

**A PROJECT REPORT ON**  
**“A STUDY OF STOCK MARKET IN INDIA”**  
**(CASE STUDY OF COMMODITY CHEMICALS SUB SECTOR)**

**SUBMITTED TO**

**ALL INDIA MANAGEMENT ASSOCIATION - CENTRE FOR MANAGEMENT EDUCATION  
15 LINK ROAD, LAJPAT NAGAR - III, NEW DELHI 110024**

**BY**

**MR. SOHAIL ABDULREHMAN SHAIKH  
ENROLLMENT NO - P552220007**

**GUIDED BY**

**MR. ASHISH KAPASIYA  
POST GRADUATE DEGREE IN DATA SCIENCE**

**FOR THE PARTIAL FULFILMENT OF  
POST GRADUATE CERTIFICATE IN MANAGEMENT WITH SPECIALIZATION IN DATA  
ANALYTICS**

## **LIST OF CONTENTS**

<b>Content</b>	<b>Page No.</b>
List of Tables	2
List of Figures	3
Acknowledgement	4
Preface	5
Introduction	6
About Commodity Chemicals Sub Sector & Statement of Problem	7
Objective of the Study	8
Research Methodology, Data Collection Method	9 - 10
Questionnaires, Number of Respondents, Time Frame of the Data Collection & Method of Sampling	11 - 12
Statistical Tools	13
Presentation & Analysis	13 - 30
Conclusion	31 - 32
Recommendation, Limitation, Area of Study & Location of Study	33
Appendices	34
Bibliography & Book References	35

## **LIST OF FIGURES**

## **PRESENTATION & ANALYSIS**

<b>Figure</b>	<b>Page No.</b>
Microsoft Excel	13
MySQL	13 - 22
Python	23 - 26
Data Visualization Through Tableau	27 - 30

## **ACKNOWLEDGEMENT**

I have a great pleasure in representing this project report entitled

**"A STUDY OF STOCK MARKET IN INDIA"**

**(CASE STUDY OF COMMODITY CHEMICALS SUB SECTOR)**

and I grab this opportunity to convey my immense regards towards all the distinguished people who have their valuable contribution in the hour of need.

I take this opportunity to thank Mr. Ashish Kapasiya, Assistant Manager (Product and Student Outcome), for giving me an opportunity to complete this project and the most needed guidance throughout the duration of the Programme.

I am extremely grateful to my project guide Miss. Rachana Bajaj for her valuable guidance and necessary support during each phase of the project. She was the source of continuous encouragement as each milestone was crossed.

I would like to express my sincere gratitude to the All India Management Association (AIMA) for entrusting me with the opportunity to undertake this project as an integral component of the academic prerequisites for the Post Graduate Certificate in Management with Specialization in Data Analytics.

SOHAIL ABDULREHMAN SHAIKH

## **PREFACE**

The Indian stock market, a dynamic and ever-evolving financial ecosystem, plays a pivotal role in the nation's economic growth. Its performance is influenced by numerous factors, including the performance of various industry sectors. The Commodity Chemicals subsector within the stock market holds particular significance due to its integral role in the Indian economy.

This project, "A Study of Stock Market in India related to case study of Commodity Chemicals Subsector," is the result of meticulous research and analysis aimed at comprehending the nuances and intricacies of this vital segment of the Indian financial landscape. The primary objective of this study is to provide a comprehensive overview of the Commodity Chemicals subsector and its impact on the broader stock market in India.

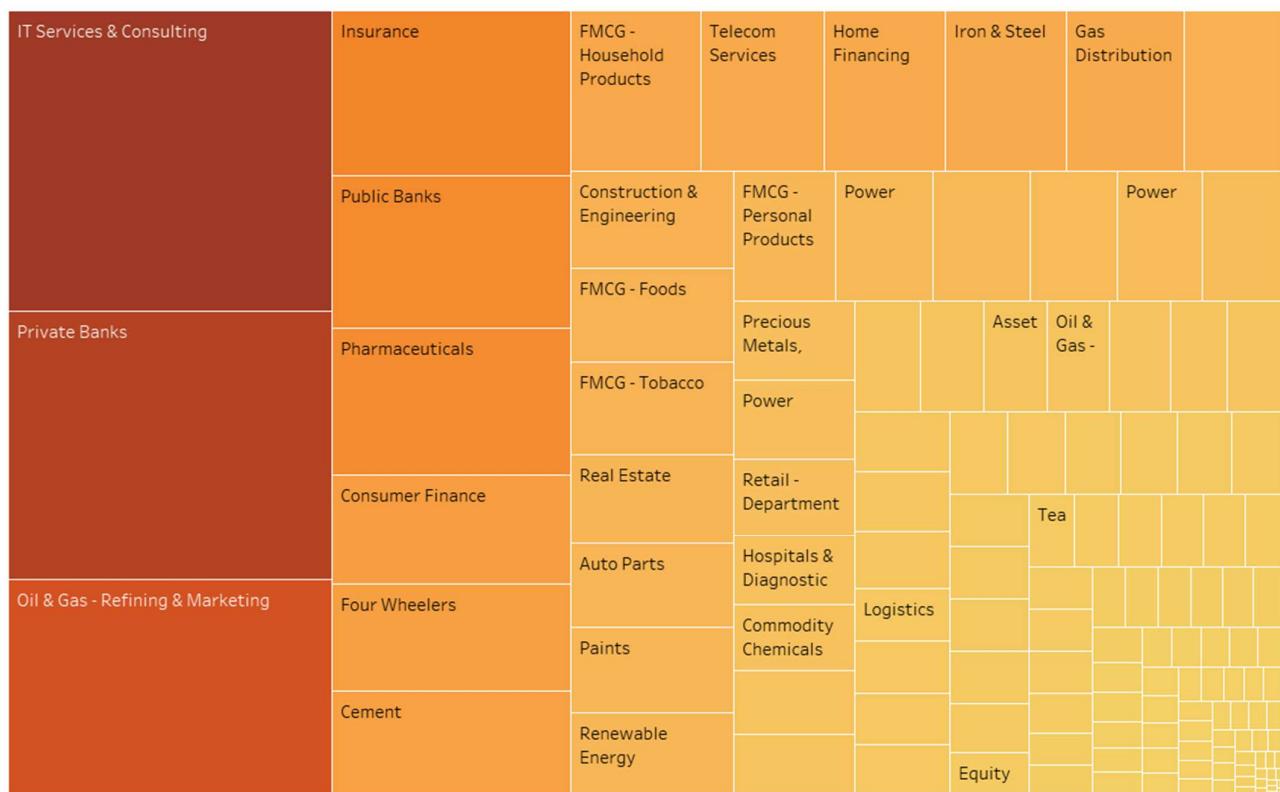
The motivation behind undertaking this project stems from the growing importance of the Commodity Chemicals Subsector within India's industrial and economic landscape. Commodity Chemicals form the foundation of various industries, from agriculture to manufacturing, and their performance in the stock market can be indicative of broader economic trends.

Furthermore, investors, analysts, policymakers, and industry stakeholders require in-depth insights into the Commodity Chemicals subsector to make informed decisions. This project seeks to bridge the gap between the financial markets and this critical sector, empowering stakeholders with valuable information

## **INTRODUCTION**

The stock market in India is one of the largest and most active in the world. It is a vital part of the Indian economy, providing a forum for companies to raise capital and for investors to participate in the growth of the economy.

The Indian stock market is divided into a number of different sectors, each of which is a home to a variety of companies. Some of the major sectors in the Indian stock market are IT Services & Consulting, Private Banks, Oil & Gas – Refining & Marketing, Insurance, Public Banks, Pharmaceuticals, Consumer Finance, Four Wheelers, Cement, FMCG – Household Products, Telecom Services, Home Financing, Iron & Steel, Gas Distribution, Construction & Engineering, FMCG – Foods & Tobacco, Real Estate, Auto Parts, Paints, Renewable Energy, FMCG – Personal Products, Precious Metals, Retail – Department, Hospitals & Diagnostic & Commodity Chemicals etc.



This case study serves as a compelling illustration of synergizing MySQL, Python & Tableau to refine stock selection within a vast dataset. With over 4500+ stocks, the objective is to suggest 5 optimal stocks within the Commodity Chemicals Sub Sector.

## **ABOUT COMMODITY CHEMICALS SUB SECTOR**

The commodity chemicals subsector in India is a significant contributor to the country's manufacturing and industrial sectors which gives a tremendous boost to Indian economy. India's diversified chemicals industry covers over **80,000 commercial products** and it is ranked **sixth** in the world in terms of sales and imports of chemicals. It accounts for 12.5% of India's total manufacturing output and 10.5% of its exports. The subsector is also playing a role of major employer which provides jobs to over 4 million people. It encompasses the production and distribution of chemicals used in various industries such as plastics, textiles, pharmaceuticals & are broadly classified into bulk chemicals, specialty chemicals, petrochemicals, agrochemicals, polymers and fertilizers. The Indian chemicals industry stood at **\$180 billion in 2020** and is projected to reach **\$304 billion by 2025**.

## **STATEMENT OF PROBLEM**

This study will help to understand the stock market in India". The fund house has some logic to pick stocks which are described on the subsequent slides. This case study is initiated in order to leverage stock market data so that it will gain insights into the specific challenges and opportunities within the Commodity Chemicals Sub Sector. The findings will enable investors, analysts and policymakers to make informed decisions, optimize investment strategies and better navigate the complexities of the stock market within the Commodity Chemicals Industry. Therefore, the primary objective of this study is to analyze historical stock market data, identify key factors that influence stock prices in the Indian market and develop predictive models that can assist in predicting stock market trends and helps in making informed investment decisions.

## **OBJECTIVES OF THE STUDY**

The objective of “STUDY OF STOCK MARKET IN INDIA” in Commodity Chemicals Sub-sector are as follows: -

1. To analyse historical stock market data in India, including price movements, trading volumes and other relevant financial indicators since this analysis will provide insights into past market trends and patterns.
2. To identify and analyse the key factors that influence stock prices in the Indian stock market. Understanding these factors is crucial for predicting and explaining stock market movements.
3. To develop predictive models using data analytics techniques & hence these models will leverage historical data and identified factors to forecast future stock price trends and market movements. The objective is to create accurate and reliable models that can assist in making informed investment decisions.
4. To evaluate the performance of developed predictive models and hence this involves comparing the model's predictions with actual stock market data to assess its accuracy and effectiveness in forecasting stock price movements.
5. To Provide Actionable Insights to investors, traders and financial institutions operating in the Indian stock market. These insights will assist stakeholders in making informed decisions, optimizing investment strategies and managing risks effectively.
6. To enhance decision-making capabilities in the Indian stock market by leveraging data analytics techniques and providing valuable insights, the project aims to empower investors and financial institutions to make more informed, data-driven decisions, thereby maximizing their investment returns and minimizing risks.

## **RESEARCH METHODOLOGY**

The research methodology for a data analytics project related to the study of the stock market in India and the commodity chemicals sub-sector could include the following steps:

1. **Data Collection:** The study will involve collecting historical stock market data from the Indian market including historical stock prices, trading volumes, news sentiment, financial statements and other relevant financial indicators. The data will be sourced from reputable financial databases, stock exchanges, news websites and regulatory bodies etc.
2. **Data Preprocessing:** The next step is to clean the data so the collected data will undergo preprocessing steps to ensure its quality and suitability for analysis. This may involve data cleaning, handling missing values, standardization and normalization. The preprocessing step aims to ensure removing errors and inconsistencies in the data so that the data is transformed into a format which is ready for analysis.
3. **Exploratory Data Analysis:** The third step is to analyze the data. The collected and pre-processed data will be subjected to exploratory data analysis (EDA). It involves using statistical and machine learning techniques to identify trends and patterns in the data which includes Feature Selection, Model Development & Model Evaluation. EDA helps in uncovering initial trends and relationships within the stock market data which involves in developing a model that can be used to predict future prices of stocks.
4. **Data visualization:** The next step is to visualize the data. This involves creating charts and graphs that can be used to communicate the findings of the analysis. The visualizations can be used to help investors to make better investment decisions.

5. Interpretation and Insight Generation: The findings from the analysis and predictive models will be interpreted and transformed into actionable insights. The objective is to provide meaningful and practical recommendations based on the data analysis to assist stakeholders in making informed investment decisions and optimizing their strategies.
6. Documentation and Report Writing: Finally, the results, insights and recommendations will be documented in a comprehensive report. The report will include the research methodology, data analysis techniques employed, findings, limitations and conclusions drawn from the study. The report will be structured and written in a clear and concise manner ensuring that it effectively communicates the research findings to the intended audience.

## **DATA COLLECTION METHOD**

The data will be collected using secondary sources viz Primary Data & Secondary Data.

**PRIMARY DATA** - All of the information will be gathered through secondary sources. The methods that will be used to collect primary data from the website.

*Web Sites:*

<https://www.nseindia.com/>  
<https://www.bseindia.com/>  
<https://www.moneycontrol.com/>  
<https://in.investing.com/>  
<https://www.cmie.com/>

**SECONDARY DATA** - Interviews, Questionnaires from different nationalized and private companies of Commodity Chemicals Sub-sector etc. Secondary data of this project on the Stock Market will be collected from

- Interviews from different categories of target groups.
- Questionnaires will be arranged to get the data.

## **QUESTIONNAIRES**

*Dear Respondents,*

I am a student of PGCM (Data Analytics). I am underlying a project named "**A STUDY OF STOCK MARKET IN INDIA (CASE STUDY OF COMMODITY CHEMICALS IN STOCK MARKET)**". So, by filling this questionnaire please help me in completing my research project.

**Name** : .....

**Age** : .....

**Address** : .....

**Gender** : .....

**Contact No.** : .....

Q.1) Are you familiar with the Commodity Chemicals Sub-sector in India?

- a. YES
- b. NO

Q.2) Have you invested in companies within the Commodity Chemicals Sub-sector?

- a. YES
- b. NO

Q.3) Do you believe that global demand for commodity chemicals significantly impacts stock prices in India?

- a. YES
- b. NO

Q.4) Are you aware of the regulatory policies that affect the commodity chemicals subsector in India?

- a. YES
- b. NO

Q.5) Do you consider raw material prices as a crucial factor influencing stock prices in the commodity chemicals subsector?

- a. YES
- b. NO

Q.6) Are you interested in utilizing data analytics techniques to analyze historical stock market data for the Commodity Chemicals Sub-sector?

- a. YES
- b. NO

Q.7) Do you believe that historical stock market data can help predict future stock price movements within the commodity chemicals subsector?

- a. YES
- b. NO

Q.8) Are you aware of the specific market indicators that are used to assess the performance of companies in the Commodity Chemicals Sub-sector?

- a. YES
- b. NO

Q.9) Do you think that technological advancements play a significant role in shaping the stock market trends for the commodity chemicals subsector?

- a. YES
- b. NO

Q.10) Would you be interested in receiving actionable insights and recommendations based on data analytics for investing in the Commodity Chemicals Sub-sector?

- a. YES
- b. NO

### **NUMBER OF RESPONDENTS**

A total sample of 40-50 respondents were contacted who responded to the questionnaires.

### **TIME FRAME OF THE DATA COLLECTION - 15 days**

### **METHOD OF SAMPLING**

The technique used for conducting the study was Stratified Random Sampling technique as a sample of respondents was chosen random.

## **STATISTICAL TOOLS**

- The tools used in this study were Microsoft Excel, MySQL, Python, Tableau/Power BI & Microsoft Word.
- Microsoft Excel is been used to collect the data from the website as well as to add “NULL” at missing values.
- MySQL is been used for removing duplicates & cleaning purposes.
- Python is used to extract the information and get an insight into the data which is used for choosing stocks.
- Tableau is used for selecting the company in which we invest our money based on the dashboard.
- Microsoft Word was used to prepare or write the whole project report.

## **PRESENTATION AND ANALYSIS**

### **1. MICROSOFT EXCEL**

- Since the data collected is in Excel format so we have opened the file and replaced all the missing values with “NULL” so that the file would be easily imported to MySQL

### **2. MySQL**

- Create the database in the name of ‘stock\_data’ & use it.

```
1 • CREATE DATABASE stock_data;
2 • USE stock_data;
3
```

Output			
Action Output	#	Time	Action
2 15:53:00 CREATE DATABASE stock_data			Message 1 row(s) affected
3 15:53:00 USE stock_data			Duration / Fetch 0.000 sec 0 row(s) affected

➤ Create Table as 'stocks' and add all the column names with proper data types.

```

4 • CREATE TABLE stocks
5   (
6     Name VARCHAR(255),
7     Ticker VARCHAR(50),
8     Sub_Sector VARCHAR(255),
9     Market_Cap DOUBLE,
10    Close_Price DOUBLE,
11    PE_Ratio DOUBLE,
12    5Y_Historical_Revenue_Growth DOUBLE,
13    1Y_Forward_Revenue_Growth DOUBLE,
14    Total_Revenue DOUBLE,
15    PBT DOUBLE,
16    Taxes_And_Other_Items DOUBLE,
17    DII_Holding_Change_3M DOUBLE,
18    FII_Holding_Change_3M DOUBLE,
19    MF_Holding_Change_3M DOUBLE,
20    Promoter_Holding_Change_3M DOUBLE,
21    Percentage_Buy_Reco DOUBLE,
22    No_of_analysts_with_buy_reco INT,
23    Percentage_Upside DOUBLE
24  );
25
26 • SELECT * from stocks;

