Family Health Survey In Villages Under RHTC Field Practice Area, Lavale (Pune).

A Project report submitted in partial fulfilment of Requirement for Degree of M.Sc. (Statistics) With specialization in Industrial Statistics

By

Mr. Patil Lalit Manohar

Ms. Patil Pooja Santosh

Mr. Mahajan Jitendra Bhikan

Project Guide

Mr. Manoj C. Patil

Department of Statistics

KBCNMU, Jalgaon

DEPARTMENT OF STATISTICS

SCHOOL OF MATHEMATICAL SCIENCES

KAVAYITRI BAHINABAI CHAUDHARI

NORTH MAHARASHTRA UNIVERSITY

JALGAON – 425001

2019-2020

CERTIFICATE

This is to certify that Mr. Lalit Manohar Patil, Ms. Pooja Santosh Patil and Mr. Jitendra Bhikan Mahajan, students of M.Sc.(Statistics) with specialization in Industrial Statistics, at Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon have successfully completed their project work entitled 'Family Health Survey' Based on the data obtained from Bharati Vidyapeeth Medical College, Pune as a part of M.Sc. (Statistics) program under my guidance and supervision during the academic year 2019-20.

Place: Jalgaon
Date: / /2020

Manoj C. Patil
(Project Guide)
Department of Statistics,
K.B.C. North Maharashtra University, Jalgaon

ACKNOWLEDGEMENT

On this note, firstly we want to convey our regards to **Prof. R. L. Shinde**, Head of Department of Statistics, Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon for giving us the permission for this project. We take this opportunity in expressing our sincere gratitude to our project guide **Mr. Manoj C. Patil** (Lecturer, Department of Statistics, Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon) for his valuable guidance. We also want to acknowledge all other teachers of department of statistics, KBCNMU, Jalgaon for their precious knowledge imported to us. We also thank **Dr. Jayashree Gothankar** and **Mr. K. H. Patil** for providing the data, their guidance and support.

Our friends also deserved a delightful thank for their continuous encouragement to us throughout the project till completion. Finally, we thank our parents for the moral supports and blessings.

Place: Jalgaon

Date: / /2020

Mr. Patil Lalit Manohar

Ms. Patil Pooja Santosh

Mr. Mahajan Jitendra Bhikan

INDEX

CHAPTER 1: INTRODUCTION

1.1 Introduction:

Health information gathering is an important part of any health system, but is often weak in low-income countries, plagued by poor quality data that are inadequate for informing health policy.

Population-based surveys are an invaluable source of health information. A key aim of these surveys is to provide high-quality data for policy development and programme planning, monitoring and evaluation. Population-based surveys have been used extensively to gather information on fertility, mortality, family planning, maternal and child health, and some other aspects of health, nutrition and health care in India.

WHO(World Health Organisation) have previously reported that the health information system in India has not kept up with the epidemiological transition. WHO assessed national health surveys conducted in India since 1992 that were designed to provide information on health indicators at subnational levels. WHO describe and compare the health information covered by these surveys over time, the availability of resulting data in the public domain and the use of these survey data in publications. Based on findings, WHO highlight the issues that need consideration to improve the usefulness of these surveys. WHO believe they should be able to provide more effective, useable and timely data on the health status of the population, given the evolving disease burden in India.

National Family Health Survey (NFHS) conducts the family health surveys in India.

The National Family Health Survey (NFHS) is a large-scale, multi-round survey conducted in a representative sample of households throughout India. Three rounds of the survey have been conducted since the first survey in 1992-93. The survey provides state and national information for India on fertility, infant and child mortality, the practice of family planning, maternal and child health, reproductive health, nutrition, anaemia, utilization and quality of health and family planning services. Each successive round of the NFHS has had two specific goals:

a) to provide essential data on health and family welfare needed by the Ministry of Health and Family Welfare and other agencies for policy and programme purposes, and

b) to provide information on important emerging health and family welfare issues. The Ministry of Health and Family Welfare (MOHFW), Government of India, designated the International Institute for Population Sciences (IIPS) Mumbai, as the nodal agency, responsible for providing coordination and technical guidance for the survey. IIPS collaborated with a number of Field Organizations (FO) for survey implementation. Each FO was responsible for conducting survey activities in one or more states covered by the NFHS.

Technical assistance for the NFHS was provided mainly by ORC Macro (USA) and other organizations on specific issues. The funding for different rounds of NFHS has been provided by USAID, DFID, the Bill and Melinda Gates Foundation, UNICEF, UNFPA, and MOHFW, GOI.

1.2 Motivation:

Being student of M.Sc. (Statistics) with specialization in industrial statistics we always heard about the applications of various statistical techniques in the social and health surveys which motivated us to perform this project.

1.3 Organization Profile:

Name of the Organization:- Bharati Vidyapeeth (Deemed to be University) Medical College, Pune.

Address:- Pune-Satara Road, Pune - 411 043

Email:- bvumedicalpune@gmail.com

Website:- mcpune.bharatividyapeeth.edu

Established Year: 1989

Founder:- Dr. Patangrao Kadam

Phone:- (0257) 2258011 2258022

Fax:- 020 - 24372175

VISION: To transform Bharati Vidyapeeth (Deemed to be University) Medical College, Pune into a **"CENTER OF EXCELLENCE"** in the field of Medical Education & Research and to cater to national as well as global health care needs.

MISSION: Social Transformation through Dynamic Education

GOAL: To train the students to be **COMPETENT, COMPASSIONATE & CONFIDENT** physicians to meet the health care delivery needs of the society.

The University and the College have been awarded 'A+' Accreditation (Third Cycle) by 'NAAC' in 2017, Category-I Deemed to be University Grade by UGC and 'A' Grade by Ministry of HRD, Government of India.

There are 28 departments in the college offering 03 Undergraduate, 27 Post Graduate degree and 08 Diploma, 05 super specialty (DM, MCh) programmes. In addition 32 Post-Doctoral Fellowships, PhD in 09 subjects, numerous short term and fellowship certificate courses and some need based programmes are also being conducted.

Overall the college provides for an excellent student friendly teaching learning environment. The attached hospital not only functions for clinical training of undergraduate and postgraduate students but also caters to the medical needs of a large segment of population of East Pune as well as the surrounding rural areas including poor and needy patients.

There is a constant thrust on faculty development, competency based curriculum, research and innovative teaching learning environment. At each step the college strives to fulfil the vision of the organization 'to be a world class university for Social Transformation through Dynamic Education'.

1.4 Objective of the Project:

- To gather information about birth rates, death rates, socio economic status health for villages under RHTC field practice area Lavale.
- To obtain information about the variables education level, socio-economic study etc. to which type of delivery (normal or cesarean) is related.
- To obtain information about child vaccination.
- To gather the information about impact of religion on various factors such as socio economic status, education etc.

1.5 Scope of the Project:

Scope of this project is to study how the Socio Economic Status (SES), education level of head of family, education level of mother, religion etc. factors impacts on family health for villages under RHTC field practice area Lavale.

CHAPTER 2: METHODOLOGY

The dataset we have used in this project is collected by using following methodology by Bharati Vidyapeeth Medical College, Pune.

2.1 Study Design:

A cross-sectional study.

2.2 Study area:

villages under RHTC field practice area Lavale, a rural area 26Km away from Pune.

2.3 Tools of data collection:

- Questionnaire- Predesigned pretested
- Weighing machine
- Glucometer for blood glucose estimation
- Sphygmomanometer for blood pressure measurement.

2.4 Methodology:

- Approval of the institutional ethical committee had taken before the start of the study.
- Informed consent of the head of the family is obtained before data collection
- Data collection by interview method by conducting house to house visit. Questionnaire was filled by trained Medical students and interns by interview technique. All the households visited marked by marker pen. If the household was locked on first visit it was visited again.
- Information on following variables were collected-Socio –demographic- age , sex, religion, education, Occupation, socio- economic Status ,vital events in

last one year, immunization status, Diabetes, Hypertension etc. .Blood pressure was taken of persons who are 40 years and above in the family.Random blood sugar was tested with the help of glucometer in symptomatic persons and in those with family history of diabetes.

2.5 Questionnaire:

Bharati Vidyapeeth University Medical College, Pune.
Department of Community Medicine
Family Care study

Sign. of verifier:

Date of visit:						
Name of Student:		Batch :		Roll No :		
Name of Informant :						
Mobile No of Informant/Head of Family(HOF)						
Address:						
A. General Information :						
1	Are you a permanent resident in this area?	1) Yes	2) No			
2	If not, How many years you are staying here?					
3	House	Own/ Rented				
4	Type of family?	Joint/Nuclear				
5	Colour of ration card	White	Yellow	Orange	No	
6	Religion	Hindu	Muslim	Christian	Buddhist	other
7	Do you belong to BPL category?	Yes /No				

Please tick (√) the correct option.

A. Family Composition								
Sr No	Name	Relation with HOF	Age	Sex	Edu.	Occ.	Marital status	OTHER (SPECIFY)
1								
2								
3								
4								
5								
6								
7								

8								
9								
10								
11								
12								
13								
14								
	Education#	illiterate/ Read & write/Prime(1-4)/middle(5-7)/secondary(8-10)/Higher Sec(11-12) / Graduate/PG/ Diploma after 10^{th}						
	Occupation §	Service/Business/Housewife/retired/unemployed/student/labourer-skilled/unskilled						
	Marital status ¤	married / unmarried/widow& widower/Divorced/Separated						
	*Other	Pregnant / lactating / malnutrition / under5 /communicable diseases / Non communicable diseases /Addictions (alcohol/tobacco/other)Specify (If any, refer to pg no.:----)						

B. Vital Events since last one year:				**Yes**	**No**		
		Date (A)	Place (B)	Registration (C)	Sex (D)	Age (E)	Cause of Death (F)
A	Birth			1) Yes 2) No 3) Don't know		--	----
B	Marriage			1) Yes 2) No 3) Don't know	--	H- W-	----
C	Death			1) Yes 2) No 3) Don't know			

C. Housing Environment:

a) Type of House : Kaccha / Pucca /semi pucca
b) No. of Rooms : 1/2/3/>3
c) Total Sq. Ft. Area :
d) Ventilation : Adequate / Inadequate
e) Light : Adequate / Inadequate
f) Disposal of Wastes : Solid: No Disposal/Collection:

Corporation/Dumped

g) Latrine : Individual/Public
h) Open Space : Present / Absent
i) Pets :Y/N.
If Y: Type: Cattle/Goat/Chicken/Dog/Cat/Others, specify :
j) Water Supply : Common Tap/Individual tap
k) Type of fuel used for cooking :LPG/Biomass/Mixed/others(specify)

D. Economic Data:

a) No. of Earning Members :
b) Family Income (per month) : Rs.
c) Loan if Any :
d) Per Capita Family Income=Total Family Income/ Total no. of family members=Rs

Modified Prasad'sclassification (Nov.2014)

Income	Class
>5775	I
5774-2887	II
2886-1733	III
1732-866	IV
<866	V

E. Social History:

a) Relationship amongst Family Members /Relatives /Neighbours – Strained/Normal
b) Working Mother (under five) : Present / Not Present
c) Care of Under Fives when mother is out for work (If mother is working): Grandparents/elder
sibling/relative/neighbour/other(specify).....
d) School Dropouts (with reason) :(applicable for <15 years old)
e) Awareness about diseases: Write as Y/N/Some

	Tuberculosis	Malaria	Dengue	Chikungunya	AIDS
Modes of transmission					
Symptoms					
Cure present absent					
Is free treatment available					
Prevention					

f) Utilization of Health Services: Allopathic-public/private (specify) --------

F. Reproductive And Child Health Care:

i) Care during pregnancy/delivery

		Current pregnancy	delivery in last one year
1	Is there any pregnant woman in house?	Yes/ No	Yes/ No
2	Registration of pregnancy? If yes, which month	Yes/ No	Yes/ No
3	Expected Due Date (E.D.D.)		---
4	Place of delivery to be done? home/Primary health centre/govt hospital/private hospital		
5	Number of T.T. injections Taken		
6	No. of Antenatal visits done		

ii) Family Welfare:

a) No. of Eligible Couples in the Family :

Eligible Couple	Married since	Use and source of contraceptive before 1st pregnancy	Interval between marriage &1st pregnancy	Contraceptive used &period of spacing between 1st & 2nd child (birth dates)	2+ child Yes/ No	Tubectomy/ vasectomy done/not done	If Not Using Family Planning Methods Reasons
First							
Second							

Third							

iii) Immunization Form (For a child of age 5 years or less)

Name of Child : _____Birth date:_____

Sr. No.	Age	Vaccine	Yes/No	Immunization Card (confirmation of Immunization)	Govt / Pvt.	Reason for not taking Vaccine
1	After Birth	BCG, Polio '0'				
2	1.5 month/ 6 wks	DTP + OPV1				
3		Hep B 1				
4		HIB 1				
5	2.5 month / 10 wks	DTP + OPV2				
6		Hep B 2				
7		HIB 2				
8	3.5 month/ 14 wks	DTP + OPV3				
9		Hep B 3				
10		HIB 3				
11	9 month	Measles I				
12	15 month	MMR/Measle II				
13	18 month / 1.5yrs	DTP + OPV 1st booster				
14		HIB booster				
15	2 yrs	Typhoid				
16	5 yrs	DTP + OPV 2nd booster				
17	Other					

(Note for students: Ask for extra immunization form if there are more than one under five year old child.)

iv) Information of recent child birth (past one year)		
Name of Child _____Name of mother:_____		
01	Birth Date	/ / 20
02	Place of Delivery	1) Govt. Hospital 2) Private Hospital 3) Home 4) On the way to hospital 5) Other_____
03	Type of Delivery	1) Normal 2) Caesarean Section 3) Forceps 4) Vaccum
04	Sex of baby	1) Male 2) Female
05	Birth weight	1) _____Kg / _____P 2) Don't Know
06	Delivery done by whom?	1) Doctor 2) Nurse 3) Trained Dai 4) untrained Dai 5) Other, specify
07	Birth Order of the child	
08	Birth Registration done?	1)Yes 2)No
	If yes, Where?	1) Grampanchayat 2) PMC
09	Whether baby cried after birth?	1) Yes 2) No

10	Colostrum given?	1) Yes 2) No
11	When breast feeding started?	1) ½ hrs 2) 4 hrs 3) 4 to 24 hrs 4) 24 to 72 hrs 5) After 72 hrs 6) Breast feeding not given(Reason)
12	Any other feed was given other than breast milk?	1) Yes 2) No
13	If yes, What?	_____
14	Exclusive Breast feeding done for how many months?	1)_____month2) N.A.
15	In which month weaning started?	1) _____ month. 2) N.A.
16	Breast feeding still continued?	1) Yes 2) No 3) N.A.
17	At what age breast feeding was stopped?	1)_____ month 2)N.A.
18	Whether child attends Anganwadi?	1) Yes 2) No
19	Do you accept facilities provided by Anganwadi for your child?	1) Yes 2) No
20	If no, why not?	

Smoking history:

Name	No. bidi/cigarette smoked per day (A)	Years of smoking (B)	Smoking index(A*B)

iv) Nutritional Data:

a) Type of Diet : Veg./ Mixed
b) Pattern of Food Purchase : Daily/ Weekly/Monthly
c) Storage of Food : Raw ----------
: Cooked----------
d) Food during Pregnancy (additional foods, foods specially consumed, foods specially avoided).

V) Morbidity Status:

Current health problem in the family – Yes / No If yes:

	Name of the patient	Diagnosis	Details of treatment
1			Place of treatment: Regular/Irregular treatment : Cost per month:
2			Place of treatment: Regular/Irregular treatment : Cost per month:
3			Place of treatment: Regular/Irregular treatment : Cost per month:
4			Place of treatment:

			Regular/Irregular treatment : Cost per month:
5			Place of treatment: Regular/Irregular treatment : Cost per month:

v) Problems in the family perceived by you.

(write in brief about medical, social, economical, nutritional and environmental problems)Describe / elaborate the problems if any

---

---

vi) Needs as perceived by the family

(Nutritional, health care, environmental sanitation, socio economic, any other).

---

vii) Advice given to the family: (write in points)

Date: / /20

Name and signature of the student
Signature of PG student

.

2.5 Some Definitions :

- **Family Health:** Family health is a state in which the family is a resource for the day-to-day living and health of its members. A family provides its individual members with key resources for healthful living, including food, clothing, shelter, a sense of self-worth, and access to medical care.
- **Socio- Economic Status (SES):** Socioeconomic status (SES) is an economic and sociological combined total measure of a person's work experience and of an individual's or family's economic and social position in relation to others. When analyzing a family's SES, the household income, earners' education, and occupation are examined, as well as combined income, whereas for an individual's SES only their own attributes are assessed. However, SES is more commonly used to depict an economic difference in society as a whole.
- **Ventilation:** Ventilation is the intentional introduction of outdoor air into a space. Ventilation is mainly used to control indoor air quality by diluting and displacing indoor pollutants; it can also be used to control indoor temperature, humidity, and air motion to benefit thermal comfort, satisfaction with other aspects of indoor environment, or other objectives.
- **Crude Birth Rate:** Crude birth rate is the number of resident live births for a specified geographic area (nation, state, county, etc.) during a specified period (usually a calendar year) divided by the total population (usually mid-year) for that area and multiplied by 1,000.
- **Crude Death Rate**: Crude death rate is the total number of deaths to residents in a specified geographic area (country, state, county, etc.) divided by the total population for the same geographic area (for a specified time period, usually a calendar year) and multiplied by 100,000.
- **Measels:** Measles is a very contagious disease caused by a virus. It spreads through the air when an infected person coughs or sneezes. Measles starts with a cough, runny nose, red eyes, and fever. Then a rash of tiny, red spots breaks out. It starts at the head and spreads to the rest of the body. Measles can be prevented with MMR vaccine.

- **Booster:** An action or substance that makes something stronger or more effective.
- **T. T. Injections:** Tetanus vaccine, also known as tetanus toxoid (TT), is an inactive vaccine used to prevent tetanus. During childhood, five doses are recommended, with a sixth given during adolescence. Additional doses every ten years are recommended. After three doses, almost everyone is initially immune.

CHAPTER 3:STATISTICAL TESTS

3.1 Chi-square test for Association:

The chi-square test for **independence,** also called **Pearson's chi-square test** or the chi-square test of **association**, is used to determine if there is any association between two variables. It is really a hypothesis test of independence.

H_0: The two variables are not associated, i. e. , **independent**.

Vs

H_1: The two variables are associated, i. e. , **dependent**.

Assumption 1: Your two variables should be measured on ordinal or nominal scale. (i.e., categorical data)

Assumption 2: Your two variables should consist of **two or more categorical, Independent groups.**

Interpretation: To determine whether the variables are independent, compare the p-value to the significance level. Usually, a significance level (denoted as α or alpha) of 0.05 works well. A significance level of 0.05 indicates a 5% risk of concluding that an association between the variables exists when there is no actual association.

P-value ≤ α: The variables have a statistically significant association (Reject H_0)

If the p-value is less than or equal to the significance level, you reject the null hypothesis and conclude that there is a statistically significant association between the variables.

P-value > α: Cannot conclude that the variables are associated (Fail to reject H_0)

If the p-value is larger than the significance level, you fail to reject the null hypothesis because there is not enough evidence to conclude that the variables are associated.

3.2 Z-test for the equality of two proportions:

Objective:

To investigate the assumption that the proportions p_1 and p_2 of elements from two populations are equal, based on two samples, one from each population.

Limitations:

The test is approximate and assumes that the number of observations in the two samples is sufficiently large (i.e. n_1, $n_2 \geq 30$) to justify the normal approximation to the binomial.

Method:

It is assumed that the populations have proportions p_1 and p_2 with the same characteristic. Random samples of size n_1 and n_2 are taken and respective proportions p_1 and p_2 calculated. The test statistic is

$$Z = \sqrt{\frac{(p_1 - p_2)}{p(1-p)(\frac{1}{n_1} + \frac{1}{n_2})}}$$

Where,

$$p = \frac{p_1 n_1 + p_2 n_2}{n_1 + n_2}$$

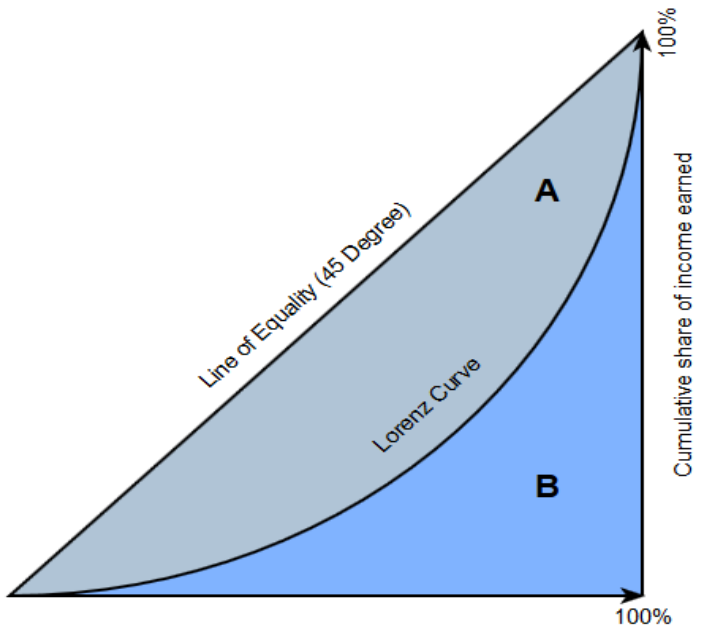
Under the null hypothesis that $p_1 = p_2$, Z is approximately distributed as a standard normal deviate and the resulting test may be either one- or two-tailed.

3.3 Lorenz Curve:

In economics, the **Lorenz curve** is a graphical representation of the distribution of income or of wealth. It was developed by Max O. Lorenz in 1905 for representing inequality of the wealth distribution.

The curve is a graph showing the proportion of overall income or wealth assumed by the bottom x % of the people, although this is not rigorously true for a finite population . It is often used to represent income distribution, where it shows for the bottom x % of households, what percentage (y %) of the total income they have. The percentage of households is plotted on the x-axis, the percentage of income on the y-axis. It can also be used to show distribution of assets. In such use, many economists consider it to be a measure of social inequality.

Points on the Lorenz curve represent statements such as, "the bottom 20% of all households have 10% of the total income."



- **Defination:** The Lorenz curve is a probability plot (a P–P plot) comparing the distribution of a parameter in a population against a hypothetical uniform distribution of that parameter. It can usually be represented by a function $L(F)$, where F, the cumulative portion of the

population, is represented by the horizontal axis, and *L*, the cumulative portion of the total wealth or income, is represented by the vertical axis.

- **Properties:**
  - A Lorenz curve always starts at (0,0) and ends at (1,1).
  - The Lorenz curve is not defined if the mean of the probability distribution is zero or infinite.
  - The Lorenz curve for a probability distribution is a continuous function. However, Lorenz curves representing discontinuous functions can be constructed as the limit of Lorenz curves of probability distributions, the line of perfect inequality being an example.
  - The information in a Lorenz curve may be summarized by the Gini coefficient and the Lorenz asymmetry coefficient.
  - The Lorenz curve cannot rise above the line of perfect equality.
  - If the variable being measured cannot take negative values, the Lorenz curve:
    - cannot sink below the line of perfect inequality,
    - is increasing.

3.4 Gini Coefficient

The Gini index or Gini coefficient is a statistical measure of distribution developed by the Italian statistician Corrado Gini in 1912. It is often used as a gauge of economic inequality, measuring income distribution or, less commonly, wealth distribution among a population. The coefficient ranges from 0 (or 0%) to 1 (or 100%), with 0 representing perfect equality and 1 representing perfect inequality. Values over 1 are theoretically possible due to negative income or wealth.

- **Key Takeways:**
  - The Gini index is a simple measure of the distribution of income across income percentiles in a population.

- A higher Gini index indicates greater inequality, with high income individuals receiving much larger percentages of the total income of the population.
- Global inequality as measured by the Gini index increased over the 19th and 20th centuries, but has declined in more recent years.
- Because of data and other limitations, the Gini index may overstate income inequality and can obscure important information about income distribution.

- **Understanding the Gini Coefficient**

  - A country in which every resident has the same income would have an income Gini coefficient of 0. A country in which one resident earned all the income, while everyone else earned nothing, would have an income Gini coefficient of 1.
  - Gini index < 0.2 represents perfect income equality, 0.2–0.3 relative equality, 0.3–0.4 adequate equality, 0.4–0.5 big income gap, and above 0.5 represents severe income gap.

- **Graphical Representation of the Gini Index**

  The Gini index is often represented graphically through the Lorenz curve, which shows income (or wealth) distribution by plotting the population percentile by income on the horizontal axis and cumulative income on the vertical axis. The Gini coefficient is equal to the area below the line of perfect equality (0.5 by definition) minus the area below the Lorenz curve, divided by the area below the line of perfect equality. In other words, it is double the area between the Lorenz curve and the line of perfect equality.

3.5 Binary Logistic Regression model

Binary logistic regression is typically used when the dependent variable is dichotomous and the independent variables are either continuous or categorical variables. One key assumption in regular binary logistic regression is that

observations are independent of each other. Violations of the assumption of independence of observations may results in incorrect statistical inferences due to biased standard errors.

The binary logistic regression model is defined as:

$$P(xi) = \frac{e^{\beta_0+\beta_1 X_{i1} +\cdots+\beta_p X_{ip}}}{1+e^{\beta_0+\beta_1 X_{i1} +\cdots+\beta_p X_{ip}}}$$

$$= \frac{e^{X\beta}}{1+e^{X\beta}} = \frac{1}{1+e^{-X\beta}}$$

Where, $\beta 0$ = the constant of the equation and βi= the coefficient of the ith predictor.

The relationship between the predictor and response variables is not a linear function in logistic regression; instead, the logarithmic transformation of equation yields the linear relationship between the predictor and response variables.

The logit transformation of $P(xi)$ given as follows:

$$logit[P(X_i)] = \log\left(\frac{P(x_i)}{1-P(x_i)}\right)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} \qquad i = 1,2,\ldots n$$

The coefficient can be interpreted as the change in the log-odds associated with a one unit change in the corresponding independent variable or the odd increases multiplicatively by eβ for every one unit change increase in x.

Logistic regression is popular in part because it over – come many of the restrictive assumption of ordinals least square (OLS) regression. Like it doesn't assume linearity, the dependent variable need not normally distributed and so on.

Logistic regressions work with odds. The odds are simply the ratio of the probabilities for the two possible outcomes. If p is the probability that the event will occur, then $1-p$ is the probability that the event will not occur: $odd = \frac{p}{1-p}$

In 2 × 2 tables, within row 1 the odds of success are $d_1 = \frac{p_1}{1-p_1}$, and within row 2 the odds of success equal $odd_2 = \frac{p_2}{1-p_2}$,

The ratio of the odds from the two rows, which we call odds ratio, will be

$$\frac{\boldsymbol{odd_1}}{\boldsymbol{odd_2}} = \frac{\frac{\boldsymbol{p_1}}{\boldsymbol{1-p_1}}}{\frac{\boldsymbol{p_2}}{\boldsymbol{1-p_2}}}$$

3.6 Crude Birth Rate (CBR) and Crude Death rate (CDR).

- CDR=$\frac{\text{Total Resident Live Births}}{\text{Total Population}} \times \mathbf{100}$
- CDR=$\frac{\text{Total Resident Live Deaths}}{\text{Total Population}} \times \mathbf{100}$

3.7 One Sample Proportion Test.

Hypothesis Tests:

- Null hypothesis H_0: p= p_0
- Alternative Hypothesis H_1: p≠ p_0

This hypothesis considers whether the population proportion is equivalent to some pre-specified value, p_0. This value might be of historical interest or a result obtained in another study that we are trying to corroborate with our study data. A rule of thumb used to perform this test is that both np_0 and $n(1-p_0)$ are greater than five.

To perform this test, we:

1. Estimate the population proportion by the sample proportion, $\hat{p}$.

   We estimate the proportion, p, as:

$$\hat{p} = \frac{X}{n}$$

Where, X is the number in the sample who have the trait or outcome of interest, and n is the size of the sample.

2. Calculate the following test statistic, which under the null hypothesis, follows approximately (dependent on the rule of thumb stated above) a Standard Normal Distribution:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Decision Rule:

Reject if $Z > Z_{\alpha/2}$, where $Z_{\alpha/2}$ is the 1-α/2 percentile of the standard normal distribution.

Chapter 4:Data Analysis and Interpretation

4.1 Chi-Square test for association

4.1.1 Chi-Square test for association between 'Education level of HOF (Head Of Family)' and 'Latrine facility'

Expected counts are printed below observed counts:

Education of HOF	Latrine		
	Individual	Public	Grand Total
Diploma	3	1	4
	3.3728395	0.62716	
Graduate	64	3	67
	56.495062	10.50494	
Higher Secondary	94	9	103
	86.850617	16.14938	
Illiterate	111	23	134
	112.99012	21.00988	
Middle	74	12	86
	72.516049	13.48395	
NA	4	0	4
	3.3728395	0.62716	
Post graduate	9	0	9
	7.5888889	1.411111	
Primary	101	23	124
	104.55802	19.44198	
R&W	2	0	2
	1.6864198	0.31358	
Secondary	221	56	277
	233.56914	43.43086	
Total	683	127	810

Hypothesis	χ^2 (Cal.)	p-value
H_0: Education level of HOF and Latrine availability are independent. VS H_1: Education level of HOF and Latrine availability are dependent.	15.9084	0.2541

Conclusion: Here, since p-value(0.2541) > level of significance(0.05).We fail to reject H_0. i.e. Education level of HOF and Latrine facility are independent.

4.1.2 Chi-Square test for association between 'Education level of HOF' and 'Waste disposal method':

Expected counts are printed below observed counts:

Education of HOF	Waste disposal method					
	Burn	Corporation	Dumped	No disposal	NR	Grand Total
Diploma	0	0	4	0	0	4
	0	0.310727	3.52651	0.138101	0	
Graduate	1	12	53	1	0	67
	0.08	5.204686	59.0691	2.313194	0	
Higher Secondary	0	6	89	5	4	104
	0.13	8.078915	91.6893	3.590629	1	
Illiterate	0	4	122	8	0	134
Middle	0	10	74	2	0	86
	0.11	6.680641	75.82	2.969174	0	
NA	0	1	3	0	0	4
Post graduate	0	0	9	0	0	9
	0.01	0.699137	7.93465	0.310727	0	
Primary	0	13	104	7	0	124
	0.15	9.632552	109.322	4.281134	1	
R&W	0	1	1	0	0	2
	0	0.155364	1.76326	0.069051	0	

Secondary	0	16	256	5	0	277
	Secondary	0.34155364	21.5179	244.2109	10	1.366215
Grand Total	1	63	715	28	4	811

Hypothesis	χ^2 (Cal.)	p-value
H_0: Education level of HOF and Waste disposal method are independent. VS H_1: Education level of HOF and Waste disposal method are dependent.	22.1570	0.0754

Conclusion: Here, since p-value(0.0754) > level of significance(0.05).We fail to reject H_0. i.e. Education level of HOF and Waste disposal method are independent.

4.1.3 Chi-Square test for association between 'Pattern of food purchase' and 'SES'

Expected counts are printed below observed counts:

Pattern of food purchase	Class I	Class II	Class III	Class IV	Class V	Total
Daily	37	42	45	24	13	161
	27.4	54	42	24.2	13.4	
Monthly	72	137	101	75	38	423
	71.9888	141.8775	110.3478	63.5813	35.2062	
Weekly	28	91	64	22	16	221
	37.611	74.1240	57.6521	33.2186	18.3937	
Total	137	307	232	119	70	805

Hypothesis	χ^2 (Cal.)	p-value
H_0: Pattern of food purchase and SES are independent. VS H_1: Pattern of food purchase and SES are dependent.	20.587	0.008

Conclusion: Here, since p-value(0.008) < level of significance(0.05).We reject H_0. i.e. Pattern of food purchase and SES are dependent.

4.1.4 Chi-Square test for association between 'Type of house' and 'Ventilation'

Expected counts are printed below observed counts:

Type of House	Adequate	Inadequate	Total
Kaccha	19	16	35
	25.4192	0.9495	
Pucca	400	90	490
	355.86	134.1307	
Semi pucca	170	116	286
	207.7115	78.2885	
Total	589	222	811

Hypothesis	χ^2 (Cal.)	p-value
H_0: Type of house and ventilation are independent. VS H_1: Type of house and ventilation are dependent.	5.34	0.00

Conclusion: Here, since p-value(0.00) < level of significance(0.05).We reject H_0. i.e. Type of house and ventilation are dependent.

4.1.5 Chi-Square test for association between 'Education leve; of HOF' and 'Presence of working mother in family':

Expected counts are printed below observed counts:

Education level of HOF	Present	Absent	Grand Total
Diploma	0	4	4
	0.356877	3.643123	
Graduate	0	67	67
	5.977695	61.0223	
Higher Secondary	11	93	104
	9.27881	94.72119	
Illiterate	9	125	134
	11.95539	122.0446	
Middle	8	78	86
	7.672862	78.32714	
NA	1	3	4
	0.356877	3.643123	
Post graduate	0	9	9
	0.802974	8.197026	
Primary	17	107	124
	11.0632	112.9368	
R&W	0	2	2
	0.178439	1.821561	
Secondary	26	247	273
	24.35688	248.6431	
Grand Total	**72**	**735**	**807**

Hypothesis	χ^2 (Cal.)	p-value
H_0: Education level of HOF and presence of working mother in family are independent. VS H_1: Education level of HOF and presence of working mother in family are dependent.	21.23	0.569902

Conclusion: Here, since p-value(0.569902) > level of significance(0.05).We fail to reject H_0. i.e. Education level of HOF and presence of working mother in family are independent.

4.2 Proportion Test:

4.2.1 Proportion test for testing the proportion of hindu families above poverty line and the prop of families of other religions above poverty line are equal.

Hypothesis	Z-value (Cal.)	p-value
H_0:Proportion of hindu families above poverty line and the prop of other families above poverty line are equal. VS H_1: Proportion of hindu families above poverty line and the prop of other families above poverty line are not equal.	1.72	0.086

Conclusion: Here, since p-value(0.086) > level of significance(0.05).We fail to reject H_0. i.e. Proportion of hindu families above poverty line and the proportion of other religion families above poverty line are equal.

4.2.2 Proportion test for testing the proportion of nuclear families in hindu religion and the proportion of nuclear families in other religion are equal.

Religion	X	n
Hindu	481	758
Others	43	51

X: Number of nuclear families.
n: Total number of families.

Hypothesis	Z-value (Cal.)	p-value
H_0:Proportion of nuclear families in hindu religion and proportion of nuclear families in other religion are equal VS H_1: Proportion of nuclear families in hindu religion and proportion of nuclear families in other religion are not equal	3.87	0.002

Conclusion: Here, since p-value(0.002) < level of significance(0.05).We reject H_0. i.e. Proportion of nuclear families in hindu religion and proportion of nuclear families in other religion are not equal

4.2.3 Proportion test for testing the proportion of Measles 1 receivers and the proportion of Measles 2 receivers are equal.

	X	n
Measles 1	119	286
Measles 2	67	246

X: Number of children received Measles vaccine.
N: Total number of Eligible children.

Hypothesis	Z-value (Cal.)	p-value
H_0: Proportion of Measles 1 receivers and the proportion of Measles 2 receivers are equal. VS H_1: Proportion of Measles 1 receivers and the proportion of Measles 2 receivers are not equal.	3.70	0.00

Conclusion: Here, since p-value(0.00) $<$ level of significance(0.05).We reject H_0. i.e. Proportion of Measles 1 receivers and the proportion of Measles 2 receivers are equal.

4.2.4 Proportion test for testing the proportion of 1st booster and the proportion of 2nd booster receivers are equal.

	X	n
Booster 1	119	286
Booster 2	67	246

X: Number of children received Booster 1 vaccine.
n: Total number of eligible children.

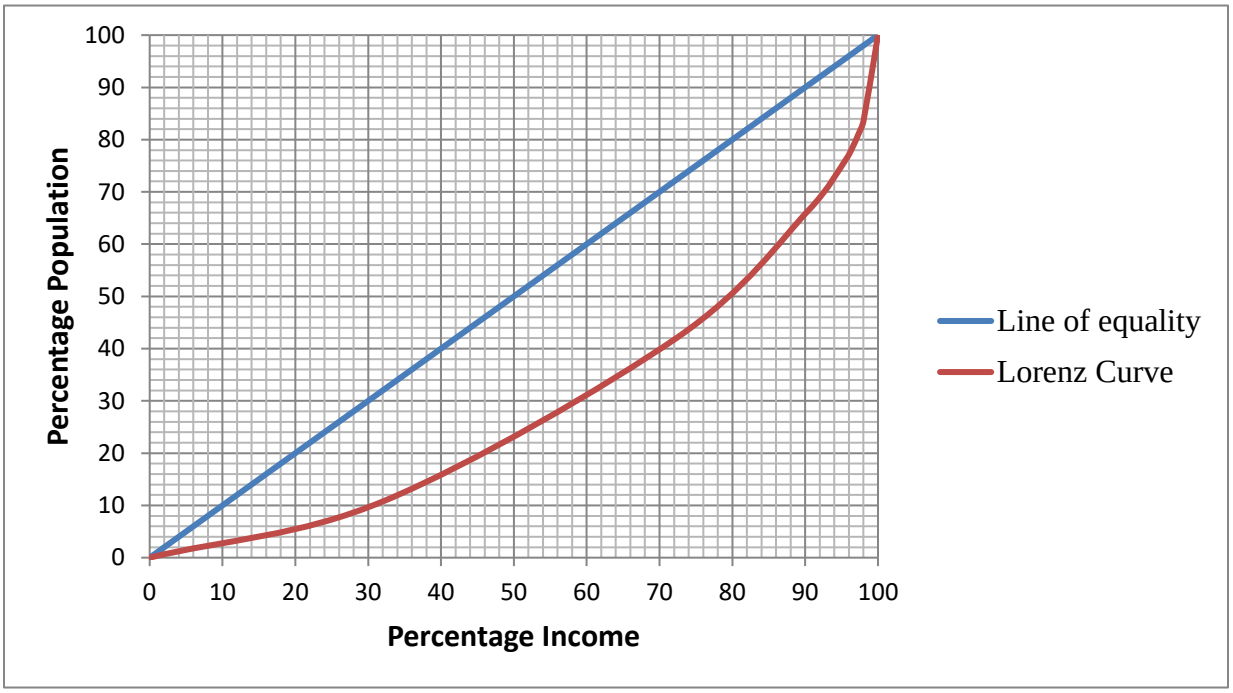
Hypothesis	Z-value (Cal.)	p-value
H_0: Proportion of 1st booster receivers and the proportion of 2nd booster receivers are equal. VS H_1: Proportion of 1st booster receivers and the proportion of 2nd booster receivers are not equal.	3.53	0.00

Conclusion: Here, since p-value(0.00) < level of significance(0.05).We reject H_0. i.e. Proportion of 1st booster receivers and the proportion of 2nd booster receivers are equal.

4.3 Lorenz curve and Gini coefficient:

4.3.1 Lorenz curve:

Income Group (In Thousands)	No. of Households	Income	Percent Population	Percent Income	Cumulative Percent Income	Cumulative Percent Population
0-10	260	1426800	32.2580	10.8955	0	0
10-20	323	4077500	40.0744	31.1371	10.8955	32.2580
20-30	146	3198000	18.1141	24.4209	42.0326	72.3325
30-40	41	1261000	5.0868	9.6294	66.4536	90.4466
40-50	7	280000	0.8684	2.1381	76.0830	95.5334
50-60	8	405000	0.9925	3.0927	78.2211	96.4019
60-70	5	302000	0.6203	2.3061	81.3139	97.3945
70 & above	16	2145000	1.9851	16.3799	83.6200	98.0148
TOTAL	806	13095300	100	100	100	100



Conclusions:

- The curve above shows the income distribution compared to a straight diagonal representing perfect equality. At the 20th income percentile, the cumulative income is near about 6%.In other words, the bottom 20% of the population takes in near about 6% of the total income. If there were a perfectly equal society, the bottom 20% would earn 20% of the total.
- At the 40th income percentile, the cumulative income is near about 16%.In other words, the bottom 40% of the population takes in near about 16% of the total income. If there were a perfectly equal society, the bottom 40% would earn 40% of the total.
- At the 60th income percentile, the cumulative income is near about 30%.In other words, the bottom 60% of the population takes in near about 30% of the total income. If there were a perfectly equal society, the bottom 60% would earn 60% of the total.
- At the 80th income percentile, the cumulative income is near about 51%.In other words, the bottom 80% of the population takes in near

about 51% of the total income. If there were a perfectly equal society, the bottom 80% would earn 80% of the total.

4.3.2 Gini Coefficient

Gini Coefficient	0.40634

Interpretation:

- G= 0 interprets the exact income equality.
- G=1 interprets the exact income inequality.
- As G increases income inequality increases.

Conclusion: Since Gini coefficient (0.40634) lies between 0.4 and 0.5 it represents a big income gap.

4.4 Binary Logistic Regression.

4.4.1 Binary logistic regression to study the effect of different levels of variables on delivery type of mother.

SR No	Edu. Group	Age Group	SES	Working Mother	Type of Delivery	SR No	Edu. Group	Age Group	SES	Working Mother	Type of Delivery
1	3	2	2	1	0	24	3	4	2	1	0
2	5	2	1	1	0	25	3	4	3	1	0
3	2	2	2	1	0	26	5	2	2	2	0
4	3	1	2	1	0	27	3	3	4	1	0
5	5	4	1	2	1	28	3	2	1	2	0
6	5	3	5	1	1	29	4	4	4	1	1
7	3	2	2	1	0	30	3	2	1	1	1
8	5	2	1	1	1	31	4	2	4	1	1
9	2	2	2	1	1	32	4	1	5	1	0
10	3	1	2	1	0	33	5	2	3	1	1
11	5	4	1	2	1	34	4	1	2	1	0

12	3	2	2	1	0	35	5	2	3	1	0
13	5	2	1	1	0	36	5	1	3	2	1
14	2	2	2	1	0	37	3	2	3	1	0
15	3	1	2	1	0	38	1	4	2	1	1
16	5	4	1	2	1	39	5	3	1	2	1
17	4	2	2	1	0	40	3	2	2	1	1
18	5	2	1	1	1	41	3	1	3	1	1
19	2	2	2	1	1	42	1	2	3	1	0
20	3	1	2	1	0	43	5	2	2	1	0
21	5	4	1	2	1	44	3	1	2	2	0
22	3	4	3	1	0	45	3	2	3	1	0
23	4	3	2	1	0	46	3	2	1	1	1

Notations:

Variable	Level	Notation
Education group of mother	Illiterate	1
	Primary	2
	Secondary	3
	High. Secondary	4
	Graduate	5
Age group of mother	18-21	1
	22-25	2
	25-28	3
	28 and above	4
SES	Class I	1
	Class II	2
	Class III	3
	Class IV	4
	Class V	5
Working mother	No	1
	Yes	2

Response	Level	Notation
Type of delivery	Normal	0
	Cesarean	1

Variable	Level	B	SE	Odds Ratio
Education group of mother	Illiterate (Ref.)			
	Primary	1.21	1.93	3.3499
	Secondary	-1.87	1.67	0.1546
	High. Secondary	-2.03	2.45	0.1316
	Graduate	-0.35	1.75	0.7073
Age group of mother	18-21 (Ref.)			
	22-25	-1.01	1.37	0.3656
	26-29	-1.05	2.09	0.3487
	29 and above	0.40	1.44	1.4892
SES	Class I (Ref.)			
	Class II	-2.80	1.31	0.0607
	Class III	-1.79	1.19	0.1672
	Class IV	0.70	2.22	2.0109
	Class V	-0.89	2.39	0.4126
Working mother	No (Ref.)			
	Yes	0.14	1.22	1.1540

Conclusions:

- Based on table, the probability of cesarean delivery for mothers having primary education is 3.3499 times more likely than mothers having no education(Illiterate).The probability of cesarean delivery for mothers having secondary education is 0.1546 times more likely than mothers having no education(Illiterate).The probability of cesarean delivery for mothers having higher secondary education is 0.1316 times more likely than mothers having no education(Illiterate).The probability of cesarean delivery graduate mothers is 0.7073 times more likely than mothers having no education(Illiterate) .

- Based on table, the probability of cesarean delivery for mothers from age group 22-25 is 0.3656 times more likely than mothers from age group 18-21.The probability of cesarean delivery for mothers from age group 26-29 is 0.3487 times more likely than mothers from age group 18-21.The probability of cesarean delivery for mothers from

age group 29 and above is 1.4892 times more likely than mothers from age group 18-21.

- Based on table, the probability of cesarean delivery for mothers having SES as Class II is 0.0607 times more likely than mothers having SES as Class I. The probability of cesarean delivery for mothers having SES as Class III is 0.1672 times more likely than mothers having SES as Class I. The probability of cesarean delivery for mothers having SES as Class IV is 2.0109 times more likely than mothers having SES as Class I. The probability of cesarean delivery for mothers having SES as Class V is 0.4126 times more likely than mothers having SES as Class I.

- Based on table, the probability of cesarean delivery for working mothers is 1.1540 times more likely than non-working mothers.

4.4.2 Binary logistic regression to study the effect of different levels of variables on presence of working mother in the family.

Variable	Level	B	SE	Odds Ratio
Education group of HOF	Illiterate (Ref.)			
	Primary	1.21	1.93	1.9715
	Secondary	-1.87	1.67	1.2354
	High. Secondary	-2.03	2.45	0.7223
Age group of mother	Other (Ref.)			
	Hindu	-1.01	1.37	4.7338
SES	Class I (Ref.)			
	Class II	-2.80	1.31	1.3269
	Class III	-1.79	1.19	0.5695
	Class IV	0.70	2.22	1.5136
	Class V	-0.89	2.39	0.1602

Conclusions:

- Based on table, the probability of presence of working mother in the family in which education level of HOF is primary is 1.9715 time more likely than family in which HOF is illiterate. The probability of presence of working mother in the family in which education level of HOF is Secondary is 1.2354 time more likely than family in which HOF is illiterate . The probability of presence of working mother in the family in which education level of HOF is Higher Secondary or above is 0.7223 time more likely than family in which HOF is illeterate .

- Based on table, the probability of presence of working mother in the hindu religion family is 4.7338 time more likely than other religion family.

- Based on table, the probability of presence of working mother in the family having SES as Class II is 1.3269 times more likely than family having SES as Class I. The probability of presence of working mother in the family having SES as Class III is 0.5695 times more likely than family having SES as Class I. The probability of presence of working mother in the family having SES as Class IV is 1.5136 times more likely than family having SES as Class I. the probability of presence of working mother in the family having SES as Class V is 0.1602 times more likely than family having SES as Class I

4.5 Crude Birth Rate and Crude Death Rate:

Birth Rate	Death Rate
11.9916	5.2137

- **Comparison with national Birth Rate (17.86)**

Hypothesis	Z-value (Cal.)	p-value
H_0: $p=p_0$ VS H_1: $p\neq p_0$ Where, p: Population Proportion P_0=0.01786	-0.4512	0.156

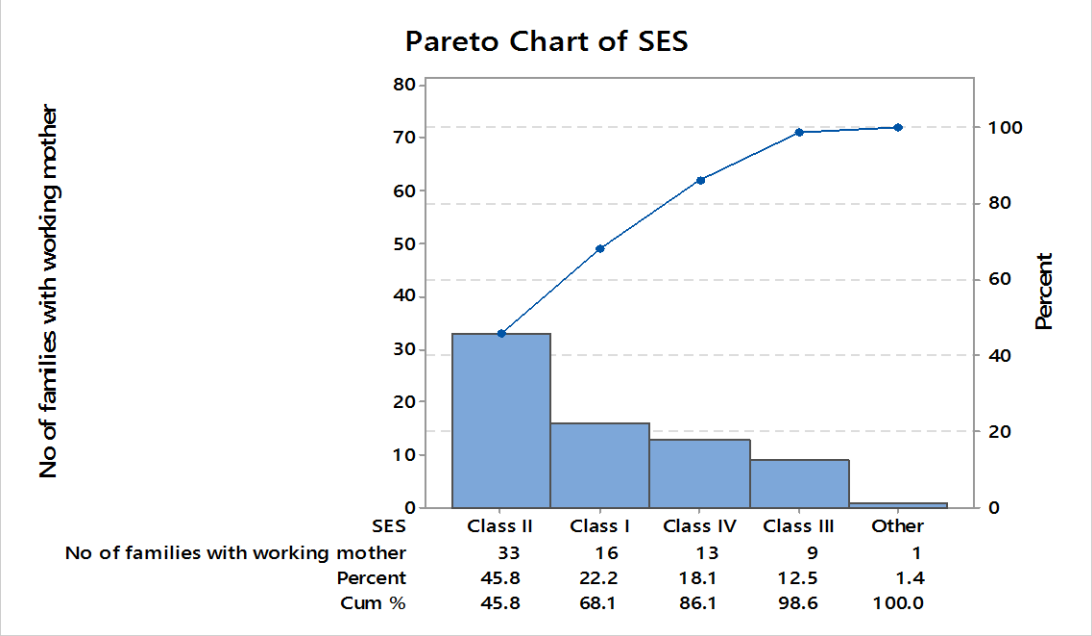
Conclusion: Here, since p-value(0.156) > level of significance(0.05).We fail to reject H_0. i.e. Their is no significant difference between national birth rate and survey birth rate.

- **Comparison with national Death Rate (7.234)**

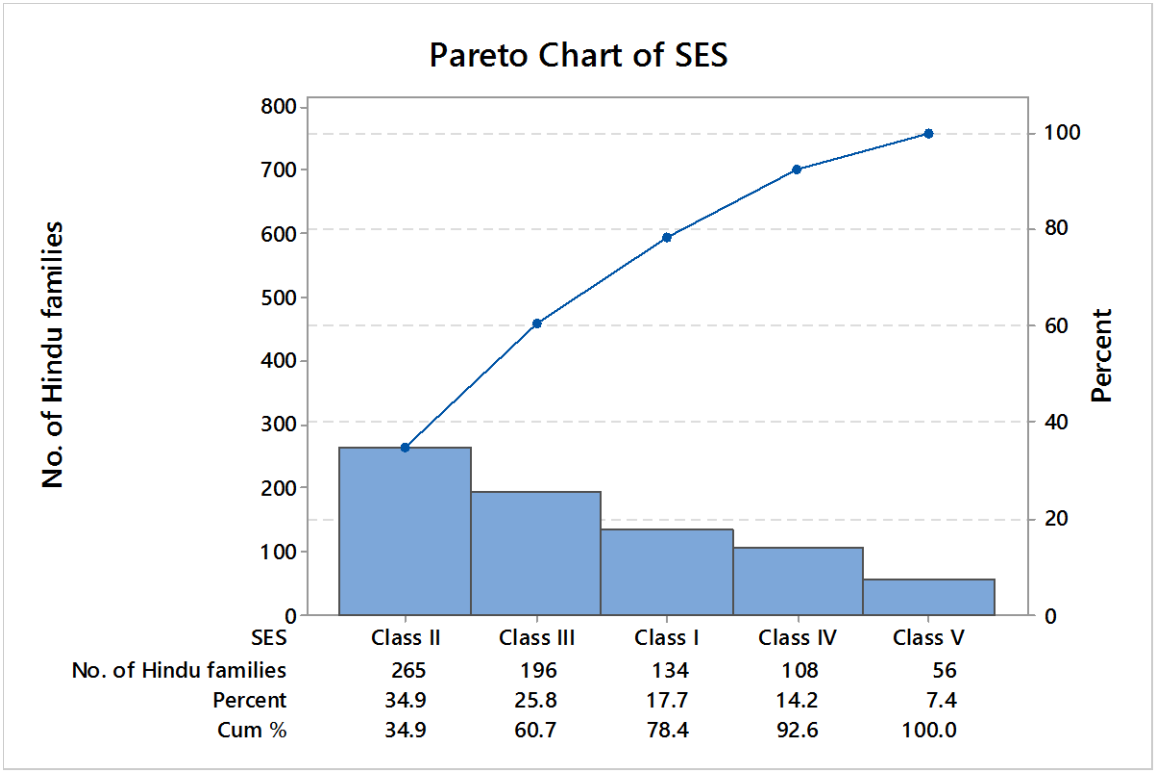
Hypothesis	Z-value (Cal.)	p-value
H_0: $p=p_0$ VS H_1: $p\neq p_0$ Where, p: Population Proportion P_0=0.007	-0.7585	0.469

Conclusion: Here, since p-value(0.156) > level of significance(0.05).We fail to reject H_0. i.e. Their is no significant difference between national death rate and survey death rate.

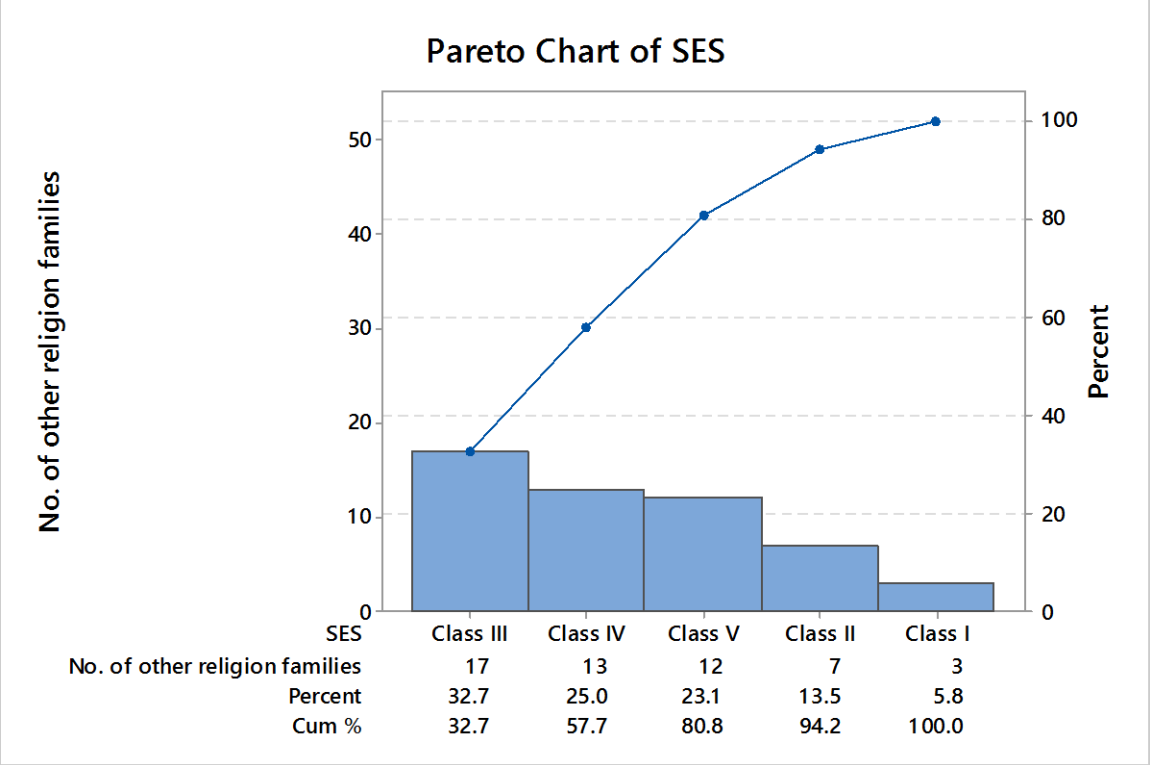
4.6 Pareto Chart:



Conclusion: 68.1% working mothers are from class II and class I and 86.1 % working mothers are from class I, class II and class IV.



Conclusion: 78.4% of the Hindu families in the survey belongs to the class I, Class II and Class III.



SES	Class III	Class IV	Class V	Class II	Class I
No. of other religion families	17	13	12	7	3
Percent	32.7	25.0	23.1	13.5	5.8
Cum %	32.7	57.7	80.8	94.2	100.0

Conclusion: 80.8% of the other religion families in the survey belongs to the class III, Class IV and Class V.

Chapter 5: Conclusions

- We can conclude that there is no association between education level of HOF and latrine facility, education level of HOF and waste disposal method, education level of HOF and presence of working mother in family.
- There is association between pattern of food purchase and socio-economic status, type of house and ventilation.
- Proportion of Hindu religion families and proportion of other religion families above poverty line are equal, proportion of measles 1 and measles 2 receivers are equal, proportion of 1^{st} booster and 2^{nd} booster receivers are equal.
- Proportion of nuclear families in Hindu religion and proportion of nuclear families in other religion are not equal.
- Lorenz curve indicates the presence of income inequality in study area and Gini coefficient (0.40634) represents that there is big income gap in study area.
- The delivery type of mother (normal or caesarean) is affected by number of factors such as education group of mother, age group of mother, socio-economic status, mother is working or not.
- Also, presence of working mother in family is affected by factors such as education group of HOF, age group of mother, socio-economic status.
- There is no significance difference between birth rate, death rate of study area and national birth rate and death rate.

References

- Gopal K. Kanji – 2006: 100 Statistical Test
- Kassahun Trueha Dumga, Ayele Gebeyehu, Gedif Mulat Alemayehu (2018) Statistical Analysis of Delivery Care Service Utilization of Women in Gurage Zone, Ethiopia.

Websites:

- bvuniversity.edu.in

Software:

- MS-Excel
- Minitab