Document on Financial Analytics
Data Analysis

Problem Statement

Without analyzing the competition, it is difficult for a business to survive. You are tasked to analyze the competition for the management to provide better results. This dataset contains information on the market capitalization of the top 500 companies in India.

Serial Number, Name, Name of Company, Mar Cap – Crore (Market Capitalization in Crores), Sales Qtr – Crore (Quarterly Sale in crores) are the columns in the dataset. Find key metrics and factors and show the meaningful relationships between attributes.

Do your own research and come up with your findings.

Introduction

In the ever-evolving landscape of business, the ability to navigate and thrive in a competitive environment is paramount for sustained success. Without a profound understanding of the competition, businesses face the risk of stagnation or even failure. Recognizing this imperative, our analysis aims to delve into the dynamics of competition within the Indian corporate landscape, focusing on the market capitalization of the top 500 companies. The dataset under scrutiny encapsulates critical information such as Serial Number, Company Name, Market Capitalization, and Quarterly Sales. By leveraging analytical methodologies, we aspire to unearth key metrics, discern influential factors, and unravel meaningful relationships between attributes. This investigation is driven by the overarching goal of equipping management with insights that pave the way for informed decision-making, strategic planning, and ultimately, superior business outcomes.

As we embark on this exploration, we anticipate unraveling the intricate tapestry of the Indian business ecosystem, shedding light on the factors that propel companies to the summit of market capitalization, and uncovering the nuances that define successful market players. The multifaceted nature of this analysis, spanning data cleansing, metric computation, visual representation, and correlation assessment, promises a comprehensive understanding of the dataset. It is our contention that these insights will not only serve as a foundational understanding of the current competitive landscape but also lay the groundwork for subsequent in-depth analyses and strategic initiatives.

This analysis assumes a holistic perspective, acknowledging the diversity and complexity inherent in the top 500 companies. Through a judicious combination of statistical measures, visualizations, and interpretative findings, we aim to deliver a compelling narrative that not only addresses the immediate requirements of the problem statement but also acts as a catalyst for continued exploration and refinement. The report's structure, characterized by modularized and methodical coding practices, is designed to facilitate clarity, reproducibility, and adaptability, ensuring that the analytical journey remains robust and insightful.

Code Demonstration

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

def load_data(file_path):
    """Load data from a CSV file and return the DataFrame."""
    df = pd.read_csv(file_path)
    return df

def handle_missing_values(data_frame):
    """Droping rows with missing values and return the cleaned
DataFrame."""
    cleaned_df = data_frame.dropna()
    return cleaned_df

def calculate_mean_metrics(data_frame):
    """Calculating mean market capitalization and mean quarterly
sales."""
    mean_mar_cap = data_frame['Mar Cap - Crore'].mean()
    mean_sales_qtr = data_frame['Sales Qtr - Crore'].mean()

    print(f"Mean Market Capitalization: {mean_mar_cap}")
    print(f"Mean Quarterly Sales: {mean_sales_qtr}")

def create_scatter_plot(data_frame):
    """Creating a scatter plot of market capitalization vs. quarterly
sales."""
    sns.scatterplot(x='Mar Cap - Crore', y='Sales Qtr - Crore',
data=data_frame)
    plt.title('Market Cap vs. Quarterly Sales')
    plt.show()

def create_correlation_matrix(data_frame):
    """Creating a correlation matrix plot."""
    correlation_matrix = data_frame.corr()
    sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
    plt.title('Correlation Matrix')
    plt.show()

def top_market_cap_companies(data_frame, n=10):
    """top N companies based on market capitalization."""
    top_companies = data_frame.nlargest(n, 'Mar Cap - Crore')
    print(f"Top {n} Companies by Market Cap:")
    print(top_companies)
```

```
def calculate_correlation(data_frame):
    """correlation between market capitalization and quarterly
sales."""
    correlation_market_cap_sales = data_frame['Mar Cap -
Crore'].corr(data_frame['Sales Qtr - Crore'])
    print(f"Correlation between Market Cap and Sales:
{correlation_market_cap_sales}")

if __name__ == "__main__":
    # Load data
    file_path = 'Financial Analytics data.csv'
    old_df = load_data(file_path)

    # Display basic information about the original dataset
    print("Original DataFrame:")
    print(old_df.head())

    # Handle missing values and continue with the cleaned DataFrame
    df = handle_missing_values(old_df)

    # Calculate mean metrics
    calculate_mean_metrics(df)

    # Create scatter plot
    create_scatter_plot(df)

    # Create correlation matrix plot
    create_correlation_matrix(df)

    # Display top companies based on market capitalization
    top_market_cap_companies(df)

    # Display correlation between market capitalization and quarterly
sales
    calculate_correlation(df)
```

Analysis Approach

1. **Data Loading and Overview**
   The analysis begins by loading the dataset from the **'Financial Analytics data.csv'** file. The initial exploration provides an overview of the dataset, allowing us to understand its structure and content.

```
# Load data
file_path = 'Financial Analytics data.csv'
old_df = load_data(file_path)

# Display basic information about the original dataset
print("Original DataFrame:")
print(old_df.head())
```

2. **Handling Missing Values**
   To ensure the quality of the analysis, missing values are handled by dropping rows with any missing data. This step ensures that the subsequent analysis is conducted on a clean dataset.

```
# Handle missing values and continue with the cleaned DataFrame
df = handle_missing_values(old_df)
```

3. **Mean Metrics Calculation**
   Key metrics, such as the mean market capitalization and mean quarterly sales, are calculated to provide a central tendency measure for the dataset.

```
# Calculate mean metrics
calculate_mean_metrics(df)
```

4. **Scatter Plot: Market Cap vs. Quarterly Sales**
   A scatter plot is created to visualize the relationship between market capitalization and quarterly sales. This plot helps identify trends and patterns in the data.

```
# Create scatter plot
create_scatter_plot(df)
```

5. **Correlation Matrix**
   A correlation matrix plot is generated to showcase the relationships between different attributes, particularly focusing on the correlation between market capitalization and quarterly sales.

```
# Create correlation matrix plot
create_correlation_matrix(df)
```

6. **Top Companies by Market Capitalization**
   The analysis includes the identification and display of the top N companies based on market capitalization. This information is valuable for understanding the market leaders.

```
    # Display top companies based on market capitalization
    top_market_cap_companies(df)
```

7. **Correlation Analysis**
   The correlation between market capitalization and quarterly sales is quantified and presented, providing insights into the strength and direction of this relationship.

```
    # Display correlation between market capitalization and quarterly
sales
    calculate_correlation(df)
```

Conclusion

This competition analysis offers valuable insights into the top 500 companies in India, focusing on market capitalization and quarterly sales. The calculated metrics, visualizations, and correlation analysis provide a comprehensive understanding of the dataset. Further exploration and in-depth analyses can be conducted based on these initial findings to support management decisions and enhance business results.