```

Result Grid											Filter Rows:	Export:	Wrap Cell Content:	Fetch Rows:	Result Grid
Name	Ticker	Sub_Sector	Market_Cap	Close_Price	PE_Ratio	5Y_Historical_Revenue_Growth	1Y_Forward_Revenue_Growth	Total_Revenue	PBT	Taxes_And_Other_Items	DII_Holding_Change_3M				
<b>stocks 1 x</b>															
Output															

- **Removing Duplicates** - Since the 'Name' column is unique so it shouldn't be having any repeated values so check whether if it has any repeated values and do remove them.

```

27
28     # REMOVING DUPLICATES
29 •   DELETE t1
30     FROM stocks t1
31     JOIN stocks t2
32     WHERE t1.Ticker = t2.Ticker AND t1.Name < t2.Name;

```

Output			
	Action	Message	Duration / Fetch
9	SHOW COLUMNS FROM 'stock_data'.stocks'	OK	0.000 sec
10	PREPARE stmt FROM INSERT INTO 'stock_data'.stocks ('Name','Ticker','Sub_Sector','Market_Cap','Close_P...	OK	0.000 sec
11	DEALLOCATE PREPARE stmt	OK	0.000 sec
12	SELECT * from stocks LIMIT 0, 1000	1000 row(s) returned	0.000 sec / 0.000 sec
13	DELETE t1 FROM stocks t1 JOIN stocks t2 WHERE t1.Ticker = t2.Ticker AND t1.Name < t2.Name	8 row(s) affected	15.219 sec

- Now let's check the count of Null values in percentage

```

37     # TO FIND NULL VALUES PRESENT IN EVERY COLUMN
38 •   select (select count(*) from stocks where Name is not NULL)/(select count(*) from stocks) * 100 Name_no_null_percent;
39 •   select (select count(*) from stocks where Ticker is not NULL)/(select count(*) from stocks) *100 Ticker_no_null_percent;
40 •   select (select count(*) from stocks where Sub_Sector is not NULL)/(select count(*) from stocks) *100 Sub_Sector_no_null_percent;
41 •   select (select count(*) from stocks where Market_Cap is not NULL)/(select count(*) from stocks) * 100 Market_Cap_no_null_percent;
42 •   select (select count(*) from stocks where Close_Price is not NULL)/(select count(*) from stocks) *100 Close_Price_no_null_percent;
43 •   select (select count(*) from stocks where PE_Ratio is not NULL)/(select count(*) from stocks) * 100 PE_Ratio_no_null_percent;
44 •   select (select count(*) from stocks where 5Y_Historical_Revenue_Growth is not NULL)/(select count(*) from stocks) * 100 5Y_Historical_Revenue_Growth_no_n...
45 •   select (select count(*) from stocks where 1Y_Forward_Revenue_Growth is not NULL)/(select count(*) from stocks) *100 1Y_Forward_Revenue_Growth_no_n...
46 •   select (select count(*) from stocks where Total_Revenue is not NULL)/(select count(*) from stocks) * 100 Total_Revenue_no_null_percent;
47 •   select (select count(*) from stocks where PBT is not NULL)/(select count(*) from stocks) * 100 PBT_no_null_percent;
48 •   select (select count(*) from stocks where Taxes_And_Other_Items is not NULL)/(select count(*) from stocks) * 100 Taxes_And_Other_Items_no_null_percent;
49 •   select (select count(*) from stocks where DII_Holding_Change_3M is not NULL)/(select count(*) from stocks) * 100 DII_Holding_Change_3M_no_null_percent;
50 •   select (select count(*) from stocks where FII_Holding_Change_3M is not NULL)/(select count(*) from stocks) * 100 FII_Holding_Change_3M_no_null_percent;
51 •   select (select count(*) from stocks where MF_Holding_Change_3M is not NULL)/(select count(*) from stocks) * 100 MF_Holding_Change_3M_no_null_percent;
52 •   select (select count(*) from stocks where Promoter_Holding_Change_3M is not NULL)/(select count(*) from stocks) * 100 Promoter_Holding_Change_3M_no...
53 •   select (select count(*) from stocks where Percentage_Buy_Reco is not NULL)/(select count(*) from stocks) * 100 Percentage_Buy_Reco_no_null_percent;
54 •   select (select count(*) from stocks where No_of_analysts_with_buy_reco is not NULL)/(select count(*) from stocks) * 100 No_of_analysts_with_buy_re...
55 •   select (select count(*) from stocks where Percentage_Upside is not NULL)/(select count(*) from stocks) * 100 Percentage_Upside_no_null_percent;

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	Name_no_null_percent			
▶	100.0000			

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	Ticker_no_null_percent			
▶	100.0000			

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	Sub_Sector_no_null_percent			

► 87.0325

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	Market_Cap_no_null_percent			

► 97.1454

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	PE_Ratio_no_null_percent			

► 95.8398

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	5Y_Historical_Revenue_Growth_no_null_percent			

► 87.2981

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	1Y_Forward_Revenue_Growth_no_null_percent			

► 14.9591

<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	Total_Revenue_no_null_percent							
▶	95.9062							
<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	PBT_no_null_percent							
▶	96.0611							
<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	Taxes_And_Other_Items_no_null_percent							
▶	96.0611							
<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	DII_Holding_Change_3M_no_null_percent							
▶	88.1832							
<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	FII_Holding_Change_3M_no_null_percent							
▶	88.1832							
<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	MF_Holding_Change_3M_no_null_percent							
▶	88.1832							
<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	Promoter_Holding_Change_3M_no_null_percent							
▶	88.1832							
<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	Percentage_Buy_Reco_no_null_percent							
▶	16.2868							
<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	No_of_analysts_with_buy_reco_no_null_percent							
▶	98.8493							
<b>Result Grid</b>			Filter Rows:	<input type="text"/>	Export:		Wrap Cell Content:	
	Percentage_Upside_no_null_percent							
▶	16.3753							

- **Filling Missing Values** - Now we need to fill the missing numerical values by using median in every column and ignore updating in columns like Name, Ticker & Sub\_Sector since they are categorical.

```

57      # FILLING NULL VALUES WITH MEDIAN FOR EVERY NUMERICAL COLUMN
58
59      # Market_Cap
60 • Ⓜ update stocks set Market_Cap = (SELECT AVG(dd.Market_Cap) as median_val
61      FROM
62      (
63          SELECT d.Market_Cap, @rownum:=@rownum+1 as `row_number`, @total_rows:=@rownum
64          FROM stocks d, (SELECT @rownum:=0) r
65          WHERE d.Market_Cap is NOT NULL
66          ORDER BY d.Market_Cap
67      ) as dd
68      WHERE dd.row_number IN ( FLOOR(@total_rows+1)/2), FLOOR(@total_rows+2)/2 ) where Market_Cap is NULL;
69
70      # Close_Price
71 • Ⓜ update stocks set Close_Price = (SELECT AVG(dd.Close_Price) as median_val
72      FROM
73      (
74          SELECT d.Close_Price, @rownum:=@rownum+1 as `row_number`, @total_rows:=@rownum
75          FROM stocks d, (SELECT @rownum:=0) r
76          WHERE d.Close_Price is NOT NULL
77          ORDER BY d.Close_Price
78      ) as dd
79      WHERE dd.row_number IN ( FLOOR(@total_rows+1)/2), FLOOR(@total_rows+2)/2 ) where Close_Price is NULL;
80
81      # PE_Ratio
82 • Ⓜ update stocks set PE_Ratio = (SELECT AVG(dd.PE_Ratio) as median_val
83      FROM
84      (
85          SELECT d.PE_Ratio, @rownum:=@rownum+1 as `row_number`, @total_rows:=@rownum
86          FROM stocks d, (SELECT @rownum:=0) r
87          WHERE d.PE_Ratio is NOT NULL
88          ORDER BY d.PE_Ratio
89      ) as dd
90      WHERE dd.row_number IN ( FLOOR(@total_rows+1)/2), FLOOR(@total_rows+2)/2 ) where PE_Ratio is NULL;
91

```

