

Problem 2: Market Risk

The dataset contains 6 years of information (weekly stock information) on the stock prices of 10 different Indian Stocks. Calculate the mean and standard deviation on the stock returns and share insights.

About Data

Table 1: Head of the dataset showing the first 5 records

	Date	Infosys	Indian Hotel	Mahindra & Mahindra	Axis Bank	SAIL	Shree Cement	Sun Pharma	Jindal Steel	Idea Vodafone	Jet Airways
0	31-03-2014	264	69	455	263	68	5543	555	298	83	278
1	07-04-2014	257	68	458	276	70	5728	610	279	84	303
2	14-04-2014	254	68	454	270	68	5649	607	279	83	280
3	21-04-2014	253	68	488	283	68	5692	604	274	83	282
4	28-04-2014	256	65	482	282	63	5582	611	238	79	243

The dataset was loaded and the head of the dataset was checked. Table 1 shows the first 5 records of the dataset. From this table, we can see the different variables or columns of the dataset.

Table 2: Information of the dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 314 entries, 0 to 313
Data columns (total 11 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Date                                  314 non-null    object
1   Infosys                              314 non-null    int64
2   Indian Hotel                          314 non-null    int64
3   Mahindra & Mahindra                  314 non-null    int64
4   Axis Bank                            314 non-null    int64
5   SAIL                                  314 non-null    int64
6   Shree Cement                         314 non-null    int64
7   Sun Pharma                           314 non-null    int64
8   Jindal Steel                         314 non-null    int64
9   Idea Vodafone                        314 non-null    int64
10  Jet Airways                          314 non-null    int64
dtypes: int64(10), object(1)
memory usage: 27.1+ KB
```

The dataset has 11 columns and 314 rows seen in Table 2. There are 314 records in all the columns. This means there are no missing records based on this initial analysis that was done.

Table 3: Data type of the columns in the dataset

```
Date                object
Infosys             int64
Indian Hotel        int64
Mahindra & Mahindra int64
Axis Bank           int64
SAIL                int64
Shree Cement        int64
Sun Pharma          int64
Jindal Steel        int64
Idea Vodafone       int64
Jet Airways         int64
dtype: object
```

The column 'Date' is of object data type while all the other columns are of integer data type. This is seen in Table 2 and Table 3. There are 10 independent variables.

The column 'Date' is detected as object data type which is changed to read the data as a time series data. Table 4 shows the data type of the columns after changing the data type of the 'Date' variable to a time series data i.e., datetime64[ns] data type.

Table 4: Data type of the columns

```
Date                datetime64[ns]
Infosys             int64
Indian Hotel        int64
Mahindra & Mahindra int64
Axis Bank           int64
SAIL                int64
Shree Cement        int64
Sun Pharma          int64
Jindal Steel        int64
Idea Vodafone       int64
Jet Airways         int64
dtype: object
```

```
The no. of rows: 314
The no. of columns: 11
```

Figure 1: Shape of the dataset

The shape of the data is (314, 11) meaning the dataset has 314 rows and 11 columns shown in Figure 1.

Table 5: Missing value of the columns in the dataset

```
Date          0
Infosys        0
Indian Hotel   0
Mahindra & Mahindra  0
Axis Bank      0
SAIL           0
Shree Cement   0
Sun Pharma     0
Jindal Steel   0
Idea Vodafone  0
Jet Airways    0
dtype: int64
```

The dataset was further checked for missing values and it is seen from Table 5 that there are no missing values in the dataset. This was also seen earlier in Table 2.

The dataset is now checked for duplicates.

