

INFORMED VOTERS: THE LOK SABHA OF INDIA

HCDE 511 Data Visualization Project

A visualization of the attendance patterns of Lok Sabha (House of the People)

Abstract

Tasked with designing an interactive visualization, our group turned to the Indian parliament elections of 2014. With its 1.27 billion residents and rapidly growing economy, India constitutes an important player in global politics and trade. However, for many decades, lack of public data availability within the country prevented any large-term political visualizations. Open government in India has come a long way since the last elections five years ago. With the recent launch of initiatives like Data Portal India, government and governance related data sets are now for the first time publicly available to all Indians on a single platform. Using user centered design methods we visualized the performance of parliament representatives during the last election term (2009 - 2014). Our goal was to let people find patterns and trends in the data both for parties as their own local representatives. The final visualization is meant to be an electoral awareness tool that empowers the people of India. User feedback was highly positive, with many commenting on the enjoyable, low-access nature of the interactive design.

Keywords

Information visualization, India, Elections, Lok Sabha

Introduction

With its 1.27 billion residents and rapidly growing economy, India constitutes an important player in global politics and trade. Characterized by rapid urban modernization, the country is quickly outgrowing its status as a developing nation. For many decades, apathy and lack of easy access to data blocked a majority from real political participation or awareness. In late 2013, the country officially inaugurated Data Portal India. Its goal was to promote transparent governance and innovation. The Hindu — a leading Indian newspaper referred to it as “The beginnings of India’s open data movement” (source). Increased access to government data allowed the citizens of India to create visualizations and use data to start fueling innovation.

India is a republic with a multi-party system, of which the Parliament is the supreme legislative body. The Indian Parliament is bicameral — the two houses are Lok Sabha (House of the People) and Rajya

Sabha (Council of States). An Indian parliamentary term lasts for five years. The people of India recently elected their representatives for the 16th Assembly – the results of which were announced on the 17th of May 2014. The Constitution of India dictates that the Parliament can have a maximum strength of 552 members. Membership is divided among the 29 States (newest state created on 2nd of June 2014) and 7 Union Territories with the ratio of seats proportional to state population. An additional 2 representatives can be appointed by the President from the Anglo-Indian community if they are underrepresented in the House. The period during which the House meets to conduct its business is called a session. The parliament conducts three sessions each year:

- Budget session (February to May)
- Monsoon session (July to September).
- Winter session (November to December).

The periods other than these when the Parliament convenes are called Special Sessions.

Our data focuses on Lok Sabha session attendance spanning a period of nearly 4 years (the first 13 sessions) from 2009 to 2013. This corresponds to the fifteenth term of Parliament, which came into being after the last elections in May 2009. Data from session 14 and 15 (November 2013 - May 2014) was not available on the data.gov.in portal at the time of creating this visualization.

Previous Work

Atre & Hoang-Le's visualization from Tableau Public

Shortly before the results for the Indian election of 2014 were to be declared, several visualizations started appearing on Tableau Public based on the data provided from Data Portal India. This particular visualization summarizes the breakdown of political parties from the 2009 and 2014 elections. Atre and Hoang-Le's visualization served as a source of inspiration in the use of the various party symbols with related color, use of a choropleth to drive the rest of the visualization, and to provide overview data after filtering with the map. This visualization creates an effective overview and comparison of political parties between the elections. It remains however very narrow in scope, limiting itself to party seat distributions. Another issue with the visualization is their decision to color the States by the party that won the largest number of constituencies in that state. This results visual distortion of data in case of an even split between parties (as observed for Arunachal Pradesh) or where the sum of others was higher than the number of seats with the parties which were represented individually on the visualization (Andhra Pradesh).

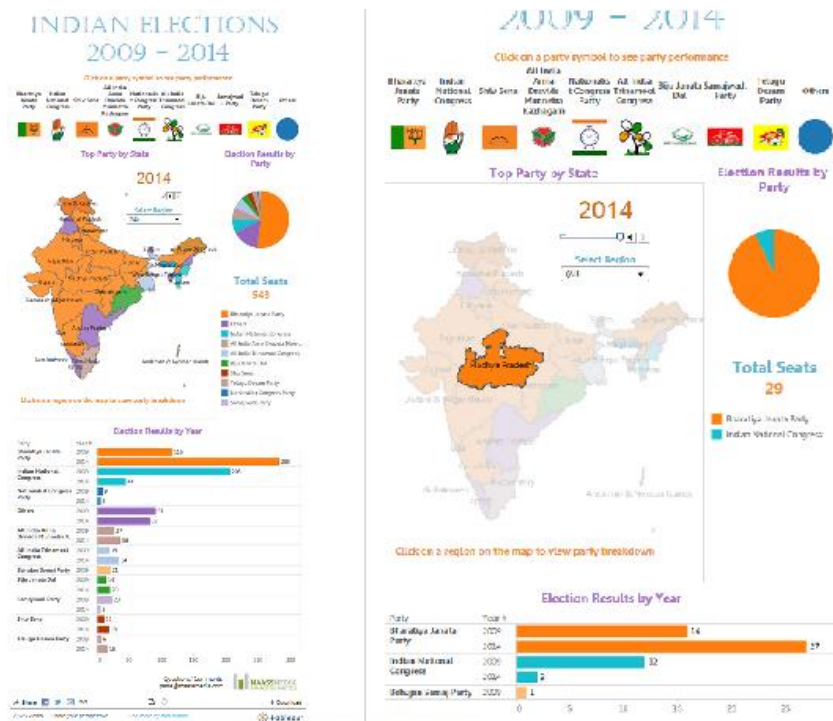


Figure 1. Priyanka Atre and Thanh-Trang Hoang-Le visualization from Tableau Public.
<http://www.tableausoftware.com/public/gallery/indian-elections>

The 15th Lok Sabha NDTV.com

During the election process a number of Indian media outlets provided their own visualizations of the parliamentary data. India's NDTV presented a multi-dashboard view that allowed users to dive deep into various aspects of parliament. The different dashboards utilize a wide variety of charts to communicate the vast array of data. One of the visualization's strengths is its multi-level approach: it allows users to zoom and adjusts the presented data accordingly. The dashboard however violates a few rules of color use. The same colors are reused between dashboards to indicate different variables. This can be misleading, as shown in figure 2. Whereas the gender break-up uses color to encode party, the break-up by age reuses the same colors to encode educational levels. This visualization also sacrifices data in some areas, highlighting only the extremes (Figure 3).

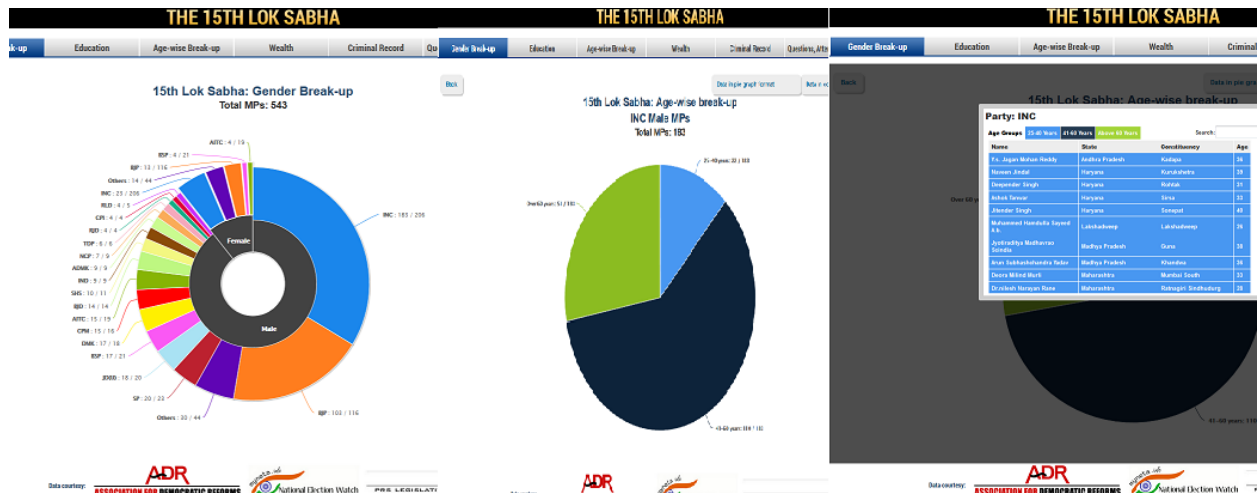


Figure 2. NDTV visualization. <http://www.ndtv.com/elections/page/charts>. Demonstration of the drill down function in the Gender and Party breakdown of Parliament.