```

92      # 5Y_Historical_Revenue_Growth
93  update stocks set 5Y_Historical_Revenue_Growth = (SELECT AVG(dd.5Y_Historical_Revenue_Growth) as median_val
94  FROM
95  (
96      SELECT d.5Y_Historical_Revenue_Growth, @rownum:="@rownum+1" as `row_number` , @total_rows:="@rownum"
97      FROM stocks d, (SELECT @rownum:=0) r
98      WHERE d.5Y_Historical_Revenue_Growth is NOT NULL
99      ORDER BY d.5Y_Historical_Revenue_Growth
100     ) as dd
101    WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where 5Y_Historical_Revenue_Growth is NULL;
102
103      # 1Y_Forward_Revenue_Growth
104  update stocks set 1Y_Forward_Revenue_Growth = (SELECT AVG(dd.1Y_Forward_Revenue_Growth) as median_val
105  FROM
106  (
107      SELECT d.1Y_Forward_Revenue_Growth, @rownum:="@rownum+1" as `row_number` , @total_rows:="@rownum"
108      FROM stocks d, (SELECT @rownum:=0) r
109      WHERE d.1Y_Forward_Revenue_Growth is NOT NULL
110      ORDER BY d.1Y_Forward_Revenue_Growth
111     ) as dd
112    WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where 1Y_Forward_Revenue_Growth is NULL;
113
114      # Total_Revenue
115  update stocks set Total_Revenue = (SELECT AVG(dd.Total_Revenue) as median_val
116  FROM
117  (
118      SELECT d.Total_Revenue, @rownum:="@rownum+1" as `row_number` , @total_rows:="@rownum"
119      FROM stocks d, (SELECT @rownum:=0) r
120      WHERE d.Total_Revenue is NOT NULL
121      ORDER BY d.Total_Revenue
122     ) as dd
123    WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where Total_Revenue is NULL;
124
125      # PBT
126  update stocks set PBT = (SELECT AVG(dd.PBT) as median_val
127  FROM
128  (
129      SELECT d.PBT, @rownum:="@rownum+1" as `row_number` , @total_rows:="@rownum"
130      FROM stocks d, (SELECT @rownum:=0) r
131      WHERE d.PBT is NOT NULL
132      ORDER BY d.PBT
133     ) as dd
134    WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where PBT is NULL;
135
136      # Taxes_And_Other_Items
137  update stocks set Taxes_And_Other_Items = (SELECT AVG(dd.Taxes_And_Other_Items) as median_val
138  FROM
139  (
140      SELECT d.Taxes_And_Other_Items, @rownum:="@rownum+1" as `row_number` , @total_rows:="@rownum"
141      FROM stocks d, (SELECT @rownum:=0) r
142      WHERE d.Taxes_And_Other_Items is NOT NULL
143      ORDER BY d.Taxes_And_Other_Items
144     ) as dd
145    WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where Taxes_And_Other_Items is NULL;

```

```

147      # DII_Holding_Change_3M
148  • update stocks set DII_Holding_Change_3M = (SELECT AVG(dd.DII_Holding_Change_3M) as median_val
149    FROM
150  • (
151    SELECT d.DII_Holding_Change_3M, @rownum:=@rownum+1 as `row_number`, @total_rows:=@rownum
152      FROM stocks d, (SELECT @rownum:=0) r
153      WHERE d.DII_Holding_Change_3M is NOT NULL
154      ORDER BY d.DII_Holding_Change_3M
155  ) as dd
156  ~ WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where DII_Holding_Change_3M is NULL;
157
158      # FII_Holding_Change_3M
159  • update stocks set FII_Holding_Change_3M = (SELECT AVG(dd.FII_Holding_Change_3M) as median_val
160    FROM
161  • (
162    SELECT d.FII_Holding_Change_3M, @rownum:=@rownum+1 as `row_number`, @total_rows:=@rownum
163      FROM stocks d, (SELECT @rownum:=0) r
164      WHERE d.FII_Holding_Change_3M is NOT NULL
165      ORDER BY d.FII_Holding_Change_3M
166  ) as dd
167  ~ WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where FII_Holding_Change_3M is NULL;
168
169      # MF_Holding_Change_3M
170  • update stocks set MF_Holding_Change_3M = (SELECT AVG(dd.MF_Holding_Change_3M) as median_val
171    FROM
172  • (
173    SELECT d.MF_Holding_Change_3M, @rownum:=@rownum+1 as `row_number`, @total_rows:=@rownum
174      FROM stocks d, (SELECT @rownum:=0) r
175      WHERE d.MF_Holding_Change_3M is NOT NULL
176      ORDER BY d.MF_Holding_Change_3M
177  ) as dd
178  ~ WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where MF_Holding_Change_3M is NULL;
179
180      # Promoter_Holding_Change_3M
181  • update stocks set Promoter_Holding_Change_3M = (SELECT AVG(dd.Promoter_Holding_Change_3M) as median_val
182    FROM
183  • (
184    SELECT d.Promoter_Holding_Change_3M, @rownum:=@rownum+1 as `row_number`, @total_rows:=@rownum
185      FROM stocks d, (SELECT @rownum:=0) r
186      WHERE d.Promoter_Holding_Change_3M is NOT NULL
187      ORDER BY d.Promoter_Holding_Change_3M
188  ) as dd
189  ~ WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where Promoter_Holding_Change_3M is NULL;
190
191      # Percentage_Buy_Reco
192  • update stocks set Percentage_Buy_Reco = (SELECT AVG(dd.Percentage_Buy_Reco) as median_val
193    FROM
194  • (
195    SELECT d.Percentage_Buy_Reco, @rownum:=@rownum+1 as `row_number`, @total_rows:=@rownum
196      FROM stocks d, (SELECT @rownum:=0) r
197      WHERE d.Percentage_Buy_Reco is NOT NULL
198      ORDER BY d.Percentage_Buy_Reco
199  ) as dd
200  ~ WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where Percentage_Buy_Reco is NULL;
201
202      # No_of_analysts_with_buy_reco
203  • update stocks set No_of_analysts_with_buy_reco = (SELECT AVG(dd.No_of_analysts_with_buy_reco) as median_val
204    FROM
205  • (
206    SELECT d.No_of_analysts_with_buy_reco, @rownum:=@rownum+1 as `row_number`, @total_rows:=@rownum
207      FROM stocks d, (SELECT @rownum:=0) r
208      WHERE d.No_of_analysts_with_buy_reco is NOT NULL
209      ORDER BY d.No_of_analysts_with_buy_reco
210  ) as dd
211  ~ WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where No_of_analysts_with_buy_reco is NULL;

```

```

213      # Percentage_Upside
214  • update stocks set Percentage_Upside = (SELECT AVG(dd.Percentage_Upside) as median_val
215    FROM
216    (
217      SELECT d.Percentage_Upside, @rownum:=@rownum+1 as `row_number`, @total_rows:="@rownum
218        FROM stocks d, (SELECT @rownum:=0) r
219       WHERE d.Percentage_Upside is NOT NULL
220      ORDER BY d.Percentage_Upside
221    ) as dd
222   WHERE dd.row_number IN ( FLOOR((@total_rows+1)/2), FLOOR((@total_rows+2)/2) ) where Percentage_Upside is NULL;

```

Output					
#	Time	Action	Message	Duration / Fetch	
33	17:02:15	update stocks set Market_Cap = (SELECT AVG(dd.Market_Cap) as median_val FROM ( SELECT d.Market_Cap, ... 129 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.156 sec			
34	17:52:26	update stocks set Market_Cap = (SELECT AVG(dd.Market_Cap) as median_val FROM ( SELECT d.Market_Cap, ... 0 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.016 sec			
35	17:52:26	update stocks set Close_Price = (SELECT AVG(dd.Close_Price) as median_val FROM ( SELECT d.Close_Price, ... 79 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.062 sec			
36	17:52:26	update stocks set PE_Ratio = (SELECT AVG(dd.PE_Ratio) as median_val FROM ( SELECT d.PE_Ratio, @rown... 188 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.109 sec			
37	17:52:27	update stocks set 5Y_Historical_Revenue_Growth = (SELECT AVG(dd.5Y_Historical_Revenue_Growth) as median... 574 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.297 sec			
38	17:52:27	update stocks set 1Y_Forward_Revenue_Growth = (SELECT AVG(dd.1Y_Forward_Revenue_Growth) as median... 3843 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.422 sec			
39	17:52:27	update stocks set Total_Revenue = (SELECT AVG(dd.Total_Revenue) as median_val FROM ( SELECT d.Total... 185 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.125 sec			
40	17:52:27	update stocks set PBT = (SELECT AVG(dd.PBT) as median_val FROM ( SELECT d.PBT, @rownum:=-@rownum... 178 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.110 sec			
41	17:52:27	update stocks set Taxes_And_Other_Items = (SELECT AVG(dd.Taxes_And_Other_Items) as median_val FROM ... 178 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.109 sec			
42	17:52:28	update stocks set DIL_Holding_Change_3M = (SELECT AVG(dd.DIL_Holding_Change_3M) as median_val FROM ... 534 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.266 sec			
43	17:52:28	update stocks set FI_Holding_Change_3M = (SELECT AVG(dd.FI_Holding_Change_3M) as median_val FROM ... 534 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.281 sec			
44	17:52:28	update stocks set MF_Holding_Change_3M = (SELECT AVG(dd.MF_Holding_Change_3M) as median_val FROM ... 534 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.281 sec			
45	17:52:28	update stocks set Promoter_Holding_Change_3M = (SELECT AVG(dd.Promoter_Holding_Change_3M) as median... 534 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.282 sec			
46	17:52:29	update stocks set Percentage_Buy_Reco = (SELECT AVG(dd.Percentage_Buy_Reco) as median_val FROM ( S... 3783 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.421 sec			
47	17:52:29	update stocks set No_of_analysts_with_buy_reco = (SELECT AVG(dd.No_of_analysts_with_buy_reco) as median... 52 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.047 sec			
48	17:52:29	update stocks set Percentage_Upside = (SELECT AVG(dd.Percentage_Upside) as median_val FROM ( SELECT... 3779 row(s) affected, 3 warning(s): 1287 Setting user variables within expressions is deprecated and will be removed... 0.407 sec			

➤ Now let's check whether we have filled all the missing values

Result Grid		Filter Rows:	Export:	Wrap Cell Content:	Fetch rows:	Result Grid	
Name	Ticker	Sub_Sector	Market_Cap	Close_Price	PE_Ratio	5Y_Historical_Revenue_Growth	1Y_Forward_Revenue_Gro
Reliance Industries Ltd	RELIANCE	Oil & Gas - Refining & Marke...	1778090.092	2712.2	29.29066951	17.9364979	30.53551161
Tata Consultancy Services Ltd	TCS	IT Services & Consulting	1241655.017	3397.35	32.39635288	9.886695907	13.29505752
HDFC Bank Ltd	HDFCBANK	Private Banks	891794.7543	1597.85	23.43575048	14.24219272	-27.60081708
Infosys Ltd	INFY	IT Services & Consulting	680667.9124	1629.05	30.78552295	11.61870846	18.12891331
ICICI Bank Ltd	ICICIBANK	Private Banks	653338.2571	946.75	26.0189325	6.796087786	-49.98364187
Hindustan Unilever Ltd	HINDUNILVR	FMCG - Household Products	592991.9281	2634.2	66.78589121	9.348279577	10.75163197
State Bank of India	SBIN	Public Banks	542852.6495	608.55	15.34614381	6.385751271	-32.3123937
Bharti Airtel Ltd	BHARTIARTL	Telecom Services	483495.4801	835.85	113.6326306	4.135738484	13.88752331
Housing Development Finance Corporation Ltd	HDFC	Home Financing	483449.4696	2665.25	21.3965967	18.86273518	-83.89215365
Adani Enterprises Ltd	ADANIENT	Commodities Trading	442774.7475	3878.4	570.1780629	13.52326532	1.797333028
ITC Ltd	ITC	FMCG - Tobacco	421780.4532	342.3	27.67105303	7.015923613	13.57888561
Bajaj Finance Ltd	BAJFINANCE	Consumer Finance	409206.0719	6734.15	58.22320441	25.92625384	-10.19079715
Life Insurance Corporation Of India	LICI	Insurance	398128.9064	636.8	96.52288437	7.266815742	11.71581862
Adani Total Gas Ltd	ATGL	Gas Distribution	396279.6026	3600	777.9340452	22.25434156	14.18749314
Kotak Mahindra Bank Ltd	KOTAKBANK	Private Banks	383187.9224	1925.55	31.69619025	11.68405012	-55.6115975
Adani Green Energy Ltd	ADANIGREEN	Renewable Energy	315159.6293	2012.05	644.4982194	57.51048975	29.71995746
Adani Transmission Ltd	ADANTRANS	Power Infrastructure	302817.6319	2782.7	251.3822998	34.05185011	-0.552776879
HCL Technologies Ltd	HCLTECH	IT Services & Consulting	302730.9039	1128.95	22.4261726	12.25812401	15.81332595
Asian Paints Ltd	ASIANPAINT	Paints	301983.2488	3135.05	99.64569332	10.85797049	20.4181838

- Since all the values seems to be filled but let's verify again the count of null in every column.

```

226 # TO CHECK WHETHER THE COUNT OF NULL VALUES IN EVERY COLUMNS AFTER FILLING WITH MEDIAN
227 • select count(*) from stocks where Name is NULL;
228 • select count(*) from stocks where Ticker is NULL;
229 • select count(*) from stocks where Sub_Sector is NULL;
230 • select count(*) from stocks where Market_Cap is NULL;
231 • select count(*) from stocks where Close_Price is NULL;
232 • select count(*) from stocks where PE_Ratio is NULL;
233 • select count(*) from stocks where 5Y_Historical_Revenue_Growth is NULL;
234 • select count(*) from stocks where Total_Revenue is NULL;
235 • select count(*) from stocks where PBT is NULL;
236 • select count(*) from stocks where Taxes_And_Other_Items is NULL;
237 • select count(*) from stocks where DII_Holding_Change_3M is NULL;
238 • select count(*) from stocks where FII_Holding_Change_3M is NULL;
239 • select count(*) from stocks where MF_Holding_Change_3M is NULL;
240 • select count(*) from stocks where Promoter_Holding_Change_3M is NULL;
241 • select count(*) from stocks where Percentage_Buy_Reco is NULL;
242 • select count(*) from stocks where No_of_analysts_with_buy_reco is NULL;
243 • select count(*) from stocks where Percentage_Upside is NULL;

```

- Since our original data count was **4527** so let's check

```

244
245 • SELECT * from stocks limit 4527;
246

```

Name	Ticker	Sub_Sector	Market_Cap	Close_Price	PE_Ratio	5Y_Historical_Revenue_Growth	1Y_Forward_Revenue_Gro
Reliance Industries Ltd	RELIANCE	Oil & Gas - Refining & Marke...	1778090.092	2712.2	29.2906951	17.9364979	30.53551161
Tata Consultancy Services Ltd	TCS	IT Services & Consulting	1241655.017	3397.35	32.39635288	9.886695907	13.29505752
HDFC Bank Ltd	HDFCBANK	Private Banks	891794.7543	1597.85	23.43575048	14.24219272	-27.60081708
Infoys Ltd	INFY	IT Services & Consulting	680667.9124	1629.05	30.78552295	11.61870846	18.12891331
ICICI Bank Ltd	ICICIBANK	Private Banks	653338.2571	946.75	26.0189325	6.796087786	-49.98364187
Hindustan Unilever Ltd	HINDUNILVR	FMCG - Household Products	592991.9281	2634.2	66.78589121	9.348279577	10.75163197
State Bank of India	SBIN	Public Banks	542852.6495	608.55	15.34614381	6.385751271	-32.3123937
Bharti Airtel Ltd	BHARTIARTL	Telecom Services	483495.801	835.85	113.6326306	4.135738484	13.88752331
Housing Development Finance Corporation Ltd	HDFC	Home Financing	483449.4696	2665.25	21.3965967	18.86273518	-83.89215365
Adani Enterprises Ltd	ADANIENT	Commodities Trading	442777.4765	3878.4	570.1780629	13.52326532	1.79733028
ITC Ltd	ITC	FMCG - Tobacco	421780.4532	342.3	27.67105303	7.015923613	13.57888561
Bajaj Finance Ltd	BAJFINANCE	Consumer Finance	409206.0719	6734.15	58.22320441	25.92625584	-10.19079715
Life Insurance Corporation Of India	LICI	Insurance	398128.9064	636.8	96.52288437	7.266815742	11.71581862
Adani Total Gas Ltd	ATGL	Gas Distribution	396279.6026	3600	777.9340452	22.25434156	14.18749314
Kotak Mahindra Bank Ltd	KOTAKBANK	Private Banks	383187.9224	1925.55	31.69619025	11.68405012	-55.6115975
Adani Green Energy Ltd	ADANIGREEN	Renewable Energy	315159.6293	2012.05	644.4982194	57.51048975	29.71995746
Adani Transmission Ltd	ADANITRANS	Power Infrastructure	302817.6319	2782.7	251.3822996	34.05185011	-0.552776879
HCL Technologies Ltd	HCLTECH	IT Services & Consulting	302730.9039	1128.95	22.4261726	12.25812401	15.81332595
Asian Paints Ltd	ASIANPAINT	Paints	301983.2488	3135.05	99.64569332	10.85797049	20.4181838
Larsen & Toubro Ltd	LT	Construction & Engineering	289970.5653	2050.55	33.44786337	7.425273362	13.74876609
Axist Bank Ltd	AXISBANK	Private Banks	273949.7378	892.55	19.40251513	8.376579249	-33.17701496
Maruti Suzuki India Ltd	MARUTI	Four Wheelers	271281.4867	8985.4	69.92692013	5.046152352	30.62189981
Bajaj Finserv Ltd	BAJAJFINSV	Insurance	262701.2313	1632	57.65075511	22.80049677	12.19410342
Avenue Supermarts Ltd	DMART	Retail - Department Stores	252991.467	3965.1	169.5028421	21.12500105	41.66282881

Now download the file by using **Export function** so that we can obtain insights from it through **Python**.

### 3. PYTHON

- We would be importing **Pandas (pd)** & **Numerical Python (NumPy)** since these libraries provide the flexibility of Data Structures, Data Cleaning, Data Selection and Filtering, Data Aggregation and Grouping, Efficient Data Storage, Array Operations, Multidimensional Arrays, Broadcasting, Random Number Generation & Linear Algebra etc. so it would be a great help in order to gain some insights.

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

- Create a Data frame naming 'Stocks'.

```
In [2]: 1 Stocks = pd.read_csv('Stock_Market_Dataset_GM100_SQL_Cleaned.csv')
```

- Let's do some basic operations like head, tail, shape, describe & info.

```
In [3]: 1 Stocks.head()
```

Out[3]:

		Name	Ticker	Sub_Sector	Market_Cap	Close_Price	PE_Ratio	5Y_Historical_Revenue_Growth	1Y_Forward_Revenue_Growth	Total_Revenue
0	Reliance Industries Ltd	RELIANCE		Oil & Gas - Refining & Marketing	1.778090e+06	2712.20	29.290670	17.936498	30.535512	718061.00 84%
1	Tata Consultancy Services Ltd	TCS		IT Services & Consulting	1.241655e+06	3397.35	32.396353	9.886696	13.295058	195772.00 51%
2	HDFC Bank Ltd	HDFCBANK		Private Banks	8.917948e+05	1597.85	23.435750	14.242193	-27.600817	167695.40 50%
3	Infosys Ltd	INFY		IT Services & Consulting	6.806679e+05	1629.05	30.785523	11.618708	18.128913	123936.00 30%
4	ICICI Bank Ltd	ICICIBANK		Private Banks	6.533383e+05	946.75	26.018932	6.796088	-49.983642	157536.32 34%

```
In [4]: 1 Stocks.tail()
```

Out[4]:

		Name	Ticker	Sub_Sector	Market_Cap	Close_Price	PE_Ratio	5Y_Historical_Revenue_Growth	1Y_Forward_Revenue_Growth	Total_Revenue
4514	Tapi Fruit Processing Ltd	TAPIFRUIT		Packaged Foods & Meats	107.848815	61.800	0.000000	7.266816	14.187493	15.22
4515	Isran International Ltd	ISHAN		Industrial Machinery	107.848815	45.000	0.000000	7.266816	14.187493	21.74
4516	Coastal Corporation Ltd Partly Paidup	COASTPP		Nan	107.848815	93.500	14.115072	7.266816	14.187493	99.92
4517	Anzen India Energy Yield Plus Trust	ANZEN		Power Transmission & Distribution	107.848815	64.635	14.115072	7.266816	14.187493	99.92
4518	Keystone Realtors Pvt Ltd	RUSTOMJEE		Real Estate	107.848815	543.050	0.000000	7.266816	14.187493	1302.97

```
In [5]: 1 Stocks.shape
```

```
Out[5]: (4519, 18)
```

```
In [6]: 1 Stocks.describe()
```

```
Out[6]:
```

	Market_Cap	Close_Price	PE_Ratio	5Y_Historical_Revenue_Growth	1Y_Forward_Revenue_Growth	Total_Revenue	PBT	Taxes_And_Other
count	4.519000e+03	4519.000000	4.519000e+03	4519.000000	4519.000000	4519.000000	4519.000000	4519.
mean	6.399541e+03	385.343542	-3.942686e+14	9.354426	64.154372	3134.938323	317.898774	94
std	4.742482e+04	2016.230695	3.883200e+16	37.086713	3358.671702	24882.725182	2811.715676	758.
min	0.000000e+00	0.120000	-2.560000e+18	-88.967448	-85.894054	0.000000	-28234.100000	-2466.
25%	2.002746e+01	17.775000	5.656455e-02	-2.043584	14.187493	11.195000	0.030000	0.
50%	1.078488e+02	64.635000	1.411507e+01	7.266816	14.187493	99.920000	3.100000	0.
75%	8.775942e+02	244.625000	3.688932e+01	14.859723	14.187493	633.185000	40.625000	9.
max	1.778090e+06	92565.150000	4.010000e+17	566.603634	225794.879700	724744.320000	84142.000000	23437.

```
In [7]: 1 Stocks.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4519 entries, 0 to 4518
Data columns (total 18 columns):
 #   Column           Non-Null Count  Dtype  
 ---  -- 
 0   Name             4519 non-null   object  
 1   Ticker            4519 non-null   object  
 2   Sub_Sector        3932 non-null   object  
 3   Market_Cap       4519 non-null   float64 
 4   Close_Price      4519 non-null   float64 
 5   PE_Ratio          4519 non-null   float64 
 6   5Y_Historical_Revenue_Growth  4519 non-null   float64 
 7   1Y_Forward_Revenue_Growth   4519 non-null   float64 
 8   Total_Revenue     4519 non-null   float64 
 9   PBT              4519 non-null   float64 
 10  Taxes_And_Other_Items 4519 non-null   float64 
 11  DII_Holding_Change_3M 4519 non-null   float64 
 12  FII_Holding_Change_3M 4519 non-null   float64 
 13  MF_Holding_Change_3M 4519 non-null   float64 
 14  Promoter_Holding_Change_3M 4519 non-null   float64 
 15  Percentage_Buy_Reco  4519 non-null   float64 
 16  No_of_analysts_with_buy_reco 4519 non-null   int64  
 17  Percentage_Upside   4519 non-null   float64 
dtypes: float64(14), int64(1), object(3)
memory usage: 635.6+ KB
```

➤ Let's find out the Null count.

```
In [8]: 1 stocks.isnull().sum()
```

```
Out[8]:
```

Name	0
Ticker	0
Sub_Sector	587
Market_Cap	0
Close_Price	0
PE_Ratio	0
5Y_Historical_Revenue_Growth	0
1Y_Forward_Revenue_Growth	0
Total_Revenue	0
PBT	0
Taxes_And_Other_Items	0
DII_Holding_Change_3M	0
FII_Holding_Change_3M	0
MF_Holding_Change_3M	0
Promoter_Holding_Change_3M	0
Percentage_Buy_Reco	0
No_of_analysts_with_buy_reco	0
Percentage_Upside	0

```
dtype: int64
```

Since there are **587** null values in **Sub\_Sector** column we can ignore it for now because we can bifurcate them while doing visualization.

➤ Let's have a reference for columns in order to find some insights.

```
In [9]: 1 Stocks.columns
```

```
Out[9]: Index(['Name', 'Ticker', 'Sub_Sector', 'Market_Cap', 'Close_Price', 'PE_Ratio',  
       '5Y_Historical_Revenue_Growth', '1Y_Forward_Revenue_Growth',  
       'Total_Revenue', 'PBT', 'Taxes_And_Other_Items',  
       'DII_Holding_Change_3M', 'FII_Holding_Change_3M',  
       'MF_Holding_Change_3M', 'Promoter_Holding_Change_3M',  
       'Percentage_Buy_Reco', 'No_of_analysts_with_buy_reco',  
       'Percentage_Upside'],  
      dtype='object')
```

➤ Finding out some insights.

```
In [10]:
```

```
1 # Finding Out Some Insights  
2  
3 Stocks['PAT (Profit After Tax)'] = Stocks['PBT'] - Stocks['Taxes_And_Other_Items'] # Profit After Tax (PAT) Formula  
4  
5 Stocks['Profit_Margin'] = (Stocks['PAT (Profit After Tax)'] / Stocks['Total_Revenue']) * 100 # Profit Margin  
6  
7 Stocks['Retail_Holding_Change'] = -(Stocks['DII_Holding_Change_3M'] + Stocks['FII_Holding_Change_3M'] + Stocks['MF_Holding_C  
8  
9 Stocks['PS (Price_To_Sales) Ratio'] = Stocks['Market_Cap'] / Stocks['Total_Revenue']
```

```
In [11]: 1 Stocks['PAT (Profit After Tax)']
```

```
Out[11]: 0      60705.00  
1      38327.00  
2      38052.75  
3      22110.00  
4      25110.11  
...  
4514      0.15  
4515      1.27  
4516      2.48  
4517      2.48  
4518      139.62  
Name: PAT (Profit After Tax), Length: 4519, dtype: float64
```

```
In [12]:
```

```
1 Stocks['Profit_Margin']
```

```
Out[12]: 0      8.45407  
1      19.577366  
2      22.691588  
3      17.839853  
4      15.939251  
...  
4514      0.985545  
4515      5.841766  
4516      2.481986  
4517      2.481986  
4518      10.715519  
Name: Profit_Margin, Length: 4519, dtype: float64
```

```
In [13]: 1 Stocks['Retail_Holding_Change']

Out[13]: 0      0.000778
1      0.258454
2     -0.299827
3    -8.934226
4   -0.285631
...
4514  -0.000000
4515  -0.000000
4516  -0.000000
4517  -0.000000
4518  -0.000000
Name: Retail_Holding_Change, Length: 4519, dtype: float64
```

```
In [14]: 1 Stocks['PS (Price_To_Sales) Ratio']

Out[14]: 0      2.476238
1      6.342352
2      5.317944
3      5.492092
4      4.147223
...
4514  7.085993
4515  4.960847
4516  1.079352
4517  1.079352
4518  0.082772
Name: PS (Price_To_Sales) Ratio, Length: 4519, dtype: float64
```

- Now let's create an additional column for 'Market\_Cap\_Size' so that it would be helpful to determine the stock falls in Large Cap, Mid Cap or Small Cap.

