In [1]:
```
!pip install sqlalchemy pymysql
```

```
Requirement already satisfied: sqlalchemy in c:\users\windows10\appdata\local\progra
ms\python\python313\lib\site-packages (2.0.39)
Requirement already satisfied: pymysql in c:\users\windows10\appdata\local\programs
\python\python313\lib\site-packages (1.1.1)
Requirement already satisfied: greenlet!=0.4.17 in c:\users\windows10\appdata\local
\programs\python\python313\lib\site-packages (from sqlalchemy) (3.1.1)
Requirement already satisfied: typing-extensions>=4.6.0 in c:\users\windows10\appdat
a\roaming\python\python313\site-packages (from sqlalchemy) (4.12.2)
[notice] A new release of pip is available: 25.0 -> 25.0.1
[notice] To update, run: python.exe -m pip install --upgrade pip
```

In [32]:
```
import pymysql
import pandas as pd
 #Datbase Connection Details
db_config = {
    "host": "localhost",
    "user": "root",
    "password": "jb#mysql@2025",
    "database": "crime_db"
}
try:
    # Establish connection
    connection = pymysql.connect(**db_config)
    cursor = connection.cursor()
    #create database if not exists
    cursor.execute("CREATE DATABASE IF NOT EXISTS crime_db")
    #to use Database
    cursor.execute("USE crime_db")
    #creating table
    create_table_query = """
    CREATE TABLE IF NOT EXISTS crime_data (
        id INT AUTO_INCREMENT PRIMARY KEY,
    state_ut VARCHAR(255),
    district VARCHAR(255),
    year INT,
    murder INT,
    attempt_to_murder INT,
    culpable_homicide_not_amounting_to_murder INT,
    rape INT,
    custodial_rape INT,
    other_rape INT,
    kidnapping_abduction INT,
    kidnapping_abduction_women_girls INT,
    kidnapping_abduction_others INT,
    dacoity INT,
    preparation_assembly_dacoity INT,
    robbery INT,
    burglary INT,
    theft INT,
    auto_theft INT,
    other_theft INT,
    riots INT,
```

```
    criminal_breach_trust INT,
    cheating INT,
    counterfeiting INT,
    arson INT,
    hurt_grievous_hurt INT,
    dowry_deaths INT,
    assault_on_women INT,
    insult_to_women INT,
    cruelty_by_husband INT,
    importation_of_girls INT,
    causing_death_by_negligence INT,
    other_ipc_crimes INT,
    total_ipc_crimes INT
    );
    """

    cursor.execute(create_table_query)
    print("Table Cancer_data successfully created in Cancer_db database")
except pymysql.MySqlError as err:
    print(f"Error: {err}")
finally:
    if connection:
        cursor.close()
        connection.close()
        print("My sql connection close")
```

```
Table Cancer_data successfully created in Cancer_db database
My sql connection close
```

In [1]:

```
import pandas as pd
import pymysql

# Load CSV
file_path= r'C:\Users\Windows10\OneDrive\Desktop\crime\01_District_wise_crimes_comm
data=pd.read_csv(file_path)

# Drop 'id' column if it exists
data = data.drop(columns=['id'], errors='ignore')  # Removes 'id' column

# Define placeholders for query
num_columns = len(data.columns)  # Should be 32 (excluding 'id')
placeholders = ', '.join(['%s'] * num_columns)

# Correct INSERT query (without 'id')
insert_query = f"""INSERT INTO crime_data (state_ut, district, year, murder, attemp
                culpable_homicide_not_amounting_to_murder, rape, custodial_rape, oth
                kidnapping_abduction, kidnapping_abduction_women_girls, kidnapping_a
                dacoity, preparation_assembly_dacoity, robbery, burglary, theft, aut
                other_theft, riots, criminal_breach_trust, cheating, counterfeiting,
                hurt_grievous_hurt, dowry_deaths, assault_on_women, insult_to_women,
                cruelty_by_husband, importation_of_girls, causing_death_by_negligenc
                other_ipc_crimes, total_ipc_crimes) VALUES ({placeholders})"""

# Connect to MySQL
connection = pymysql.connect(host="localhost", user="root", password="jb#mysql@2025
cursor = connection.cursor()
```

```
# Insert Data
for row in data.itertuples(index=False, name=None):
    try:
        cursor.execute(insert_query, row)  # Ensure row matches column count
    except pymysql.MySQLError as err:
        print(f"Error inserting row {row}: {err}")

# Commit and Close Connection
connection.commit()
cursor.close()
connection.close()

print("Data inserted successfully")
```

```
Data inserted successfully
```

In [3]:
```
#Display 5 rows
data.head()
```

Out[3]:

	STATE/UT	DISTRICT	YEAR	MURDER	ATTEMPT TO MURDER	CULPABLE HOMICIDE NOT AMOUNTING TO MURDER	RAPE	CUSTODIAL RAPE
0	ANDHRA PRADESH	ADILABAD	2001	101	60	17	50	0
1	ANDHRA PRADESH	ANANTAPUR	2001	151	125	1	23	0
2	ANDHRA PRADESH	CHITTOOR	2001	101	57	2	27	0
3	ANDHRA PRADESH	CUDDAPAH	2001	80	53	1	20	0
4	ANDHRA PRADESH	EAST GODAVARI	2001	82	67	1	23	0

5 rows × 33 columns

In [4]:
```
data.tail()
```

Out[4]:

	STATE/UT	DISTRICT	YEAR	MURDER	ATTEMPT TO MURDER	CULPABLE HOMICIDE NOT AMOUNTING TO MURDER	RAPE	CU
9012	LAKSHADWEEP	LAKSHADWEEP	2012	0	0	0	0	
9013	LAKSHADWEEP	TOTAL	2012	0	0	0	0	
9014	PUDUCHERRY	KARAIKAL	2012	5	6	2	6	
9015	PUDUCHERRY	PUDUCHERRY	2012	24	21	10	7	
9016	PUDUCHERRY	TOTAL	2012	29	27	12	13	

5 rows × 33 columns

In [6]:
```
print(f"Number of Rows: {data.shape[0]}")
print(f"Number of Columns: {data.shape[1]}")
```

```
Number of Rows: 9017
Number of Columns: 33
```

In [7]:
```
#information about dataset
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9017 entries, 0 to 9016
Data columns (total 33 columns):
 #   Column                                          Non-Null Count  Dtype 
---  ------                                          --------------  ----- 
 0   STATE/UT                                        9017 non-null   object
 1   DISTRICT                                        9017 non-null   object
 2   YEAR                                            9017 non-null   int64 
 3   MURDER                                          9017 non-null   int64 
 4   ATTEMPT TO MURDER                               9017 non-null   int64 
 5   CULPABLE HOMICIDE NOT AMOUNTING TO MURDER       9017 non-null   int64 
 6   RAPE                                            9017 non-null   int64 
 7   CUSTODIAL RAPE                                  9017 non-null   int64 
 8   OTHER RAPE                                      9017 non-null   int64 
 9   KIDNAPPING & ABDUCTION                          9017 non-null   int64 
 10  KIDNAPPING AND ABDUCTION OF WOMEN AND GIRLS     9017 non-null   int64 
 11  KIDNAPPING AND ABDUCTION OF OTHERS              9017 non-null   int64 
 12  DACOITY                                         9017 non-null   int64 
 13  PREPARATION AND ASSEMBLY FOR DACOITY            9017 non-null   int64 
 14  ROBBERY                                         9017 non-null   int64 
 15  BURGLARY                                        9017 non-null   int64 
 16  THEFT                                           9017 non-null   int64 
 17  AUTO THEFT                                      9017 non-null   int64 
 18  OTHER THEFT                                     9017 non-null   int64 
 19  RIOTS                                           9017 non-null   int64 
 20  CRIMINAL BREACH OF TRUST                        9017 non-null   int64 
 21  CHEATING                                        9017 non-null   int64 
 22  COUNTERFIETING                                  9017 non-null   int64 
 23  ARSON                                           9017 non-null   int64 
 24  HURT/GREVIOUS HURT                              9017 non-null   int64 
 25  DOWRY DEATHS                                    9017 non-null   int64 
 26  ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY  9017 non-null   int64 
 27  INSULT TO MODESTY OF WOMEN                      9017 non-null   int64 
 28  CRUELTY BY HUSBAND OR HIS RELATIVES             9017 non-null   int64 
 29  IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES     9017 non-null   int64 
 30  CAUSING DEATH BY NEGLIGENCE                     9017 non-null   int64 
 31  OTHER IPC CRIMES                                9017 non-null   int64 
 32  TOTAL IPC CRIMES                                9017 non-null   int64 
dtypes: int64(31), object(2)
memory usage: 2.3+ MB
```

In [8]:
```
data.isnull()
```

Out[8]:

	STATE/UT	DISTRICT	YEAR	MURDER	ATTEMPT TO MURDER	CULPABLE HOMICIDE NOT AMOUNTING TO MURDER	RAPE	CUSTODIAL RAPE
0	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...
9012	False	False	False	False	False	False	False	False
9013	False	False	False	False	False	False	False	False
9014	False	False	False	False	False	False	False	False
9015	False	False	False	False	False	False	False	False
9016	False	False	False	False	False	False	False	False

9017 rows × 33 columns

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

Out[9]:
```
STATE/UT                                                0
DISTRICT                                                0
YEAR                                                    0
MURDER                                                  0
ATTEMPT TO MURDER                                       0
CULPABLE HOMICIDE NOT AMOUNTING TO MURDER               0
RAPE                                                    0
CUSTODIAL RAPE                                          0
OTHER RAPE                                              0
KIDNAPPING & ABDUCTION                                  0
KIDNAPPING AND ABDUCTION OF WOMEN AND GIRLS             0
KIDNAPPING AND ABDUCTION OF OTHERS                      0
DACOITY                                                 0
PREPARATION AND ASSEMBLY FOR DACOITY                    0
ROBBERY                                                 0
BURGLARY                                                0
THEFT                                                   0
AUTO THEFT                                              0
OTHER THEFT                                             0
RIOTS                                                   0
CRIMINAL BREACH OF TRUST                                0
CHEATING                                                0
COUNTERFIETING                                          0
ARSON                                                   0
HURT/GREVIOUS HURT                                      0
DOWRY DEATHS                                            0
ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY     0
INSULT TO MODESTY OF WOMEN                              0
CRUELTY BY HUSBAND OR HIS RELATIVES                     0
IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES             0
CAUSING DEATH BY NEGLIGENCE                             0
OTHER IPC CRIMES                                        0
TOTAL IPC CRIMES                                        0
dtype: int64
```

In [10]:
```
data.nunique()
```

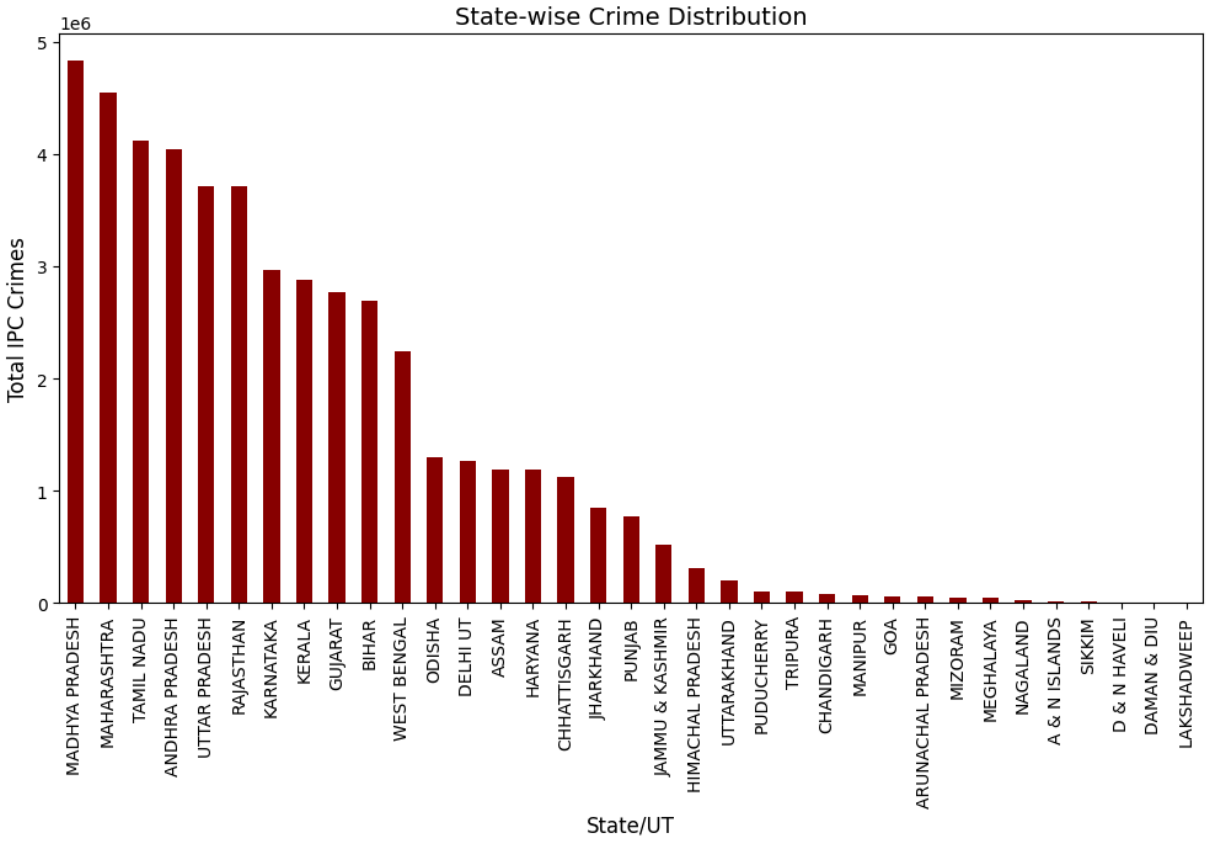
Out[10]:
```
STATE/UT                                                35
DISTRICT                                               808
YEAR                                                    12
MURDER                                                 477
ATTEMPT TO MURDER                                      514
CULPABLE HOMICIDE NOT AMOUNTING TO MURDER              171
RAPE                                                   420
CUSTODIAL RAPE                                           5
OTHER RAPE                                             419
KIDNAPPING & ABDUCTION                                 534
KIDNAPPING AND ABDUCTION OF WOMEN AND GIRLS            460
KIDNAPPING AND ABDUCTION OF OTHERS                     307
DACOITY                                                227
PREPARATION AND ASSEMBLY FOR DACOITY                   191
ROBBERY                                                455
BURGLARY                                               922
THEFT                                                 1751
AUTO THEFT                                            1114
OTHER THEFT                                           1343
RIOTS                                                  804
CRIMINAL BREACH OF TRUST                               411
CHEATING                                               824
COUNTERFIETING                                         167
ARSON                                                  295
HURT/GREVIOUS HURT                                    1855
DOWRY DEATHS                                           256
ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY     553
INSULT TO MODESTY OF WOMEN                             343
CRUELTY BY HUSBAND OR HIS RELATIVES                    846
IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES             37
CAUSING DEATH BY NEGLIGENCE                            854
OTHER IPC CRIMES                                      3173
TOTAL IPC CRIMES                                      5010
dtype: int64
```

In [13]:
```
print(data.columns)
```

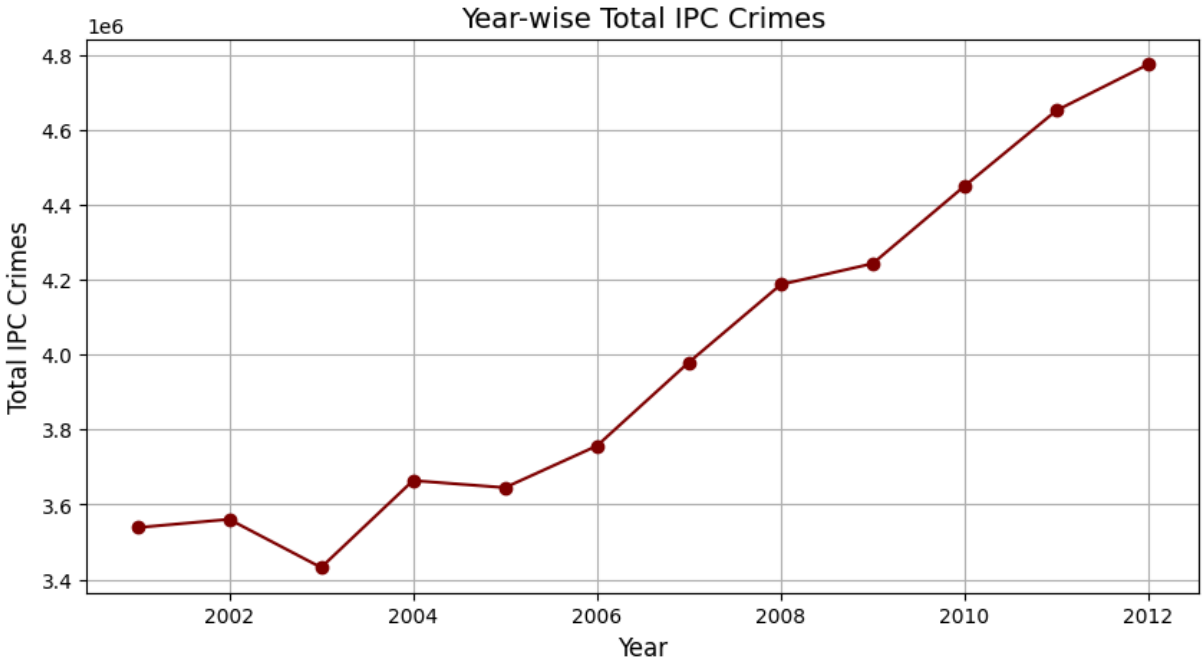
```
Index(['STATE/UT', 'DISTRICT', 'YEAR', 'MURDER', 'ATTEMPT TO MURDER',
       'CULPABLE HOMICIDE NOT AMOUNTING TO MURDER', 'RAPE', 'CUSTODIAL RAPE',
       'OTHER RAPE', 'KIDNAPPING & ABDUCTION',
       'KIDNAPPING AND ABDUCTION OF WOMEN AND GIRLS',
       'KIDNAPPING AND ABDUCTION OF OTHERS', 'DACOITY',
       'PREPARATION AND ASSEMBLY FOR DACOITY', 'ROBBERY', 'BURGLARY', 'THEFT',
       'AUTO THEFT', 'OTHER THEFT', 'RIOTS', 'CRIMINAL BREACH OF TRUST',
       'CHEATING', 'COUNTERFIETING', 'ARSON', 'HURT/GREVIOUS HURT',
       'DOWRY DEATHS', 'ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY',
       'INSULT TO MODESTY OF WOMEN', 'CRUELTY BY HUSBAND OR HIS RELATIVES',
       'IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES',
       'CAUSING DEATH BY NEGLIGENCE', 'OTHER IPC CRIMES', 'TOTAL IPC CRIMES'],
      dtype='object')
```

In [16]:
```
# 1. State-wise Crime Distribution
plt.figure(figsize=(12, 6))
data.groupby('STATE/UT')['TOTAL IPC CRIMES'].sum().sort_values(ascending=False).plo
plt.title('State-wise Crime Distribution', fontsize=14, color='black')
plt.xlabel('State/UT', fontsize=12, color='black')
```

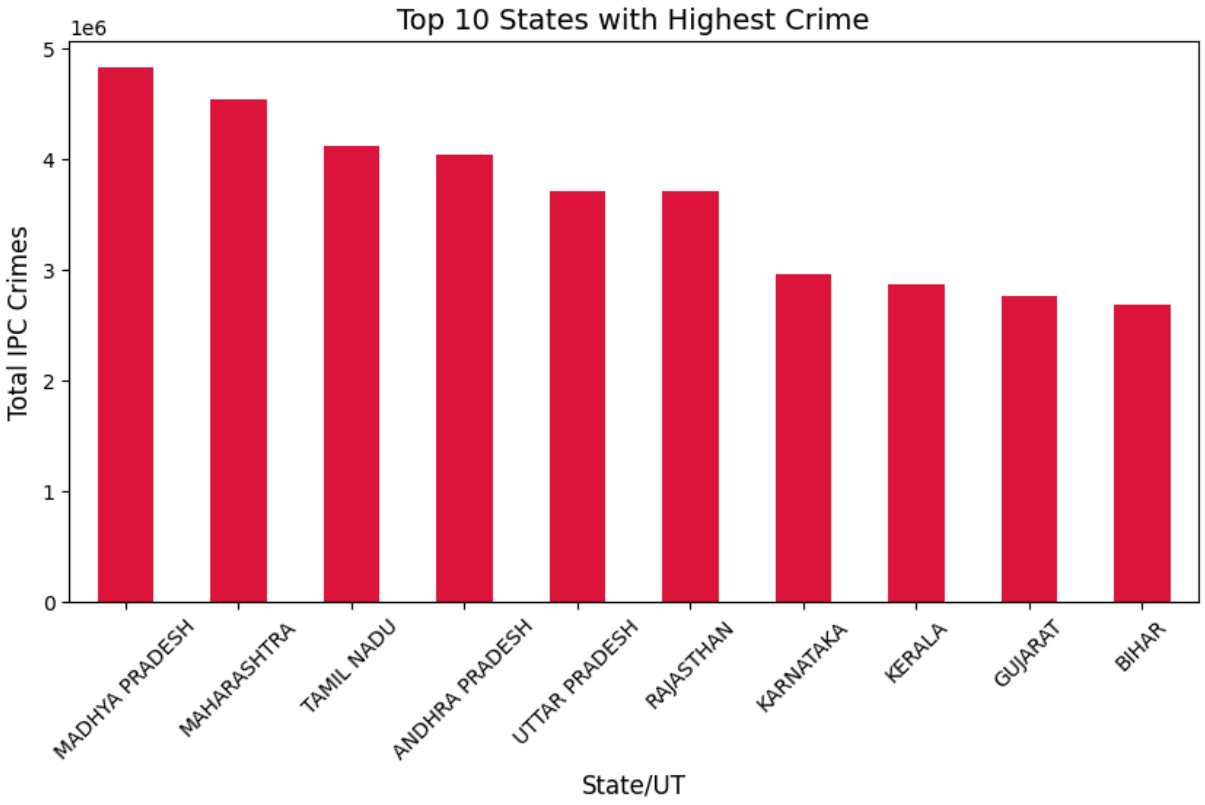
```
plt.ylabel('Total IPC Crimes', fontsize=12, color='black')
plt.xticks(rotation=90, color='black')
plt.yticks(color='black')
plt.show()
```



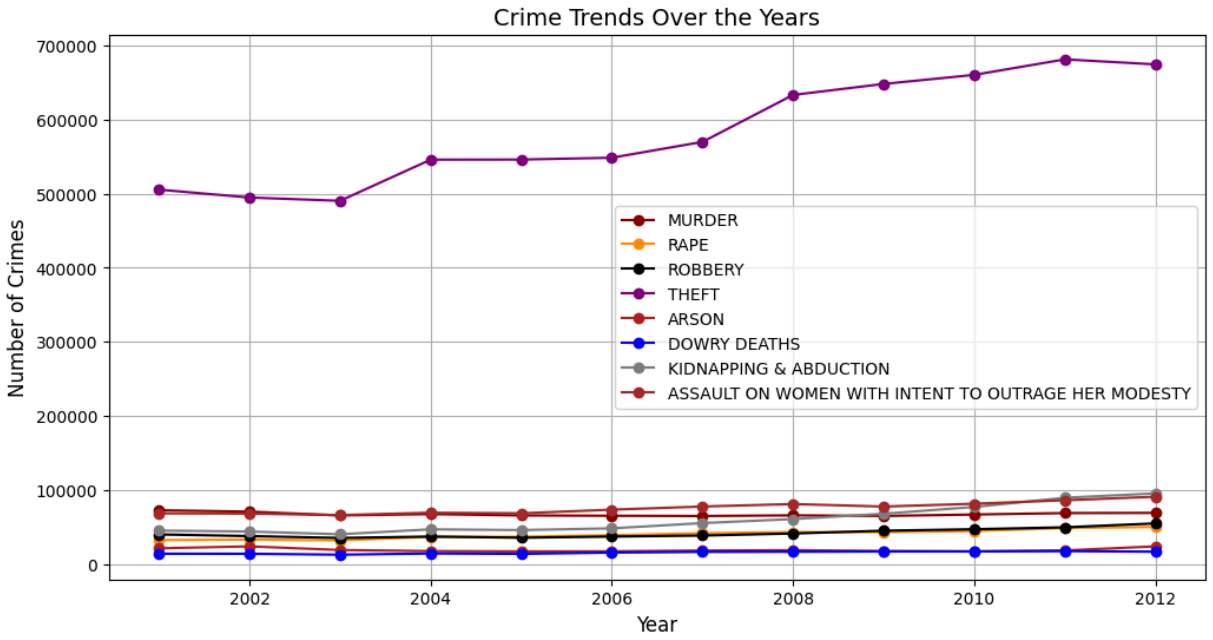
In [17]:
```
# 2. Year-wise Total IPC Crimes
plt.figure(figsize=(10, 5))
data.groupby('YEAR')['TOTAL IPC CRIMES'].sum().plot(marker='o', linestyle='-', colo
plt.title('Year-wise Total IPC Crimes', fontsize=14, color='black')
plt.xlabel('Year', fontsize=12, color='black')
plt.ylabel('Total IPC Crimes', fontsize=12, color='black')
plt.grid()
plt.show()
```



In [18]:
```
# 3. Top 10 States with Highest Crime
top_states = data.groupby('STATE/UT')['TOTAL IPC CRIMES'].sum().nlargest(10)
plt.figure(figsize=(10, 5))
top_states.plot(kind='bar', color='crimson')
plt.title('Top 10 States with Highest Crime', fontsize=14, color='black')
plt.xlabel('State/UT', fontsize=12, color='black')
plt.ylabel('Total IPC Crimes', fontsize=12, color='black')
plt.xticks(rotation=45, color='black')
plt.yticks(color='black')
plt.show()
```



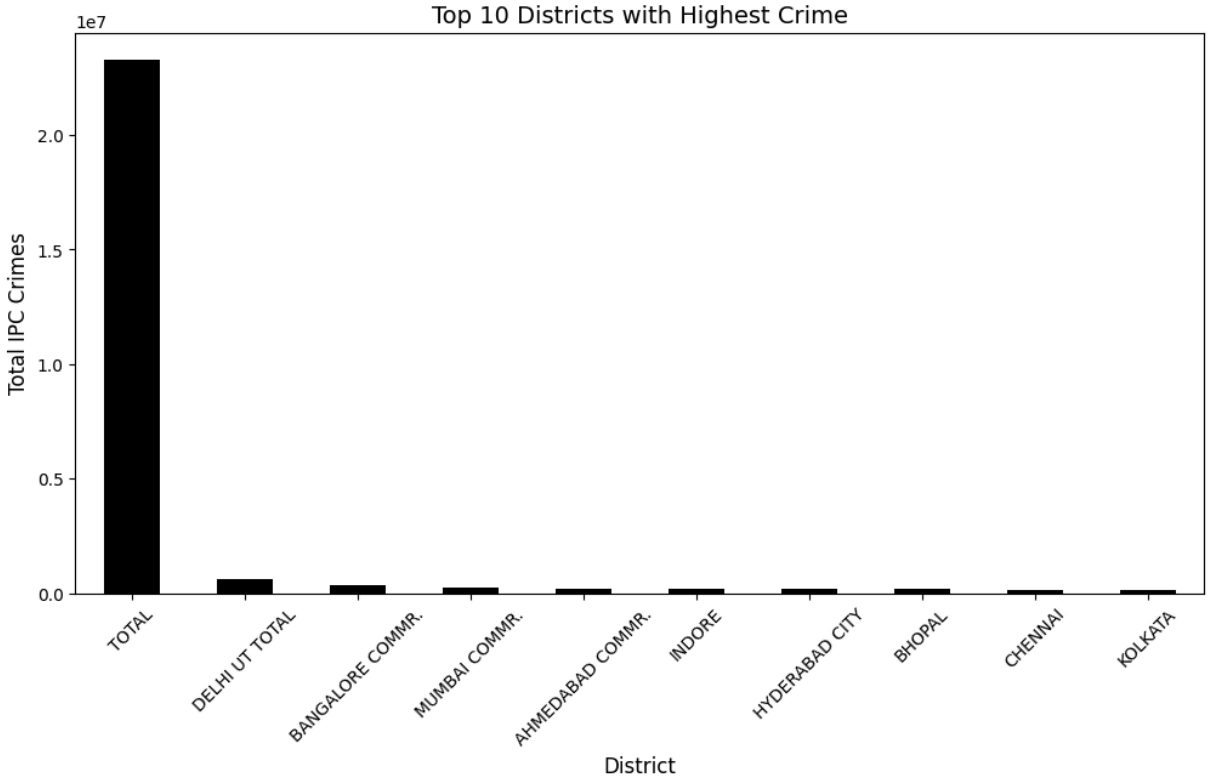
In [27]:
```
# 5. Crime Trends Over the Years
plt.figure(figsize=(12, 6))
crime_colors = {
    'MURDER': 'darkred', 'RAPE': 'darkorange', 'ROBBERY': 'black', 'THEFT': 'purple
    'DOWRY DEATHS': 'blue', 'KIDNAPPING & ABDUCTION': 'grey', 'ASSAULT ON WOMEN WIT
}
for crime in crime_colors.keys():
    plt.plot(data.groupby('YEAR')[crime].sum(), marker='o', linestyle='-', label=cr
plt.title('Crime Trends Over the Years', fontsize=14, color='black')
plt.xlabel('Year', fontsize=12, color='black')
plt.ylabel('Number of Crimes', fontsize=12, color='black')
plt.legend()
plt.grid()
plt.show()
```



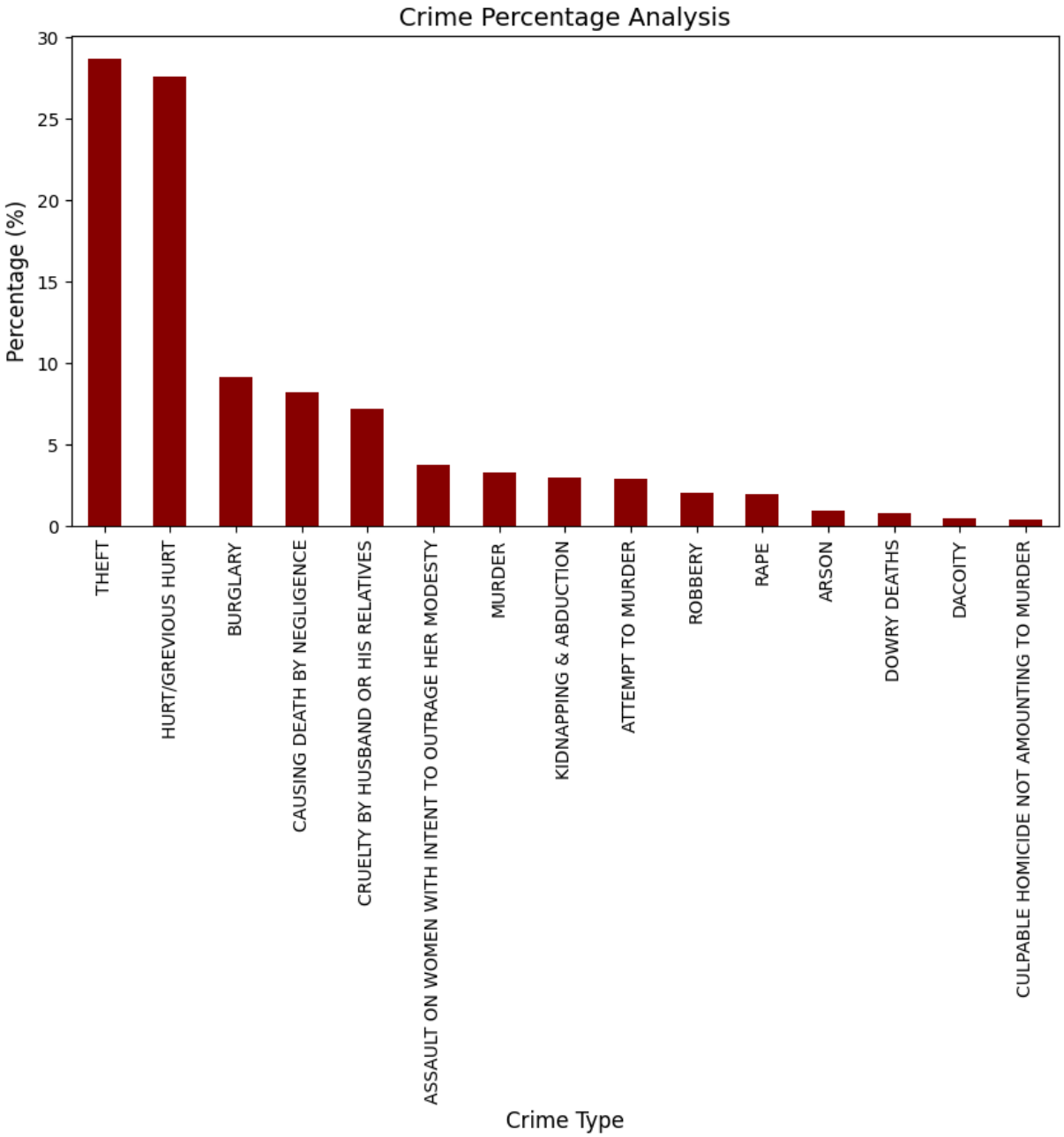
In [28]:
```
# 6. Correlation Heatmap of Crime Data
plt.figure(figsize=(12, 6))
sns.heatmap(data.select_dtypes(include=['number']).corr(), annot=True, cmap='Reds',
plt.title('Crime Correlation Heatmap', fontsize=14, color='black')
plt.show()
```

In [29]:
```
# 7. Crime Distribution by District(Bar Chart)
plt.figure(figsize=(12, 6))
top_districts = data.groupby('DISTRICT')['TOTAL IPC CRIMES'].sum().nlargest(10)
```

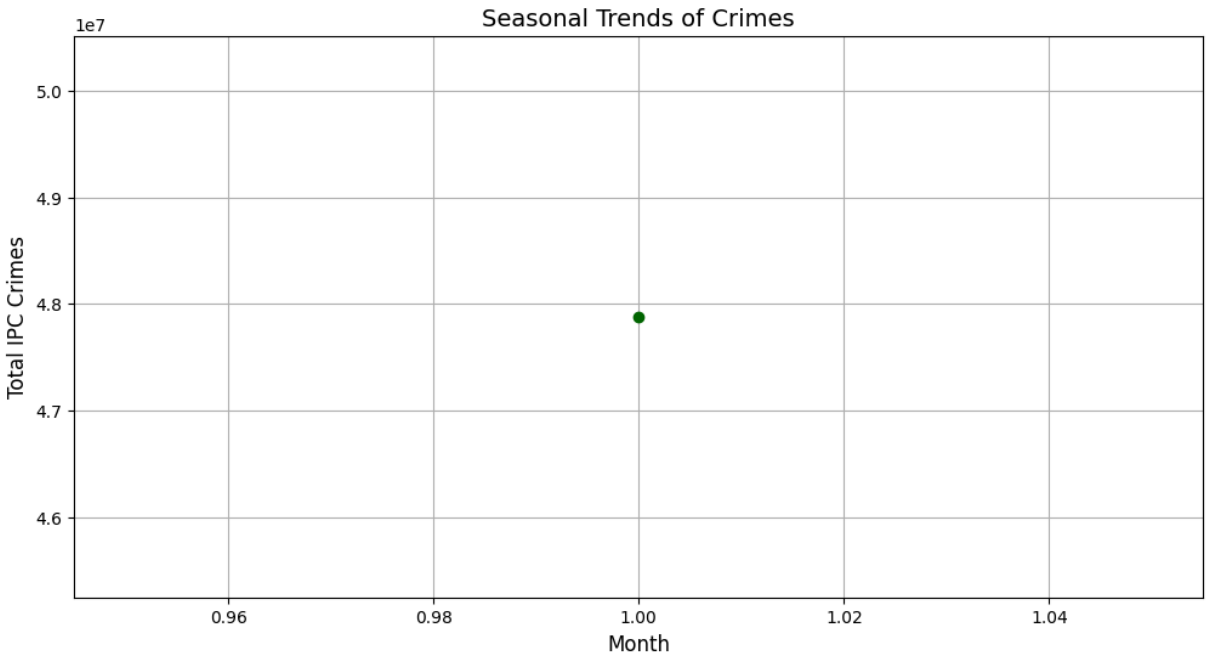
```
top_districts.plot(kind='bar', color='black')
plt.title('Top 10 Districts with Highest Crime', fontsize=14, color='black')
plt.xlabel('District', fontsize=12, color='black')
plt.ylabel('Total IPC Crimes', fontsize=12, color='black')
plt.xticks(rotation=45, color='black')
plt.yticks(color='black')
plt.show()
```



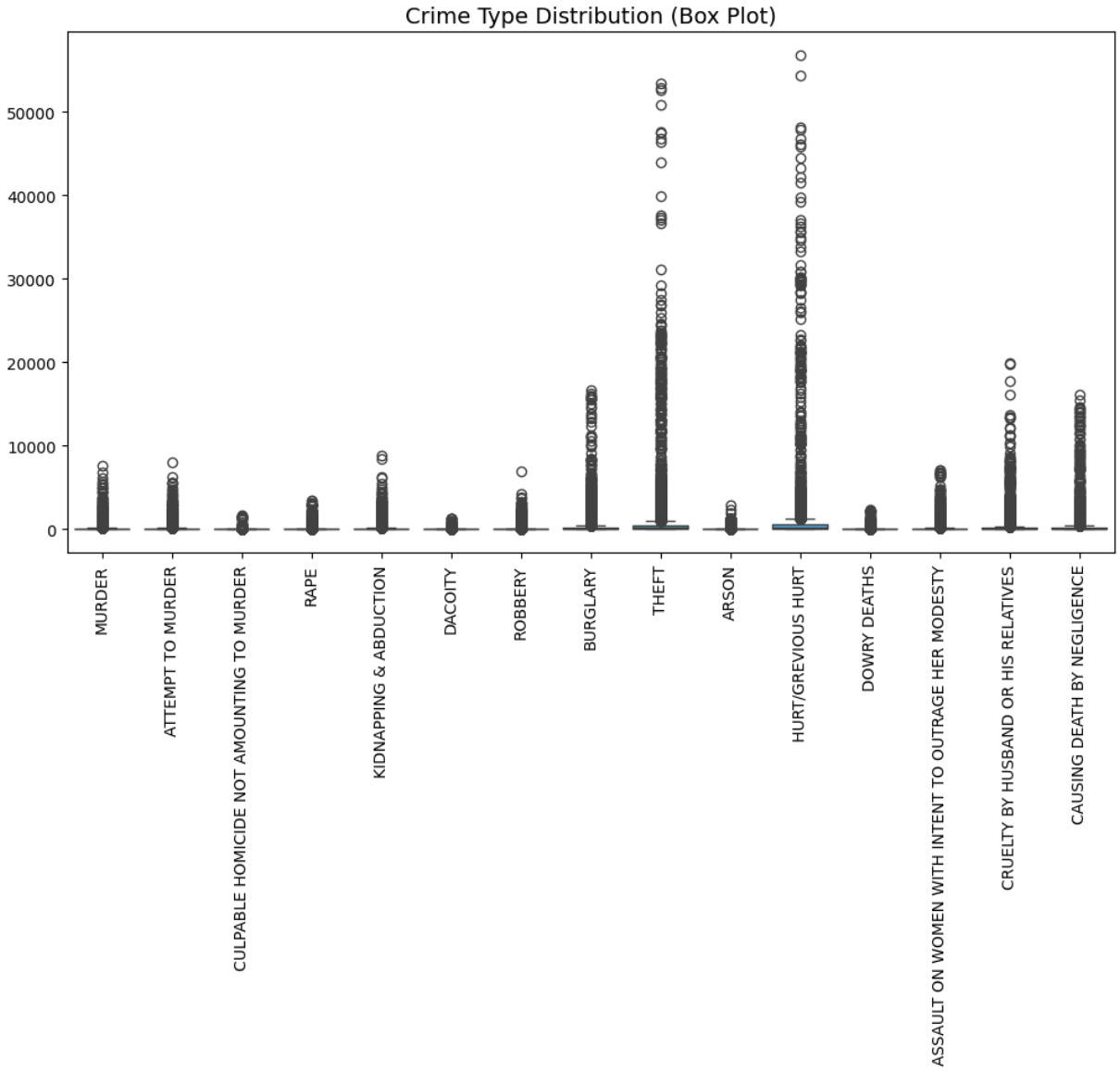
In [30]:
```
# 8. Crime Percentage Analysis
crime_percent = (crime_counts / crime_counts.sum()) * 100
plt.figure(figsize=(10, 5))
crime_percent.sort_values(ascending=False).plot(kind='bar', color='darkred')
plt.title('Crime Percentage Analysis', fontsize=14, color='black')
plt.xlabel('Crime Type', fontsize=12, color='black')
plt.ylabel('Percentage (%)', fontsize=12, color='black')
plt.xticks(rotation=90, color='black')
plt.yticks(color='black')
plt.show()
```



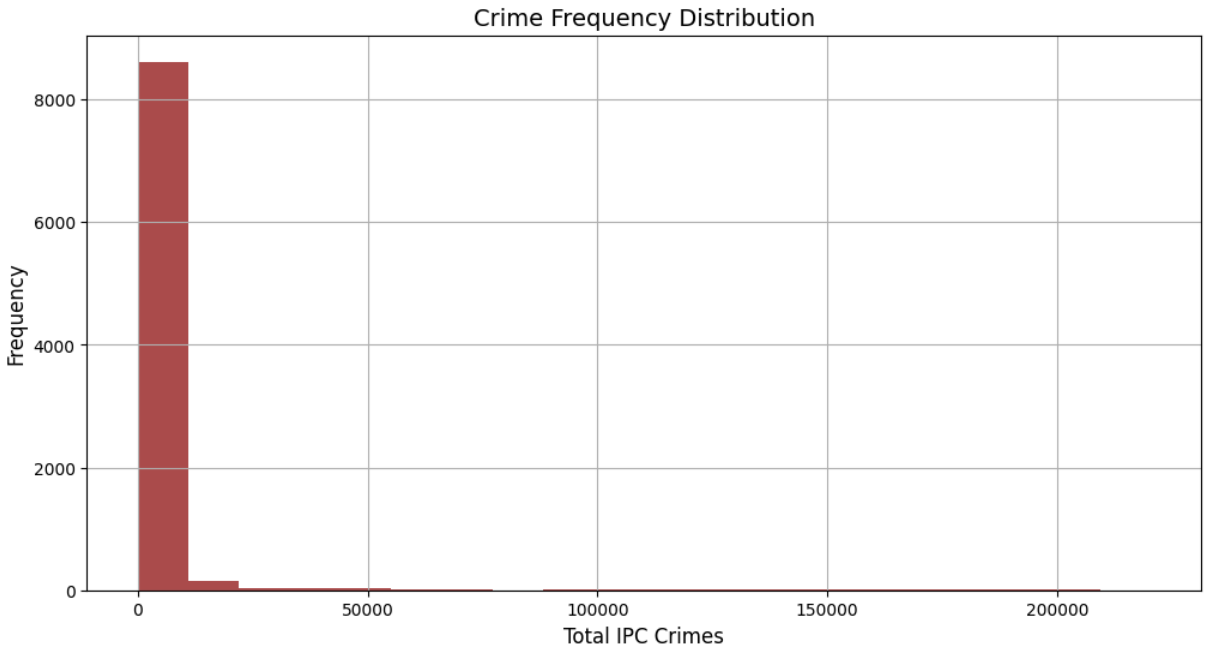
In [35]:
```
# 8. Seasonal Trends of Crimes
plt.figure(figsize=(12, 6))
data['MONTH'] = pd.to_datetime(data['YEAR'], format='%Y').dt.month  # Assuming mont
if 'MONTH' in data.columns:
    monthwise_crime = data.groupby('MONTH')['TOTAL IPC CRIMES'].sum()
    monthwise_crime.plot(marker='o', linestyle='-', color='darkgreen')
    plt.title('Seasonal Trends of Crimes', fontsize=14, color='black')
    plt.xlabel('Month', fontsize=12, color='black')
    plt.ylabel('Total IPC Crimes', fontsize=12, color='black')
    plt.grid()
    plt.show()
```



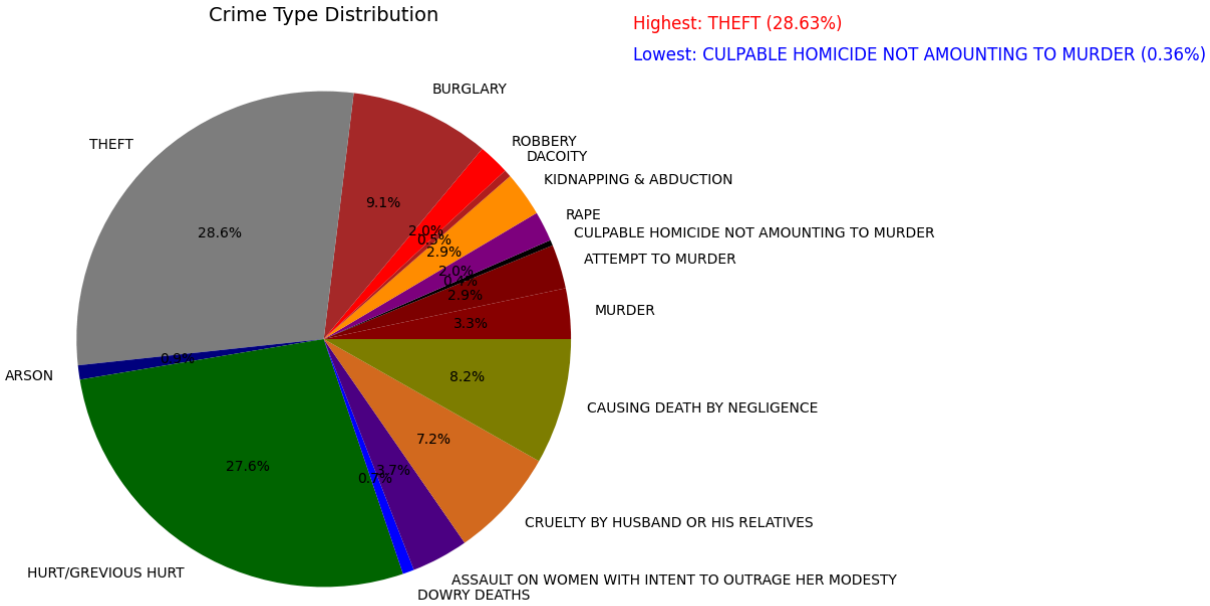
In [39]:
```
# 10. Box Plot for Crime Types
plt.figure(figsize=(12, 6))
sns.boxplot(data=data[crime_types])
plt.xticks(rotation=90)
plt.title('Crime Type Distribution (Box Plot)', fontsize=14, color='black')
plt.show()
```



In [43]:
```
plt.figure(figsize=(12, 6))
plt.hist(data['TOTAL IPC CRIMES'], bins=20, color='darkred', alpha=0.7)
plt.title('Crime Frequency Distribution', fontsize=14, color='black')
plt.xlabel('Total IPC Crimes', fontsize=12, color='black')
plt.ylabel('Frequency', fontsize=12, color='black')
plt.grid()
plt.show()
```



In [45]:
```
# 4. Crime Type Distribution (Pie Chart)
crime_types = [
    'MURDER', 'ATTEMPT TO MURDER', 'CULPABLE HOMICIDE NOT AMOUNTING TO MURDER', 'RA
    'DACOITY', 'ROBBERY', 'BURGLARY', 'THEFT', 'ARSON', 'HURT/GREVIOUS HURT', 'DOWR
    'CRUELTY BY HUSBAND OR HIS RELATIVES', 'CAUSING DEATH BY NEGLIGENCE'
]
crime_counts = data[crime_types].sum()
fig, ax = plt.subplots(figsize=(8, 8))
colors = ['darkred', 'maroon', 'black', 'purple', 'darkorange', 'firebrick', 'red',
wedges, texts, autotexts = ax.pie(crime_counts, labels=crime_counts.index, autopct=
plt.title('Crime Type Distribution', fontsize=14, color='black')

# Adding annotation for highest and lowest crime percentages
highest_crime = crime_counts.idxmax()
lowest_crime = crime_counts.idxmin()
highest_percent = (crime_counts.max() / crime_counts.sum()) * 100
lowest_percent = (crime_counts.min() / crime_counts.sum()) * 100
plt.annotate(f'Highest: {highest_crime} ({highest_percent:.2f}%)', xy=(1, 1), xycoo
plt.annotate(f'Lowest: {lowest_crime} ({lowest_percent:.2f}%)', xy=(1, 0.95), xycoo
plt.show()
```


Crime Type Distribution
Highest: THEFT (28.63%)
Lowest: CULPABLE HOMICIDE NOT AMOUNTING TO MURDER (0.36%)
BURGLARY
THEFT
ROBBERY
DACOITY
KIDNAPPING & ABDUCTION
RAPE
CULPABLE HOMICIDE NOT AMOUNTING TO MURDER
ATTEMPT TO MURDER
MURDER
CAUSING DEATH BY NEGLIGENCE
CRUELTY BY HUSBAND OR HIS RELATIVES
ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY
DOWRY DEATHS
HURT/GREVIOUS HURT
ARSON
9.1%
28.6%
2.0%
0.5%
2.9%
2.0%
0.4%
2.9%
3.3%
8.2%
7.2%
3.7%
0.7%
27.6%

In [ ]:

In [64]:
```
!pip install sqlalchemy pymysql
```

```
Requirement already satisfied: sqlalchemy in c:\users\admin\appdata\local\package
s\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\py
thon311\site-packages (2.0.39)
Requirement already satisfied: pymysql in c:\users\admin\appdata\local\packages\p
ythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packages\pytho
n311\site-packages (1.1.1)
Requirement already satisfied: greenlet!=0.4.17 in c:\users\admin\appdata\local\p
ackages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\local-packa
ges\python311\site-packages (from sqlalchemy) (3.1.1)
Requirement already satisfied: typing-extensions>=4.6.0 in c:\users\admin\appdata
\local\packages\pythonsoftwarefoundation.python.3.11_qbz5n2kfra8p0\localcache\loc
al-packages\python311\site-packages (from sqlalchemy) (4.12.2)
[notice] A new release of pip is available: 24.0 -> 25.0.1
[notice] To update, run: C:\Users\Admin\AppData\Local\Microsoft\WindowsApps\Pytho
nSoftwareFoundation.Python.3.11_qbz5n2kfra8p0\python.exe -m pip install --upgrade
pip
```

In [2]:
```
import pandas as pd
from sqlalchemy import create_engine

# MySQL connection details
host = 'localhost'
user = 'root'
password = 'mysql28/may/2003'
database = 'crimeDB'

# Create a connection to MySQL
engine = create_engine(f"mysql+pymysql://{user}:{password}@{host}/{database}")

# Read the CSV file
file_path = '42_District_wise_crimes_committed_against_women_2014.csv'
data = pd.read_csv(file_path)

# Load the data into MySQL table
table_name = 'district_wise_crimes'
data.to_sql(table_name, con=engine, if_exists='replace', index=False)

# Confirm the upload
result = pd.read_sql(f"SELECT COUNT(*) FROM {table_name};", con=engine)
print(f"Data uploaded successfully! Total rows: {result.iloc[0, 0]}")
```

```
Data uploaded successfully! Total rows: 837
```

In [8]:
```
import pandas as pd
import numpy as np

print("Loading data...")
data = pd.read_csv(file_path)
print("Data loaded successfully!")
print(data.head())

# Load the CSV file
file_path = r'C:\Users\Admin\42_District_wise_crimes_committed_against_women_201
data = pd.read_csv(file_path)
```

```
Loading data...
Data loaded successfully!
       States/UTs          District  Year  Rape  Custodial Rape  \
0  Andhra Pradesh         Anantapur  2014    35               0
1  Andhra Pradesh          Chittoor  2014    32               0
2  Andhra Pradesh          Cuddapah  2014    28               0
3  Andhra Pradesh     East Godavari  2014    85               0
4  Andhra Pradesh  Guntakal Railway  2014     0               0

   Custodial_Gang Rape  Custodial_Other Rape  Rape other than Custodial  \
0                    0                     0                         35
1                    0                     0                         32
2                    0                     0                         28
3                    0                     0                         85
4                    0                     0                          0

   Rape_Gang Rape  Rape_Others  ...  Commission of Sati Prevention Act, 1987  \
0               0           35  ...                                       0
1               1           31  ...                                       0
2               0           28  ...                                       0
3               0           85  ...                                       0
4               0            0  ...                                       0

   Protection of Women from Domestic Violence Act, 2005  \
0                                                  0
1                                                  0
2                                                  0
3                                                  0
4                                                  0

   Immoral Traffic Prevention Act  ITP Under Section 5  ITP Under Section 6  \
0                               0                    0                    0
1                               4                    4                    0
2                               5                    0                    0
3                              16                    0                    0
4                               0                    0                    0

   ITP Under Section 7  ITP Under Section 8  ITP Under Other Sections  \
0                    0                    0                         0
1                    0                    0                         0
2                    0                    0                         5
3                    0                    0                        16
4                    0                    0                         0

   Other SLL Crimes against Women  Total Crimes against Women
0                               0                       1097
1                               0                        607
2                               0                        609
3                               0                       1277
4                               0                          4

[5 rows x 62 columns]
```

In [9]:
```
# Total crimes against women
total_crimes = data['Total Crimes against Women'].sum()
print(f"\nTotal crimes against women in 2014: {total_crimes}")
```

```
Total crimes against women in 2014: 914348
```

In [10]:
```
# Mean, median, min, and max for total crimes
print(f"Mean: {np.mean(data['Total Crimes against Women'])}")
print(f"Median: {np.median(data['Total Crimes against Women'])}")
print(f"Min: {np.min(data['Total Crimes against Women'])}")
print(f"Max: {np.max(data['Total Crimes against Women'])}")
```

```
Mean: 1092.4109916367981
Median: 393.0
Min: 0
Max: 57101
```

In [11]:
```
# ---------------- STATE-WISE ANALYSIS ----------------
# Total crimes by state
state_crime = data.groupby('States/UTs')['Total Crimes against Women'].sum().sor
print("\n--- Total Crimes by State ---")
print(state_crime.head(10))
```

```
--- Total Crimes by State ---
States/UTs
Madhya Pradesh    114202
Uttar Pradesh     104818
Maharashtra        98334
West Bengal        96788
Rajasthan          64802
Delhi UT           44510
Assam              39218
Odisha             37128
Bihar              36416
Karnataka          35584
Name: Total Crimes against Women, dtype: int64
```

In [12]:
```
# Top 5 districts with highest and lowest crimes
print("\n--- Top 5 Districts with Highest Crimes ---")
print(data[['District', 'Total Crimes against Women']].nlargest(5, 'Total Crimes

print("\n--- Top 5 Districts with Lowest Crimes ---")
print(data[['District', 'Total Crimes against Women']].nsmallest(5, 'Total Crime
```

```
--- Top 5 Districts with Highest Crimes ---
    District  Total Crimes against Women
404    Total                       57101
755    Total                       52409
451    Total                       49167
800    Total                       48394
605    Total                       32401

--- Top 5 Districts with Lowest Crimes ---
                  District  Total Crimes against Women
23             Crime Branch                           0
113  Economic Offences Unit                           0
114    Anti Terrorist Squad                           0
159          C. I. D. Crime                           0
214      Irrigation & Power                           0
```

In [18]:
```
# CRIME TYPE BREAKDOWN
# Total number of each type of crime
crime_type_totals = data.iloc[:, 3:-1].sum().sort_values(ascending=False)
print("\n--- Total Number of Each Type of Crime ---")
print(crime_type_totals)
```

```
--- Total Number of Each Type of Crime ---
Cruelty by Husband or his Relatives                          246326
Assault on Women with intent to outrage her Modesty_Total    164822
Other IPC Crimes                                             129398
Kidnapping & Abduction_Total                                 116478
Others                                                        97268
Rape                                                          77356
Rape other than Custodial                                     76958
Rape_Others                                                   72254
Kidnapping & Abduction of Women to compel her for marriage    62368
Sexual Harassment                                             43940
Kidnaping & Abduction                                         34652
Other SLL Crimes against Women                                25114
Robbery                                                       22262
Dowry Prohibition Act, 1961                                   20112
Grievous Hurt                                                 19902
Hurt                                                          19546
Insult to the Modesty of Women_Total                          19476
Kidnaping & Abduction_Others                                  18860
In other Places                                               18182
Murder                                                        17488
Dowry Deaths                                                  16916
Attempt to commit Murder                                      13088
Assault on women with intent to Disrobe                       12850
Stalking                                                       9404
Attempt to commit Rape                                         8620
Abetment of Suicides of Women                                  7470
Rape_Gang Rape                                                 4704
Immoral Traffic Prevention Act                                 4140
ITP Under Other Sections                                       2320
Voyeurism                                                      1360
ITP Under Section 5                                            1210
Culpable Homicide not amounting to Murder                       960
Arson                                                           952
In places related to work                                       938
HumanTrafficking                                                924
Protection of Women from Domestic Violence Act, 2005            850
Attempt to commit Culpable Homicide                             686
Dacoity_Total                                                   572
Other Dacoity                                                   554
Kidnaping & Abduction in order to Murder                        416
Custodial Rape                                                  398
Custodial_Other Rape                                            384
Acid attack                                                     276
In Public Transport system                                      242
ITP Under Section 7                                             240
UnNatural Offences                                              218
ITP Under Section 8                                             208
Kidnapping for Ransom                                           182
ITP Under Section 6                                             162
At Office premises                                              114
Indecent Representation of Women (P) Act, 1986                   94
Causing miscarriage without consent of women                     90
Attempt to Acid Attack                                           80
Importation of Girls from Foreign Country                        28
Dacoity with Murder                                              18
Custodial_Gang Rape                                              14
Deaths caused with intent to cause miscarriage                    6
Commission of Sati Prevention Act, 1987                           0
dtype: int64
```

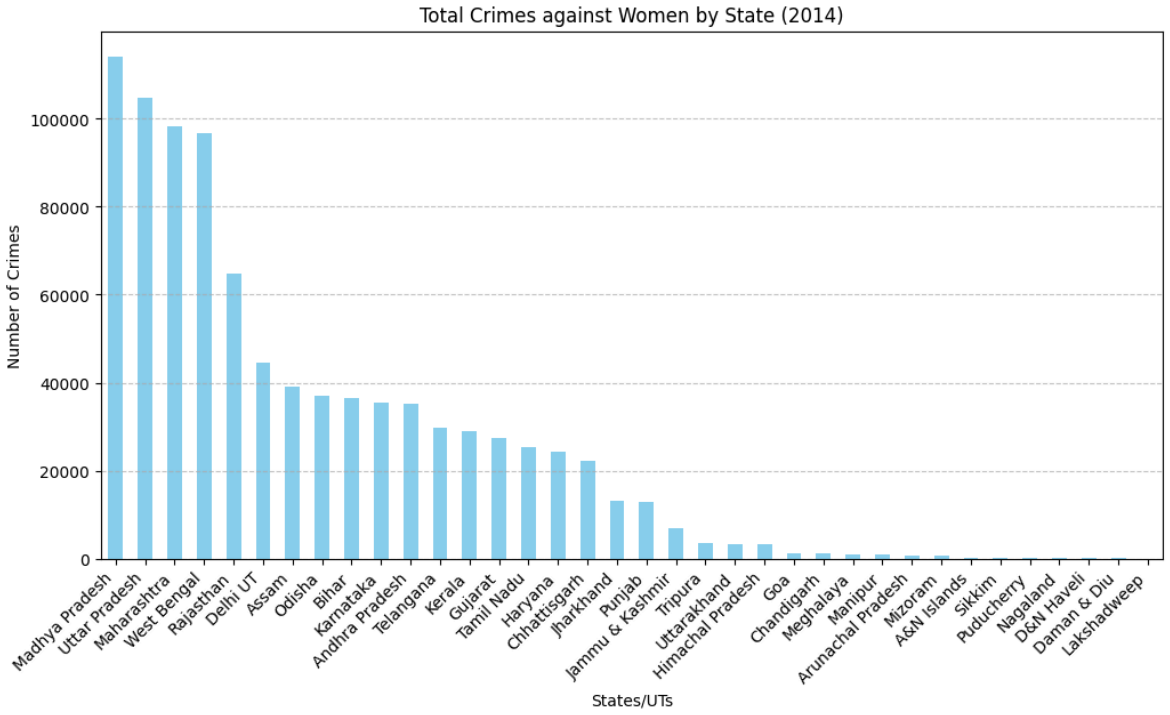
In [14]:
```
# Percentage contribution of each type of crime
crime_type_percent = (crime_type_totals / total_crimes) * 100
print("\n--- Percentage Contribution of Each Type of Crime ---")
print(crime_type_percent)
```

```
--- Percentage Contribution of Each Type of Crime ---
Cruelty by Husband or his Relatives                              26.940071
Assault on Women with intent to outrage her Modesty_Total        18.026178
Other IPC Crimes                                                 14.151942
Kidnapping & Abduction_Total                                     12.738913
Others                                                           10.637963
Rape                                                              8.460236
Rape other than Custodial                                         8.416708
Rape_Others                                                       7.902243
Kidnapping & Abduction of Women to compel her for marriage        6.821035
Sexual Harassment                                                 4.805610
Kidnaping & Abduction                                             3.789804
Other SLL Crimes against Women                                    2.746657
Robbery                                                           2.434740
Dowry Prohibition Act, 1961                                       2.199600
Grievous Hurt                                                     2.176633
Hurt                                                              2.137698
Insult to the Modesty of Women_Total                              2.130042
Kidnaping & Abduction_Others                                      2.062672
In other Places                                                   1.988521
Murder                                                            1.912620
Dowry Deaths                                                      1.850061
Attempt to commit Murder                                          1.431402
Assault on women with intent to Disrobe                           1.405373
Stalking                                                          1.028492
Attempt to commit Rape                                            0.942748
Abetment of Suicides of Women                                     0.816976
Rape_Gang Rape                                                    0.514465
Immoral Traffic Prevention Act                                    0.452782
ITP Under Other Sections                                          0.253733
Voyeurism                                                         0.148740
ITP Under Section 5                                               0.132335
Culpable Homicide not amounting to Murder                         0.104993
Arson                                                             0.104118
In places related to work                                         0.102587
HumanTrafficking                                                  0.101056
Protection of Women from Domestic Violence Act, 2005              0.092962
Attempt to commit Culpable Homicide                               0.075026
Dacoity_Total                                                     0.062558
Other Dacoity                                                     0.060590
Kidnaping & Abduction in order to Murder                          0.045497
Custodial Rape                                                    0.043528
Custodial_Other Rape                                              0.041997
Acid attack                                                       0.030185
In Public Transport system                                        0.026467
ITP Under Section 7                                               0.026248
UnNatural Offences                                                0.023842
ITP Under Section 8                                               0.022748
Kidnapping for Ransom                                             0.019905
ITP Under Section 6                                               0.017718
At Office premises                                                0.012468
Indecent Representation of Women (P) Act, 1986                    0.010281
Causing miscarriage without consent of women                      0.009843
Attempt to Acid Attack                                            0.008749
Importation of Girls from Foreign Country                         0.003062
Dacoity with Murder                                               0.001969
Custodial_Gang Rape                                               0.001531
Deaths caused with intent to cause miscarriage                    0.000656
Commission of Sati Prevention Act, 1987                           0.000000
dtype: float64
```

In [19]:
```
# DATA CLEANING
# Handle missing values (if any)
data.fillna(0, inplace=True)

# Remove duplicates (if any)
data.drop_duplicates(inplace=True)

print("\nData cleaned!")
```

```
Data cleaned!
```

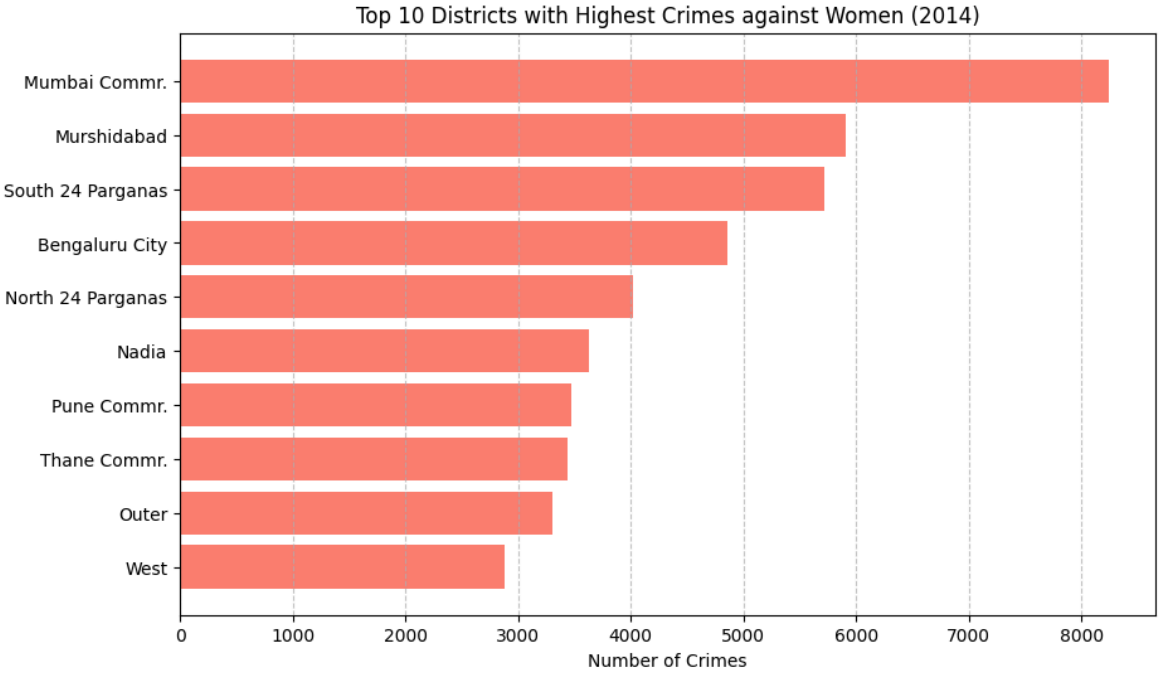
In [20]:
```
# MATPLOTLIB VISUALIZATION
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
# STATE-WISE CRIME DISTRIBUTION
state_crime = data.groupby('States/UTs')['Total Crimes against Women'].sum().sor

plt.figure(figsize=(12, 6))
state_crime.plot(kind='bar', color='skyblue')
plt.title('Total Crimes against Women by State (2014)')
plt.ylabel('Number of Crimes')
plt.xticks(rotation=45, ha='right')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```



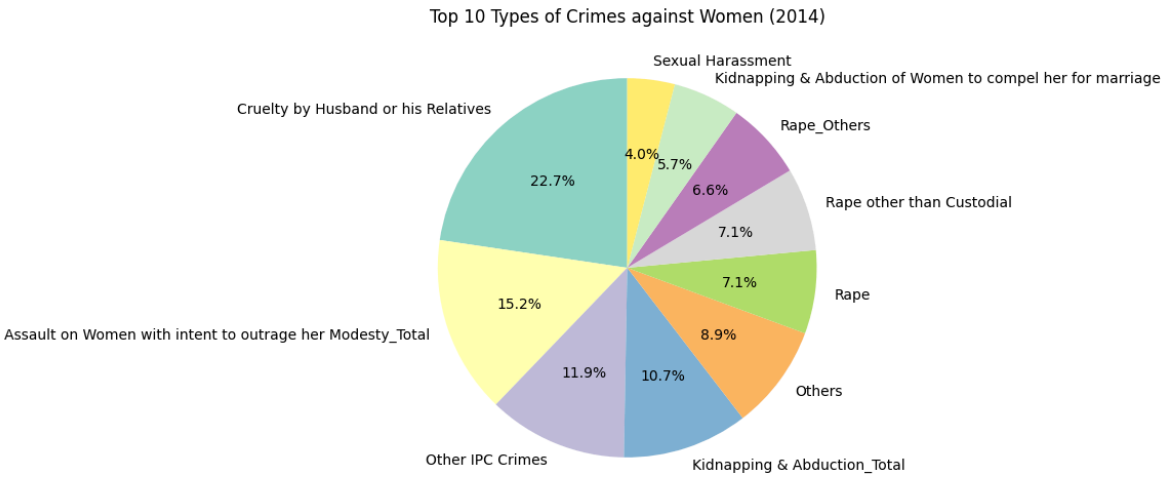
In [23]:
```
# TOP 10 DISTRICTS WITH HIGHEST CRIMES
print(top_districts)
```

	District	Total Crimes against Women
404	Total	57101
755	Total	52409
451	Total	49167
800	Total	48394
605	Total	32401
831	Total	22255
68	Total	19609
533	Total	18564
115	Total	18208
326	Total	17792

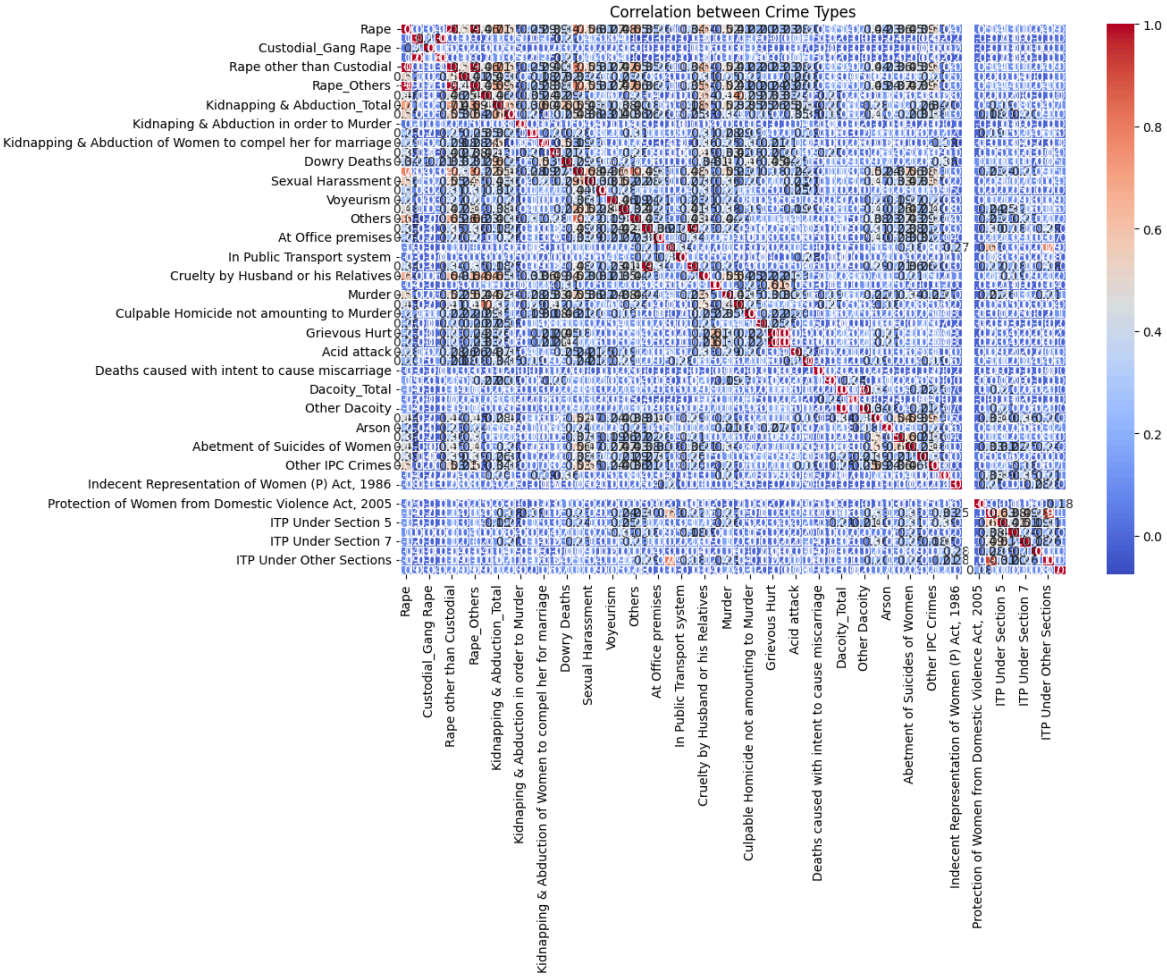
In [26]:
```
data = data[data['District'] != 'Total']

top_districts = data[['District', 'Total Crimes against Women']].nlargest(10, 'T

plt.figure(figsize=(10, 6))
plt.barh(top_districts['District'], top_districts['Total Crimes against Women'],
plt.title('Top 10 Districts with Highest Crimes against Women (2014)')
plt.xlabel('Number of Crimes')
plt.gca().invert_yaxis()
plt.grid(axis='x', linestyle='--', alpha=0.7)
plt.show()
```



In [27]:
```
# CRIME TYPE BREAKDOWN
crime_type_totals = data.iloc[:, 3:-1].sum().sort_values(ascending=False)[:10]

plt.figure(figsize=(10, 6))
crime_type_totals.plot(kind='pie', autopct='%1.1f%%', startangle=90, colormap='S
plt.title('Top 10 Types of Crimes against Women (2014)')
plt.ylabel('')
plt.show()
```

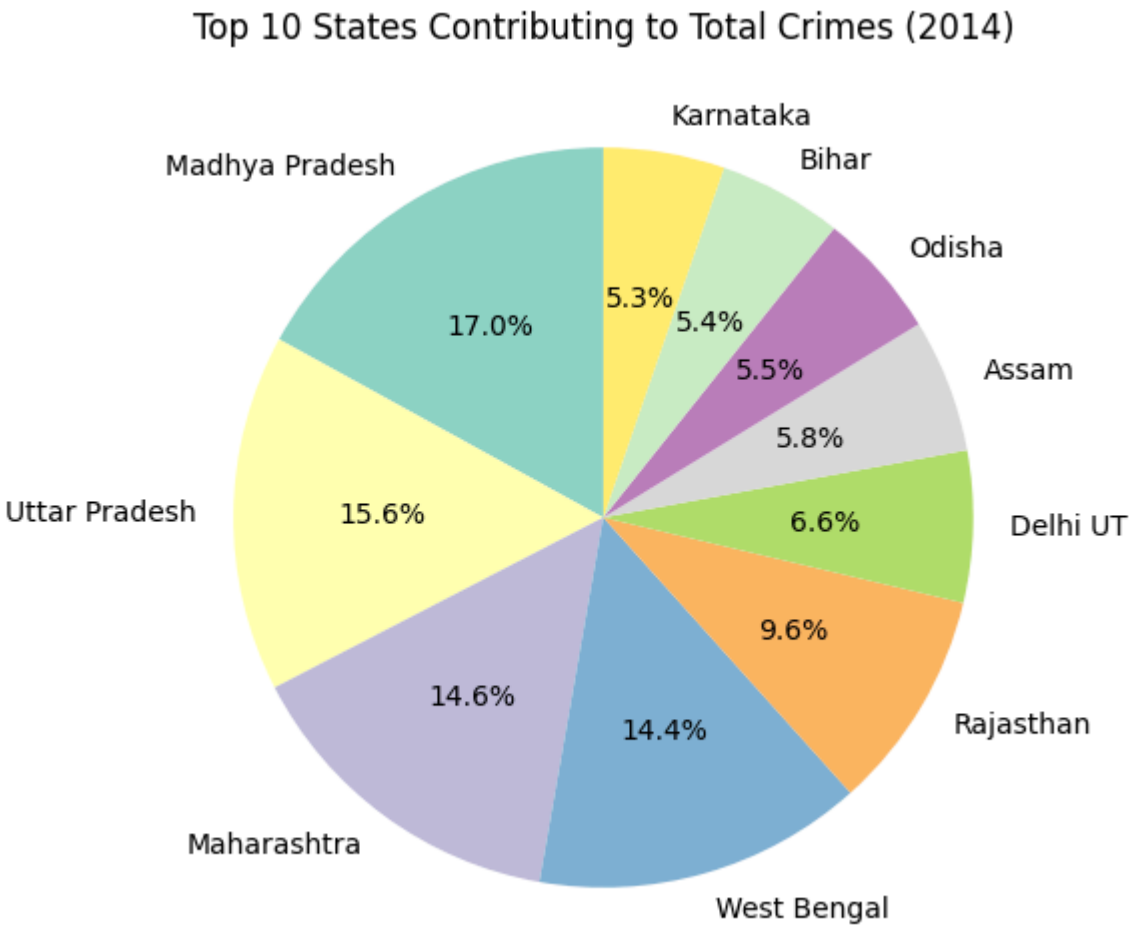


In [28]:
```
# CORRELATION HEATMAP
plt.figure(figsize=(12, 8))
sns.heatmap(data.iloc[:, 3:-1].corr(), cmap='coolwarm', annot=True, fmt='.2f', l
plt.title('Correlation between Crime Types')
plt.show()
```

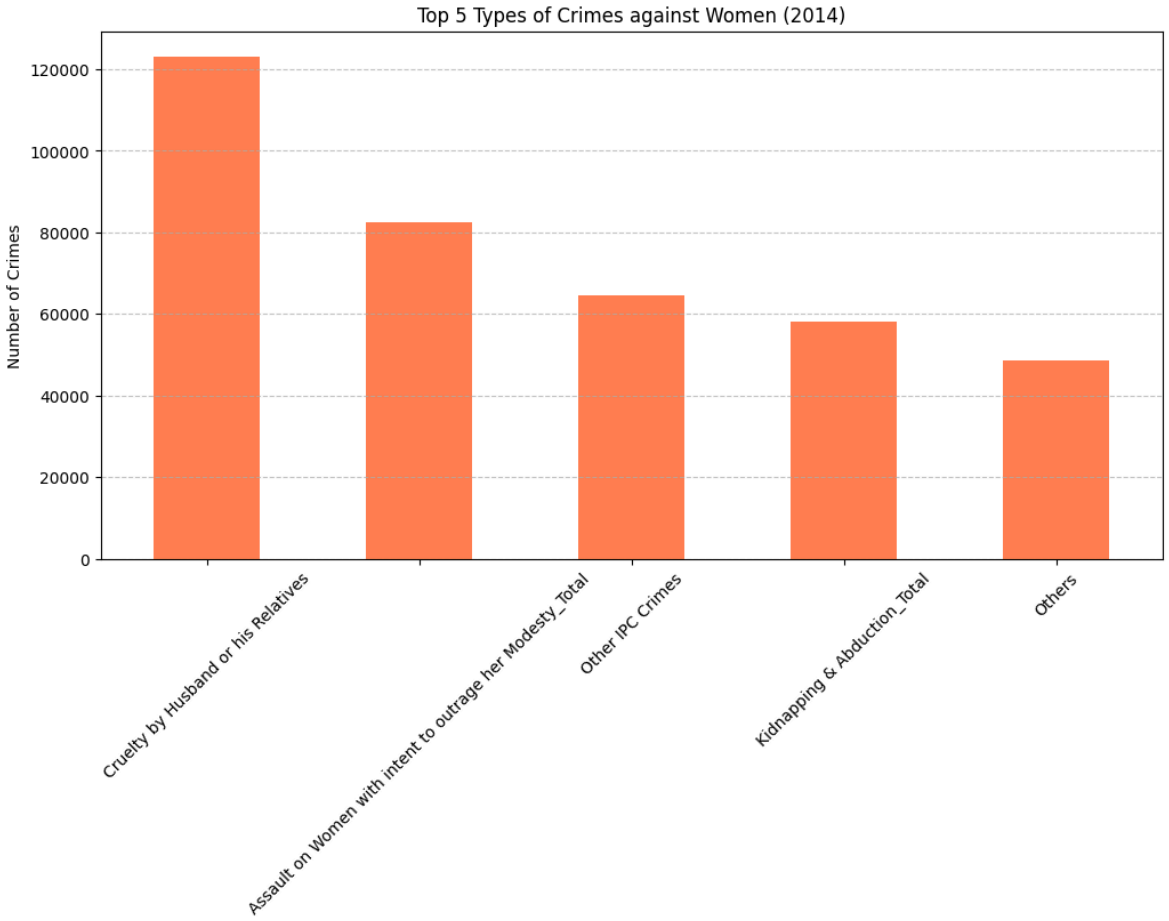


In [29]:
```
# STATE CONTRIBUTION TO TOTAL CRIMES
state_contribution = data.groupby('States/UTs')['Total Crimes against Women'].su

plt.figure(figsize=(10, 6))
state_contribution.plot(kind='pie', autopct='%1.1f%%', startangle=90, colormap='
plt.title('Top 10 States Contributing to Total Crimes (2014)')
```

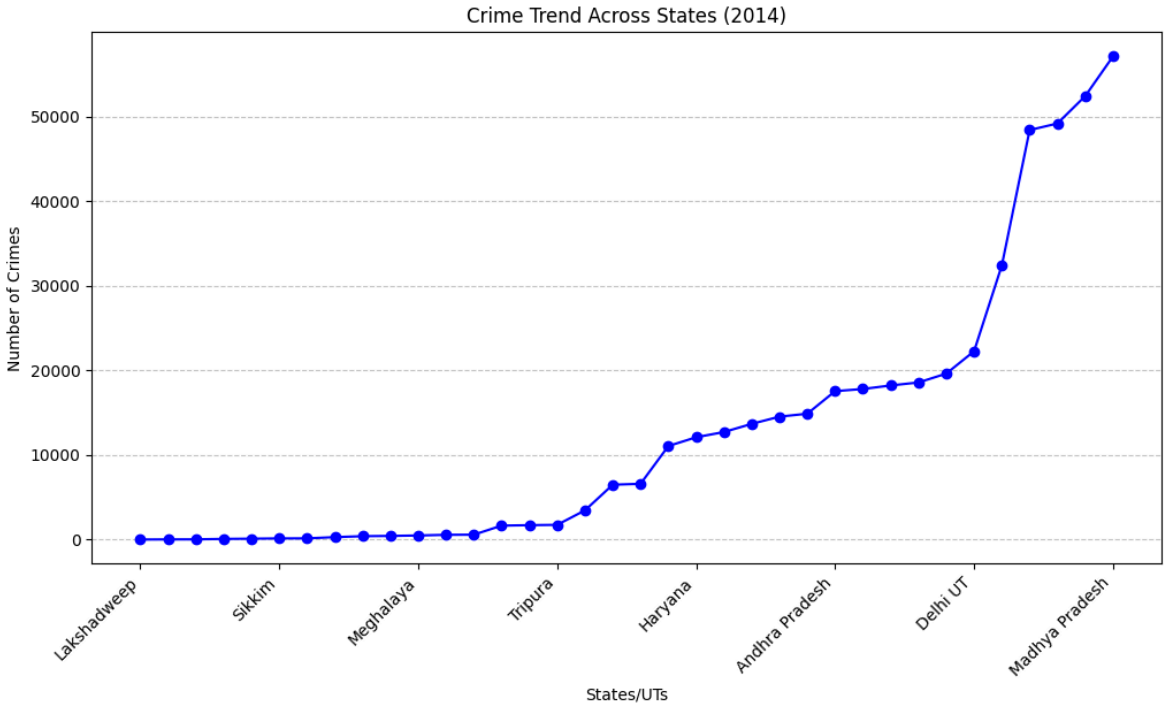
```
plt.ylabel('')
plt.show()
```



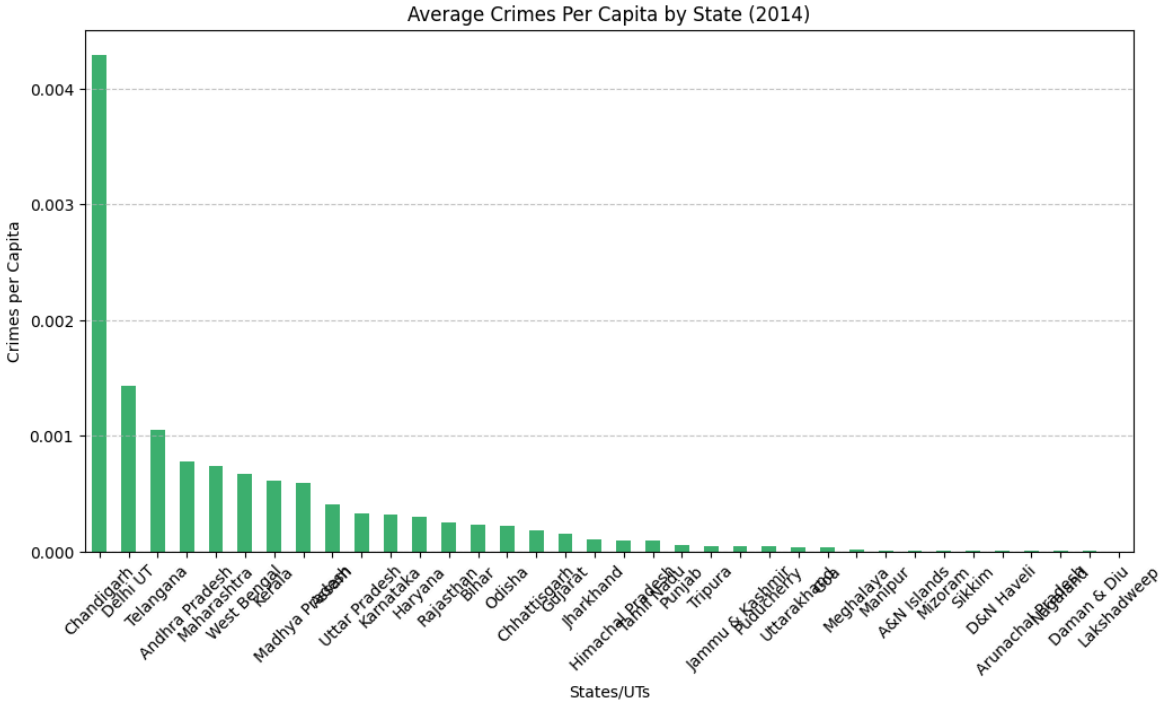
In [30]:
```
# CRIMES BY TYPE (TOP 5)
top_5_crimes = data.iloc[:, 3:-1].sum().nlargest(5)

plt.figure(figsize=(12, 6))
top_5_crimes.plot(kind='bar', color='coral')
plt.title('Top 5 Types of Crimes against Women (2014)')
plt.ylabel('Number of Crimes')
plt.xticks(rotation=45)
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```



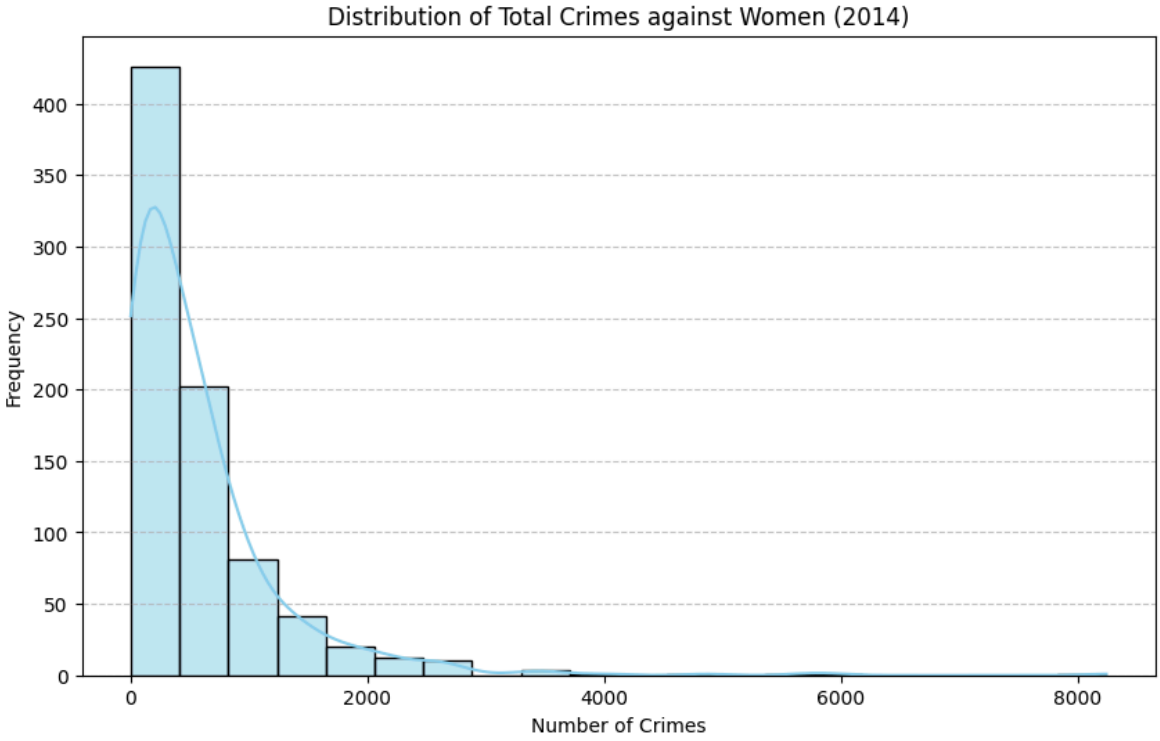
In [31]:
```
# CRIME TREND ACROSS STATES
state_crime_trend = data.groupby('States/UTs')['Total Crimes against Women'].sum

plt.figure(figsize=(12, 6))
state_crime_trend.plot(marker='o', color='blue')
plt.title('Crime Trend Across States (2014)')
plt.ylabel('Number of Crimes')
plt.xticks(rotation=45, ha='right')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```



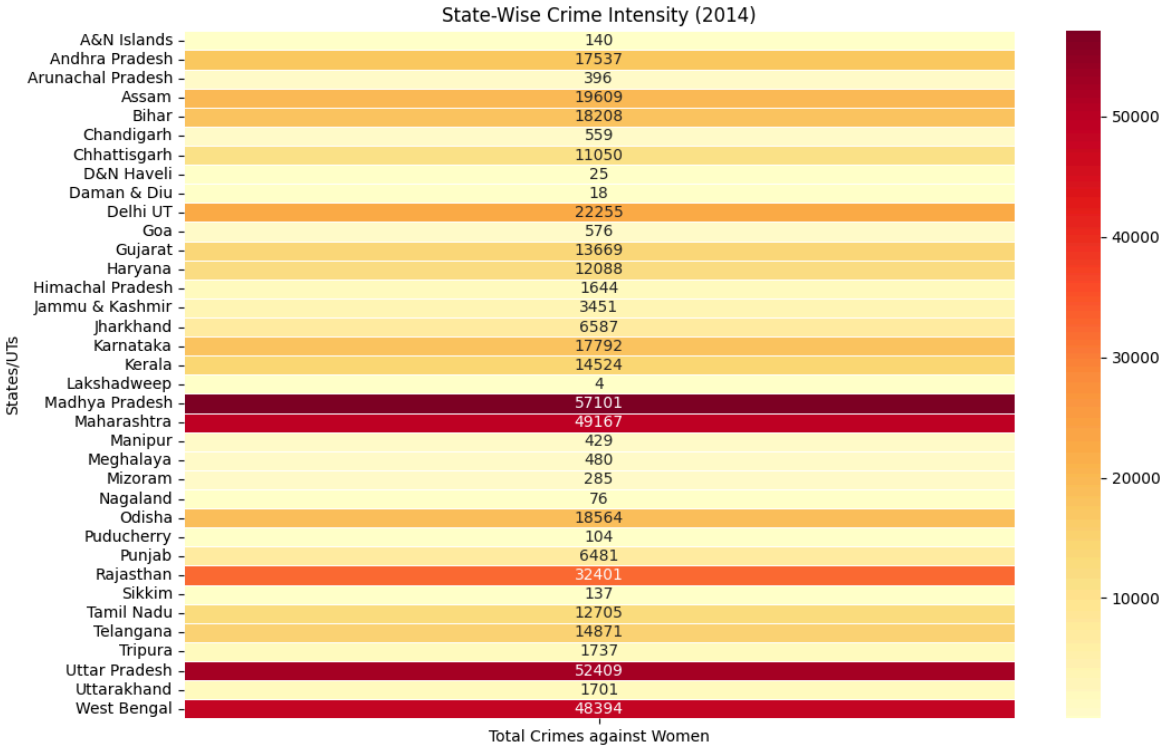
In [39]:
```
# CRIMES PER CAPITA (since population available)
# Assuming hypothetical population data for demonstration:
data['Population'] = np.random.randint(100000, 10000000, size=len(data))

data['Crimes per Capita'] = data['Total Crimes against Women'] / data['Populatio

state_crime_per_capita = data.groupby('States/UTs')['Crimes per Capita'].mean().

plt.figure(figsize=(12, 6))
state_crime_per_capita.plot(kind='bar', color='mediumseagreen')
plt.title('Average Crimes Per Capita by State (2014)')
plt.ylabel('Crimes per Capita')
plt.xticks(rotation=45)
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```



In [40]:
```
# DISTRIBUTION OF TOTAL CRIMES
plt.figure(figsize=(10, 6))
sns.histplot(data['Total Crimes against Women'], kde=True, color='skyblue', bins
plt.title('Distribution of Total Crimes against Women (2014)')
plt.xlabel('Number of Crimes')
plt.ylabel('Frequency')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```

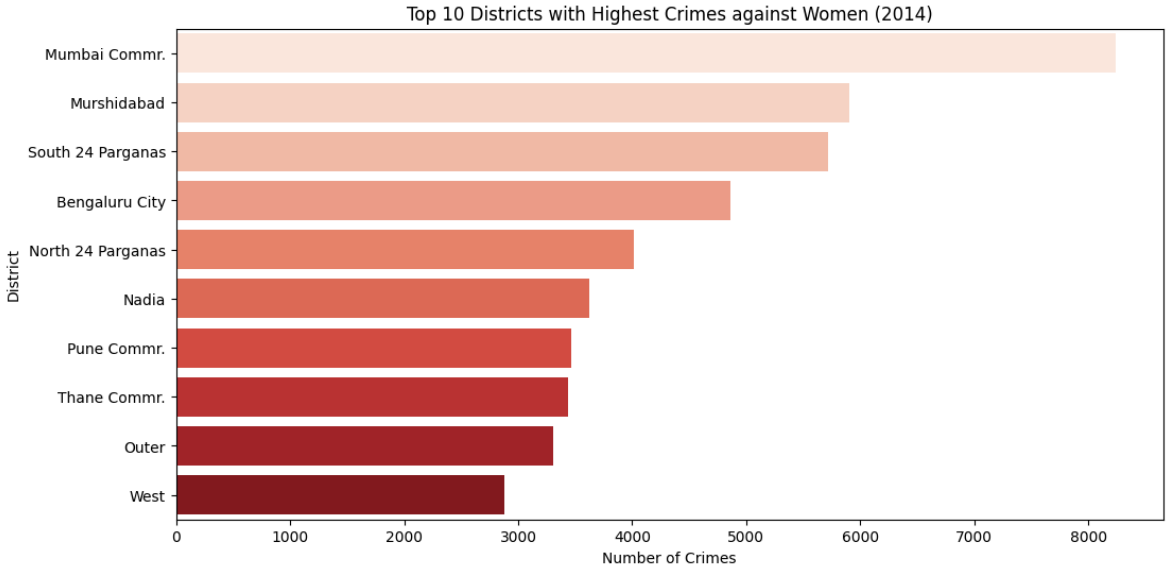


In [41]:
```
# STATE-WISE CRIME INTENSITY (HEATMAP)
state_crime_matrix = data.pivot_table(index='States/UTs', values='Total Crimes a

plt.figure(figsize=(12, 8))
sns.heatmap(state_crime_matrix, cmap='YlOrRd', annot=True, fmt='.0f', linewidths
plt.title('State-Wise Crime Intensity (2014)')
plt.show()
```

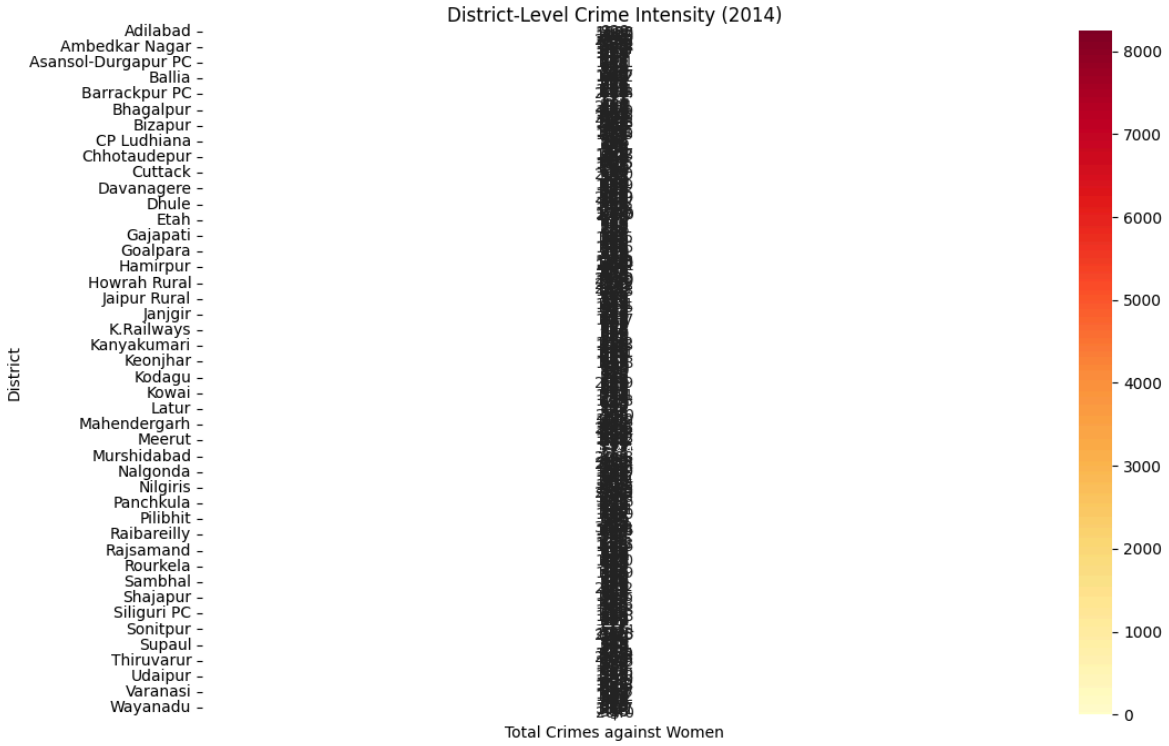


In [42]:
```
# Remove rows where District is 'Total'
data = data[data['District'] != 'Total']
# TOP 10 DISTRICTS BY TOTAL CRIMES
top_districts = data[['District', 'Total Crimes against Women']].nlargest(10, 'T

plt.figure(figsize=(12, 6))
sns.barplot(data=top_districts, x='Total Crimes against Women', y='District', pa
plt.title('Top 10 Districts with Highest Crimes against Women (2014)')
plt.xlabel('Number of Crimes')
plt.ylabel('District')
plt.show()
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_3720\1671144253.py:7: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `y` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.barplot(data=top_districts, x='Total Crimes against Women', y='District', p
alette='Reds')
```

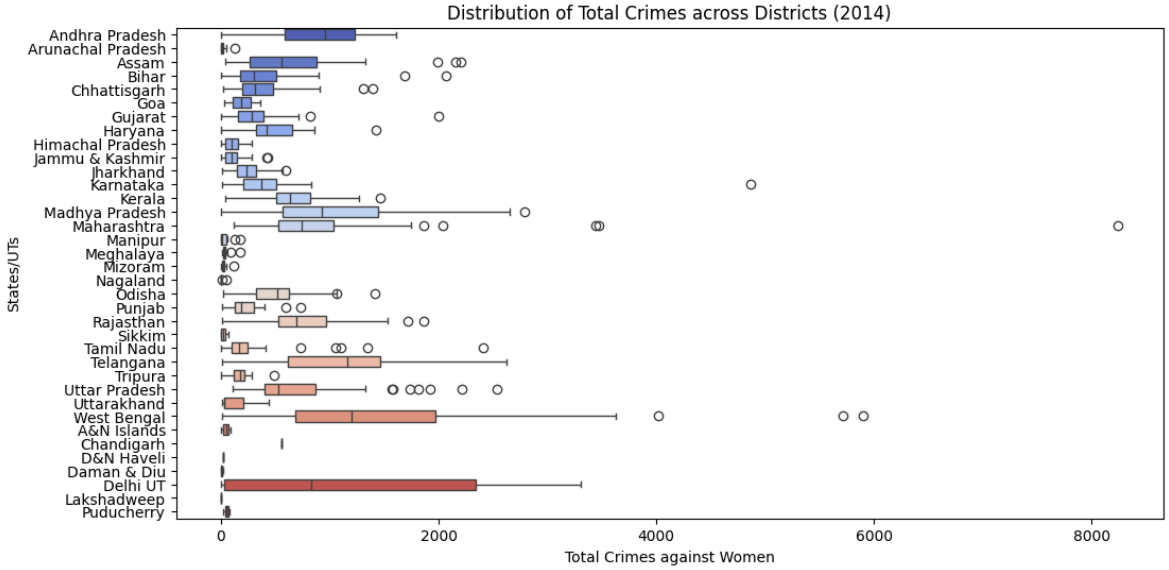


In [47]:
```
#  DISTRICT-LEVEL HEATMAP
district_crime_matrix = data.pivot_table(index='District', values='Total Crimes

plt.figure(figsize=(12, 8))
sns.heatmap(district_crime_matrix, cmap='YlOrRd', annot=True, fmt='.0f', linewid
plt.title('District-Level Crime Intensity (2014)')
plt.show()
```



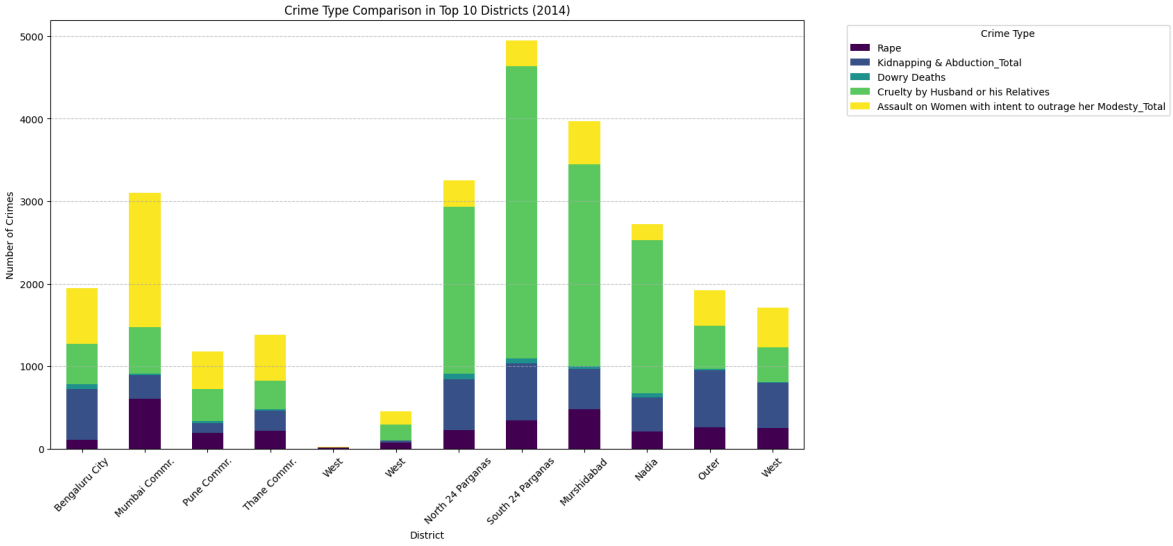
In [48]:
```
# CRIME DISTRIBUTION ACROSS DISTRICTS
plt.figure(figsize=(12, 6))
sns.boxplot(data=data, x='Total Crimes against Women', y='States/UTs', palette='
plt.title('Distribution of Total Crimes across Districts (2014)')
plt.show()
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_3720\1122401964.py:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `y` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.boxplot(data=data, x='Total Crimes against Women', y='States/UTs', palette
='coolwarm')
```



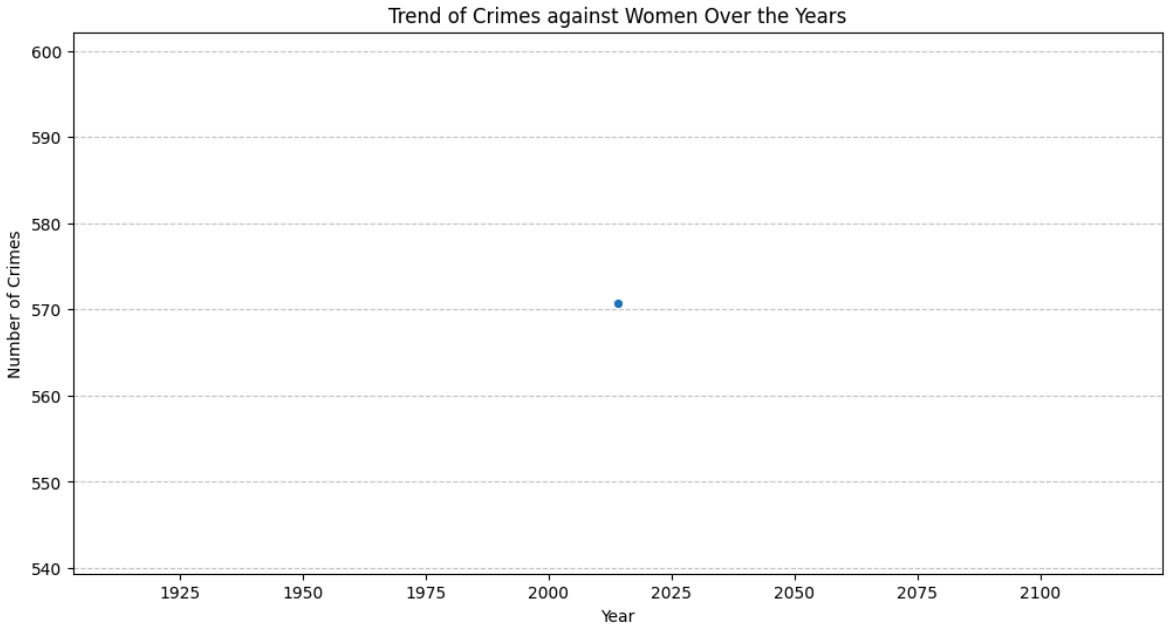
In [50]:

```
# DISTRICT-WISE CRIME TYPE COMPARISON (STACKED BAR CHART)
crime_types = ['Rape', 'Kidnapping & Abduction_Total', 'Dowry Deaths',
               'Cruelty by Husband or his Relatives',
               'Assault on Women with intent to outrage her Modesty_Total']

top_districts_crime = data.loc[data['District'].isin(top_districts['District']),

top_districts_crime.set_index('District').plot(kind='bar', stacked=True, figsize
plt.title('Crime Type Comparison in Top 10 Districts (2014)')
plt.ylabel('Number of Crimes')
plt.xticks(rotation=45)
plt.legend(title='Crime Type', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```



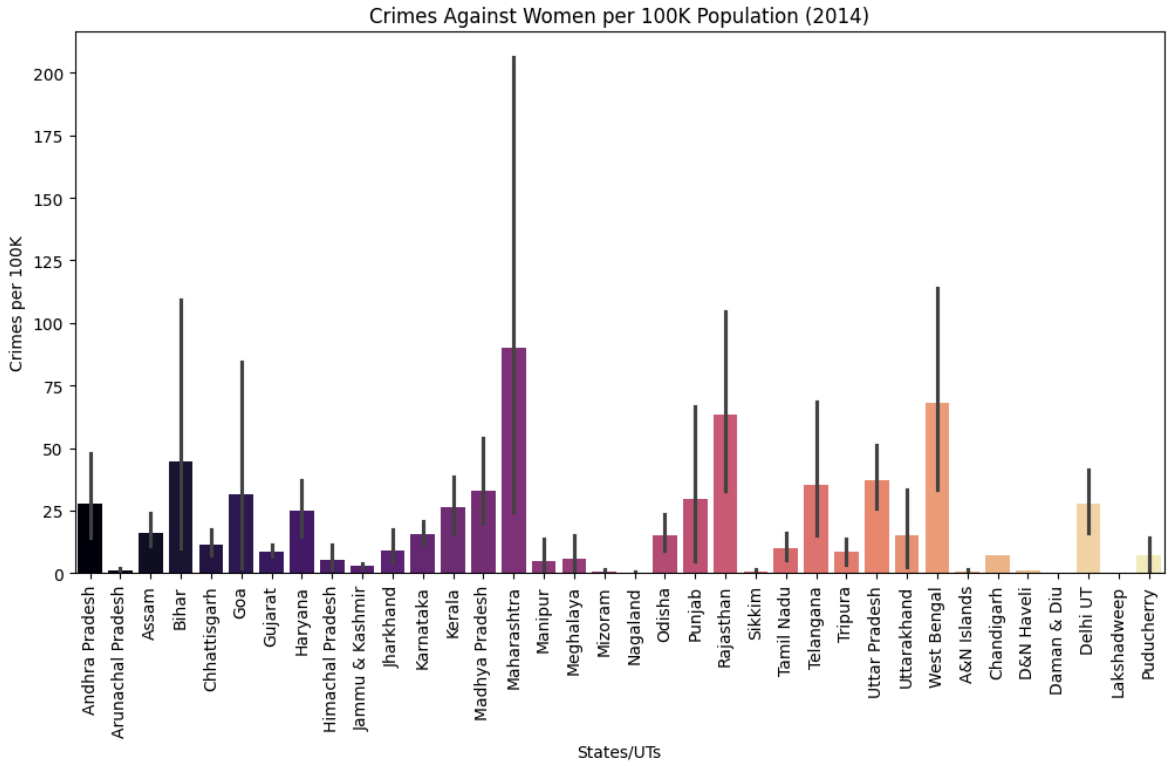
In [51]:
```
# CRIME TREND OVER TIME
plt.figure(figsize=(12, 6))
sns.lineplot(data=data, x='Year', y='Total Crimes against Women', marker='o', ci
plt.title('Trend of Crimes against Women Over the Years')
plt.ylabel('Number of Crimes')
plt.xlabel('Year')
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_3720\42406162.py:3: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

  sns.lineplot(data=data, x='Year', y='Total Crimes against Women', marker='o', c
i=None)
```



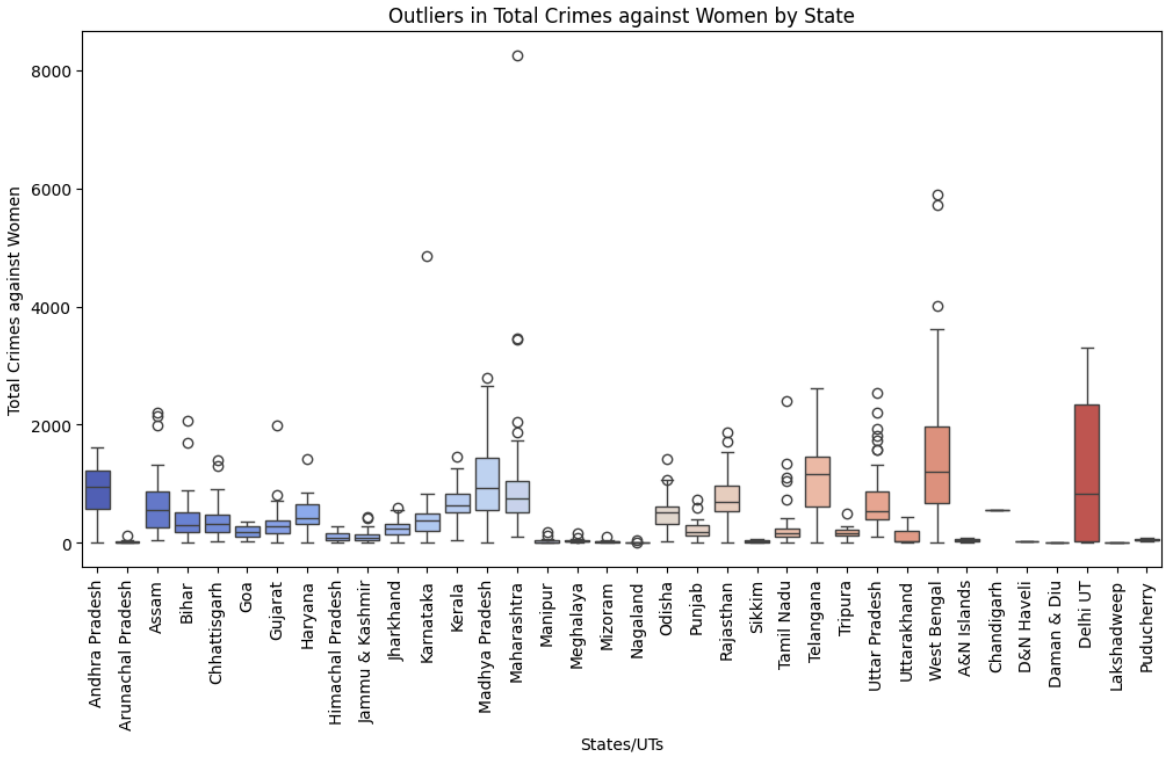
In [53]:
```
# CRIMES AGAINST WOMENBY POPULATION
data['Population'] = np.random.randint(100000, 10000000, size=len(data))  # Hypo
data['Crimes per 100K'] = (data['Total Crimes against Women'] / data['Population

plt.figure(figsize=(12, 6))
sns.barplot(data=data, x='States/UTs', y='Crimes per 100K', palette='magma')
plt.title('Crimes Against Women per 100K Population (2014)')
plt.xticks(rotation=90)
plt.ylabel('Crimes per 100K')
plt.show()
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_3720\3971166026.py:6: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.barplot(data=data, x='States/UTs', y='Crimes per 100K', palette='magma')
```



In [54]:
```
# HIGHLIGHTING OUTLIERS
plt.figure(figsize=(12, 6))
sns.boxplot(data=data, x='States/UTs', y='Total Crimes against Women', palette='
plt.title('Outliers in Total Crimes against Women by State')
plt.xticks(rotation=90)
plt.show()
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_3720\3767720324.py:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.boxplot(data=data, x='States/UTs', y='Total Crimes against Women', palette
='coolwarm')
```



In [57]:
```
# FINAL STATS
# Summary of total crimes
total_crimes_summary = data['Total Crimes against Women'].describe()
print(total_crimes_summary)
```

```
count     801.000000
mean      570.754057
std       718.395196
min         0.000000
25%       139.000000
50%       379.000000
75%       731.000000
max      8244.000000
Name: Total Crimes against Women, dtype: float64
```

In [59]:
```
# Top 5 States by Total Crimes
top_states = data.groupby('States/UTs')['Total Crimes against Women'].sum().nlar
print(top_states)
```

```
States/UTs
Madhya Pradesh    57101
Uttar Pradesh     52409
Maharashtra       49167
West Bengal       48394
Rajasthan         32401
Name: Total Crimes against Women, dtype: int64
```

In [60]:
```
# Bottom 5 States by Total Crimes
bottom_states = data.groupby('States/UTs')['Total Crimes against Women'].sum().n
print(bottom_states)
```

```
States/UTs
Lakshadweep      4
Daman & Diu     18
D&N Haveli      25
Nagaland        76
Puducherry     104
Name: Total Crimes against Women, dtype: int64
```

In [61]:
```
# State with Maximum and Minimum Crimes
max_state = top_states.idxmax()
min_state = bottom_states.idxmin()

print(f"State with highest crimes: {max_state}")
print(f"State with lowest crimes: {min_state}")
```

```
State with highest crimes: Madhya Pradesh
State with lowest crimes: Lakshadweep
```

In [62]:
```
# Crime Type Contribution (%)
crime_types = ['Rape', 'Kidnapping & Abduction_Total', 'Dowry Deaths',
               'Cruelty by Husband or his Relatives',
               'Assault on Women with intent to outrage her Modesty_Total']

crime_contribution = data[crime_types].sum() / data['Total Crimes against Women'
print(crime_contribution)
```

```
Rape                                                 8.460236
Kidnapping & Abduction_Total                        12.738913
Dowry Deaths                                         1.850061
Cruelty by Husband or his Relatives                 26.940071
Assault on Women with intent to outrage her Modesty_Total    18.026178
dtype: float64
```

In [63]:
```
# District with Highest and Lowest Crimes
max_district = data.loc[data['Total Crimes against Women'].idxmax(), ['District'
min_district = data.loc[data['Total Crimes against Women'].idxmin(), ['District'

print(f"District with highest crimes:\n{max_district}")
print(f"District with lowest crimes:\n{min_district}")
```

```
District with highest crimes:
District                      Mumbai Commr.
Total Crimes against Women             8244
Name: 423, dtype: object
District with lowest crimes:
District                      Crime Branch
Total Crimes against Women               0
Name: 23, dtype: object
```

In [ ]:
```

```