WEST BENGAL STATE UNIVERSITY

SUBJECT-STATISTICS

PAPER-STSA PAPER VIII PROJECT

ROLL -3202307 NO - 24207

REGISTRATION NO-1081721400222 OF2017

Problems and Prospects of Tea Industry in India:(A Case study of various teas)

SUMMARY

this project is an analysis of the tea export and production in the world.

First there is a brief introduction about the Indian as well as world tea industry and production.

The second part contains the analysis part of the project analyzing the production scenario of india.here,we have two case studies,one of NORTH and other of SOUTH INDIA.

The third and fourth part consists of the discussion of results and problems of tea industry respectively.

The fifth and sixth part give the scenario of tea export to different parts of the world and the prospects of the tea industries respectively.

The seventh and eighth section respectively conclude the project and list the references of data used in the making the project.

AIM OF THE PROJECT

The project is about the Indian tea industry in general and the tea industry in West Bengal in particular.

- *It is based on systematic collection of data and information from the Tea Board of India (Kolkata), Tea Research Association (Dooars), Nya Sylee Tea Estate (Dooars), Happy Valley Tea Estate (Darjeeling) and Makai Bari Tea Estate (Kurseong).*
- *The project includes analysis of both production of Indian Tea.*
- *It also considered ESTE(Political Factors, Economical Factors. Social Factors, Technological Factors, Environmental Factors) analysis to understand the Indian Tea Industry comprehensively.*
- *The forecasting of production (million kg) through trend analysis method depicts encouraging future trend.*
- *Some problem are identified here and the prospects thus identified primarily based on the survey informations.*

1. INTRODUCTION:

*The Indian tea industry is nearly 200 years old. **Robert Bruce**, a British national discovered tea plants growing in the upper **Brahmaputra valley** in **Assam** and adjoining areas. In 1838, Indian tea that was grown in Assam was sent to the UK for the **first** time, for public sale.Tea in India is grown primarily in **Assam**, **West Bengal**, **Tamil Nadu** and **Kerala**. Apart from this, it is also grown in small quantities in **Karnataka**, **HP**, **Tripura**, **Uttaranchal**, **Arunachal Pradesh**, **Manipur**, **Sikkim** and **Meghalaya**. India has a dual tea base, unlike most other tea exporting countries.The tea industry provides direct employment to over 1 million persons. Through its forward and backward linkages another **10 million** persons derive their livelihood from tea. In Northeast India alone, the tea industry employs around **900,000** persons on permanent rolls. It is one of the largest employers of women amongst organized industries*

*In India, Women constitute nearly 51% of the total workforce. The tea estates in the **North Eastern India** are located in industrially backward areas.*

The three most distinct known varieties of tea in India are:

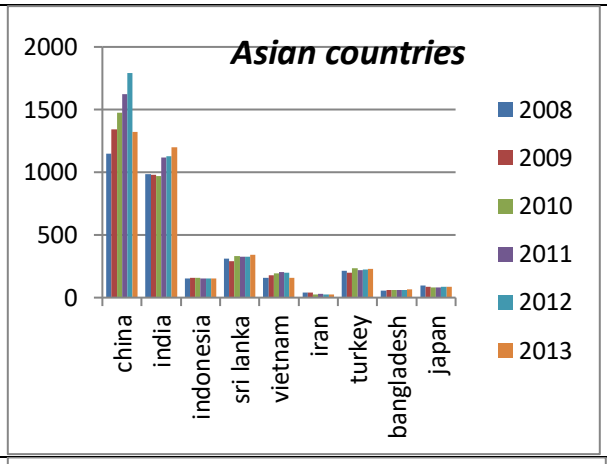
a) Assam tea (grown in Assam and other parts of NE India)

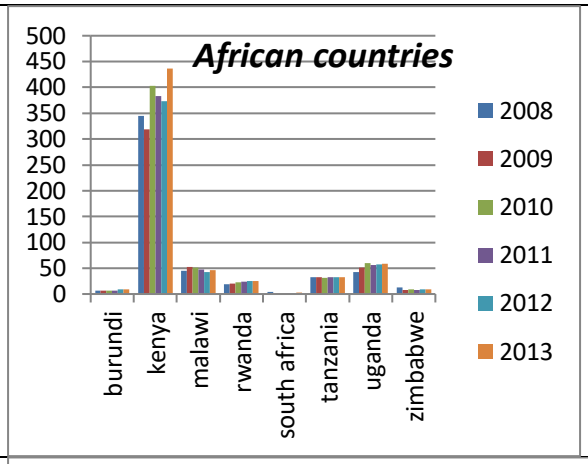
b) Darjeeling tea (grown in Darjeeling and other parts of West Bengal)

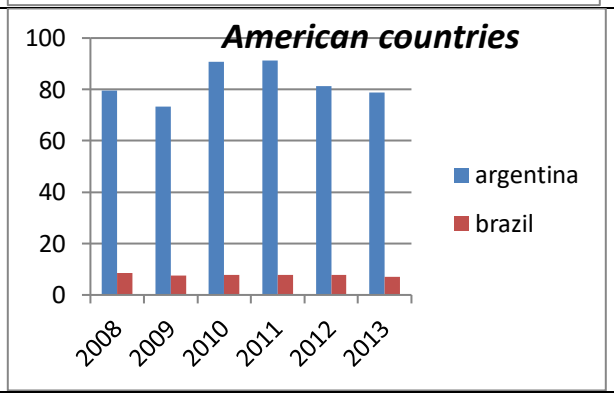
c) Nilgiri tea (grown in the Nilgiri hills of Tamil Nadu)

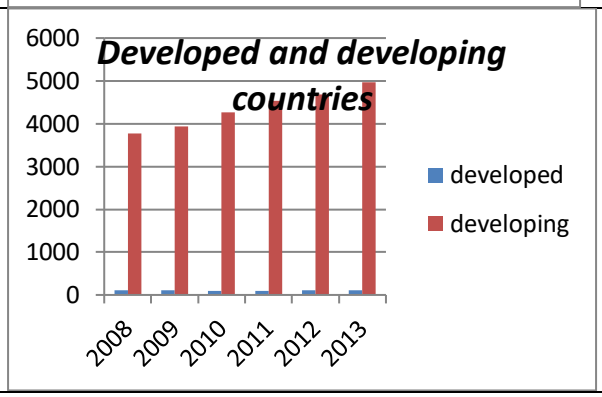
1.2. WORLD TEA PRODUCTION:

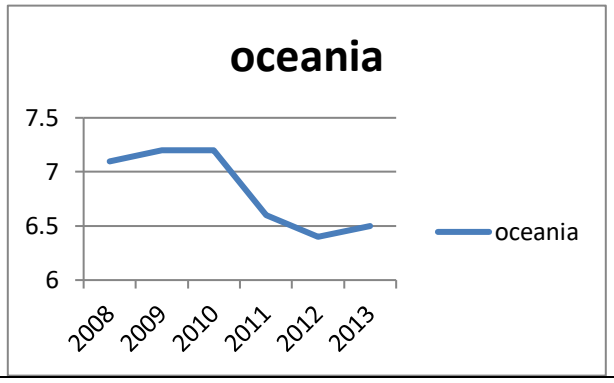
*India is the second largest producer of tea in the world, producing an average 900,000 tonnes each year.India produces large quantities, which is important because they're a nation of over **one billion** tea drinkers, with over **70%** of the tea produced in the nation being consumed within the nation instead of exported.*











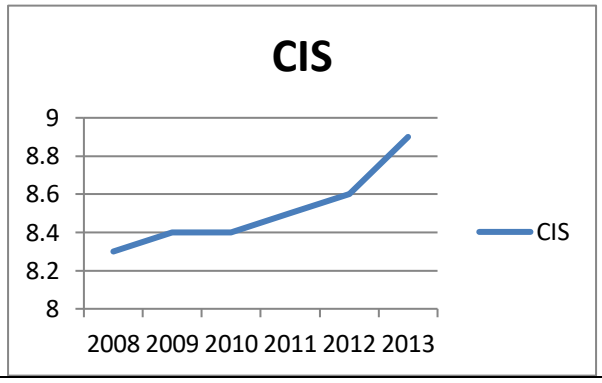


Figure1 world tea-production

1.3.INDIAN SCENARIO:

***India** is the **largest producer** and **consumer** of tea in the **world**. India also leads in global R&D in tea industry. India is the **largest manufacturer** and **exporter** of tea machinery. Other major tea producers (also developing nations) source equipment and technology from India.The annual per capita consumption in India is low at **650gm** compared to other countries like **Pakistan**(950gm),**Sri Lanka**(1.2kg),UK(2.5kg) and **Ireland** (3.16kg). In India, tea cultivation on commercial scale was first started in Assam in 1839 and then it was extended to other parts of the country between 50's and 60's of the 19thcentury.Due to certain specific soil and climatic requirements its cultivation was confined to only certain parts of the country*

2.TEA PRODUCTIVITY:

- *The Tea Board had released the **September** (2016) production data which shows that India **lost** 44.17 million kg (M.kg) over September 2016 to total 145.83 M.kg.*
- ***North India** lost 42.14 M.kg to produce 127.54 M.kg while **South India** lost 2.03 M.kg to produce 18.29 M.kg", Consequently, our compilation shows that at the end of the first three quarters of 2016, India's production has fallen to 912.55 M.kg from 923.27 M.kg, marking a marginal fall of 1.16 per cent".*
- *In the ten months, North India lost 25.79 M.kg to produce 737.87 M.kg. Assam **continued** to top India's production table despite a lower output this year. The state produced 464.29 M.kg – as much as 23.33 M.kg less than the relevant months of 2016.*
- ***Darjeeling** also reported a **lower** production – 2.07 M.kg against 6.88 M.kg last year.As of 2017, industry sources estimate the country's production in the current calendar to be around 1257 M.kg in 2017 – some 10 M.kg less than 2016.*
- *But in 2017 India's production had increased to 1,278.83 M.kg from 1,267.36 M.kg in 2016. This increase of 11.47 M.kg marks a marginal growth of 0.91 per cent. North Indian overall output in 2017 dropped to 1,046.42 M.kg from 1,054.51 M.kg due to adverse weather conditions in many months, marking a marginal loss of 0.77 per cent·*

Data table

	2014													
	January	February	March	April	May	June	July	August	Septembe	October	Novembe	Decembe	Jan	February
ASAM	1.51	2.51	3.51	4.51	5.51	6.51	7.51	8.51	9.51	10.51	11.51	12.51	1.18	0.2
WEST BEN	3.26	1.32	19.9	13.41	21.64	37.88	43.15	46.4	45.36	44.99	32.2	19.95	2.86	1.9
NORTH IN	4.96	1.79	36.13	42.08	69.29	113.43	140.97	149.84	137.54	144.64	82.76	41.77	4.18	2.2
TAMILNAI	11.78	9.99	12.86	11.82	16.81	21.41	18.06	10.5	13.91	17.09	13.8	11.76	11.37	9.5
KERALA	4.9	3.49	4.2	4.11	8.16	8.06	5.37	5.35	5.26	7.63	5.66	5.36	5.34	4.0
KARNATA	0.42	0.45	0.46	0.57	0.8	0.74	0.51	0.28	0.62	0.71	0.69	0.49	0.45	0.3
SOUTH INI	17.1	13.93	17.52	16.5	25.77	30.21	23.94	16.13	19.79	25.43	20.15	17.61	17.16	13.9

Table whose 41 columns have been omitted.

2.1. DATA ANALYSIS

*In this project,an attempt has been made to represent the production of tea by analyzing monthly tea production data of india from 2014 to 2019 and to forecast the tea production in the subsequent **three** years*

2.1.1.OBJECTIVES

The study and analysis has been conducted with a view to fulfill the following objectives:

- *To represent the linear trend of the tea production by analyzing monthly tea production data of the india from **2014** to **2019**.*
- *To **forecast** the tea production for the subsequent years of **2018,2019,2020**.*

2.1.2.METHODOLOGY

For analysing the tea production in india we collected the required dataset for tea production in all the states containing a part of tea garden in various areas of india from the year 2014 to2019. From within all these states,we select 5 states of geographical importance i.e containing a large area of the tea garden.

On plotting the time series plot of the 5 states we come to the conclusion that they can be broadly classified into catagories based on similarities:

- *GroupA: (assam,west Bengal)-north india*
- *GroupB:(tamilnadu,kerala,Karnataka)-south india*

For analysis,we select north india and south india as the representatives of GroupI and GroupII respectively.

2.2.CASE I:NORTH INDIA

First we provide a visual representation of the data obtained from NORTH INDIA in a time series plot.

>northindia <-read.table("d:/northindia.csv",header=TRUE)
>library("ggplot2")
>northindia1<-ts(northindia,frequency=12,start=c(2014,1))
>northindia1
>plot.ts(northindia1)

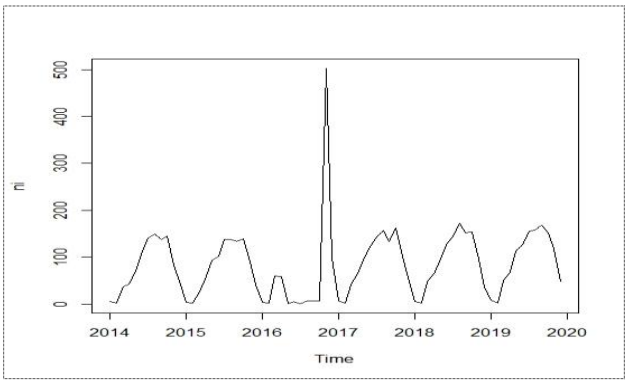


FIGURE2:TIME SERIES PLOT OF NORTH INDIA

From the plot we can see that the data might have seasonality with it.So we assume the additive model of the time series i.e.

$$Y_t=T_t+S_t+C_t+I_t$$

Where

Y_t = observations of the time series

T_t= the value attributable to secular trend
S_t= the value attributable to seasonal component
C_t= the value attributable to cyclic component
I_t=the value attributable to variation component
And decompose the additive time series into its trend component using R.

>northindiacomp<-decompose(northinda1)
>northindiacomp
>plot(northindiacomp)
>plot.ts(northindiacomp$trend)
>plot.ts(northindiacomp$seasonal)

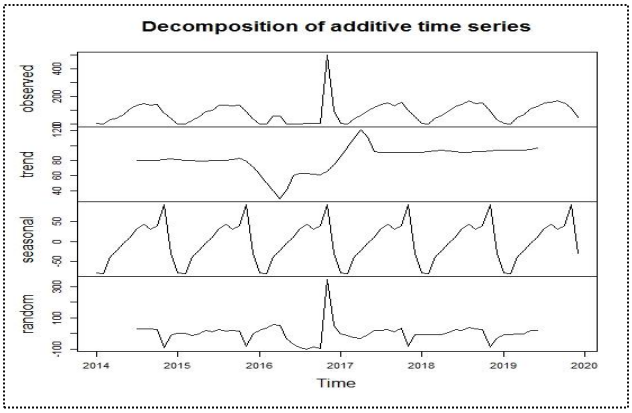


FIGURE2:DECOMPOSITION OF ADDITIVE TIME SERIES

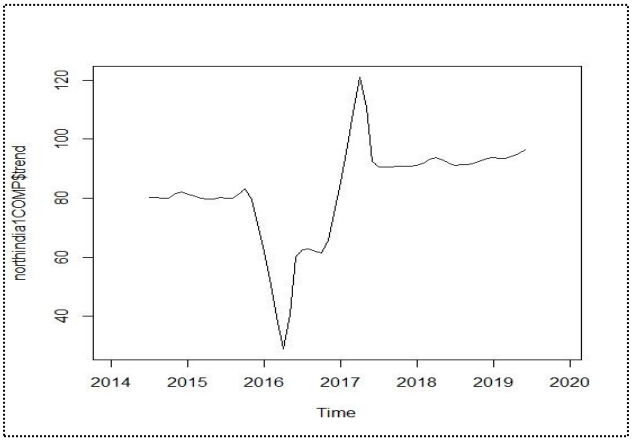


FIGURE4:DECOMPOSITION OF TIME SERIES IN ITS TREND COMPONENT

*Here we can see the trend variation of the no. of tea producory states in north india represented by the uniform crest and troughs over the years.*We can also **decompose** the time series into its seasonal trend.

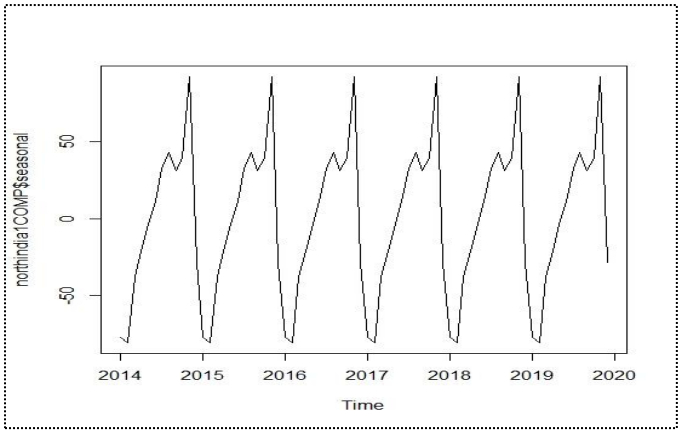


FIGURE5:DECOMPOSITION OF TIME SERIES IN ITS SEASONAL TREND

2.2.1.FITTING AND FORECASTING

Here,we fit the data on no. of tea producory states in north india from 2014 to 2019 into a suitable model using auto.arima() function in R.

*An **ARIMA** model is in the form of **ARIMA**(p,d,q)model,where,*

p=number of nonseasonal autoregressive terms
q=number of non seasonal moving average terms
d=number of non seasonal differences
Here the l ARIMA model used is (2,0,2)with non-zero mean.
***Coefficient**:*

$MA_1(\boldsymbol{\theta_1})$	-1.4353
$MA_2(\boldsymbol{\theta_2})$	0.6248
$AR_1(\boldsymbol{\varphi_1})$	1.6472
$AR_2(\boldsymbol{\varphi_2})$	-0.8965

MEAN=83.2553
AIC=817.13,AICc=818.42,BIC=830.79
*The model is so chosen by the auto.arima() function as it has the lowest AIC value.**[refer to appendix1]**.*
*The chosen model indicates that the data for northindia is seasonal and non-stationary.[**refer to appendix2**]*
Now the forecasting model is given by

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

Using which we forecast the no. of monthly tea production for the years 2020,2021,2022 in R.

```
>library("forecast")
>northindia2<-auto.arima(northindia1)
>northindia3<-forecast(northindia2,h=36)
>plot(northindia3)
```

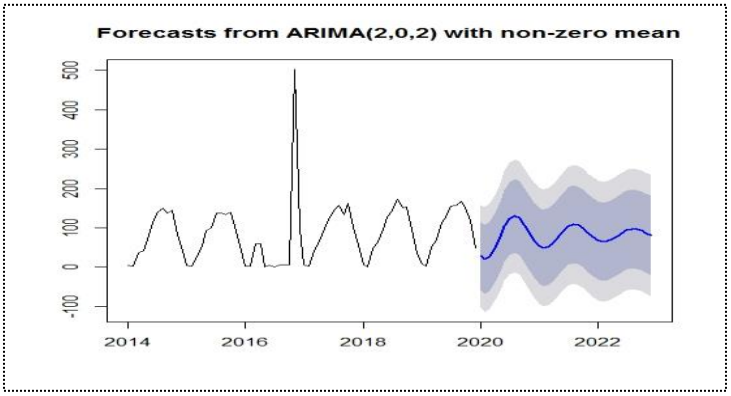


FIGURE6:FORECASTING

Here the forecast and 95% limits are given from for years of 2020,,2021, 2021.Hence we conclude the analysis of north india and generalize their conclusion for group A as a whole.

2.3.CASE II:SOUTH INDIA

First we provide a visual representation of the data obtained from SOUTH INDIA in a time series plot.

```
>southindia <-read.table("d:/southindia.csv",header=TRUE)
>library("ggplot2")
>southindia1<-ts(southindia,frequency=12,start=c(2014,1))
>southindia1
>plot.ts(southindia1)
```

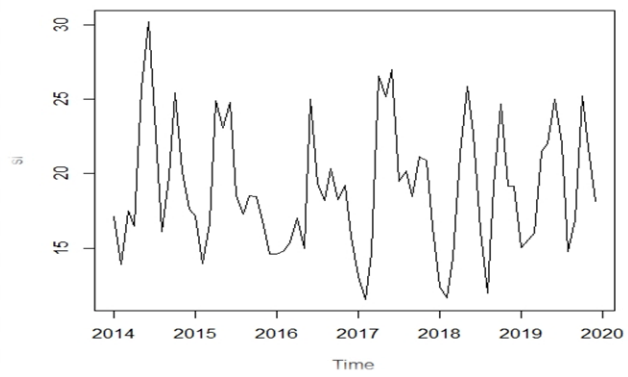


FIGURE7:TIME SERIES

From the plot we can see that the data might have seasonality with it.So we assume the additive model of the time series i.e.

$$Y_t=T_t+S_t+C_t+I_t$$

Where

Y_t = observations of the time series
T_t= the value attributable to secular trend
S_t= the value attributable to seasonal component
C_t= the value attributable to cyclic component
I_t=the value attributable to variation component
And decompose the additive time series into its trend component using R.

```
>southindiacomp<-decompose(southindia1)
>southindiacomp
>plot(southindiacomp)
>plot.ts(southindiacomp$trend)
>plot.ts(southindiacomp$seasonal)
```

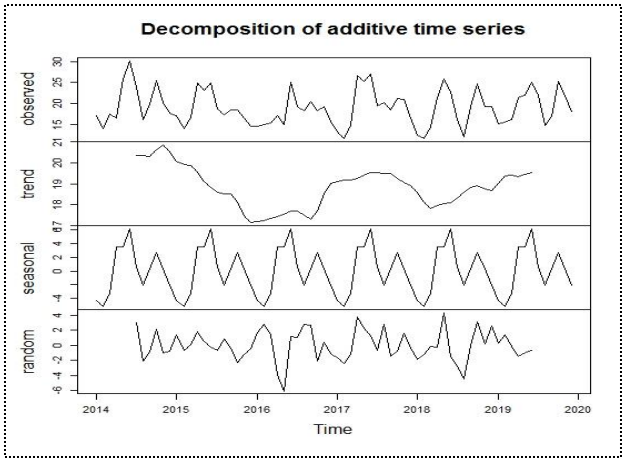


FIGURE8:DECOMPOSITION OF ADDITIVE TIME SERIES

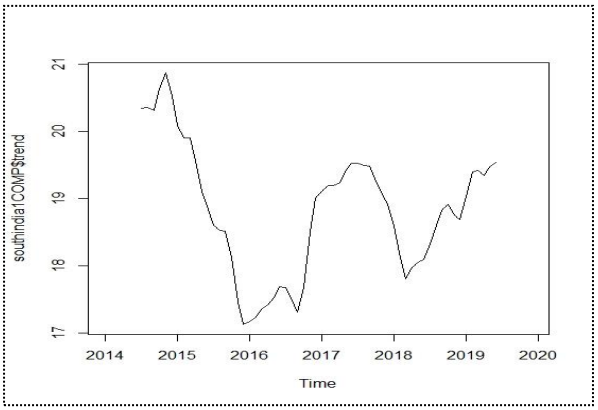


FIGURE9: DECOMPOSITION OFTIME SERIES INTO ITS TREND COMPONENT

*Here we can see the trend variation of the no. of tea producory states in north india represented by the uniform crest and troughs over the years.*We can also **decompose** the time series into its seasonal trend.

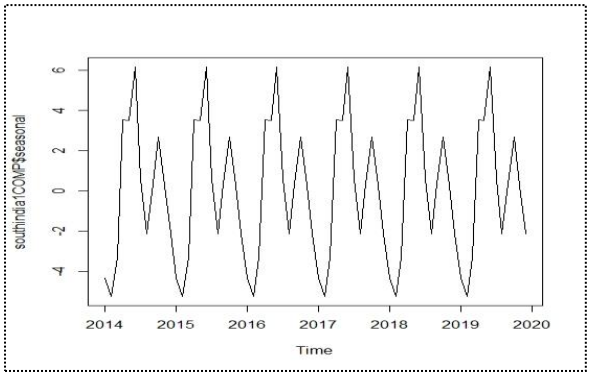


FIGURE10:DECOMPOSITION OF TIME SERIES INTO ITS SEASONAL TREND

2.3.1.FITTING AND FORECASTING

Here,we fit the data on no. of tea producory states in north india from 2014 to 2019 into a suitable model using AUTO. ARIMA model() function in R.
Here the SARIMA(seasonal ARIMA)model selected is (0,0,1),(2,1,0) [12]
*A **SARIMA** model is in the form of **ARIMA**(p,d,q)*(P,D,Q)model,where,*
*A **seasonal ARIMA** model is classified as an **ARIMA**(p,d.q)*(P,D,Q) model,where*
P=number of seasonal autoregressive(SAR)terms
D= numbers of seasonal differences
Q=number of seasonal moving averages(SMA)terms
p=number of nonseasonal autoregressive terms
q=number of non seasonal moving average terms
d=number of non seasonal differences

***Coefficient**:*

$MA_1(\boldsymbol{\theta_1})$	0.2910
$SAR1(\boldsymbol{\Phi_1})$	-0.5779
$SAR2(\boldsymbol{\Phi_2})$	-0.2642

AIC=314.27,AICc=315,BIC=322.65
*The model is so chosen by the auto.arima() function as it has the lowest AIC value.**[refer to appendix1]**.*
*The chosen model indicates that the data for southindia is seasonal and non-stationary.[**refer to appendix2**]*

Now the forecasting model is given by

$$y_t = y_{t-12} + \Phi_1 y_{t-12} + \Phi_2 y_{t-13} - \Phi_1 y_{t-24} - \Phi_2 y_{t-25} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Using which we forecast the no. of monthly tea production for the years 2020,2021,2022 in R.

```
>library("forecast")
>southindia2<-auto.arima(southindia1)
>southindia3<-forecast(southindia2,h=36)
>plot(southindia3)
```

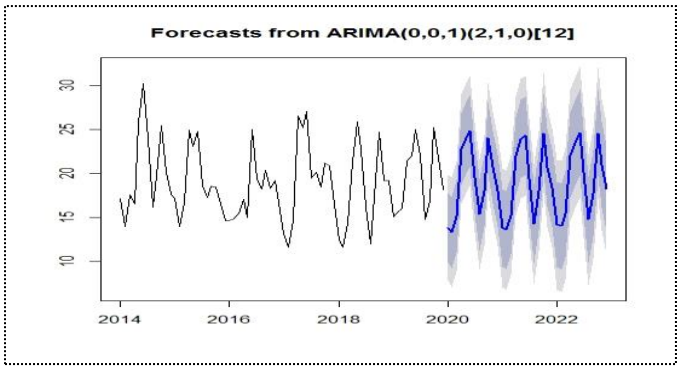


FIGURE11:FORECASTING

Here the forecast and 95% limits are given from for years of 2020,,2021, 2021.Hence we conclude the analysis of north india and generalize their conclusion for group A as a whole.

3. RESULTS AND DISCUSSION:

*Area under tea in West Bengal has varied between **82705 hectares** to **560000 hectares** with an average of **321351.72 hectar**es, registering a simple growth rate of almost 0.75 % per annum as against compound growth rate of 0.70%.*

- *The value of **skewness** indicates that there has been shift of area in favour of tea during the early phase of study period.So far about the productions of tea during study period is concernened.*
- *An average production of 154422.74-thousand-kilogram associate with a simple growth rate of almost 4.69 % per annum.*
- ***Skewness** is positive (0.70) while kurtosis is negative (-0.06) indicates that that there has been increase in production during early half of the study period and it remains steady for a long duration. Average productivity of tea in West Bengal is 1524.33 kg^{-1} ha and ranges between 998 to 3986 kg^{-1} ha By and large the productivity of tea has recorded a 2.40% compound growth rate during the whole period under study.*
- *Both **skewness** and **kurtosis** were **positive** indicating been increase in productivity during early half of the study period could not be sustained for longer period*

4. PROBLEMS OF TEA INDUSTRY:

*The North East tea industry is gradually **sinking** due to various problems being faced by it. There are many tea gardens which have closed down in recent years due to various problems affecting the industry. Some of these are **Red Bank**, **Dharanipur**, **Dheklapara**, **Surendranagar** and **Bundapani** of the **Dooars** and **Raipur Tea estate** in **West Bengal**. In **Assam** too, many tea estates have closed down or are operating just minimally. The Problem areas of tea industry in north-Bengal have been identified through survey work.*

4.1.1 Economical Factors:

- *The wage component constitutes over 60% of the total cost of production. Escalating input costs of electricity, fuel, pesticides and agro-chemicals and irrigation, etc. make Indian tea non-competitive in international market..*
- *Inconsistent investment is not ideal for Tea industry. Most of the labourers come from socially and economically weaker sections of the society and from their native place lying far away.*

4.1.2.Social Factors:

- *Leaving aside the natives, most of the workers in the tea gardens in Assam and Bengal are the ones who have migrated from states like **Bihar**, **Orissa** and **Madhya Pradesh**, since the later part of the 19th century. These workers are solely dependent on the tea industry for their daily income and livelihood..*
- *labourers have to stay in **over-crowded** and **unhygienic living conditions** in the residential colonies. They live in **poor socio-economic** conditions and most of these labourers remain ignorant due to **illiteracy**. The labourers suffer from various **diseases** and **health problems** .*
- *Most of the tea garden labourers are basically from the "Adivasi" community and the local people. These people are the most exploited ones in the tea gardens. Most of them work as part-time labourers as per production requirement and peak season*

4.1.3. Technological Factors

- *The technology that is used in tea production is **not** very **advanced**. New technology innovation is not very advanced the basic product i.e. the tea. Tea industry is not an industry where overnight changes have to be made or production is affected due to new technology.*
- *Being highly labour intensive technological factors is least affecting the tea industry. India is the largest manufacturer and exporter of tea machinery.Other major tea producers (also developing nations) source equipment and technology from India.*
- *In the Assam-Darjeeling region, the transportation of goods is highly **time-consuming** and **costly** because of the poor condition of the narrow roads.*
- *The problem of **storing** premium quality tea has always been there. Due to delay in transportation and lack of storage facilities, the processed tea gains moisture from the atmosphere and deteriorates in quality.*

4.1.4. Environmental Factors

- ***Change** in **climatic** condition is a major cause of the variation in tea production. In **south India**, tea productivity is better because of almost **uniform climatic** condition. In **North India**, winter is **cold**, therefore, the tea bushes go to dormant stage and as a result productivity goes down during winter months.*
- ***Erratic rainfall** pattern causes frequent landslides in the hilly **terrain**, causing huge damage and heavy losses to the estates.*

5. TEA EXPORT:

*Having determined the importance of the tea industry in the Indian industry, we also realize that exports are essential for the sustenance of this industry. Indian tea exports have been losing out to competition from **Kenya**, **Sri Lanka**, **China** and other new-age tea exporting nations.Here we show the tea export of **Asian countries** (**Bangladesh,Srilanka,China,Indonesia,Vietnam,India**) in figure 12 and tea export of **African countries(Kenya,Malawi,Zimbabwe,Rwanda,Tanzania,Uganda**)in figure 13.*

	2008	2009	2010	2011	2012	2013
banglades	7.9	2.1	0.9	1.5	0.6	0.5
sri lanka	303.5	279.9	305.8	303.2	306.1	311
china	291	303	302.4	322.6	321.8	329.7
india	200.2	180.5	182.7	205.3	199.1	209.2
indonesia	91.7	92.3	87.1	75.5	70.1	70.8
vietnam	108.2	134.1	138.4	122.6	145	133.5
asia	1002.5	991.9	1017.3	1030.7	1042.7	1054.7
kenya	301	281.1	362.3	347.5	349.9	415.9
malawi	42.9	46.8	48.9	44.9	41.8	40.5
zimbabwe	9.1	4.5	5.1	5.7	5.9	5.9
rwanda	17.4	18.8	21.5	23.2	23	23.5
tanzania	26.4	24.4	26.1	27.1	27.8	26.2
uganda	39.6	47.9	53.7	47.9	52.3	56.7
AFRICA	436.4	423.5	517.6	496.3	500.7	568.7

Data table

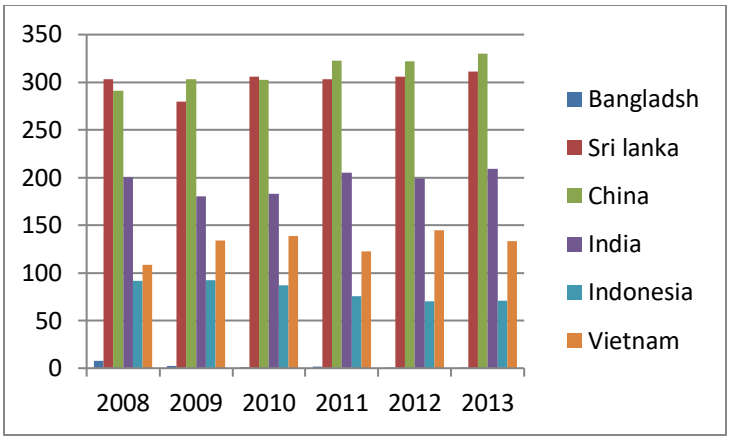


FIGURE12:TEA EXPORT IN ASIAN COUNTRIES

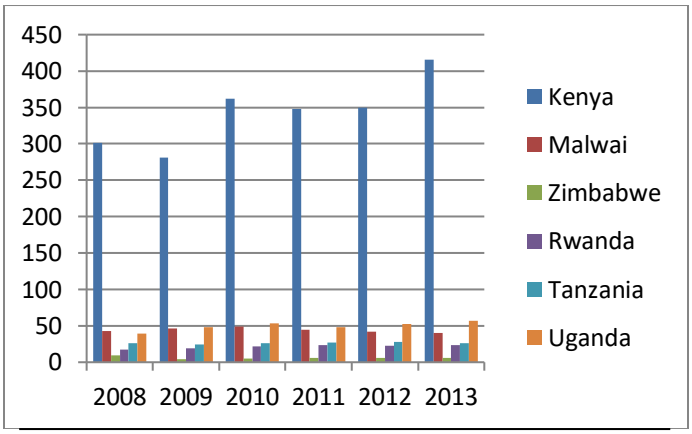


FIGURE13:TEA EXPORT IN AFRICAN COUNTRIES

Exports are essential to earn foreign exchange for the nation and Indian tea had traditionally been a major contributor in this regards.

5.1. PROBLEMS OF TEA EXPORT:

a) Unfavourable age profile of significant proportion of India's tea gardens

b) Tariff and non-tariff measures

c) Changing Consumer Preferences

d) Shift in India's production

e) Decline in quality

f) Spurious varieties

g) Lack of marketing initiatives

5.2.TEA EXPORT FORECASTING:

- *Tea in **Zimbabwe** posted slower growth than other countries due to a dip in production of black tea due to poor weather conditions. Some of the smaller players had to increase their prices, with this resulting in a small drop in consumption.*
- *In 2013, the number of tea houses in Asia and African countries in the world increased.The countries have witnessed a change in consumer perception, with tea drinking becoming popular among younger generations.*
- *Teas offering health benefits are growing thanks to consumers increasingly adopting healthier lifestyles.*
- *Companies such as Tata Global Beverages have introduced Ayurveda-inspired teas containing including tulsi, basil and ginger.*

*We can **forecast** the tea export of asian and African countries for 2014,2015,2016 using R.*

5.2.1.CASE1:ASIAN COUNTRIES:

Fitting and Forecasting

*Here we fit the data on tea export in **Asian countries(Bangladesh,Srilanka,China,Indonesia,Vietnam,India)** from 2008 to 2013 into an **ARIMA** model using R*

Here the seasonal ARIMA model used is (0,1, 0)

Now we forecast the tea export for the years 2014,2015,2016 using the given equation in R

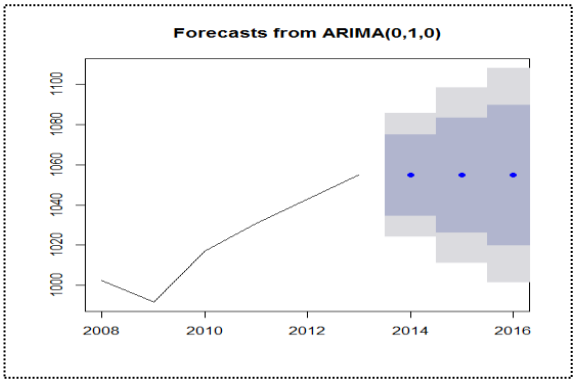


FIGURE14:FORECASTING

Here the forecast and 95% limits are given from for years of 2014,,2015, 2016.Hence we conclude the

analysis of Asian countries tea export and generalize their conclusion for case 1 as a whole.

5.2.2.CASE2:AFRICAN COUNTRIES:

Fitting and Forecasting

*Here we fit the data on tea export in **African countries(Kenya,Malawi,Zimbabwe,Rwanda,Tanzania,Uganda)** from 2008 to 2013 into an **ARIMA** model using R.*

Here the seasonal ARIMA model used is (0,1,0)

*Now we **forecast** the tea export for the years 2014,2015,2016 using the given equation in R*

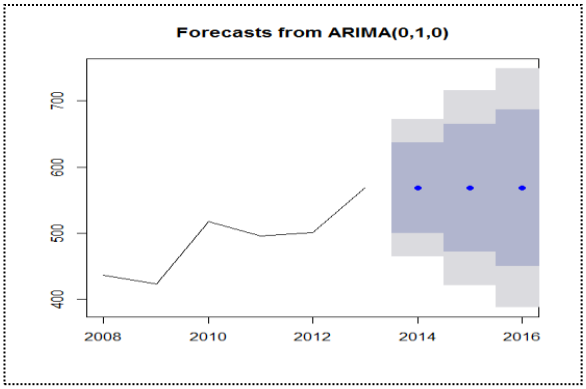


FIGURE15:FORECASTING

6. PROSPECTS OF TEA INDUSTRY:

- *Concerted efforts have to be made by different stakeholders to increase the productivity of tea plantations. The extension services should be strengthened to disseminate the technical know-how to the small tea growers located in remote areas.*
- *The country can get back its place among our traditional tea markets of Russia and other CIS nations by diversification of tea products, quality upgradation and aggressive brand and logo campaign. These efforts would also help in realization of better prices thus improving our competitiveness and profitability of tea industry.*

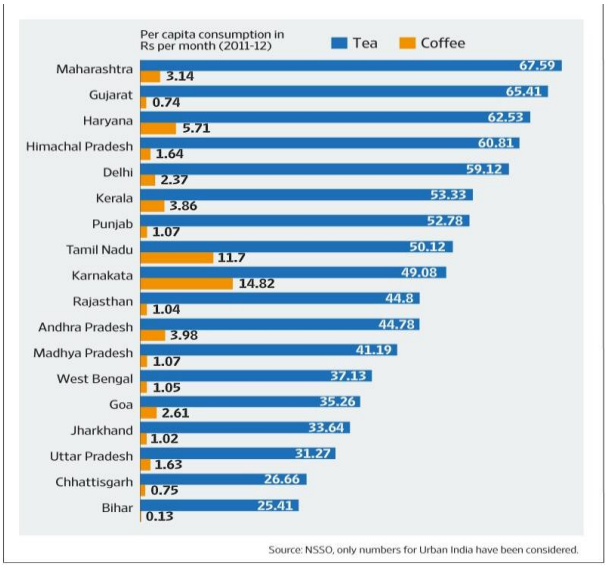


FIGURE16:STATE WISE PER CAPITA TEA CONSUMPTION IN INDIA

*This Figure depicts with the help of a **multiple bar diagram** the per capita consumption of tea for different states in India in decreasing order of magnitude. Here emphasis should be laid on those states for which per capita tea consumption is relatively on the lower side. So the states need to be considered from the right most side of the bar diagram. From this perspective, focus needs to be given for increasing per capita consumption of tea on Bihar, Chhattisgarh, UP, Jharkhand, Goa, West Bengal.. so on and so forth.*

- *we need to consolidate our distribution channels in these markets and look at developing more attractive retail options to appeal to the consumer.*
- *As far as our long term strategy is concerned, clearly there is much benefit associated with **diversifying** in **non-traditional** yet fast growing tea consuming nations . This is because these markets are not saturated and the consumers will thus be willing to try out new options.Because of the nascent stage of tea penetration, these markets are still hooked onto black tea and this opportunity must be harnessed. The market for organic and green tea is still comparatively underdeveloped, which augurs favorably for Indian exporters of black tea.*
    - *However, the **major** competition in these markets will not be from other **developing nations** like **Kenya**, **Sri Lanka** or **China**, but the major **threat** will be from the re-exporters of tea within **UK**, **Germany** and **Poland**. However, it is essential for us to move up the value chain by packaging and marketing the tea according to local specifications as these will grant our exporters higher margins.There is a need to reduce the import of tea in India.*
- *There is a need for creating **awareness** among the people, especially among the rural people, regarding the health benefits of tea consumption.*

7. CONCLUSION

At the end of this project it can be concluded that:

- *In most cases,the seasonal peak in the tea production corresponded with the dry seasons of **Northern** and **Southern** India where the states are located i.e. the first flush starts in late March.the second flush occurs during end of May to June.this season produces the best quality tea leaves with gold tips,the raw material for the prized "**tippy tea**".*
- *The global scenario as well as Indian Scenario of tea in terms of area, production,exports and imports indicated overall increase in the quantity of tea in the Indian market over last two decades and the trend is increasing.*
- *General consumption of tea, health benefit effects of tea need to be promoted more vigorously to trap the nonconventional areas of tea in the world for an expansion in the consumption.*
- *In West Bengal some tea gardens, plagued with low yield, can partially or fully opt,instead of cultivating tea. It is reported that West Bengal Government is extending its helping hand for this switch over.Through these ways, the Indian tea industry can think of sustainability in future keeping in mind the climate change across the globe.*

ACKNOWLEDGMENT

"It is not possible to complete a project without the assistance and encouragement of other people.this one is certainly not the exception."

On the very outset of this project,I would like to extend my sincere and heartfelt obligation towards all the personages who have helped me a lot in this endeavor.Without their active guidance,help,cooperation and encouragement,I would not have made headway in this project.

I am ineffably indebted to our head of the statistics department Sir.Arup Kumar Hait for conscientious guidance and encouragement to accompalish this assignment.

I am extremely thankful and pay my gratitude to my faculty Sir Kiranmoy chatterjee,Sir Suryasish Chatterjee and Sir Soumyadeep das for their valuable guidance and support on completion of this project in its presently.

I extend my gratitude to my college Bidhannagar Govt. College for giving me this opportunity.

I also acknowledge with a deep sense of reverence,my gratitude towards my parents who have always supported me morally as well as economically.

At last but not the least gratitude goes to all of my friends who directly or indirectly helped me to complete this project.

APPENDIX

Appendix 1-

*The R package which enables the plot function is "**ggplot2**".the forecast is the another of the R packages which contains functions such as*

- *auto.arima()-----for fitting an arima model to the data.*
- *forecast()-------for forecasting a time series based on a model.*

*Both the R packages can be downloaded from **https:\\cran.rproject.org**.*

A suitable model selection by auto.arima()-

***Forecast package** provides function **auto.arima**() for the automatic selection of arima model.the auto.arima in R uses a combination of unique root test,minimization of the **AIC**(**akaike's information criterion**) and **MLE**(**maximum likelihood estimation**) to obtain an arima model.*

A model is chosen and vaue of AIC is calculated.the best model is the one with the lowest AIC.the best model considered so far becomes the new current model.now this process is repeated until no lower AIC can be found.

AIC,AICc and BIC:

*AIC is an **estimator** of out of sample prediction error and thereby relative quality of statistical model for a given set of data.given a collection of models for the data,AIC estimates the quality of each model,relative to the each of the other model.Thus AIC provides a mean for model selection.AIC is founded on information theory.when a statistical model is used to represent the process that generated the data,the representation will almost never be exact.so some information will be lost by using the model to represent the process.AIC estimates the relative amount of information lost by a given model.the less information a model looses,the higher the quality of model.in estimating the amount of information lost by a model,AIC deals with the trade off between the **goodness** of **fit** of the model and the simplicity of the model.in other words,AIC deals with both the risk of **overfitting** and the risk of **underfitting**.*

$$AIC = 2k - 2\ln(L)$$

Let, k be the number of estimated parameter in the model.

L be the maximum value of the likelihood function for the model.

AICc is AIC with a correction for small sample sizes.the formula for AICc depends upon the statistical model.assuming that the model is univariate,is linear in its parameters,and has normally distributed residuals then the formula for AICc is as followed:

$$AICc = AIC + (2k^2 + 2k)/(n-k-1)$$ *(Where n is sample size).*

the formula for the BIC is similar to the formula for AIC but with a different penalty for te number of parameters.with AIC the penalty is 2k,whereas with BIC ,the penalty is ln(n)k.
AIC is asymptotically optimal or selecting the model with least mean squared error ,under the assumption that the "true model" is not in the candidate set.BIC is not asymptotically optimal under the assumption.

Appendix 2-

Different stochastic model:-

- ***AR(autoregressive) and MA(moving average)***

In an AR(p) model ,the future value of a variable is assumed to be a linear combination of p past observation and a random error together with a constant term.

- ***AR(p) model :***

$$y_t = c + \sum_{i=1}^{p} \varphi_i y_{t-i} + \varepsilon_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t$$

i.e. $\varepsilon_t = \varphi(L) y_t$

where L is a lag operator such that $Ly_t = y_{t-1}$

On the other hand,MA(q) model,the future value of a variable is assumed to be a linear combination of q past error and the random error together with the mean of the series.

- ***MA(q) model:***

$$y_t = \mu + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

i.e. $y_t = \theta(L) \varepsilon_t$

ARMA model:

Auto regressive(AR) and moving average(MA) model can be effectively combine together to form a general and useful class of time series model,known as the ARMA model.

ARMA(p,q)model:

$$y_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

i.e. $\varphi(L) y_t = \theta(L) \varepsilon_t$

Here, $\varphi(L) = 1 - \sum_{i=1}^{p} \varphi_i L^i$ **and** $\theta(L) = 1 + \sum_{j=1}^{Q} \theta_j L^j$

Now,ARMA model can only be used for stationary time series data.however in practice many time series show non stationary behavior.

ARIMA MODEL :

Time series which contain trend and seasonal pattern are also non stationary in nature.thus from an application view point ARMA models are inadequate to properly describe non stationary time series,which are frequently encountered in practice.for this reason the ARIMA model is proposed which is a generalization of an ARMA model to include the case of non stationarity as well.

In ARIMA model nonstionary time series is made stationary by applying finite differencing of the data point.

ARIMA(p,d,q):

$$\varphi(L)(1 - L)^d y_t = \theta(L) \varepsilon_t$$

i.e. $\left(1 - \sum_{i=1}^{p} \varphi_i L^i\right)(1 - L)^d y_t = \left(1 + \sum_{j=1}^{q} \theta_j L^j\right) \varepsilon_t$

SARIMA MODEL:

The ARIMA model is for non seasonal,non stationary data.Box and Jenkins have generalized this model to deal with seasonality.their proposed model is known as the seasonal arima(SARIMA) model.in this model,seasonal differencing of appropriate order is used to remove nonstationarity from the series.

$$\Phi_P(L^s) \varphi_p (1 - L)^d (1 - L^s)^d y_t = \vartheta_Q(L^s) \theta_q(L) \varepsilon_t$$

REFERENCE

- *A.M.Gun,M.K.Gupta,B.Dasgupta,Fundamentals of Statistics Voulme 1*
- *A.M.Gun,M.K.Gupta,B.Dasgupta,Fundamentals of statistics Volume 2*
- *Gupta, Kapoor,Mathematical statistics Volume 2*
- *Kaggle.com*
- https://www.goldsteinresearch.com/report/india-tea-industry-outlook*www.forewatch.org*
- *www.wwf.org*
- https://en.wikipedia.org/wiki/Darjeeling_tea*www.reuters.com*