CHAPTER I

INTRODUCTION

Local government is the government of the village and district level. It is the government closest to common people that involves in day-to-day life and attempt to resolve problems of ordinary citizens. Democracy is in fact about meaningful participation and also about accountability. Hence, strong and vibrant local governments ensure about active participation and purposeful accountability. Local Self Government is the management of local affairs by such local bodies who have been elected by the local people. The local self-Government includes both rural and urban government. It is the third level of the government. There are 2 types of local government in operation – panchayats in rural areas and Municipalities in urban areas. Our Constitution has envisaged LSGIs as institutions of local governance. In accordance with their status, the Kerala Panchayat Raj Act, 1994 and the Kerala Municipality Act, 1994 have devolved many functions, powers and transferred institutions to the LSGIs, to enable them to provide optimum service to the people. Panchayats/Urban local bodies should emerge as such bodies that can satisfy the needs of the people invoking the functions and powers assigned to them. It is the bounden duty of the elected representatives to steer the LSGIs in such a direction.

The purposes of local government in India are:

- To build the infrastructure of development like road, transport etc.
- To build and maintain community assets;
- To promote agricultural development through management and control of minor irrigation and water management; soil conservation and land improvement;
- To promote social forestry and animal husbandry, dairy and poultry;
- To promote the development of village industry; and
- To manage and control education and health at the local level.

In nutshell, the local bodies are institutions of empowering people for self-government.

Grievance Redressal Mechanism

Grievance Redress Mechanism is part and parcel of the machinery of any administration. No administration can claim to be accountable, responsive and user-friendly unless it has established an efficient and effective grievance redress mechanism. In fact, the grievance redress mechanism of an organization is the gauge to measure its efficiency and effectiveness as it provides important feedback on the working of the administration. The grievances of public are received at various points in the Government of India. There are primarily two designated nodal agencies in the Central Government handling these grievances. These agencies are: -

(i) Department of Administrative Reforms and Public Grievances, Ministry of Personnel, Public Grievances and Pensions
(ii) Directorate of Public Grievances, Cabinet Secretariat

Department of Administrative Reforms & Public Grievances
Department of Administrative Reforms & Public Grievances is the nodal agency in respect of policy initiatives on public grievances redress mechanism and citizen centric initiatives. The role of Department of Administrative Reforms and Public Grievances consists primarily to undertake such citizen-centric initiatives in the fields of administration reforms and public grievances in the Government so as to enable the Government machinery to deliver quality public services to the citizen in a hassle-free manner and eliminate the causes of grievance. The grievances received by the Department are forwarded to the concerned Ministries/Departments/State Governments/UTs, who are dealing with the substantive function linked with the grievance for redress under intimation to the complainant. The Department 'takes up' about 1000 grievances every year depending upon the seriousness of the grievance and follows them regularly till their final disposal. This enables the Department to evaluate the effectiveness of the grievance redress machinery of the concerned government agency. On the basis of the grievances received, Department identifies the problem areas in Government which are complaint-prone. These problem areas are then subjected to studies and remedial measures are suggested to the Department/Organisation concerned.

Directorate of Public Grievances (DPG)
Based on the review of the public grievances redress machinery in Government of India carried out in 1987, the Directorate of Public Grievances was set up in the Cabinet Secretariat with effect from 01.04.88. This Directorate was set up initially to look into individual

complaints pertaining to four Central Government Departments which were more prone to public complaints. Subsequently, more Departments having larger public interface were added to its purview and presently this Directorate is handling grievances pertaining to 16 Central Government Organisations. The Directorate was envisaged as an appellate body investigating grievances selectively and particularly those where the complainant had failed to get redress at the hands of internal machinery and the hierarchical authorities. Unlike the Department of AR&PG, Directorate of Public Grievances has been empowered to call for the files and officers for discussion to see that grievance handling has been done in a fair, objective and just manner. Wherever the Directorate is satisfied that the grievance has not been dealt in such a manner, it makes suitable recommendations for consideration and adoption by the concerned Ministry/Department which are required to be implemented within a period of one month. The empowered and enlightened citizenry of today is far more demanding and the government, therefore, has to develop, evolve and enable itself to meet the evolving demands of the society that it has to serve. The society today is impatient with the old system of governance which is not coming up to its expectations. To them, a government employee is perceived as insensitive, aloof, corrupt and overall the administrative system as autocratic, opaque and with no work culture This requires a paradigm shift in governance to a system where the citizen is in the centre and he is consulted at various stages of formulation and implementation of public policy. To achieve this objective, India needs a public service which is capable, innovative and forward looking. The traditional role of civil service which was of administrator, service provider and controller of development activities has to make way for the new roles of facilitator and regulator so as to create best environment and conditions in the country for building a nation of excellence.

