

Import pandas and required dataset csv

```
import pandas as pd
```

```
data = pd.read_csv("/workspace/nvidia-examples/cnn/data/NYPD_Shooting_Inci
```

Show the values of the dataset

```
data.head()
```



	INCIDENT_KEY	OCCUR_DATE	OCCUR_TIME	BORO	PRECINCT	JURISDICTION_CODE	I
0	24050482	08/27/2006	05:35:00	BRONX	52		0.0
1	77673979	03/11/2011	12:03:00	QUEENS	106		0.0
2	203350417	10/06/2019	01:09:00	BROOKLYN	77		0.0
3	80584527	09/04/2011	03:35:00	BRONX	40		0.0
4	90843766	05/27/2013	21:16:00	QUEENS	100		0.0

Broader scope of the dataset

```
date = data.sort_values(by=['OCCUR_DATE'])
date.max
```

```
<bound method NDFrame.add_numeric_operations.<locals>.max of
INCIDENT_KEY  OCCUR_DATE  OCCUR_TIME  BORO  PRECINCT  \
4437          9953247  01/01/2006  03:30:00  BROOKLYN    67
16838         139716503  01/01/2006  12:30:00  BROOKLYN    77
14553          9953246  01/01/2006  05:51:00    BRONX    44
21448          9953250  01/01/2006  02:34:00   QUEENS   114
21506          9953245  01/01/2006  02:00:00    BRONX    48
...          ...          ...          ...    ...    ...
12562         206891917  12/31/2019  20:14:00  BROOKLYN    73
22540         222473262  12/31/2020  23:45:00  MANHATTAN    33
```

13915	222466833	12/31/2020	19:27:00	QUEENS	113
11511	222468112	12/31/2020	14:59:00	QUEENS	103
6996	222446417	12/31/2020	00:42:00	BRONX	44

	JURISDICTION_CODE	LOCATION_DESC	STATISTICAL_MURDER_FLAG	\
4437	0.0	NaN	False	
16838	0.0	PVT HOUSE	True	
14553	0.0	NONE	False	
21448	0.0	BAR/NIGHT CLUB	True	
21506	0.0	NONE	False	
...	
12562	0.0	NaN	False	
22540	0.0	NaN	False	
13915	0.0	NaN	True	
11511	0.0	NaN	False	
6996	0.0	SMALL MERCHANT	False	

	PERP_AGE_GROUP	PERP_SEX	PERP_RACE	VIC_AGE_GROUP	VIC_SEX	\
4437	UNKNOWN	U	UNKNOWN	18-24	M	
16838	NaN	NaN	NaN	25-44	M	
14553	25-44	M	WHITE HISPANIC	18-24	M	
21448	25-44	M	BLACK	25-44	M	
21506	18-24	M	BLACK	<18	M	
...	
12562	NaN	NaN	NaN	25-44	M	
22540	25-44	M	BLACK HISPANIC	18-24	M	
13915	NaN	NaN	NaN	25-44	M	
11511	18-24	M	BLACK	18-24	M	
6996	NaN	NaN	NaN	25-44	M	

	VIC_RACE	X_COORD_CD	Y_COORD_CD	Latitude	Longitude	\
4437	BLACK	9.993160e+05	176460.000000	40.651014	-73.945707	
16838	BLACK	9.964416e+05	184160.359375	40.672154	-73.956052	
14553	WHITE HISPANIC	1.007418e+06	243859.218750	40.835990	-73.916276	
21448	BLACK	1.002577e+06	221583.453125	40.774861	-73.933833	
21506	BLACK	1.013405e+06	251800.750000	40.857770	-73.894607	
...	
12562	BLACK	1.007235e+06	179931.906250	40.660527	-73.917156	
22540	BLACK HISPANIC	9.999233e+05	244406.562500	40.837509	-73.943360	
13915	BLACK	1.041343e+06	189493.562500	40.686617	-73.794141	
11511	BLACK	1.040590e+06	191768.156250	40.692865	-73.796836	
6996	WHITE HISPANIC	1.009943e+06	244586.390625	40.837979	-73.907148	