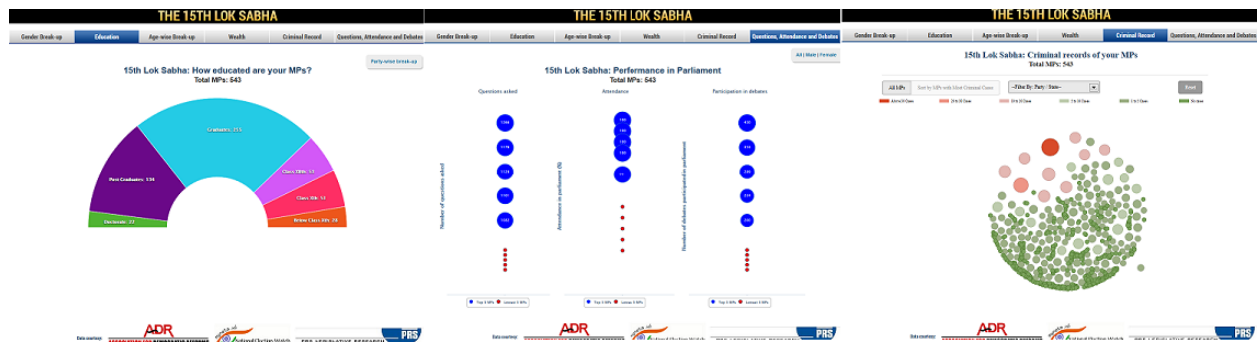


Figure 3. NDTV visualization. <http://www.ndtv.com/elections/page/charts>. Sample of various dashboard views.

Mirroar visualization: Summary of parties by constituency from the 2009 and 2014 elections

Similar to the visualization of Atre & Hoang-Le, Mirroar was created using Tableau. It places a series of polygons overlaid on the map of India. Party information is dual encoded with color and the actual party symbol. The primary map lacks interactivity, the only change that can be made is filtering gender. Selecting the second dashboard shows the results of the 2014 election and hovering over a constituency gives a large amount of data on demand. This visualization does well to create a constituency level of information — something that was lacking in the other visualizations presented. However, it does not support levels of comparison between constituencies and is very limited in the amount of data presented (Figure 4).

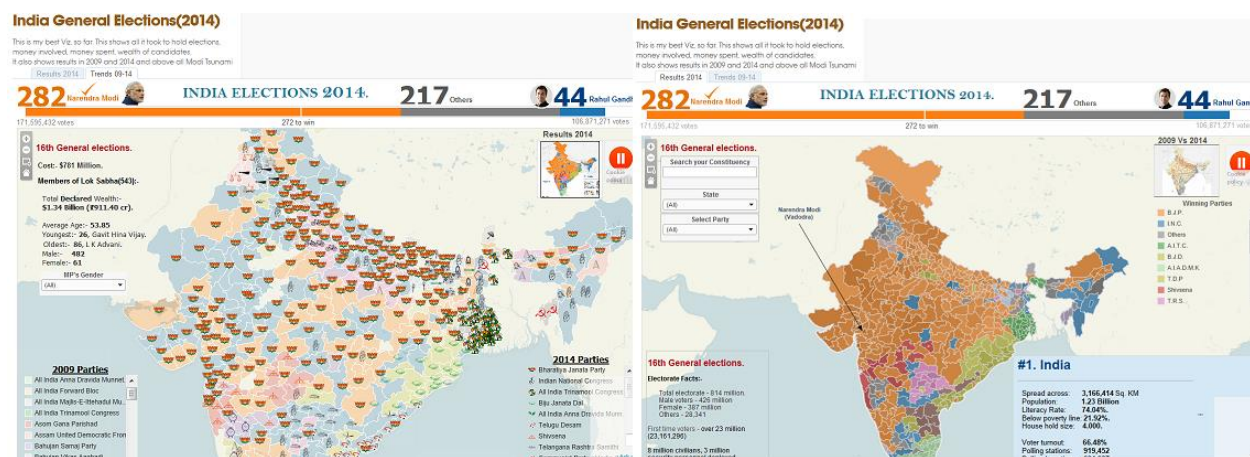


Figure 4. India General Elections visualization created on the Mirroar blog. Demonstration of the two dashboards used. http://blog.mirroar.com/post/86478270394/india-general-elections-2014?utm_content=bufferf9d83&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer.

Method

Dataset

The data that was used to create our final dataset was gathered from several different and disparate data sources. Sources included the Government Data Portal data.gov.in, websites of Government Institutions such as the Indian Parliament and Election Commission of India. Data was also requested from the Indian civil society group Association for Democratic Reforms.

Table 1

Proportion of missing data: variable completeness

Variable name	Size (N)	Type	Missing (%)
Session ID	13	Ordinal	0.00%
Session Duration	13	Ratio	0.00%
Session Average Attendance (for entire Parliament)	13	Ratio	0.00%
Session Start Date	13	Interval	0.00%
Session End Date	13	Interval	0.00%
Session Passes Issued	13	Ratio	0.00%
Session 0 Hr Matters Raised	13	Ratio	7.70%
Constituency ID	545	Nominal	0.00%
Constituency State	545	Nominal	0.00%
Constituency State Abbr.	545	Nominal	0.00%
Constituency	545	Nominal	0.00%
Fund Entitlement	545	Ratio	0.37%
Fund Released	545	Ratio	0.37%

Fund Available with Interest	545	Ratio	0.37%
Fund Recommended	545	Ratio	0.37%
Fund Sanctioned	545	Ratio	0.37%
Fund Expenditure	545	Ratio	0.37%
Fund % Utilization	545	Ratio	0.37%
Fund Unspent	545	Ratio	0.37%
Member ID	560	Nominal	0.00%
Member Name	560	Nominal	0.00%
Start Date as MP	560	Interval	0.00%
End Date as MP	560	Interval	0.00%
Current Status as MP	560	Nominal	0.00%
Gender	560	Nominal	0.00%
Age	560	Ratio	0.53%
Marital Status	560	Nominal	1.42%
Was Freedom Fighter?	560	Nominal	0.00%
Registered Criminal Cases	560	Ratio	5.00%
Education Level	560	Ordinal	0.53%
Education Bins	560	Ordinal	0.53%
Total Assets	560	Ratio	5.00%
Total Liabilities	560	Ratio	5.00%
Attendance	560	Ratio	0.00%
Is a Minister?	560	Nominal	0.00%
Party ID	39	Nominal	0.00%
Party Name	39	Nominal	0.00%
Party Name Abbreviation	39	Nominal	0.00%
Party Bins	39	Nominal	0.00%
Party Influence	39	Nominal	0.00%
Party Stance	39	Nominal	0.00%
Party Alliance	39	Nominal	0.00%

In his book “Now you see it”, Few (2009) describes ideal data to be of high volume, historical, multivariate, atomic, clean, clear, dimensionally structured, richly segmented and of known pedigree. We used these guiding principles to come up with our final dataset.

A dimensional model was created that linked each entity such as constituencies, MPs, parties, sessions in a database which was first populated from the raw data that we received. It was then used to generate our final CSV file that was used to create visualizations. The dataset underwent six major revisions – each time adding new sources, dimensions, cleaning, purging incorrect information. The Table 1 provides an overview of all variables that became a part of the final dataset.

The only manipulation that we had to perform on the data was regarding attendances declared for the seventh session which reported was twice for several days. This resulted in reported attendances being higher than the total number of days the parliament was in session. In these cases the data was cut-off at the maximum value.

Tool design

Filters

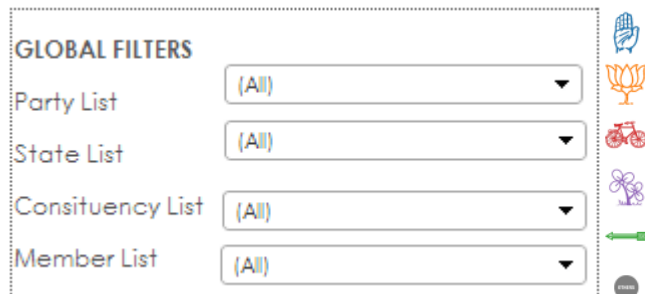


Figure 5. Global filters and party selector

The different levels of political familiarity within India necessitate a couple of different entry methods. Early exploratory interviews with contacts in India revealed that familiarity with party name, party symbol and political color strongly varies for different users. We addressed this challenge by designing a multi-access system, with inputs depending on whatever aspect the user is most acquainted with (Fig 5).

Other users were unfamiliar with one or multiple political elements of the viz. Political parties in India are traditionally associated to not just one, but two or three colors. That made it crucial to select an appropriate Tableau color palette. Also, parties like AITC (purple daisy symbol) and JDU (green arrow symbol) are regional parties and not well known in other parts of India.

The main purpose of these global filters is however to allow users to drill down to specific regions or candidates. With over 39 parties, 28 states (the state of Telangana was formed after the dissolution of the 15th Lok Sabha), 7 union territories and 545 parliament representatives in India there is no way to represent each in a purely visual way. The global filters provide a solution to this problem by listing all in an alphabetical way. A user may start typing a name to narrow the list length until a specific match is found. Alternatively a user might limit list length by selecting a particular party or state from any of the other driver elements, and then drilling down further through global filters. This might be useful in situation where the user does not know the names of his/her representatives.

Global filters may aid each other, or they can take input from any other visual elements like the party distribution (pie chart), geographical map (choropleth). It is possible to select 'female members' from the gender graph, click one's state, select an age group and have the global member list reduce to only a few members. This allows the user a maximum amount of freedom to find the information he/she is looking for.

Line graph

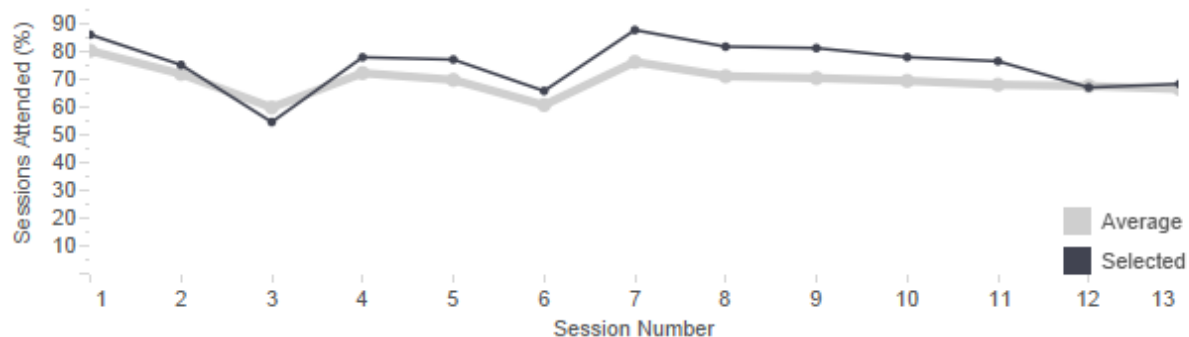


Figure 6. Parliament attendance of selected members

Our visualization originally focused on just a single depend factor: parliament attendance. A parliament term can be divided into 15 sessions. These sessions roughly correspond to a trimester system, with fifteen sessions forming the five year term between elections. Each session has an undetermined amount of representative meetings.

Parliament attendance is one of the most important factors in evaluating a representative's performance. Users want to make sure that their representative regularly attends meetings to voice their needs. A visualization of attendance has remained a prominent element in our design ever since the first iteration.

Attendance data is given per representative as the number of meetings he/she attended per session. The total amount of meeting per session is also given. Given that the amount of meetings per session is not fixed, it stands to reason to convert absolute numbers to percentages. Given the fact that only one number is given per session, a bar chart would be a natural choice. However we felt a line graph more accurately showed the underlying temporal continuity (even if the data itself was aggregated).

Pie chart

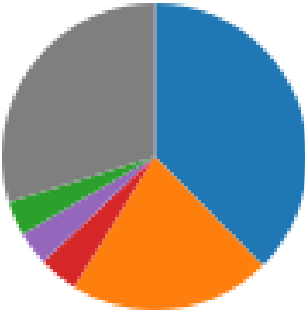


Figure 7. Party seat distribution

Much previous work into political visualizations rely on pie charts to convey party seating (Fig 1, Fig 2, Fig 3, Fig 4). Parliamentary sessions are commonly held in half-circular or circular shaped halls, so that pie charts form a direct reflection of the reality they're trying to represent. Much thought was given to whether percentages ought to be superimposed on the chart. This would be to make up for the inaccurate comparison of area (Bertin, 1983). However, percentages proved hard to do since the chart would dynamically update and some parties only hold a narrow hold. The thin slices of these parties always made the text illegible. It cluttered the view, and held people's attention while they tried to make out the text.

Ultimately we made a decision to present the exact percentage on hover instead. This made sure the exact information remained available as details on demand, without placing too much emphasis on the pie chart itself.

Choropleth map



Figure 8. Choropleth map of India

A map of India was included since the start of the design process. While initially just serving as a visual reference, we quickly found the map could be used as a driver to provide region-specific breakdowns of representatives. One difficult decision focused on whether to show constituencies or states. There is a lot to be said for constituency map. The base geographical element in India are constituencies: small clusters of villages that form a single election district. Representatives' data is ordered by the constituency they represent. Indian users are more interested in their local representatives than their state as a whole. However, the boundaries of constituencies are fairly arbitrary with unfamiliar names. This confused users during early user testing, when they saw upon zoom in that the map was not a simple city map. In addition, a Tableau map of constituencies would prevent users from selecting a state as a whole. Thus, we decided to opt for state map instead and provide the drill down to constituency through alternative means.

A second issue involved what information ought to be presented on the map. While we all agreed a map was a useful geographical filter, its surface area could at the same time be used to show an additional variable. We experimented with showing party seating per state, and population density. For the final design however we returned to a blank map to avoid confusion in the user. Putting data on the map drew away from its main purpose.

(Some link to class literature and the strength/weaknesses of choropleth maps)

Bubble chart

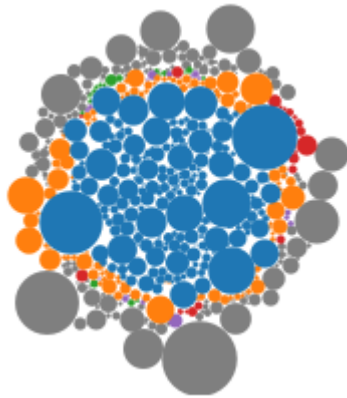


Figure 9. Criminal cases

Figure 9 represents the number of criminal cases for each representative. Every representative is one bubble (colored by party), with the radius representing the number of crimes. Bubble charts carry a number of flaws for data visualization. The relative size of the bubbles is difficult for users to accurately estimate (Heer, Bostock & Ogievetsky, 2010) and small entities are hard to select. It is however the only Tableau visualization that shows individual data points (other than a scatterplot). A scatterplot would however been inappropriate since only one measure is being represented. We made similar designs for representatives' total assets and total liabilities.

Scatter Plot

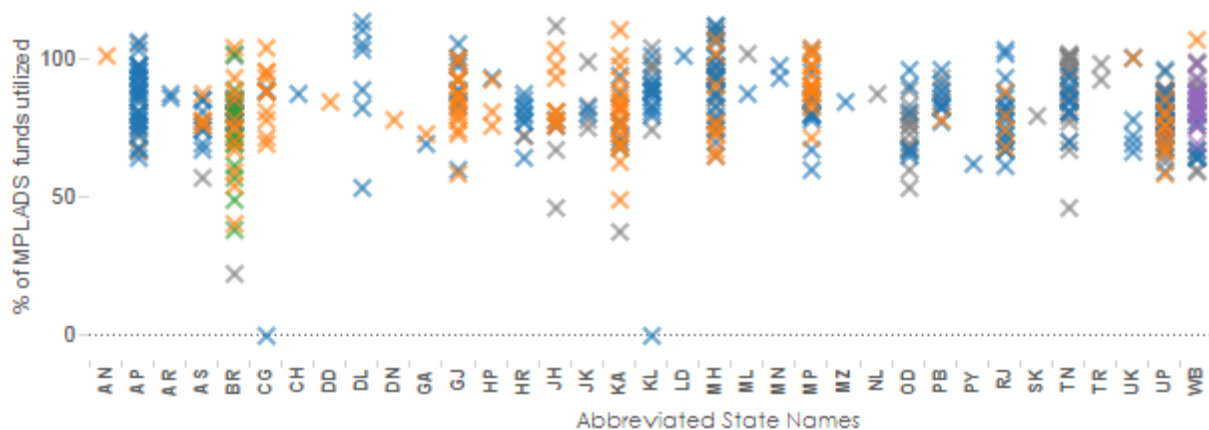


Figure 10. Member of Parliament Local Area Development Scheme (MPLADS) spending details

The MPLADS spendings is the second performance metric of our visualization, in addition to parliament attendance. For the 15th Lok Sabha a net amount of 19 Crore Rupees (around \$ 3.2 Million) were assigned to each representative, to be used in the development of his/her

constituency. This scatter plot shows the percentage of that money that was spent during this term of Lok Sabha.

The scatterplot went through a rather long design process. In its initial form the spending were presented on the state level as a bar chart. We abandoned that design in favor of something that also showed the range within each state (box plots), but quickly found that the large number of states made that approach impractical. Since 'state' functions as a categorical variable, we decided to use a scatterplot. The vertical arrangement of data points allows the user to still visually assess the range, at the expense of information on the mean. The conveyed information is however enriched by the coloring according to political affiliation. This already creates an immediate visual impression per party. The individual data points also work better because they can be dynamically filtered.

A number of graphical elements were tried out: dots, diamonds and crosses. Data occlusion remains an issue for unfiltered views. We found that crosses minimize the level of overlap between data points, thus at least limiting occlusion to a necessary minimum.

Note that a number of representatives also spend more than hundred percent of their assigned funding. This is an artefact of how the MPLADS website calculates this metric. Portion of a released amount to a representative may lie in a bank account and accrue interest till disbursement. High interest rates in India – 4-6% for checking accounts - causes the available funds to increase over time. In calculating % funds utilized, the expended amount is compared with what was originally released, causing this disparity.

Bar charts

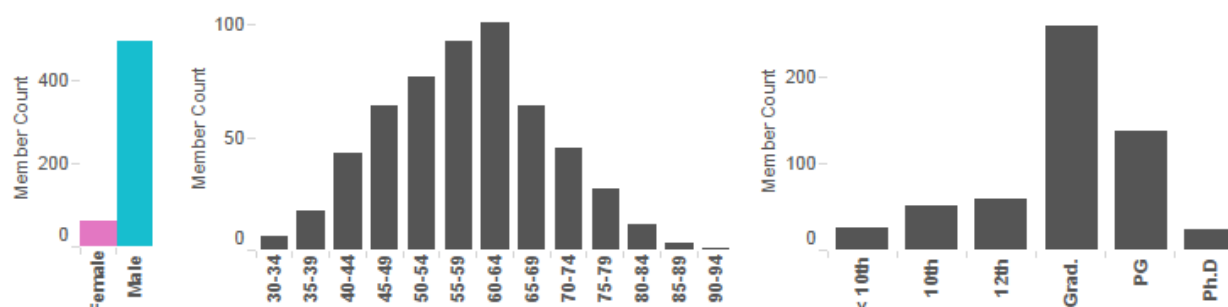


Figure 11. Distributions of gender, age and education

Nearly all of our visualizations originally started off as bar charts. Bar charts reliance on line length makes them a good visualization tool, as judgments of length tend to be highly accurate (Bertin, 1983). A good default because line width is easy to see. Most users are already familiar with this type of visualization, which makes them easier to read (Tufte, 1983). At the same time the proliferation of this technique in different media makes them less eye catching. The bar

charts for our visualization are located in the bottom right corner. There they are able to function in a more supporting role while more colorful graphs draw the users' attention. Color is a sensitive issue for this viz. — it is strongly associated with political affiliation. Grey coloring makes sure these graphs remain politically neutral.

Of course, one obvious alternative would be to forgo this view in favor of a more detailed political differentiation. Earlier iterations included gender, age and education as stacked bar charts, colored according to party. It produced a visual effect reminiscent of our colored scatter plot. While this solution worked out nicely for the scatterplot, it did not work well for the bar charts. As noted by Heer, Bostock & Ogievetsky (2010) stacked bar charts lack a common baseline which makes them harder to read. Furthermore the split per party was rather uninteresting, with similar proportions for the different age and education groups. Finally, the abundance of color distracted first users with all elements being equally prominent. By keeping the bars grey we introduced a visual hierarchy that none the less remains capable of still delivering the same information. Users can click any category of gender, age or education to restrict the data to members of that group.

As these bar charts all represent individual variables, they are deliberately placed together in the final design. Details in their style have been homogenized: vertical category labelling, identical y- axis labels. The axes themselves could not be fixed because of significant amount differences.

Note that the gender graph does include color. It forms a significant exception on this viz. to color for party affiliation. We decided that these hues are so well-known that the exception was justified. It adds to the playful nature of the overall design. None the less this aspect was extensively user tested, partially to rule out confusion but also to make sure the connotation held up within an Indian cultural context. The coloring of the gender graph makes it a natural gaze attractor and was therefore positioned close to another attractor, the colored bubble chart. This limits the attraction to only a few areas.

Interface design

The Lok Sabha visualization uses smart positioning to group variables in natural categories. Primary drivers (filters, map, and party selection) are provided in close proximity to each other in the top left corner. Additionally, we separated individual variables (criminal cases, assets, liabilities, gender, age and education) from performance metrics (parliament attendance, MPLADS spending). This serves to create a less cluttered feel, and allow users to quickly get started.

The design also contains a lot of different visualization types. Pie charts consistently came up during both our literature review and user testing as a strongly associated way of representing parliament info. In addition users needed a way to quickly deduce that the graphs were about India. The most prominent visual elements therefore had to be a map and the representation of seating. These two elements were the primary visualization drivers during most of our design process. Clicking a state (map) or party (pie chart) would filter the remaining information such as attendance. Global filters would have to be provided for smaller areas that were harder to select (Figure 12).



Figure 12. Early design showing a prominent map and party selector

The richness of different visualizations originates from a large amount of variables. We ran into trouble early on having only a limited amount of Tableau real estate to represent them. Even our final visualization is quite information dense with more than twenty-five individual elements. Our early work explored different means of making that amount more manageable. Two of the ideas we explored were tabs within the dashboard (Fig 13) and different dashboards linked together from a top menu (Fig 13).

