

# **CSP301: Assignment 2**

## **State of the Indian Lok Sabha**

Submitted by:

Shashank Jain  
2011CS10254

Nipun Gupta  
2011CS50289

Abhishek Bansal  
2011CS50271

# Contents

<b>1</b>	<b>Overview</b>	<b>2</b>
1.1	Objective . . . . .	2
1.2	Database . . . . .	2
1.3	Implementation . . . . .	3
1.4	Website . . . . .	3
<b>2</b>	<b>Trends Observed</b>	<b>4</b>
2.1	Political Parties . . . . .	4
2.2	Alliance . . . . .	4
2.3	States . . . . .	5
2.4	Age . . . . .	5
2.5	Education . . . . .	5
2.6	Attendance . . . . .	5
2.7	Gender . . . . .	6
2.8	Debates Participated . . . . .	6
2.9	Questions Raised . . . . .	7
2.10	Private Member Bills passed . . . . .	7
2.11	MP Rating . . . . .	7
<b>3</b>	<b>Hypotheses</b>	<b>9</b>
3.1	Introduction to Welch t-test . . . . .	9
3.2	Welch t-test procedure . . . . .	9
3.3	List of Hypotheses . . . . .	10

# Chapter 1

## Overview

### 1.1 Objective

Use the MP Track data from PRS website on the 15th Lok Sabha to show a series of visualizations that capture the state of the Indian Lok Sabha.

Our main idea is to allow exploration and mashup of different parameters together. The main goal in the statistics part is to mathematically describe any interesting trends that have been spotted in the visualizations, and to verify them using Welch t-test.

### 1.2 Database

Our main database of the 15th Lok Sabha has been taken from the PRS website. The data is in Excel and contains several things such as the political party to which the MPs belong, their educational qualifications, age, and most interestingly their activity in parliament – attendance, debates in which they participated, questions they raised in parliament, etc.

In order to obtain interesting visualizations, we have also taken data about 14th Indian Lok Sabha and Rajya Sabha from the same PRS website, and then compared it with our main database. To make things more interesting, some worldwide data has also been collected from the World Bank website and visualized.

## 1.3 Implementation

We have used Google Developers Tools which includes Google Charts to make various data visualizations (pie charts, Scatter plots, histograms, column charts, geomaps, treemaps, etc.), and Google Query Language to read data from Google Spreadsheets. Apart from this, we have used d3 library of JavaScript to create various mashups. The website which contains all these visualizations has been coded using HTML, JavaScript, and CSS.

Various hypotheses were constructed by observing the trends in our visualizations, and were then tested statistically, using Welch t-test.

## 1.4 Website

Our website offers the user to see and explore our original data in every possible way.

There is a Search Engine where the entire database table is stored and the user can search the Member of Parliament(MP) on various parameters like name, state-wise, political party-wise, gender and age. The table gives the apt result and displays the entire information about the targeted MPs. In addition to this, table also shows the rating of MPs<sup>1</sup>.

All the hypotheses are also shown. For every hypothesis, there is a corresponding visualization which shows a trend. Below it, it has been statistically proved via Welch t-test and the test results are shown.

Data Visualization has been done extensively and systematically. Visualizations have been categorized into nation-wide, state-wise, and party-wise data. Pie charts, column charts, India maps, histograms, tables, and tree maps have been used to visualize the data. Also, comparison between 14th Lok Sabha, 15th Lok Sabha, and Rajya Sabha has also been visualised via pie charts and tables. And comparison with the worldwide data of sex ratios in the parliament of different countries has also been visualized via the world map.

---

<sup>1</sup>See MP rating trends, 2.10

# Chapter 2

## Trends Observed

### 2.1 Political Parties

Given data provides the political party of each MP. Party with the highest number of seats is ‘Indian National Congress (207 seats)’, which is the current ruling party followed by ‘Bhartiya Janta Party (117 seats)’, which forms the opposition. Overall there are 39 different political parties which have at least 1 member in 15th Indian Lok sabha.

On comparison with 14th Lok Sabha, we found that Indian National Congress not only remained the party in power but also grew in the number of elected MPs from a count of 153 to 207. It is also the leading party in Rajya Sabha with 88 seats.

### 2.2 Alliance

Alliance is a group of parties that come together to form the government. Currently we are having three major alliances namely ‘UPA’, ‘NDA’, and ‘Third Front’. Apart from these, we have some independent MPs and non aligned parties too. UPA currently forms the government and is led by Indian National Congress.

