

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')

from google.colab import files
uploaded = files.upload()

<IPython.core.display.HTML object>

Saving FDI data.csv to FDI data.csv

df = pd.read_csv('FDI data.csv')
df.head()

{
  "summary": {
    "name": "df",
    "rows": 63,
    "fields": [
      {
        "column": "Sector",
        "properties": {
          "dtype": "string",
          "num_unique_values": 63,
          "samples": [
            "CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects",
            "TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)",
            "METALLURGICAL INDUSTRIES"
          ],
          "semantic_type": "\",
          "description": "\n\n",
          "column": "2000-01",
          "properties": {
            "dtype": "number",
            "std": 112.22785996479905,
            "min": 0.0,
            "max": 832.07,
            "num_unique_values": 41,
            "samples": [
              { "60.04": 1.42, "177.69": 1 },
              { "60.04": 1.42, "177.69": 1 }
            ],
            "semantic_type": "\",
            "description": "\n\n",
            "column": "2001-02",
            "properties": {
              "dtype": "number",
              "std": 157.87873650170664,
              "min": 0.0,
              "max": 873.23,
              "num_unique_values": 46,
              "samples": [
                { "32.12": 5.28, "32.12": 16.7 },
                { "32.12": 5.28, "32.12": 16.7 }
              ],
              "semantic_type": "\",
              "description": "\n\n",
              "column": "2002-03",
              "properties": {
                "dtype": "number",
                "std": 86.6064393738354,
                "min": 0.0,
                "max": 419.96,
                "num_unique_values": 51,
                "samples": [
                  { "11.01": 29.13, "11.01": 0.04 },
                  { "11.01": 29.13, "11.01": 0.04 }
                ],
                "semantic_type": "\",
                "description": "\n\n",
                "column": "2003-04",
                "properties": {
                  "dtype": "number",
                  "std": 67.65373542948477,
                  "min": 0.0,
                  "max": 368.32,
                  "num_unique_values": 53,
                  "samples": [
                    { "47.54": 0.11, "47.54": 31.12 },
                    { "47.54": 0.11, "47.54": 31.12 }
                  ],
                  "semantic_type": "\",
                  "description": "\n\n",
                  "column": "2004-05",
                  "properties": {
                    "dtype": "number",
                    "std": 101.93487250276844,
                    "min": 0.0,
                    "max": 201.93487250276844,
                    "num_unique_values": 55,
                    "samples": [
                      { "201.93487250276844": 101.93487250276844 },
                      { "201.93487250276844": 101.93487250276844 }
                    ]
                  }
                }
              }
            ]
          }
        }
      }
    ]
  }
}

```



```

1466.23,\n          214.8,\n          401.46\n        ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    }\n  },\\n  {\n    \\"column\\": \"2013-14\",\\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\\n      \\"std\\": 658.429943568493,\n      \\"min\\": 0.0,\n      \\"max\\": 3982.89,\n      \\"num_unique_values\\": 61,\n      \\"samples\\": [\n        567.63,\n        112.23,\n        285.85\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    }\n  },\\n  {\n    \\"column\\": \"2014-15\",\\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\\n      \\"std\\":\n      837.7870598371759,\n      \\"min\\": 0.0,\n      \\"max\\": 4443.26,\n      \\"num_unique_values\\": 62,\n      \\"samples\\": [\n        78.86,\n        1.43,\n        359.34\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    }\n  },\\n  {\n    \\"column\\": \"2015-16\",\\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\\n      \\"std\\": 1335.3077062190832,\n      \\"min\\": 0.0,\n      \\"max\\": 6889.46,\n      \\"num_unique_values\\": 59,\n      \\"samples\\": [\n        456.31,\n        103.02,\n        505.88\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    }\n  },\\n  {\n    \\"column\\": \"2016-17\",\\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\\n      \\"std\\": 1411.9653538026591,\n      \\"min\\": 0.0,\n      \\"max\\": 8684.07,\n      \\"num_unique_values\\": 58,\n      \\"samples\\": [\n        1440.18,\n        180.4,\n        727.22\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    }\n  }\n},\"type\":\"dataframe\",\"variable_name\":\"df\"}

df.tail()

{
  \"summary\": {
    \"name\": \"df\",\n    \"rows\": 5,\n    \"fields\": [
      {\n        \"column\": \"Sector\",\\n        \"properties\": {\n          \\"dtype\\": \"string\",\\n          \\"num_unique_values\\": 5,\n          \\"samples\\\": [\n            \"COIR\",\\n            \"MISCELLANEOUS INDUSTRIES\",\\n            \"CONSTRUCTION (INFRASTRUCTURE) ACTIVITIES\"\n          ],\n          \\"semantic_type\\": \"\",\\n          \\"description\\": \"\"\n        }\n      },\\n      {\n        \"column\": \"2001-02\",\\n        \"properties\": {\n          \\"dtype\\": \"number\",\\n          \\"std\\": 369.54304275686206,\n          \\"min\\": 0.0,\n          \\"max\\": 832.07,\n          \\"num_unique_values\\": 3,\n          \\"samples\\\": [\n            0.0,\n            24.33,\n            832.07\n          ],\n          \\"semantic_type\\": \"\",\\n          \\"description\\": \"\"\n        }\n      },\\n      {\n        \"column\": \"2002-03\",\\n        \"properties\": {\n          \\"dtype\\": \"number\",\\n          \\"std\\": 95.86948487396812,\n          \\"min\\": 0.0,\n          \\"max\\": 221.37,\n          \\"num_unique_values\\": 3,\n          \\"samples\\\": [\n            0.0,\n            51.75,\n            221.37\n          ],\n          \\"semantic_type\\": \"\",\\n          \\"description\\": \"\"\n        }\n      }
    ]
  }
}

```

```

0.0,\n      \\"max\\": 218.76,\n      \\"num_unique_values\\": 4,\n      \\"samples\\": [\n          0.0,\n          218.76,\n          6.3\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n  },\n  {\n    \\"column\\": \"2003-04\",\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\n      \\"std\\": 102.10298957425292,\n      \\"min\\": 0.0,\n      \\"max\\": 235.48,\n      \\"num_unique_values\\": 3,\n      \\"samples\\": [\n          0.0,\n          47.04,\n          235.48\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    },\n    \\"column\\": \"2004-05\",\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\n      \\"std\\": 75.66992421034925,\n      \\"min\\": 0.0,\n      \\"max\\": 152.06,\n      \\"num_unique_values\\": 5,\n      \\"samples\\": [\n          0.47,\n          121.83,\n          0.0\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    },\n    \\"column\\": \"2005-06\",\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\n      \\"std\\": 108.12742746408055,\n      \\"min\\": 0.59,\n      \\"max\\": 228.71,\n      \\"num_unique_values\\": 5,\n      \\"samples\\": [\n          0.59,\n          164.76,\n          0.93\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    },\n    \\"column\\": \"2006-07\",\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\n      \\"std\\": 592.174893481647,\n      \\"min\\": 0.04,\n      \\"max\\": 1392.95,\n      \\"num_unique_values\\": 5,\n      \\"samples\\": [\n          0.04,\n          304.87,\n          64.06\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    },\n    \\"column\\": \"2007-08\",\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\n      \\"std\\": 1668.0891793995909,\n      \\"min\\": 0.01,\n      \\"max\\": 3887.33,\n      \\"num_unique_values\\": 5,\n      \\"samples\\": [\n          0.01,\n          528.42,\n          182.92\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    },\n    \\"column\\": \"2008-09\",\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\n      \\"std\\": 1993.945600444004,\n      \\"min\\": 0.0,\n      \\"max\\": 4657.51,\n      \\"num_unique_values\\": 5,\n      \\"samples\\": [\n          0.0,\n          1549.7,\n          172.7\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    },\n    \\"column\\": \"2009-10\",\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\n      \\"std\\": 2317.3550553745536,\n      \\"min\\": 0.25,\n      \\"max\\": 5466.13,\n      \\"num_unique_values\\": 5,\n      \\"samples\\": [\n          0.25,\n          1147.56,\n          324.56\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    },\n    \\"column\\": \"2010-11\",\n    \\"properties\\": {\n      \\"dtype\\": \"number\",\n      \\"std\\": 780.3020834266688,\n      \\"min\\": 0.1,\n      \\"max\\": 1663.03,\n      \\"num_unique_values\\": 5,\n      \\"samples\\": [\n          0.1,\n          1475.97,\n          675.07\n      ],\n      \\"semantic_type\\": \"\",\\n      \\"description\\": \"\"\n    },\n    \\"column\\": \"2011-12\"\n  }

```

```
\"properties\": {\n    \"Sector\" : {\n        \"dtype\" : \"number\", \"std\" : 1306.327732703398,\n        \"min\" : 0.55, \"max\" : 3140.78,\n        \"num_unique_values\" : 5, \"samples\" : [\n            0.55, 813.38, 386.28\n        ],\n        \"semantic_type\" : \"\",\n        \"description\" : \"\"\n    },\n    \"2000-01\" : {\n        \"column\" : \"2012-13\", \"properties\" : {\n            \"dtype\" : \"number\", \"std\" : 551.5447816088918,\n            \"min\" : 0.15, \"max\" : 1332.49,\n            \"num_unique_values\" : 5, \"samples\" : [\n                0.15, 229.49, 283.89\n            ],\n            \"semantic_type\" : \"\",\n            \"description\" : \"\"\n        },\n        \"column\" : \"2013-14\", \"properties\" : {\n            \"dtype\" : \"number\", \"std\" : 479.19680532115405,\n            \"min\" : 0.54, \"max\" : 1226.05,\n            \"num_unique_values\" : 5, \"samples\" : [\n                0.54, 468.74, 485.37\n            ],\n            \"semantic_type\" : \"\",\n            \"description\" : \"\"\n        },\n        \"column\" : \"2014-15\", \"properties\" : {\n            \"dtype\" : \"number\", \"std\" : 421.73950422031845,\n            \"min\" : 1.36, \"max\" : 870.25,\n            \"num_unique_values\" : 5, \"samples\" : [\n                1.36, 765.88, 870.25\n            ],\n            \"semantic_type\" : \"\",\n            \"description\" : \"\"\n        },\n        \"column\" : \"2015-16\", \"properties\" : {\n            \"dtype\" : \"number\", \"std\" : 1933.7405055022248,\n            \"min\" : 0.0, \"max\" : 4510.71,\n            \"num_unique_values\" : 5, \"samples\" : [\n                0.0, 668.77, 4510.71\n            ],\n            \"semantic_type\" : \"\",\n            \"description\" : \"\"\n        },\n        \"column\" : \"2016-17\", \"properties\" : {\n            \"dtype\" : \"number\", \"std\" : 789.2715126431463,\n            \"min\" : 0.0, \"max\" : 1860.73,\n            \"num_unique_values\" : 5, \"samples\" : [\n                0.0, 296.4, 1860.73\n            ],\n            \"semantic_type\" : \"\",\n            \"description\" : \"\"\n        }\n    },\n    \"type\" : \"dataframe\"\n}\n\ndf.shape\n(63, 18)\n\ndf.info()\n\n<class 'pandas.core.frame.DataFrame'>\nRangeIndex: 63 entries, 0 to 62\nData columns (total 18 columns):\n #   Column   Non-Null Count  Dtype \n--- \n 0   Sector    63 non-null    object \n 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

```

##All the datatypes are correct

```

df.describe()

{"summary": {"name": "df", "rows": 8, "fields": [
    {"column": "2000-01", "properties": {
        "dtype": "number", "std": 284.6279789234439, "min": 0.0, "max": 832.07, "num_unique_values": 7, "samples": [37.75730158730158, 23.50999999999998], "semantic_type": "\\", "description": "\n"}, "column": "2001-02", "properties": {
        "dtype": "number", "std": 296.4451362780846, "min": 0.0, "max": 873.23, "num_unique_values": 7, "samples": [63.93158730158729, 44.83], "semantic_type": "\\", "description": "\n"}, "column": "2002-03", "properties": {
        "dtype": "number", "std": 139.71592781330386, "min": 0.0, "max": 419.96, "num_unique_values": 8, "samples": [42.92571428571428, 11.01], "semantic_type": "\\", "description": "\n"}, "column": "2003-04", "properties": {
        "dtype": "number", "std": 122.50335531707434, "min": 0.0, "max": 368.32, "num_unique_values": 8, "samples": [34.72777777777774, 6.37], "semantic_type": "\\", "description": "\n"}, "column": "2004-05", "properties": {
        "dtype": "number", "std": 122.50335531707434, "min": 0.0, "max": 368.32, "num_unique_values": 8, "samples": [34.72777777777774, 6.37], "semantic_type": "\\", "description": "\n"}]
}
```

```
{\n      "dtype": "number",\n      "std":\n176.5583464852197,\n      "min": 0.0,\n      "max": 527.9,\n      "num_unique_values": 8,\n      "samples": [\n        9.09,\n        63.0\n      ],\n      "semantic_type": "",\n      "description": "\\"\n      }\n    },\n    {\n      "column": "2005-06",\n      "properties":\n      {\n        "dtype": "number",\n        "std":\n463.08795573536133,\n        "min": 0.0,\n        "max": 1359.97,\n        "num_unique_values": 8,\n        "samples": [\n          22.62,\n          63.0\n        ],\n        "semantic_type": "",\n        "description": "\\"\n      }\n    },\n    {\n      "column": "2006-07",\n      "properties":\n      {\n        "dtype": "number",\n        "std":\n1627.5256592816488,\n        "min": 0.0,\n        "max": 4713.78,\n        "num_unique_values": 8,\n        "samples": [\n          25.82,\n          63.0\n        ],\n        "semantic_type": "",\n        "description": "\\"\n      }\n    },\n    {\n      "column": "2007-08",\n      "properties":\n      {\n        "dtype": "number",\n        "std":\n2402.0117528616856,\n        "min": 0.0,\n        "max": 6986.17,\n        "num_unique_values": 8,\n        "samples": [\n          58.82,\n          63.0\n        ],\n        "semantic_type": "",\n        "description": "\\"\n      }\n    },\n    {\n      "column": "2008-09",\n      "properties":\n      {\n        "dtype": "number",\n        "std":\n2110.9962555869183,\n        "min": 0.0,\n        "max": 6183.49,\n        "num_unique_values": 8,\n        "samples": [\n          84.88,\n          63.0\n        ],\n        "semantic_type": "",\n        "description": "\\"\n      }\n    },\n    {\n      "column": "2009-10",\n      "properties":\n      {\n        "dtype": "number",\n        "std":\n1866.8912653928687,\n        "min": 0.0,\n        "max": 5466.13,\n        "num_unique_values": 8,\n        "samples": [\n          69.74,\n          63.0\n        ],\n        "semantic_type": "",\n        "description": "\\"\n      }\n    },\n    {\n      "column": "2010-11",\n      "properties":\n      {\n        "dtype": "number",\n        "std":\n1115.838226017203,\n        "min": 0.0,\n        "max": 3296.09,\n        "num_unique_values": 8,\n        "samples": [\n          58.07,\n          63.0\n        ],\n        "semantic_type": "",\n        "description": "\\"\n      }\n    },\n    {\n      "column": "2011-12",\n      "properties":\n      {\n        "dtype": "number",\n        "std":\n1761.043514093648,\n        "min": 0.0,\n        "max": 5215.98,\n        "num_unique_values": 8,\n        "samples": [\n          129.36,\n          63.0\n        ],\n        "semantic_type": "",\n        "description": "\\"\n      }\n    },\n    {\n      "column": "2012-13",\n      "properties":\n      {\n        "dtype": "number",\n        "std":\n557.4726984126983,\n        "min": 0.0,\n        "max": 1359.97,\n        "num_unique_values": 8,\n        "samples": [\n          22.62,\n          63.0\n        ],\n        "semantic_type": "",\n        "description": "\\"\n      }\n    }\n  ]\n}
```

```

1648.3880813782573, \n          \\"min\\": 0.0, \n          \\"max\\": 4832.98, \n          \\"num_unique_values\\": 8, \n          \\"samples\\": [\n            355.93000000000006, \n              95.41, \n                63.0\n              ], \n            \\"semantic_type\\": \"\", \n              \\"description\\": \"\"\n            }, \n            {\n              \\"column\\": \"2013-14\", \n                \\"properties\\": {\n                  \\"dtype\\": \"number\", \n                    \\"std\\": 1343.4886091439434, \n                      \\"min\\": 0.0, \n                        \\"max\\": 3982.89, \n                          \\"num_unique_values\\": 8, \n                            \\"samples\\": [\n                              385.7034920634921, \n                                113.78, \n                                  63.0\n                                  ], \n                                \\"semantic_type\\": \"\", \n                                  \\"description\\": \"\"\n                                }, \n                                {\n                                  \\"column\\": \"2014-15\", \n                                    \\"properties\\": {\n                                      \\"dtype\\": \"number\", \n                                        \\"std\\": 1491.0659401286234, \n                                          \\"min\\": 0.0, \n                                            \\"max\\": 4443.26, \n                                              \\"num_unique_values\\": 8, \n                                                \\"samples\\": [\n                                                  490.95984126984126, \n                                                    177.22, \n                                                      63.0\n                                                      ], \n                                                      \\"semantic_type\\": \"\", \n                                                        \\"description\\": \"\"\n                                                        }, \n                                                        {\n                                                          \\"column\\": \"2015-16\", \n                                                            \\"properties\\": {\n                                                              \\"dtype\\": \"number\", \n                                                                \\"std\\": 2340.8235985173587, \n                      \\"min\\": 0.0, \n                        \\"max\\": 6889.46, \n                          \\"num_unique_values\\": 8, \n                            \\"samples\\": [\n                              634.9363492063493, \n                                159.13, \n                                  63.0\n                                  ], \n                                  \\"semantic_type\\": \"\", \n                                    \\"description\\": \"\"\n                                    }, \n                                    {\n                                      \\"column\\": \"2016-17\", \n                                        \\"properties\\": {\n                                          \\"dtype\\": \"number\", \n                                            \\"std\\": 2958.664929487532, \n                                              \\"min\\": 0.0, \n                                                \\"max\\": 8684.07, \n                                                  \\"num_unique_values\\": 8, \n                                                    \\"samples\\": [\n                                                      690.131111111112, \n                                                        110.86, \n                                                          63.0\n                                                          ], \n                                                          \\"semantic_type\\": \"\", \n                                                            \\"description\\": \"\"\n                                                            }\n                                                            ]\n                                                        ], \n                                                        \\"type\\": \"dataframe\"}

```

Here from above we can understand that the data of FDI is largliy skewed specifically right skewed as mean is greater than median(50%)

```
df.isnull().sum()
```

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
```

##There are no null values

```
df.duplicated().sum()
0
```

##From above we can conclude that there are no duplicates involved here!!!

```
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')

df1 = df.set_index('Sector') #####making Sector as index column
df1.head()

{"summary": {"name": "df1", "rows": 63, "fields": [
    {"column": "Sector", "properties": {"dtype": "string", "num_unique_values": 63, "samples": ["CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects", "TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)", "METALLURGICAL INDUSTRIES"], "semantic_type": "\\", "description": "\n", "properties": {"2000-01": {"dtype": "number", "std": 112.22785996479905, "min": 0.0, "max": 832.07, "num_unique_values": 41, "samples": [60.04, 1.42, 177.69]}, "2001-02": {"dtype": "number", "std": 157.87873650170664, "min": 0.0, "max": 873.23, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}, "2002-03": {"dtype": "number", "std": 86.6064393738354, "min": 0.0, "max": 321.2, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}], "semantic_type": "\\", "description": "\n", "properties": {"2000-01": {"dtype": "number", "std": 112.22785996479905, "min": 0.0, "max": 832.07, "num_unique_values": 41, "samples": [60.04, 1.42, 177.69]}, "2001-02": {"dtype": "number", "std": 157.87873650170664, "min": 0.0, "max": 873.23, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}, "2002-03": {"dtype": "number", "std": 86.6064393738354, "min": 0.0, "max": 321.2, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}}], "column": "Sector", "properties": {"dtype": "string", "num_unique_values": 63, "samples": ["CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects", "TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)", "METALLURGICAL INDUSTRIES"], "semantic_type": "\\", "description": "\n", "properties": {"2000-01": {"dtype": "number", "std": 112.22785996479905, "min": 0.0, "max": 832.07, "num_unique_values": 41, "samples": [60.04, 1.42, 177.69]}, "2001-02": {"dtype": "number", "std": 157.87873650170664, "min": 0.0, "max": 873.23, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}, "2002-03": {"dtype": "number", "std": 86.6064393738354, "min": 0.0, "max": 321.2, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}], "semantic_type": "\\", "description": "\n", "properties": {"2000-01": {"dtype": "number", "std": 112.22785996479905, "min": 0.0, "max": 832.07, "num_unique_values": 41, "samples": [60.04, 1.42, 177.69]}, "2001-02": {"dtype": "number", "std": 157.87873650170664, "min": 0.0, "max": 873.23, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}, "2002-03": {"dtype": "number", "std": 86.6064393738354, "min": 0.0, "max": 321.2, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}}], "column": "2000-01", "properties": {"dtype": "number", "num_unique_values": 41, "samples": [60.04, 1.42, 177.69], "semantic_type": "\\", "description": "\n", "properties": {"2000-01": {"dtype": "number", "std": 112.22785996479905, "min": 0.0, "max": 832.07, "num_unique_values": 41, "samples": [60.04, 1.42, 177.69]}}, "semantic_type": "\\", "description": "\n", "properties": {"2000-01": {"dtype": "number", "std": 112.22785996479905, "min": 0.0, "max": 832.07, "num_unique_values": 41, "samples": [60.04, 1.42, 177.69]}]}, "column": "2001-02", "properties": {"dtype": "number", "num_unique_values": 46, "samples": [32.12, 5.28, 16.7], "semantic_type": "\\", "description": "\n", "properties": {"2001-02": {"dtype": "number", "std": 157.87873650170664, "min": 0.0, "max": 873.23, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}}, "semantic_type": "\\", "description": "\n", "properties": {"2001-02": {"dtype": "number", "std": 157.87873650170664, "min": 0.0, "max": 873.23, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}]}, "column": "2002-03", "properties": {"dtype": "number", "num_unique_values": 46, "samples": [32.12, 5.28, 16.7], "semantic_type": "\\", "description": "\n", "properties": {"2002-03": {"dtype": "number", "std": 86.6064393738354, "min": 0.0, "max": 321.2, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}}, "semantic_type": "\\", "description": "\n", "properties": {"2002-03": {"dtype": "number", "std": 86.6064393738354, "min": 0.0, "max": 321.2, "num_unique_values": 46, "samples": [32.12, 5.28, 16.7]}}]
```

```
  "max": 419.96, "num_unique_values": 51, "samples": [{"column": "2003-04", "description": "11.01", "properties": {"min": 0.0, "max": 368.32, "std": 67.65373542948477}, "semantic_type": "number"}, {"column": "2004-05", "description": "47.54", "properties": {"min": 0.0, "max": 527.9, "std": 101.93487250276844}, "semantic_type": "number"}, {"column": "2005-06", "description": "139.0", "properties": {"min": 0.0, "max": 1359.97, "std": 206.4369666349658}, "semantic_type": "number"}, {"column": "2006-07", "description": "94.33", "properties": {"min": 0.0, "max": 4713.78, "std": 686.7831152739163}, "semantic_type": "number"}, {"column": "2007-08", "description": "169.94", "properties": {"min": 0.0, "max": 6986.17, "std": 1026.2499354787049}, "semantic_type": "number"}, {"column": "2008-09", "description": "1175.75", "properties": {"min": 0.0, "max": 6183.49, "std": 1134.6490399861423}, "semantic_type": "number"}, {"column": "2009-10", "description": "349.29", "properties": {"min": 0.0, "max": 5466.13, "std": 926.8146261766324}, "semantic_type": "number"}, {"column": "2010-11", "description": "419.88", "properties": {"min": 0.0, "max": 3296.09, "std": 627.1411390805016}, "semantic_type": "number"}], "num_unique_values": 61, "samples": [29.13, 0.04, 31.12, 200.38, 102.78, 149.13, 12.09, 27.58, 0.39, 5.01, 278.89, 265.53, 274.84, 556.43]
```

```

    "semantic_type": "\",\n      "description": \"\\n      \"},\n      {\n        \"column\": \"2011-12\", \"properties\": {\n          \"dtype\": \"number\", \"std\": 1031.4740556792376,\n          \"min\": 0.0, \"max\": 5215.98,\n          \"num_unique_values\": 60,\n          \"samples\": [\n            1786.14, 2029.98, 222.08\n          ]},\n        \"semantic_type\": "\",\n        \"description\": \"\\n      \"},\n        {\n          \"column\": \"2012-13\", \"properties\": {\n            \"dtype\": \"number\", \"std\": 778.0913683930326,\n            \"min\": 0.0, \"max\": 4832.98,\n            \"num_unique_values\": 59,\n            \"samples\": [\n              1466.23, 214.8, 401.46\n            ]},\n          \"semantic_type\": "\",\n          \"description\": \"\\n      \"},\n          {\n            \"column\": \"2013-14\", \"properties\": {\n              \"dtype\": \"number\", \"std\": 658.429943568493,\n              \"min\": 0.0, \"max\": 3982.89,\n              \"num_unique_values\": 61,\n              \"samples\": [\n                567.63, 112.23, 285.85\n              ]},\n            \"semantic_type\": "\",\n            \"description\": \"\\n      \"},\n            {\n              \"column\": \"2014-15\", \"properties\": {\n                \"dtype\": \"number\", \"std\": 837.7870598371759,\n                \"min\": 0.0, \"max\": 4443.26,\n                \"num_unique_values\": 62,\n                \"samples\": [\n                  78.86, 1.43, 359.34\n                ]},\n              \"semantic_type\": "\",\n              \"description\": \"\\n      \"},\n              {\n                \"column\": \"2015-16\", \"properties\": {\n                  \"dtype\": \"number\", \"std\": 1335.3077062190832,\n                  \"min\": 0.0,\n                  \"max\": 6889.46,\n                  \"num_unique_values\": 59,\n                  \"samples\": [\n                    456.31, 103.02,\n                    505.88\n                  ]},\n                \"semantic_type\": "\",\n                \"description\": \"\\n      \"},\n                {\n                  \"column\": \"2016-17\", \"properties\": {\n                    \"dtype\": \"number\", \"std\": 1411.9653538026591,\n                    \"min\": 0.0,\n                    \"max\": 8684.07,\n                    \"num_unique_values\": 58,\n                    \"samples\": [\n                      1440.18, 180.4,\n                      727.22\n                    ]},\n                  \"semantic_type\": "\",\n                  \"description\": \"\\n      \"},\n                  {\n                    \"column\": \"2017-18\", \"properties\": {\n                      \"dtype\": \"number\", \"std\": 1411.9653538026591,\n                      \"min\": 0.0,\n                      \"max\": 8684.07,\n                      \"num_unique_values\": 58,\n                      \"samples\": [\n                        1440.18, 180.4,\n                        727.22\n                      ]}\n                    }\n                  },\n                  \"type\": \"dataframe\", \"variable_name\": \"df1\"}\n\ndf['Sector'].nunique()

```

63

Univariate Plots

Distrubtion of FDI For Each Year

```
def histplots(df1):
    columns = df1.columns
    n_cols = 3
    n_rows = -(-len(columns) // n_cols)

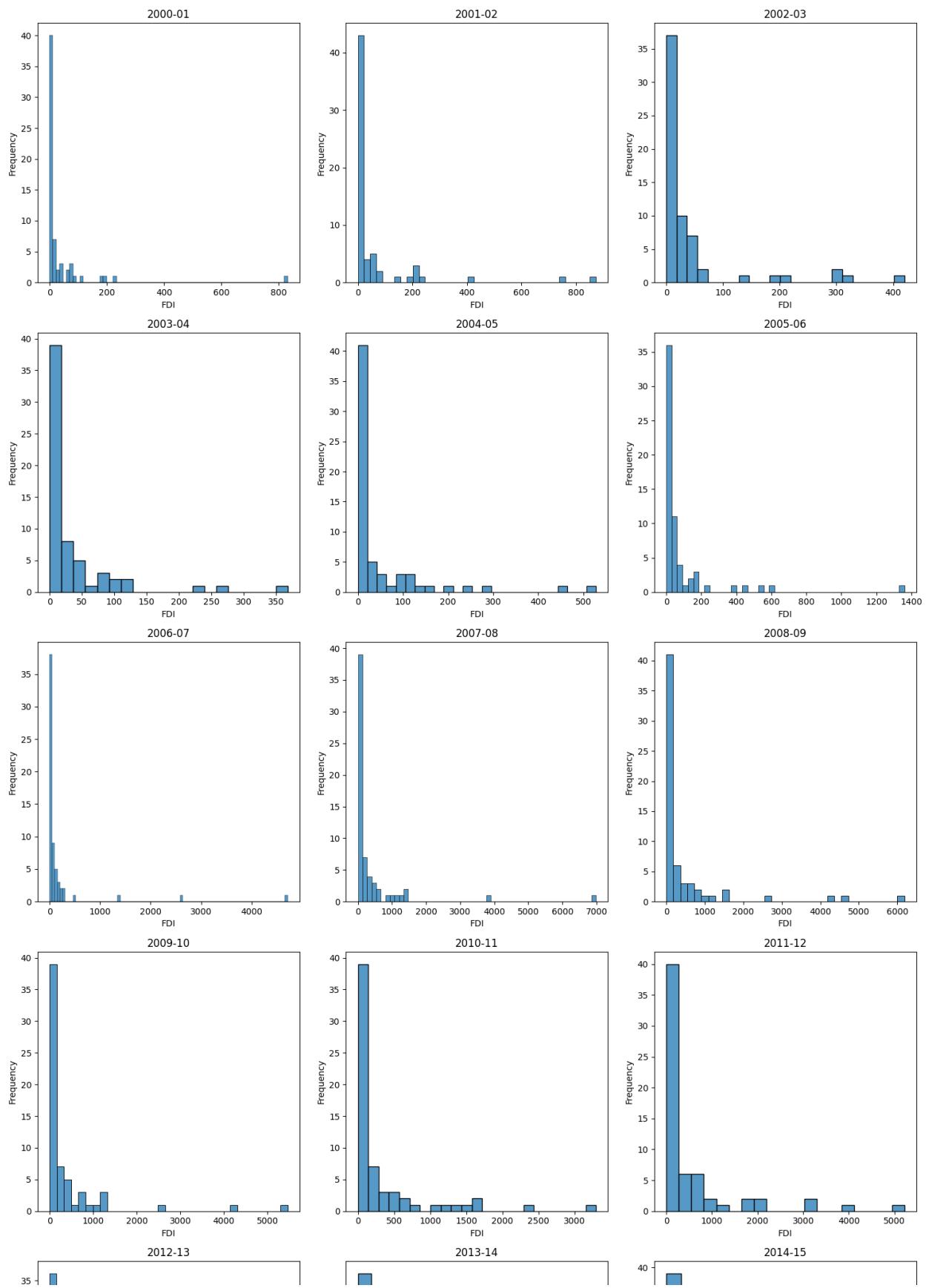
    fig, axes = plt.subplots(n_rows, n_cols, figsize=(15, 5*n_rows))
    axes = axes.flatten()

    for i, col in enumerate(columns):
        sns.histplot(df1[col], ax=axes[i])
        axes[i].set_title(col)
        axes[i].set_xlabel("FDI")
        axes[i].set_ylabel('Frequency')

    # Remove empty subplots
    for j in range(i + 1, n_rows * n_cols):
        fig.delaxes(axes[j])

    plt.tight_layout()
    plt.show()

histplots(df1)
```



```
df2 = df1.T
df2.head()

{"type": "dataframe", "variable_name": "df2"}
```

df2.columns

Index(['METALLURGICAL INDUSTRIES', 'MINING', 'POWER',
 'NON-CONVENTIONAL ENERGY', 'COAL PRODUCTION', 'PETROLEUM &
NATURAL GAS',
 'BOILERS AND STEAM GENERATING PLANTS',
 'PRIME MOVER (OTHER THAN ELECTRICAL GENERATORS)',
 'ELECTRICAL EQUIPMENTS', 'COMPUTER SOFTWARE & HARDWARE',
'ELECTRONICS',
 'TELECOMMUNICATIONS',
 'INFORMATION & BROADCASTING (INCLUDING PRINT MEDIA)',
 'AUTOMOBILE INDUSTRY', 'AIR TRANSPORT (INCLUDING AIR FREIGHT)',
 'SEA TRANSPORT', 'PORTS', 'RAILWAY RELATED COMPONENTS',
 'INDUSTRIAL MACHINERY', 'MACHINE TOOLS', 'AGRICULTURAL
MACHINERY',
 'EARTH-MOVING MACHINERY',
 'MISCELLANEOUS MECHANICAL & ENGINEERING INDUSTRIES',
 'COMMERCIAL, OFFICE & HOUSEHOLD EQUIPMENTS',
 'MEDICAL AND SURGICAL APPLIANCES', 'INDUSTRIAL INSTRUMENTS',
 'SCIENTIFIC INSTRUMENTS',
 'MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS',
'FERTILIZERS',
 'CHEMICALS (OTHER THAN FERTILIZERS)', 'PHOTOGRAPHIC RAW FILM
AND PAPER',
 'DYE-STUFFS', 'DRUGS & PHARMACEUTICALS',
 'TEXTILES (INCLUDING DYED,PRINTED)',
 'PAPER AND PULP (INCLUDING PAPER PRODUCTS)', 'SUGAR',
 'FERMENTATION INDUSTRIES', 'FOOD PROCESSING INDUSTRIES',
 'VEGETABLE OILS AND VANASPATI',
 'SOAPs, COSMETICS & TOILET PREPARATIONS', 'RUBBER GOODS',
 'LEATHER,LEATHER GOODS AND PICKERS', 'GLUE AND GELATIN',
'GLASS',
 'CERAMICS', 'CEMENT AND GYPSUM PRODUCTS', 'TIMBER PRODUCTS',
 'DEFENCE INDUSTRIES', 'CONSULTANCY SERVICES',
 'SERVICES SECTOR (Fin.,Banking,Insurance,Non
Fin/Business,Outsourcing,R&D,Courier,Tech. Testing and Analysis,
Other)',
 'HOSPITAL & DIAGNOSTIC CENTRES', 'EDUCATION', 'HOTEL &
TOURISM',
 'TRADING', 'RETAIL TRADING', 'AGRICULTURE SERVICES',
 'DIAMOND,GOLD ORNAMENTS',
 'TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)',
 'PRINTING OF BOOKS (INCLUDING LITHO PRINTING INDUSTRY)',
'COIR',
 'CONSTRUCTION (INFRASTRUCTURE) ACTIVITIES',

```

'CONSTRUCTION DEVELOPMENT: Townships, housing, built-up
infrastructure and construction-development projects',
'MISCELLANEOUS INDUSTRIES'],
dtype='object', name='Sector')

df2.rename(columns={'CONSTRUCTION DEVELOPMENT: Townships, housing,
built-up infrastructure and construction-development
projects':'CONSTRUCTION DEVELOPMENT',"SERVICES SECTOR
(Fin.,Banking,Insurance,Non Fin/Business,Outsourcing,R&D,Courier,Tech.
Testing and Analysis, Other)": "SERVICES SECTOR"}, inplace=True)
df2.head()

{"type": "dataframe", "variable_name": "df2"}

```

FDI Over The Years For Each Sector

```

def scatterp(df2):
    columns = df2.columns
    n_cols = 3
    n_rows = -(-len(columns) // n_cols)

    fig, axes = plt.subplots(n_rows, n_cols, figsize=(20, 6*n_rows))
    axes = axes.flatten()

    for i, col in enumerate(columns):
        axes[i].plot(df2.index.values, df2[col])
        axes[i].set_title("FDI for " + col)
        axes[i].set_xlabel("Year")
        axes[i].set_xticklabels(df2.index.values, rotation=45)
        axes[i].set_ylabel("FDI")
        axes[i].grid(True)

    # Remove empty subplots
    for j in range(i + 1, n_rows * n_cols):
        fig.delaxes(axes[j])

    plt.tight_layout()
    plt.show()

scatterp(df2)

```



The above plot shows us the how FDI changed over the years for each sector

Year wise total FDI

```
year_total = df2.sum(axis=1)
year_total = pd.DataFrame(year_total, columns=['Total_FDI'])
year_total.sort_values(by='Total_FDI', ascending=False)

{"summary": {"name": "year_total", "rows": 17,
 "fields": [{"column": "Total_FDI", "properties": {"dtype": "number", "std": 14221.119667993022, "min": 2187.85, "max": 43478.26, "num_unique_values": 17, "samples": [43478.26, 40000.990000000005, 25834.380000000005], "semantic_type": "\\", "description": "\""}, "type": "dataframe"}]}
```

#Year wise Average FDI

```
year_avg = df2.mean(axis=1)
year_avg = pd.DataFrame(year_avg, columns=['Avg_FDI'])
year_avg.sort_values(by='Avg_FDI', ascending=False)

{"summary": {"name": "year_avg", "rows": 17,
 "fields": [{"column": "Avg_FDI", "properties": {"dtype": "number", "std": 225.73205822211145, "min": 34.72777777777774, "max": 690.131111111112, "num_unique_values": 17, "samples": [690.131111111112, 634.9363492063493, 410.0695238095239], "semantic_type": "\\", "description": "\""}, "type": "dataframe"}]}
```

Top 10 Sector wise total FDI

```
df3 = df2.T
sec_total = df3.sum(axis=1)
sec_total = pd.DataFrame(sec_total, columns=['Total_FDI'])
sec_total_top = sec_total.sort_values(by='Total_FDI', ascending=False).head(10)
sec_total_top

{"summary": {"name": "sec_total_top", "rows": 10,
 "fields": [{"column": "Sector", "properties": {"semantic_type": "\\", "description": "\""}, "type": "dataframe"}]}
```

```

  "properties": {
    "dtype": "string",
    "num_unique_values": 10,
    "samples": [
      "POWER",
      "COMPUTER SOFTWARE & HARDWARE",
      "DRUGS & PHARMACEUTICALS"
    ],
    "semantic_type": "\",
    "description": "\n      }},\n    {\n      \"column\": \"Total_FDI\",
      \"properties\": {\n        \"dtype\": \"number\",
        \"std\": 14447.683699221052,
        \"min\": 10330.539999999999,
        \"max\": 59476.489999999999,
        \"num_unique_values\": 10,
        \"samples\": [
          11589.130000000001,
          24669.489999999998,
          14706.900000000001
        ]
      },
      "semantic_type": "\",
      "description": "\n      }\n    ]\n  }",
    "type": "dataframe",
    "variable_name": "sec_total_top"
  }

```

From above data we can see that the service sector was able to draw more FDI than other sectors

Bottom 10 Sector wise total FDI

```

sec_total_tail = sec_total.sort_values(by='Total_FDI',
                                       ascending=False).tail(10)
sec_total_tail

{
  "summary": {
    "name": "sec_total_tail",
    "rows": 10,
    "fields": [
      {
        "column": "Sector",
        "properties": {
          "dtype": "string",
          "num_unique_values": 10,
          "samples": [
            "DEFENCE INDUSTRIES",
            "GLUE AND GELATIN",
            "PHOTOGRAPHIC RAW FILM AND PAPER"
          ],
          "semantic_type": "\",
          "description": "\n      }},\n    {\n      \"column\": \"Total_FDI\",
        \"properties\": {\n          \"dtype\": \"number\",
            \"std\": 55.07763176745266,
            \"min\": 4.0600000000000005,
            \"max\": 157.67999999999998,
            \"num_unique_values\": 10,
            \"samples\": [
              5.12,
              128.39,
              67.28
            ],
            \"semantic_type\": "\",
            \"description\": "\n      }\n    ]\n  }",
        "type": "dataframe",
        "variable_name": "sec_total_tail"
      }
    ]
  }
}

```

From above data we can see that the COIR sector was able to draw least FDI than other sectors

Top 10 Sector wise Average FDI

```

sec_avg = df3.mean(axis=1)
sec_avg = pd.DataFrame(sec_avg, columns=['Avg_FDI'])
sec_avg_top = sec_avg.sort_values(by='Avg_FDI',

```

```
ascending=False).head(10)
sec_avg_top

{"summary": {"name": "sec_avg_top", "rows": 10, "fields": [{"column": "Sector", "dtype": "string", "num_unique_values": 10, "samples": ["POWER", "COMPUTER SOFTWARE & HARDWARE", "DRUGS & PHARMACEUTICALS"]}, {"column": "Avg_FDI", "properties": {"number": 607.6788235294117, "std": 849.8637470130031, "min": 607.6788235294117, "max": 3498.617058823529, "num_unique_values": 10, "samples": [681.7135294117647, 865.1117647058825]}], "semantic_type": "\\", "description": "\n"}, {"column": "Avg_FDI", "properties": {"number": 1451.1464705882352, "std": 1451.1464705882352, "min": 1451.1464705882352, "max": 1451.1464705882352, "num_unique_values": 10, "samples": [1451.1464705882352, 1451.1464705882352]}], "semantic_type": "\\", "description": "\n"}}, "type": "dataframe", "variable_name": "sec_avg_top"}
```

Bottom 10 Sector wise Average FDI

```
sec_avg_bottom = sec_avg.sort_values(by='Avg_FDI',  
ascending=False).tail(10)  
sec_avg_bottom  
  
{ "summary": { \n    \"name\": \"sec_avg_bottom\", \n    \"rows\": 10, \n    \"fields\": [ \n        { \n            \"column\": \"Sector\", \n            \"properties\": { \n                \"dtype\": \"string\", \n                \"num_unique_values\": 10, \n                \"samples\": [ \n                    \"DEFENCE INDUSTRIES\", \n                    \"GLUE AND GELATIN\", \n                    \"PHOTOGRAPHIC RAW FILM AND PAPER\" \n                ], \n                \"semantic_type\": \"\", \n                \"description\": \"\", \n                \"column\": \"Avg_FDI\", \n                \"properties\": { \n                    \"dtype\": \"number\", \n                    \"std\": 3.2398606922030972, \n                    \"min\": 0.23882352941176474, \n                    \"max\": 9.275294117647057, \n                    \"num_unique_values\": 10, \n                    \"samples\": [ \n                        0.3011764705882353, \n                        7.55235294117647, \n                        3.9576470588235293 \n                    ], \n                    \"semantic_type\": \"\", \n                    \"description\": \"\" \n                } \n            } \n        } \n    ] \n}, \n    \"type\": \"dataframe\", \n    \"variable name\": \"sec avg bottom\" \n}
```

Yearwise Total FDI in INDIA

```
sns.set(style="whitegrid")  
plt.figure(figsize=(10, 6))
```

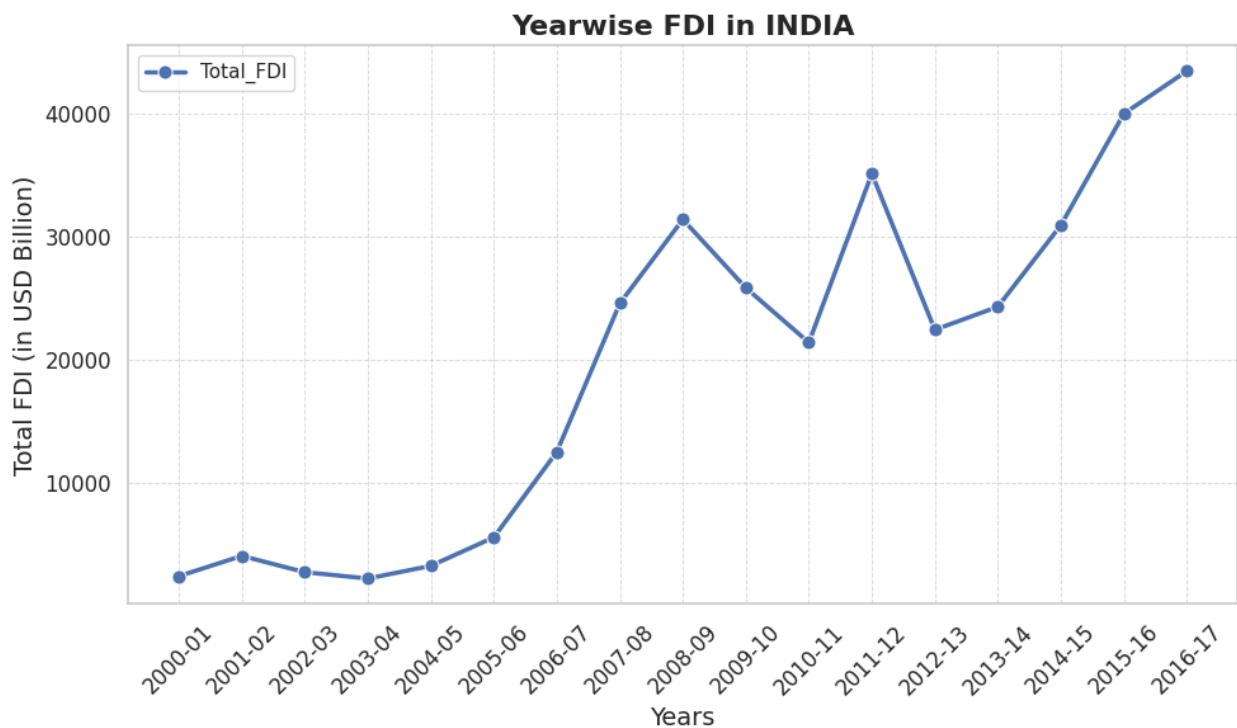
```

sns.lineplot(data=year_total, marker='o', color="#1f77b4",
markersize=8, linewidth=2.5)

plt.title('Yearwise FDI in INDIA', fontsize=16, fontweight='bold')
plt.xlabel('Years', fontsize=14)
plt.ylabel('Total FDI (in USD Billion)', fontsize=14)
plt.xticks(rotation=45, fontsize=12)
plt.yticks(fontsize=12)

plt.grid(True, linestyle='--', linewidth=0.7, alpha=0.7)
plt.tight_layout()
plt.show()

```



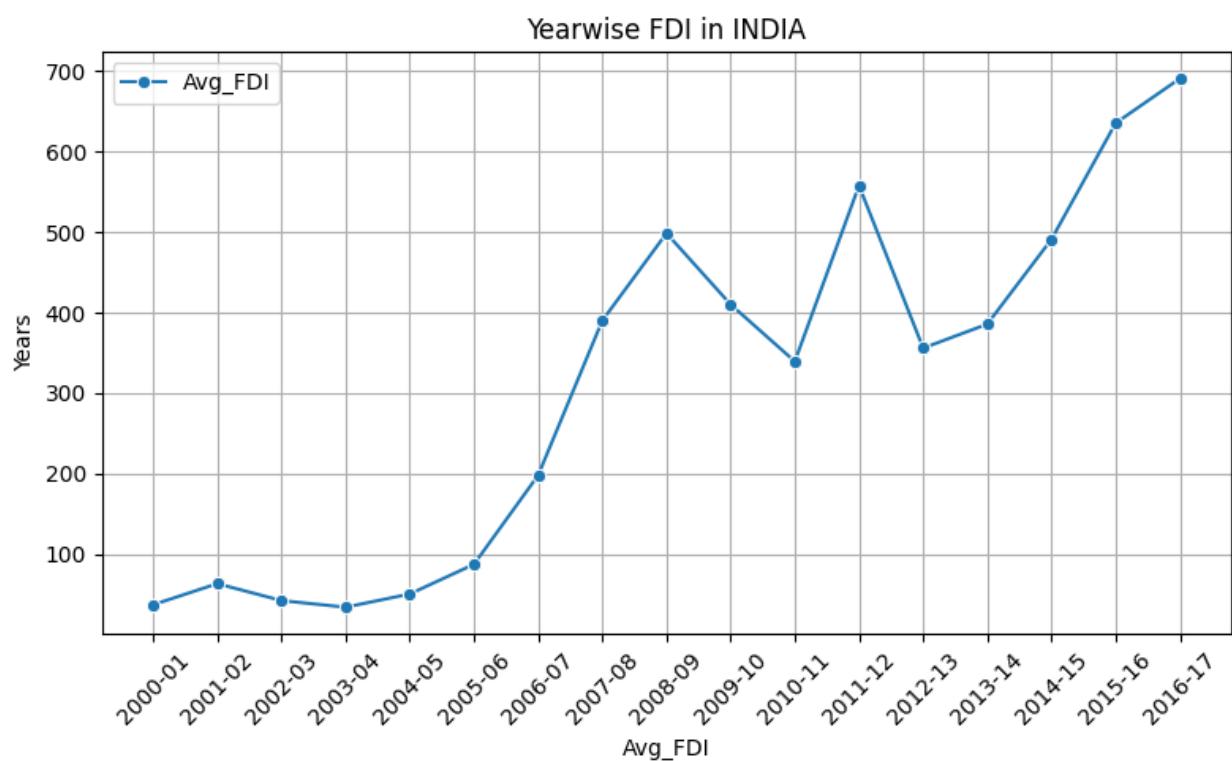
Conclusions:

- 1.The combination of economic reforms, liberalization policies, robust economic growth, and sector-specific developments led to a significant increase in FDI in India on 2006 onwards till 2008-09.
- 2.Again there was steep decrease in FDI for couple of years due to global financial crisis.
- 3.Again Following increase and decrease trend, from 2012-13 till today FDI is increasing rapidly.

Yearwise Avg_FDI in INDIA

```
plt.figure(figsize=(8, 5))
sns.lineplot(data=year_avg, marker='o', color='b')
plt.title('Yearwise FDI in INDIA')
plt.xlabel('Avg_FDI')
plt.ylabel('Years')
plt.grid(True)
plt.xticks(rotation=45)

plt.tight_layout()
plt.show()
```



Top10Sectors FDI in INDIA

```
sns.set(style="whitegrid")

plt.figure(figsize=(13, 6))

sns.lineplot(data=sec_total_top, marker='o', color='#ff7f0e',
markersize=8, linewidth=2.5)

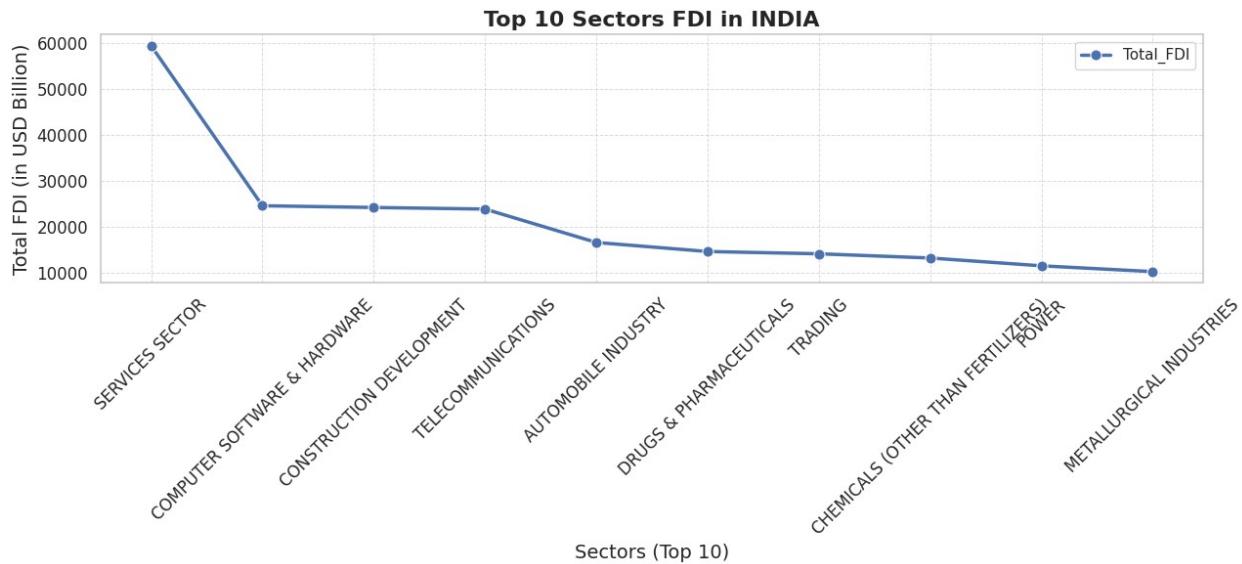
plt.title('Top 10 Sectors FDI in INDIA', fontsize=16,
```

```

fontweight='bold')
plt.xlabel('Sectors (Top 10)', fontsize=14)
plt.ylabel('Total FDI (in USD Billion)', fontsize=14)
plt.xticks(rotation=45, fontsize=12)
plt.yticks(fontsize=12)

plt.grid(True, linestyle='--', linewidth=0.7, alpha=0.7)
plt.tight_layout()
plt.show()

```



Bottom 10 Sectors FDI in INDIA

```

sns.set(style="whitegrid")

plt.figure(figsize=(13, 8))

sns.lineplot(data=sec_total_tail, marker='o', color='#2ca02c',
markersize=8, linewidth=2.5)

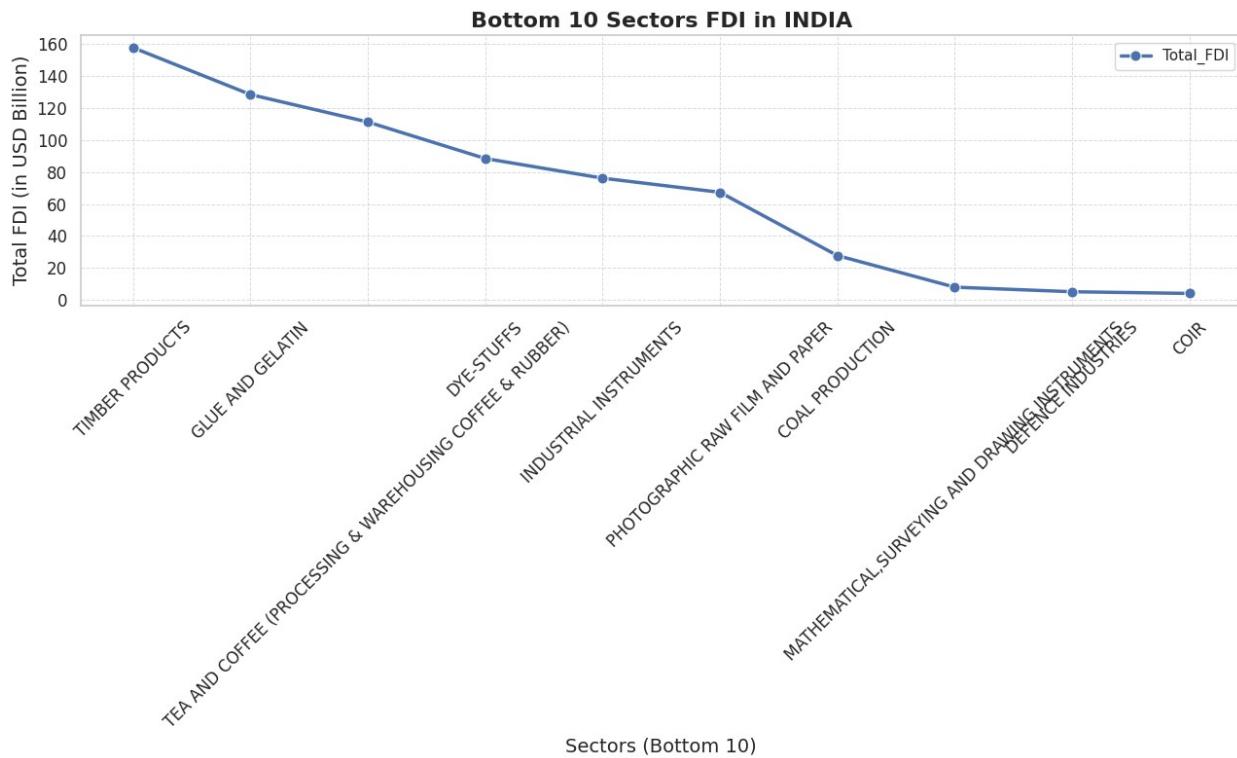
plt.title('Bottom 10 Sectors FDI in INDIA', fontsize=16,
fontweight='bold')
plt.xlabel('Sectors (Bottom 10)', fontsize=14)
plt.ylabel('Total FDI (in USD Billion)', fontsize=14)

plt.xticks(rotation=45, fontsize=12)
plt.yticks(fontsize=12)

plt.grid(True, linestyle='--', linewidth=0.7, alpha=0.7)

```

```
plt.tight_layout()
plt.show()
```



Key observations from the sector-wise analysis:

Services sector dominance: The services sector, including financial, banking, insurance, and outsourcing, has received the highest FDI.

Technology focus: Computer Software & Hardware and Telecommunications are in the top 3, indicating strong foreign interest in India's tech industry.

Infrastructure importance: Construction Development and Automobile Industry feature prominently, highlighting the focus on infrastructure and manufacturing.

Diverse appeal: The top 10 list includes sectors from various industries, showing India's diverse economic appeal to foreign investors.

```
df3.head()
```

```
{"summary": {"\n    "name": "df3",\n    "rows": 63,\n    "fields": [\n        {\n            "column": "Sector",\n            "properties": {\n                "dtype": "string",\n                "num_unique_values": 63,\n                "samples": ["CONSTRUCTION DEVELOPMENT"]\n            }\n        }\n    ]\n}
```

```

\"TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)\",\n
\"METALLURGICAL INDUSTRIES\"\n      ],\n      \"semantic_type\":\n      \"\",,\n      \"description\": \"\\n      }\\n      },\\n      {\n      \"column\": \"2000-01\",\\n      \"properties\": {\n      \"number\":,\n      \"std\": 112.22785996479905,\\n      \"min\":\n      0.0,\\n      \"max\": 832.07,\\n      \"num_unique_values\": 41,\\n      \"samples\": [\n      60.04,\\n      1.42,\\n      177.69\n      ],\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n      {\n      \"column\": \"2001-02\",\\n      \"properties\": {\n      \"dtype\": \"number\",\\n      \"std\":\n      157.87873650170664,\\n      \"min\": 0.0,\\n      \"max\": 873.23,\\n      \"num_unique_values\": 46,\\n      \"samples\": [\n      32.12,\\n      5.28,\n      16.7\n      ],\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n      {\n      \"column\": \"2002-03\",\\n      \"properties\": {\n      \"dtype\": \"number\",\\n      \"std\": 86.6064393738354,\\n      \"min\": 0.0,\\n      \"max\": 419.96,\\n      \"num_unique_values\": 51,\\n      \"samples\": [\n      29.13,\\n      0.04,\\n      11.01\n      ],\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n      {\n      \"column\": \"2003-04\",\\n      \"properties\": {\n      \"dtype\": \"number\",\\n      \"std\":\n      67.65373542948477,\\n      \"min\": 0.0,\\n      \"max\": 368.32,\\n      \"num_unique_values\": 53,\\n      \"samples\": [\n      47.54,\\n      0.11,\n      31.12\n      ],\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n      {\n      \"column\": \"2004-05\",\\n      \"properties\": {\n      \"dtype\": \"number\",\\n      \"std\": 101.93487250276844,\\n      \"min\": 0.0,\\n      \"max\": 527.9,\\n      \"num_unique_values\": 58,\\n      \"samples\": [\n      200.38,\\n      102.78,\\n      139.0\n      ],\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n      {\n      \"column\": \"2005-06\",\\n      \"properties\": {\n      \"dtype\": \"number\",\\n      \"std\": 206.4369666349658,\\n      \"min\": 0.0,\\n      \"max\": 1359.97,\\n      \"num_unique_values\": 57,\\n      \"samples\": [\n      149.13,\\n      12.09,\\n      94.33\n      ],\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n      {\n      \"column\": \"2006-07\",\\n      \"properties\": {\n      \"dtype\": \"number\",\\n      \"std\":\n      686.7831152739163,\\n      \"min\": 0.0,\\n      \"max\": 4713.78,\\n      \"num_unique_values\": 56,\\n      \"samples\": [\n      169.94,\\n      87.71,\n      27.58\n      ],\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n      {\n      \"column\": \"2007-08\",\\n      \"properties\": {\n      \"dtype\": \"number\",\\n      \"std\":\n      1026.2499354787049,\\n      \"min\": 0.0,\\n      \"max\": 6986.17,\\n      \"num_unique_values\": 61,\\n      \"samples\": [\n      1175.75,\\n      1405.04,\n      0.39\n      ],\n      \"semantic_type\": \"\",\\n      \"description\": \"\\n      }\\n      {\n      \"column\": \"2008-09\",\\n      \"properties\":\n
```



```

0      0.000000
1      0.000000
2      0.000000
3      0.000000
4      0.000000
...
1066   -0.567055
1067   0.000000
1068   -0.587486
1069   -0.065837
1070   -0.556798
Name: Growth Rate, Length: 1071, dtype: float64

df_long['Moving Average'] = df_long.groupby('Sector')
[FDI].transform(lambda x: x.rolling(window=3, min_periods=1).mean())
df_long['Moving Average']

0      22.690000
1      1.320000
2      89.420000
3      0.000000
4      0.000000
...
1066   82.853333
1067   0.453333
1068   2413.896667
1069   328.943333
1070   577.016667
Name: Moving Average, Length: 1071, dtype: float64

def identify_trend(series):
    return 'Uptrend' if series.mean() > 0 else 'Downtrend'

trend_df = df_long.groupby('Sector').apply(lambda x:
identify_trend(x['Growth Rate'])).reset_index(name='Trend')
trend_df

{"summary": {"name": "trend_df", "rows": 63,
"fields": [{"column": "Sector", "dtype": "string",
"num_unique_values": 63, "samples": ["TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)", "AGRICULTURAL MACHINERY"], "semantic_type": "\", \"description\": \"\\n\", \"properties\": {\"Trend\": {\"column\": \"Trend\", \"properties\": {\"dtype\": \"category\", \"num_unique_values\": 1, \"samples\": [\"Uptrend\", \"Downtrend\"]}}, \"semantic_type\": \"\\n\", \"description\": \"\\n\"}}}], "type": "dataframe", "variable_name": "trend_df"}}

```

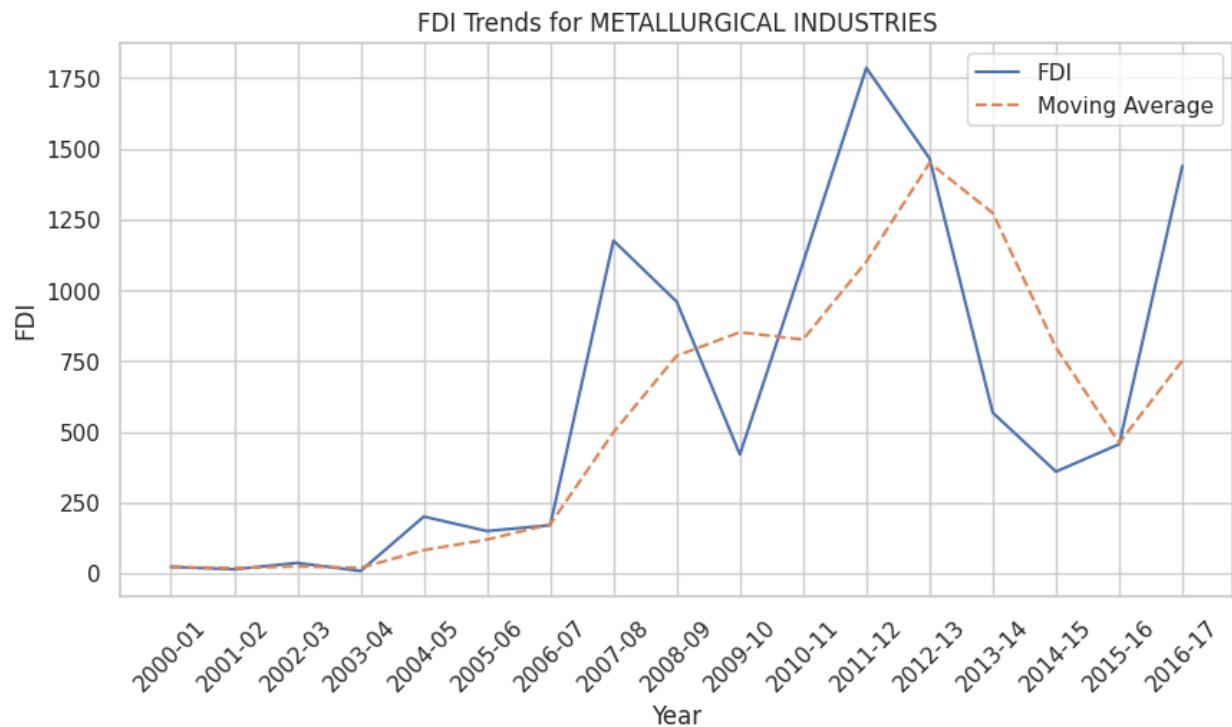
```

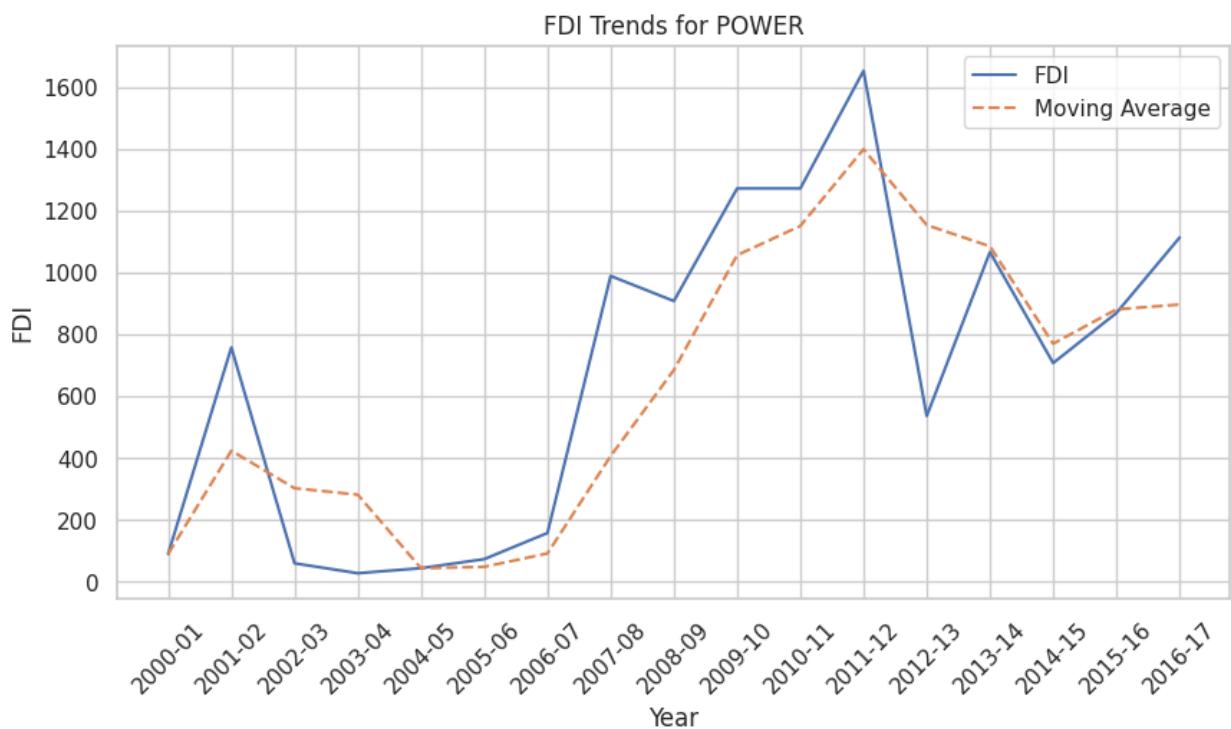
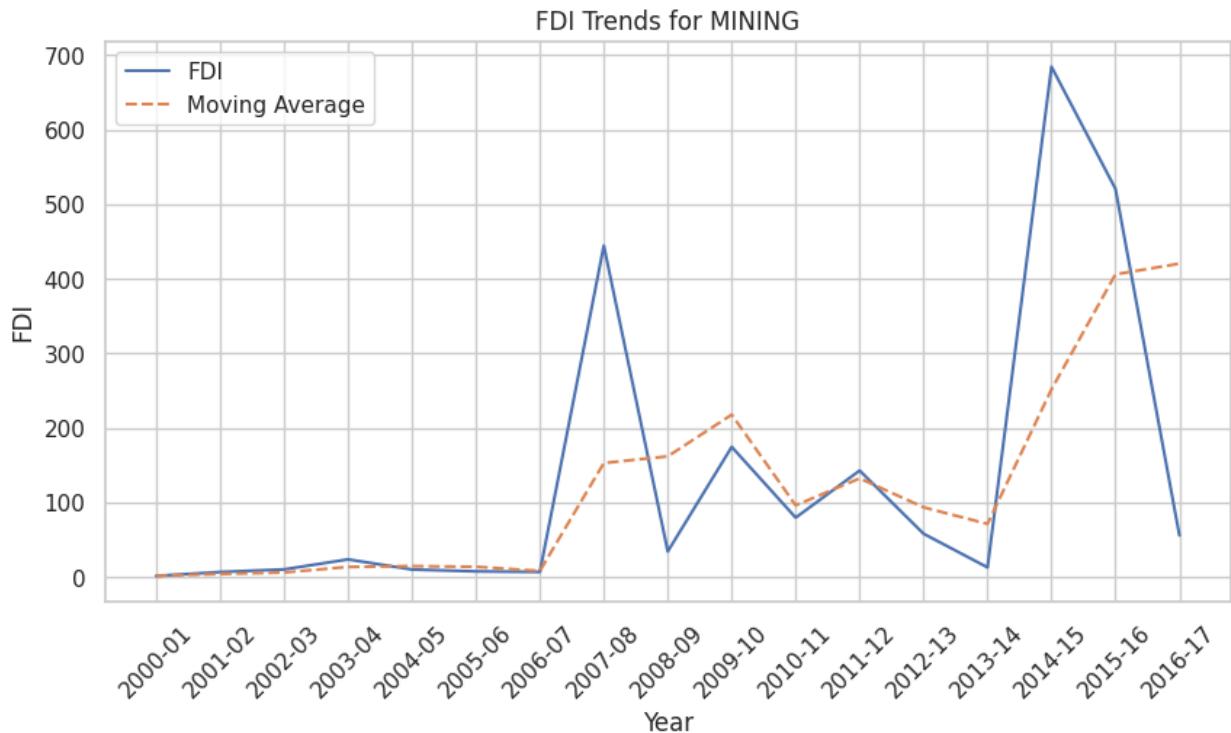
import matplotlib.pyplot as plt

sectors = df_long['Sector'].unique()

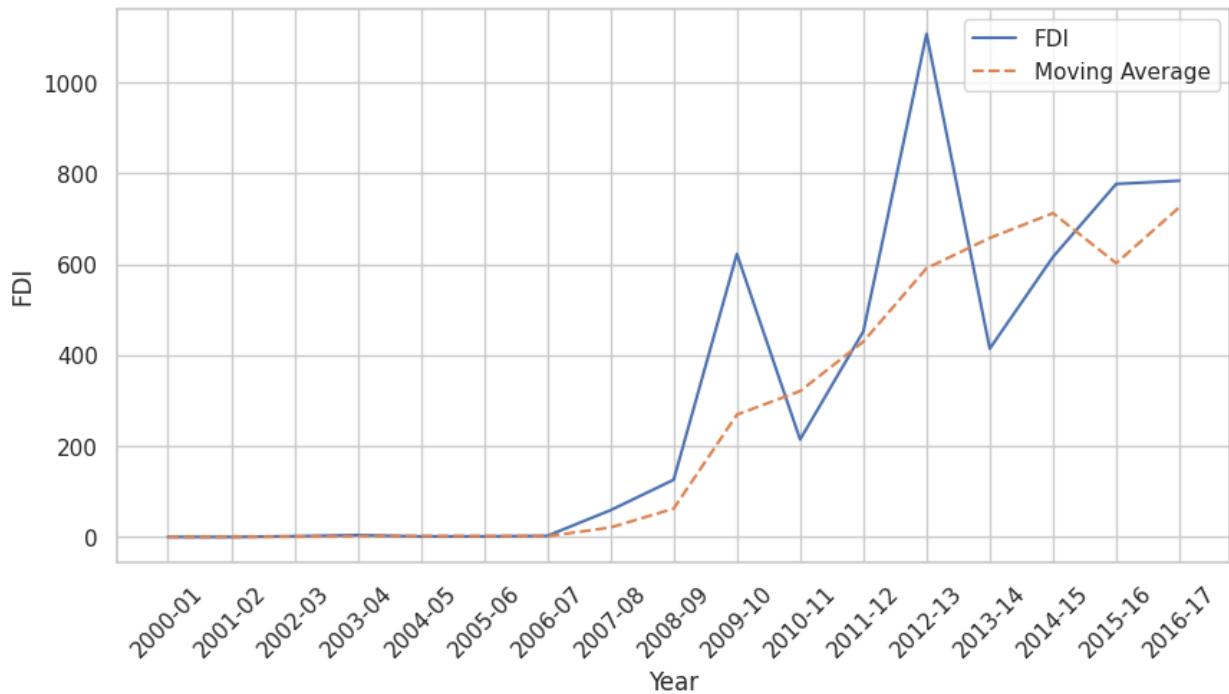
for sector in sectors:
    sector_data = df_long[df_long['Sector'] == sector]
    plt.figure(figsize=(10, 5))
    plt.plot(sector_data['Year'], sector_data['FDI'], label='FDI')
    plt.plot(sector_data['Year'], sector_data['Moving Average'], '--',
label='Moving Average')
    plt.title(f'FDI Trends for {sector}')
    plt.xlabel('Year')
    plt.xticks(rotation=45)
    plt.ylabel('FDI')
    plt.legend()
    plt.grid(True)
    plt.show()

```

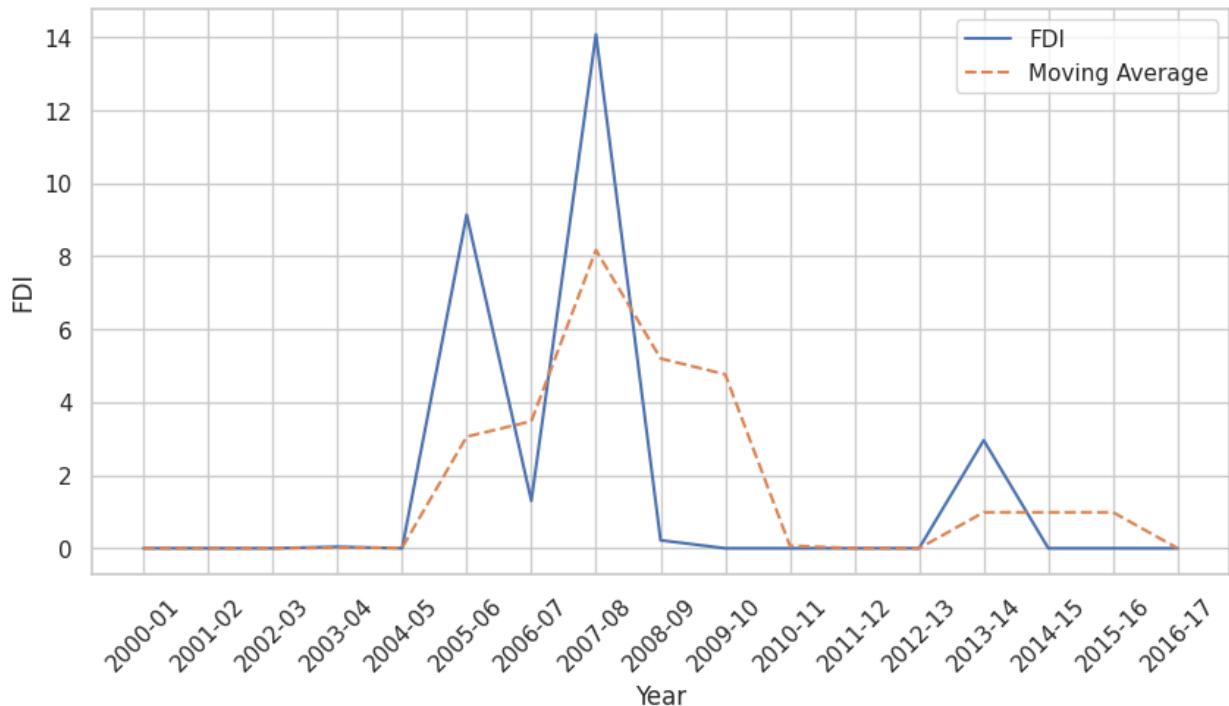


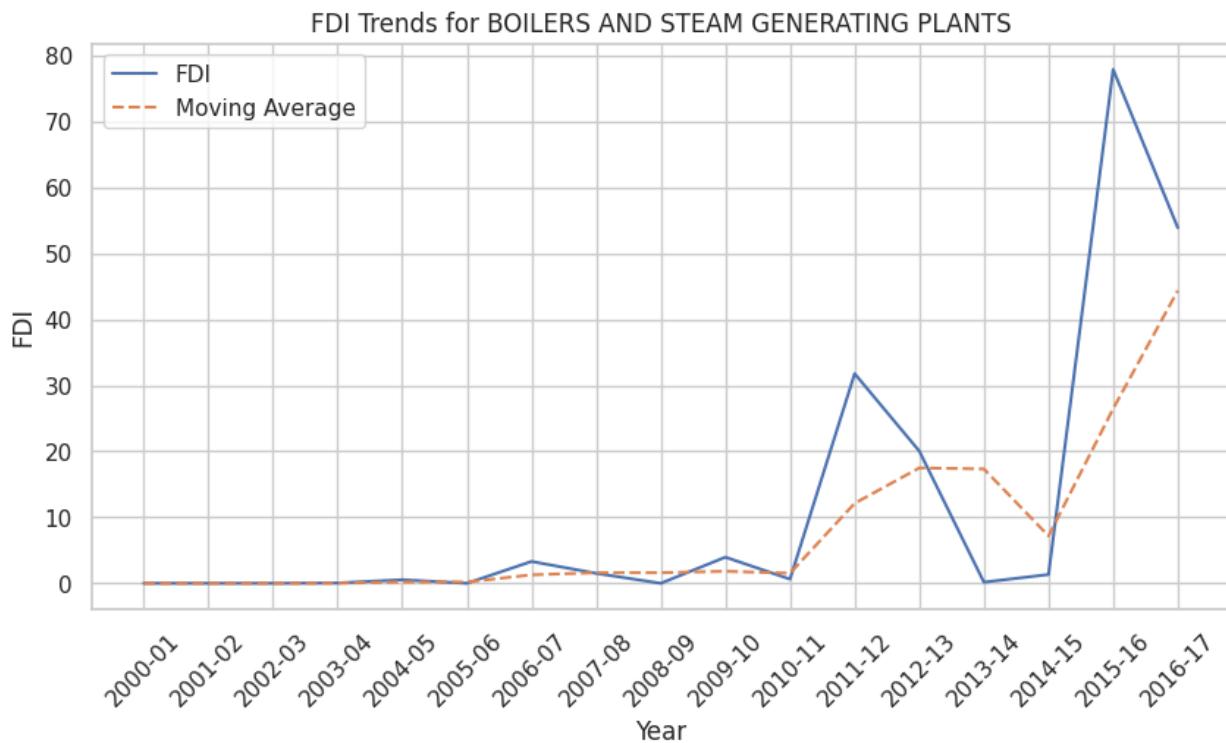
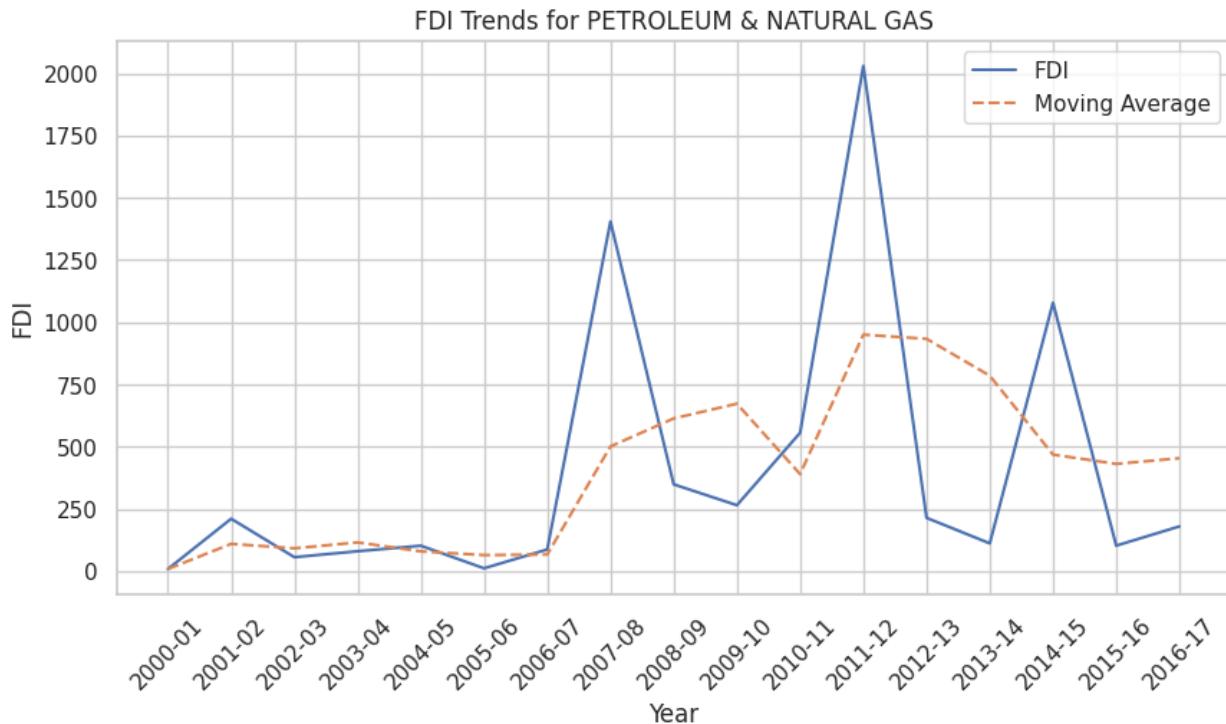


FDI Trends for NON-CONVENTIONAL ENERGY

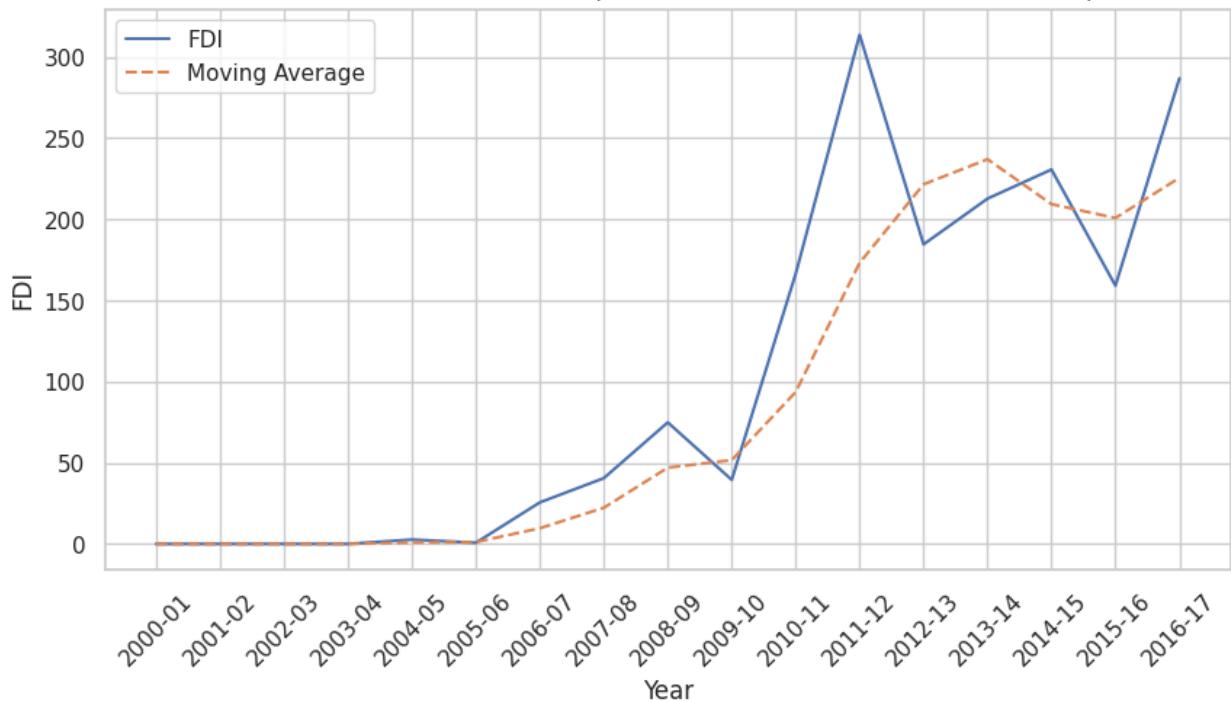


FDI Trends for COAL PRODUCTION

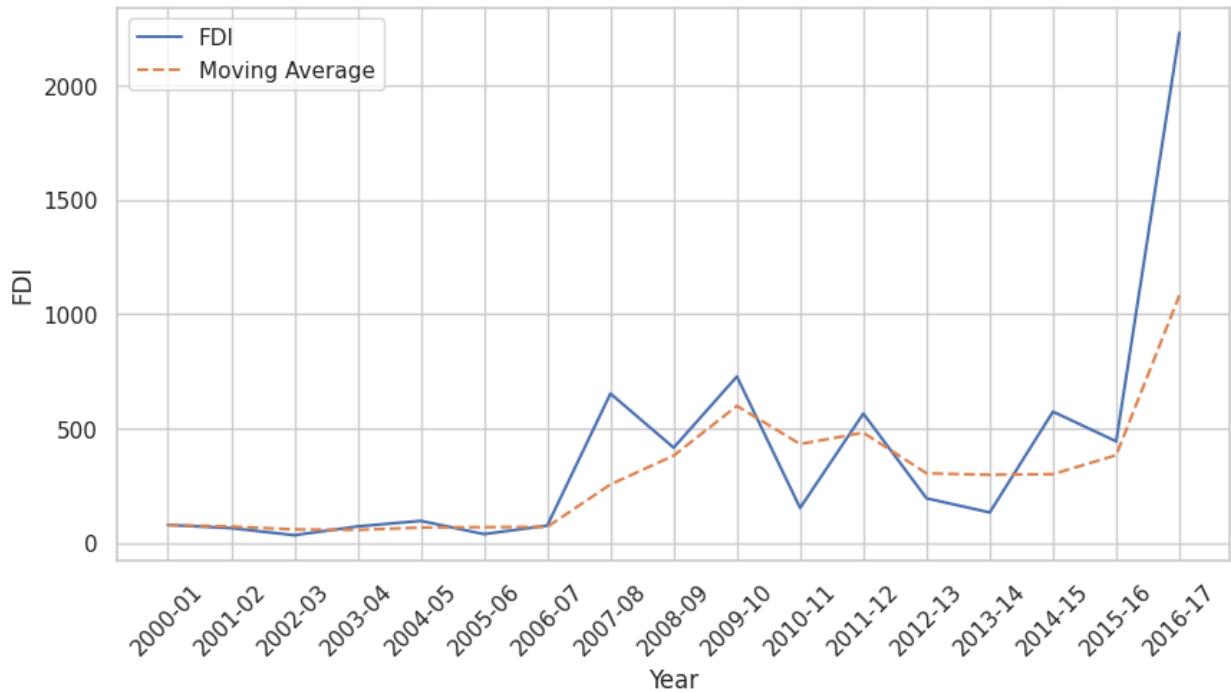




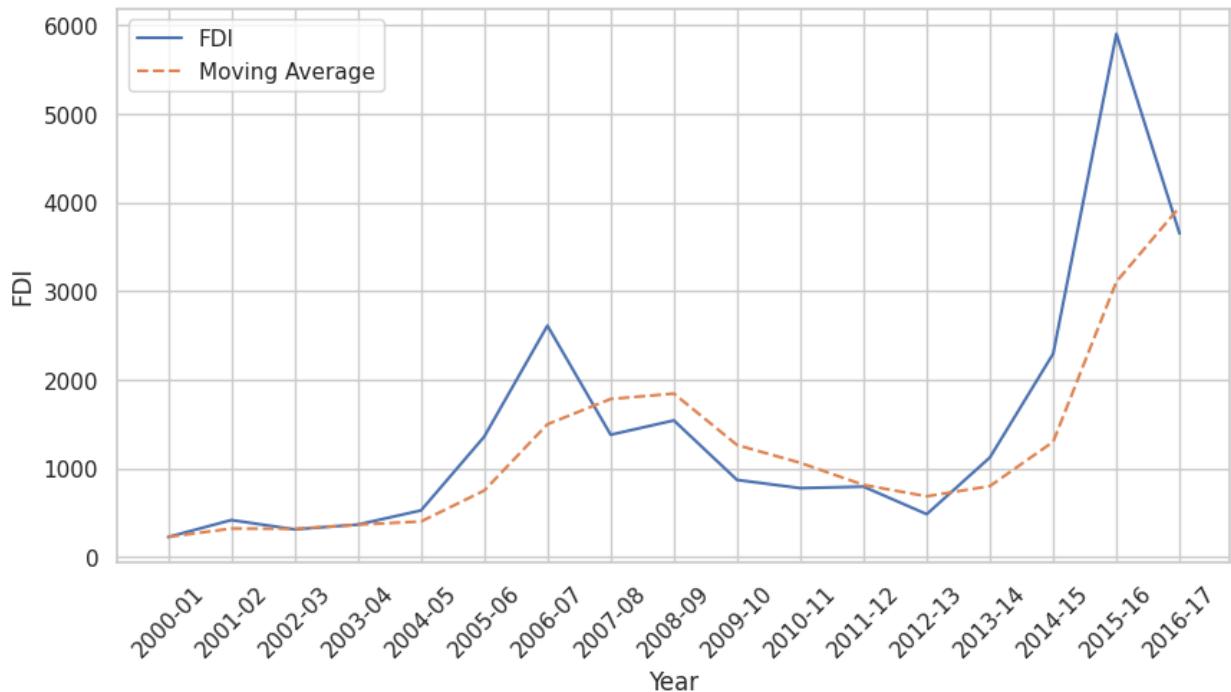
FDI Trends for PRIME MOVER (OTHER THAN ELECTRICAL GENERATORS)



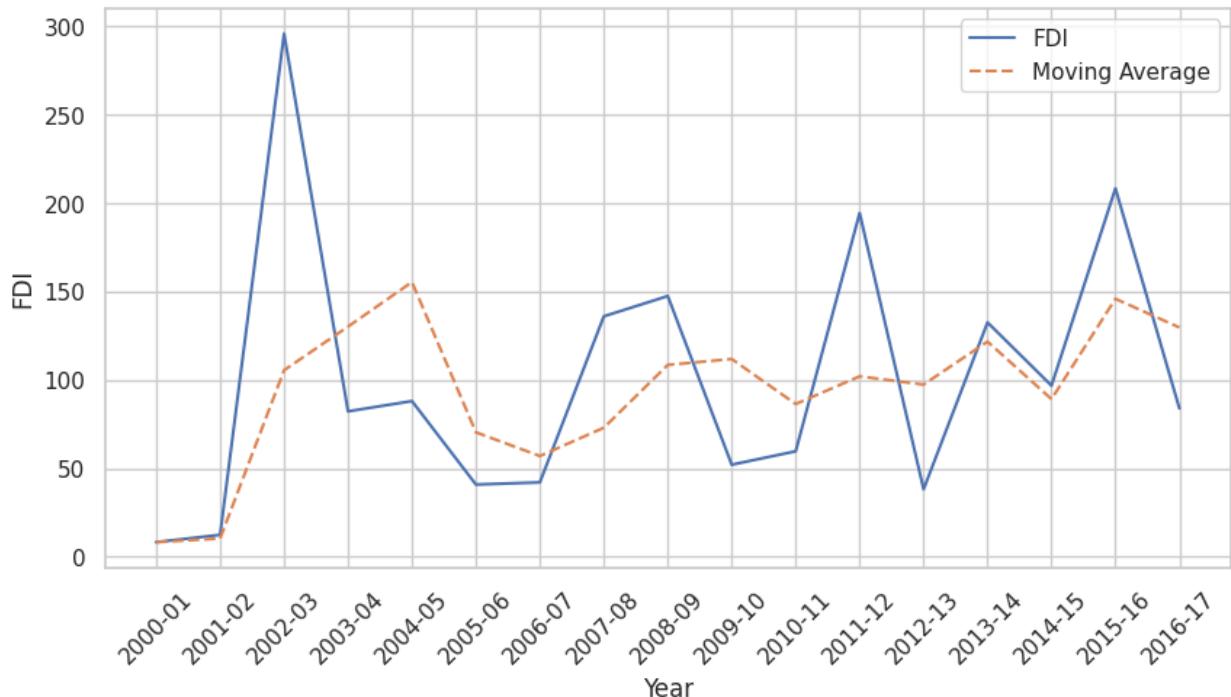
FDI Trends for ELECTRICAL EQUIPMENTS



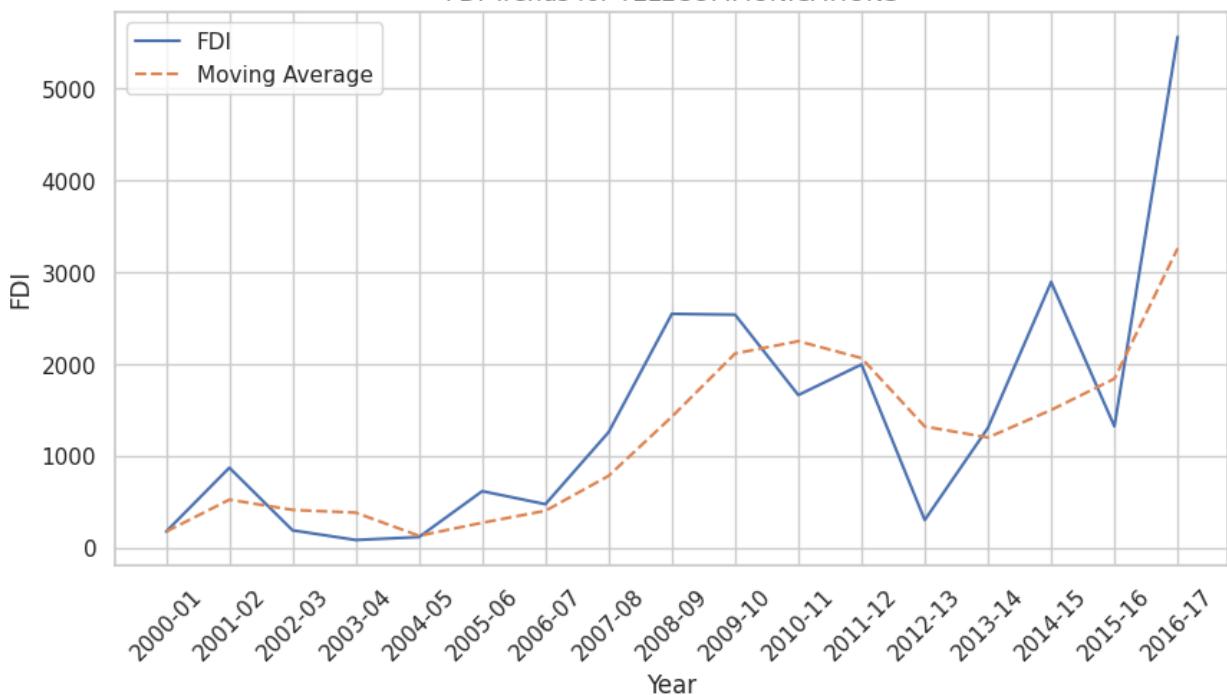
FDI Trends for COMPUTER SOFTWARE & HARDWARE



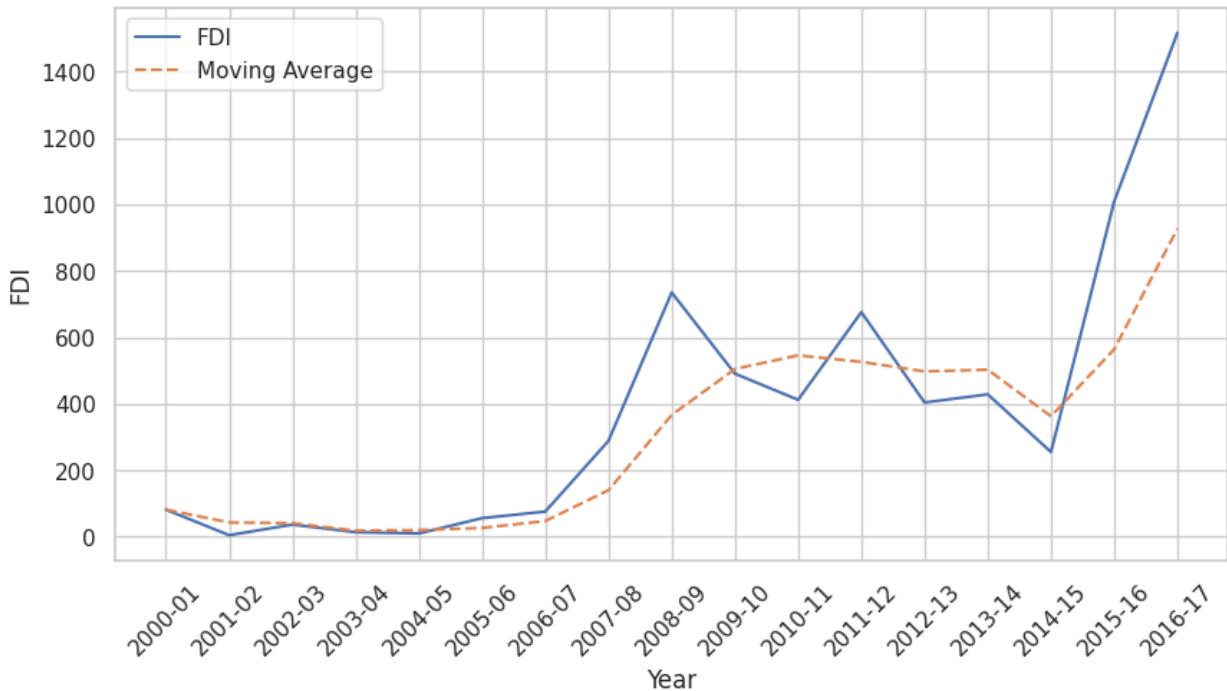
FDI Trends for ELECTRONICS



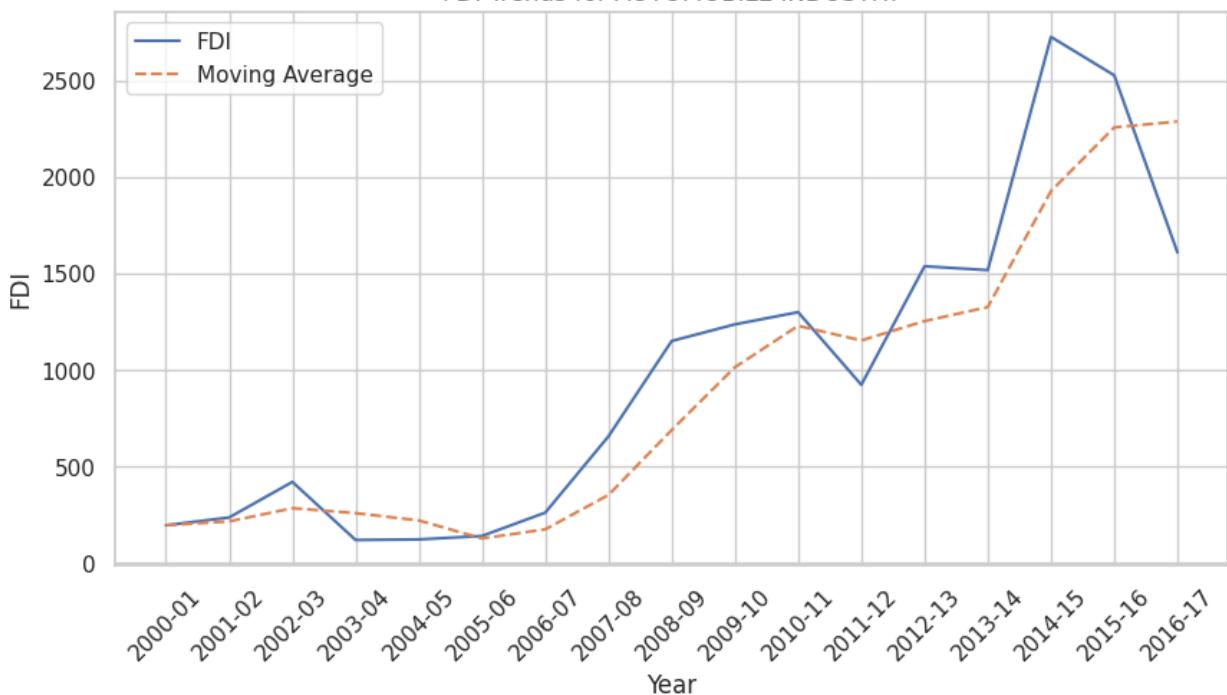
FDI Trends for TELECOMMUNICATIONS



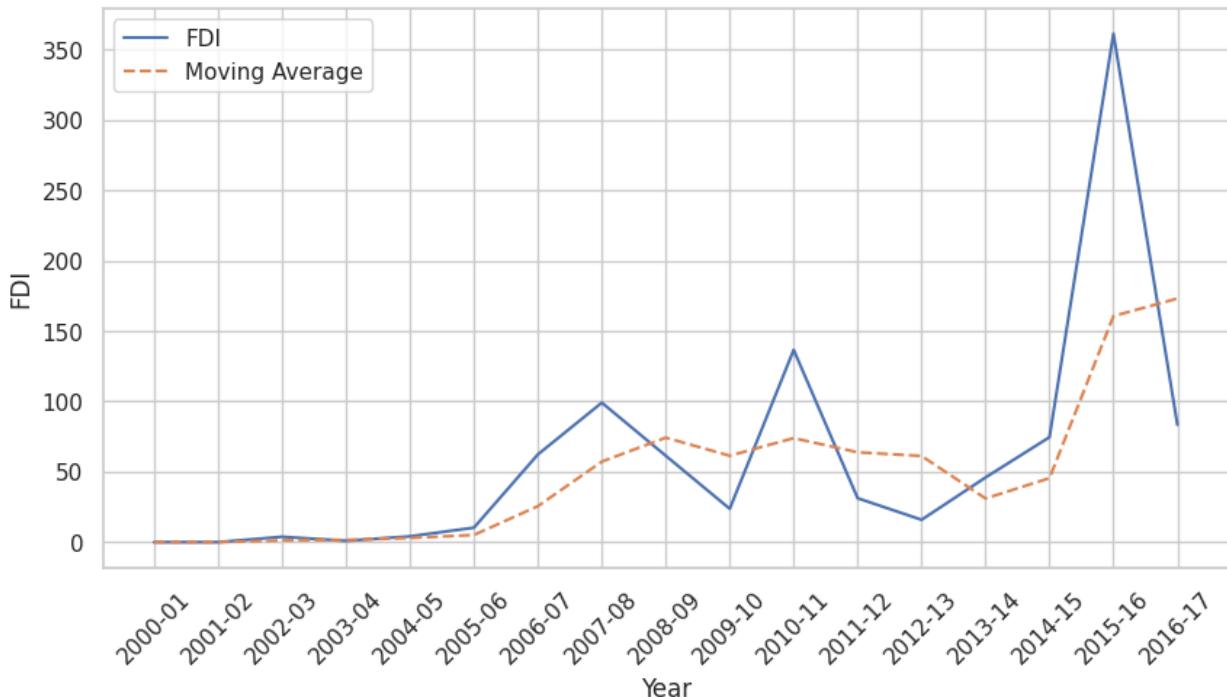
FDI Trends for INFORMATION & BROADCASTING (INCLUDING PRINT MEDIA)



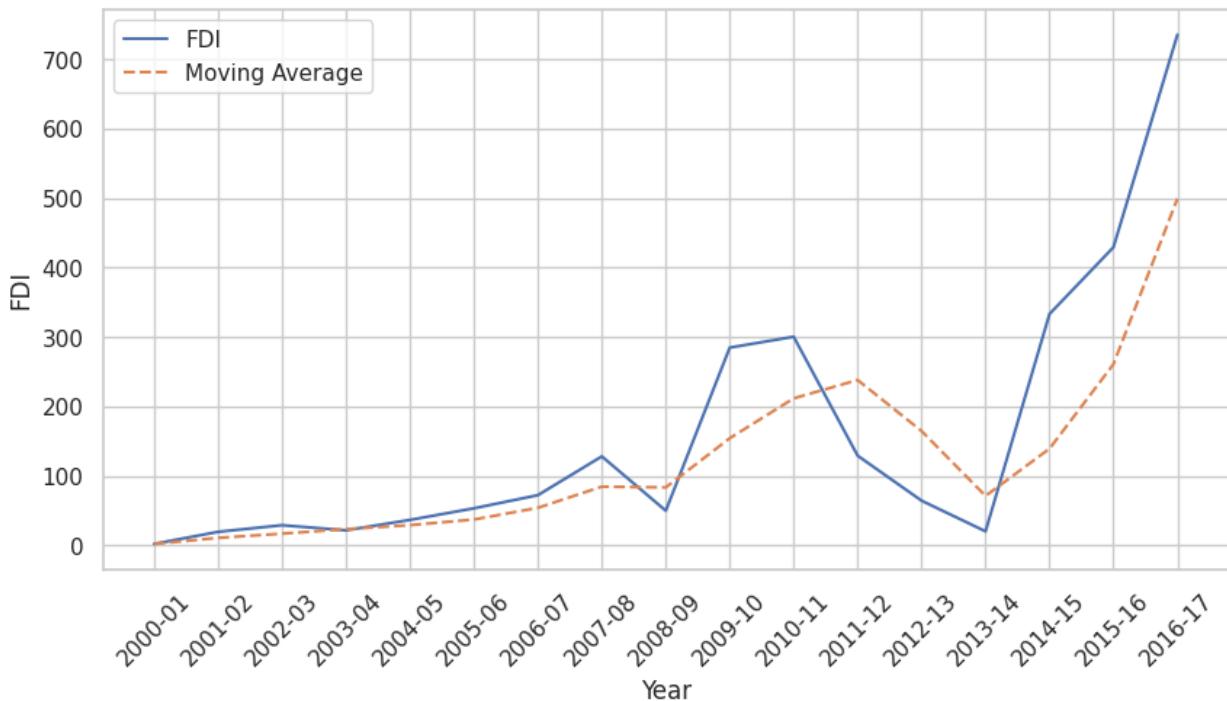
FDI Trends for AUTOMOBILE INDUSTRY



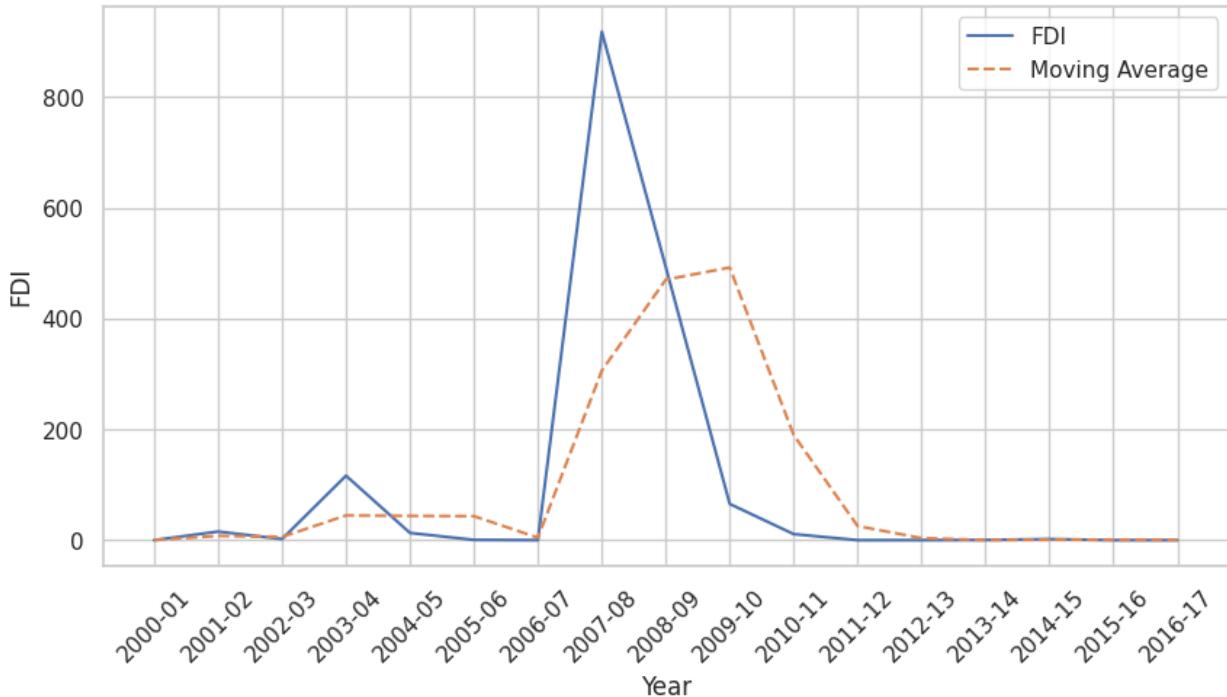
FDI Trends for AIR TRANSPORT (INCLUDING AIR FREIGHT)



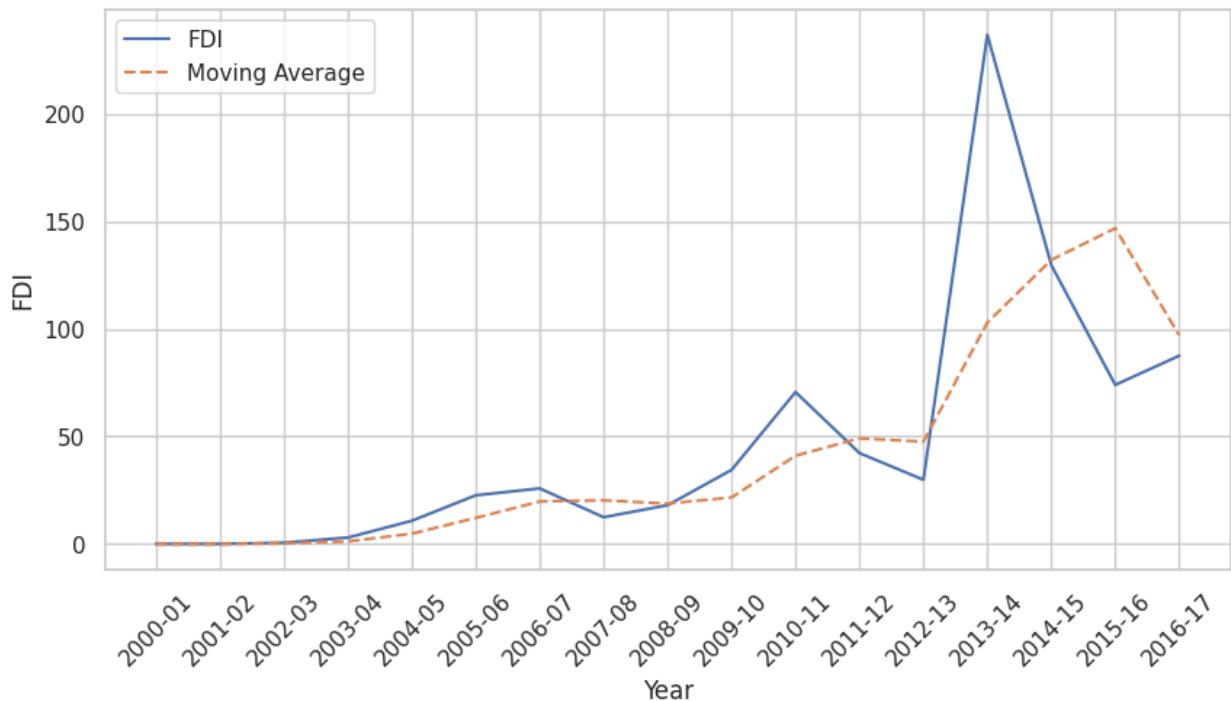
FDI Trends for SEA TRANSPORT



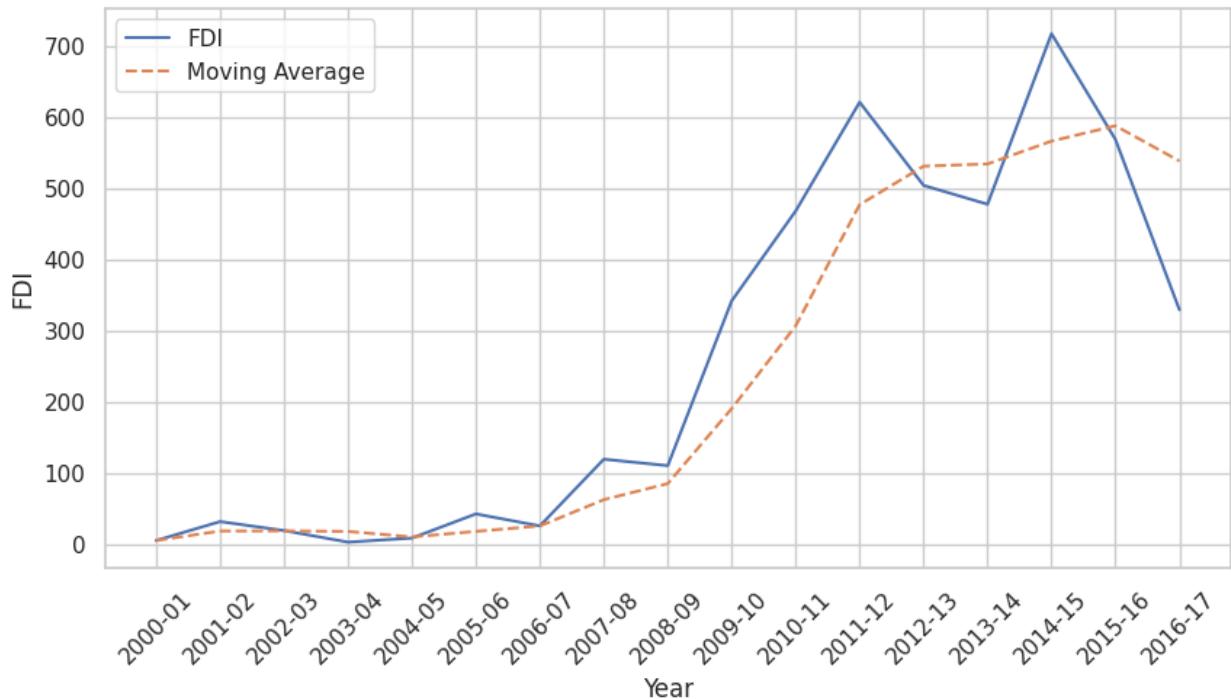
FDI Trends for PORTS

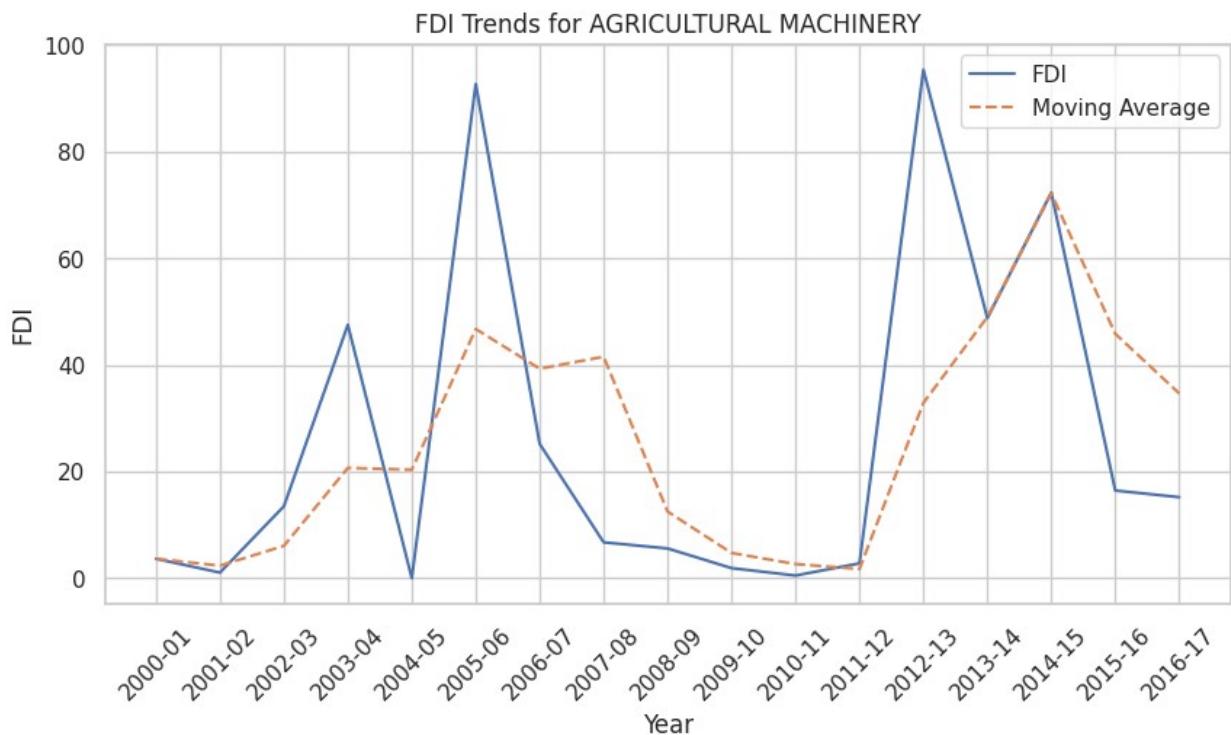
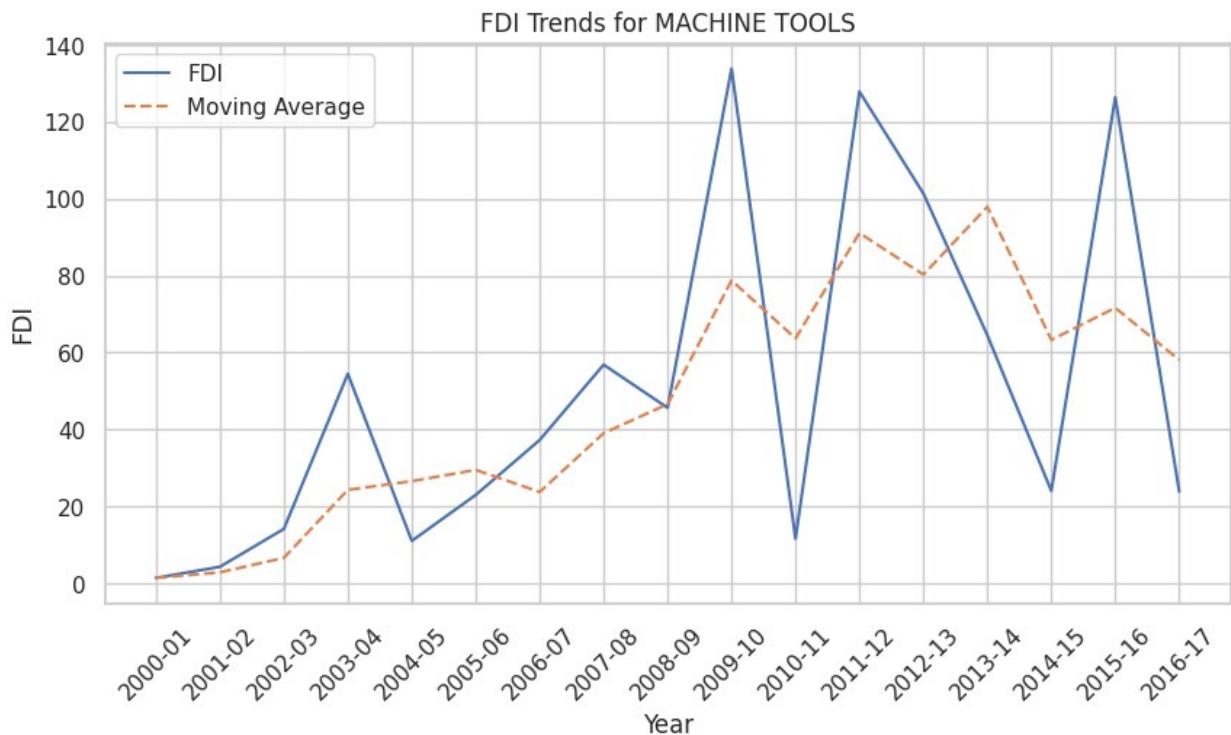


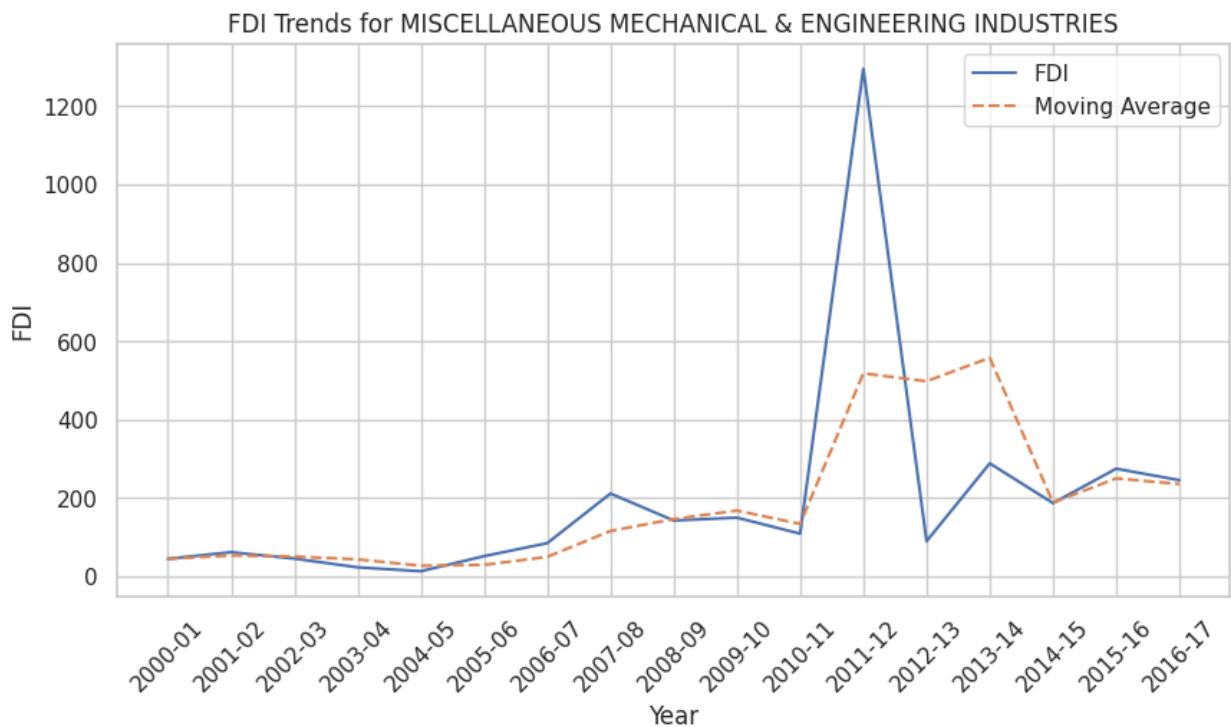
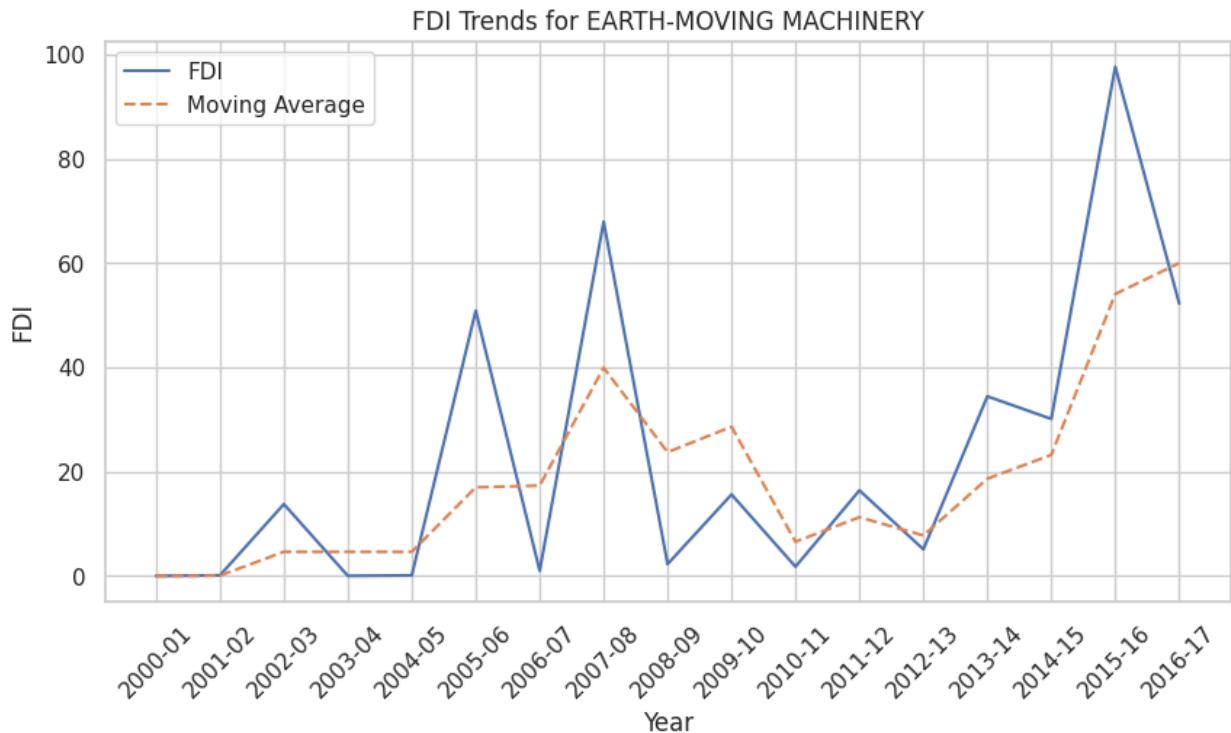
FDI Trends for RAILWAY RELATED COMPONENTS



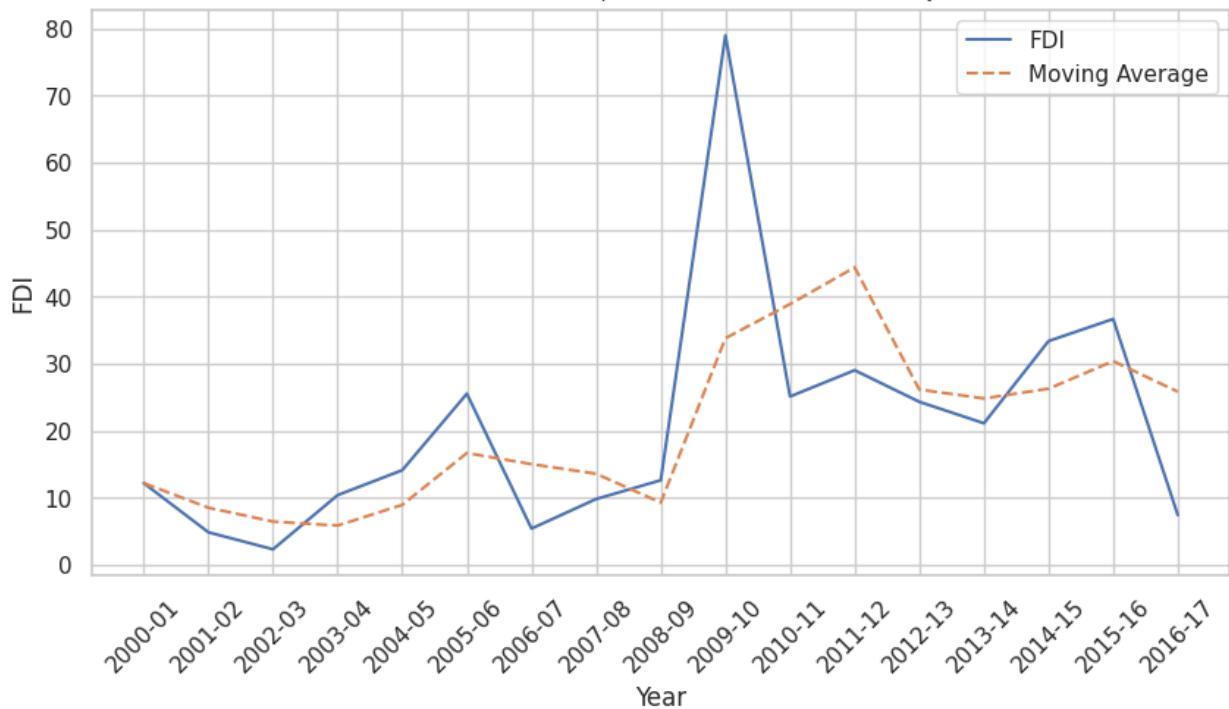
FDI Trends for INDUSTRIAL MACHINERY



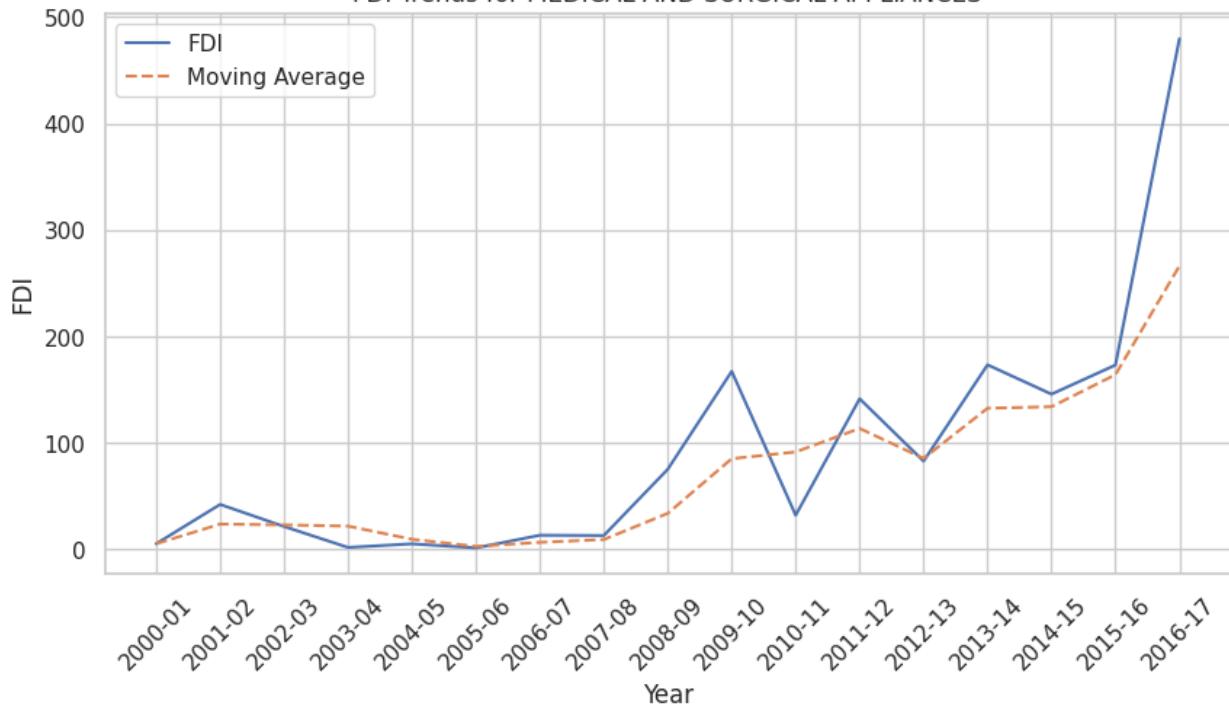




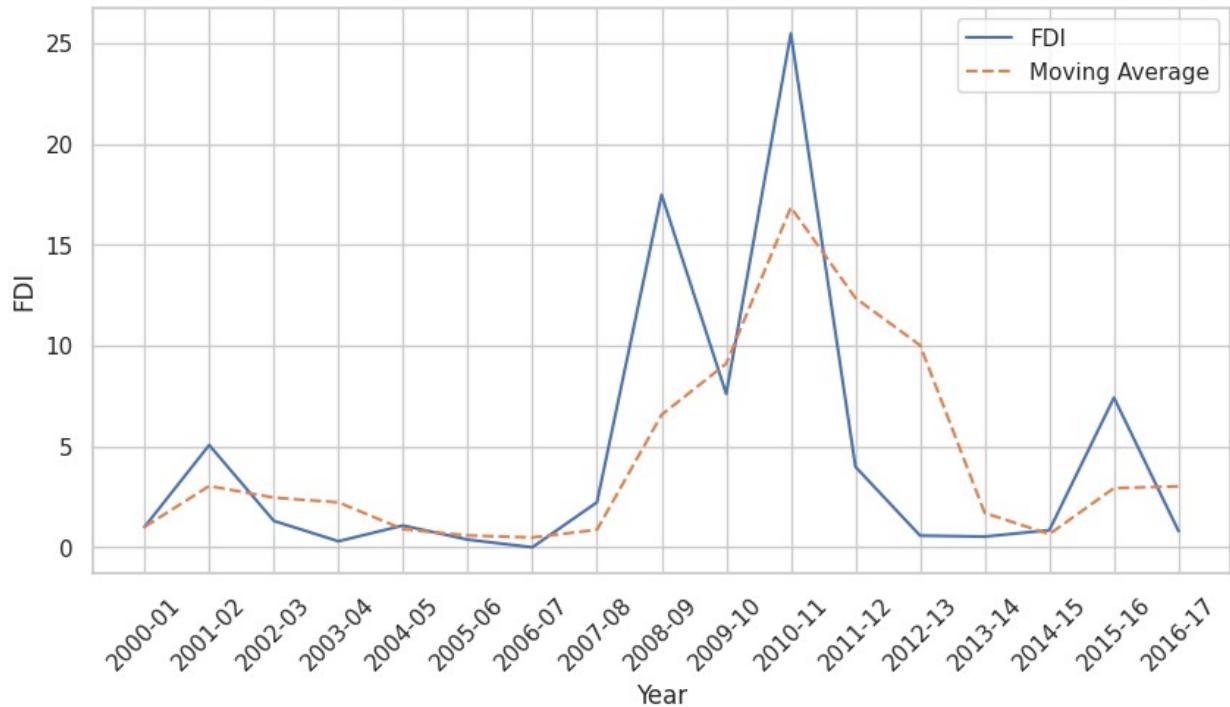
FDI Trends for COMMERCIAL, OFFICE & HOUSEHOLD EQUIPMENTS



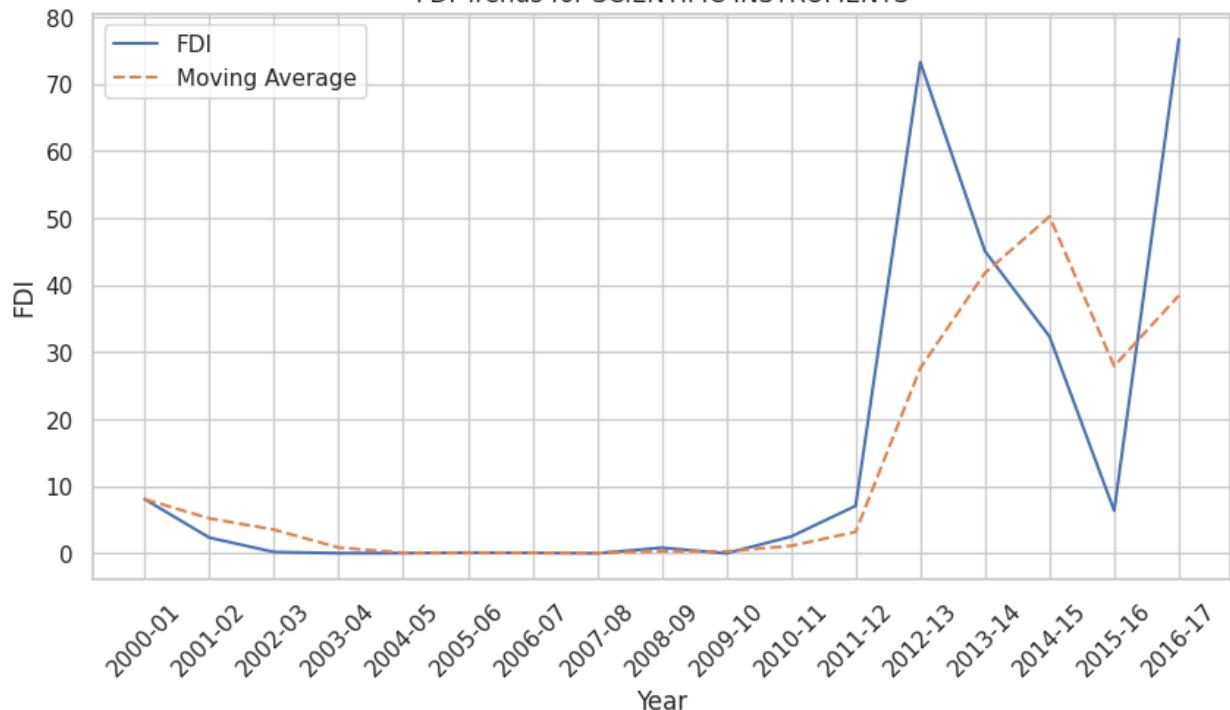
FDI Trends for MEDICAL AND SURGICAL APPLIANCES

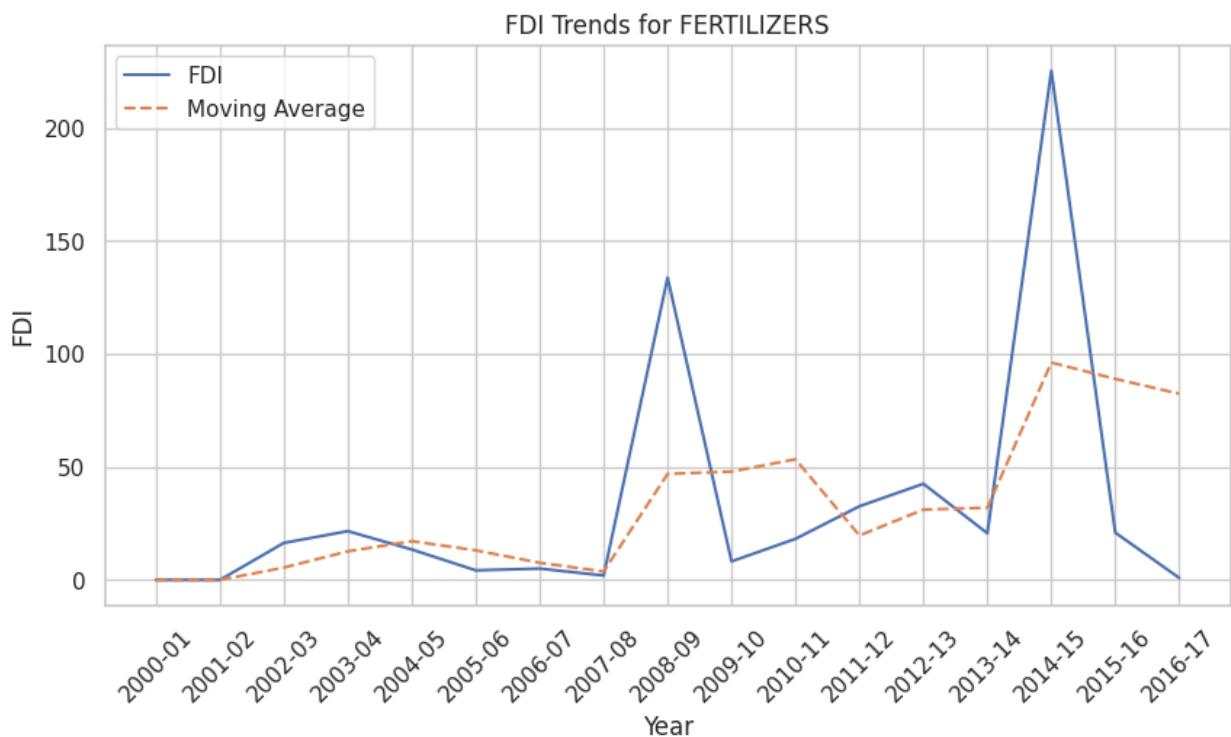
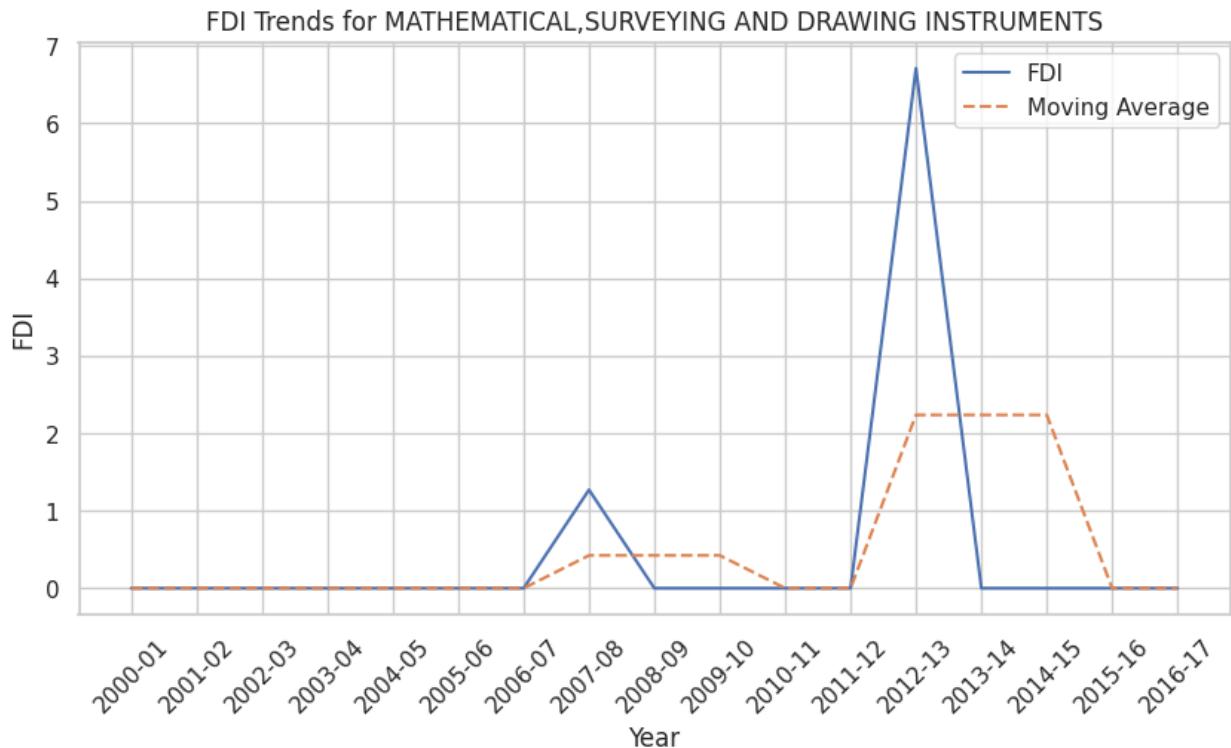


FDI Trends for INDUSTRIAL INSTRUMENTS

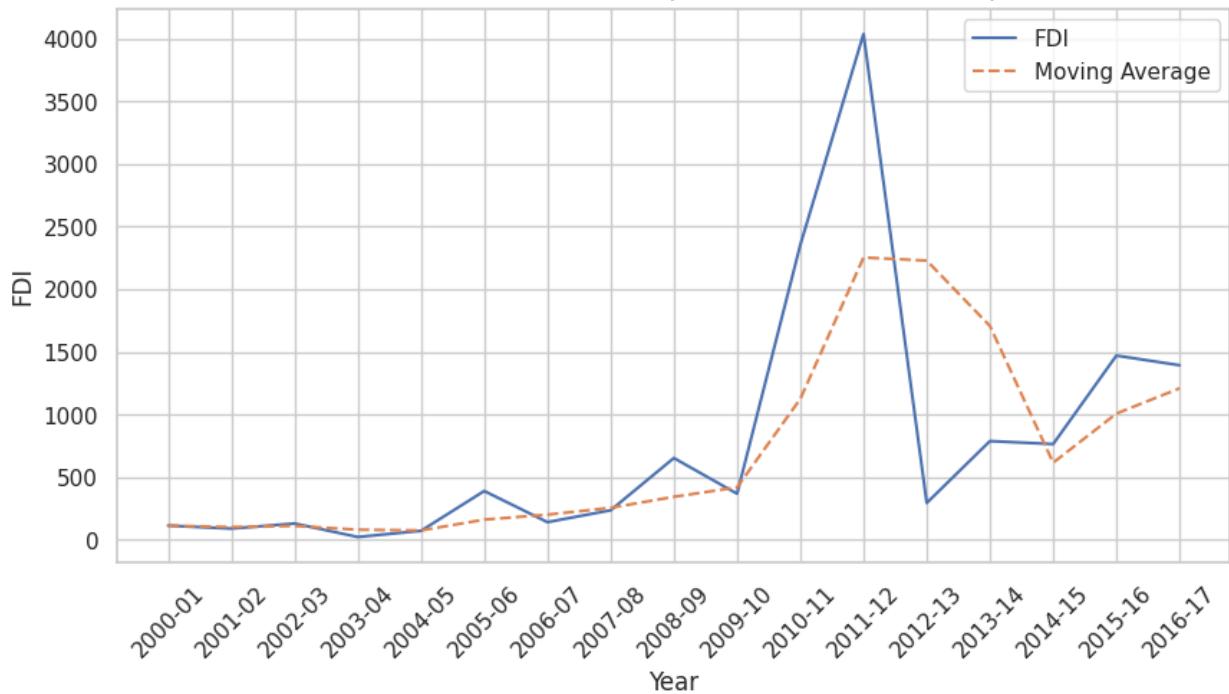


FDI Trends for SCIENTIFIC INSTRUMENTS

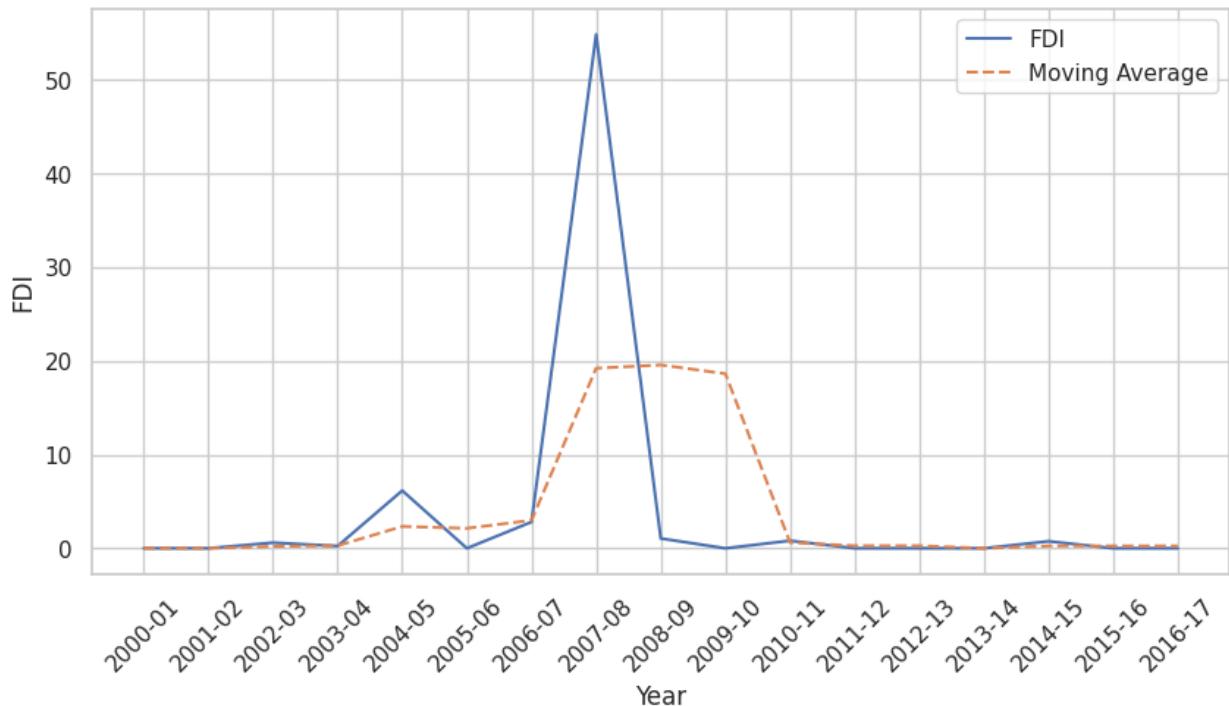




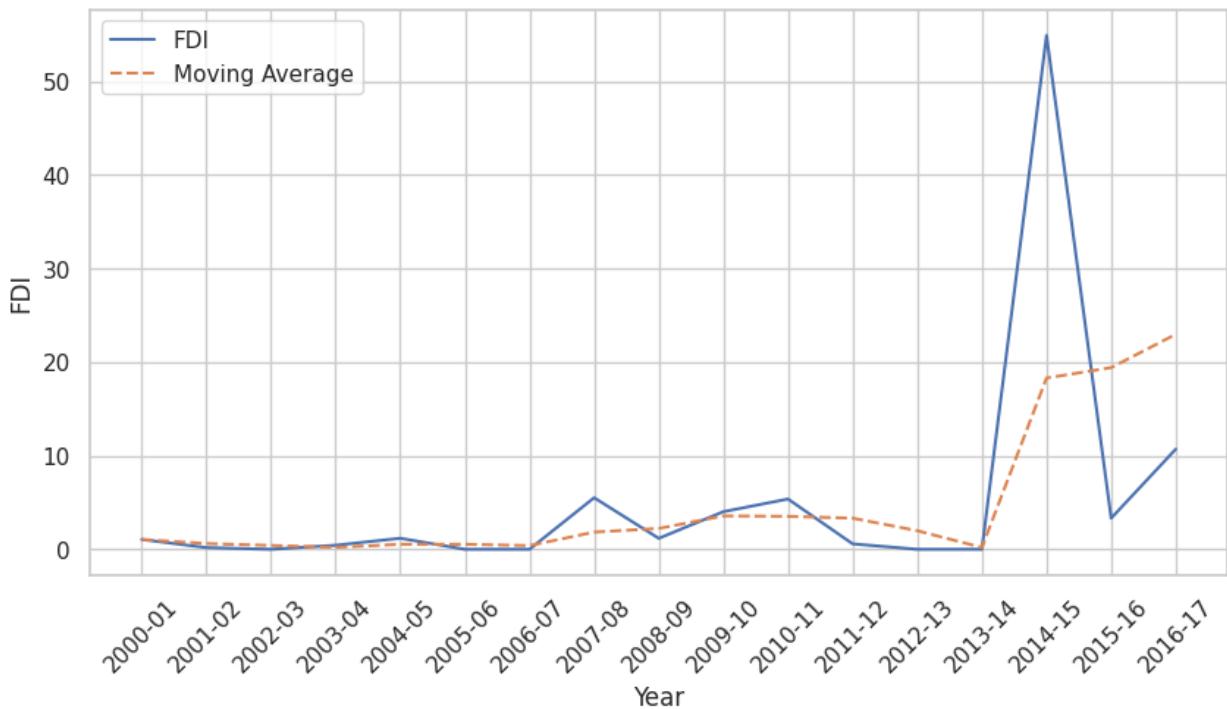
FDI Trends for CHEMICALS (OTHER THAN FERTILIZERS)



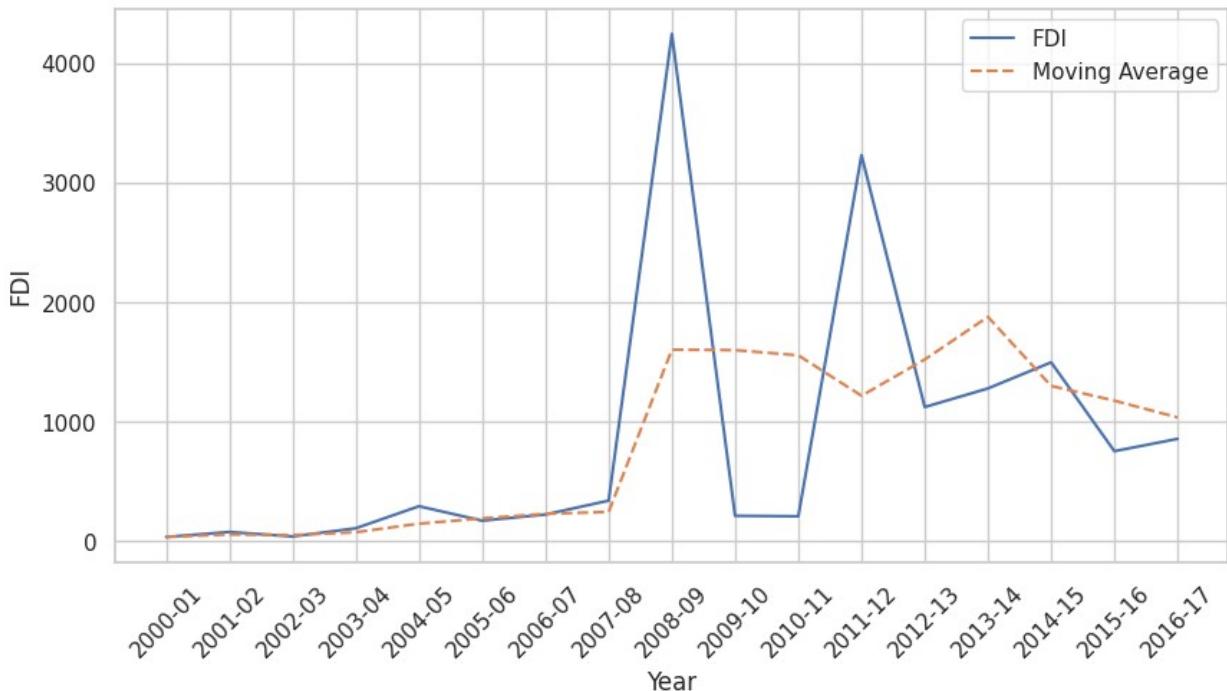
FDI Trends for PHOTOGRAPHIC RAW FILM AND PAPER



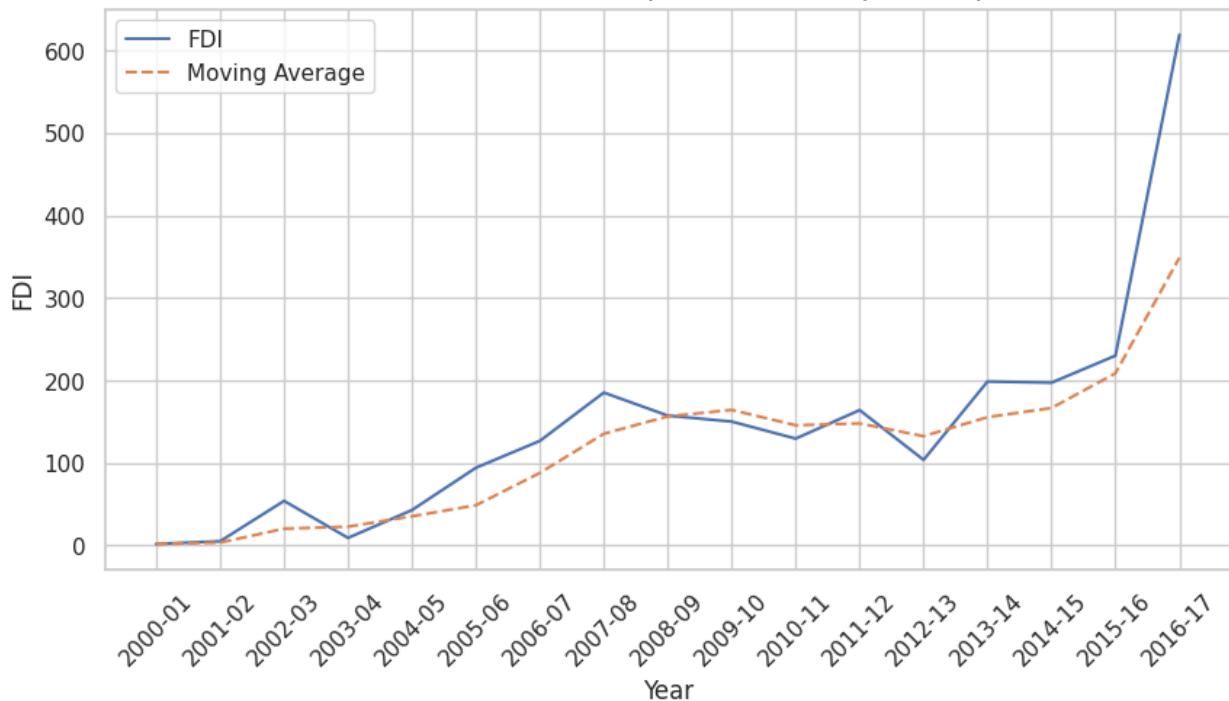
FDI Trends for DYE-STUFFS



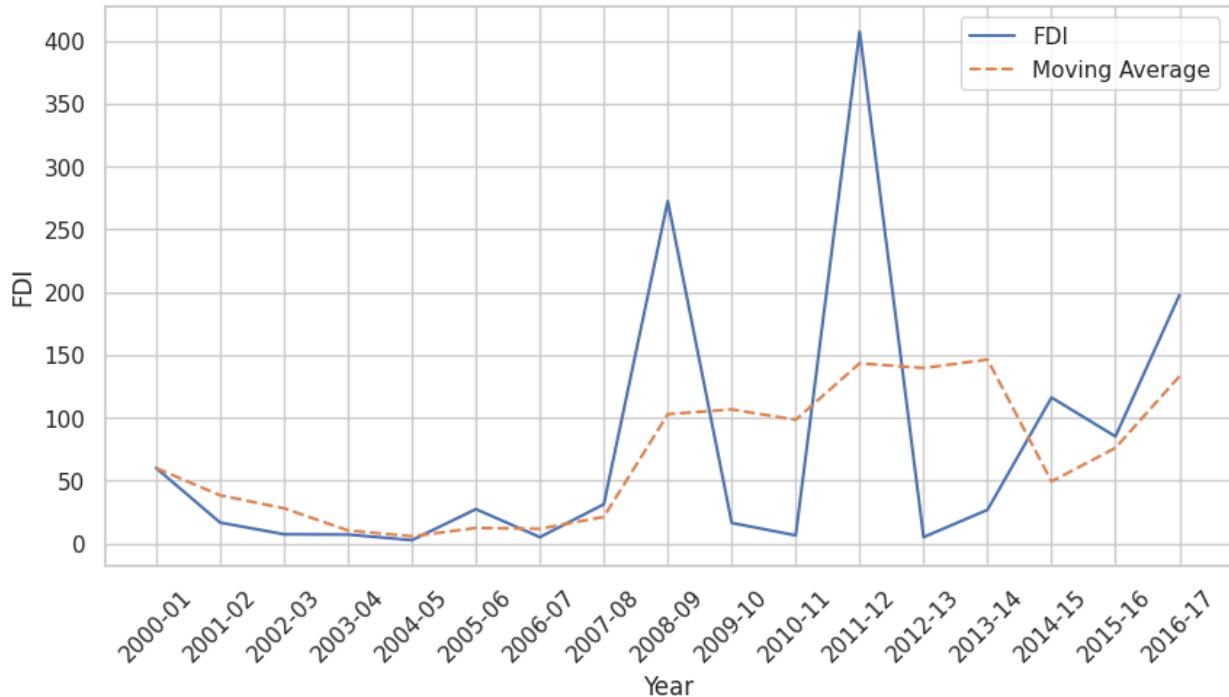
FDI Trends for DRUGS & PHARMACEUTICALS



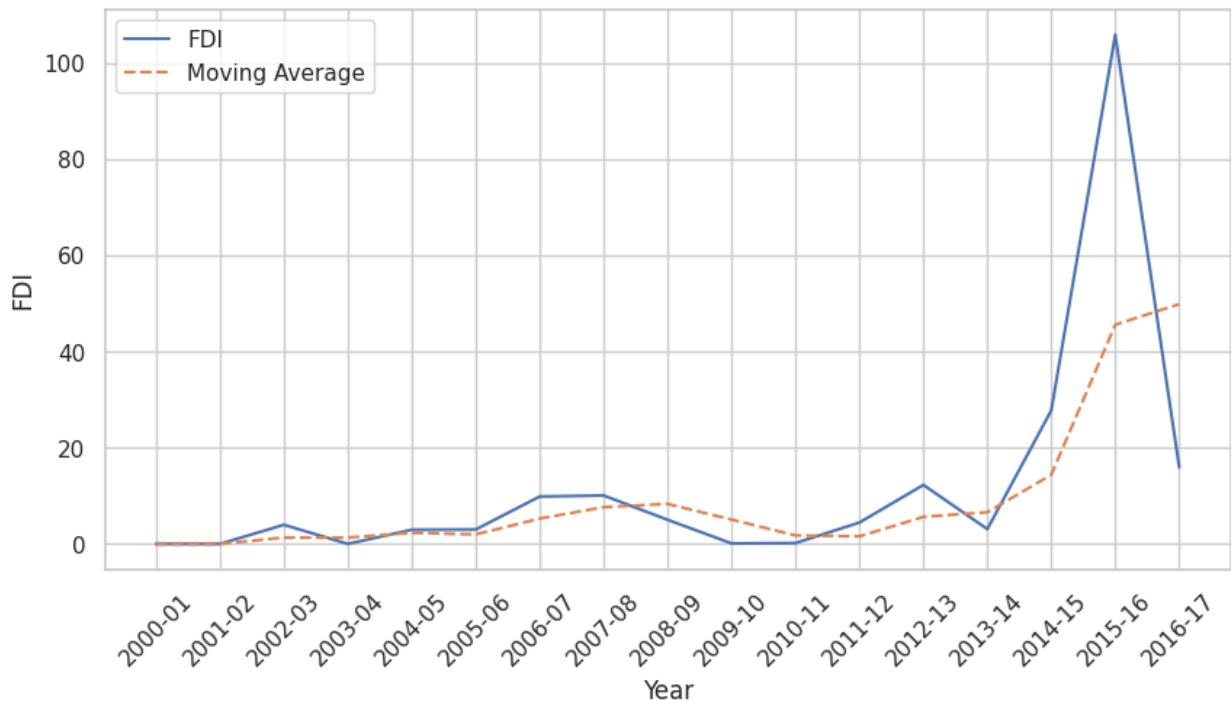
FDI Trends for TEXTILES (INCLUDING DYED,PRINTED)



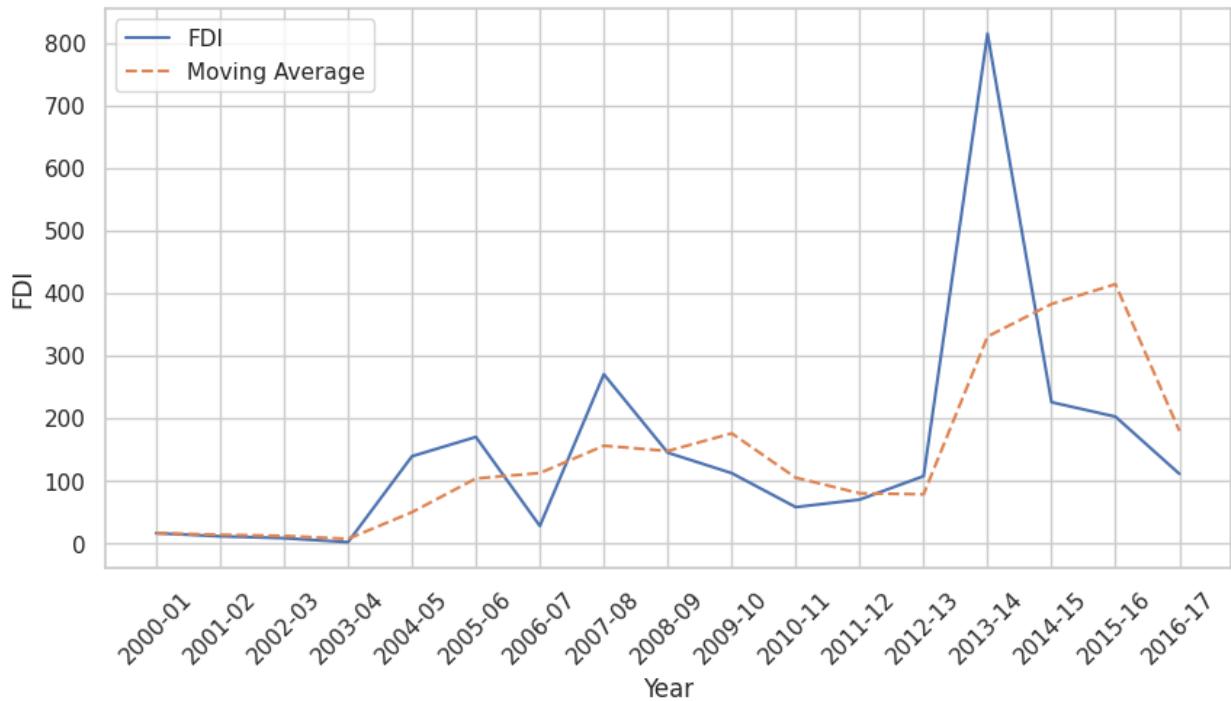
FDI Trends for PAPER AND PULP (INCLUDING PAPER PRODUCTS)