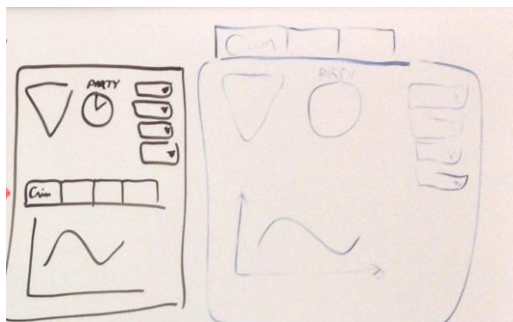


Figure 13. Early design showing tab menus

However none of those solutions offered the same freedom to explore links. We decided to limit the amount of data instead to only the most interesting variables. Information on wealth (total assets, total liabilities) was initially removed to relieve some of the information load. It

was re-introduced later as extra options within the bubble chart drop down menu. This option nicely solved the problem, given that only one bubble chart can be displayed at any particular time. (Fig 14).

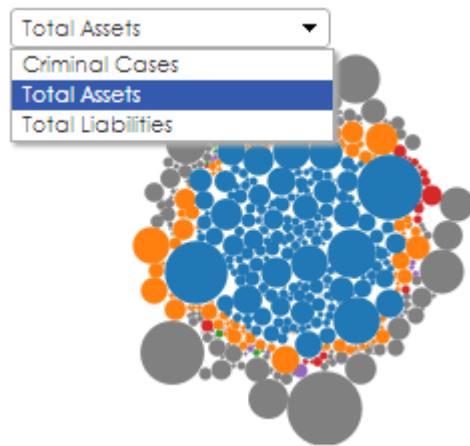


Figure 14. Multiple bubble charts, accessed through a dropdown menu

Usability study design

The goal of our usability test was (1) to observe how study participants interacted with the Lok Sabha visualization and pinpoint areas or pain-points where functionality could be improved (2) to gain insights relating to how scope and presentation could be altered for better knowledge acquisition, and (3), to observe and understand the interactions and motivations of our particular user group

User Personas

As an electoral awareness tool, our user base includes all Indian citizens of legal age. The country's social demographics are however very diverse. To model that kind of complexity, we created three personas based on 2011 census data for India as well as demographic information for three representative regions. We retrieved information on gender, age, education, profession and income as well as specific information about the states represented. In addition we collected cultural information from a number of independent agencies and government portals. Sources and background research for our user personas may be found as part of the appendix.

**Arnav, Age 29**

Farmer, Bihar

Goals:

- Determine attendance habits of incumbents
- Determine incumbent's work habits through use of constituency funds

**"I want to know how the candidate I voted for is performing... Are they doing the job?"**

Arnav is a married farmer from Bihar. He has access to internet, but does not consider himself highly internet savvy. Arnav is a hard worker and wants to see if the incumbents from his region are performing well.

Technical Skills:

- Check email
- Read online news sources

Motivation:

"The media and papers are biased, I want to find things out for myself"

Frustration Points:

"I don't like it when things take too long to happen on the internet"

Typical Resources:

"I usually get news from the local papers"

**Navya, Age 22**

Student, Dehli

Goals:

- Determine education levels of incumbents
- Compare incumbents based on demographics and location

**"I'm interested in learning about the incumbents lives, what are they like?"**

Navya is a student living in New Dehli. She has a very active social life while not in her studies. Navya is very tech savvy. She is heavily influenced by social media.

Technical Skills:

- Create facebook pages and groups
- Conducts detailed research online for school

Motivation:

"I want to compare people from my region to the rest of the country without exhaustive searching"

Frustration Points:

"I don't get frustrated with much online, just as long as it works properly!"

Typical Resources:

"I get information from news websites, social media, blogs, and television"



Sanjay, Age 46

Hydroelectric Plant Foreman, Andhra Pradesh

Goals:

- Determine financial status of incumbents
- Compare incumbents performance based on demographics and party



“I want to see how my party is performing. Party level politics are very important to me”

Sanjay is a hydroelectric plant foreman in Andhra Pradesh. Sanjay is a habitual voter and focuses on party divides. Sanjay has to use computers and technology for work on a regular basis.

Technical Skills:

- Send email and navigate the web
- Interact with complex computer programs for work

Motivation:

“I want to learn more about my political party and how it is represented”

Frustration Points:

“I don’t like it when I have to go multiple places for one thing”

Typical Resources:

“I get information from news websites, newspapers, and television”

Participants

A total of seven individuals were recruited for the usability testing sessions that were conducted as a part of this project. Two individuals helped evaluate the low fidelity prototypes while the rest were involved in evaluating the Tableau visualizations. All participants were Indians and were recruited through personal connections. Three out of seven were residing in India during the time of the study. The rest were residing in the US for a median of 10 months and a maximum of 2 years. For participants that weren’t collocated with the researchers, Tableau public was used to show the visualization and instructions were delegated remotely using Skype. The participants’ self-assessed knowledge of Indian politics ranged from being “completely ignorant” to being “in continuous touch with the latest happenings”.

Pilot studies

In the first week of the project we conducted interviews to understand the diversity and political knowledge of our users. We also gathered data from them regarding our low fidelity prototypes. This phase of user testing resulted in a number of design principles, which we tied in to Shneiderman’s infovis tasks (Schneiderman, 1996). For example, most users were interested in only a few specific variables which had to be instantly available as part of an overview. People were often only familiar with parties from their resident area, which necessitates a way to efficiently zoom from a high-level overview into specific regions.

Our investigation of other visualizations had shown many visualizations neglect to tie the data views together. Allowing users to relate different trends to each other would have to be one of our goals.

Our first user evaluations were made using low fidelity paper prototypes. These included three possible design frames, along with post-its representing different graph types. The different graph types could easily be edited, swapped or moved to different locations on the frame (Fig 15).

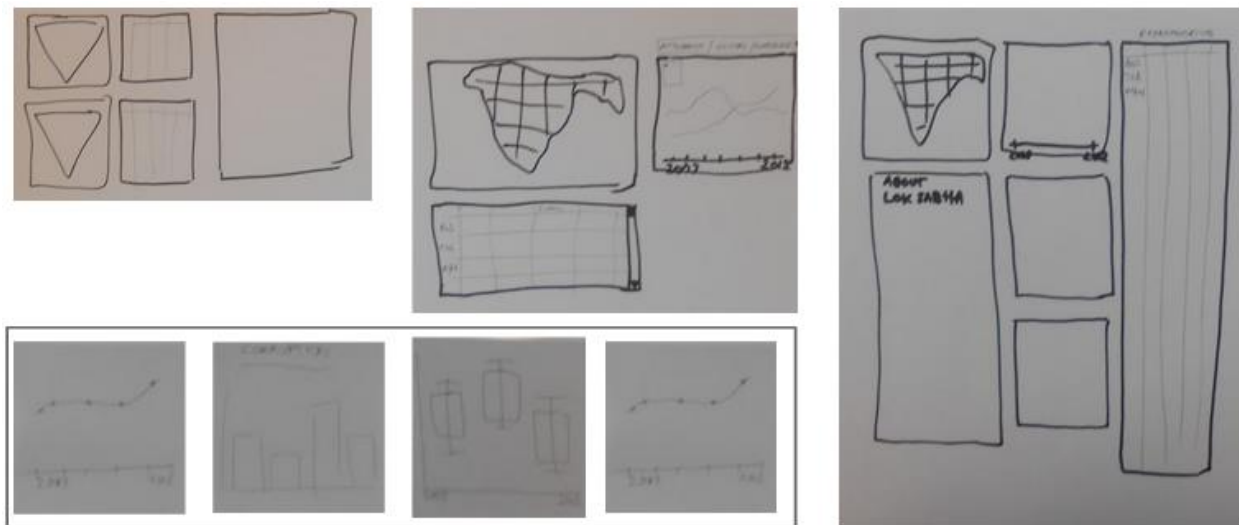


Figure 15. Paper prototypes

We also gathered user feedback on different graph prototypes in tableau itself. Most of those efforts concentrated on the choropleth map: what variable would you expect to see? Which color palette makes more sense? Users had trouble telling the difference between different shades of a single color. A subsequent heat map (adjusted for color blindness) provided better results, however significant confusion remained over color value and the mid-range white colored states (Fig 16). The final design ultimately reverted the design back into a plain geographical map.

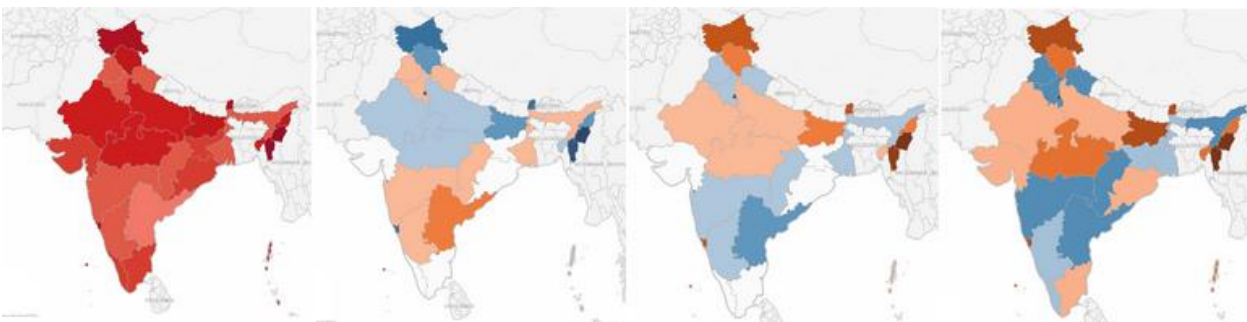


Figure 16. Choropleth maps, trying out different color schemes.

Finally, we asked participants to draw on a whiteboard what they expected the interaction flow to be like. Both participants started out by drawing maps of India then drilling down to state and then constituency level to find their representative. The information that they were interested to see on their representative included demographic information such as age, education level, criminal history and performance data such as developmental schemes implemented in constituency, attendance in parliament and issues raised.

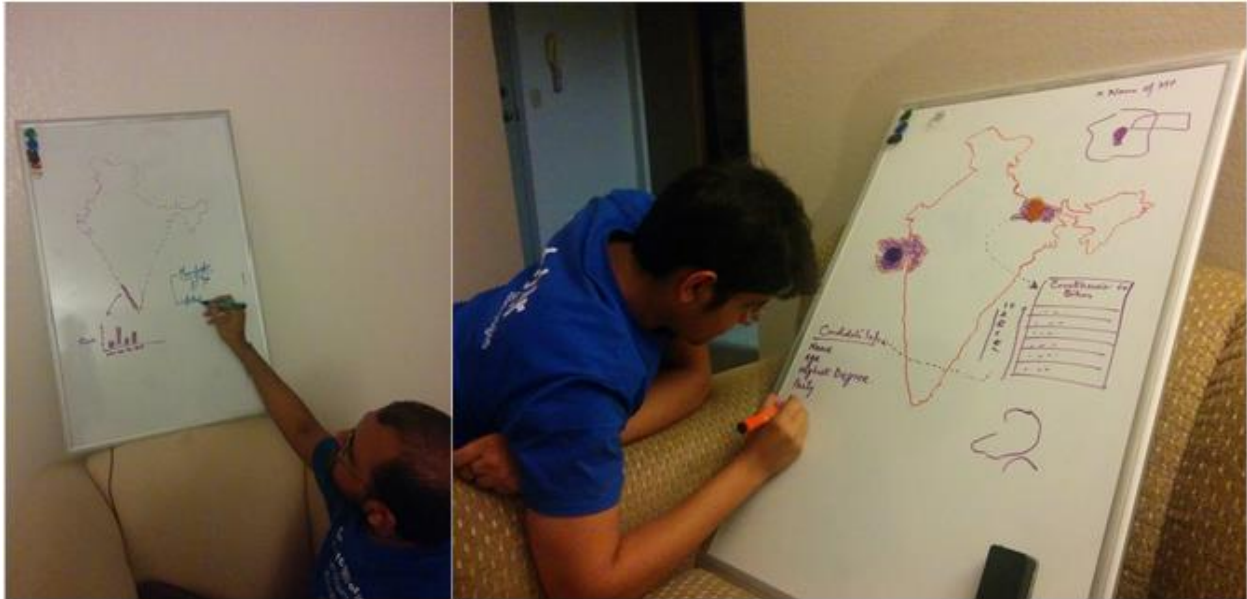


Figure 17. User white boarding

Redesigns/Iterations

Usability testing on two alternative designs.

Based on user feedback we developed two high quality Tableau prototypes. Both versions were fully functional and explored a number of designer choices. We tested these designs with five participants (All Indian, three of which located in India).

The first design is characterized by a lot of textual instructions, acting as clues to each visualization. It includes one global filter (State) and a choropleth map displaying population density (Figure 18). All graphs provide further detail on demand.

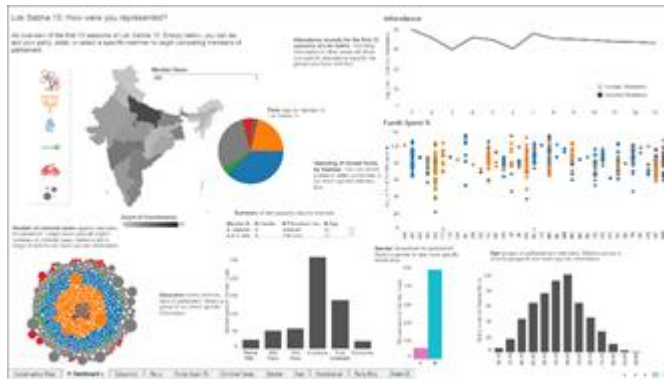


Figure 18. First hi-fidelity prototype for testing

The second design adds more visual reference points at the expense of text instructions. It tries to grasp the audience's interest by including elements such as a national banner and political caricatures of prominent Indian politicians. Though not strictly necessary from an infovis point of view ('chart junk', Tufte 1983), these elements enliven the display and help guide the user's gaze in a better way than white space would.

The second display adds a second global filter (member list), and reverts the choropleth to a geographical map of India. The map is also larger and more centrally placed to draw a user's attention. It introduces a table near the top with more detailed information, depending on which representatives are selected. Finally, it introduces a number of smaller esthetic changes such as a more coherent style for bar charts and a repositioning of different visual elements (Figure 19).

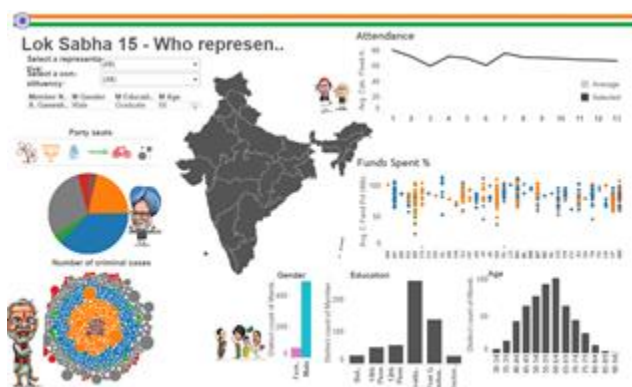


Figure 19. Second hi-fidelity prototype for testing

Challenges during usability testing included the time zone difference between Seattle (GMT - 7) and India (GMT + 5:30). Usability sessions were scheduled accordingly. Poor local internet connectivity also resulted in a non-continuous session flow. The connection was lost a few times and small bandwidth slowed down mouse over actions for users in India. However, we do not believe this influenced the usability results in any major way.