Department of Administrative Reforms & Public Grievances is the nodal agency in Government of India for formulation and implementation of such policies and strategic initiatives so as to enable and equip the government machinery to meet the challenges involved in achieving this objective. Department of Administrative Reforms and Public Grievances is the driving engine of reforms in administration and governance. The Department proposes to introduce and lead Change to establish a public service of quality, efficiency, integrity and effectiveness and modernize the public service. It is the nodal agency in government for facilitating administrative improvements and reengineering of processes across the government. Citizen’s Charter initiative, Public Grievance Policy, Quality Management in Government, e-Governance, Review of Administrative Laws etc.

Documentation and Dissemination of Best Practices, Organisation & Methods, Information & Facilitation Counters, Civil Services Reforms are some of the areas under the ambit of Department of Administrative Reforms & Public Grievances.

Binary Logistic Regression

Regression analysis is a process that estimates the probability of the target variable given some linear combination of the predictors. Binary logistic regression (LR) is a regression model where the target variable is binary, that is, it can take only two values, 0 or 1. It is the most utilized regression model in readmission prediction, given that the output is modelled as readmitted (1) or not readmitted (0).

Survival models, on the other hand, relate the features to the time that passes before the event (i.e. readmission) occurs.

In recent years, machine learning and data mining have emerged as approaches that can potentially improve the prediction ability of readmission risk prediction models. Those techniques introduce classification algorithms widely used in multiple predictive modelling fields. These are not limited to the classifier itself, as they also encompass a wider set of techniques such as feature selection, variable discretization or missing value imputation among others.

Logistic regression sometimes called the logistic model or logit model, analyses the relationship between multiple independent variables and a categorical dependent variable, and estimates the probability of occurrence of an event by fitting data to a logistic curve. There are two models of logistic regression, binary logistic regression and multinomial logistic regression. Binary logistic regression is typically used when the dependent variable is dichotomous and the independent variables are either continuous or categorical. When the dependent variables not dichotomous and is comprised of more than two categories, a multinomial logistic regression can be employed.

Kerala Institute of Local Administration (KILA)

Kerala Institute of Local Administration, abbreviated as KILA, is an autonomous training, research and consultancy organisation constituted under the Ministry of Local Self Government, Government of Kerala, registered as per Travancore – Cochin Literacy, Scientific and Charitable Societies Act-1955. KILA was established in 1990, in the pattern of a national institute, with the main objective of strengthening decentralization and local governance. The internship was held at KILA and Here, the data is collected from the Requirement Assessment Forum of the Newly Elected Grama panchayat Representatives

online training section conducted by KILA. The data taken under study is a cross sectional data with 17388 observations with dependent variable as the awareness of the elected representatives about the knowledge in the local government activities they represent and independent variables as age and awareness about grievance redressal mechanism. The powerful statistical software Stata is used here which enables users to analyse, manage, and produce graphical visualizations of data. It is widely used in the field of economics, biomedicine, and political science to examine data patterns.

CHAPTER II

REVIEW OF LITERATURE

This chapter reviews available literature on the powers and functioning of local government bodies, literature on the logistic regression using STATA software.

D Narayana (2005) made study on local governance without capacity building with the passing of the 73rd and 74th constitutional amendments in India the structure of governance has changed permanently from a two-tier to a three-tier system with union, state and panchayats/nagar palikas. However, a necessary condition for the transformation of panchayats into local governments is devolution of powers, resources and functions to them and capacity building among the elected representatives. This paper analyses the functioning of the elected representatives at the gram panchayat level in Madhya Pradesh, Tamil Nadu and Kerala. In Madhya Pradesh and Tamil Nadu, elected ward members show poor awareness of powers and responsibilities, but Kerala is different. In Madhya Pradesh and Tamil Nadu, panchayats are perceived as agents of state governments, whereas in Kerala they are taken as local governments. In all three states, panchayat presidents understand the legislation better but planning for development is a far cry as little effort seems to have gone into capacity building and devolution of powers and resources. Unless larger powers and resources are devolved and elected representatives are trained, local government will be a dream.

Richard Williams (2006) This article describes the gologit2 program for generalized ordered logit models. gologit2 is inspired by Vincent Fu's gologit routine (Stata Technical Bulletin Reprints 8: 160–164) and is backward compatible with it but offers several additional powerful options. A major strength of gologit2 is that it can fit three special cases of the generalized model: the proportional odds/parallel-lines model, the partial proportional odds model, and the logistic regression model. Hence, gologit2 can fit models that are less restrictive than the parallel-lines models fitted by ologit (whose assumptions are often violated) but more parsimonious and interpretable than those fitted by a nonordinal method, such as multinomial logistic regression (i.e., mlogit). Other key advantages of gologit2 include support for linear constraints, survey data estimation, and the computation of estimated probabilities via the predict command.

Xing Liu, Hari Koirala (2012) The proportional odds (PO) assumption for ordinal regression analysis is often violated because it is strongly affected by sample size and the number of covariate patterns. To address this issue, the partial proportional odds (PPO) model and the generalized ordinal logit model were developed. However, these models are not typically used in research. One likely reason for this is the restriction of current statistical software packages: SPSS cannot perform the generalized ordinal logit model analysis and SAS requires data restructuring. This article illustrates the use of generalized ordinal logistic regression models to predict mathematics proficiency levels using Stata and compares the results from fitting PO models and generalized ordinal logistic regression models.

KS Pushpa (2016) had conducted a study to know the legal knowledge of the elected members of Panchayat Raj Institutions (PRIs) in Tumkur district Karnataka State. The introduction of the community development programme (CDP) in the first five-year plan gave a new dimension as well as impetus to the participation of people in development. Development implies on overall positive change in the physical quality of life. This positive change for the better encompasses economic as well as social aspects. In other words, Development implies growth with justice; it means an improvement in the quality of life through better Health, Education, Housing and Welfare.

Victorovna, Bikram Biswas (2019) Local government is a fundamental part of Government and undeniable part of development in all spheres of a country. While most of the constructions in the worlds pronounce various provisions of having local government, it's got more significance in recent time. Moreover, without overall participation of the population of a country, development of all sectors can't be achieved. Participation of grassroots people are encouraged and ensured by local institutions. It also provides the opportunities to the population oriented of making policies and implementing decision better for them which is ultimately a step forwarding sustainable development. Since the time of independence from Pakistan in 1971, Bangladesh has local government system in the constitution of the country. Various amendments regarding local government have been brought basically when regimes changed. Thus, till to date, the constitutional provisions have been far dream from the reality. In the very root level, the local government established maintaining the constitutional provisions can extend the practice of democracy. Thus, following constitutional scheme and

to enhance good governance at all levels of the society, it's ripe time for the concerned authority to be more conscious to establish local government.

Joydeep Guha and Bhaskar Chakrabarti (2019) presented paper about the role of local democracy and governance to achieve the sustainable development goals (SDGs). Increased reliance on locally generated revenue, difficulties in managing networks of actors with diverse goals and objectives, imperfect flow of information, and trust deficit in stakeholders pose major challenges to achieving SDGs locally. By doing a systematic review of the recent literature on decentralisation with examples from different local governments, the paper outlines ways in which these challenges could be addressed. The paper also highlights the need for enhancing local leadership capabilities and demarcation of responsibilities among local politicians and bureaucrats, a point missed in the SDG agenda.

Ernest Yeboah Boateng, Daniel A. Abaye (2019) This study explored and reviewed the logistic regression (LR) model, a multivariable method for modelling the relationship between multiple independent variables and a categorical dependent variable, with emphasis on medical research. Logistic regression concepts such as odds, odds ratio, logit transformation, logistic curve, assumption, selecting dependent and independent variables, model fitting, reporting and interpreting were presented.

CHAPTER III
INTERNSHIP PROBLEM

The local government refers specifically to a level of administration that is both geographically-localised and has limited powers. One of the most common challenges for local government associations (LGAs) around the world is to have enough financial resources to fulfil their core functions. These core functions are to represent the interest of their members, the local governments of the country; to provide services to these local governments and to offer them a platform for exchange and mutual learning. Financial management is the ability to prioritise and reconcile unlimited demands and needs with limited financial resources. In this paper we deal with

- Binary logistic regression using Stata
- Interpretation using coefficients and odds ratio
- Probability and predicted probability(margins)
- Does the elected representatives have enough knowledge about local government activities body they represent?
- Does the elected representatives have enough knowledge about grievance redressal mechanism?

CHAPTER IV
DATA AND METHODOLOGY

The data used for the study is the primary data set collected from the Requirement Assessment Forum of the Newly Elected Grama panchayat Representatives online training section conducted at KILA. The data taken under study is a cross sectional data with 17388 observations with dependent variable as the awareness of the elected representatives about knowledge in local government activities they represent and independent variables as age and awareness about the grievance redressal mechanism. Initial data set had some missing data sets and after correcting it converted the dependent variable into dichotomous variable. One dependent variable is binary and two independent variables which includes one continuous and one categorical variable. The powerful statistical software Stata is used here which enables users to analyse, manage, and produce graphical visualizations of data. It is widely used in the field of economics, biomedicine, and political science to examine data patterns.

Logistic regression is a statistical technique used in research designs that call for analysing the relationship of an outcome or dependent variable to one or more predictors or independent variables when the dependent variable is either (a) dichotomous, having only two categories, for example, whether one uses illicit drugs (no or yes); (b) unordered polytomous, which is a nominal scale variable with three or more categories, for example, political party identification (Democrat, Republican, other, or none); or (c) ordered polytomous, which is an ordinal scale variable with three or more categories, for example, level of education completed (e.g., less than elementary school, elementary school, high school, an undergraduate degree, or a graduate degree). Here, the basic logistic regression model for dichotomous outcomes is examined, noting ...

CHAPTER V

DATA ANALYSIS AND INTERPRETATION

A binomial logistic regression is used to predict a dichotomous dependent variable based on one or more continuous or nominal independent variables. It is the most common type of logistic regression and is often simply referred to as logistic regression. In Stata they refer to **binary outcomes** when considering the binomial logistic regression. In many ways a binomial logistic regression can be considered as a multiple linear regression, but for a dichotomous rather than a continuous dependent variable. Here the data is collected from the Requirement Assessment Forum of the Newly Elected Grama panchayat Representatives. The data taken under study is a cross sectional data with 17388 observations with dependent variable as the awareness of representatives about knowledge of local government system activities they represent and independent variables as age and awareness about the grievance redressal mechanism. We consider first the case where the response yi is binary, assuming only two values that for convenience we code as one or zero.

For example, we could define

yi = 1 if they are aware about local government system activities they represent

yi = 0 otherwise

Let:

$$P_{i=} P_r (Y=1|X=x_i)$$

Then we can write the model:

$$\text{Log } (P_i / 1-P_i) = \text{logit } P_{i=} \beta_1+\beta_2 x_i+\beta_3 z_i$$

Where Pi, is the probability of awareness of representatives about knowledge of local government activities they represent

X_i = age

Z_i = awareness about grievance redressal mechanism.

β_1 = intercept or constant

β 2= log odds of x_i

β 3= log odds of z_i

We can write the model in terms of odds as:

$$Pi/(1-Pi) = \exp(\beta_1+\beta_2 x_i+\beta_3 z_i)$$

Or in terms of the probability of the outcome occurring as:

$$Pi= \exp(\beta_1+\beta_2 x_i+\beta_3 z_i)/(1 + \exp(\beta_1+\beta_2 x_i+\beta_3 z_i))$$

Conversely the probability of the outcome not occurring is

$$1- Pi = 1/ (1 + \exp(\beta_1+\beta_2 x_i+\beta_3 z_i))$$

Any factor that affects the probability will alter not just the mean but also the variance of the observations. This suggest that a linear model that allows the predictors to affect the mean but assumes that the variance is constant will not be adequate for the analysis of binary data.

A brief summary of the given data is shown below.

Overall summary

Variable	Obs	Mean	Std. Dev.	Min	Max
Age	17,388	45.6501	9.81292	21	78
q11	17,388	.6017943	.4895423	0	1
q22	17,388	.5822981	.4931947	0	1

Dependent variable q11

Q11	Freq.	Percent	Cum.
0	6,924	39.82	39.82
1	10,464	60.18	100.00
Total	17,388	100.00	

Independent variable q22

Q22	Freq.	Percent	Cum.
0	7,263	41.77	41.77
1	10,125	58.23	100.00
Total	17,388	100.00	

Independent variable age

AGE	Freq.	Percent	Cum.
21	29	0.17	0.17
22	44	0.25	0.42
23	82	0.47	0.89
24	78	0.45	1.34
25	91	0.52	1.86
26	114	0.66	2.52
27	144	0.83	3.35
28	162	0.93	4.28
29	178	1.02	5.30
30	203	1.17	6.47
31	267	1.54	8.01
32	235	1.35	9.36
33	299	1.72	11.08
34	336	1.93	13.01
35	405	2.33	15.34
36	425	2.44	17.78
37	470	2.70	20.49

38	531	3.05	23.54
39	596	3.43	26.97
40	581	3.34	30.31
41	633	3.64	33.95
42	656	3.77	37.72
43	630	3.62	41.34
44	675	3.88	45.23
45	760	4.37	49.60
46	755	4.34	53.94
47	707	4.07	58.01
48	757	4.35	62.36
49	617	3.55	65.91
50	611	3.51	69.42
51	632	3.63	73.06
52	509	2.93	75.98
53	451	2.59	78.58
54	412	2.37	80.95
55	385	2.21	83.16
56	423	2.43	85.59
57	394	2.27	87.86
58	368	2.12	89.98
59	299	1.72	91.70
60	249	1.43	93.13
61	252	1.45	94.58
62	144	0.83	95.40
63	170	0.98	96.38
64	127	0.73	97.11

65	110	0.63	97.75
66	92	0.53	98.27
67	64	0.37	98.64
68	66	0.38	99.02
69	58	0.33	99.36
70	30	0.17	99.53
71	26	0.15	99.68
72	11	0.06	99.74
73	13	0.07	99.82
74	15	0.09	99.90
75	7	0.04	99.94
76	5	0.03	99.97
77	1	0.01	99.98
78	4	0.02	100.00
Total	17,388	100.00	

Where, q11 is the dependent variable representing the awareness of representatives about the knowledge of local government activities they represent and q22 represents awareness about grievance redressal mechanism.

ASSUMPTIONS OF LOGISTIC REGRESSION

Logistic regression does not require many of the principle assumptions of linear regression models that are based on ordinary least squares method–particularly regarding linearity of relationship between the dependent and independent variables, normality of the error distribution, homoscedasticity of the errors, and measurement level of the independent variables. Logistic regression can handle non-linear relationships between the dependent and independent variables, because it applies a non-linear log transformation of the linear regression. The error terms (the residuals) do not need to be multivariate normally distributed–although multivariate normality yields a more stable solution. The variance of

errors can be heteroscedastic for each level of the independent variables. Logistic regression can handle not only continuous data but also discrete data as independent variables. However, some other assumptions still applied.

There are six assumptions that underpin binomial logistic regression. If any of these six assumptions are not met, you might not be able to analyse your data using a binomial logistic regression because you might not get a valid result.

ASSUMPTIONS

1. Dependent variable should consist of two categorical, independent (unrelated) groups (i.e., a dichotomous variable). The two categories of the dependent variable need to be mutually exclusive and exhaustive. Since the dependent variable in the model is a binomial variable with values 0 and 1, thus this assumption is satisfied.

2. There should be two or more independent variables, which should be measured at the continuous or nominal level. Here our independent variables are age and other one awareness about grievance redressal mechanism is a binomial variable hence it satisfies the assumption.

3. There should be independence of observations, which means that there should not be any relationship between the observations. Here, the observations are collected from different individuals, so, we don't need to use any further test to check the independence of observations.
4. Data must not show multicollinearity, which occurs when you have two or more independent variables that are highly correlated with each other. To check the multicollinearity, we use spearman's ratio

Number of obs = 17388

Spearman's rho = 0.0259

Test of Ho : age and q22 are independent

Prob > |t| = 0.0006

Here the value of spearman's ratio .0259 indicates that there is no or negligible correlation between the independent variables.

5. There needs to be a linear relationship between any continuous independent variables and the logit transformation of the dependent variable. The Box-Tidwell test is used to check for linearity between the predictors and the logit. This is done by adding log transformed interaction terms between the continuous independent variables and their corresponding natural log into the model.

   Null hypothesis H_o: there is linear relationship between any continuous independent variables and the logit transformation of the dependent variable.

   Alternative hypothesis H_1: there is no linear relationship between any continuous independent variables and the logit transformation of the dependent variable.

Box-Tidwell regression model

```
Iteration 0:  Deviance =  19841.18
Iteration 1:  Deviance =  19841.18 (change = -2.90e-06)
-> gen double Iage__1 = X^0.1210-1.201624332 if e(sample)
-> gen double Iage__2 = X^0.1210*ln(X)-1.824571357 if e(sample)
   (where: X = age/10)

[Total iterations: 1]

Box-Tidwell regression model

Logistic regression                     Number of obs     =     17,388
                                        LR chi2(3)        =    3537.94
                                        Prob > chi2       =     0.0000
Log likelihood = -9920.5912             Pseudo R2         =     0.1513
```

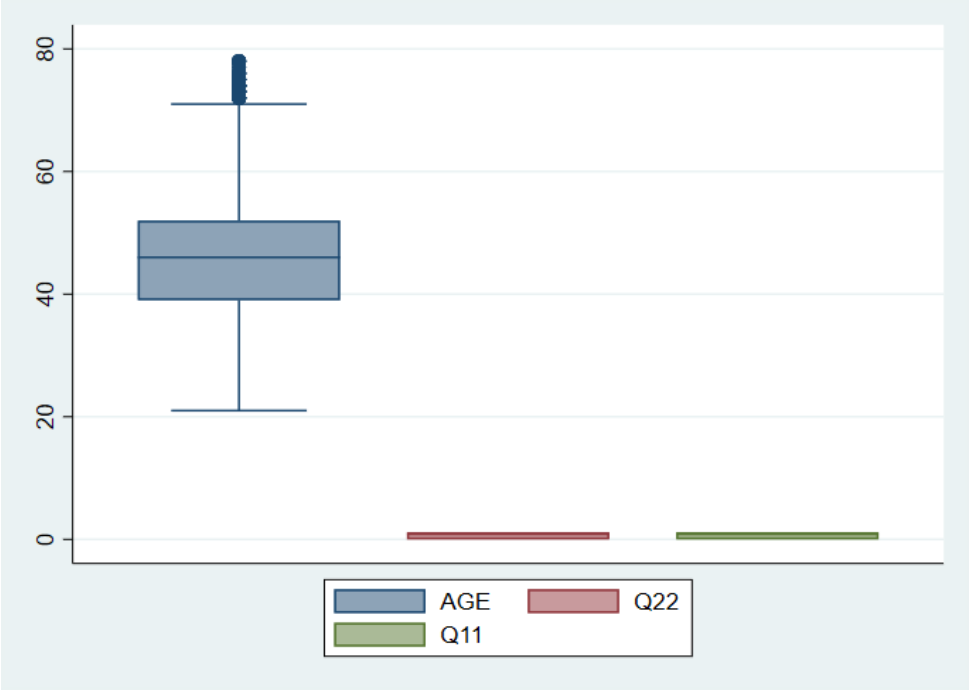
q11	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
Iage__1	1.861571	33.26609	0.06	0.955	-63.3388	67.06191
Iage_p1	0.00861	3.426941	0	0.998	-6.70807	6.725291
q22	1.946622	0.034628	56.21	0	1.878752	2.014492
_cons	-0.63689	0.027335	-23.3	0	-0.69046	-0.58331

Age	0.00636	0.001767	3.6	Nonlin.	dev.	0.409	P=(0.522)
P1	0.120964	1.605776					

The p value of the continuous variable age is greater than the five percent (0.05) that means it is statistically insignificant, therefore, we accept the null hypothesis which actually fulfils the assumption of linearity.

6. There should be no significant outliers, high leverage points or highly influential points, which represent observations in data set that are in some way unusual. These can have a very negative effect on the binomial logistic regression equation that is used to predict the value of the dependent variable based on the independent variables. You can check for outliers, leverage points and influential points using box plot.

BOXPLOT



Here, there is no significant outliers, high leverage points or highly influential points, hence we can say that this assumption is also satisfied.

Now let's move on with binary logistic regression since it satisfies all the six assumptions of binomial logistic regression

BINOMIAL LOGISTIC REGRESSION

The following output will present the results needed to ascertain whether the independent variables statistically significantly predict the awareness of the elected representatives about knowledge of local government system activities they represent.

There is a direct relationship between the coefficients produced by logit and the odds ratios produced by logistic. First, let's define what is meant by a logit: A logit is defined as the log base e (log) of the odds.:

[1] logit(p) = log(odds) = log(p/q)

```
. logit q11 age q22

Iteration 0:   log likelihood =  -11689.56
Iteration 1:   log likelihood =      -9934
Iteration 2:   log likelihood = -9920.8096
Iteration 3:   log likelihood = -9920.7959
Iteration 4:   log likelihood = -9920.7959

Logistic regression                             Number of obs     =     17,388
                                                LR chi2(2)        =    3537.53
                                                Prob > chi2       =     0.0000
Log likelihood = -9920.7959                     Pseudo R2         =     0.1513
```

a. This is a listing of the log likelihoods at each iteration. (Remember that logistic regression uses maximum likelihood, which is an iterative procedure.) The first iteration (called iteration 0) is the log likelihood of the "null" or "empty" model; that is, a model with no predictors. At the next iteration, the predictor(s) are included in the model. At each iteration, the log likelihood increases because the goal is to maximize the log likelihood. When the difference between successive iterations is very small, the model is said to have "converged", the iterating is stopped and the results are displayed.

