

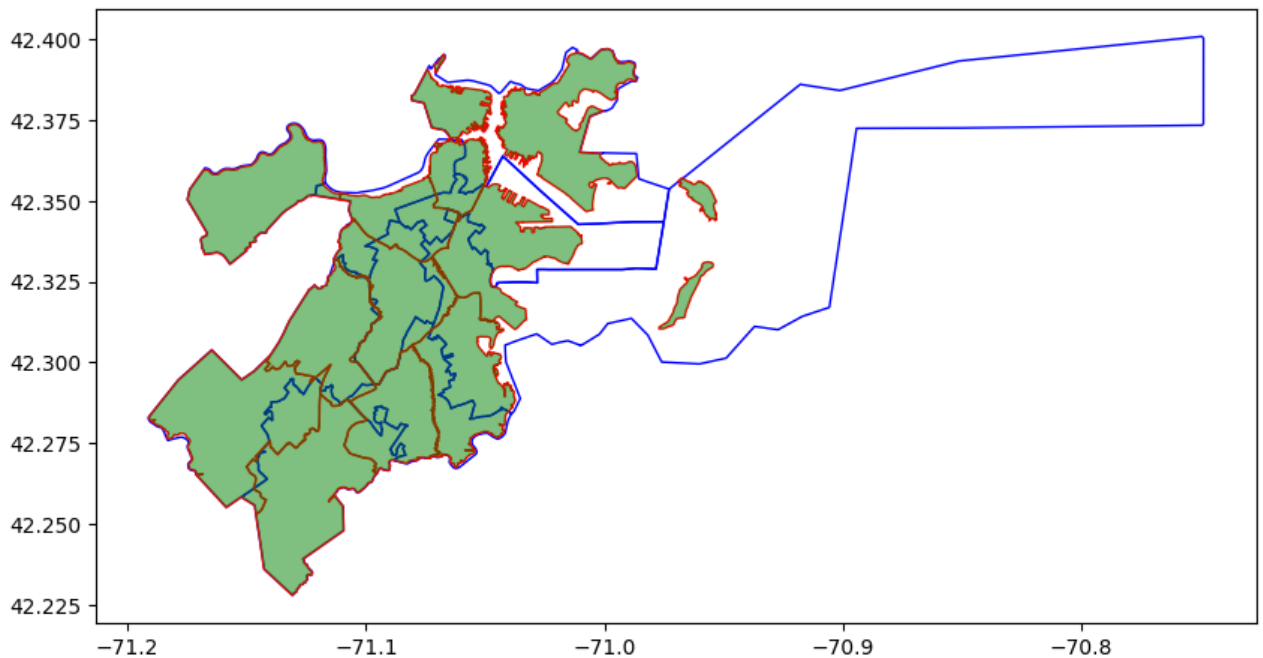
What is the rate of gun violence in District 4? How does this compare to the rest of the city?

```
In [59]: import geopandas as gpd
import matplotlib.pyplot as plt

# Load the shapefiles
city_council = gpd.read_file('City_Council_Districts_-_Effective_for_the_202
BPD = gpd.read_file('Police_Districts/Police_Districts.shp')

# Overlay the two shapefiles
overlay = gpd.overlay(city_council, BPD, how='intersection')

# Plot the result
fig, ax = plt.subplots(figsize=(10, 10))
city_council.plot(ax=ax, facecolor='none', edgecolor='blue')
BPD.plot(ax=ax, facecolor='none', edgecolor='red')
overlay.plot(ax=ax, facecolor='green', alpha = 0.5)
plt.show()
```



In []:

```
In [48]: shooting_incident = df[df['SHOOTING'] == 1]
shooting_incident
```

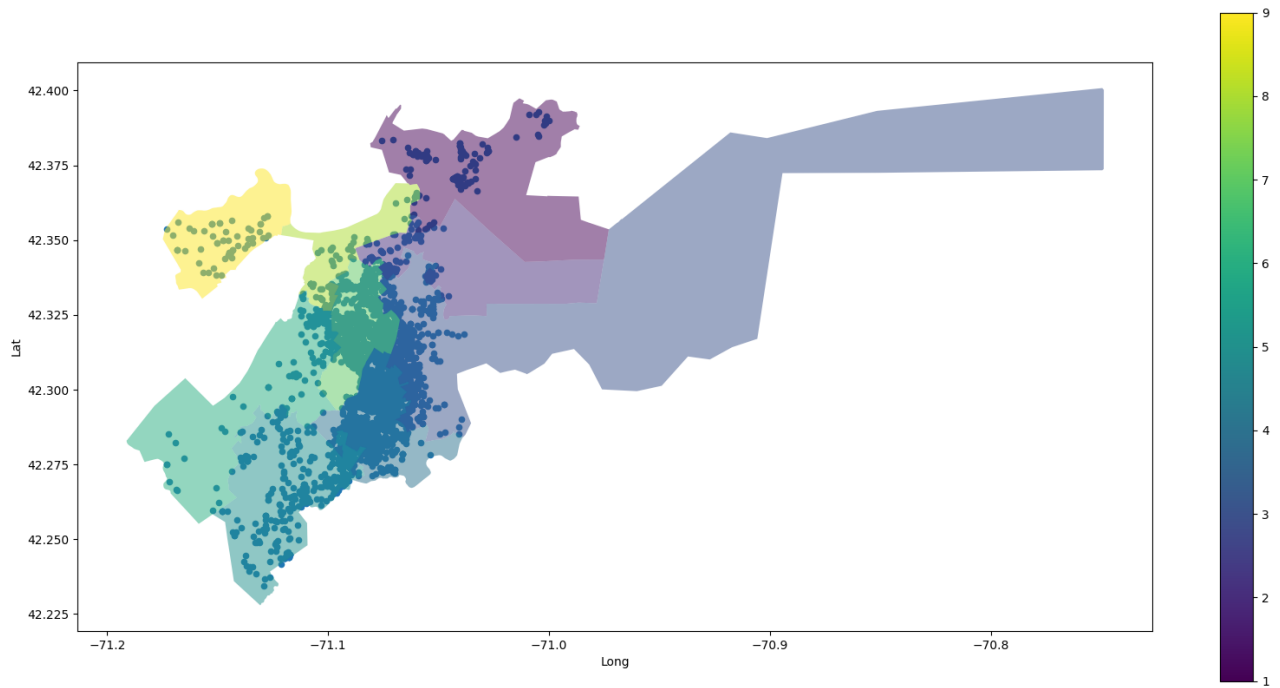
Out [48]:

| | INCIDENT_NUMBER | OFFENSE_CODE | OFFENSE_CODE_GROUP | OFFENSE_DESCRIPTION |
|-------------|-----------------|--------------|--------------------|---|
| 34 | 192000011 | 3114 | NaN | INVESTIGATE PROPERTY |
| 72 | 192000030 | 1402 | NaN | VANDALISM |
| 248 | 192000225 | 3114 | NaN | INVESTIGATE PROPERTY |
| 321 | 192000312 | 423 | NaN | ASSAULT - AGGRAVATED |
| 494 | 192000560 | 3114 | NaN | INVESTIGATE PROPERTY |
| ... | ... | ... | ... | .. |
| 8202 | 232011624 | 2662 | NaN | BALLISTICS EVIDENCE/FOUND |
| 8230 | 232011907 | 111 | NaN | MURDER, NON- NEGLIGENT MANSLAUGHTER |
| 8426 | 232011954 | 3114 | NaN | INVESTIGATE PROPERTY |
| 8449 | 232011974 | 3115 | NaN | INVESTIGATE PERSON |
| 8601 | 232012234 | 3114 | NaN | INVESTIGATE PROPERTY |

3645 rows × 17 columns

```
In [49]: ax = shooting_incident.plot(x="Long", y="Lat", kind="scatter",
    figsize=(20, 10))
    city_council.plot(ax = ax, column = 'DISTRICT', legend = True, figsize=(20,
```

```
Out [49]: <Axes: xlabel='Long', ylabel='Lat'>
```



```
In [75]: district_4 = city_council[city_council['FID'] == 4]
```

```
In [56]: from turfpy.measurement import boolean_point_in_polygon
from geojson import Point, Polygon, Feature

shooting_incident['point'] = [Point(xy) for xy in zip(shooting_incident.Long

district_4.iloc[0]['geometry'].geom_type
geometry = gpd.GeoSeries([district_4.iloc[0]['geometry']]).__geo_interface__
features = geometry['features']
polygon = Polygon(features[0]['geometry']['coordinates'])

shooting_incident['district_4'] = [boolean_point_in_polygon(shooting_incident
```

```

/var/folders/3s/4bvxjynd2_s4dkh__461f5hr0000gn/T/ipykernel_64926/1088657021.
py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/
stable/user_guide/indexing.html#returning-a-view-versus-a-copy
    shooting_incident['point'] = [Point(xy) for xy in zip(shooting_incident.Lo
ng, shooting_incident.Lat)]
/var/folders/3s/4bvxjynd2_s4dkh__461f5hr0000gn/T/ipykernel_64926/1088657021.
py:11: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/
stable/user_guide/indexing.html#returning-a-view-versus-a-copy
    shooting_incident['district_4'] = [boolean_point_in_polygon(shooting_incid
ent.iloc[i]['point'], polygon) for i in range(len(shooting_incident))]

```

```

In [57]: shooting_incident_4 = shooting_incident[shooting_incident['district_4'] == T
shooting_incident_4

