Wireframe Document

Investment Analytics

Revision Number - 1.2
Last Date of Revision - 16/02/2023

Siddharth Jain

Document Control

Date	Version	Description	Author
10/02/2023	1.0	Introduction, Problem Statement	Siddharth Jain
11/02/2023	1.1	Dataset Information, Architecture Description	Siddharth Jain
16/02/2023	1.2	Final Revision	Siddharth Jain

Introduction& Problem Statement

Investment analysts are financial professionals who evaluate securities, stocks, bonds, and other financial assets. They identify potential investments for purchasing, research business financials, and advise clients. They also contribute to financial decision-making through in-depth analyses of market trends and performance. Investment is a game of understanding historic data of investment objects under different events but it is still a game of chances to minimize the risk we apply analytics to find the equilibrium investment. To understand the Foreign direct investment in India for the last 17 years from 2000-01 to 2016-17. This dataset contains sector and financial year-wise data of FDI in India

- Sector-wise investment analysis
- Year-wise investment analysis
- Find key metrics and factors and show the meaningful relationships between attributes.
- Do your own research and come up with your findings

Required libraries:
1. Pandas
2. Matplotlib

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

1. Retrieving dataset of FDI

```
#Printing the Top 10 data

df.head(10)
```

✓ 0.1s

Python

	Sector	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17
0	METALLURGICAL INDUSTRIES	22.69	14.14	36.61	8.11	200.38	149.13	169.94	1175.75	959.94	419.88	1098.14	1786.14	1466.23	567.63	359.34	456.31	1440.18
1	MINING	1.32	6.52	10.06	23.48	9.92	7.40	6.62	444.36	34.16	174.40	79.51	142.65	57.89	12.73	684.39	520.67	55.75
2	POWER	89.42	757.44	59.11	27.09	43.37	72.69	157.15	988.68	907.66	1271.79	1271.77	1652.38	535.68	1066.08	707.04	868.80	1112.98
3	NON-CONVENTIONAL ENERGY	0.00	0.00	1.70	4.14	1.27	1.35	2.44	58.82	125.88	622.52	214.40	452.17	1106.52	414.25	615.95	776.51	783.57
4	COAL PRODUCTION	0.00	0.00	0.00	0.04	0.00	9.14	1.30	14.08	0.22	0.00	0.00	0.00	0.00	2.96	0.00	0.00	0.00
5	PETROLEUM & NATURAL GAS	9.35	211.07	56.78	80.64	102.78	12.09	87.71	1405.04	349.29	265.53	556.43	2029.98	214.80	112.23	1079.02	103.02	180.40
6	BOILERS AND STEAM GENERATING PLANTS	0.00	0.00	0.00	0.04	0.54	0.00	3.31	1.51	0.00	3.96	0.63	31.79	20.05	0.17	1.33	77.91	53.91
7	PRIME MOVER (OTHER THAN ELECTRICAL GENERATORS)	0.00	0.00	0.00	0.00	2.66	0.74	25.57	40.53	74.88	39.50	166.44	313.75	184.60	212.78	230.70	159.13	286.88
8	ELECTRICAL EQUIPMENTS	79.76	65.76	34.71	73.20	97.40	39.50	76.85	653.74	417.35	728.27	153.90	566.39	195.87	134.31	574.83	444.88	2230.69
9	COMPUTER SOFTWARE & HARDWARE	228.39	419.39	314.24	368.32	527.90	1359.97	2613.33	1382.25	1543.34	871.86	779.81	796.35	485.96	1126.27	2296.04	5904.36	3651.71