```
The no. of duplicated rows is: 0
```

Figure 2: Number of duplicates in the dataset

The dataset is checked for duplicates and it was found that there were no duplicate rows as seen in Figure 2.

Table 6: Descriptive statistics of the numerical columns in the dataset

	count	mean	std	min	25%	50%	75%	max
Infosys	314.0	511.340764	135.952051	234.0	424.00	466.5	630.75	810.0
Indian Hotel	314.0	114.560510	22.509732	64.0	96.00	115.0	134.00	157.0
Mahindra & Mahindra	314.0	636.678344	102.879975	284.0	572.00	625.0	678.00	956.0
Axis Bank	314.0	540.742038	115.835569	263.0	470.50	528.0	605.25	808.0
SAIL	314.0	59.095541	15.810493	21.0	47.00	57.0	71.75	104.0
Shree Cement	314.0	14806.410828	4288.275085	5543.0	10952.25	16018.5	17773.25	24806.0
Sun Pharma	314.0	633.468153	171.855893	338.0	478.50	614.0	785.00	1089.0
Jindal Steel	314.0	147.627389	65.879195	53.0	88.25	142.5	182.75	338.0
Idea Vodafone	314.0	53.713376	31.248985	3.0	25.25	53.0	82.00	117.0
Jet Airways	314.0	372.659236	202.262668	14.0	243.25	376.0	534.00	871.0

Table 6 shows the description or the summary of the numerical columns in the dataset. The values of mean, standard deviation, minimum and maximum, 25th, 50th and 75th percentile is mentioned in the above table. It can be seen that there are no missing values in the columns. By looking at Table 6, we are able to deduce that the mean and standard deviation value of each variable has vast differences. This is probably because each column or variable are the stock prices of different stocks at different time points. People are interested in the stock returns rather than the stock prices. Therefore, it may not be very meaningful to analyse this right now.

2.1 Draw Stock Price Graph (Stock Price vs Time) for any 2 given stocks with inference

Stock Price graph of Infosys

The stock price graph of 'Infosys' is plotted and shown in Figure 3.

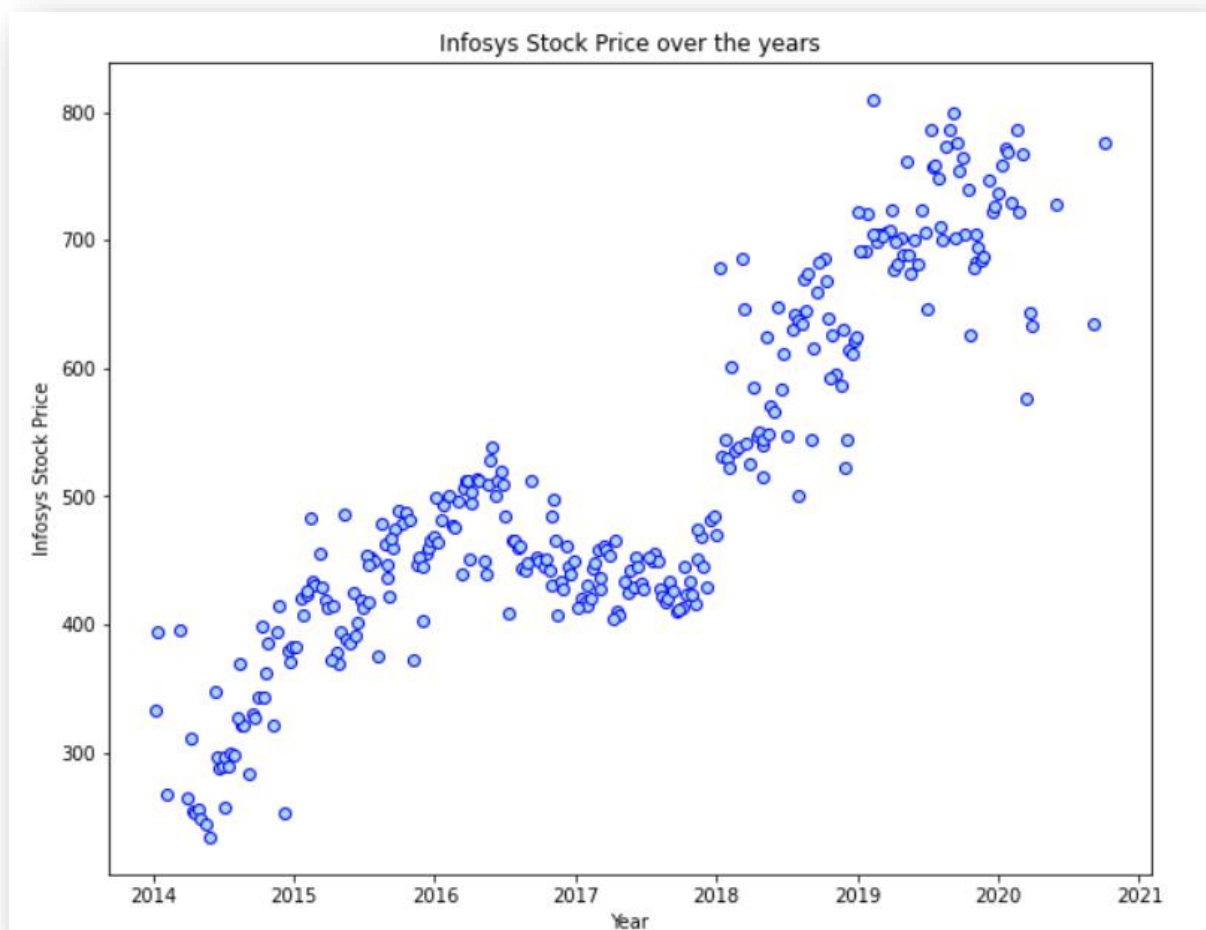


Figure 3: Infosys stock price graph

From Figure 3, it is seen that the stock price of 'Infosys' used to be around 250-400 in 2014 and 2015 whereas in 2020, the stock price has increased as high as 700-800. Overall, the stock price seems to have shown an upward movement with a slight dip in the year 2017.

Stock Price graph of Sun Pharma

The stock price graph of 'Sun Pharma' is plotted and shown in Figure 4.

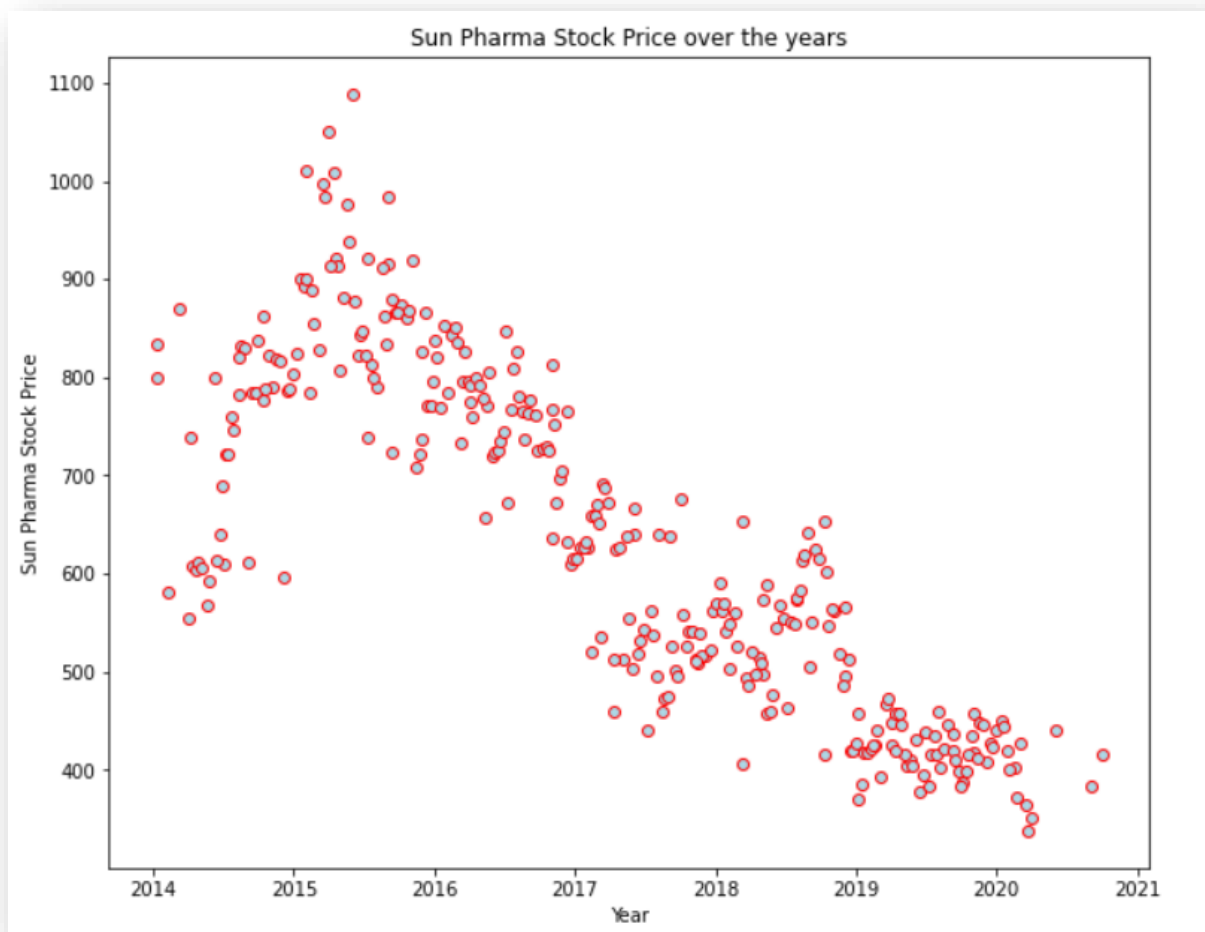


Figure 4: Sun Pharma stock price graph

From Figure 4, it is seen that the stock price of 'Sun Pharma' used to be around 600 in 2014. After that the stock prices showed a huge increase till 2016 where the prices of the stocks were as high as 1000 at a point. After 2016, the stock prices showed a declining trend. In 2019 and 2020, the prices of stocks were around 400. Overall, the stock price seems to have shown an upward movement till 2016 and a downward movement from 2016 to 2020.

2.2 Calculate Returns for all stocks with inference

Returns

Returns is the change in the stock price compared to what the stock price previously was. Returns is denoted by R.

The returns can be calculated as the ratio of the difference in stock price to the stock price at a time (t).

$$\text{Returns } (R_t) = \frac{\text{Price } (t) - \text{Price } (t-1)}{\text{Price } (t-1)}$$

It can also be calculated as the difference of log of price at time (t) and the log of price at time (t-1).

$$\text{Returns } (R_t) = \log \text{Price } (t) - \log \text{Price } (t-1)$$

The returns of all the stocks are calculated and the head of the data is shown in Table 7 (only the head of the data is shown due to space constraints). Table 7 shows week over week returns for all the stocks. The first row is displayed as 'NaN' as this observation did not have a previous value to be converted into a return. The values following this have all been calculated and the returns are shown.

Table 7: Stock Returns (head)

	Infosys	Indian Hotel	Mahindra & Mahindra	Axis Bank	SAIL	Shree Cement	Sun Pharma	Jindal Steel	Idea Vodafone	Jet Airways
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	-0.026873	-0.014599	0.006572	0.048247	0.028988	0.032831	0.094491	-0.065882	0.011976	0.086112
2	-0.011742	0.000000	-0.008772	-0.021979	-0.028988	-0.013888	-0.004930	0.000000	-0.011976	-0.078943
3	-0.003945	0.000000	0.072218	0.047025	0.000000	0.007583	-0.004955	-0.018084	0.000000	0.007117
4	0.011788	-0.045120	-0.012371	-0.003540	-0.076373	-0.019515	0.011523	-0.140857	-0.049393	-0.148846

2.3 Calculate Stock Means and Standard Deviation for all stocks with inference

Returns can be measured by means and standard deviation.

Returns measured by means

Expected returns measures the amount of money that an investor would make on an average over a period of time. Expected returns is denoted as $E(R)$.

$$E(R) = \frac{1}{T} \sum_{t=1}^T R_t$$

The stock means are calculated and shown in Table 8. Stock Means are average returns that the stock is making on a week-to-week basis.

Table 8: Stock Means

Shree Cement	0.003681
Infosys	0.002794
Axis Bank	0.001167
Indian Hotel	0.000266
Sun Pharma	-0.001455
Mahindra & Mahindra	-0.001506
SAIL	-0.003463
Jindal Steel	-0.004123
Jet Airways	-0.009548
Idea Vodafone	-0.010608
dtype:	float64

From Table 8, it is seen that 'Shree Cement' has the highest returns and 'Idea Vodafone' has the lowest returns.

Returns measured by Standard Deviation

The risk or variability of the returns is measured by the standard deviation. Stock standard deviation is a measure of volatility meaning the more a stock's returns vary from the stock's average return, the more volatile the stock.

$$\text{Standard deviation of returns (R)} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (R_t - E(R))^2}$$

The standard deviation of returns is calculated and shown in Table 9.

Table 9: Standard Deviation of Stocks

Idea Vodafone	0.104315
Jet Airways	0.097972
Jindal Steel	0.075108
SAIL	0.062188
Indian Hotel	0.047131
Axis Bank	0.045828
Sun Pharma	0.045033
Mahindra & Mahindra	0.040169
Shree Cement	0.039917
Infosys	0.035070
dtype: float64	

From Table 9, it is seen that 'Infosys' has the lowest risk factor. Therefore, it would be the least risky investment option while 'Idea Vodafone' has the highest risk factor.

2.4 Draw a plot of Stock Means vs Standard Deviation and state your inference

The stock means (average) and standard deviation (volatility) that were calculated were merged into one dataframe and is shown in Table 10. These values were then plotted and shown in Figure 5.

Table 10: Dataframe with Stock Means (Average) and Standard Deviation (Volatility)

	Unnamed: 0	Average	Volatility
0	Infosys	0.002794	0.035070
1	Indian Hotel	0.000266	0.047131
2	Mahindra & Mahindra	-0.001506	0.040169
3	Axis Bank	0.001167	0.045828
4	SAIL	-0.003463	0.062188
5	Shree Cement	0.003681	0.039917
6	Sun Pharma	-0.001455	0.045033
7	Jindal Steel	-0.004123	0.075108
8	Idea Vodafone	-0.010608	0.104315
9	Jet Airways	-0.009548	0.097972

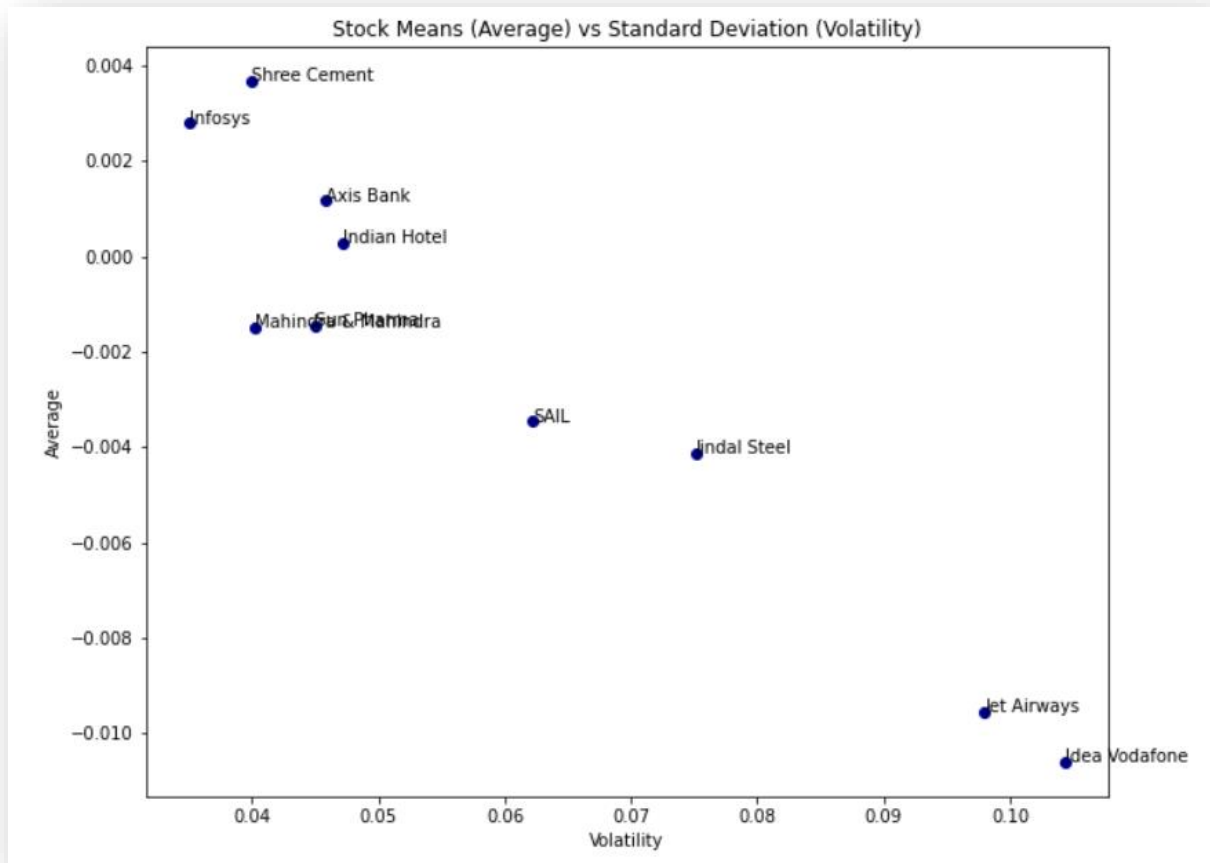


Figure 5: Scatter plot of Stock Means (Average) vs Standard Deviation (Volatility)

In Figure 5, average is plotted on the y-axis and volatility is plotted on the x-axis. Therefore, stocks on the left most have lower risk and vice versa. Stocks on the top most of the graph have high returns and stocks on the bottom indicate low returns.