Visualisation on India map showed that UPA is very strong in Northern and South-eastern states of India while NDA is strong in Central and South-western states of India.

## 2.3 States

We have 28 states and 7 union territories in India, and each of them have different number of representatives in Lok Sabha depending on the population and issues in that region. The highest representation is from Uttar Pradesh(81 MPs), which is the highest populated state of India, followed by states like Maharashtra, Andhra Pradesh, Tamil Nadu, West Bengal and Bihar.

## 2.4 Age

Most of the MPs are from 50 to 60 years of age while average age of MPs is around 56 years. Age of MPs ranges from 30 years(Hamdullah Sayeed from Lakshadweep) to 91 years(Ram Sundar Das from Bihar). 11.8% of the MPs(65 out of 552) are above 70 years of age while 8.7% (around 48 MPs) are below 40 years of age.

## 2.5 Education

Education qualification of MPs provides us with interesting facts. Only 26 out of 552 MPs (around 4.7%) are doctorate. 27% MPs are Post Graduate while 46.6% are Graduate. 21.7% of the MPs are below graduate(i.e around 120 MP's), out of which 19 are Under-Matric.

In order to mathematically compare education of MPs among different states and Political Parties, we have defined an *Education Quotient*. It is calculated by taking the mean of all the education values of MPs involved in a particular state or a party.

The education values are assigned according to education level of the MPs Under Matric are given 10 points, under graduates are given 20, those with diplomas are given 45, graduates are given 60 points, post graduates a healthy 75 points and those holding Ph.D. have 90 points to their name.

## 2.6 Attendance

One of the most important attribute of an MP is his attendance in the Lok Sabha, which is also the reason for it having the highest weight in our MP

rating. Though the average attendance<sup>2</sup> is 77%, which looks reasonably good, but the individual attendance of some MPs is really disappointing. Almost 53% MPs have attendance greater than 80% with 4 MPs having 100% attendance. While on the other hand 26% of the MPs have attendance less than 70% with 27 MPs having less than 50% attendance. *Baliram Kashyap* from Chhattisgarh has the lowest attendance of only 1%.

Visualisation on India map shows that MPs from South Indian states have less attendance than those who are from North Indian states. This may be because of the fact that MPs from South India repeatedly travel to their states during parliament sessions.

A very interesting point to note is that the average attendance of Lok Sabha jumped from 70% of 14th Lok Sabha to 77% of 15th Lok Sabha, almost a jump of 7%. While the average attendance of Rajya Sabha is also 77%.

## 2.7 Gender

The most disappointing data was that of Gender ratio in Indian Lok Sabha. Only 11% of the MPs, which is around 61 out of 552 MPs are female. Another disappointing fact was that the percentage of female MPs from South Indian states was very less despite of high sex ratio and high education status of people in these states.

On comparison with the world wide data on *Percentage females in Law Making Body of each Country*, we found that India is much behind from a very large number of countries in this field, the world average is at 20% and the highest percentage of females being 56.3% in Rwanda.

## 2.8 Debates Participated

Another important attribute of an MP is the number of debates in which he participated. It reflects the attentiveness and activity rate of the MP in the parliament. Almost 37.4% of the MPs have less than 10 debates with 32 MPs having involved in no debates. Only 22 MPs have greater than 100 debates among which *Arjun Ram Meghwal* of Bikaner has the highest number of debates at 323.

---

<sup>2</sup>Data of attendance, debates, questions raised etc. of around 60 MPs was not available.

The average debates in which an MP participated fell from 30 in the 14th Lok Sabha to 27 in the current one. However the average number of debates is just 18 in the Rajya Sabha.

## 2.9 Questions Raised

Similar to debates, questions raised by the MPs is also important. It reflects the attention of an MP towards his constituency's woes and also his awareness towards the functioning of the government. The average questions raised by an MP is 207 with 20% of the MPs raising more than 400 questions. On the other hand, 76 MPs have raised less than 10 questions in comparison to 886 questions raised by *Anandrao Adsul* from Maharashtra.

The average questions asked by an MP jumped from 179 in 14th Lok Sabha to 207 in 15th Lok Sabha. In Rajya Sabha the average number of questions raised was 141.

## 2.10 Private Member Bills passed

In the Indian Parliament, an MP can propose a bill even if he is not in the executive government. A total of 288 bills were proposed in the Lok Sabha taking the average to be about 0.58 bills per MP. In the previous Lok Sabha too about 300 bills were proposed.

However it is not necessary that these bills get passed every time they are introduced. Of the 300 odd Private Members' Bills introduced in the 14th Lok Sabha, barely 4% were discussed; 96% lapsed without even a single debate in the House. Till date, Parliament has passed 14 Private Members' Bills. Six of these were passed in 1956 alone and The last Private Members' Bill passed by Parliament was 'The Supreme Court (Enlargement of Criminal Appellate Jurisdiction) Bill, 1968' that became an Act on 9 August 1970. No Private Member's Bill has been passed by Parliament since 1970<sup>3</sup>.

## 2.11 MP Rating

To effectively visualize the MP activity in the parliament we defined an *MP Rating* of each member of parliament on the basis of the education, atten-

---

<sup>3</sup>data collected from Wikipedia.com



dance, debates participated, questions raised and bills passed. The rating is generated by taking the weighted mean of each individual rating.

$$MPRating = \frac{(0.8 * edur) + (1.2 * depr) + (1.4 * billr) + (1.1 * quesr) + (1.5 * attenr)}{6.0}$$

where,

*edur* - Education Rating

*debr* - Debate Rating

*billr* - Bill Rating

*quesr* - Question Rating

*attenr* - Attendance Rating

All the individual ratings are calculated based on the deviation of individual fields from their mean value. Each weight is decided by their relevance in the activity of the MP. So an MP who is not highly educated but has a high attendance and more debates is rated more than the MP who is highly educated but neglects his parliamentary duties and skips the assembly sessions.

The lowest rated MP is *Baliram Kashyap* of Chhattisgarh at just 0.47. He is among the 23 MPs who were rated below 1. The highest rated MP is *P.L. Punia* of Uttar Pradesh at a solid 5 out of 5 rating points. He and 15 others were the only MPs from about 500 MPs to be rated above 4. The average MP rating came out to be nearly 2.5 with MPs divided equally on both sides of the mean.

# Chapter 3

## Hypotheses

### 3.1 Introduction to Welch t-test

Based on the trends observed in the visualizations, various hypotheses were constructed and tested statistically using Welch t-tests.

The t-test assesses whether the means of two groups are statistically different from each other. Welch t-test is intended for use with two samples having possibly unequal variances.

Difference between the Welch t-test and other t-tests :

- It does not require the assumption of equal population variances.
- It is an approximate method and not exact.
- This procedure has different standard errors and degrees of freedom than pooled-variance t procedures.

Welch t-test assumptions :

- Samples are independent and random
- Population is normally distributed

### 3.2 Welch t-test procedure

For doing a Welch t-test on the data samples we denote:

$s_1$  = standard deviation of sample-1

$s_2$  = standard deviation of sample-2

$n_1$  = size of sample-1

$n_2$  = size of sample-2

$\bar{X}_1$  = mean of sample-1

$\bar{X}_2$  = mean of sample-2

Null Hypothesis,  $H_0 : \bar{X}_1 = \bar{X}_2$

Alternate Hypothesis,  $H_A : \bar{X}_1 > \bar{X}_2$

t = T-value

$\nu$  = degrees of freedom

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$
$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}}$$

After getting the value of t and  $\nu$ , we check the t-table for  $t_{critical}$  at a confidence level of 95% (  $\alpha(1 \text{ tail}) = 0.05$ ), and verify our results.

If  $|t| > t_{critical}$ , then our null hypothesis  $H_0$  can be rejected with 95% confidence level, otherwise,  $H_0$  cannot be rejected. If  $t < 0$ , then our alternate hypothesis  $H_A$  can also be rejected.

In our hypothesis testing,  $H_0$  and  $H_A$  are taking complementary to the proposed hypothesis. So, when  $H_0$  and  $H_A$  are rejected, then we can say our proposed hypothesis can be accepted.

### 3.3 List of Hypotheses

All the hypotheses and the corresponding t-values and results are listed below.

**Hypothesis 1: More experienced MPs maintain better attendance.**

$$t = -3.1814299415896183$$

$$t_{critical} = 1.6525$$

Hypothesis-1 can be accepted with 95% confidence level.

**Hypothesis 2: MPs from small states maintain better attendance.**

$$t = -0.025563012603867976$$

$$t_{critical} = 1.6525$$

Hypothesis-2 cannot be accepted.

**Hypothesis 3: MPs from UPA(in power) or NDA(opposition) maintain better attendance.**

$$t = -1.7948203853123752$$

$$t_{critical} = 1.6607$$

Hypothesis-3 can be accepted with 95% confidence level.

**Hypothesis 4: MPs from North India maintain better attendance than MPs from South India.**

$$t = -3.862235634397904$$

$$t_{critical} = 1.6525$$

Hypothesis-4 can be accepted with 95% confidence level.

**Hypothesis 5: More educated MPs maintain better attendance.**

$$t = -1.9812357265969112$$

$$t_{critical} = 1.6538$$

Hypothesis-5 can be accepted with 95% confidence level.

**Hypothesis 6: Female MPs maintain better attendance.**

$$t = -0.5837488550262022$$

$$t_{critical} = 1.6683$$

Hypothesis-6 cannot be accepted.

**Hypothesis 7: MPs with greater number of debates maintain better attendance.**

$$t = -10.720810077053816$$

$$t_{critical} = 1.6525$$

Hypothesis-7 can be accepted with 95% confidence level.

**Hypothesis 8: MPs with greater number of questions maintain better attendance.**

$$t = -3.7747072708036096$$
$$t_{critical} = 1.6525$$

Hypothesis-8 can be accepted with 95% confidence level.

**Hypothesis 9: Female MPs are better educated than male MPs.**

$$t = 0.26172874338878926$$
$$t_{critical} = 1.6694$$

Hypothesis-9 cannot be accepted.

**Hypothesis 10: MPs from small states are better educated.**

$$t = -1.293594585666722$$
$$t_{critical} = 1.6525$$

Hypothesis-10 cannot be accepted.

**Hypothesis 11: MPs from UPA(in power) or NDA(opposition) are better educated.**

$$t = 2.264719710973411$$
$$t_{critical} = 1.6583$$

Hypothesis-11 cannot be accepted.

**Hypothesis 12: MPs with greater number of questions are better educated.**

$$t = -4.807716498159002$$
$$t_{critical} = 1.6525$$

Hypothesis-12 can be accepted with 95% confidence level.

**Hypothesis 13: MPs with greater number of debates are better educated.**

$$t = -1.2843810098925774$$

$$t_{critical} = 1.6525$$

Hypothesis-13 cannot be accepted.

**Hypothesis 14: MPs from small states participate more in debates.**

$$t = 3.154825210155718$$

$$t_{critical} = 1.6525$$

Hypothesis-14 cannot be accepted.

**Hypothesis 15: MPs from UPA/NDA participate more in debates.**

$$t = 0.0975442247030194$$

$$t_{critical} = 1.6601$$

Hypothesis-15 cannot be accepted.

**Hypothesis 16: MPs with greater number of questions participate more in debates.**

$$t = -5.726582316694883$$

$$t_{critical} = 1.6525$$

Hypothesis-16 can be accepted with 95% confidence level.

# Bibliography

- “PRS Legislative Research”, <http://www.prsindia.org/>
- “Wikipedia”, [http://en.wikipedia.org/wiki/15th\\_Lok\\_Sabha](http://en.wikipedia.org/wiki/15th_Lok_Sabha)
- “The World Bank”, <http://data.worldbank.org/>
- “W3 Schools”, <http://www.w3schools.com/>
- “Google Developers ”, <https://developers.google.com/chart/>
- “Scott Murray, code artist”, <http://alignedleft.com/tutorials/d3/making-a-scatterplot/>
- “Wikipedia”, [http://en.wikipedia.org/wiki/Welch%27s\\_t\\_test](http://en.wikipedia.org/wiki/Welch%27s_t_test)
- “Chapter 8, Part III : Probability and the Foundations of Inferential Statistics”, [http://www.sagepub.com/upm-data/40007\\_Chapter8.pdf](http://www.sagepub.com/upm-data/40007_Chapter8.pdf)
- “Khan Academy”, <http://www.khanacademy.org/math/statistics/>
- “t Test : table of critical t values”, [http://www.stattools.net/tTest\\_Tab.php](http://www.stattools.net/tTest_Tab.php)
- “Github”, [https://github.com/just-abhishek/CSP301\\_ass2.git](https://github.com/just-abhishek/CSP301_ass2.git)