

```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

import plotly.offline as pyo
import plotly.express as px
```

```
In [2]: df = pd.read_csv('FDI data.csv')
```

```
In [3]: df.head()
```

Out[3]:

	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

```
In [4]: df.describe()
```

Out[4]:

	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12
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
mean	37.757302	63.931587	42.925714	34.727778	51.090317	87.932540	198.281905	390.085714	498.348571	410.069524	339.413810	557.472691
std	112.227860	157.878737	86.606439	67.653735	101.934873	206.436967	686.783115	1026.249935	1134.649040	926.814626	627.141139	1031.474051
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
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
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
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
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

```
In [5]: len(df)
```

Out[5]: 63

```
In [6]: df.isnull().sum()
```

Out[6]:

```
Sector      0
2000-01     0
2001-02     0
2002-03     0
2003-04     0
2004-05     0
2005-06     0
2006-07     0
2007-08     0
2008-09     0
2009-10     0
2010-11     0
2011-12     0
2012-13     0
2013-14     0
2014-15     0
2015-16     0
2016-17     0
dtype: int64
```

```
In [7]: df.columns[1:]
```

Out[7]:

```
Index(['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')
```

In [8]:

```
Year = ['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']
Sectors = ['Sector']
```

In [9]:

```
FDI_USD = df.copy()
FDI_USD.head()
```

Out[9]:

	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

Unpivoting DataFrames from wide to long format:

In [10]:

```
melt = pd.melt(FDI_USD, id_vars = Sectors, value_vars = Year, var_name='Year',value_name='FDI (USD Million)',
               ignore_index=True)
melt
```

Out[10]:

	Sector	Year	FDI (USD Million)
0	METALLURGICAL INDUSTRIES	2000-01	22.69
1	MINING	2000-01	1.32
2	POWER	2000-01	89.42
3	NON-CONVENTIONAL ENERGY	2000-01	0.00
4	COAL PRODUCTION	2000-01	0.00
...
1066	PRINTING OF BOOKS (INCLUDING LITHO PRINTING IN...	2016-17	53.17
1067	COIR	2016-17	0.00
1068	CONSTRUCTION (INFRASTRUCTURE) ACTIVITIES	2016-17	1860.73
1069	CONSTRUCTION DEVELOPMENT: Townships, housing, ...	2016-17	105.14
1070	MISCELLANEOUS INDUSTRIES	2016-17	296.40

1071 rows × 3 columns

In [11]:

```
sorted_FDI = melt.sort_values(['Sector','Year'], ignore_index = True)
sorted_FDI
```

Out[11]:

	Sector	Year	FDI (USD Million)
0	AGRICULTURAL MACHINERY	2000-01	3.64
1	AGRICULTURAL MACHINERY	2001-02	1.04
2	AGRICULTURAL MACHINERY	2002-03	13.48
3	AGRICULTURAL MACHINERY	2003-04	47.54
4	AGRICULTURAL MACHINERY	2004-05	0.00
...
1066	VEGETABLE OILS AND VANASPATI	2012-13	108.39
1067	VEGETABLE OILS AND VANASPATI	2013-14	21.55
1068	VEGETABLE OILS AND VANASPATI	2014-15	148.34
1069	VEGETABLE OILS AND VANASPATI	2015-16	34.22
1070	VEGETABLE OILS AND VANASPATI	2016-17	108.45

1071 rows × 3 columns

Converting longe sector name into Sort name

```
In [12]: Sorted = sorted_FDI[['Sector','Year','FDI (USD Million)'],
                             ,]].replace(["CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects",
                             ,"SERVICES SECTOR (Fin.,Banking,Insurance,Non Fin/Business,Outsourcing,R&D,Courier,Tech. Testing and Analysis)",
                             ,"TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)"],
                             ,["CONSTRUCTION DEVELOPMENT","SERVICES SECTOR","TEA AND COFFEE"])
```

Sector wise total FDI 2000 - 2017

```
In [13]: sectorWise = Sorted.groupby('Sector').sum().sort_values(by = 'FDI (USD Million)',ascending = False)
sectorWise
```

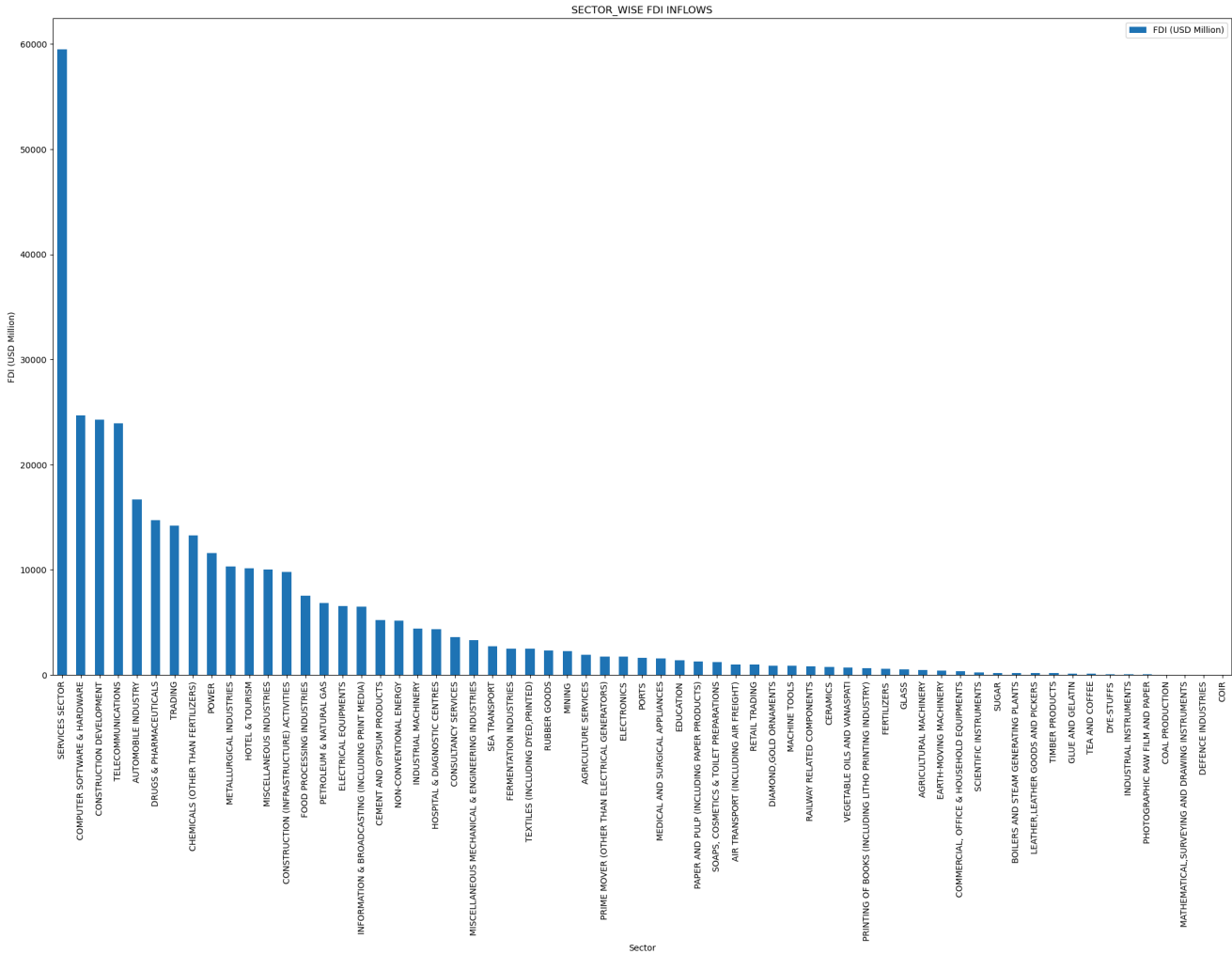
Out[13]:

	Year	FDI (USD Million)
Sector		
SERVICES SECTOR	2000-012001-022002-032003-042004-052005-062006...	59476.49
COMPUTER SOFTWARE & HARDWARE	2000-012001-022002-032003-042004-052005-062006...	24669.49
CONSTRUCTION DEVELOPMENT	2000-012001-022002-032003-042004-052005-062006...	24293.09
TELECOMMUNICATIONS	2000-012001-022002-032003-042004-052005-062006...	23946.01
AUTOMOBILE INDUSTRY	2000-012001-022002-032003-042004-052005-062006...	16673.92
...
PHOTOGRAPHIC RAW FILM AND PAPER	2000-012001-022002-032003-042004-052005-062006...	67.28
COAL PRODUCTION	2000-012001-022002-032003-042004-052005-062006...	27.74
MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS	2000-012001-022002-032003-042004-052005-062006...	7.98
DEFENCE INDUSTRIES	2000-012001-022002-032003-042004-052005-062006...	5.12
COIR	2000-012001-022002-032003-042004-052005-062006...	4.06

