

Declaration on Plagiarism

This form must be filled in and completed by the student(s) submitting an assignment

Name:	Ajit Kumar
Student Number:	19210438
Programme:	Data Analytics
Module Code:	CA682
Assignment Title:	Data Visualisation
Submission Date:	13 Dec 2019
Module Coordinator:	Dr Suzanne Little

I declare that this material, which I now submit for assessment, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work. I understand that plagiarism, collusion, and copying are grave and serious offences in the university and accept the penalties that would be imposed should I engage in plagiarism, collusion or copying. I have read and understood the Assignment Regulations. I have identified and included the source of all facts, ideas, opinions, and viewpoints of others in the assignment references. Direct quotations from books, journal articles, internet sources, module text, or any other source whatsoever are acknowledged and the source cited are identified in the assignment references. This assignment, or any part of it, has not been previously submitted by me or any other person for assessment on this or any other course of study.

I have read and understood the referencing guidelines found at <http://www.dcu.ie/info/regulations/plagiarism.shtml>, <https://www4.dcu.ie/students/az/plagiarism> and/or recommended in the assignment guidelines

Name: Ajit Kumar

Date: 13th December 2019

A walk-through Indian Election

Abstract

My topic for data visualization assignment is on 'A walk through Indian Election'. I have very keen interest in democracy and Indian democracy is one of the largest in the world that is why I have chosen this topic. Election in India is mainly divided into two categories Rajya Sabha and Lok Sabha and I will be focusing on Lok Sabha Elections. Lok Sabha elections take place once in 5 years to elect 543 members for the Lok Sabha (Lower house). Candidates winning the Lok Sabha elections are appointed 'Member of Parliament' and hold their seats for five years or until the president dissolves the body on the advice of the minister's council. My purpose for this visualization was to identify the number of different states won parties like the National Democratic Alliance (NDA), United Progressive Alliance (UPA) in order to form the central government in the last 6 Lok Sabha elections. I have also tried to analyse the relationship between the candidate's educational qualification and their chances of winning the Lok Sabha elections. I have collected data on the last 6 Lok Sabha elections from various sources. Dataset was short so I have merged various datasets so that different charts can be plotted.

Dataset

- ❖ I have used the below mentioned link to fetch the data for candidates of different party winning the election from their respective area.

Lok Sabha Election 2019	http://www.indiavotes.com/pc/info?eid=17&state=0
Lok Sabha Election 2014	http://www.indiavotes.com/pc/info?eid=16&state=0
Lok Sabha Election 2009	http://www.indiavotes.com/pc/info?eid=15&state=0
Lok Sabha Election 2004	http://www.indiavotes.com/pc/info?eid=14&state=0
Lok Sabha Election 1999	http://www.indiavotes.com/pc/info?eid=13&state=0
Lok Sabha Election 1998	http://www.indiavotes.com/pc/info?eid=12&state=0
Lok Sabha Election 1996	http://www.indiavotes.com/pc/info?eid=11&state=0
Educational Qualification	http://myneta.info/LokSabha2019/

- ❖ Dataset of Lok Sabha election 2014 & 2009 was directly available in CSV format whereas for years 2019,2004,1999,1998 and 1996 I have manually extracted the data from the website and pasted in excel.
- ❖ Data of educational qualification is extracted manually by opening manually each constituency and extracting the details of each candidates.
- ❖ The aspect of big data 'Variety and Veracity' is present in the data set.

- ❖ Last 6 Lok Sabha election dataset are present in excel format which comprises of 11 attributes named (PC Name, No, Type, State, Winning Candidate, Party Electors, Votes, Turnout, Margin, Margin %) and 543 rows each and all 6 excel files are merged.
- ❖ Data set educational qualification are present in excel format which comprises of 12 attributes (Age, Candidate, Criminal_case, District, Education, Party, State, Winner, Year, Liabilities, Assest) and 7930 rows.

Cleaning

- ❖ Data cleaning was primarily done in python.
- ❖ Unwanted text like [2000 Onwards], [2014 Onwards], [1977 Onwards], [1977 Onwards], [1947 - 1999] are removed from state column of all 6 csv of Lok Sabha Election.
- ❖ Name of the party is mapped based on the alliance of which they were part of either NDA, UPA or Other year wise.
- ❖ Some of the states like Uttarakhand which is formed from Uttar Pradesh in the year 2000, Jharkhand which is formed from Bihar in the year 2000, Chhattisgarh which formed from Madhya Pradesh in the year 2000 and Telangana formed from Andhra Pradesh their constituency was missing in the file they were mapped accordingly in all different Lok Sabha Elections CSV.

