



Comprehensive Report
on

“Data Analytics and Data Visualization with Tableau”
(Summer Course)

Submitted by:

Name: Hemanth Prasad Hegde
SRN: PES2UG19CS148
SEM: 7TH

Faculty Handling the Course:

Dr.Sudeepa Roy Dey
Dr.Prajwala TR
Prof.Ruby Dinakar

4th -9th, July 2022

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
PES UNIVERSITY, EC campus

(Established under Karnataka Act No. 16 of 2013)
Electronic City, Hosur Road, Bengaluru – 560 100, Karnataka, Ind

Table of Contents:

| <u>Content</u> | <u>Page no.</u> |
|--|------------------------|
| Data Analytics <ul style="list-style-type: none">- Dataset- Identify the nature of dataset whether it is monotonic- Find correlation among the features using Pearson Correlations- Applying random forest | 2 |
| Data Visualization <ul style="list-style-type: none">- Nifty50 over the years- The Trend and Forecast- Financial Ratios-Volume and Turnover | 4 |
| Insights | 7 |

Data Analytics:

Dataset:

```
[1] import pandas as pd
```

```
[2] df=pd.read_csv("/content/NIFTY 50.csv")
```

```
[3] df
```

| | Date | Open | High | Low | Close | Volume | Turnover | P/E | P/B | Div Yield |
|------|------------|----------|----------|----------|----------|-----------|----------|-------|------|-----------|
| 0 | 2000-01-03 | 1482.15 | 1592.90 | 1482.15 | 1592.20 | 25358322 | 884.15 | 25.91 | 4.63 | 0.95 |
| 1 | 2000-01-04 | 1594.40 | 1641.95 | 1594.40 | 1638.70 | 38787872 | 1973.69 | 26.67 | 4.76 | 0.92 |
| 2 | 2000-01-05 | 1634.55 | 1635.50 | 1555.05 | 1595.80 | 62153431 | 3084.79 | 25.97 | 4.64 | 0.95 |
| 3 | 2000-01-06 | 1595.80 | 1639.00 | 1595.80 | 1617.60 | 51272875 | 2531.18 | 26.32 | 4.70 | 0.94 |
| 4 | 2000-01-07 | 1616.60 | 1628.25 | 1597.20 | 1613.30 | 54315945 | 1914.63 | 26.25 | 4.69 | 0.94 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 5488 | 2022-01-21 | 17613.70 | 17707.60 | 17485.85 | 17617.15 | 277645373 | 25941.64 | 24.29 | 4.44 | 1.18 |
| 5489 | 2022-01-24 | 17575.15 | 17599.40 | 16997.85 | 17149.10 | 323847388 | 28387.73 | 23.63 | 4.32 | 1.21 |
| 5490 | 2022-01-25 | 17001.55 | 17309.15 | 16836.80 | 17277.95 | 326515896 | 31344.46 | 23.47 | 4.35 | 1.20 |
| 5491 | 2022-01-27 | 17062.00 | 17182.50 | 16866.75 | 17110.15 | 395596577 | 36296.64 | 23.09 | 4.31 | 1.21 |
| 5492 | 2022-01-28 | 17208.30 | 17373.50 | 17077.10 | 17101.95 | 355284285 | 29799.89 | 23.10 | 4.31 | 1.21 |

5493 rows x 10 columns

Identify the nature of dataset whether it is monotonic:

```
[4] df['Date'].is_monotonic
```

True

```
[5] df['Open'].is_monotonic
```

False

```
[6] df['High'].is_monotonic
```

False

```
[7] df['Close'].is_monotonic
```

False

```
[8] df['Volume'].is_monotonic
```

False

```
[9] df['Turnover'].is_monotonic
```

False

```
[10] df['P/E'].is_monotonic
```

False

```
[11] df['P/B'].is_monotonic
```

False

```
[12] df['Div Yield'].is_monotonic
```

False

Find correlation among the features using Pearson Correlations:

[13] df.corr(method='pearson')