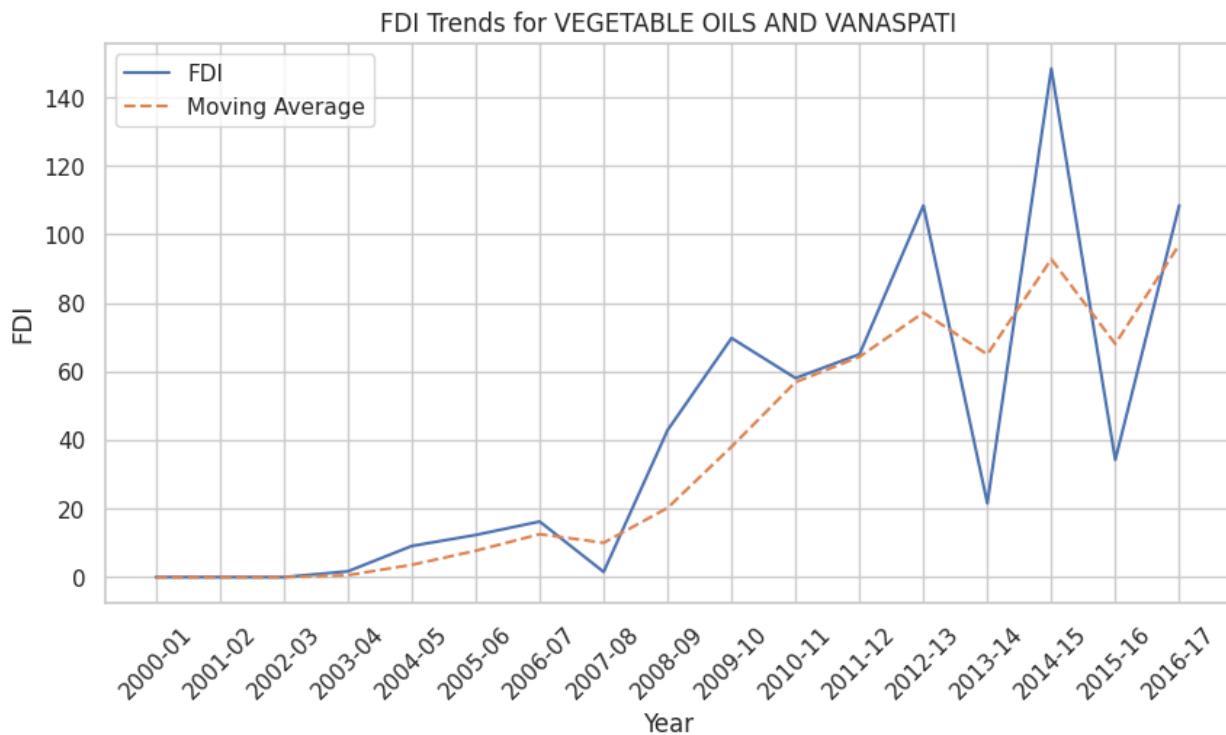
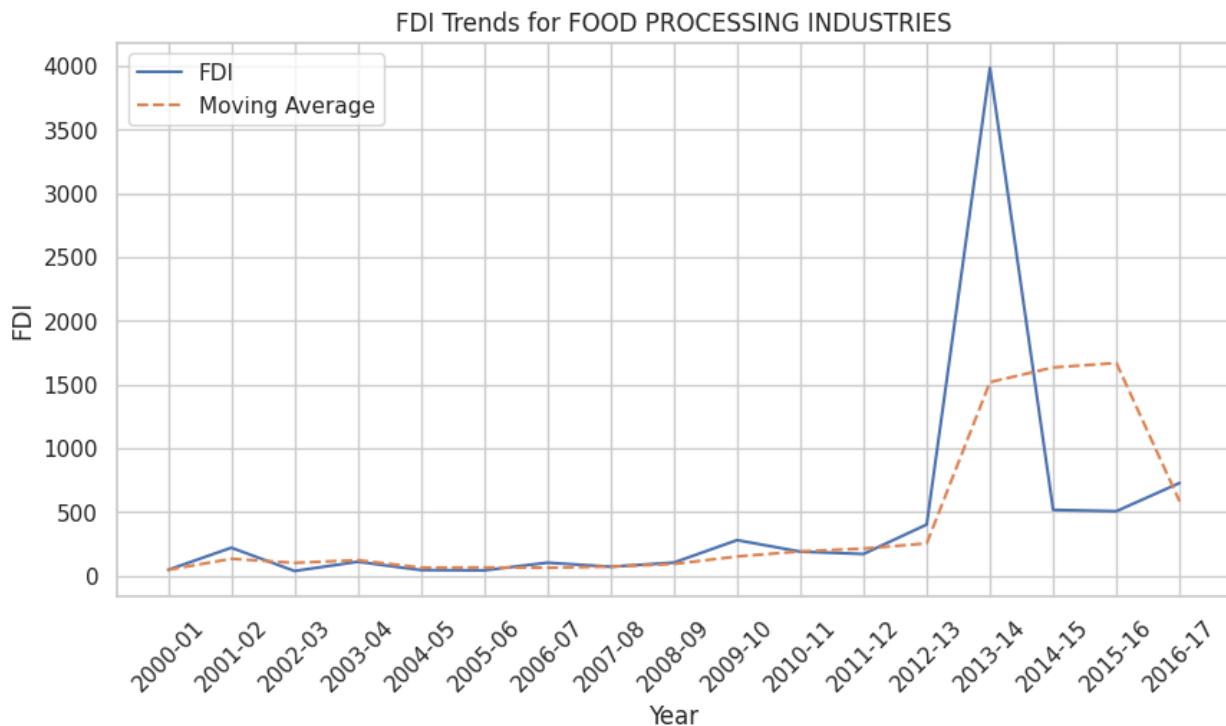


FDI Trends for SUGAR

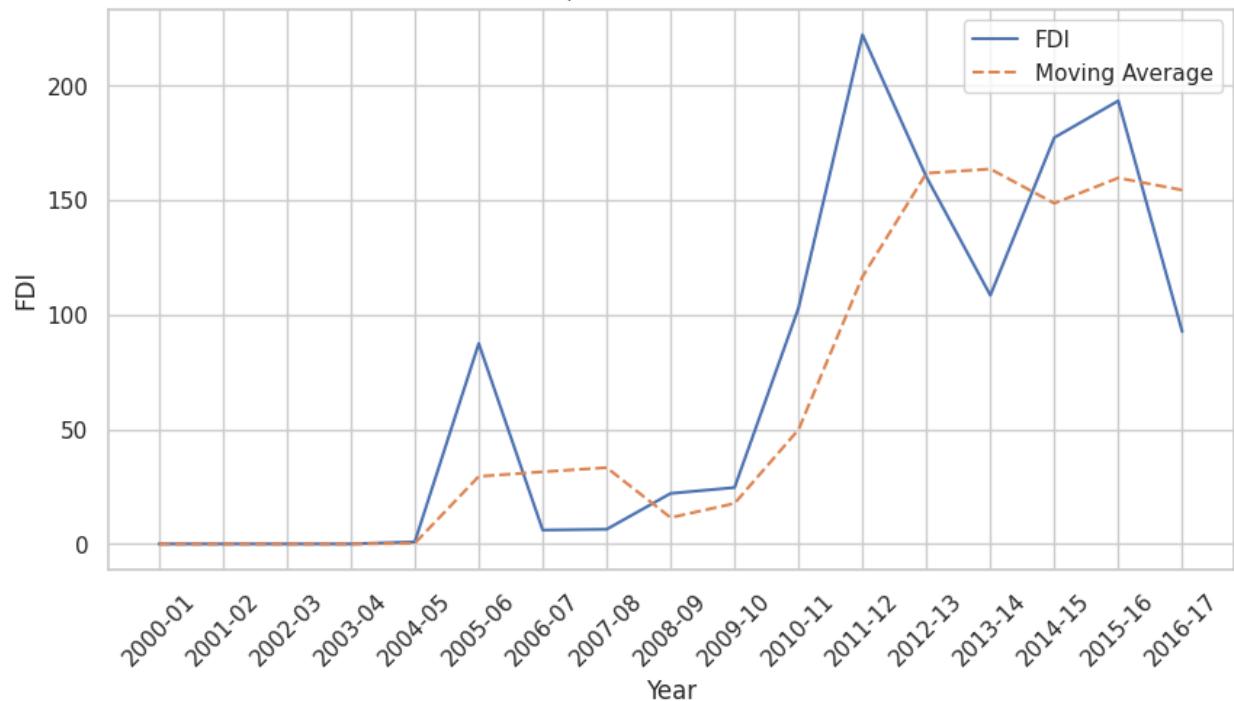


FDI Trends for FERMENTATION INDUSTRIES

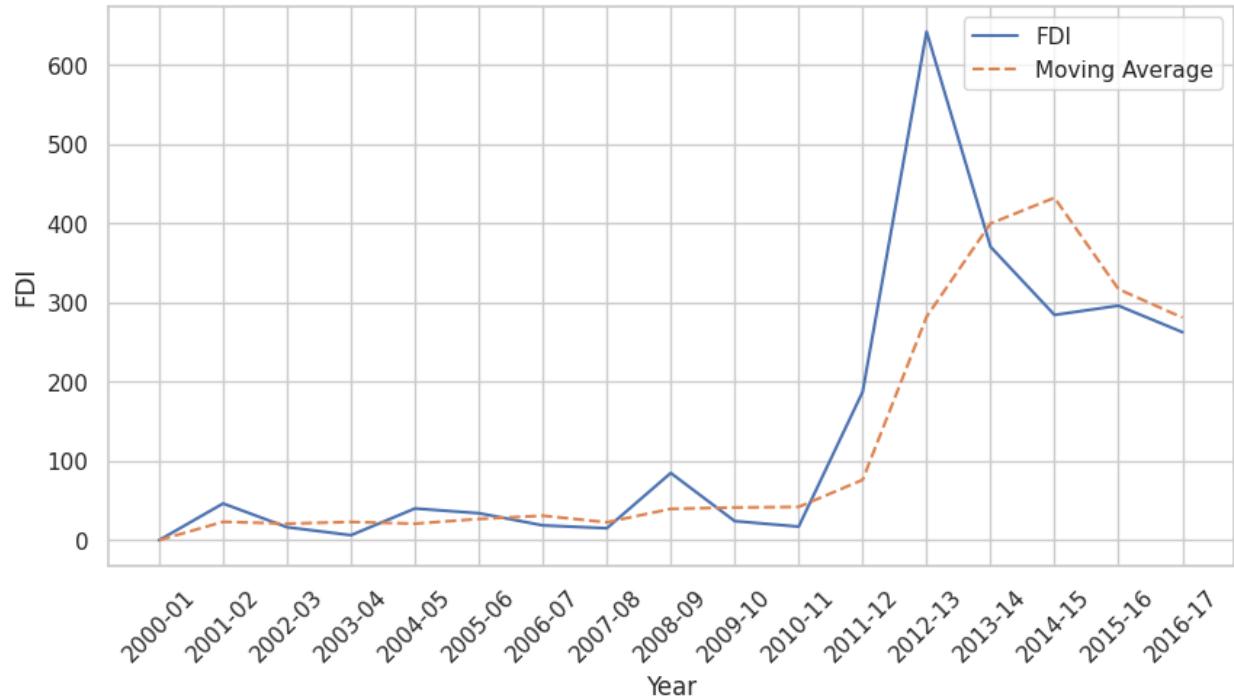




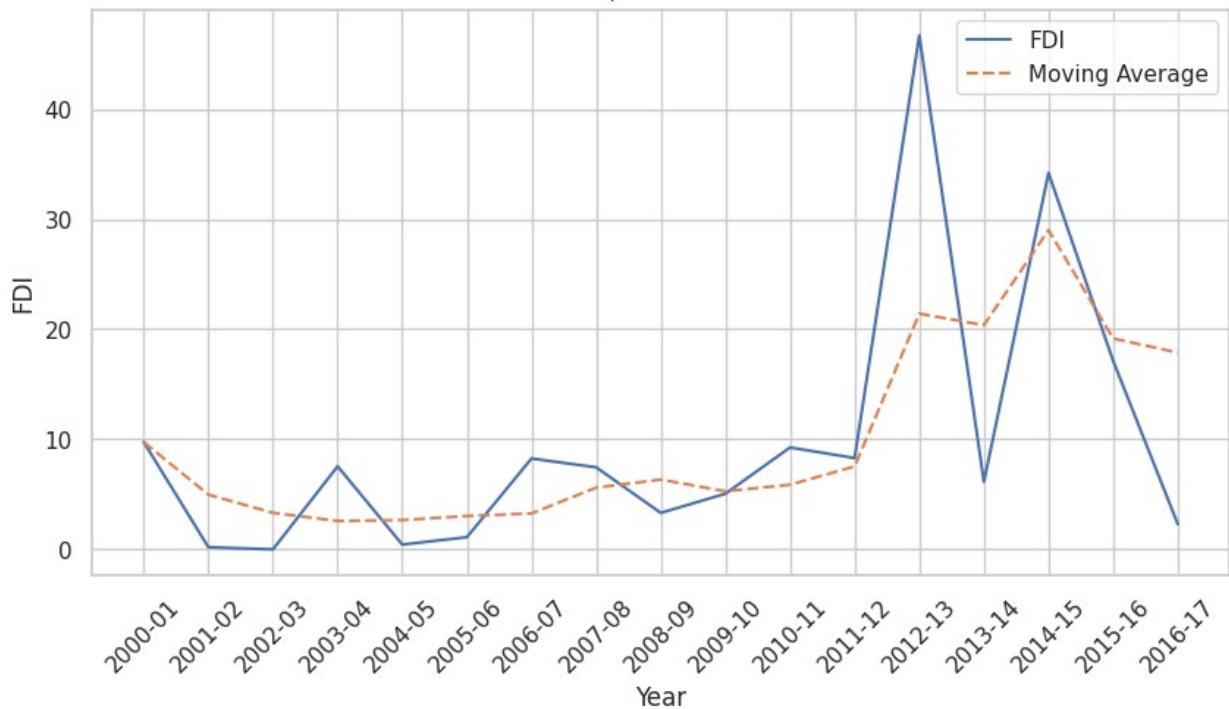
FDI Trends for SOAPS, COSMETICS & TOILET PREPARATIONS



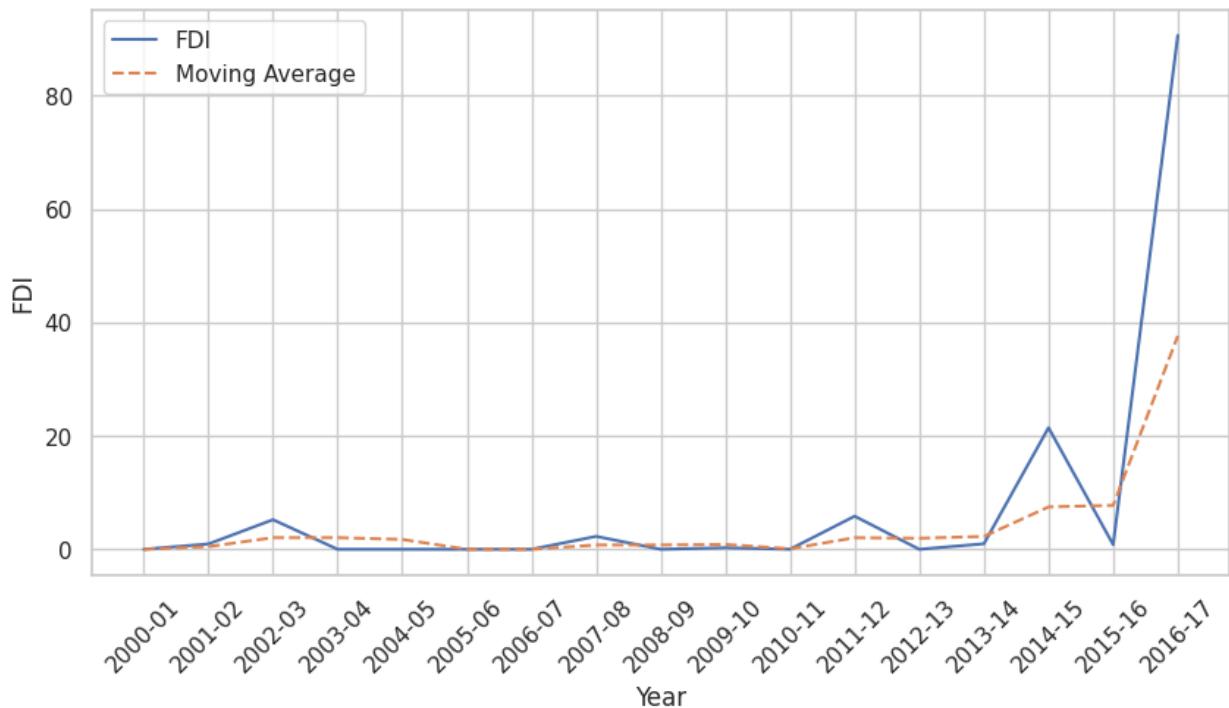
FDI Trends for RUBBER GOODS



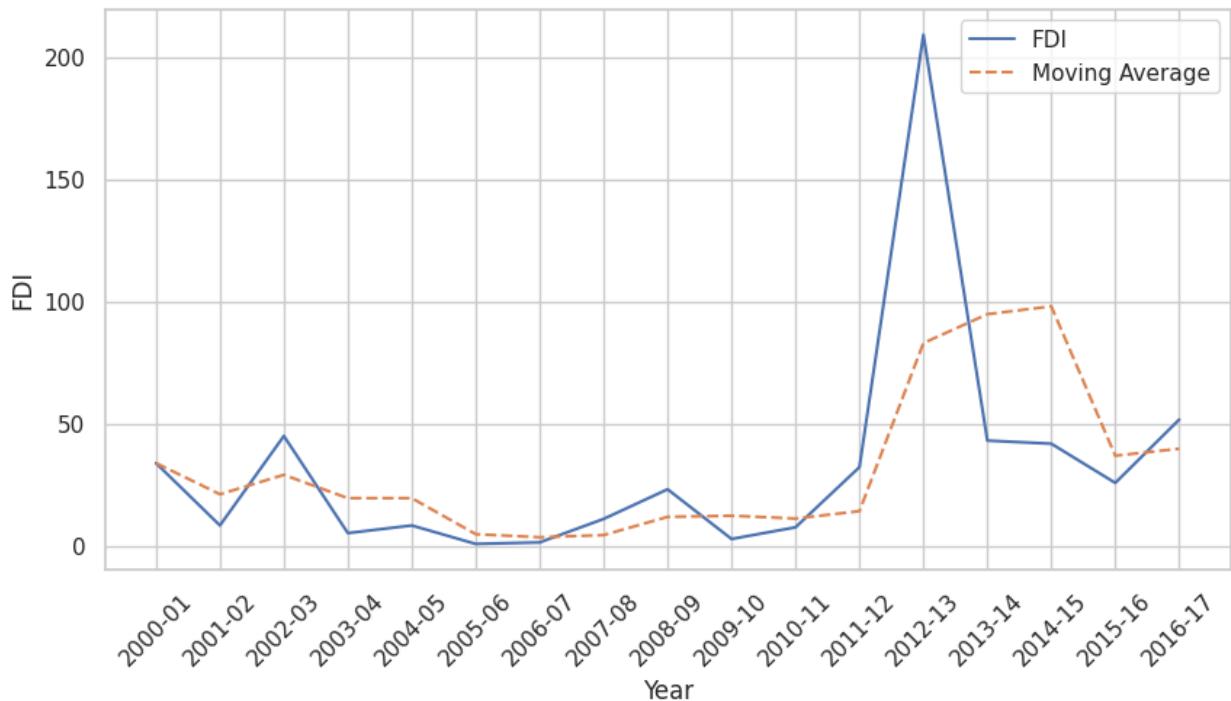
FDI Trends for LEATHER,LEATHER GOODS AND PICKERS



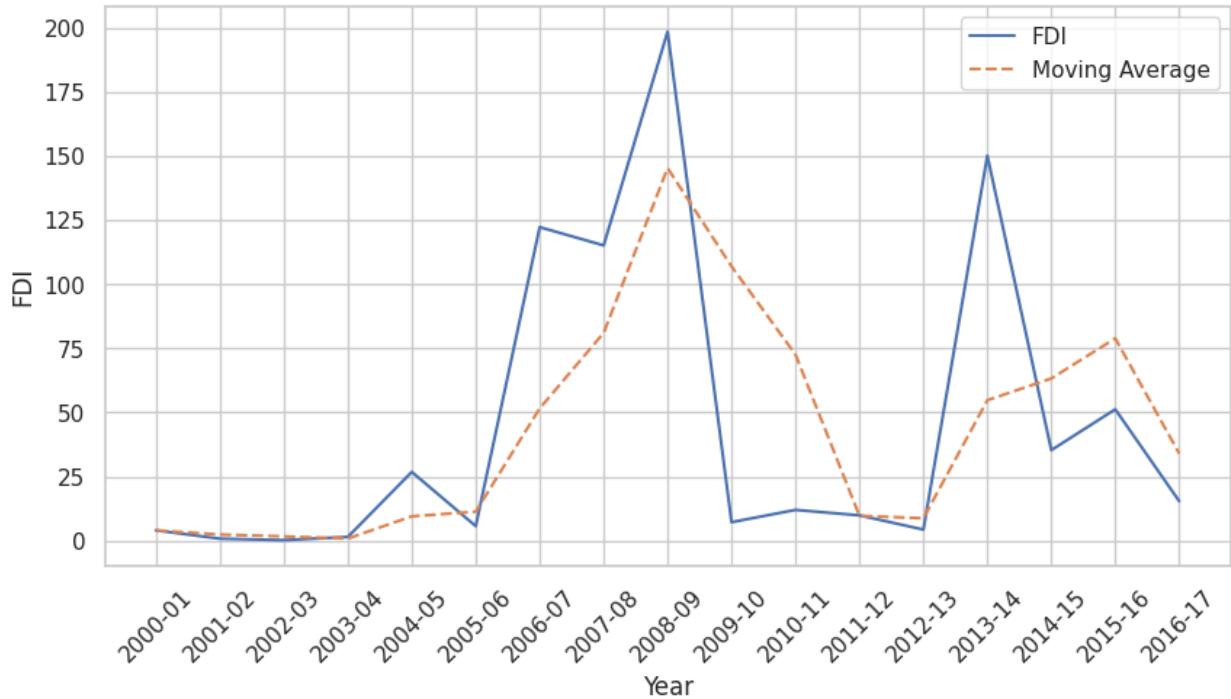
FDI Trends for GLUE AND GELATIN



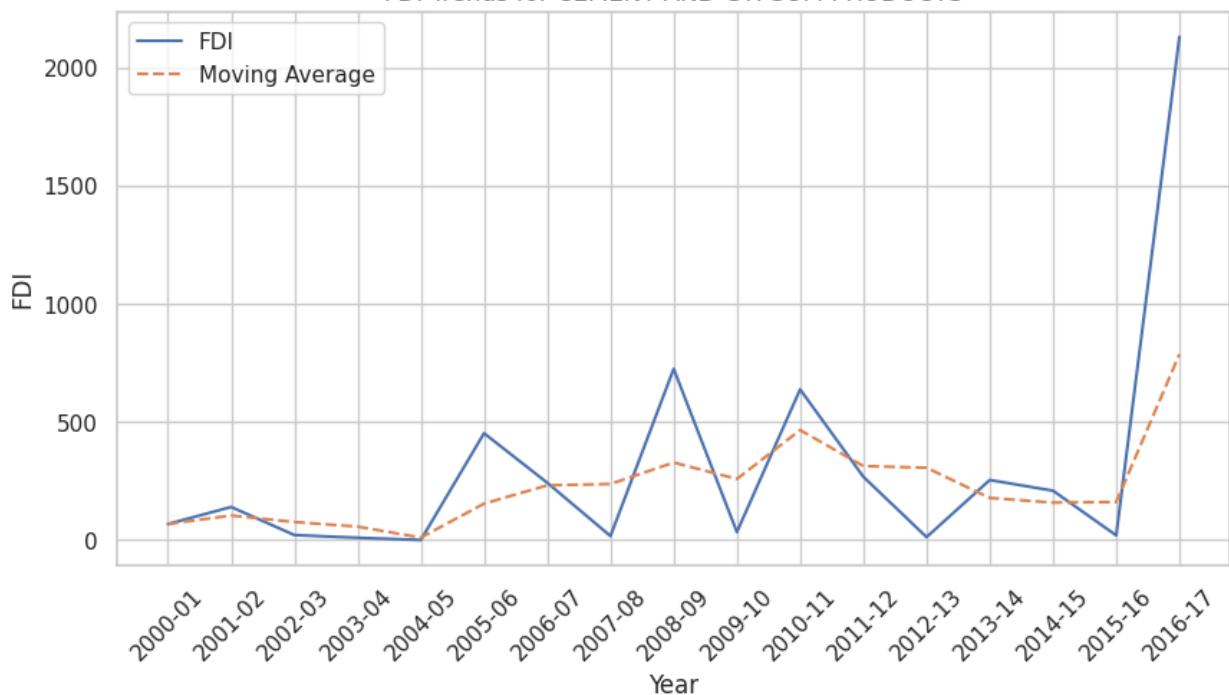
FDI Trends for GLASS



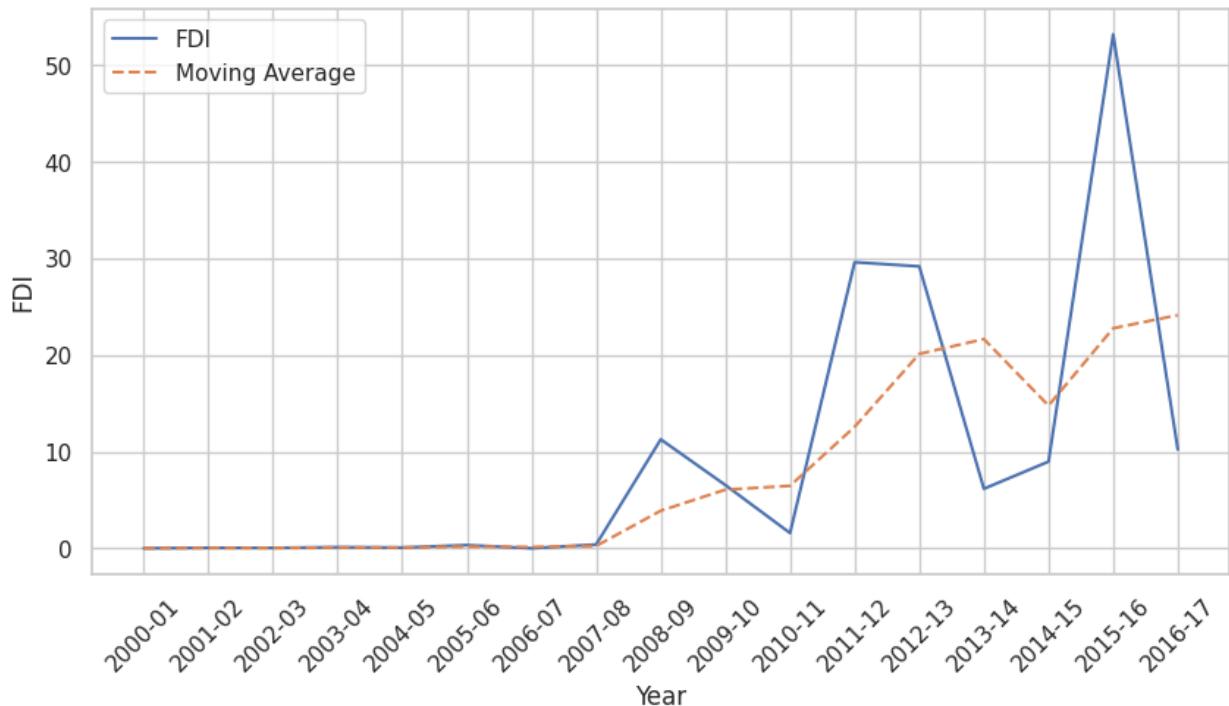
FDI Trends for CERAMICS

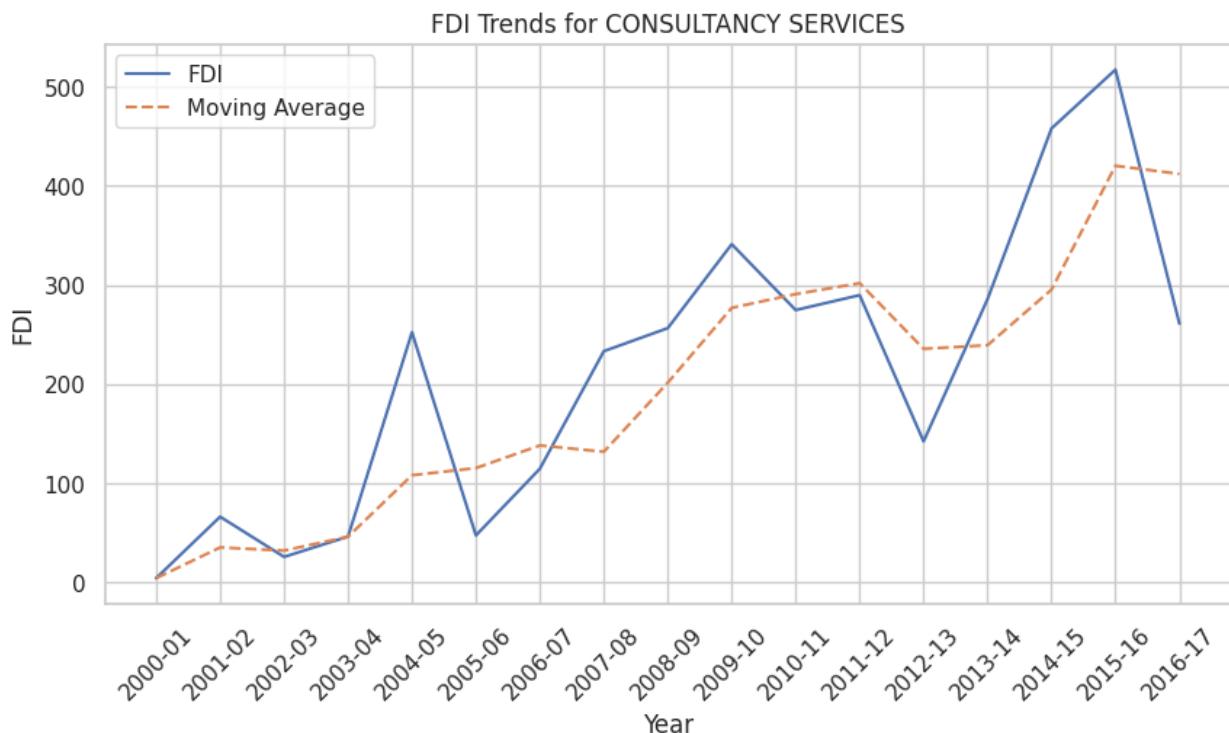
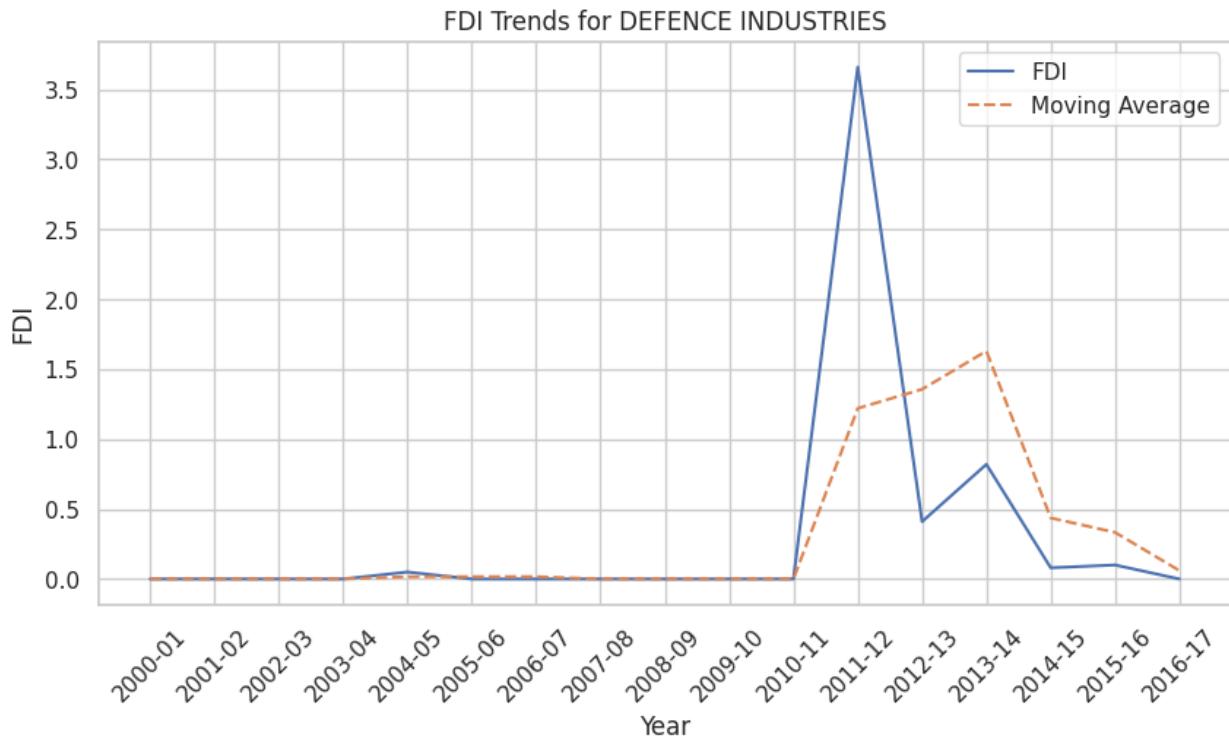


FDI Trends for CEMENT AND GYPSUM PRODUCTS

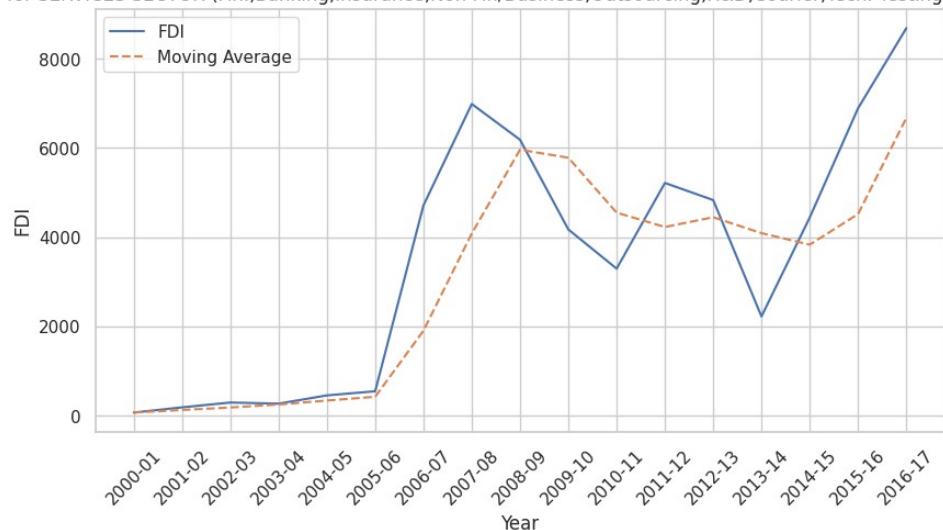


FDI Trends for TIMBER PRODUCTS

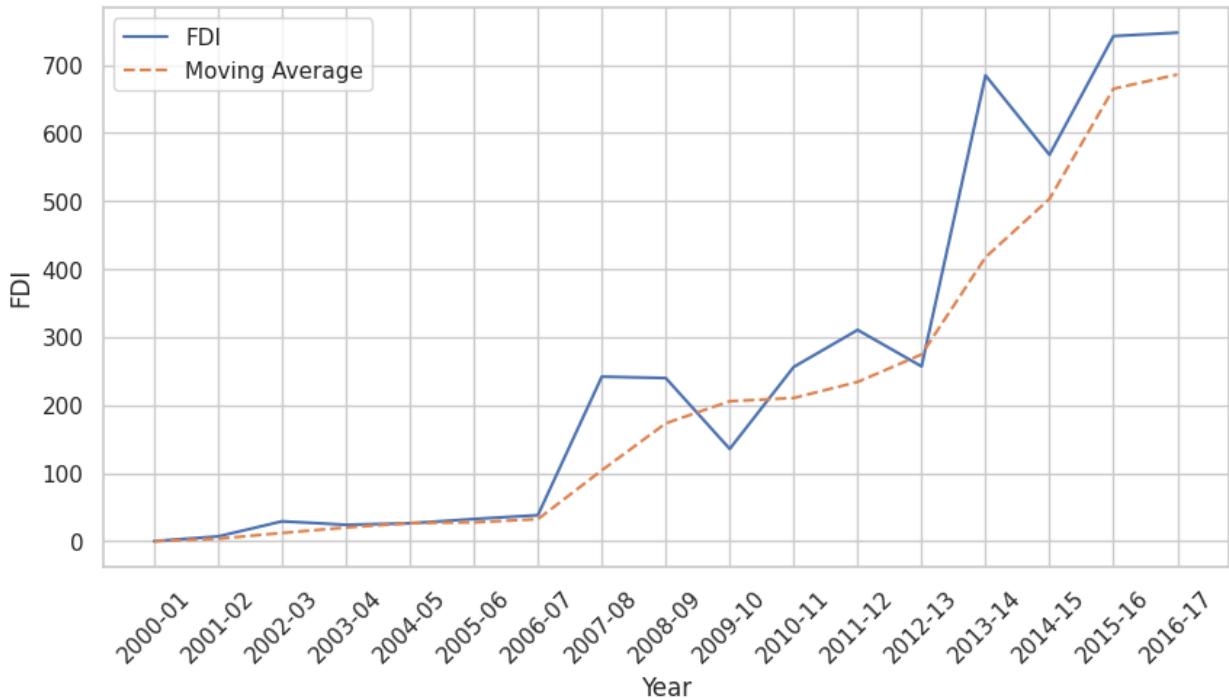




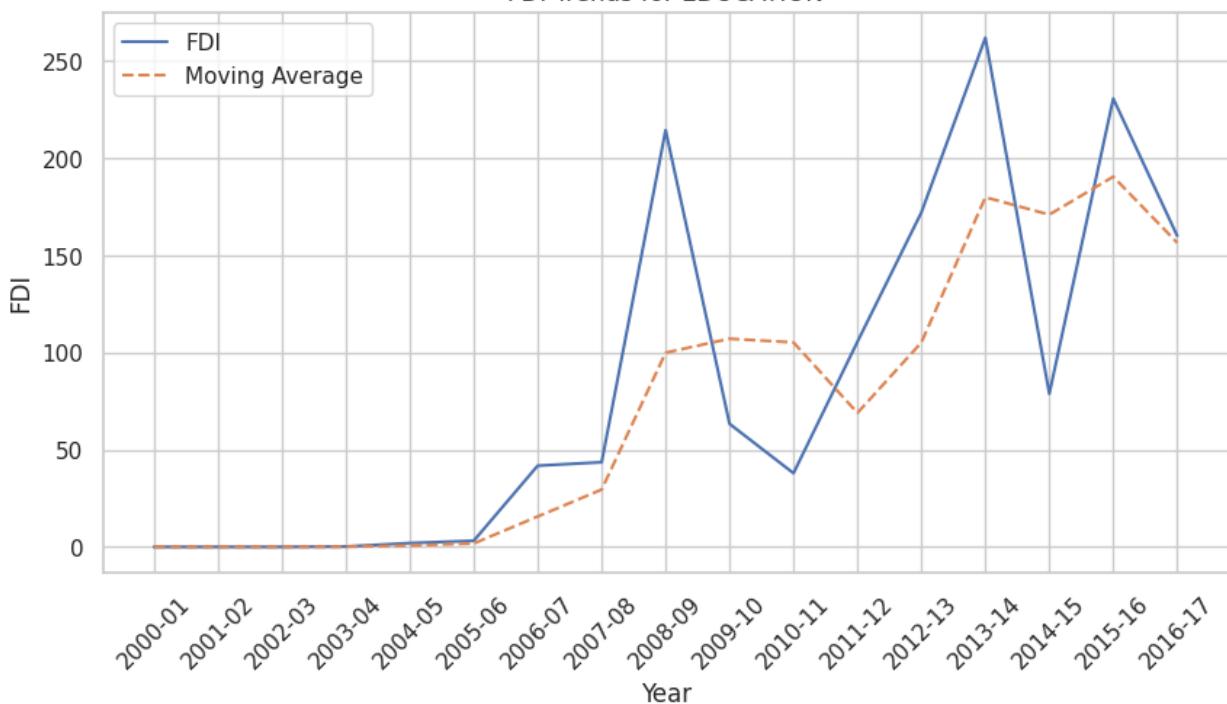
FDI Trends for SERVICES SECTOR (Fin.,Banking,Insurance,Non Fin/Business,Outsourcing,R&D,Courier,Tech. Testing and Analysis, Other)



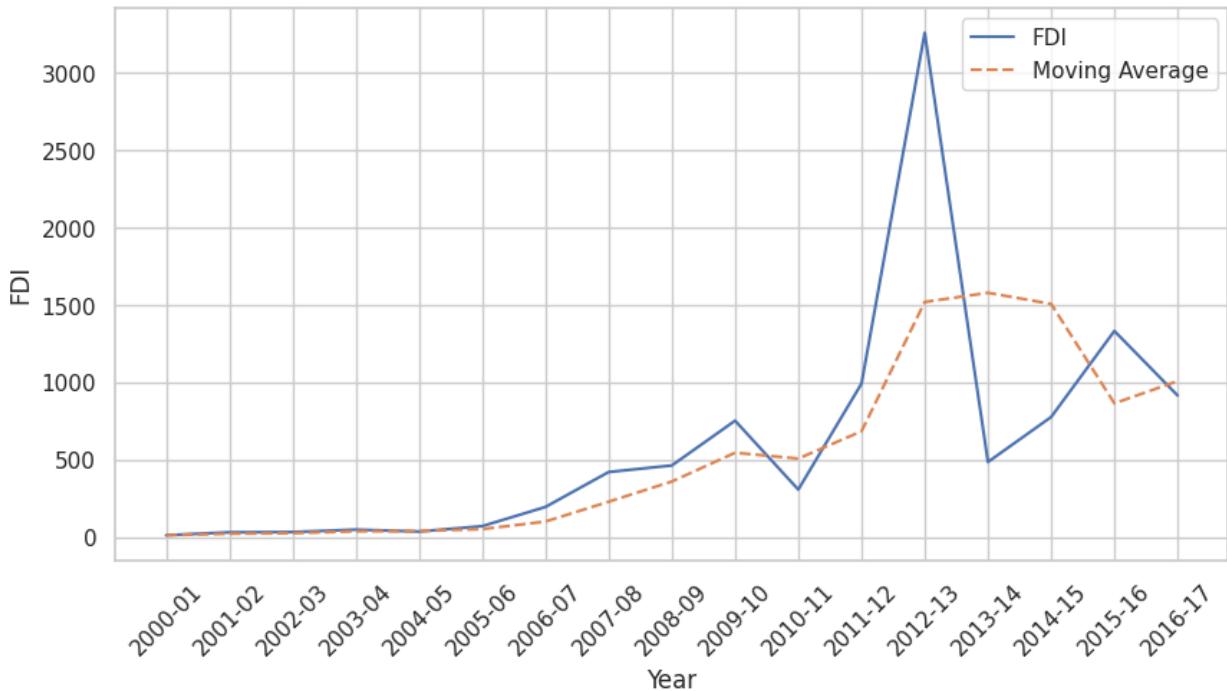
FDI Trends for HOSPITAL & DIAGNOSTIC CENTRES

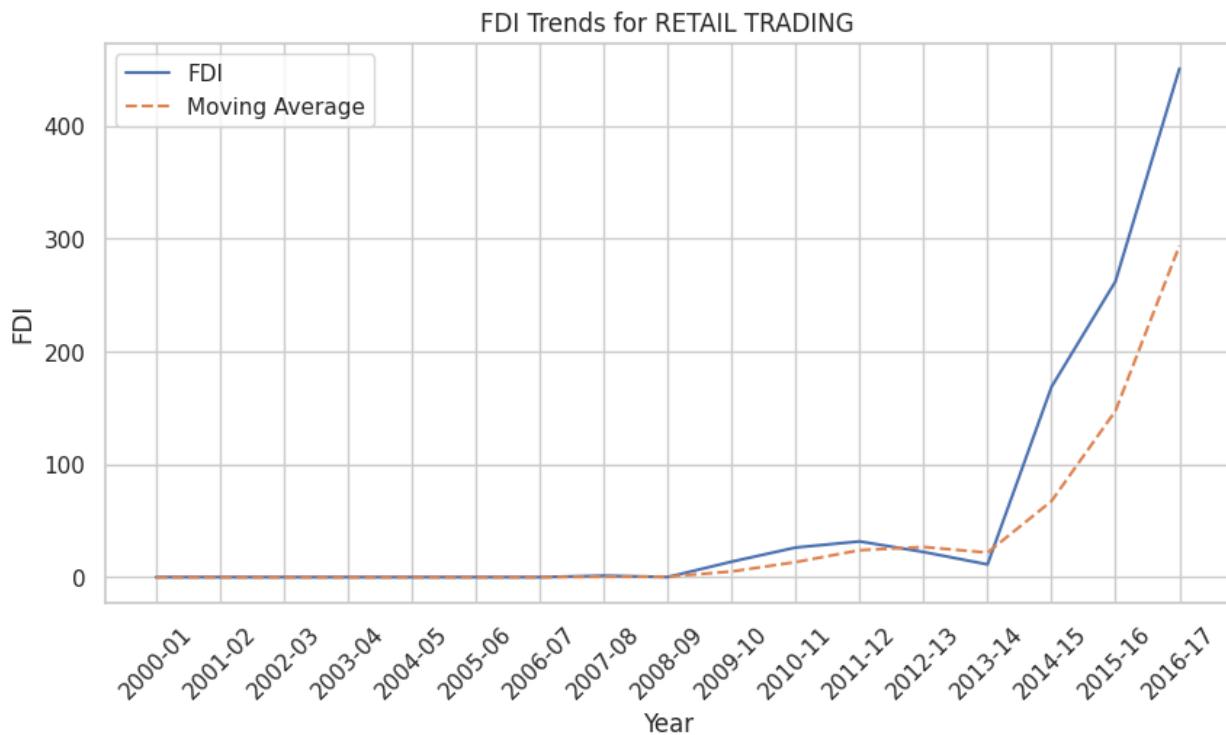
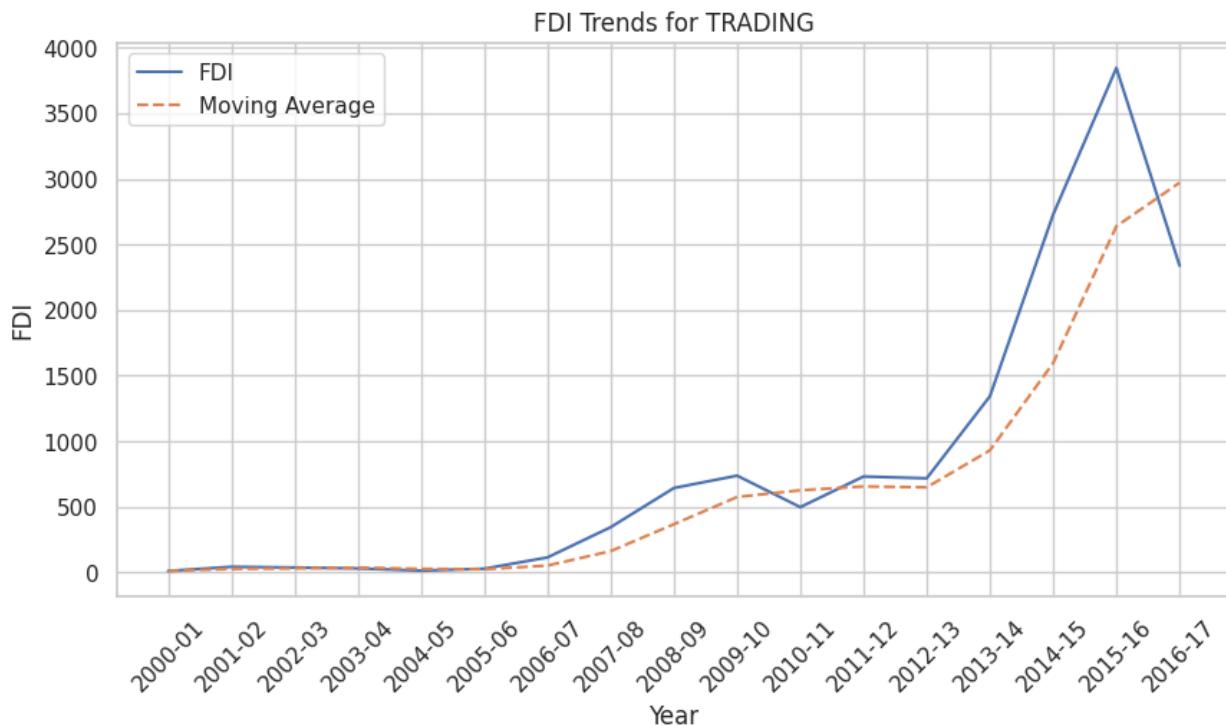


FDI Trends for EDUCATION

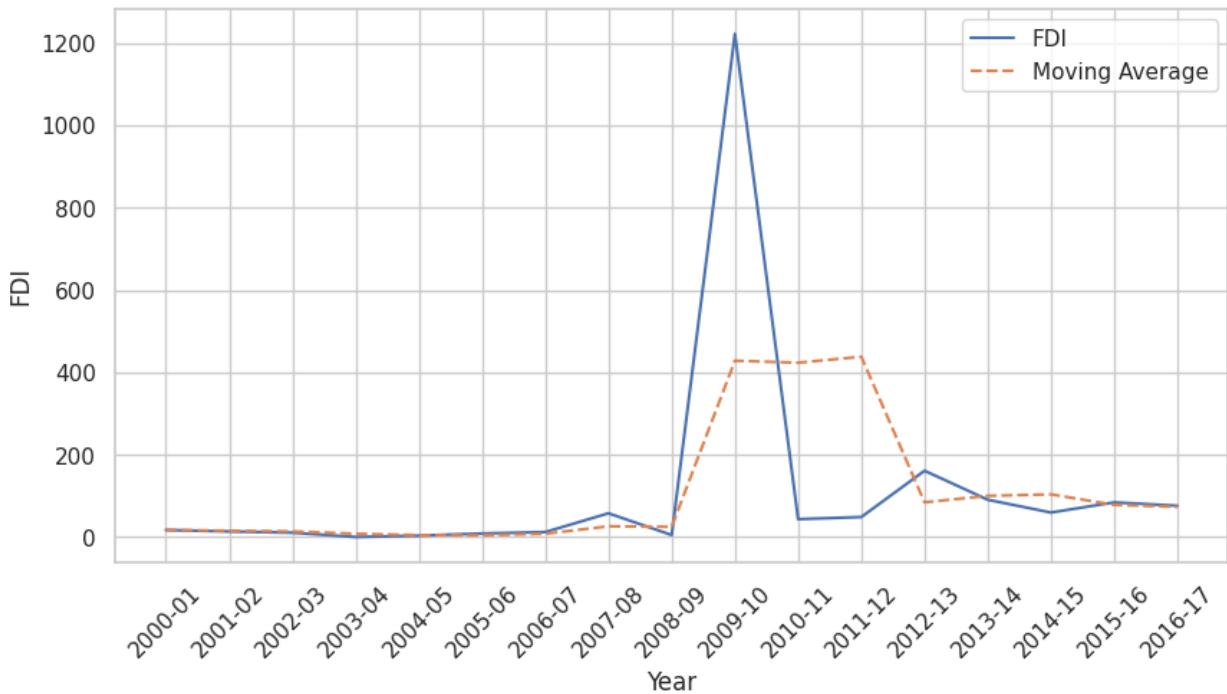


FDI Trends for HOTEL & TOURISM

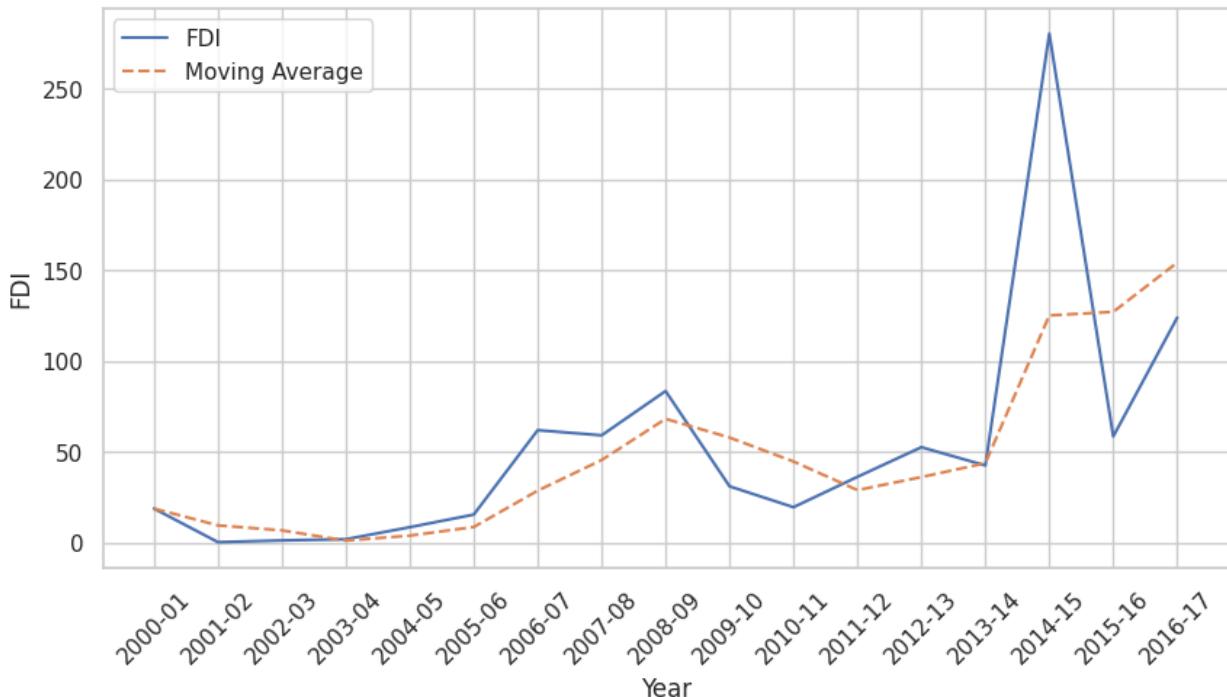




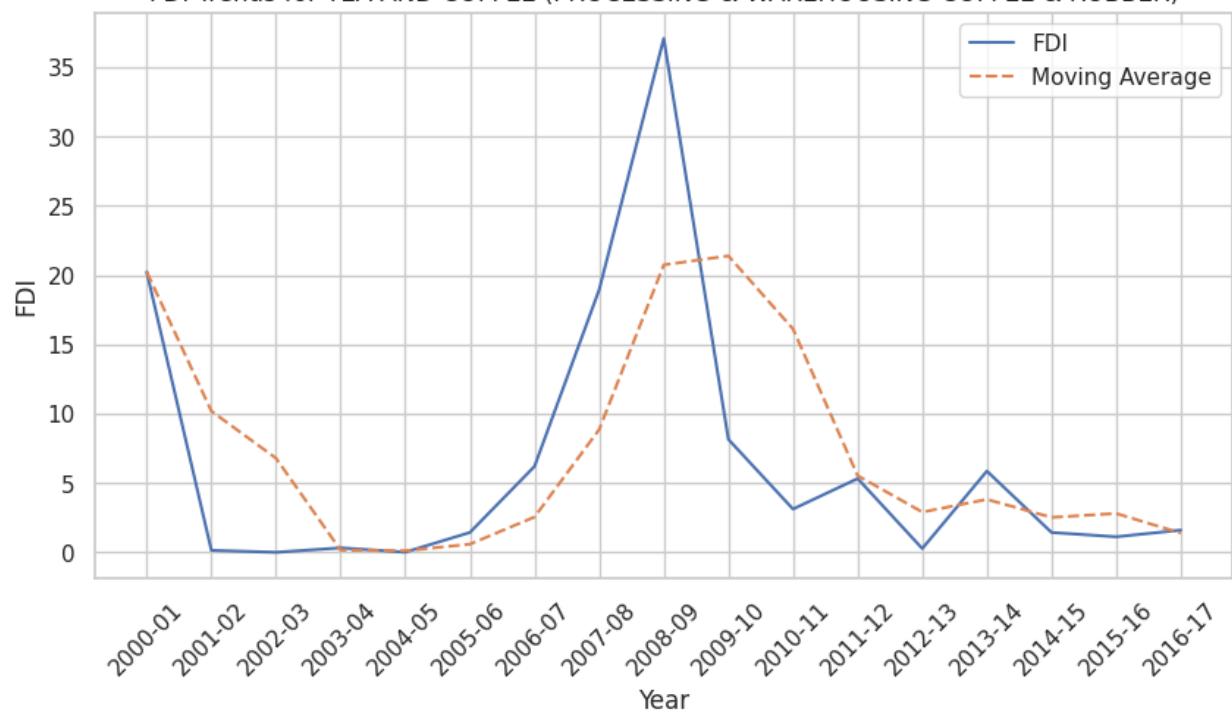
FDI Trends for AGRICULTURE SERVICES



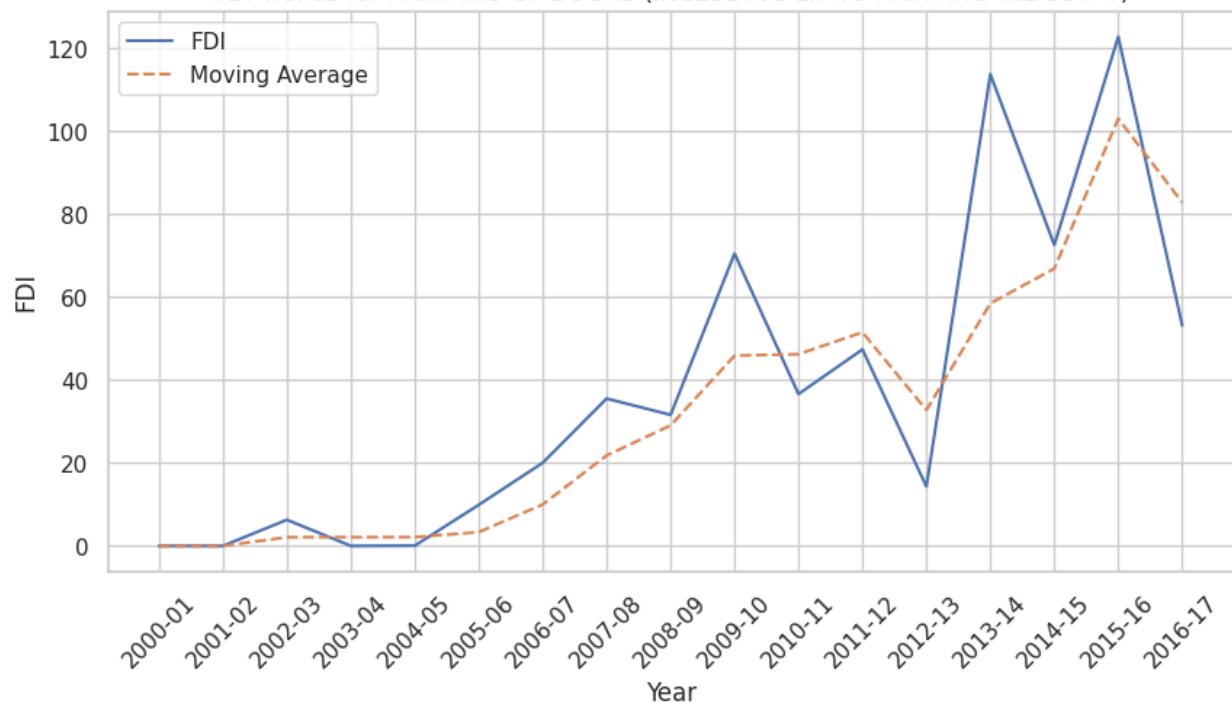
FDI Trends for DIAMOND,GOLD ORNAMENTS

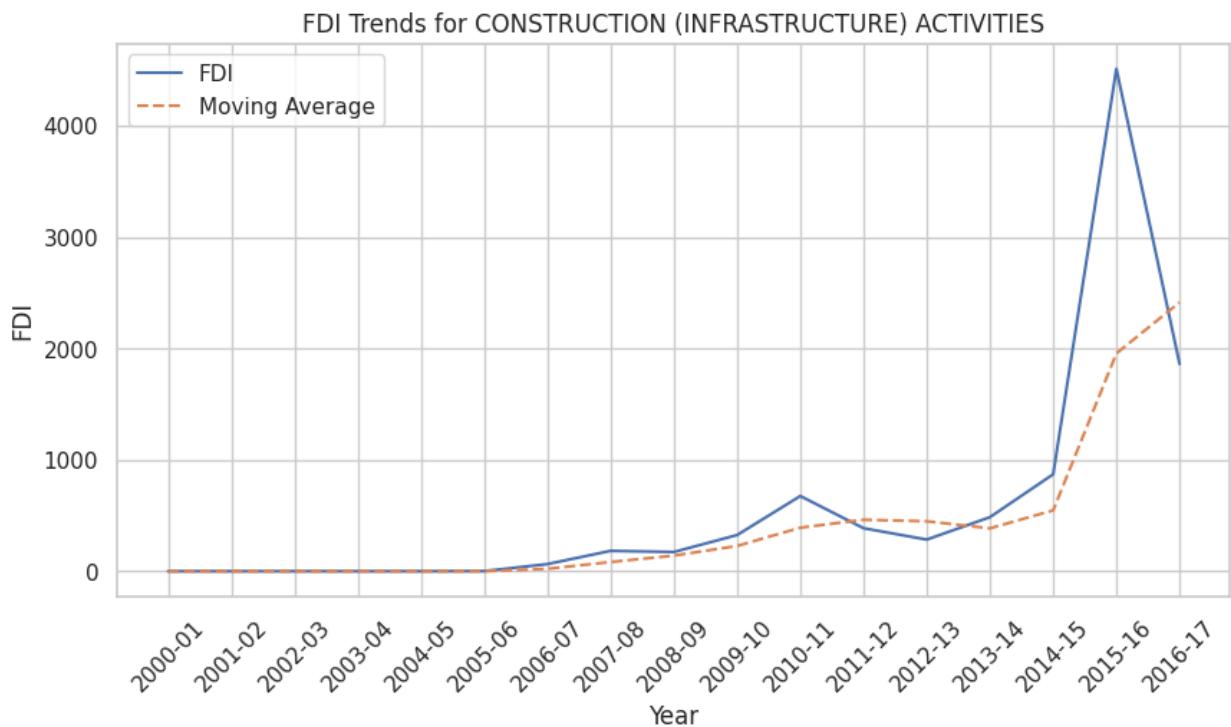
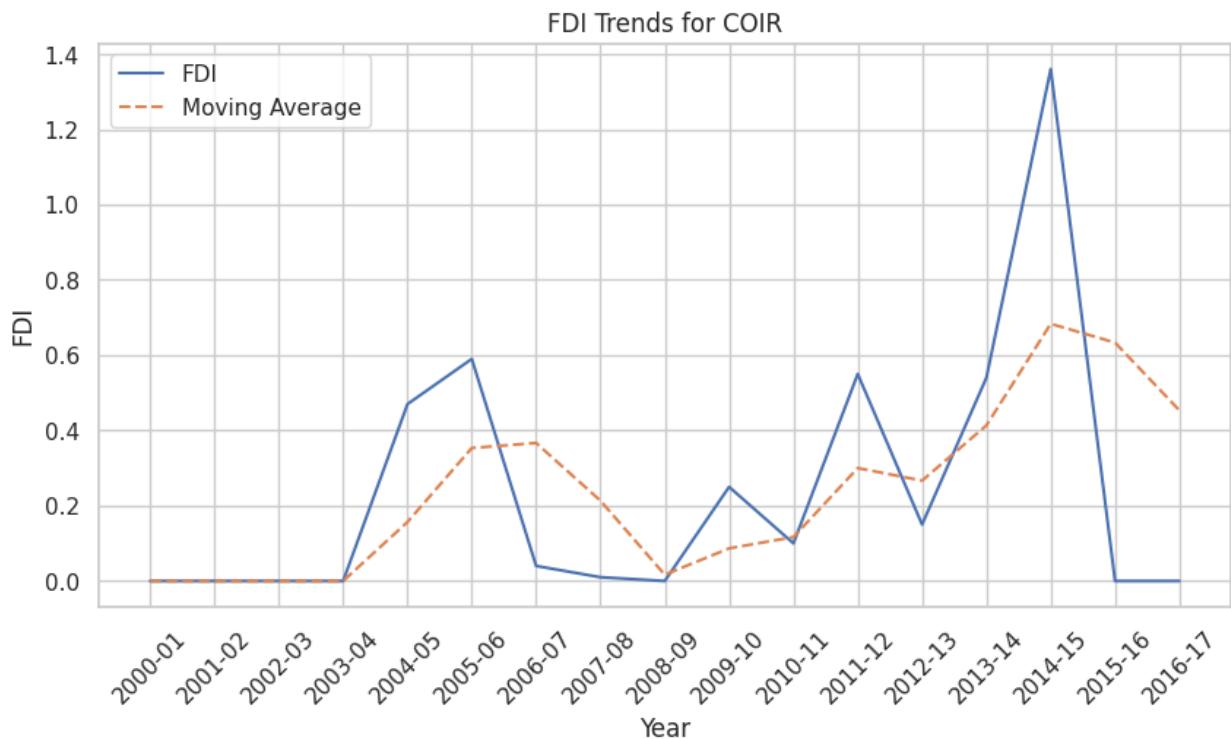


FDI Trends for TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)

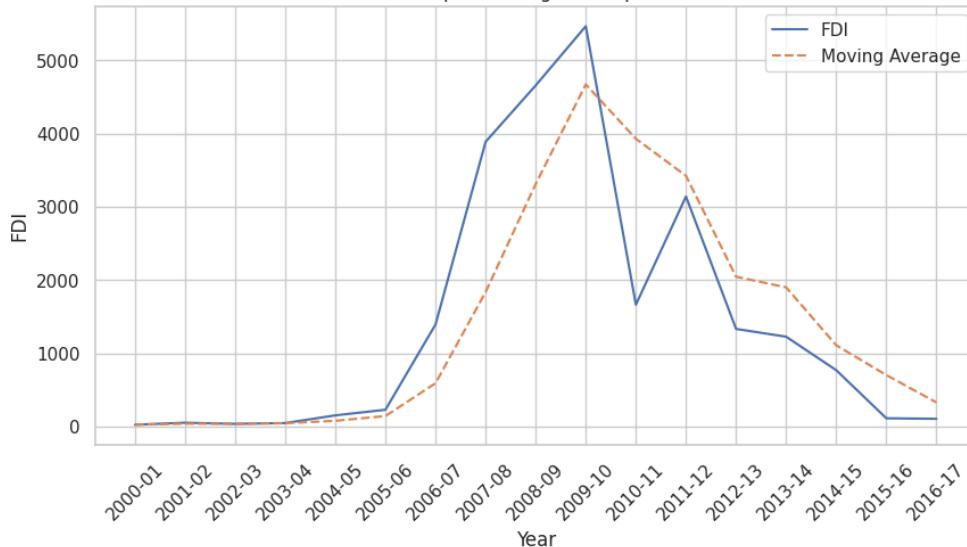


FDI Trends for PRINTING OF BOOKS (INCLUDING LITHO PRINTING INDUSTRY)

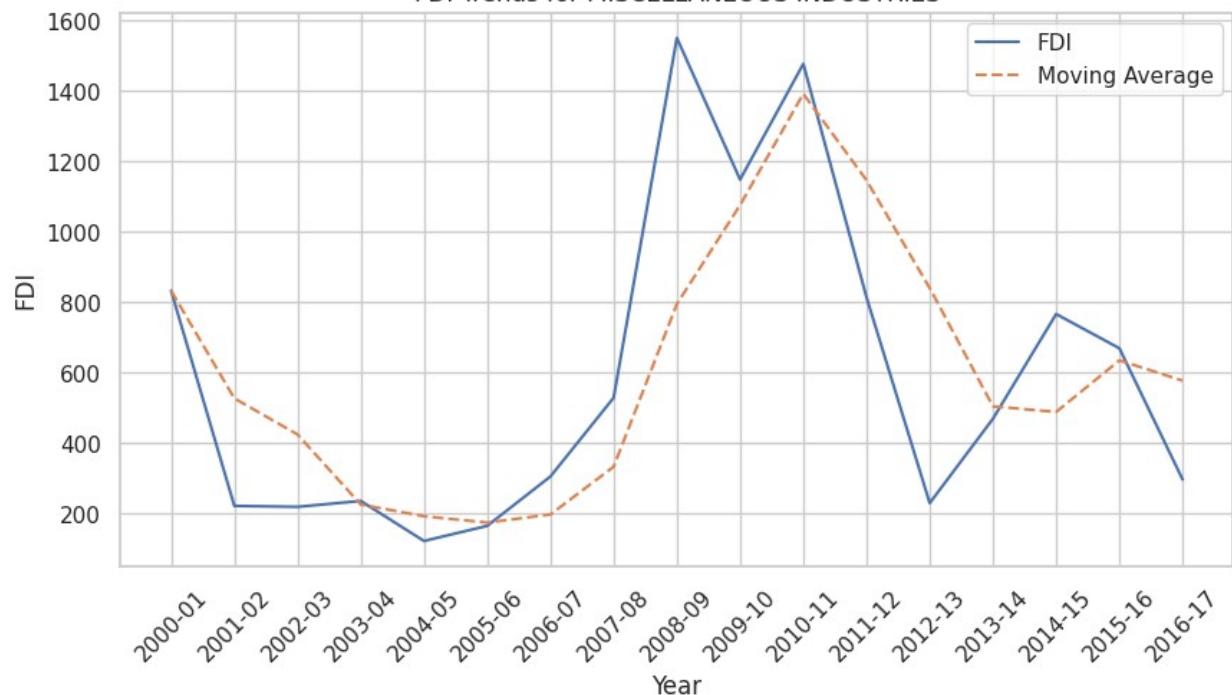




FDI Trends for CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects



FDI Trends for MISCELLANEOUS INDUSTRIES



```

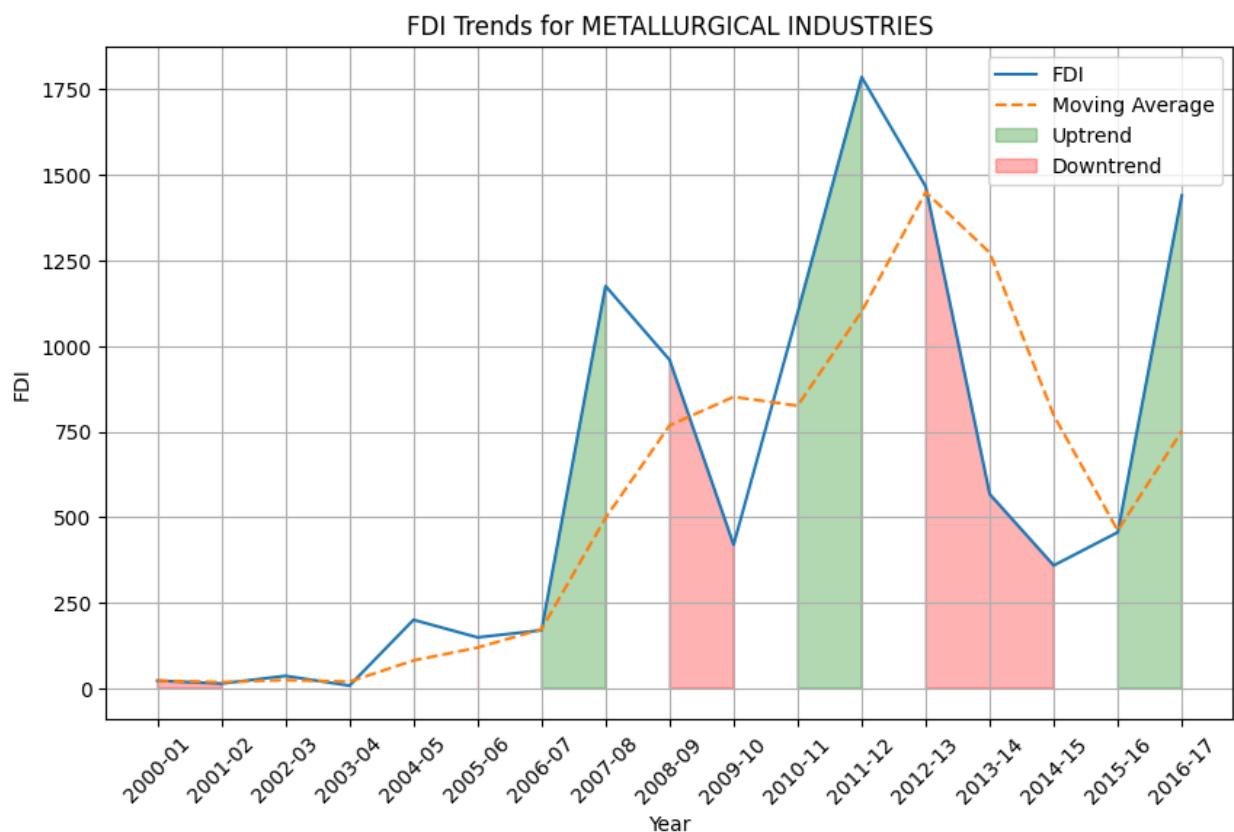
for sector in sectors:
    sector_data = df_long[df_long['Sector'] == sector]
    plt.figure(figsize=(10, 6))
    plt.plot(sector_data['Year'], sector_data['FDI'], label='FDI')
    plt.plot(sector_data['Year'], sector_data['Moving Average'], '--',
             label='Moving Average')
    plt.fill_between(sector_data['Year'], sector_data['FDI'],
                    where=sector_data['Growth Rate'] > 0, color='green', alpha=0.3,
                    label='Uptrend')

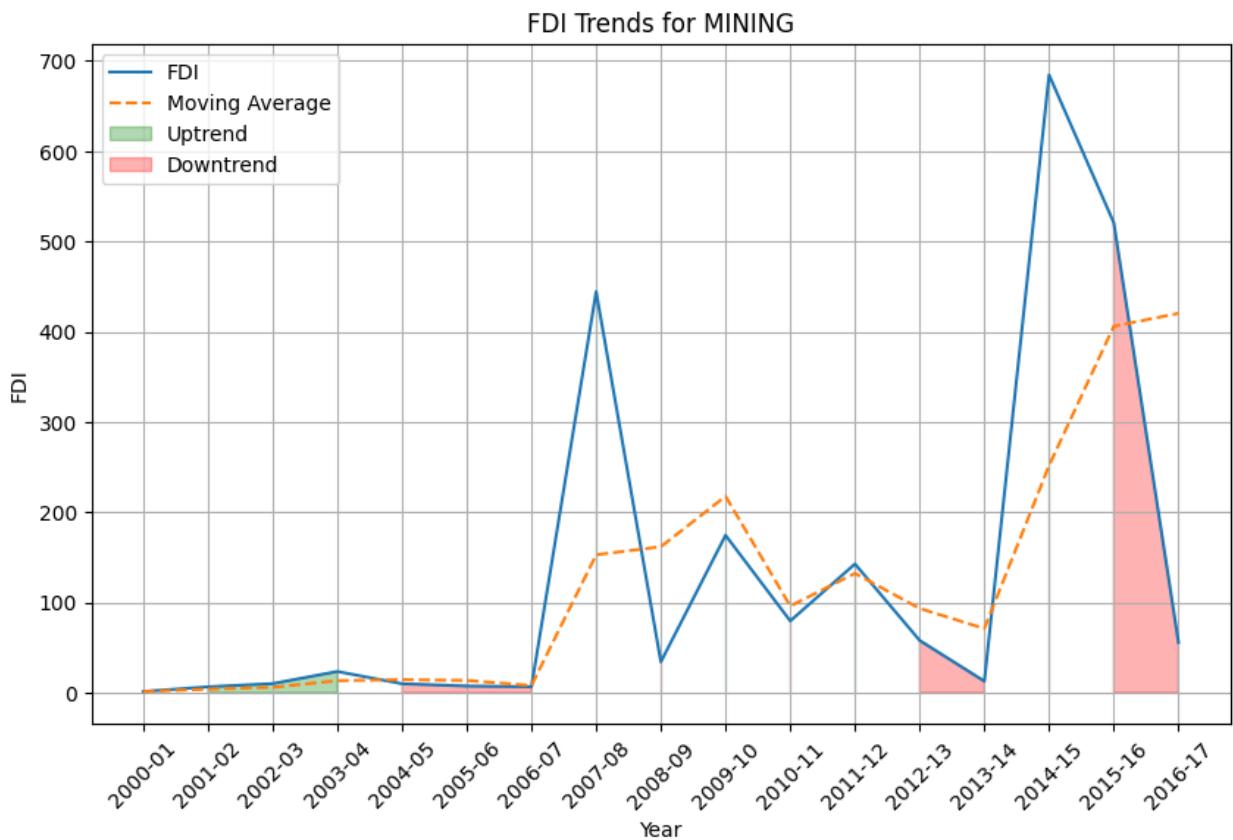
```

```

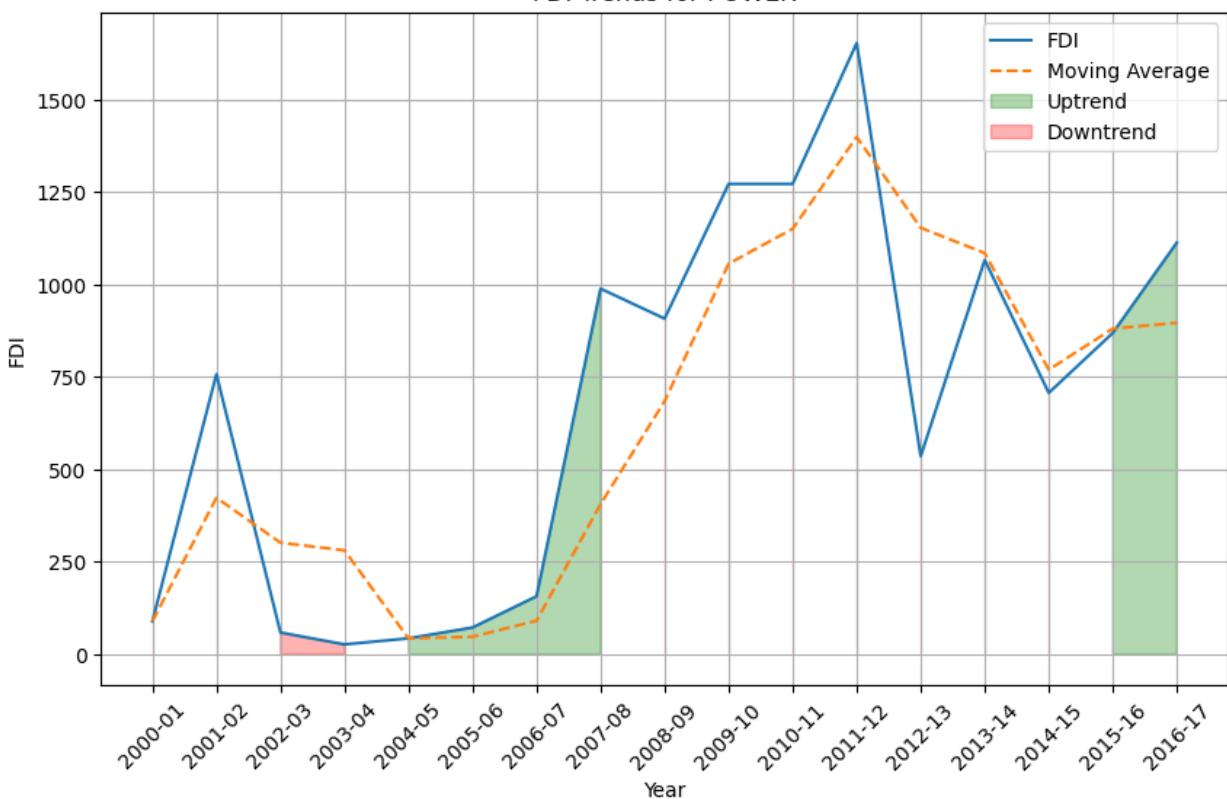
plt.fill_between(sector_data['Year'], sector_data['FDI'],
where=sector_data['Growth Rate'] <= 0, color='red', alpha=0.3,
label='Downtrend')
plt.title(f'FDI Trends for {sector}')
plt.xlabel('Year')
plt.xticks(rotation=45)
plt.ylabel('FDI')
plt.legend()
plt.grid(True)
plt.show()

```

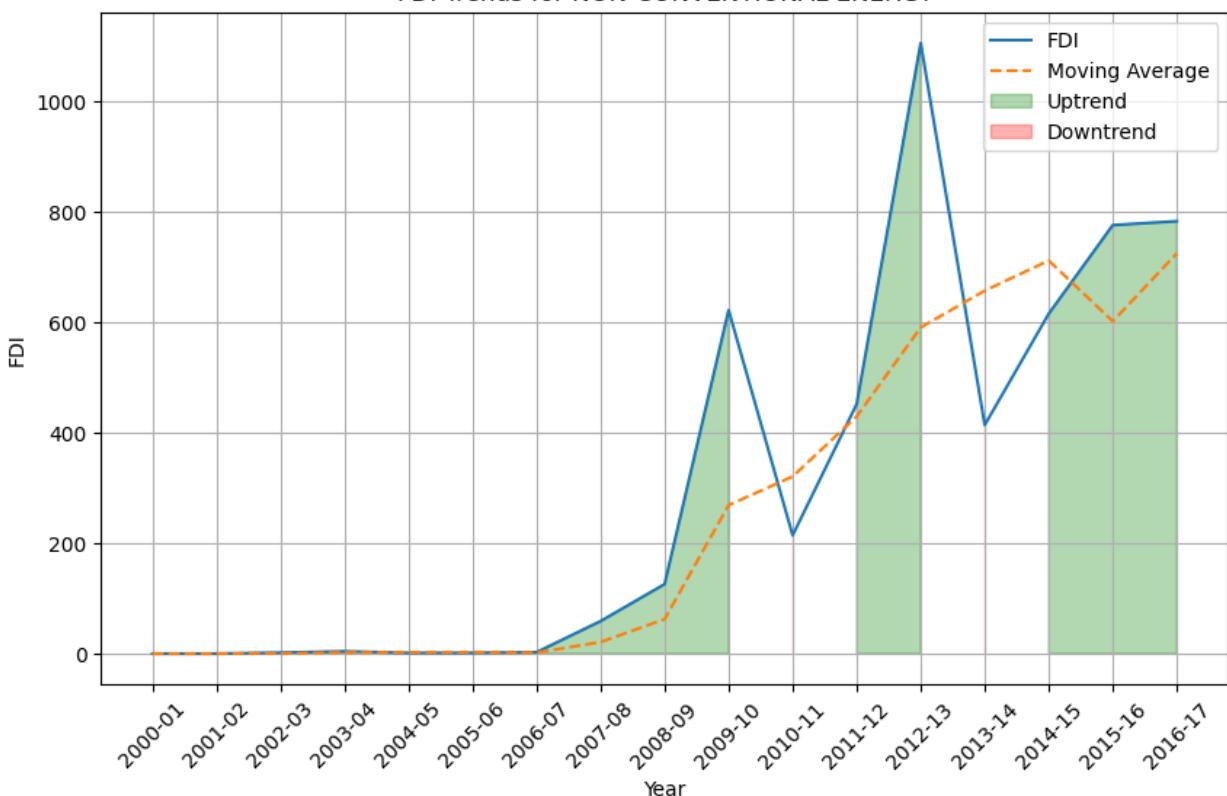




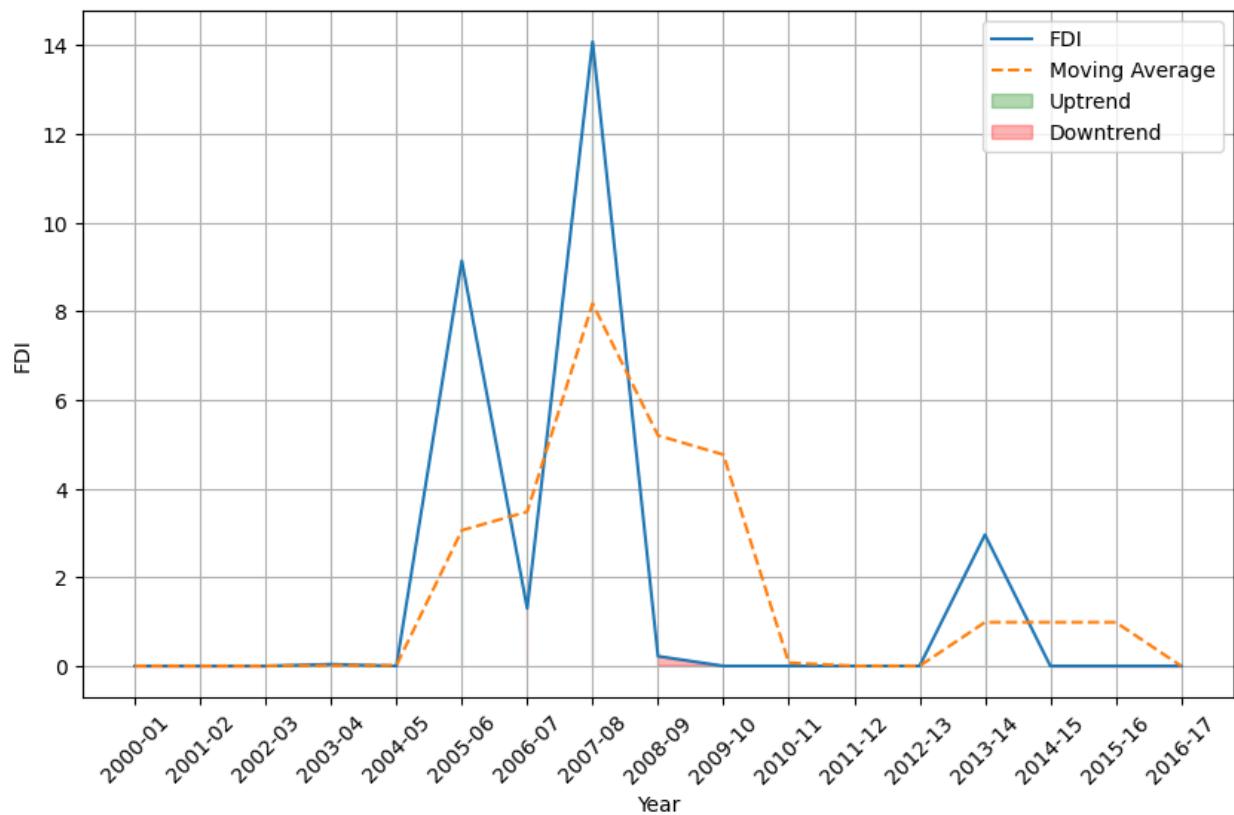
FDI Trends for POWER



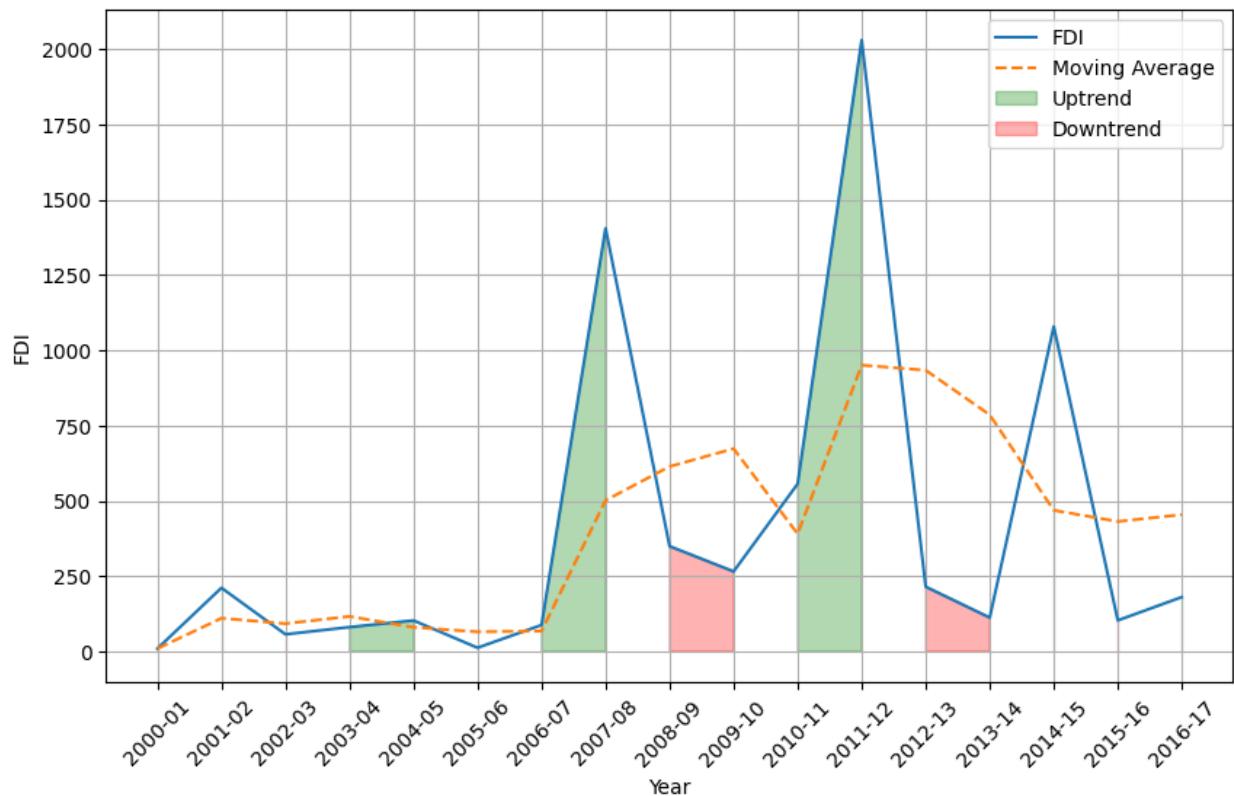
FDI Trends for NON-CONVENTIONAL ENERGY



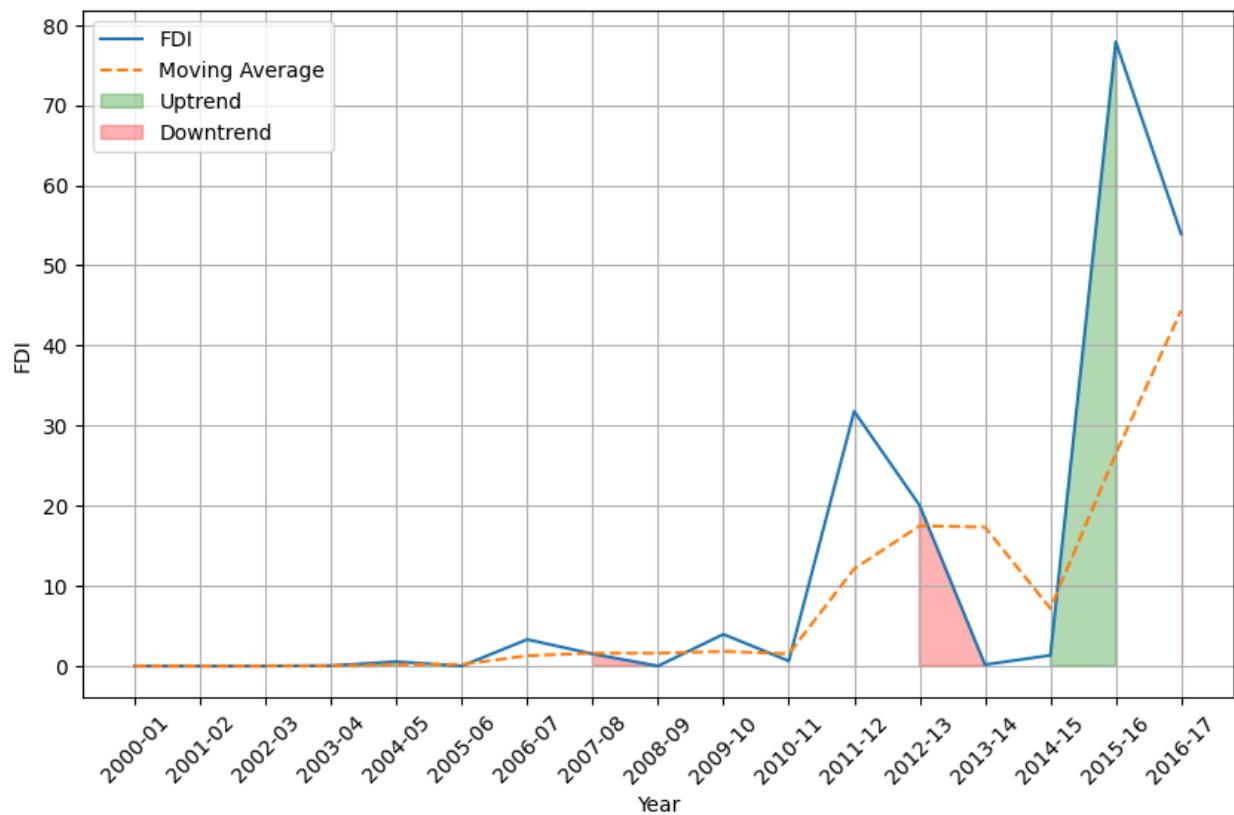
FDI Trends for COAL PRODUCTION



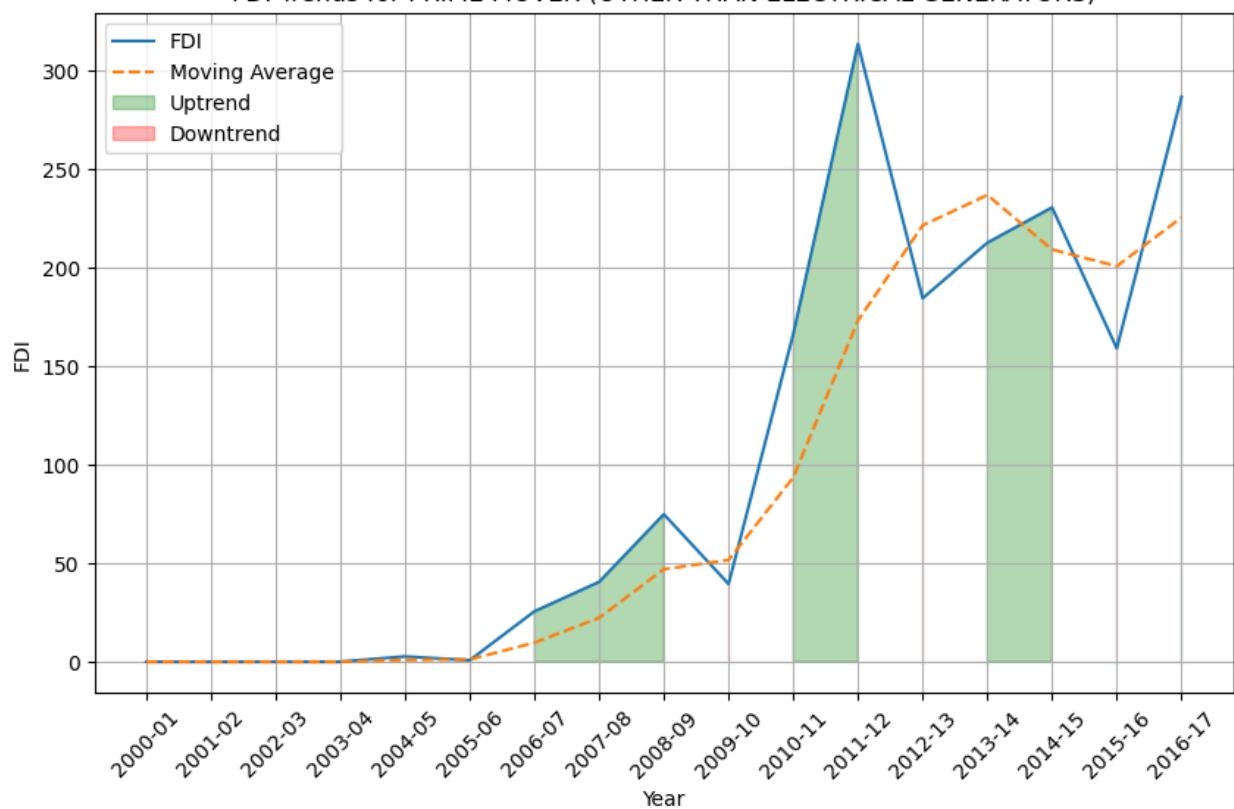
FDI Trends for PETROLEUM & NATURAL GAS



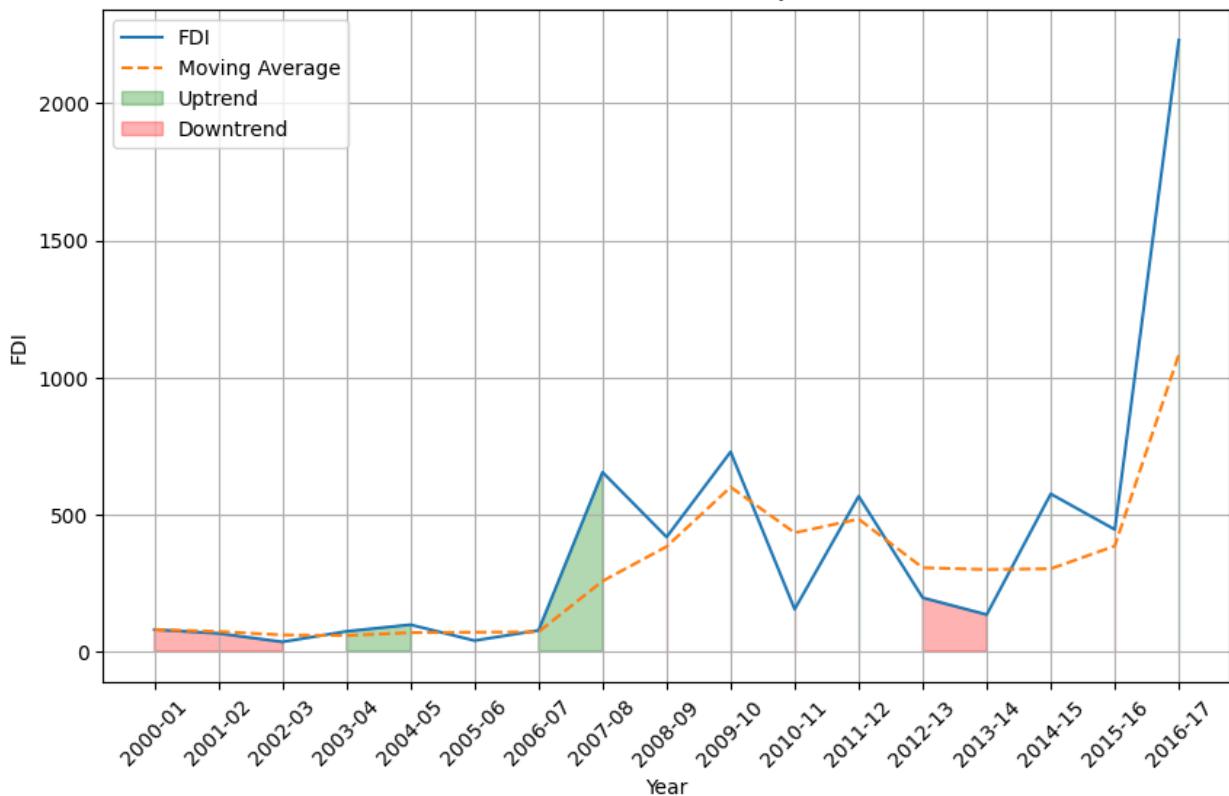
FDI Trends for BOILERS AND STEAM GENERATING PLANTS



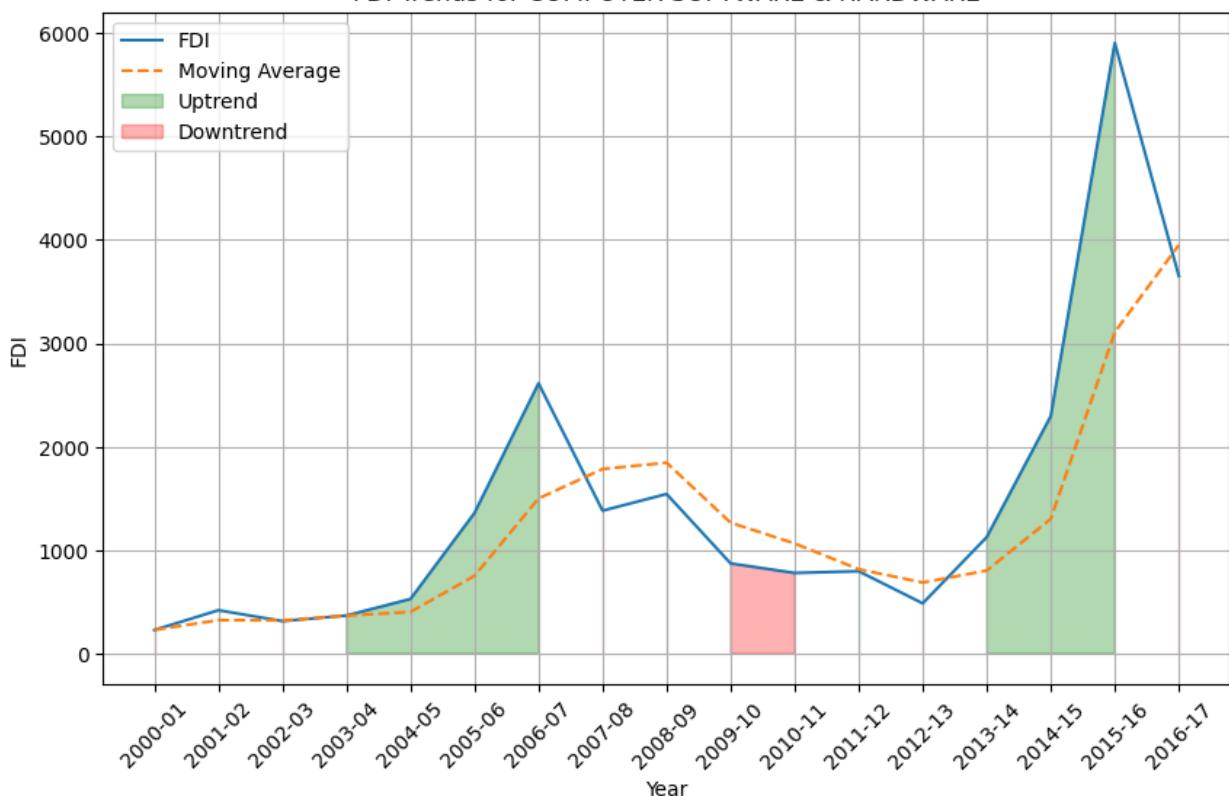
FDI Trends for PRIME MOVER (OTHER THAN ELECTRICAL GENERATORS)

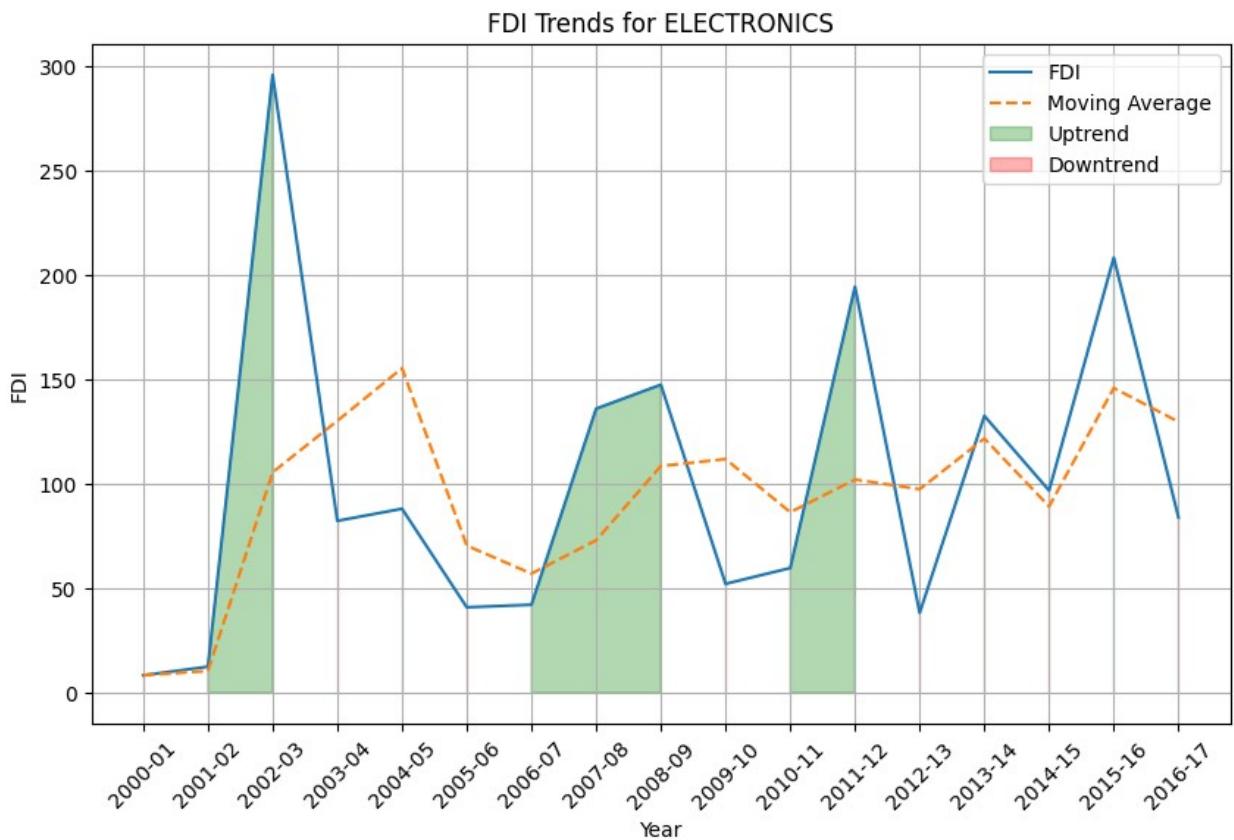


FDI Trends for ELECTRICAL EQUIPMENTS

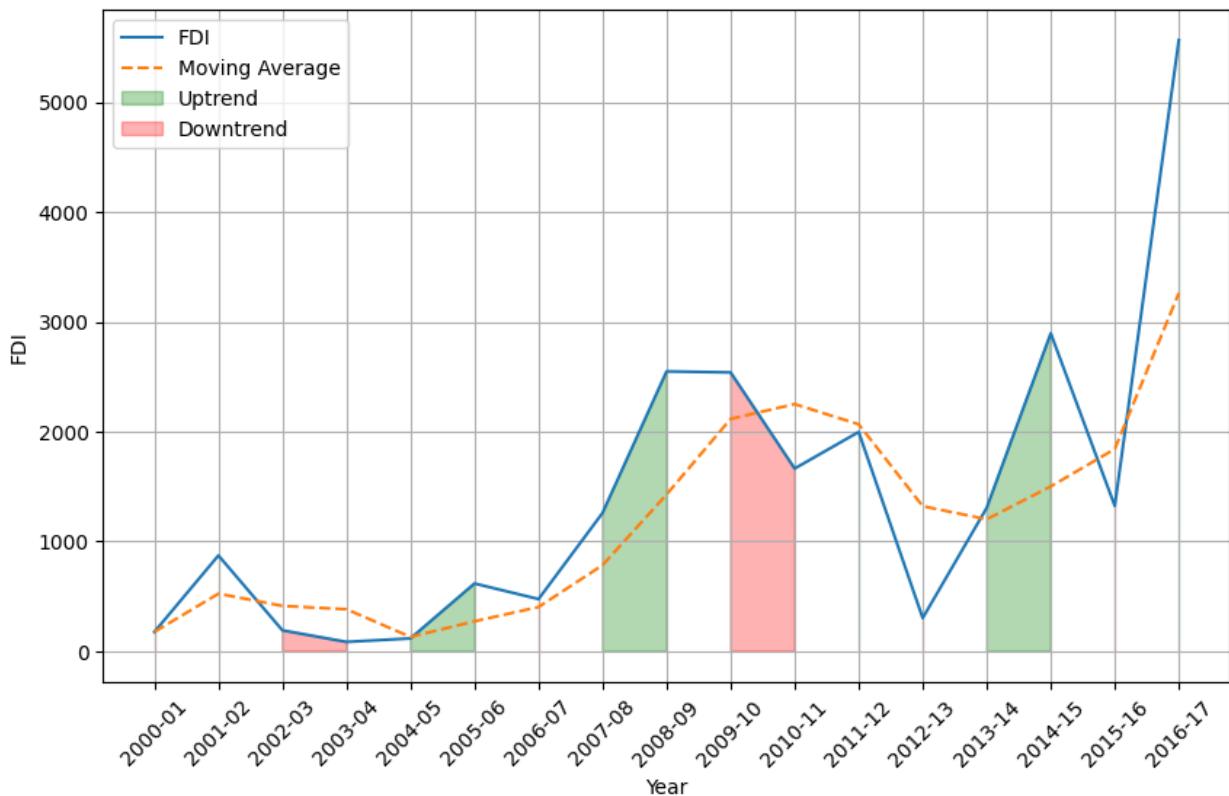


FDI Trends for COMPUTER SOFTWARE & HARDWARE

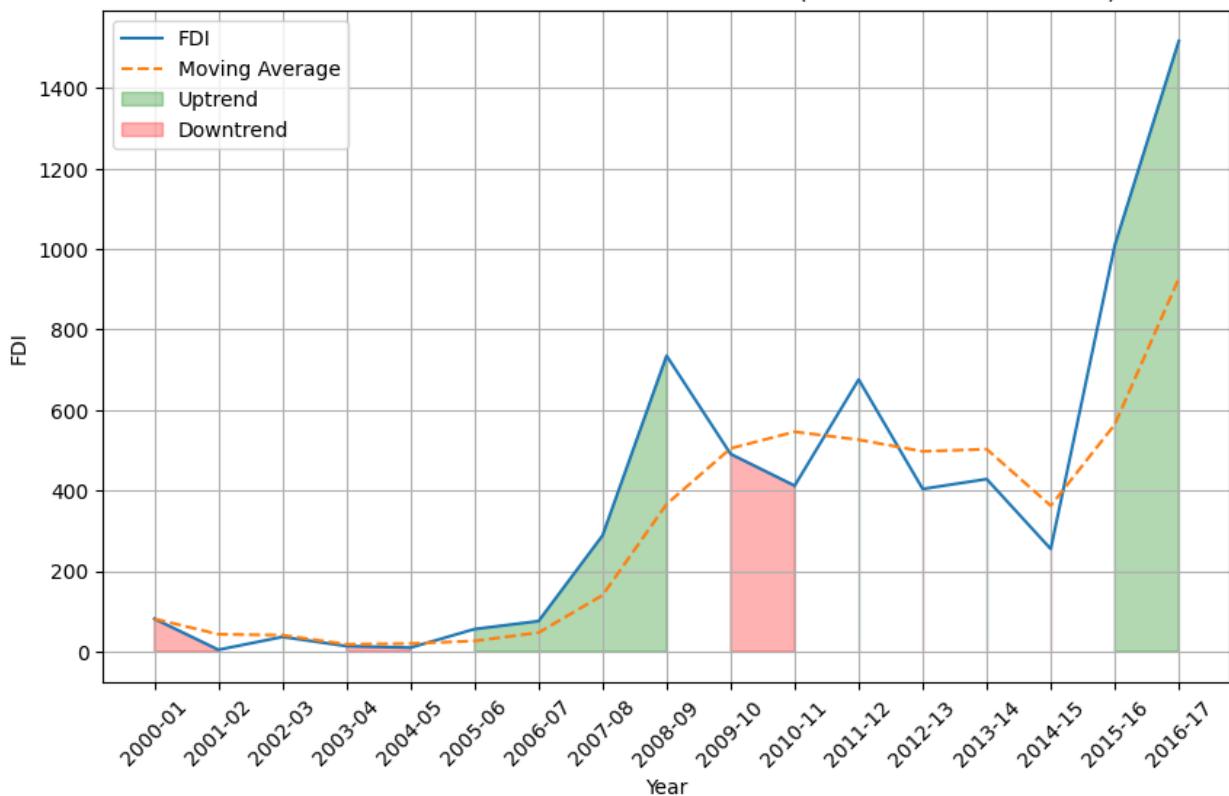




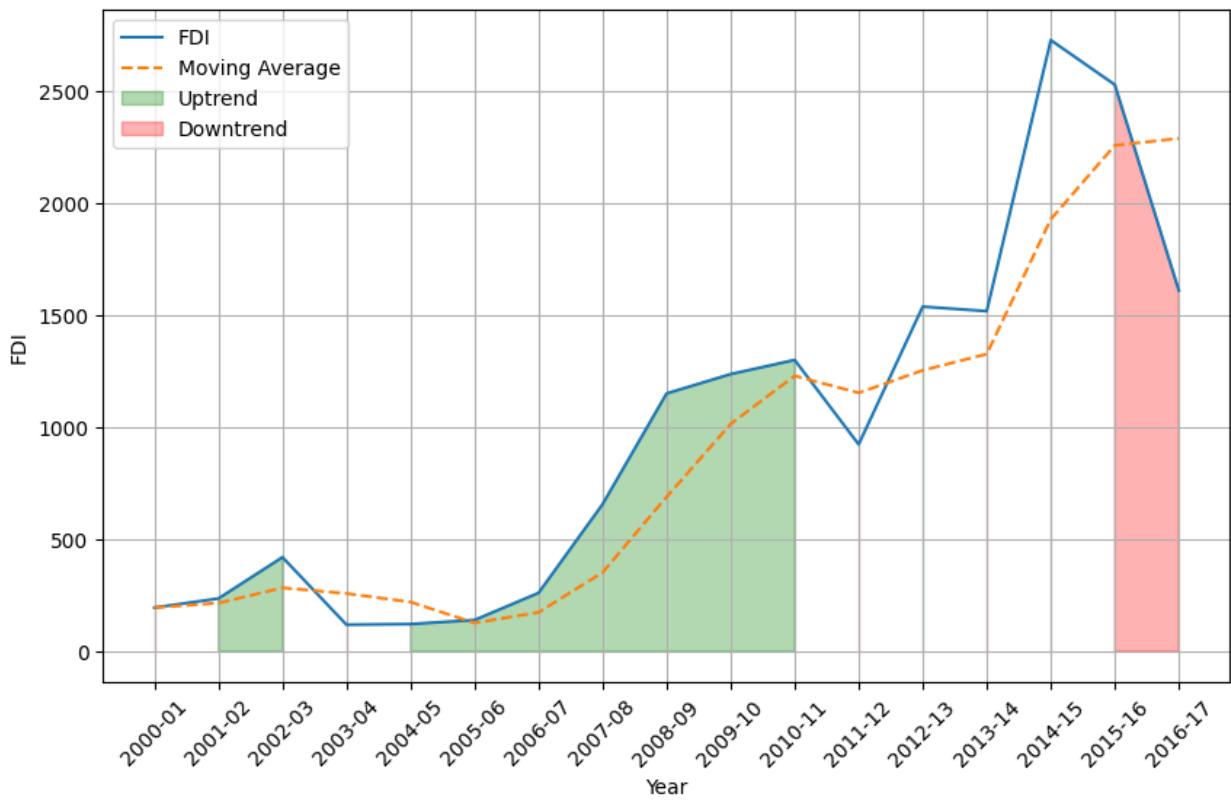
FDI Trends for TELECOMMUNICATIONS



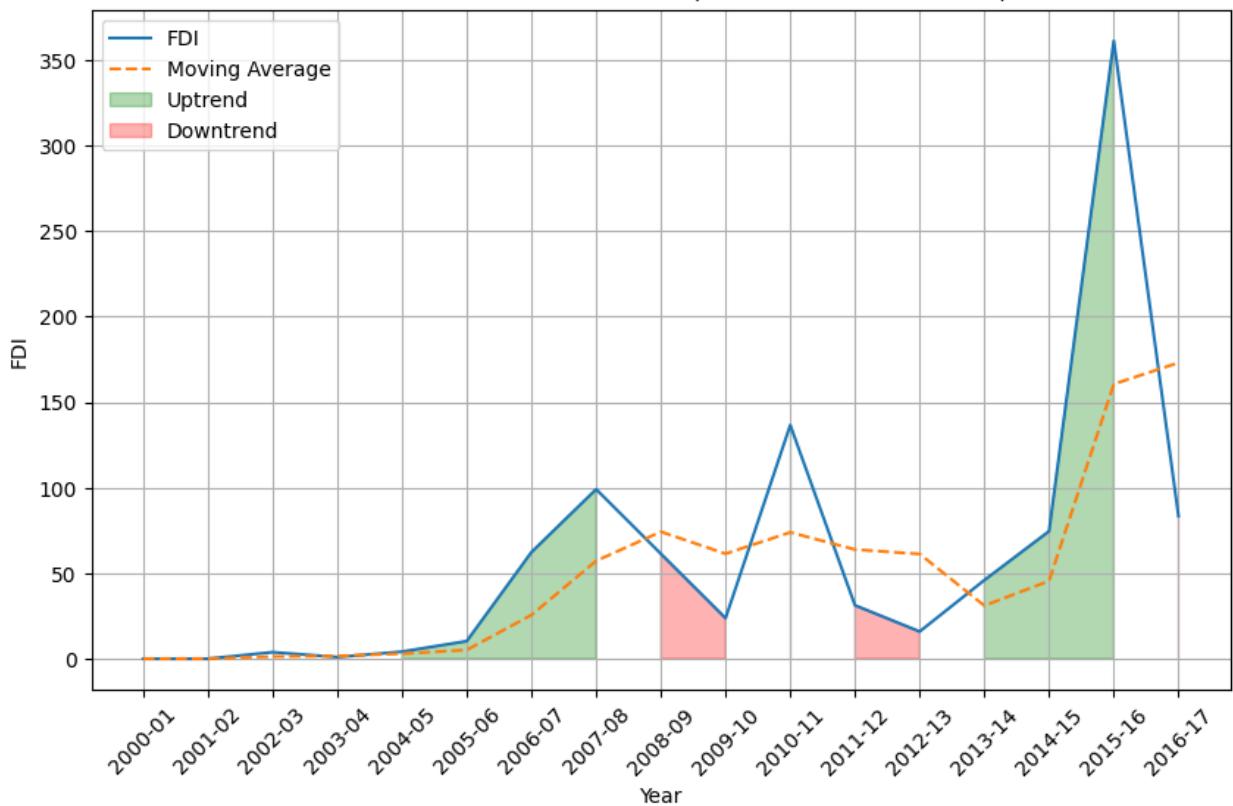
FDI Trends for INFORMATION & BROADCASTING (INCLUDING PRINT MEDIA)



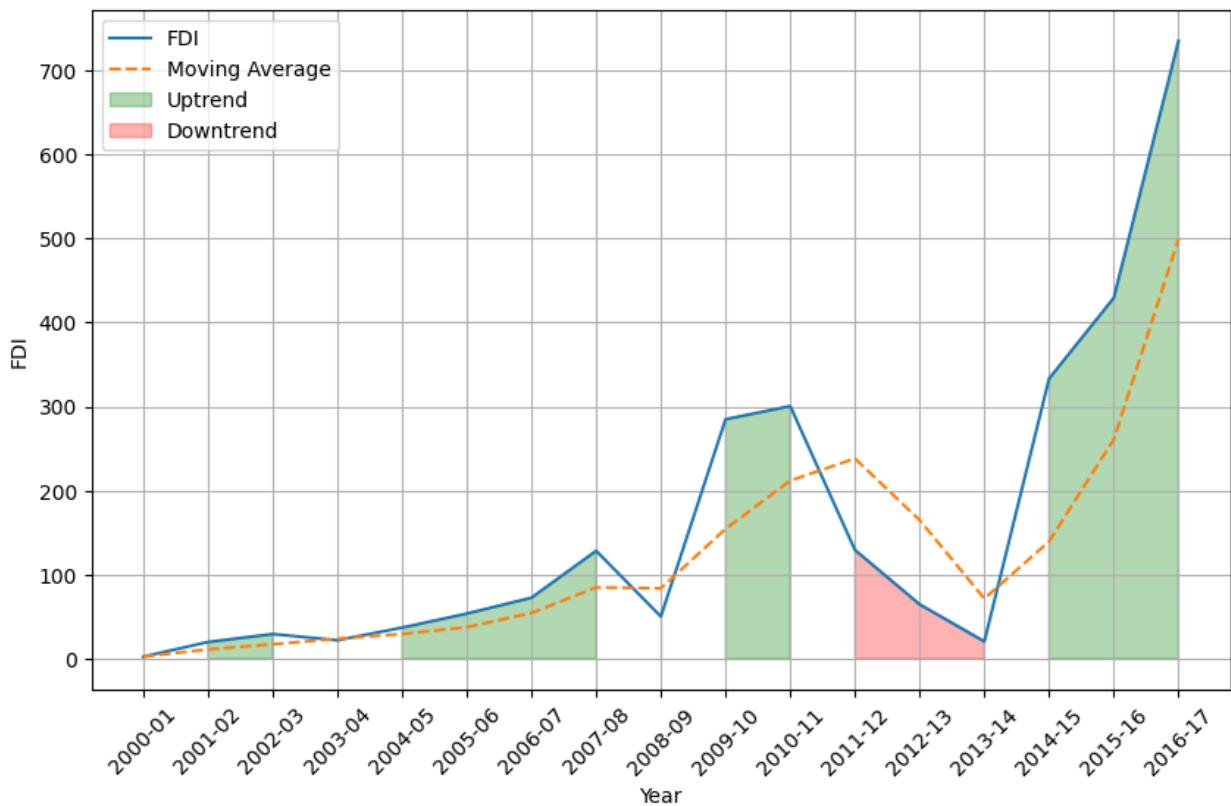
FDI Trends for AUTOMOBILE INDUSTRY



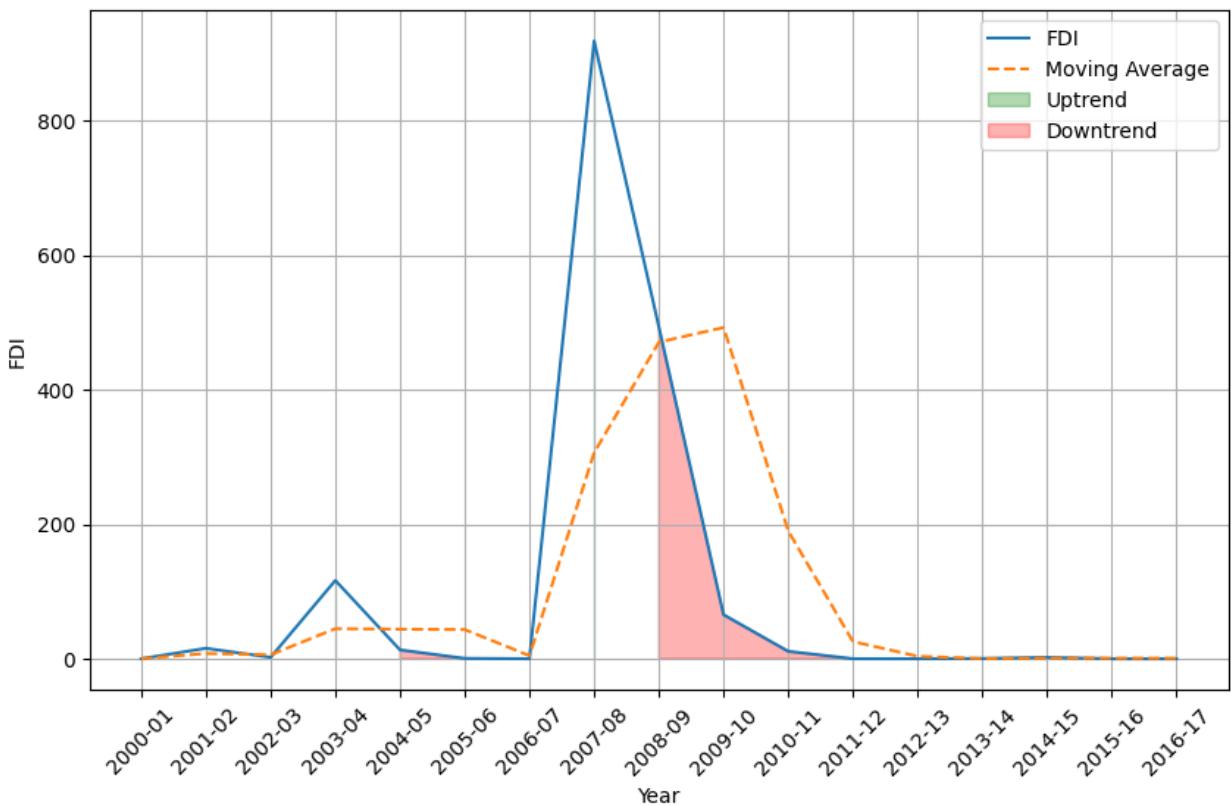
FDI Trends for AIR TRANSPORT (INCLUDING AIR FREIGHT)



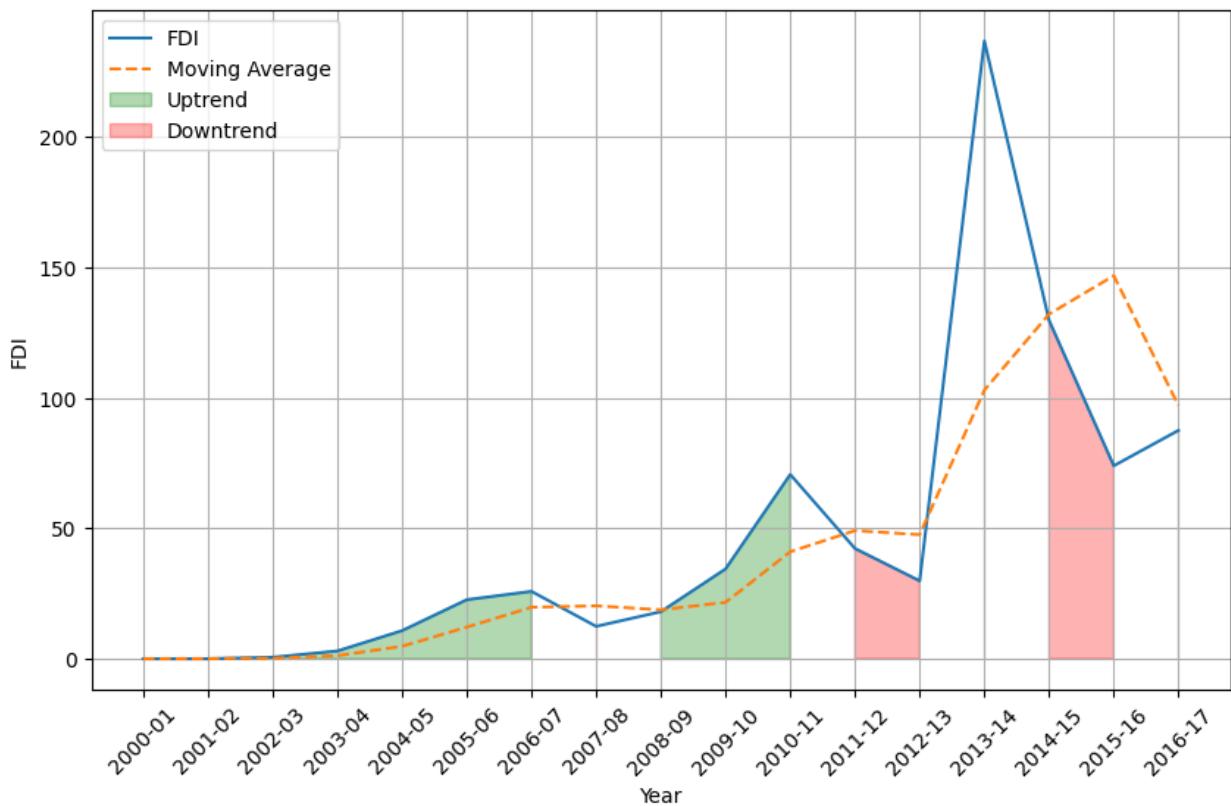
FDI Trends for SEA TRANSPORT



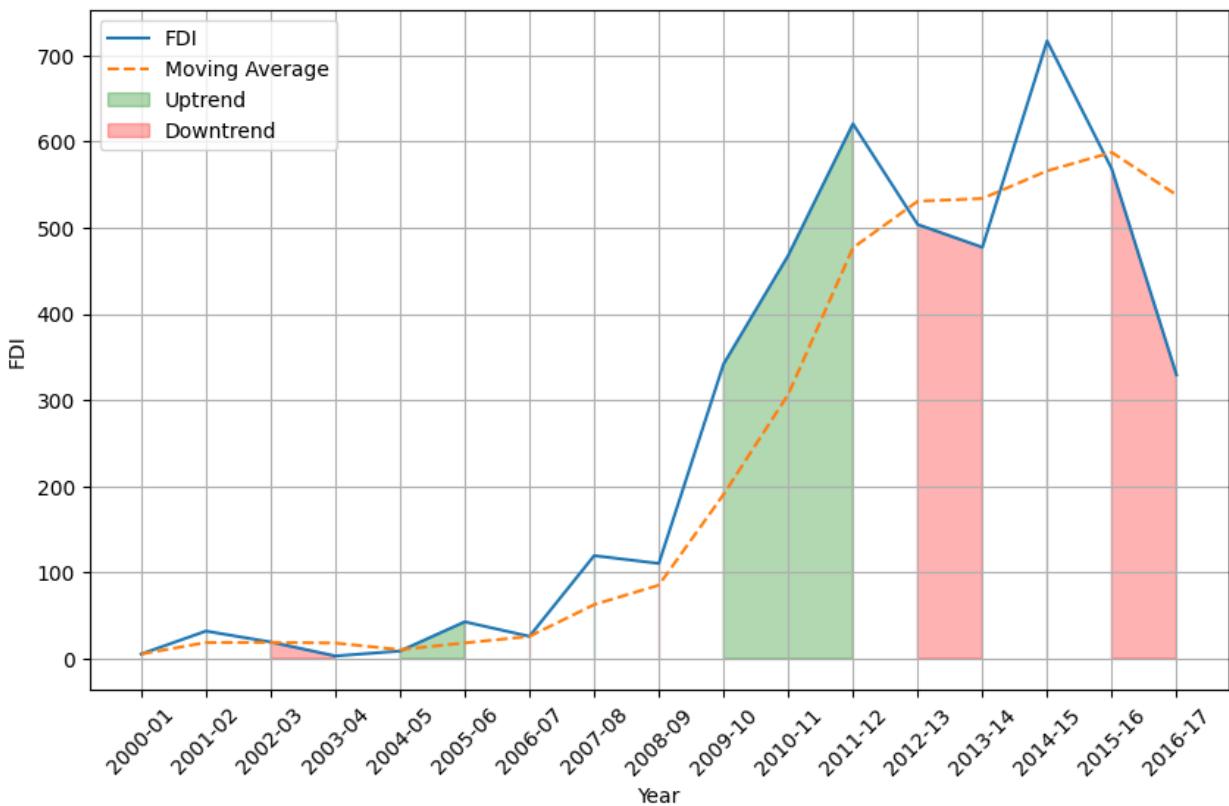
FDI Trends for PORTS

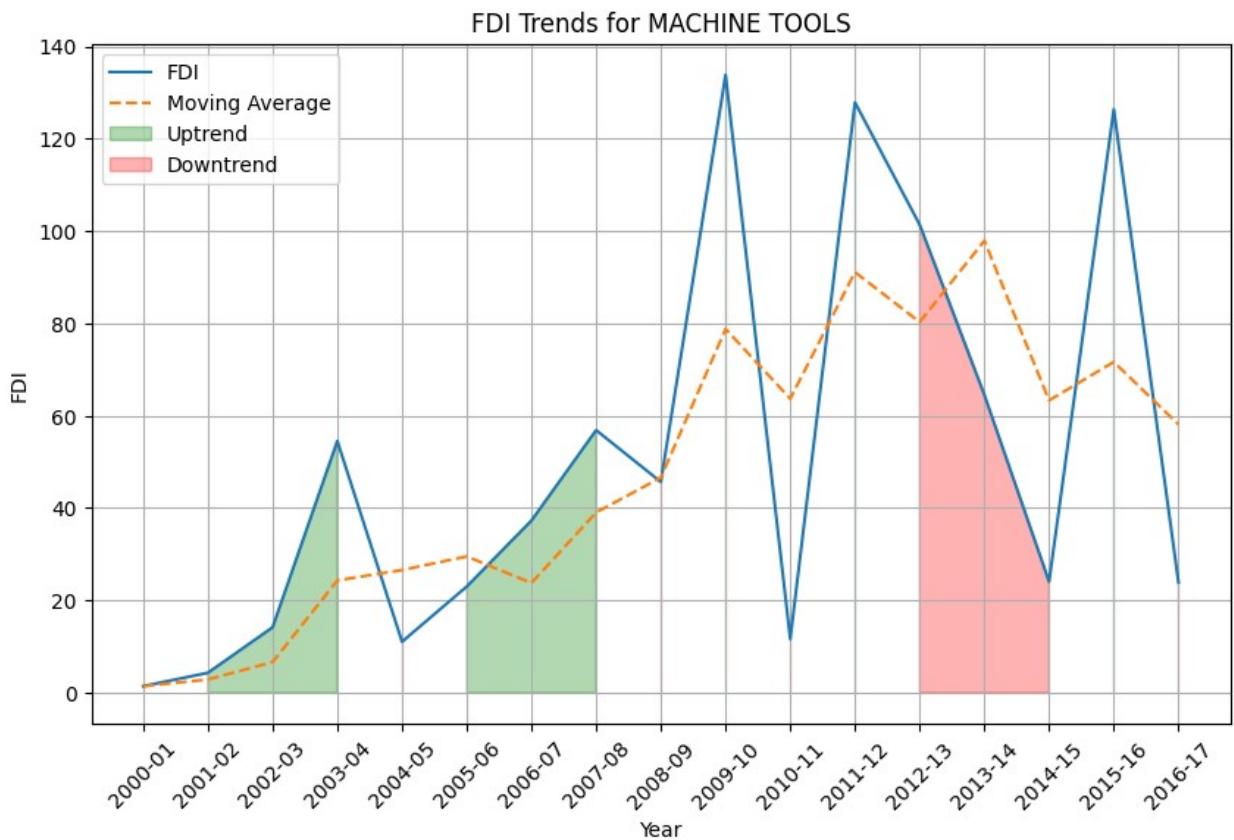


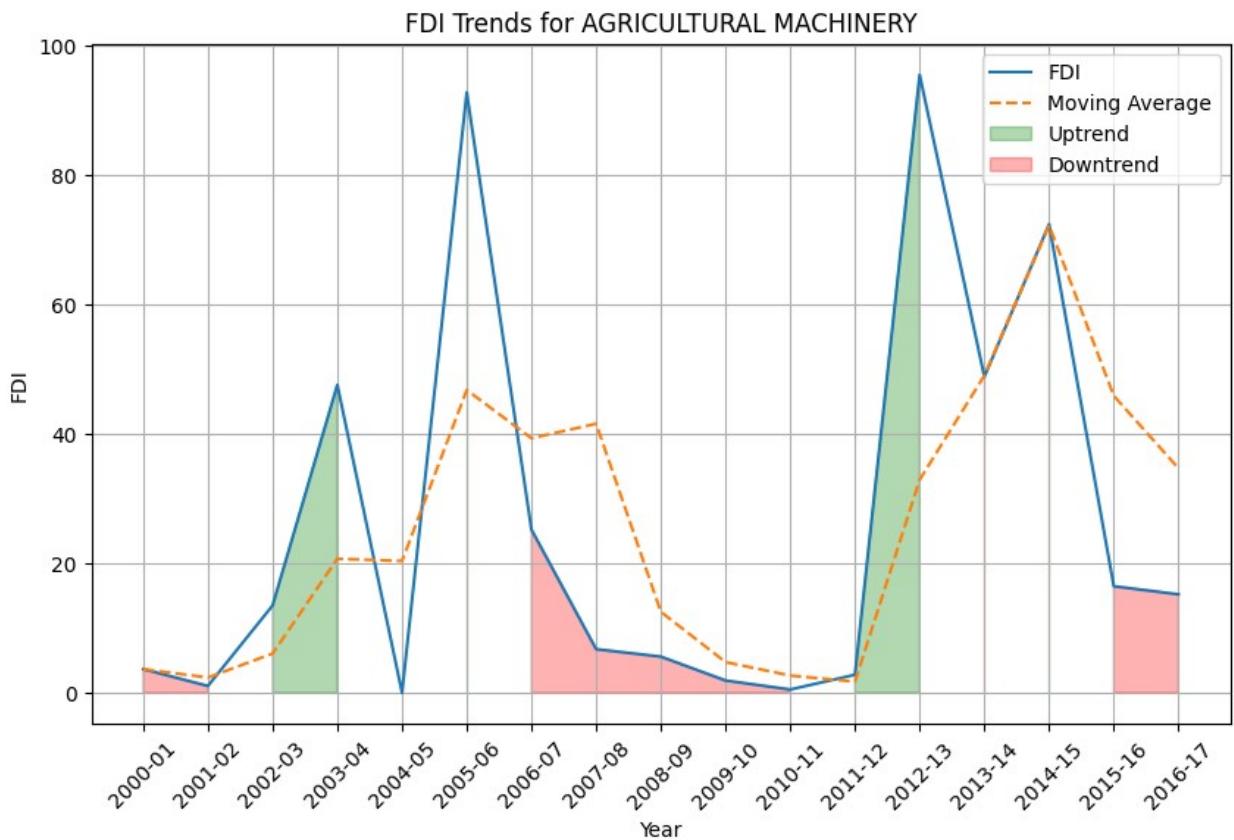
FDI Trends for RAILWAY RELATED COMPONENTS



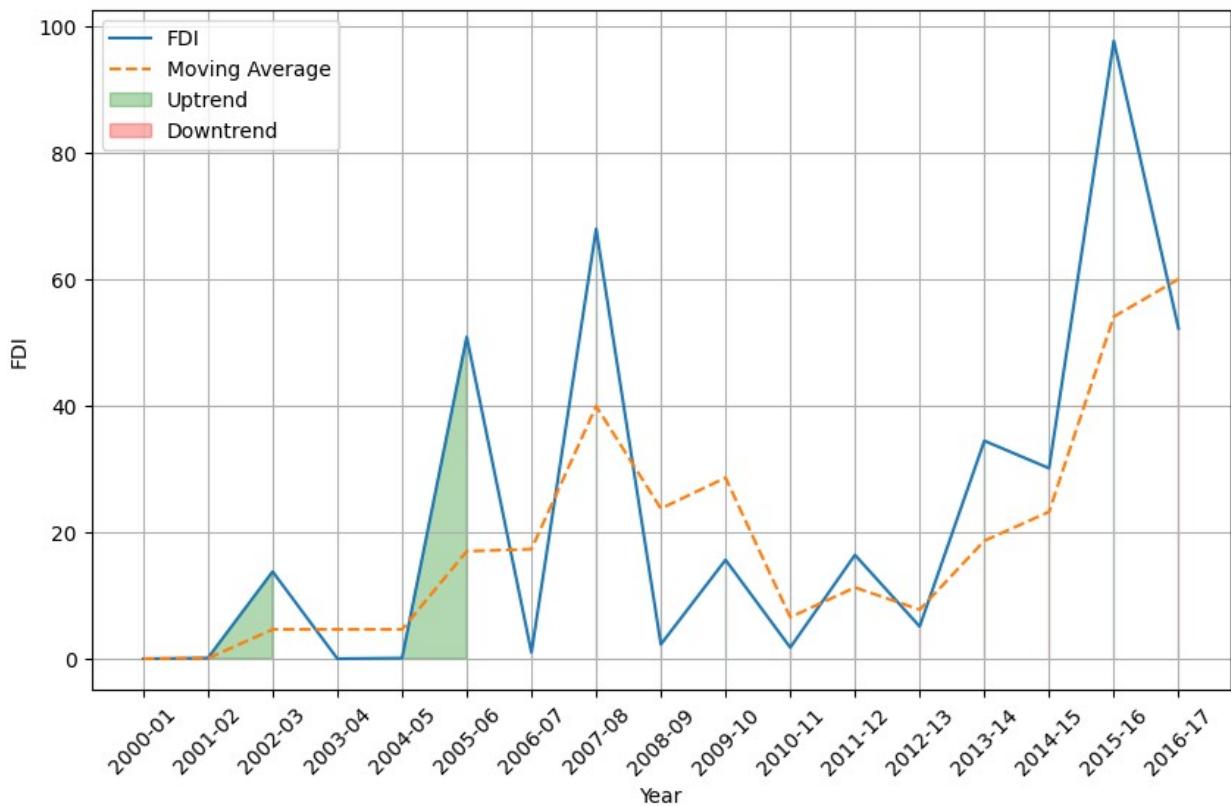
FDI Trends for INDUSTRIAL MACHINERY



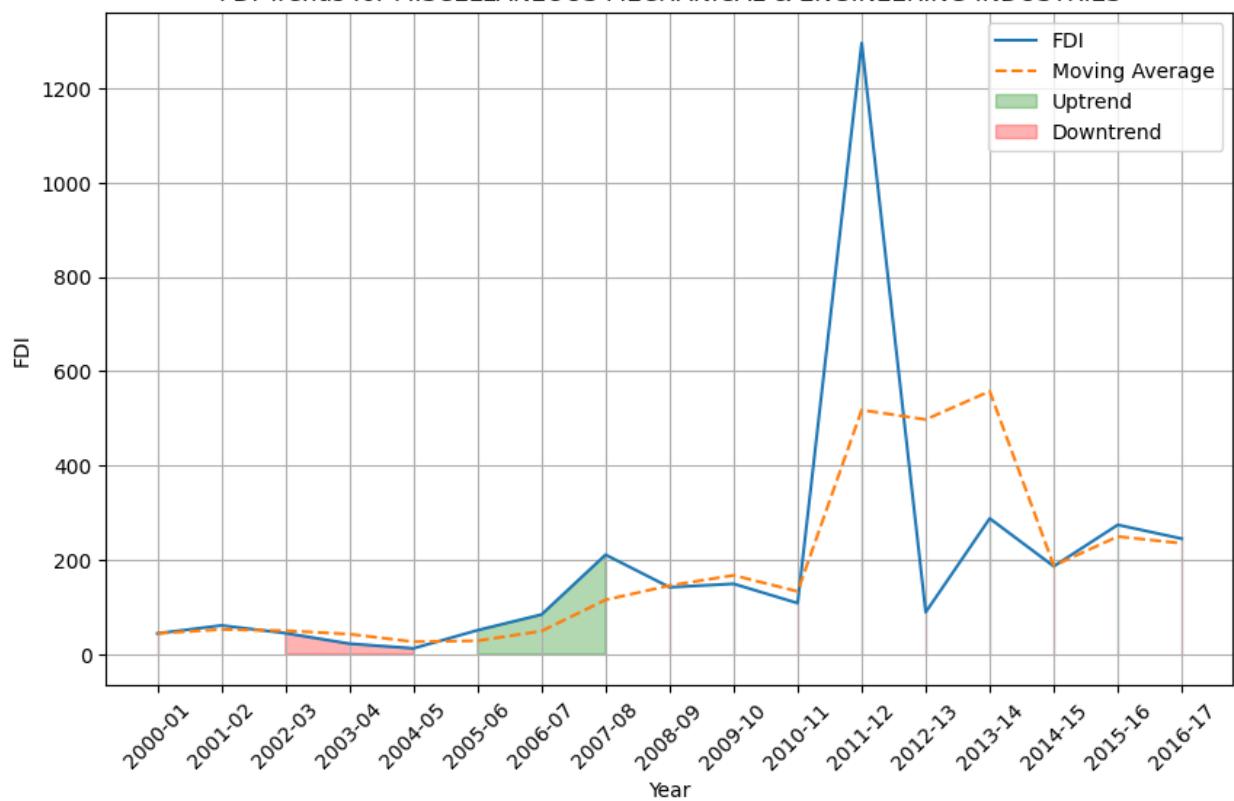




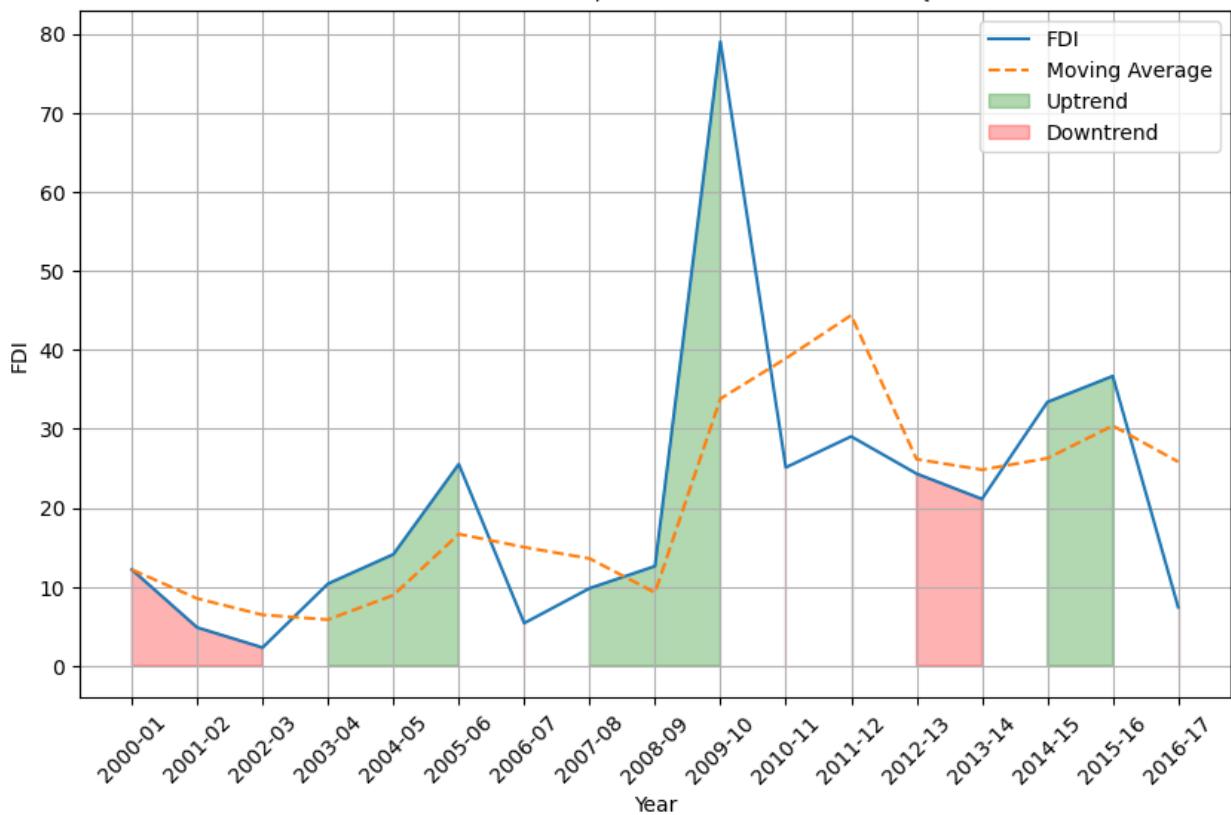
FDI Trends for EARTH-MOVING MACHINERY

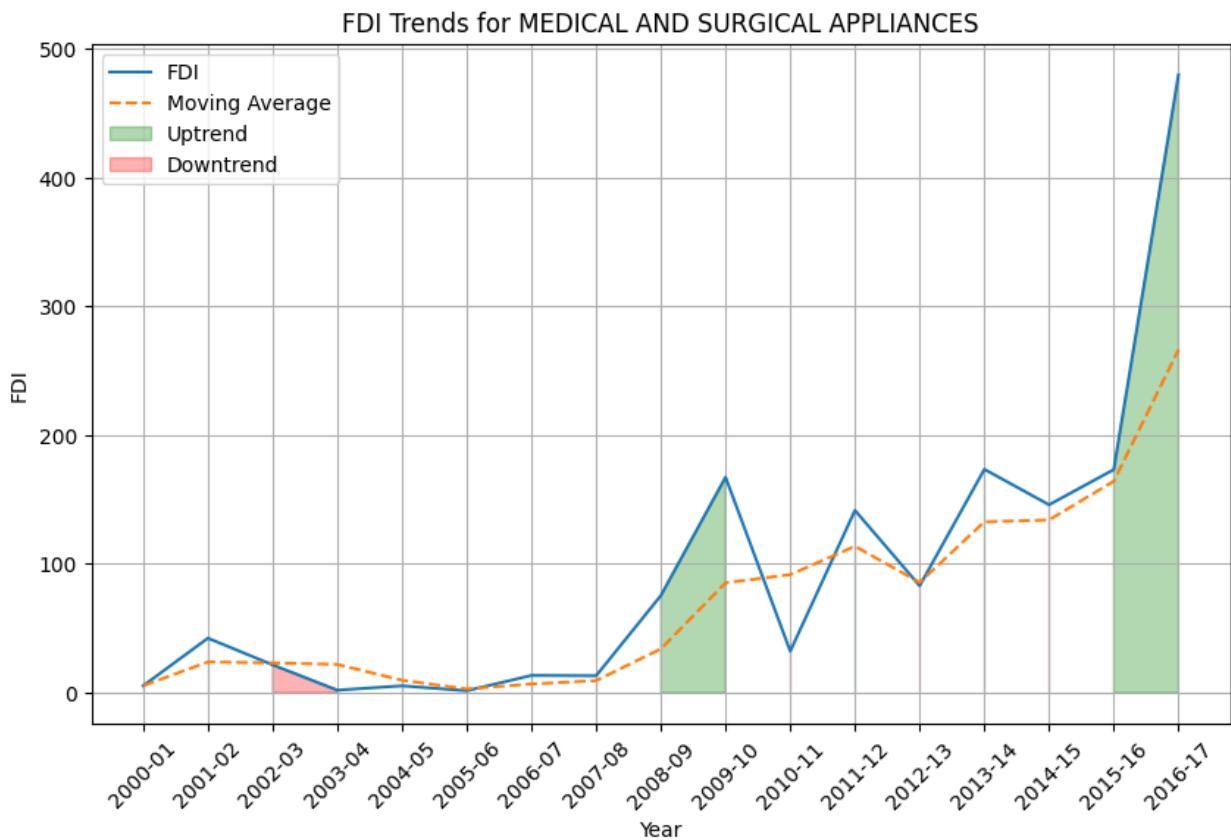


FDI Trends for MISCELLANEOUS MECHANICAL & ENGINEERING INDUSTRIES

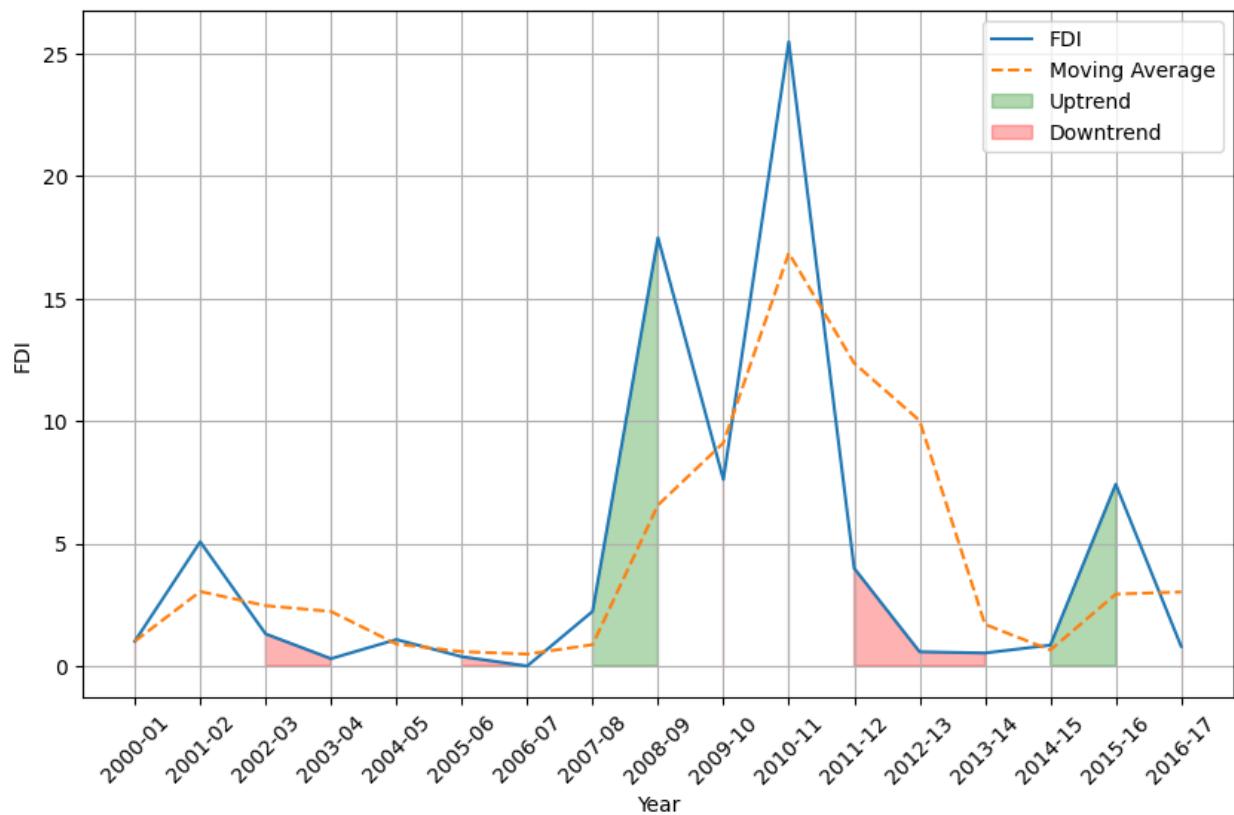


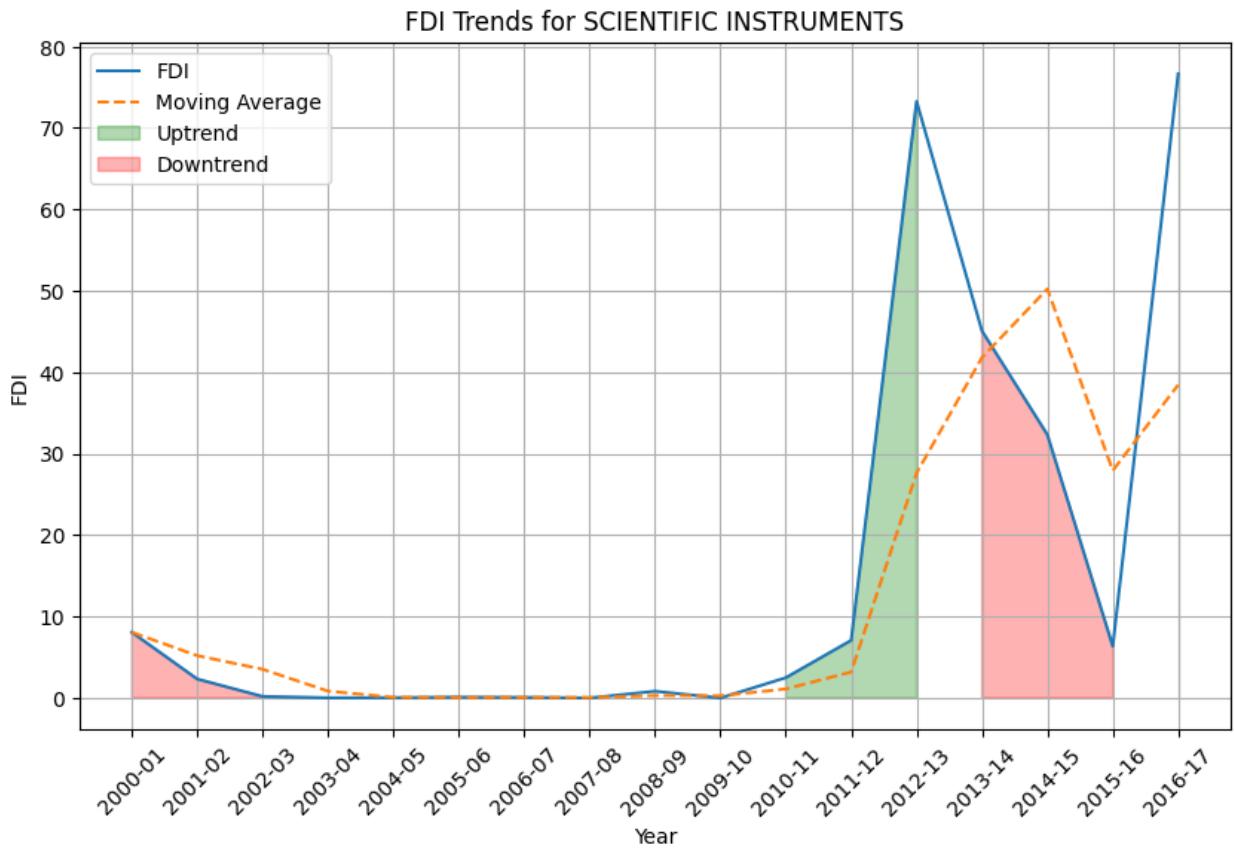
FDI Trends for COMMERCIAL, OFFICE & HOUSEHOLD EQUIPMENTS



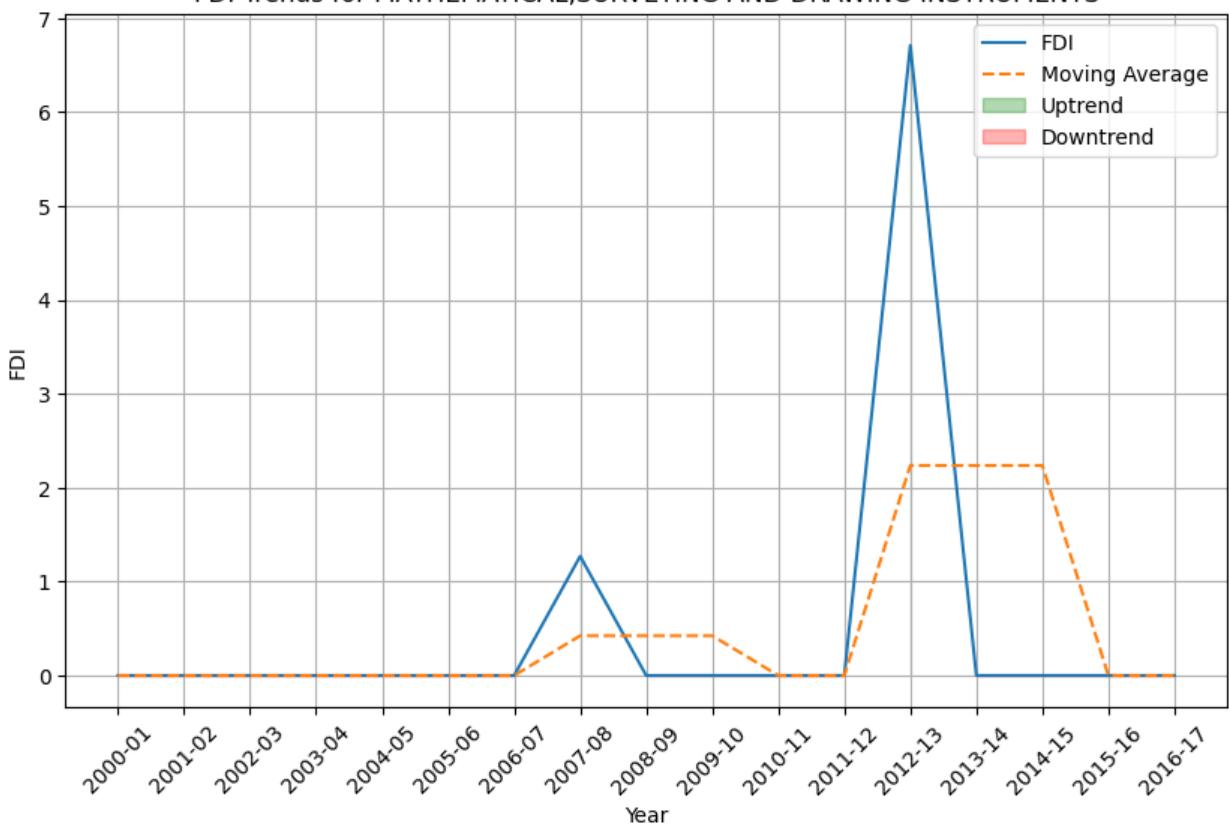


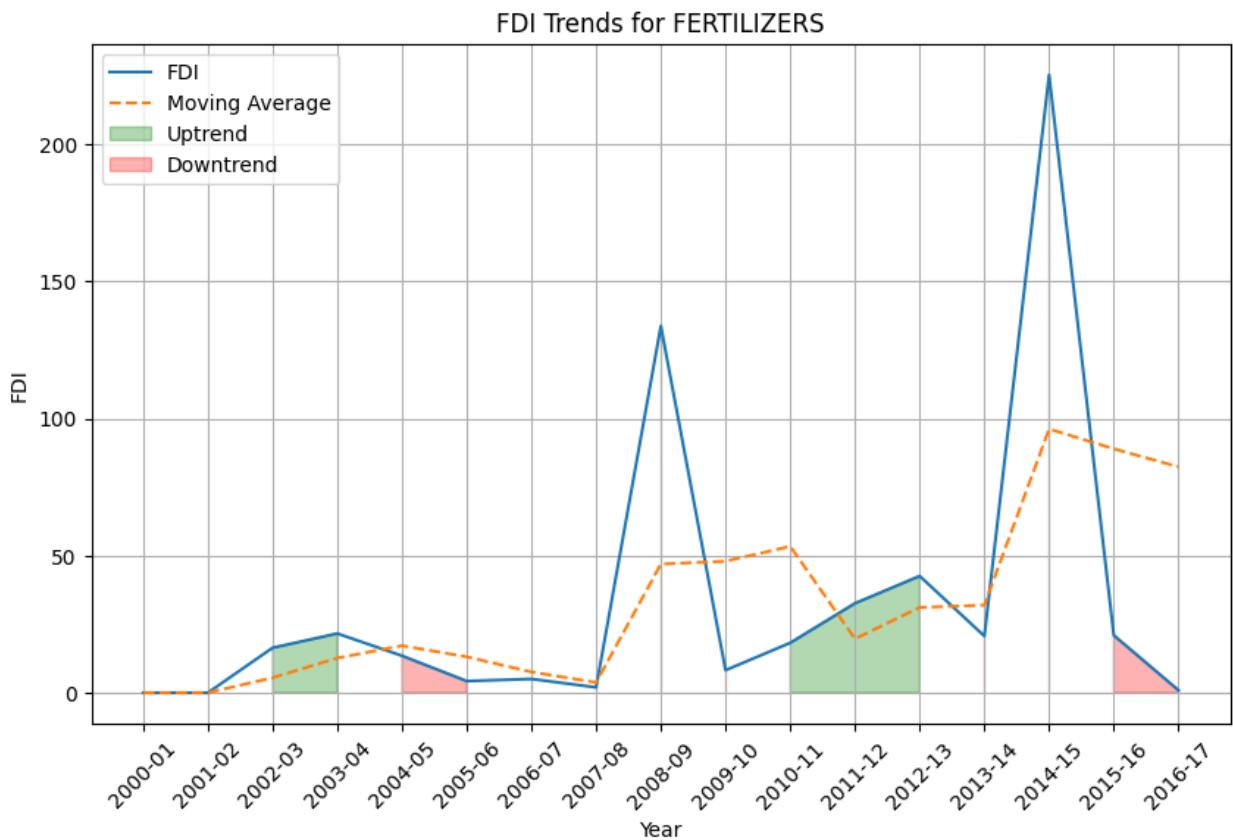
FDI Trends for INDUSTRIAL INSTRUMENTS



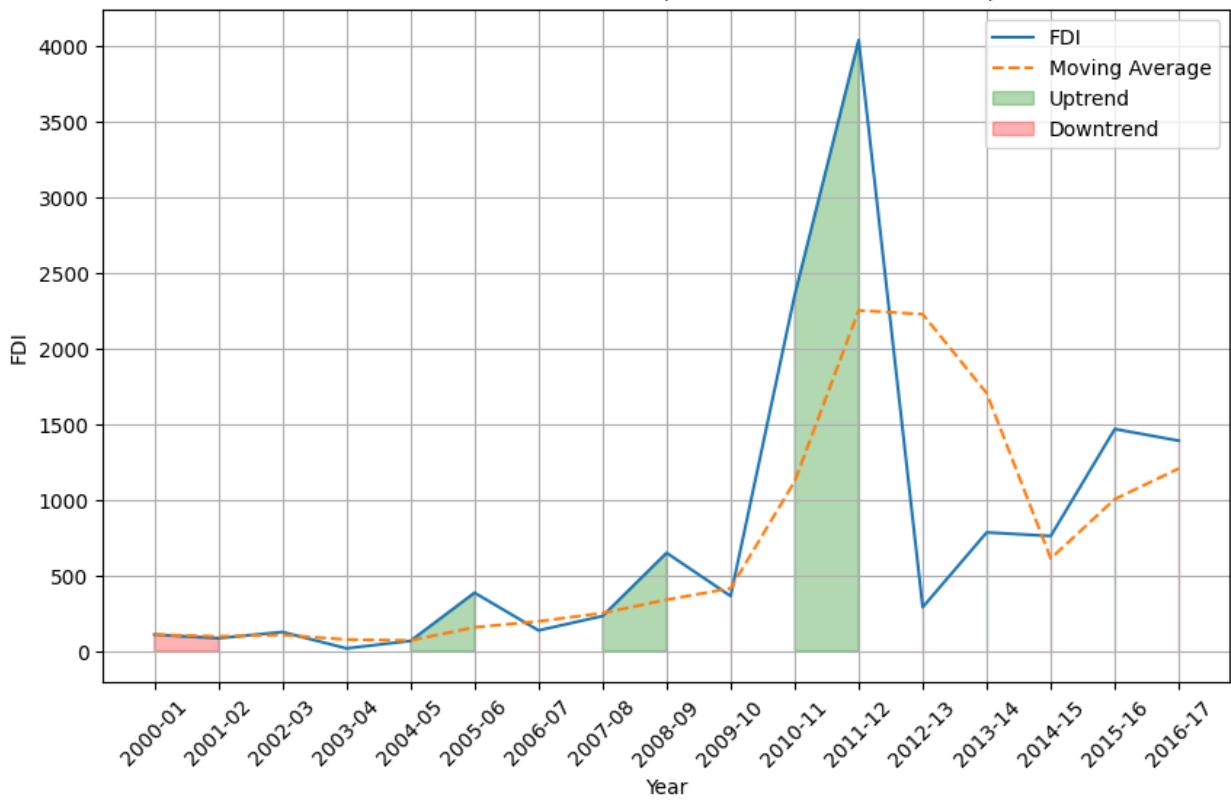


FDI Trends for MATHEMATICAL,SURVEYING AND DRAWING INSTRUMENTS

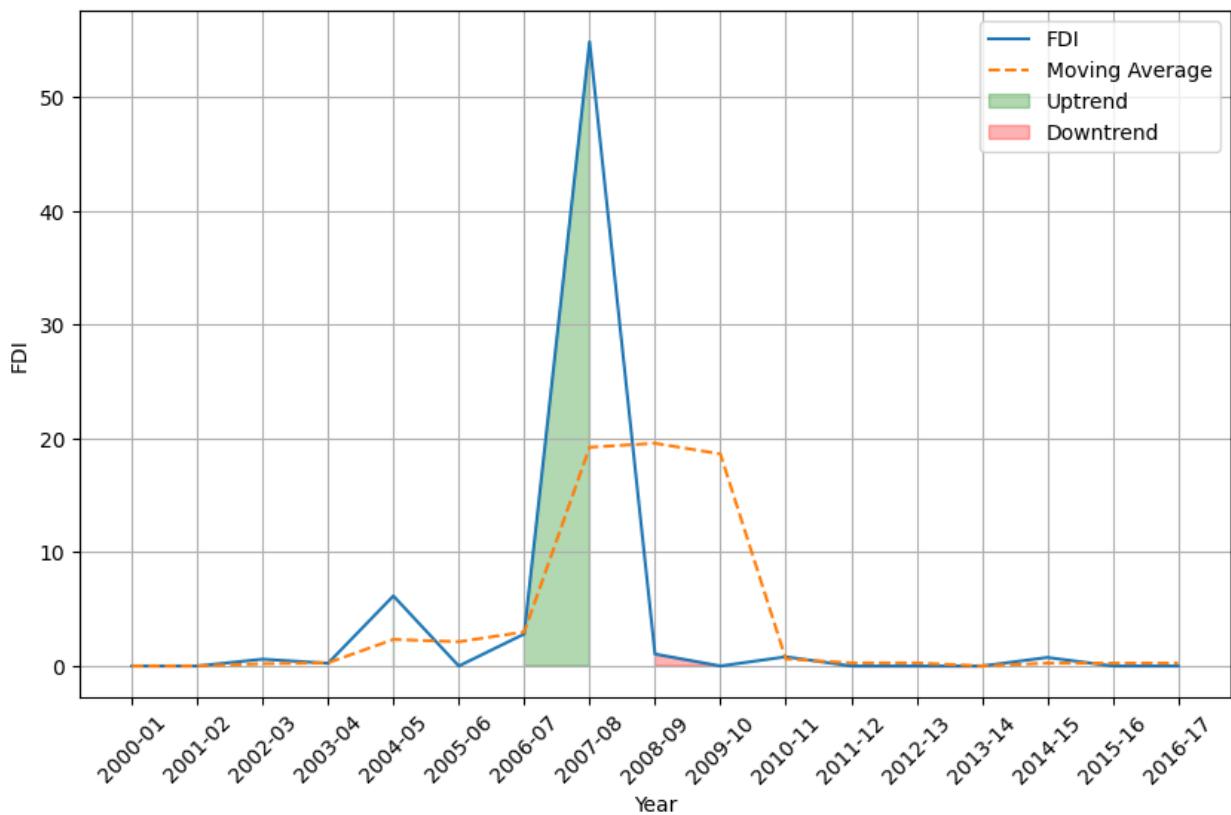




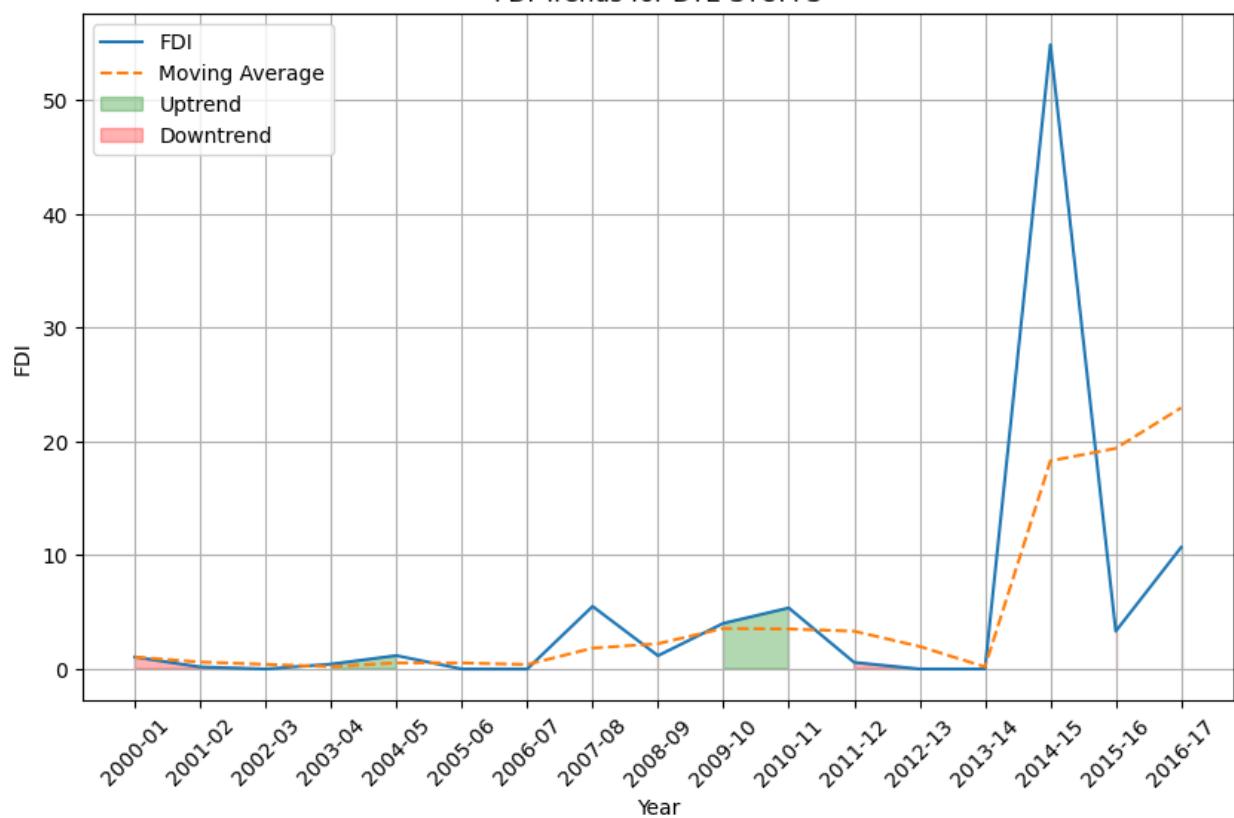
FDI Trends for CHEMICALS (OTHER THAN FERTILIZERS)



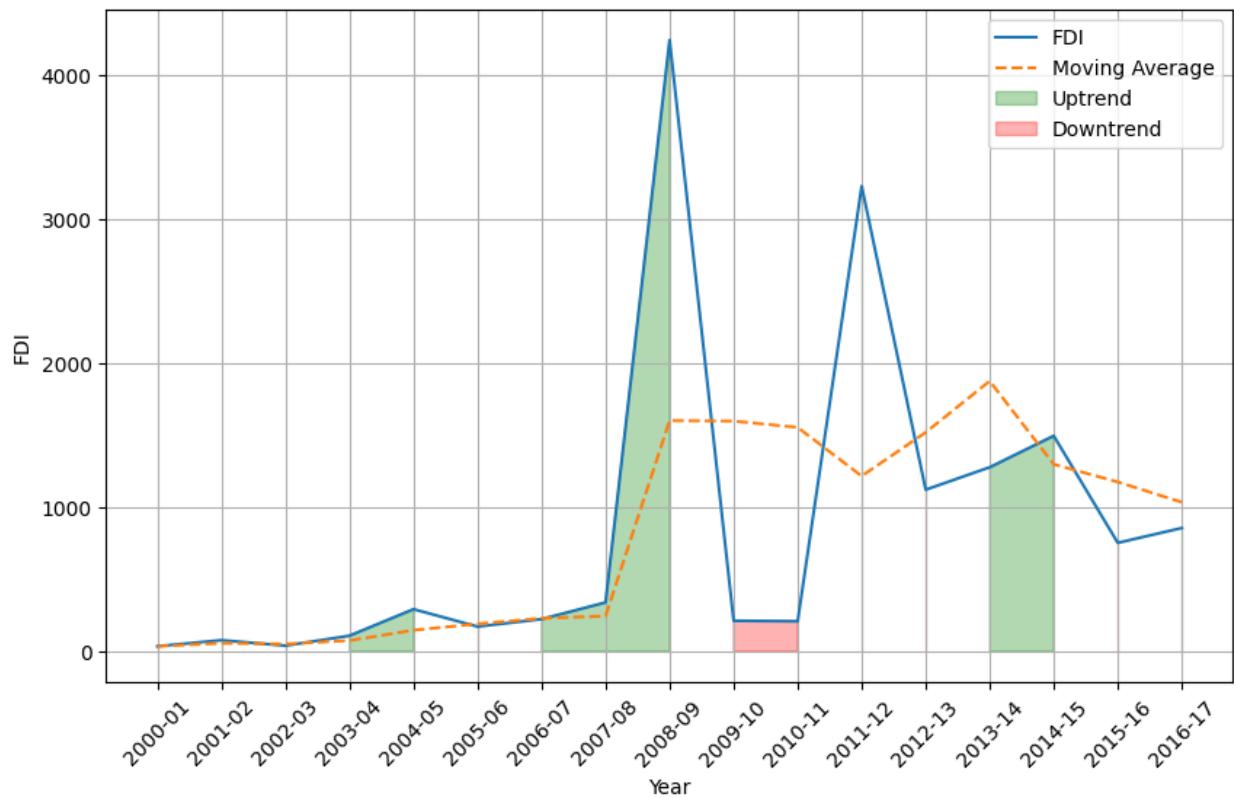
FDI Trends for PHOTOGRAPHIC RAW FILM AND PAPER



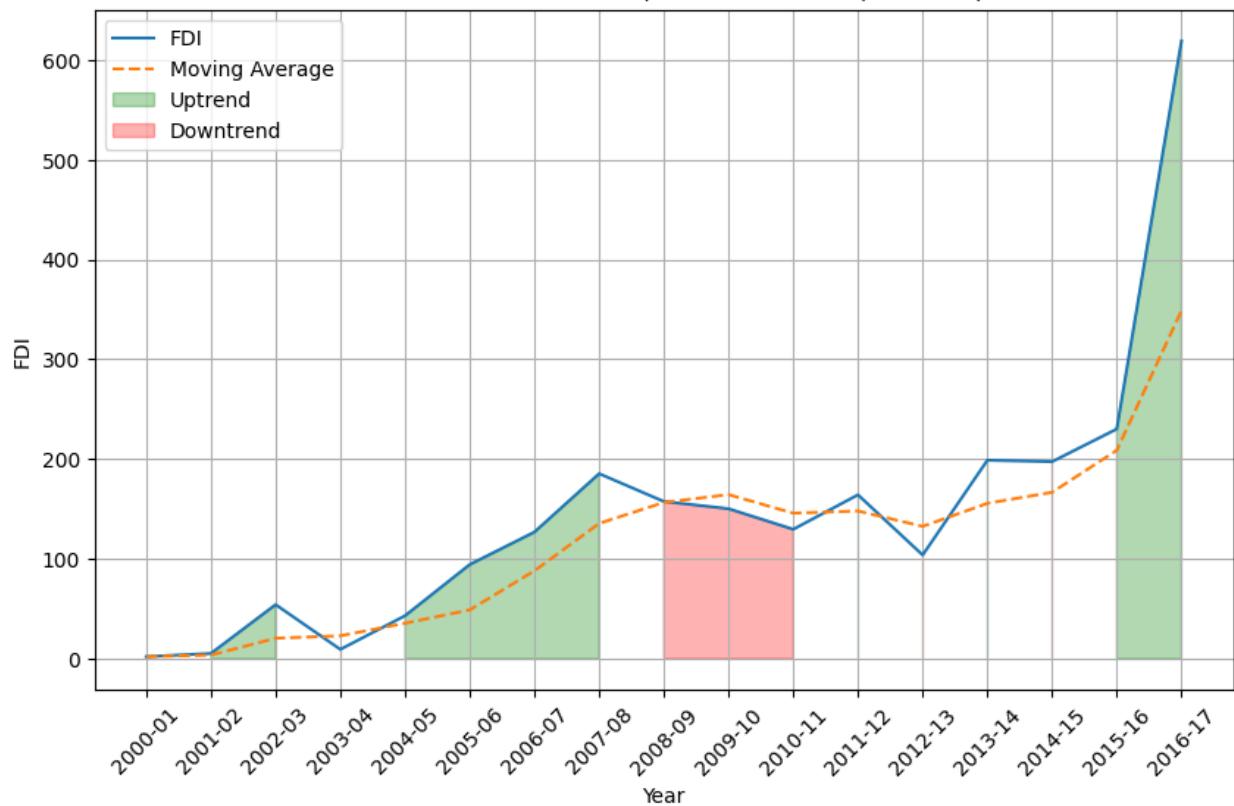
FDI Trends for DYE-STUFFS



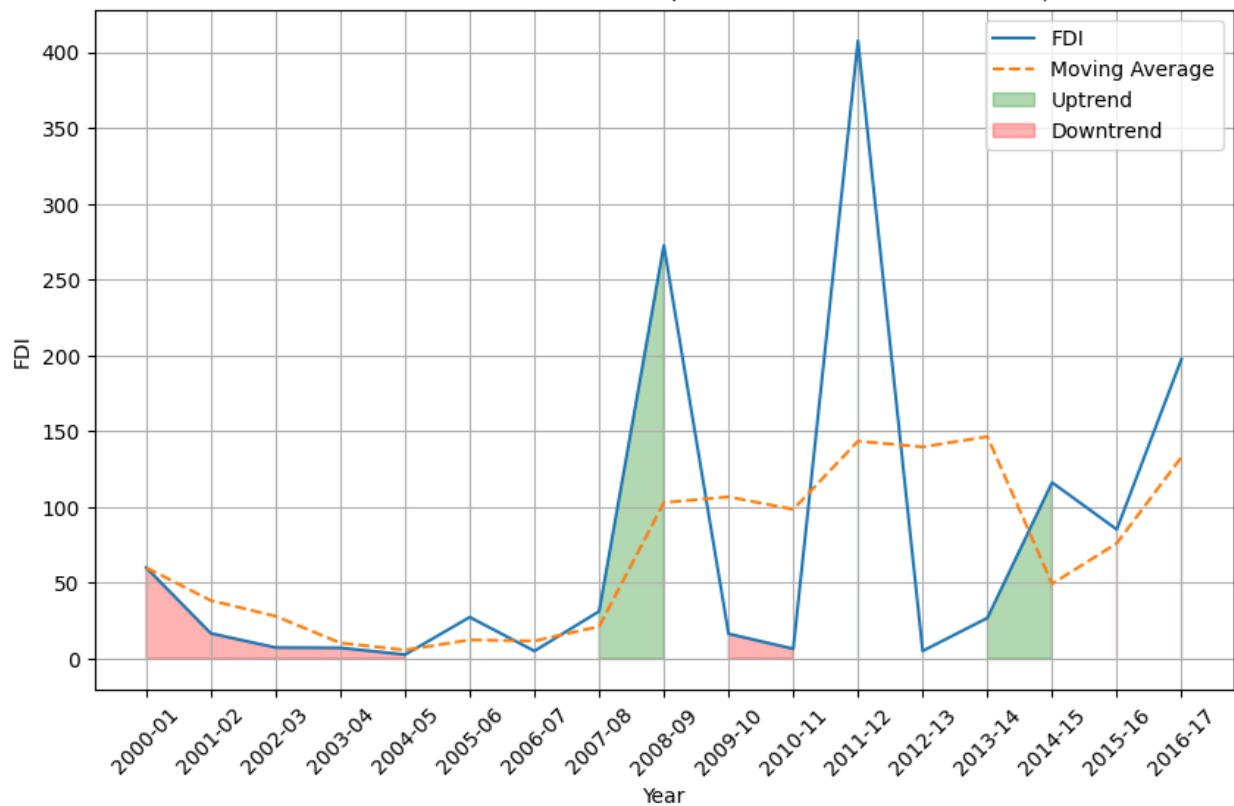
FDI Trends for DRUGS & PHARMACEUTICALS



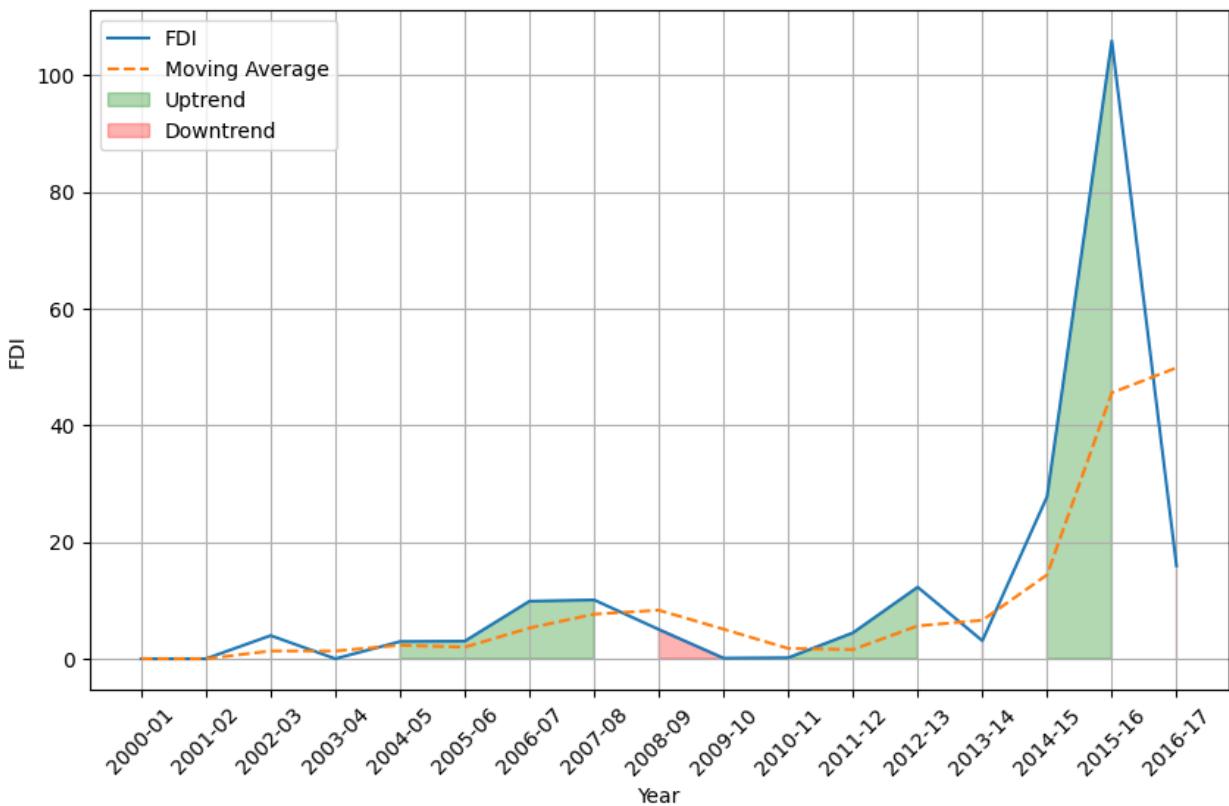
FDI Trends for TEXTILES (INCLUDING DYED,PRINTED)



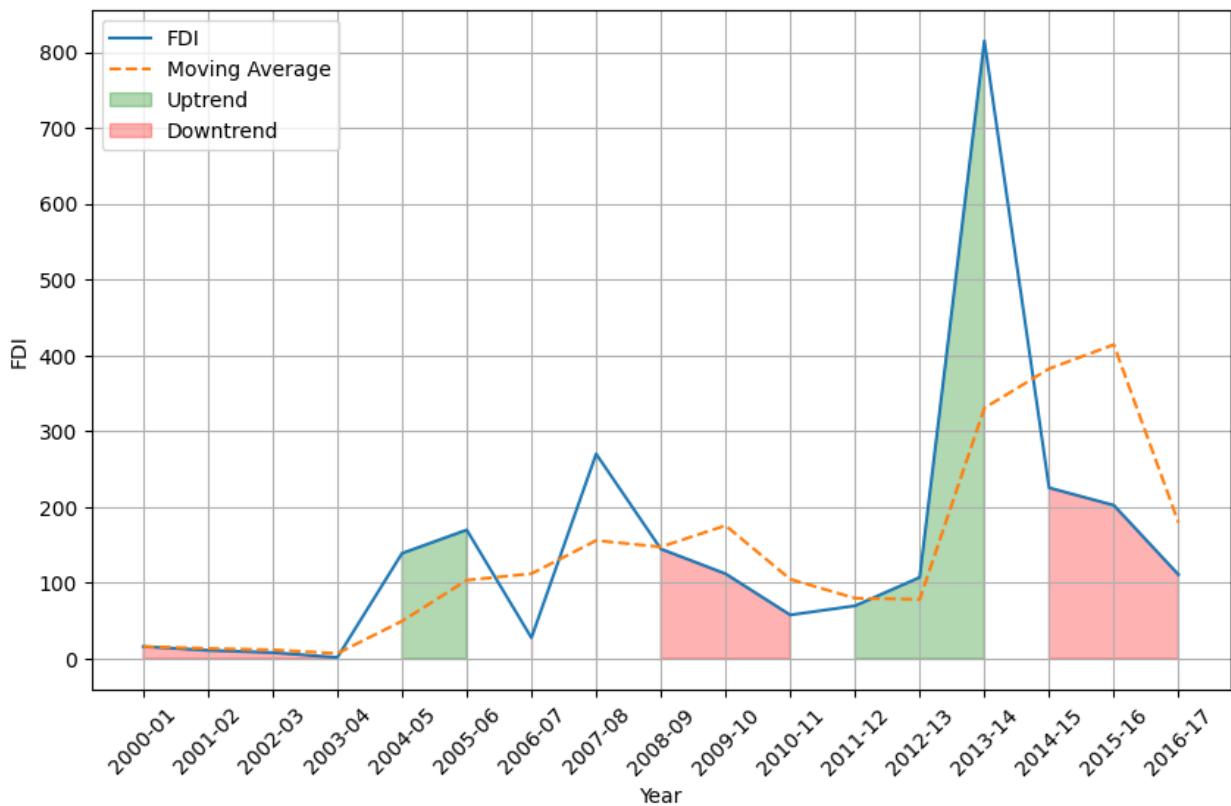
FDI Trends for PAPER AND PULP (INCLUDING PAPER PRODUCTS)



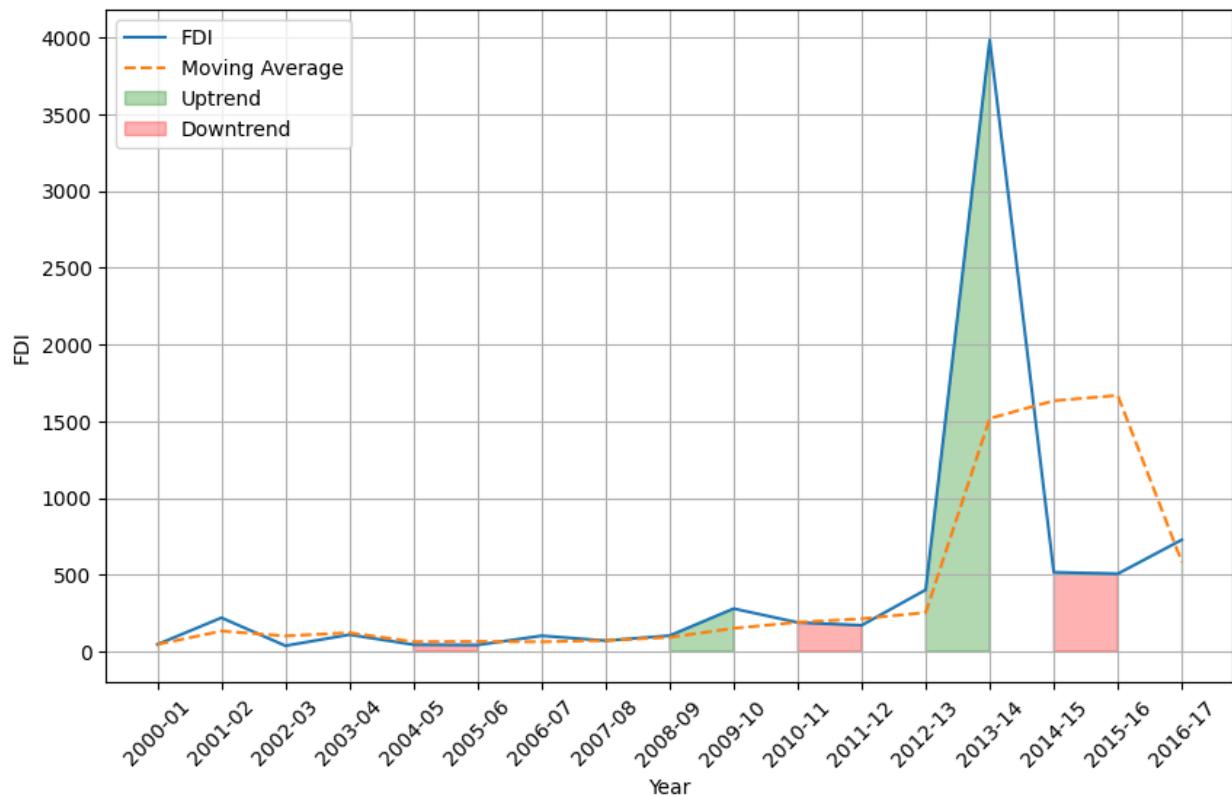
FDI Trends for SUGAR



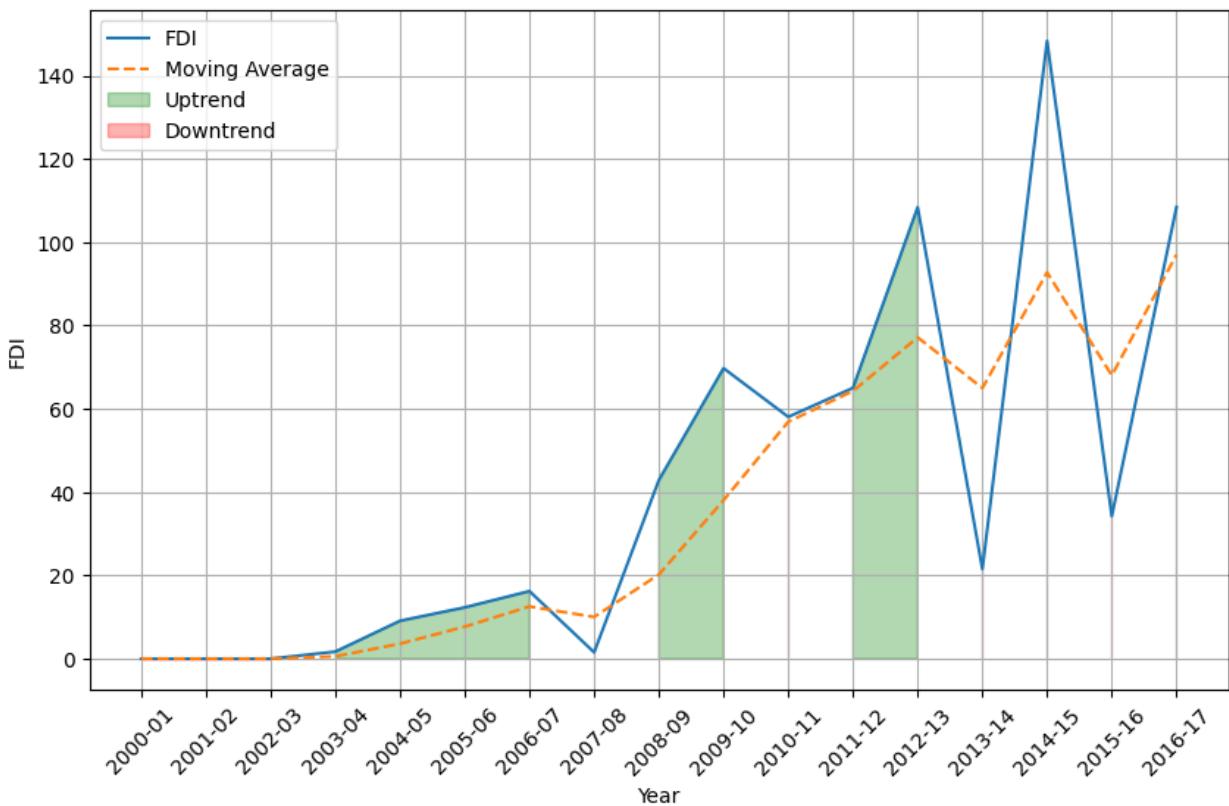
FDI Trends for FERMENTATION INDUSTRIES



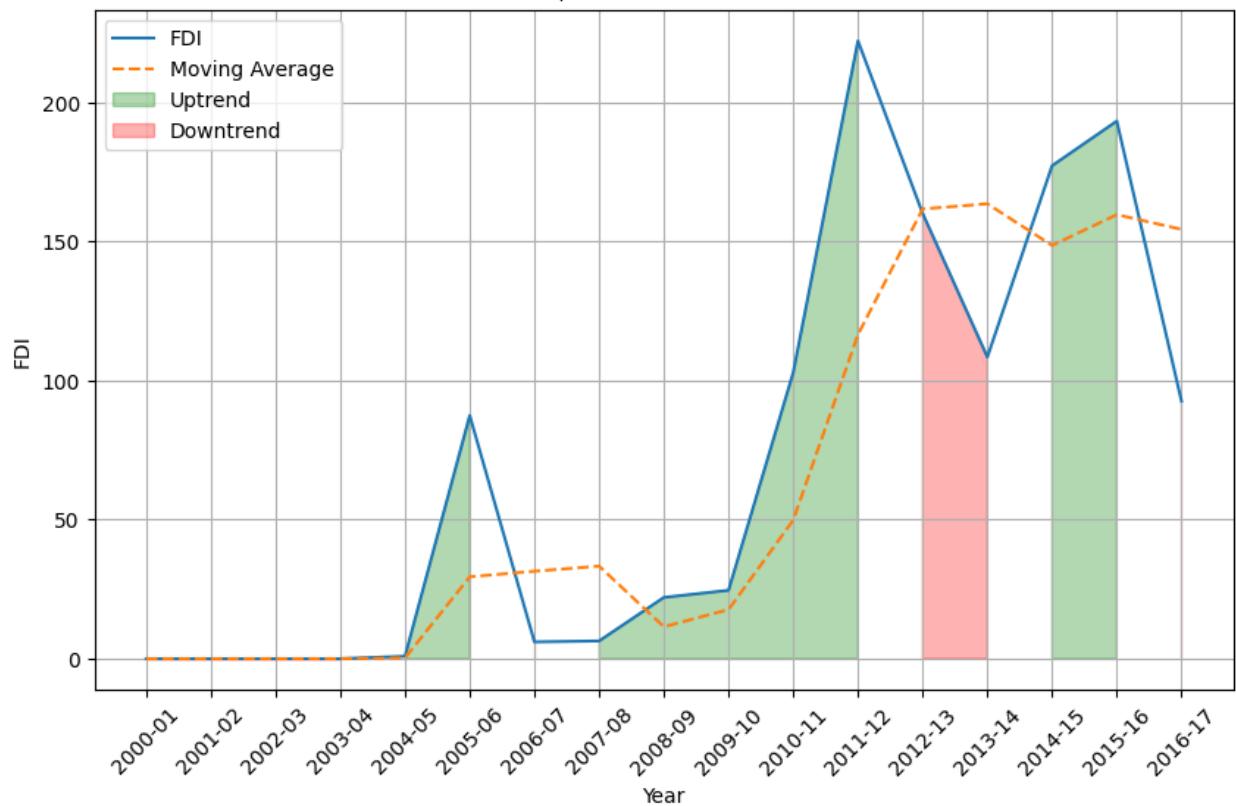
FDI Trends for FOOD PROCESSING INDUSTRIES



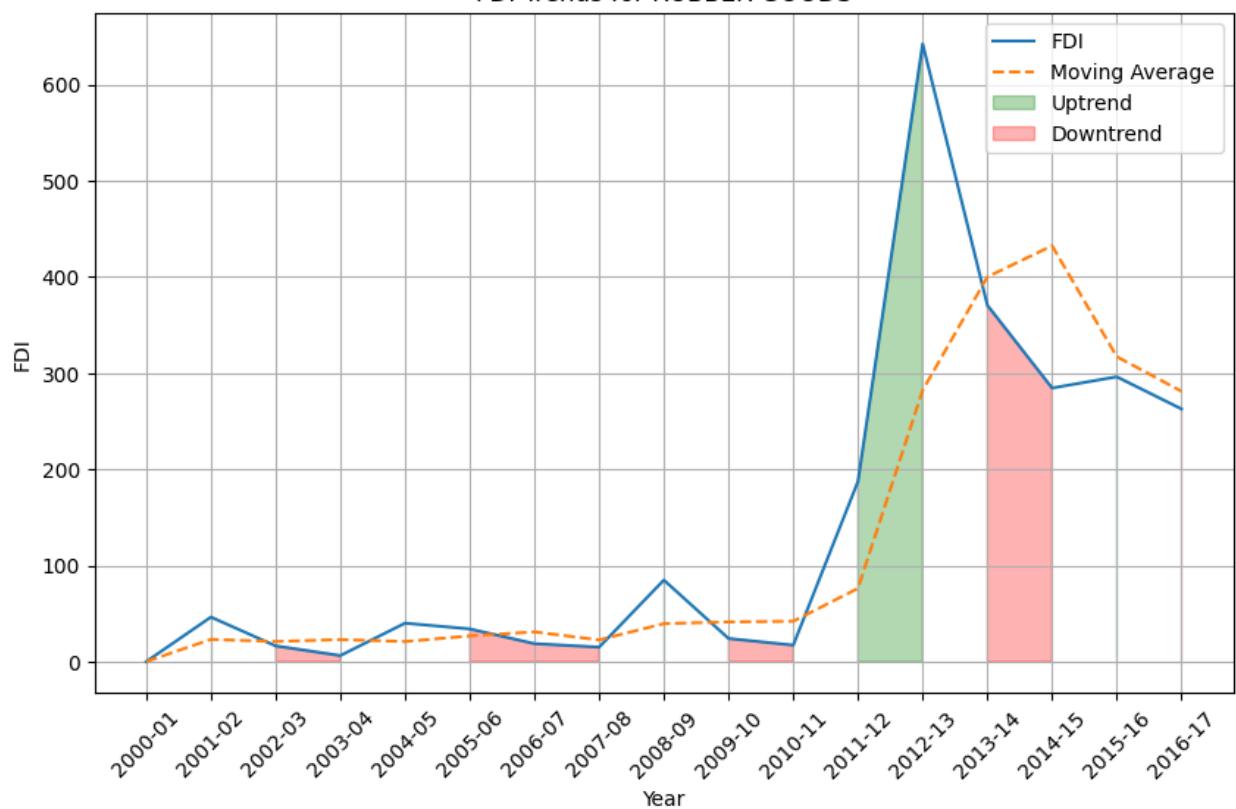
FDI Trends for VEGETABLE OILS AND VANASPATI



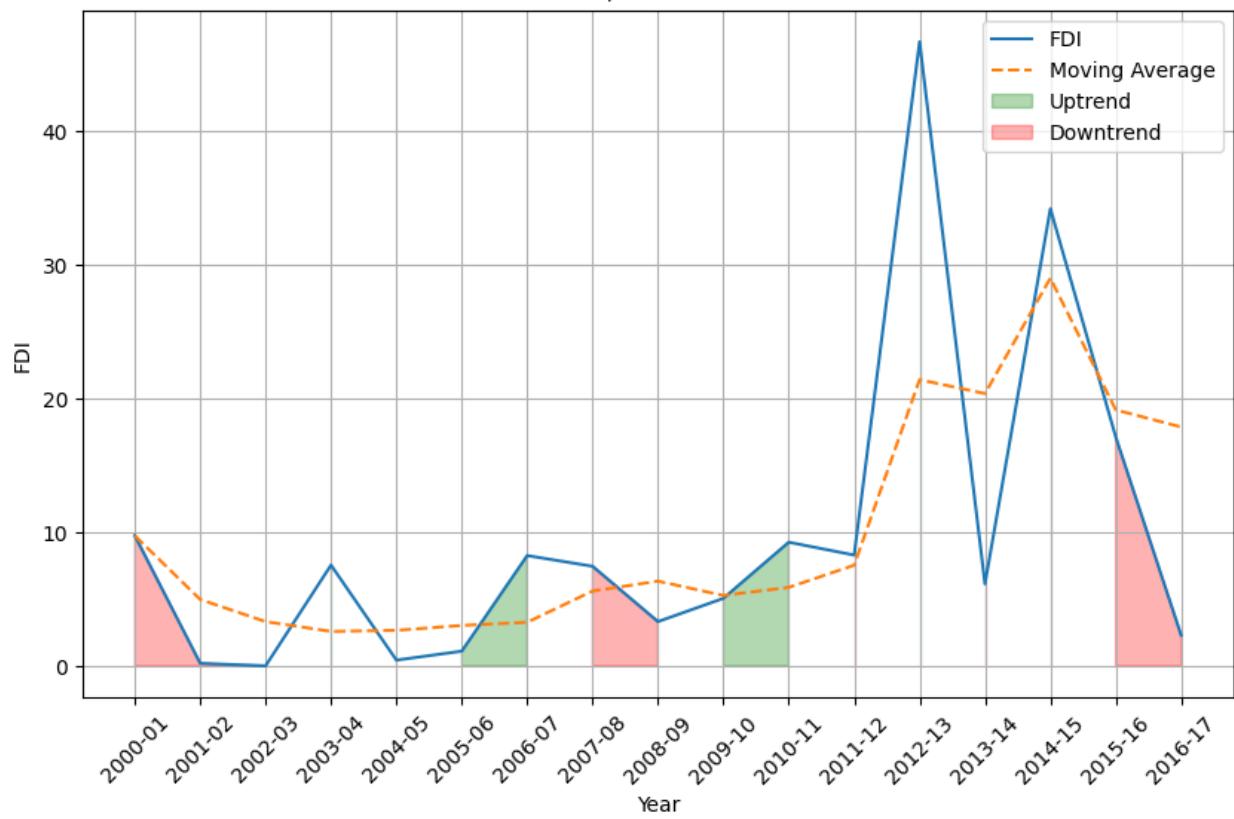
FDI Trends for SOAPS, COSMETICS & TOILET PREPARATIONS



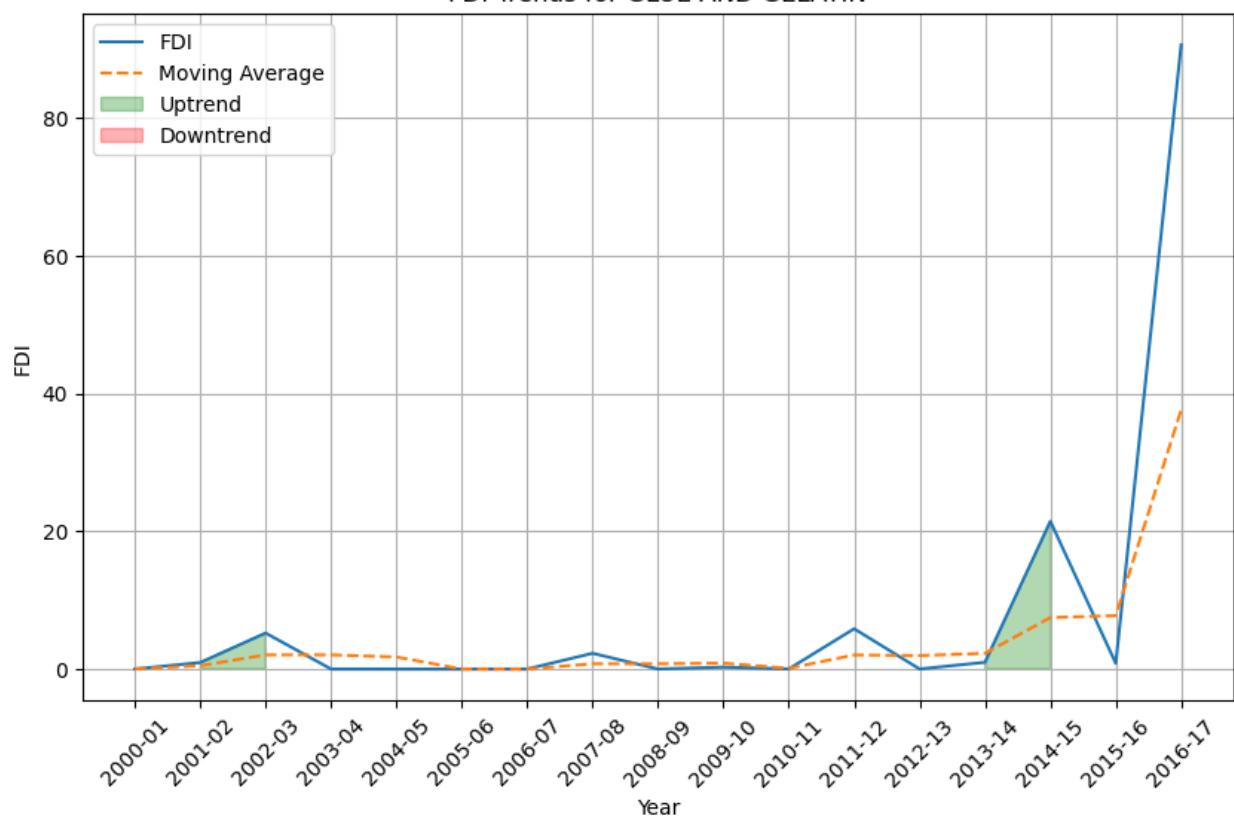
FDI Trends for RUBBER GOODS

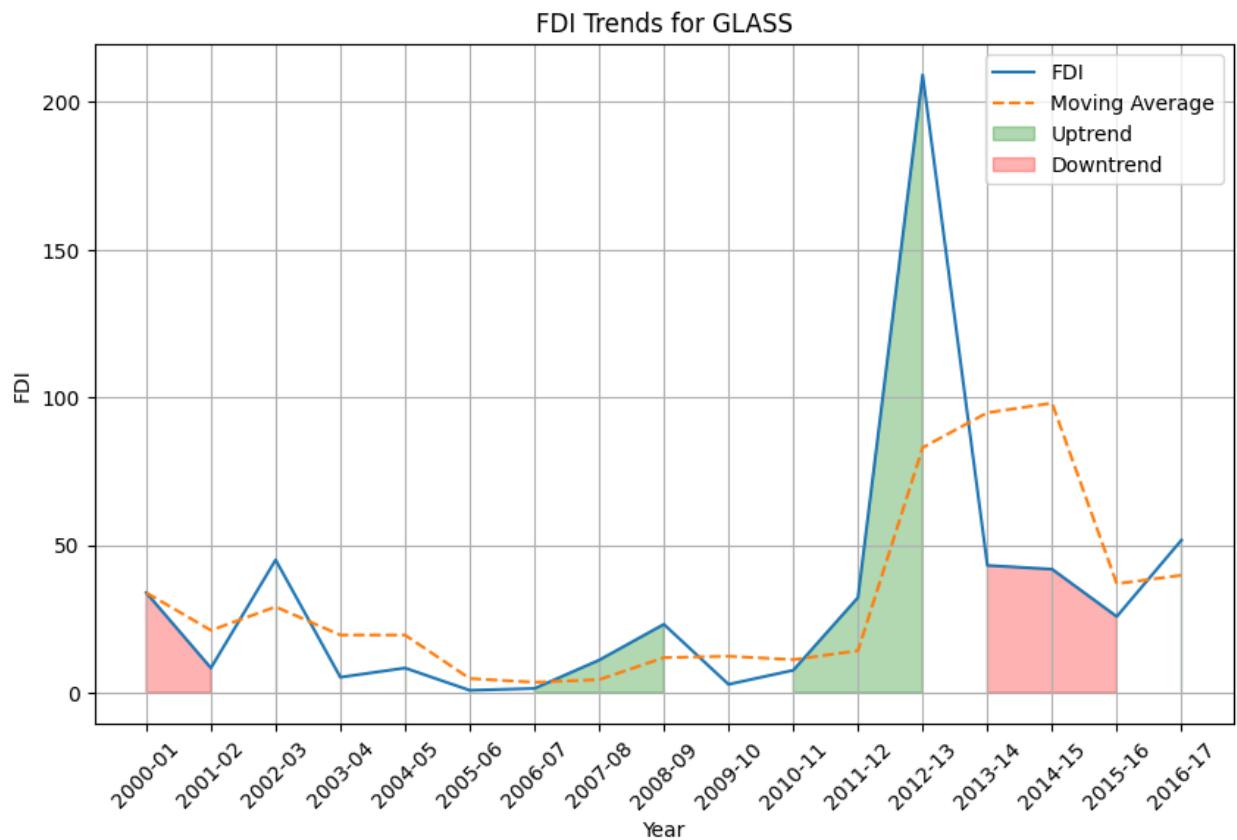


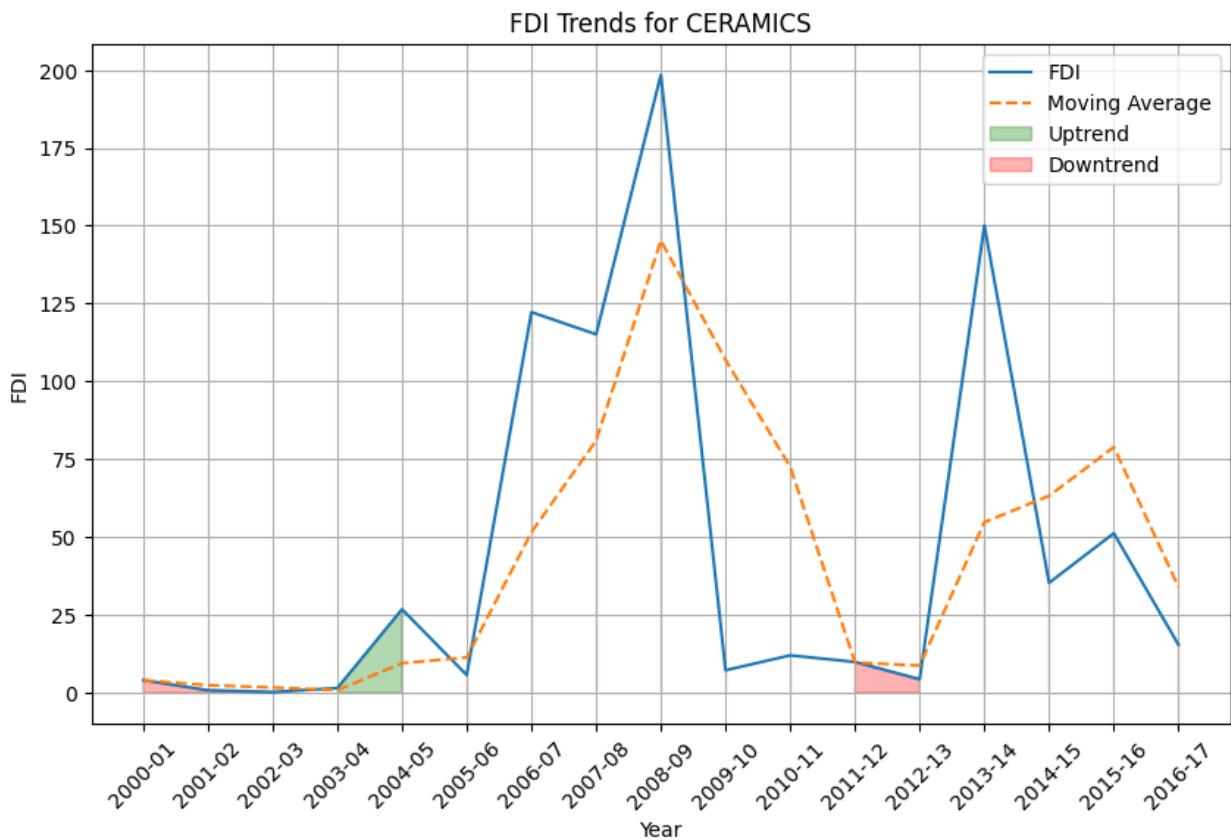
FDI Trends for LEATHER,LEATHER GOODS AND PICKERS



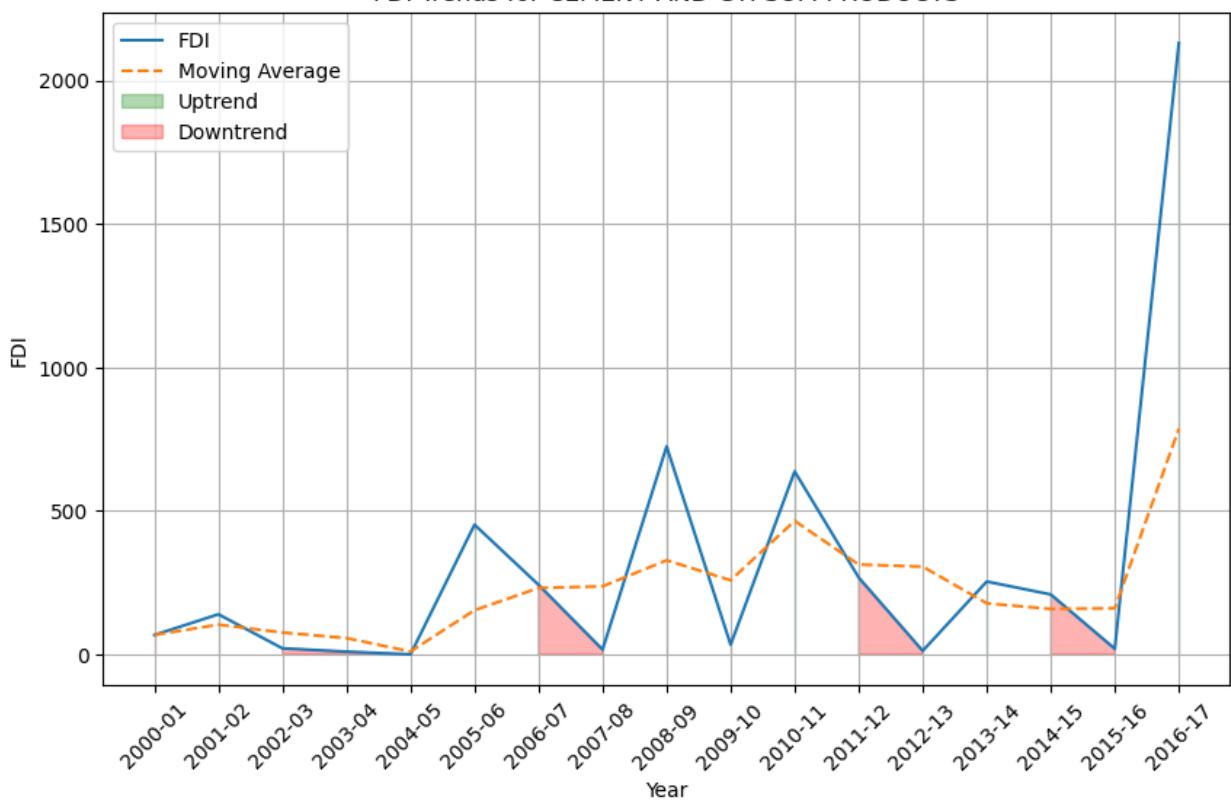
FDI Trends for GLUE AND GELATIN



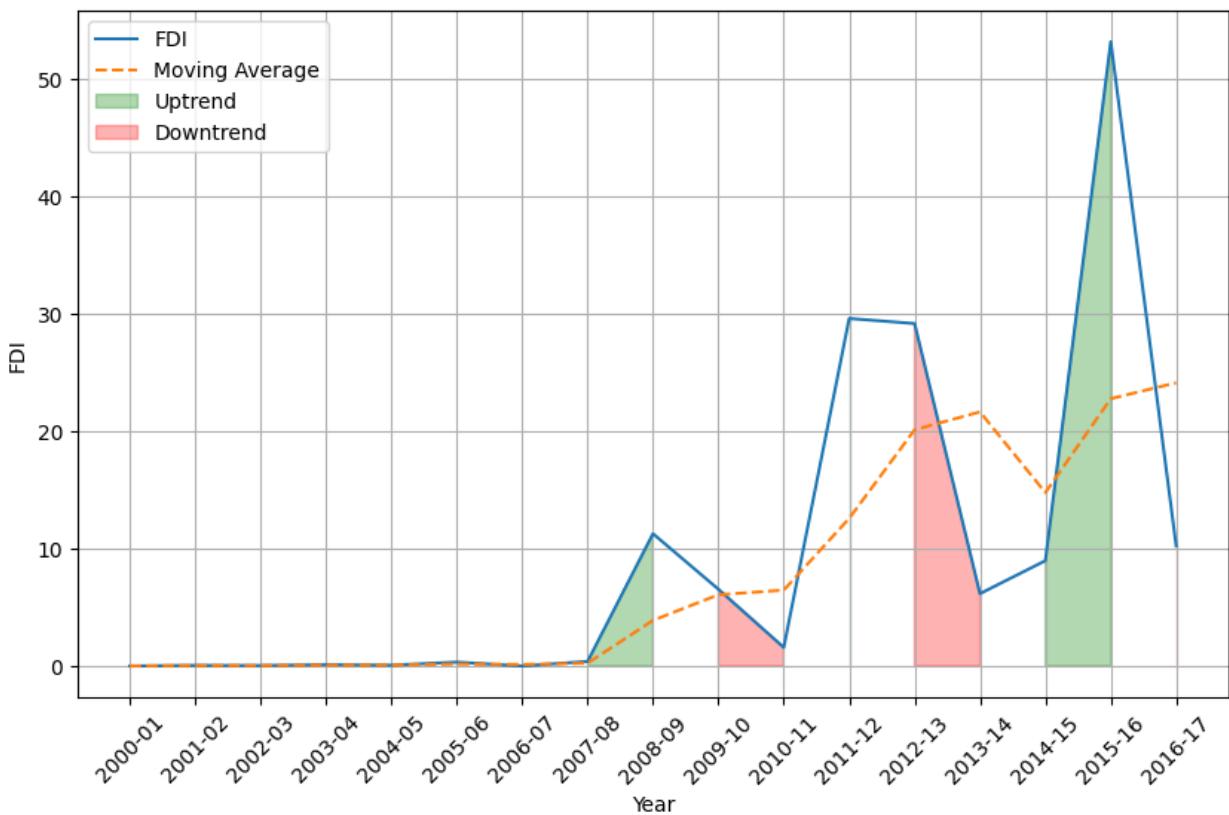




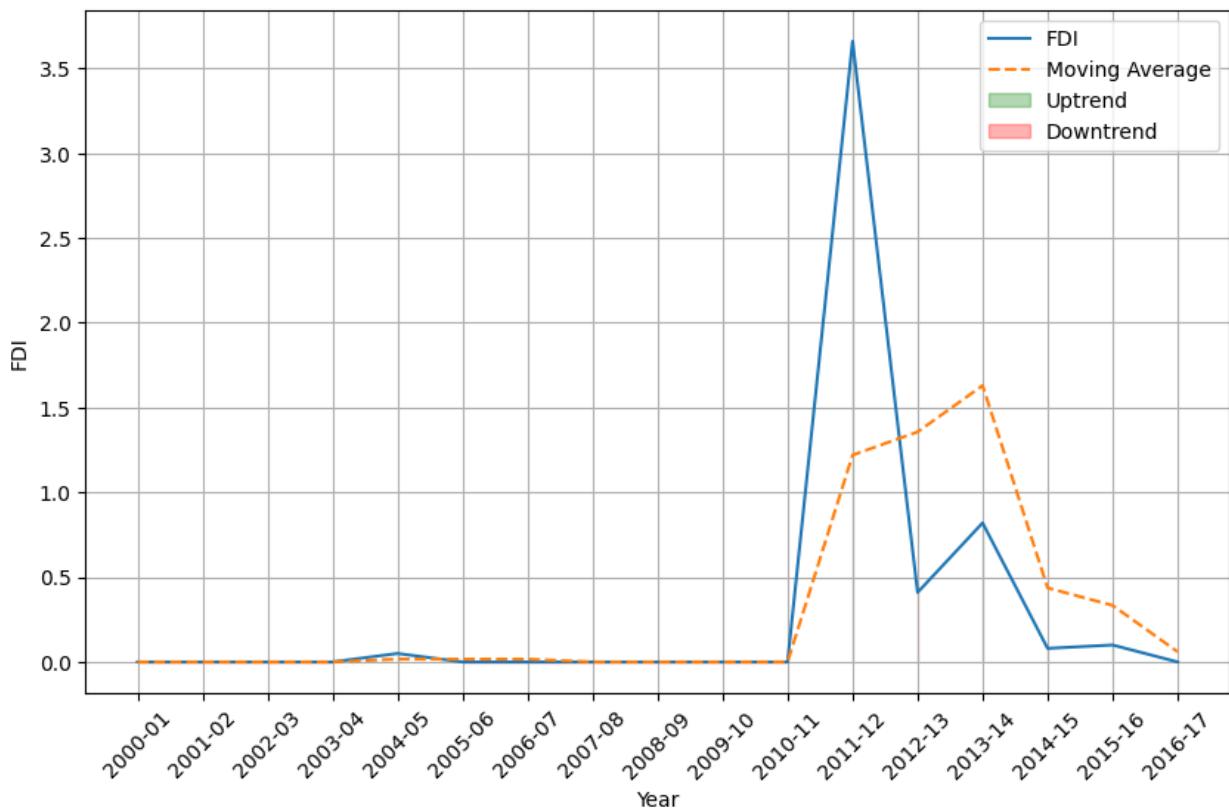
FDI Trends for CEMENT AND GYPSUM PRODUCTS

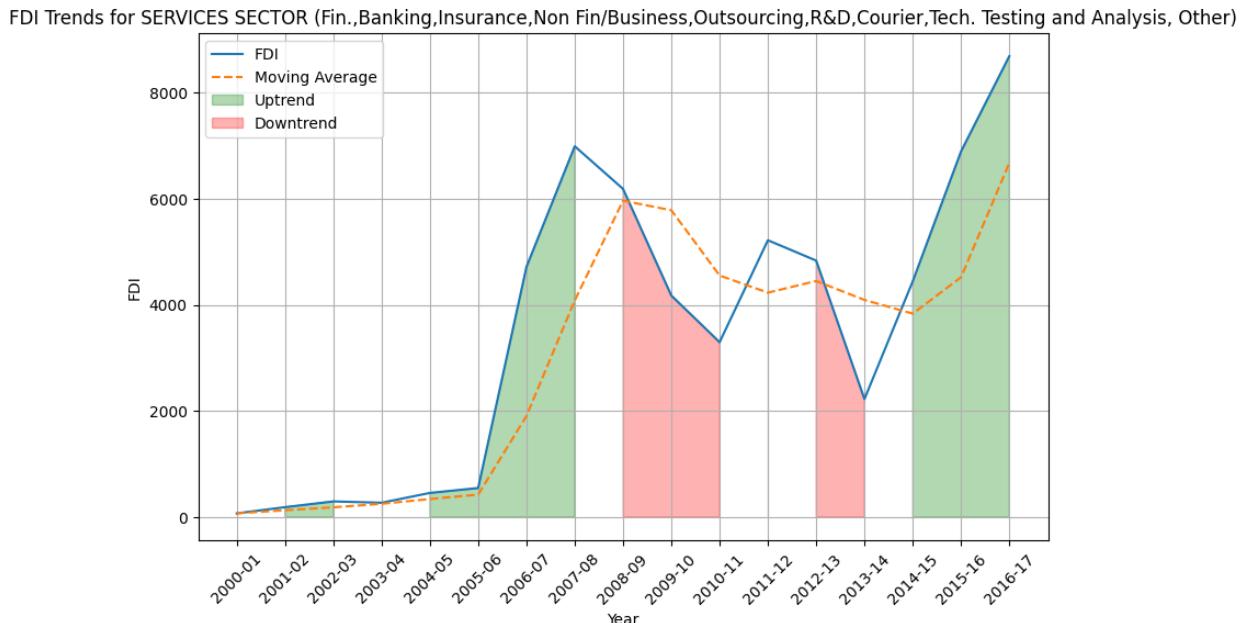
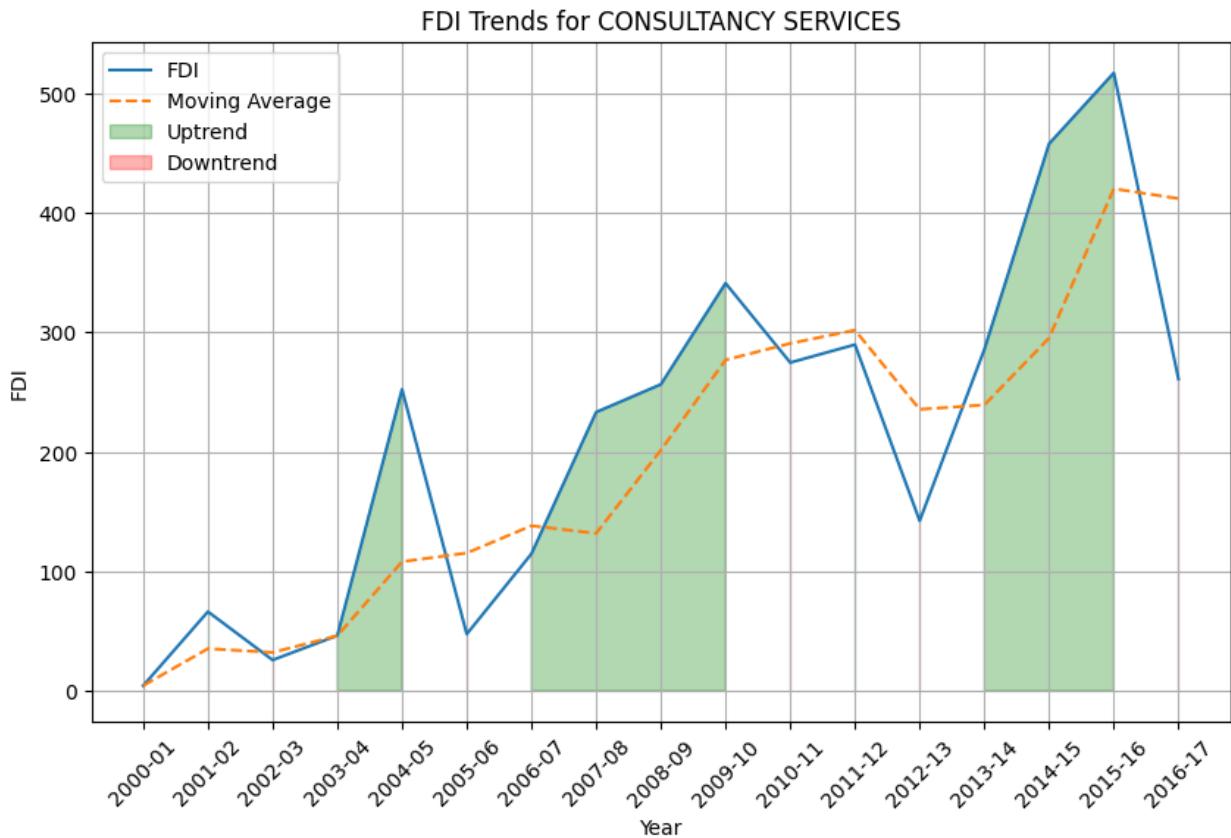


FDI Trends for TIMBER PRODUCTS

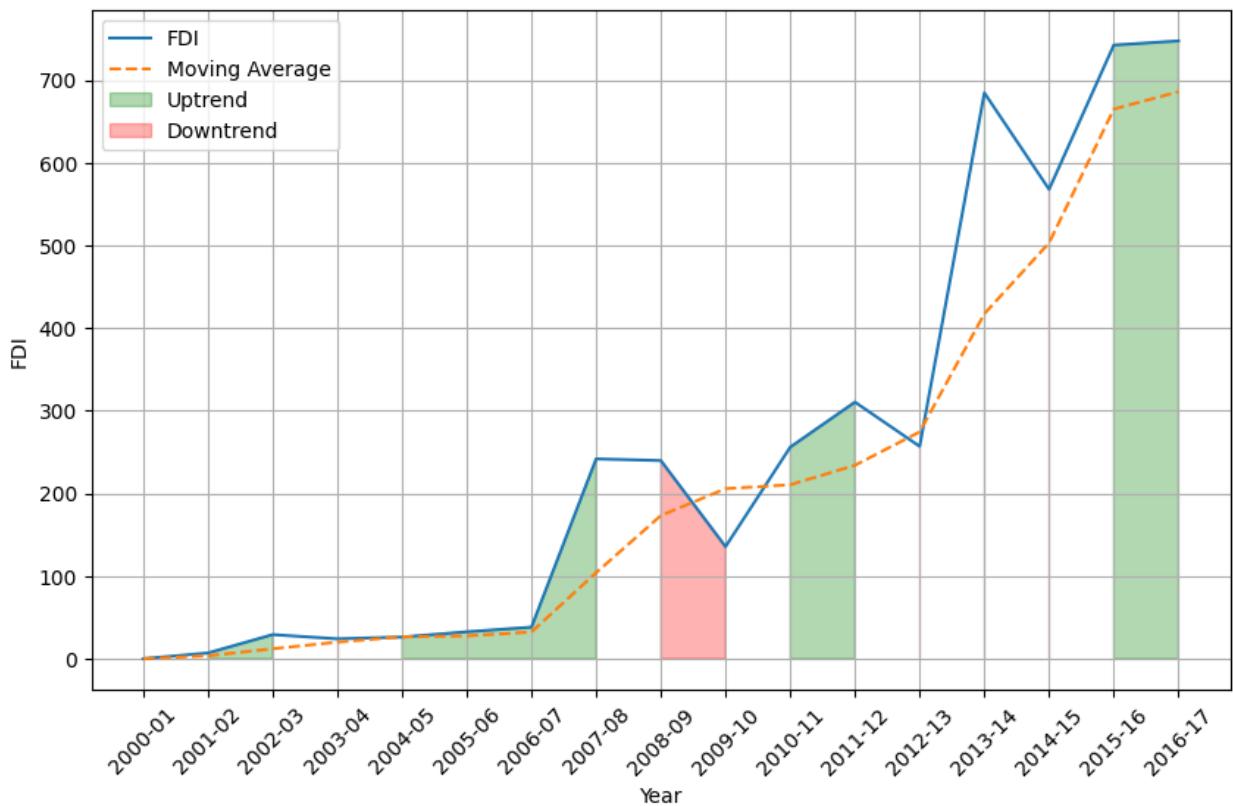


FDI Trends for DEFENCE INDUSTRIES

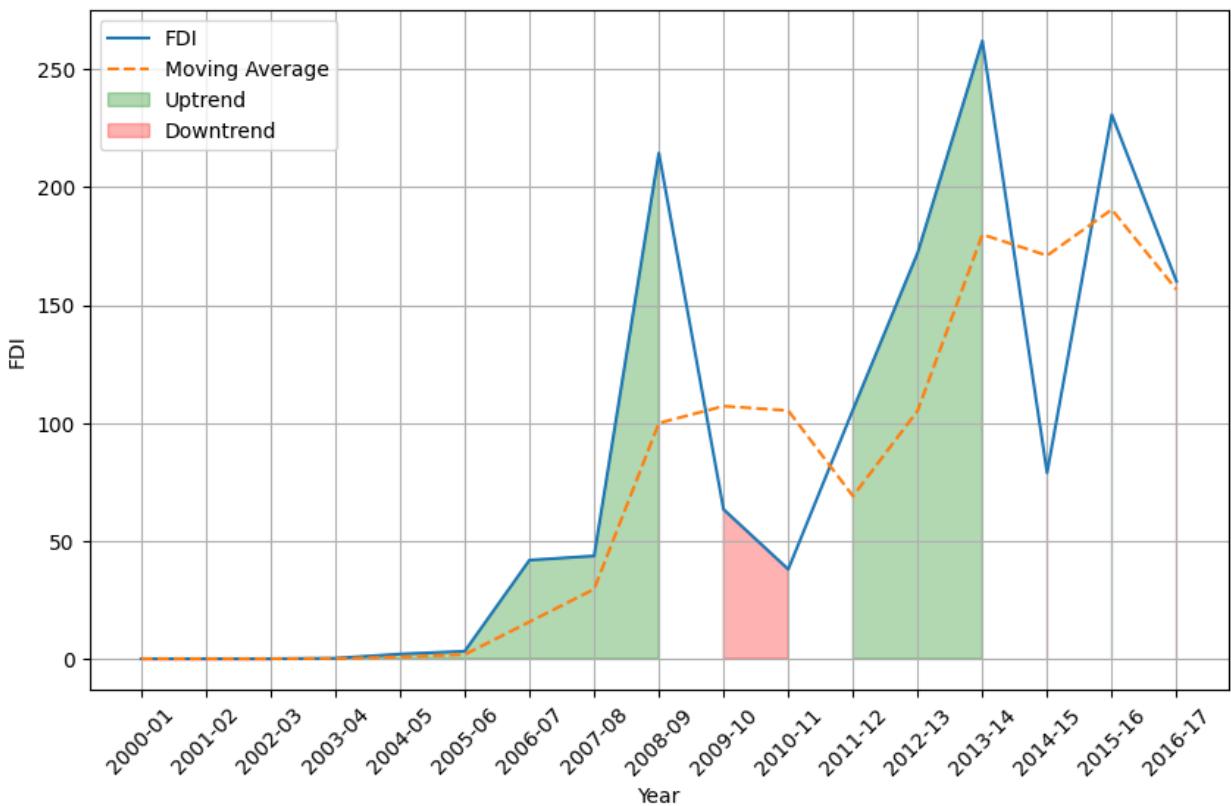




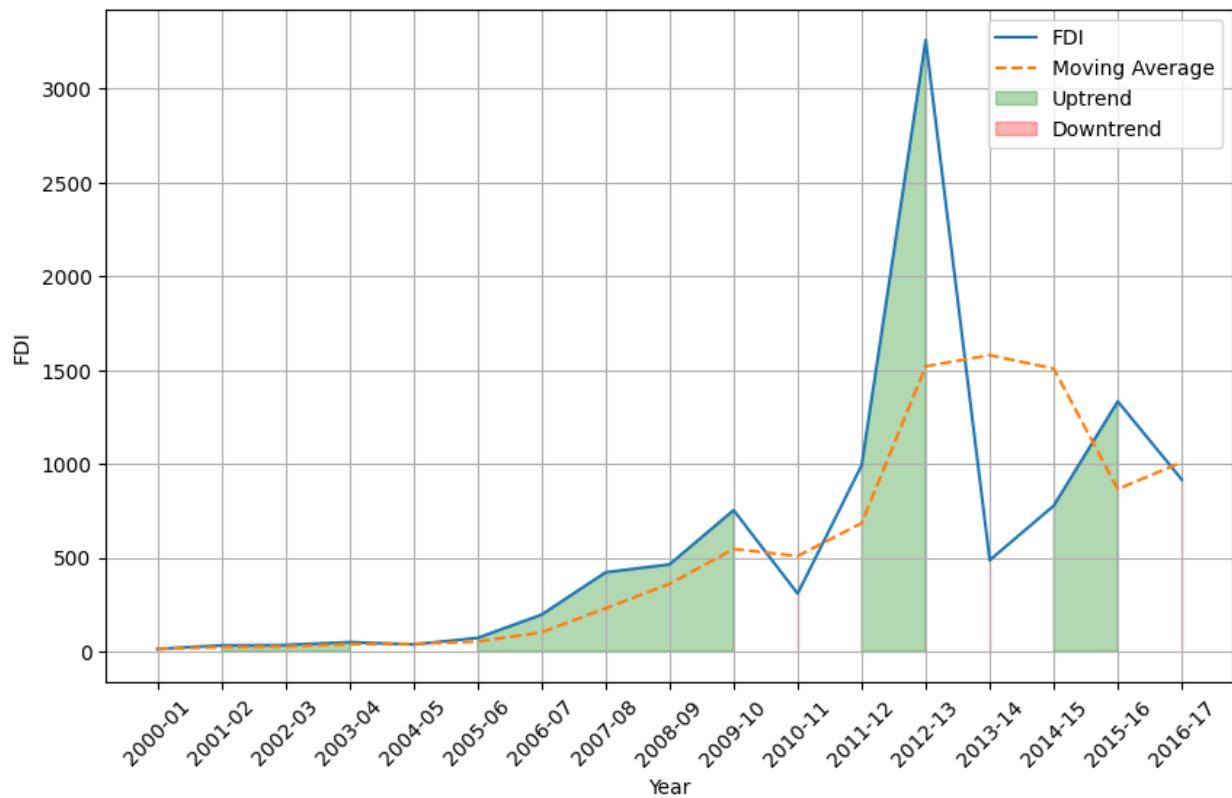
FDI Trends for HOSPITAL & DIAGNOSTIC CENTRES



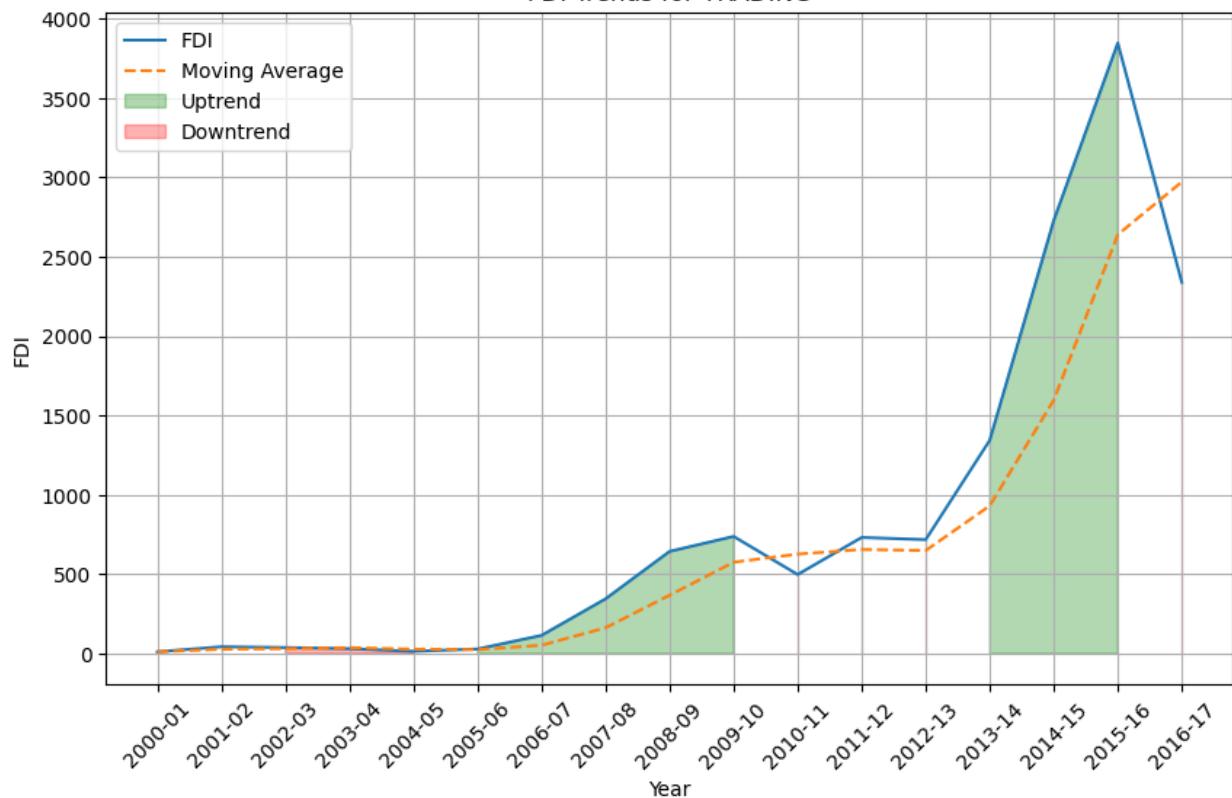
FDI Trends for EDUCATION



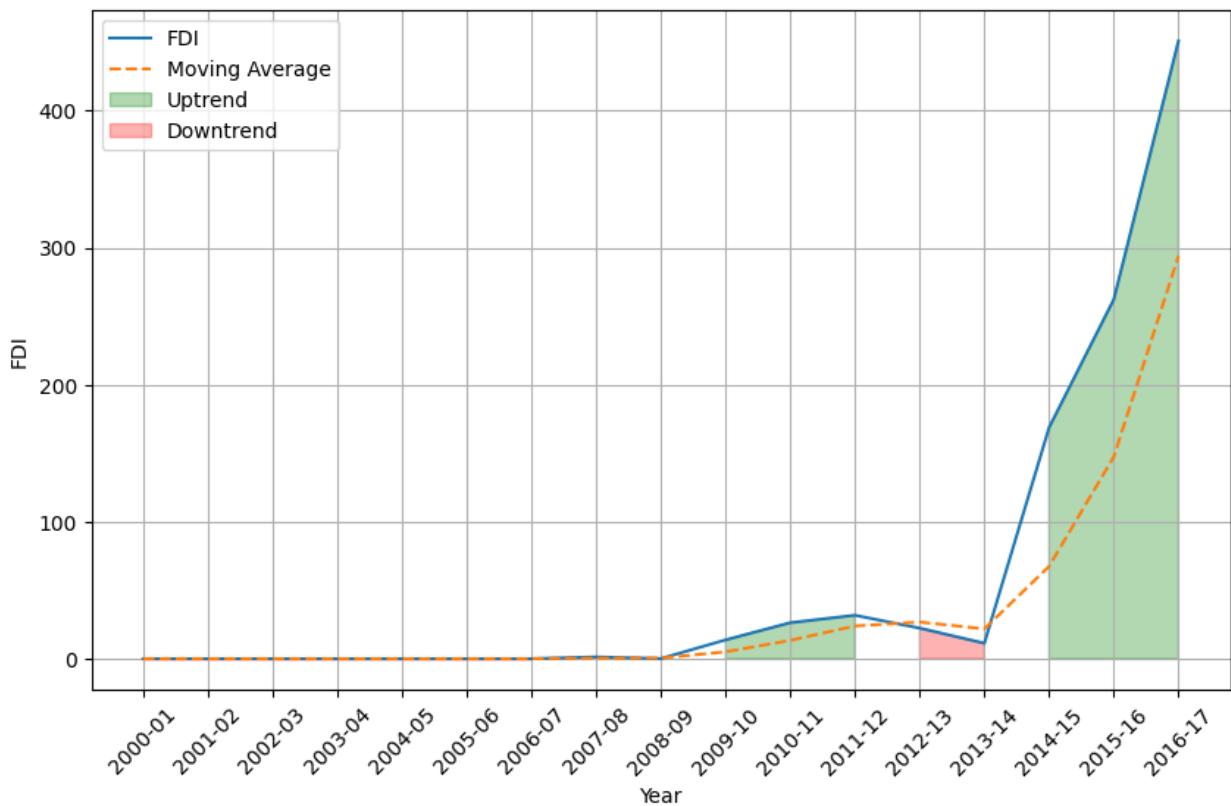
FDI Trends for HOTEL & TOURISM



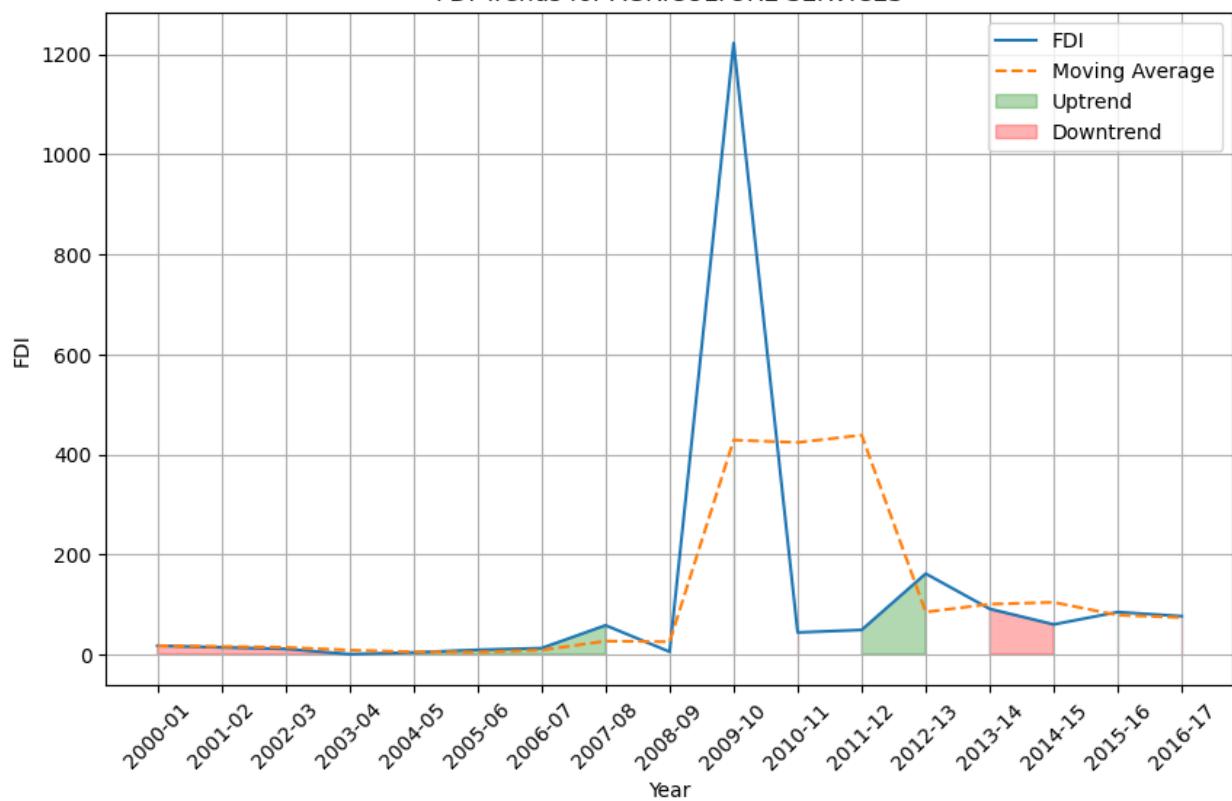
FDI Trends for TRADING



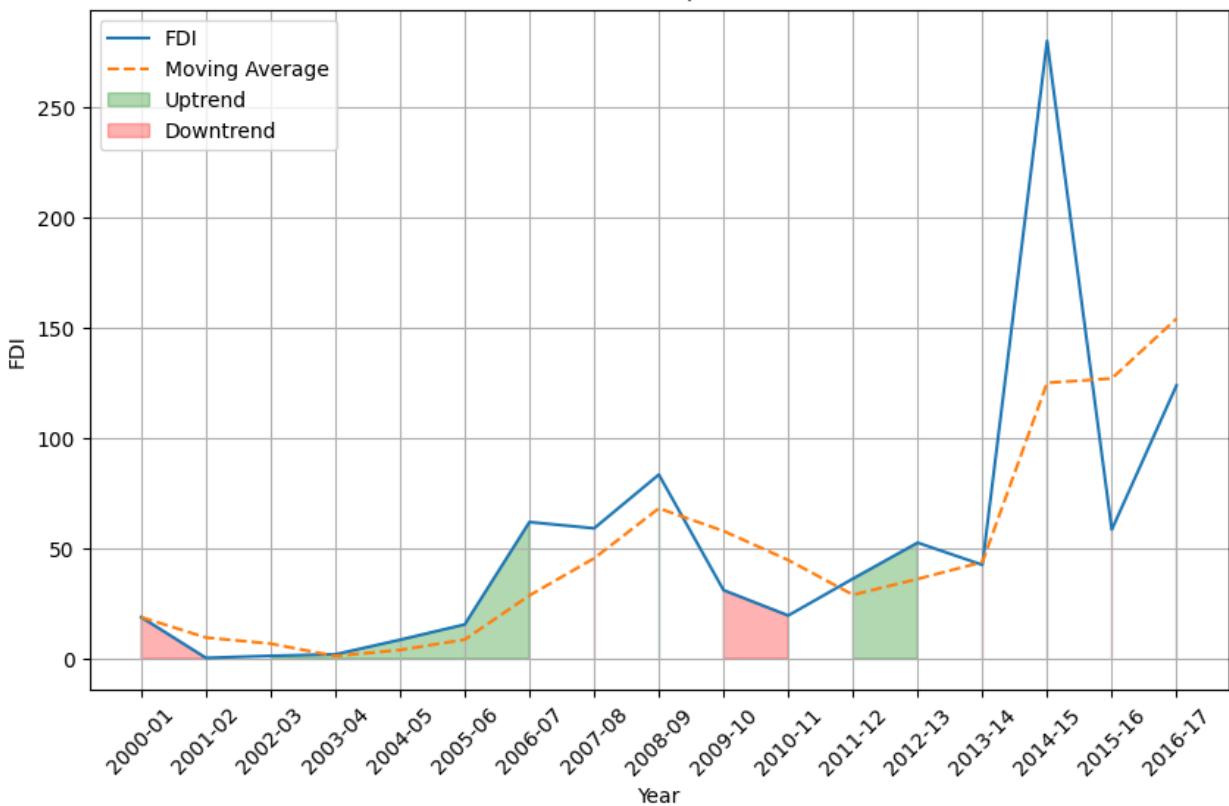
FDI Trends for RETAIL TRADING



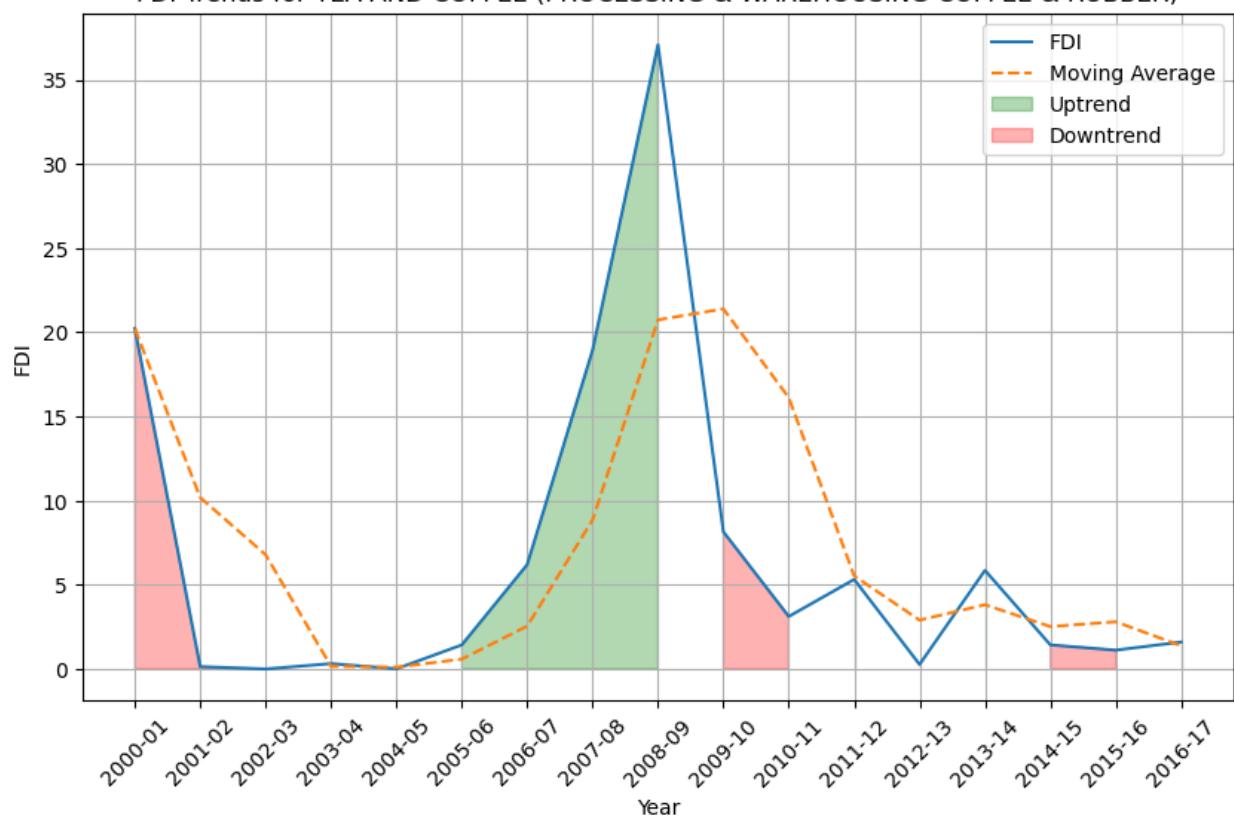
FDI Trends for AGRICULTURE SERVICES



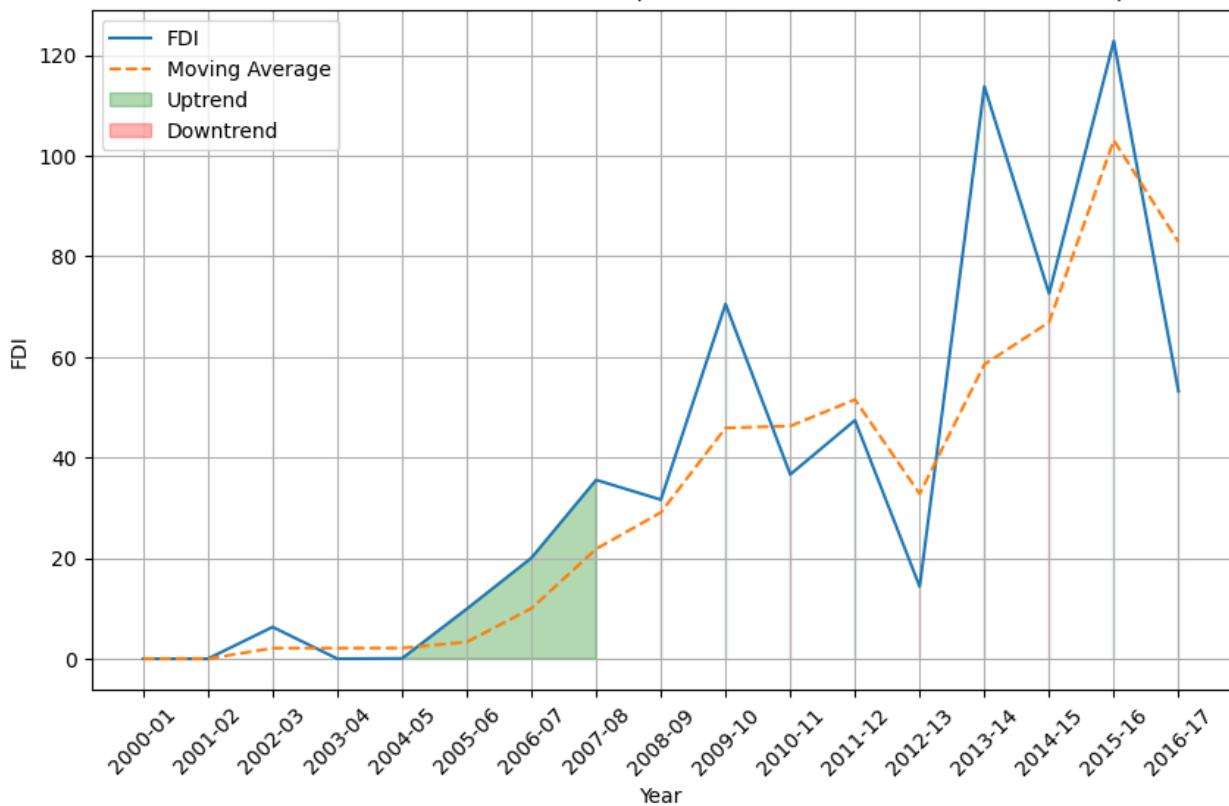
FDI Trends for DIAMOND,GOLD ORNAMENTS

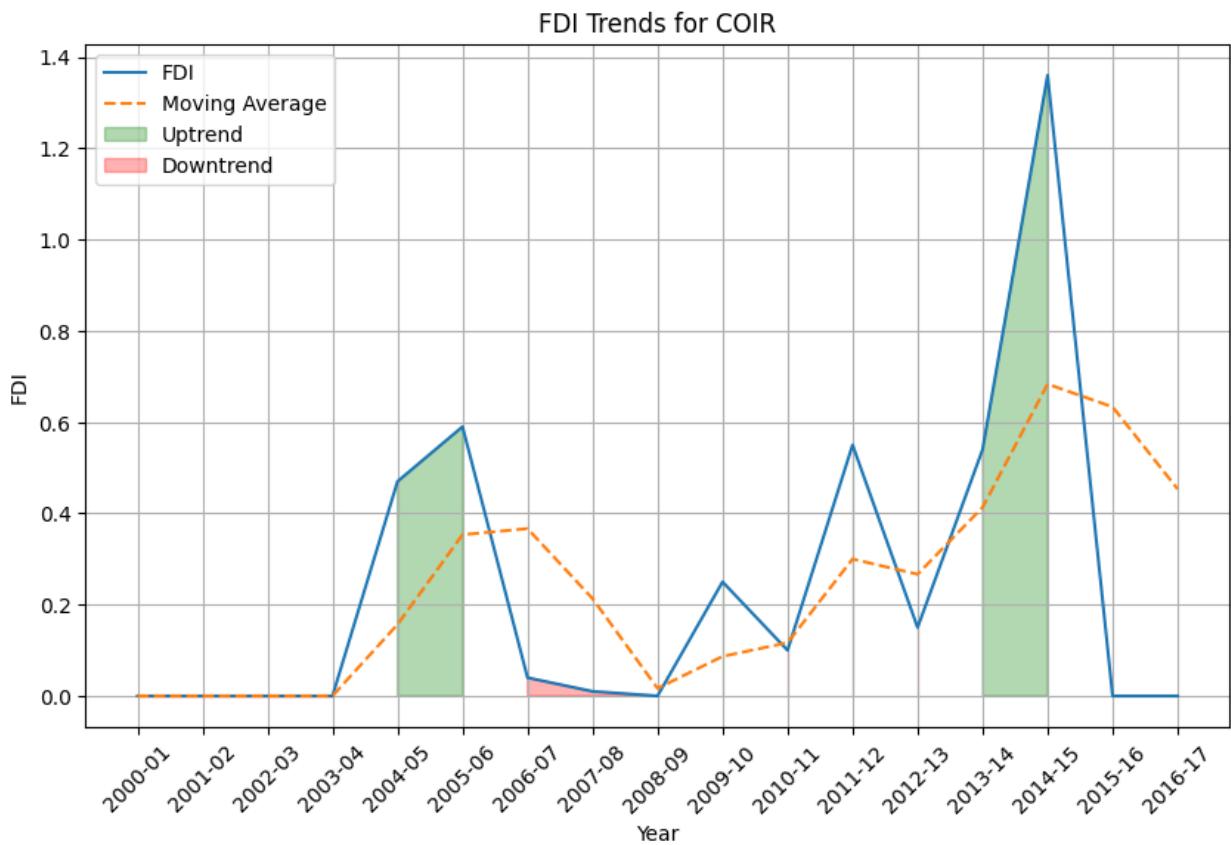


FDI Trends for TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)

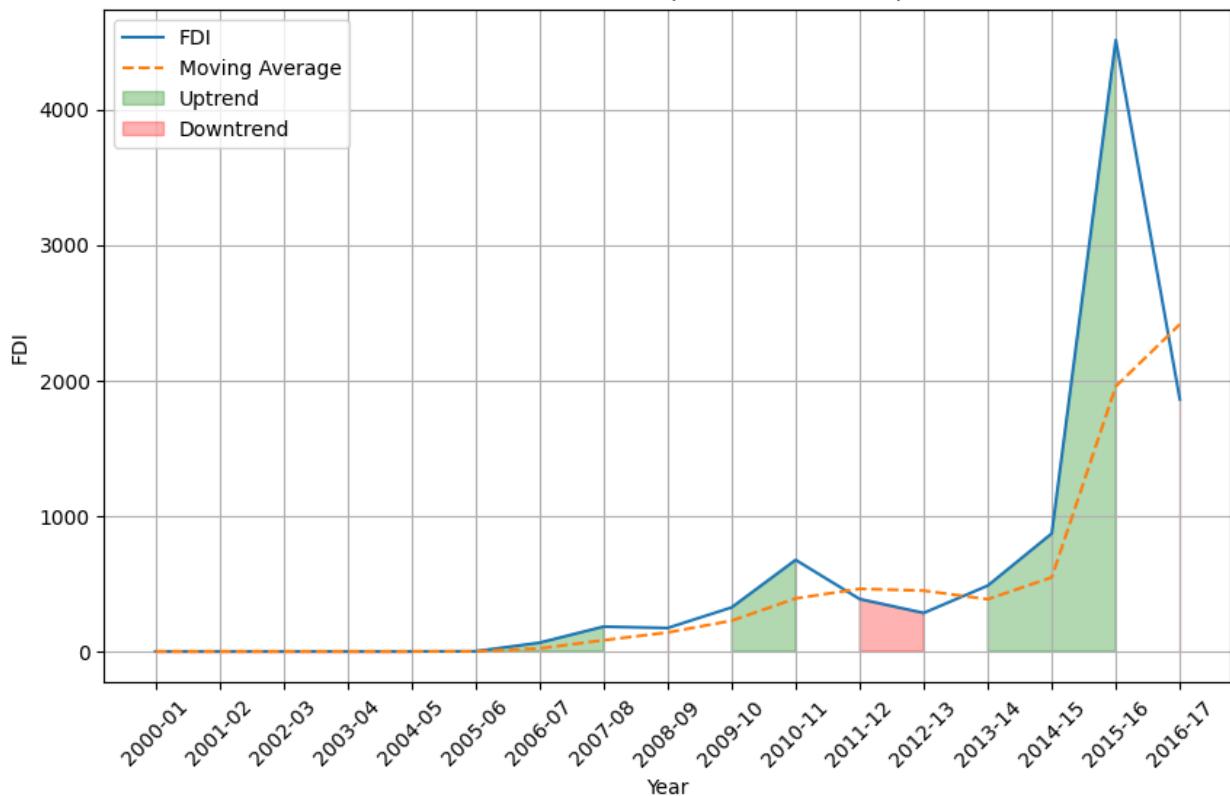


FDI Trends for PRINTING OF BOOKS (INCLUDING LITHO PRINTING INDUSTRY)

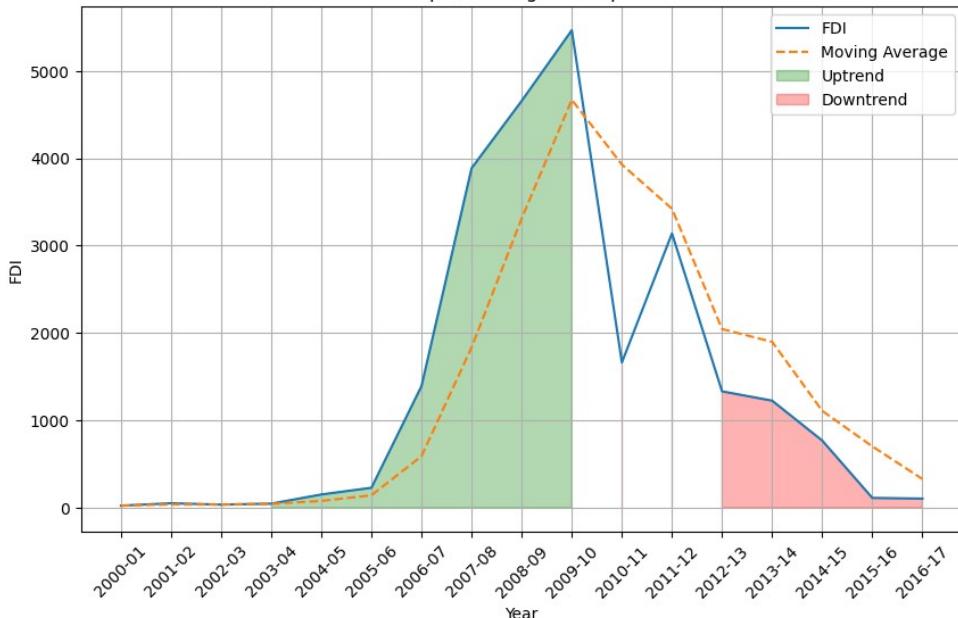


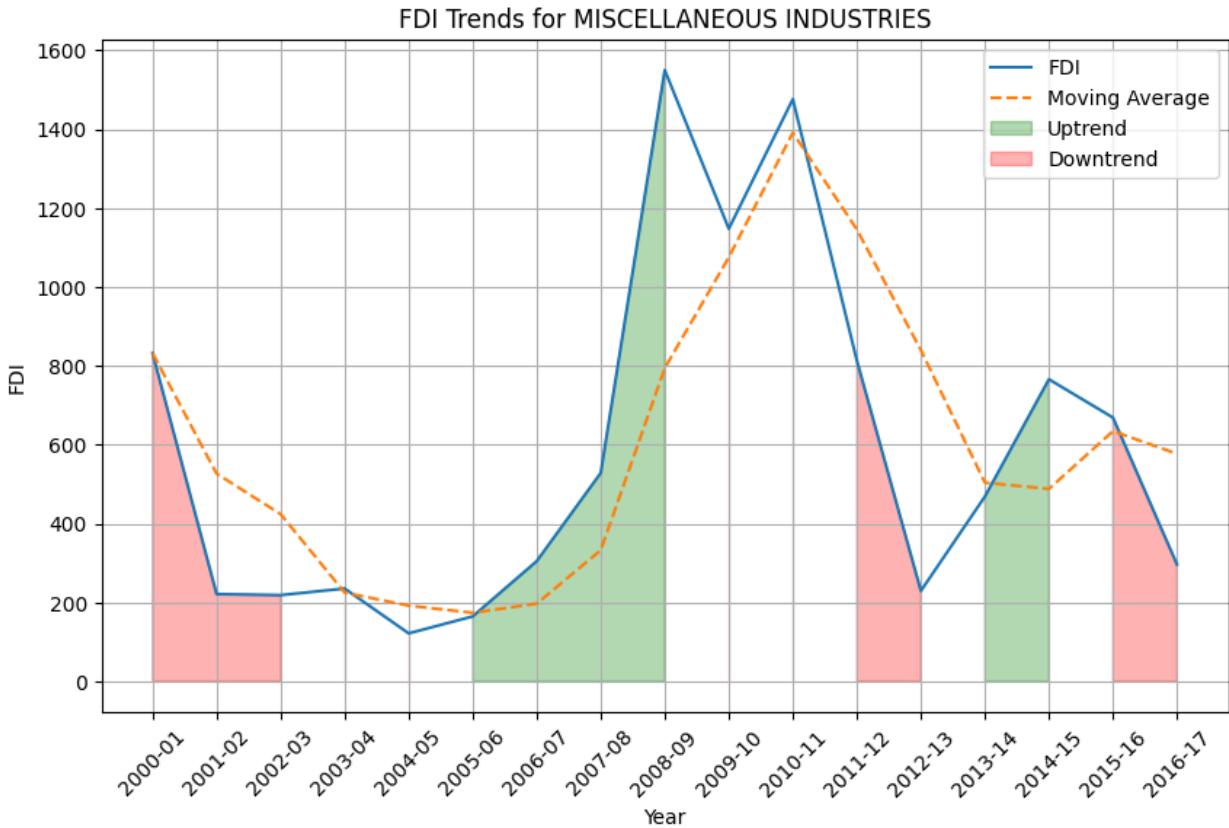


FDI Trends for CONSTRUCTION (INFRASTRUCTURE) ACTIVITIES



FDI Trends for CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects





```

year_total.drop('level_0', axis=1, inplace=True)
year_total

{
  "summary": {
    "name": "year_total",
    "rows": 17,
    "fields": [
      {
        "column": "index",
        "properties": {
          "dtype": "string",
          "num_unique_values": 17,
          "samples": [
            "2000-01",
            "2001-02",
            "2005-06"
          ],
          "semantic_type": "\",
          "description": "\n",
          "column": "Total_FDI",
          "properties": {
            "dtype": "number",
            "std": 14221.119667993022,
            "min": 2187.85,
            "max": 43478.26,
            "num_unique_values": 17,
            "samples": [
              2378.709999999996,
              4027.689999999996,
              5539.750000000001
            ],
            "semantic_type": "\",
            "description": "\n"
          }
        }
      }
    ],
    "type": "dataframe",
    "variable_name": "year_total"
  }
}

year_total['Year'] = year_total['index'].astype(str) # Ensure 'Year' is treated as a string
year_total = year_total.sort_values(by='Year') # Sort by year if not already sorted

# Calculate the growth rate

```

```

year_total['Growth Rate (%)'] = year_total['Total_FDI'].pct_change() * 100
year_total

{
  "summary": {
    "name": "year_total",
    "rows": 17,
    "fields": [
      {
        "column": "index",
        "properties": {
          "dtype": "string",
          "num_unique_values": 17,
          "samples": [
            "2000-01",
            "2001-02",
            "2005-06"
          ],
          "semantic_type": "\",
          "description": "\n"
        }
      },
      {
        "column": "Total_FDI",
        "properties": {
          "dtype": "number",
          "std": 14221.119667993022,
          "min": 2187.85,
          "max": 43478.26,
          "num_unique_values": 17,
          "samples": [
            2378.709999999996,
            4027.689999999996,
            5539.750000000001
          ],
          "semantic_type": "\",
          "description": "\n"
        }
      },
      {
        "column": "Year",
        "properties": {
          "dtype": "string",
          "num_unique_values": 17,
          "samples": [
            "2000-01",
            "2001-02",
            "2005-06"
          ],
          "semantic_type": "\",
          "description": "\n"
        }
      },
      {
        "column": "Growth Rate (%)",
        "properties": {
          "dtype": "number",
          "std": 47.95377737315086,
          "min": -36.15292712747266,
          "max": 125.49320817726435,
          "num_unique_values": 16,
          "samples": [
            69.32244788141473,
            -32.856799,
            125.49320817726435
          ],
          "semantic_type": "\",
          "description": "\n"
        }
      }
    ]
  },
  "type": "dataframe",
  "variable_name": "year_total"
}

year_total.dropna(inplace=True)

# Display the DataFrame with the calculated growth rate
print(year_total[['Year', 'Total_FDI', 'Growth Rate (%)']])

```

	Year	Total_FDI	Growth Rate (%)
1	2001-02	4027.69	69.322448
2	2002-03	2704.32	-32.856799
3	2003-04	2187.85	-19.097962
4	2004-05	3218.69	47.116576
5	2005-06	5539.75	72.111946
6	2006-07	12491.76	125.493208
7	2007-08	24575.40	96.732886
8	2008-09	31395.96	27.753607
9	2009-10	25834.38	-17.714317
10	2010-11	21383.07	-17.230179
11	2011-12	35120.78	64.245733
12	2012-13	22423.59	-36.152927
13	2013-14	24299.32	8.364985
14	2014-15	30930.47	27.289447

```

15 2015-16 40000.99      29.325516
16 2016-17 43478.26      8.692960

import plotly.express as px

fig = px.bar(year_total, x='Year', y='Growth Rate (%)', text='Growth
Rate (%)',
             title='Year-over-Year Growth Rates of FDI in India',
             labels={'Growth Rate (%)': 'Growth Rate (%)', 'Year':
'Year'})

# Customize the hover template to show the growth rate value
fig.update_traces(texttemplate=' %{text:.2f}%', textposition='outside',
hovertemplate='Year: %{x}<br>Growth Rate: %{y:.2f}%')

# Customize the layout for better readability
fig.update_layout(title_font_size=16, xaxis_title='Year',
yaxis_title='Growth Rate (%)',
xaxis_tickangle=-45, height=500, width=900)

# Show the plot
fig.show()

```

#Key observations from the year-wise growth analysis: **Volatile growth:** FDI growth rates have been highly volatile, ranging from -36% to 125%.

Peak growth periods: The highest growth rates were observed in 2005-06 (72%), 2006-07 (125%), and 2007-08 (97%).

Global financial crisis impact: Negative growth rates in 2009-10 and 2010-11 coincide with the global financial crisis.

Recent stabilization: Growth rates have been positive but more moderate in recent years (2013-14 to 2016-17).

Based on our analysis, some key metrics and factors influencing FDI in India include:

- a. Total annual FDI inflow
- b. Sector-wise FDI distribution
- c. Year-over-year growth rates
- d. Global economic conditions (e.g., the 2008 financial crisis)
- e. Policy changes and economic reforms in India
- f. Technological advancements and digital transformation

g. Infrastructure development

h. Skilled workforce availability, especially in the services and tech sectors