b. Log likelihood – This is the log likelihood of the final model. The value has no meaning in and of itself; rather, this number can be used to help compare nested models.

c. Number of obs – This is the number of observations that were used in the analysis.

d. LR chi2(2) – This is the likelihood ratio (LR) chi-square test. The likelihood chi-square test statistic can be calculated by hand as 2*(11689.56 – 9920.7959) = 3537.53. This is minus two (i.e., -2) times the difference between the starting and ending log likelihood. The number in the parenthesis indicates the number of degrees of freedom. In this model, there are two predictors, so there are two degrees of freedom.

e. Prob > chi2 – This is the probability of obtaining the chi-square statistic given that the null hypothesis is true. In other words, this is the probability of obtaining this chi-square statistic (3537.53) if there is in fact no effect of the independent variables, taken together, on the dependent variable. This is, of course, the p-value, which is compared to a critical value,

perhaps .05 or .01 to determine if the overall model is statistically significant. In this case, the model is statistically significant because the p-value is less than .0006.

f. Pseudo R2 – This is the pseudo-R-squared. Logistic regression does not have an equivalent to the R-squared that is found in OLS regression; however, many people have tried to come up with one. There are a wide variety of pseudo-R-square statistics. Because this statistic does not mean what R-square means in OLS regression (the proportion of variance explained by the predictors).

q11	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
age	0.00636	0.001767	3.6	0	0.002897	0.009822
q22	1.946452	0.034625	56.22	0	1.878589	2.014316
_cons	-0.93313	0.084205	-11.08	0	-1.09817	-0.76809

g. q_{11}– This is the dependent variable in our logistic regression. The variables listed below it are the independent variables.

h. Coef. – These are the values for the logistic regression equation for predicting the dependent variable from the independent variable. They are in log-odds units. Similar to OLS regression, the prediction equation is

$\log(p/1-p) = \beta_1+\beta_2 age+\beta_3 q22$

where p is the probability of q_{11}. Expressed in terms of the variables used in this example, the logistic regression equation is

$\log(p/1-p) = \text{Logit } P_i = -0.93313 - 0.00636age+1.946452q22$

These estimates tell you about the relationship between the independent variables and the dependent variable, where the dependent variable is on the logit scale. These estimates tell the amount of increase in the predicted log odds of $q_{11} = 1$ that would be predicted by a 1 unit increase in the predictor, holding all other predictors constant. Note: For the independent variables which are not significant, the coefficients are not significantly different from 0, which should be taken into account when interpreting the coefficients. (The columns with the z-values and p-values regarding tests whether the coefficients are statistically significant).

Because these coefficients are in log-odds units, they are often difficult to interpret, so they are often converted into odds ratios.

Age – The coefficient (or parameter estimate) for the variable age is 0.00636. This means that for a one-unit increase in age, we expect a 0.00636increase in the log-odds of the dependent variable q_{11}, holding all other independent variables constant. q_{22} – For every one-unit increase in q_{22}, we expect 1.946452 an increase in the log-odds of q_{11}, holding all other independent variables constant. constant – This is the expected value of the log-odds of q_{11} when all of the predictor variables equal zero. In most cases, this is not interesting. Also, oftentimes zero is not a realistic value for a variable to take.

i. Std. Err. – These are the standard errors associated with the coefficients. The standard error is used for testing whether the parameter is significantly different from 0; by dividing the parameter estimate by the standard error you obtain a z-value (see the column with z-values and p-values). The standard errors can also be used to form a confidence interval for the parameter, as shown in the last two columns of this table.

j. z and P>|z| – These columns provide the z-value and 2-tailed p-value used in testing the null hypothesis that the coefficient (parameter) is 0. If you use a 2-tailed test, then you would compare each p-value to your preselected value of alpha. Coefficients having p-values less than alpha are statistically significant. For example, if you chose alpha to be 0.05, coefficients having a p-value of 0.05 or less would be statistically significant. Here, the p value of both age and q_{22} are nearly zero. Hence, they are both statistically significant.

k. [95% Conf. Interval] – This shows a 95% confidence interval for the coefficient. This is very useful as it helps you understand how high and how low the actual population value of the parameter might be. The confidence intervals are related to the p-values such that the coefficient will not be statistically significant if the confidence interval includes 0.

Logistic regression

Logistic regression is in reality an ordinary regression using the logit as the response variable. The logit transformation allows for a linear relationship between the response variable and the coefficients:

[1] logit(p) = a + bX

or

[2] log(p/q) = a + bX

This means that the coefficients in a simple logistic regression are in terms of the log odds.

```
. logistic q11 age   q22

Logistic regression                             Number of obs     =     17,388
                                                LR chi2(2)        =    3537.53
                                                Prob > chi2       =     0.0000
Log likelihood = -9920.7959                     Pseudo R2         =     0.1513
```

q11	Odds Ratio	Std. Err.	z	P>z	[95% Conf.	Interval]
age	1.00638	0.001778	3.6	0	1.002901	1.009871
q22	7.003794	0.242506	56.22	0	6.544261	7.495596
_cons	0.39332	0.03312	-11.08	0	0.333481	0.463897

The exponentiated value of the coefficient of q 22 results in the odds ratio for q 11 to q 22 . ie, $e^{1.946452}$ =7.003794. which implies that the odds of being aware about the local

government activities is 7.003794. times that of awareness about grievance redressal mechanism.

Margins

The margins and prediction packages are a combined effort to port the functionality of Stata's (closed source) margins command to (open source) R. These tools provide ways of obtaining common quantities of interest from regression-type models. margins provides "marginal effects" summaries of models and prediction provides unit-specific and sample average predictions from models. Marginal effects are partial derivatives of the regression equation with respect to each variable in the model for each unit in the data; average marginal effects are simply the mean of these unit-specific partial derivatives over some sample. In ordinary least squares regression with no interactions or higher-order term, the estimated slope

coefficients are marginal effects. In other cases and for generalized linear models, the coefficients are not marginal effects at least not on the scale of the response variable. Margins therefore provides ways of calculating the marginal effects of variables to make these models more interpretable.

```
Predictive margins                              Number of obs     =     17,388
Model VCE    : OIM

Expression   : Pr(q11), predict()
```

		Delta-method				
	Margin	Std. Err.	z	P>z	[95% Conf.	Interval]
q22						
0	0.344746	0.005577	61.81	0	0.333815	0.355677
1	0.786259	0.004075	192.96	0	0.778273	0.794245

The values in the column headed Margin are the predicted probabilities for q 22 We also get standard errors z-statistics and p-values testing the difference from zero and a 95% confidence interval for each predicted probability.

CHAPTER VI
SUMMARY AND CONCLUSION

For our study we used a cross sectional data of the elected representatives. The Statistical software Stata was used in this study to run the logit and logistic regression. Here, one dependent variable which is a binary variable and two independent variables in which one is categorical variable and other is a continuous variable. Dependent variable as the awareness of the elected representatives about knowledge of local government activities they represent and independent variables as age and awareness about the grievance redressal mechanism are selected. Assumptions of the binary logistic regression is tested and satisfied. Henceforth, conducted the logit and logistic regression which gave the following results;

- The age and awareness about the grievance redressal mechanism statistically significantly predicted the awareness of the elected representatives about the knowledge of local government activities they represent by consulting the "P>|z|" column for the "age" & "q22" respectively. The "P>|z|" column contains the p-value for each coefficient and the constant (both expressed as odds ratios). You can see that age was statistically significant (i.e., p =0 .0006) and also, the awareness about the grievance redressal mechanism statistically significant (i.e., p = 0.000).
- The mean age of most of the respondents is 45.
- Even though we know that the knowledge increase with age, the tabs of each variables make it clear.
- The coefficient (or parameter estimate) for the variable age is 0.00636. This means that for a one-unit increase in age, we expect a 0.00636 increase in the log-odds of the dependent variable q11, holding all other independent variables constant.
- For every one-unit increase in q22, we expect1.946452 a increase in the log-odds of q11, holding all other independent variables constant.
- The exponentiated value of the coefficient of q22 results in the odds ratio for q11 to q22. i.e., $e^{1.946452}$=7.003794. which implies that the odds of being aware about the power is 7.003794. times that of awareness about policies and plans.
- The predicted probability of q22 was obtained with the help of margin even though we know that the knowledge increase with age.

- Most of the elected representatives have adequate knowledge about their local government activities, grievance redressal mechanism etc.