| | Open | High | Low | Close | Volume | Turnover | P/E | P/B | Div Yield |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Open | 1.000000 | 0.999933 | 0.999901 | 0.999848 | 0.659885 | 0.799974 | 0.794407 | 0.137718 | -0.439410 |
| High | 0.999933 | 1.000000 | 0.999870 | 0.999926 | 0.662015 | 0.801748 | 0.794609 | 0.138844 | -0.439911 |
| Low | 0.999901 | 0.999870 | 1.000000 | 0.999923 | 0.657046 | 0.797678 | 0.794504 | 0.137306 | -0.439531 |
| Close | 0.999848 | 0.999926 | 0.999923 | 1.000000 | 0.659543 | 0.799808 | 0.794764 | 0.138430 | -0.439880 |
| Volume | 0.659885 | 0.662015 | 0.657046 | 0.659543 | 1.000000 | 0.902435 | 0.595562 | -0.078315 | -0.169601 |
| Turnover | 0.799974 | 0.801748 | 0.797678 | 0.799808 | 0.902435 | 1.000000 | 0.736733 | 0.067043 | -0.271825 |
| P/E | 0.794407 | 0.794609 | 0.794504 | 0.794764 | 0.595562 | 0.736733 | 1.000000 | 0.398311 | -0.629573 |
| P/B | 0.137718 | 0.138844 | 0.137306 | 0.138430 | -0.078315 | 0.067043 | 0.398311 | 1.000000 | -0.555012 |
| Div Yield | -0.439410 | -0.439911 | -0.439531 | -0.439880 | -0.169601 | -0.271825 | -0.629573 | -0.555012 | 1.000000 |

Applying Random forest/Regression:

Weka Explorer

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier: Choose **RandomForest** -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Test options:

- ☐ Use training set
- ☐ Supplied test set
- ☒ Cross-validation Folds **10**
- ☐ Percentage split % **66**

More options...

(Num) Volume

Start Stop

Result list (right-click for options)

00:08:33 - trees.RandomForest

Classifier output

Scheme: weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Relation: NIFTY 50

Instances: 5493

Attributes:

- Date
- Open
- High
- Low
- Close
- Volume
- Turnover
- P/E
- P/B
- Div Yield

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 2.75 seconds

=== Cross-validation ===

=== Summary ===

| | |
|-----------------------------|----------------|
| Correlation coefficient | 0.8822 |
| Mean absolute error | 74428521.5027 |
| Root mean squared error | 115043872.9172 |
| Relative absolute error | 63.7274 % |
| Root relative squared error | 66.6899 % |
| Total Number of Instances | 5493 |

Status: OK

Log x 0

Data Visualization:

Nifty50 over the years:

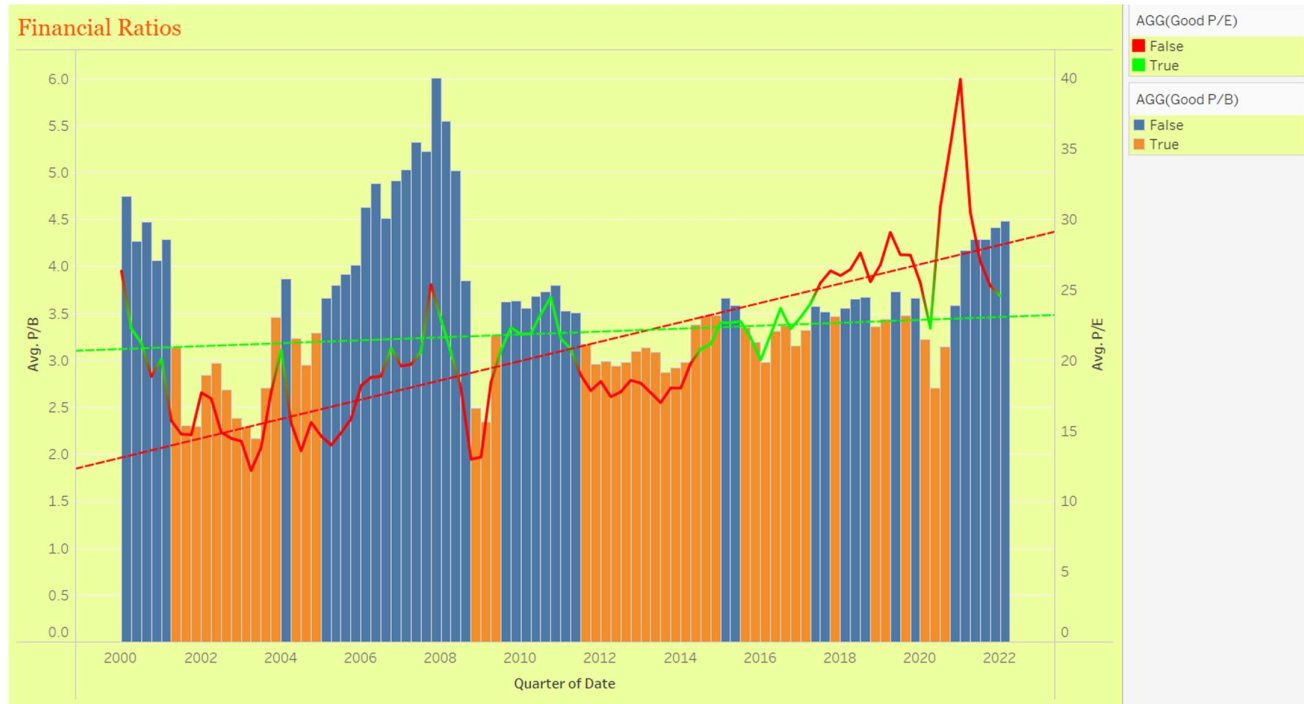


The Trend and Forecast:



On the trendline Nifty shows a good uptrend, which means it is safe to say in the coming years Nifty may give good gains. The shows an estimate for the trend of Nifty in the future interval of time.