```

```

Out[57]:      INCIDENT_NUMBER  OFFENSE_CODE  OFFENSE_CODE_GROUP  OFFENSE_DESCRIPTION

```

| | | | | |
|-------------|-----------|------|-----|-----------------------------------|
| 770 | 192000890 | 3114 | NaN | INVESTIGATE PROPERTY |
| 1255 | 192001459 | 3114 | NaN | INVESTIGATE PROPERTY |
| 2540 | 192003016 | 413 | NaN | ASSAULT - AGGRAVATED - BATTERY |
| 2605 | 192003087 | 2662 | NaN | BALLISTICS EVIDENCE/FOUND |
| 2988 | 192003523 | 2662 | NaN | BALLISTICS EVIDENCE/FOUND |
| ... | ... | ... | ... | ... |

| | | | | |
|-------------|-----------|------|-----|----------------------|
| 7213 | 232010222 | 3114 | NaN | INVESTIGATE PROPERTY |
|-------------|-----------|------|-----|----------------------|

| | | | | |
|-------------|-----------|------|-----|----------------------|
| 7253 | 232010259 | 3114 | NaN | INVESTIGATE PROPERTY |
|-------------|-----------|------|-----|----------------------|

| | | | | |
|-------------|-----------|------|-----|----------------------|
| 8032 | 232011364 | 3114 | NaN | INVESTIGATE PROPERTY |
|-------------|-----------|------|-----|----------------------|

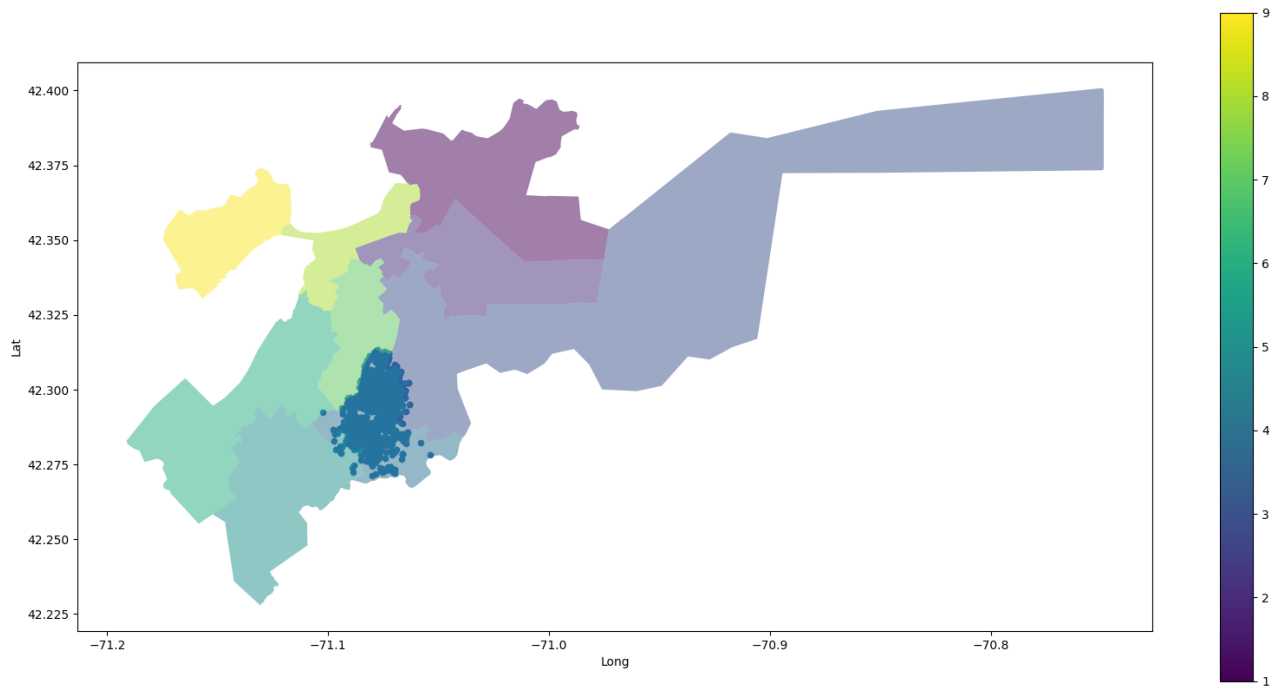
| | | | | |
|-------------|-----------|------|-----|------------------------------|
| 8202 | 232011624 | 2662 | NaN | BALLISTICS EVIDENCE/FOUND |
|-------------|-----------|------|-----|------------------------------|

| | | | | |
|-------------|-----------|------|-----|----------------------|
| 8426 | 232011954 | 3114 | NaN | INVESTIGATE PROPERTY |
|-------------|-----------|------|-----|----------------------|

1289 rows × 19 columns

```
In [58]: ax = shooting_incident_4.plot(x="Long", y="Lat", kind="scatter",
    figsize=(20, 10))
    city_council.plot(ax = ax, column = 'DISTRICT', legend = True, figsize=(20,
```

```
Out[58]: <Axes: xlabel='Long', ylabel='Lat'>
```