63 rows × 2 columns

Sector wise total FDI 2000 - 2017

```
In [14]: sectorWise.plot(kind='bar',y='FDI (USD Million)',figsize = (25,14), legend= True, title='SECTOR_WISE FDI INFLOWS',ylabel='FDI (USD M
Out[14]: <Axes: title={ 'center': 'SECTOR_WISE FDI INFLOWS'}, xlabel='Sector', ylabel='FDI (USD Million)'>
```



Best and Worst Performing Sectors

Share among TOP 10 Sectors in FDI

```
In [15]: Top_10_Sectors = sectorWise.nlargest(10,['FDI (USD Million)'])

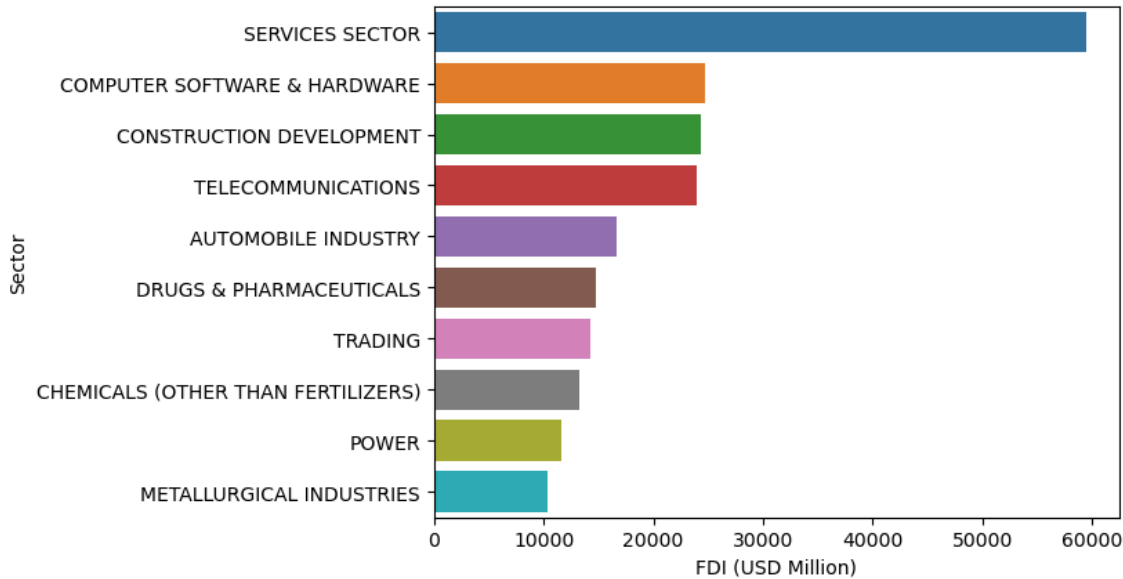
Top_10_Sectors
```

		Year	FDI (USD Million)
Sector			
SERVICES SECTOR	2000-012001-022002-032003-042004-052005-062006...		59476.49
COMPUTER SOFTWARE & HARDWARE	2000-012001-022002-032003-042004-052005-062006...		24669.49
CONSTRUCTION DEVELOPMENT	2000-012001-022002-032003-042004-052005-062006...		24293.09
TELECOMMUNICATIONS	2000-012001-022002-032003-042004-052005-062006...		23946.01
AUTOMOBILE INDUSTRY	2000-012001-022002-032003-042004-052005-062006...		16673.92
DRUGS & PHARMACEUTICALS	2000-012001-022002-032003-042004-052005-062006...		14706.90
TRADING	2000-012001-022002-032003-042004-052005-062006...		14210.88
CHEMICALS (OTHER THAN FERTILIZERS)	2000-012001-022002-032003-042004-052005-062006...		13293.09
POWER	2000-012001-022002-032003-042004-052005-062006...		11589.13
METALLURGICAL INDUSTRIES	2000-012001-022002-032003-042004-052005-062006...		10330.54

Visualization

```
In [16]: sns.barplot(data = Top_10_Sectors, y = Top_10_Sectors.index, x = 'FDI (USD Million)',orient = 'h' )
```

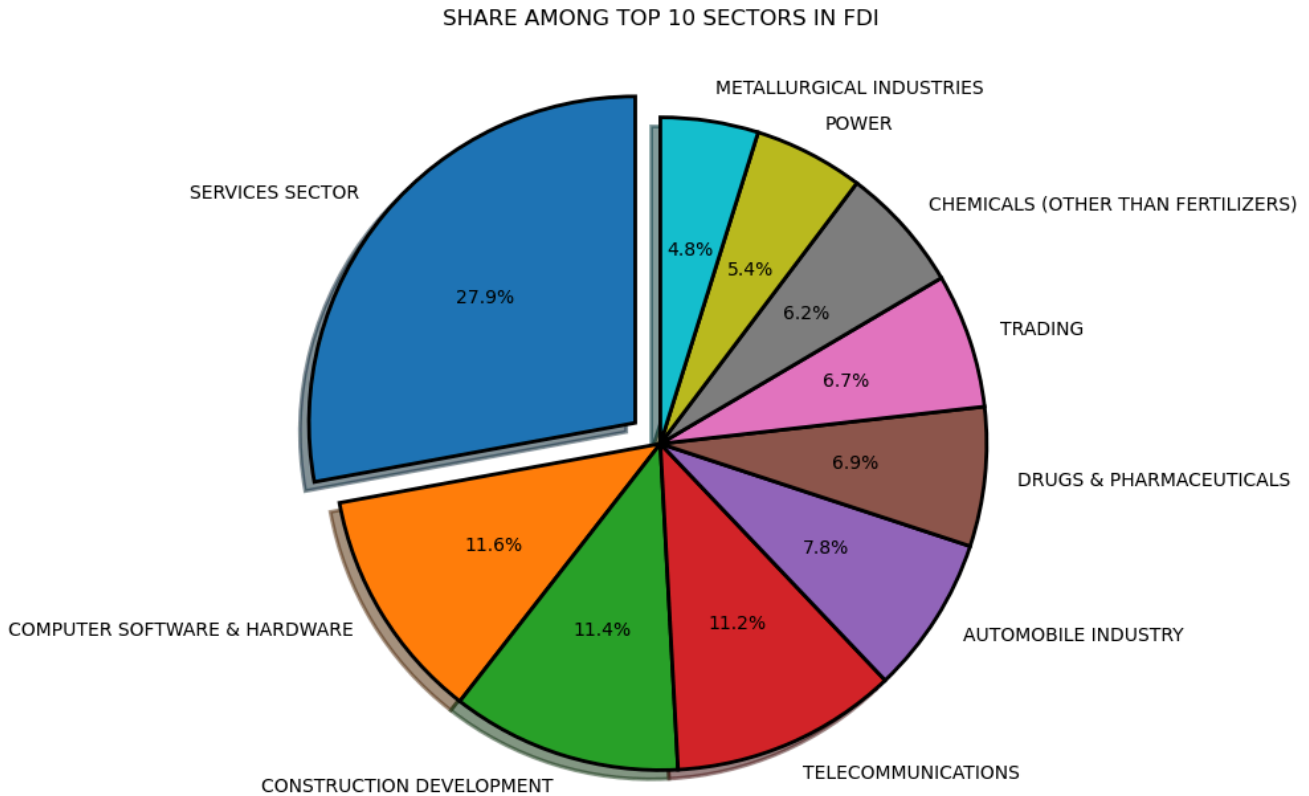
```
plt.show()
```



```
In [17]: explode = [.1,0,0,0,0,0,0,0,0,0]

plt.figure(figsize=(15,8))
plt.pie(Top_10_Sectors['FDI (USD Million)'],labels=Top_10_Sectors.index,autopct='%1.1f%%',explode=explode,shadow = True,startangle=90,
        wedgeprops = {'edgecolor':'black', 'linewidth': 2}
        )

plt.title('SHARE AMONG TOP 10 SECTORS IN FDI')
plt.show()
```



SHARE AMONG LOWEST 5 SECTORS IN FDI

```
In [18]: Bottom_5_Sectors = sectorWise.nsmallest(5,['FDI (USD Million)'])

Bottom_5_Sectors
```

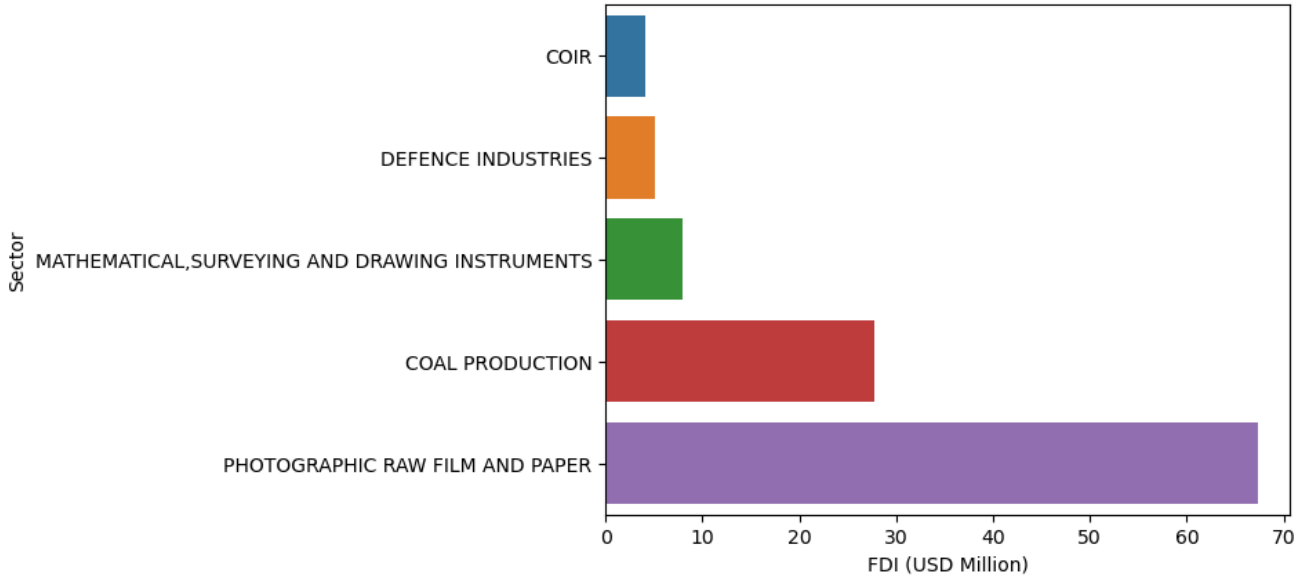
Out[18]:

		Year	FDI (USD Million)
Sector			
	COIR	2000-012001-022002-032003-042004-052005-062006...	4.06
	DEFENCE INDUSTRIES	2000-012001-022002-032003-042004-052005-062006...	5.12
	MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS	2000-012001-022002-032003-042004-052005-062006...	7.98
	COAL PRODUCTION	2000-012001-022002-032003-042004-052005-062006...	27.74
	PHOTOGRAPHIC RAW FILM AND PAPER	2000-012001-022002-032003-042004-052005-062006...	67.28

Visualization

```
In [19]: sns.barplot(data = Bottom_5_Sectors, y = Bottom_5_Sectors.index, x = 'FDI (USD Million)',orient = 'h' )

plt.show()
```

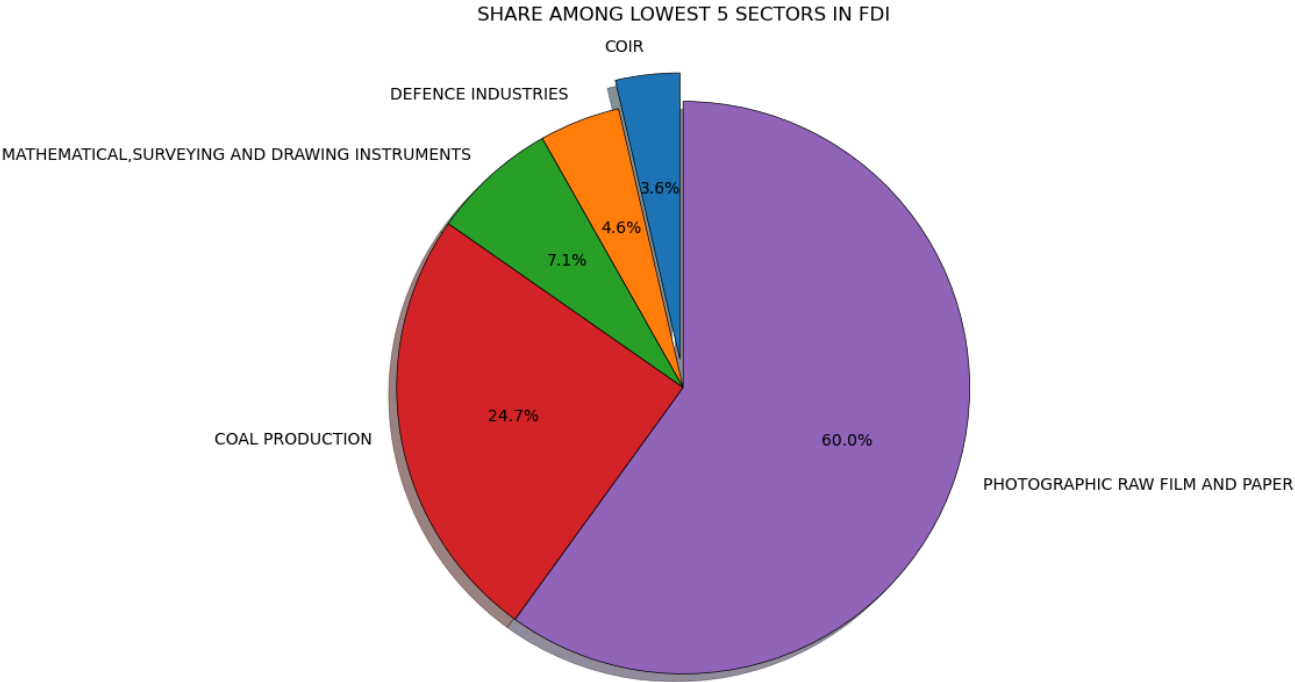


```
In [20]: explode = [.1,0,0,0,0]

plt.figure(figsize=(15,8))
plt.pie(Bottom_5_Sectors['FDI (USD Million)'],labels=Bottom_5_Sectors.index,autopct='%1.1f%%',explode=explode,shadow = True,startang
        wedgeprops = {'edgecolor':'black', 'linewidth': .5}

        )

plt.title('SHARE AMONG LOWEST 5 SECTORS IN FDI')
plt.show()
```



Year Wise Performance

Year wise Inflow

```
In [21]: melt_2 = melt[['Year', 'FDI (USD Million)']]
melt_2=round(melt_2.groupby('Year').sum(),2)

try:
    melt_2['% growth over previous year'] = round(melt_2.pct_change()*100,2)
except:
    melt_2 = np.nan()

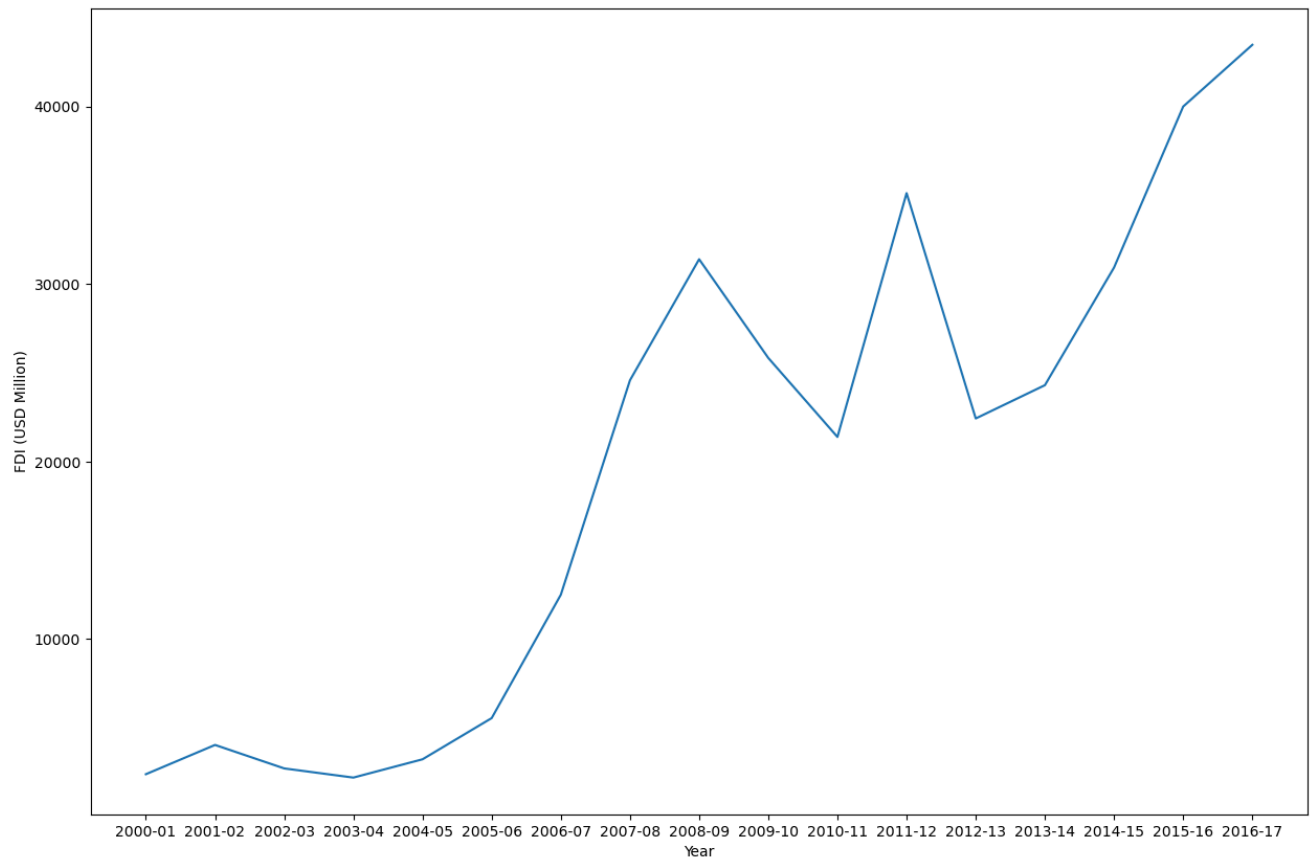
melt_2
```

Out[21]:

	FDI (USD Million)	% growth over previous year
Year		
2000-01	2378.71	NaN
2001-02	4027.69	69.32
2002-03	2704.32	-32.86
2003-04	2187.85	-19.10
2004-05	3218.69	47.12
2005-06	5539.75	72.11
2006-07	12491.76	125.49
2007-08	24575.40	96.73
2008-09	31395.96	27.75
2009-10	25834.38	-17.71
2010-11	21383.07	-17.23
2011-12	35120.78	64.25
2012-13	22423.59	-36.15
2013-14	24299.32	8.36
2014-15	30930.47	27.29
2015-16	40000.99	29.33
2016-17	43478.26	8.69

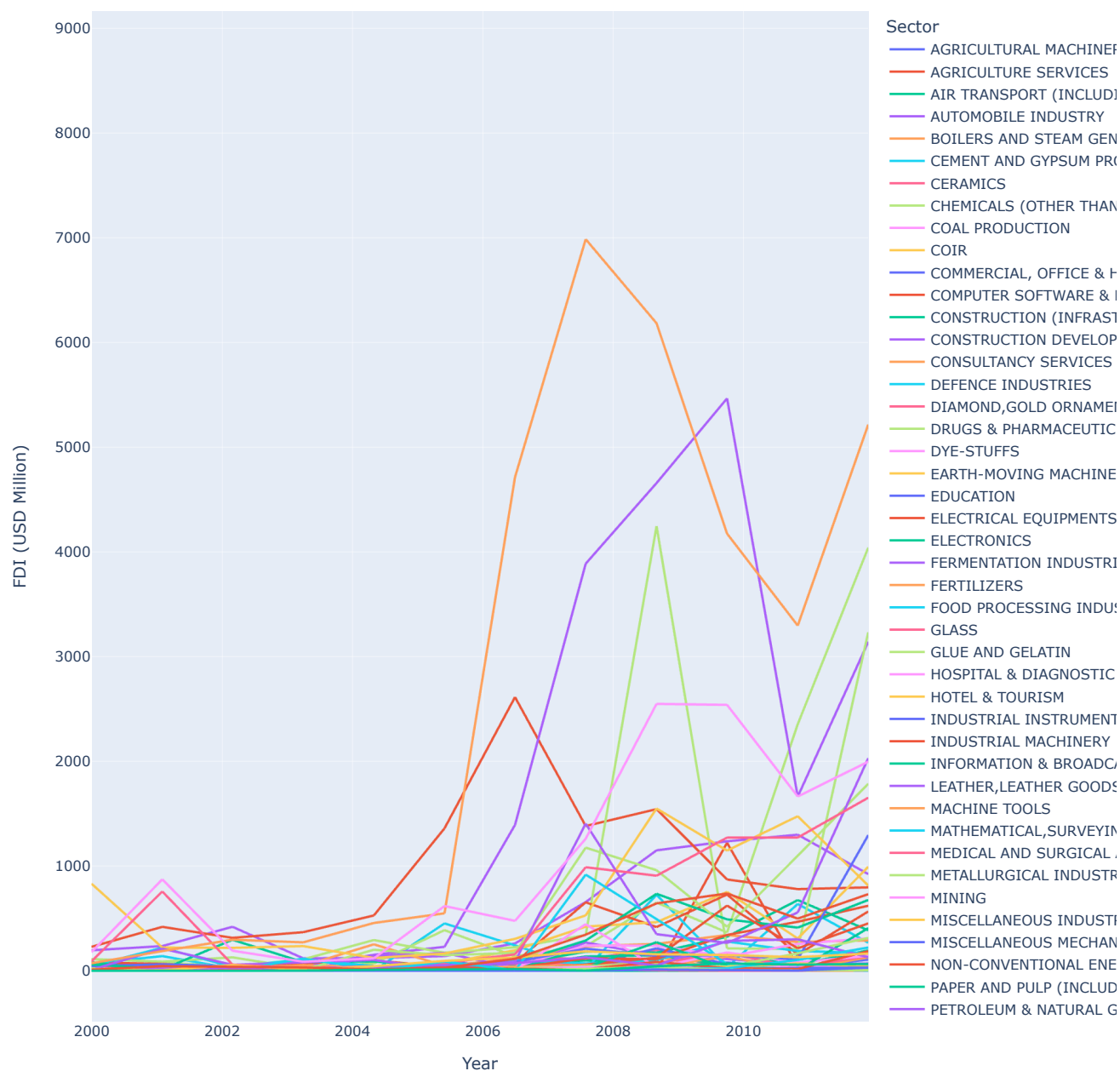
Visualization

```
In [22]: plt.figure(figsize=(15,10))  
  
sns.lineplot(data = melt_2, x = melt_2.index, y = 'FDI (USD Million)')  
  
plt.show()
```



Year wise FDI of each Sector

```
In [23]: fig = px.line(Sorted, x = 'Year', y = 'FDI (USD Million)', color = 'Sector', width=1200,  
                    height=1000, )  
  
fig
```

Sector wise % of Total Inflows for 2000 - 17

In [24]: sectorWise['Year'] = '2000-17'

```
sectorWise = sectorWise[['Year', 'FDI (USD Million)']]
sectorWise['% of Total Inflows'] = round((sectorWise['FDI (USD Million)'] / sectorWise['FDI (USD Million)'].sum()) * 100, 3)
sectorWise
```

Out[24]:

	Year	FDI (USD Million)	% of Total Inflows
Sector			
SERVICES SECTOR	2000-17	59476.49	17.915
COMPUTER SOFTWARE & HARDWARE	2000-17	24669.49	7.431
CONSTRUCTION DEVELOPMENT	2000-17	24293.09	7.317
TELECOMMUNICATIONS	2000-17	23946.01	7.213
AUTOMOBILE INDUSTRY	2000-17	16673.92	5.022
...
PHOTOGRAPHIC RAW FILM AND PAPER	2000-17	67.28	0.020
COAL PRODUCTION	2000-17	27.74	0.008
MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS	2000-17	7.98	0.002
DEFENCE INDUSTRIES	2000-17	5.12	0.002
COIR	2000-17	4.06	0.001

63 rows × 3 columns

```
In [25]: melt.to_csv('FDI-INDIA-2000-2017.csv')

In [ ]:

In [ ]:

In [ ]:
```