**PROJECT REPORT ON**

**A STATISTICAL ANALYSIS ON IMPORT AND EXPORT**



**DEPARTMENT OF STATISTICS**

**FACULTY OF SCIENCE**

**THE MAHARAJA SAYAJIRAO UNIVERSITY OF BARODA**

**PRESENTED BY: GUIDED BY:**

GHODASARA KRIYANSHI Dr. DEEPA KANDAPAL

ROHIT SHARMA

VAISHNAV SRIVASTAVA

DEEP SEKH

KIRAN KHATIK

NEEHAR SINGH

**ACKNOWLEDGEMENT**

We would like to express our profound gratitude to Dr. Deepa Kandpal for her constant support, help and encouragement throughout the project work. She helped us in selecting the data and the methods and prediction we can work on.

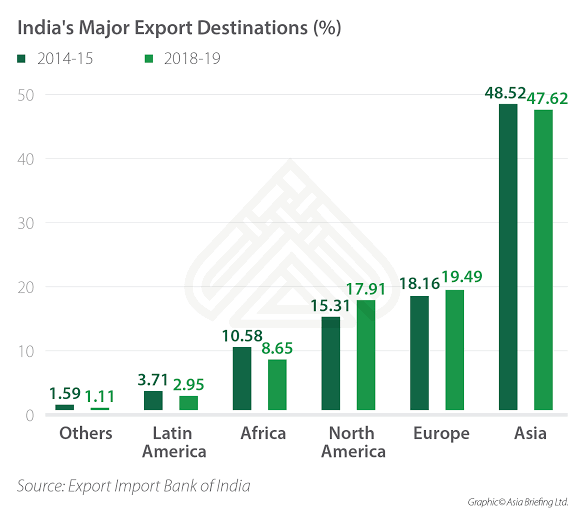
We are highly indebted to our head: Prof. Vipul A. Kalamkar for his guidance and support as well as for providing necessary information regarding the project and helpful feedback.

Our thanks and appreciation also go to our colleagues in developing the project and people who have willingly helped us out with our abilities. Even the other faculty member who guided us in the project work.

**INDEX**

1. INTODUCTION
2. OBJECTIVES
3. DESCRIPTIVE TECHNIQUES
4. CHART REPRESENTATION
5. METHOD TO ANALYSE THE FOLLOWING DATA
6. METHOD OF LEAST SQUARE
7. ARIMA MODEL
8. CONCLUSION

**INTRODUCTION**



Exports are the goods and services that are produced domestically, but then sold to customers residing in other countries. It leads to an inflow of funds to the seller’s country since export transactions involve selling domestic goods and services to foreign buyers.

Imports are the goods and services that are purchased from the rest of the world by a country’s residents, rather than buying domestically produced items. It leads to an outflow of funds from the country since import transactions involve payments to sellers residing in another country.

**What is Gross Domestic Product?**

Gross Domestic Product is the gross market value of the total goods and services produced within the domestic boundaries of a country during the given period of the time. It is also known as National Income. Total imports and exports are essential components of a country’s GDP. They are taken into account as “Net Exports”.

There are some ways by which a nation can decrease imports or increase exports:

1. ***Taxes and quotas***-Government decrease excessive import activity by imposing tariffs and quotas on imports. The tariffs make importing goods and services more expensive than purchasing them domestically. Imposing tariffs is one way a country can work to improve its balance of trade.
2. ***Subsidies***-Governments provide subsidies to domestic business in order to reduce their business costs. This helps bring down the price of domestic goods and services, hopefully, encouraging consumers to buy domestic rather than imported goods. By enabling domestic producers to produce goods less expensively and, thus, lower their prices, subsidies may also increase exports as the cheaper goods become more attractive to foreign buyers.
3. ***Trade agreements***-It typically focus on the exchange of different types of products. High volume of both imports and exports, by entering into a trade agreement with another country. Such agreement is aimed at stimulating trade and supporting economic growth for both countries involved.
4. ***Currency devaluation***- Increasing exports and decreasing imports is by devaluing the domestic currency. Governments devalue their currency with the aim of bringing down the prices of domestic goods and services, the ultimate goal being to increase net exports.

**Importance of Imports and Exports:**

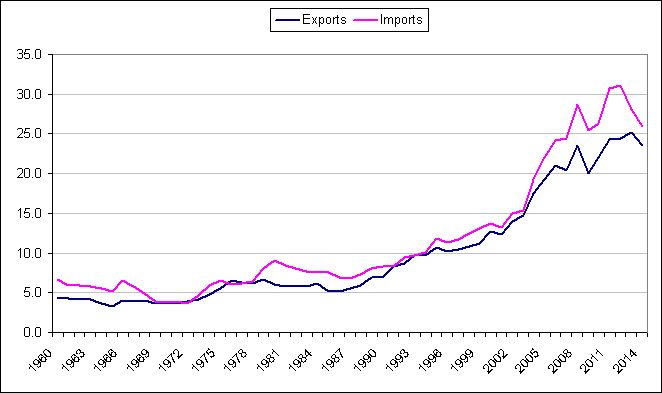
Countries vary considerably with regard to how important imports and exports and their overall balance of trade is to their economies. Maintaining a high level of exports is also very important to the economies of the U.K. and Australia. The growth of economies of developing countries is often fueled by massive exports of commodities and raw materials to developed nations. For this reason, mining is commonly a key industry in such countries.

India’s major exports included petroleum products (9201.76 million US$), gems and jewelry (3236.98 million US$), drug formulation (2054.24 million US$), engineering goods (9201.76 million US$), organic and inorganic chemicals (2447.89 million US$), cotton yarn (1389.46 million US$), electronic goods (1364.63 million US$), plastic and linoleum (844.47 million US$), rice (813.75 million US$) in January 2022. Exports from India increased 24.2% year-on-year to $38.19 billion in April of 2022.India’s merchandise export in January 2022 was USD 34.06 billion, an increase of nearly 24% over USD 27.54 billion in January 2021 and an increase of 31.75% over USD 25.85 billion in January 2020. The top 5 destination of India’s export are US, UAE, Bangladesh and Hong Kong.

While on the other hand the major imported products in India for January 2022 are Crude and petroleum products (11431.59 million US$), gold(2404.19 million US$), pearls and precious stones(2332.04 million US $), artificial resins(1762.08 million US$), vegetable oil(1858.58 million US$) and non-ferrous metals(1831.41 million US$).The largest share of India’s import came from China accounting for over 16 percent in the year 2021.The major countries from which we import are China, Saudi Arabia, Switzerland, Iraq and South Korea.

The government of India levies a customs duty on all the imports within and some of the exports from the country. The amount to be paid as customs duty can be determined by several factors such as value, weight, dimension and many more. The duty varies from not only form product to product but also from the bonding with the other nation. Custom duty has been replaced by the Goods and Services Tax (GST) from 1 July 2017. The customs duty is imposed by the Government of India under the Indian Customs Act formulated in 1962 under the Constitution of India under Article 265.

**IMPORT AND EXPORT WITH INDIA IN US$ million FROM 1980 TO 2014**



Thus, it can be observed that after independence till 2014 how the import as well as export increases. By this even the other field has improved a lot and thus India has been developing at a great pace and the quality of citizens is also improving.

**OBJECTIVES**

1. Exploratory data analysis of import and export in India from 2011 to 2021.
2. To obtain the simple descriptive measures of main properties of time series of the product.
3. To understand the deeper mechanism and interrelation amongst country, export and import as well as the commodities.
4. To analyze that whether the process is in control or not.

**DESCRIPTIVE TECHNIQUES:**

Traditional methods of Time series analysis are mainly concerned with decomposing the variation in a series into: Trend, Seasonal variation, Cyclic variation and Random fluctuations. This approach is NOT always the best but particularly useful when the variations are dominated by Trend and/or Seasonality.

1. ***Seasonal effect***- Many time series such as sales of rice/umbrella/ice-cream or temperature reading which are annual in nature during the period. This yearly variation is easy to understand, it can be measured explicitly and removed from the data to deseasonalize the data.
2. ***Cyclic changes***-Some time series exhibit variation at a fixed period due to some other physical cause.in addition to this some data is oscillating which do not have fixed period but can be predicted to some extent.
3. ***Trend***-May be loosely defined as long term change in mean level, a difficulty may arise in defining the series. We must take into account the number of observation available and make a subjective assembly which is long term.
4. ***Irregular fluctuation***-After removing trend and cyclic variation have been removed from a set of data, we are left with a series of residuals which may or may not be random.

***Stationary Time series***: A time series is said to be stationary if there is no systematic change in mean or variance and strictly periodic variation have been removed.

**TIME SERIES EXTRA DETAILS**

A Time series is a data set that tracks a sample over time. In particular, a time series allow one to see what factors influence certain variables from period to period. Time series analysis can be useful to see how a given assets, security or economic variable change over time.

BASIC TYPES OF ‘TIME DOMAIN’ MODELS: The analysis of mathematical functions, physical signals or time series of economic or environmental data with respect to time.

1. Models that relate the present value of a series to past values and past prediction errors- that are called ARIMA models (for Autoregressive integrated moving average).
2. Ordinary regression models that use time indices as x- variable.

LIMITATION OF TIME SERIES:

Time series analysis suffers from a number of weaknesses, including problems with generalization from a single study, difficulty in obtaining appropriate measures and problem with accurately identifying the correct model to represent the data.

**GRAPHICAL REPRESENTATION OF THE DATA**

1. COMPARING RICE AND AGRO-CHEMICALS EXPORT FROM 2012 TO 2021 IN (US$ million)

|  |  |  |
| --- | --- | --- |
| YEAR | RICE EXPORT | AGRO EXPORTS |
| 2012 | 3,216.98 | 1427.86 |
| 2013 | 3564.04 | 1740.07 |
| 2014 | 4864.69 | 1922.23 |
| 2015 | 4516.275453 | 1951.767036 |
| 2016 | 3477.980959 | 1965.705797 |
| 2017 | 3222.347293 | 2146.913949 |
| 2018 | 4169.52 | 2558.99 |
| 2019 | 4712.443 | 3156.573 |
| 2020 | 4372.11 | 3350.35 |
| 2021 | 4018.4 | 3578.34 |

***Conclusion***- Rice is exported more as compared to Agro-chemicals and the export of it are increasing yearly. Seasonal effect is observed in export of rice as production of rice is more in the particular season.

**2**. **YEARLY IMPORT AND EXPORT OF ALL THE COMMODITIES**

|  |  |  |
| --- | --- | --- |
| YEAR | EXPORT | IMPORT |
| 2012 | 305963.92 | 4,89,319.49 |
| 2013 | 300400.58 | 4,90,736.65 |
| 2014 | 314405.3 | 4,50,199.79 |
| 2015 | 310352.0092 | 448033.4087 |
| 2016 | 262290.1254 | 381006.6293 |
| 2017 | 276547.0012 | 382740.9433 |
| 2018 | 607046.38 | 465579.74 |
| 2019 | 660155.996 | 514078.371 |
| 2020 | 626729.47 | 474707.52 |
| 2021 | 291808.48 | 3,94,435.88 |

***CONCLUSION***-The import in 2012 and 2020 remained nearly same, while the export gradually increased from 2017 till 2020 and drastically decreased due to pandemic situation.

***3.PERCENTAGE WISE EXPORT OF FOUR PARTICULAR PRODUCT FROM 2015-2020***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| PRODUCT | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
| RICE | 62.67 | 55.94 | 52.31 | 54.52 | 53.726 | 50.82 |
| AGRO CHEMICALS | 9.46 | 31.62 | 34.855 | 33.463 | 35.9877 | 38.95 |
| MICA | 0.77 | 0.0084 | 0.00908 | 1.06392 | 0.00812 | 0.611 |
| TEA | 27.08 | 11.582 | 11.9206 | 10.9492 | 9.4733 | 9.6 |

***CONCLUSION***-Mica is exported is negligible amount and maximum exported is rice as compared to four products.

***4.COMPARING EXPORT OF ALL COMMODITIES OF CHINA AND HONG KONG FROM 2012 TO 2020***

|  |  |  |
| --- | --- | --- |
| YEAR | CHINA | HONGKONG |
| 2012 | 18,076.55 | 12,931.90 |
| 2013 | 13,534.88 | 12,279.20 |
| 2014 | 14,824.36 | 12,731.74 |
| 2015 | 11,934.25 | 13,599.88 |
| 2016 | 9013.537495 | 12092.21323 |
| 2017 | 10203.30233 | 14157.17466 |
| 2018 | 13334.43 | 14690.28 |
| 2019 | 16752.794 | 13001.983 |
| 2020 | 16,612.75 | 10,967.12 |
| 2021 | 21,187.15 | 10,162.40 |

***CONCLUSION***-India exported highest to China in 2021 while to Hongkong the export was nearly half as compared to that of the China.

1. ***COMPARING EXPORT OF RICE FOR USA AND TURKEY FROM 2015 TO 2020***

|  |  |  |
| --- | --- | --- |
| YEAR | USA | TURKEY |
| 2015 | 132.074678 | 24.253656 |
| 2016 | 131.552178 | 30.393395 |
| 2017 | 102.389198 | 45.53697 |
| 2018 | 149.01 | 51.64 |
| 2019 | 168.736 | 40.82 |
| 2020 | 181.26 | 33.26 |

***CONCLUSION***- As compared to Turkey, USA has been exported higher amount of rice in all the years.

1. ***OBSERVING TOTAL EXPORT WITH TOTAL(IMPORT+EXPORT)***

|  |  |  |
| --- | --- | --- |
| YEAR | EXPORT | TOTAL |
| 2012 | 305963.92 | 795283.41 |
| 2013 | 300400.58 | 791137.23 |
| 2014 | 314405.3 | 764605.09 |
| 2015 | 310352.0062 | 758385.418 |
| 2016 | 262290.1254 | 643296.7547 |
| 2017 | 276547.0012 | 659287.9445 |
| 2018 | 607046.38 | 1072626.12 |
| 2019 | 660155.996 | 1174234.367 |
| 2020 | 626729.47 | 1101436.99 |
| 2021 | 291808.48 | 686244.36 |

***CONCLUSION***-Export was maximum in the year 2019 that is before covid and in 2021 several factors affected the export to nation and thus we had to suffer great economic loss as compared to previous years. It has been observed that only in the year the import was also less.

1. ***FOUR COUNTRIES IMPORT OF CRUDE AND GOLD FROM 2011-2021***

|  |  |
| --- | --- |
| Countries | Total Import |
| SAUDI ARAB | 200306.5532 |
| IRAQ | 168312.5109 |
| NIGERIA | 98680.6711 |
| U ARAB EMTS | 93377.9198 |
| IRAN | 71981.50098 |

***CONCLUSION***-India imported maximum crude from Saudi Arab and minimum from Iran because Arab countries have highest mine of it.

|  |  |
| --- | --- |
| Countries | Total |
| SWITZERLAND | 193327.4697 |
| U ARAB EMTS | 49618.91095 |
| GHANA | 13339.99142 |
| U S A | 18620.02942 |
| SOUTH AFRICA | 26587.09659 |

***CONCLUSION***-Maximum gold was imported from Switzerland and least from Ghana, it may be because of bonding amongst the country and tax laid on it.

1. ***COMPAQRING TOTAL IMPORT OF CRUDE AND GOLD FROM 2011 TO 2021***

|  |  |  |
| --- | --- | --- |
| YEAR | GOLD IMPORT | CRUDE IMPORT |
| 2011-12 | 56501.71 | 134154.97 |
| 2012-13 | 53820.63 | 144519.72 |
| 2013-14 | 28704.12 | 143638.53 |
| 2014-15 | 34407.17 | 116442.8567 |
| 2015-16 | 31770.74 | 65922.9837 |
| 2016-17 | 27509.77 | 70716.92198 |
| 2017-18 | 33657.20 | 87372.23 |
| 2018-19 | 32910.06 | 114042.113 |
| 2019-20 | 28229.67 | 102748.73 |
| 2020-21 | 34603.93 | 59477.65 |

***CONCLUSION***-From 2012 till 2021 the amount of crude India import is nearly double and more than gold and the price increment in gold was one of the factors purchase of gold decline in rapid manner as compared in 2012 and 2013.

1. ***MARKET PERCENTAGE OF GOLD AND CRUDE OF PARTICULAR COUNTRIES***

|  |  |
| --- | --- |
| Countries | MARKET % |
| SWITZERLAND | 53.388417 |
| U ARAB EMTS | 13.702528 |
| GHANA | 3.6839101 |
| U S A | 5.1420209 |
| SOUTH AFRICA | 7.3421692 |
| OTHER | 16.740954 |

***CONCLUSION***-Half of the Gold imported in India comes from Switzerland that is 53%.

|  |  |
| --- | --- |
| Countries | market % |
| SAUDI ARAB | 19.2781 |
| IRAQ | 16.1989 |
| NIGERIA | 9.497323 |
| U ARAB EMTS | 8.98697 |
| IRAN | 6.927715 |
| other | 39.11099 |

***CONCLUSION***-All the countries shared random number of crude imports for India.

**ANALYSING THE DATA BY DIFFERENT METHODS**

1. ***METHOD OF LEAST SQUARE***

* FITTING OF EXPONENTIAL TREND

The general equation for an exponential trend is Y=a.b t

Here ‘a’ and ‘b’ are constants that needs to determine from the data.

Taking log on both the sides we get log Y= log a + log b, this is the linear equation in log Y.

Let A=log a and B= log b thus by substituting it in the equation we get log Y= A +Bt.

By selecting a suitable origin such that its summation will be equal to 0.

Now Y=anti(A+BX)

=anti (log a + X log b)

=a.b.x



We have taken A=2017

Now lny = lna + xlnb

Let, U=lny, A=lna, B=lnb

U=nA + B and

n=6

=77.71769/6

=12.9524

=42.3935/19

=2.2312

Now lna=A

a=antilog(A)=422075.7615

b=antilog(B)=9.3113

***FORECASTING***-The trend line shows that how drastically the export has been increased in in 2019-2020. Thus, by using this method we can calculate the export number of upcoming years. For example, in the year 2025 the export will be

**ARIMA MODEL**

An autoregressive integrated moving average-ARIMA model is if aa generalization of simple autoregressive moving average-ARMA model. Both of these models are used to forecast or predict future points in the time series data. It indicates the strength of a dependent variable relative to other changing variables.

The final objective of the model is to predict future time series movement by examining the differences between values in the series instead through actual values. They are applied in the cases where the data show evidence of non-stationarity. In time series analysis, non-stationary data are always transformed into stationary data.

The general representation of the ARIMA model is (p,d,q) and parameter p,d,q are defined as follow:

p: the lag order or the number of time lag of autoregressive model AR(p)

d: degree of differencing or the number of times the data have had subtracted with past value

q: the order of moving average model MA(q)

***ADVANTAGES OF ARIMA MODEL:***

1. If one is forecasting a large number of time series.
2. Avoids a problem that occurs sometimes with multivariate model

***DISADVANTAGES OF ARIMA MODEL:***

1. It is computationally expensive.
2. Poorer performance for long term forecast.
3. Cannot be used for seasonal time series.
4. Less explainable than exponential smoothing.

**METHOD TO OBTAIN ARIMA MODEL**

**Step 1:** Plot the data as time series.

**Step 2:** Difference data to make data stationary on mean (remove trend)

**Step 3:** Log transform data to make data stationary on variance

**Step 4:** Difference log transform data to make data stationary on both mean and variance

**Step 5:** Plot ACF and PACF to identify potential AR and MA model

**Step 6:** Identification of best fit of ARIMA model

**Step 7:** Forecast using the best ARIMA model

**Step 8:** Plot ACF and PACF for residuals of ARIMA model to ensure no more information is left for extraction

ARIMA MODELLING ON IMPORT DATA(1997-2021) AND FORECASTING FOR NEXT FIVE YEAR(USING R)

##CODE

> data=read.csv("data2.csv")

> data

year TOTAL.IMPORT

1 1997 39132.41

2 1998 41484.49

3 1999 42388.71

4 2000 49738.06

5 2001 50536.45

6 2002 51413.28

7 2003 61412.14

8 2004 78149.11

9 2005 111517.43

10 2006 149165.73

11 2007 185735.24

12 2008 251654.01

13 2009 303696.31

14 2010 288372.88

15 2011 369769.13

16 2012 489319.49

17 2013 490736.65

18 2014 450199.79

19 2015 448033.41

20 2016 381006.63

21 2017 382740.94

22 2018 465579.74

23 2019 514078.37

24 2020 474707.52

25 2021 394435.88

> d1 = ts(data$TOTAL.IMPORT, start = min(data$year), end = max(data$year), frequency = 1)

> class(d1)

[1] "ts"

> library(forecast)

> library(tseries)

> d2 = log(d1)

> adf.test(d2)

Augmented Dickey-Fuller Test

data: d2

Dickey-Fuller = -0.11354, Lag order = 2, p-value = 0.99

alternative hypothesis: stationary

Warning message:

In adf.test(d2) : p-value greater than printed p-value

> plot(d2)

Chart, line chart

Description automatically generated

> import\_model = auto.arima(d2, ic="aic", trace = TRUE)

ARIMA(2,1,2) : -21.61764

ARIMA(0,1,0) : -18.53403

ARIMA(1,1,0) : -23.64073

ARIMA(0,1,1) : -27.2256

ARIMA(0,1,0) : -12.37667

ARIMA(1,1,1) : -25.32579

ARIMA(0,1,2) : -25.32837

ARIMA(1,1,2) : -23.57396

ARIMA(0,1,1) : -24.74373

Best model: ARIMA(0,1,1)

> acf(import\_model$residuals)

Chart, timeline

Description automatically generated

> pacf(import\_model$residuals)

Timeline

Description automatically generated

> forecast\_import = forecast(import\_model, level = c(95), h=5\*1)

> forecast\_import

Point Forecast Lo 95 Hi 95

2022 12.85880 12.61488 13.10273

2023 12.95160 12.46265 13.44054

2024 13.04439 12.39737 13.69141

2025 13.13718 12.36375 13.91062

2026 13.22998 12.34806 14.11189

> plot(forecast\_import)

Chart, line chart, histogram

Description automatically generated

ARIMA MODELLING ON IMPORT DATA(1997-2021) AND FORECASTING FOR NEXT FIVE YEAR(USING R)

##CODE

> data=read.csv("data3.csv")

> data

ï..year TOTAL.EXPORT

1 1997 33469.95

2 1998 34784.98

3 1999 33218.72

4 2000 36822.49

5 2001 44560.29

6 2002 43826.72

7 2003 52719.43

8 2004 63842.55

9 2005 83535.94

10 2006 103090.53

11 2007 126414.05

12 2008 136132.18

13 2009 185295.36

14 2010 178751.42

15 2011 249815.65

16 2012 305963.92

17 2013 300400.58

18 2014 314405.30

19 2015 310352.01

20 2016 262290.13

21 2017 276547.00

22 2018 607046.38

23 2019 660156.00

24 2020 626729.47

25 2021 291808.48

> d = ts(data$TOTAL.EXPORT, start = min(data$ï..year), end = max(data$ï..year), frequency = 1)

> class(d1)

[1] "ts"

> library(forecast)

> library(tseries)

> d\_new = log(d)

> adf.test(d\_new)

Augmented Dickey-Fuller Test

data: d\_new

Dickey-Fuller = -1.5875, Lag order = 2, p-value = 0.7295

alternative hypothesis: stationary

> plot(d\_new)

Chart, line chart

Description automatically generated

> import\_model = auto.arima(d\_new, ic="aic", trace = TRUE)

ARIMA(2,1,2) : Inf

ARIMA(0,1,0) : 6.89411

ARIMA(1,1,0) : 8.566424

ARIMA(0,1,1) : 8.504826

ARIMA(0,1,0) : 7.683483

ARIMA(1,1,1) : 10.50299

Best model: ARIMA(0,1,0)

> acf(import\_model$residuals) Chart, timeline

Description automatically generated

> pacf(import\_model$residuals) Chart, timeline

Description automatically generated

> forecast\_import = forecast(import\_model, level = c(95), h=5\*1)

> forecast\_import

Point Forecast Lo 95 Hi 95

2022 12.67408 12.15950 13.18866

2023 12.76431 12.03658 13.49203

2024 12.85453 11.96326 13.74581

2025 12.94476 11.91560 13.97392

2026 13.03499 11.88435 14.18562

> plot(forecast\_import) Chart, line chart, histogram

Description automatically generated