

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
import plotly.graph_objects as go
```

```
In [2]: import warnings
warnings.filterwarnings('ignore')
```

```
In [5]: df = pd.read_csv(r"E:\PYTHON DS\PROJECTS\Palestine Body Count.csv")
```

```
In [6]: df.head()
```

```
Out[6]:
```

	Year	Month	Palestinians Injuries	Israelis Injuries	Palestinians Killed	Israelis Killed
0	2000.0	DECEMBER	781	NaN	51	8
1	2000.0	NOVEMBER	3838	NaN	112	22
2	2000.0	OCTOBER	5984	NaN	104	10
3	2000.0	SEPTEMBER	NaN	NaN	16	1
4	2001.0	DECEMBER	304	NaN	67	36

```
In [7]: df.tail()
```

```
Out[7]:
```

	Year	Month	Palestinians Injuries	Israelis Injuries	Palestinians Killed	Israelis Killed
246	2021.0	MARCH	NaN	NaN	4	0
247	2021.0	APRIL	NaN	NaN	1	0
248	2021.0	MAY	NaN	NaN	26	3
249	NaN	NaN	NaN	NaN	NaN	NaN
250	NaN	NaN	111,475	5,160	10,000	1,275

```
In [8]: df.shape
```

```
Out[8]: (251, 6)
```

```
In [9]: df.columns
```

```
Out[9]: Index(['Year', 'Month', 'Palestinians Injuries', 'Israelis Injuries',  
              'Palestinians Killed', 'Israelis Killed'],  
             dtype='object')
```

```
In [10]: df.duplicated().sum()
```

```
Out[10]: 0
```

```
In [11]: df.isnull().sum()
```

```
Out[11]: Year                2  
Month                2  
Palestinians Injuries    55  
Israelis Injuries       118  
Palestinians Killed      1  
Israelis Killed          1  
dtype: int64
```

```
In [12]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 251 entries, 0 to 250  
Data columns (total 6 columns):  
#   Column                Non-Null Count  Dtype  
---  -  
0   Year                  249 non-null   float64  
1   Month                 249 non-null   object  
2   Palestinians Injuries 196 non-null   object  
3   Israelis Injuries     133 non-null   object  
4   Palestinians Killed    250 non-null   object  
5   Israelis Killed        250 non-null   object  
dtypes: float64(1), object(5)  
memory usage: 11.9+ KB
```

```
In [13]: df = df.fillna({
    'Palestinians Injuries': '0',
    'Israelis Injuries': '0'
})
```

```
In [14]: numerical_columns = ['Palestinians Injuries', 'Israelis Injuries', 'Palestinians Killed', 'Israelis Killed']
df[numerical_columns] = df[numerical_columns].replace({' ': ''}, regex=True)
df[numerical_columns] = df[numerical_columns].apply(pd.to_numeric, errors='coerce')
```

```
In [15]: df
```

```
Out[15]:
```

	Year	Month	Palestinians Injuries	Israelis Injuries	Palestinians Killed	Israelis Killed
0	2000.0	DECEMBER	781.0	0.0	51.0	8.0
1	2000.0	NOVEMBER	3838.0	0.0	112.0	22.0
2	2000.0	OCTOBER	5984.0	0.0	104.0	10.0
3	2000.0	SEPTEMBER	0.0	0.0	16.0	1.0
4	2001.0	DECEMBER	304.0	0.0	67.0	36.0
...
246	2021.0	MARCH	0.0	0.0	4.0	0.0
247	2021.0	APRIL	0.0	0.0	1.0	0.0
248	2021.0	MAY	0.0	0.0	26.0	3.0
249	NaN	NaN	0.0	0.0	NaN	NaN
250	NaN	NaN	111475.0	5160.0	10000.0	1275.0

251 rows × 6 columns

In [16]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 251 entries, 0 to 250
Data columns (total 6 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Year                  249 non-null   float64
1   Month                 249 non-null   object
2   Palestinians Injuries 249 non-null   float64
3   Israelis Injuries     249 non-null   float64
4   Palestinians Killed   250 non-null   float64
5   Israelis Killed       250 non-null   float64
dtypes: float64(5), object(1)
memory usage: 11.9+ KB
```

In [17]: df.describe()

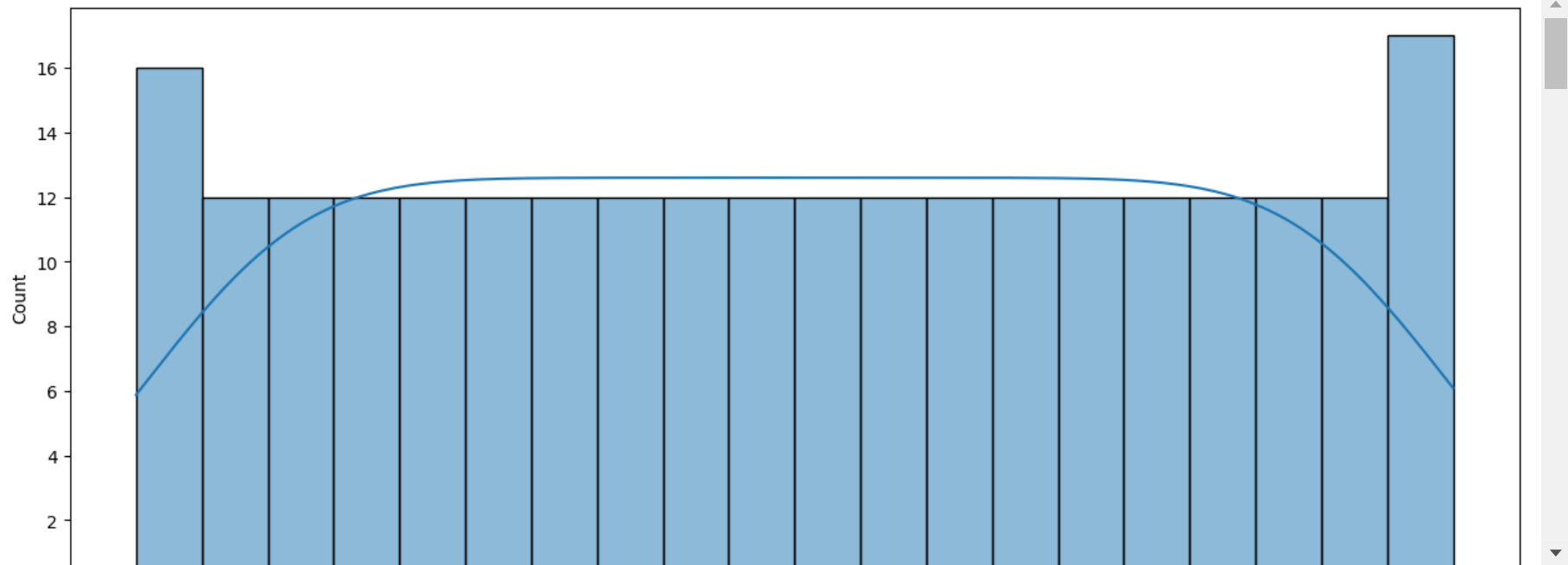
Out[17]:

	Year	Palestinians Injuries	Israelis Injuries	Palestinians Killed	Israelis Killed
count	249.000000	249.000000	249.000000	250.000000	250.000000
mean	2010.542169	895.381526	41.445783	80.000000	10.200000
std	6.014702	7187.086268	358.476500	642.966593	81.151898
min	2000.000000	0.000000	0.000000	0.000000	0.000000
25%	2005.000000	60.000000	0.000000	4.000000	0.000000
50%	2011.000000	161.000000	3.000000	12.000000	1.000000
75%	2016.000000	303.000000	15.000000	37.000000	5.000000
max	2021.000000	111475.000000	5160.000000	10000.000000	1275.000000

```
In [18]: df.nunique()
```

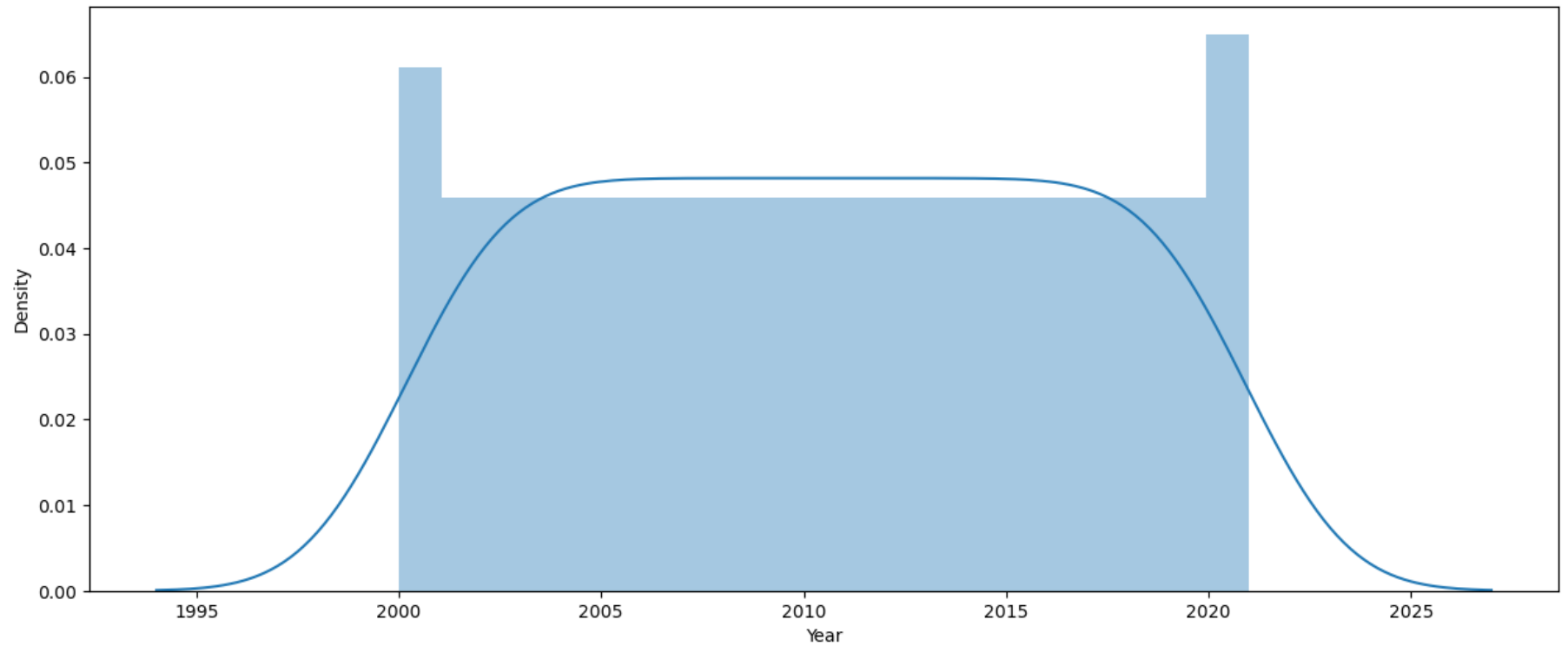
```
Out[18]: Year                22  
Month                14  
Palestinians Injuries    170  
Israelis Injuries        54  
Palestinians Killed      78  
Israelis Killed          34  
dtype: int64
```

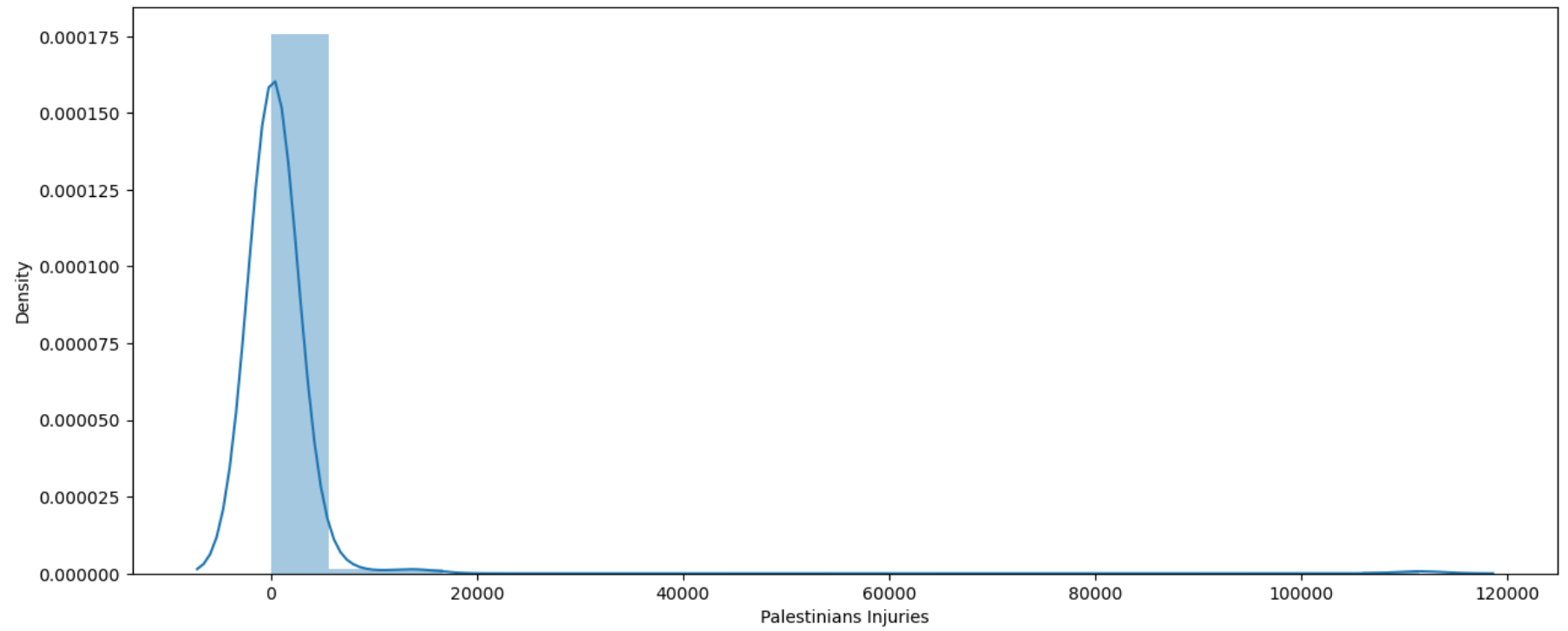
```
In [19]: for i in df.columns:  
plt.figure(figsize=(15,6))  
sns.histplot(df[i], kde = True, bins = 20, palette = 'hls')  
plt.xticks(rotation = 90)  
plt.show()
```

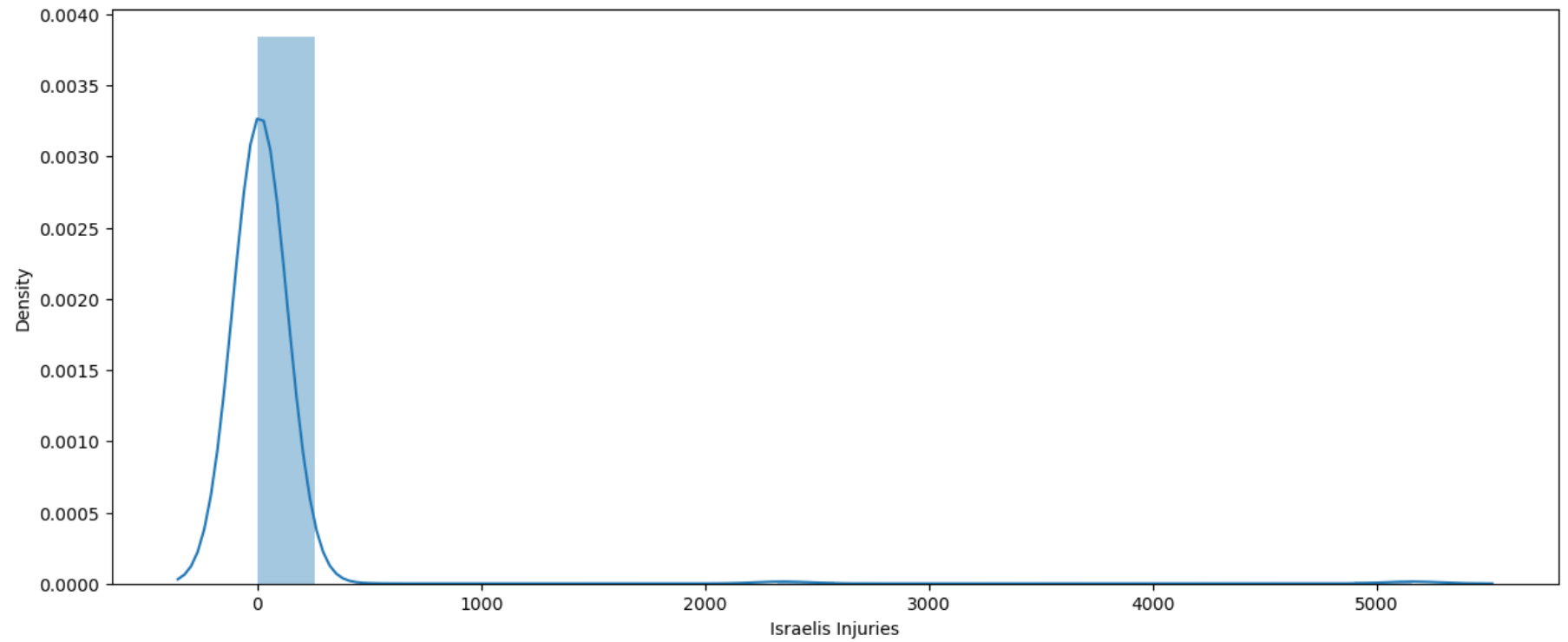


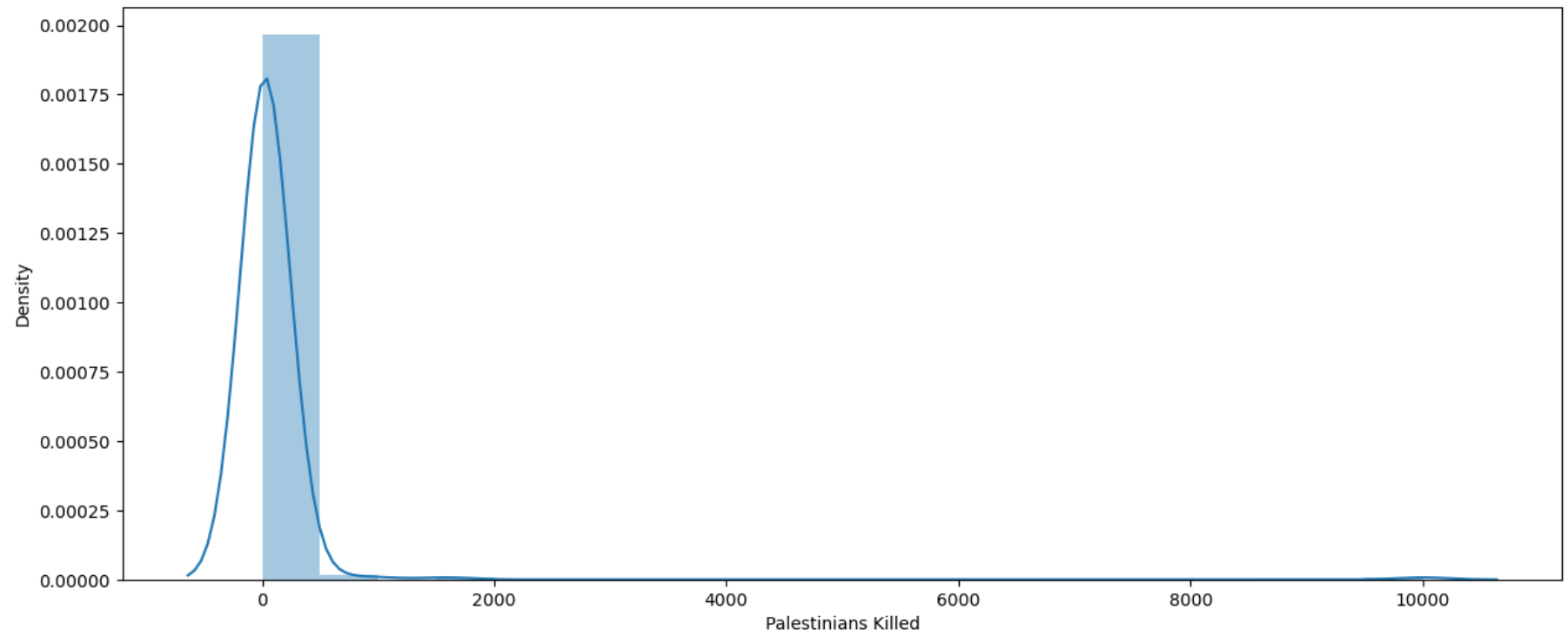
```
In [29]: df1 = df[['Year', 'Palestinians Injuries', 'Israelis Injuries', 'Palestinians Killed',  
                'Israelis Killed']]
```

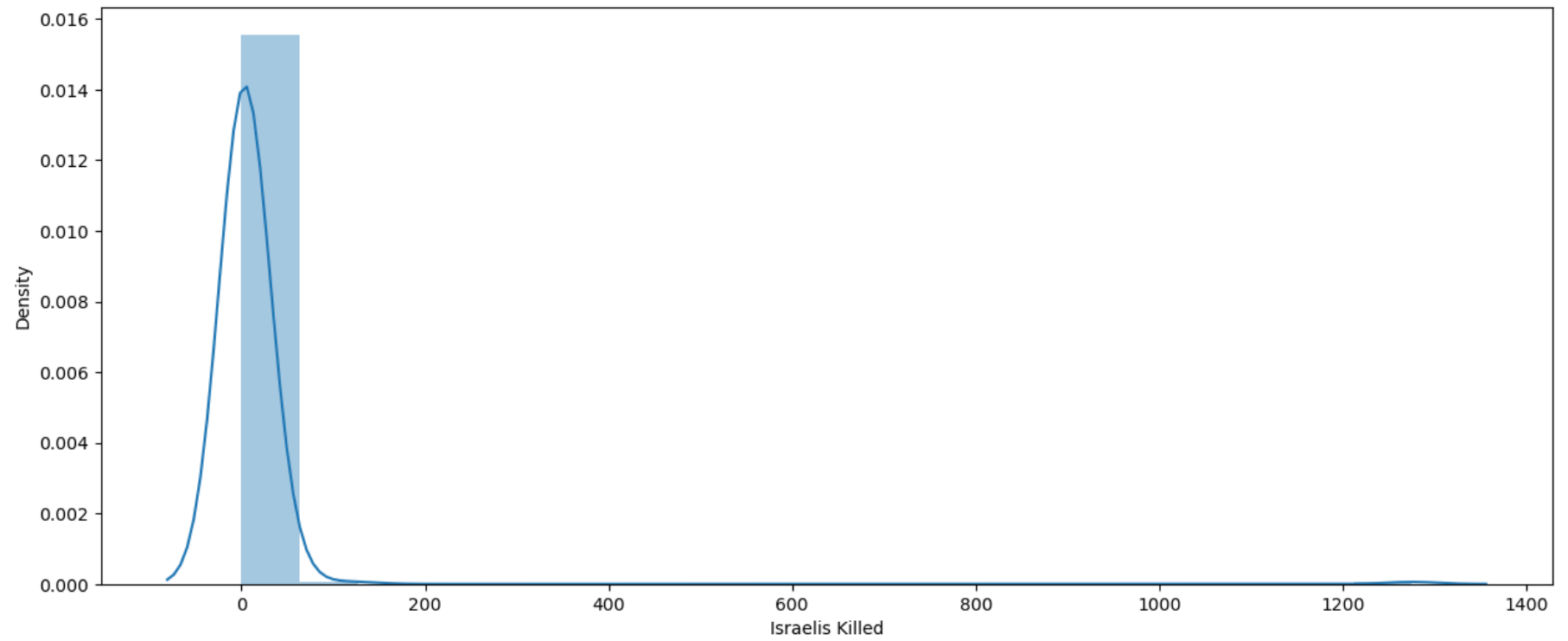
```
In [30]: for i in df1.columns:  
    plt.figure(figsize=(15,6))  
    sns.distplot(df1[i], kde = True, bins = 20)  
    plt.xticks(rotation = 0)  
    plt.show()
```