Results

Both alternatives

- X and Y axes labels for most graphs needs to be reworded. Many were undecipherable (*"what does 1 to 13 mean?"*), or were cut off because they were too long to display.
- Data labels for detail on demand were unclear.
- The label 'Funds %' needed to be reworded to a more intuitive alternative.
- Occlusion problem for the scatterplot
- Some party symbols were unclear (*"I don't know what the arrow or the four dots represent"*) – The arrow refers to a regional party, the four dots was a symbol created by us to denote an 'others' category of smaller parties.
- The term "15th Lok Sabha" was confusing
- Color associations were easily recognized

First alternative (Text focused)

- Much easier to understand interactivity when starting out. (*"Better navigation with more info"*)
- The title ('How were you represented?') was more understandable than the second alternative ('who represented you?')
- The design felt less structured and information heavy (*"Looks cramped."*)
- Users thought the choropleth was just a map with no info; next they thought it was a criminal distribution. Users did not see the population density legend until pointed out.

Second Alternative (Visually focused)

- Looked more organized compared to the first alternative.
- Users did not understand what most of the graphs were about.
- Users did not understand right away that the graphs are interactive (*"thought it was a one way view"*)
- The global filter for constituencies contains names of state instead of constituencies.
- Users confused the party distribution (pie chart) to be a distribution of criminal cases (bubble chart)
- Users did not notice the members table until very late in the session.

Discussion

During the usability testing, we discovered that Tableau distorts the dashboard upon uploading to the Tableau Public server. This issue has to do with the user's browser configuration and screen resolution. From users' feedback we concluded that written instructions are highly

appreciated, along with a thoughtful visual arrangement of the graphs. Both alternatives were criticized for having non-intuitive variable labelling, along with some user confusion concerning certain graphs like the choropleth or the scatterplot.

Usability testing on combined, final design

We combined the strengths of both prototypes into a final design. We then addressed the user concerns identified in the previous phase and iterated upon this design. The number of filters was increased from two to five. More written instructions were added to provide context to users about each graph's function. Axis labelling was reviewed and changed. Finally, each of the graph elements was redesigned to address user concerns.

Results

The final visualization is seen in figure 20.

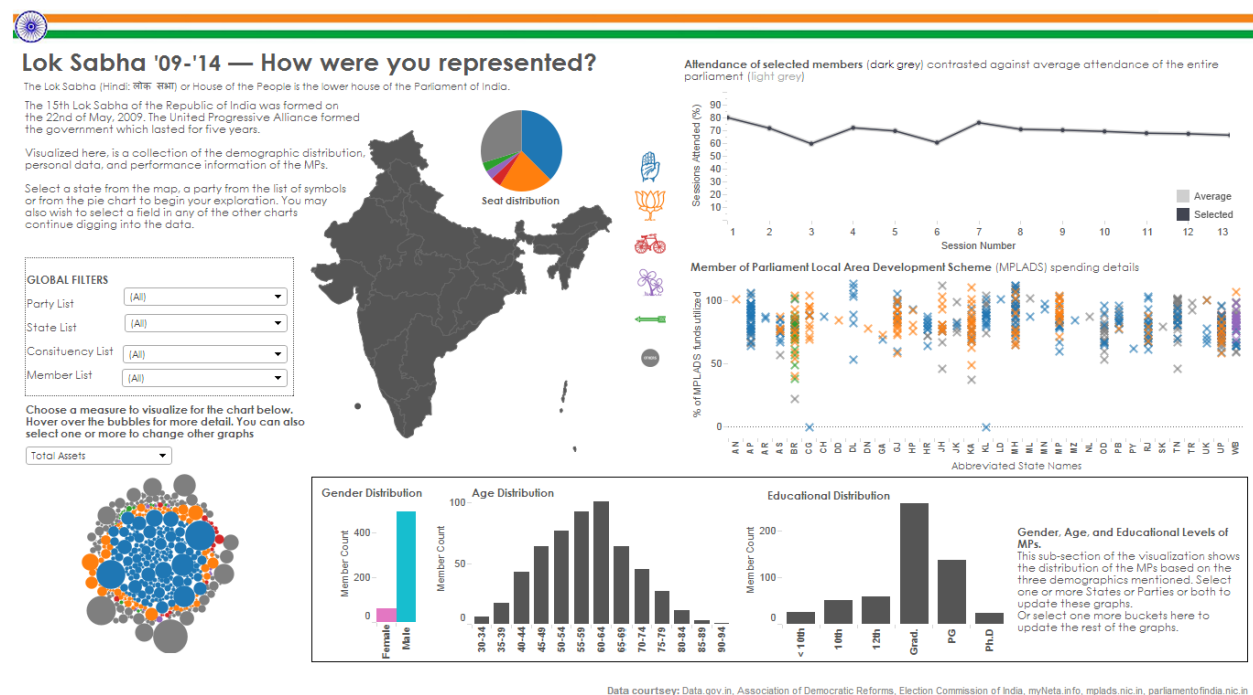


Figure 20. Final visualization

Discussion

Along with making changes as per user feedback, we added a couple of features to the visualization. We updated the bubble chart to display different metrics such as criminal cases, total assets, total liabilities based on user input. We also added the functionality of highlighting the relevant bins in the demographics charts and MPLADS chart instead of updating the view for each selection made on the bubble chart. The same interaction was also made possible for

selections made on the MPLADS graph. The demographic bar charts could be used to filter and highlight other graphs too. The only graph that was excluded from driving other graphs was the attendance graph as it was the only visualization containing time series data.

Evaluation

Our usability study showed that participants were highly satisfied with the final design. Many noted on the intuitive interface and how they felt ‘guided’ through the process. While they did feel like a lot of information was presented to them, it was not perceived as overwhelming. One participant noted: “The amount and the way in which the information has been presented in a limited screen space is quite impressive.” One user referred to this as “leaving a trail of breadcrumbs”.

Further Work

State-to-state or Party-to-Party comparisons

As it stands, our current visualization does not allow for comparison of one or more states or parties. A user may select one state or one party, then make a separate selection. Multiple selections of states although possible through the Map only result in an aggregated view of the data in all other graphs. Currently, users can see a general overview of demographic and performance data. The only chart on the visualization that allows for a direct state-to-state or party-to-party comparison is the MPLADS chart. Moving forward, we would like to add an option to select two or three states or parties and demonstrate that data in the other charts such that each data point can be uniquely associated with a state. This would likely be achieved by creating a second dashboard that is linked to the one we have created. The current dashboard lacks space required for comparison data as mentioned since we would likely require 2 of each chart.

Party bins

The final version of our visualization severely limited the number of parties we could show. India has 1616 parties registered with the election commission out of which 39 parties were represented in the Parliament (independent candidates are grouped as “Independents”), which makes encoding extremely difficult. Tableau’s pre-loaded color sets allow for a maximum of 20 colors to be used, many of these appear similar to one another making it very difficult to distinguish between parties with similar colors. Furthermore, many of the parties repeat colors, which would make encoding party to color very difficult for users to interpret. Similar issues would occur with shape as an encoding tool. For future versions of this visualization, we would like to be able to create a drill down method to allow users to select from national versus local parties, then further down through levels of party size and power within the government. Improvements on our current design would likely require expansion beyond the current five plus one other grouping that we have, and increasing to between 8 and 11 parties with a group of

others. Users would be able to interact with the “others” group to drill down into the other groups of parties contained within it.

