**A STATISTICAL STUDY ON ORGANIC FARMING IN WAYANAD**

A Project Report submitted to **UNIVERSITY OF CALICUT** in partial fulfilment of the requirements for the award for Degree of

**BACHELOR OF SCIENCE IN STATISTICS**

**2017-2020**

Submitted By

NAME: REG.NO:

AVELINE MARIYA SHAJI - OSARSST001

LINTA STEPHEN - OSARSST002

MOHAMMED SHIYAS T - OSARSST003

SREEHARI N.M - OSARSST004

Under the guidance and supervision of

Ms. VARSHA C

(Asst. Professor, Department of Statistics)



DEPARTMENT OF STATISTICS

ORIENTAL SCHOOL OF HOTEL MANAGEMENT

LAKKIDI, WAYANAD

**DECLARATION**

We hereby declare that the project, “A STATISTICAL STUDY ON ORGANIC FARMING IN WAYANAD”, is an original work done by us under the guidance and supervision of Ms. VARSHA C, Asst. Professor, Department of Statistics, Oriental School of Hotel Management Lakkidi. This project is submitted to UNIVERSITY OF CALICUT in partial fulfilment of the requirements for the award of degree of Bachelor of Science in Statistics. We also declare that this project has not been previously formed the basis of the award for any other degree/diploma/other similar courses.

AVELINE MARIYA SHAJI

LINTA STEPHEN

MOHAMMED SHIYAS T

SREEHARI N.M

**CERTIFICATE**

This is to certify that the bona-fide students of this institution, Ms. AVELINE MARIYA SHAJI with register number OSARSST001, Ms. LINTA STEPHEN with register number OSARSST002, Mr. MOHAMMED SHIYAS T with register number OSARSST003 and Mr. SREEHARI N.M with register number OSARSST004 has done the project work entitled A STATISTICAL STUDY ON ORGANIC FARMING IN WAYANAD for the partial fulfilment of the requirement for the award of degree of Bachelor of Science in Statistics of Oriental School of Hotel Management under my supervision and guidance.

Ms. VARSHA C Mrs. DHANYA K G

Research Guide Research Coordinator

Dr. K C ROBBINS

Principal

Submitted for the Research verification and viva-voice held on ……………………...

INTERNAL EXAMINER EXTERNAL EXAMINER

DATE :

COLLEGE SEAL:

**ACKNOWLEDGEMENT**

It is the profound gratitude that we submit this project for the kind consideration of all those concerned. The success and final outcome of this project required a lot of guidance and assistance from many people and we are extremely privileged to have got this all along the completion of our project. All that we have done is only due to such supervision and assistance and we would like to extend our thanks to all of them. Hence let us place on record of all those who supported and assisted us in completion of this project.

We thank GOD, the Almighty, under whose blessing we did everything we could for the completion of this project.

We respect, thank and owe our deep gratitude to Dr. K C ROBBINS, Principal, Oriental School of Hotel Management Lakkidi, for providing us an opportunity to do the project work and giving all support and guidance which help in bringing out this project.

We are extremely thankful to Mr. VINU GEORGE, Vice Principal, Oriental School of Hotel Management Lakkidi for his valuable support.

We extent our heartfelt thanks and gratitude to Mrs. SIAVUDEEN V S, Add. Vice Principal, Oriental School of Hotel Management Lakkidi, for his valuable advices, support and encouragement to do the project work.

We are fortunate and extremely thankful to Mrs. DHANYA K G, Head of the Department of Statistics, for her constant supervision as well as for providing necessary information regarding the project and also for the support in completing the project.

We owe our deep gratitude to our project guide Ms. VARSHA C, Asst. Professor, Department of Statistics, who took keen interest on this project work and guided all along, till the completion of this project work by providing all necessary information for developing a good project report.

We are extremely thankful to Mrs. CHINCHU VARGHESE, Asst. Professor, Department of Statistics, for her valuable support and advice for helping us to complete this project on time

Above all, we heartily thank our parents for helping us to complete this entire project and for their support. And we also thank our friends and family for their support and co-operation.

**ABSTRACT**

Statistics plays an important role in virtually all aspects of human life. This role is used to analyse a project, which aimed at analysing the level of organic farming practice in Wayanad district, Kerala. The main objective of this project is to analyse the data collected directly from people of Wayanad and carry-out a statistical analysis on the development of Organic farming in Wayanad, Kerala.

Chapter 1 contains the introduction and aims and objectives of the study. Chapter 2 consists a note of various studies about organic farming. This chapter holds the Review of Literature for this project. Chapter 3 comprises of the information regarding the different terms used in statistics and all the statistical tools that has been used for the successful completion of this project work. The analysis procedures and results of the analysed data which is obtained by the use of Chi-square test, ANOVA test, Wilcoxon rank sum test and also figures and plots such as bar diagram, pie-chart and line plots are included in Chapter 4. The final chapter, Chapter 5 contains conclusion and findings, reference and questionnaire.

We hope that there will be something to comprehend and learn out of this project, which will stand useful to understand as a source of data analysis.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| CHAPTERS | TITLE | PAGE NUMBER |
| **Chapter 1** | **INTRODUCTION** | 1-4 |
| 1.1 INTRODUCTION | 2 – 3 |
| 1.2 AIMS AND OBJECTIVES OF THE STUDY | 4 |
| **Chapter 2** | **REVIEW OF LITERATURE** | 5 – 8 |
| 2.1 AGRICULTURE AND ENVIRONMENT | 6 |
| 2.2 SUSTAINABLE FARMING | 6 |
| 2.3 INTERNATIONAL STUDIES ON ORGANIC FARMING | 7 |
| 2.4 STUDIES ON ORGANIC FARMING IN INDIA | 8 |
| **Chapter 3** | **METHODOLOGY** | 9 – 18 |
| 3.1 SOURCE OF DATA | 10 |
| 3.2 DEFINITION OF ITEMS | 10 – 11 |
| 3.3 TYPE OF STATISTICAL DATA USED | 11 – 12 |
| 3.4 SAMPLE AND SAMPLING TECHNIQUE | 12 |
| 3.5 STATISTICAL TESTS USED | 12 – 15 |
| 3.6 DIAGRAMS AND GRAPHS | 15 – 16 |
| 3.7 SOFTWARE USED | 16 – 18 |
| **Chapter 4** | **DATA ANALYSIS** | 19 – 38 |
| 4.1 INTRODUCTION | 20 |
| 4.2 SOCIO-DEMOGRAPHIC CHARACTERISTICS | 20 |
| 4.3 DESCRIPTIVE ANALYSIS | 21 - 30 |
| 4.4 ANALYSIS OF RELATION AMONG VARIABLES | 30 - 38 |
| **Chapter 5** | **FINDINGS, CONCLUSION AND SUGGESTIONS** | 39- 41 |
| 5.1 FINDINGS AND CONCLUSION | 40 |
| 5.2 BIBLIOGRAPHY | 41 |
|  | ANNEXTURE | 42-47 |

**LIST OF TABLES**

**CHAPTER 1**

**INTRODUCTION**

**INTRODUCTION**

Sustainable development has caught the imagination and action all over the world for more than a decade. Sustainable agriculture is necessary to attain the goal of sustainable development. According to the Food and Agriculture Organization (FAO), sustainable agriculture "is the successful management of resources for agriculture to satisfy changing human needs while maintaining or enhancing the quality of environment and conserving natural resources". All definitions of sustainable agriculture lay great emphasis on maintaining an agriculture growth rate, which can meet the demand for food of all living things without draining the basic resources.

Organic farming is one of the several approaches found to meet the objectives of sustainable agriculture. Many techniques used in organic farming like inter-cropping, mulching and integration of crops and livestock are not alien to various agriculture systems including the traditional agriculture practiced in old countries like India. However, organic farming is based on various laws and certification programmes, which prohibit the use of almost all synthetic inputs, and health of the soil is recognised as the central theme of the method.

Adverse effects of modern agricultural practices not only on the farm but also on the health of all living things and thus on the environment have been well documented all over the world. Application of technology, particularly in terms of the use of chemical fertilizers and pesticides all around us has persuaded people to think aloud. Their negative effects on the environment are manifested through soil erosion, water shortages, soil contamination, genetic erosion, etc.

The farming being practiced for the last three decades in India has increasingly been found non-sustainable. The system is oriented towards high production without much concern for ecology and the very existence of man himself.

Organic farming is an alternative agricultural system which originated early in the 20th century in reaction to rapidly changing farming practices. Certified organic agriculture accounts for 70 million hectares globally, with over half of that total in Australia. Organic farming continues to be developed by various organizations today. It is defined by the use of fertilizers of organic origin such as compost manure, green manure, and bone meal and places emphasis on techniques such as crop rotation and companion planting. Biological pest control, mixed cropping and the fostering of insect predators are encouraged. Organic standards are designed to allow the use of naturally occurring substances while prohibiting or strictly limiting synthetic substances.

Kerala is going through a minor revolution in what it eats. As urban residents across the state have grown anxious over fruits and vegetables doused in pesticide, an organic kitchen garden movement has been quietly gaining momentum. As the organic food movement has grown, it is drawn support from the media, the government, political organisations and many other groups. But it all started with one institution – the Kerala Agriculture University (KAU).It was the KAU’s studies in the past few years, on the levels of insecticides and pesticides in most commonly consumed vegetables that sparked public fears, and began receptivity towards organic food.

Biology, physics, chemistry, meteorology, sociology, communication, and even information technology all use statistics. For many of these categories, the use of statistics in that field involves collecting data, analysing it, coming up with a hypothesis, and testing that hypothesis.

Statistics is playing an important role in virtually all the aspects of human life. This role can be used to analyse any kind of data.

Wayanad being predominantly a district with majority of people following agriculture, it is relevant to analysis the growth of organic farming practice in Wayanad.

**AIMS AND OBJECTIVES OF THE STUDY**

The main objective of the study is to statistically analyse the development of the organic farming in Wayanad.

In general, the study aimed at the followings:

* To analyse whether there is any relation between the crops cultivated and organic farming.
* To analyse whether there is any relation between external financial support and organic farming.
* To analyse whether there is an increase or decrease in the number of Certified Organic Farmers.
* To show and study which age group follows organic farming more.
* To check whether there is a relationship between education and organic farming.
* To analyse whether there is any relation between the area of land and organic farming.
* To show the increase or decrease happened in practising organic farming in Wayanad.

