In [1]:
```
import pandas as pd
import numpy as np
```

In [2]:
```
df=pd.read_csv(r"C:\Users\Admin\Downloads\road accident dataset(Sheet1).csv")
df
```

Out[2]:

	Million Plus Cities	Cause category	Cause Subcategory	Outcome of Incident	Count
0	Agra	Traffic Control	Flashing Signal/Blinker	Grevously Injured	0.0
1	Agra	Traffic Control	Flashing Signal/Blinker	Minor Injury	0.0
2	Agra	Traffic Control	Flashing Signal/Blinker	Persons Killed	0.0
3	Agra	Traffic Control	Flashing Signal/Blinker	Total Injured	0.0
4	Agra	Traffic Control	Flashing Signal/Blinker	Total number of Accidents	0.0
...	...	...	...	...	...
9545	Vizag	Weather	Sunny/Clear	Grevously Injured	561.0
9546	Vizag	Weather	Sunny/Clear	Minor Injury	252.0
9547	Vizag	Weather	Sunny/Clear	Persons Killed	176.0
9548	Vizag	Weather	Sunny/Clear	Total number of Accidents	1207.0
9549	Vizag	Weather	Sunny/Clear	Total Injured	813.0

9550 rows × 5 columns

In [3]:
```
df.shape
```

Out[3]: (9550, 5)

In [4]:
```
df.size
```

Out[4]: 47750

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

Out[5]:

	Million Plus Cities	Cause category	Cause Subcategory	Outcome of Incident	Count
0	Agra	Traffic Control	Flashing Signal/Blinker	Grevously Injured	0.0
1	Agra	Traffic Control	Flashing Signal/Blinker	Minor Injury	0.0
2	Agra	Traffic Control	Flashing Signal/Blinker	Persons Killed	0.0
3	Agra	Traffic Control	Flashing Signal/Blinker	Total Injured	0.0
4	Agra	Traffic Control	Flashing Signal/Blinker	Total number of Accidents	0.0

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

Out[6]:

	Million Plus Cities	Cause category	Cause Subcategory	Outcome of Incident	Count
9545	Vizag	Weather	Sunny/Clear	Grevously Injured	561.0
9546	Vizag	Weather	Sunny/Clear	Minor Injury	252.0
9547	Vizag	Weather	Sunny/Clear	Persons Killed	176.0
9548	Vizag	Weather	Sunny/Clear	Total number of Accidents	1207.0
9549	Vizag	Weather	Sunny/Clear	Total Injured	813.0

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9550 entries, 0 to 9549
Data columns (total 5 columns):
 #   Column                 Non-Null Count  Dtype  
---  ------                 --------------  -----  
 0   Million Plus Cities    9550 non-null   object 
 1   Cause category         9550 non-null   object 
 2   Cause Subcategory      9550 non-null   object 
 3   Outcome of Incident    9550 non-null   object 
 4   Count                  9550 non-null   float64
dtypes: float64(1), object(4)
memory usage: 373.2+ KB
```

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

Out[8]:
```
Index(['Million Plus Cities', 'Cause category', 'Cause Subcategory',
       'Outcome of Incident', 'Count '],
      dtype='object')
```

In [9]:
```
df.isnull().sum()
```

Out[9]:
```
Million Plus Cities    0
Cause category         0
Cause Subcategory      0
Outcome of Incident    0
Count                  3
dtype: int64
```

In [10]:
```
df.fillna(0)
```

Out[10]:

	Million Plus Cities	Cause category	Cause Subcategory	Outcome of Incident	Count
0	Agra	Traffic Control	Flashing Signal/Blinker	Grevously Injured	0.0
1	Agra	Traffic Control	Flashing Signal/Blinker	Minor Injury	0.0
2	Agra	Traffic Control	Flashing Signal/Blinker	Persons Killed	0.0
3	Agra	Traffic Control	Flashing Signal/Blinker	Total Injured	0.0
4	Agra	Traffic Control	Flashing Signal/Blinker	Total number of Accidents	0.0
...	...	...	...	...	...
9545	Vizag	Weather	Sunny/Clear	Grevously Injured	561.0
9546	Vizag	Weather	Sunny/Clear	Minor Injury	252.0
9547	Vizag	Weather	Sunny/Clear	Persons Killed	176.0
9548	Vizag	Weather	Sunny/Clear	Total number of Accidents	1207.0
9549	Vizag	Weather	Sunny/Clear	Total Injured	813.0

9550 rows × 5 columns

In [11]:
```
df["Million Plus Cities"].value_counts()
```

Out[11]:
```
Million Plus Cities
Agra                 191
Ahmedabad            191
Allahabad (Prayagraj)191
Amritsar             191
Asansol Durgapur     191
Aurangabad           191
Bengaluru            191
Bhopal               191
Chandigarh           191
Chennai              191
Coimbatore           191
Delhi                191
Dhanbad              191
Faridabad            191
Ghaziabad            191
Gwalior              191
Hyderabad            191
Indore               191
Jabalpur             191
Jaipur               191
Jamshedpur           191
Jodhpur              191
Kannur               191
Kanpur               191
Kozhikode            191
Kochi                191
Kolkata              191
Kollam               191
Kota                 191
Lucknow              191
Ludhiana             191
Madurai              191
Meerut               191
Mumbai               191
Nagpur               191
Nashik               191
Patna                191
Pune                 191
Raipur               191
Rajkot               191
Ranchi               191
Surat                191
Thiruvanthapuram     191
Thrissur             191
Tiruchirapalli       191
Vadodra              191
Varanasi             191
Vijayawada city      191
Vizag                191
Name: count, dtype: int64
```

In [12]:
```
df["Cause Subcategory"].value_counts()
```

Out[12]:
```
Cause Subcategory
Unclassified                                    1650
Flashing Signal/Blinker                          250
Police Controlled                                250
Stop Sign                                        250
Traffic Light Signal                             250
Uncontrolled                                     250
Four arm Junction                                250
Round about Junction                             250
Staggered Junction                               250
T-                                               250
Y                                                250
Driving on Wrong side                            250
Drunken Driving/ Consumption of alcohol and drug 250
Jumping Red Light                                250
Over                                             250
Use of Mobile Phone                              250
Bridge                                           250
Culvert                                          250
Curved Road                                      250
Ongoing Road Works/Under Construction            250
Pot Holes                                        250
Steep Grade                                      250
Straight Road                                    250
Fog and Misty                                    250
Rain/Sleet                                       250
Rainy                                            250
Sunny/Clear                                      250
Cars, Taxis, Vans and LMV                        200
Buses                                            200
Bicycles                                         200
Three Wheelers                                   200
Two Wheelers                                     200
Trucks/Lorries                                   200
Pedestrian                                       200
Other Non                                        200
Name: count, dtype: int64
```

In [13]:
```
df["Outcome of Incident"].value_counts()
```

Out[13]:
```
Outcome of Incident
Grevously Injured    2000
Minor Injury         2000
Persons Killed       2000
Total number of Accidents  2600
Total Injured        1150
Name: count, dtype: int64
```

In [14]:
```
df=pd.read_csv(r"C:\Users\Admin\Downloads\road accident dataset(Sheet1).csv",index_col="Million Plus Cities")
df
```

Out[14]:

Million Plus Cities	Cause category	Cause Subcategory	Outcome of Incident	Count
Agra	Traffic Control	Flashing Signal/Blinker	Grevously Injured	0.0
Agra	Traffic Control	Flashing Signal/Blinker	Minor Injury	0.0
Agra	Traffic Control	Flashing Signal/Blinker	Persons Killed	0.0
Agra	Traffic Control	Flashing Signal/Blinker	Total Injured	0.0
Agra	Traffic Control	Flashing Signal/Blinker	Total number of Accidents	0.0
...	...	...	...	...
Vizag	Weather	Sunny/Clear	Grevously Injured	561.0
Vizag	Weather	Sunny/Clear	Minor Injury	252.0
Vizag	Weather	Sunny/Clear	Persons Killed	176.0
Vizag	Weather	Sunny/Clear	Total number of Accidents	1207.0
Vizag	Weather	Sunny/Clear	Total Injured	813.0

9550 rows × 4 columns

In [15]:
```
df.sort_index(ascending=False)
```

Out[15]:

Million Plus Cities	Cause category	Cause Subcategory	Outcome of Incident	Count
Vizag	Weather	Sunny/Clear	Total Injured	813.0
Vizag	Junction	Y	Grevously Injured	25.0
Vizag	Traffic Violation	Over	Minor Injury	277.0
Vizag	Traffic Violation	Over	Grevously Injured	590.0
Vizag	Traffic Violation	Others	Total Injured	304.0
...	...	...	...	...
Agra	Traffic Violation	Use of Mobile Phone	Grevously Injured	8.0
Agra	Traffic Violation	Use of Mobile Phone	Minor Injury	3.0
Agra	Traffic Violation	Use of Mobile Phone	Total number of Accidents	16.0
Agra	Traffic Violation	Use of Mobile Phone	Persons Killed	9.0
Agra	Traffic Control	Flashing Signal/Blinker	Grevously Injured	0.0

9550 rows × 4 columns

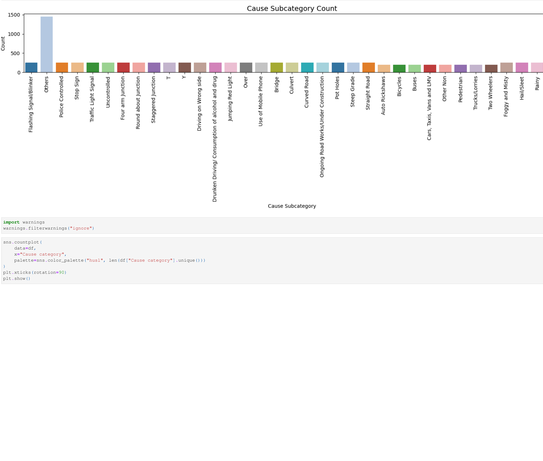
In [16]:
```
df
```

Out[16]:

Million Plus Cities	Cause category	Cause Subcategory	Outcome of Incident	Count
Agra	Traffic Control	Flashing Signal/Blinker	Grevously Injured	0.0
Agra	Traffic Control	Flashing Signal/Blinker	Minor Injury	0.0
Agra	Traffic Control	Flashing Signal/Blinker	Persons Killed	0.0
Agra	Traffic Control	Flashing Signal/Blinker	Total Injured	0.0
Agra	Traffic Control	Flashing Signal/Blinker	Total number of Accidents	0.0
...	...	...	...	...
Vizag	Weather	Sunny/Clear	Grevously Injured	561.0
Vizag	Weather	Sunny/Clear	Minor Injury	252.0
Vizag	Weather	Sunny/Clear	Persons Killed	176.0
Vizag	Weather	Sunny/Clear	Total number of Accidents	1207.0
Vizag	Weather	Sunny/Clear	Total Injured	813.0

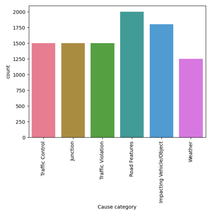
9550 rows × 4 columns

In [17]:
```
import seaborn as sns
import matplotlib.pyplot as plt

plt.figure(figsize=(14,6))
sns.countplot(data=df, x="Cause Subcategory", hue="Cause Subcategory", palette="tab10", legend=False)
plt.xticks(rotation=90)
plt.title("Cause Subcategory Count", fontsize=14)
plt.xlabel("Cause Subcategory")
plt.ylabel("Count")
plt.tight_layout()
plt.show()
```



In [18]:
```
import warnings
warnings.filterwarnings("ignore")
```

In [19]:
```
sns.countplot(
    data=df,
    x="Cause category",
    palette=sns.color_palette("husl", len(df["Cause category"].unique()))
)

plt.xticks(rotation=90)
plt.show()
```


2000
1750
1500
1250
1000
750
500
250
0
count
Traffic Control
Junction
Traffic Violation
Road Features
Impacting Vehicle/Object
Weather
Cause category