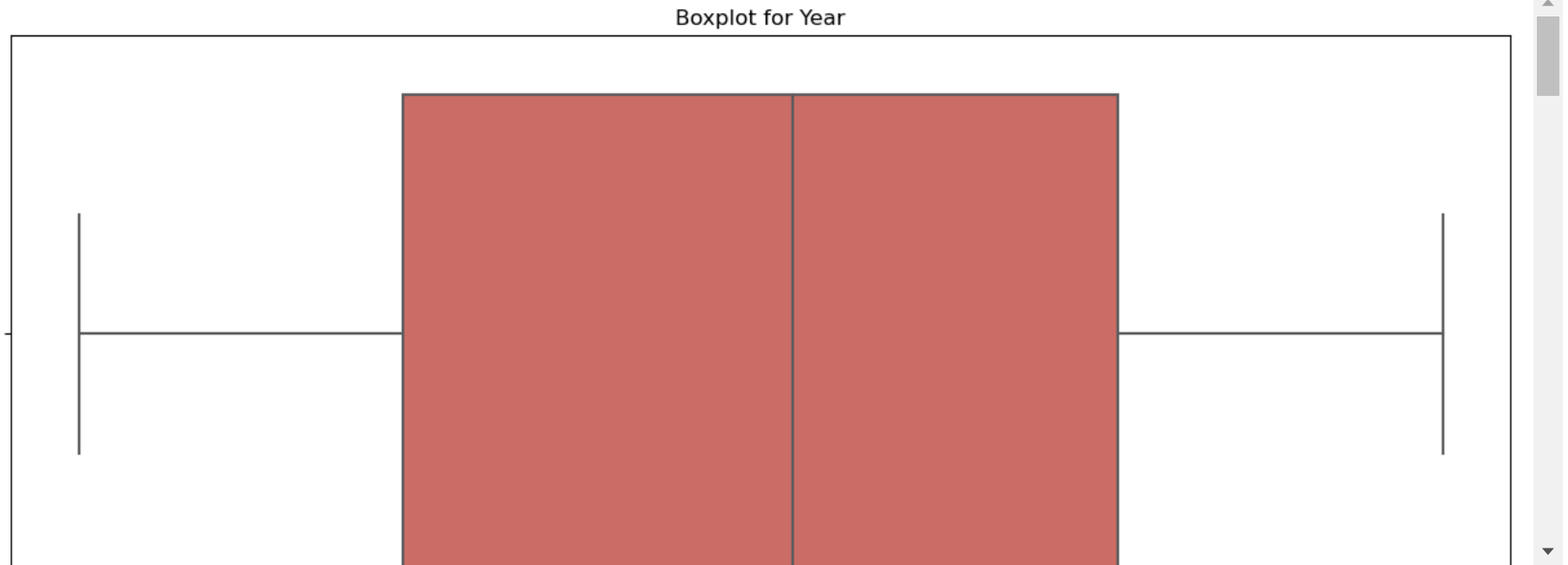






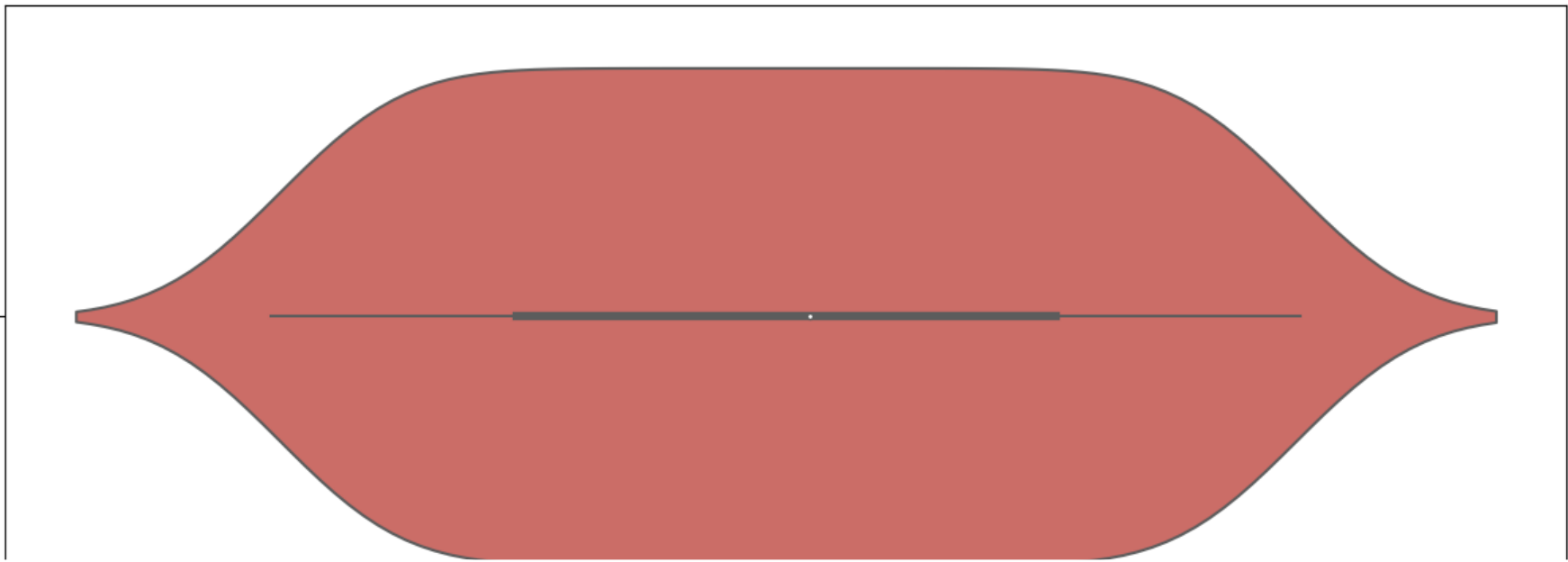


```
In [32]: for column_name in df1.columns:  
    plt.figure(figsize=(15, 6))  
    sns.boxplot(x=df1[column_name], palette='hls')  
    plt.xticks(rotation=90)  
    plt.title(f'Boxplot for {column_name}')  
    plt.show()
```

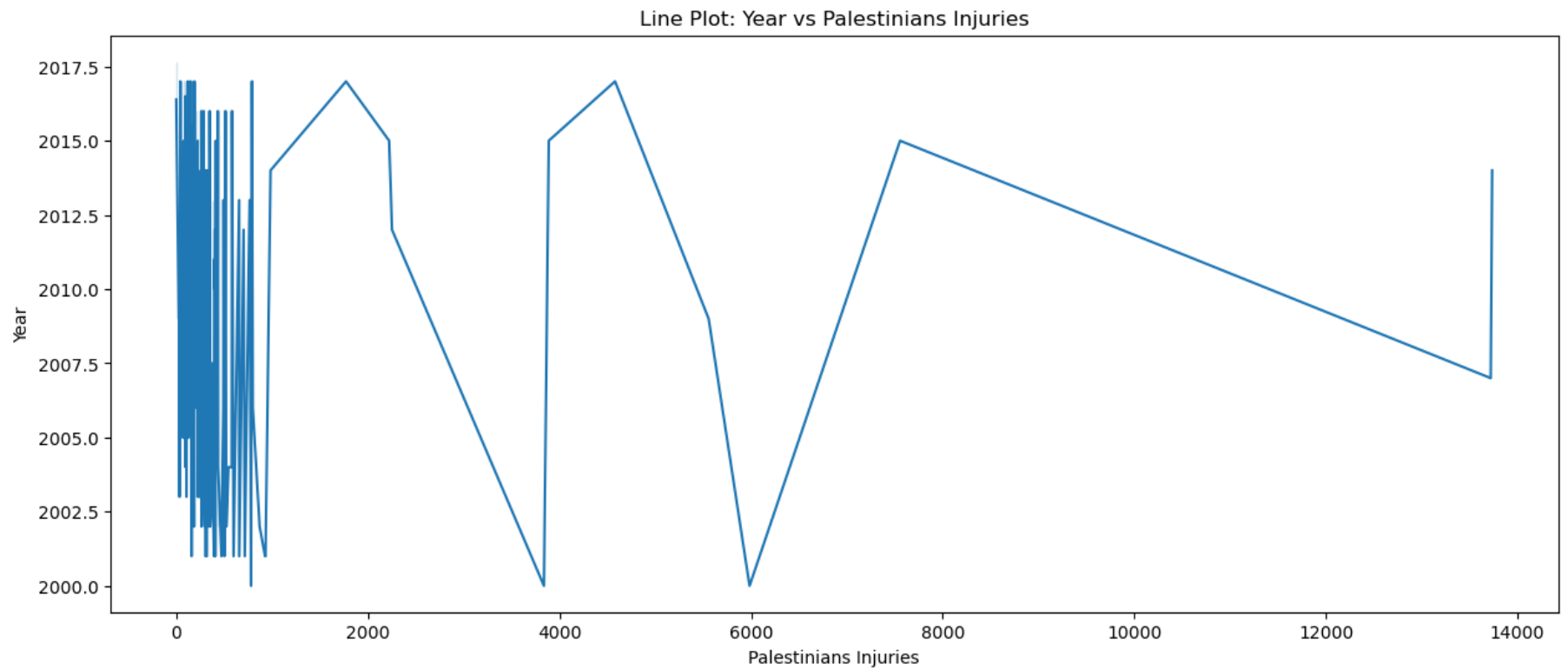


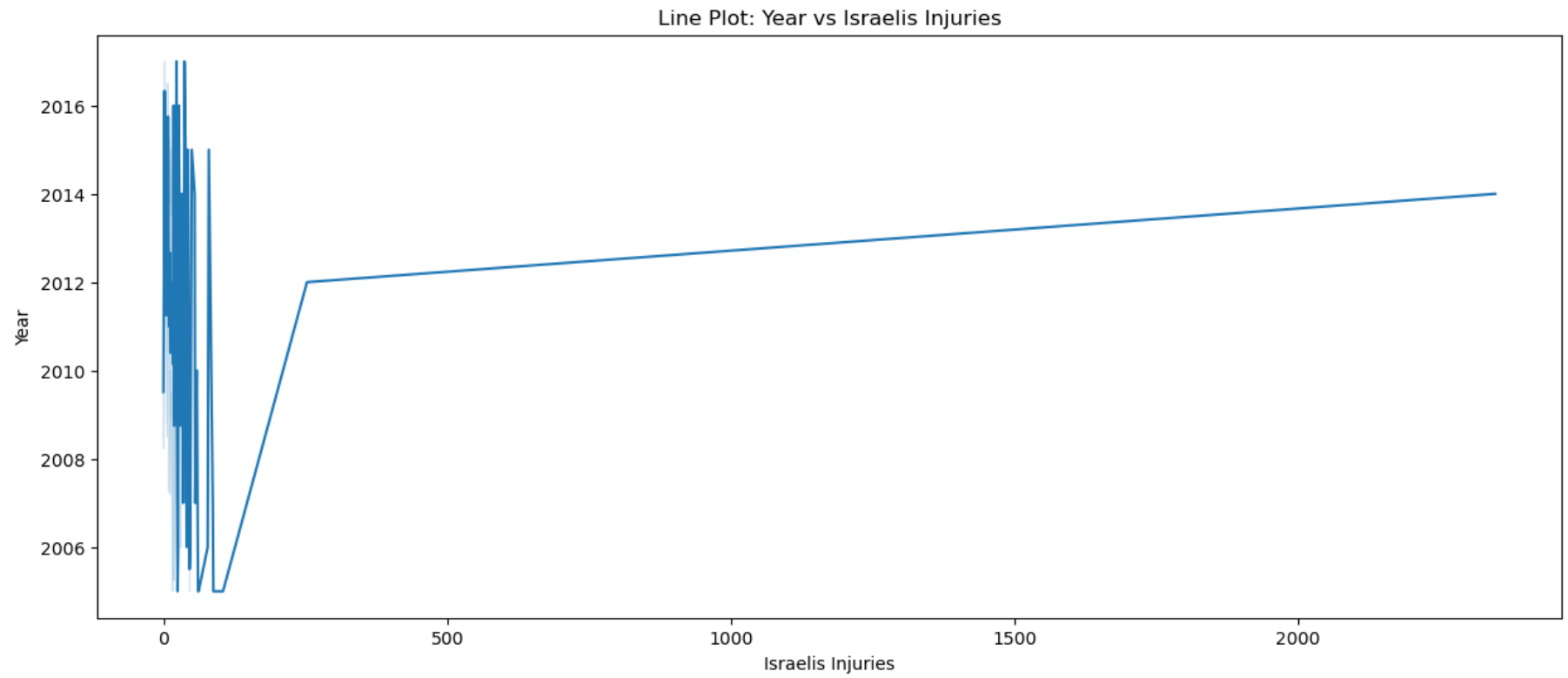
```
In [34]: for column_name in df1.columns:
plt.figure(figsize=(15, 6))
sns.violinplot(x=df1[column_name], palette='hls')
plt.xticks(rotation=90)
plt.title(f'Violin Plot for {column_name}')
plt.show()
```

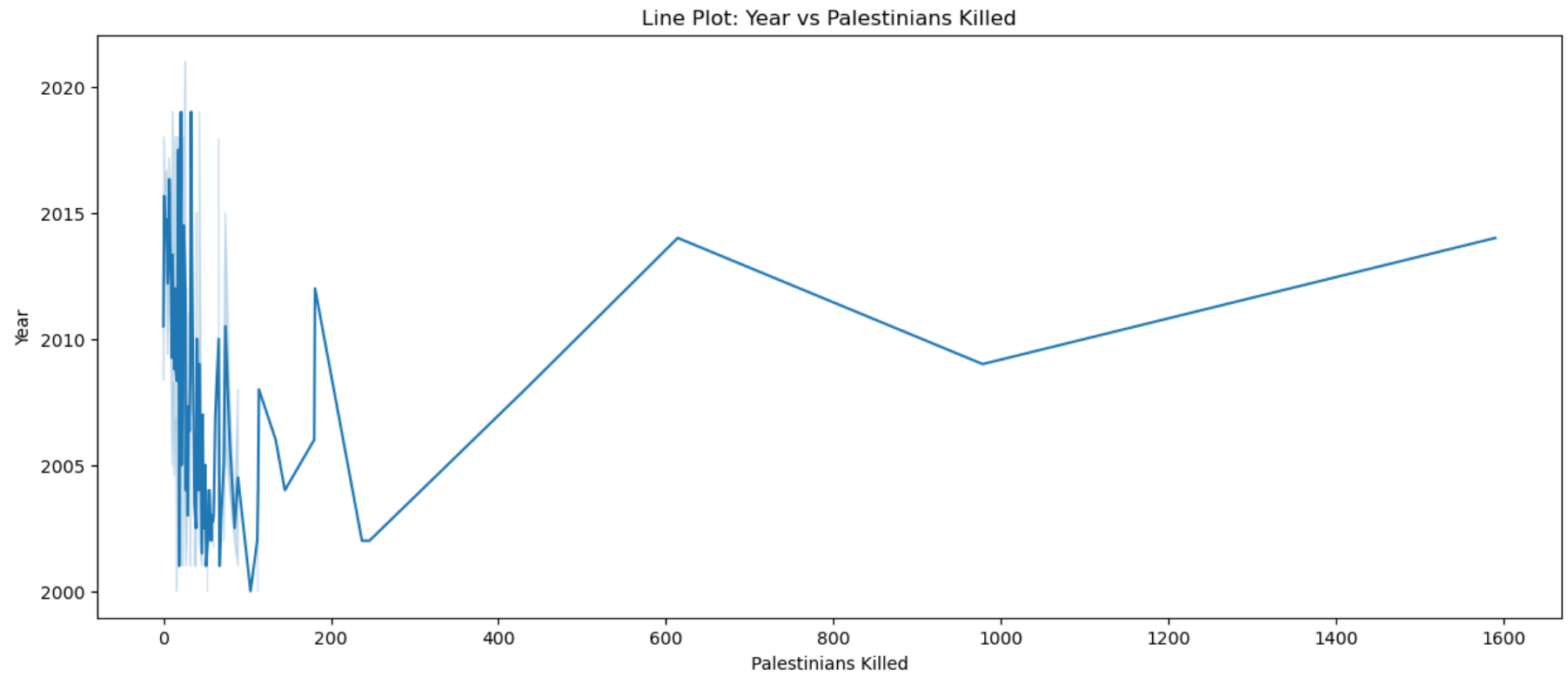
Violin Plot for Year

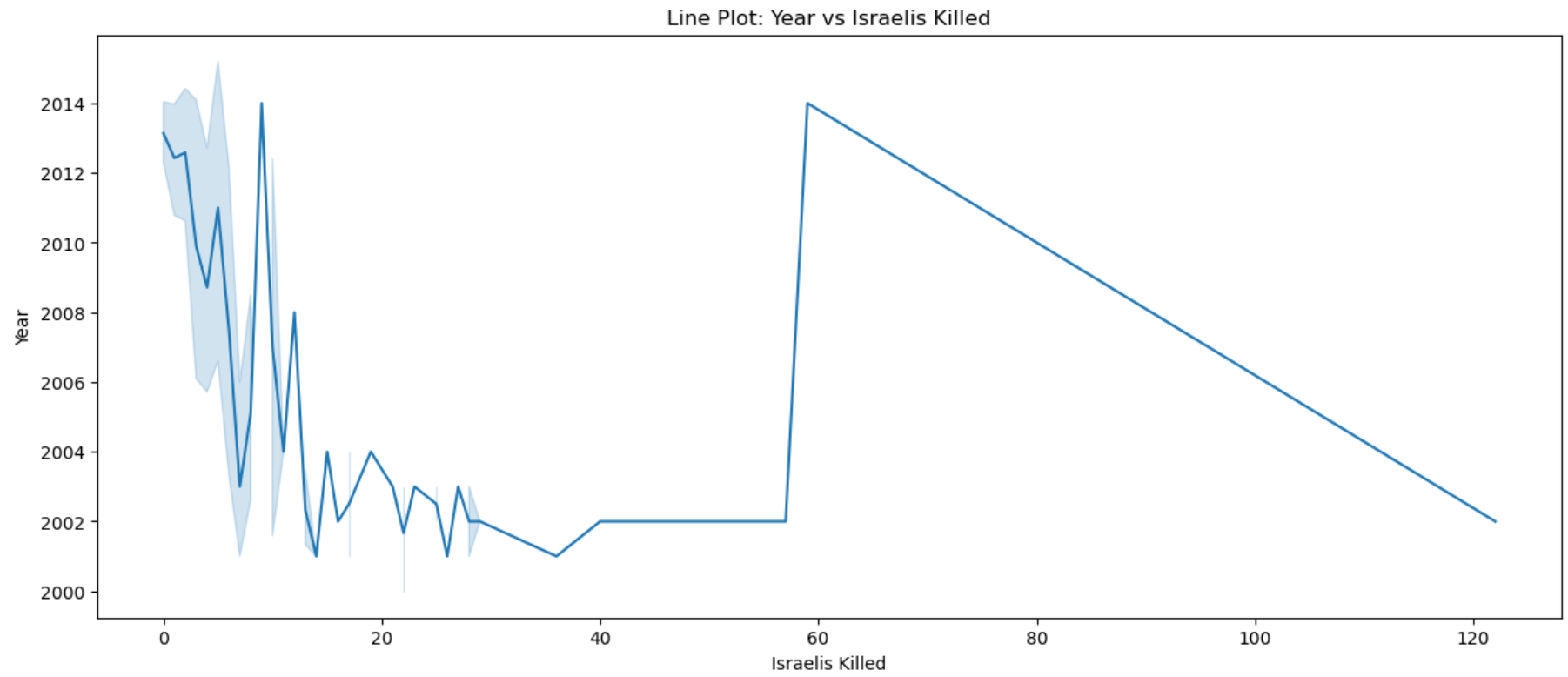


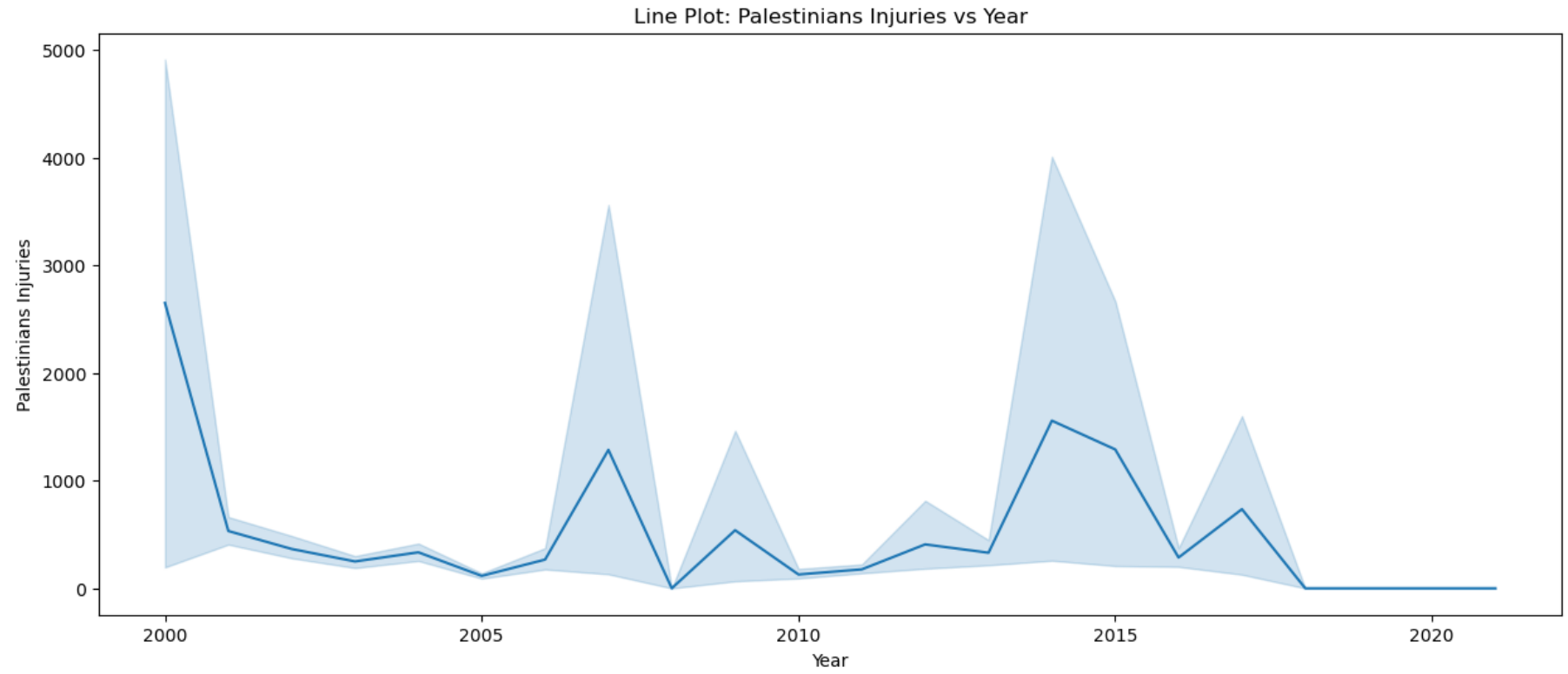
```
In [39]: for i in df1.columns:
         for j in df1.columns:
             if i != j:
                 plt.figure(figsize=(15, 6))
                 sns.lineplot(x=df1[j], y=df1[i], data=df1, palette='hls')
                 plt.title(f'Line Plot: {i} vs {j}')
                 plt.show()
```

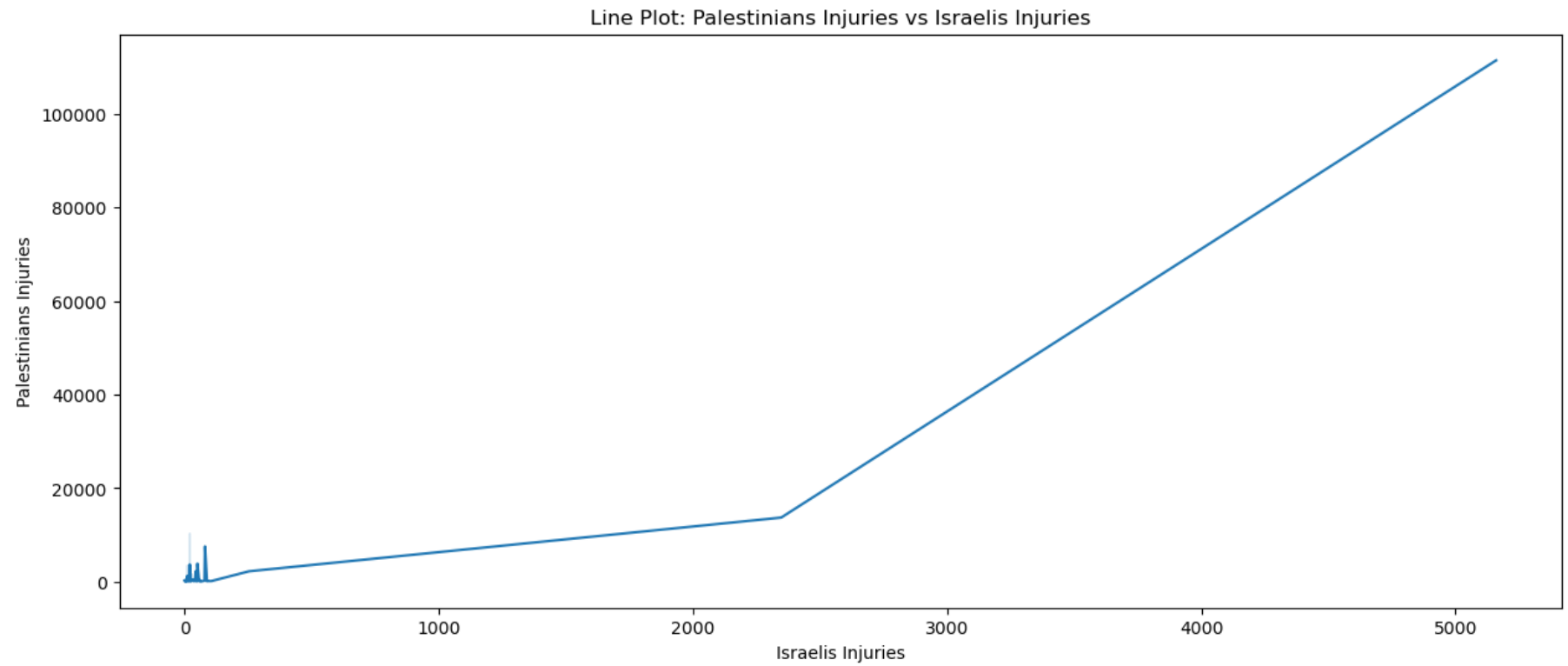


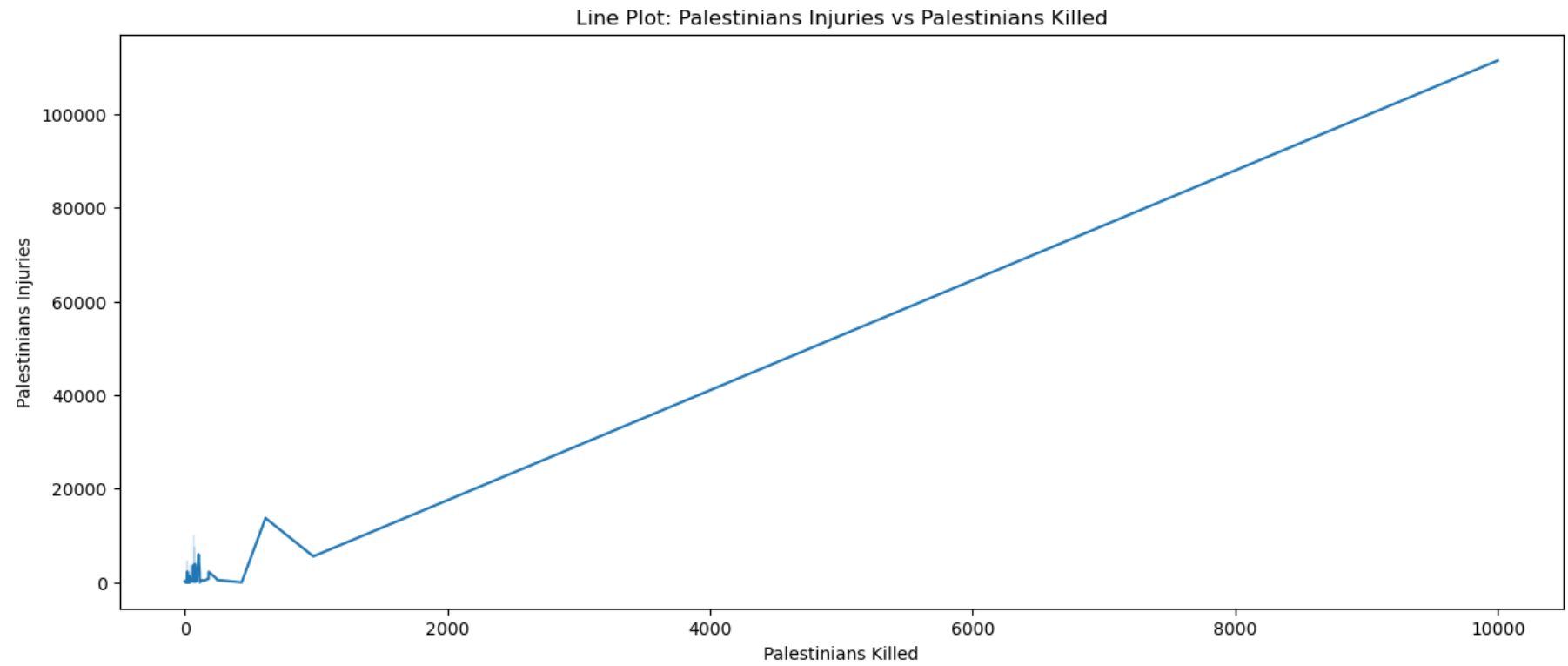


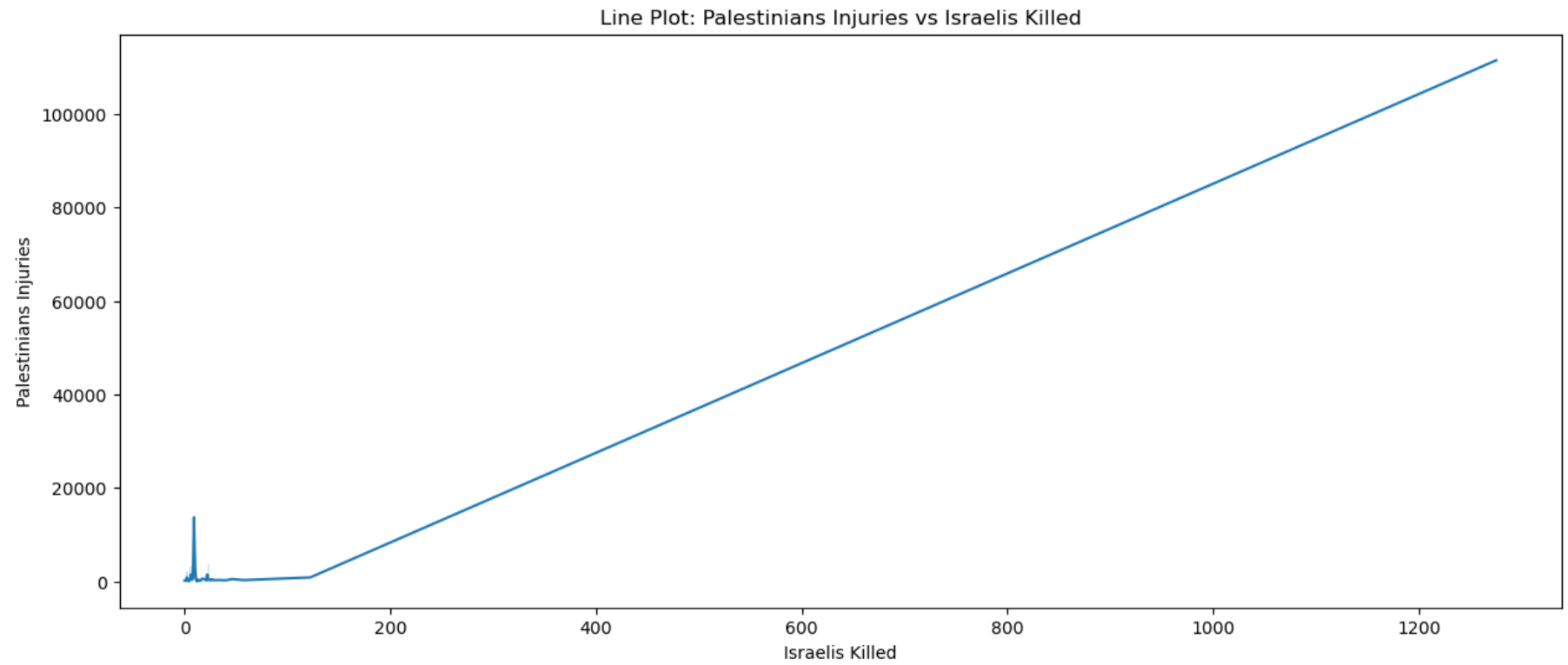


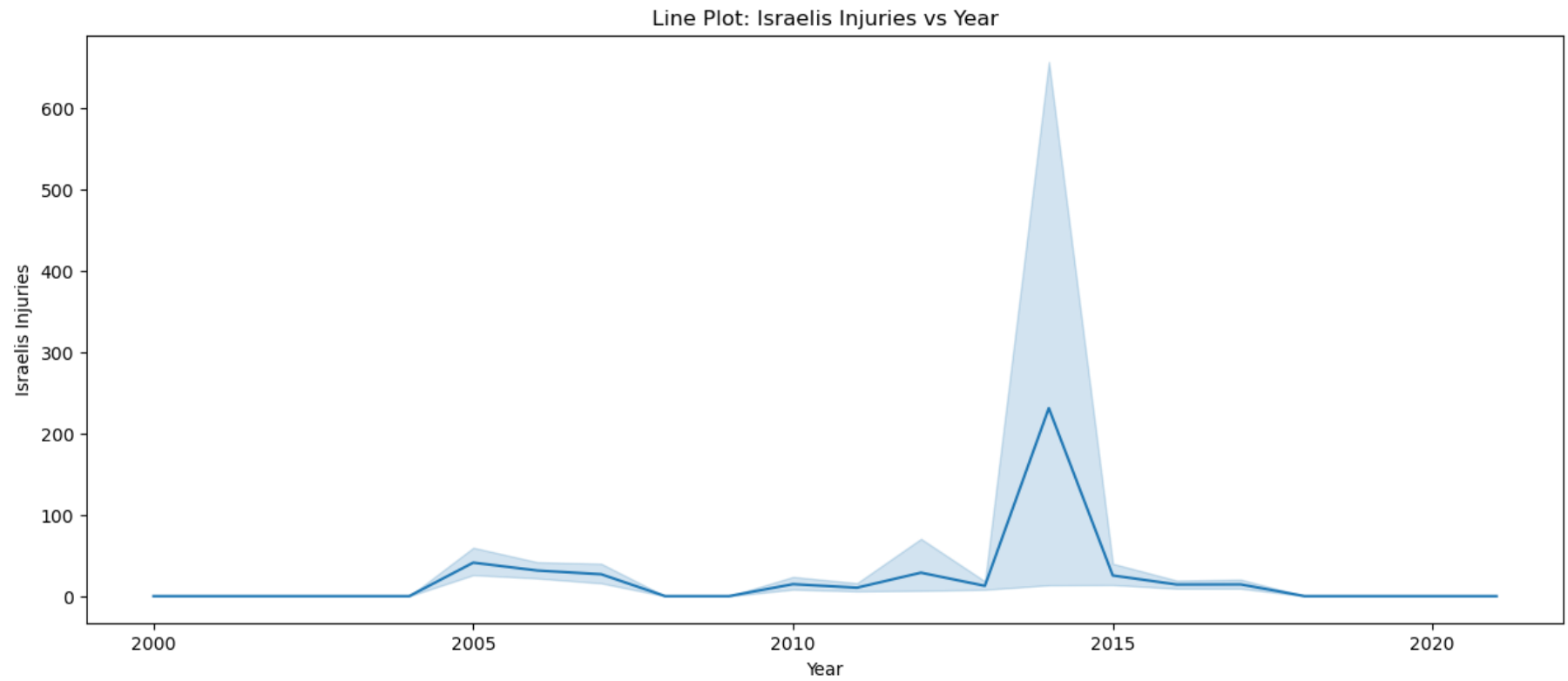


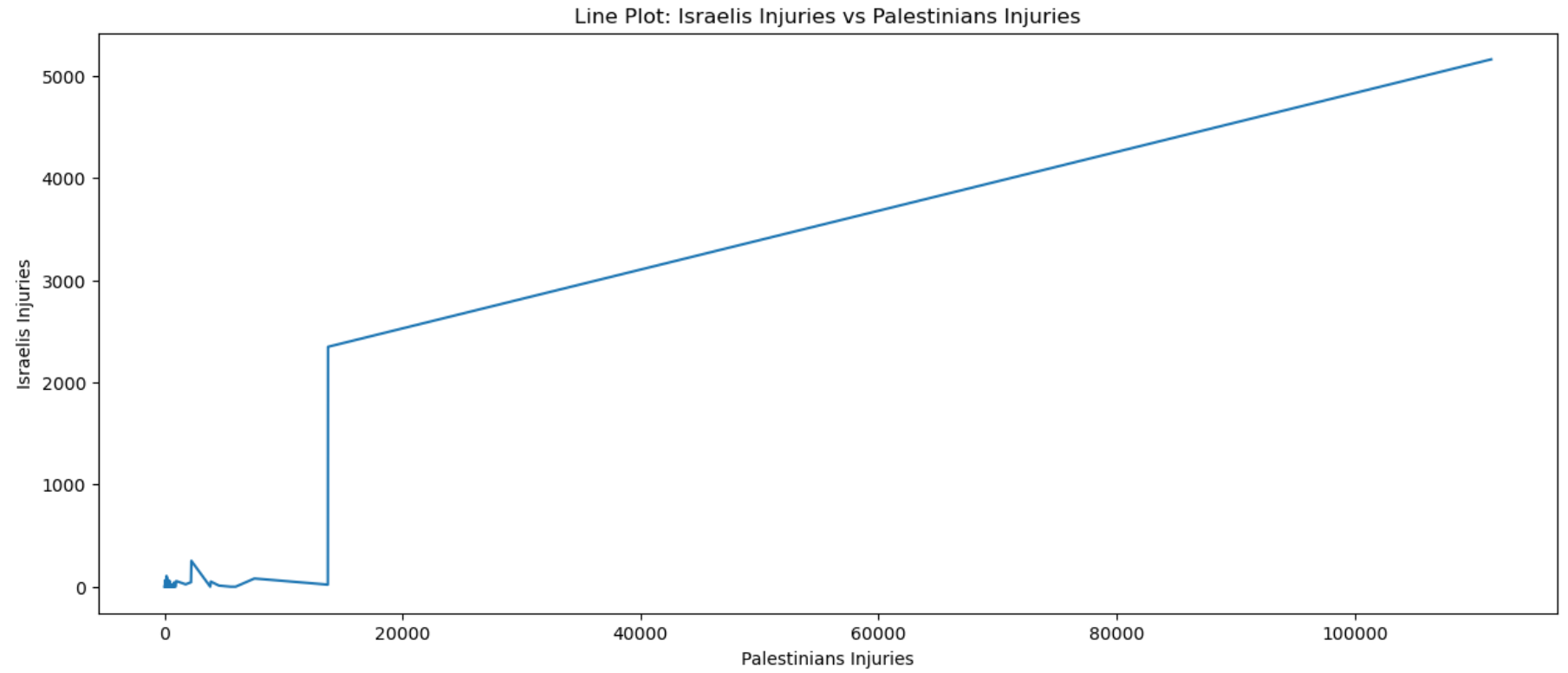


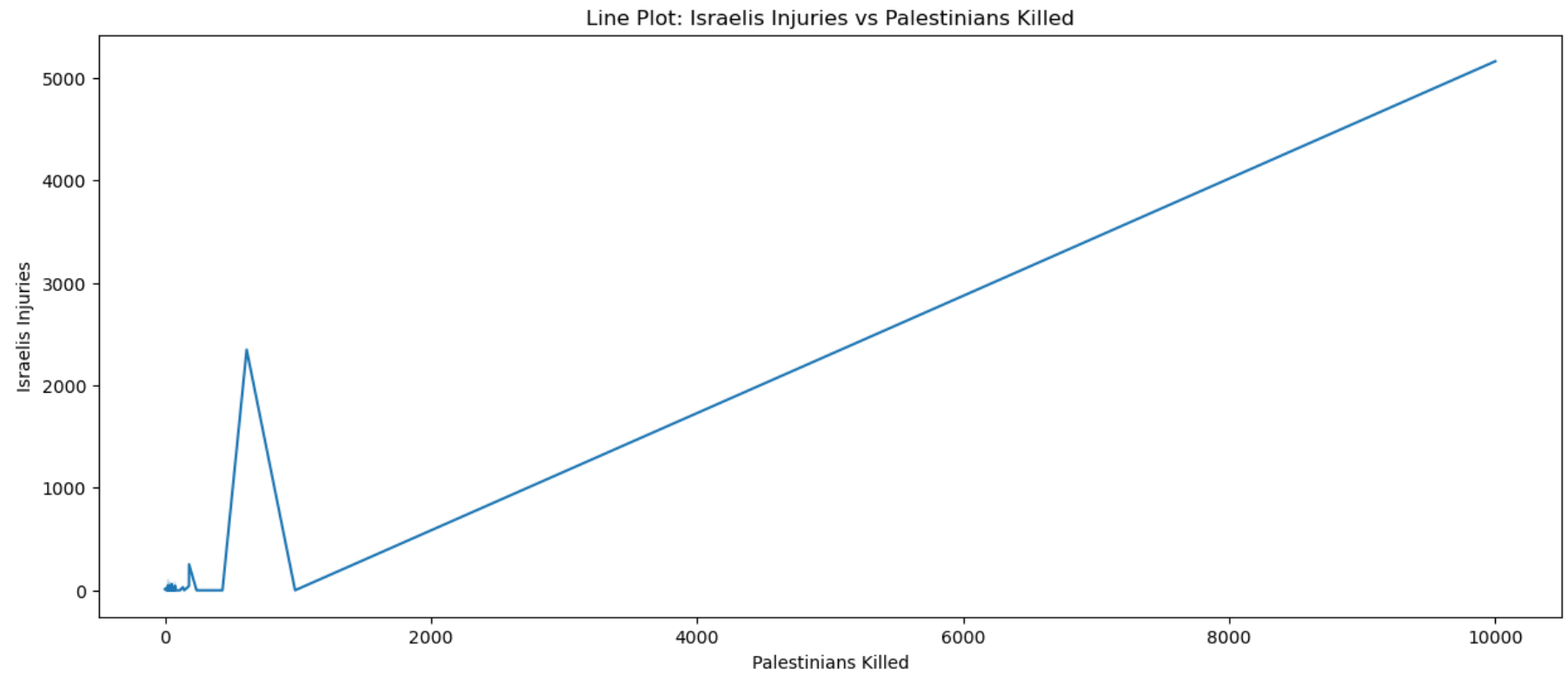


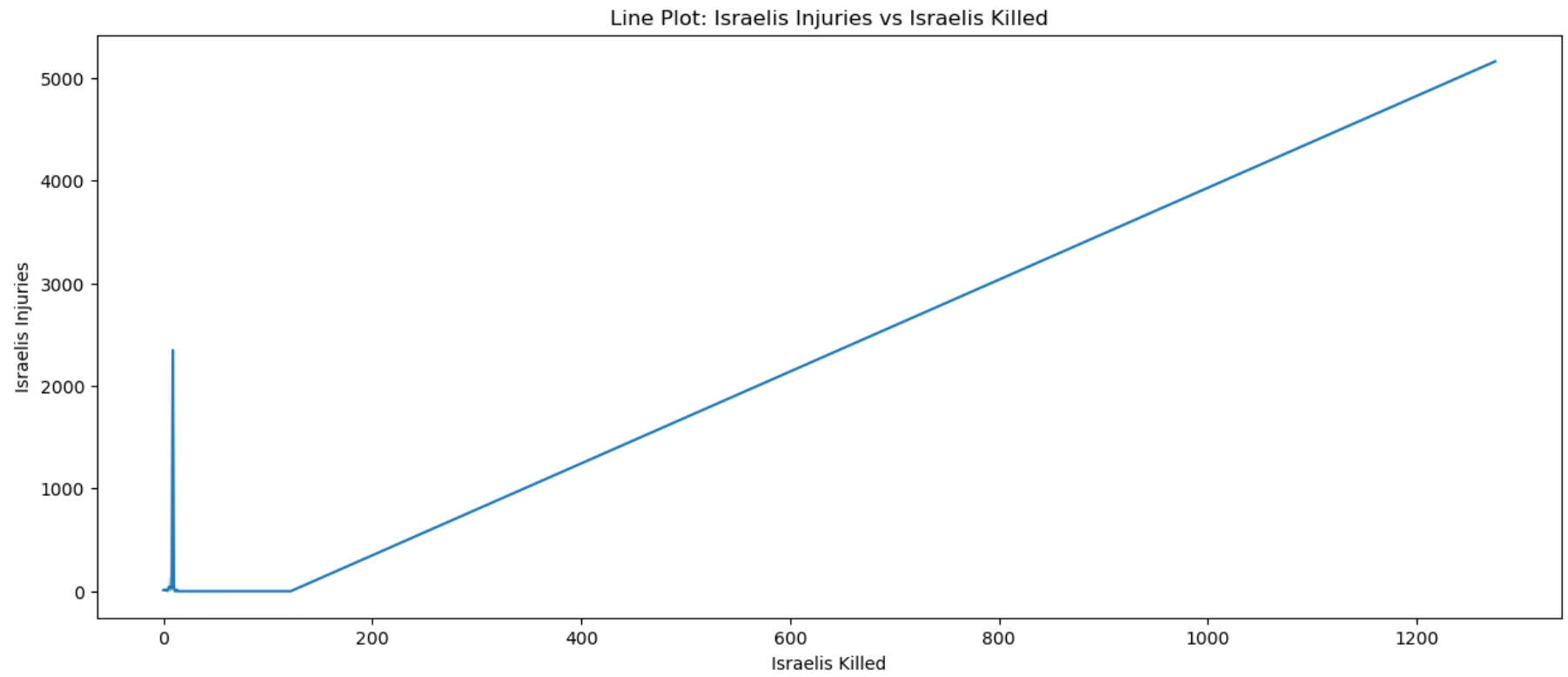


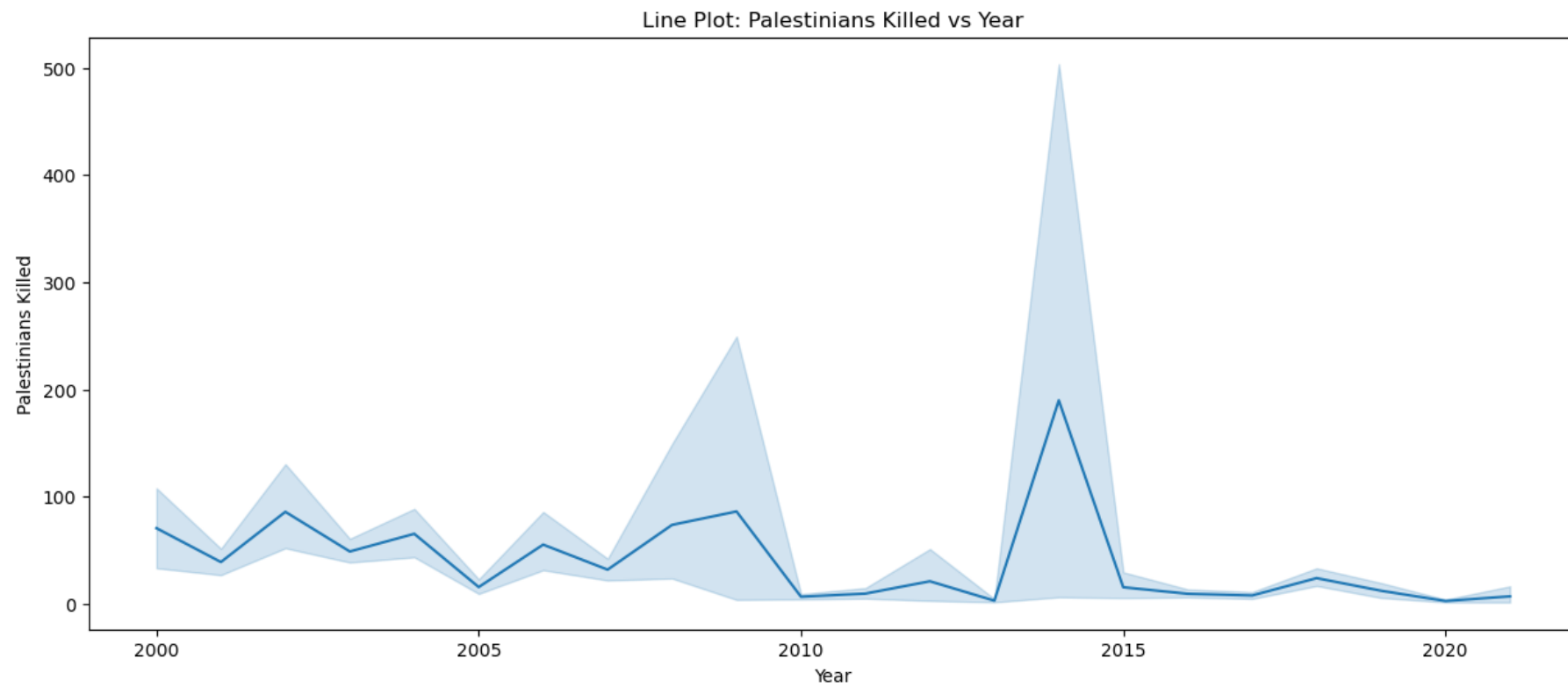


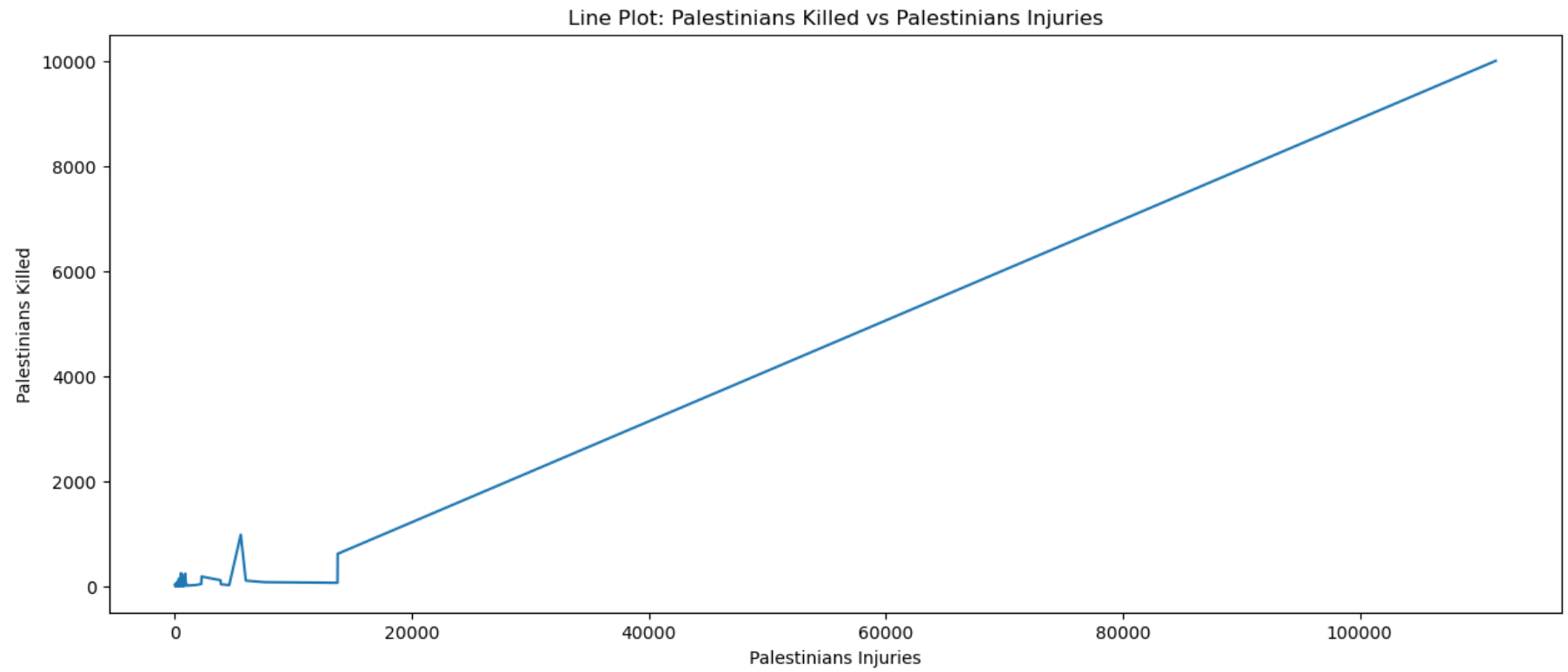


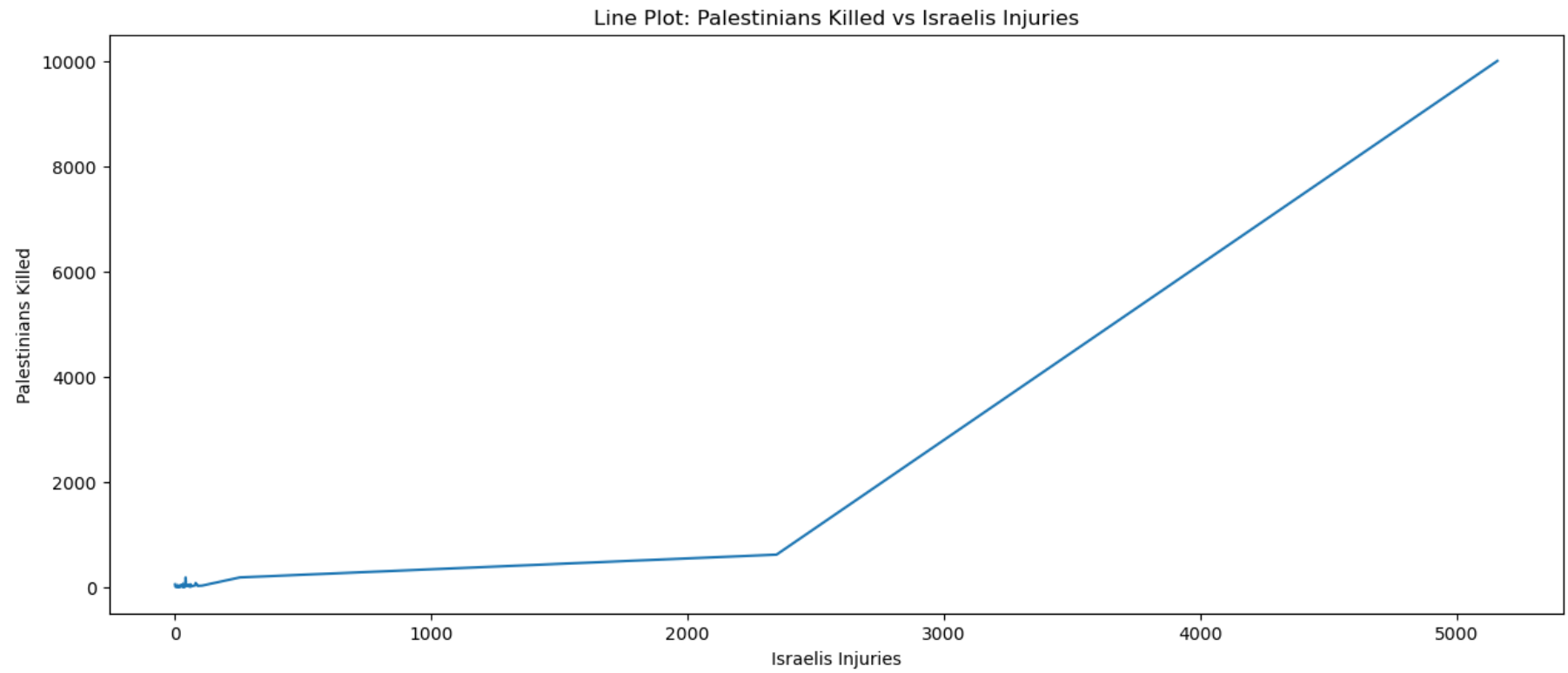


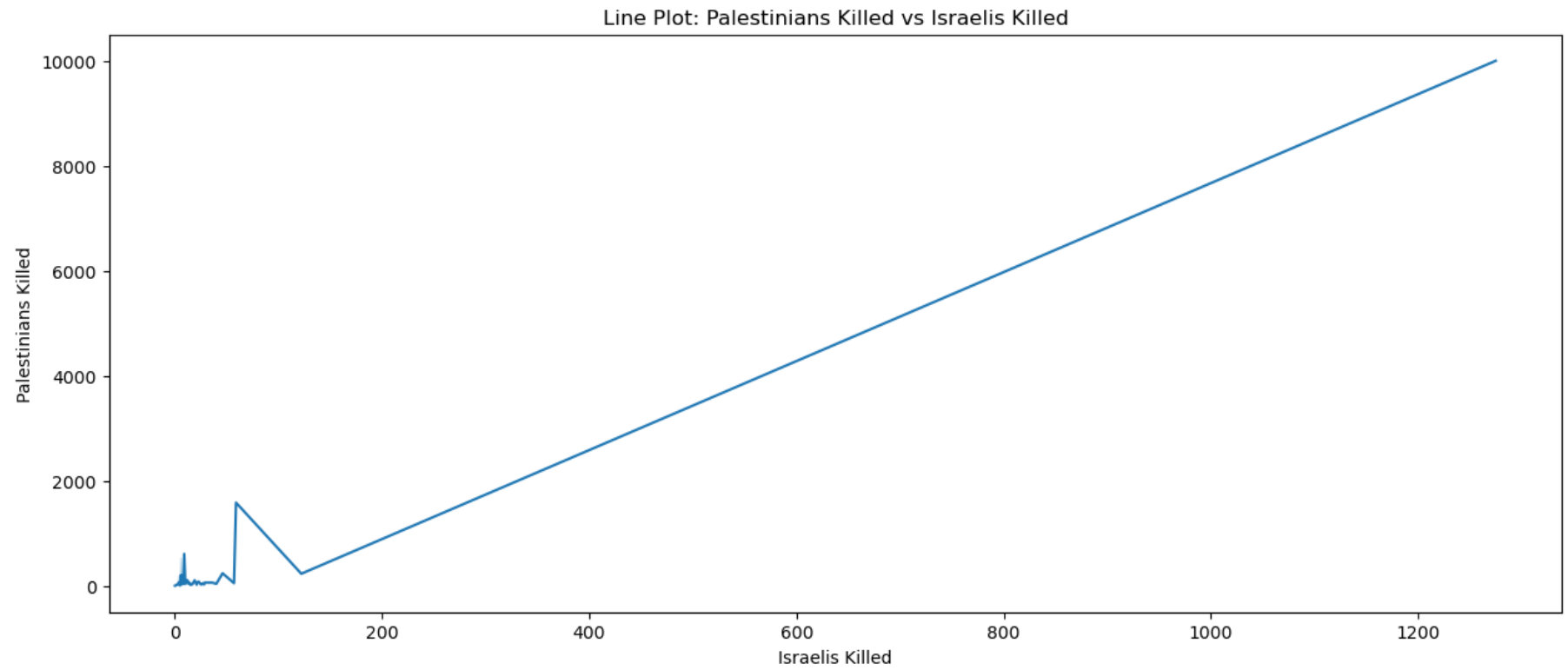


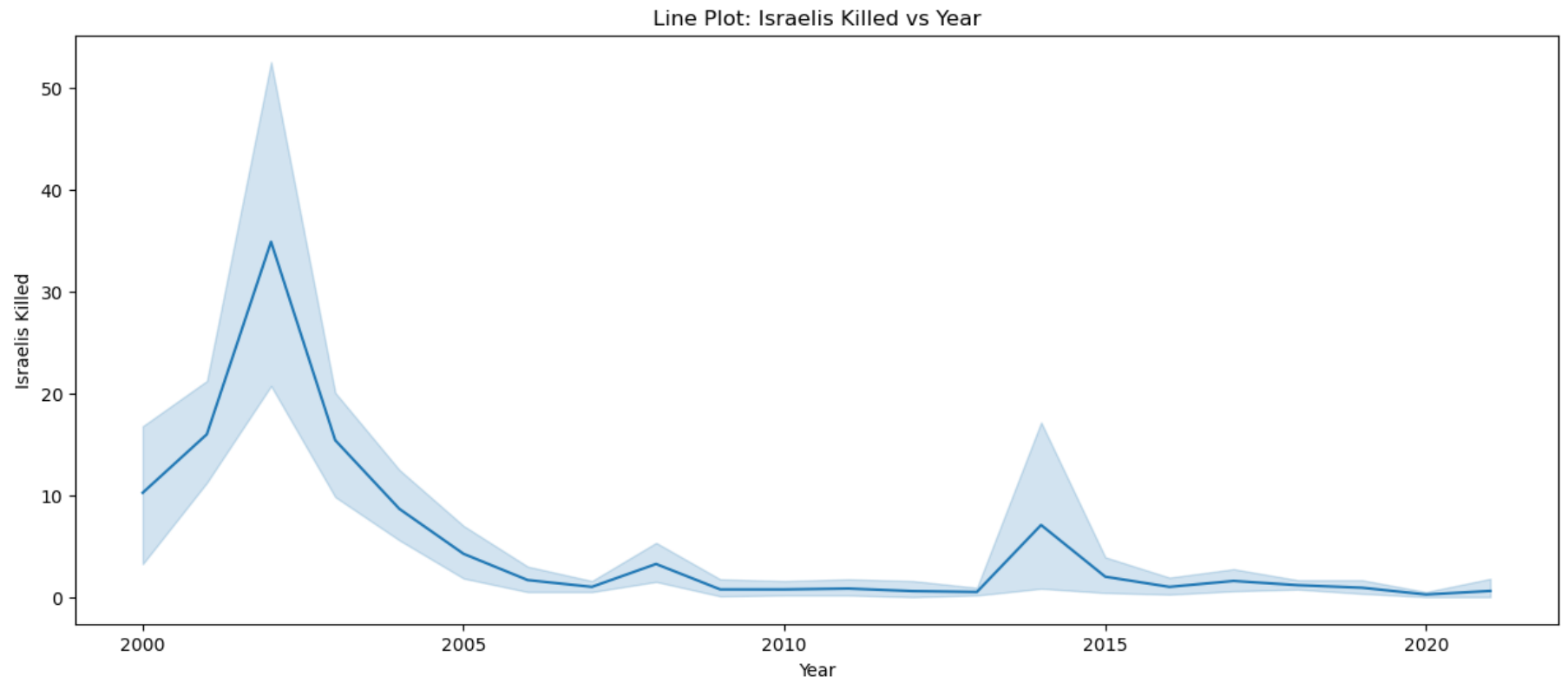


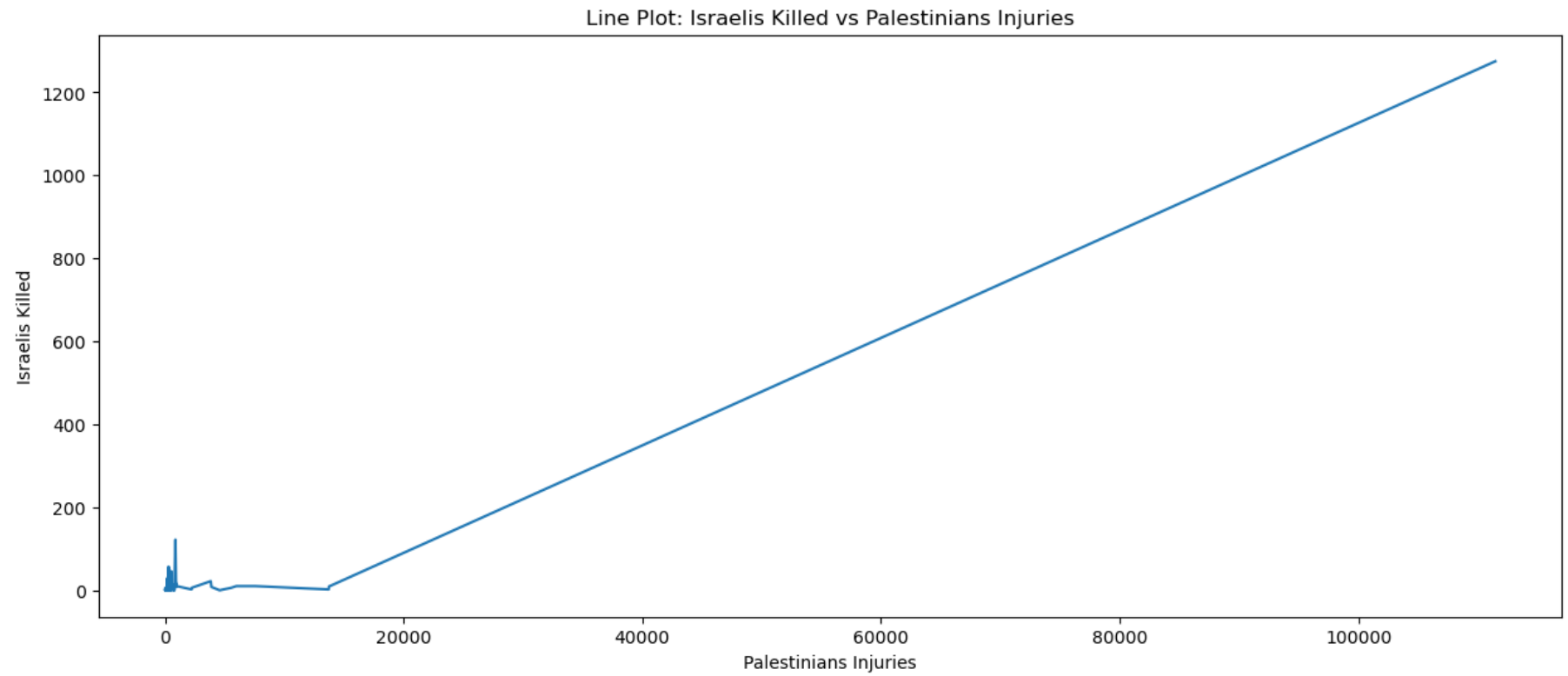


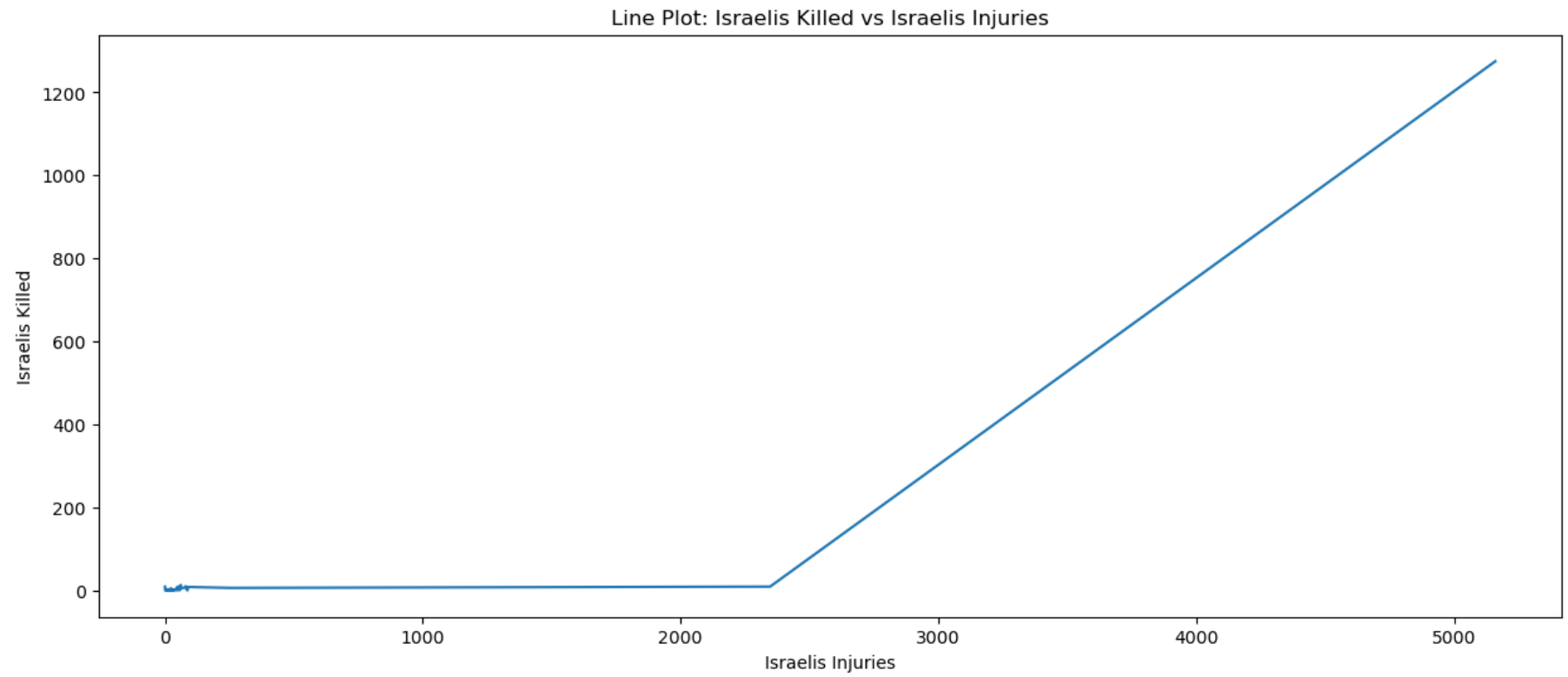


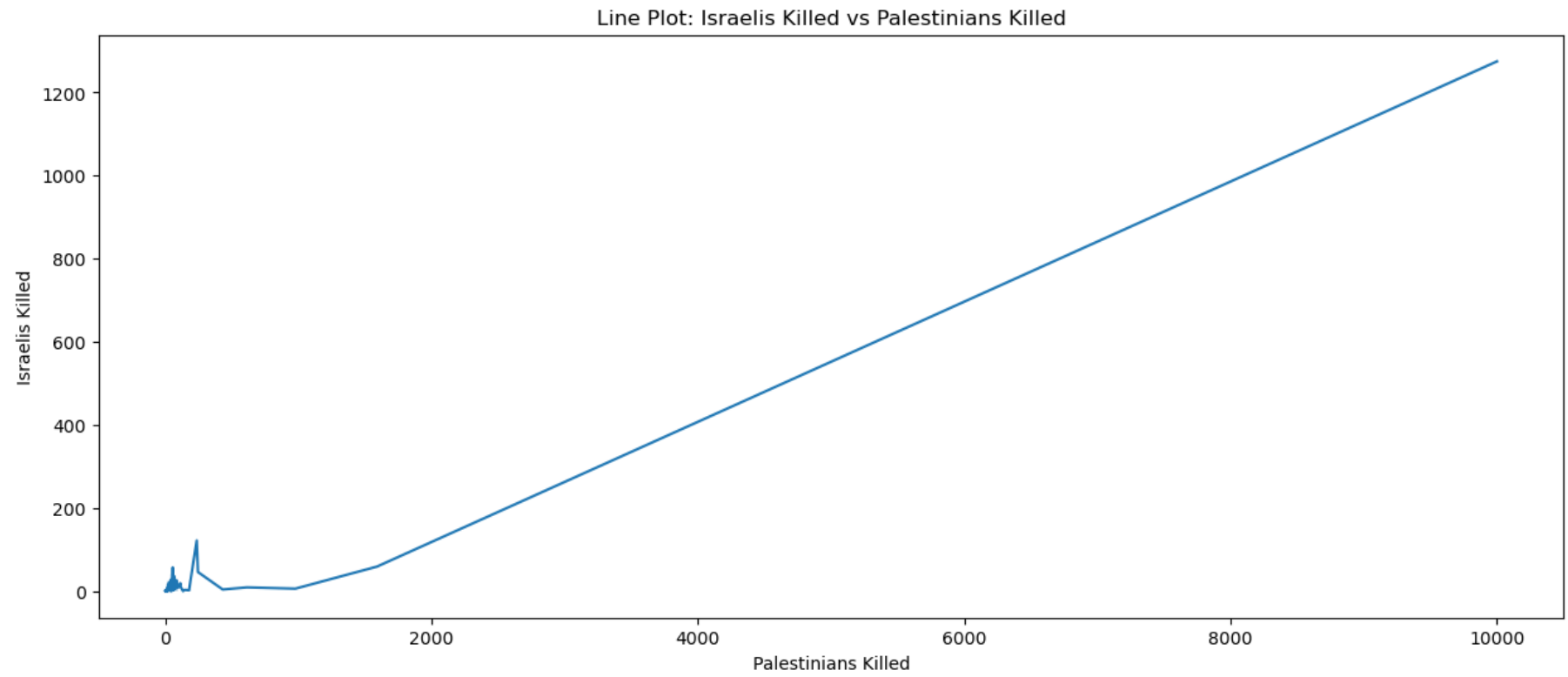




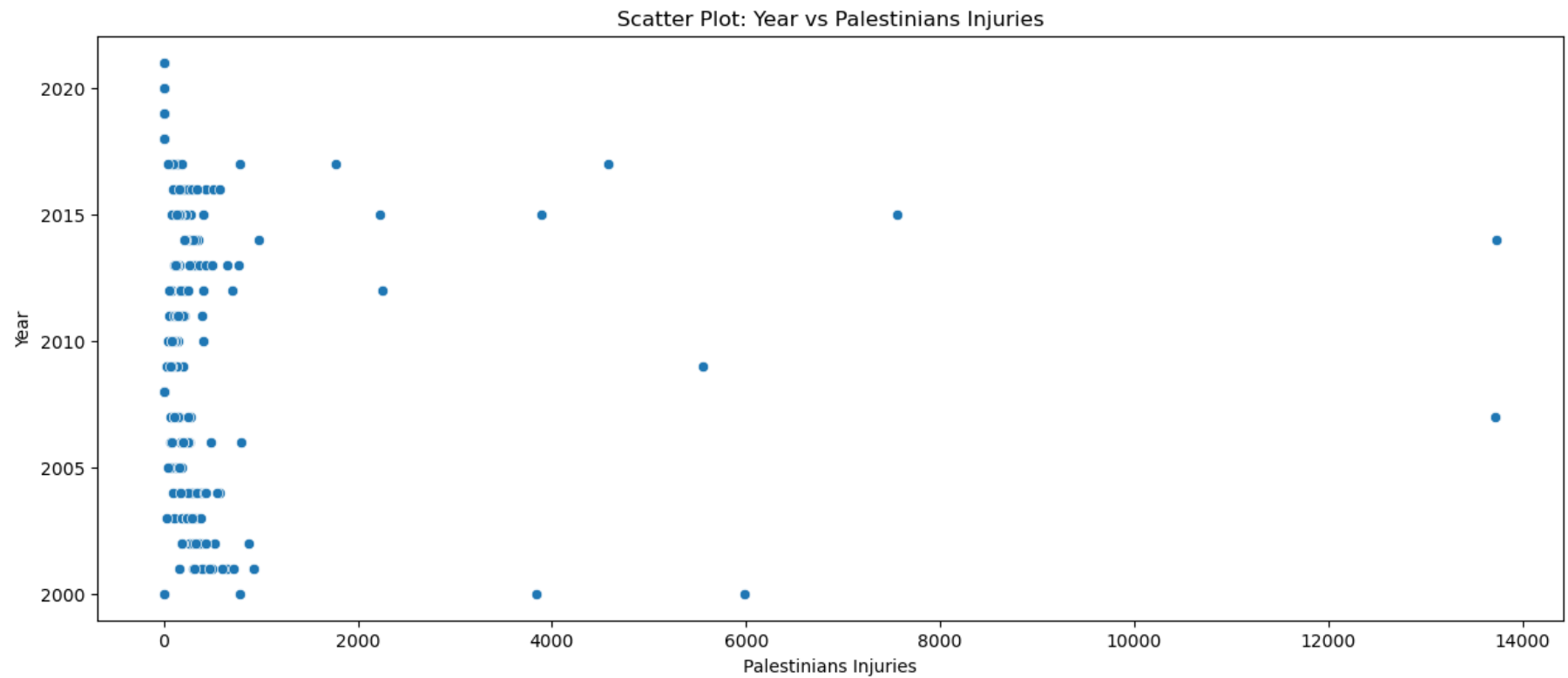


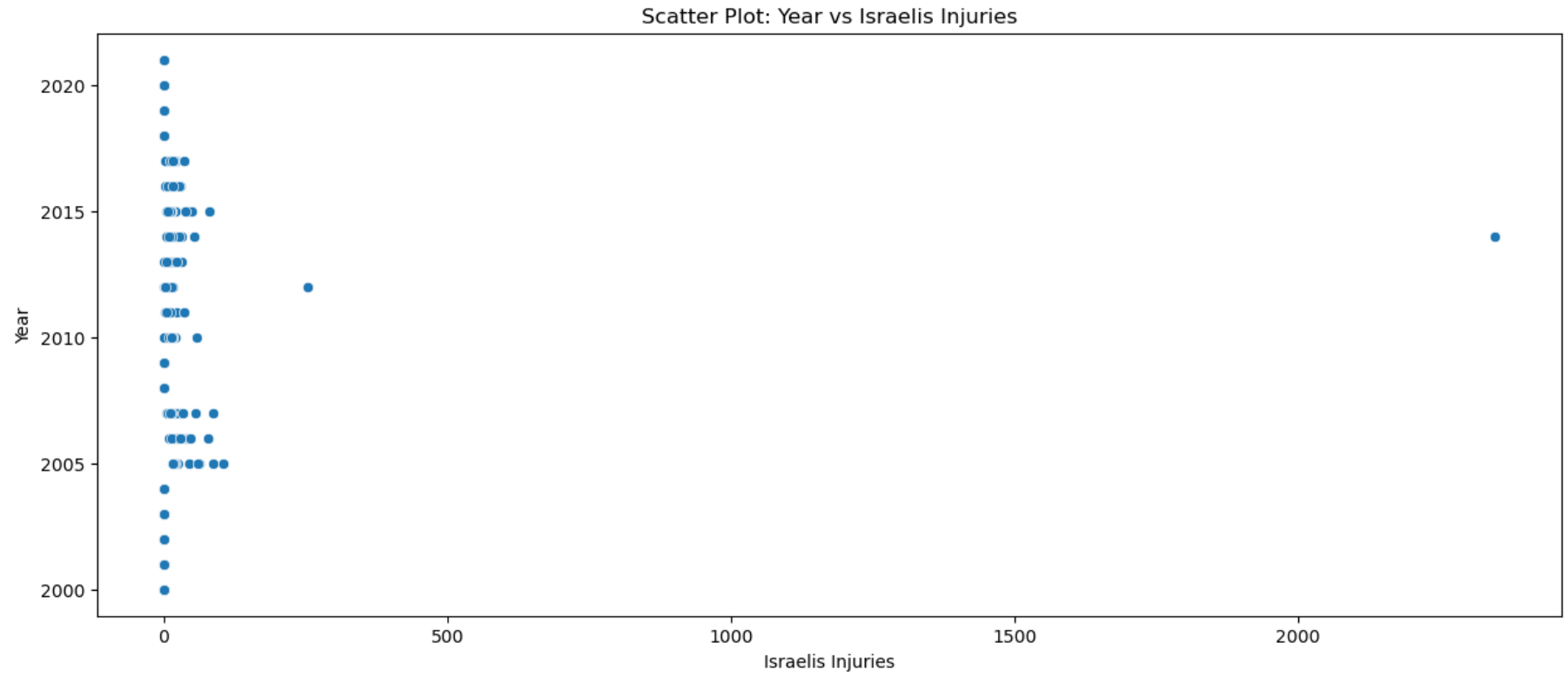




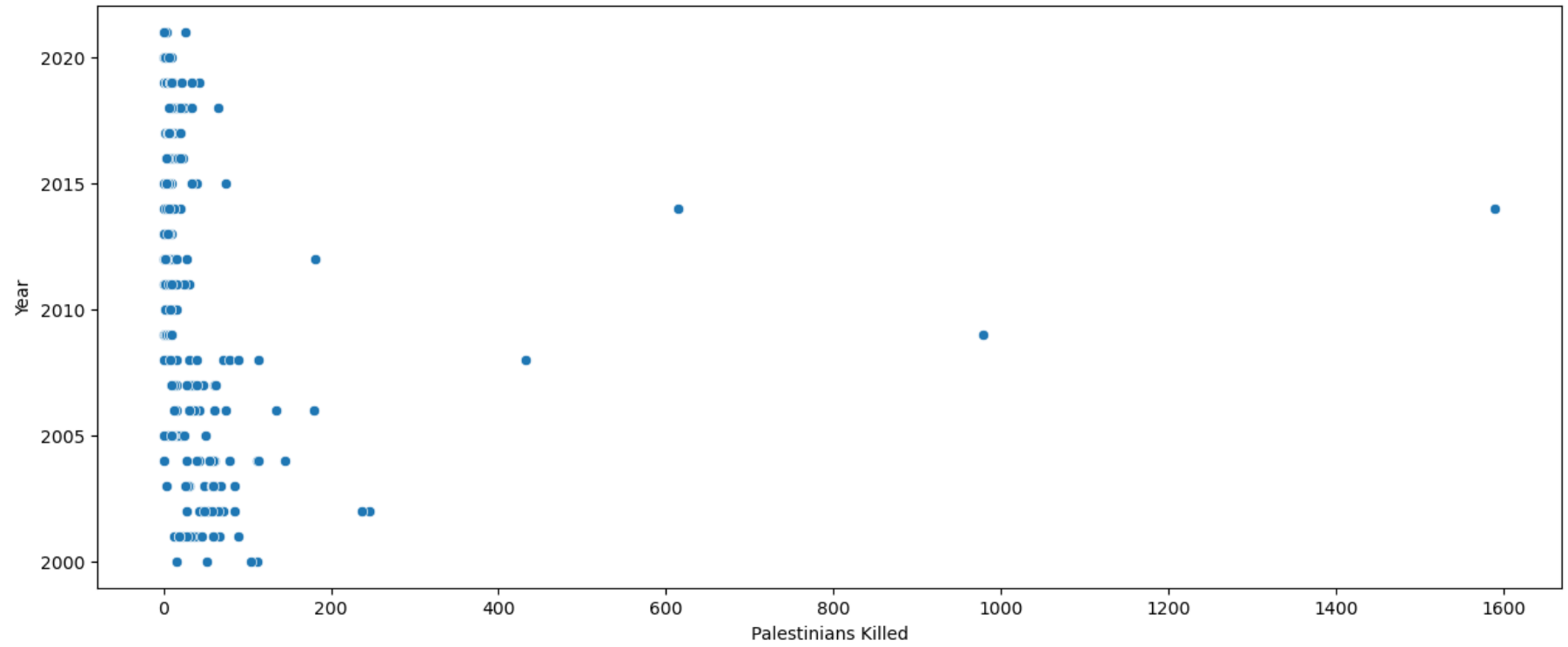


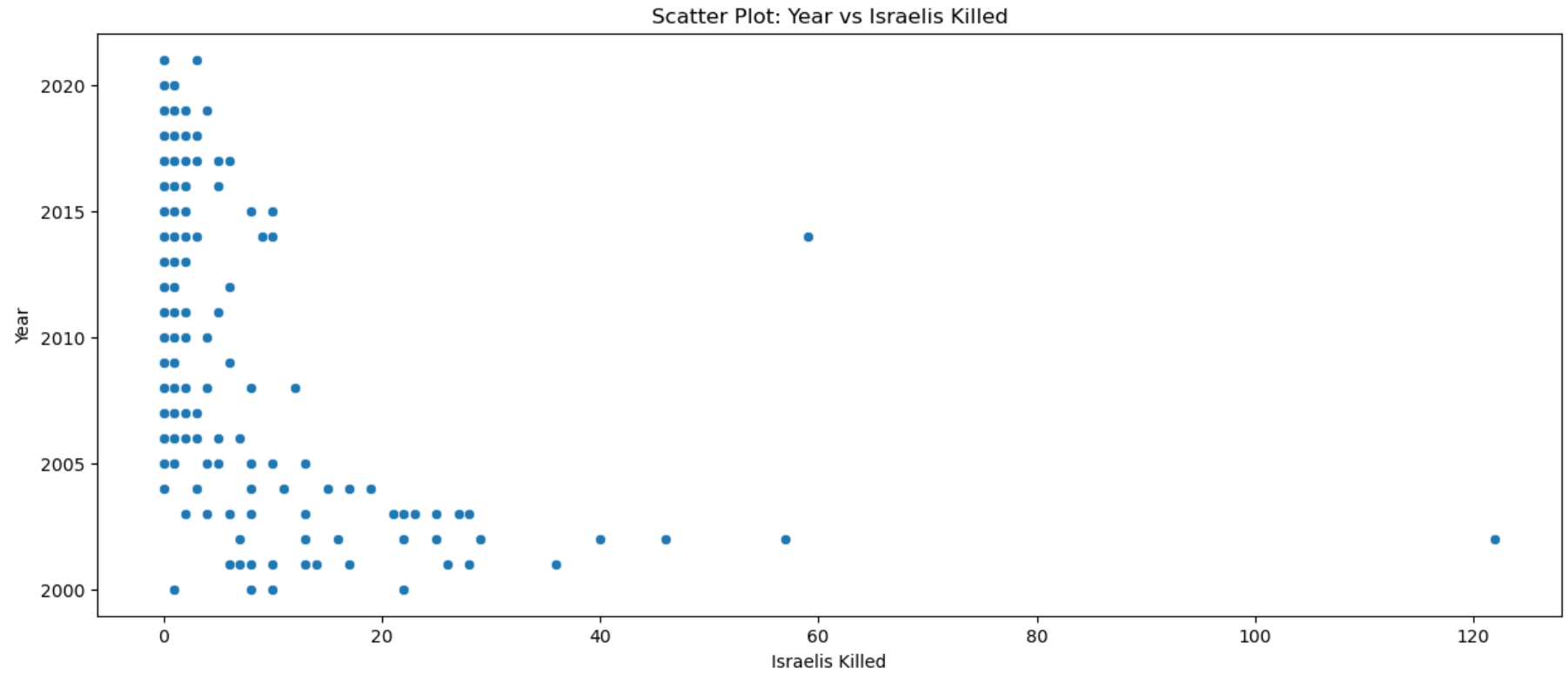
```
