Maragathavalli C S

Data Science Intern

Prodigy Info Tech

Task:5

Analyze traffic accident data to identify patterns related to road conditions, weather, and time of day. Visualize accident hotspots and contributing factors.

Importing Necessary Libraries

In [11]:

```
pip install folium
```

```
Collecting foliumNote: you may need to restart the kernel to use updated packages.

  Downloading folium-0.19.7-py2.py3-none-any.whl.metadata (4.1 kB)
Collecting branca>=0.6.0 (from folium)
  Downloading branca-0.8.1-py3-none-any.whl.metadata (1.5 kB)
Requirement already satisfied: jinja2>=2.9 in c:\users\cskes\anaconda3\lib\site-packages (from folium) (3.1.4)
Requirement already satisfied: numpy in c:\users\cskes\anaconda3\lib\site-packages (from folium) (1.26.4)
Requirement already satisfied: requests in c:\users\cskes\anaconda3\lib\site-packages (from folium) (2.32.2)
Requirement already satisfied: xyzservices in c:\users\cskes\anaconda3\lib\site-packages (from folium) (2022.9.0)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\cskes\anaconda3\lib\site-packages (from jinja2>=2.9->folium) (2.1.3)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\cskes\anaconda3\lib\site-packages (from requests->folium)
(2.0.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\cskes\anaconda3\lib\site-packages (from requests->folium) (3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\cskes\anaconda3\lib\site-packages (from requests->folium) (2.2.2)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\cskes\anaconda3\lib\site-packages (from requests->folium) (2024.
7.4)
Downloading folium-0.19.7-py2.py3-none-any.whl (112 kB)
   ---------------------------------------- 0.0/112.5 kB ? eta -:--:--
   --- ------------------------------------ 10.2/112.5 kB ? eta -:--:--
   -------------- ------------------------- 41.0/112.5 kB 667.8 kB/s eta 0:00:01
   ------------------------------------ --- 102.4/112.5 kB 1.2 MB/s eta 0:00:01
   -------------------------------------- 112.5/112.5 kB 937.9 kB/s eta 0:00:00
Downloading branca-0.8.1-py3-none-any.whl (26 kB)
Installing collected packages: branca, folium
Successfully installed branca-0.8.1 folium-0.19.7
```

In [13]:

```python
# Necessary Libraries
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import folium
from folium.plugins import HeatMap
```

Loading the Dataset

In [16]:

```python
df=pd.read_csv(r"C:\Users\cskes\Downloads\us_accidents_sample.csv")
```

In [18]:

```python
print(df.head())
```

```
   ID  Severity           Start_Time             End_Time  Start_Lat  \
0  A0         3  2021-03-24 02:00:00  2021-01-22 15:00:00  36.641328   
1  A1         4  2021-08-30 22:00:00  2021-08-16 15:00:00  31.151036   
2  A2         1  2021-09-13 22:00:00  2021-10-29 09:00:00  37.684957   
3  A3         3  2021-04-28 11:00:00  2021-02-15 22:00:00  31.979007   
4  A4         3  2022-01-14 22:00:00  2021-01-25 21:00:00  41.500050   

     Start_Lng         City State Weather_Condition  Visibility(mi)  \
0  -86.534642    Charlotte    NC              Snow        6.088934   
1 -113.138399    Charlotte    AZ               Fog        1.464781   
2 -100.886555  Los Angeles    CA               Fog        3.577159   
3  -93.644221       Dallas    AZ            Cloudy        6.234069   
4  -91.516228      Houston    NC               Fog        0.627830   

   Temperature(F)  Humidity(%)  Pressure(in)  Wind_Speed(mph) Sunrise_Sunset  \
0       31.384765    67.146340     29.522971        24.474251          Night   
1       48.253771    48.987362     30.443441        20.546759            Day   
2       83.547598    93.946125     30.302680         0.772268            Day   
3       42.137568    41.574706     30.771660         0.214517          Night   
4       37.801858    34.066033     29.366949         3.414106            Day   

   Traffic_Signal  Amenity  
0           False    False  
1            True     True  
2            True    False  
3            True    False  
4           False    False  
```

Convert Start_Time to Datetime format

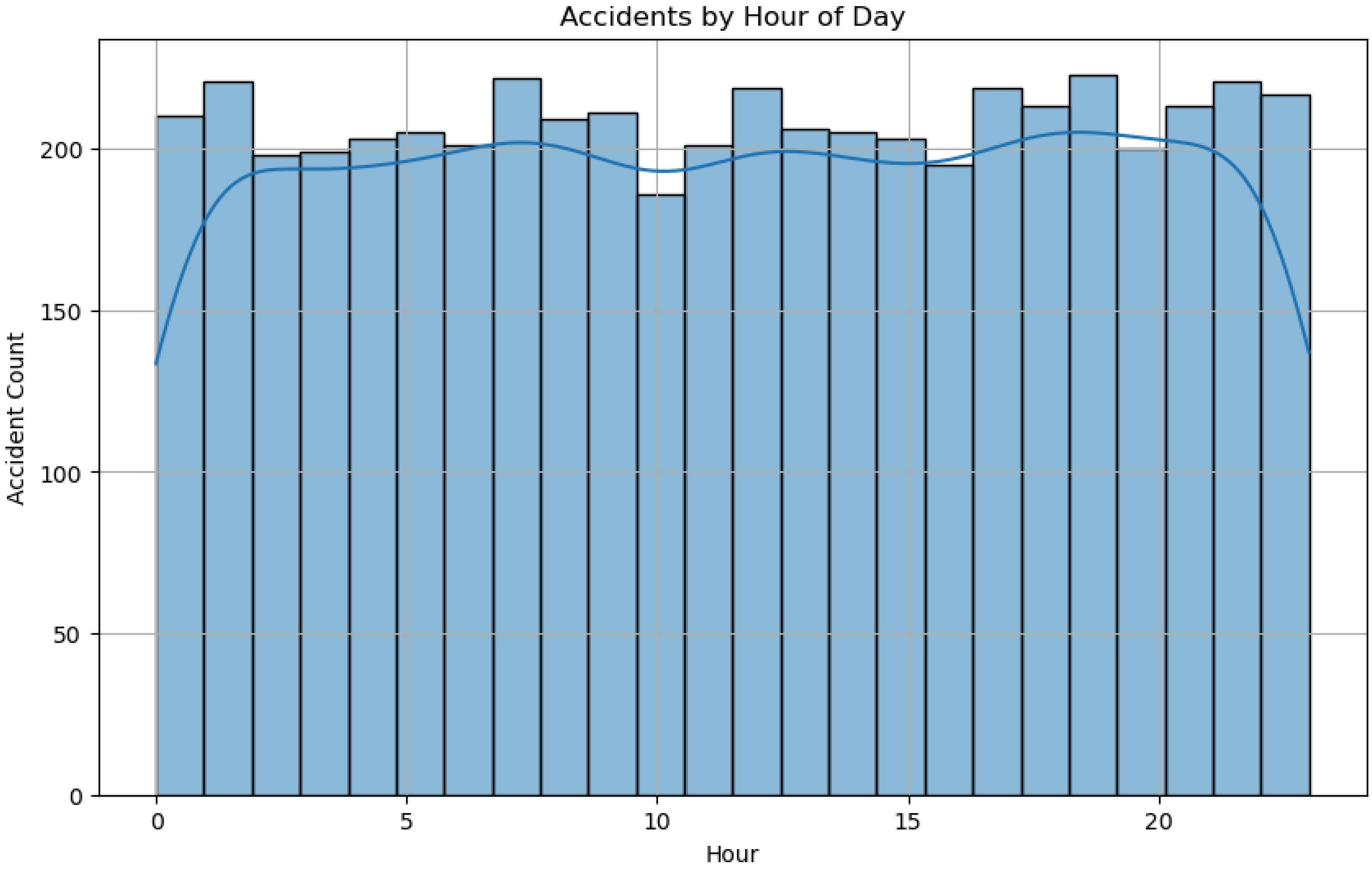
In [21]:

```
df['Start_Time'] = pd.to_datetime(df['Start_Time'])
df['Hour'] = df['Start_Time'].dt.hour
df['Day'] = df['Start_Time'].dt.day_name()
```

