SVKM'S NMIMS
Deemed to be UNIVERSITY

SUNANDAN DIVATIA
SCHOOL OF SCIENCE

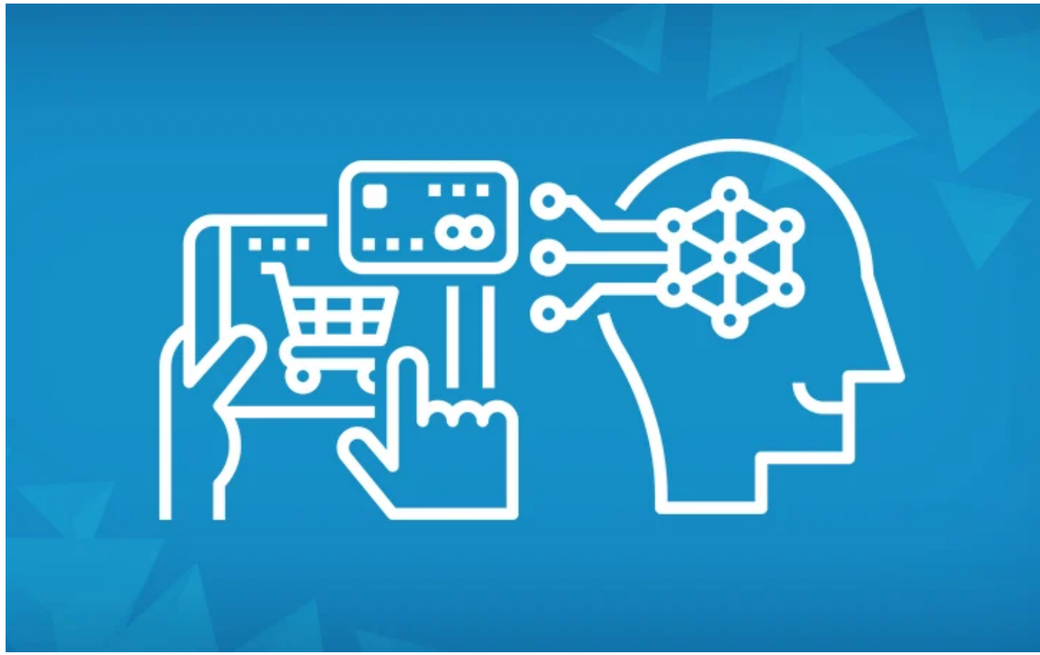
Stock Recommendation System

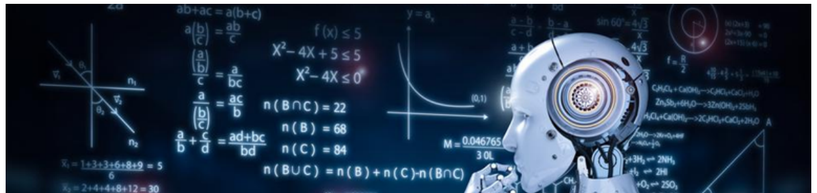
Project By:
Nikita Pathak - A003,
Amaan Multani - A005,
Kanika Malkani - A009

Contents

Introduction

- Information filtering technique
- Provides users with information which they may be interested in.
- Examples : Netflix, LinkedIn, Spotify
- Systems that are designed to recommend things to the user based on many different factors.
- These systems recommend products to the user they are most likely to purchase.





Objectives

01.

Conduct thorough research to build portfolios for people who have just started investing.

02.

Recommend a user-defined number of stocks according to the inputted stock.

Types of Recommendation Systems

Collaborative Filtering

- Collaborative filtering uses similarities between users and items simultaneously to provide recommendations.
- Collaborative filtering models can recommend an item to user A based on the interests of a similar user B.
- For example, Netflix, Hotstar, Spotify and many more platforms use this system to recommend content to the users.

Content-Based Filtering

- Content-based filtering uses item features to recommend other items similar to what the user likes, based on their previous actions or explicit feedback.
- In this, items are ranked according to their relevancy and the most relevant ones are recommended to the user.
- For example, if a user listens to rock music every day, his Youtube recommendation feed will get full of rock music and music of related genres.

The Indian Stock Market

- Aggregation of buyers and sellers of stocks (also called shares),
- Represent ownership claims on businesses
- Secondary market where buying and selling between investors takes place. (Trading)
- Need an exchange platform for the trading to take place. The two exchanges in India are:
    - NSE (National Stock Exchange)
    - BSE (Bombay Stock Exchange)

BSE – Bombay Stock Exchange

- Dalal Street Mumbai
- Approximately 5439 stocks in the market,
- 8th largest stock exchange
- With an overall market capitalization in the world of more than ₹276.713 lakh crore
- Account for only around 4% of the Indian economy

Recommendation System

Dataset

- 4186 Rows
- 19 Columns

Size
(4186, 19)

df - DataFrame

Index	S.No.	Name	CMP Rs.	P/E	Ind PE	Div Yld%	Debt/Eq	EPS 12M Rs.	Profit q
0	1	R J Bio-Tech	7.5	nan	21.21	0	nan	-4.28	69.3
1	2	Forbes & Co	688.4	0.19	33.94	0	3.32	3392.22	8231
2	3	Nexus Surgical	14.4	18.33	32.87	0	0	0.79	530
3	4	SBEC Systems	9.43	6.99	32.87	0	nan	nan	542.
4	5	Sri Lak.Sar.Arni	51.8	11.2	7.73	0	nan	4.62	-48.
5	6	Baroda Rayon	266.05	1.68	9.61	0	11.81	163.37	420.
6	7	Standard Inds.	29.8	1.03	32.87	8.39	0.18	28.87	3245
7	8	EKI Energy	1699.15	10.28	24.98	0.29	0	165.31	1955
8	9	Dynavision	185	14.5	24.98	0	0	12.76	-3.3

Ratios used in the data set.

P/E = is the ratio for valuing a company that measures its current share price relative to its per-share earnings.

Ind P/E = jst like above but of the overall industry to compare

Div Yld = percentage of the share price given as dividend

Debt/equity = parameter to measure companies debt

Profit grwth % = percentage of the growth of profit annually

Ratios used in the data set.

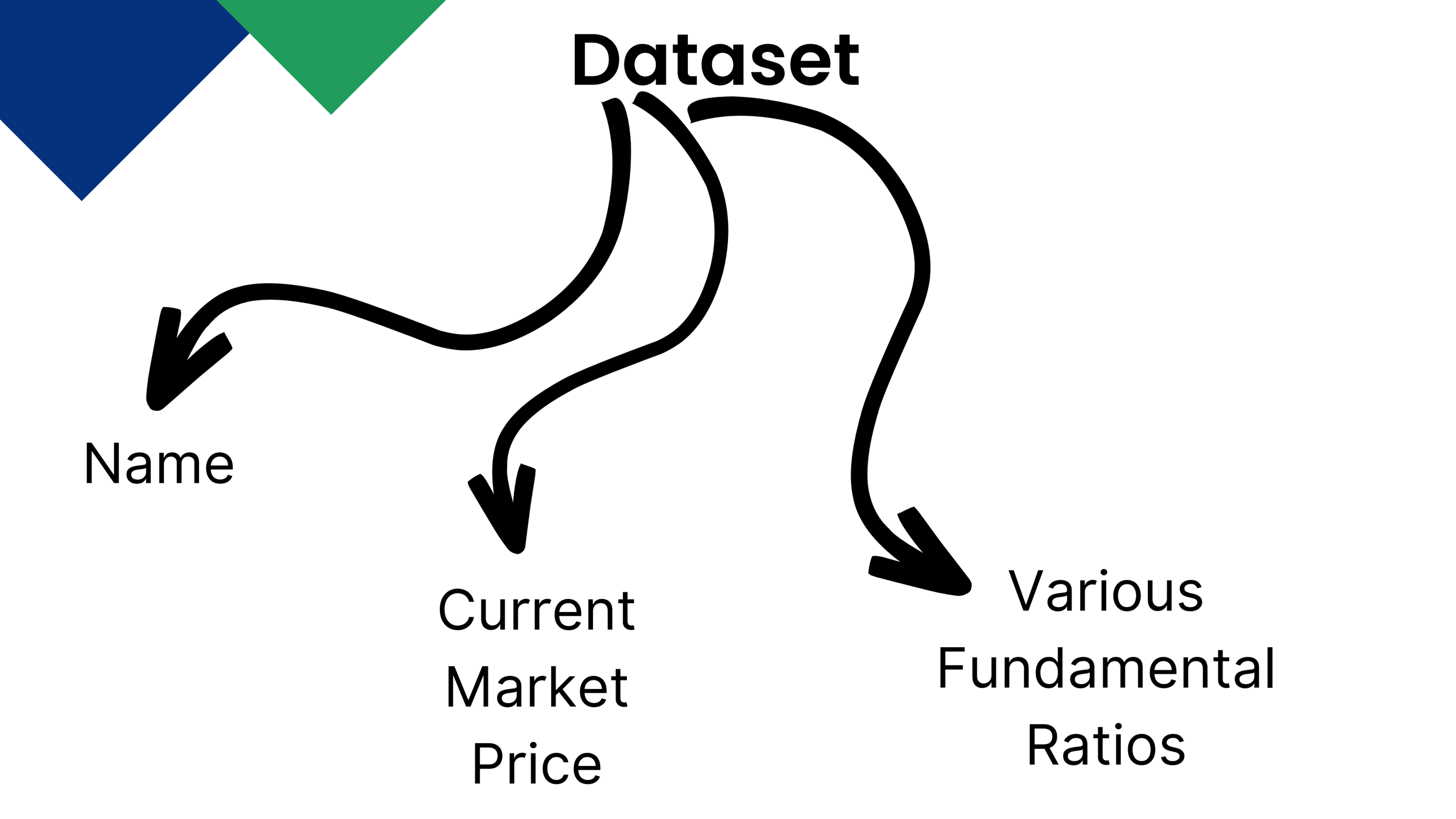
EV/EBITDA = valuation tool to compare the value of a company, debt included

sales var 10yrs= var in sales of past 10yrs

Profit var 10yrs = var in profit of past 10 yrs

ROE = return on equity

ROCE= return on capital employed.


Dataset
Name
Current Market Price
Various Fundamental Ratios

Web Scraping

- Automatic method to obtain large amounts of data from websites
- Data is unstructured data in an HTML format
- Converted into structured data in a spreadsheet or a database
- Online services, particular API’s or even creating your code for web scraping from scratch.
- Moneycontrol
- Financial Ratios
- Library used: Beautiful Soup

Data Pre-processing

- Importing Dataset
- Checking for Null Values
- Imputing Null Values

```
In [18]: print("% of NULL : ",df.isnull().sum()/len(df)*100)
% of NULL :  S.No.                    0.000000
Name                      0.000000
CMP Rs.                   0.000000
P/E                       0.000000
Ind PE                    0.000000
Div Yld%                  0.430005
Debt/Eq                   9.268992
EPS 12M Rs.               0.477783
Profit growth%            1.505017
PEG                       0.000000
EV/EBITDA                 0.453894
Sales Var 10Yrs%          0.000000
EBIDT Ann Rs.Cr.          0.023889
Profit Var 10Yrs%         0.000000
Avg PAT 10Yrs Rs.Cr.     22.240803
ROE%                     10.033445
Prom. Hold.%              0.477783
Sales Var 3Yrs%          11.657907
ROCE%                     3.702819
dtype: float64
```

Before

```
In [20]: print("% of NULL : ",df.isnull().sum()/len(df)*100)
% of NULL :  S.No.                    0.0
Name                      0.0
CMP Rs.                   0.0
P/E                       0.0
Ind PE                    0.0
Div Yld%                  0.0
Debt/Eq                   0.0
EPS 12M Rs.               0.0
Profit growth%            0.0
PEG                       0.0
EV/EBITDA                 0.0
Sales Var 10Yrs%          0.0
EBIDT Ann Rs.Cr.          0.0
Profit Var 10Yrs%         0.0
Avg PAT 10Yrs Rs.Cr.      0.0
ROE%                      0.0
Prom. Hold.%              0.0
Sales Var 3Yrs%           0.0
ROCE%                     0.0
dtype: float64
```

After

Content Based Recommendation System

Cosine Similarity

- It is a metric, helpful in determining, how similar the data objects are irrespective of their size

- Data objects in a dataset are treated as a vector.

- It's formula is:
  Cos(x, y) = x . y / ||x|| * ||y||
  where,
- x . y = product (dot) of the vectors ‘x’ and ‘y’.
- ||x|| and ||y|| = length of the two vectors ‘x’ and ‘y’.
- ||x|| * ||y|| = cross product of the two vectors ‘x’ and ‘y’.

Content Based Recommendation System

Cosine Similarity

Stock Name	R J Bio-Tech	Forbes & Co	Nexus Surgical	SBEC Systems	Sri Lak.Sar.Arni
R J Bio-Tech	1	0.143105	0.758555	0.902039	0.99084
Forbes & Co	0.143105	1	0.379515	0.472768	0.076129
Nexus Surgical	0.758555	0.379515	1	0.822166	0.718931

Content Based Recommendation System

Cosine Similarity

USER A

HDFC Bank Ltd.
Sector: Banks - Private Sector

Tata Consultancy Services Ltd.
Sector: Computers - Software

Cipla Ltd.
Sector: Pharmaceuticals

Mahindra and Mahindra Ltd.
Sector: Auto - Cars & Jeeps

Content Based Recommendation System

Cosine Similarity

USER A

Which Stock would you like to pick?

```
Which stock would you like to pick?
HDFC Bank
```


HDFC Bank
Stored as variable "name"

Content Based Recommendation System

Cosine Similarity

USER A

How many number of similar stocks do you want to find?

```
How many number of similar stocks do you want to find?
10
```


10
Stored as variable "n"

Output

Cosine Values for 1st 10 stocks in the dataset

Stock Name	HDFC Bank
R J Bio-Tech	0.051516
Forbes & Co	0.112999
Nexus Surgical	0.076263
SBEC Systems	0.184029
Sri Lak.Sar.Arni	0.082247
Baroda Rayon	0.207064
Standard Inds.	0.083489
EKI Energy	0.259046
Dynavision	0.327087
Ghosts	0.685

Index	HDFC Bank
0	0.051516
1	0.112999
2	0.076263
3	0.184029
4	0.082247
5	0.207064

Rearranging this in descending order

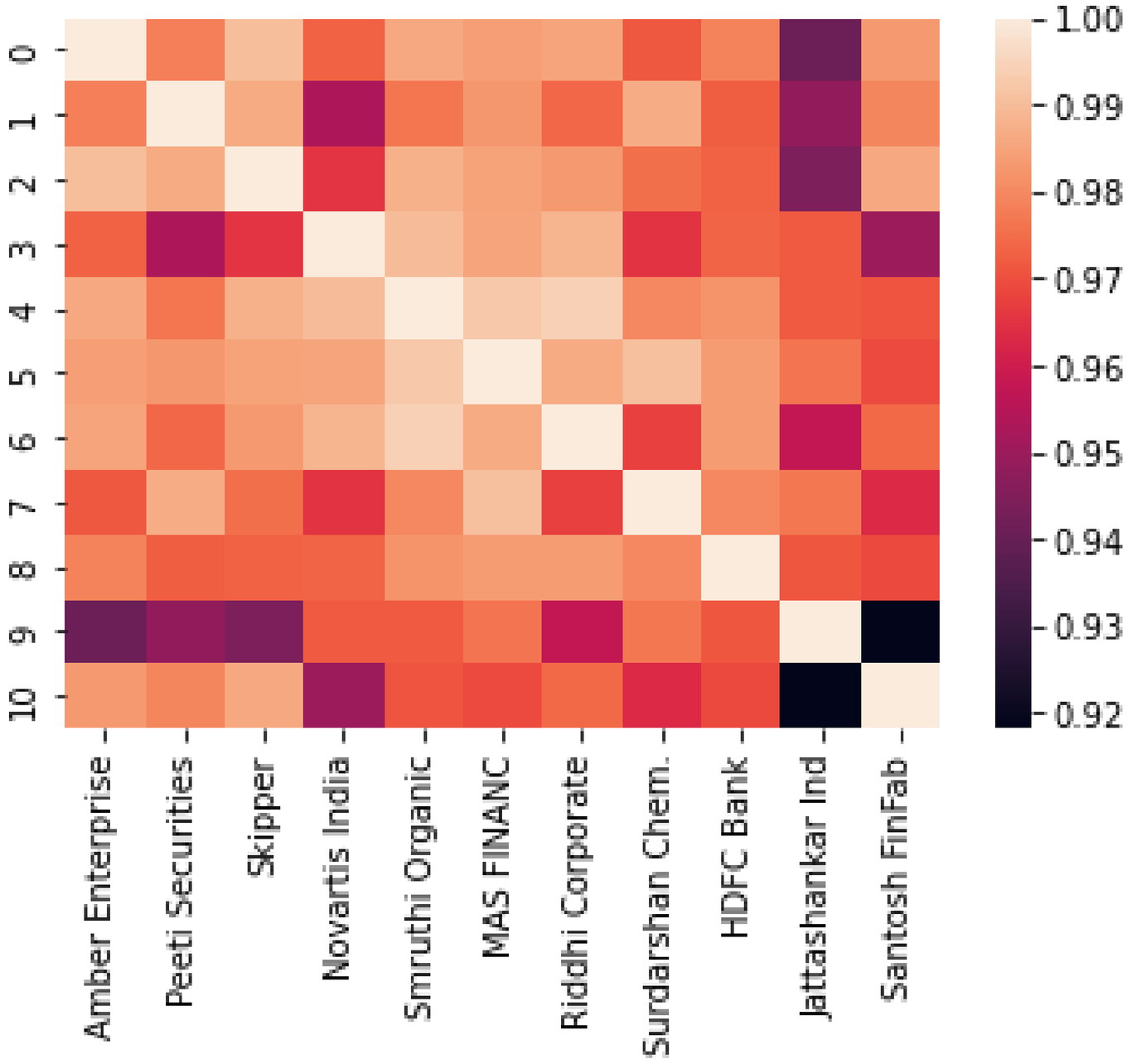
Index	HDFC Bank
2356	1
1990	0.98352
1782	0.983427
1762	0.98184
2323	0.979711
760	0.97884

Output

```
for i in n_similar:
    print(df._get_value(i, 'Name'))
```

Stock Name
Amber Enterp.
Peeti Securities
Skipper
Novartis India
Smruthi Organic
MAS FINANC SER
Riddhi Corporate
Sudarshan Chem.
Jattashankar Ind
Santosh Fin Fab

Output


0
1
2
3
4
5
6
7
8
9
10
Amber Enterprise
Peeti Securities
Skipper
Novartis India
Smruthi Organic
MAS FINANC
Riddhi Corporate
Surdarshan Chem.
HDFC Bank
Jattashankar Ind
Santosh FinFab
1.00
0.99
0.98
0.97
0.96
0.95
0.94
0.93
0.92

Conclusion

- Successfully recommended stocks using cosine similarity
- Web Scraping to obtain the dataset
- 'n' number of similar stocks can be invested in based on fundamental analysis

Limitations

- Industry Wise segregation
- Lack of Accuracy

Future Scope

- Creation of a Bot to filter stocks industry wise and recommend them
- Stock Price prediction

Thank You!