In [41]: for i in df1.columns:
         for j in df1.columns:
             if i != j:
                 plt.figure(figsize=(15, 6))
                 sns.scatterplot(x=df1[j], y=df1[i], data=df1, palette='hls')
                 plt.title(f'Scatter Plot: {i} vs {j}')
                 plt.show()
```

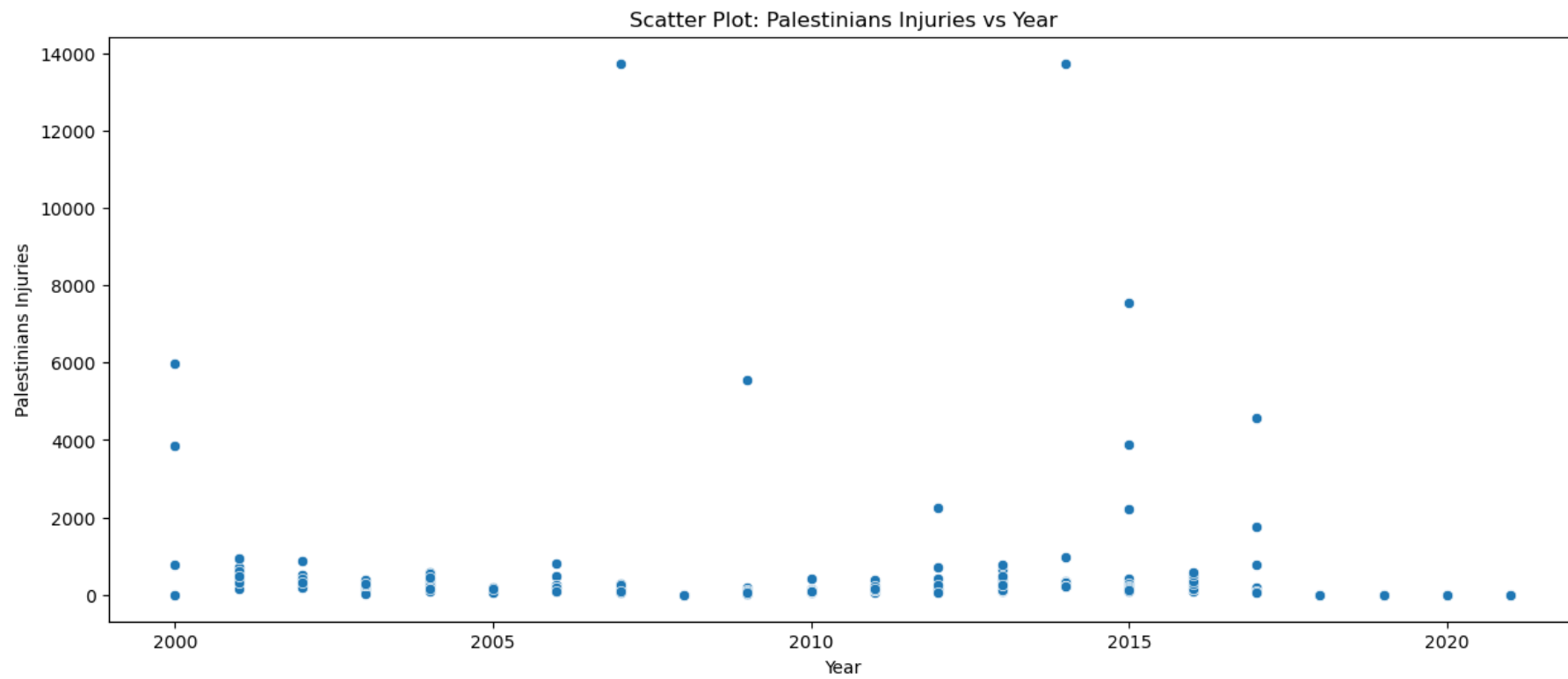




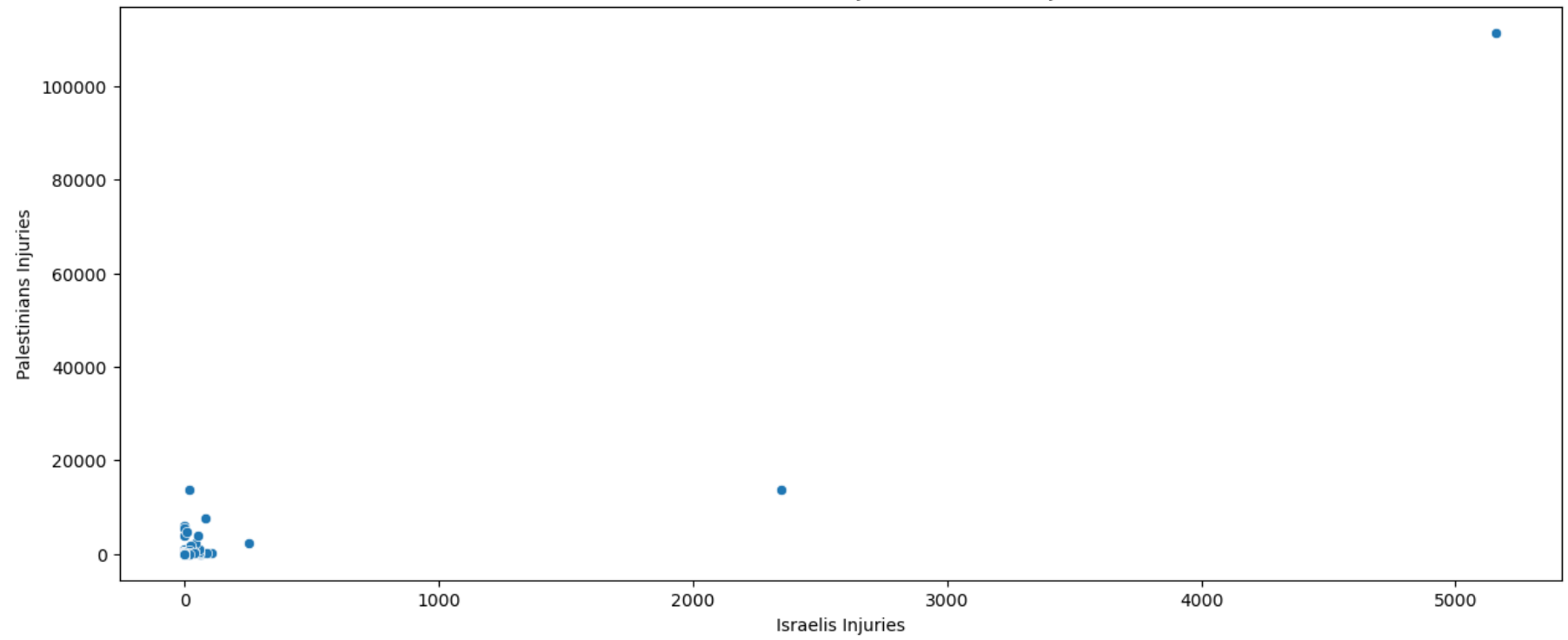
Scatter Plot: Year vs Palestinians Killed

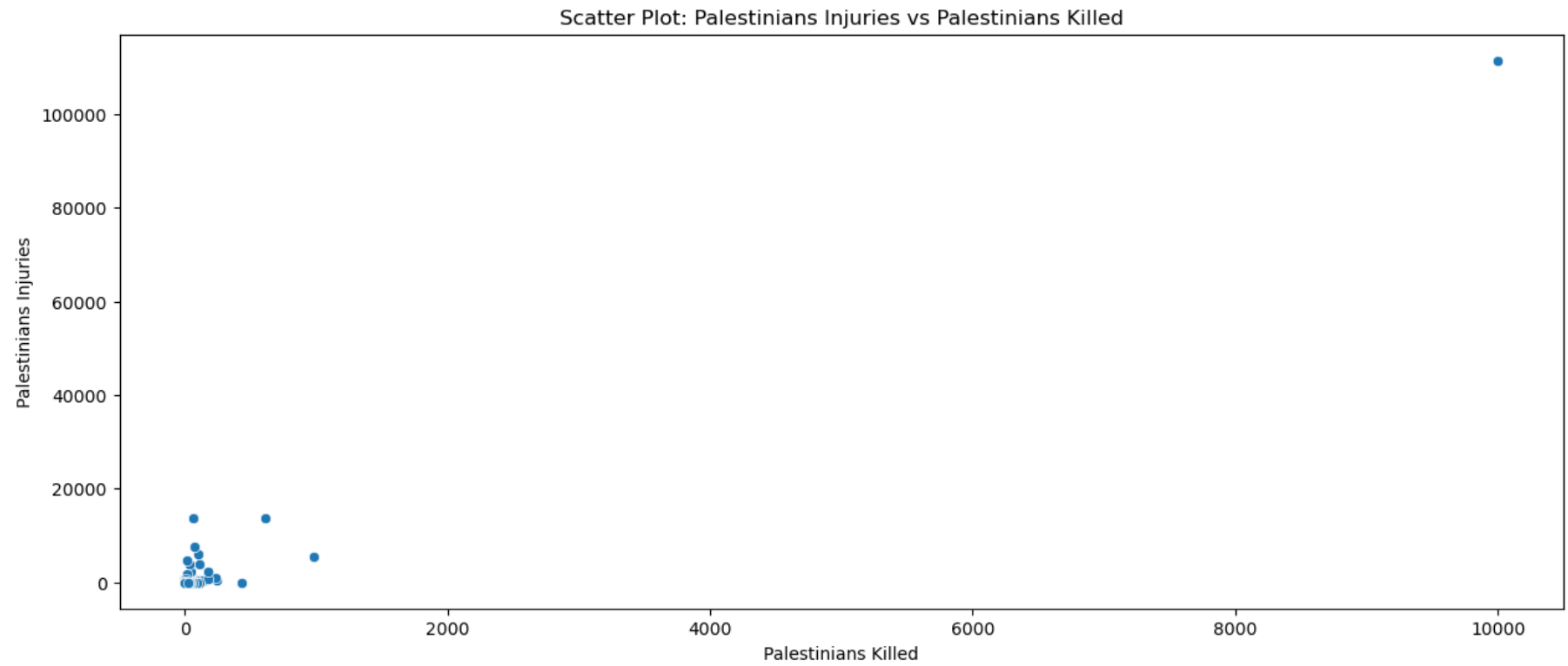


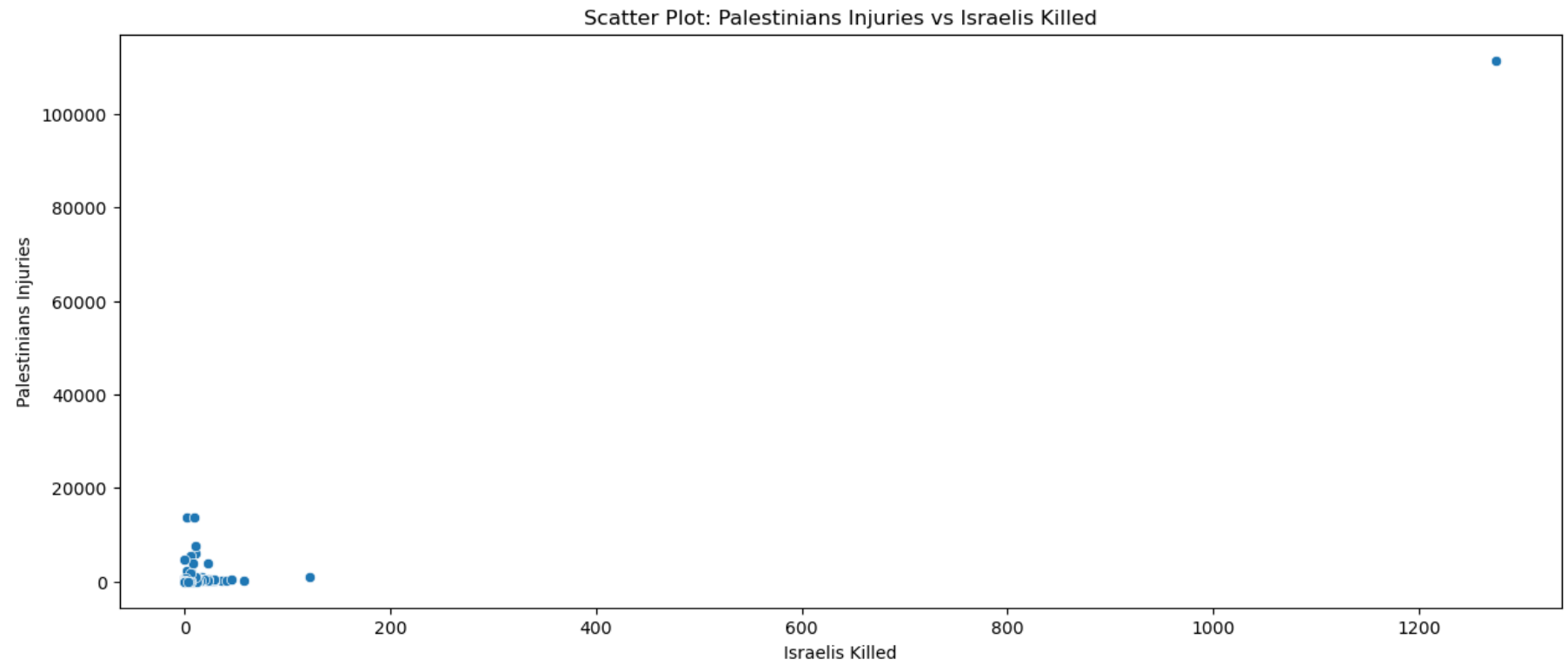


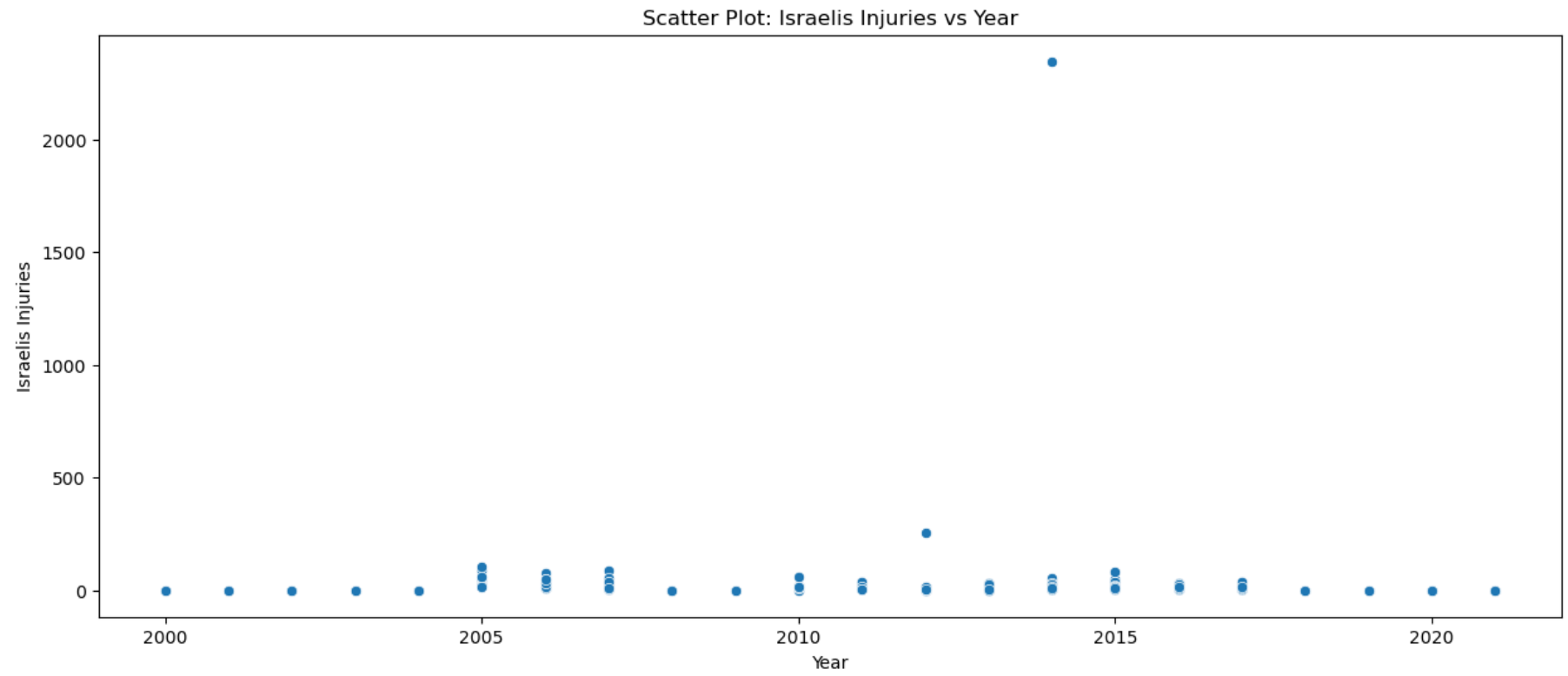


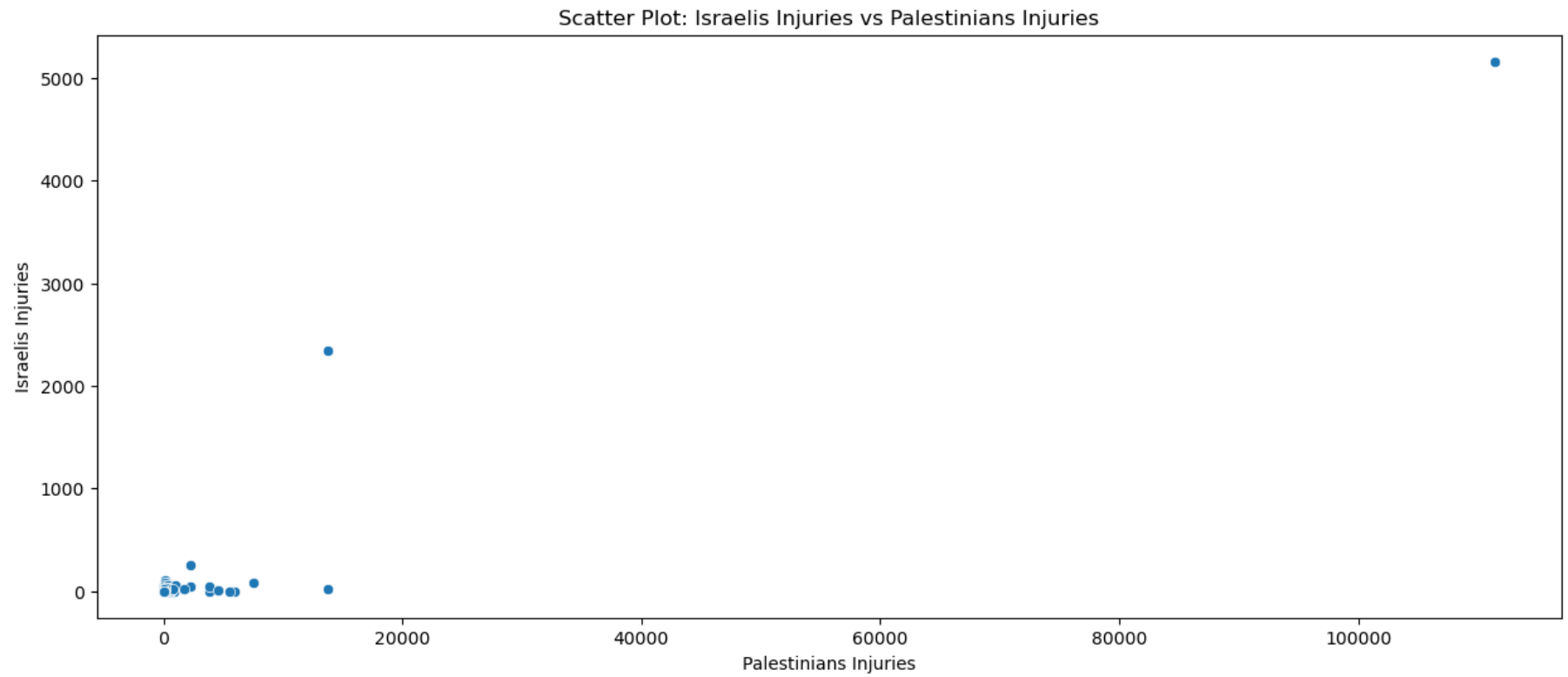
Scatter Plot: Palestinians Injuries vs Israelis Injuries

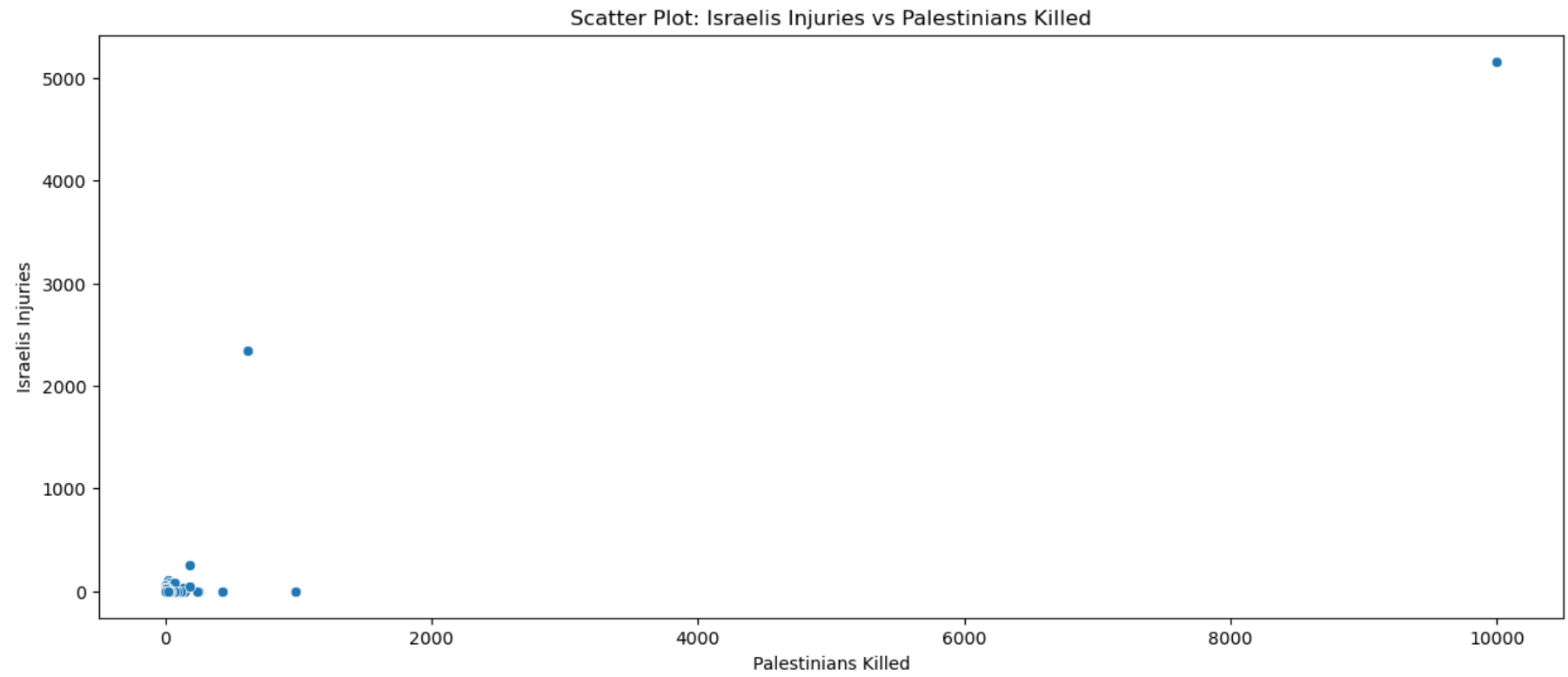




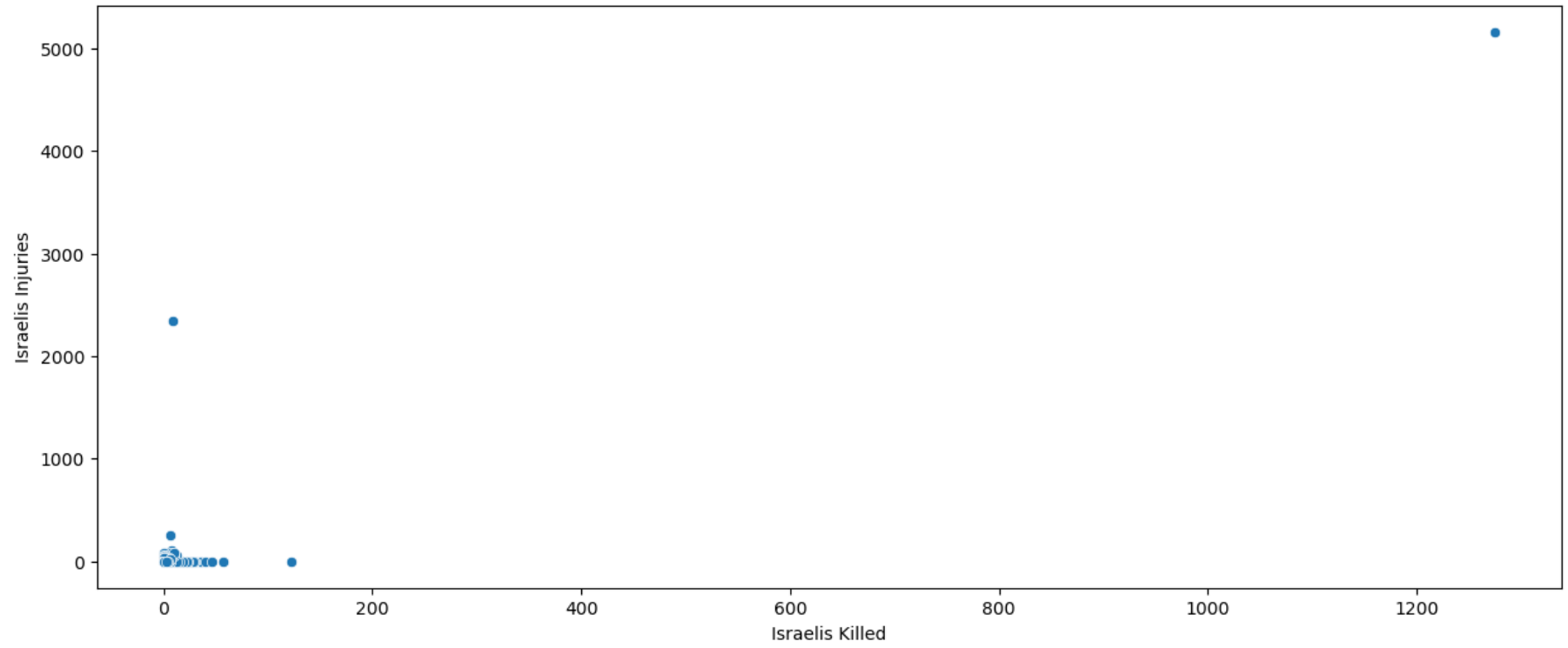


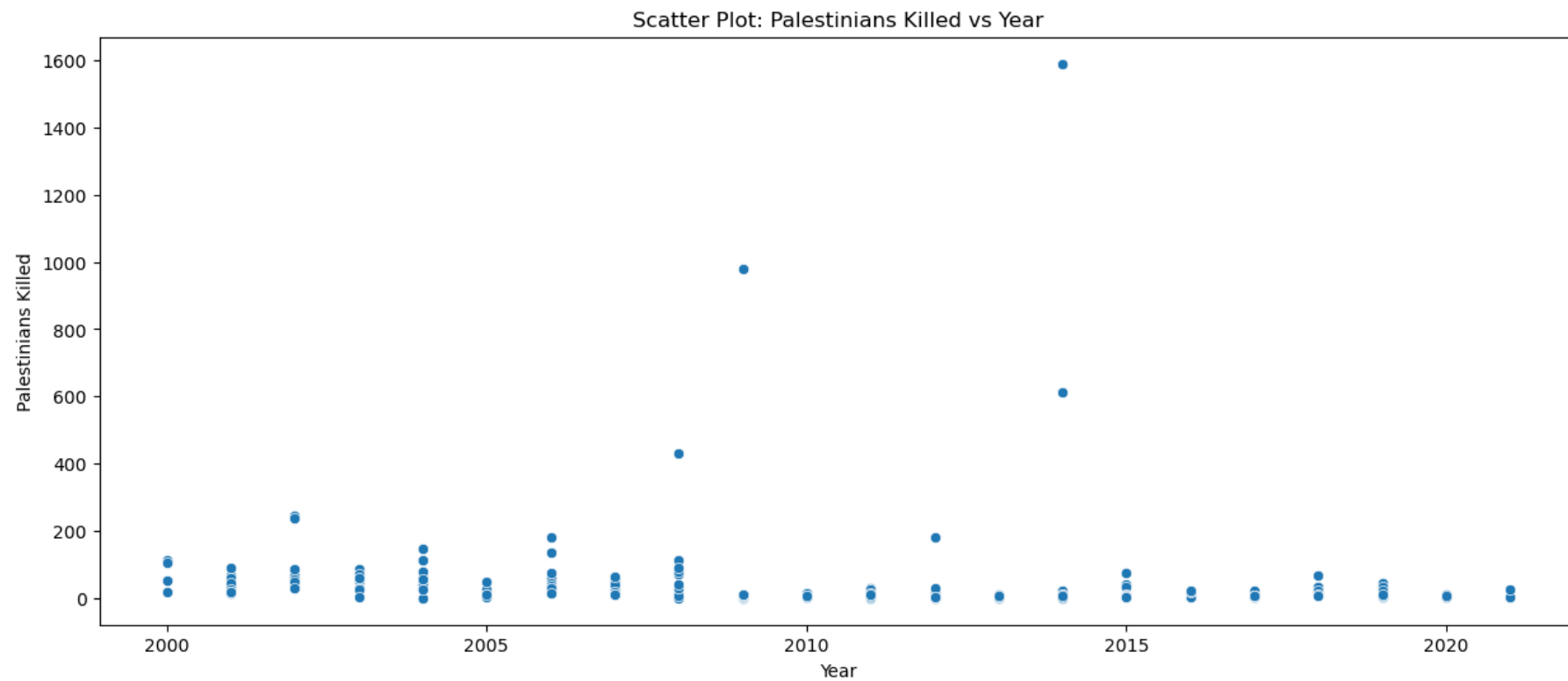




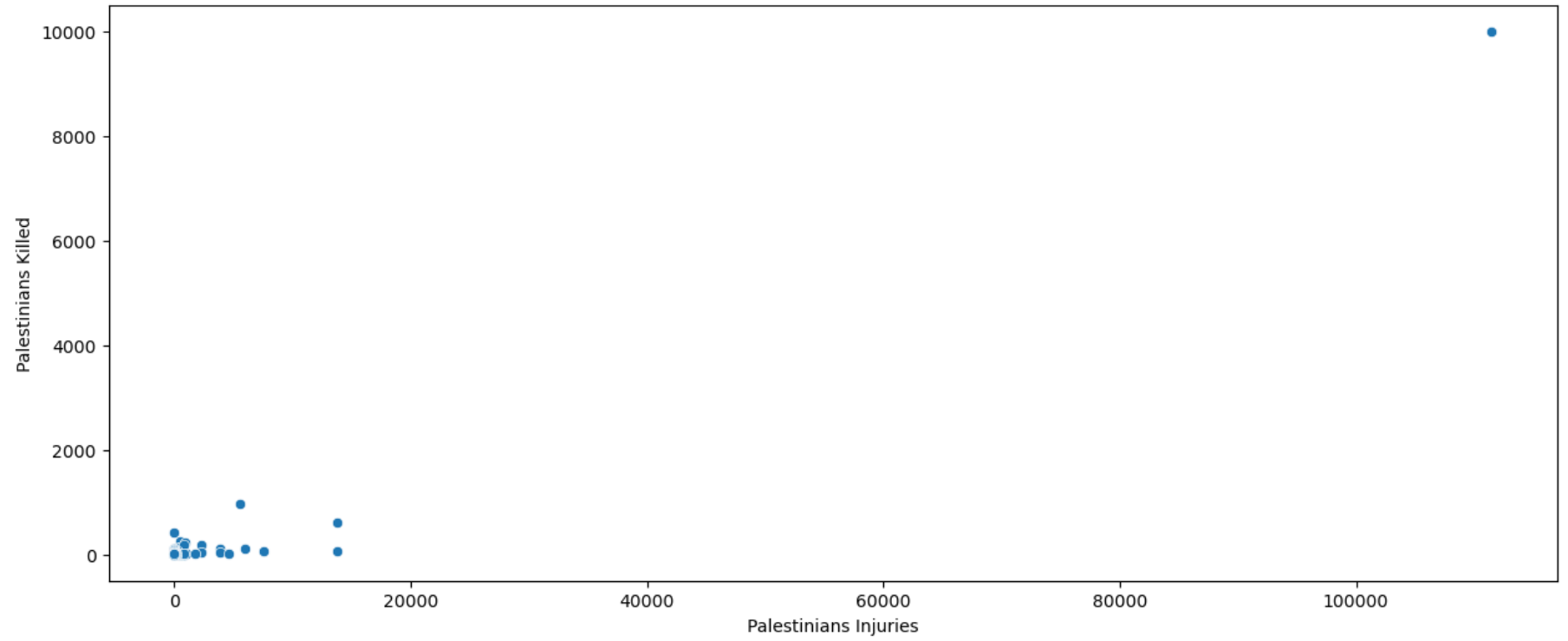


Scatter Plot: Israelis Injuries vs Israelis Killed

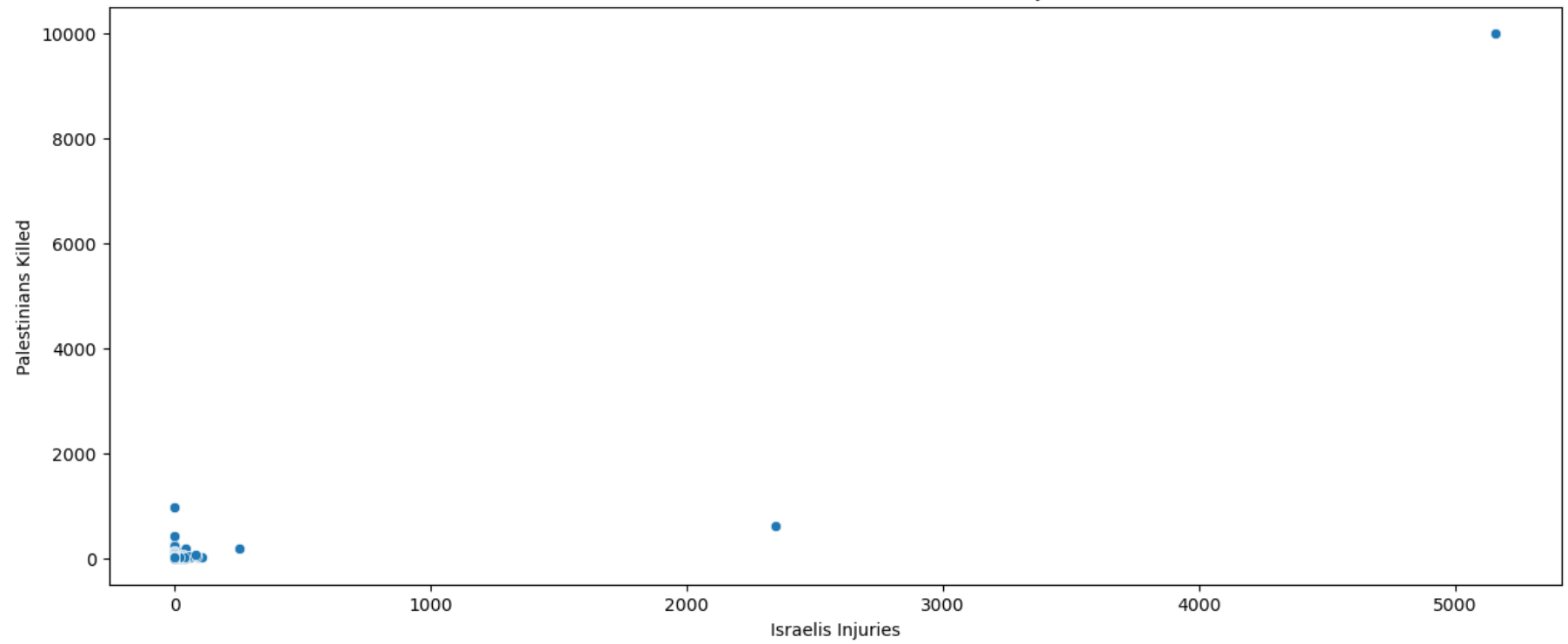




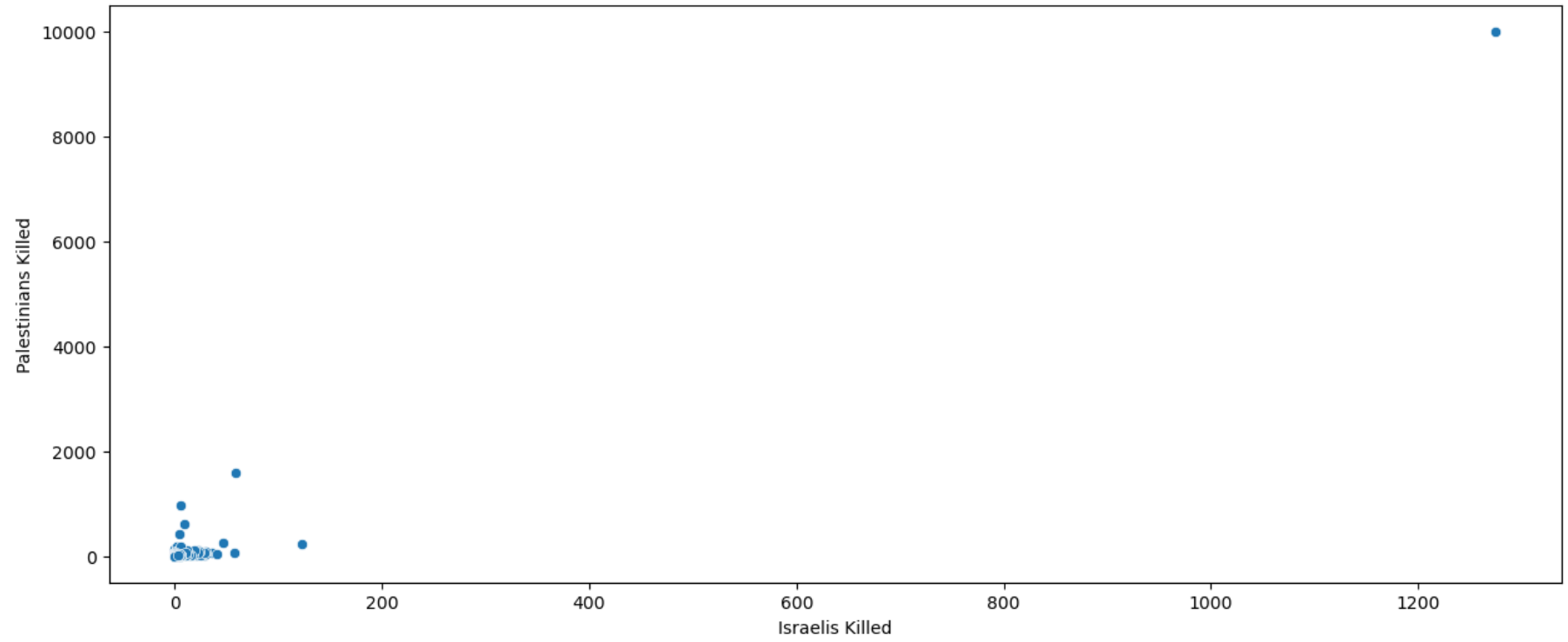
Scatter Plot: Palestinians Killed vs Palestinians Injuries

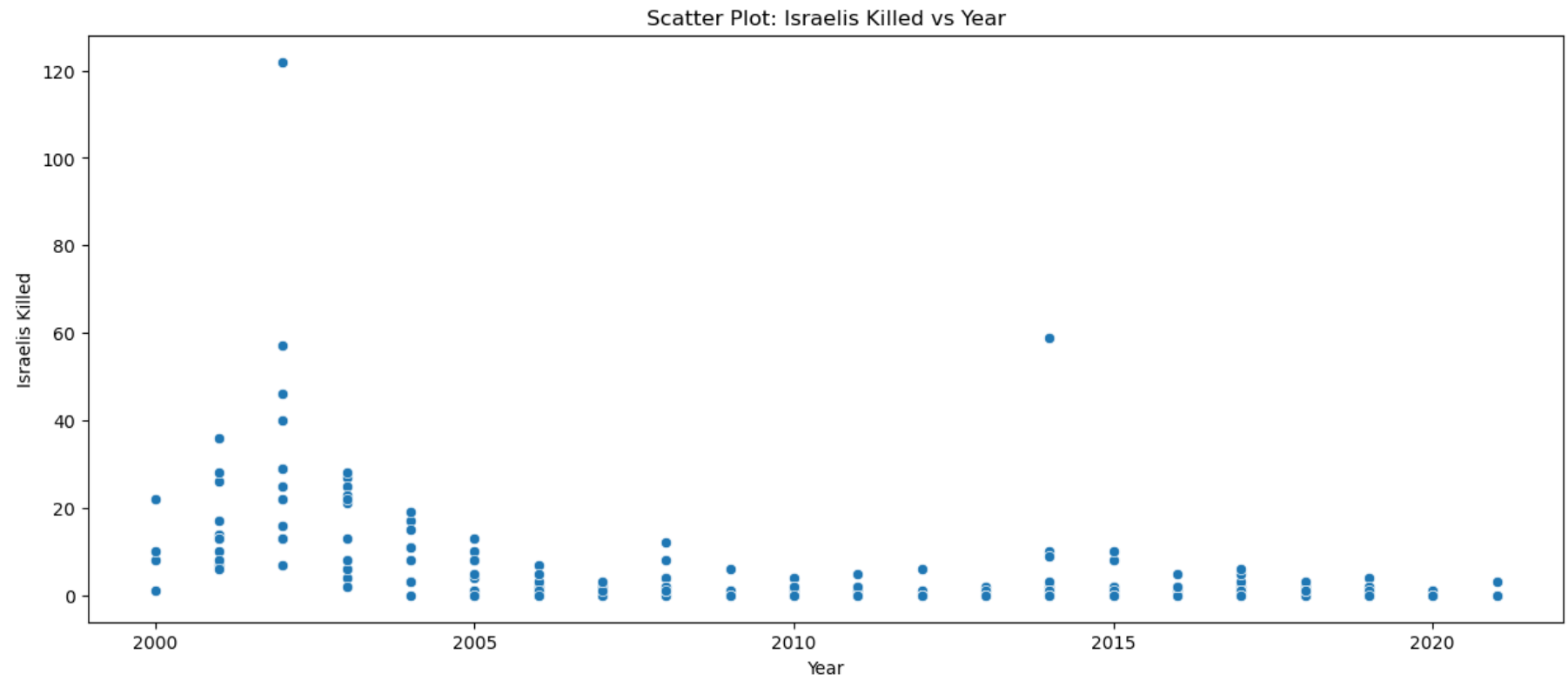


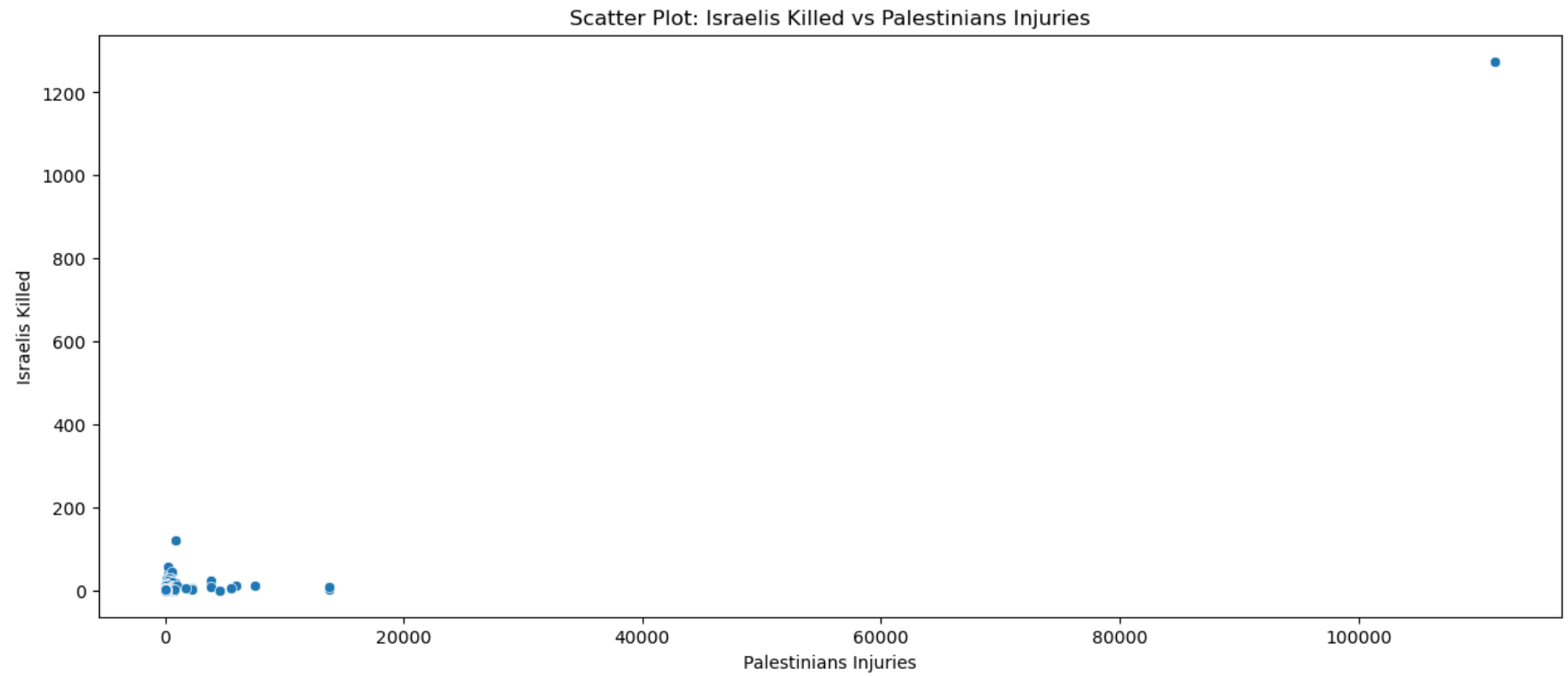
Scatter Plot: Palestinians Killed vs Israelis Injuries



Scatter Plot: Palestinians Killed vs Israelis Killed

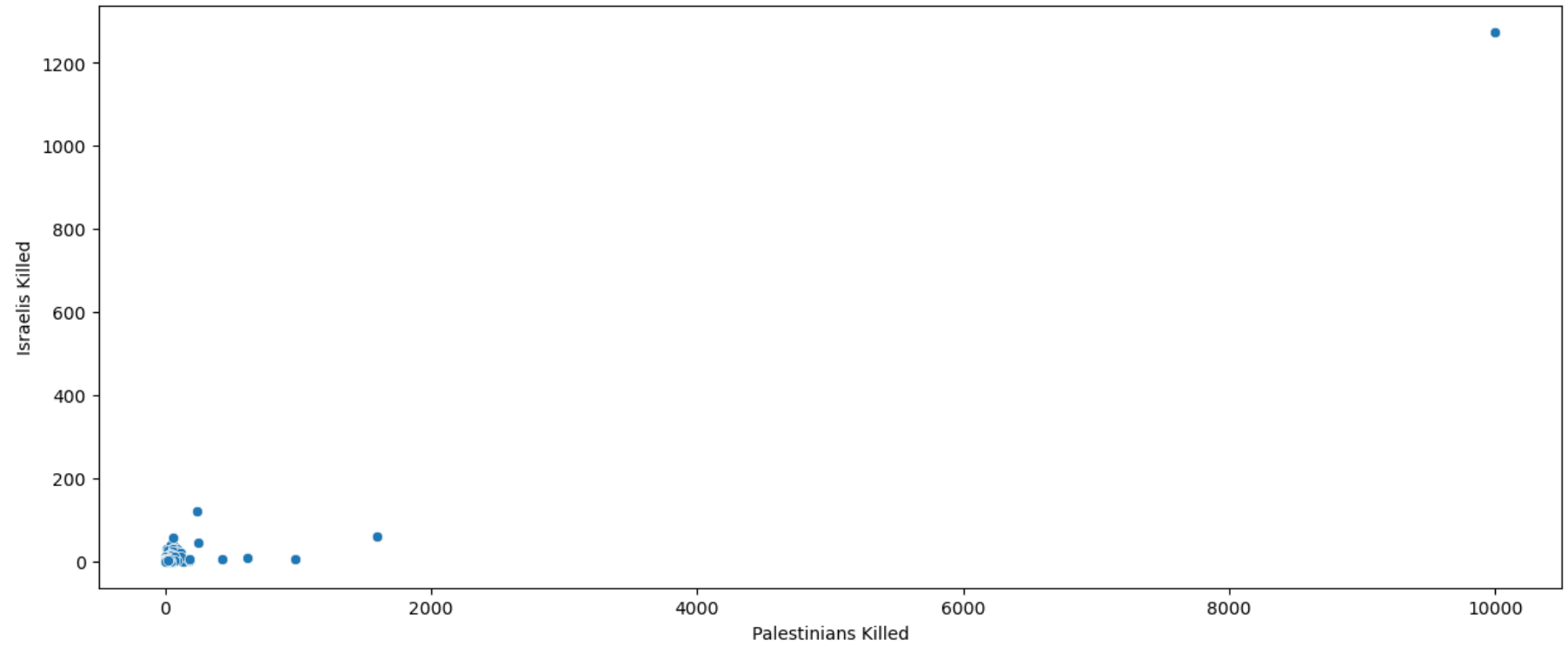




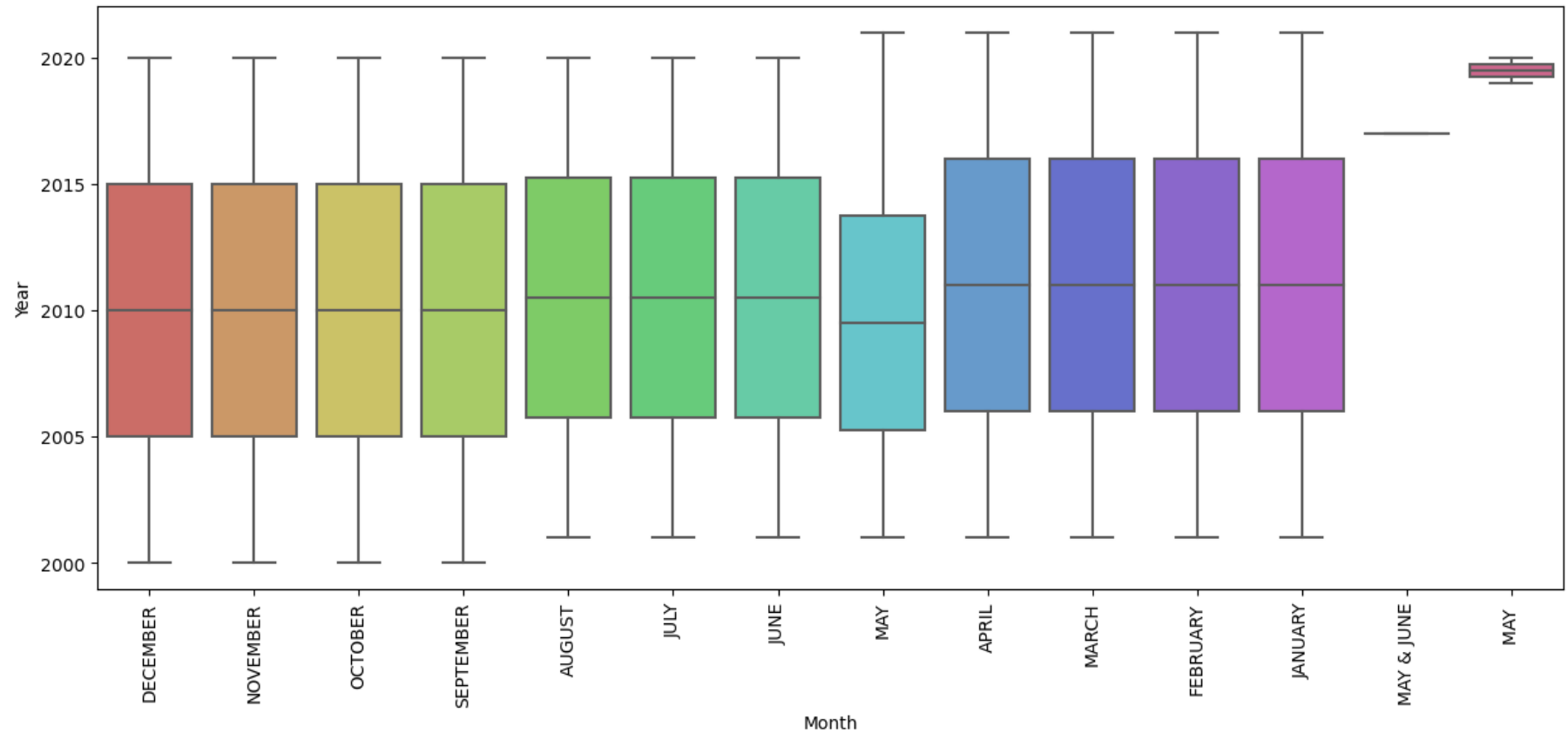


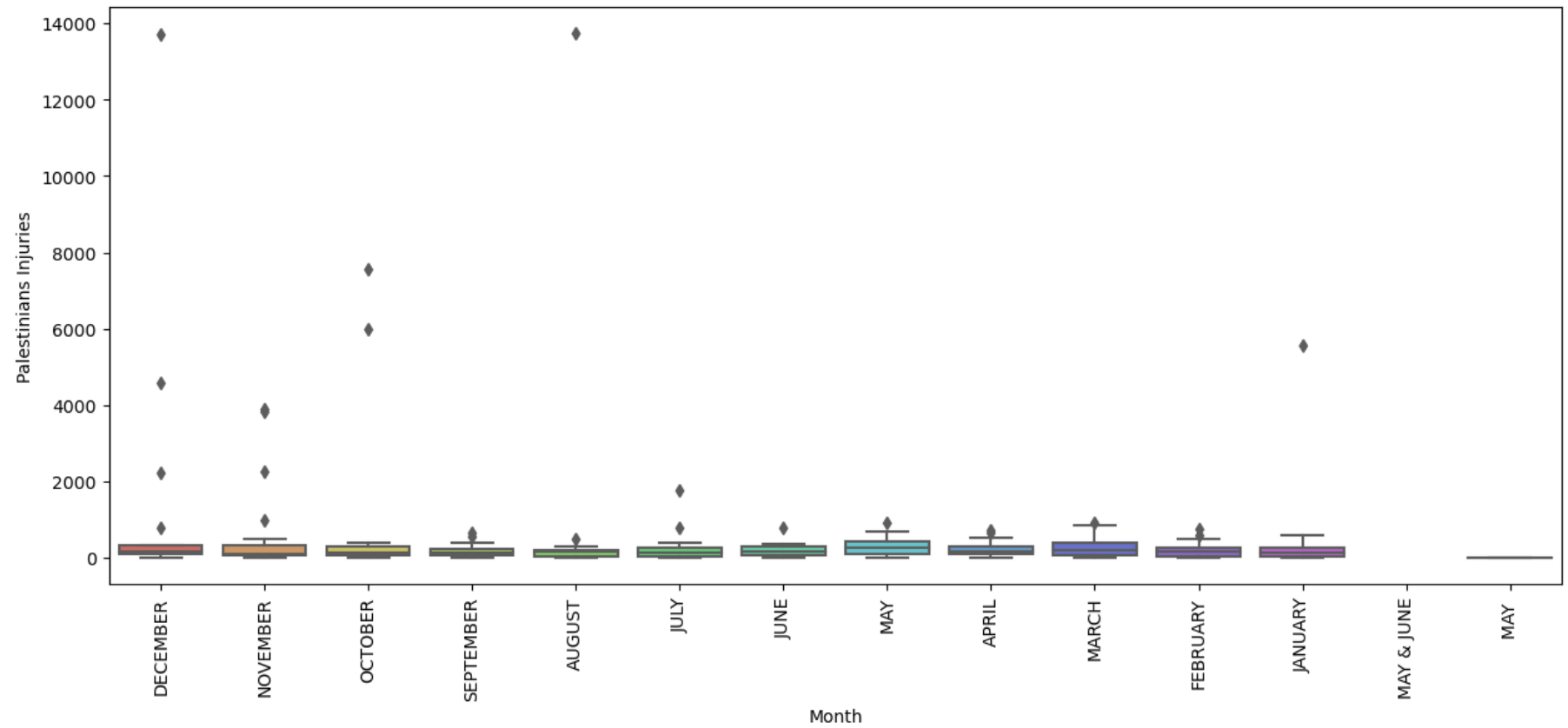


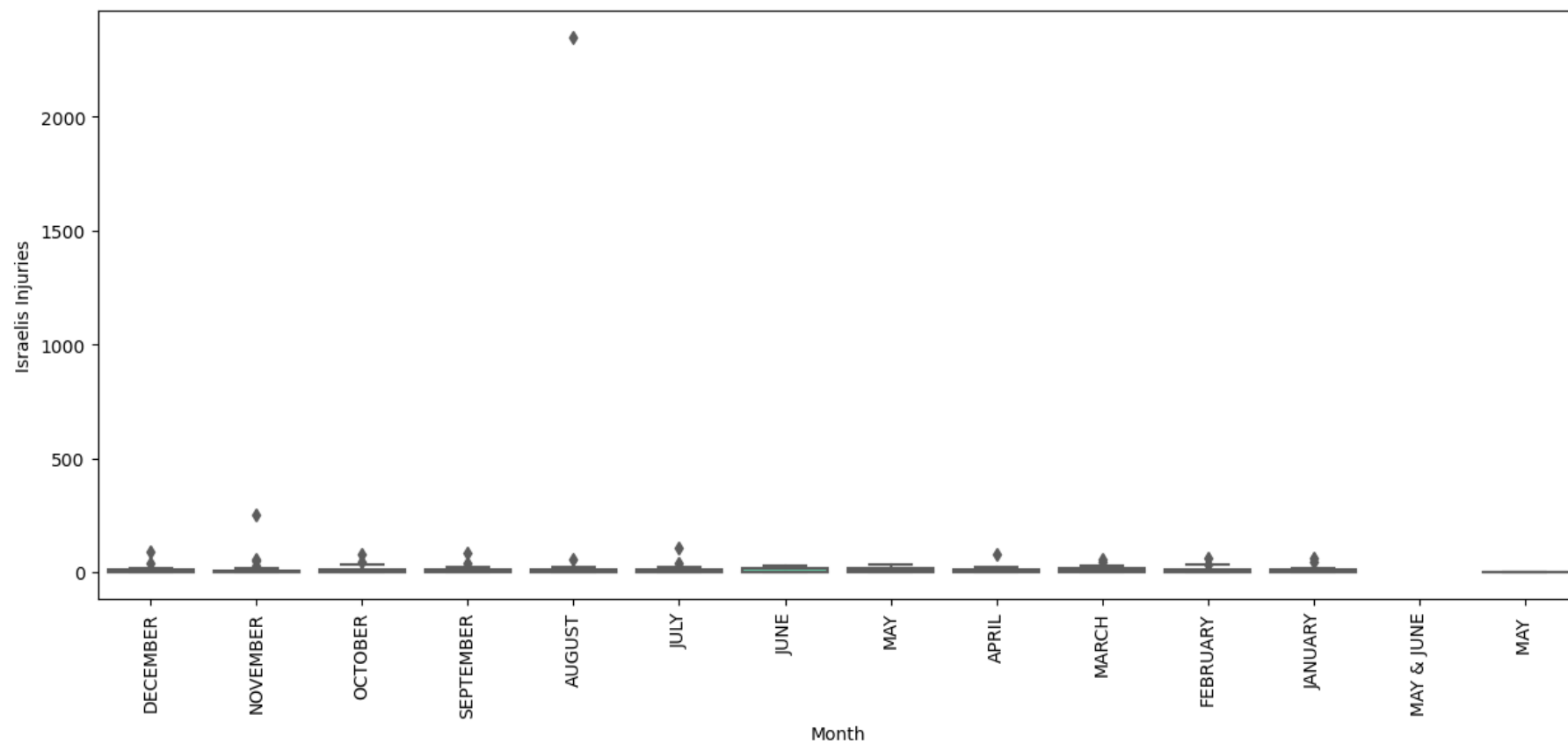
Scatter Plot: Israelis Killed vs Palestinians Killed

