```
from google.colab import drive
drive.mount('/content/drive')
```

Suggested code may be subject to a license |

```
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

```
df = pd.read_csv('/content/drive/MyDrive/FDI_in_India.csv')
df
```

	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
58	PRINTING OF BOOKS (INCLUDING LITHO PRINTING IN...	0.00	0.00	6.30	0.00	0.06	9.90	20.04	35.54	31.61	70.51	36.63	47.39	14.34	113.78	72.58	122.81	53.17
59	COIR	0.00	0.00	0.00	0.00	0.47	0.59	0.04	0.01	0.00	0.25	0.10	0.55	0.15	0.54	1.36	0.00	0.00
60	CONSTRUCTION (INFRASTRUCTURE) ACTIVITIES	0.00	0.00	0.00	0.00	0.00	0.93	64.06	182.92	172.70	324.56	675.07	386.28	283.89	485.37	870.25	4510.71	1860.73
61	CONSTRUCTION DEVELOPMENT: Townships, housing	24.33	51.75	36.10	47.04	152.06	228.71	1392.95	3887.33	4657.51	5466.13	1663.03	3140.78	1332.49	1226.05	769.14	112.55	105.14

Next steps: Generate code with df View recommended plots

```
df.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
```

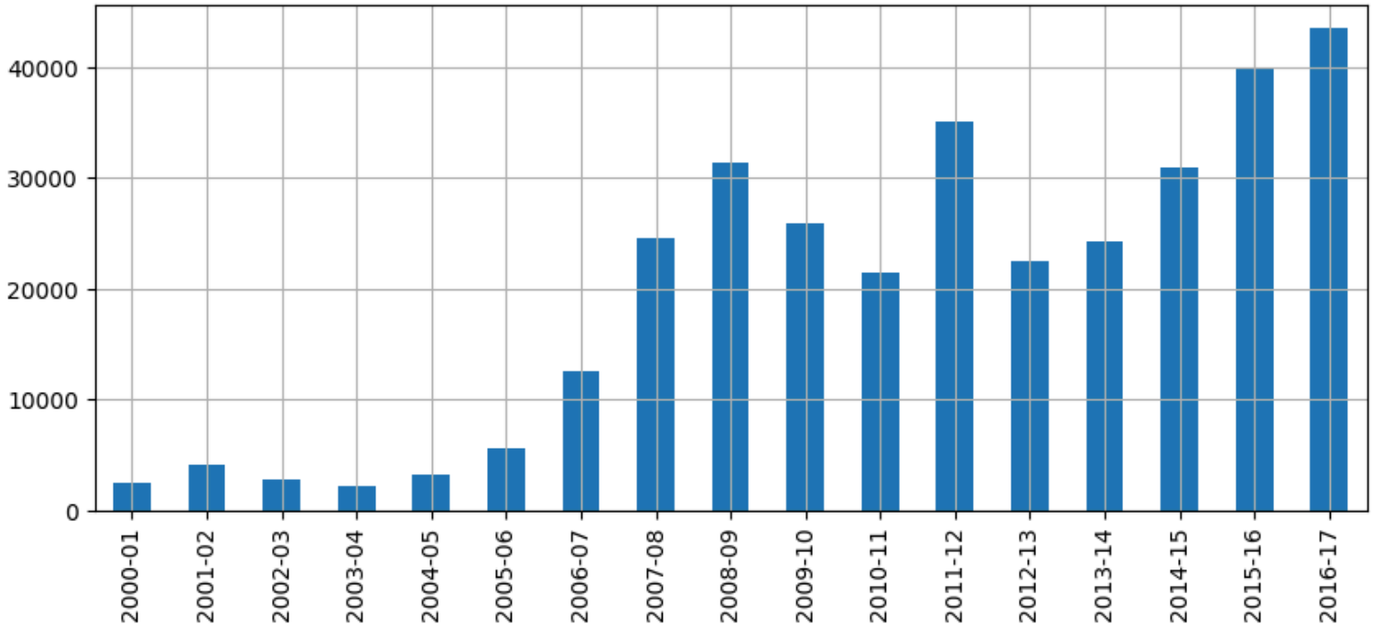
Descriptive Statistics of the FDI Dataset

```
df.describe()
```

	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2
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.
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.
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.
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.
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
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.
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.
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.

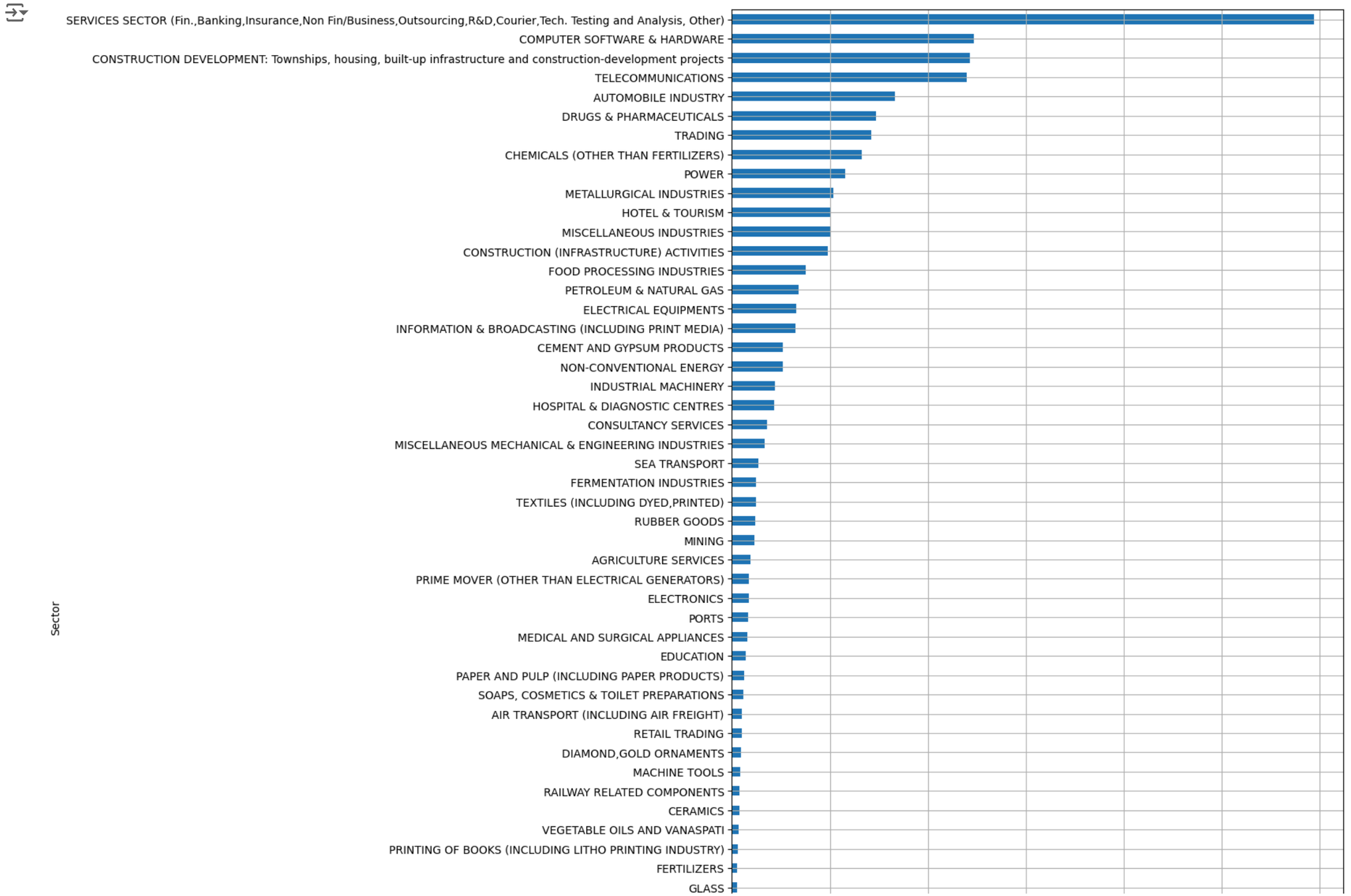
Year-wise investment

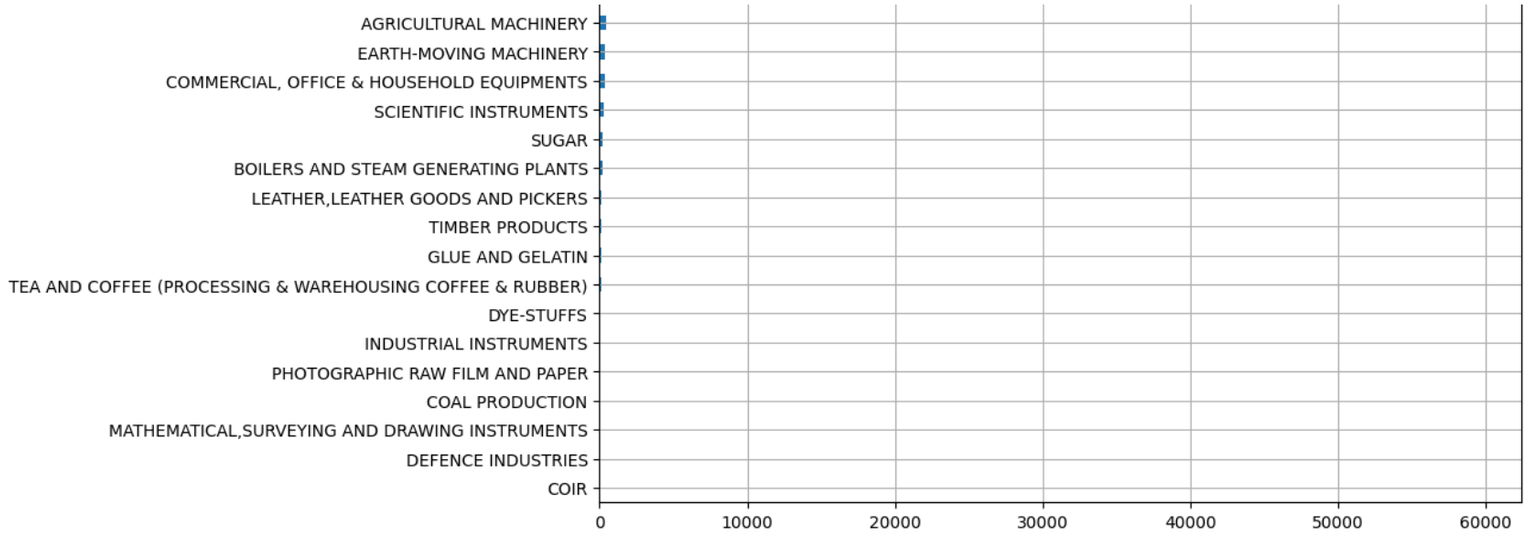
```
df.sum(axis=0).plot(kind='bar', figsize=(10,4))
plt.grid()
plt.show()
```



Sector-wise Investment over the years

```
df.sum(axis=1).sort_values().plot(kind = 'barh', figsize=(10,20))
plt.grid()
plt.show()
```

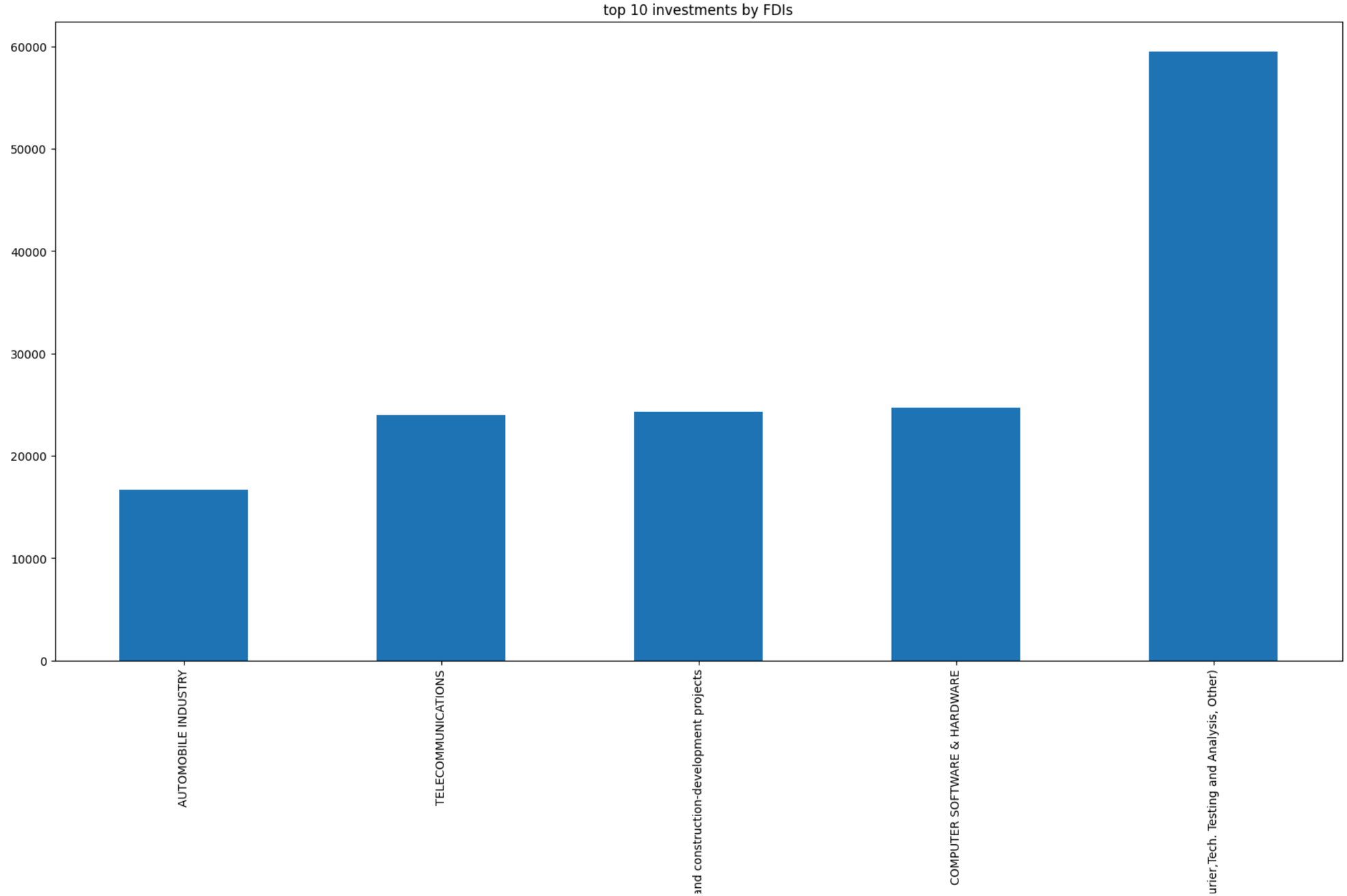

SERVICES SECTOR (Fin.,Banking,Insurance,Non Fin/Business,Outsourcing,R&D,Courier,Tech. Testing and Analysis, Other)
COMPUTER SOFTWARE & HARDWARE
CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects
TELECOMMUNICATIONS
AUTOMOBILE INDUSTRY
DRUGS & PHARMACEUTICALS
TRADING
CHEMICALS (OTHER THAN FERTILIZERS)
POWER
METALLURGICAL INDUSTRIES
HOTEL & TOURISM
MISCELLANEOUS INDUSTRIES
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
Sector



Top 5 Investment Sector

```
df_trans = df.transpose()
df_trans.sum().sort_values()[-5:].plot(figsize=(20,10),kind='bar', title ='top 10 investments by FDIs')
```

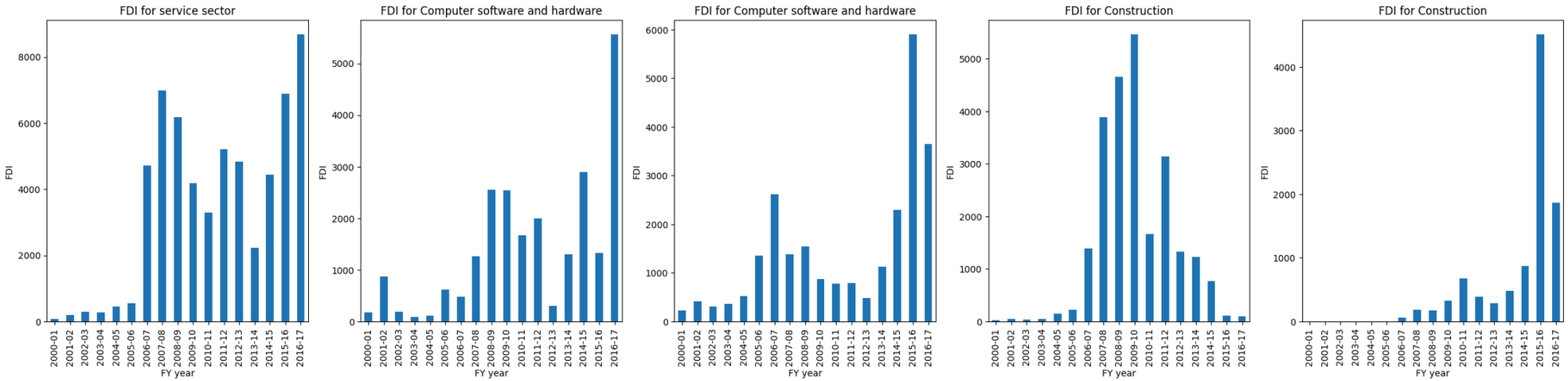
```
<Axes: title={'center': 'top 10 investments by FDIs'}, xlabel='Sector'>
```



Year on Year Top 5 best performing sector

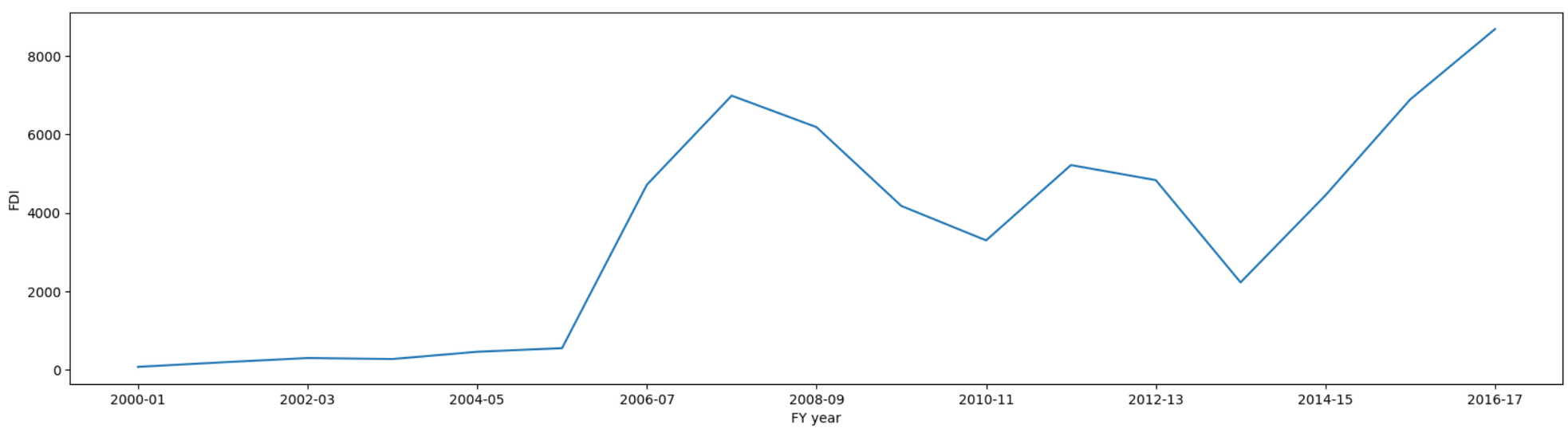
```
plt.subplot(1, 5, 1)
df.loc['SERVICES SECTOR (Fin.,Banking,Insurance,Non Fin/Business,Outsourcing,R&D,Courier,Tech. Testing and Analysis, Other)'].plot(kind='bar',figsize=(30,6))
plt.title('FDI for service sector')
plt.xlabel('FY year')
plt.ylabel('FDI')
plt.subplot(1, 5, 2)
df.loc['TELECOMMUNICATIONS'].plot(kind='bar',title='FDI 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 Construc
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')
```



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()
```




```
X = list(df.columns)
# X.remove('Sector')
X
```

```
['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']
```

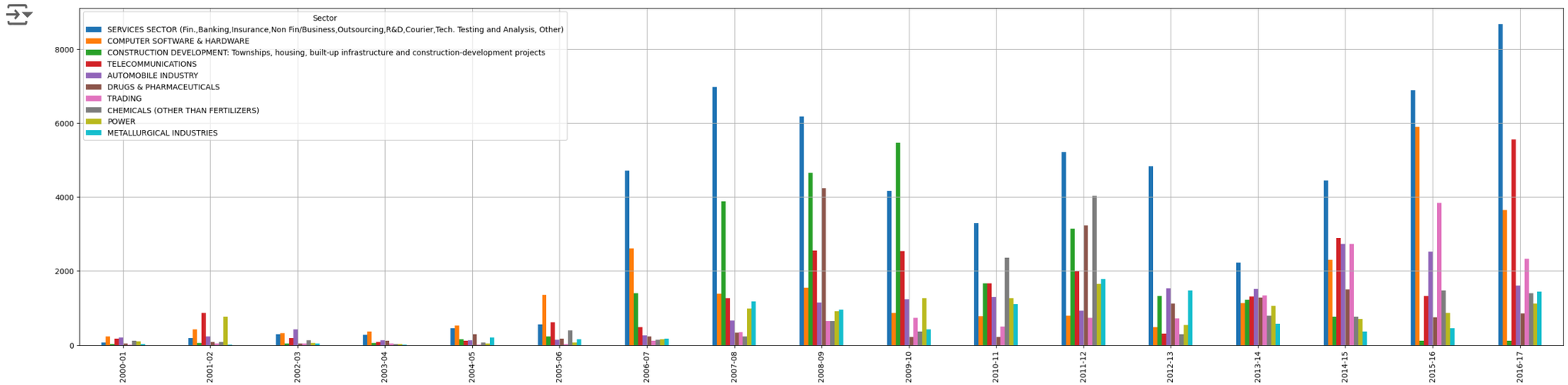
Top and Bottom Sectors*

Top 10 Sectors

```
df.sum(axis=1).nlargest(10)
```

```
Sector
SERVICES SECTOR (Fin.,Banking,Insurance,Non Fin/Business,Outsourcing,R&D,Courier,Tech. Testing and Analysis, Other)    59476.49
COMPUTER SOFTWARE & HARDWARE                                                                                          24669.49
CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects           24293.09
TELECOMMUNICATIONS                                                                                                    23946.01
AUTOMOBILE INDUSTRY                                                                                                   16673.92
DRUGS & PHARMACEUTICALS                                                                                               14706.90
TRADING                                                                                                               14210.88
CHEMICALS (OTHER THAN FERTILIZERS)                                                                                    13293.09
POWER                                                                                                                 11589.13
METALLURGICAL INDUSTRIES                                                                                              10330.54
dtype: float64
```

```
df_1 = df.loc[['SERVICES SECTOR (Fin.,Banking,Insurance,Non Fin/Business,Outsourcing,R&D,Courier,Tech. Testing and Analysis, Other)',
'COMPUTER SOFTWARE & HARDWARE',
'CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects',
'TELECOMMUNICATIONS',
'AUTOMOBILE INDUSTRY',
'DRUGS & PHARMACEUTICALS',
'TRADING',
'CHEMICALS (OTHER THAN FERTILIZERS)',
'POWER',
'METALLURGICAL INDUSTRIES'   ], X]
df_1.transpose().plot(kind = 'bar', figsize=(35,8))
plt.grid()
plt.show()
```

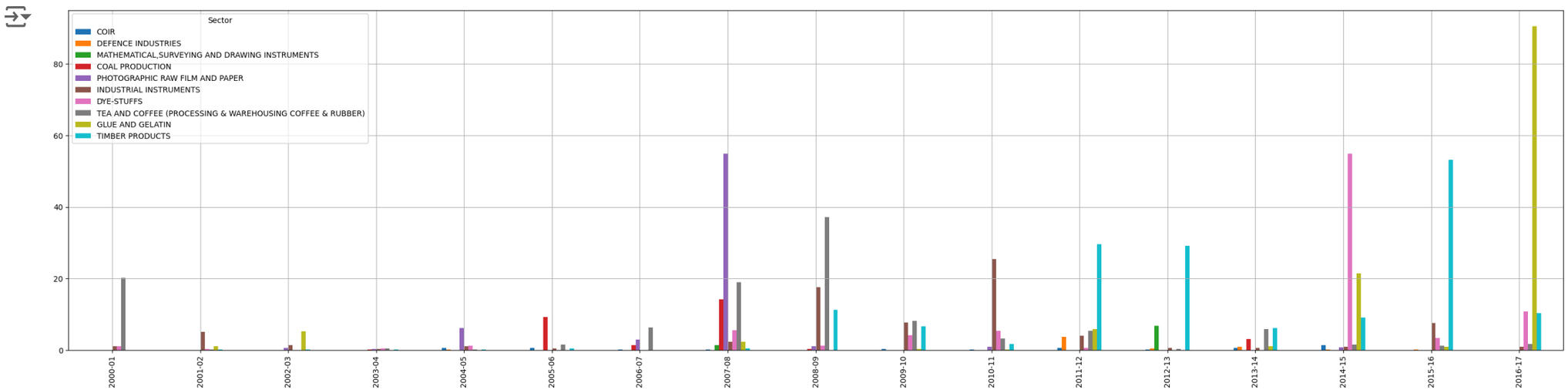


Bottom 10 Sectors

```
s = df.sum(axis=1).sort_values(ascending=True)
s.head(10)
```

```
Sector
COIR                                                  4.06
DEFENCE INDUSTRIES                                    5.12
MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS        7.98
COAL PRODUCTION                                      27.74
PHOTOGRAPHIC RAW FILM AND PAPER                      67.28
INDUSTRIAL INSTRUMENTS                               76.12
DYE-STUFFS                                           88.40
TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)   111.22
GLUE AND GELATIN                                    128.39
TIMBER PRODUCTS                                     157.68
dtype: float64
```

```
df_1 = df.loc[['COIR',
'DEFENCE INDUSTRIES',
'MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS',
'COAL PRODUCTION',
'PHOTOGRAPHIC RAW FILM AND PAPER',
'INDUSTRIAL INSTRUMENTS',
'DYE-STUFFS',
'TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)',
'GLUE AND GELATIN',
'TIMBER PRODUCTS'], X]
df_1.transpose().plot(kind = 'bar', figsize=(35,8))
plt.grid()
plt.show()
```



Maximum Investment in Sectors Year-wise

```
df.idxmax()
```

```
2000-01                        MISCELLANEOUS INDUSTRIES
2001-02                             TELECOMMUNICATIONS
```

```
2002-03                                AUTOMOBILE INDUSTRY
2003-04                          COMPUTER SOFTWARE & HARDWARE
2004-05                          COMPUTER SOFTWARE & HARDWARE
2005-06                          COMPUTER SOFTWARE & HARDWARE
2006-07    SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...
2007-08    SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...
2008-09    SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...
2009-10    CONSTRUCTION DEVELOPMENT: Townships, housing, ...
2010-11    SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...
2011-12    SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...
2012-13    SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...
2013-14                            FOOD PROCESSING INDUSTRIES
2014-15    SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...
2015-16    SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...
2016-17    SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...
dtype: object
```

Minimum Investment in Sectors Year-wise

```
df.idxmin()
```

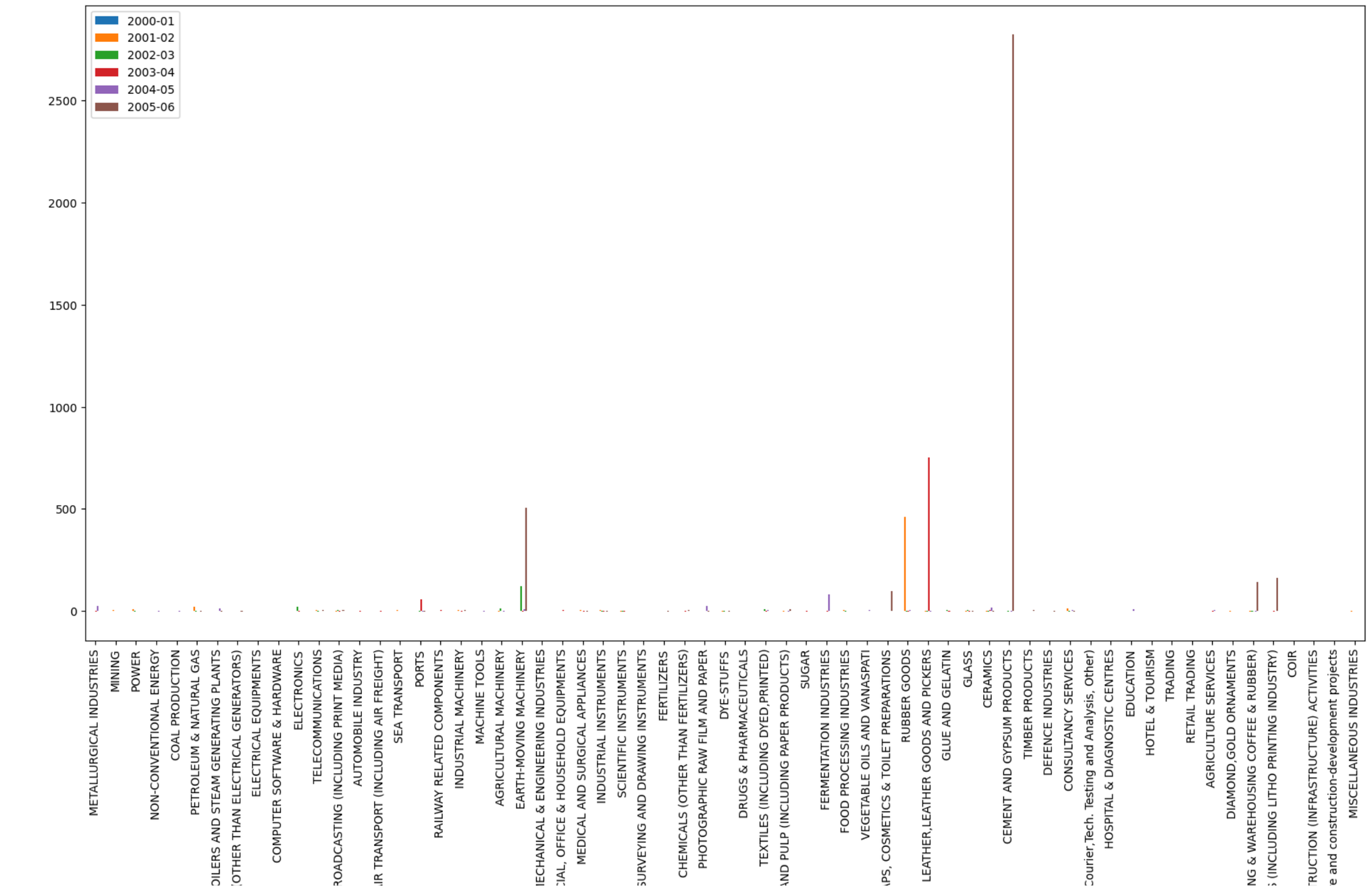
```
2000-01                            NON-CONVENTIONAL ENERGY
2001-02                            NON-CONVENTIONAL ENERGY
2002-03                                    COAL PRODUCTION
2003-04    PRIME MOVER (OTHER THAN ELECTRICAL GENERATORS)
2004-05                                    COAL PRODUCTION
2005-06                BOILERS AND STEAM GENERATING PLANTS
2006-07                                             PORTS
2007-08                             SCIENTIFIC INSTRUMENTS
2008-09                BOILERS AND STEAM GENERATING PLANTS
2009-10                                    COAL PRODUCTION
2010-11                                    COAL PRODUCTION
2011-12                                    COAL PRODUCTION
2012-13                                    COAL PRODUCTION
2013-14    MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS
2014-15                                    COAL PRODUCTION
2015-16                                    COAL PRODUCTION
2016-17                                    COAL PRODUCTION
dtype: object
```

Quinquennial percentage changes in investment over the sectors

Period 2000-2005

```
df.iloc[:, 0:6].pct_change(axis=1).plot(kind = 'bar', figsize=(20,10))
```

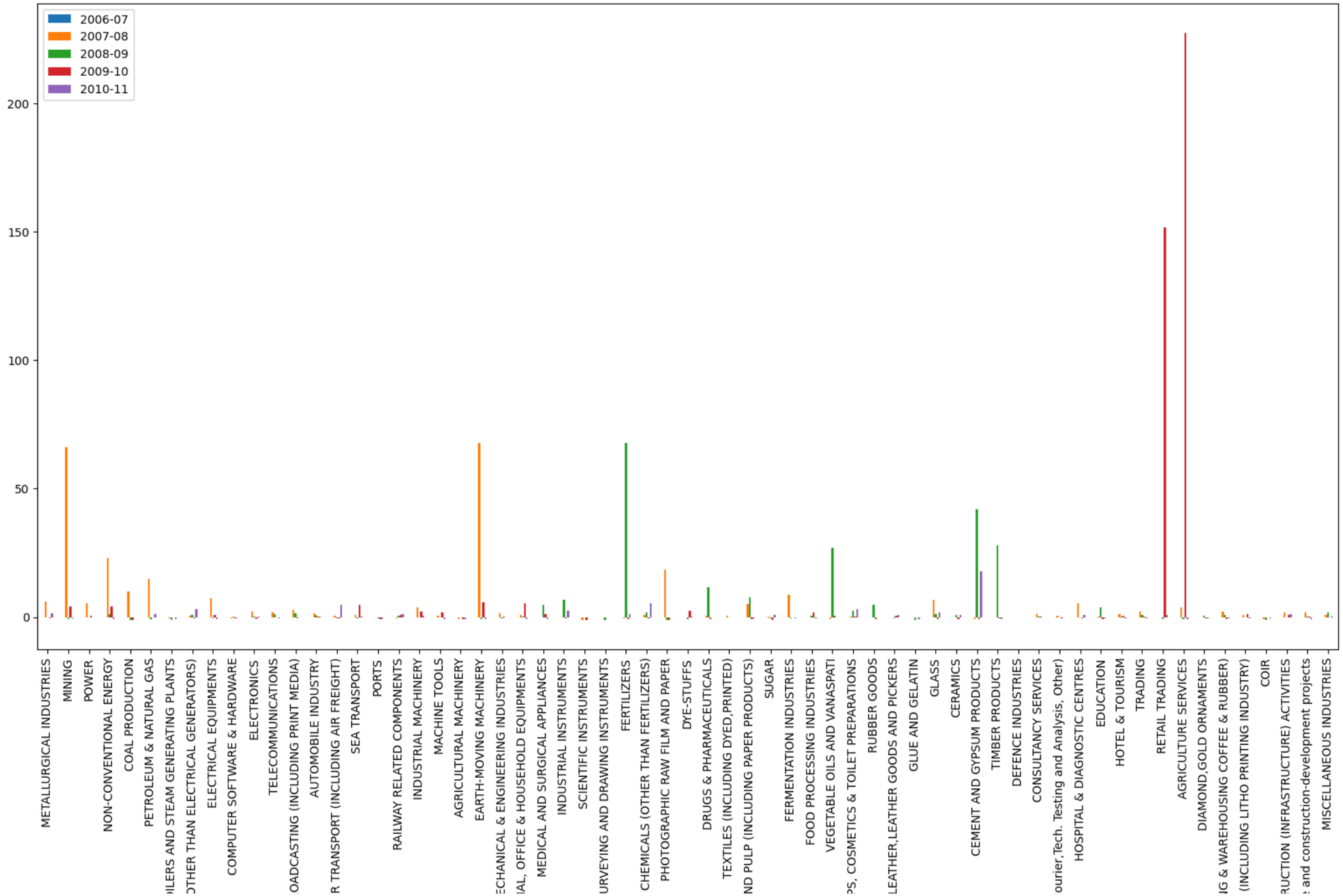
<Axes: xlabel='Sector'>

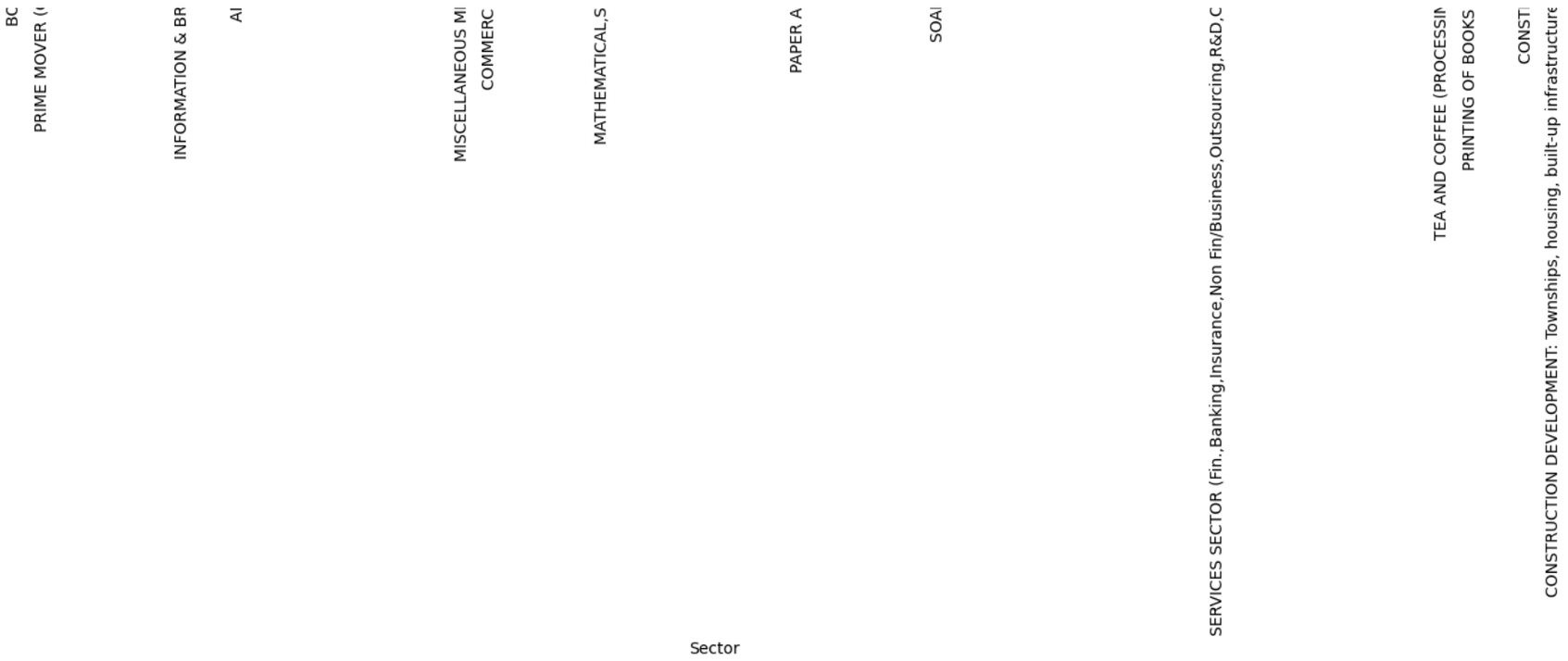


Period 2006-2010

```
df.iloc[:, 6:11].pct_change(axis=1).plot(kind = 'bar', figsize=(20,10))
```

```
<Axes: xlabel='Sector'>
```





Period 2011-2016

```
df.iloc[:, 11:].pct_change(axis=1).plot(kind = 'bar', figsize=(20,10))
```

<Axes: xlabel='Sector'>