	Lon_Lat
4437	POINT (-73.94570651699996 40.65101399800005)
16838	POINT (-73.95605150499995 40.67215420900004)
14553	POINT (-73.91627635899994 40.83599040100006)
21448	POINT (-73.93383258499995 40.77486094100004)

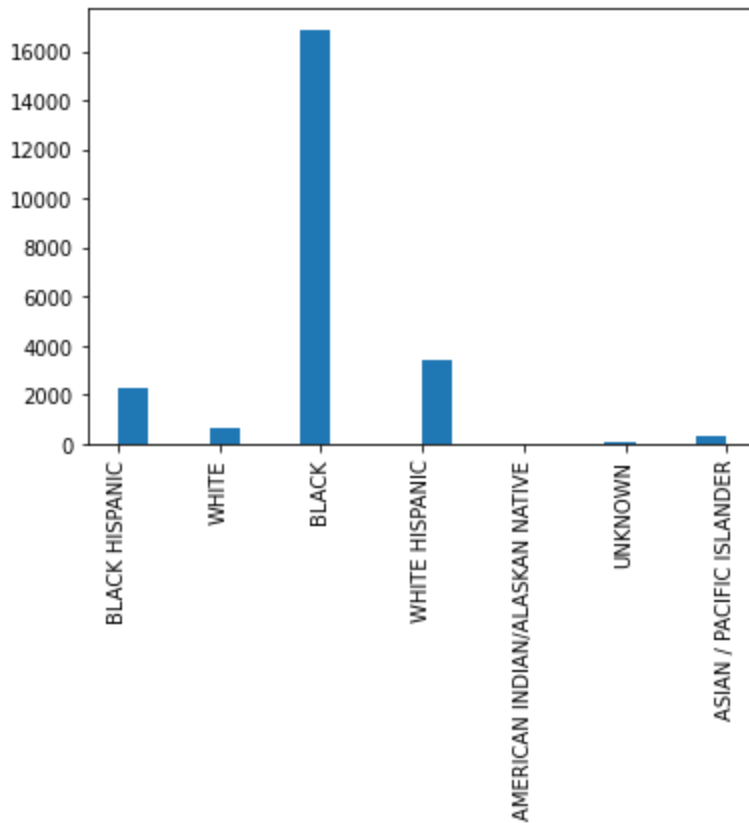
Plot histogram of VIC_RACE

```
import matplotlib.pyplot as plt
import numpy as np
%matplotlib inline
```

```
race = data['VIC_RACE']
```

Plot histogram of the Races

```
p = plt.hist(race, density=False, bins=20)
plt.xticks(rotation='vertical')
plt.show()
```



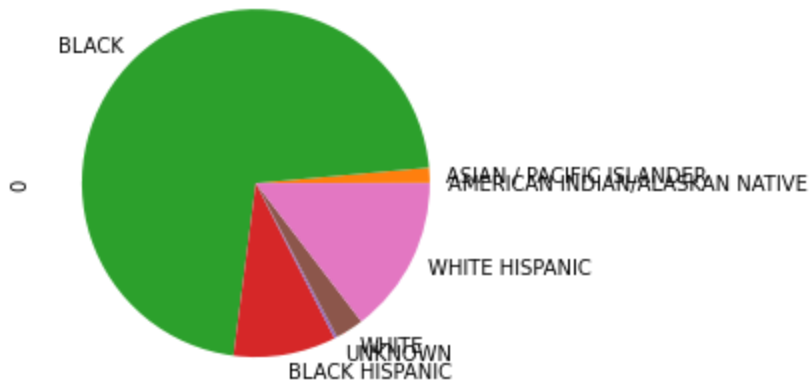
```
pieC = data.groupby(['VIC_RACE']).size()
pieC = pd.DataFrame(pieC)
pieC
```

0

Pie Chart of the Race.

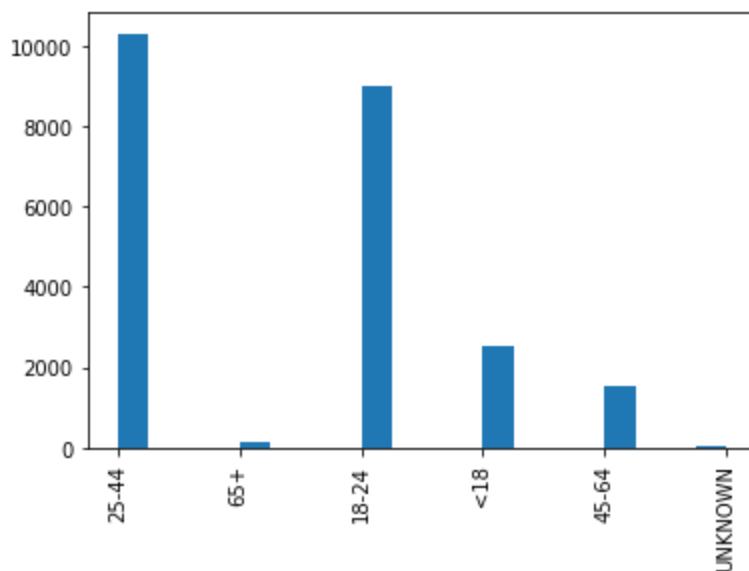
```
pieC.groupby(['VIC_RACE']).sum().plot(kind='pie', y=0, legend=None)
```

<AxesSubplot:ylabel='0'>



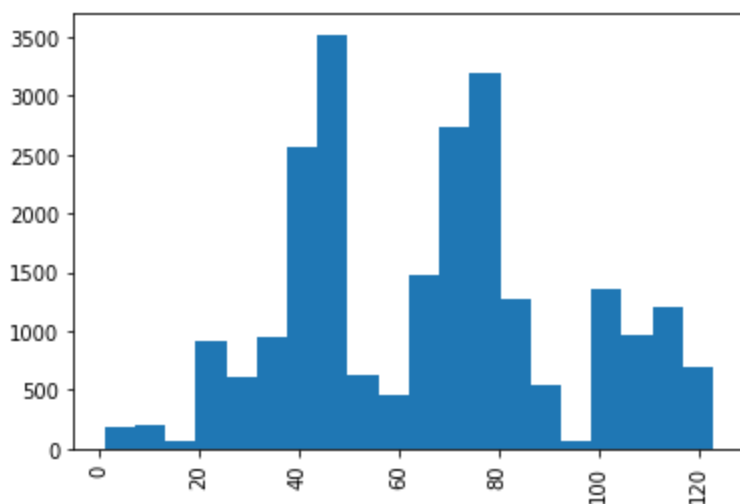
```
age = data['VIC_AGE_GROUP']
```

```
p_age = plt.hist(age, density=False, bins=20)
plt.xticks(rotation='vertical')
plt.show()
```



```
precinct = data['PRECINCT']
```

```
p_precinct = plt.hist(precinct, density=False, bins=20)
plt.xticks(rotation='vertical')
plt.show()
```



```
import pandas as pd
```

```
# Load the dataset
```

```
file_path = './NYPD_Shooting_Incident_Data__Historic_.csv'
shooting_data = pd.read_csv(file_path)
```

```
# Display the first few rows of the dataset
```

```
shooting_data.head()
```

	INCIDENT_KEY	OCCUR_DATE	OCCUR_TIME	BORO	LOC_OF_OCCUR_DESC	PRECINCT	JUI
0	228798151	05/27/2021	21:30:00	QUEENS	NaN	105	
1	137471050	06/27/2014	17:40:00	BRONX	NaN	40	
2	147998800	11/21/2015	03:56:00	QUEENS	NaN	108	
3	146837977	10/09/2015	18:30:00	BRONX	NaN	44	
4	58921844	02/19/2009	22:58:00	BRONX	NaN	47	

5 rows × 21 columns

```

import matplotlib.pyplot as plt
import seaborn as sns

# Convert 'OCCUR_DATE' to datetime format and extract year and month
shooting_data['OCCUR_DATE'] = pd.to_datetime(shooting_data['OCCUR_DATE'])
shooting_data['YEAR'] = shooting_data['OCCUR_DATE'].dt.year
shooting_data['MONTH'] = shooting_data['OCCUR_DATE'].dt.month

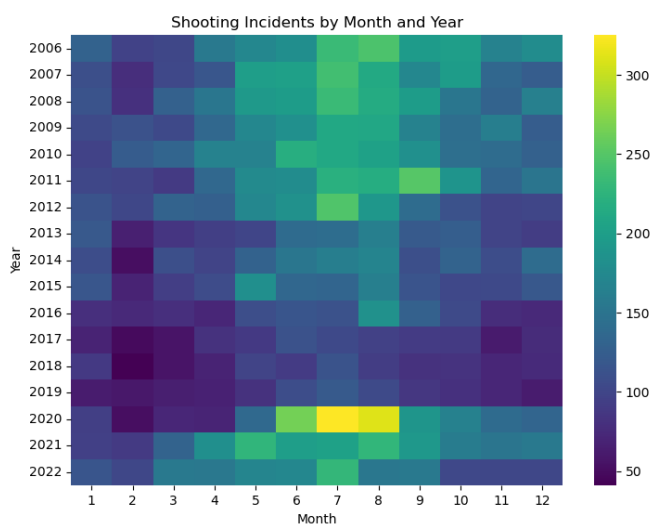
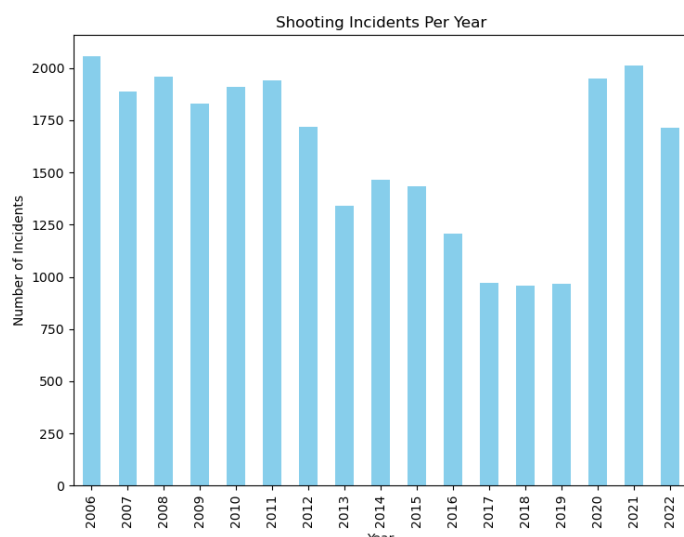
# Aggregate data by year and month
yearly_counts = shooting_data['YEAR'].value_counts().sort_index()
monthly_counts = shooting_data.groupby('YEAR')['MONTH'].value_counts().unstack().fillna(0)

# Plotting
plt.figure(figsize=(15, 6))
plt.subplot(1, 2, 1)
yearly_counts.plot(kind='bar', color='skyblue')
plt.title('Shooting Incidents Per Year')
plt.xlabel('Year')
plt.ylabel('Number of Incidents')

plt.subplot(1, 2, 2)
sns.heatmap(monthly_counts, cmap='viridis')
plt.title('Shooting Incidents by Month and Year')
plt.xlabel('Month')
plt.ylabel('Year')

plt.tight_layout()
plt.show()