Constituency Level representation

We were limited by Tableau, in that Tableau does not recognize constituency borders outside of the US. Generally, one Member of Parliament is elected per constituency. With limitations presented by Tableau, we were forced to represent our data at the state level. This created several problems for comparison and limited us to broad comparison in the majority of our charts. We found several visualizations where designers created polygonal overlays to represent constituencies. Having time to do this would allow us to use the choropleth more effectively and efficiently. We would be able to use filters and other selection tools to give users the power to select the data represented. This would also allow us to more effectively use some of the comparison methods mentioned above without creation of a second dashboard.

Acknowledgements

We would like to thank all participants and contacts in India for their time and cooperation during the testing of this product. Data sets for this visualization were kindly provided by the following sources:

- Data.gov.in (Attendance data)
- Association of Democratic Reforms (Attendance data)
- Election Commission of India (Attendance data)
- MyNeta.info (Attendance data)
- Mplads.nic.in (Attendance data)
- Parliamentofindia.nic.in (Attendance data)

A joined dataset is available online as part of our Tableau Public workbook.

References

Atre, Priyanka; Hoang-Le, Thanh-Trang. (n.d.) Indian Elections 2009-2014. Retrieved from:
<http://www.tableausoftware.com/public/gallery/indian-elections>.

Bertin, J. (1983). *Semiology of graphics: Diagrams, networks, maps* (WJ Berg, Trans.). Madison, WI: The University of Wisconsin Press, Ltd.

Biswas, R. (2012, April 25). State wise Internet Users in India Census 2011. Retrieved from:
<http://updateox.com/india/state-wise-internet-users-in-india-census-2011/>

Census of India 2011 (Final Population Totals). (n.d.) Census Provisional Population Totals 2011.
Retrieved from: <http://www.censusindia.gov.in/2011census/censusinfodashboard/index.html>

Few, S. (2009) *Now You See It: Simple Visualization Techniques for Quantitative Analysis*, Analytics Press

Government of AP. (2008). AP State Portal: Official Portal of Andhra Pradesh Government. Retrieved from: <http://www.ap.gov.in/Pages/PortalHome.aspx>

Government of India. (2014). Open Government Data (OGD) Platform India. Retrieved from: <http://data.gov.in/>

Government of NCT Delhi. (n.d.) Delhi Government Portal. Retrieved from: <http://delhi.gov.in/wps/wcm/connect/doi/Delhi+Govt/Delhi+Home>

Heer, J., Bostock, M., & Ogievetsky, V. (2010). A tour through the visualization zoo. *Commun. ACM*, 53(6), 59-67.

Mirroar. (2014, May 22). India General Elections 2014. [Blog website]. Retrieved from: http://blog.mirroar.com/post/86478270394/india-general-elections-2014?utm_content=bufferf9d83&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer

NDTV. (n.d.). The Fifteenth Lok Sabha. Retrieved from: <http://www.ndtv.com/elections/page/charts>

NIC Bihar State Centre, Government of India. (n.d.) Bihar Government Web Site. Retrieved from: <http://gov.bih.nic.in/Profile/default.htm>

Shneiderman, B. (1996, September). The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings. IEEE Symposium on* (pp. 336-343). IEEE.

Tufte, E. R., & Graves-Morris, P. R. (1983). *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.

Appendices

Appendix 1: User Persona research information

Information used to research user personas. We selected three territories due to wide variance of per capita income in the country. Geographic location was also factored in as the areas tend to have widely different cultural makeups.



Figure X. per Capita Income of States in India. Bihar is marked in light blue, Delhi is marked by the red square, and Andhra Pradesh is highlighted in yellow on the Southeastern coast.

Population size was also considered. We selected three regions that represented average populations for small, medium, and average sized territories.

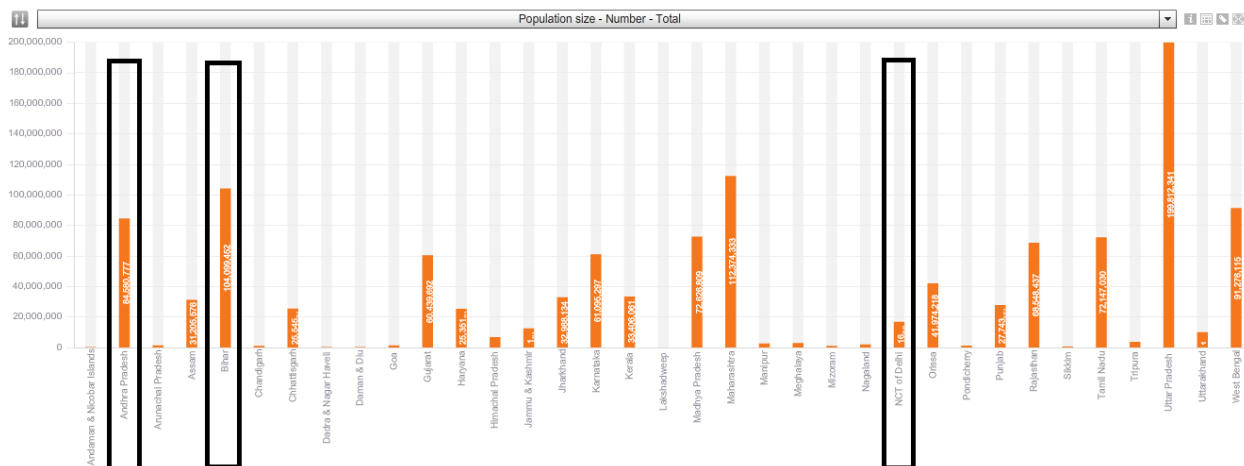


Figure X. Population size per state. Bihar, Andhra Pradesh, and the territory of Delhi are highlighted with black boxes.

Internet connectivity is not widespread in India. We selected areas that represent low, medium, and high levels of internet connectivity, and adjusted persona behavior accordingly. Despite the relatively low numbers of internet users in India, we had to assume that our users would have access either through home, work, or school.

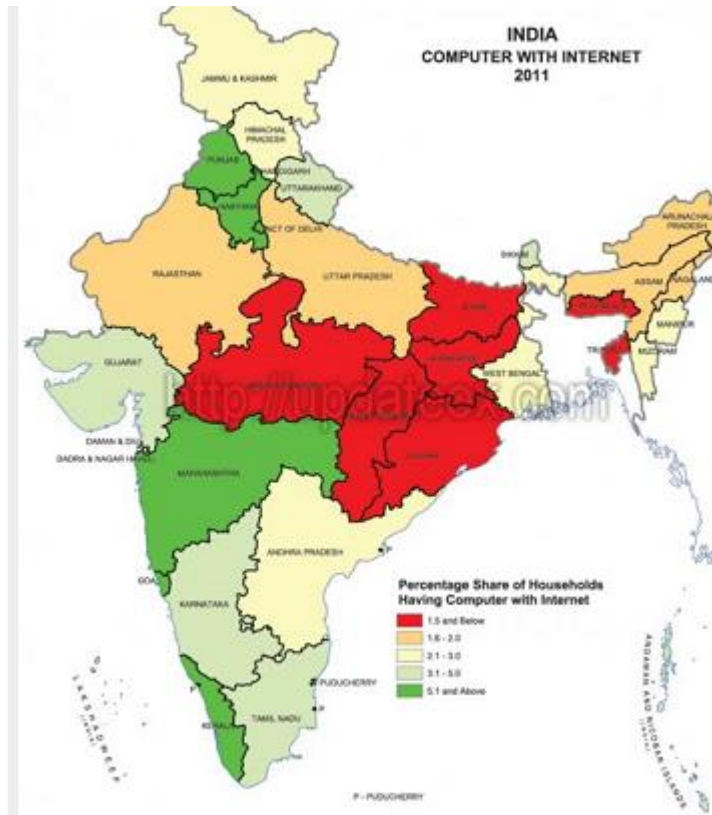


Figure X. Internet connectivity rates by percentage per State or Territory. Delhi has a rate of 5.1% or above with internet access, Andhra Pradesh has a rate of 2.1-3.0%, and Bihar has a rate of below 1.5%.