From Figure 5, we can deduce that 'Shree Cement' and 'Infosys' have low risk and high returns. 'Jet Airways' and 'Idea Vodafone' stocks have very high risk and low returns. Stocks like 'SAIL' and 'Jindal Steel' have medium risk and medium returns.

From the perspective of returns, 'Shree Cement' is the best as it has the highest returns. This is followed by 'Infosys' and 'Axis Bank'.

From the perspective of risk (as measured by standard deviation), 'Infosys' has the lowest risk followed by 'Shree Cement' and 'Mahindra & Mahindra'.

This is a very useful graph to find a balance between risk and reward when it comes to investing in different companies.

2.5 Conclusion and Recommendations

Stock with a lower mean and higher standard deviation do not play a role in a portfolio compared to a stock that has more returns and less risk. Thus, for the data we have here, we are only left few stocks:

- Ones with higher return for a comparative or lower risk

- Ones with medium returns and medium risk
- Ones with low returns and high risk

Ones with higher return for a comparative or lower risk are considered better for investment purposes.

From the perspective of returns, 'Shree Cement' is the best as it has the highest returns. This is followed by 'Infosys' and 'Axis Bank'.

From the perspective of risk (as measured by standard deviation), 'Infosys' has the lowest risk followed by 'Shree Cement' and 'Mahindra & Mahindra'.

Risk to reward ratio can be assessed using the values of stock means and standard deviation. More volatile stock might give short term gains but might not be a good investment in the long term as the risk factor is high. A low volatile stock might not be a good investment in short term, but might give a good return in the long term. Hence, based on the type of investment that the investor is looking for, and inferences can be derived from the stock means vs standard deviation graph.