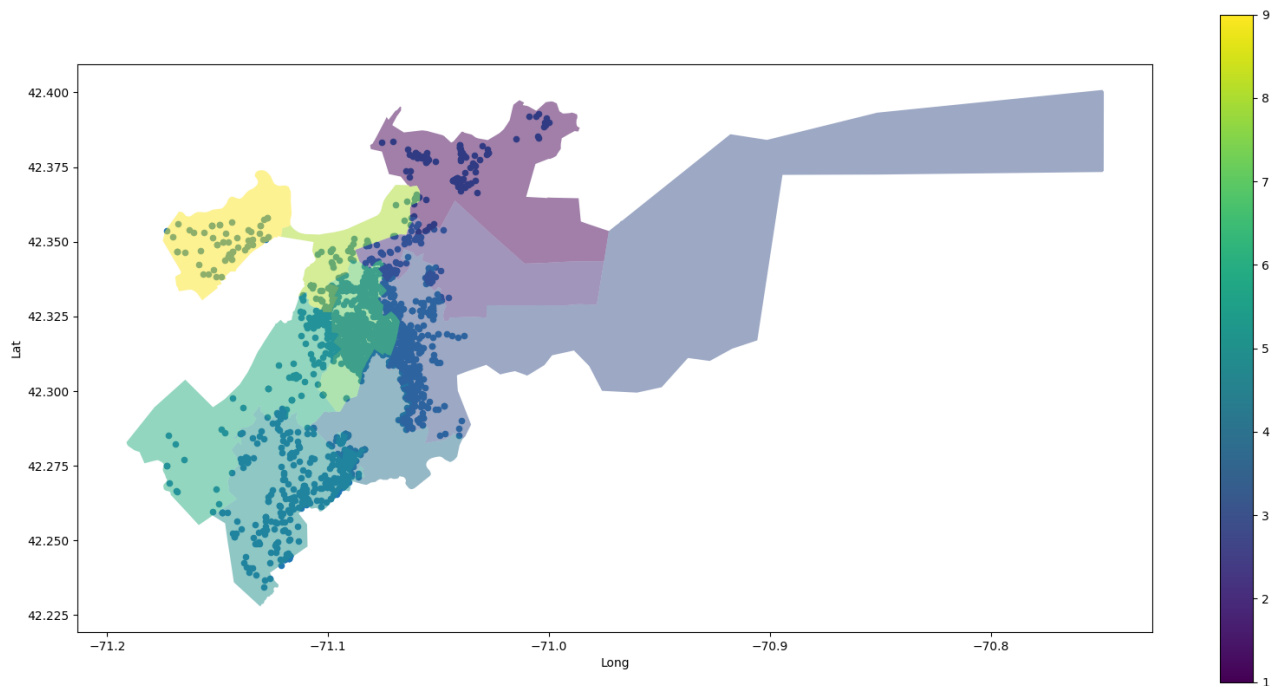


```
In [79]: shootings_not_4 = shooting_incident[shooting_incident['district_4'] == False]

# Plot the scatter plot for all districts except District 4
ax = shootings_not_4.plot(x='Long', y='Lat', kind='scatter', figsize=(20, 10))

# Overlay the city council districts on the scatter plot
city_council.plot(ax=ax, column='DISTRICT', legend=True, figsize=(20, 10), a
```

```
Out[79]: <Axes: xlabel='Long', ylabel='Lat'>
```



```
In [67]: # Calculate the number of incidents in District 4

num_district_4 = len(shooting_incident_4)

# Calculate the total number of incidents in the city
num_city = len(shooting_incident)

# Calculate the rate of gun violence in District 4
rate_district_4 = num_district_4 / num_city

# Print the results
print('Number of incidents in District 4:', num_district_4)
print('Total number of incidents in the city:', num_city)
print('Rate of gun violence in District 4:', rate_district_4)
```

```
Number of incidents in District 4: 1289
Total number of incidents in the city: 3645
Rate of gun violence in District 4: 0.35363511659807956
```

What is the rate of gun violence in District 4?