Financial Ratios:

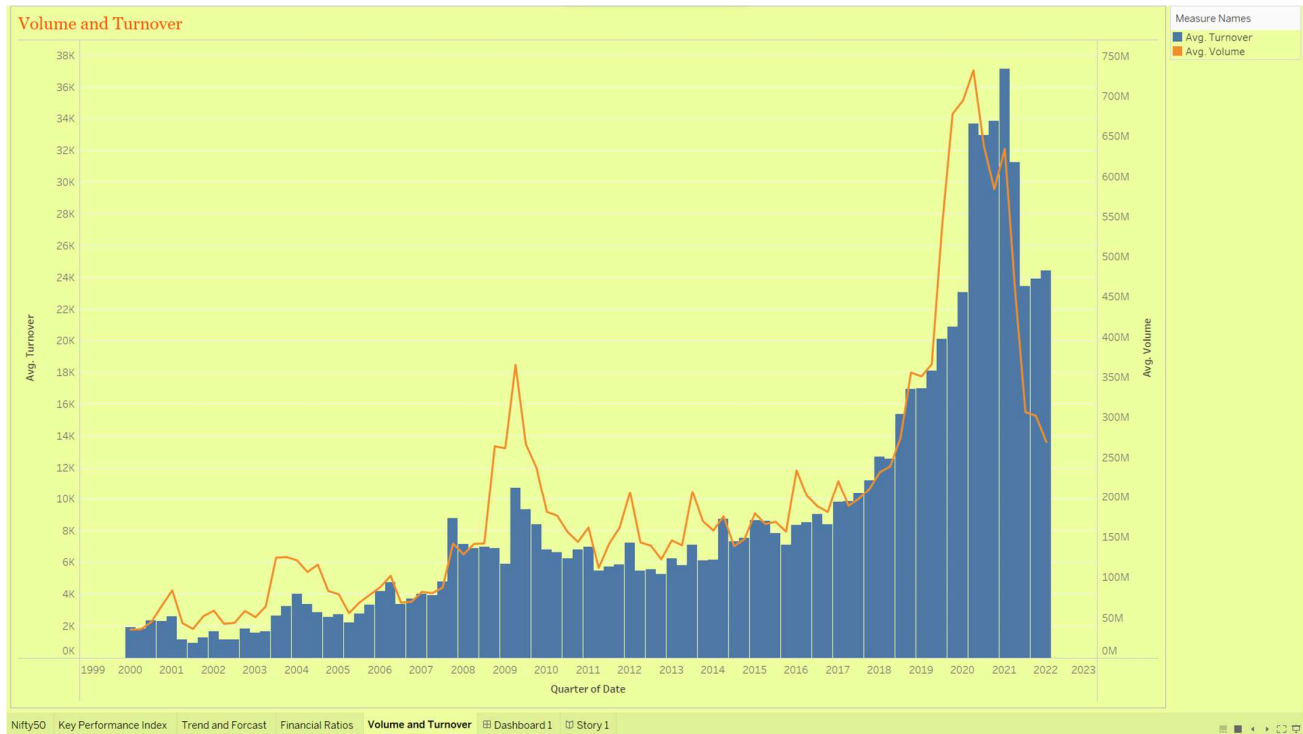


The price-to-book ratio, or P/B ratio, is a financial ratio used to compare a company's current market value to its book value (where book value is the value of all assets minus liabilities owned by a company). The PB ratio is essential for value investors- investors looking to purchase undervalued stock with the presumption that in the future, the market value of the stock will rise and they can sell off their shares at a profit. Some value investors and financial analysts also consider any value under 3.5 as a good PB ratio.

The Price Earnings Ratio (P/E Ratio) is the relationship between a company's stock price and earnings per share (EPS). Earnings are important when valuing a company's stock because investors want to know how profitable a company is and how profitable it will be in the future. Many value investors would refer to 20 to 25 as the average P/E ratio range.

On The plot for P/B and P/E v/s Each Quarter the current trend shows no major uptrend or downtrend which is a good thing because the Price Earnings ratio is considered good P/E within the range of 20-25 and lower the price-to-book ratio the better.

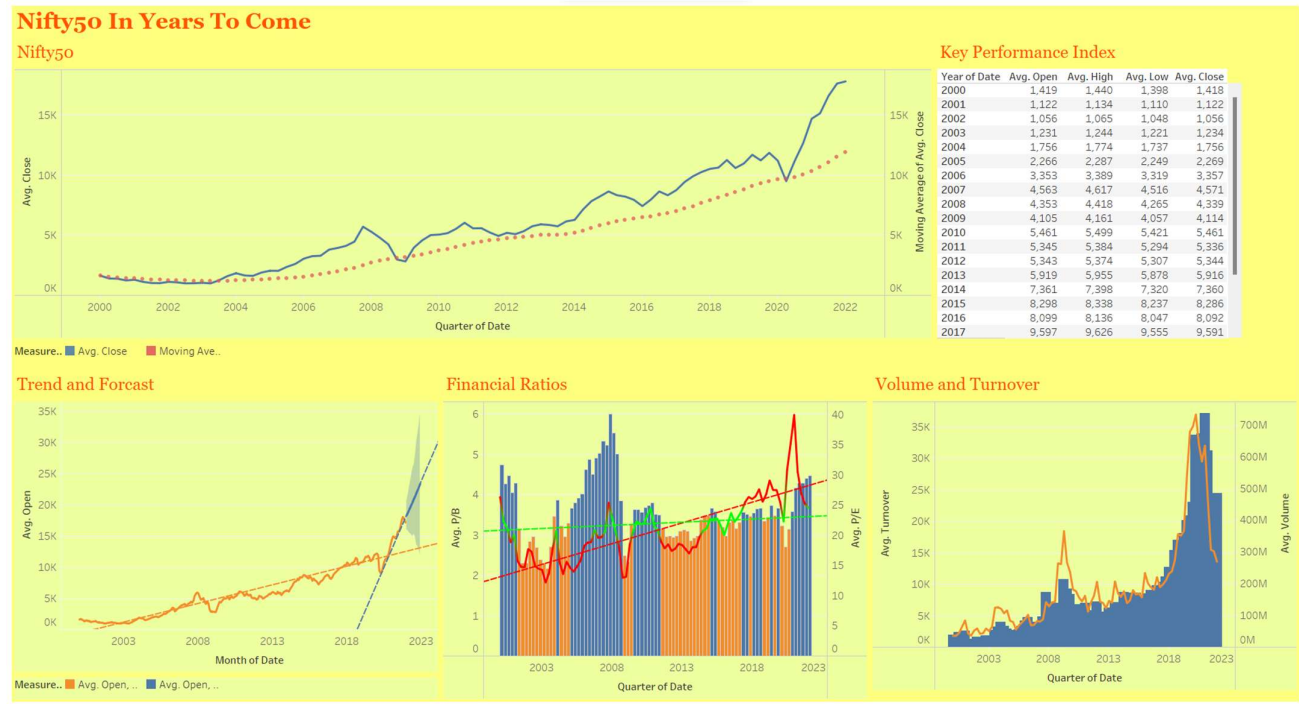
Volume and Turnover:



Volume is an indicator that means the total number of shares that have been bought or sold in a specific period of time. Turnover is sometimes referred to as ‘gross revenue’ or ‘income’. This is different to profit, which is a measure of earnings. It’s an important measure of performance.

In this plot which shows volume and turnover of Nifty in each quarter indicates and good uptrend meaning Nifty has a very good performance and how well volume and turnover are correlated.

Insights:



From this data analysis and data visualization of Nifty50 dataset of year 2000-2022 we learnt about how its trend was and how Nifty has been performing over the years. We had look at the its financials and turnovers and also made a forecast on the Nifty50 trend.

From this Data Visualization following Insights can be inferred:

1. During the time period of almost 20 years Nifty50 has been in a strong uptrend.
2. The forecast of Nifty shows a very good estimate that the value of Nifty50 can be even higher in the future.
3. The value of P/B in latest quarters have been high so on the short-term view investing in Nifty may not be profitable
4. Turnover in the recent quarters is decreasing. This may infer a fall in Nifty for a short-term.
5. On a long-term view the trend shows a good uptrend and Nifty would be profitable.

Tableau Public Link:

https://public.tableau.com/views/Nifty50_16572993732550/Dashboard1?:language=en-GB&:display_count=n&:origin=viz_share_link

Dataset Link:

<https://drive.google.com/file/d/1fg6Nj2ZR1ZwFncqC3XxDdK4y2wR7POiX/view?usp=sharing>