**CHAPTER 2**

**METHODOLOGY**

* SOURCE OF DATA

The research titled “A statistical study on the organic farming in Wayanad” made use of the primary data collected from the farmers in Wayanad. The primary data was collected using questionnaire and doing direct surveys along different places in Wayanad.

* DEFENITION OF ITEMS

* + - POPULATION: A population is an aggregate of items, units or objects that are under reference for the study. In statistics, a population is the entire pool from which a statistical sample is drawn. A population may refer to an entire group of people, objects, events, hospital visits or measurements. A population can thus be said to be an aggregate observation of subjects grouped together by a common feature. It may consist of finite or infinite number of units. Hence collected data with one variable is called “univariate population”, two variables are called “bivariate population” and the population is called “multivariate population” when more than two variables are studied.
    - SAMPLE: A sample is a random selection of members of a population. It is a smaller group drawn from the population that has the characteristics of the entire population. The observations and the conclusions made against the sample data are attributed to the population. The information obtained from the statistical sample allows statisticians to develop hypothesis about the larger population. Number of objects in the sample is called “the size of the sample”. The process of obtaining suitable samples from a population is called “sampling”.
    - POPULATION PARAMETERS: A parameter is data based on an entire population. Statistics such as averages and standard deviations. When taken from populations, are referred to as population parameters. The standard deviation is the variation in the population inferred from the variation in the sample. When the standard deviation is divided by the square root of the number of observations in the sample, the result is referred to as the standard error of the mean.
    - DATA: A set of observations (a set of possible outcomes); most data can be put into two groups: qualitative (an attribute whose value is indicated by a label) or quantitative (an attribute whose value is indicated by a number). Quantitative data can be separated into two subgroups: discrete and continuous. Data is discrete if it is the result of counting. Data is continuous if it is the result of measuring.
    - AVERAGE: It is also called as mean. A number that describes the central tendency of the data is known as the Average.
    - PROBABLITY: A number between zero and one that gives the likelihood that a specific event will occur or not.
    - VARIABLE: Variable is a characteristic of interest for each person or object in a population. A variable is called a numerical variable if that takes on values that are indicated by numbers.
* TYPES OF STATISTICAL DATA :

The data collected may be:

* PRIMARY DATA: it means the raw data which has just been collected from the sources and has not gone any kind of statistical treatment like sorting and tabulation. The term primary data may sometimes be used to refer as first-hand information. The source of primary data are primary units such as basic experimental units, individuals, households. Following methods are used to collect data from primary unit.

SOURCES OF PRIMARY DATA:

PERSONAL INVESTIGATION: The researcher conducts the experiment or survey himself/herself and collected data from it. The collected data is generally accurate and reliable.

THROUGH INVESTIGATORS: The trained investigators are employed to collect the required data.

THROUGH QUESTIONNAIRE: The required information is obtained by sending a questionnaire to the selected individuals who fill in the questionnaire and return it to the investigator.

THROUGH LOCAL SOURCES: The local representatives or agents are asked to send requisite information who provide the information based upon their own experience.

THROUGH TELEPHONE: The information may be obtained by contacting the individuals on telephone.

THROUGH INTERNET: The people may be contacted through internet and the individuals may be asked to provide the information.

* SECONDARY DATA: Data which has already been collected by someone. May be sorted, tabulated and has undergone a statistical treatment. It is fabricated or tailored data.

SOURCES OF SECONDARY DATA:

GOVERNMENT ORGANIZATIONS

SEMI-GOVERNMENTAL ORGANIZATIONS

TEACHIND AND RESEARCH ORGANIZATIONS

RESEARCH JOURNALS AND NEWSPAPERS

INTERNET.

* SAMPLE SURVEYS

A sample survey is a method for collecting data from or about the members of a population so that inferences about the entire population can be obtained from a subset, or sample, of the population members. Sample surveys are commonly used for data collection in many academic areas, particularly the social sciences. The data collected are then used to estimate statistics of interest and the sample sizes are determined so that those statistics are estimated with an acceptable sampling error.

 The three parts of survey sampling are:

* Sample selection.
* [Data Collection](https://www.statisticshowto.datasciencecentral.com/data-collection-methods/): collecting the data through mail, phone, or in-person.
* Estimation: using [estimators](https://www.statisticshowto.datasciencecentral.com/statistic/)from the collected data to make inferences about the population as a whole.

THREE TECHNIQUES COMMONLY USED:

* **QUESTIONNAIRES**: written questions (on paper or on a computer), which can include [open ended questions](http://www.oucom.ohiou.edu/fd/Open%20Ended%20Questions.htm), [closed ended questions](https://facultyinnovate.utexas.edu/teaching/feedback/mid-semester/closed-ended), multiple choice questions, or questions that require the respondent to rate something on a scale (i.e. 1 to 10).
* **INTERVIEWS**: oral questions (in person or via a phone or computer) — closed-ended or open-ended, multiple choice or on a scale.
* **SURVEYS**: brief interviews about a specific topic. Like questionnaires and interviews, these can also come in a variety of question formats.

* SIMPLE RANDOM SAMPLING

Simple random sampling (SRS) is a method of selection of a sample comprising of n number of sampling units out of the population having N number of sampling units such that every sampling unit has an equal chance of being chosen. The samples can be drawn in two possible ways. The sampling units are chosen without replacement in the sense that the units once chosen are not placed back in the population. The sampling units are chosen with replacement in the sense that the chosen units are placed back in the population.

**1. Simple random sampling without replacement (SRSWOR):**

SRSWOR is a method of selection of n units out of the N units one by one such that at any stage of selection, anyone of the remaining units have same chance of being selected, i.e. 1/N.

**2. Simple random sampling with replacement (SRSWR):**

SRSWR is a method of selection of n units out of the N units one by one such that at each stage of selection each unit has equal chance of being selected, i.e., 1/N.

PROCEDURE OF SELECTION OF A RANDOM SAMPLE:

The procedure of selection of a random sample follows the following steps:

1. Identify the N units in the population with the numbers 1 to N.

2. Choose any random number arbitrarily in the random number table and start reading numbers.

3. Choose the sampling unit whose serial number corresponds to the random number drawn from the table of random numbers.

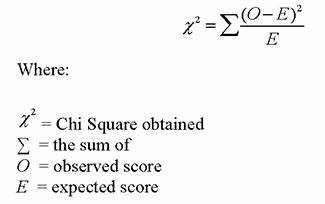
4. In case of SRSWR, all the random numbers are accepted ever if repeated more than once.

In case of SRSWOR, if any random number is repeated, then it is ignored and more numbers are draw. Such process can be implemented through programming and using the discrete uniform distribution. Any number between 1 and N can be generated from this distribution and corresponding unit can be selected into the sample by associating an index with each sampling unit. Many statistical softwares like R, SAS, etc. have inbuilt functions for drawing a sample using SRSWOR or SRSWR.

* **STATISTICAL TEST USED**
* **THE CHI SQUARE TEST**

The Chi square test is a statistical test which measures the association between two categorical variables. The null hypothesis of the Chi-Square test is that no relationship exists on the categorical variables in the population; they are independent.  
 The Chi-Square statistic is most commonly used to evaluate Tests of Independence when using a cross-tabulation (also known as a bivariate table).  Cross-tabulation presents the distributions of two categorical variables simultaneously, with the intersections of the categories of the variables appearing in the cells of the table.  The Test of Independence assesses whether an association exists between the two variables by comparing the observed pattern of responses in the cells to the pattern that would be expected if the variables were truly independent of each other.  Calculating the Chi-Square statistic and comparing it against a critical value from the Chi-Square distribution allows the researcher to assess whether the observed cell counts are significantly different from the expected cell counts.

The calculation of the Chi-Square statistic is quite straight-forward and intuitive:



As depicted in the formula, the Chi-Square statistic is based on the difference between what is actually observed in the data and what would be expected if there was truly no relationship between the variables.

**Hypothesis:**

**Null hypothesis:** Assumes that there is no association between the two variables.

**Alternative hypothesis:**Assumes that there is an association between the two variables.

**Hypothesis testing:**

Hypothesis testing for the chi-square test of independence as it is for other tests like [ANOVA](http://www.statisticssolutions.com/data-analysis-plan-repeated-measures-anova/), where a test statistic is computed and compared to a critical value.  The critical value for the chi-square statistic is determined by the level of significance (typically .05) and the degrees of freedom.  The degrees of freedom for the chi-square are calculated using the following formula: *df* = (r-1)(c-1) where r is the number of rows and c is the number of columns. If the observed chi-square test statistic is greater than the critical value, the null hypothesis can be rejected.

CHI-SQUARE TEST USING R SOFTWARE:

**Syntax of a chi-square test:** chisq.test ( )

Particularly in this test, we have to check the p-values. The p-value is the probability of the observing sample statistics as extreme as the statistic. Since the test statistic is chi-square, use the chi-square distribution to calculate, to assess the probability associated with the test statistic. Moreover, like all statistical tests, we assume this test as a null hypothesis and an alternate hypothesis. The main thing is, we reject the null hypothesis if the p-value that comes out in the result is less than a predetermined significance level, which is 0.05 usually and then we reject the null hypothesis.

H0: The two variables are independent.

H1: The two variables relate to each other.

In the case of a null hypothesis, a chi-square test is to test the two variables that are independent. In case of an alternative hypothesis, a chi-square test is to test the dependency of two variables.

* **ANOVA**

Analysis of variance (ANOVA) is a collection of [statistical models](https://en.wikipedia.org/wiki/Statistical_model) and their associated estimation procedures (such as the "variation" among and between groups) used to analyse the differences among group means in a [sample](https://en.wikipedia.org/wiki/Sample_(statistics)). ANOVA was developed by [statistician](https://en.wikipedia.org/wiki/Statistician) and [evolutionary biologist](https://en.wikipedia.org/wiki/Evolutionary_biology) [Ronald Fisher](https://en.wikipedia.org/wiki/Ronald_Fisher). The ANOVA is based on the [law of total variance](https://en.wikipedia.org/wiki/Law_of_total_variance), where the observed [variance](https://en.wikipedia.org/wiki/Variance) in a particular variable is partitioned into components attributable to different sources of variation. In its simplest form, ANOVA provides a [statistical test](https://en.wikipedia.org/wiki/Statistical_test) of whether two or more population [means](https://en.wikipedia.org/wiki/Mean) are equal, and therefore generalizes the [*t*-test](https://en.wikipedia.org/wiki/Student%27s_t-test#Independent_two-sample_t-test) beyond two means.

Researchers and students use ANOVA in many ways.  The use of ANOVA depends on the research design.  Commonly, ANOVAs are used in three ways: [one-way ANOVA](http://www.statisticssolutions.com/data-analysis-plan-one-way-anova/), two-way ANOVA[,](http://www.statisticssolutions.com/academic-solutions/resources/directory-of-statistical-analyses/factorial-anova/) and N-way ANOVA.

One-Way ANOVA:

A one-way ANOVA has just one independent variable.  For example, difference in IQ can be assessed by Country, and County can have 2, 20, or more different categories to compare.

Two-Way ANOVA:

A two-way ANOVA refers to an ANOVA using two independent variables.  Expanding the example above, a 2-way ANOVA can examine differences in IQ scores (the dependent variable) by Country (independent variable 1) and Gender (independent variable 2).  Two-way ANOVA can be used to examine the interaction between the two independent variables. Interactions indicate that differences are not uniform across all categories of the independent variables.

ANOVA USING R:

Syntax used for ANOVA test in R: aov ( )

* DIAGRAMS AND GRAPHS

One of the most effective and interesting alternative ways to present statistical data is through diagrams and graphs. There are several ways in which statistical data may be displayed pictorially, such as different types of graphs and diagrams.   
  
a. Multiple bar diagram

In a multiple bars diagram two or more sets of inter-related data are represented (multiple bar diagram facilitates comparison between more than one phenomena). The technique of making a simple bar chart is used to draw this diagram but the difference is that we use different shades, colours, or dots to distinguish between different phenomena.

b. 3D Pie diagram

A pie chart displays data, information, and statistics in an easy-to-read 'pie-slice' format with varying slice sizes telling you how much of one data element exists. The bigger the slice, the more of that particular data was gathered.

A 3D Pie - chart or 3D Pie diagram is nothing but a 3 dimensional representation of pie the regular pie chart.

c. Line plot

It is a graph that represents data along the x-axis and y-axis. It is simple and easy to understand.

**SOFTWARE USED**

* MICROSOFT EXCEL

Excel is a spreadsheet program from Microsoft, a component of its Office product group for business applications. In common with similar products, such as Lotus 1-2-3 and Corel's QuattroPro, Excel is an automated version of the paper-based spreadsheet that makes it easier to manipulate, process, and view the data. Although Lotus 1-2-3 was reportedly the first spreadsheet to introduce cell names and macros, Excel implemented a graphical user interface and the ability to point and click using a mouse

Some features offered by MS Excel:

**Conditional Formatting:** You can format cells/data according to condition and can get glimpse of entire data within seconds.

**VBA, Macros & automation:** This would make computer work for you. You can automate mundane job for you. (Ping me if you want to know how)

**Pivot Tables:** Worried about how you’ll interpret data. Let excel do it for you. This is to must know feature of excel.

**Lookup Formulas:** This formula is deadly. Combination of it with match, index can help in doing wonders.

**Power Pivot:** It helps in querying the huge data effectively. Has been introduced in recent version (2010 later).

* MICROSOFT WORD

Microsoft Word (or MS Word) is one of the most widely used word-processing programs. Word-processor programs primarily allow users to create and edit text documents. Typical use may include writing an essay or report, creating a resume, or writing notes. In addition, work can be presented in the form of inserted tables, diagrams or pictures. The main function of Microsoft Word is to write and edit documents. This will allow you to draft your text content and include special characters and fonts.

Some features offered by MS Word:

* WYSIWYG (what-you-see-is-what-you-get) display: It ensures that everything displayed on screen appears the same way when printed or moved to another format or program.
* Spell check: Word features a built-in dictionary for spell checking; miss spelled words are marked with a red squiggly underline. Sometimes, Word auto-corrects an obviously misspelled word or phrase.
* Text-level features such as bold, underline, italic and strike-through.
* Page-level features such as indentation, paragraphing and justification.
* External support: Word is compatible with many other programs the most common being the other members of the Office suite.
* R PROGRAMMING

R acts as an alternative to traditional statistical packages such as SPSS, SAS, and Stata such that it is an extensible, open-source language and computing environment for Windows, Macintosh, UNIX, and Linux platforms. Such software allows for the user to freely distribute, study, change, and improve the software under the [Free Software Foundation](https://my.fsf.org/associate/support_freedom?referrer=4052)'s [GNU General Public License](http://www.r-project.org/COPYING). It is a free implementation of the S programming language, which was originally created and distributed by Bell Labs. However, most code written in S will run successfully in the R environment. R performs a wide variety of basic to advanced statistical and graphical techniques at little to no cost to the user. These advantages over other statistical software encourage the growing use of R in cutting edge social science research.

Installation files for Windows, Mac, and Linux can be found at the website for the Comprehensive R Archive Network, <http://cran.r-project.org/>. The site also contains documentation for downloading and installing the software on different operating systems. There is no cost for downloading and using R.

R and its libraries implement a wide variety of statistical and [graphical](https://en.wikipedia.org/wiki/Graphical) techniques, including [linear](https://en.wikipedia.org/wiki/Linear) and [nonlinear](https://en.wikipedia.org/wiki/Nonlinear) modelling, classical statistical tests, [time-series analysis](https://en.wikipedia.org/wiki/Time-series_analysis), classification, clustering, and others. R is easily extensible through functions and extensions, and the R community is noted for its active contributions in terms of packages. Many of R's standard functions are written in R itself, which makes it easy for users to follow the algorithmic choices made. For computationally intensive tasks, [C](https://en.wikipedia.org/wiki/C_(programming_language)), [C++](https://en.wikipedia.org/wiki/C%2B%2B), and [FORTRAN](https://en.wikipedia.org/wiki/Fortran) code can be [linked](https://en.wikipedia.org/wiki/Linking_(computing)) and called at run time. Advanced users can write C, C++, [Java](https://en.wikipedia.org/wiki/Java_(programming_language)), [NET](https://en.wikipedia.org/wiki/.NET_Framework) or [Python](https://en.wikipedia.org/wiki/Python_(programming_language)) code to manipulate R objects directly. R is highly extensible through the use of user-submitted packages for specific functions or specific areas of study. Due to its [S](https://en.wikipedia.org/wiki/S_(programming_language)) heritage, R has stronger [object-oriented programming](https://en.wikipedia.org/wiki/Object-oriented_programming) facilities than most statistical computing languages. Extending R is also eased by its [lexical scoping](https://en.wikipedia.org/wiki/Lexical_scoping) rules

**CHAPTER 4**

**DATA ANALYSIS**

* **CHI-SQUARE TEST OF INDEPENDENCE**

1. TESTING WHETHER THE METHOD OF FARMING AND THE CROPS CULTIVATED ARE RELATED.

**HYPOTHESIS:**

H0: There is no significance relation between the method of farming and the crops cultivated.

H1: There is significant relation between the method of farming and the crops cultivated.

TABLE 3.1 :

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **1Column1** | PADDY | FRU & VEG | COFFEE | PEPPER | BANANA | COCCO | RUBBER | COCONUT | ROOT VEG | OTHER |
| Org | 54 | 54 | 82 | 75 | 45 | 13 | 16 | 19 | 11 | 32 |
| Inorg | 19 | 9 | 49 | 41 | 33 | 6 | 17 | 14 | 7 | 21 |

In the above table,

Org = number of organic farmers

Inorg= number of inorganic farmers

**OUTPUT:**

>crop\_reltn

paddy fruit&veg coffee pepper bananna cocco rubber coconut root veg

ORGANIC 54 54 82 75 45 13 16 19 11

INORGANIC 19 9 49 41 33 6 17 14 7

others

ORGANIC 32

INORGANIC 21

> chisq.test(crop\_reltn)

Pearson's Chi-squared test

data: crop\_reltn

X-squared = 22.104, df = 9, p-value = 0.008554

**CONCLUSION:**

From the above output, since the calculated p-value is less than the significance level 0.05(p-value = 0.008554 < 0.05), we reject the null hypothesis H0 at 5% significance level. Hence, we conclude that there is dependency of farming practice and crop cultivated.

2. TESTING WHETHER EXTERNAL FINANCIAL SUPPORT AND ORGANIC FARMING ARE RELATED.

2.1 Testing whether the number of bank loans and the number of organic farmers are related

**HYPOTHESIS:**

H0: There is no significant relation between number of organic farmers and the number of bank loans.

H1 There is significant relation between number of organic farmers and the number of bank loans. .

TABLE 3.2:

|  |  |  |
| --- | --- | --- |
|  | AVAILING BANK LOAN | NO BANK LOAN |
| ORGANIC | 66 | 47 |
| INORGANIC | 26 | 24 |

**OUTPUT:**

>BANK\_RLTN

YES NO

ORG 66 47

INORG 26 24

> chisq.test(BANK\_RLTN)

Pearson's Chi-squared test with Yates' continuity correction

data: BANK\_RLTN

X-squared = 0.3475, df = 1, p-value = 0.5555

**CONCLUSION:**

From the above output, since the calculated p-value is greater than the significance level 0.05 (p-value=0.5555>0.05), we accept the null hypothesis H0 at 5% level of significance. Hence, we conclude that the number of agriculture bank loans and the number of organic farmers are independent.

2.2 Testing whether the number of government financial support and the number of organic farmers are related

**HYPOTHESIS:**

H0: There is no significant relation between the number of government financial support and the number of organic farmers.

H1: There is significant relation between the number of government financial support and the number of organic farmers.

TABLE 3.3 :

|  |  |  |
| --- | --- | --- |
|  | NUMBER OF GOVT. FINANCIAL SUPPORT | NO GOVT. FINANCIAL SUPPORT |
| ORGANIC | 50 | 63 |
| INORGANIC | 17 | 33 |

**OUTPUT:**

>GFS\_RLTN

YES NO

OR 50 63

INOR 17 33

> chisq.test(GFS\_RLTN)

Pearson's Chi-squared test with Yates' continuity correction

data: GFS\_RLTN

X-squared = 1.1101, df = 1, p-value = 0.2921

**CONCLUSION:**

From the above output, since the calculated p-value is greater than the significance level 0.05 (p-value=0.2921>0.05), we accept the null hypothesis H0 at 5% level of significance. Hence, we conclude that the number of government financial support for farmers and the number of organic farmers are independent.

**INFERENCE:**

From the above two tests 2.1 and 2.2, we draw the inference that neither bank loans nor financial support for farmers by government affect the number of organic farmers. In the above two sessions the p-value obtained from the Chi-Square test is greater than the significance level 0.05. Hence we accept H0 and conclude that external financial support does not have any relation with organic farming.

3. TESTING WHETHER THE AGE AND ORGANIC FARMING ARE RELATED

**HYPOTHESIS:**

H0: The number of organic farmers and age are independent.

H1: The number of organic farmers and age are dependent.

TABLE 3.4

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| AGE GROUP | BELOW 40 | 40-49 | 50-59 | 60-69 | 70-79 |
| FARMERS | 11 | 56 | 79 | 45 | 9 |
| ORGANIC FARMERS | 5 | 36 | 42 | 23 | 7 |

OUTPUT:

> rltn

Below 40 40-49 50-59 60-69 70-79

people 11 56 79 45 9

org 5 36 42 23 7

> chisq.test(rltn)

Pearson's Chi-squared test

data: rltn

X-squared = 1.1893, df = 4, p-value = 0.8799

**CONCLUSION:**

From the above output, since the calculated p-value is greater than the significance level 0.05 (p-value=0.8799>0.05), we accept the null hypothesis H0 at 5% level of significance. Hence, we conclude that the number of organic farmers doesn’t depend upon their age.

**GRAPHICAL REPRESENTATION OF NUMBER OF FARMERS AND ORGANIC FARMERS UPON AGE.**

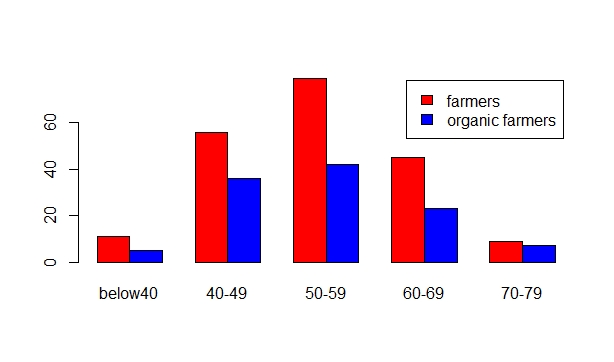


FIG 3.4.1: MULTIPLE BAR DIAGRAM SHOWING THE NUMBER OF FARMERS AND ORGANIC FARMERS IN WAYANAD

**INFERENCE**

The above graph shows that only 5.5% of the farmers are below 40 years old and out of these farmers 45% farmers follow organic farming. Between 40-49 years old there lies 28% of total farmers and 64.29% of these people are organic farmers. The highest percent of farmers lies between the age group 52-59. This age group consist of 39.5% of farmers in Wayanad. Out of these farmers 53.16% of the farmers’ population 53.16% are organic farmers.22.5% of the farmers are between the age group 60-69. 51.11% of the farmers of these age group are organic farmers. The age group with least number of farmers is 70-79 years old age group with 4.5% farmers. But out of the total population of this age group 77.8% of farmers follows organic farming showing the highest percent of organic farming as compared to other age groups.

* **ANOVA (RANDAMIZED BLOCK DESIGN)**

1. THE EFFECT OF EDUCATION ON THE NUMBER OF FARMERS

**HYPOTHESIS:**

H0: The number of farmers are same in all education levels (treatment)

H1: The number of farmers are not same in all education levels (treatment)

TABLE 3.5 - ANOVA TABLE:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SOURCE OF VARIANCE | DEGREE OF FREEDOM | SUM OF SQUARES (S.S) | MEAN SUM OF SQUARES | F |
| TREATMENT | 3 | 227.67 | 75.89 | 0.0101 |
| BLOCK | 2 | 30.50 | 15.25 | 0.2221 |
| ERROR | 6 | 46.83 | 7.81 |  |
| TOTAL | 5 | 305 | 61 |  |

**OUTPUT:**

summary(av)

Df Sum Sq Mean Sq F value Pr(>F)

tm 3 227.67 75.89 9.722 0.0101 \*

blk 2 30.50 15.25 1.954 0.2221

Residuals 6 46.83 7.81

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**CONCLUSION:**

From the above output, since the calculated p-value (F-value) is less than the significance level 0.05 (p-value=0.0101<0.05), we reject the null hypothesis H0 at 5% level of significance. Therefore education affects the number of people choosing agriculture as job.

2.THE EFFECT OF EDUCATION ON THE NUMBER OF ORGANIC FARMERS

**HYPOTHESIS:**

H0: The number of organic farmers are same in all education levels (treatment)

H1: The number of organic farmers are not same in all education levels (treatment)

TABLE 3.6 - ANOVA TABLE:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SOURCE OF VARIANCE | DEGREE OF FREEDOM | SUM OF SQUARES (S.S) | MEAN SUM OF SQUARES | F |
| TREATMENT | 3 | 224.25 | 74.75 | 0.012 |
| BLOCK | 2 | 1.17 | 0.58 | 0.932 |
| ERROR | 6 | 49.50 | 8.25 |  |
| TOTAL | 5 | 274.92 | 54.984 |  |

**OUTPUT:**

summary(av)

Df Sum Sq Mean Sq F value Pr(>F)

tm 3 224.25 74.75 9.061 0.012 \*

blk 2 1.17 0.58 0.071 0.932

Residuals 6 49.50 8.25

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**CONCLUSION:**

From the above output, since the calculated p-value (F-value) is less than the significance level 0.05 (p-value=0.012<0.05), we reject the null hypothesis H0 at 5% level of significance. Therefore education affects the number of people choosing organic farming.

* **MULTIPLE BAR DIAGRAM**
  1. EDUCATIONAL QUALIFICATION AND THE FARMING PRACTICE

TABLE 3.7

|  |  |  |
| --- | --- | --- |
|  | INORGANIC FARMER | ORGANIC FARMER |
| BELOW 10 TH | 18 | 20 |
| 10TH | 14 | 36 |
| 12TH | 12 | 45 |
| DEGREE & ABOVE | 6 | 12 |

FIGURE 3.7.1: Multiple bar diagram showing educational qualification and the farming practice chosen

**INFERENCE:**

From the above multiple bar diagram it is clear that people who are educated above 12th are not much interested in farming. But the figure also shows that in all the groups of the educational qualifications, majority of the farmers are following organic farming in Wayanad.

2. NUMBER OF PEOPLE IN DIFFERENT OCCUPATIONS IN DIFFERENT TALUKS IN WAYANAD

TABLE 3.8

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | AGRICULTURE | BUSSINESS | EMPLOYED | OTHERS |
| SBY | 60.65 | 6.56 | 16.39 | 16.39 |
| MNDY | 64.4 | 3.39 | 11.86 | 20.34 |
| VYTHIRI | 63.75 | 7.5 | 11.25 | 17.5 |

FIGURE 3.8.1: Multiple bar diagram showing the occupation of people in wayanad

in different taluks.

**INFERENCE:**

From the above multiple bar diagram we see that majority of the people in Wayanad is takes agriculture for occupation irrespective of the Taluks they are distributed. In all the taluks above 60% of the total population follows agriculture. So we can draw the conclusion that Wayanad is a land of Farmers. Within the different taluks, Mananthavady taluk is having the highest percentage of farmers with 64.4%, followed by Vythiri with 63.75% of farmers and finally Sulthan Bathery with 60.65%.

* **PIE-DIAGRAM**

TABLE 3.9: CERTIFIED OGRANIC FARMERS

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **YEAR** | **2000-2004** | **2005-2009** | **2010-2014** | **2015-2019** | **NOT** |
| NO. | 2 | 5 | 6 | 8 | 94 |

FIGURE 3.9.1: 3-D pie-diagram showing certified organic farmers with respect to the year they were certified.

**INFERENCE:**

The Pie-Chart compares the numerical proportion of certified organic farmers during different span of time and non-certified organic farmers. It can be clearly seen that majority of the organic farmers are not certified. This data also reveals that the proportion of farmers who get certified increase over years. Among all the organic farmers only 18.26 % people are only certified. And within the certified 38.09 % of people got certified during 2015-2019. Finally from the above data we can draw the conclusion that majority of the organic farmers remain non-certified in Wayanad.

* **LINE PLOT**

1. THE LINE PLOT SHOWING THE NUMBER OF ORGANIC AND INORGANIC FARMERS WITH RESPECT TO THE LAND THEY OWN.

FIGURE 3.10:The figure shows the relation between the land and the farming practice.

**INFERENCE:**

The above figure shows that majority of the (97.3%) of the farmers who own land more than 5 acres follow organic method for farming. And moreover in all kind of farmers majority of the farmers follow organic farming.

**CHAPTER 4**

**CONCLUSIONS AND FINDINGS**

As per the statistical analysis on the primary data collected from peoples from different parts of Wayanad, the major findings of the project are:

* The method of farming is dependent on the crop which they cultivated. During the survey, many farmers do mixed farming, that is for some crops they prefer organic and for some other crops they add fertilizers.
* The external financial support for a farmer comes in mainly through two ways; one is agricultural bank loans and the other is the financial assistance provided by the government. Both of these doesn’t make any effect on their decision on which farming practice. All farmers irrespective of whether organic or inorganic avail external financial support.
* The data collected reveals that majority of the farmers are above 40 years old. But the age does not have any effect on the farming practice they choose. The data also revealed the upcoming decrease in the number of farmers.
* The education attained by the people have an effect on their choice of occupation and also the education has an effect on the farming practice the farmers choose.
* Majority of the organic farmers are not certified. The data reveals the slow growth of certified organic farmers in Wayanad.
* Majority of the people who own land more than 5 acres prefer to do organic farming. The people who live in rented house doesn’t follow any kind of farming.
* Finally, the number of organic farmers are high in Wayanad.