

DATA ANALYTICS PROJECT

ON

ELECTION PREDICTION

Under the guidance of –

Dr. Chidambaran G Iyer,

PES UNIVERSITY, Bangalore.

TEAM MEMBERS:

DEVARAJ V NAYAK -01FB15ECS092

KRISHNAKUMAR HEGDE – 01FB15ECS153

NIKHIL Y DIXIT – 01FB15ECS190

INTRODUCTION

The current Election system that is being followed only considers the number of votes gained by the candidates to decide the winner ignoring all other vital aspects. These vital aspects include various important factors like the representative's attendance in the Parliament, development factors like road infrastructure and funds utilised in his/her constituency, his/her experience in the political field i.e. no of terms the representative got elected. Also we also consider the participation of the representative in the Parliament such as number of questions asked during their tenure. If people ignore these factors while voting then there are chances of wrong candidates getting selected for high & important positions in Government offices and other organizations. Hence, we believe that votes given to a candidate just on the basis of the spurious claims made during their Election campaigns without knowing his or her achievements, previous records and eligibility are not justifiable.

The outcome of a person contesting for elections can either be a win or a loss. Since the dependent variable i.e. the outcome of the election is taking only two values, we have designed our model based on **Logistic Regression**. Winning an election is considered to be 1 and losing to be 0.

We have collected data for the state of **Karnataka** for our prediction.

OBJECTIVE

The main objective of this project is to predict whether a person gets elected in an election or not. This prediction is based on previous data for that person. Analysis with logistic regression enhances the accuracy of our prediction since the dependent variable takes up only two discrete values. Obtaining summary statistics for independent variables corresponding to the elected representatives to get an overview of their performance in their tenure.

DATA INFORMATION

The data has been collected from various forums available on government portals as well as other related websites. The data set contains five random variables namely, attendance, Percentage of MPLAD funds utilised, Road infrastructure, number of terms served including the present term, number of questions asked in the Parliament. The outcome of the election is a limited dependent variable as it takes values 0 and 1. We have collected the data for Members of Parliament of Karnataka state.

We have collected the data for three terms separately:

1)14TH Lok Sabha

2)15th Lok Sabha

3)16th Lok Sabha

Given below is the list of sources for the data set collected:-

<http://www.prsindia.org/mptrack/16loksabha>

<http://www.mplads.gov.in/MPLADS/UploadedFiles/HTML/15ls/lsstat10.htm>

www.data.gov.in

http://www.loksabha.nic.in/Members/statedetailar.aspx?state_name=Karnataka&lsno=13

www.karnataka.com

www.wikipedia.com

SUMMARY STATISTICS

(i) Comparison across different Lok Sabha elections

| | | <u>14th LOK SABHA</u> | | <u>15th LOK SABHA</u> | | <u>16th LOK SABHA</u> | | |
|--|----------------------------|-----------------------|-------------|-----------------------|-------------|-----------------------|-------------|--|
| | Variable | Mean | S.D. | Mean | S.D. | Mean | S.D. | |
| | ATTENDANCE | 57.75 | 16.59 | 71.75 | 15.64 | 83.04 | 15.96 | |
| | MPFUNDinLacs | 1952 | 58.25 | 1879 | 157.7 | 1359 | 332.4 | |
| | OFMPLADFUNDSTILISED | 97.6 | 2.912 | 93.93 | 7.886 | 67.94 | 16.62 | |
| | ROADINFRASTRUCTUREinLacs | 3985 | 2672 | 5631 | 3367 | 8078 | 4469 | |
| | NUMBEROFTERMS | 1.857 | 1.177 | 1.929 | 1.489 | 2.464 | 1.774 | |
| | NUMBEROFQUESTIONSASKEDINPA | 223.7 | 322.5 | 435.2 | 323.5 | 382.1 | 237.5 | |

Figure showing the summary statistics of random variables across different Lok Sabha terms.

- The average attendance of MPs has increased over every term which implies that MPs are becoming more regular to the Parliament.
- Percentage of MPLAD funds utilised has decreased by around 3.5% in 15th Lok Sabha as compared to 14th Lok Sabha. As the 16th Lok Sabha is not yet over, we can't really draw conclusions on funds utilised until this term gets over.
- Road infrastructure has increased substantially over the years which denotes that MPs are more concerned about development of roads in their constituencies.
- The average number of terms has gone up from 1.857 in 14th Lok Sabha to 2.464 in 16th Lok Sabha which means that the numbers of experienced members in 16th Lok Sabha are more. In other words, more experienced representatives are getting elected to the Parliament.

- Number of questions asked by representatives in Parliament has also increased. This means that MPs are actively taking part in debates in Parliament and are more bothered about raising issues in Parliament regarding their constituencies.

(ii) Summary Statistics by Lok Sabha term-wise

gretl: summary statistics

| | Mean | Median | S.D. | Min | Max |
|-------------------|-------|--------|-------|-------|-------|
| ATTENDANCE | 57.75 | 54.50 | 16.59 | 27.00 | 100.0 |
| MPFUNDinLacs | 1952 | 1989 | 58.25 | 1817 | 2000 |
| OFMPLADFUNDUTIL~ | 97.61 | 99.44 | 2.912 | 90.83 | 100.0 |
| ROADINFRASTRUCTU~ | 3985 | 3035 | 2672 | 1940 | 11650 |
| NUMBEROFTERMS | 1.857 | 1.000 | 1.177 | 1.000 | 5.000 |
| NUMBEROFQUESTION~ | 223.7 | 59.50 | 322.5 | 1.000 | 975.0 |

Figure showing Summary Statistics for 14th Lok Sabha

- Mean for attendance is 57.75% which implies that on an average MP's attend just over half the sessions held in the Parliament. Least attendance is 27% which corresponds to M H Ambareesh. K H Muniyappa has 100% attendance.
- Every MP is entitled Rs.20 Crore in their respective constituencies as MPLAD funds. On an average, an MP spends about Rs.1952 Lacs (Rs.19.52 Cr) for the constituency. Minimum utilisation is done by K Virupakshappa. 13 members have utilised their funds to the fullest.
- 97.61% is the average of the MPLAD funds utilised by MP's. This means that on an average, MPs utilise about 97.61% of funds released to their respective constituencies. Also, Standard Deviation is 2.912% which

implies almost uniform utilisation of funds by MP's. K Virupakshappa of Koppal constituency has least utilization of funds at 90.83%. 12 MPs have completely utilised all the funds released to their constituencies.

- An average of Rs.3985 Lacs is spent by the MPs on roads. Also, Rs.2672 Lacs is the standard deviation which depicts the non-uniformity of infrastructure spent on roads. Kunnur Manjunath Channappa has spent least money on roads (Rs.1940 Lacs) and R L Jallappa of Chikkaballapur has spent maximum funds on road development.
- Mean for no of terms doesn't exactly depict the average no of terms since it takes only integer values. 16 MPs are elected for the first time while K H Muniyappa has won maximum no of time among all MPs (5 terms).
- On an average, each MP asks about 223 questions in the Parliament sessions. Three MPs have asked only one question while Karunakar G Reddy of Bellary has asked maximum questions (975).

gretl: summary statistics

| | Mean | Median | S.D. | Min | Max |
|-------------------|-------|--------|-------|-------|-------|
| ATTENDANCE | 71.75 | 70.00 | 15.64 | 42.00 | 100.0 |
| MPFUNDinLacs | 1879 | 1952 | 157.7 | 1351 | 2000 |
| OFMPLADFUNDSUTIL~ | 93.93 | 97.59 | 7.886 | 67.57 | 100.0 |
| ROADINFRASTRUCTU~ | 5631 | 5105 | 3367 | 1950 | 15000 |
| NUMBEROFTERMS | 1.929 | 1.000 | 1.489 | 1.000 | 6.000 |
| NUMBEROFQUESTION~ | 435.2 | 338.0 | 323.5 | 5.000 | 1009 |

Figure showing Summary Statistics for 15th Lok Sabha

- Mean for attendance is 71.75% which implies that MPs have better attendance than that in 14th Lok Sabha. Least attendance is 42% which

corresponds to D V Sadananda Gowda. 4 members have 100% attendance.

- Every MP is entitled Rs.20 Crore in their respective constituencies as MPLAD funds. On an average, an MP spends about Rs.1879 Lacs (Rs.18.79 Cr) for the constituency. Minimum utilisation is done by K H Muniyappa of Kolar. 10 people have utilised the funds fully.
- 93.93% is the average of the MPLAD funds utilised by MP's. This means that on an average, MPs utilise about 93.93% of funds released to their respective constituencies and it is less than previous term. Also, Standard Deviation is 7.886% which is more than that of previous term. K H Muniyappa of Kolar constituency has least utilization of funds at 90.83%. 10 MPs have completely utilised all the funds released to their constituencies.
- An average of Rs.5631 Lacs is spent by the MPs on roads. Also, Rs.3367 Lacs is the standard deviation which depicts the non-uniformity of infrastructure spent on roads. N Cheluva Rayaswamy has spent least money on roads (Rs.1950 Lacs) and Ananth Kumar Hegde of Uttara Kannada has spent maximum funds on road development.(Rs.15000 Lacs)
- 17 MPs are elected for the first time while K H Muniyappa has won maximum no of time among all MPs (6 terms).
- On an average, each MP asks about 435 questions in the Parliament sessions. This is fairly larger than no of questions asked in 14th Lok Sabha. Janaardhan Swamy has asked least no of questions (9) while K H Muniyappa has asked maximum no of questions (1009).

gretl: summary statistics

| | Mean | Median | S.D. | Min | Max |
|-------------------|-------|--------|-------|--------|-------|
| ATTENDANCE | 83.04 | 89.50 | 15.96 | 36.00 | 100.0 |
| MPFUNDinLacs | 1359 | 1424 | 332.4 | 458.0 | 1990 |
| OFMPLADFUNDSUTIL~ | 67.94 | 71.19 | 16.62 | 22.90 | 99.50 |
| ROADINFRASTRUCTU~ | 8078 | 7300 | 4469 | 0.0000 | 16550 |
| NUMBEROFTERMS | 2.464 | 2.000 | 1.774 | 1.000 | 7.000 |
| NUMBEROFQUESTION~ | 382.1 | 415.0 | 237.5 | 0.0000 | 845.0 |

Figure showing Summary Statistics for 16th Lok Sabha

- Mean for attendance is 83.04% which implies that MPs have better attendance than that in 15th Lok Sabha. Least attendance is 36% which corresponds to Prakash Babbanna Hukkeri of Chikkodi . 2 members have 100% attendance.
- Every MP is entitled Rs.20 Crore in their respective constituencies as MPLAD funds. On an average, an MP spends about Rs.1359 Lacs (Rs.13.59 Cr) for the constituency. Minimum utilisation is done by Ramesh Jiganjinagi of Bijapur. He has spent only Rs.458 Lacs. D V Sadananda Gowda of Bangalore North has utilised maximum funds at Rs.1990 Lacs.
- 67.94% is the average of the MPLAD funds utilised by MP's. This means that on an average, MPs utilise about 67.94% of funds released to their respective constituencies and it is less than previous term. Also, Standard Deviation is 16.62% which is more than that of previous term. None of the MPs have utilised the funds released to them completely.
- An average of Rs.8078 Lacs is spent by the MPs on roads. Also, Rs.4469Lacs is the standard deviation. B V Nayak of Raichur has not spent any money on road infrastructure in his constituency and Ananth

Kumar Hegde of Uttara Kannada has spent maximum funds on road development.(Rs.16550 Lacs)

- 10 MPs are elected for the first time while K H Muniyappa has won maximum no of time among all MPs (7 terms). This tells us that number of new MPs elected in the 16th Lok Sabha is less than 15th Lok Sabha.
- On an average, each MP asks about 382 questions in the Parliament sessions. This is less than the no of questions asked in 15th Lok Sabha. Prakash Babbanna Hukkeri of Chikkodi has not asked even a single question in Parliament while Ananth Kumar of Bangalore has asked maximum no of questions (845) in the Parliament sessions.

ANALYSIS

Analysis for the above data is done based on Logistic Regression as the dependent variable takes up only two values.

Logistic regression is a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. The outcome is measured with a binary variable (in which there are only two possible outcomes).

We have used GRETl tool for analysis of our data. After analysing the data set we found out that only MP funds, Number of terms, Road infrastructure turn out to be the important factors for determining the outcome of elections. Practically speaking this abstraction of random variables makes sense. This is because railways don't have well connected network across Karnataka and hence infrastructure on roads becomes a vital factor for deciding the outcome of election in Karnataka. As the number of term increases, experience of the representative increases and therefore an experienced people's representative has more tendency to perform better during his tenure in his constituency. It is quite natural that more the funds utilised, more is the constituency developed. Also, the co-efficients for these variables have small p-values indicating the significance of these random variables.

gretl: model 1

File Edit Tests Save Graphs Analysis LaTeX

Model 1: Logit, using observations 1-28
Dependent variable: Yi
Standard errors based on Hessian

| | coefficient | std. error | z | p-value | |
|--------------------|-------------|--------------------|----------|---------|----|
| const | -21.2627 | 12.5296 | -1.697 | 0.0897 | * |
| MPFUNDinLacs | 0.00886588 | 0.00599861 | 1.478 | 0.1394 | |
| ROADINFRASTRUCTU~ | 0.000528917 | 0.000249825 | 2.117 | 0.0342 | ** |
| NUMBEROFTERMS | 1.45117 | 0.736142 | 1.971 | 0.0487 | ** |
| Mean dependent var | 0.607143 | S.D. dependent var | 0.497347 | | |
| McFadden R-squared | 0.402599 | Adjusted R-squared | 0.189382 | | |
| Log-likelihood | -11.20740 | Akaike criterion | 30.41479 | | |
| Schwarz criterion | 35.74361 | Hannan-Quinn | 32.04387 | | |

Figure showing the p-values for various RV's

The above figure shows the p-values for MP Fund, Road infrastructure and number of terms indicating that they have significant co-efficients. Therefore, applying Logistic Regression Model, we have

$$L_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

where L_i = limited dependent variable

$\beta_0, \beta_1, \beta_2, \beta_3$ are the co-efficients

X_1 = MP Funds ; X_2 = Road Infrastrucutre & X_3 = No of terms

Let P_i be the probability that a person wins in election. Therefore, $1-P_i$ represents the probability that a person loses in election.

We define, Odds Ratio = $P_i / (1-P_i)$

Also, $\log(\text{Odds Ratio}) = L_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$

(i) Analysis for 15th Lok Sabha

gretl: model 1

File Edit Tests Save Graphs Analysis LaTeX

Model 1: Logit, using observations 1-28
Dependent variable: Yi
Standard errors based on Hessian

| | coefficient | std. error | z | p-value | |
|--------------------|-------------|--------------------|----------|---------|----|
| const | -21.2627 | 12.5296 | -1.697 | 0.0897 | * |
| MPFUNDinLacs | 0.00886588 | 0.00599861 | 1.478 | 0.1394 | |
| ROADINFRASTRUCTU~ | 0.000528917 | 0.000249825 | 2.117 | 0.0342 | ** |
| NUMBEROFTERMS | 1.45117 | 0.736142 | 1.971 | 0.0487 | ** |
| Mean dependent var | 0.607143 | S.D. dependent var | 0.497347 | | |
| McFadden R-squared | 0.402599 | Adjusted R-squared | 0.189382 | | |
| Log-likelihood | -11.20740 | Akaike criterion | 30.41479 | | |
| Schwarz criterion | 35.74361 | Hannan-Quinn | 32.04387 | | |

Figure showing logit model for 15th Lok Sabha elections

Co-efficient values are as follows:

$$\beta_0 = -21.2627 \quad \beta_1 = 0.00886588 \quad \beta_2 = 0.000528917 \quad \beta_3 = 1.45117$$

The L_i and odds ratio values are shown below:

| SI No | NAME | Yi | Li | ODD'S RATIO |
|-------|-----------------------------------|----|--------------|-------------|
| 1 | Adagooru H Vishwanath | 0 | 2.09625863 | 8.135674333 |
| 2 | Anant Kumar Hegde | 1 | 6.773819568 | 874.6462929 |
| 3 | Ananth Kumar | 1 | 2.383540822 | 10.84322891 |
| 4 | D.B. Chandre Gowda | 0 | -0.002653172 | 0.997350345 |
| 5 | D.K. Suresh | 1 | 1.063845186 | 2.897490988 |
| 6 | D.V. Sadananda Gowda | 1 | 2.6871154 | 14.68924216 |
| 7 | Dhruva Narayana Rangaswamy | 1 | 0.73657211 | 2.088763178 |
| 8 | Gangasandra Siddappa Basavaraj | 0 | -2.877338536 | 0.056284362 |
| 9 | Gowdar Mallikarjunappa Siddeswara | 1 | 1.59225721 | 4.914830221 |
| 10 | H.D. Devegowda | 1 | 8.710060208 | 6063.607555 |
| 11 | J. Shantha | 0 | -0.7739392 | 0.461192755 |
| 12 | Janardhana Swamy | 0 | -0.37725145 | 0.685743622 |
| 13 | K.H. Muniyappa | 1 | 3.164519232 | 23.67735798 |
| 14 | Katti Ramesh Vishwanath | 0 | -0.90616845 | 0.404069474 |
| 15 | M. Veerappa Moily | 1 | -1.136474358 | 0.320948579 |
| 16 | Mallikarjun Kharge | 1 | 0.601245 | 1.824388751 |
| 17 | N Cheluvarya Swamy | 0 | -3.145082578 | 0.043063368 |
| 18 | N.Dharam Singh | 0 | -1.669427628 | 0.188354844 |
| 19 | Nalin Kumar Kateel | 1 | 2.2408877 | 9.40167345 |
| 20 | P. C. Mohan | 1 | -0.74749335 | 0.473552096 |
| 21 | P.C. Gaddigoudar | 1 | 2.205722268 | 9.076805087 |
| 22 | Pralhad Venkatesh Joshi | 1 | 3.37999667 | 29.37067331 |
| 23 | Raghavendra Yeddyurappa | 0 | 0.71231757 | 2.038710642 |
| 24 | Ramesh Chandappa Jigajinagi | 1 | 7.001027616 | 1097.760655 |
| 25 | S. Pakkhirappa | 0 | -1.209442262 | 0.298363642 |
| 26 | Shivaramagouda Shivanagouda | 0 | -1.600057068 | 0.201884996 |
| 27 | Suresh Channabasappa Angadi | 1 | 3.278962874 | 26.54822456 |
| 28 | Udasi Shivkumar Chanabasappa | 1 | -0.387228164 | 0.678936168 |

Odd's ratio greater than 1 implies that probability of MP getting elected in the next election is more, in other words odds are in favour of MP getting re-elected. By observing the above figure we can conclude that most of the odds ratio values are in accordance with the actual values.

For example,

Consider the odds ratio for D K Suresh that is 2.897490988 which is greater than 1. This means that D K Suresh has a higher probability of getting re-elected which is in fact correct as $Y_i = 1$.

There are a few data points which don't comply with their Odds ratio because of external factors which are not in the scope of our analysis.

For example,

Consider the odds ratio for D V Sadananda Gowda that is 0.49418093 which is less than 1. This means that he has a lower probability of getting re-elected which does not agree with the actual value.

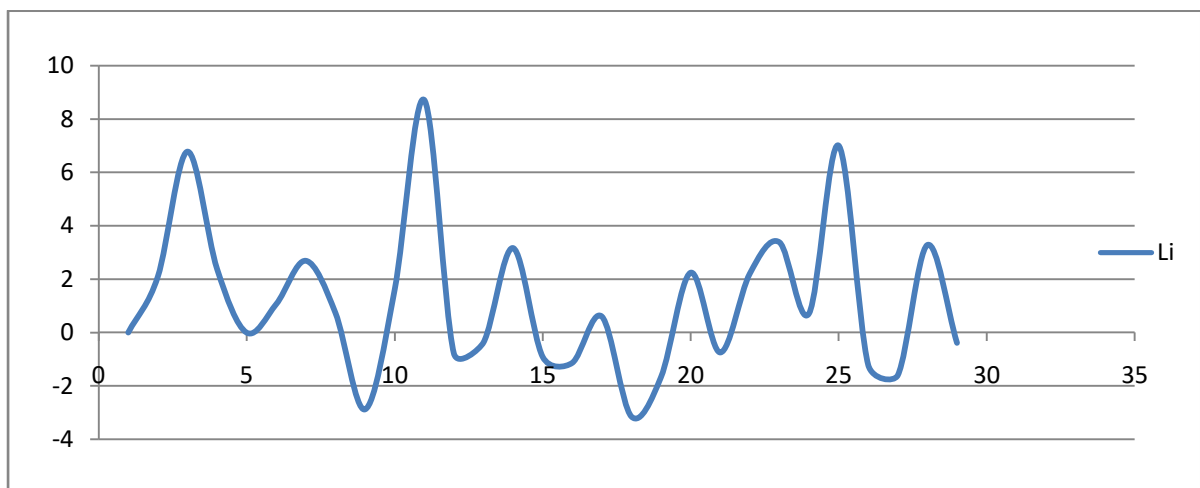


Figure showing Scatter Plot of Li VS Names of MPs

We are getting 9 peaks for the above scatter plot as 9 MPs have high odds ratio of getting re-elected. These MPs have performed better than others in their period.

(ii) Analysis for 14th Lok Sabha

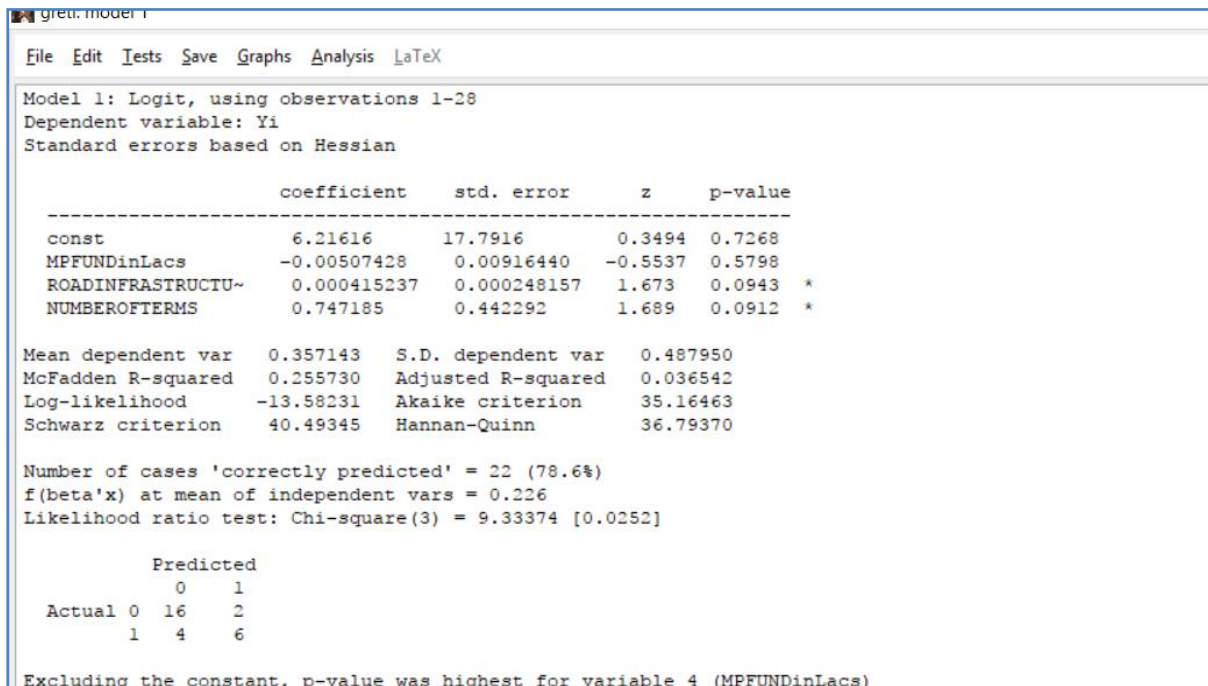


Figure showing logit model for 14th Lok Sabha elections

Co-efficient values are as follows:

$$\beta_0 = 6.21616 \quad \beta_1 = -0.00507428 \quad \beta_2 = 0.000415237 \quad \beta_3 = 0.747185$$

The L_i and odds ratio values are shown below:

| SI No | NAME | Yi | Li | ODD'S RATIO |
|-------|-----------------------------------|----|-------------|-------------|
| 1 | A. Venkatesh Naik | 0 | -0.60707643 | 0.544941718 |
| 2 | Anant Kumar Hegde | 1 | 2.926404548 | 18.6604171 |
| 3 | Ananth Kumar | 1 | 0.003336548 | 1.00334212 |
| 4 | Basangouda Patil | 0 | -1.43730883 | 0.23756623 |
| 5 | C.H. Vijayashankar | 0 | -1.53281334 | 0.215927335 |
| 6 | D.C. Srikantappa^ | 0 | -1.61793693 | 0.1983074 |
| 7 | D.V. Sadananda Gowda | 1 | -0.70485357 | 0.49418093 |
| 8 | Gowdar Mallikarjunappa Siddeswara | 1 | -1.59900966 | 0.202096563 |
| 9 | H.D. Devegowda | 1 | 2.93465358 | 18.81498412 |
| 10 | H.T. Sangliana^ | 0 | -1.8494984 | 0.157316056 |
| 11 | Iqbal Ahmed Saradgi | 0 | -2.36304574 | 0.094133081 |
| 12 | K. Virupakshappa | 0 | -1.17082348 | 0.310111466 |
| 13 | K.H. Muniyappa | 1 | 1.45616826 | 4.289491781 |
| 14 | Karunakara G. Reddy^ | 0 | -1.79513071 | 0.166105739 |
| 15 | Kunnur Manjunath Channappa^ | 0 | -2.37965522 | 0.092582493 |
| 16 | M. Shivanna | 0 | -2.20940805 | 0.109765605 |
| 17 | M.H. Ambareesh | 0 | -1.10096686 | 0.332549399 |
| 18 | Manorama Madhwaraj^ | 0 | -1.15041076 | 0.316506734 |
| 19 | N.Y. Hanumanthappa | 0 | -1.85230423 | 0.156875273 |
| 20 | Narsing Hulla Suryawanshi | 0 | -2.04746562 | 0.129061581 |
| 21 | P.C. Gaddigoudar | 1 | -0.7193555 | 0.487066071 |
| 22 | Pralhad Venkatesh Joshi | 1 | -0.70541125 | 0.493905414 |
| 23 | R.L. Jalappa | 0 | 2.318041586 | 10.15576562 |
| 24 | Ramesh Chandappa Jigajinagi | 1 | 0.603380694 | 1.828289251 |
| 25 | S. Bangarappa^ | 0 | 0.102988736 | 1.108478923 |
| 26 | S. Mallikarjunaiiah | 0 | -0.68008332 | 0.506574785 |
| 27 | Suresh Chanabasappa Angadi | 1 | 0.3788296 | 1.460574133 |
| 28 | Tejasvini Gowda | 0 | -1.33832236 | 0.262285319 |

Odd's ratio greater than 1 implies that probability of MP getting elected in the next election is more, in other words odds are in favour of MP getting re-elected. By observing the above figure we can conclude that most of the odds ratio values are in accordance with the actual values.

Consider the odds ratio for H D Devegowda that is 18.81498412 which is greater than 1. This means that H D Devegowda has a higher probability of getting re-elected which is actually true as $Y_i = 1$.

As mentioned in the 15th Lok Sabha analysis, there exist some members who have defying odds ratio. This can be due to various reasons like MP might not have contested for the next election.

Consider the odds ratio for D V Sadananda Gowda that is 0.49418093 which is less than 1. This means that he has a lower probability of getting re-elected which does not agree with the actual value.

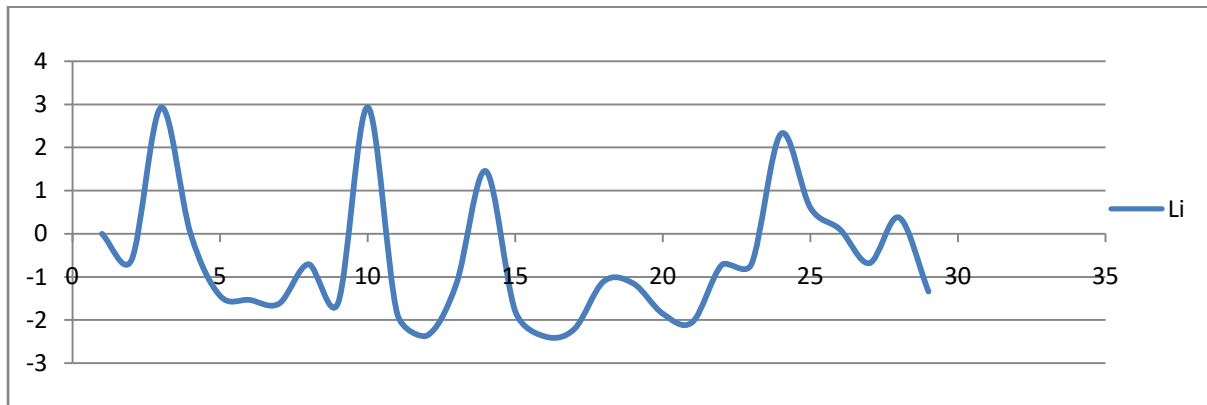


Figure showing Scatter Plot of Li VS Names of MPs for

We are getting 6 peaks for the above scatter plot as 6 MPs have high odds ratio of getting re-elected. These MPs have performed better than others in their period.

(iii) Analysis for 16th Lok Sabha

As we don't have Y_i values for 16th Lok Sabha, we use the co-efficients of 15th Lok Sabha and plug in the values of 16th Lok Sabha in our Logistic regression to get the odds ratio for each MP. Using odds ratio, we then predict the Y_i value for all the members of 16th Lok Sabha. An MP having Y_i as 1 is likely to get re-elected in the upcoming elections while $Y_i=0$ for an MP indicates that he/she is likely to lose in the upcoming elections.

Co-efficient values are as follows:

$$\beta_0 = -21.2627 \quad \beta_1 = 0.00886588 \quad \beta_2 = 0.000528917 \quad \beta_3 = 1.45117$$

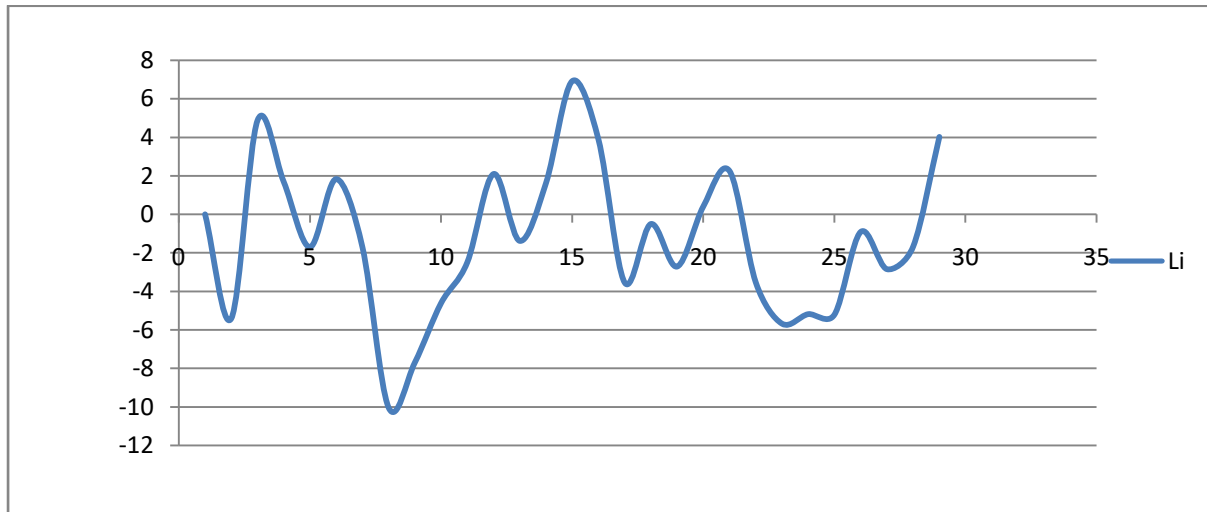
The L_i and odds ratio values are shown below:

| SL NO | NAME | LI | ODDS RATIO | PREDICTED YI |
|-------|-------------------------------|--------------|-------------|--------------|
| 1 | Ananth Kumar | -5.40723428 | 0.004484025 | 0 |
| 2 | Ananthkumar Hegde | 4.874300966 | 130.8826298 | 1 |
| 3 | Angadi Suresh Channabasappa | 1.685341284 | 5.394291604 | 1 |
| 4 | B. N. Chandrappa | -1.699868322 | 0.182707581 | 0 |
| 5 | B. S. Yeddyurappa | 1.831194314 | 6.241336324 | 1 |
| 6 | B. Sriramulu | -1.705539734 | 0.181674304 | 0 |
| 7 | B. V. Nayak | -10.03681741 | 4.37588E-05 | 1 |
| 8 | Bhagwanth Khuba | -7.692713284 | 0.000456139 | 0 |
| 9 | C. S. Puttaraju | -4.608235724 | 0.009969392 | 0 |
| 10 | D. K. Suresh | -2.490441278 | 0.082873388 | 0 |
| 11 | D. V. Sadananda Gowda | 2.0926337 | 8.10623647 | 1 |
| 12 | G.M. Siddeswara | -1.381011206 | 0.251324284 | 0 |
| 13 | Gaddigoudar P. C. | 1.621845076 | 5.062422259 | 1 |
| 14 | H. D. Devegowda | 6.91442956 | 1006.696604 | 1 |
| 15 | K. H. Muniyappa | 3.885535374 | 48.69300445 | 1 |
| 16 | Karadi Sanganna Amarappa | -3.54276392 | 0.028933247 | 0 |
| 17 | M. Veerappa Moily | -0.502621252 | 0.604942872 | 0 |
| 18 | Malikarjun Kharge | -2.71235731 | 0.066380144 | 0 |
| 19 | Muddahanumegowda | 0.413575915 | 1.512215683 | 1 |
| 20 | Nalin Kumar Kateel | 2.243545332 | 9.426692869 | 1 |
| 21 | P. C. Mohan | -3.537760342 | 0.02907838 | 0 |
| 22 | Prakash Babanna Hukkeri | -5.677571216 | 0.003421859 | 0 |
| 23 | Pralhad Joshi | -5.182528126 | 0.005613796 | 0 |
| 24 | Pratap Simha | -5.206265162 | 0.00548211 | 0 |
| 25 | R. Dhruvanarayana | -0.91358414 | 0.401084103 | 1 |
| 26 | Ramesh Jigajinagi | -2.84880501 | 0.057913486 | 0 |
| 27 | Shivkumar Channabasappa Udasi | -1.675256322 | 0.187260174 | 1 |
| 28 | Shobha Karandlaje | 4.018274035 | 55.60505056 | 1 |

From the above figure, we can thus predict that the following MPs have a higher chances of getting re-elected in the next elections.

- 1) Ananthkumar Hegde
- 2) Angadi Suresh Cannabasappa
- 3) B S Yeddyurappa
- 4) B V Nayak
- 5) D V Sadananda Gowda
- 6) Gaddigoudar P C
- 7) H D Devegowda
- 8) K H Muniyappa
- 9) Muddahanumegowda
- 10) Nalin Kumar Kateel
- 11) R Dhruvanarayana
- 12) Shivkumar Channabasappa Udasi

13) Shobha Karandlaje



Plot showing the relation between Li with corresponding MPs where crests indicate higher odds ratio and trough indicate lower odds ratio

Conclusion

We have calculated summery statistics for the data collected. After applying Logistic Regression Model we were able to predict the outcome of next elections of Lok Sabha for Karnataka state. Odds ratio for each MP was also calculated for 14th and 15th Lok Sabha. We analysed our prediction with actual data by plotting scatter plots of Li with Names of MPs.

In this way, our model predicts the outcome for each constituency by considering vital factors and past performance of MPs. This kind of analysis is more accurate than considering only the number of votes secured by the candidate in the previous election.

Finally, we predicted the candidates who have higher chances of winning the upcoming Lok Sabha elections by using Logistic Regression model.

