

```
In [2]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

# Load the dataset
file_path = 'D:\\\\Internship\\\\Financial Analytics\\\\Financial Analytics data.csv'
data = pd.read_csv(file_path)

# Preview the dataset
print("First few rows of the dataset:")
print(data.head())
print("\\nColumns in the dataset:")
print(data.columns)

# --- Data Cleaning ---
# Handle missing values by filling them with a placeholder
data.fillna('Missing', inplace=True)

# Drop unnecessary columns (if any)
data.drop(columns=['Unnamed: 4'], inplace=True, errors='ignore')

# Convert relevant columns to numeric for analysis
data['Mar Cap - Crore'] = pd.to_numeric(data['Mar Cap - Crore'], errors='coerce')
data['Sales Qtr - Crore'] = pd.to_numeric(data['Sales Qtr - Crore'], errors='coerce')

# --- Basic Statistics ---
# Display basic statistics to check for data anomalies
print("\\nBasic Statistics of the Dataset:")
print(data.describe())

# --- Key Metrics Analysis ---
# Calculate and display key metrics for Market Capitalization and Sales
average_mar_cap = data['Mar Cap - Crore'].mean()
total_sales = data['Sales Qtr - Crore'].sum()
median_mar_cap = data['Mar Cap - Crore'].median()
median_sales = data['Sales Qtr - Crore'].median()
std_dev_mar_cap = data['Mar Cap - Crore'].std()
std_dev_sales = data['Sales Qtr - Crore'].std()

print(f"\\nKey Metrics:")
print(f"Average Market Capitalization: {average_mar_cap:.2f} Crore")
print(f"Total Quarterly Sales: {total_sales:.2f} Crore")
print(f"Median Market Capitalization: {median_mar_cap:.2f} Crore")
print(f"Median Quarterly Sales: {median_sales:.2f} Crore")
print(f"Standard Deviation of Market Capitalization: {std_dev_mar_cap:.2f} Crore")
print(f"Standard Deviation of Quarterly Sales: {std_dev_sales:.2f} Crore")

# --- Correlation Analysis ---
# Check the correlation between Market Capitalization and Quarterly Sales
correlation = data[['Mar Cap - Crore', 'Sales Qtr - Crore']].corr()
print("\\nCorrelation between Market Capitalization and Quarterly Sales:")
print(correlation)

# --- Data Visualization ---

# Scatter plot of Market Capitalization vs Quarterly Sales
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Mar Cap - Crore', y='Sales Qtr - Crore', data=data)
plt.title('Market Capitalization vs Quarterly Sales')
plt.xlabel('Market Capitalization (Crore)')
plt.ylabel('Quarterly Sales (Crore)')
plt.grid(True)
plt.savefig('D:\\\\Internship\\\\sol1\\\\Market_Cap_vs_Sales.png')
plt.show()

# Distribution of Market Capitalization and Quarterly Sales
plt.figure(figsize=(14, 6))

plt.subplot(1, 2, 1)
sns.histplot(data['Mar Cap - Crore'], dropna(), kde=True, color='blue')
plt.title('Distribution of Market Capitalization')
plt.xlabel('Market Capitalization (Crore)')
plt.ylabel('Frequency')

plt.subplot(1, 2, 2)
sns.histplot(data['Sales Qtr - Crore'], dropna(), kde=True, color='green')
plt.title('Distribution of Quarterly Sales')
plt.xlabel('Quarterly Sales (Crore)')
plt.ylabel('Frequency')

plt.tight_layout()
plt.savefig('D:\\\\Internship\\\\sol1\\\\Distribution_Mar_Cap_and_Sales.png')
plt.show()

# Box plots to detect outliers
plt.figure(figsize=(14, 6))

plt.subplot(1, 2, 1)
sns.boxplot(x=data['Mar Cap - Crore'])
plt.title('Box Plot of Market Capitalization')

plt.subplot(1, 2, 2)
sns.boxplot(x=data['Sales Qtr - Crore'])
plt.title('Box Plot of Quarterly Sales')

plt.tight_layout()
plt.savefig('D:\\\\Internship\\\\sol1\\\\Box_Plots_Mar_Cap_and_Sales.png')
plt.show()

# --- Top 10 Companies by Market Capitalization ---
# Extract the top 10 companies by Market Capitalization
top_companies = data[['Name', 'Mar Cap - Crore']].sort_values(by='Mar Cap - Crore', ascending=False).head(10)

print("\\nTop 10 Companies by Market Capitalization:")
print(top_companies)

# Bar plot for the top 10 companies by Market Capitalization
plt.figure(figsize=(12, 8))
sns.barplot(x='Mar Cap - Crore', y='Name', data=top_companies)
plt.title('Top 10 Companies by Market Capitalization')
plt.xlabel('Market Capitalization (Crore)')
plt.ylabel('Company Name')
plt.grid(True)
plt.savefig('D:\\\\Internship\\\\sol1\\\\Top_10_Companies_by_Market_Cap.png')
plt.show()

# --- Save Key Metrics to a CSV File ---
# Store the calculated key metrics in a CSV file for reporting
metrics = pd.DataFrame({
    'Metric': ['Average Market Capitalization', 'Total Quarterly Sales', 'Median Market Capitalization', 'Median Quarterly Sales', 'Standard Deviation of Market Capitalization', 'Standard Deviation of Quarterly Sales'],
    'Value': [average_mar_cap, total_sales, median_mar_cap, median_sales, std_dev_mar_cap, std_dev_sales]
})

metrics.to_csv('D:\\\\Internship\\\\sol1\\\\Key_Metrics.csv', index=False)

# --- Additional Analysis (Optional) ---
# Uncomment the following lines if sector/industry analysis is needed
# sector_analysis = data.groupby('Sector')[['Mar Cap - Crore', 'Sales Qtr - Crore']].mean()
# print("\\nSector-wise Analysis:")
# print(sector_analysis)
```

First few rows of the dataset:

S.No.	Name	Mar Cap - Crore	Sales Qtr - Crore	Unnamed: 4
0	1 Reliance Inds.	583436.72	99810.00	NaN
1	2 TCS	563709.84	30904.00	NaN
2	3 HDFC Bank	482953.59	20581.27	NaN
3	4 ITC	320985.27	9772.02	NaN
4	5 H D F C	289497.37	16840.51	NaN

Columns in the dataset:

```
Index(['S.No.', 'Name', 'Mar Cap - Crore', 'Sales Qtr - Crore', 'Unnamed: 4'], dtype='object')
```

Basic Statistics of the Dataset:

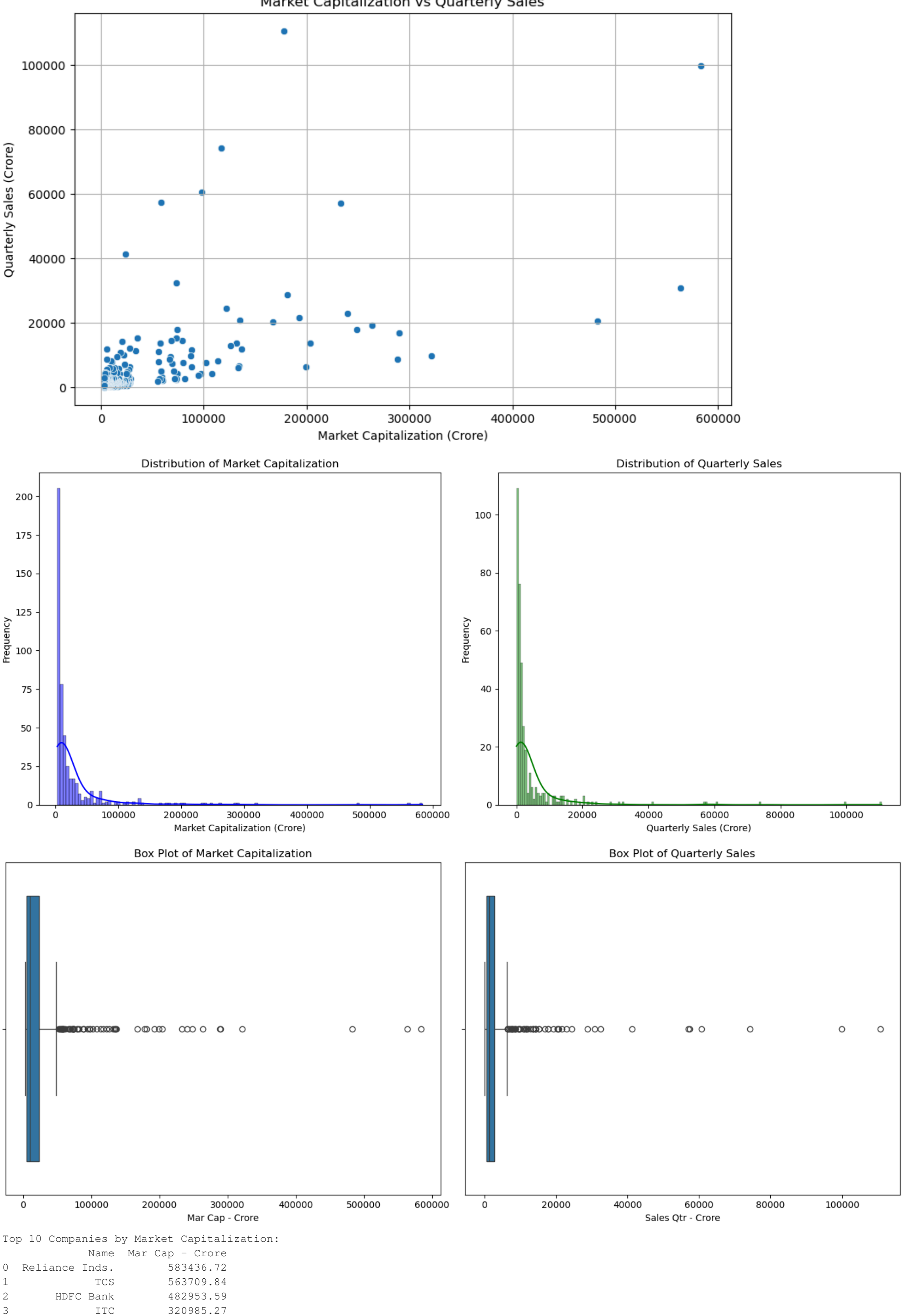
	S.No.	Mar Cap - Crore	Sales Qtr - Crore
count	488.000000	479.000000	365.000000
mean	251.508197	28043.857119	4395.976849
std	145.884078	59464.615831	11092.206185
min	1.000000	3017.070000	47.240000
25%	122.750000	4843.575000	593.740000
50%	252.500000	9885.050000	1278.300000
75%	378.250000	23549.900000	2840.750000
max	500.000000	583436.720000	110666.930000

Key Metrics:

Average Market Capitalization: 28043.86 Crore
Total Quarterly Sales: 1604531.55 Crore
Median Market Capitalization: 9885.05 Crore
Median Quarterly Sales: 1278.30 Crore
Standard Deviation of Market Capitalization: 59464.62 Crore
Standard Deviation of Quarterly Sales: 11092.21 Crore

Correlation between Market Capitalization and Quarterly Sales:

	Mar Cap - Crore	Sales Qtr - Crore
Mar Cap - Crore	1.000000	0.620702
Sales Qtr - Crore	0.620702	1.000000



Top 10 Companies by Market Capitalization:

	Name	Mar Cap - Crore
0	Reliance Inds.	583436.72
1	TCS	563709.84
2	HDFC Bank	482953.59
3	ITC	320985.27
4	H D F C	289497.37
5	Hind. Unilever	288265.26
6	Maruti Suzuki	263493.81
7	Infosys	248320.35
8	O N G C	239981.50
9	St Bk of India	232763.33

Top 10 Companies by Market Capitalization

Company Name	Market Capitalization (Crore)
Reliance Inds.	583436.72
TCS	563709.84
HDFC Bank	482953.59
ITC	320985.27
H D F C	289497.37
Hind. Unilever	288265.26
Maruti Suzuki	263493.81
Infosys	248320.35
O N G C	239981.50
St Bk of India	232763.33