```

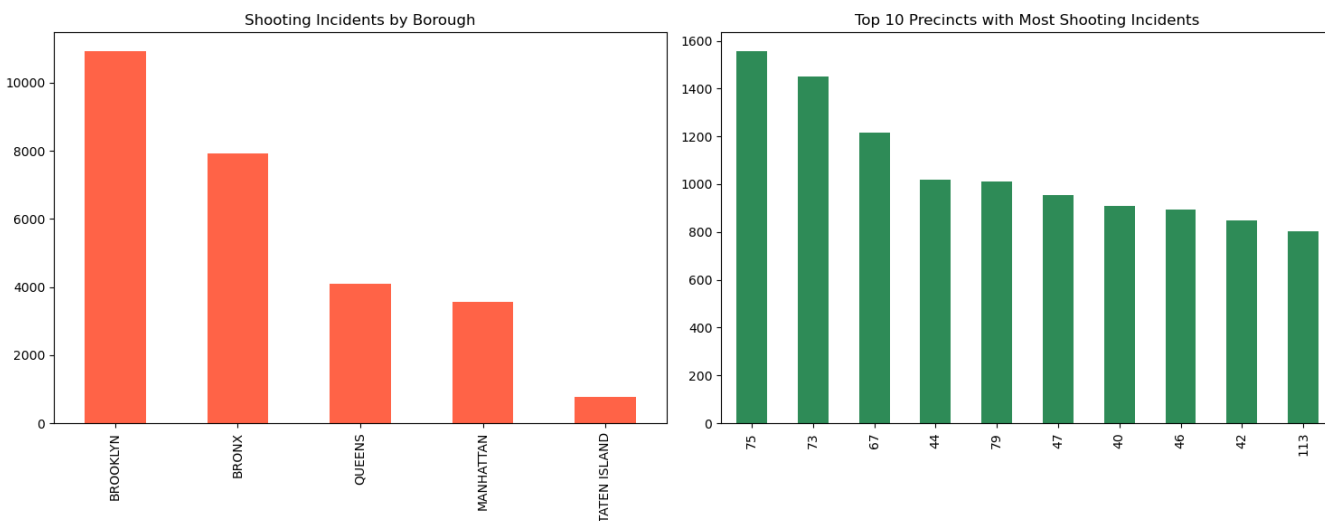


```
# Count of incidents by borough and precinct
borough_counts = shooting_data['BORO'].value_counts()
precinct_counts = shooting_data['PRECINCT'].value_counts().head(10)

# Plotting
plt.figure(figsize=(15, 6))
plt.subplot(1, 2, 1)
borough_counts.plot(kind='bar', color='tomato')
plt.title('Shooting Incidents by Borough')

plt.subplot(1, 2, 2)
precinct_counts.plot(kind='bar', color='seagreen')
plt.title('Top 10 Precincts with Most Shooting Incidents')

plt.tight_layout()
plt.show()
```



```
# Demographic analysis
vic_age_group_counts = shooting_data['VIC_AGE_GROUP'].value_counts()
vic_sex_counts = shooting_data['VIC_SEX'].value_counts()
vic_race_counts = shooting_data['VIC_RACE'].value_counts()

perp_age_group_counts = shooting_data['PERP_AGE_GROUP'].value_counts()
perp_sex_counts = shooting_data['PERP_SEX'].value_counts()
perp_race_counts = shooting_data['PERP_RACE'].value_counts()

# Plotting
fig, axes = plt.subplots(3, 2, figsize=(15, 15))
vic_age_group_counts.plot(kind='bar', ax=axes[0, 0], title='Victim Age Group')
vic_sex_counts.plot(kind='bar', ax=axes[0, 1], title='Victim Sex')
vic_race_counts.plot(kind='bar', ax=axes[1, 0], title='Victim Race')

perp_age_group_counts.plot(kind='bar', ax=axes[1, 1], title='Perpetrator Age Group')
perp_sex_counts.plot(kind='bar', ax=axes[2, 0], title='Perpetrator Sex')
perp_race_counts.plot(kind='bar', ax=axes[2, 1], title='Perpetrator Race')

plt.tight_layout()
plt.show()
```