```
In [15]: 1 def cat_market(x):
2     if x >= 50000:
3         return 'Large_Cap'
4     elif x > 20000:
5         return 'Mid_Cap'
6     else :
7         return 'Small_Cap'
8
9 Stocks['Market_Cap_Size'] = Stocks['Market_Cap'].map(cat_market)

In [16]: 1 Stocks['Market_Cap_Size']

Out[16]: 0      Large_Cap
1      Large_Cap
2      Large_Cap
3      Large_Cap
4      Large_Cap
...
4514  Small_Cap
4515  Small_Cap
4516  Small_Cap
4517  Small_Cap
4518  Small_Cap
Name: Market_Cap_Size, Length: 4519, dtype: object
```

- Now download the updated csv file from python by using below code so that we can do visualization on it through Tableau.

```
In [17]: 1 Stocks.to_csv('C:\\\\Users\\\\Sohail\\\\Sohail_Work\\\\Stock_Market_Dataset_GM100_Python_Insights.csv', index = False)
```

#### **4. DATA VISUALIZATION THROUGH TABLEAU**

Tableau is a powerful data visualization tool that offers several advantages, making it a popular choice for data analysis and visualization in various industries. Some of the benefits are listed below :

- 1) **Ease of Use:** Tableau provides a user-friendly and intuitive interface that allows users with varying levels of technical expertise to create visualizations and analyse data effectively. Users can create interactive dashboards and reports using a drag-and-drop approach, reducing the need for extensive coding or scripting.
- 2) **Wide Range of Data Sources:** Tableau can connect to a wide variety of data sources, including databases, spreadsheets, cloud-based data sources, web services, and more. This flexibility enables users to work with data from diverse sources and combine them into meaningful visualizations.
- 3) **Interactive Visualizations:** Tableau allows for the creation of highly interactive and dynamic visualizations, including charts, graphs, maps, and dashboards. Users can explore data, drill down into details, apply filters, and perform ad-hoc analysis within the visualizations themselves.
- 4) **Real-Time Data Updates:** Tableau supports real-time data connections, enabling users to visualize and analyse data as it changes or updates. This is valuable for businesses that rely on up-to-the-minute information for decision-making.
- 5) **Scalability:** Tableau is scalable, making it suitable for both small businesses and large enterprises. It can handle large datasets and complex analytics tasks. Organizations can grow their Tableau deployment as their data needs expand.
- 6) **Sharing and Collaboration:** Tableau allows users to share visualizations, dashboards, and reports with others within the organization or externally. Collaboration features make it easy for teams to work together on data projects.
- 7) **Advanced Analytics:** Tableau supports advanced analytics and statistical functions, enabling users to perform complex calculations and predictive modelling within the tool. Integration with Excel and Python allows for even more advanced analysis.

**Story**   **Layout**

New story point

Blank   Duplicate

Total Count of ...  
 Count of Sub ...  
 Count of ...  
 Top Sub ...  
 PE Ratio  
 Top Sub ...  
 Top Sub ...  
 Top Companie...  
 Top Companie...  
 Top Companie...  
 % Buy Reco v...  
 Close Price  
 5Y vs 1Y

A Drag to add text

Show title

Size

Automatic

## Stock Market Analysis for Commodity Chemicals Sub Sector

Sub Sectors of Stock Market

**Count of Sub Sectors**   **Total Count of Companies**   **Count of Companies in Commodity Chemicals Sub Sector**

140      4,519

**Top Sub Sectors**

Sub Sector	Count of Name
P..	12
P..	11
Pr..	10
In..	8
IT..	8
C..	7
R..	7
C..	6
In..	5
A..	5
Ir..	4
F..	4
Di..	3
S..	3
S..	3
O..	2
Oi..	2
H..	2
G..	2
Fo..	2
C..	2

**Top Sub Sectors - Mid Market Cap**

Sub Sector	Market Cap
Phar..	~300K
Cons..	~200K
Real..	~180K
Auto..	~160K
Indu..	~140K
Publ..	~130K
Priv..	~120K
Soft..	~110K
Com..	~100K
Spec..	~90K

**Story**   **Layout**

New story point

Blank   Duplicate

Total Count of ...  
 Count of Sub ...  
 Count of ...  
 Top Sub ...  
 PE Ratio  
 Top Sub ...  
 Top Sub ...  
 Top Companie...  
 Top Companie...  
 Top Companie...  
 % Buy Reco v...  
 Close Price  
 5Y vs 1Y

A Drag to add text

Show title

Size

Automatic

## Top Companies In Commodity Chemicals Sub Sector

**Top Companies - Large Market Cap**

Name	Market Cap
RELIANCE	~2.5M
TCS	~1.8M
HDFCBANK	~1.5M
INFY	~1.2M
ICICIBANK	~1.0M
HINDUNILVR	~1.0M
SBIN	~800K
BHARTIARTL	~600K
HDFC	~600K
ADANIET	~600K

**Top Companies - Mid Market Cap**

Name	Market Cap
MOTHERSON	~80K
LODHA	~60K
ACC	~55K
CONCOR	~50K
INDOFET	~45K
PATANJALI	~40K
HDFCAMS	~35K
PGHH	~30K
SCHAFFLER	~25K
UBL	~25K

**Top Companies In Market Cap For Commodity Chemicals**

Sub Sector / Ticker / Name	Market Cap
SOLARIS	~15K
DEEPAL	~10K
LINDE	~10K
SUMITRA	~10K
CASTROL	~5K
ACI	~5K
RAIN	~5K
GUJAL	~5K
POLYPL	~5K
GHCL	~5K

**Story**   **Layout** <

**New story point**

Blank    Duplicate

Pricing of Companies belonging to Commodity Chemicals Sub Sector

**5Y vs 1Y**

**Sub Sector / Ticker / Name**

Commodity Chemicals				
SOLARINDS	GULFOILLUB	SUMICHEM	DEEPAKNTR	POLYPLEX
Mid_Cap	Small_Cap	Mid_Cap	Mid_Cap	Small_Cap
Mid_Cap	Small_Cap	Mid_Cap	Mid_Cap	Small_Cap
Solar Industri...	Gulf Oil Lubric...	Sumitomo Che...	Deepak Nitrite...	Polyplex Corp...

1YFo.SY His..

**Total Revenue**

**Sub Sector / Ticker / Name**

Commodity Chemicals				
DEEPAKNTR	POLYPLEX	SOLARINDS	SUMICHEM	RAIN
Mid_Cap	Small_Cap	Mid_Cap	Mid_Cap	Small_Cap
5K	10K	15K	5K	15K
Deepak Nitrit...	Polyplex Corp...	Solar Industri...	Sumitomo Che...	Rain Industri...

Total Revenue

**A** Drag to add text

Show title

**Size**

Automatic

**Story**   **Layout** <

**New story point**

Blank    Duplicate

Predictions of Companies belonging to Commodity Chemicals Sub Sector

**PE Ratio**

**PE Ratio (Bins)**

Count of PE Ratio

PE Ratio (Bins)

-59.7, -11.6, 31.6

**% Buy Reco v/s % Upside**

**Sub Sector**   **Name** \*

Sub Sector	Name
Chemicals	Valiant O.
	Transpek
	Bodal Ch.

Avg. Percentage Buy Reco \*

**Close Price**

**Sub Sector / Ticker / Name**

Commodity Chemicals				
POLSON	SOLARINDS	LINDEINDIA	DEEPAKNTR	POLYPLEX
Small_Cap	Mid_Cap	Mid_Cap	Mid_Cap	Small_Cap
10K	5K	0K	5K	0K
Polson Ltd	Solar Industri...	Linde India Ltd	Deepak Nitrite...	Polyplex Corp...

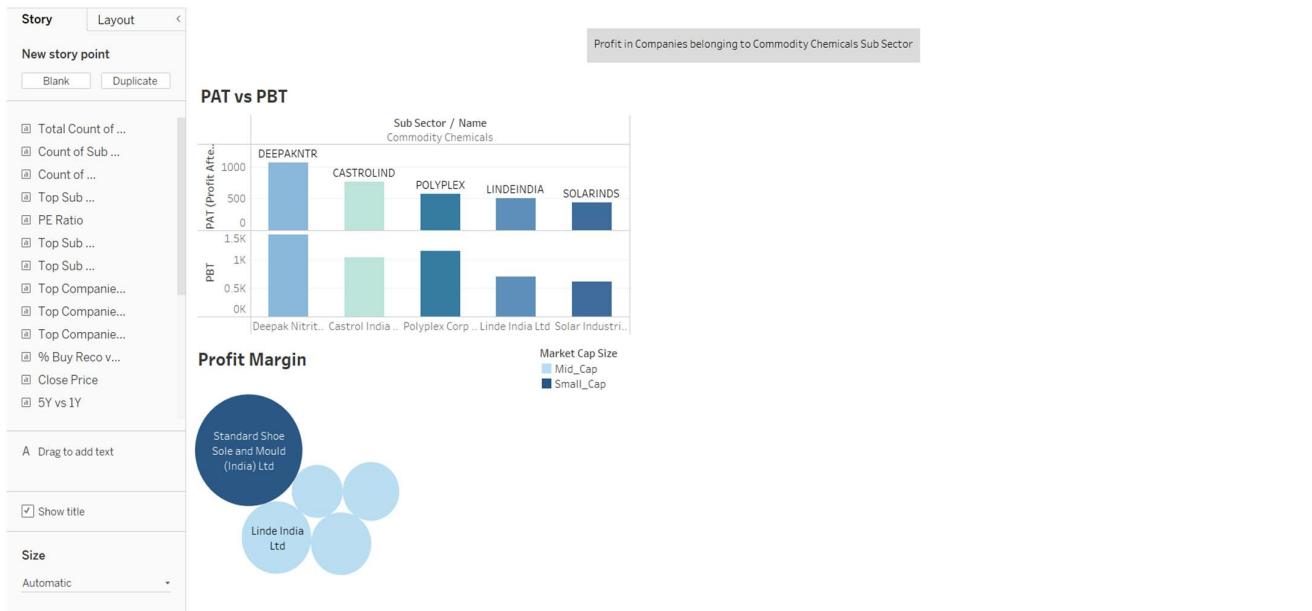
Avg.Close Price

**A** Drag to add text

Show title

**Size**

Automatic



## Dashboard Link:

[https://public.tableau.com/views/StockMarketAnalysis\\_16949338904110/SubSectorsofStockMarket?:language=en-US&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/StockMarketAnalysis_16949338904110/SubSectorsofStockMarket?:language=en-US&:display_count=n&:origin=viz_share_link)

## **CONCLUSION**

The Commodity Chemicals Sub Sector accounts for only 2.61% of the dataset with a market capitalization of 0.8%. It is further divided into three categories: Small Cap (114 Stocks), Mid Cap (4 Stocks) & Large Cap (0 Stocks).

After detailed analysis, we have identified five stocks in the Commodity Chemicals Sub Sector that we believe are the best for investors. These stocks met all of our expectations and outperformed other stocks in the sector.

The Five Stocks are:

1. **Solar Industries India Limited** - Solar Industries India Limited is a Public incorporated on 24 February 1995. It is classified as Non Govt company and is registered at Registrar of Companies, Mumbai. Solar Industries India Limited has entered into the Business of Propulsion system for space application, which is synergistic with current business of ammunition, and has decided to make an equity investment in a startup "Skyroot Aerospace" for manufacturing of Space Launch Vehicles.

Link - <https://solargroup.com/#home>

2. **Deepak Nitrite Limited** - Deepak Nitrite Ltd. is an Indian chemical manufacturing company with around 50 years of rich heritage and legacy, has been one of the earliest adopters of the 'Make in India' philosophy. Over the years, they have made significant efforts to diversify in refining processes, fostering stronger relationships, adopting sustainable practices to shareholders, investors, partners, customers, employees which leads to flourishing profits.

Link - <https://www.godeepak.com/>

3. **Linde India Limited** - Linde stands for over 130 years of technological progress. Technology, innovation and an inventive spirit characterized this company from the very beginning. Shareholders constitute the most important stakeholders in Linde India Limited and they do value their long-term relationship. Increasing shareholder value is a key long-term objective. In the 2017 financial year, The Linde Group generated revenue of EUR 17.113 billion & it's also a leading global industrial Gases and Engineering company with sales of \$33 billion in 2022, making it one of the leading gases and engineering companies in the world with approximately 58,000 employees working in more than 100 countries worldwide.

Link - <https://www.linde.in/en/index.html>

4. **Sumitomo Chemical India Limited** - Sumitomo Chemical India Limited is an India based company. The Company manufactures, imports and markets products for crop protection, grain fumigation, rodent control, bio pesticides, environmental health, professional pest control and feed additive. Sumitomo Chemicals is been rated as a good company to invest.

Link - <https://sumichem.co.in/index.php>

5. **Polyplex Corporation Limited** - Polyplex Corporation Ltd. is an Indian multinational company which produces biaxially oriented polyester film for packaging, electrical and various industrial applications. The company is a major exporter of PET film to the United States, Europe, Southeast Asia, South America and Australia and is considered as semi strong company by investors for purchasing of stocks.

Link - <https://www.polyplex.com/>

## **RECOMMENDATION**

We recommend that investors should consider investing in the Commodity Chemicals Sub Sector. This Sub Sector is poised for growth in the foreseeable future, and the five stocks highlighted in this report present investors with the potential for attractive favorable returns.

## **LIMITATION**

This report is for informational purposes only and should not be construed as investment advice. Past performance is not indicative of future results. Investors should carefully consider their own investment objectives and risk tolerance before making any investment decisions.

The analysis is based on historical dataset, which may not be indicative of future results. The analysis does not take into account all factors that could affect the performance of the stocks, such as changes in the macroeconomic environment or the introduction of new technologies. The analysis is based on a limited dataset, which may not be representative of the entire Commodity Chemicals Sub Sector. The analysis is subjective and may not be universally shared by all investors.

## **AREA OF STUDY**

Stock Market in India (Commodity Chemicals Sub Sector)

## **LOCATION OF STUDY**

Mumbai

## **APPENDICES**

### **PROFILE OF PROJECT GUIDE**

**Name:** ASHISH KAPASIYA

**Educational Qualification:** Postgraduate Degree in Data Science

**Professional Experience:** 5 Years in Electrical Engineering, 1.5 years in Data Science

**Organization:** IMS Proschool Private Limited

**Current Designation:** Assistant Manager – Product and Student Outcome

**Brief Profile:** An electrical engineer turned data scientist, Ashish Kapasiya is a data science trainer and mentor at IMS Proschool, Delhi. He is also the Assistant Manager of Product and Student Outcome. He conducts data science, machine learning, and database systems workshops across schools in Delhi and also online nationally.

**Address:** Proschool Skill Development Pvt Ltd, 10th floor, Tolstoy House 15

Street: Tolstoy Road

City: Janpath, Connaught Place, New Delhi State: Delhi

Country: India

Phone Number (Office): 7304957442

Mobile Number (10 Digits): 8097594008

Email: [delhi.academics.analytics@proschoolonline.com](mailto:delhi.academics.analytics@proschoolonline.com)

## **BIBLIOGRAPHY**

- <https://economictimes.indiatimes.com/markets>
- <https://chemicals.gov.in/>
- <https://www.bseindia.com/>
- <https://stockedge.com/>
- <https://in.investing.com/>
- <https://www.ris.org.in/>
- <https://www.cii.in/>
- <https://ficci.in/api/home>
- <https://www.cmie.com/>
- <https://www.indianchemicalcouncil.com/>

## **BOOK REFERENCES**

- **The Chemical Industry in India: Investment Opportunities, Growth, and Challenges** by Bishwanath Ghosh
- **Indian Petrochemicals Industry: A Strategic Outlook** by Nitya Nanda
- **India's Chemical Industry: Policies, Practices, and Performance** by Anil K. Sinha
- **Chemical Industry in India: Policies, Opportunities and Challenges** by Bibek Debroy and Kirit S. Parikh
- **Chemicals and Petrochemicals Industry in India: Policy and Performance** by P.R. Brahmananda and Arun K. Misra

**END OF REPORT**

**!!!!THANK YOU!!!!**