

Report on
DEVELOPING A DIAGNOSTIC TOOL AND
TIME SERIES ANALYSIS IN RESPECT OF
SOYABEAN PRICES

Submitted By

Bhavya Kapoor

(2K16/MBA/11)

Under the Guidance of

Dr. Pradeep Kumar Suri

Professor



DELHI SCHOOL OF MANAGEMENT

Delhi Technological University

Bawana Road Delhi 110042

Jan - May 2018

CERTIFICATE

This is to certify that the dissertation report titled “**Developing a Diagnostic Tool and Time Series Analysis in Respect of Soyabean Prices**” is a bonafide work carried out by **Ms. Bhavya Kapoor** of **MBA 2016-18** and submitted to Delhi School of Management, Delhi Technological University, Bawana Road, Delhi-42 in partial fulfillment of the requirement for the award of the Degree of Masters of Business Administration.

Signature of Guide

Signature of Head (DSM)

Seal of Head

Place:

Date:

DECLARATION

I, **Bhavya Kapoor**, student of **MBA 2016-18** of Delhi School of Management, Delhi Technological University, Bawana Road, Delhi – 42, hereby declare that the dissertation report “**Developing a Diagnostic Tool and Time Series Analysis in Respect of Soyabean Prices**” submitted in partial fulfillment of Degree of Masters of Business Administration is the original work conducted by me.

The information and data given in the report is authentic to the best of my knowledge.

This report is not being submitted to any other University, for award of any other Degree, Diploma or Fellowship.

Place:

Bhavya Kapoor

Date:

ACKNOWLEDGEMENT

First of all, I would like to thank Dr. Pradeep Kumar Suri, who was my guide and my mentor during this project, for providing me with the opportunity to do this dissertation under his guidance and expertise.

I would also like to extend my gratitude to my Parents, Mrs. Sapna Kapoor and Mr. Sanjeev Kapoor, who provided me with their love and support whenever I needed it.

Lastly, I would like to thank God for giving me the skills and the patience because of which this project was made possible.

Bhavya Kapoor

ABSTRACT

AGMARKNET is an initiative towards the globalization of Indian Agriculture by the Ministry of Agriculture under Government of India. The website provides detailed information on the agriculture marketing reform, marketing research, market profiles, prices, agricultural commodities, commodity arrivals, contracting farmers etc. Details of commodities and agricultural markets are provided. Consumers can benefit too as this will bring transparency in arrivals, curbing hoarding.

The information that this portal provides is critical to all the stakeholders associated with agriculture. AGMARKNET aims to deliver correct and complete information to the farmers in order to strengthen their economic position. Farmers are the main stakeholders of this project. With effective use of this portal, the farmers can have an estimate of the prices of the commodity that they are going to sell in the markets. It tells them the prices of the commodity in detail such as arrival price of each and every day.

This study aims to develop a diagnostic for major markets across India where Soyabean arrivals are high. This diagnostic tool will help the data entry operators to have various checks on the entire process of data entry. Thereby, improving the data quality by monitoring the data reporting process continuously. Statistical Process Control (SPC) can help the correct reporting of data. Microsoft Excel 2016 was used for diagnostic tool development.

The second objective of the study is to forecast the modal prices of Soyabean for using time series modeling. Data filtering, sorting and cleaning are the essential tasks conducted as part of this process. The time series modeling has been attempted to forecast soyabean prices in 2017. Microsoft Excel 2016 and Eviews 9 have been used for time series analysis.

Table of Contents

Chapter 1 Introduction.....	3-7
1.1 Background	3
1.1.1 Directorate of Marketing and Inspection	5
1.1.2 National Informatics Centre (NIC)	5
1.2 Significance of Study	6
1.3 Objectives of the Study.....	7
Chapter 2 Review of the Literature	8-21
2.1 What Is Agricultural Marketing?	8
2.2 Agricultural Marketing Information System.....	10
2.3 Data Reporting on AGMARKNET portal	11
2.4 Overview of Soyabean Production	12
2.4.1 Soyabean Statistics	13
2.4.2 Major Soyabean Producing States (2017).....	14
2.5 Statistical Process Control	14
2.5.1 SPC and Non-Manufacturing Sector	15
2.5.2 Individuals and Moving Range Chart	17
2.6 Time Series Forecasting	18
2.6.1 Exponential Smoothing Models.....	18
2.6.2 Making Time Series Stationary	20
2.7 Concluding Remarks	21
Chapter 3 Methodology	22-26
3.0 Introduction to Methodology	22
3.1 Project Phases	22
3.1.1 Understanding Phase	22
3.1.2 Defining Phase	23
3.1.3 Data Preparation and Analysis Phase.....	24
3.1.4 Developing Phase.....	25
3.1.5 Testing Phase.....	25
3.1.6 Findings and Recommendations Phase.....	25

3.2 Concluding Remarks	26
Chapter 4 Data Analysis	27-43
4.0 Introduction to Data Analysis	27
4.1 Development of a Diagnostic Tool	27
4.1.1 Choosing a Tool to Analyze Data	27
4.1.2 Selection of Market for Analysis	28
4.1.3 Checking for Modal Price Reporting Frequency	30
4.1.4 Data Cleaning of Modal Prices	30
4.1.5 Developing a Control Limits and Control Charts.....	31
4.1.6 Developing the Interface of the diagnostic tool.....	32
4.2 Forecasting the Modal Prices of Soyabean	34
4.2.1 Choosing the Tools of Analyses.....	34
4.2.2 Drawing Inferences from the Arrivals Data.....	34
4.2.3 Data Cleaning of Modal Prices	36
4.2.4 Choosing a Forecast Model and Technique.....	36
4.2.5 Forecast Errors for the Months of the 2016.....	40
4.3 Findings	41
4.4 Limitations	43
4.5 Concluding Remarks and Recommendations	43
Chapter 5 Conclusion of the Study	45-46
Bibliography	

Chapter 1

Introduction

1.1 Background

Information & Technology has transformed many aspects of human endeavor and has helped to create systems for responding to a wide range of societal needs. Indeed, transportation, communication, national security, and health systems are completely reliant on IT to perform even basic functions. Information, and its automated technological embodiment, has also impacted agriculture to the same level.

The Importance of Agriculture

Agriculture is a major sector which is vital for the survival of modern man. Plants are the producers in the food chain, and without them, the life cycle would just not be possible. Agricultural produce, though highly perishable compared to other food sources, is essential for survival. Crops are used to produce several food sources by themselves or through by-products such as bread, powders, organic additives to other goods and the like.

The produce from agriculture drives trade from one country to another, brings income for farmers, makes productive use of otherwise idle land, and brings food on the table. It is such an important part of everyone's daily life, although it may not be seen as a direct factor since the produce goes a long way before reaching the hands of everyone who benefits from it. Because of its importance to society, it's must to evolve with the times and adjust to meet the needs of modern people. By adapting and making use of IT to help improve agricultural progress, everyone benefits from the union of these sectors.

India particularly has one third of its population directly or indirectly dependent on agriculture and contributing 25 per cent to the GDP. The basic definition of Agricultural Marketing is the buying and selling of agricultural products.

Marketing information is a very important aspect for agricultural stakeholders in order to plan their crop produce and marketing of the produce. The stakeholders involved in the agriculture markets also require market information to make and take well informed decisions. So, it is necessary that the marketing information conveyed to the stakeholders is complete and correct. Information & Technology and IT enabled services have enabled to manage and disseminate information even in the remote corners of the country. It was the need of the hour to setup a marketing information network for agricultural commodities grown throughout the country.

About AGMARKNET

The Central Sector Scheme project of Agricultural Marketing Information Network (AGMARKNET) was launched in March 2000 by the Ministry of Agriculture under Government of India. The AGMARKNET website (<http://www.agmarknet.nic.in>) is a G2C e-governance portal that caters to the needs of various stakeholders such as farmers, industry, policy makers and academic institutions by providing agricultural marketing related information from a single window. The portal has helped to reach farmers who do not have sufficient resources to get adequate market information. It facilitates web- based information flow, of the daily arrivals and prices of commodities in the agricultural produce markets spread across the country. The data transmitted from all the markets is available on the AGMARKNET portal in 8 regional languages and English. It displays Commodity-wise, Variety-wise daily prices and arrivals information from all wholesale markets. Various types of reports can be viewed including trend reports for prices and arrivals for important commodities. The project is technically supported by the National Informatics Centre (NIC).

Every unit is furnished with important PC equipment, alongside web network. "AGMARK," an easy to understand portal has been produced to encourage assemblage and transmission of information at market level. The revealing framework is currently web empowered. It additionally gives week by week information to value incline for real markets and significant agribusiness products. Online Exchange Portals are connected to the AGMARKNET entryways to give spot and future costs.

1.1.1 Directorate of Marketing and Inspection

The Directorate of Marketing and Inspection (DMI), under Ministry of Agriculture & Farmers Welfare, was set up in the year 1935. The Agricultural Marketing Adviser to the Government of India (AMA) heads the Directorate.

Its prime aim is to implement agricultural policies and plans so that an integrated development of marketing information pertaining to agriculture can take place. By maintaining a close liaison with Central Government and State Government DMI also thrives to safeguard the interest of both farmers and consumers (**source:** <http://dmi.gov.in/About.aspx>).

1.1.2 National Informatics Centre (NIC)

Since its establishment in the year 1976, National Informatics Centre (NIC) has been the major builder applications pertaining to the domain of e-governance having a reach up to the grassroots levels. "NICNET", connects all the Ministries and Departments of the Central Government, links 36 of the State Governments or Union Territories, and spans about 688 district administrations across India. NIC aims to implement Information Communication Technology based applications to derive competitive advantage in social as well as public administration.

NIC also has under its belt various initiatives such as Government eProcurement System (GePNIC), Office Management Software (eOffice),

Hospital Management System (eHospital), Government Financial Accounting Information System (eLekha).

For the Agricultural Marketing Network Scheme, NIC had provided computer hardware, developed the software, provided training to market personnel towards the operation of the hardware and software systems and provided internet connectivity. It has also developed the integration between the software packages developed by the various states with AGMARKNET to bring about seamless uniformity in the database.

For AGMARKNET, the State Government/Marketing Boards gave the list of markets to be secured under the AGMARKNET project. The chosen markets were to give site to establishment including offices for computer installation, phone network and computer operating personnel.

Showcase Committees/Controlling specialists of AGMARKNET hub at market level were doled out to gather pertinent information and data, encourage it and transmit it to the State level and AGMARKNET entry. NIC had additionally prepared suitable people from every hub in working on a computer and dealing with programming bundle.

At each market hub, there is a man appointed to gather information and transmit it. A motivation plot has been acquainted with reward information section administrators for keeping up execution guidelines consistently.

1.2 Significance of Study

The Agricultural Marketing Information System Network (AGMARKNET) which is based on the "NICNET" aims at linking all the important Agricultural Produce Market Committees (APMC), State Marketing Boards and Directorate of Marketing and Inspection regional offices. So, that there can be seamless exchange of information.

The purpose of the study is to help AGMARKNET to improve the data quality of the reporting prices. As data quality of an agricultural commodity is of prime

importance to the stakeholders particularly when it comes to issue such as price rise, hence with the study it was aimed to improve upon the quality of price reported. More over forecasting the future prices of an agricultural commodity can help the stakeholders to be better prepared for price rise or fall. Through timely awareness, forecasted values can provide the much-needed leverage of time to adapt and perform accordingly. Data filtering, sorting and data gap filling can help us to draw meaningful inferences from the agricultural data and perform activities such as forecasting and diagnostic tool development using that data for distinct markets.

1.3 Objectives of the Study

The arrivals and prices data of all the commodities is easily available on the AGMARKNET portal. However, the reported data has many short comings when it comes to the quality. Additionally, the daily price data suffered from irregularities because of lack of uploading data on a daily basis. It was also observed that he daily reporting makes the data prone to errors which might be unintentional or deliberate in nature. The objective of the is to find out gaps in the data reporting, and then to bridge those gaps in order to improve upon the data quality. The study was primarily around the agricultural commodity Soyabean and focuses on the study of modal prices. After the analysis of data, the study aims to fulfil two major objectives:

- To develop a diagnostic tool which can report irregularities in price for major Soyabean markets
- Forecast the future prices of Soyabean for major markets

Chapter-2

Review of Literature

2.1 What Is Agricultural Marketing?

India is an agricultural country and one third population depends on the agricultural sector directly or indirectly. Agriculture remains as the main stay of the Indian economy since times immemorial. Indian agriculture contribution to the national gross domestic product (GDP) is about 17 per cent. With food being the crowning need of mankind, much emphasis has been on commercializing agricultural production. For this reason, adequate production and even distribution of food has of late become a high priority global concern.

Agricultural marketing is mainly the buying and selling of agricultural products. Agricultural marketing includes all those activities which are mostly related to the procurement, grading, storing, transporting and selling of the agricultural produce. It is an efficient way by which the farmers can dispose their surplus produce at a fair and reasonable price. Improvement in the condition of farmers and their agriculture depends to a large extent on the elaborate arrangements of agricultural marketing.

Agricultural marketing system in developing countries including India can be understood to compose of two major sub-systems viz., product marketing and input (factor) marketing. The actors in the product marketing sub-system include farmers, village/primary traders, wholesalers, processors, importers, exporters, marketing cooperatives, regulated market committees and retailers. The input sub-system includes input manufacturers, distributors, related associations, importers, exporters and others who make available various farm production inputs to the farmers

As indicated by the National Commission on Agriculture (XII Report, 1976), agricultural marketing is a procedure which begins with a choice to create a

saleable farm product, and it includes every one of the parts of market structure or framework, both practical and institutional, in view of specialized and monetary contemplations, and incorporates pre-and post-gather operations, amassing, evaluating, capacity, transportation and circulation. Effective agricultural marketing requires the ability to analyze complex market data, use that data to identify changes in demand, and develop persuasive arguments for multiple audiences in order to increase demand.

Marketing information not only helps farmers make profitable decisions in the short term on when and where to market produce, and what price to expect. It also has another vital function, of helping farmers decide what to produce. Since there may be a time delay of months or years between investing in a crop or herd and getting a profit, farmers who understand market trends and market opportunities have a better chance of succeeding than those who do not.

In fact, marketing information plays a vital role in the functioning of the whole market, by regulating the competitive marketing process. By helping ensure that produce goes to markets where there is a demand for it, it shortens marketing channels and cuts down on transport costs. It helps ensure that each marketing transaction is a fair one, and that all participants share the risks and benefits. However, this does not happen if marketing information is distributed unequally, as is generally the case when many small-scale farmers in Asia are selling to a relatively few large-scale dealers. The farmers then end up bearing the greater part of the risk, while the dealers end up with the greater part of the profits.

Recent advances in information technology are making it more feasible to provide small-scale farmers with the marketing information they need. However, farmers may not benefit from sophisticated facilities, if the system is poorly managed or not designed for their needs. It is not enough for marketing information to be collected: it must also be disseminated in a form accessible to farmers and adapted to their needs.

2.2 Agricultural Marketing Information System

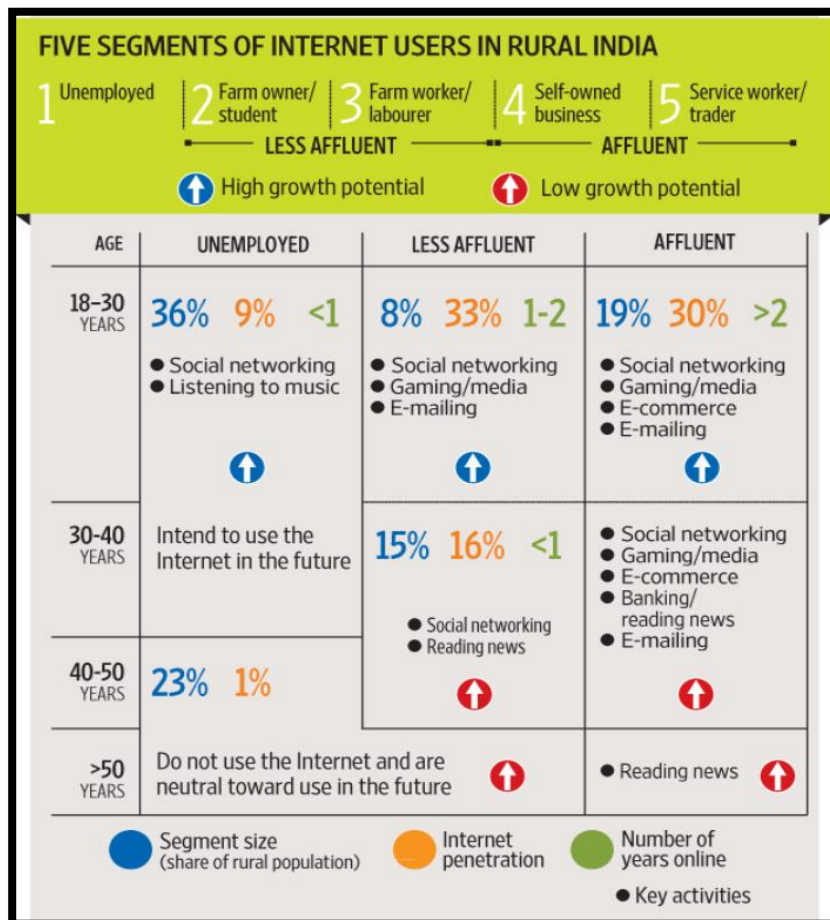
An Agricultural Marketing Information System is a system in which marketing data is formally gathered, stored, analysed and distributed to managers in accordance with their informational needs on a regular basis." In nations like India, the different activities especially that of marketing data is prepared by government as a piece of agricultural marketing and agribusiness procedure advancement. Productive market data arrangement can be shown to have positive advantages for farmers, dealers and the government. Up-to-date, or current, market data strengthens the farmer's position to consult with dealers from a place of a better comparative advantage. It also encourages spatial dispersion of items from country territories to urban zones and between urban markets by sending the representation of data in its true form from urban shoppers to rural producers in regards to quantities and varieties required.

Current interchanges innovations open up the conceivable outcomes for market data administrations to enhance data conveyance through SMS on mobile phones and the fast development of FM radio stations in many creating nations offers the possibilities of more localized information services. Radio projects like Kisanvani on All India Radio and Kisan Suvidha on versatile give farmers the genuinely necessary data in regards to farm and farm produce delivery.

Also, there has been an increase in penetration of internet over the last 16 years. As a result, there is rapid flow of information. AGMARKNET is one such portal giving agricultural related information of various varieties of crop to its stakeholders.

Figure 2.1: Showing Percentage of Population Using Internet

(Source: BCG CCI Rural Digital Influence Study)



2.3 Data Reporting on AGMARKNET portal

Data on the AGMARKNET portal is primarily reported as commodity Arrivals and Prices. The arrival data shows the amount of agricultural produce received in tonnes. Arrival data is reported on yearly, monthly, weekly and even daily basis which is used to compile reports. One can easily see which state, district or market receives maximum arrivals of a particular agricultural commodity.

On the other hand, the price data is reporting in the unit of Rs/Quintal. There are three types of prices reported on the portal they are:

- Maximum Price

The maximum prices paid for an agricultural commodity during several transactions held in a day in a market is termed at maximum price.

- Minimum Price

The minimum price paid for an agricultural commodity during several transactions held in a day in a market is termed at minimum price.

- Modal Price

The modal price is the price at which most of the commodity was sold in a market during the day.

Along with these three prices one can also get the details of state name, district name, group to which agricultural commodity belongs, variety and grade.

Thus, the information being disseminated through the AGMARKNET portal is quite comprehensive and vital to all the stakeholders of agriculture.

2.4 Overview of Soyabean Production

Soybean is a legume classified by FAO as an oil crop. Soybean was originally used for crop rotation, as a nitrogen fixing plant, and for the production of fermented food products. Until fairly recently, cultivation was limited largely to Asian countries and only gained popularity in the Americas and Europe in the early 1900s.

Although in some countries soybean is transformed at the local level into food products for direct human consumption, much of today's production is crushed at industrial scale to obtain soybean oil and soybean meal, which account for about one third and two thirds of the crop's economic value respectively. Soy-oil and soy-meal are consumed worldwide as food and animal feedstuff respectively. In recent years, soybean oil also started to be used for the production of bio-diesel.

Commercial production of genetically modified soybeans has increased fast in recent years, with important implications for consumption and trade.

The climatic conditions suitable for growing Soyabean are given below:

- Temperature: The crop is grown under warm conditions in the tropics, subtropics and temperate climates. Soybean is relatively resistant to low and very high temperatures but growth rates decrease above 35°C and below 18°C. In some varieties, flowering may be delayed at temperatures below 24°C. Minimum temperatures for growth are about 10°C and for crop production about 15°C.
- Water Requirement: Water requirements for maximum production vary between 450 and 700 mm/ season depending on climate and length of growing period.

2.4.1 Soyabean Statistics

Soybean is known as the “GOLDEN BEAN”. Although Soybean is a legume crop, yet it is widely used as oilseed. Due to very poor cookability and digestibility on account of inherent presence of trypsin inhibitor, it cannot be utilized as a pulse. It is now the second largest oilseed in India after groundnut. It grows in varied agro-climatic conditions. Due to its popularity, several countries such as Japan, China, Indonesia, Philippines, and European countries are importing Soybean to supplement their domestic requirement for human consumption and cattle feed. Soybean has great potential as an exceptionally nutritive and very rich protein food. It can supply the much-needed protein to human diets, because it contains above 40 per cent protein of superior quality and all the essential amino acids particularly glycine, tryptophan and lysine, similar to cow’s milk and animal proteins. Soybean also contains about 20 per cent oil with an important fatty acid, lecithin and Vitamin A and D. The 4 percent mineral salts of Soybeans are fairly rich in phosphorous and calcium.

2.4.2 Major Soyabean Producing States (2017)

The Major Soyabean producing states according to Directorate of Economics and Statistics, Ministry of Agriculture are:

Table 2.2: Major Soyabean Producing States 2017

(Source: Directorate of Economics & Statistics)

State	Area	% to All - India	Production	% to All - India	Yield
(1)	(2)	(3)	(4)	(5)	(6)
Madhya Pradesh	5.91	50.65	4.91	57.13	831
Maharashtra	3.77	32.37	2.10	24.46	557
Rajasthan	1.20	10.33	1.00	11.63	829
Karnataka	0.24	2.08	0.14	1.57	558
Andhra Pradesh	0.00	0.01	0.00	0.01	1000
Others	0.53	4.56	0.45	5.19	@
All India	11.66	100.00	8.59	100.00	737

Thus, it is seen that when it comes to agriculture then the quality of data reported becomes of prime importance. Data reporting of arrivals and prices on a daily basis becomes a cumbersome task often involving errors. Thus, the entire data reporting process keeps on running without any checks and controls.

2.5 Statistical Process Control

There are primarily seven key tools of statistical process control, they are:

- Check Sheets

The check sheet may be a type (document) used to assemble data progressively at the range the place the data is processed. The majority of the data it catches might make quantitative and alternately subjective. In that perspective, the point when some information is quantitative, then a check sheet is sometimes called a count sheet.

- Pareto Chart

A Pareto chart, also called a Pareto distribution diagram, is a vertical bar graph where values are plotted in decreasing order of relative frequency from left to right.

- Histograms

Introduced by Karl Pearson, a histogram graphically represents the distribution of data in which is in numerical form.

- Scatter Diagrams

Values of two variables are plotted along two axes on a single graph, the pattern of the resulting points revealing any correlation present.

- Defect Concentration Diagrams

The defect concentration diagram (also problem concentration diagram) is graphical apparatus that is helpful in breaking down the reasons for the item or part defects.

- Control Charts

The control chart is a chart used to study how a procedure changes after some time. Information are plotted in time arrange. A control graph dependably has a central for the average, an upper line for the upper control limit and a lower line for the lower control limit. These lines are developed from past information.

2.5.1 SPC and Non-Manufacturing Sector

Since its being statistical process control was thought to be applicable to only manufacturing processes with the objective of reducing waste or scrap and it was thought impractical to apply statistical process control to non-manufacturing processes. But in the year 1988, Software Engineering Institute suggested that statistical process control could be applied to non-manufacturing processes, such as software engineering processes. The

Level 4 and Level 5 practices of the Capability Maturity Model Integration (CMMI) use this concept. Also in his book titled “Statistical Process Control”, Leonard A. Doty has said that any of the control charts can be applied in the non-manufacturing sector (like education, health care, politics, family life and self-improvement) in the similar way as they are applied in the manufacturing process. Thus, there is enough evidence that statistical process control can be applied to non-manufacturing processes as well.

Due to the above evidence, it was clear that information relating to statistical process control tools particularly that of control charts needs to be gathered and comprehended. Control charts are basically graphs to show how a process changes over time. With three lines namely, Upper Control Limit (UCL), Lower Control Limit (LCL) and a Control Line (CL) determined by the previous data available. These charts can help in drawing conclusions about process variations.

Table 2.3: Various Types of Control Charts

(Source: <https://www.isixsigma.com/>)

Attribute Data Chart	Variable Data Chart
p Chart (Fraction Defective)	\bar{X} and R (Average and Range Chart)
np Chart (Number Defective)	\bar{X} and R_m (Individuals and Moving Range Chart)
c Chart (Number of Defects)	
u Chart (Number of Defects per Unit)	

Attribute data chart is used when the data is counted as discrete events, whereas, in variable data chart data is measured on a continuous scale. The type of chart used is determined by the type of data collected as well as the subgroup size of the data. Since the aim is to collect price data over

a period of year, variable data chart will be most suitable for the process. When it comes to control charts it is important to understand the subgroups i.e. a group of measured units under the same conditions. The subgroup restrictions for Individuals and Moving Range Chart are that the data must have a subgroup size equals 1 (i.e. data cannot be grouped and each measurement is unique), whereas subgroup size for Average and Range Chart is greater than 1 (i.e. data can be grouped and each measurement is not unique).

2.5.2 Individuals and Moving Range Chart

An I-MR chart is a plot of individual observations (I chart) and moving ranges (MR chart) over time for variables data. The moving range is defined as $MR_i = |X_i - X_{i-1}|$, which is also the absolute value of the first difference.

Table 2.4: Control Limits

Lower Control Limit	Control Limit	Upper Control Limit
$\bar{X} - \frac{3\overline{MR}}{d_2}$	\bar{X}	$\bar{X} + \frac{3\overline{MR}}{d_2}$

The value of d_2 is 1.128 (Shewhart constants) for moving range value that equals 2 (since we are taking the absolute first level difference). So, substituting this value in the control limits equations, the equation becomes $\bar{X} \pm 2.66\overline{MR}$, where \bar{X} is the average of the series and \overline{MR} is the moving range.

According to Nancy R. Tague's 'The Quality Toolbox', the control chart basic procedure is to firstly choose the appropriate control chart for the

data. Secondly, determine the appropriate time period for collecting and plotting data. And third, collect data, construct the chart and analyze the data. Once the chart is plotted one should look for “out-of-control signals” on the control chart. Continue to plot data as they are generated. As each new data point is plotted, check for new out-of-control signals.

2.6 Time Series Forecasting

A time series is a series of information focuses filed (or recorded or diagramed) in time arrange. Most normally, a time series is an arrangement taken at progressive similarly separated focuses in time. Forecasting is a technique that is utilized widely in time series examination to foresee a reaction variable, for example, month to month benefits, stock execution, or unemployment figures, for a predetermined timeframe. Conjectures depend on examples in existing information. One can utilize an assortment of time series techniques, such as trend analysis, decomposition, or single exponential smoothing, to model examples in the information and extrapolate those examples to what's to come.

Ramasubramanian V. of Indian Agricultural Statistics Research Institute(IASR) with expertise in Agricultural Economics, Aquaculture, Artificial Neural Network has discussed various time series models for agricultural forecasting in “Forecasting Techniques in Agriculture” namely:

- Exponential Smoothing Models
- Auto Regressive Integrated Moving Average (ARIMA) Models

So, from the above references it is seen that there is evidence that forecasting of time series has been done in agriculture sector to predict the future crop yields.

2.6.1 Exponential Smoothing Models

Exponential smoothing weights past observations with exponentially decreasing weights to forecast future values. For any time period t , the smoothed value S_t is found by computing where S_i stands for smoothed observation and Y stands for the original value.

$$S_t = \alpha Y_{t-1} + (1-\alpha) S_{t-1}, \quad 0 < \alpha \leq 1, \quad t \geq 3.$$

Alpha (α) is known as the smoothing constant. The optimized value of alpha can be chosen using the value of Mean Absolute Percentage Error (MAPE) i.e. choose the value of alpha for which the value of MAPE is least.

Another form of Exponential Smoothing method is ETS (Error-Trend-Seasonality) exponential smoothing method. The ETS modelling framework was developed in 2002 IJF paper (with Hyndman, Koehler, Snyder and Grose), and in 2008 Springer book (with Koehler, Ord and Snyder). Exponential smoothing methods were originally classified by ‘Pegels’ (1969). This was later extended by Gardner (1985), modified by Hyndman et al. (2002), and extended again by Taylor (2003), giving a total of fifteen methods seen in the following table.

Table 2.5: The fifteen exponential smoothing methods

(**Source:** Automatic Time Series Forecasting: The forecast Package for R)

Trend Component	Seasonal Component		
	N (None)	A (Additive)	M (Multiplicative)
N (None)	N,N	N,A	N,M
A (Additive)	A,N	A,A	A,M
A _d (Additive damped)	A _d ,N	A _d ,A	A _d ,M
M (Multiplicative)	M,N	M,A	M,M
M _d (Multiplicative damped)	M _d ,N	M _d ,A	M _d ,M

According to Rob J. Hyndman and Yeasmin Khandakar, “the cells (N, N) describes the simple exponential smoothing (or SES) method, cell (A, N) describes Holt’s linear method, and cell (A_d, N) describes the damped trend method. The additive Holt-Winters’ method is given by cell (A, A) and the

multiplicative Holt-Winters' method is given by cell (A, M). Equations for method (A, A), the Holt-Winters' additive method.”

Figure 2.6: Equations in ETS Models

(**Source:**Automatic Time Series Forecasting:The forecast Package for R)

Level:	$\ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1})$
Growth:	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$
Seasonal:	$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}$
Forecast:	$\hat{y}_{t+h t} = \ell_t + b_t h + s_{t-m+h_m^+}$

Also, according to Rob J. Hyndman and Yeasmin Khandakar, in the above equation “where m is the length of seasonality, t represents the level of the series, b_t denotes the growth, s_t is the seasonal component, $\hat{y}_{t+h|t}$ is the forecast for h periods ahead. Some interesting special cases can be obtained by setting the smoothing parameters to extreme values. For example, if $\alpha = 0$, the level is constant over time; if $\beta^* = 0$, the slope is constant over time; and if $\gamma = 0$, the seasonal pattern is constant over time.”

2.6.2 Making Time Series Stationary

To check a whether a time series is stationary or not a test known as Breakpoint Unit Root Test is used. This test has a null hypothesis that H_0 : *Series has a unit root.*

If a series has unit root then it a non-stationary series. To check the hypothesis, p-value is seen i.e. if p-value is less than 0.05 then hypothesis is rejected and time series is stationary, where as if p-value is greater than 0.05 then hypothesis is accepted and the time series is non-stationary.

One approach to make a time series stationary is to find the differences between the data recorded. This is known as differencing. Changes, for

example, logarithms can balance out the variations of a time series. Differencing can help balance out the mean of a time series.

- Random walk model

The differenced series is formed by subtracting the next observation from the previous one, and can be written as

$$y'_t = y_t - y_{t-1}$$

The differenced series will have only $t-1$ values since it is not possible to calculate a difference for the first observation. Once the series is stationary one can run time series analysis to perform forecasting.

2.7 Concluding Remarks

Through the review, it has been brought to light that extensive literature is available online regarding marketing information. Also, how AGMARK reports prices and arrivals to deliver timely information was also noted. There are evidences regarding the applicability of statistical process control in non-manufacturing sector and how time series forecasting has been done on the agricultural sector as well. Now, the methodology involved in the study shall be discussed.

Chapter-3

Methodology

3.0 Introduction to Methodology

Analytical research was embraced keeping in mind the end goal to study this venture. Analytical research is a particular sort of research that includes basic deduction abilities and the assessment of truths and data in respect to the research being directed. Inside analytical research articles, information and other essential truths that relate to a venture is gathered; after the data is gathered and assessed, the sources are utilized to demonstrate a speculation or bolster a thought. A person can bring out small details to form greater assumptions about the material by using critical thinking skills effectively.

The research in this project focusses on secondary data obtained from the AGMARKNET.GOV.IN portal. Through this analysis, the study aims to bring to forward the data gaps in order to develop an effective diagnostic tool and a forecasting model.

In this chapter, the various phases involved in the study shall be discussed.

3.1 Project Phases

3.1.1 Understanding Phase

Step 1: To comprehend the broader purpose(s) of the project.

The project aims to develop a diagnostic tool for a particular commodity in order to improve the quality of the data reported in future. Developing a diagnostic tool will help farmers and citizens of the country to have correct information regarding the crop commodity as it will help address the issues such as modal prices reported as zeros, modal prices reported as averages of maximum

and minimum price and modal prices with extreme values. Thereby reducing the error in reporting.

The second objective of the project is to forecast the prices of the chosen commodity - Soyabean. This activity will help farmers and government to understand the prices in the coming months of 2018.

Step 2: To understand the data reporting fields of AGMARKNET portal.

The AGMARKNET portal reports Arrivals (in tonnes) and Prices (in Rs/Quintal) along with crop commodities' state, district, variety, grade and date of reporting.

3.1.2 Defining Phase

Step 1: To choose a commodity for analysis.

The commodity chosen for the project was Soyabean.

Step 2: To define a time frame of the data reporting period for that commodity.

The time frame for the data so that the diagnostic tool can be developed was 01 January, 2017 to 31 December, 2017. Whereas, the time period for data forecasting for the year 2018 was chosen to be 01 January,2014 to 31 December,2017.

Step 3: To identify the key data term(s) to be worked upon for analysis from amongst maximum price, modal price and minimum price.

The data reporting on the AGMARKNET portal consists of maximum price, minimum price and modal price. Out of them modal prices

were chosen for the analysis part as this is the price at which most of the commodity is transacted in a market.

3.1.3 Data Preparation and Analysis Phase

Step 1: Tools chosen for Analysis

The software chosen for development of diagnostic tool was Microsoft Excel 2016, where the softwares chosen for time series forecasting were Microsoft Excel 2016 and Eviews 9.

Step 2: To find the state, district and market with maximum arrivals within the reporting period.

The state with maximum arrivals for soyabean was found out. Then within that state the district and market with maximum arrivals for soyabean were found.

Step3: Check for the reporting frequency of the data fields (modal price and date).

The plot of prices reported per month was done for the whole reporting period to check the consistency in the data.

Step 4: Clean the data for better analysis.

The data gaps were identified and they were filled using the average method in order to have a consistently reported past data which can be worked upon easily to develop diagnostic tool and forecast model.

Step 4: Analyze the data to develop diagnostic tool and forecast model in the next phase.

The data was analyzed for data gaps, zeros and reporting of averages of maximum prices and minimum prices in place of modal prices.

3.1.4 Developing Phase

Step 1: To develop a diagnostic tool incorporating statistical process control charts.

To develop a diagnostic tool, it was necessary to keep the process statistically controlled. For this control chart was used. Various literature pertaining to control charts was studied and the correct chart was chosen according to the best data fit. The chart helped us to get the upper and lower control limits.

Step 2: Develop a time series forecast model for the modal price of the chosen commodity.

To forecast the time series data into 2018, exponential smoothing method was used.

3.1.5 Testing Phase

Step 1: Test the diagnostic tool on other markets reporting the same commodity.

The diagnostic tool thus developed for one market was applied on other markets as well by using their respective control limits to check the applicability of the tool across markets for the same commodity.

Step 2: Check the accuracy of the forecast model.

To check the accuracy of the forecast various parameters like Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) were used.

3.1.6 Findings and Recommendations Phase

After extensive analysis, one will be able to bring out some findings and recommendations through which the AGMARKNET portal can benefit. The

limitations of the diagnostic tool and forecasting model are also be highlighted.

3.2 Concluding Remarks

The approach towards the study of the project and the project methodology was discussed in this chapter. Now in the next chapter the data analysis, findings, limitations and recommendations regarding the study shall be discussed.

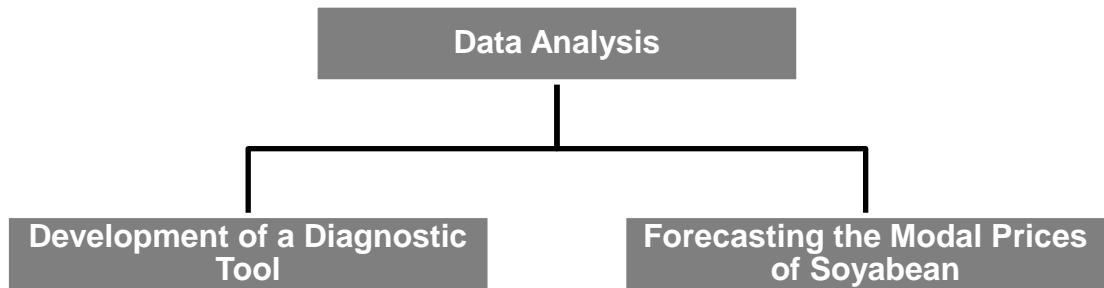
Chapter-4

Data Analysis

4.0 Introduction to Data Analysis

The commodity chosen for analysis in the understanding phase was Soyabean. The commodity data for soyabean obtained from the AGMARKNET portal was usually in .xls format and various operations such as data filtering, data sorting and data cleaning was done in order to draw and make sense out of the compiled data. Since, the project had two major objectives the entire analytics process was classified into two major headings as shown in Figure 4.1

Figure 4.1: Components of Data Analysis



4.1 Development of a Diagnostic Tool

The development of diagnostic tool was a multi phased process. The phases were as follows:

4.1.1 Choosing a Tool to Analyze Data

Since, the data obtained from AGMARKNET was in excel format, so, Microsoft Excel 2016 was chosen to analyze the data and develop a diagnostic tool also.

4.1.2 Selection of Market for Analysis

The arrival data for soyabean for the past one year i.e. from 01 January,2017 to 31 December,2017 was obtained for all the states from the AGMARKNET portal. The arrivals were for each state and the percentage share of total arrivals for each state was found. It was clear to see that Madhya Pradesh was the leader in soyabean arrivals with 3909358.5 tonnes. Hence, the data of the state Madhya Pradesh was analyzed to find the district and market showcasing maximum arrivals in tonnes. The district that gave maximum arrivals was Ujjain with 153554 tonnes of arrival and in that district Ujjain market showed the highest arrivals of 500784.95 tonnes. So, it was decided to develop the diagnostic tool for Ujjain market of Ujjain district in Madhya Pradesh. The arrivals were seen throughout the year making the market appropriate for further analysis.

Table 4.2: Arrival Data for Soyabean 2017 - State
(**Source:** Own Analysis on AGMARKNET Data)

State	Arrivals (Tonnes)
Total	6,635,485.59
Madhya Pradesh	3909358.5
Maharashtra	1792113
Rajasthan	598870.91
Chattisgarh	158582.7
Karnataka	71006
Gujarat	63331.56
Telangana	27185.58
Uttar Pradesh	10714
Uttrakhand	3067.42
Manipur	1216.24
Nagaland	33
Mizoram	6.68

Table 4.2.1: Arrival Data for Soyabean 2017 - District

(Source: Own Analysis on AGMARKNET Data)

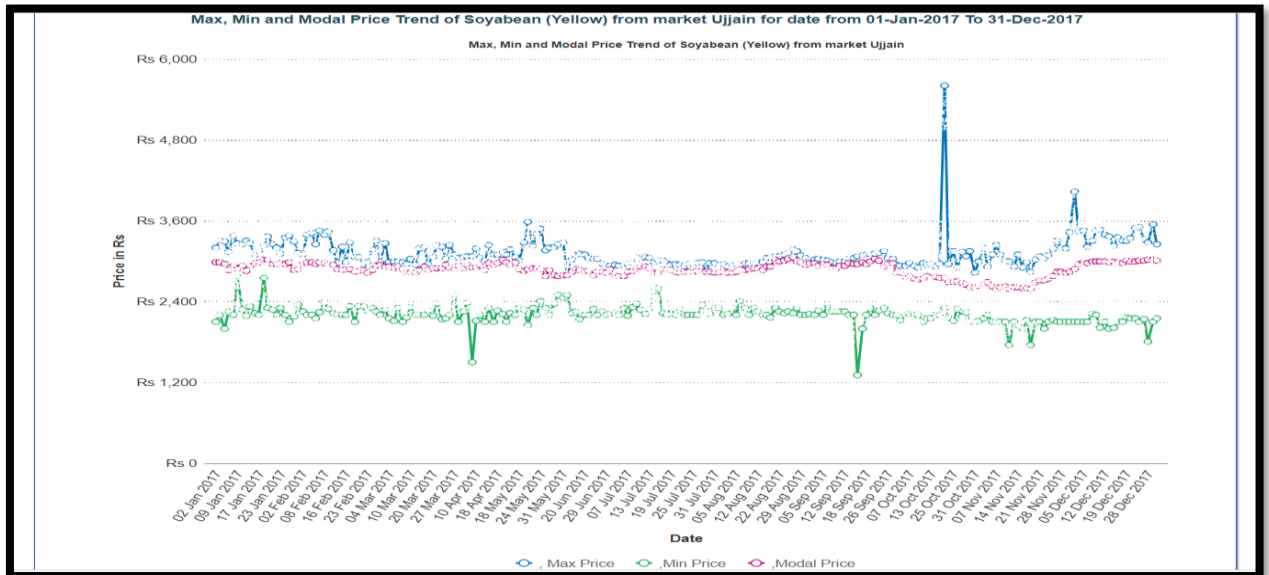
District	Arrival (tonnes)
Ujjain	500784.95
Neemuch	383410.6
Dewas	329310.15
Shajapur	258609.86
Rajgarh	257714.5
Sehore	223234.69
Dhar	217189.79
Indore	214121
Ratlam	196831.89
Mandsaur	185200.59
Sagar	178159.76
Harda	110720.07
Ashoknagar	101874.81
Guna	99581.78
Vidisha	94181.81
Khandwa	78930.75
Khargone	72069.84
Hoshangabad	59756.9
Bhopal	50586.31

Table 4.2.2: Arrival Data for Soyabean 2017 - Market Wise)

(Source: Own Analysis on AGMARKNET Data)

Market	Arrival (tonnes)
Ujjain	153554
Unhel	93132
Badnagar	87236
Mahidpur	68860.2
Tarana	50915
Khachrod	28202.7
Nagda	18884.9

Graph 4.3: Arrival Data for Soyabean for Madhya Pradesh 2017
(Source: AGMARKNET)



4.1.3 Checking for Modal Price Reporting Frequency

Along with arrivals it became necessary to see the frequency with which prices were reported for Ujjain market in Madhya Pradesh. We found the price reporting to be good throughout the year as it reported for 307 days out of 366 days.

4.1.4 Data Cleaning of Modal Prices

The price reporting sheet for Ujjain market was analyzed on the following points

- Modal Price is Average

In the 307 entries, it was checked that how many entries had modal price equal to average of maximum and minimum price by using the “IF” statement in MS Excel. It was found that there were 50 entries in all where modal price was actually average.

- Modal Price is Zero

The reported entries were checked for where the modal price is reported to be zero. No entry was found to be reporting modal price zero.

The entries having modal price as average of maximum and minimum prices were deleted. 257 entries were now reported after deletion and by using VLOOKUP command PIVOT TABLE the daily data gaps were filled by the average modal price of that month. Thereby, having a continuous series.

4.1.5 Developing a Control Limits and Control Charts

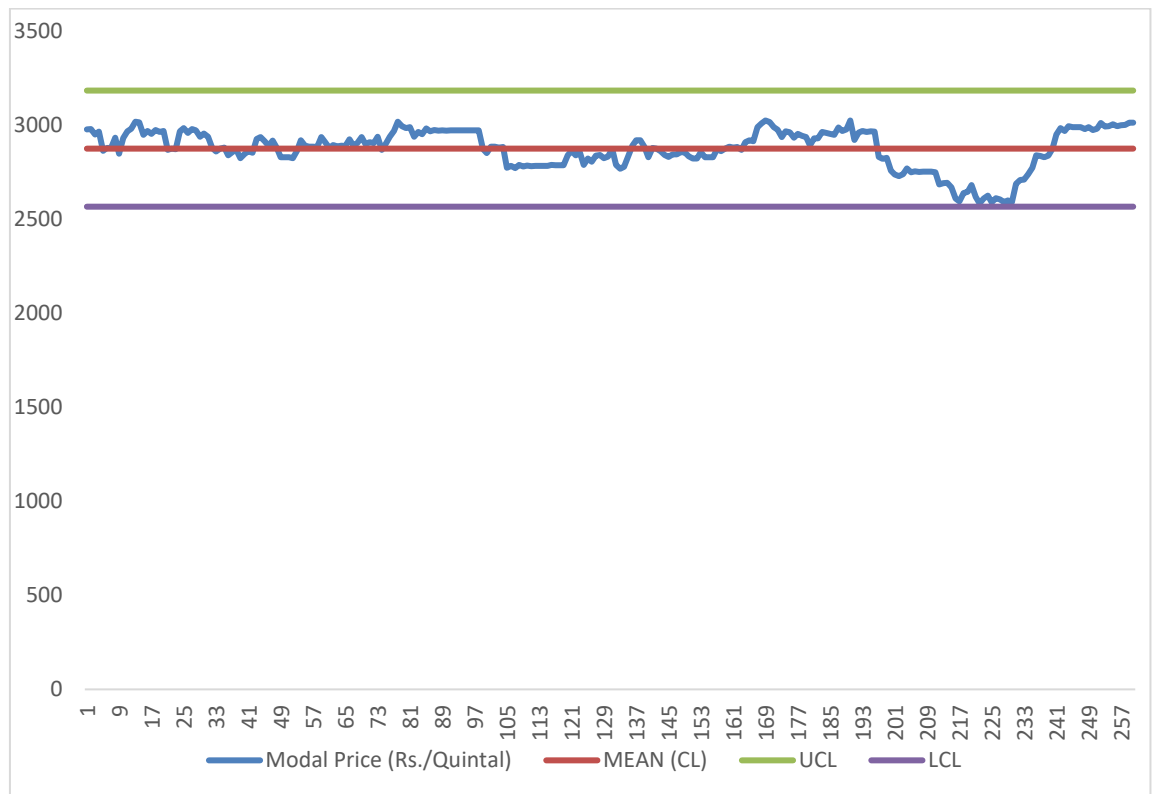
Once the data was properly cleaned it was necessary to develop control limits. For this statistical process control charts were referred. Since, the modal prices were of variable nature, measured on a continuous scale of daily dates and the data set had a subgroup size of one (as each day's price data cannot be combined with the next day's data), so, the Individual Moving Range Chart was chosen to get the upper control limit and lower control limit. The equation for control limits for I-MR chart is:

$$\bar{X} \pm \frac{3\overline{MR}}{d_2}, \quad \text{Where } E_2 = \frac{3}{d_2} \text{ and } d_2 = 1.128.$$

Therefore, the equation becomes,

$$\bar{X} \pm 2.66\overline{MR}$$

Graph 4.4: Control Chart for Ujjain Market 2017
 (Source: Own Analysis on AGMARKNET Data)



To reduce the effect of seasonality, factors were developed for each month of the entire year 2017.

4.1.6 Developing the Interface of the diagnostic tool

The diagnostic tool was a standalone excel sheet where users can enter the data for modal price to check whether the price values were within the range or not. The excel sheet also consisted of dynamic Control Charts, reflecting new data value entered dynamically.

If the data entry operator wished to proceed with the reporting prices, then, two checks were developed. These checks would keep the error messages on record if the data entry operator still proceeded with the price reporting.

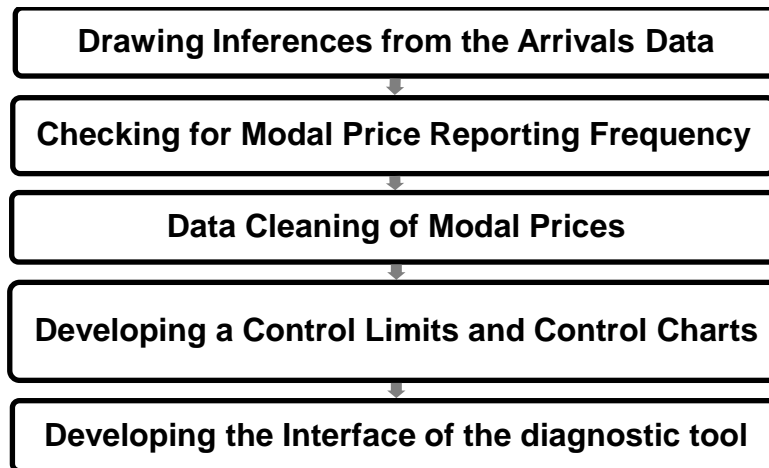
Check 1 took care of the error if the modal price thus reported was out of bounds and gave a red flag “Check Modal” or if the price is within limits then a green flag “Go”. Similarly, Check 2 took care if the modal price was the average of maximum and minimum prices by giving a red flag “Modal is Average” or giving a green flag “Go” if modal is not average, dynamically.

The tool also showed the percentage of modal price data reported being within limits and the number of entries where modal price is the average of maximum and minimum prices.

Figure 4.5: Look of the diagnostic tool
(Source: Own Analysis on AGMARKNET Data)

	A	B	C	D	E	F	G	H	I	J	K
	Min Price (Rs./Quin tal)	Max Price (Rs./Quin tal)	Modal Price (Rs./Quin tal)	Price Date	Check 1	Check 2					
1											
2	2211	3425	3027	01-Jan-18	Go	Go					
3	2200	3289	3001	02-Jan-18	Go	Go		LCL	UCL	Within Limits	Modal is Average
4	2100	3451	3041	03-Jan-18	Go	Go		2478.148	3200.831	0.1	0
5	2151	3333	3050	04-Jan-18	Go	Go					
6	2101	3521	3140	05-Jan-18	Go	Go					
7	1600	3353	3140	06-Jan-18	Go	Go					
8	1850.5	3437	3140	07-Jan-18	Go	Go					
9	2130	3570	3140	08-Jan-18	Go	Go					
10	2351	3521	3177	09-Jan-18	Go	Go					
11	2200	3599	3216	10-Jan-18	Check Modal	Go					
12	2050	3627	3229	11-Jan-18	Check Modal	Go					
13	2100	3601	3216	12-Jan-18	Check Modal	Go					
14	2075	3614	3222.5	13-Jan-18	Check Modal	Go					
15	2087.5	3607.5	3219.25	14-Jan-18	Check Modal	Go					
16	2225	3586	3256	15-Jan-18	Check Modal	Go					
17	2156.25	3596.75	3237.625	16-Jan-18	Check Modal	Go					
18	2114	3765	3241	17-Jan-18	Check Modal	Go					
19	1726	3801	3280	18-Jan-18	Check Modal	Go					
20	1971	3829	3290	19-Jan-18	Check Modal	Go					
21	2005	3831	3318	20-Jan-18	Check Modal	Go					
22	1988	3830	3304	21-Jan-18	Check Modal	Go					
23	1996.5	3830.5	3311	22-Jan-18	Check Modal	Go					
24	2126	3756	3413	23-Jan-18	Check Modal	Go					
25	2162	3832	3480	24-Jan-18	Check Modal	Go					
26	2155	3772	3585	25-Jan-18	Check Modal	Go					

Figure 4.6: Diagnostic Tool Development



4.2 Forecasting the Modal Prices of Soyabean

In order to forecast the soyabean modal prices for the year 2018 based on the modal prices of the previous years following steps were followed:

4.2.1 Choosing the Tools of Analyses

To develop a better forecast model two tools were used

- Microsoft Excel 2016: For sorting, filtering and cleaning data
- Eviews 9: To help choose a forecast model

4.2.2 Drawing Inferences from the Arrivals Data

It was aimed to develop forecast model for one of the top Soyabean producing markets of Madhya Pradesh, and for that it became necessary to look into the arrivals data of Madhya Pradesh's markets.

Table 4.7: District Leading in Soyabean Arrivals 2017
(Source: Own Analysis on AGMARKNET Data)

District	Arrival (tonnes)
Ujjain	500784.95
Neemuch	383410.6
Dewas	329310.15
Shajapur	258609.86
Rajgarh	257714.5
Sehore	223234.69
Dhar	217189.79
Indore	214121
Ratlam	196831.89
Mandsaur	185200.59
Sagar	178159.76
Harda	110720.07
Ashoknagar	101874.81
Guna	99581.78
Vidisha	94181.81
Khandwa	78930.75
Khargone	72069.84
Hoshangabad	59756.9
Bhopal	50586.31

During the period of 01 January 2014 to 31 December 2017, Ujjain reported only one variety of soyabean, which was “Yellow.

Table 4.8: Market Leading in Soyabean Arrivals 2017
(Source: Own Analysis on AGMARKNET Data)

Market	Arrival (tonnes)
Ujjain	153554
Unhel	93132
Badnagar	87236
Mahidpur	68860.2
Tarana	50915
Khachrod	28202.7
Nagda	18884.9

4.2.3 Data Cleaning of Modal Prices

To forecast the data, it was necessary to fill the data gaps and improve the quality of data. The months which had more than 15 days of reporting of modal prices had their remaining non-reported days filled with the averages of the past 2 days modal prices. Additionally, there were instances where the modal prices for Saturdays & Sundays were reported. As the mandis are closed on these 2 days, the data of these days was removed.

4.2.4 Choosing a Forecast Model and Technique

Since, the series is a time series a popular technique of forecasting was used viz. Exponential Smoothing Method has been used for forecasting. To choose the right fit model the software EViews version 9 has been used.

First of all, it was checked that was the series stationary or not. Using the Breakpoint Unit Root Test which is based on the Dickey-Fuller test, the p-value came out to be 0.4271 which means that the null hypothesis "Series has a unit root" was not rejected, this acceptance means that the time series is non-stationary in nature.

So, in order to make the time series stationary first level difference was taken i.e. $\text{Difference}_{n+1} = \text{Price}_{n+1} - \text{Price}_n$. We again ran the unit root test on it based on Dickey-Fuller test to get the p-value. The p-value this time was found to be less than 0.01, hence the null hypothesis was rejected, meaning that it does not have a unit root and thus is stationary in nature.

Figure 4.9: Test for Stationarity After Difference
 (Source: Own Analysis on AGMARKNET Data Using EViews)

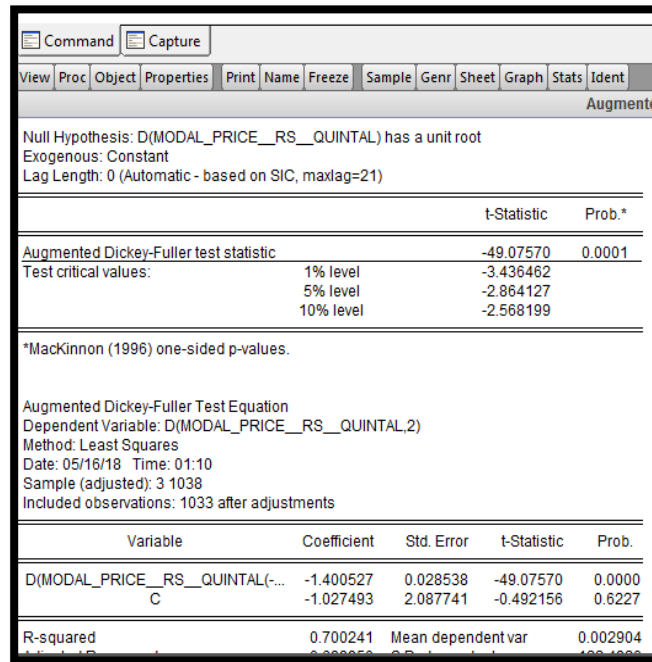
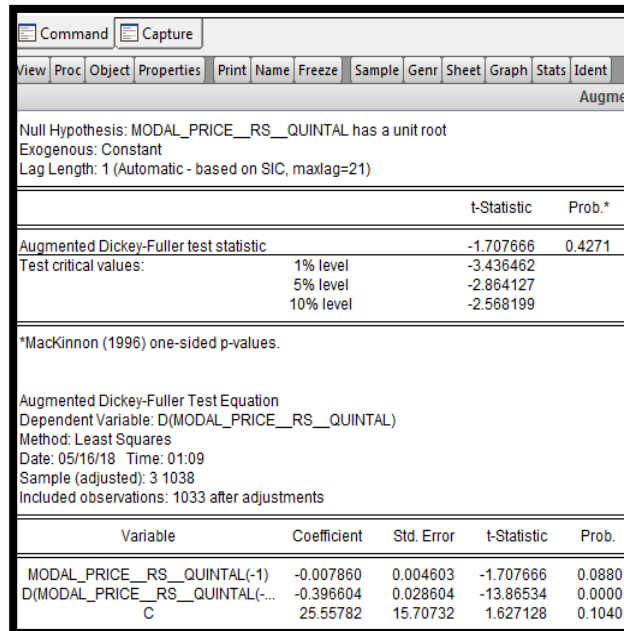


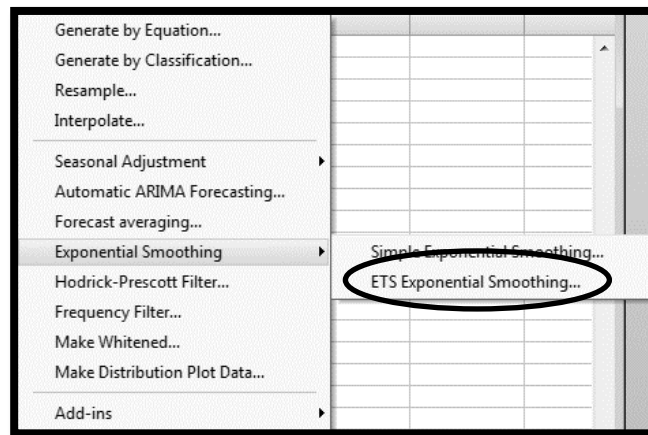
Figure 4.10: Stationary Series Plot
 (Source: Own Analysis on AGMARKNET Data)



Thus, a stationary series at first level of difference was found.

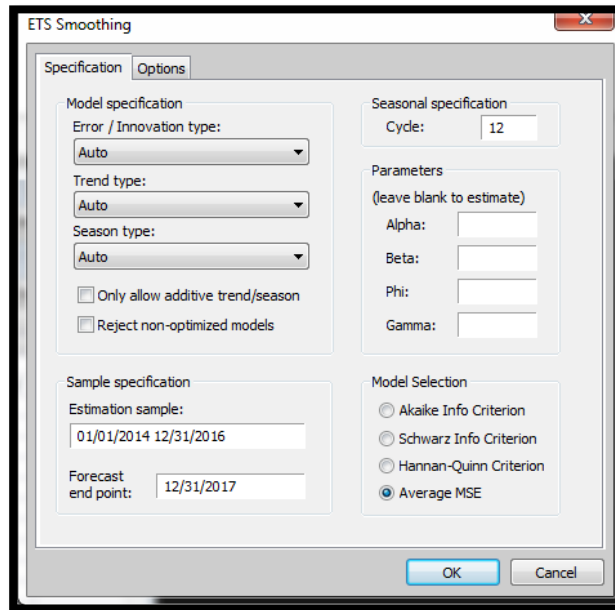
Now, EViews has a special command titled “Exponential Smoothing” which helps users to find the best fit model out of the given models. The ETS(Error-Trend-Seasonality) Exponential Smoothing of EViews gives the best fit model according to the select criteria and justifies the chosen model statistically.

Figure 4.11: ETS Exponential Smoothing-EViews
(Source: EViews)



The ETS exponential smoothing method was first employed on the period 01 January 2014 to 31 December 2016 to forecast the modal prices for the month January to December 2017. Since the data for 2017 was already present it was easy to check the model's accuracy by finding out various errors.

Figure 4.12: ETS Model Options
(Source: EViews)



In order to find the appropriate model, the value of Error, Trend and Seasonality was set to “AUTO”. The model selection criteria were set to Average MSE option. To further optimize the model the objective was set to Average MSE the cyclicity for the forecast was set as 12, since it focusses on 12 months a year. After running the operation, it was found that Eviews by using ETS Exponential Smoothing Technique gave the model A, A, A. It gave the value of alpha as 0.7200, beta as 0.0100. The MSE comparison made it evident (as visible in the screenshots below) that the A, A, A model has the least MSE as compared to other ETS exponential smoothing models. Thus, the model was applied on the modal price data from 01 January, 2014-31 December 2016, to get the forecast for 2017.

Figure 4.13: ETS Model Parameters and Comparisons
(Source: EViews)

Parameters:	Alpha	0.7200
	Beta	0.0100

4.2.5 Forecast Errors for the Months of the 2017

Once the model was applied and the forecast of the model prices were obtained month wise, then, the mean absolute percentage error and mean absolute error were calculated.

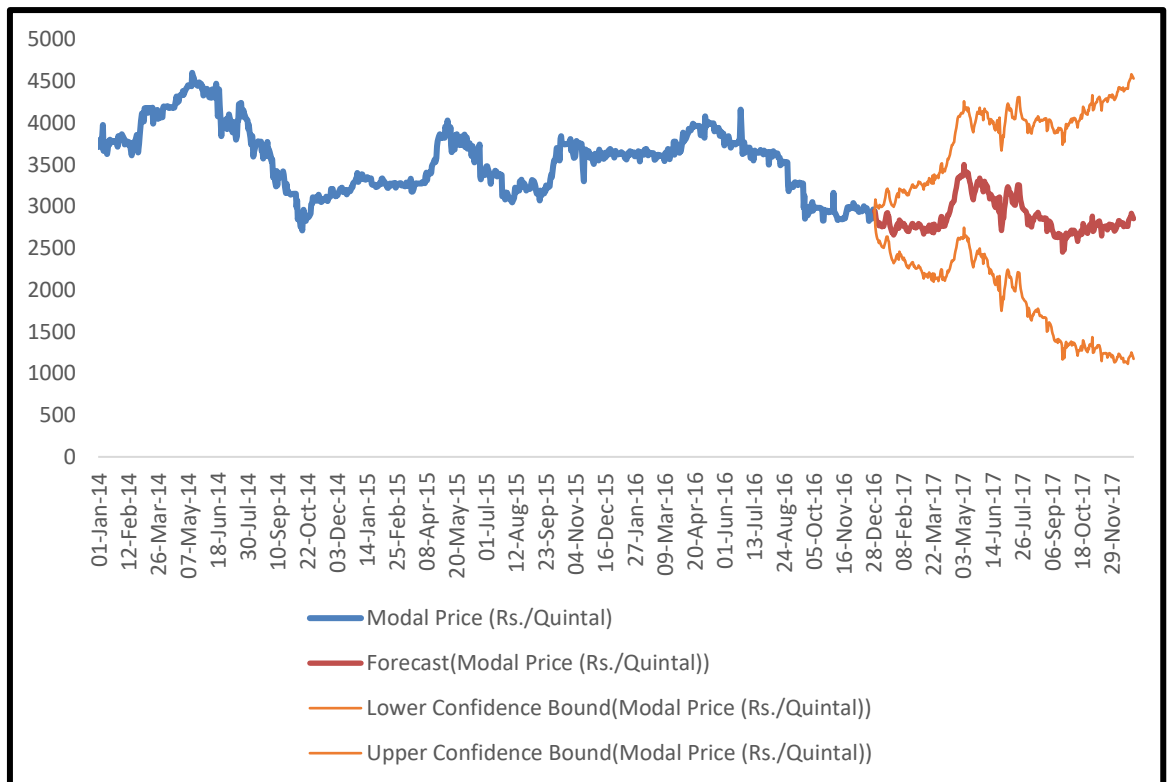
Table 4.14: Forecast Errors

(Source: Own Analysis on AGMARKNET Data)

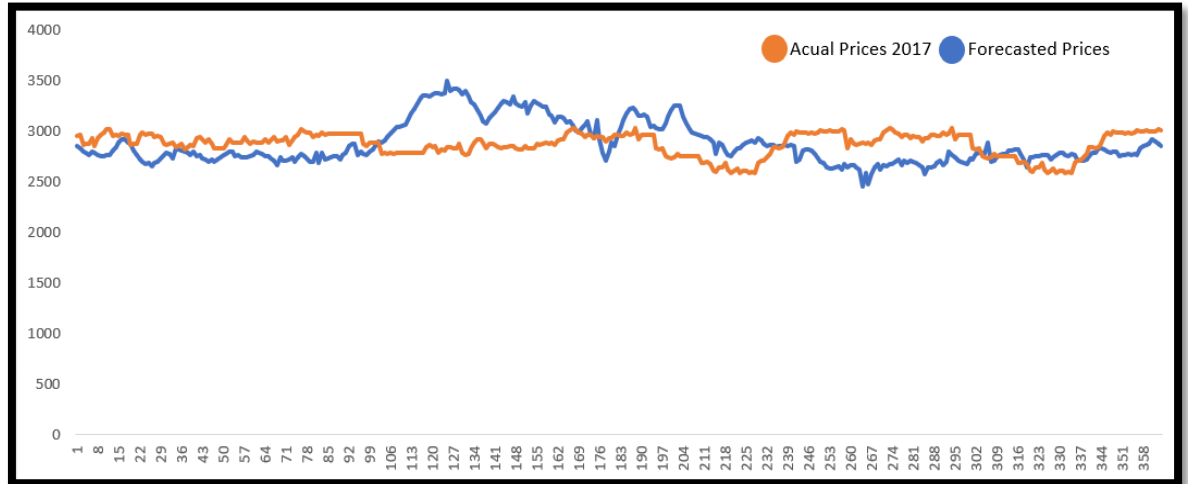
Errors	Value
Mean Absolute Percentage Error(MAPE)	3.99%
Mean Absolute Error(MAE)	71

Graph 4.15: Forecast for 2017

(Source: Own Analysis on AGMARKNET Data)



Graph 4.16: Comparison for Actual Values & Forecasted Values for 2017
 (Source: Own Analysis on AGMARKNET Data)



The forecasted values were close to the actual values of model price, and hence the data from 01 January 2014 to 31 December 2016 was used to forecast the values of various months of 2017 using the same model constraints obtained earlier.

4.3 Findings

- Once the diagnostic tool was prepared it was seen that what percentage of values of the months January, February and March for the year 2017 actually fell within limits. It was also seen whether the seasonal correction improved the tool or not. It was evident that the control limits were increased marginally to include nominally larger or smaller modal prices. Thus, the diagnostic tool was improved.

Table 4.17: Control Limits

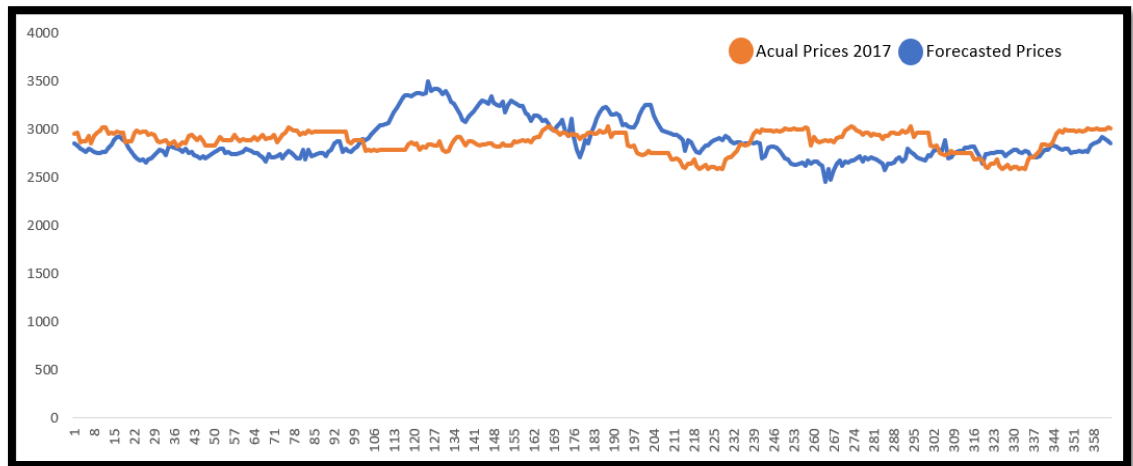
(Source: Own Analysis on AGMARKNET Data)

LCL	UCL
2478.148	3200.831

- The forecast model (A, A, A) obtained after ETS exponential smoothing was applied to the previous data to get forecasts for the months of January, February and March 2017. The MAPE for first three months of 2017 was 7.92% whereas the MAE was Rs 136/Quintal. The Market was Ujjain and soyabean variety Yellow.

Graph 4.18: Actual and Forecast Values

(Source: Own Analysis on AGMARKNET Data)



- It was seen that the analysis can go bad if the data is being changed continuously. The data range for arrivals was from 01 January 2017 to 31 December 2017.
- The total of arrivals of soyabean on AGMARKNET portal for the state of Madhya Pradesh was reported to be 944262.23 tonnes, whereas, upon downloading the file the total was 7178623.57 tonnes which was the actual value obtained after using the sum rows formula in excel.

Thus, higher arrivals were being shown on the AGMARKNET portal before downloading.

4.4 Limitations

- The analysis was done only on modal price and not on maximum and minimum prices.
- The data used for the development of the diagnostic tool was only of one year i.e. of 2017. To develop a better tool the data horizon can be widened to include previous year data.
- The diagnostic tool needs to be developed for each market separately which can make the task cumbersome.
- The data used for forecasting the modal prices of 2018 was taken from 01 January 2014 to 31 December 2017, giving a forecast horizon of only four years which can be made large.
- The forecast was done only for one market and a specific variety of Soyabean. Thus, it can only be applied to that market and that variety of soyabean.

4.5 Concluding Remarks and Recommendations

The time series forecasting for Yellow variety of Soyabean for a chosen market (Ujjain-Madhya Pradesh) was done. To conclude the data analysis of the project, here are some recommendations.

- The diagnostic tool developed as part of the project work can be accommodated alongside the data entry reporting portal of soyabean prices i.e. http://agmarknet.nic.in/market_online/. Integration with the

portal will give the checks and raise alerts then and there itself, making the data entry process effective and thereby improving the data quality of reporting.

- The price data reporting frequency should be monitored daily as it can help provide the actual data. This actual data can in turn help in developing robust diagnostic tools and accurate forecast models.
- The time series forecast in the case of this study was done for a specific variety of soyabean and for a specific market but instead the forecast model needs to be developed for all varieties of soyabean across various markets.
- The variety terms of Soyabean on AGMARKNET portal like “Others” needs to be classified as well to maintain effective tracking on the Soyabean arrivals.
- In the findings of the study it was seen that the data reported historically pertaining to an agricultural commodity can be edited and changed. The editing of already reported data can disturb and hamper the entire data analysis of the project and thus, the changing of the data should be done in a restrictive manner to ensure integrity of reported data.

Chapter 5

Conclusion of the Study

The project began by understanding the agencies associated with AGMARKNET. Also, the importance of the study and objectives of the study were highlighted which gave a direction to develop a well-planned methodological approach to the project study.

Further on in the project, the review of literature had highlighted the importance of the information particularly when it comes to the aspect of agricultural marketing information. The review of literature had also brought to light the need manage and improve the quality of data being reported. Various definitions of marketing information were obtained. The working of AGMARKNET, an agricultural commodity data reporting portal which can help in bringing insightful information to the farmers, government and consumers was understood. It was also seen how the data is being reported on this portal. The agricultural commodity's data worked on in this project was that of Soyabean. Time series forecasting was done for the Yellow variety of soyabean. It was also necessary to understand some facts related to soyabean like major soyabean producing states, temperature and rainfall conditions and trade policies, and the same was done in this study.

The study had also highlighted how statistical process control can be used in non-manufacturing sector to continuously monitor the data being reported. In this project, it was also emphasized that which statistical process control charts can be used to have mechanism of check on the prices being reported on the AGMARKNET portal for a particular commodity such as soyabean. The equations of upper control limit, lower control limit and control limit were highlighted for variable process control chart i.e. individual moving range chart. It was also discussed that how the control limits can be improved by developing seasonal factors for each month so that marginally neglected prices can be accommodated. The diagnostic tool developed on the platform of Microsoft Excel was on the basis of statistical process control charts for one major Soyabean arrival market of

Madhya Pradesh. To check whether the approach adopted for developing the diagnostic tool was scalable to other markets or not, the same approach was applied to other four markets of Madhya Pradesh. The diagnostic tool also took care of the modal prices that were being reported as average of maximum and minimum prices and reported a warning whenever modal prices was average.

Also, the study of this project emphasized the importance of time series forecasting particularly when it comes to the prices of agricultural commodities. Through the review of the available literature it was seen that what were the different models and techniques to analyze the time series. With a well-defined mechanism for data gap filling and after applying data filtering and data sorting in Microsoft Excel 2016, EViews 9 was used to develop a suitable model to forecast the modal prices for Soyabean for the year 2017. Using the ETS exponential smoothing, EViews gave a particular model, this model was applied to a subset of the dataset and was checked for errors. Then the ETS exponential model was applied on the entire data set to get the forecasted values for the year 2017. Till March 2018, the mean absolute error percentage was 7.92%. The forecasting of time series was done for one particular market of Madhya Pradesh and for one specific variety of yellow - Ujjain.

During the study, it was also seen that AGMARKNET had some data reporting inconsistencies. Along with the findings, the study also highlighted the findings and recommendations of the study.

Thus, we see that by developing a diagnostic tool for various markets can help in improving the quality of data being reported as it will create a mechanism of checks and balances. The need for developing these kinds of diagnostic tools across all markets and for all varieties of agricultural commodities was emphasized in the study. Whereas, the forecasted modal prices will assure the farmers, government agencies and consumers to get a prior information regarding the agricultural commodities. It was also highlighted that with a wider range of data set the forecasted model thus developed can be more accurate.

Bibliography

- Biswas, A. (2004, JULY 13). *domain-b*. Retrieved from domain-b: http://www.domain-b.com/economy/agriculture/20040713_marketing.html
- Department of Agriculture Cooperation and Farmers . (2017, March 10). *Commodity Profile of Soyabean for March 2017*. Retrieved from <http://agricoop.nic.in/>: <http://agricoop.nic.in/sites/default/files/Soyabean.pdf>
- Hyndman, R. J., & Khandakar, Y. (2008). Automatic Time Series Forecasting: The Forecast Package for R. *Journal of Statistical Software*, 2-3.
- InfinityQS International, I. (2014, January). A Practical Guide to Selecting the Right Control Charts. Fairfax, Virginia, United States of America.
- Milivojevic, A. (2015, December 29). *Control Chart Constants – How to Derive A2 and E2*. Retrieved from andrewmilivojevic.com: <https://andrewmilivojevic.com/control-chart-constants-how-to-derive-a2-and-e2/>
- Ramasubramanian. (n.d.). *Forecasting Techniques in Agriculture*. New Delhi: Indian Agricultural Statistics Research Insitute.
- S. , A. S. (2008). *Agricultural Marketing In India*. In A. S. S., *Agricultural Marketing In India* (pp. 1-8). Oxford & Ibh.
- Spinks, G. R. (n.d.). *Review of Marketing and Uses and Methods of Crop Forecasting*.
- *Statistical Process Control (SPC)*. (n.d.). Retrieved from www.moresteam.com: <https://www.moresteam.com/toolbox/statistical-process-control-spc.cfm>
- (1995). *The State of Food and Agriculture*. Rome: Food and Agriculture Organization of the United Nations.
- Thomsen, F. L. (1951). *Agricultural Marketing*.
- UNIVERSITY, T. N. (2011). *eagri.tnau.ac.in*. Retrieved from eagri.tnau.ac.in/eagri50/AECO242/pdf/lec01.pdf
- *www.internetlivestats.com*. (2017, March). Retrieved from [internetlivestats](http://internetlivestats.com).