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

In [2]:
```python
fd=pd.read_excel("C:\\Users\\USER\\Downloads\\projects\\FDI data.xlsx")
fd
```

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	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
62	MISCELLANEOUS INDUSTRIES	832.07	221.37	218.76	235.48	121.83	164.76	304.87	528.42	1549.70	1147.56	1475.97	813.38	229.49	468.74	765.88	668.77	296.40

63 rows × 18 columns

In [3]:
```python
fd.isna().any()
```

Out[3]:
```
Sector     False
2000-01    False
2001-02    False
2002-03    False
2003-04    False
2004-05    False
2005-06    False
2006-07    False
2007-08    False
2008-09    False
2009-10    False
2010-11    False
2011-12    False
2012-13    False
2013-14    False
2014-15    False
2015-16    False
2016-17    False
dtype: bool
```

In [4]:
```python
fd.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
```

In [5]:
```python
fd.describe()
```

Out[5]:

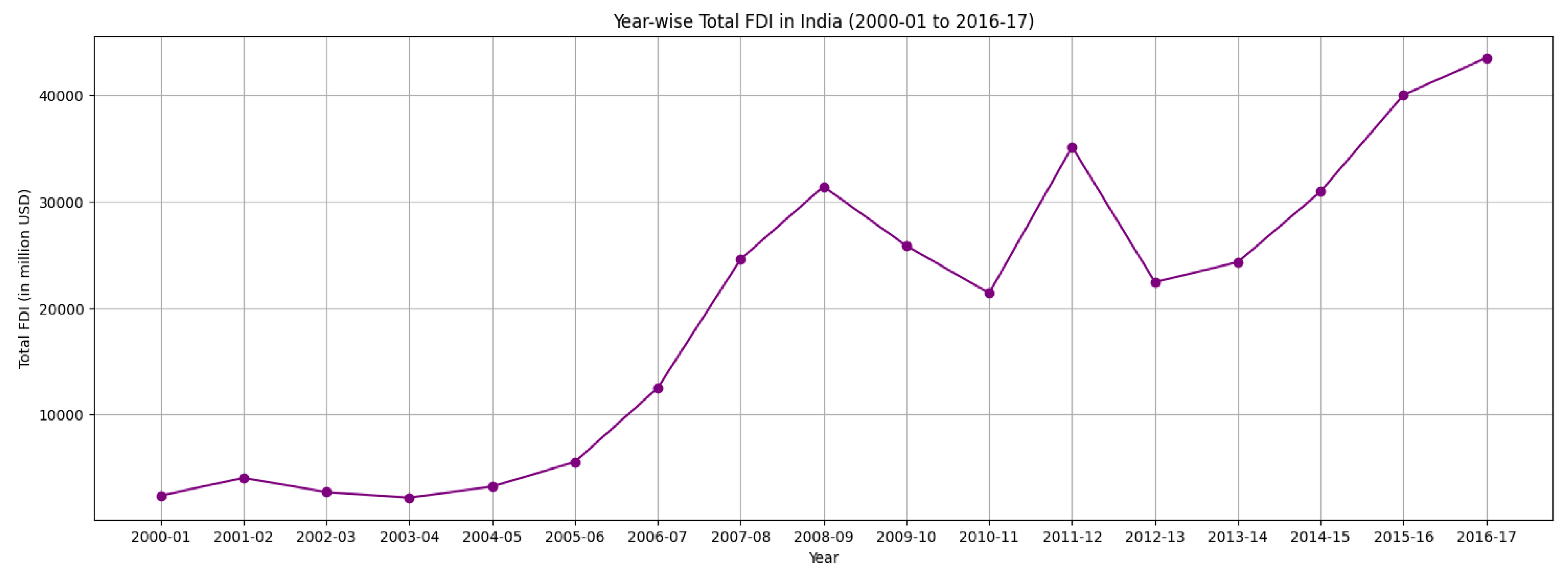
	2000-01	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014
count	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.000000	63.0000
mean	37.757302	63.931587	42.925714	34.727778	51.090317	87.932540	198.281905	390.085714	498.348571	410.069524	339.413810	557.472698	355.930000	385.703492	490.9598
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	837.7870
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	0.0000
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	33.8000
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	177.2200
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	595.3900
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	4443.2600

In [9]:
```python
print("Descriptive Statistics")
fd['Total_FDI'] = fd.iloc[:, 1:].sum(axis=1)
fd['Average_FDI'] = fd.iloc[:, 1:-1].mean(axis=1)
fd['StdDev_FDI'] = fd.iloc[:, 1:-1].std(axis=1)
print(fd[['Sector', 'Total_FDI', 'Average_FDI', 'StdDev_FDI']].sort_values(by='Total_FDI', ascending=False).head(10))
```

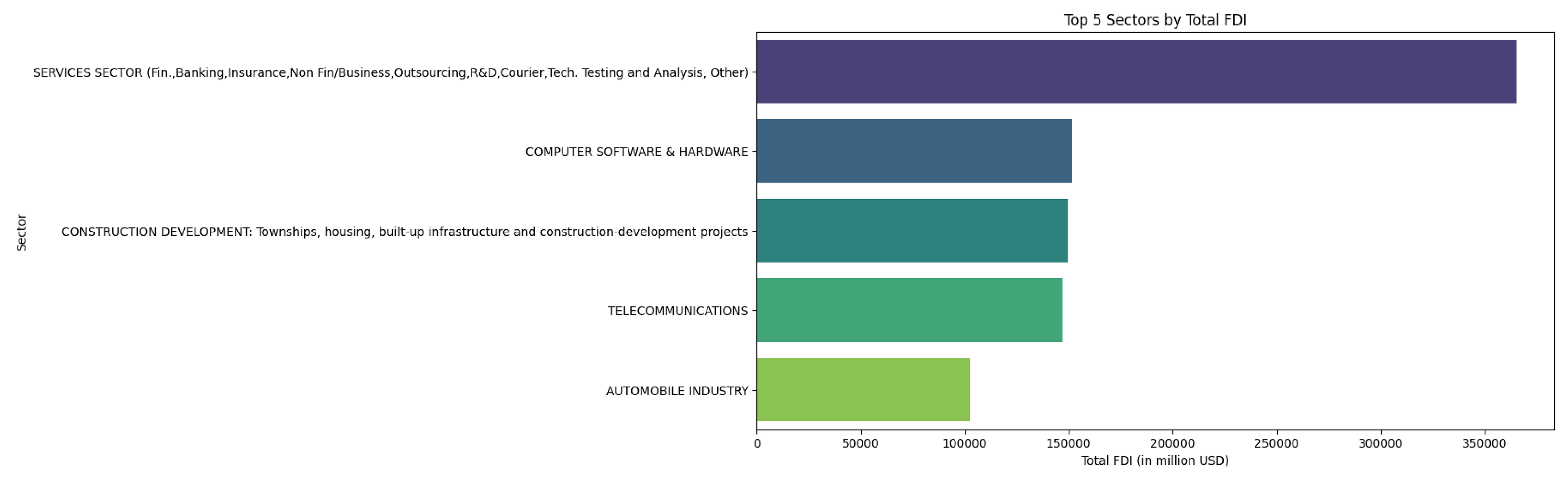
```
Descriptive Statistics
                                               Sector      Total_FDI  \
49  SERVICES SECTOR (Fin.,Banking,Insurance,Non Fi...  365320.430090   
9                         COMPUTER SOFTWARE & HARDWARE  151688.824472   
61  CONSTRUCTION DEVELOPMENT: Townships, housing, ...  149586.579479   
11                                  TELECOMMUNICATIONS  147227.799272   
13                                 AUTOMOBILE INDUSTRY  102431.437334   
32                             DRUGS & PHARMACEUTICALS   90643.813617   
53                                             TRADING   87563.437554   
29                 CHEMICALS (OTHER THAN FERTILIZERS)   81912.225005   
2                                                POWER   71170.756919   
0                             METALLURGICAL INDUSTRIES   63504.526579   

     Average_FDI    StdDev_FDI  
49  23206.128630  82922.463086  
9    9634.237267  34441.826711  
61   9498.833388  33978.525881  
11   9351.014043  33428.105264  
13   6506.578022  23251.490701  
32   5755.180896  20595.350325  
53   5559.806636  19893.906005  
29   5200.950071  18610.246119  
2    4521.069809  16153.899480  
0    4033.518517  14417.996059  
```

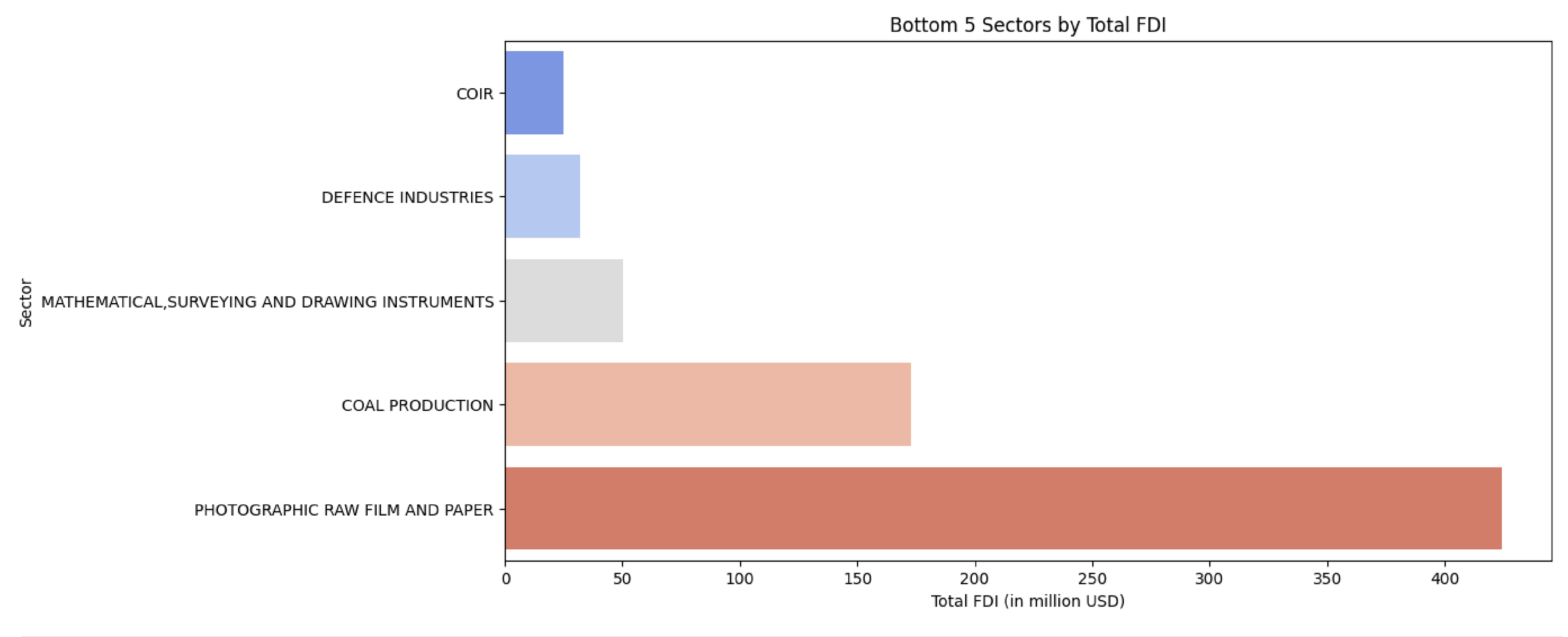
In [15]:
```python
yearly_fdi = fd.iloc[:, 1:-3].sum(axis=0)
plt.figure(figsize=(18, 6))
plt.plot(yearly_fdi.index, yearly_fdi.values, marker='o', color="purple")
plt.title('Year-wise Total FDI in India (2000-01 to 2016-17)')
plt.xlabel('Year')
plt.ylabel('Total FDI (in million USD)')
plt.grid(True)
```



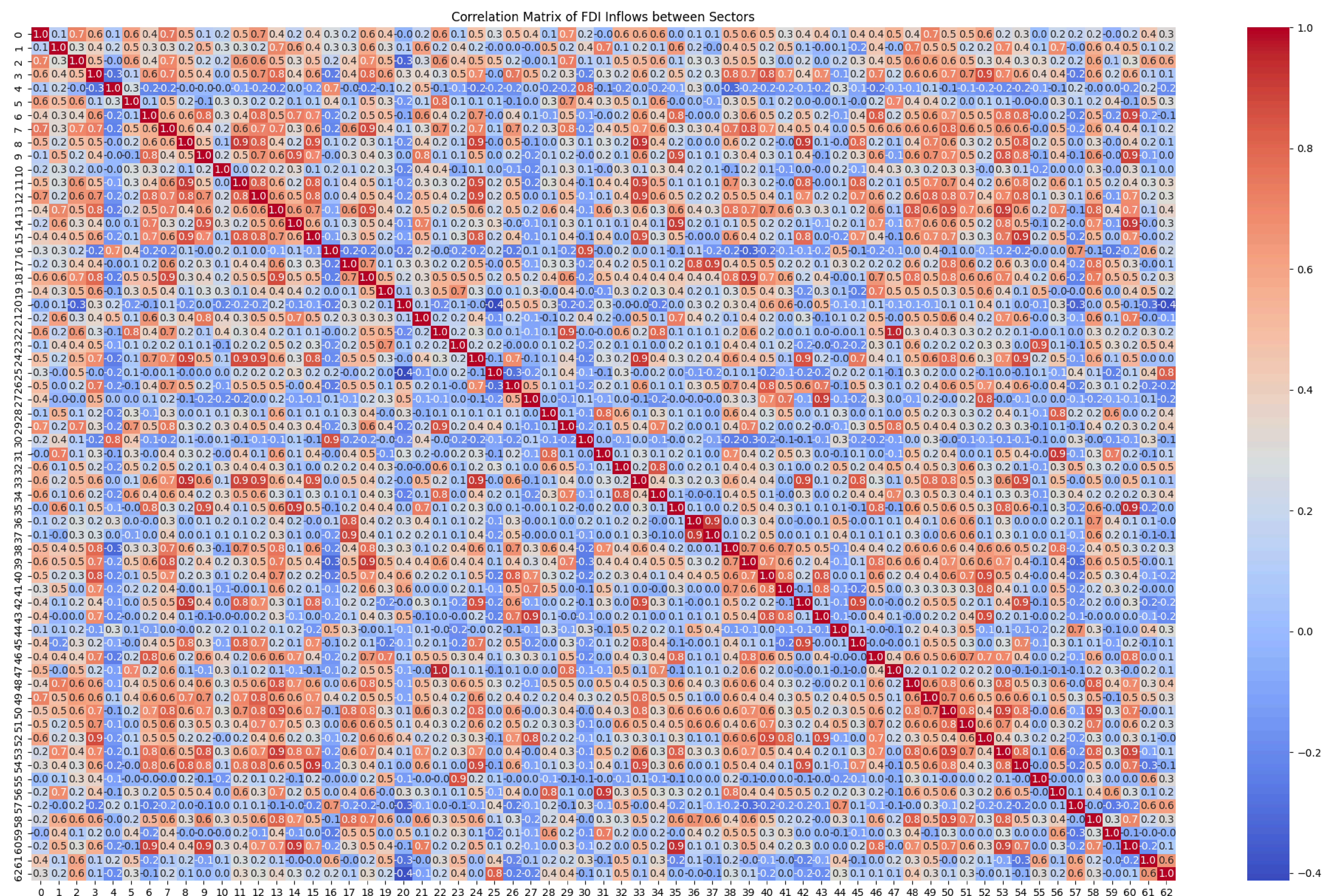
In [23]:
```python
top_5_sectors = fd[['Sector', 'Total_FDI']].sort_values(by='Total_FDI', ascending=False).head(5)
plt.figure(figsize=(12, 6))
sns.barplot(data=top_5_sectors, x='Total_FDI', y='Sector',hue='Sector', palette="viridis")
plt.title('Top 5 Sectors by Total FDI')
plt.xlabel('Total FDI (in million USD)')
plt.ylabel('Sector')
plt.show()
```



In [24]:
```python
bottom_5_sectors = fd[['Sector', 'Total_FDI']].sort_values(by='Total_FDI', ascending=True).head(5)
plt.figure(figsize=(12, 6))
sns.barplot(data=bottom_5_sectors, x='Total_FDI', y='Sector',hue='Sector', palette="coolwarm")
plt.title('Bottom 5 Sectors by Total FDI')
plt.xlabel('Total FDI (in million USD)')
plt.ylabel('Sector')
plt.show()
```



In [25]:
```python
fdi_corr = fd.iloc[:, 1:-3].transpose().corr()
plt.figure(figsize=(25, 15))
sns.heatmap(fdi_corr, annot=True, cmap='coolwarm', fmt='.1f')
plt.title('Correlation Matrix of FDI Inflows between Sectors')
plt.show()
```



In [ ]: