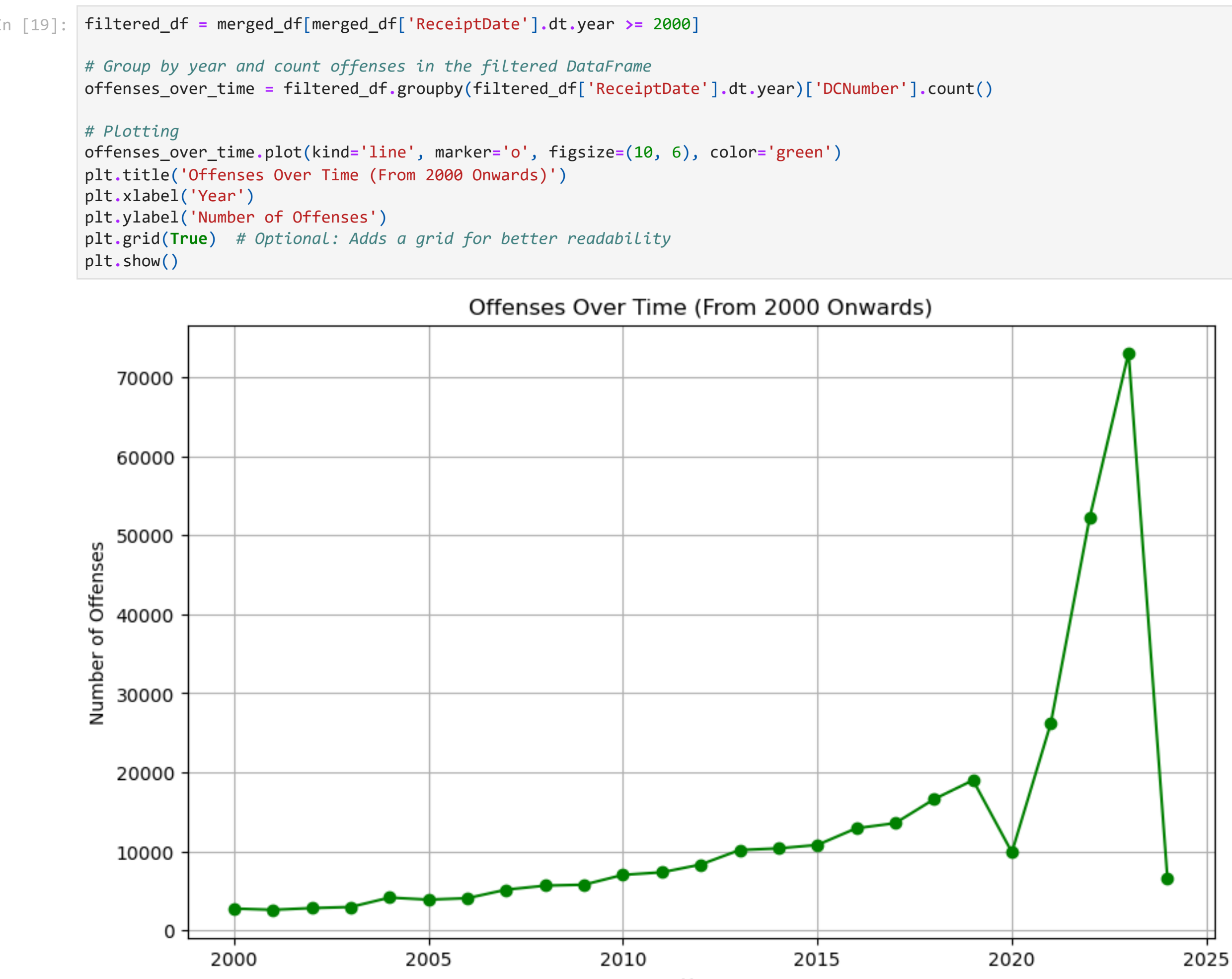
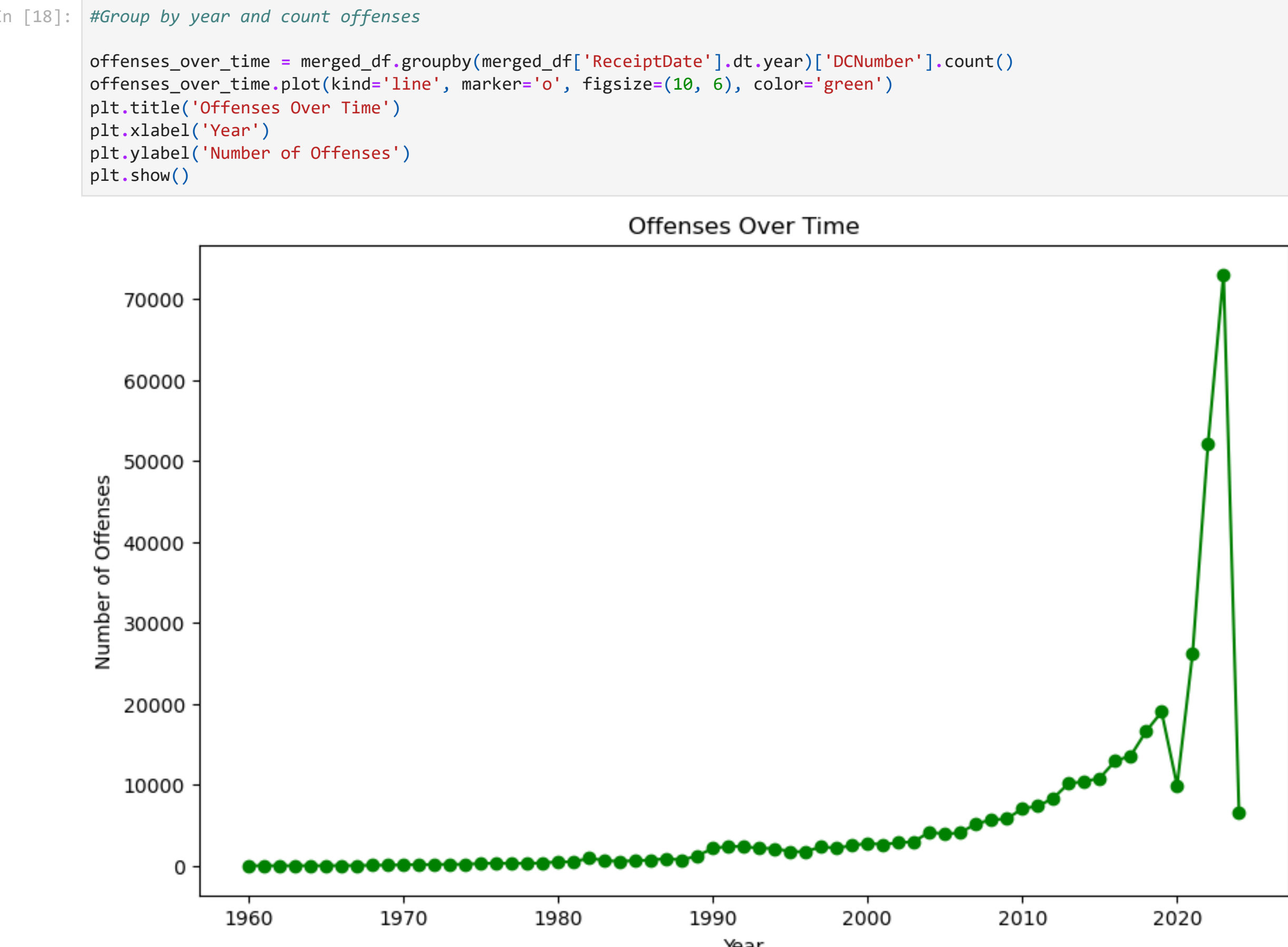
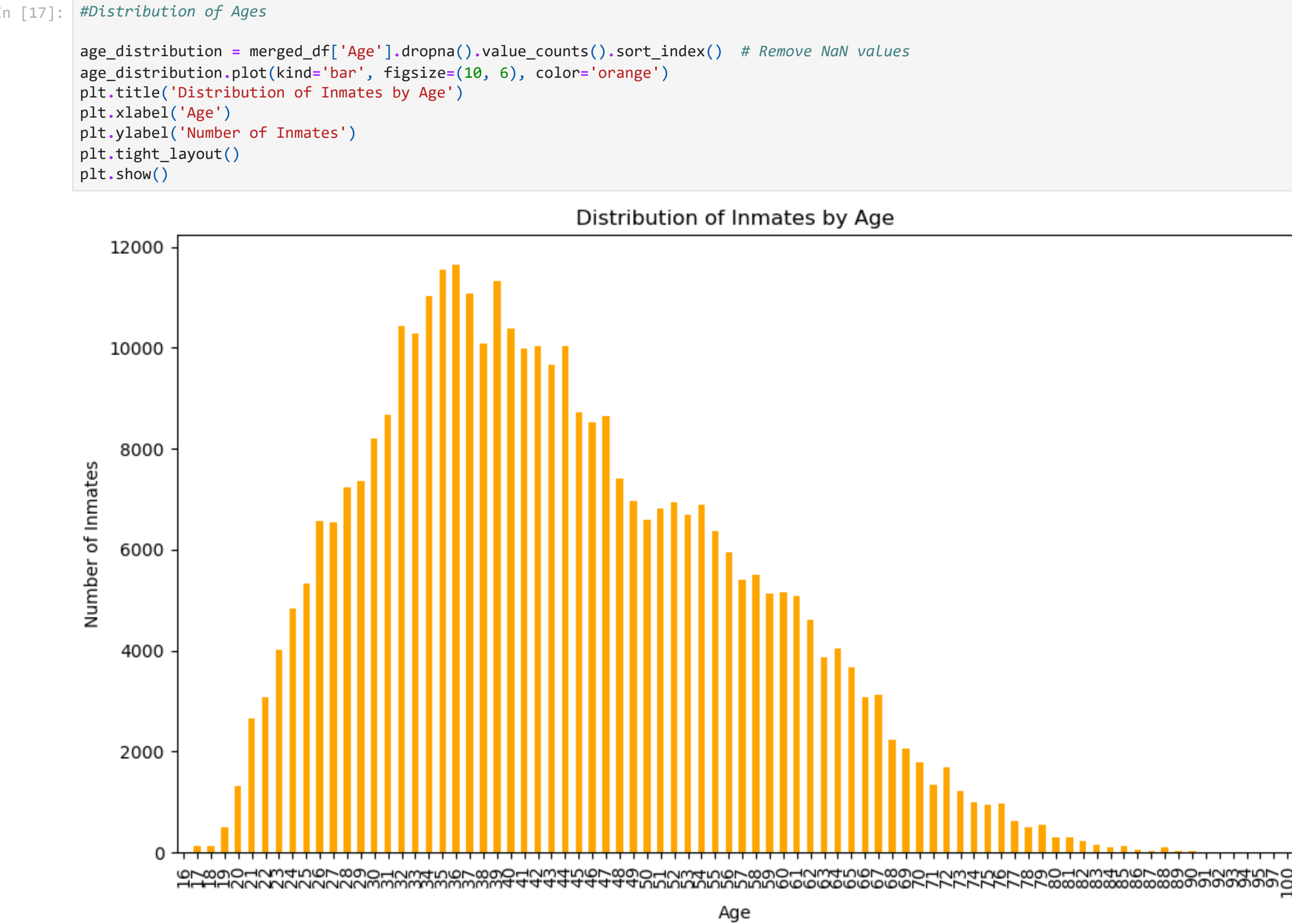
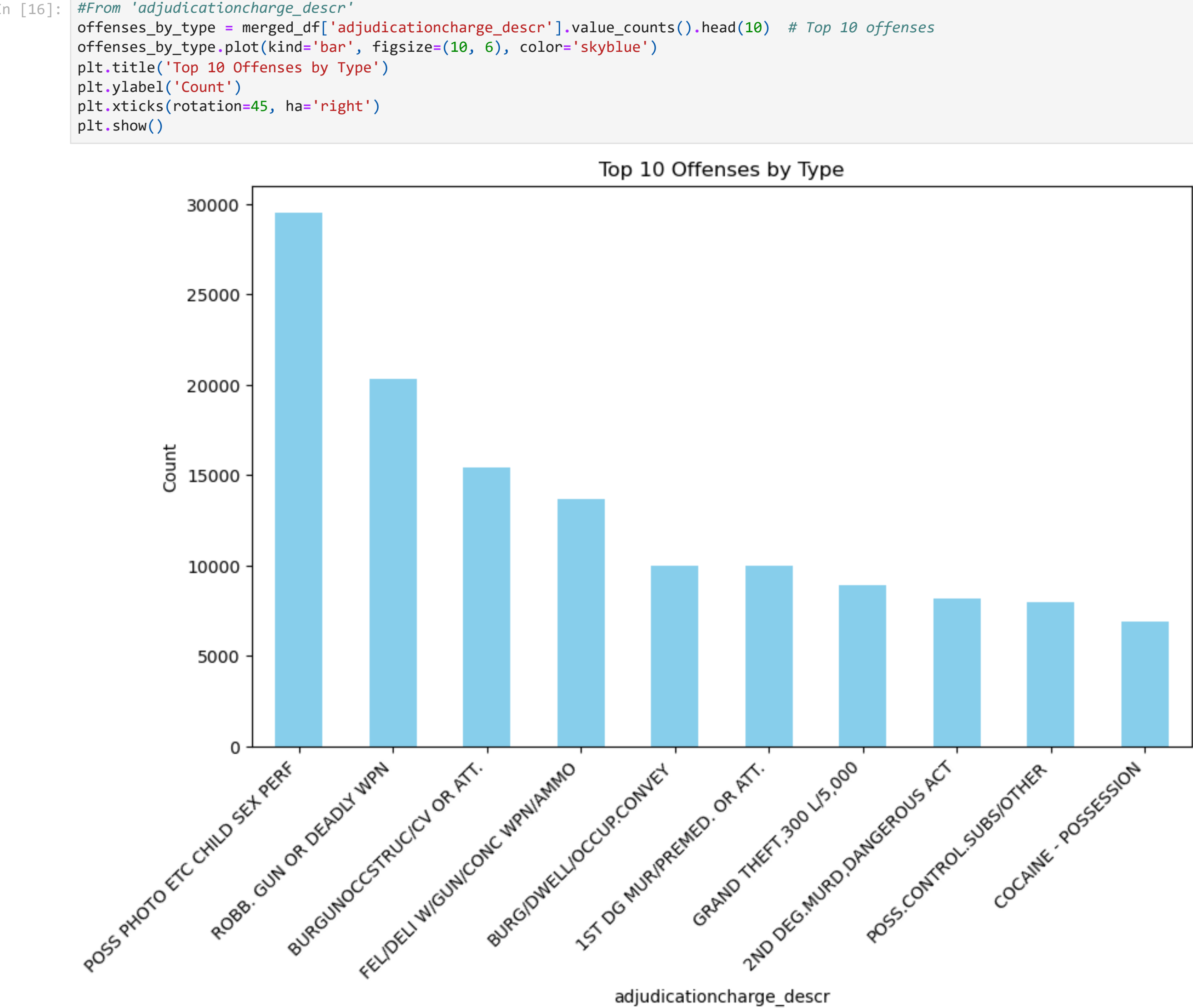
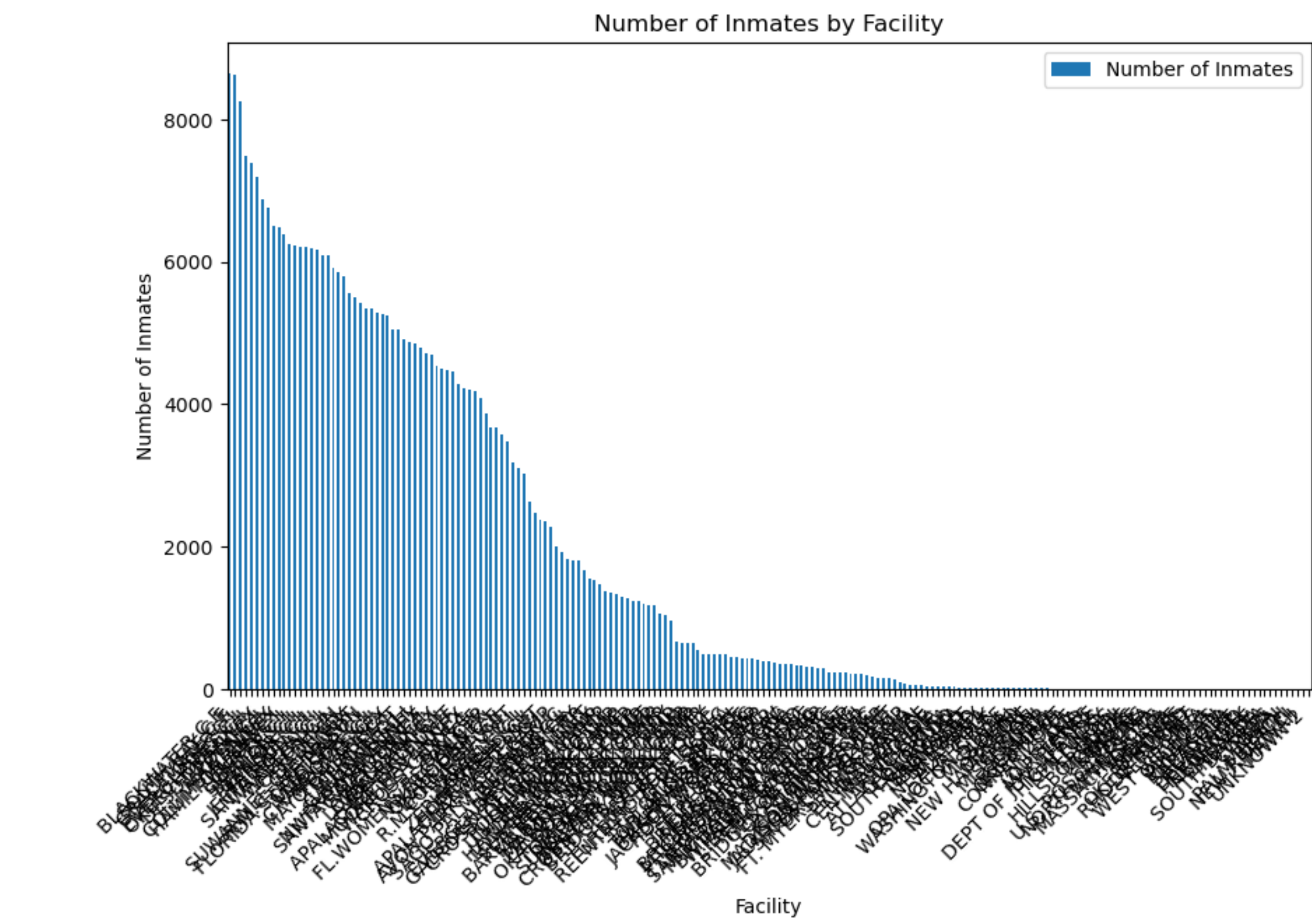
In [13]: #Distribution of Inmates by Age
from datetime import datetime

merged_df['BirthDate'] = pd.to_datetime(merged_df['BirthDate'])
merged_df['Age'] = merged_df['BirthDate'].apply(lambda x: datetime.now().year - x.year)
age_distribution = merged_df['Age'].value_counts().sort_index()

In [14]: #Distribution of Offenses Over Time
merged_df['ReceiptDate'] = pd.to_datetime(merged_df['ReceiptDate'])
offenses_over_time = merged_df.resample('YE', on='ReceiptDate').size()

In [15]: import matplotlib.pyplot as plt

#Plotting number of inmates by facility
inmates_by_facility.plot(kind='bar', x='Facility', y='Number of Inmates', figsize=(10, 6))
plt.title('Number of Inmates by Facility')
plt.ylabel('Number of Inmates')
plt.xticks(rotation=45, ha='right')
plt.show()
```



In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: