

Analysis of FDA Dataset

Problem Statement

FDI in various sectors is impacted by several factors, such as government policies, economic reforms, or global events.

The objective is to:

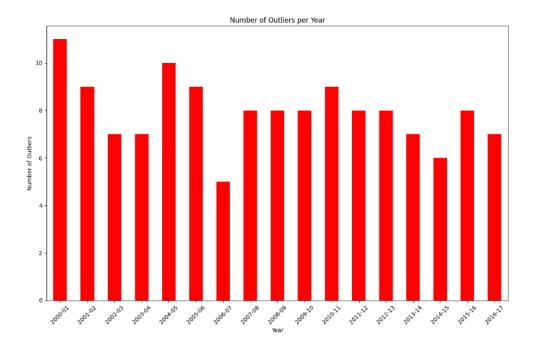
- 1. analyze trends and patterns in FDI inflows across different sectors over the years to identify sectors attracting more or less investments.
- 2. Forecasting or predicting the future FDI values for different sectors for the year 2017-2018 based on historical data to aid investment planning and policymaking.
- 3. Currently the dataset has values only upto 2016-2017. Comparing FDI inflows across sectors to understand the relative attractiveness and growth potential of different industries in India.

DATA CLEANING AND TRANSFORMATION

- Converted the header for columns(years) into readable and concise format. Eg: 2000-2001 to 2000_01 for neat plotting purposes
- Added and additional column Total FD to ease in plotting certain values
- Converted the 'Sector' column into index (reset it as index)
- Almost all the columns had outliers due to the wise range of variables from the mean. Treated the outliers with the capping method of 99 percentile

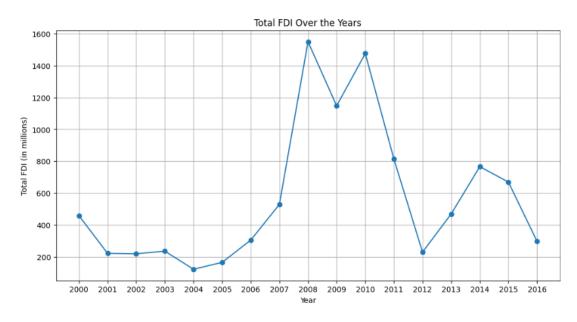
• Descriptive analysis & observations

- i. Each year consistently has data for 63 sectors.
- ii. Mean: The average investment per sector each year, which shows a general upward trend over the years.
- iii. Here are some specific details for selected years: There is a high difference between mean and median coupled with high standard deviation for almost all years, which indicates the presence of outliers in the dataset. In most cases, there is a big difference between the mean and the max values, the mean is greater than the median, which indicates right skewed dataset, with outliers on the higher end. The high SD also shows the spread of data away from the mean



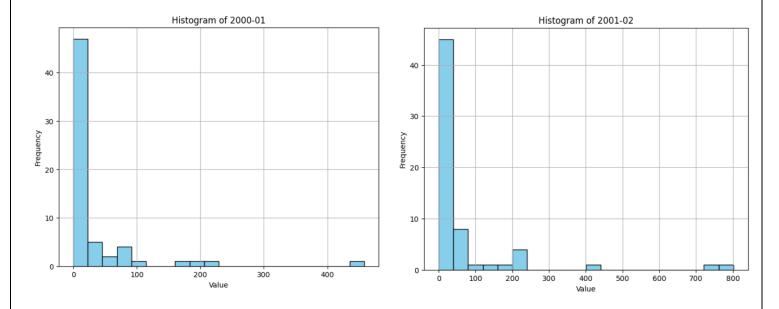
DATA INSIGHTS

 The total FDA was at its peak during the year 2008 (about 1580 units), followed by the year 2010. It was at its lowest in 2004 with 100 units. The FDA seems to be declining the last two years.

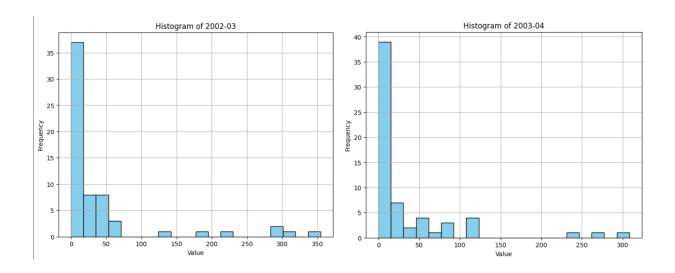


FDA distribution by year

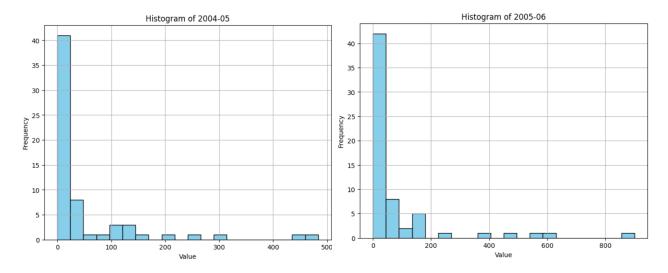
i. The year 2000-01 is dense around 0-120 units, the peak being between 0-40 units. The year 2001-02 is dense around 0-200 units, the peak being between 0-80 units



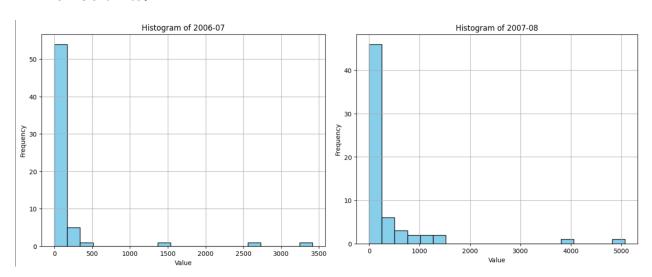
ii. The year 2002-03 is dense around 0-60 units, the peak being between 0-20 units. The year 2003-04 is dense around 0-90 units, the peak being between 0-20 units.



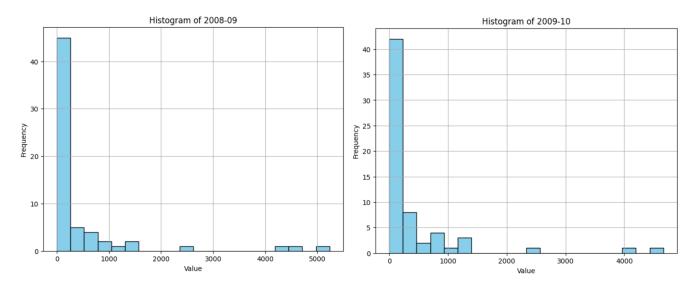
iii. The year 2004-05 is dense around 0-180 units, the peak being between 0-30 units. The year 2005-06 is dense around 0-190 units, the peak being between 0-20 units.



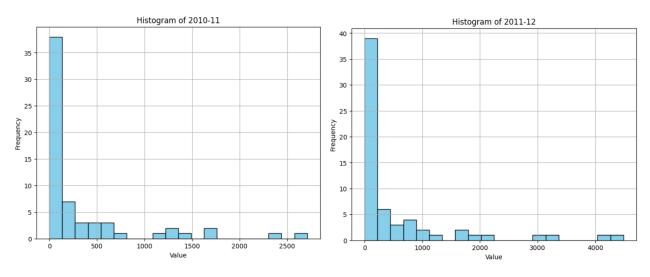
iv. The year 2006-07 is dense around 0-200 units, the peak being between 0-170 units. The year 2007-08 is dense around 0-1500 units, the peak being between 0-250 units.



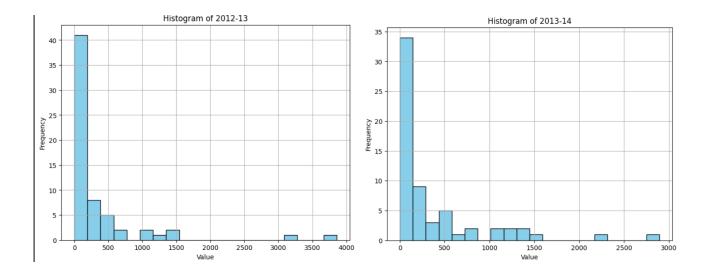
v. The year 2008-09 is dense around 0-1500 units, the peak being between 0-250 units. The year 2009-10 is dense around 0-1400 units, the peak being between 0-250 units.



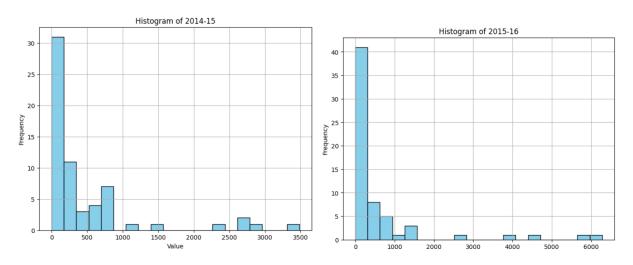
vi. The year 2010-11 is dense around 0-800 units, the peak being between 0-125 units. The year 2011-12 is dense around 0-1300 units, the peak being between 0-250 units.



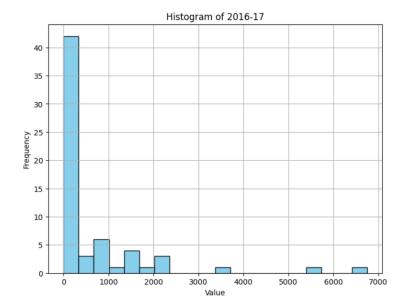
vii. The year 2010-11 is dense around 0-700 units, the peak being between 0-250 units. The year 2011-12 is dense around 0-900 units, the peak being between 0-250 units.



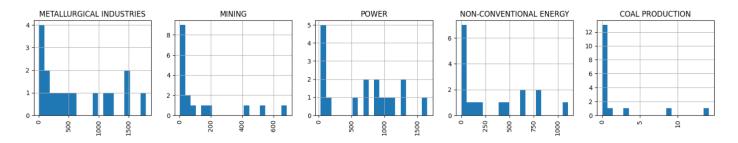
viii. The year 2014-15 is dense around 0-900 units, the peak being between 0-200 units. The year 2015-16 is dense around 0-1500 units, the peak being between 0-250 units.



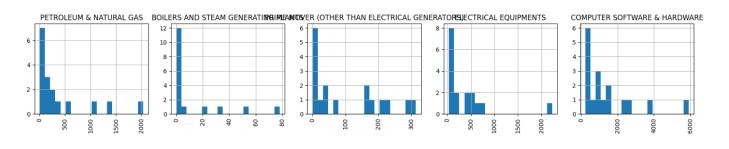
ix. The year 2016-17 is dense around 0-2200 units, the peak being between 0-250 units.



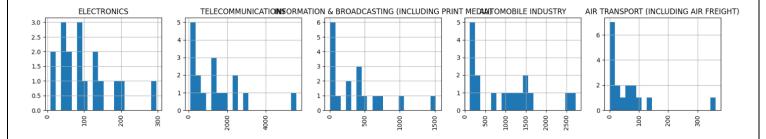
- Metallurgical industries: FDA frequency is dense between 0-600 units, the peal being at 0-250 units
- Mining: FDA is dense between 0-100 units, the peak being at 0-25 units
- Power: The peak being at 0-150 units



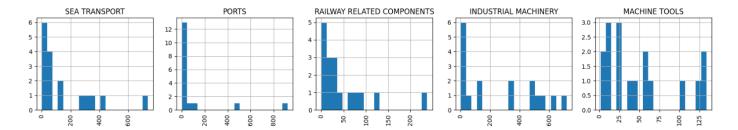
- Petroleum and Natural gas: FDA frequency is dense between 0-400 units, the peak being at 0-100 units
- BOILERS AND STEAM GENERATING PLANTS: FDA is dense between 0-15 units, the peak being at 0-5 units
- PRIME MOVER (OTHER THAN ELECTRICAL GENERATORS): The peak is at 0-10 units
- ELECTRICAL EQUIPMENTS: The peak is at 0-100 units
- COMPUTER SOFTWARE & HARDWARE: FDA is dense between 0-1800 units, the peak being at 0-250 units



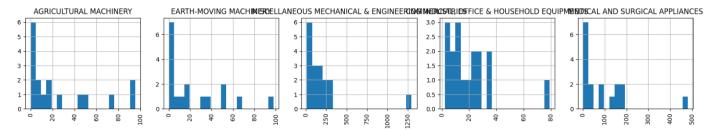
- Electronics: FDA frequency is mostly uniform, the peak being at 40-50 and 90-100 units
- TELECOMMUNICATION: FDA is dense between 0-300 units, the peak being at 0-250 units
- INFORMATION & BROADCASTING (PRINT MEDIA): The peak distribution is at 0-63 units
- AUTOMOBILE INDUSTRY: The peak distribution is at 0-125 units
- AIR TRANSPORT(INCLUDING AIR FREIGHT): The peak distribution is at 0-12.5 units



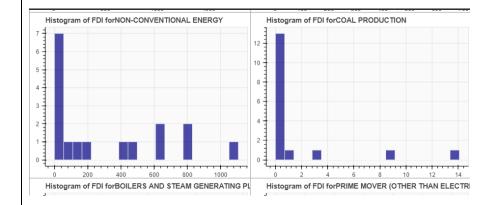
- SEA TRANSPORT: The peak distribution is at 0-50 units
- PORTS: The peak distribution is at 0-25 units
- RAILWAY RELATED COMPONENTS: The peak distribution is at 0-12.5 units
- INDUSTRIAL MACINERY: The peak distribution is at 0-25 units
- AIR TRANSPORT(INCLUDING AIR FREIGHT): The peak distribution is at 0-12.5 units

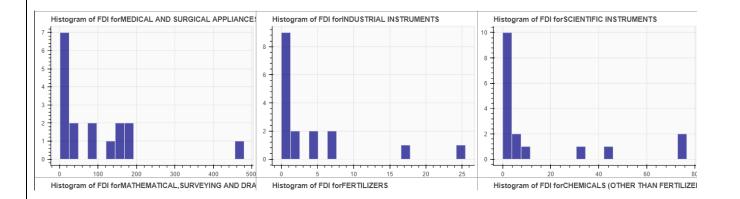


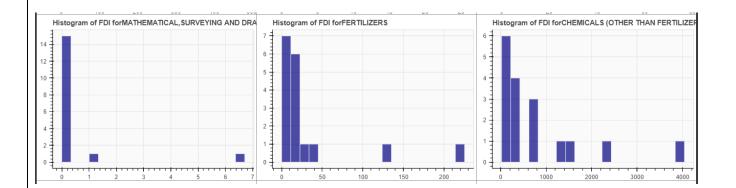
- AGRICULTURAL MACHINERY: The peak distribution is at 0-5 units
- EARTH MOVING MACINERY: The peak distribution is at 0-25 units
- MISCELLANEOUS MECHANICAL & ENGINEERING: The peak distribution is at 0-62 units. It is most dense at 0-300 units
- OFFICE AND HOUSEHOLD EQUIPMENTS: The two peak distribution one at 0-5 units and one at 10-15 units. The dense distribution is between 0-30 units
- NUMERICAL AND SURGICAL APPLIANCES: The peak distribution is at 0-50 units



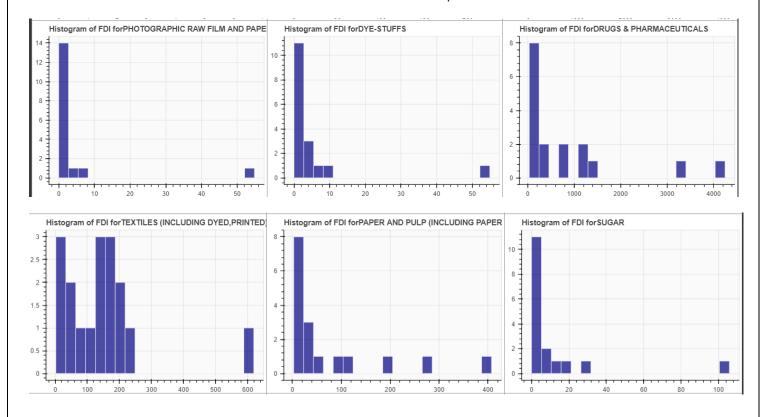
- AGRICULTURAL MACHINERY: The peak distribution is at 0-5 units
- EARTH MOVING MACINERY: The peak distribution is at 0-25 units
- MISCELLANEOUS MECHANICAL & ENGINEERING: The peak distribution is at 0-62 units. It is most dense at 0-300 units
- OFFICE AND HOUSEHOLD EQUIPMENTS: The two peak distribution one at 0-5 units and one at 10-15 units. The dense distribution is between 0-30 units
- NUMERICAL AND SURGICAL APPLIANCES: The peak distribution is at 0-50 units

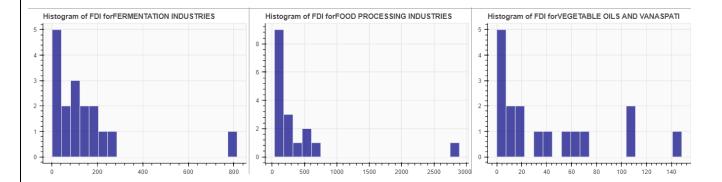




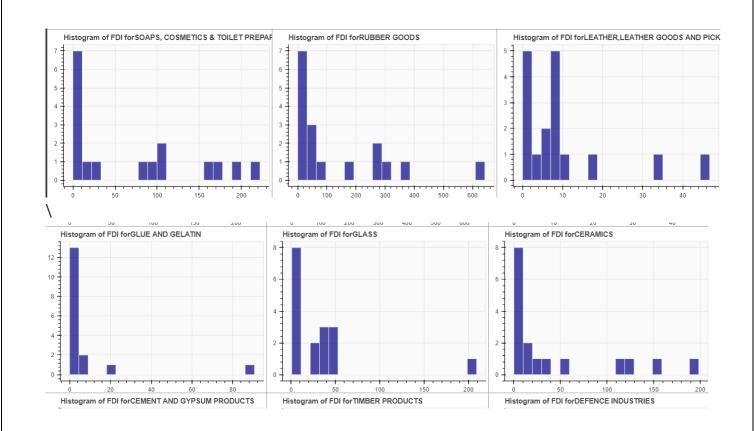


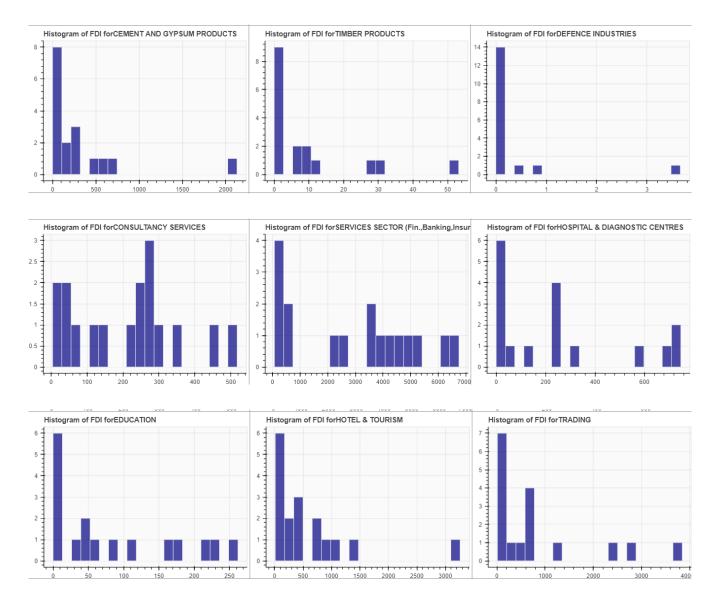
- AGRICULTURAL MACHINERY: The peak distribution is at 0-5 units
- EARTH MOVING MACINERY: The peak distribution is at 0-25 units
- MISCELLANEOUS MECHANICAL & ENGINEERING: The peak distribution is at 0-62 units. It is most dense at 0-300 units
- OFFICE AND HOUSEHOLD EQUIPMENTS: The two peak distribution one at 0-5 units and one at 10-15 units. The dense distribution is between 0-30 units
- NUMERICAL AND SURGICAL APPLIANCES: The peak distribution is at 0-50 units



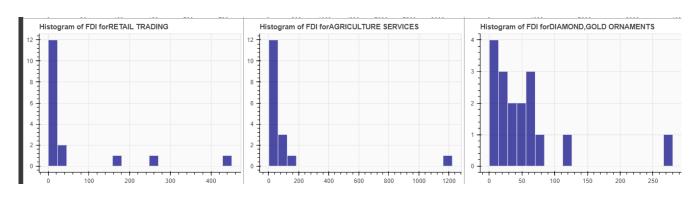


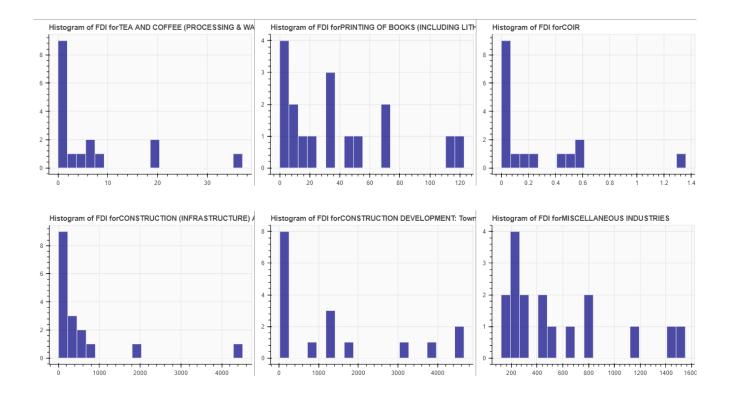
- AGRICULTURAL MACHINERY: The peak distribution is at 0-5 units
- EARTH MOVING MACINERY: The peak distribution is at 0-25 units
- MISCELLANEOUS MECHANICAL & ENGINEERING: The peak distribution is at 0-62 units. It is most dense at 0-300 units
- OFFICE AND HOUSEHOLD EQUIPMENTS: The two peak distribution one at 0-5 units and one at 10-15 units. The dense distribution is between 0-30 units
- NUMERICAL AND SURGICAL APPLIANCES: The peak distribution is at 0-50 units



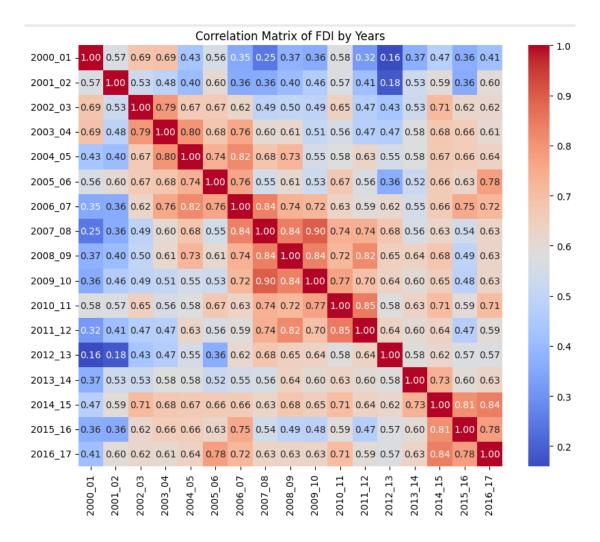


- AGRICULTURAL MACHINERY: The peak distribution is at 0-5 units
- EARTH MOVING MACINERY: The peak distribution is at 0-25 units
- MISCELLANEOUS MECHANICAL & ENGINEERING: The peak distribution is at 0-62 units. It is most dense at 0-300 units
- OFFICE AND HOUSEHOLD EQUIPMENTS: The two peak distribution one at 0-5 units and one at 10-15 units. The dense distribution is between 0-30 units
- NUMERICAL AND SURGICAL APPLIANCES: The peak distribution is at 0-50 units





CORRELATION MATRIX



70 and above 70 correlation values:

```
2000_01 - 2003_04

2002_03 - 2003_04

2003_04 - 2004_05, 2006_07

2004_05 - 2006_07

2006_07 - 2004_05

2007_08 - 2006_07, 2008_09, 2009_10, 2010_11, 2011_2012

2008_2009 - 2004_05, 2006_07, 2009_10, 2010_11, 2011_12

2009_10 - 2006_07, 2010_11

2014_15 - 2002_03, 2013_14, 2010_11, 2015_16, 2016_17

2016_17 - 2006_07, 2010_11, 2005_06, 2015_16

2011_12 - 2010_11
```

• Top 5 industries with highest total FDI. These sectors have the highest cumulative FDI contributions over the years.

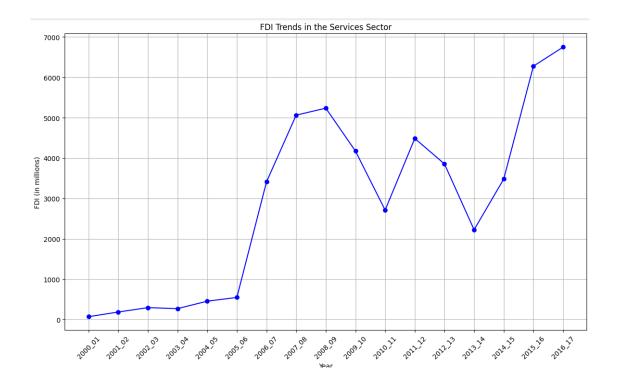
SECTOR	FDI
SERVICES SECTOR	49513.1148
(Fin/Business,Outsourcing,R&D,Courier,Tech.	
Testing and Analysis, Other)	
COMPUTER SOFTWARE & HARDWARE	24104.7258
TELECOMMUNICATIONS	23874.2202
CONSTRUCTION DEVELOPMENT: Townships,	23492.2980
housing, built-up infrastructure and	
construction-development projects	
AUTOMOBILE INDUSTRY	16608.3736

 Top 5 industries with highest total FDI. These sectors have the highest cumulative FDI contributions over the last 5 years

SECTOR	FDI
SERVICES SECTOR	22593.6774
(Fin/Business,Outsourcing,R&D,Courier,Tech.	
Testing and Analysis, Other)	
COMPUTER SOFTWARE & HARDWARE	13464.3400
TELECOMMUNICATIONS	11393.8500
TRADING	10972.8700
AUTOMOBILE INDUSTRY	9916.3400

Note: The order of the sector has not changed when we compae it with the toytal FDI values over the entire years, except instead of 'CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects' (rank 4 above with total FDI for all years) is not among the top 5 and is replaced by TRADING.

- FDI DISTRIBUTION BY SECTOR & YEAR
 - 1. SERVICES SECTOR (Fin/Business, Outsourcing, R&D, Courier, Tech. Testing and Analysis, Other) by year: The service sector seems to be growing in the recent years, the peak being in the tear 2016 17.

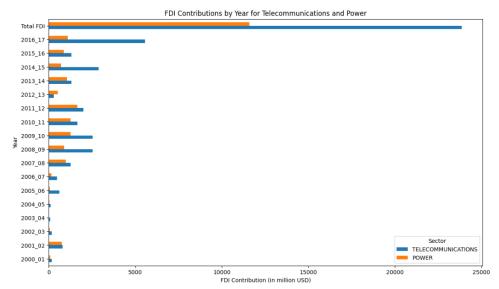


2011 vs 2016

COAL PRODUCTION, COIR, DEFENCE INDUSTRIES have started at 0 value in 20012 and also ended with 0 in the year 2015_16

COMMERCIAL, OFFICE & HOUSEHOLD EQUIPEMENTS, INDUSTRIAL INSTRUMENTS, LEATHER, LEATHER GOODS AND PICKERS, MISCELLANEOUS INDUSTRIES and TEA AND COFFEE(PROCESSING AND WAREHOUSING COFFEE & RUBBER) actually have a higher trend in the year 2011 as compared to 2015_16

POWER VS TELECOMMUNICATION TREND OVER YEARS



Overall, TELECOMMUNICATIONS seems to have a higher FDI as compared to POWER. TELECOMMUNICATIONS seems to peak the highest during the year 2016.

PREDICTING 2017 0r 2017 18

Used the lag method to predict the values in 2017

Following industries had the highest FDI in 2017

- SERVICES SECTOR
 (Fin.,Banking,Insurance,NonFin/Business,Outsourcing,R&D,Courier,Tech.
 Testing and Analysis, Other)
- TELECOMMUNICATIONS
- COMPUTER SOFTWARE & HARDWARE
- TRADING
- ELECTRICAL EQUIPMENTS
- CEMENT AND GYPSUM PRODUCTS
- CONSTRUCTION (INFRASTRUCTURE) ACTIVITIES
- AUTOMOBILE INDUSTRY
- INFORMATION & BROADCASTING (INCLUDING PRINT MEDIA)
- CHEMICALS (OTHER THAN FERTILIZERS)

Following industries had the lowest FDI in 2017

- LEATHER, LEATHER GOODS AND PICKERS
- TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)
- FERTILIZERS
- INDUSTRIAL INSTRUMENTS

- DEFENCE INDUSTRIES
- PORTS
- COAL PRODUCTION
- PHOTOGRAPHIC RAW FILM AND PAPER
- MATHEMATICAL, SURVEYING AND DRAWING INSTRUMENTS
- COIR

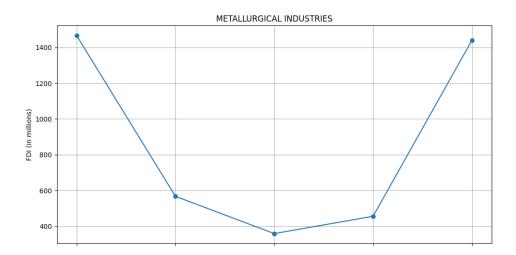
INDUSTRIES PREDICTED IN YEAR 2017, HAVING A HIGHER VALUE THAN YEAR 2016

- COAL PRODUCTION
- PETROLEUM & NATURAL GAS
- BOILERS AND STEAM GENERATING PLANTS
- PRIME MOVER (OTHER THAN ELECTRICAL GENERATORS)
- ELECTRONICS
- AIR TRANSPORT (INCLUDING AIR FREIGHT)
- 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
- PHOTOGRAPHIC RAW FILM AND PAPER
- DYE-STUFFS
- PAPER AND PULP (INCLUDING PAPER PRODUCTS)
- SUGAR
- FERMENTATION INDUSTRIES
- VEGETABLE OILS AND VANASPATI
- SOAPS, COSMETICS & TOILET PREPARATIONS
- RUBBER GOODS
- LEATHER, LEATHER GOODS AND PICKERS
- GLUE AND GELATIN
- GLASS51.6900

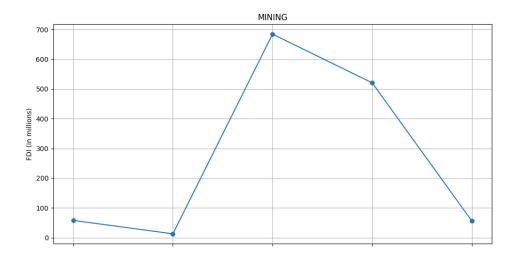
- CERAMICS
- TIMBER PRODUCTS
- DEFENCE INDUSTRIES
- CONSULTANCY SERVICES
- SERVICES SECTOR
 (Fin., Banking, Insurance, NonFin/Business, Outsourcing, R&D, Courier, Tech.
 Testing and Analysis, Other)
- EDUCATION
- RETAIL TRADING
- AGRICULTURE SERVICES
- DIAMOND, GOLD ORNAMENTS
- TEA AND COFFEE (PROCESSING & WAREHOUSING COFFEE & RUBBER)
- PRINTING OF BOOKS (INCLUDING LITHO PRINTING INDUSTRY)
- COIR
- CONSTRUCTION DEVELOPMENT: Townships, housing, built-up infrastructure and construction-development projects
- MISCELLANEOUS INDUSTRIES

TOP FDI VALUES IN LAST 5 YEARS: **2012_13, 2013_14, 2014_15, 2015_16, 2016_17**

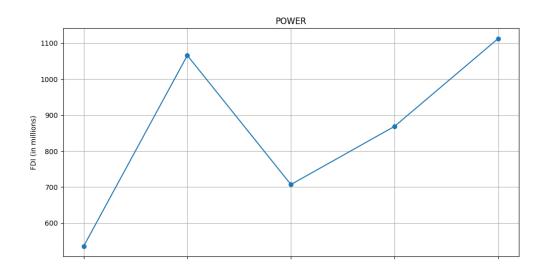
METALLURGICAL INDUSTRIES: The FDI has seen a sudden peak in the last year 2016_2017, although the highest still remain the year **2012_13**



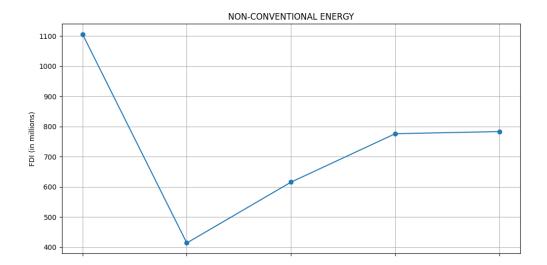
MINING: (2012, 2013, 2014, 2015, 2016): The FDI has seen a sudden peak in the year 2014_2015, after which it has seen a continuous decline. The lowest point was observed on the year 2013_14, followed h the last year 1015_16



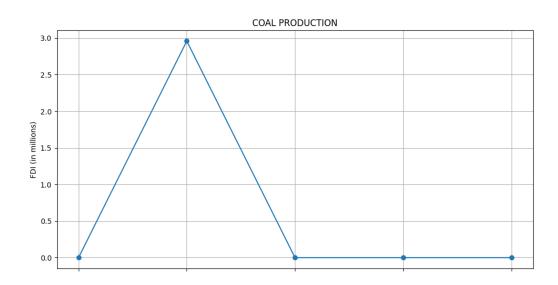
POWER: (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2015_2016 after 2 years of decline. The lowest point was observed on the year 2014_15



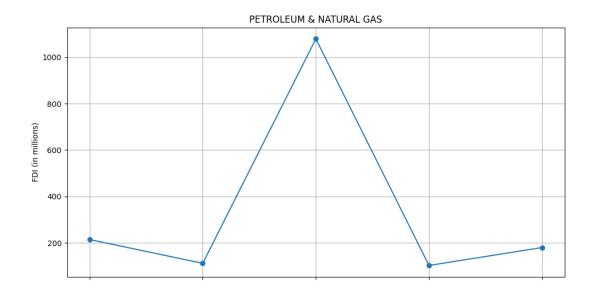
NON CONVENTIONAL ENERGY: (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2012_2013 (1100), followed by the lowest point was observed on the year 2013_14, although after which it begins to rise in the recent years but it is not as much as the year 2012_13.



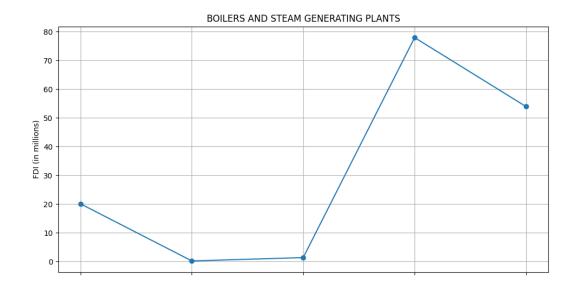
COAL PRODUCTION: (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2013_2014, followed by a continuous decline and constant value of 0 in the last 3 years.



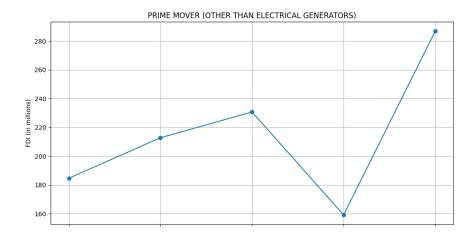
PETROLEUM & NATURAL GAS: (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest value being in the year 2014_15



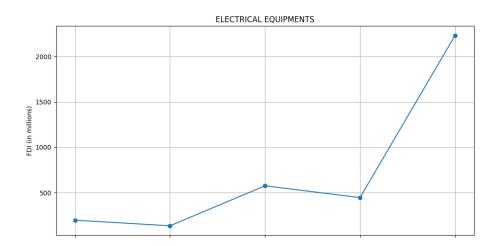
BOILERS & STEAM GENERATION PLANT: (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest value being in the year 2013_14.



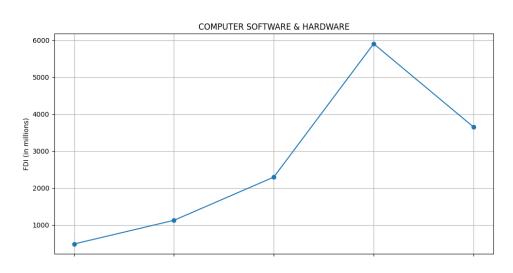
PRIME MOVER (OTHER THAN ELECTRICAL GENERATORS): (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2015_2016, the lowest value being in the year 2014_15. The FDI seems to have a sudden rise in the last recent year



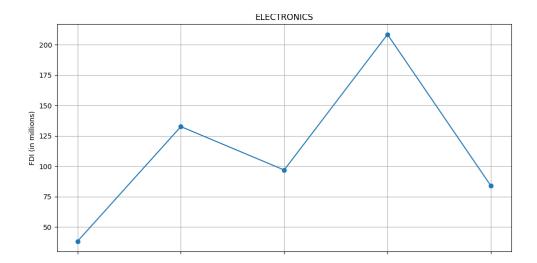
ELECTRICAL EQUIPMENTS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2015_2016, the lowest value being in the year 2012_13. The FDI seems to have been increasing over the past recent years.



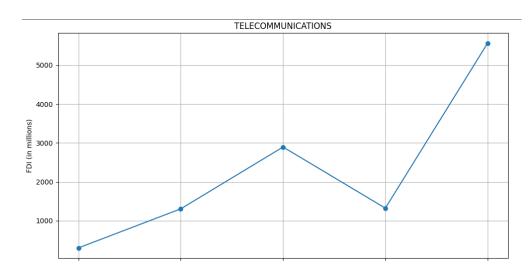
COMPUTER SOFTWARE HARDWARE (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest value being in the year 2012_13.



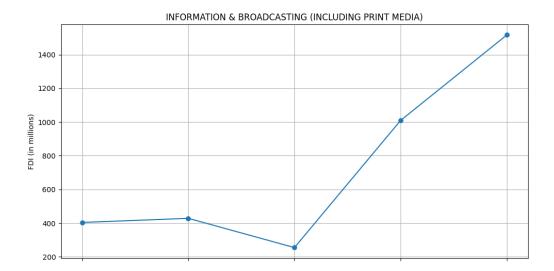
ELECTRONICS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest value being in the year 2012_13.



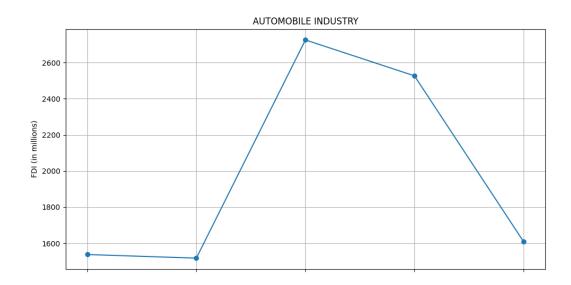
TELECOMMUNICATIONS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2015_2016, the lowest value being in the year 2012_13. The FDI seems to have been increasing over the past recent years.



INFORMATION & BROADCASTING (INCLUDING PRINT MEDIA) (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2015_2016, the lowest value being in the year 2014_15. The FDI seems to have been increasing over the past three recent years.



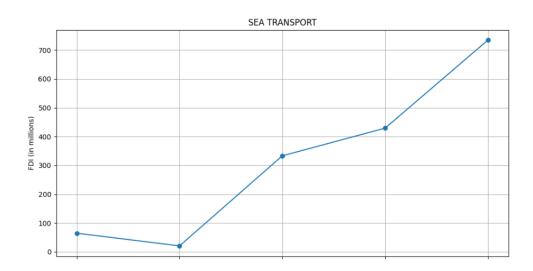
AUTOMOBILE (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest value being in the year 2013_14. The FDI seems to have been decreasing over the past three recent years.



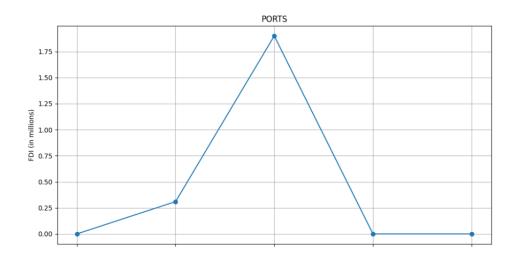
AIR TRANSPORT(INCLUDING AIR FREIGHT) (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest value being in the year 2012_13.



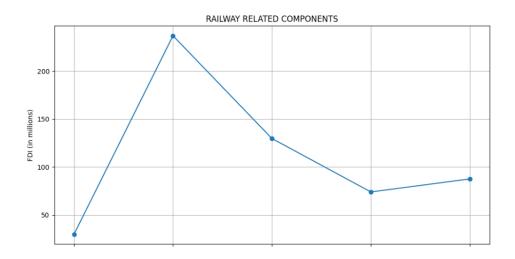
SEA TRANSPORT(INCLUDING AIR FREIGHT) (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2015_2016, the lowest value being in the year 2013_14.



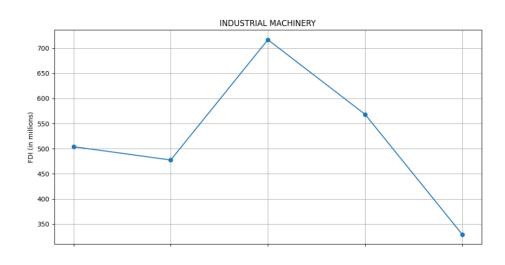
PORTS(2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest values being 0 in the year 2012_13, 2014_15 and 2015_16.



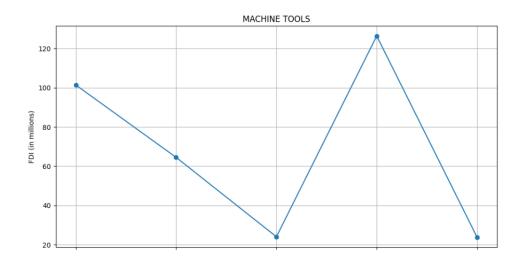
RAILWAY RELATED COMPONENETS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2013_2014, the lowest being in the year 2012_13



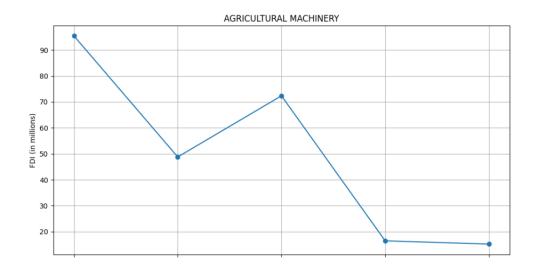
INDUSTRIAL MACHINERY (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest being in the year 2015_16



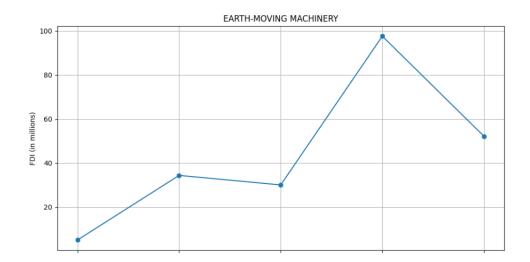
MACHINE TOOLS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest being in the year 2015_16



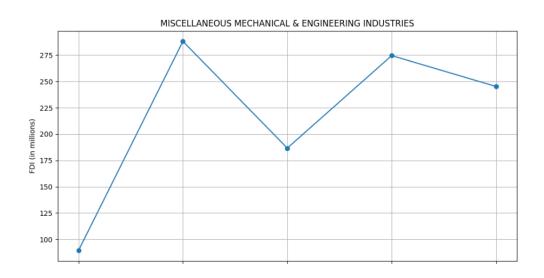
AGRICULTURAL MACHINERY (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2012_2013, the lowest being in the year 2015_16. The trend is seen to be continuously diminishing.



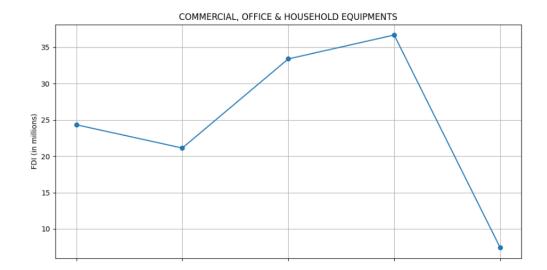
EARTH MOVING MACHINERY (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest being in the year 2012_13.



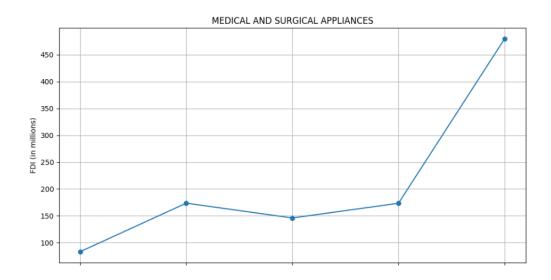
MISCELLAEANOUS MECHANICAL& ENGINEERING INDUSTRIES (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2013_2014, the lowest being in the year 2012_13.



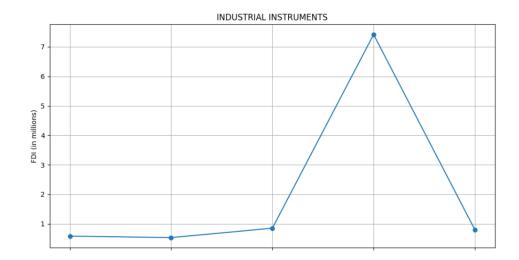
COMMERCIAL, OFFICE & HOUSEHOLD EQUIPMENTS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest being in the year 2015_16.



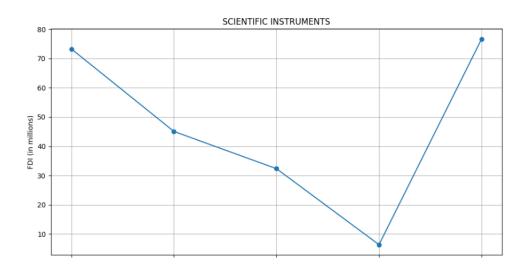
COMMERCIAL, OFFICE & HOUSEHOLD EQUIPEMENTS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2015_2016, the lowest being in the year 2012_13.



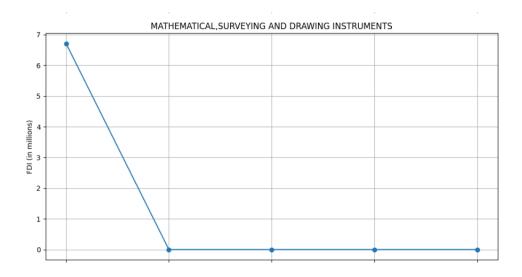
INDUSTRIAL INSTRUMENTS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest being in the year 2013_14.



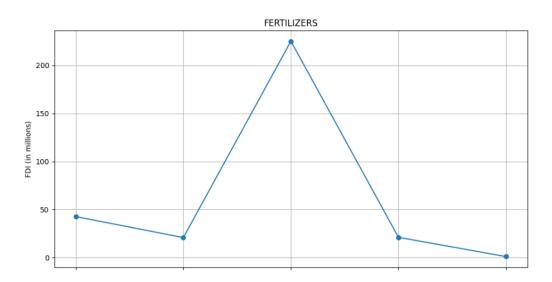
SCIENTIFIC INSTRUMENTS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest being in the year 2013_14.



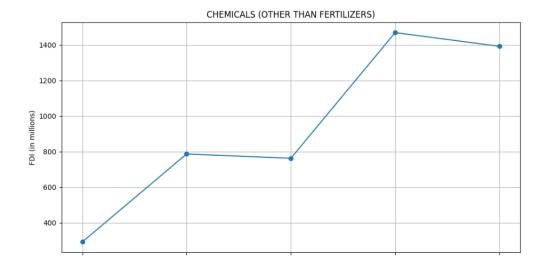
MATHEMATICAL SURVEYING AND DRAWING INSTRUMENTS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest and stagnant in the years 2013_14, 2014_15, 2015_16 at '0'



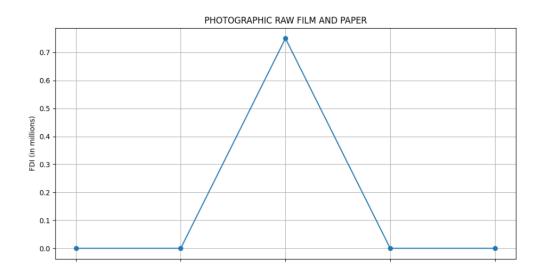
FERTILIZERS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2013_2014, the lowest in the year 2015_16 at '0'



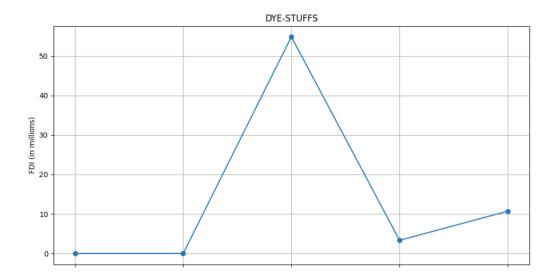
CHEMICALS(OTHER THAN FERTILIZERS) (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest in the year 2012_13



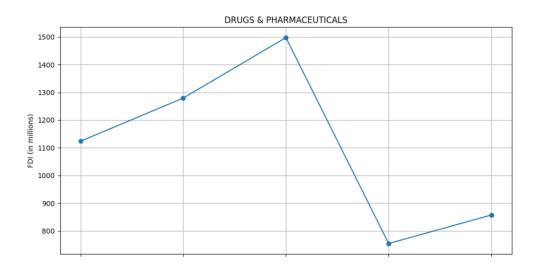
PHOTOGRAPHIC RAW FILM AND PAPER (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest in the years 2012_13, 2013_14, 2015_16 at '0'. The peak can only be seen in the year 2014_15.



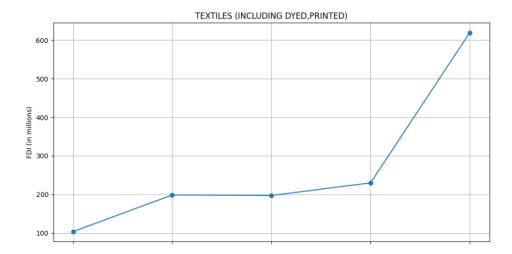
DYE STUFF (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest in the years 2012_13, 2013_14. The sudden peak can be seen in the year 2014_15.



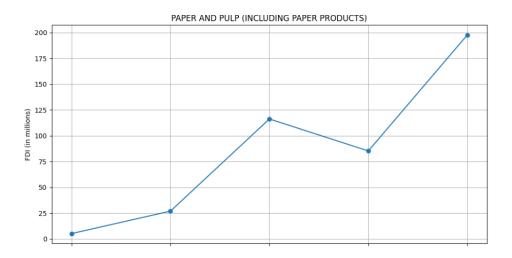
DRUGS & PHARMACEUTICALS (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2014_2015, the lowest in the year 2014_15.



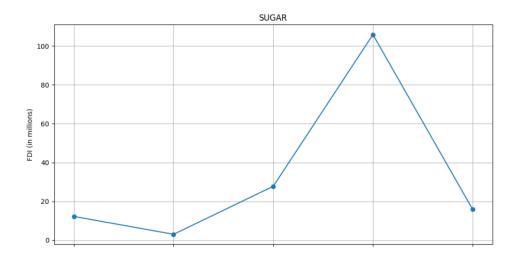
TEXTILES (INCLUDIN DYED AND PRINTED) (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2015_2016, the lowest in the year 2012_13. The trend is upward in the recent years.



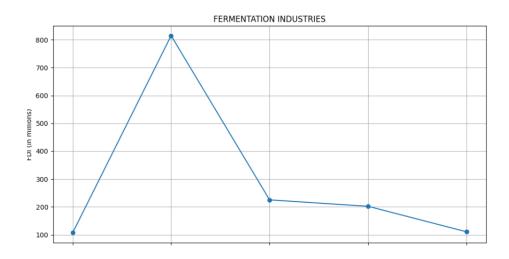
PAPER AND PULP (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2015_2016, the lowest in the year 2012_13. The trend is upward in the recent years.



SUGAR (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2013_2014, the lowest in the year 2014_15.



SUGAR (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2013_2014, the lowest in the year 2012_13 and 2015_16.



FOOD PROCESSING INDUSTRIES (2012, 2013, 2014, 2015, 2016): The FDI has seen a peak in the year 2013_2014, the lowest in the year 2012_13 & **2015_16**