1. Time of Day Analysis

In [24]:

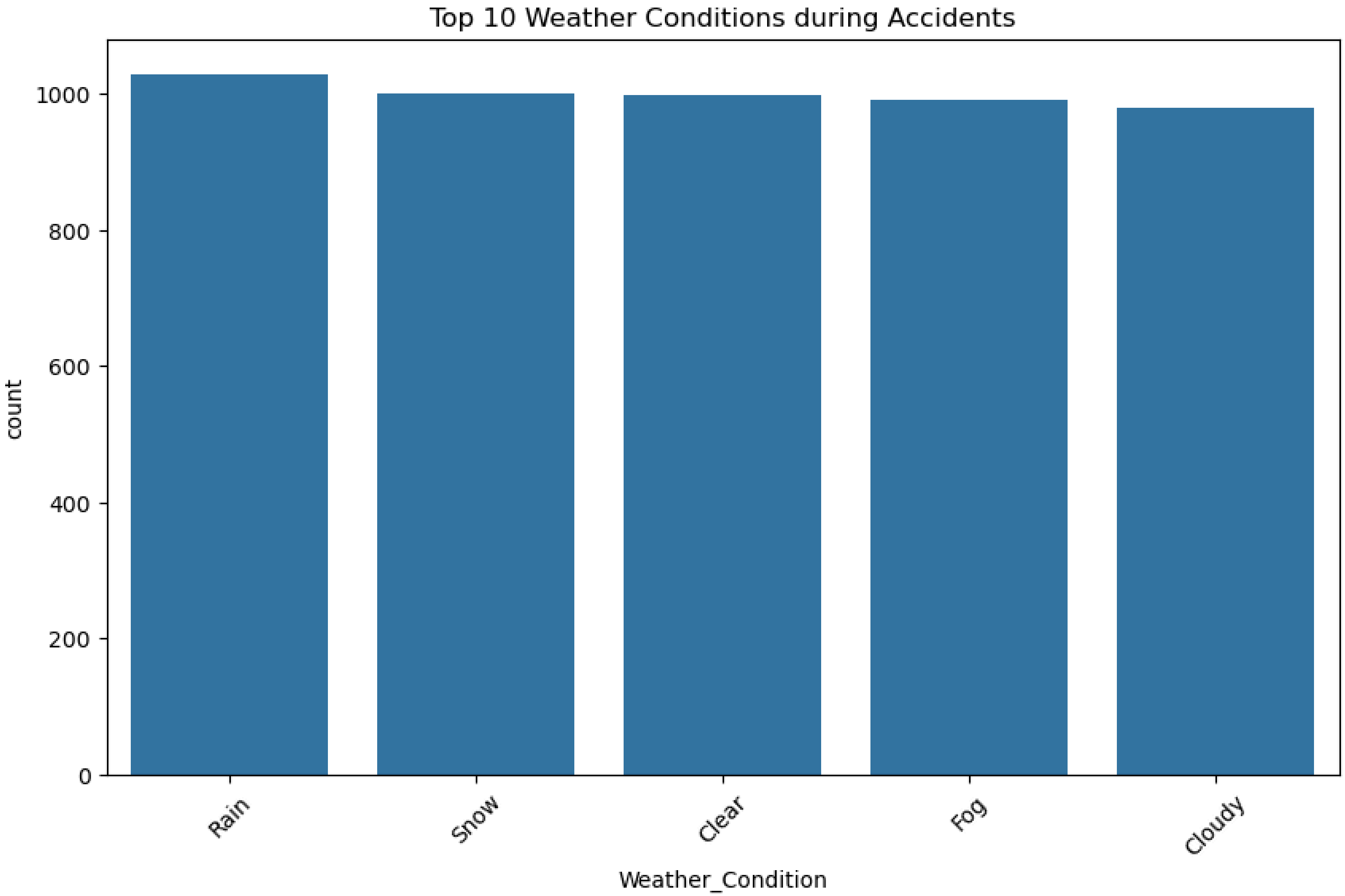
```
plt.figure(figsize=(10,6))
sns.histplot(df['Hour'],bins=24,kde=True)
plt.title('Accidents by Hour of Day')
plt.xlabel('Hour')
plt.ylabel('Accident Count')
plt.grid(True)
plt.show()
```



2. Weather Condition Analysis

In [28]:

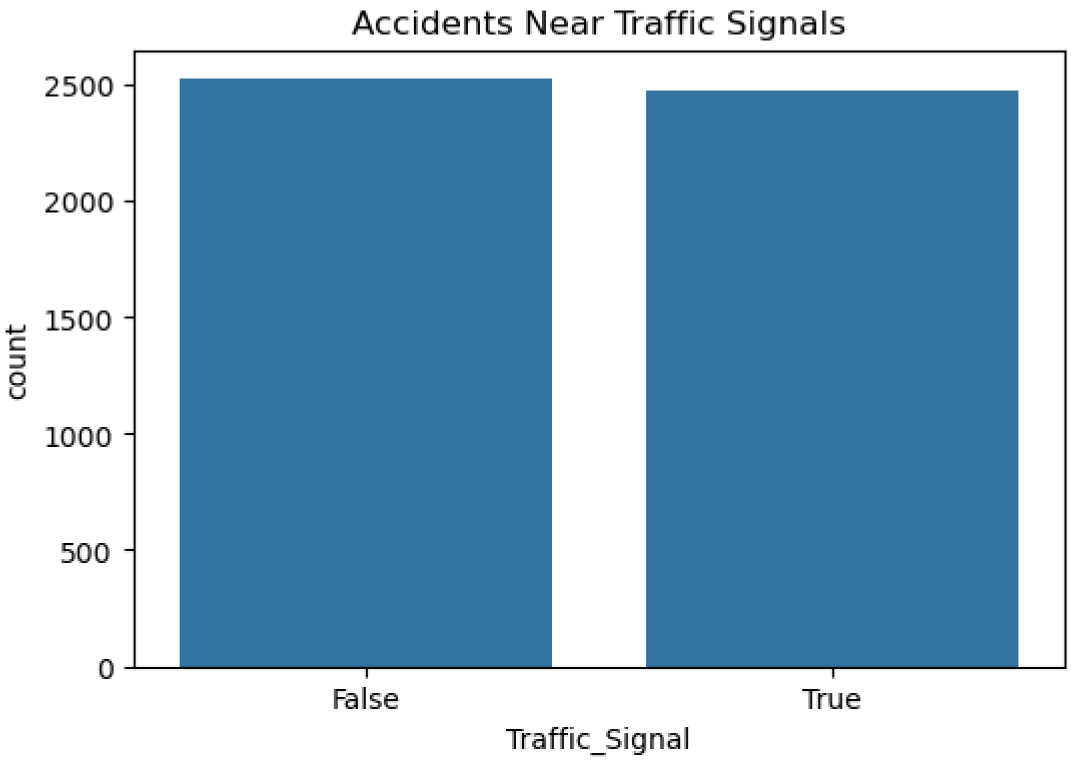
```
plt.figure(figsize=(10,6))
sns.countplot(data=df, x='Weather_Condition', order=df['Weather_Condition'].value_counts().iloc[:10].index)
plt.title('Top 10 Weather Conditions during Accidents')
plt.xticks(rotation=45)
plt.show()
```



3. Road Condition Indicators

In [31]:

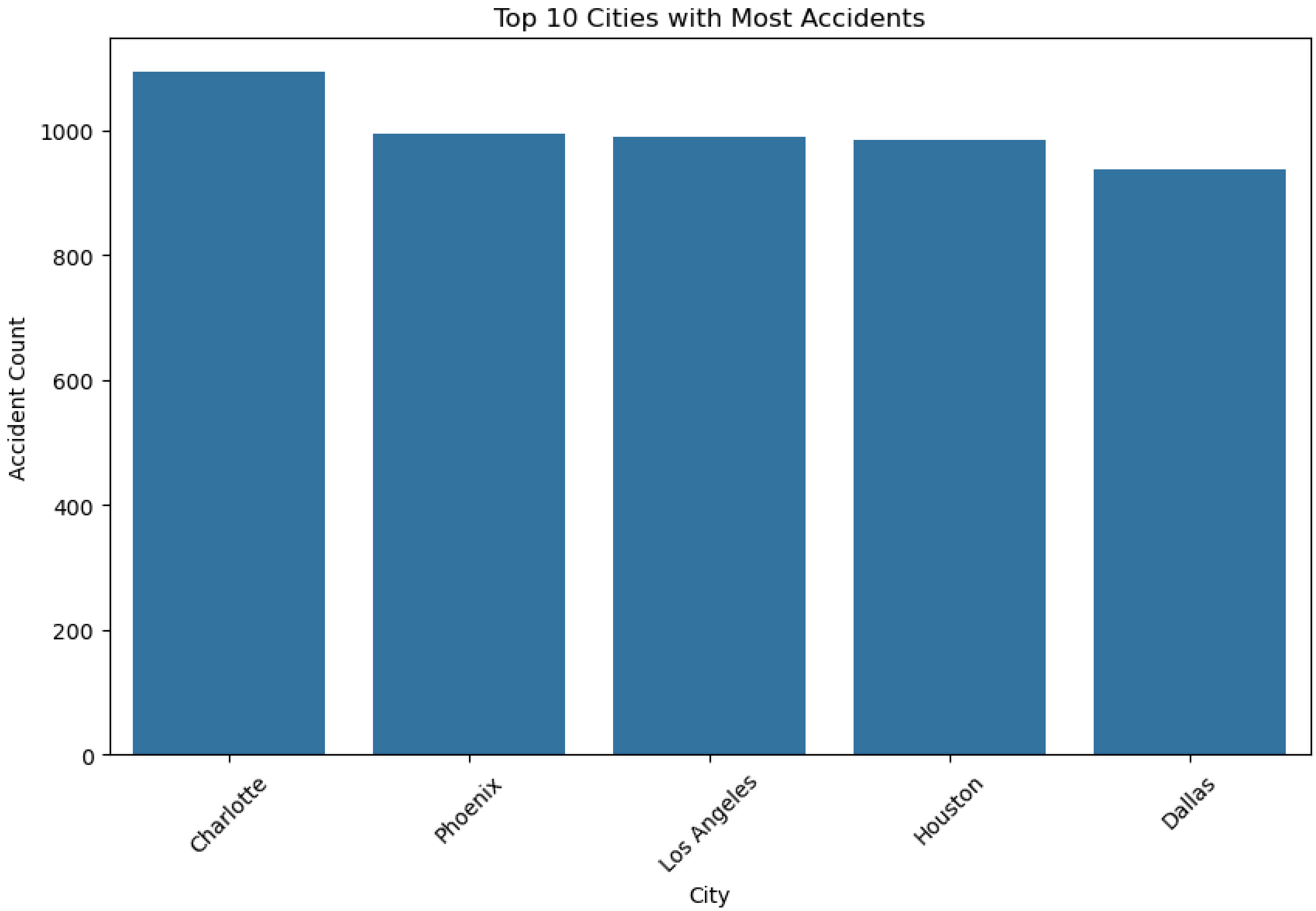
```
plt.figure(figsize=(6,4))
sns.countplot(x='Traffic_Signal',data=df)
plt.title("Accidents Near Traffic Signals")
plt.show()
```



4. Accident Hotspots (Top Cities)

In [38]:

```
plt.figure(figsize=(10,6))
top_cities = df['City'].value_counts().head(10)
sns.barplot(x=top_cities.index, y=top_cities.values)
plt.title("Top 10 Cities with Most Accidents")
plt.xticks(rotation=45)
plt.ylabel("Accident Count")
plt.show()
```



