

```
In [ ]: import pandas as pd
import matplotlib.pyplot as plt
import plotly_express as px
import os
import numpy as np
# csv to dataframe
df = pd.read_csv('combined.csv')

df = df.drop(['OffenseCode', 'LocationStreetNumber', 'LocationDirectional', 'Ca

# if the offense description is empty, fill it with the string " no offense des
```

```
In [ ]: df.head(5)
```

```
Out[ ]:
```

	IncidentFromDate	IncidentFromTime	IncidentToDate	IncidentToTime	Offense Description	CaseDisp
0	1/1/2010	1:55:00	1/1/2010	2:58:00	Miscellaneous Offenses	
1	1/1/2010	1:36:00	NaN	NaN	Miscellaneous Offenses	
2	1/1/2010	2:00:00	1/1/2010	4:00:00	Larceny – From Building	
3	1/2/2010	6:15:00	NaN	NaN	Miscellaneous Offenses	
4	1/3/2010	1:29:00	NaN	NaN	Miscellaneous Offenses	Pros [

```
In [ ]: # # if there's no locationlatitude and longitude, fill it with NaN
df['LocationLatitude'] = df['LocationLatitude'].fillna(0)
df['LocationLongitude'] = df['LocationLongitude'].fillna(0)

# combine incident from date with incident from time to get the date and time
df['DateTime'] = df['IncidentFromDate'] + ' ' + df['IncidentFromTime']

# move incident fromdate to the datetime value
df['IncidentFromDate'] = df['DateTime']
df.drop(['IncidentFromTime'], axis=1, inplace=True)
df.drop(['DateTime'], axis=1, inplace=True)
```

```
In [ ]: df.head(5)
```

Out [ ]:

	IncidentFromDate	IncidentToDate	IncidentToTime	Offense Description	CaseDisposition	LocationCo
0	1/1/2010 1:55:00	1/1/2010	2:58:00	Miscellaneous Offenses	NaN	ONCAMF
1	1/1/2010 1:36:00	NaN	NaN	Miscellaneous Offenses	NaN	NONC
2	1/1/2010 2:00:00	1/1/2010	4:00:00	Larceny – From Building	NaN	NONC
3	1/2/2010 6:15:00	NaN	NaN	Miscellaneous Offenses	NaN	ONC
4	1/3/2010 1:29:00	NaN	NaN	Miscellaneous Offenses	Prosecution Declined	F

In [ ]:

```
#convert incidentfromdate to datetime
df['IncidentFromDate'] = pd.to_datetime(df['IncidentFromDate'])

#replace Offense Description value 1 with no offense description
df['Offense Description'] = df['Offense Description'].replace(1, 'no offense de

#using plotly, plot the bar graph of the number of crimes per year, and count
fig = px.histogram(df, x=df['IncidentFromDate'].dt.year, text_auto=True, color_
fig.update_layout(bargap=0.2 )
fig.write_image('CrimePerYear.png')
```

In [ ]:

```
# filter by crime type received from 'Offense Description'
df['Offense Description'] = df['Offense Description'].str.lower()

# change nan with no offense description
df.loc[df['Offense Description'].isna(), 'Offense Description'] = 'no offense c

# replace specials with -
df['Offense Description'] = df['Offense Description'].str.replace('0', '-')

Crime_Type_List = sorted(list(set(df['Offense Description'].values.tolist())))
print(Crime_Type_List)
```

['1', 'aggravated assault', 'aggravated assault/battery – gun', 'aggravated assault/battery – other weapon', 'aggravated assault/battery – gun', 'aggravated assault/battery – other weapon', 'all other larceny', 'all other offenses', 'amphetamine – possession of', 'amphetamine – sales of', 'amphetamine – possession of', 'arson', 'arson offense free text', 'assault/battery offense free text', 'bad checks', 'bomb threat', 'burglary – forced entry – nonresidence', 'burglary – forced entry – residence', 'burglary – forced entry – nonresidence', 'burglary – forced entry – residence', 'burglary – no forced entry – residence', 'burglary – no forced entry – nonresidence', 'burglary – no forced entry – residence', 'burglary – no forced entry – nonresidence', 'burglary/breaking & entering', 'carrying a concealed weapon', 'carrying a prohibited weapon', 'child neglect', 'cocaine – possession of', 'cocaine free text', 'cocaine – possession of', 'contributing to the delinquency of a minor', 'counterfeiting free text', 'counterfeiting/forgery', 'credit card/automatic teller machine fraud', 'criminal trespass', 'cruelty to animals', 'curfew violation', 'curfew/loitering/vagrancy violations', 'damage to property – business', 'damage to property – private', 'damage to property – public', 'damage to property offense free text', 'damage to property – business', 'damage to property – private', 'damage to property – public', 'dangerous drugs offense free text', 'destruction/damage/vandalism of property', 'disorderly conduct', 'disturbing the peace', 'driving under the influence', 'driving under the influence of alcohol', 'driving under the influence of drugs', 'drug equipment violations', 'drug/narcotic violations', 'drunkenness', 'embezzlement', 'embezzlement/theft of conversion – business property', 'embezzlement/theft of conversion – business property', 'escape', 'espionage', 'extortion offense free text', 'extortion/blackmail', 'false fire alarm', 'false pretenses/swindle/confidence game', 'family offense free text', 'family offenses', 'firing a weapon', 'flight to avoid prosecution/confinement/wanted person', 'fondling', 'forgery offense free text', 'forgery – of check', 'forgery – of other object', 'fraud – bad/worthless check', 'fraud – illegal use of credit/financial transaction cards', 'fraud – impersonation', 'fraud – swindle', 'fraud – bad/worthless check', 'fraud – illegal use of credit/financial transaction cards', 'fraud – impersonation', 'fraud – swindle', 'fraudulent activity offense free', 'habitual violator offense free text', 'hacking/computer invasion', 'hallucinogen – possession of', 'hallucinogen – distribution of', 'hallucinogen – possession of', 'harassing communication', 'heroin free text', 'heroin – possession of', 'identity theft', 'impersonation', 'indecent exposure', 'intimidation', 'invasion of privacy offense free text', 'kidnapping offense free text', 'kidnapping/abduction', 'larceny – articles from vehicle', 'larceny – bicycle', 'larceny – from building', 'larceny – from coin machine', 'larceny – from mail', 'larceny – of government property', 'larceny – parts from vehicle', 'larceny – purse snatching', 'larceny offense free text', 'larceny – articles from vehicle', 'larceny – bicycle', 'larceny – from building', 'larceny – from coin machine', 'larceny – from mail', 'larceny – of government property', 'larceny – parts from vehicle', 'larceny – sudden snatching', 'liquor – possession of', 'liquor law violation offense free text', 'liquor law violations', 'liquor – manufacturing', 'liquor – possession of', 'liquor – sales of', 'loitering violation', 'making a false report/giving false information', 'marijuana – possession of', 'marijuana – sales of', 'marijuana free text', 'marijuana – possession of', 'marijuana – sales of', 'marijuana – smuggling', 'miscellaneous offenses', 'misrepresentation of age', 'motor vehicle theft', 'no offense description', 'non-crime (not reported to state)', 'obscene communication', 'obscenity offense free text', 'obstructing justice', 'obstructing officer/criminal investigation', 'obstructing police offense free text', 'opium or derivative – possession of', 'opium or derivative – possession of', 'peeping tom', 'pocket picking', 'pornography/obscene material', 'possession of burglary tools', 'possession of drug related items/narcotic equipment', 'possession of explosives/fireworks', 'possession of other weapon', 'posse

ssion of stolen property', 'possession of stolen vehicle', 'prostitution', 'public peace offense free text', 'rape', 'rape - strongarm', 'rape - other weapon', 'rape - strongarm', 'resisting officer/arrest', 'robbery', 'robbery - residence - strongarm', 'robbery - street - gun', 'robbery - street - strongarm', 'robbery offense free text', 'robbery - business - strongarm', 'robbery - residence - strongarm', 'robbery - residence - gun', 'robbery - street - gun', 'robbery - street - strongarm', 'robbery - street - other weapon', 'runaway juvenile', 'sex offense free text', 'sexual assault offense free text', 'shoplifting', 'simple assault', 'simple assault/battery', 'sodomy', 'statutory rape', 'stolen property offenses', 'synthetic narcotic - possession of', 'synthetic narcotic - possession of', 'terroristic threats/intimidation', 'theft by receiving stolen property', 'theft by receiving stolen vehicle', 'theft from building', 'theft from coin-operated machine or device', 'theft from motor vehicle', 'theft of automobile', 'theft of motor vehicle parts or accessories', 'theft of other vehicle', 'theft of truck/van/bus', 'traffic offense (describe offense)', 'trespass of real property', 'vagrancy offense free text', 'weapon law violations', 'weapons violation offense free text', 'wire fraud']

```
In [ ]: # map year 2021 based on crimetype

# for this data to be valuable, get rid of the crimetype 'non-crime', no offense

df1 = df[df['Offense Description'] != 'non-crime (not reported to state)']
df1 = df1[df1['Offense Description'] != 'no offense description']
df1 = df1[df1['Offense Description'] != 'non-crime']
df1 = df1[df1['Offense Description'] != 'all other offenses']
df1 = df1[df1['Offense Description'] != 'miscellaneous offenses']

fig2 = px.histogram(df1, x=df1['Offense Description'], color_discrete_sequence=
fig2.update_layout(
    margin=dict(l=20, r=20, t=20, b=20),
    paper_bgcolor="LightSteelBlue",
)

fig2.write_image('OffenseType.png')
```