the rate of gun violence in District 4 is 35.36%

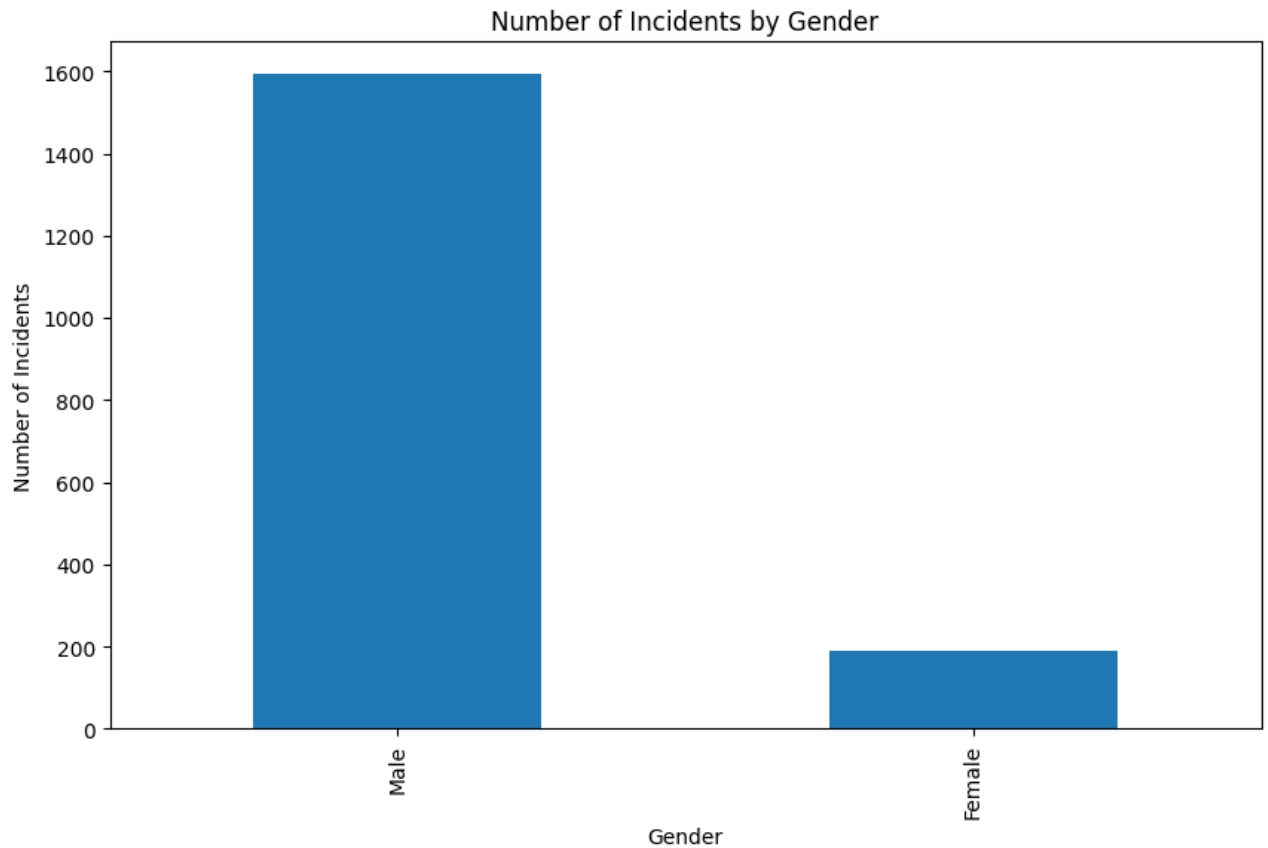
How does this compare to the rest of the city?

we can see from the scatter graph above, district 4 is denser than the other area, so the rate of gun violence in District 4 is much higher than the overall rate of gun violence in the city.

Other analysis for the rest of data:

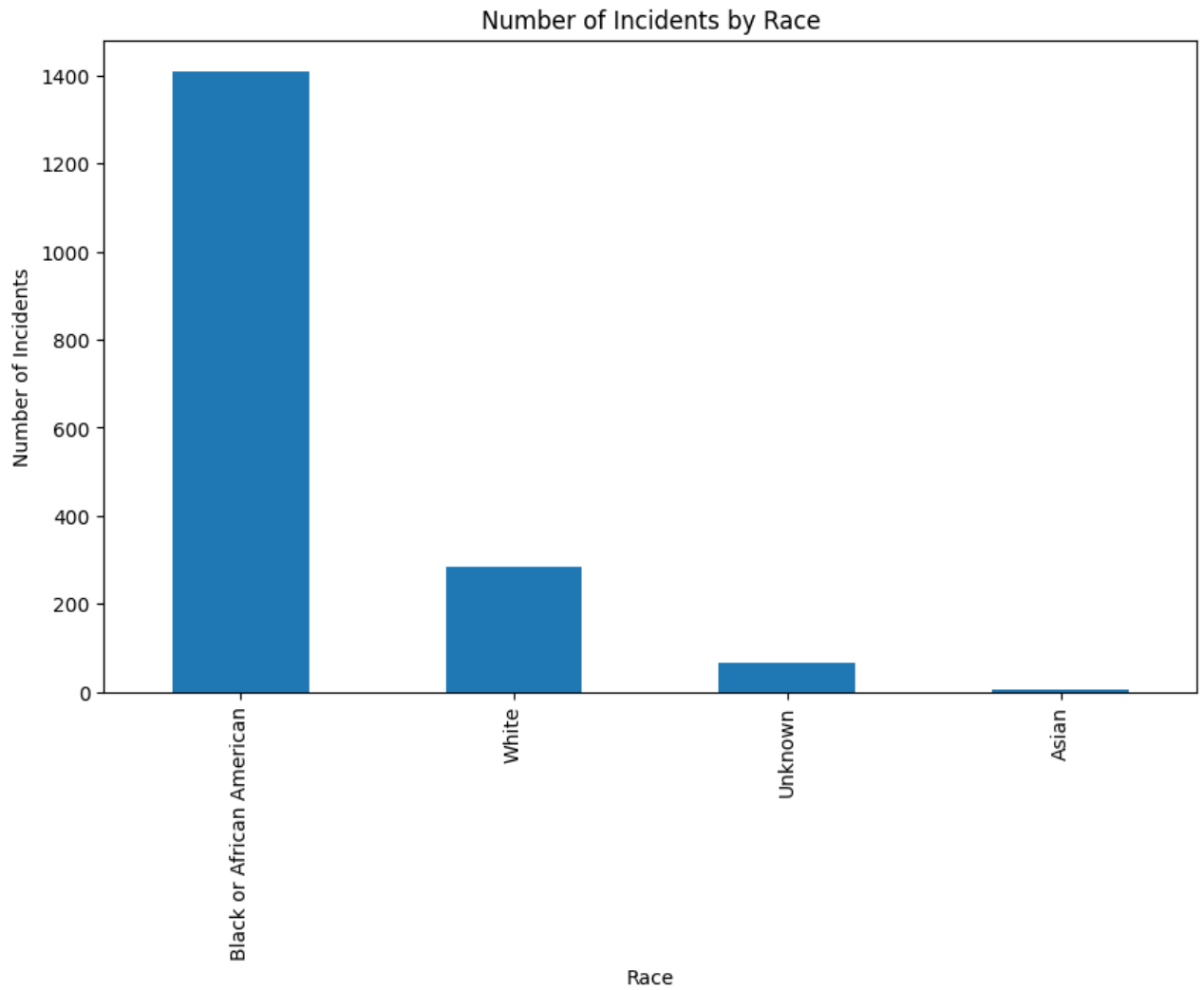
```
In [82]: shootings = pd.read_csv('SHOOTINGS.csv')
shootings['victim_gender'].value_counts().plot(kind='bar',figsize=(10, 6)).s

Out[82]: [Text(0.5, 1.0, 'Number of Incidents by Gender'),
Text(0.5, 0, 'Gender'),
Text(0, 0.5, 'Number of Incidents')]
```