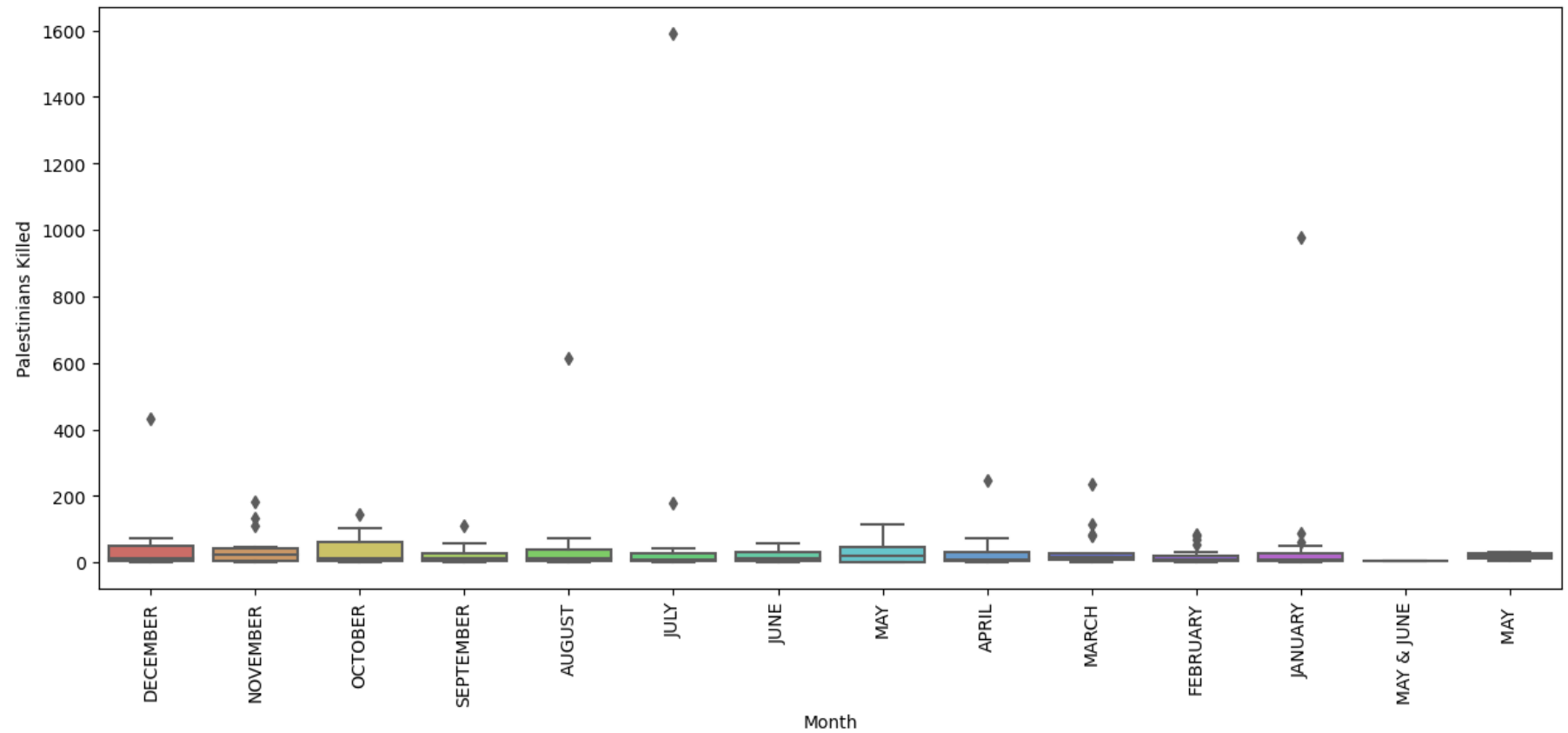


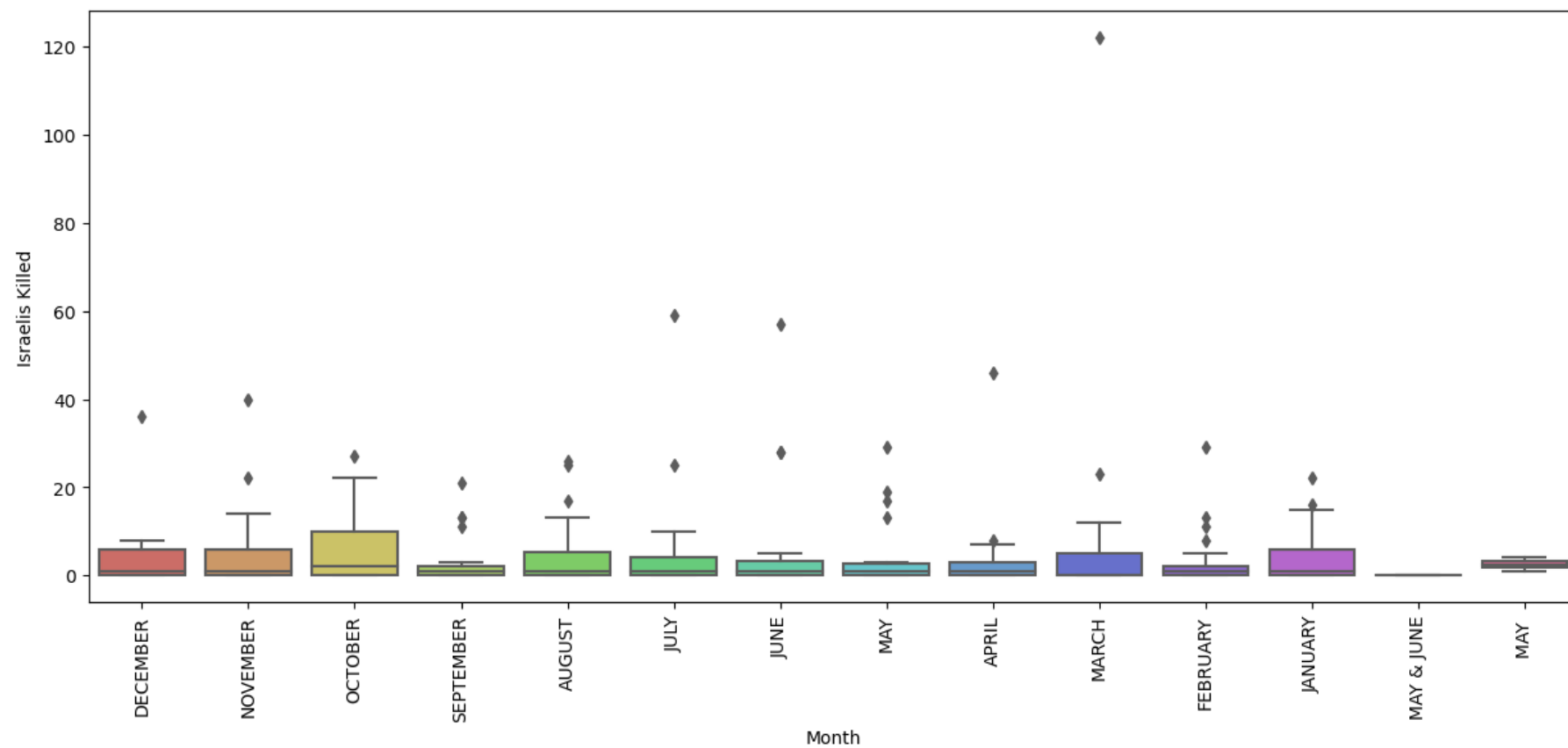
```
In [46]: for i in df1.columns:
plt.figure(figsize=(15,6))
sns.boxplot(x = df['Month'] , y = df1[i] , data = df, palette = 'hls')
plt.xticks(rotation = 90)
plt.show()
```



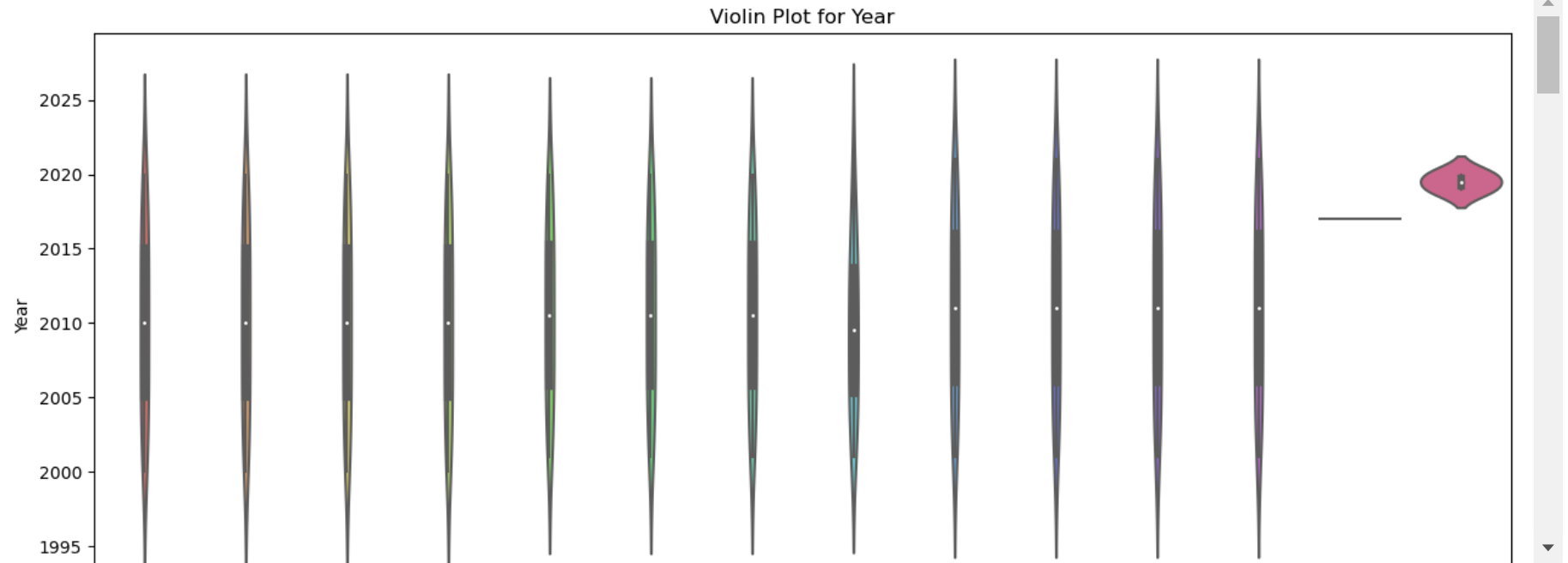








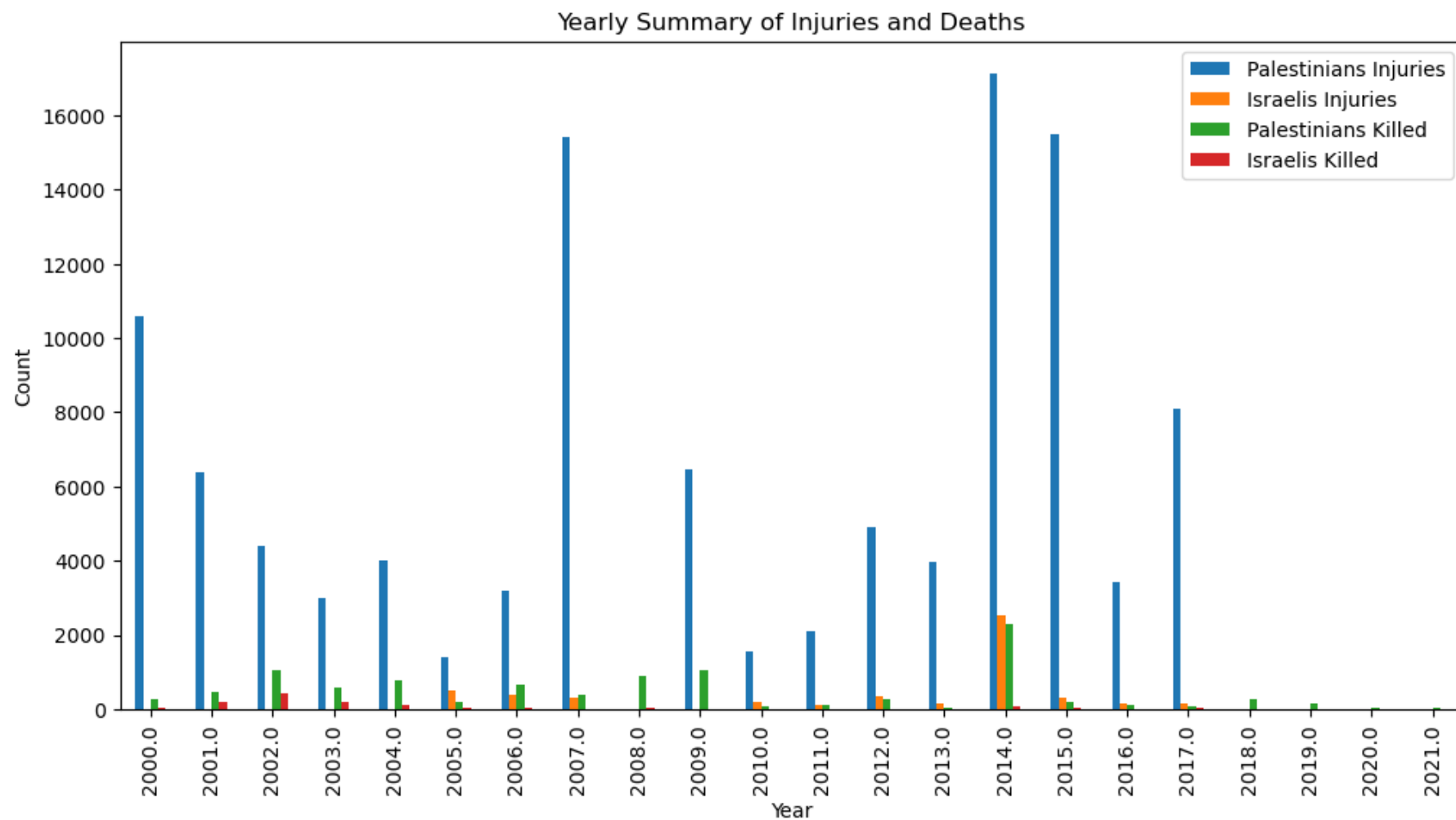
```
In [47]: for i in df1.columns:
plt.figure(figsize=(15, 6))
sns.violinplot(x=df['Month'], y=df1[i], data=df1, palette='hls')
plt.xticks(rotation=90)
plt.title(f'Violin Plot for {i}')
plt.show()
```



```
In [52]: import plotly.express as px
```

```
In [ ]: fig = px.line(df, x='Month', y=['Palestinians Injuries', 'Israelis Injuries', 'Palestin Killed', 'Israelis Killed'])
fig.show()
```

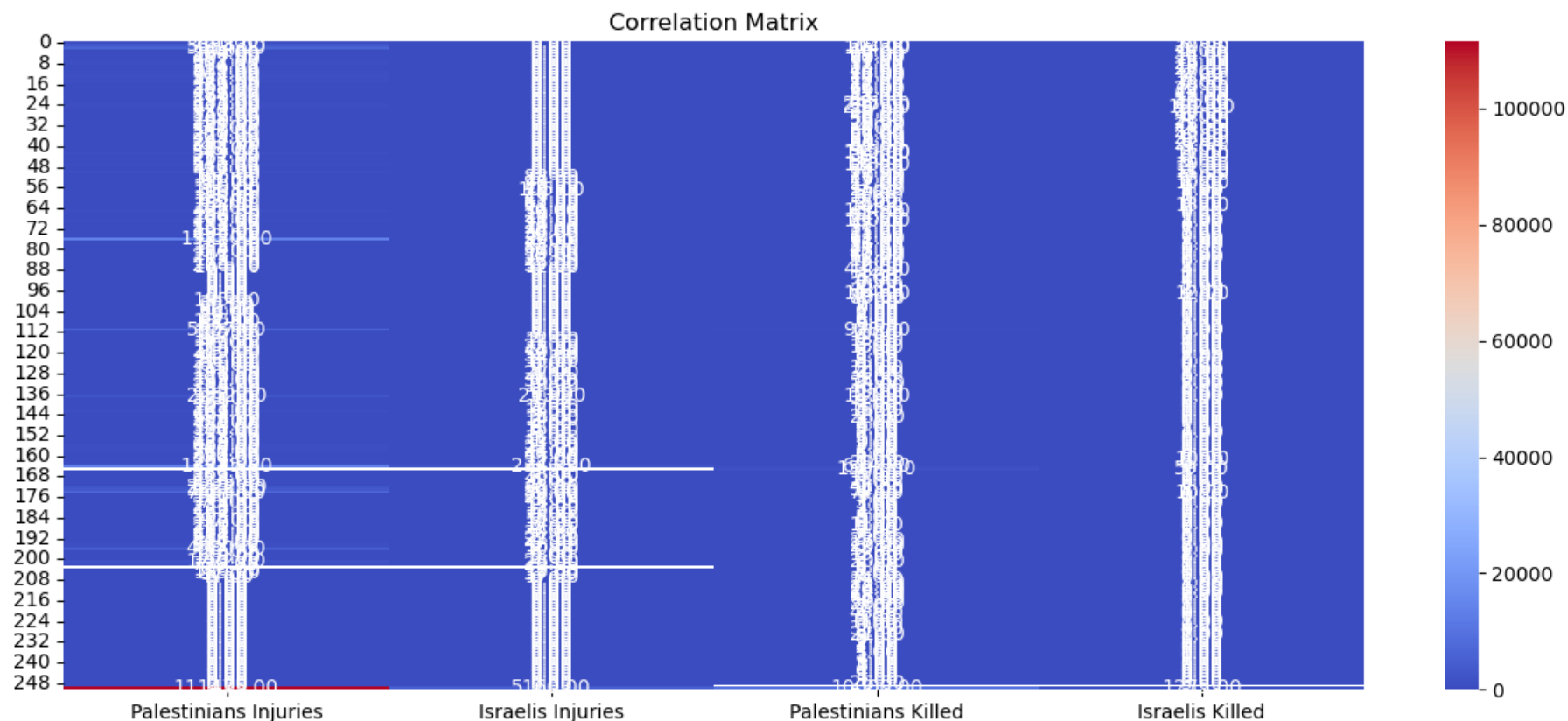
```
In [70]: yearly_summary = df.groupby('Year')[['Palestinians Injuries', 'Israelis Injuries', 'Palestinians Killed', 'Israelis Killed']]  
  
ax = yearly_summary.plot(kind='bar', figsize=(12, 6))  
ax.set_title('Yearly Summary of Injuries and Deaths')  
ax.set_xlabel('Year')  
ax.set_ylabel('Count')  
plt.show()
```



```
In [71]: fig = go.Figure()
         for col in yearly_summary.columns:
             fig.add_trace(go.Bar(x=yearly_summary.index, y=yearly_summary[col], name=col))
         fig.update_layout(
             title='Yearly Summary of Injuries and Fatalities',
             xaxis=dict(title='Year'),
             yaxis=dict(title='Count'),
             barmode='stack'
         )
         fig.show()
```



```
In [75]: plt.figure(figsize=(15,6))
correlation_matrix = df[['Palestinians Injuries', 'Israelis Injuries', 'Palestinians Killed', 'Israelis Killed']]
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt='.2f')
plt.title('Correlation Matrix')
plt.show()
```



```
In [77]: df['Total Incidents'] = df['Palestinians Injuries'] + df['Israelis Injuries'] + df['Palestinians Killed'] + df['Israelis Killed']
df['Palestinian Fatality Rate'] = df['Palestinians Killed'] / df['Total Incidents']
df['Israeli Fatality Rate'] = df['Israelis Killed'] / df['Total Incidents']
```

```
In [78]: total_incidents = df['Total Incidents'].sum()
total_palestinian_incidents = df['Total Incidents'].sum()
total_israeli_incidents = df['Total Incidents'].sum()
palestinian_proportion = total_palestinian_incidents / total_incidents
israeli_proportion = total_israeli_incidents / total_incidents
```

```
In [79]: df
```

```
Out[79]:
```

	Year	Month	Palestinians Injuries	Israelis Injuries	Palestinians Killed	Israelis Killed	Total Incidents	Palestinian Fatality Rate	Israeli Fatality Rate
0	2000.0	DECEMBER	781.0	0.0	51.0	8.0	840.0	0.060714	0.009524
1	2000.0	NOVEMBER	3838.0	0.0	112.0	22.0	3972.0	0.028197	0.005539
2	2000.0	OCTOBER	5984.0	0.0	104.0	10.0	6098.0	0.017055	0.001640
3	2000.0	SEPTEMBER	0.0	0.0	16.0	1.0	17.0	0.941176	0.058824
4	2001.0	DECEMBER	304.0	0.0	67.0	36.0	407.0	0.164619	0.088452
...
246	2021.0	MARCH	0.0	0.0	4.0	0.0	4.0	1.000000	0.000000
247	2021.0	APRIL	0.0	0.0	1.0	0.0	1.0	1.000000	0.000000
248	2021.0	MAY	0.0	0.0	26.0	3.0	29.0	0.896552	0.103448
249	NaN	NaN	0.0	0.0	NaN	NaN	NaN	NaN	NaN
250	NaN	NaN	111475.0	5160.0	10000.0	1275.0	127910.0	0.078180	0.009968

251 rows × 9 columns

```
In [80]: df.columns
```

```
Out[80]: Index(['Year', 'Month', 'Palestinians Injuries', 'Israelis Injuries',
               'Palestinians Killed', 'Israelis Killed', 'Total Incidents',
               'Palestinian Fatality Rate', 'Israeli Fatality Rate'],
              dtype='object')
```



```
In [82]: import plotly.graph_objs as go
```

```
In [90]: df['Group'] = df.apply(lambda row: 'Palestinian' if row['Palestinians Killed'] > row['Israelis Killed'] else 'Israeli'
total_incidents_by_group = df.groupby('Group')[['Total Incidents', 'Palestinians Killed', 'Israelis Killed']].sum()
fatality_rate_by_group = df.groupby('Group')[['Palestinian Fatality Rate', 'Israeli Fatality Rate']].mean()

print("Total Incidents, Palestinians Killed, Israelis Killed by Group:\n", total_incidents_by_group)
print("\nFatality Rates by Group:\n", fatality_rate_by_group)
```

Total Incidents, Palestinians Killed, Israelis Killed by Group:

	Total Incidents	Palestinians Killed	Israelis Killed
Group			
Israeli	4911.0	90.0	115.0
Palestinian	249254.0	19910.0	2435.0

Fatality Rates by Group:

	Palestinian Fatality Rate	Israeli Fatality Rate
Group		
Israeli	0.016757	0.080655
Palestinian	0.279927	0.025674

```
In [91]: total_incidents_by_group
```

```
Out[91]:
```

	Total Incidents	Palestinians Killed	Israelis Killed
Group			
Israeli	4911.0	90.0	115.0
Palestinian	249254.0	19910.0	2435.0

```
In [92]: fatality_rate_by_group
```

```
Out[92]:
```

	Palestinian Fatality Rate	Israeli Fatality Rate
Group		
Israeli	0.016757	0.080655
Palestinian	0.279927	0.025674

```
In [104]: import pandas as pd

yearly_trends = df.groupby('Year') [['Total Incidents', 'Palestinians Killed', 'Israelis Killed']]
```

```
In [105]: yearly_trends
```

```
Out[105]: <pandas.core.groupby.generic.DataFrameGroupBy object at 0x0000023EF2AC4A10>
```

```
In [98]: print("Total Incidents, Palestinians Killed, Israelis Killed by Group:\n", total_incidents_by_group.sum())
print("\nFatality Rates by Group:\n", fatality_rate_by_group.mean())
```

Total Incidents, Palestinians Killed, Israelis Killed by Group:

Total Incidents 254165.0

Palestinians Killed 20000.0

Israelis Killed 2550.0

dtype: float64

Fatality Rates by Group:

Palestinian Fatality Rate 0.148342

Israeli Fatality Rate 0.053165

dtype: float64

```
In [100]: monthly_summary = df.groupby('Month')[['Total Incidents', 'Palestinians Killed', 'Israelis Killed']]
```

```
In [103]: print(monthly_summary)
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at 0x0000023EF2AC4A50>
```

```
In [106]: print('Total Incidents:', total_incidents)
print('Total Palestinian Incidents:', total_palestinian_incidents)
print('Total Israeli Incidents:', total_israeli_incidents)
print('Proportion of Palestinian Incidents:', palestinian_proportion)
print('Proportion of Israeli Incidents:', israeli_proportion)
```

Total Incidents: 254165.0

Total Palestinian Incidents: 254165.0

Total Israeli Incidents: 254165.0

Proportion of Palestinian Incidents: 1.0

Proportion of Israeli Incidents: 1.0

```
In [107]: total_incidents = df['Total Incidents'].sum()
total_palestinian_incidents = df[df['Group'] == 'Palestinian']['Total Incidents'].sum()
total_israeli_incidents = df[df['Group'] == 'Israeli']['Total Incidents'].sum()
palestinian_proportion = total_palestinian_incidents / total_incidents
israeli_proportion = total_israeli_incidents / total_incidents
```

```
In [ ]: fig = go.Figure()
fig.add_trace(go.Scatter(x=df['Year'], y=df['Palestinian Fatality Rate'], mode='lines',
fig.add_trace(go.Scatter(x=df['Year'], y=df['Israeli Fatality Rate'], mode='lines', nam
fig.update_layout(title='Fatality Rates Over Time', xaxis_title='Year', yaxis_title='Fa
fig.show()
```

```
In [108]: import plotly.graph_objs as go

# Assuming df is your DataFrame
fig = go.Figure()

# Add a scatter trace for Palestinian Fatality Rate
fig.add_trace(go.Scatter(x=df['Year'], y=df['Palestinian Fatality Rate'], mode='lines', name='Palestinian Fatality Rat

# Add another scatter trace for Israeli Fatality Rate
fig.add_trace(go.Scatter(x=df['Year'], y=df['Israeli Fatality Rate'], mode='lines', name='Israeli Fatality Rate'))

fig.update_layout(
    title='Fatality Rates Over Time',
    xaxis_title='Year',
    yaxis_title='Fatality Rate'
)

fig.show()
```

