Foreign Direct Investment Analysis
Importing required Libraries

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

```python
In [2]: FDI=pd.read_csv(r"C:\Users\kruna\OneDrive\Desktop\internship\FDI\FDI data (1).csv")
        FDI.head()
```

Out[2]:

	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
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
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
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
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
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

```python
In [3]: FDI.columns
```

```
Out[3]: 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')
```

Extracting Detailed Information

```python
In [4]: FDI.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 63 entries, 0 to 62
Data columns (total 18 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   Sector   63 non-null     object
 1   2000-01  63 non-null     float64
 2   2001-02  63 non-null     float64
 3   2002-03  63 non-null     float64
 4   2003-04  63 non-null     float64
 5   2004-05  63 non-null     float64
 6   2005-06  63 non-null     float64
 7   2006-07  63 non-null     float64
 8   2007-08  63 non-null     float64
 9   2008-09  63 non-null     float64
 10  2009-10  63 non-null     float64
 11  2010-11  63 non-null     float64
 12  2011-12  63 non-null     float64
 13  2012-13  63 non-null     float64
 14  2013-14  63 non-null     float64
 15  2014-15  63 non-null     float64
 16  2015-16  63 non-null     float64
 17  2016-17  63 non-null     float64
dtypes: float64(17), object(1)
memory usage: 9.0+ KB
```

Checking the null Values

```python
In [5]: FDI.isnull().sum()
```

```
Out[5]: 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
```

duplicate values in data

```python
In [20]: FDI.duplicated().sum()
```

```
In [20]: FDI.duplicated().sum()
```

```
Out[20]: 0
```

```
In [13]: FDI.columns.isnull()
```

```
Out[13]: array([False, False, False, False, False, False, False, False, False,
                False, False, False, False, False, False, False, False])
```

make Sector & Yearwise investment

```
In [17]: sector = ['Sector']
         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']
```

```
In [18]: FDI.describe().round(2)
```

Out[18]:

	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	201
count	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63
mean	37.76	63.93	42.93	34.73	51.09	87.93	198.28	390.09	498.35	410.07	339.41	557.47	355.93	385.70	490.96	634.
std	112.23	157.88	86.61	67.65	101.93	206.44	686.78	1026.25	1134.65	926.81	627.14	1031.47	778.09	658.43	837.79	1335.
min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
25%	0.00	0.00	0.20	0.22	0.72	1.23	4.16	9.95	11.95	7.88	8.43	22.72	15.12	16.61	33.80	30.
50%	4.03	5.07	11.01	6.37	9.09	22.62	25.82	58.82	84.88	69.74	58.07	129.36	95.41	113.78	177.22	159.
75%	23.51	44.83	36.56	38.66	43.20	63.86	108.32	279.27	383.32	341.60	304.28	593.52	288.02	473.06	595.39	519.
max	832.07	873.23	419.96	368.32	527.90	1359.97	4713.78	6986.17	6183.49	5466.13	3296.09	5215.98	4832.98	3982.89	4443.26	6889.

Sort vlaues in Numerial values in this dataset are in US$ Millions Data is of yearwise investment in all sectors

```
In [23]: FDI_1 = pd.melt(FDI,id_vars = sector, value_vars = year, var_name='year', value_name = 'FDI in US$ Millions')
         FDI_1
```

Out[23]:

	Sector	year	FDI in US$ Millions
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 [24]: sorted = FDI_1.sort_values(['Sector','year'])
         sorted
```

Out[24]:

	Sector	year	FDI in US$ Millions
20	AGRICULTURAL MACHINERY	2000-01	3.64
83	AGRICULTURAL MACHINERY	2001-02	1.04
146	AGRICULTURAL MACHINERY	2002-03	13.48
209	AGRICULTURAL MACHINERY	2003-04	47.54
272	AGRICULTURAL MACHINERY	2004-05	0.00
...	...	...	...
794	VEGETABLE OILS AND VANASPATI	2012-13	108.39
857	VEGETABLE OILS AND VANASPATI	2013-14	21.55
920	VEGETABLE OILS AND VANASPATI	2014-15	148.34
983	VEGETABLE OILS AND VANASPATI	2015-16	34.22
1046	VEGETABLE OILS AND VANASPATI	2016-17	108.45

1071 rows × 3 columns

Description of sectors in new format

```
In [26]: FDI_1.describe().round(2)
```

Out[26]:

	FDI in US$ Millions
count	1071.00
mean	309.98
std	819.04
min	0.00
25%	3.14
50%	37.94
75%	213.74
max	8684.07

Rename sectors

```
In [27]: sorted = sorted.replace(["CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and constructio
                                  ,"SERVICES SECTOR (Fin.,Banking,Insurance,Non Fin/Business,Outsourcing,R&D,Courie
                                  ,'TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)']
                                 ,["CONSTRUCTION DEVELOPMENT","SERVICES SECTOR",'TEA AND COFFEE'])
```

sector wise total FDI from 2000-17

```
In [28]: total_fdi_sector = sorted.groupby('Sector').sum().sort_values(by = 'FDI in US$ Millions',ascending = False)
         total_fdi_sector
```
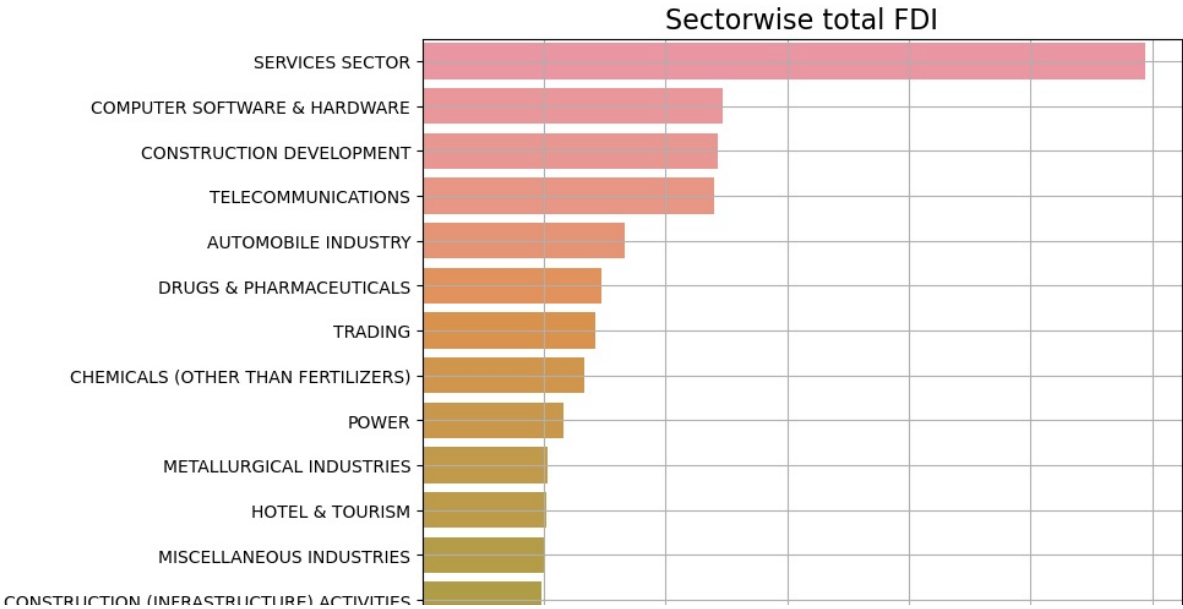
Out[28]:

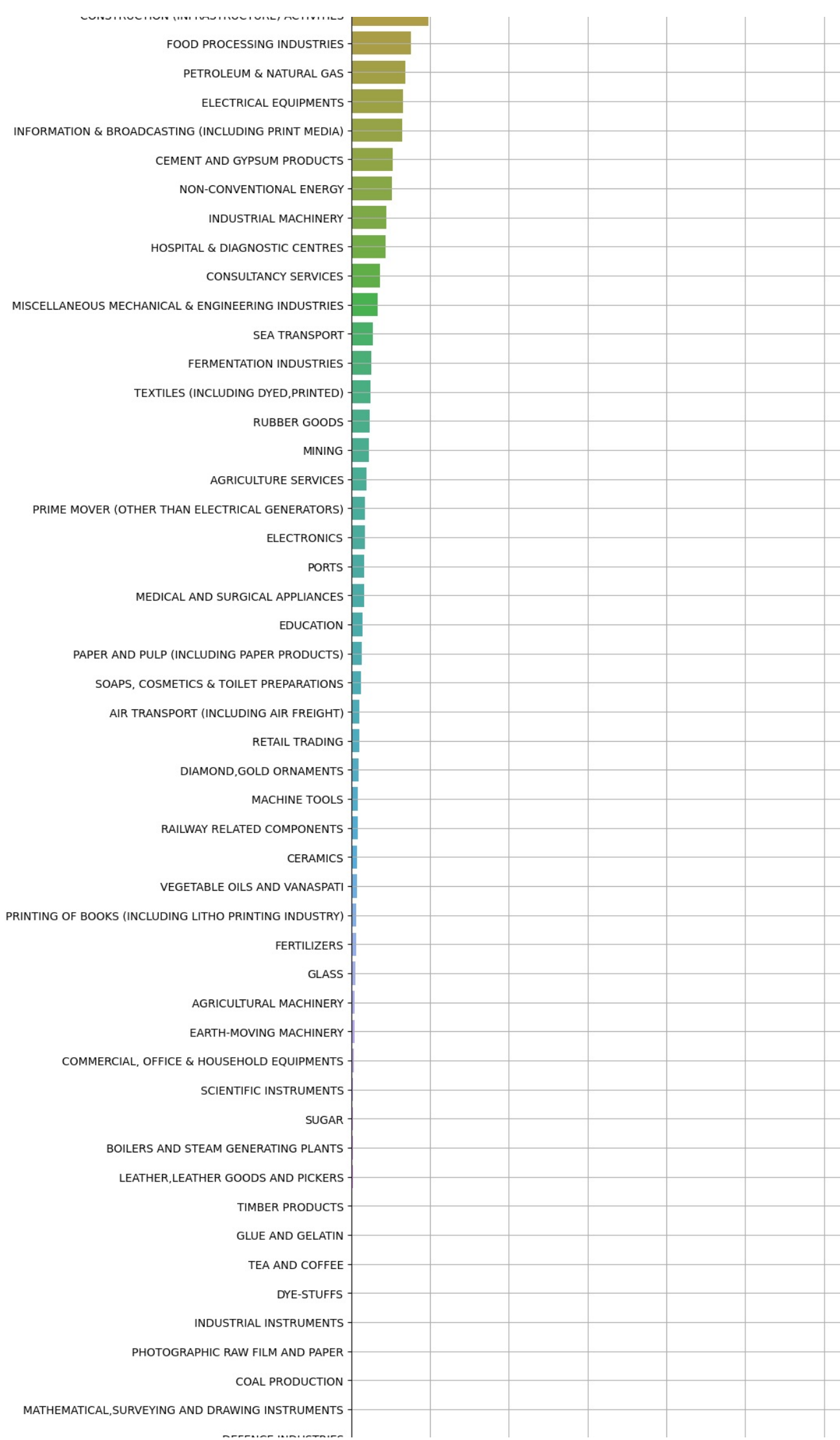|	year	FDI in US$ Millions
Sector		
SERVICES SECTOR	2000-012001-022002-032003-042004-052005-062006...	59476.49
COMPUTER SOFTWARE & HARDWARE	2000-012001-022002-032003-042004-052005-062006...	24669.49
CONSTRUCTION DEVELOPMENT	2000-012001-022002-032003-042004-052005-062006...	24293.09
TELECOMMUNICATIONS	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

Visualisation of sectorwise total FDI

```
In [29]: plt.figure(figsize=(8,30))
         sns.barplot(x='FDI in US$ Millions',y = total_fdi_sector.index ,data = total_fdi_sector)
         plt.xlabel('FDI in US$ Millions',fontsize=13)
         plt.ylabel('Sectors',fontsize=13)
         plt.title('Sectorwise total FDI',fontsize=16)
         plt.grid()
         plt.show()
```

Sectors (y-axis, top to bottom):
- CONSTRUCTION (INFRASTRUCTURE) ACTIVITIES
- FOOD PROCESSING INDUSTRIES
- PETROLEUM & NATURAL GAS
- ELECTRICAL EQUIPMENTS
- INFORMATION & BROADCASTING (INCLUDING PRINT MEDIA)
- CEMENT AND GYPSUM PRODUCTS
- NON-CONVENTIONAL ENERGY
- INDUSTRIAL MACHINERY
- HOSPITAL & DIAGNOSTIC CENTRES
- CONSULTANCY SERVICES
- MISCELLANEOUS MECHANICAL & ENGINEERING INDUSTRIES
- SEA TRANSPORT
- FERMENTATION INDUSTRIES
- TEXTILES (INCLUDING DYED,PRINTED)
- RUBBER GOODS
- MINING
- AGRICULTURE SERVICES
- PRIME MOVER (OTHER THAN ELECTRICAL GENERATORS)
- ELECTRONICS
- PORTS
- MEDICAL AND SURGICAL APPLIANCES
- EDUCATION
- PAPER AND PULP (INCLUDING PAPER PRODUCTS)
- SOAPS, COSMETICS & TOILET PREPARATIONS
- AIR TRANSPORT (INCLUDING AIR FREIGHT)
- RETAIL TRADING
- DIAMOND,GOLD ORNAMENTS
- MACHINE TOOLS
- RAILWAY RELATED COMPONENTS
- CERAMICS
- VEGETABLE OILS AND VANASPATI
- PRINTING OF BOOKS (INCLUDING LITHO PRINTING INDUSTRY)
- FERTILIZERS
- GLASS
- AGRICULTURAL MACHINERY
- EARTH-MOVING MACHINERY
- COMMERCIAL, OFFICE & HOUSEHOLD EQUIPMENTS
- SCIENTIFIC INSTRUMENTS
- SUGAR
- BOILERS AND STEAM GENERATING PLANTS
- LEATHER,LEATHER GOODS AND PICKERS
- TIMBER PRODUCTS
- GLUE AND GELATIN
- TEA AND COFFEE
- DYE-STUFFS
- INDUSTRIAL INSTRUMENTS
- PHOTOGRAPHIC RAW FILM AND PAPER
- COAL PRODUCTION
- MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS
- DEFENCE INDUSTRIES

DEFENCE INDUSTRIES
COIR

0   10000   20000   30000   40000   50000   60000
FDI in US$ Millions

Best and worst performing sectors

```
In [50]: top_10 = total_fdi_sector.head(10)
         top_10
```

Out[50]:

	year	FDI in US$ Millions
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

check worst 10 performing secotrs

```
In [54]: least_10 = total_fdi_sector.sort_values(by='FDI in US$ Millions', ascending=True).head(10)
         least_10
```

Out[54]:

	year	FDI in US$ Millions
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
INDUSTRIAL INSTRUMENTS	2000-012001-022002-032003-042004-052005-062006...	76.12
DYE-STUFFS	2000-012001-022002-032003-042004-052005-062006...	88.40
TEA AND COFFEE	2000-012001-022002-032003-042004-052005-062006...	111.22
GLUE AND GELATIN	2000-012001-022002-032003-042004-052005-062006...	128.39
TIMBER PRODUCTS	2000-012001-022002-032003-042004-052005-062006...	157.68

pie chart to visualise percentage share of FDI among top 10 sectors

```
In [51]: plt.figure(figsize=(20,8))
         colors = sns.color_palette('pastel')[0:7]
         textprops = {'fontsize':13}
         plt.pie(top_10['FDI in US$ Millions'],labels=top_10.index,colors=colors, autopct="%1.1f%%",textprops = textprop
         plt.axis('equal')
         plt.title('SHARE AMONG TOP 10 SECTORS', fontsize = 20)
         plt.show()
```
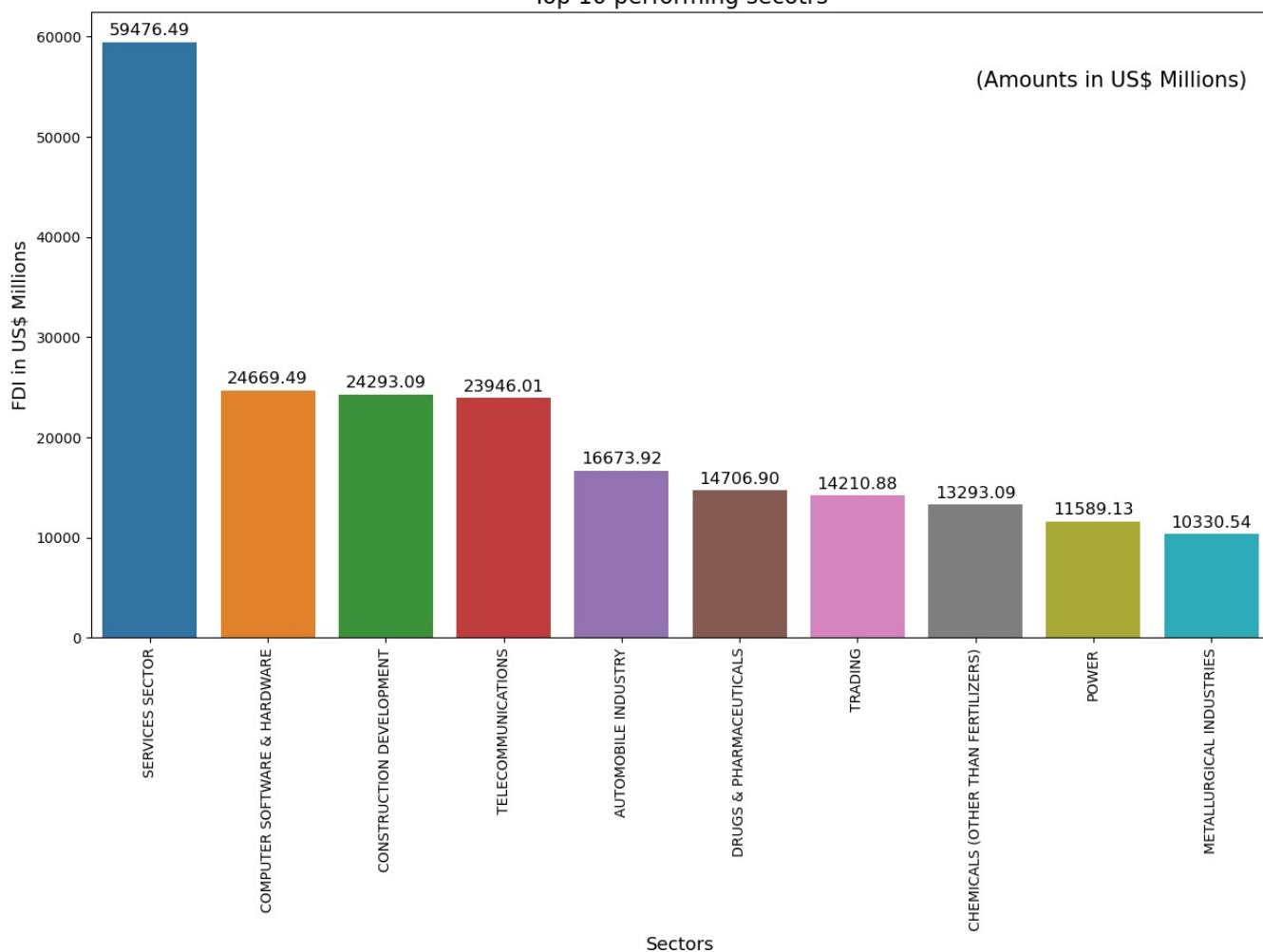
SHARE AMONG TOP 10 SECTORS
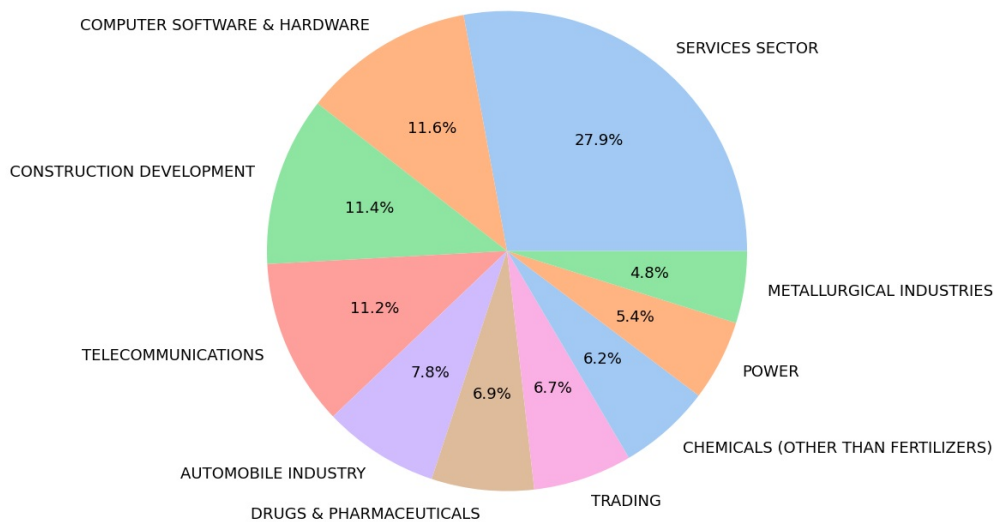
visualise Top 10 performing secotrs

```
In [52]:  plt.figure(figsize=(15,8))
          plots = sns.barplot(x = top_10.index, y='FDI in US$ Millions' ,data = top_10)
          for bar in plots.patches:
            plots.annotate(format(bar.get_height(),'.2f'),
                           (bar.get_x() + bar.get_width() / 2,
                            bar.get_height()), ha='center', va='center',
                           size=12, xytext=(0, 8),
                           textcoords='offset points')
          plt.xlabel('Sectors',fontsize=13)
          plt.ylabel('FDI in US$ Millions',fontsize=13)
          plt.xticks(rotation = 90)
          plt.text(7,55000, "(Amounts in US$ Millions)", fontsize = 15)
          plt.title('Top 10 performing secotrs',fontsize=16)
          plt.show()
```


visualise worst 10 performing secotrs

```
In [56]:  plt.figure(figsize=(15,8))
          plots = sns.barplot(x = least_10.index, y='FDI in US$ Millions' ,data = least_10)
```

```
for bar in plots.patches:
    plots.annotate(format(bar.get_height(),'.2f'),
                   (bar.get_x() + bar.get_width() / 2,
                    bar.get_height()), ha='center', va='center',
                   size=12, xytext=(0, 8),
                   textcoords='offset points')
plt.xlabel('Sectors',fontsize=13)
plt.ylabel('FDI in US$ Millions',fontsize=13)
plt.xticks(rotation = 90)
plt.text(0,150, "(Amounts in US$ Millions)", fontsize = 15)
plt.title('Worst 10 secotrs',fontsize=16)
plt.show()
```
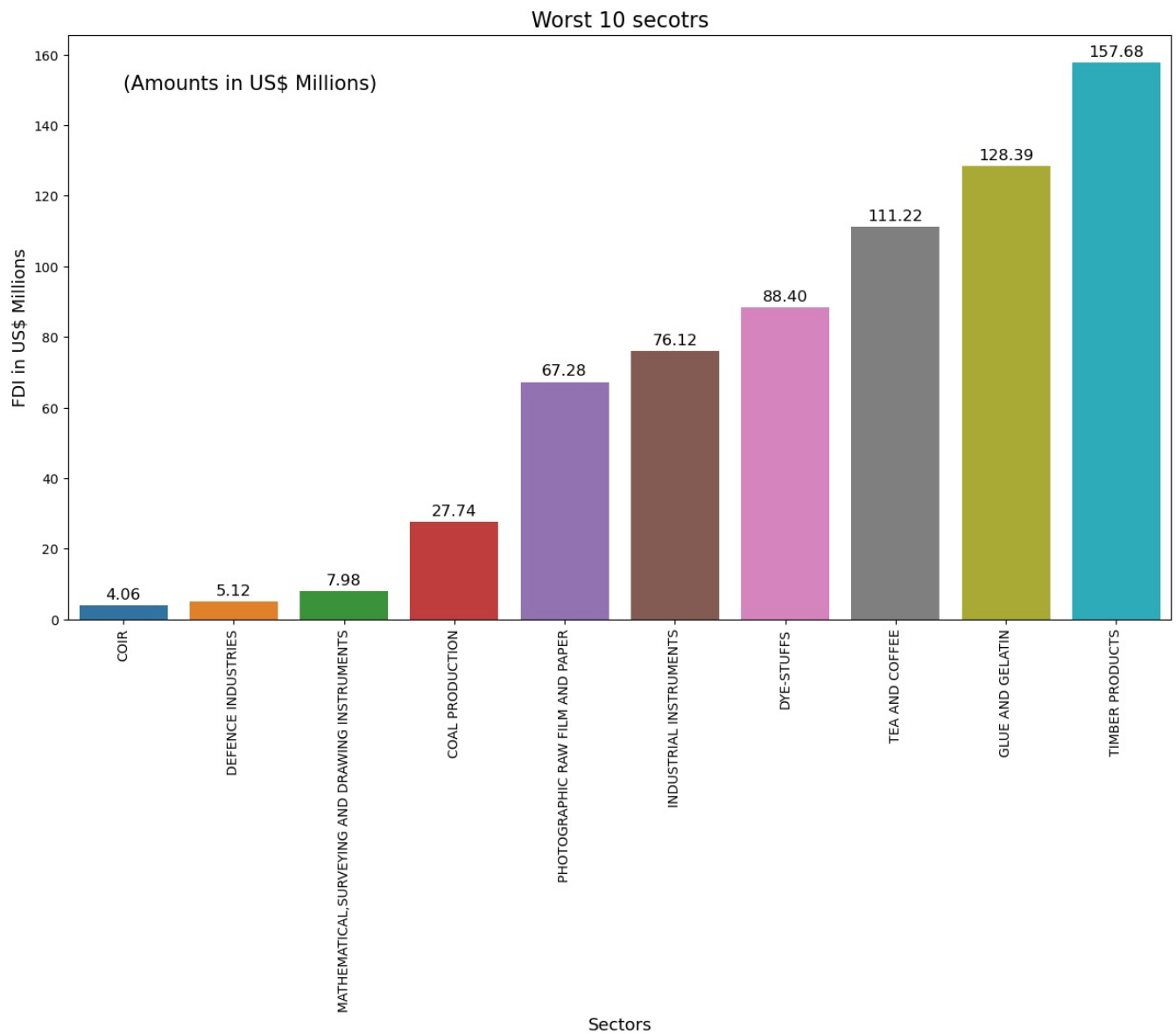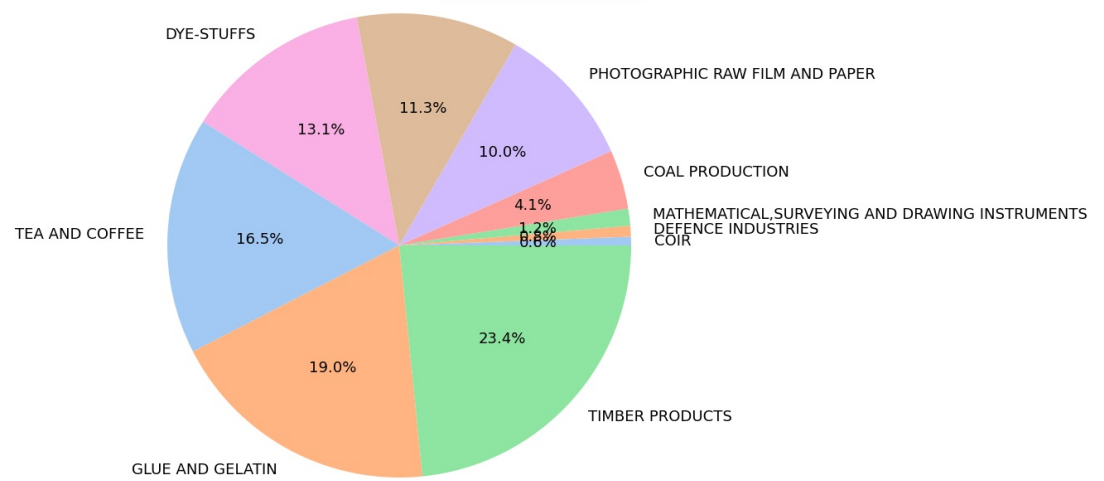

pie chart to visualise percentage share of FDI among worst 10 sectors

```
In [57]: plt.figure(figsize=(20,8))
         colors = sns.color_palette('pastel')[0:7]
         textprops = {'fontsize':13}
         plt.pie(least_10['FDI in US$ Millions'],labels=least_10.index,colors=colors, autopct="%1.1f%%",textprops = text
         plt.axis('equal')
         plt.title('SHARE AMONG WORST 10 SECTORS', fontsize = 20)
         plt.show()
```

SHARE AMONG WORST 10 SECTORS


FDI Inflow Yearwise

```
In [59]: year_inflow = FDI_1[['year', 'FDI in US$ Millions']]
         year_inflow = year_inflow.groupby('year').sum()
         year_inflow['% growth over previous year'] = round(year_inflow.pct_change()*100,2)
         year_inflow
```

Out[59]:

year	FDI in US$ Millions	% growth over previous 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

plotting to show Year by Year FDI Inflow

```
In [60]: year_inflow.plot.line(y='FDI in US$ Millions',figsize = (20,9))
         plt.xlabel('Year', fontsize = 13)
         plt.ylabel('FDI in US$ Millions', fontsize = 13)
         plt.title('FDI Inflow', fontsize = 15)
         plt.show()
```
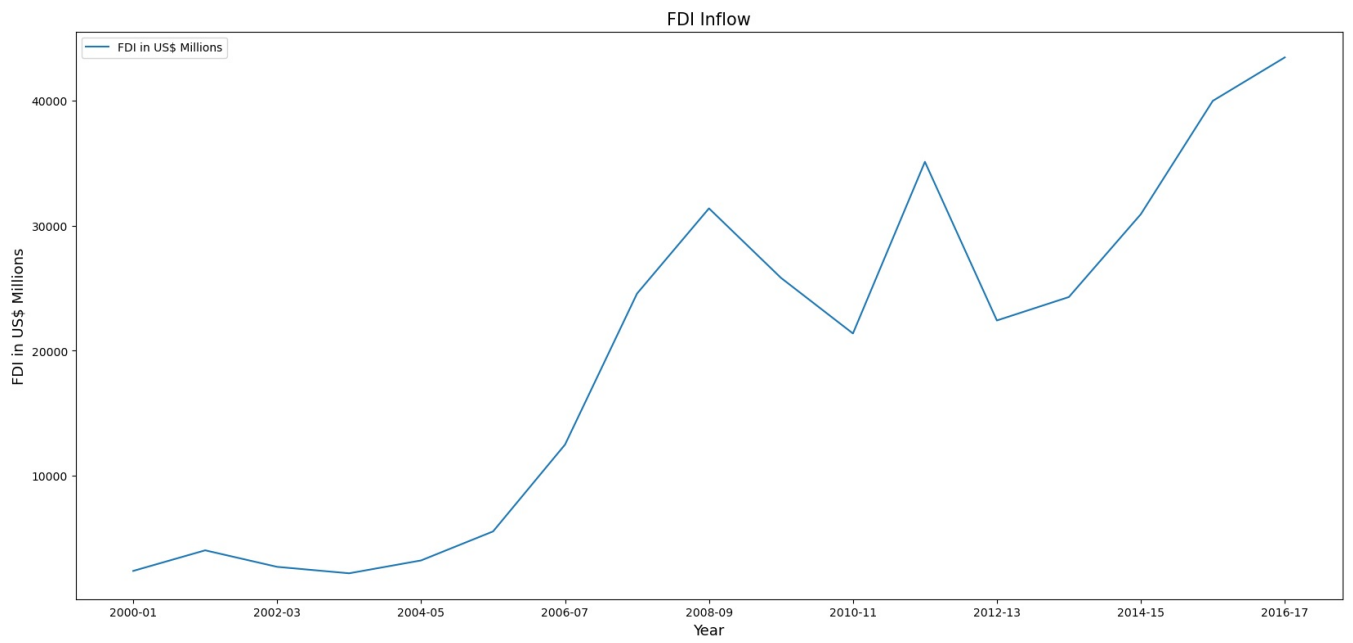
FDI Inflow

```
In [ ]:  2000-01 upto 2005-06 there is not much inflow in India.
```

```
In [ ]:  2000-01 upto 2005-06 there is not much inflow in India.
```

```
In [ ]:  2008-09 to 2010-11 some downfall
```

```
In [ ]:  after 2012-13 come back flow in India.
```

```
In [ ]:  overall there is growth in FDI inflows.
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js