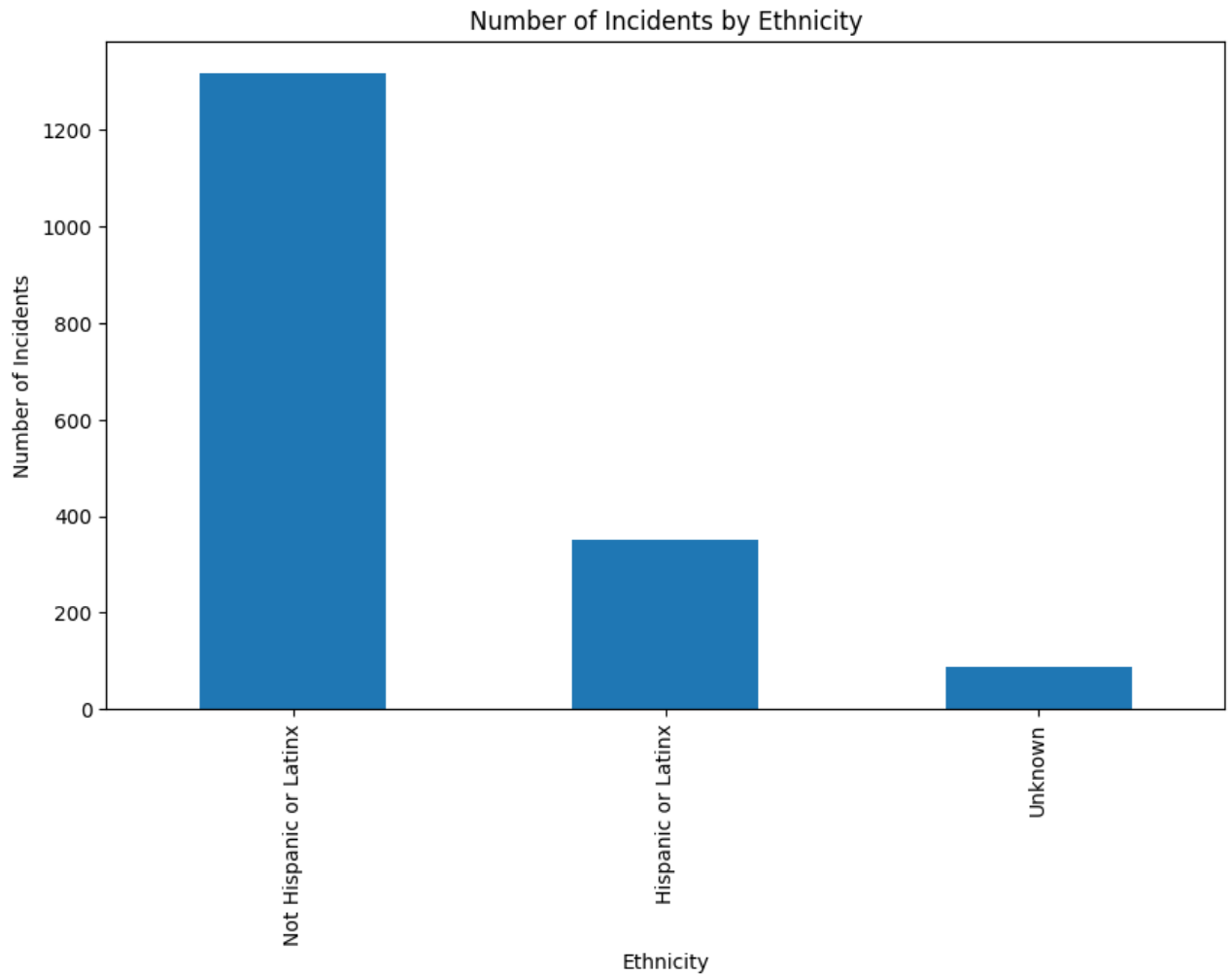
```
In [83]: shootings['victim_race'].value_counts().plot(kind='bar',figsize=(10, 6)).set
```

```
Out[83]: [Text(0.5, 1.0, 'Number of Incidents by Race'),  
Text(0.5, 0, 'Race'),  
Text(0, 0.5, 'Number of Incidents')]
```