2. Descriptive statistics of the Dataset

```
df.describe()
```

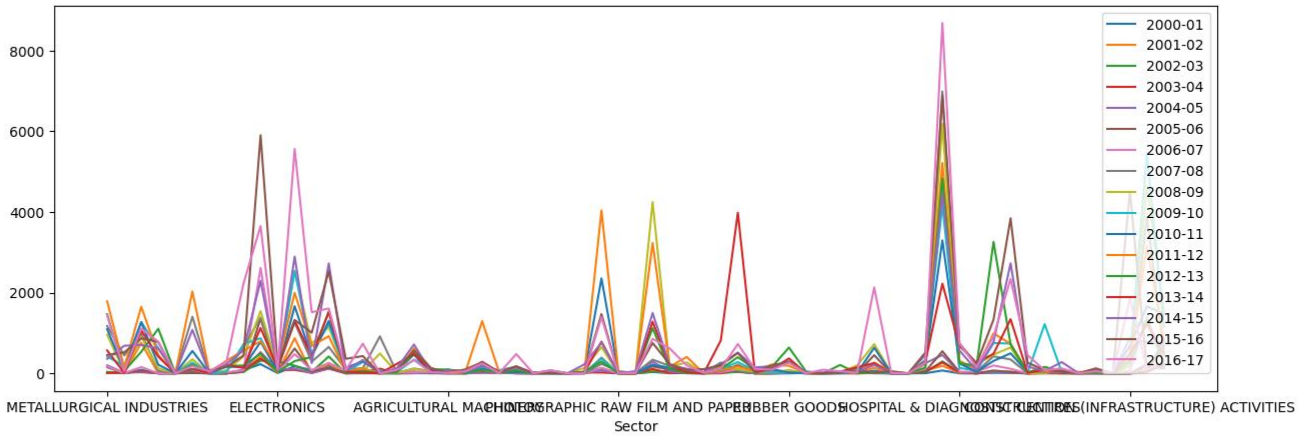
Python

	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-1
count	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.00000
mean	37.757302	63.931587	42.925714	34.727778	51.090317	87.932540	198.281905	390.085714	498.348571	410.069524	339.413810	557.472698	355.930000	385.703492	490.95984
std	112.227860	157.878737	86.606439	67.653735	101.934873	206.436967	686.783115	1026.249935	1134.649040	926.814626	627.141139	1031.474056	778.091368	658.429944	837.78706
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00000
25%	0.000000	0.000000	0.200000	0.215000	0.715000	1.230000	4.160000	9.950000	11.950000	7.880000	8.430000	22.720000	15.115000	16.610000	33.80000
50%	4.030000	5.070000	11.010000	6.370000	9.090000	22.620000	25.820000	58.820000	84.880000	69.740000	58.070000	129.360000	95.410000	113.780000	177.22000
75%	23.510000	44.830000	36.555000	38.660000	43.205000	63.855000	108.325000	279.270000	383.320000	341.595000	304.280000	593.525000	288.025000	473.060000	595.39000
max	832.070000	873.230000	419.960000	368.320000	527.900000	1359.970000	4713.780000	6986.170000	6183.490000	5466.130000	3296.090000	5215.980000	4832.980000	3982.890000	4443.26000

```
#Columns in the file
df.columns
```

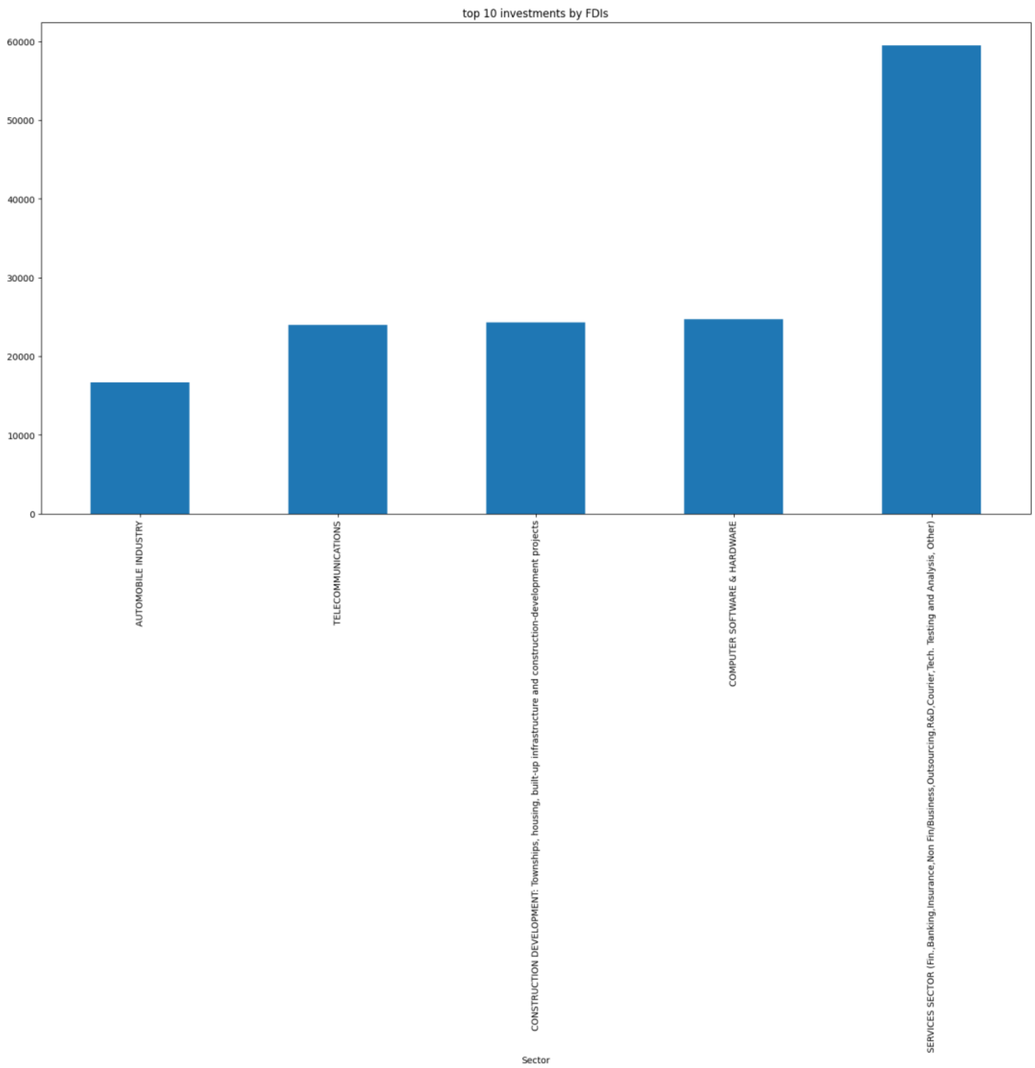
```
Index(['Sector', '2000-01', '2001-02', '2002-03', '2003-04', '2004-05',
       '2005-06', '2006-07', '2007-08', '2008-09', '2009-10', '2010-11',
       '2011-12', '2012-13', '2013-14', '2014-15', '2015-16', '2016-17'],
      dtype='object')
```

3. Visualization of the FDI dataset



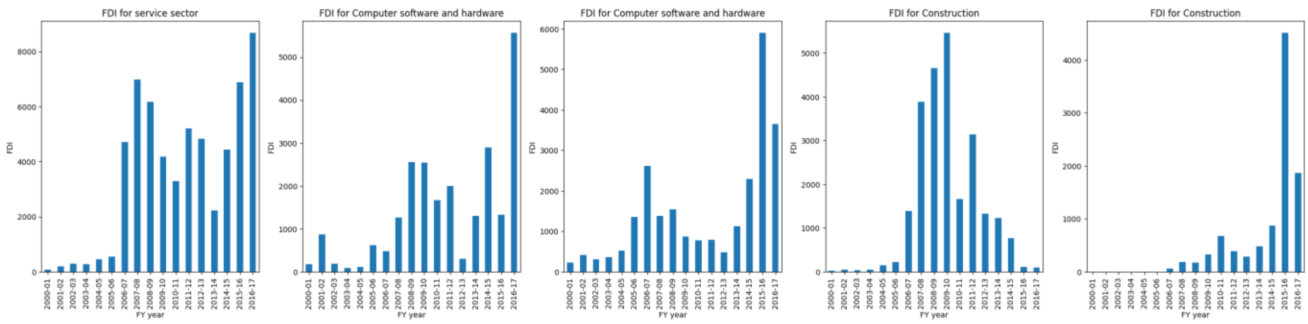
4. Top 5 Investment Sectors by FDI

```
df_trans = df.transpose()
df_trans.sum().sort_values()[-5:].plot(figsize=(20,10),kind='bar',
title ='top 10 investments by FDIs')
<AxesSubplot:title={'center':'top 10 investments by FDIs'},
xlabel='Sector'>
```



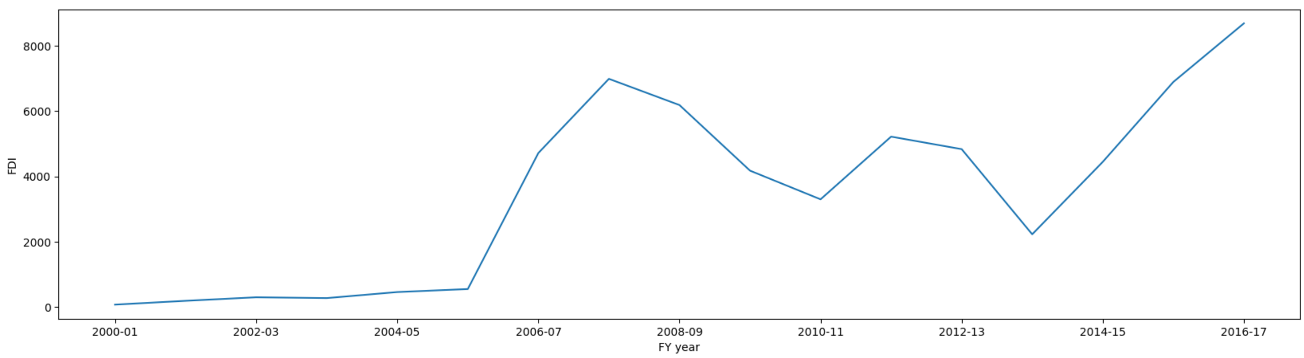
5. Detailed FDI year on year for the top 5 best performing sector

```
plt.subplot(1, 5, 1)
df.loc['SERVICES SECTOR (Fin.,Banking,Insurance,Non
Fin/Business,Outsourcing,R&D,Courier,Tech. Testing and Analysis,
Other)'].plot(kind='bar',figsize=(30,6))
plt.title('FDI for service sector')
plt.xlabel('FY year')
plt.ylabel('FDI')
plt.subplot(1, 5, 2)
df.loc['TELECOMMUNICATIONS'].plot(kind='bar',title='FDI for Computer
software and hardware',figsize=(40,5)) plt.xlabel('FY year')
plt.ylabel('FDI') plt.subplot(1, 5, 3)
df.loc['COMPUTER SOFTWARE & HARDWARE'].plot(kind='bar',title='FDI for
Computer software and hardware',figsize=(40,5))
plt.xlabel('FY year') plt.ylabel('FDI')
plt.subplot(1, 5, 4)
df.loc['CONSTRUCTION DEVELOPMENT: Townships, housing, built-up
infrastructure and construction-development
projects'].plot(kind='bar',title='FDI for
Construction',figsize=(30,6))
plt.xlabel('FY year')
plt.ylabel('FDI')
plt.subplot(1, 5,5)
df.loc['CONSTRUCTION (INFRASTRUCTURE)
ACTIVITIES'].plot(kind='bar',title='FDI for
Construction',figsize=(30,6))
plt.xlabel('FY year')
plt.ylabel('FDI')
Text(0, 0.5, 'FDI')
```



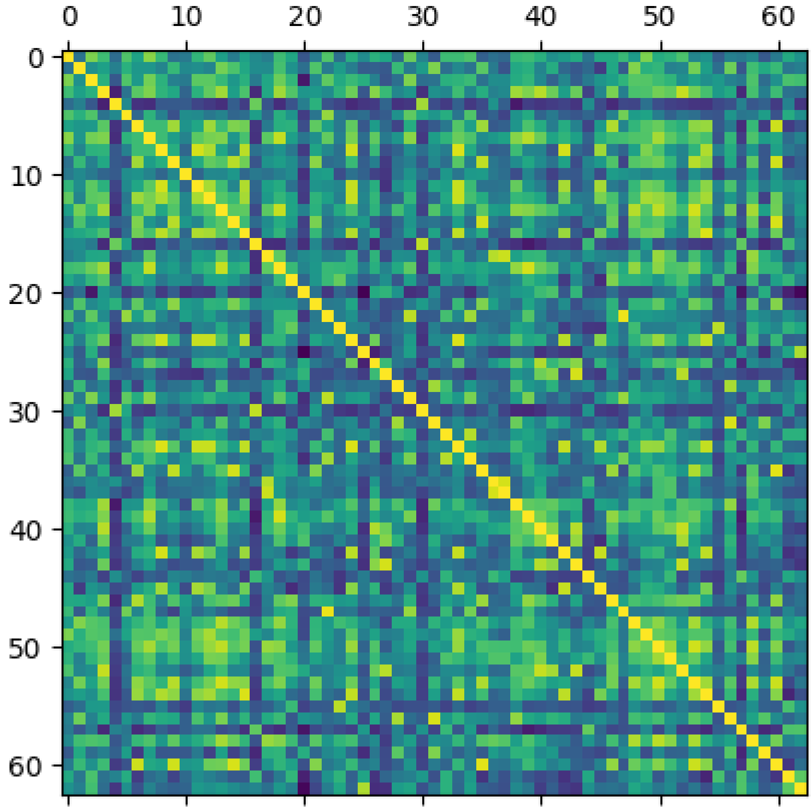
6. FDI overall growth of the best performing Sector from FY 2000-2016

```
df.loc['SERVICES SECTOR (Fin.,Banking,Insurance,Non
Fin/Business,Outsourcing,R&D,Courier,Tech. Testing and Analysis,
Other)'].plot(figsize=(20,5))
plt.xlabel('FY year')
plt.ylabel('FDI') plt.show()
```



7. Visualizing the correlation among sectors

```
corr = df_trans.corr()
plt.matshow(corr)
<matplotlib.image.AxesImage at 0x7fde4962a700>
```

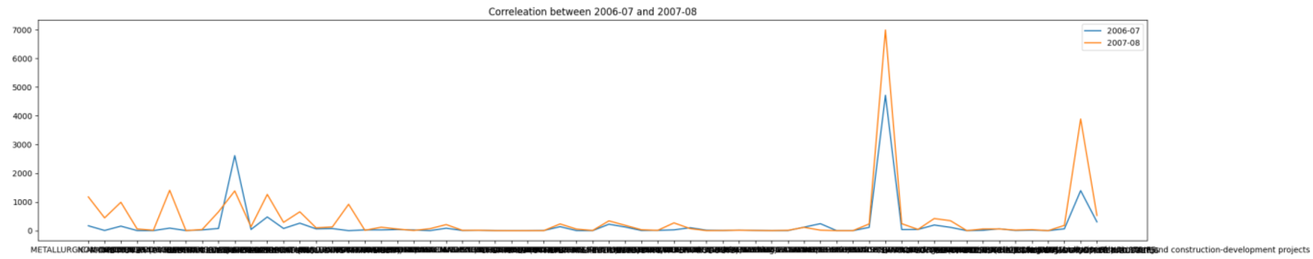


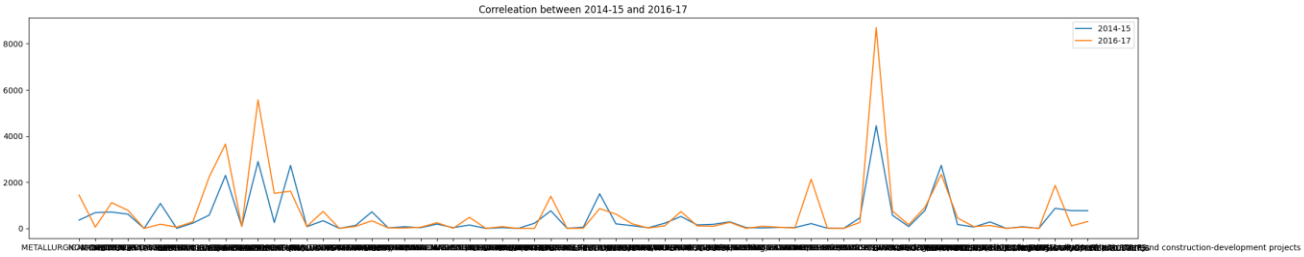
8. Top 10 correlated sectors

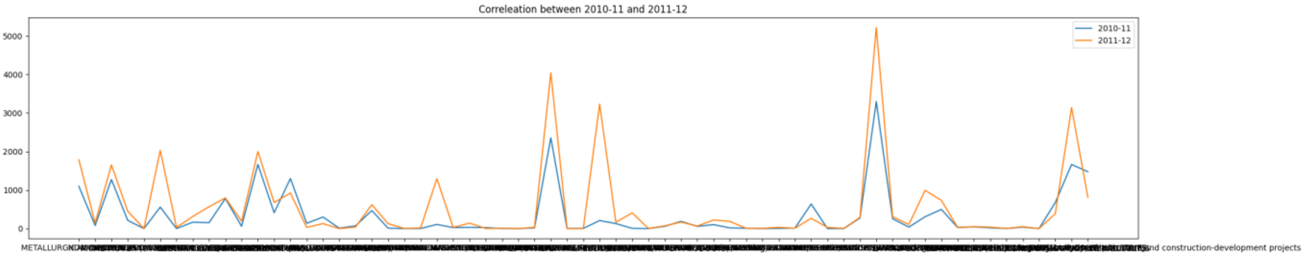
```
Top Absolute Correlations

Sector                                              Sector
MISCELLANEOUS MECHANICAL & ENGINEERING INDUSTRIES   DEFENCE INDUSTRIES                            0.958449
SUGAR                                               CONSTRUCTION (INFRASTRUCTURE) ACTIVITIES      0.937258
ELECTRICAL EQUIPMENTS                               TEXTILES (INCLUDING DYED,PRINTED)             0.926705
MEDICAL AND SURGICAL APPLIANCES                     TEXTILES (INCLUDING DYED,PRINTED)             0.919642
SEA TRANSPORT                                       RETAIL TRADING                                0.918936
DYE-STUFFS                                          DIAMOND,GOLD ORNAMENTS                        0.916723
AIR TRANSPORT (INCLUDING AIR FREIGHT)               CONSTRUCTION (INFRASTRUCTURE) ACTIVITIES      0.916622
FERMENTATION INDUSTRIES                             FOOD PROCESSING INDUSTRIES                    0.910990
ELECTRICAL EQUIPMENTS                               GLUE AND GELATIN                              0.908833
MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS      GLASS                                         0.908687
dtype: float64
```

9. 3 most correlated years




Correleation between 2014-15 and 2016-17
2014-15
2016-17
8000
6000
4000
2000
0
and construction-development projects


Correleation between 2010-11 and 2011-12
2010-11
2011-12
5000
4000
3000
2000
1000
0
and construction-development projects