Processing

All different 6 Lok Sabha CSV are processed and each of them is transformed in one CSV each for Lok Sabha Election which comprises of 7 Columns State, Seats, NDA, UPA, Other, Result and year.

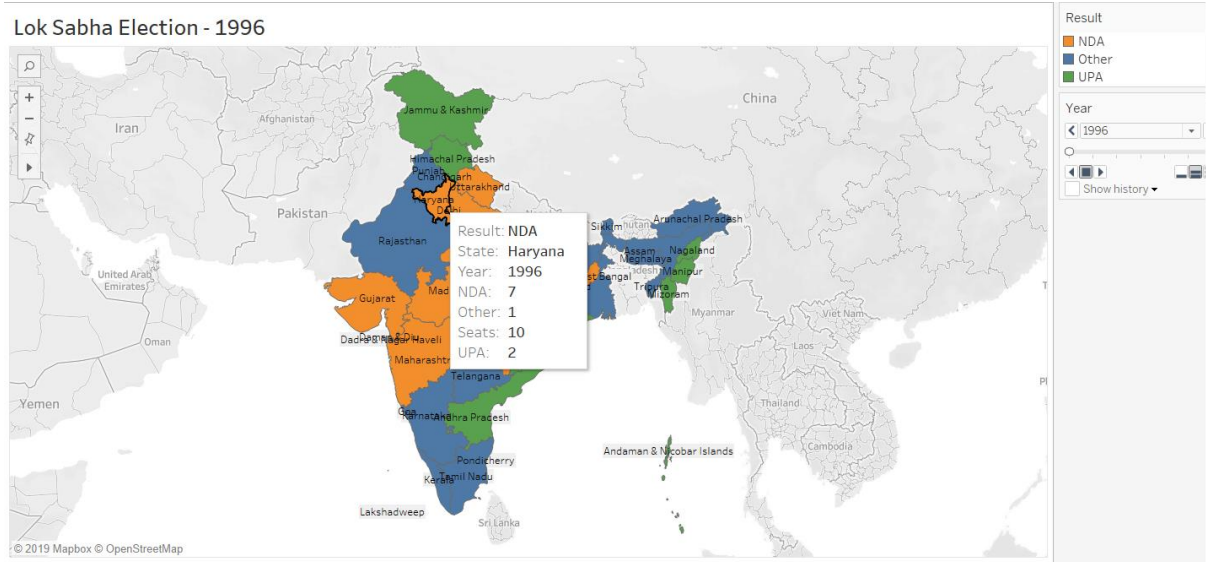
- State column have the name of all different states present in India.
- Seats consists of total number of seats present in each state.
- NDA consists of cumulative sum of total seats won by NDA alliance state wise.
- UPA consists of cumulative sum of total seats won by UPA alliance state wise.
- Other consists of cumulative sum of total seats won by Other Parties state wise.

After all, the different CSV files are combined to form single file. Educational qualification CSV are integrated with Lok Sabha Election 2019 CSV files in order to find the relation between qualification of candidates and winning Lok Sabha election for the year 2019.

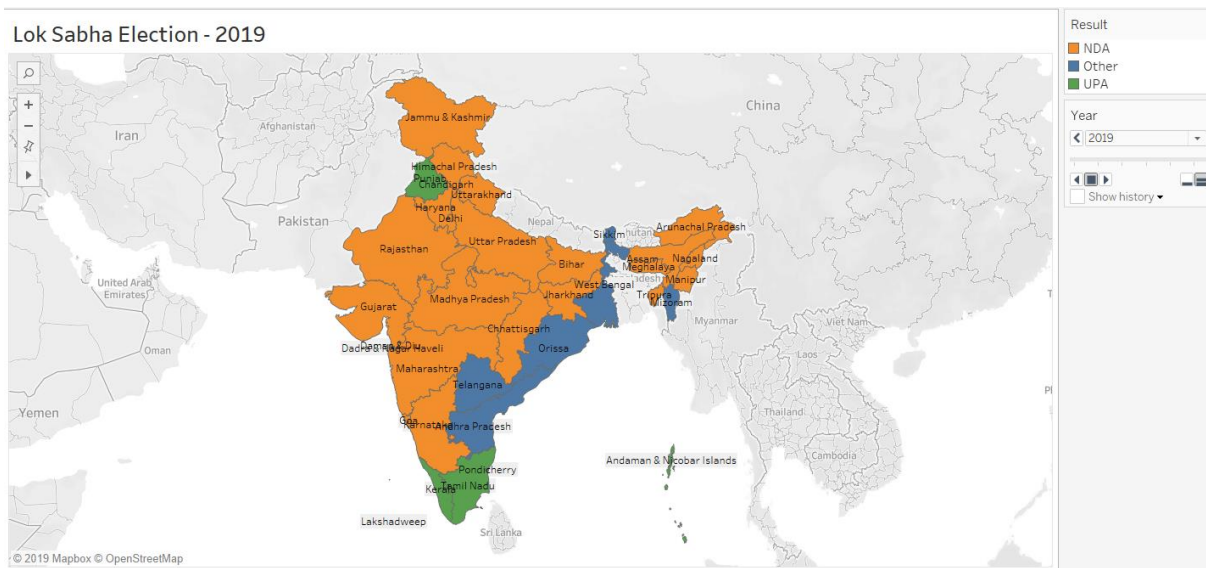
Visualisation

❖ Symbol Maps

Tool used: Tableau



to



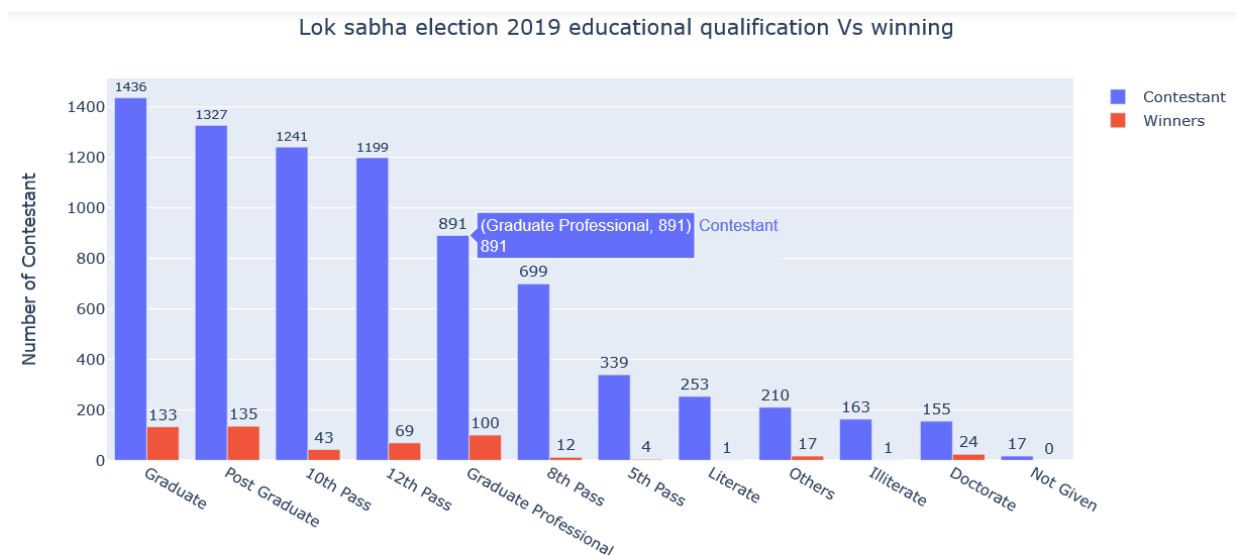
My first visualization is done using tableau, I have used this because it provides us with a greater ability to more easily understand the number of seats won by different parties in the different states. In this orange color represents seats won by NDA and green color represents seats won by UPA and blue color represents the states won by other parties right from 1996 to 2019. 6 different pages were joined based on year to form this animation. I have used orange for NDA and green for UPA because these are mainly used by the party itself and blue color for other parties due to the high contrast between which can be easily identified by the human eye. This symbols map gives us a clear picture of all the last 6 Lok Sabha elections. In the year 1996, neither of the top two parties have gained the majority to form the government which forced the country back to the election in 1998. After the government collapsed in 1996, the election of 1998 took place. The outcome of this election was also indecisive as no alliance

was in majority to form the government which led to fresh election in 1999, in which NDA alliance came into the majority and formed a stable government for a period of 5 years thus ending a period of political instability in the country which has been characterized by the last two Lok Sabha elections. In 2004 and 2009 UPA government came to power for two successive times. In 2014 and 2019 NDA came in majority by winning 336 and 353 out of 543 respectively this was the biggest win of a party in the history of Indian democracy.

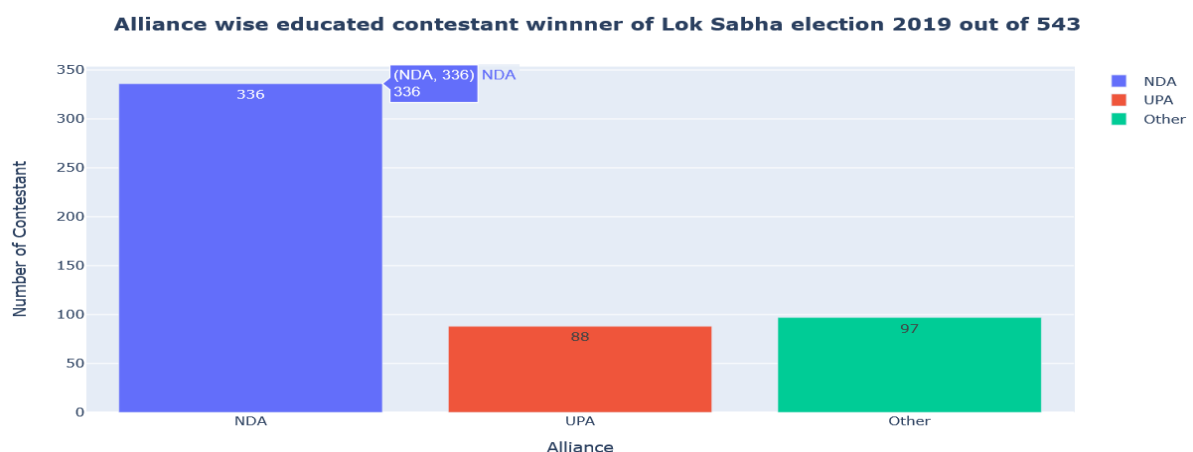
❖ Interactive Bar graphs

Created using: Plotly using python on Jupyter notebook

Libraries used: plotly.graph_objects



This is an interactive bar is graph drawn in group mode. The X-axis represents educational qualification and Y-axis represents the number of contestants. I have used this bar graph because it is good to way to show relative size. we can easily get the information about the number of qualified contestants who participated in the election and how many of them have won the election. The simple color scheme has been used to clearly distinguish between the contestant and winners. X-axis title has been tilted so that entire qualification name fits in the frame.



This is an interactive graph drawn in normal mode. The X-axis alliance and Y-axis represents the number of contestants. I have used this graph because the bar graph is best to show a comparison between the different attributes. A simple color scheme are used to distinguish between them. This graph gives us information about the number of educated candidates who are winners present at different parties.

Following things have been taken care while creating visualization.

- Legend displayed properly.
- Color combination to clearly distinguished between the attributes.
- Scale for each axis.
- Correct font size so that name of states should be clearly visible in tableau.

Challenges faced.

- Data gathering was really challenging because I have to manually gather the data which I want to use.
- Initially tried to create map of India using python libraries but was not satisfies hence dropped the plan and used tableau for creating the same.
- The Major challenge faced while creating the map in tableau as this is the first time, I am using tableau for this I have gone through some videos in YouTube to create the same.

Conclusion

Indian democracy is the largest in the world. This modern era trends have changed people thinking where people analyse fruit and loss before casting their votes to anyone and they are not wasting their votes. From the current analysis, I can conclude that contestant is getting elected only if they are educated and have the vision for the development of the society. Parties are only getting the majority in the central government based on their history of good work and decision which is beneficial for the people.

Improvement can be done by sending only the highly educated member with skills to parliament so that they can play a vital role in decision making.

There were no effects or functionality that were not achieved.

References

- [1] Plotly bar graph
<https://plot.ly/python/bar-charts/>
- [2] Creating district level maps in tableau
<https://www.youtube.com/watch?v=hN5Q2bve3n4>