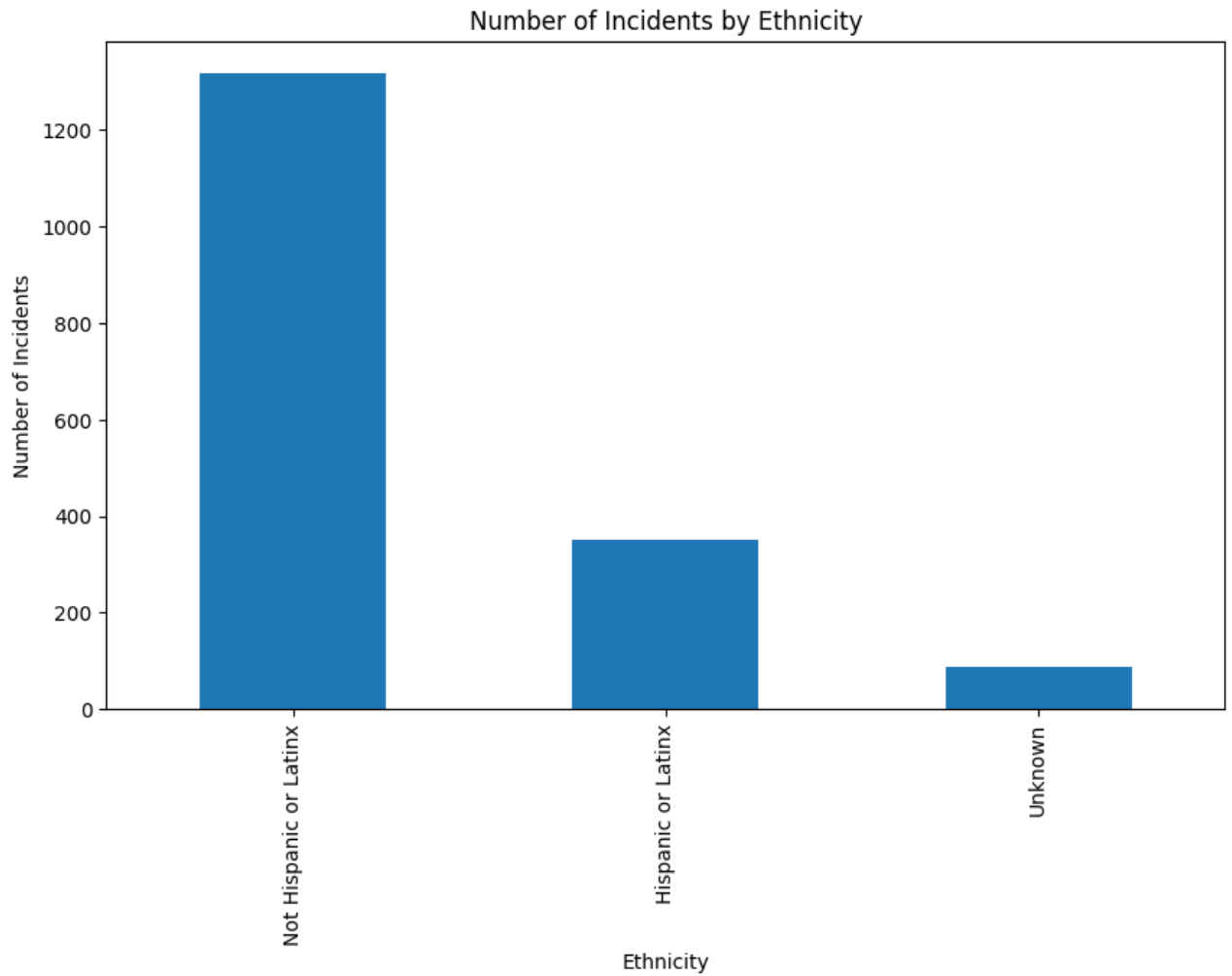
```
In [84]: shootings['victim_ethnicity_NIBRS'].value_counts().plot(kind='bar',figsize=(
```

```
Out[84]: [Text(0.5, 1.0, 'Number of Incidents by Ethnicity'),  
Text(0.5, 0, 'Ethnicity'),  
Text(0, 0.5, 'Number of Incidents')]
```



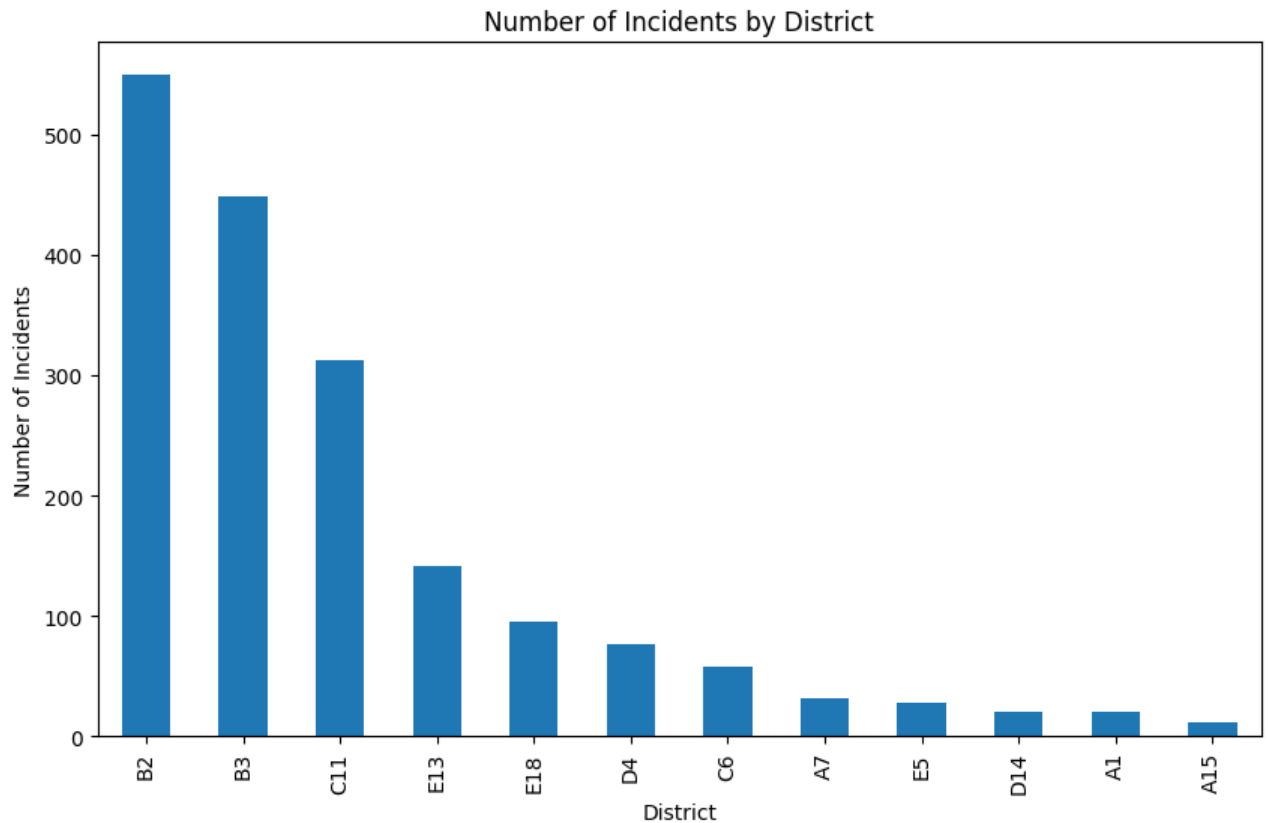
```
In [87]: shootings['victim_ethnicity_NIBRS'].value_counts().plot(kind='bar',figsize=(
```

```
Out[87]: [Text(0.5, 1.0, 'Number of Incidents by Ethnicity'),  
Text(0.5, 0, 'Ethnicity'),  
Text(0, 0.5, 'Number of Incidents')]
```



```
In [88]: shootings['district'].value_counts().plot(kind='bar',figsize=(10, 6)).set(ti
```

```
Out[88]: [Text(0.5, 1.0, 'Number of Incidents by District'),  
Text(0.5, 0, 'District'),  
Text(0, 0.5, 'Number of Incidents')]
```



```
In [86]: district_rates = shootings['district'].value_counts(normalize=True)

# Create a bar plot to compare the rates
district_rates.plot(kind='bar').set(title='Gun Violence Rates by District',

# Add text labels to the bars
for i, v in enumerate(district_rates):
    plt.text(i - 0.1, v + 0.006, str(round(v * len(shootings))), fontsize=8)
```