5. HeatMap of Accidents

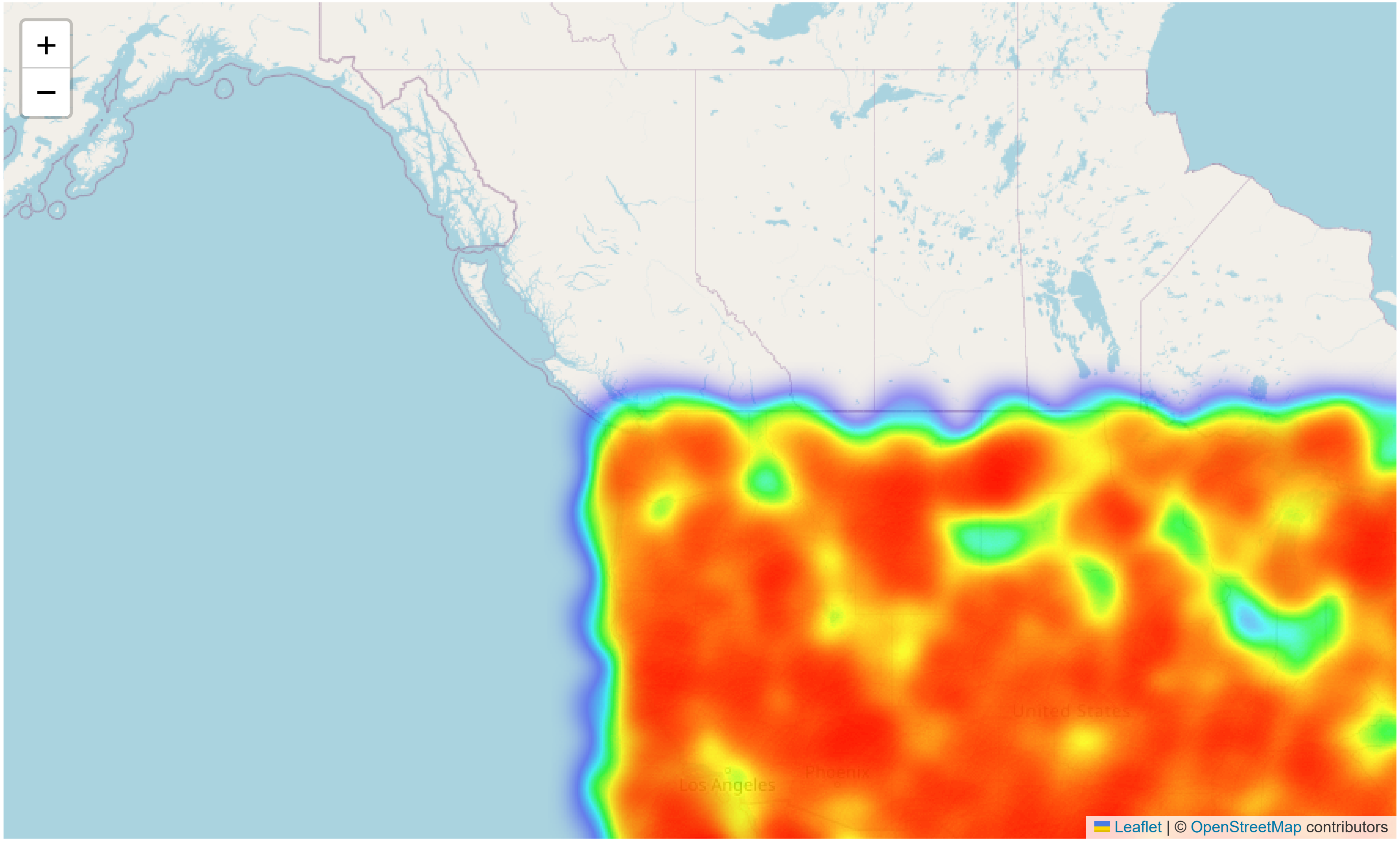
In [41]:

```
sample_df = df[['Start_Lat', 'Start_Lng']].dropna().sample(1000)
heatmap = folium.Map(location=[39.5, -98.35], zoom_start=4)
HeatMap(data=sample_df).add_to(heatmap)
heatmap.save("accident_heatmap.html")    # Open this file in browser
```

In [73]:

```
heatmap
```

Out[73]:

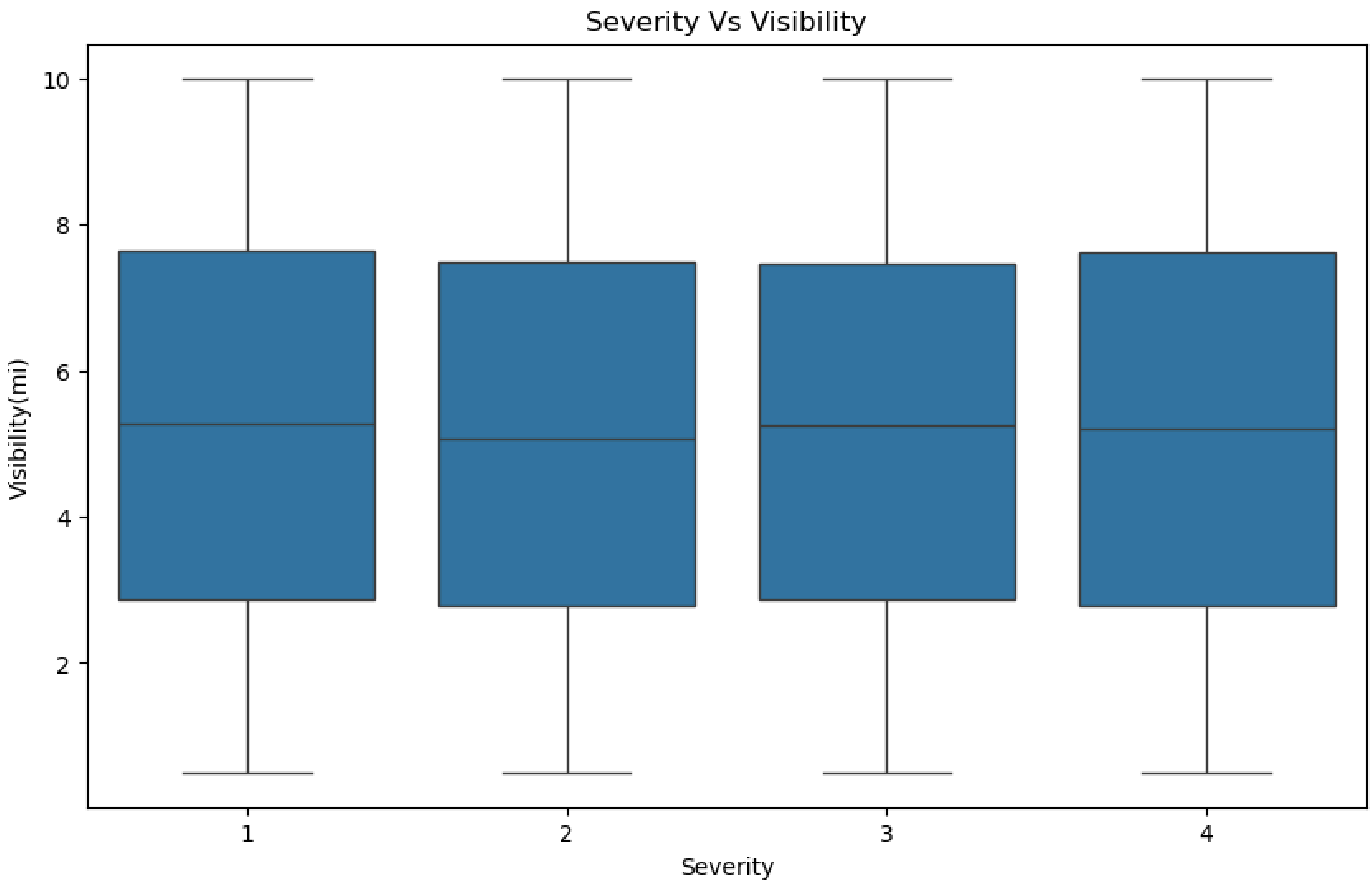


6. Contributing Factors

(i) Visibility Vs Severity

In [50]:

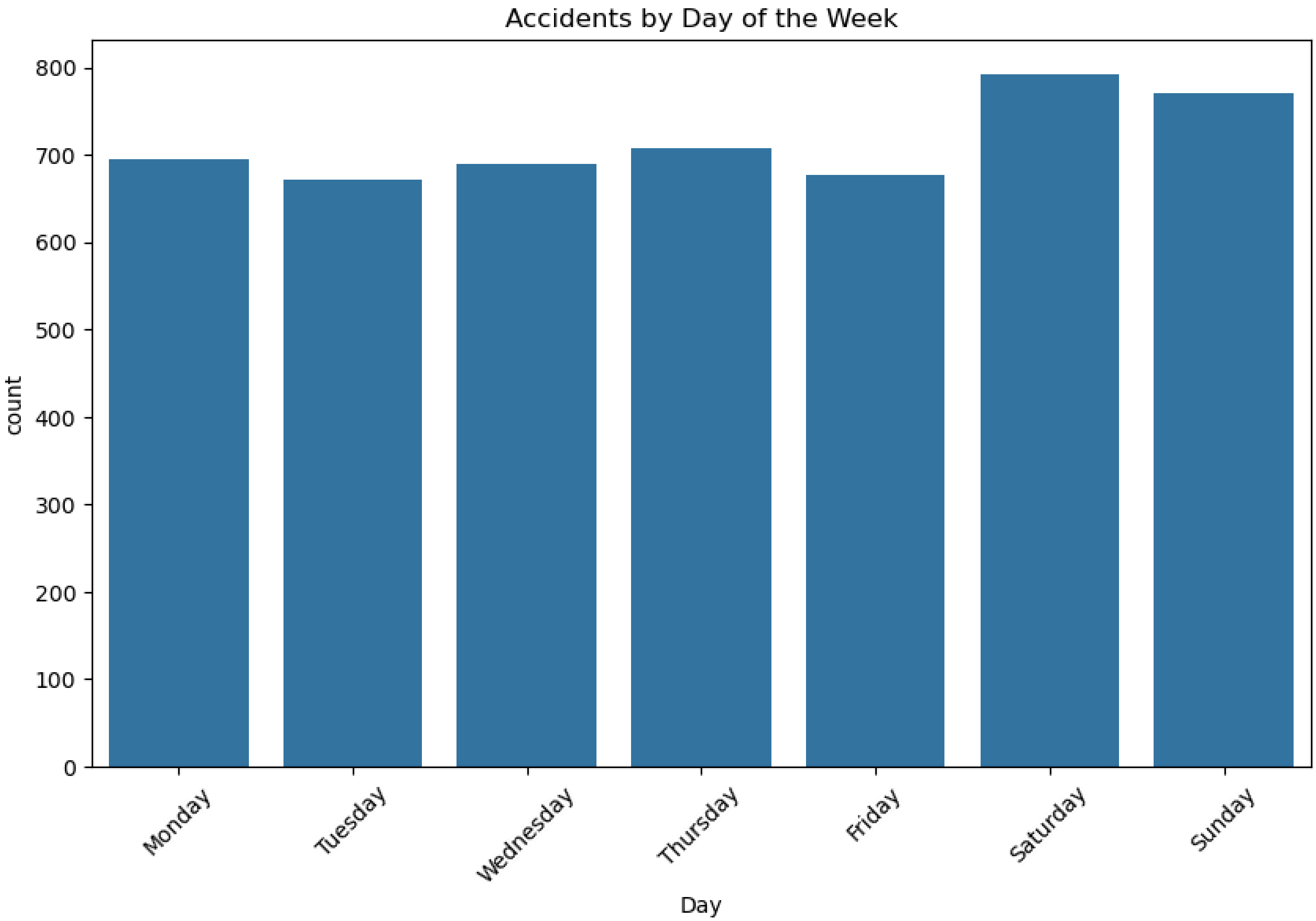
```
# Visibility Vs Severity
plt.figure(figsize=(10,6))
sns.boxplot(x='Severity', y='Visibility(mi)', data=df)
plt.title('Severity Vs Visibility')
plt.show()
```



(ii) Day of Week Analysis

In [63]:

```
# Day of Week Analysis
plt.figure(figsize=(10,6))
sns.countplot(x='Day', data=df, order=['Monday','Tuesday','Wednesday','Thursday','Friday','Saturday','Sunday'])
plt.title("Accidents by Day of the Week")
plt.xticks(rotation=45)
plt.show()
```



7. Conclusion

Key Findings:

(i) Most accidents occur during everning hours (4-6 PM).

(ii) 'Clear' and 'Rain' are common weather conditions.

(iii) Cities like Los Angeles and Dallas have high accident counts.

(iv) Visibility seems lower for high severity accidents.

(v) Traffic signals are present in many accidents spots.