```
Import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load the dataset
file_path = '/mnt/data/Financial Analytics data.csv'
df = pd.read_csv(file_path)

# Display the first few rows of the dataset
print("First few rows of the dataset:")
print(df.head())

# Check for missing values
print("\nMissing values in each column:")
print(df.isnull().sum())

# Summary statistics of the dataset
print("\nSummary statistics:")
print(df.describe())

# Preprocess the data
# Drop rows with missing values (if any)
df.dropna(inplace=True)

# Convert columns to appropriate data types
df['Mar Cap – Crore'] = df['Mar Cap – Crore'].str.replace(',', '').astype(float)
df['Sales Qtr – Crore'] = df['Sales Qtr – Crore'].str.replace(',', '').astype(float)

# Exploratory Data Analysis (EDA)
# Correlation matrix
corr_matrix = df.corr()

# Plot the correlation matrix
plt.figure(figsize=(10, 8))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()

# Scatter plot: Market Capitalization vs. Quarterly Sales
plt.figure(figsize=(10, 6))
sns.scatterplot(data=df, x='Sales Qtr – Crore', y='Mar Cap – Crore', hue='Name of
Company', legend=False)
plt.title('Market Capitalization vs. Quarterly Sales')
plt.xlabel('Quarterly Sales (in Crores)')
plt.ylabel('Market Capitalization (in Crores)')
plt.show()

# Top 10 companies by market capitalization
top_10_companies = df.nlargest(10, 'Mar Cap – Crore')
```

```
# Bar plot: Top 10 companies by market capitalization
plt.figure(figsize=(12, 6))
sns.barplot(data=top_10_companies, x='Name of Company', y='Mar Cap – Crore')
plt.title('Top 10 Companies by Market Capitalization')
plt.xlabel('Company Name')
plt.ylabel('Market Capitalization (in Crores)')
plt.xticks(rotation=45)
plt.show()

# Distribution of market capitalization
plt.figure(figsize=(10, 6))
sns.histplot(df['Mar Cap – Crore'], kde=True, bins=30)
plt.title('Distribution of Market Capitalization')
plt.xlabel('Market Capitalization (in Crores)')
plt.ylabel('Frequency')
plt.show()

# Distribution of quarterly sales
plt.figure(figsize=(10, 6))
sns.histplot(df['Sales Qtr – Crore'], kde=True, bins=30)
plt.title('Distribution of Quarterly Sales')
plt.xlabel('Quarterly Sales (in Crores)')
plt.ylabel('Frequency')
plt.show()
```