052215 - Visualization and Visual Data Analysis (VU)

ANALYZING PARLIAMENT SPEECHES AND HOW THEY HAVE CHANGED OVER TIME

MILESTONE IV

Homepage:

http://homepage.univie.ac.at/a1349198

Public Tableau:

https://public.tableau.com/views/project_77/OverviewofParties?:embed=y&:display_count=yes

Group 6

Team: Bacic Ines, Mesanovic Tarik, Lascsak Christian, Altaf Umair

Table of Content

MOTIVATION

Every month the parliament has many meetings on which it decides about, for the present or the future, important subjects or new laws. Its members meet every day, write the content of the meeting down and upload it online for the world to see. Of course, for many analysts this kind of available information is of great importance and can be used to create different statistics that can be put to use in particular fields.

With every day, the amount of data grows rapidly what causes new and more complicated needs and demands. Not all informations in the data are significant and can be used therefore filtering of the data is required. After applying distinct and for the cause important filters, the data can be reduces to a much more understandable and more usefull level where also the people without previous knowledge can work with it (not only the anaylists).
On the other hand, sometimes it is very important to understand the data and its meaning in order to work with it efficiently. For example, if a person without any knowledge of the data and its purpose sits down and tries to make something out of it, he might remove or not consider some important features that are significant. Therefore our goal groups are political analysts or different research groups with have to work with this type of information almost every day.

On the regular basis, the data is stored, collected, processed and analyzed. All this steps are done by anaylists and other experts. On the website from the parliament speeches, every meeting or speech is given in a pdf format(textual format). It takes lots of time to extract to most important information from the text and convert it into something more useful. Only then, we can think about visualizing and putting the information to its use.

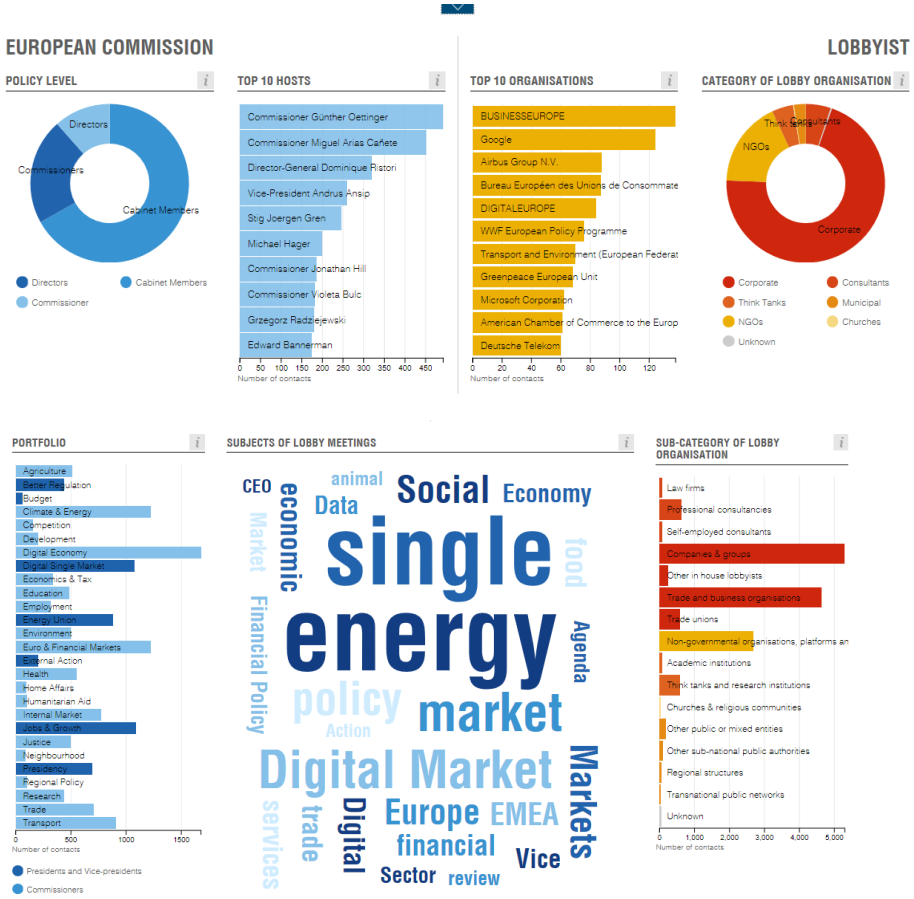
Therefore, we have decided to make a useful product for our goal user groups that can do their job more efficiently with it. We want to, not only present data in any way, but to think thoroughly about the future use of the product. Our project will therefore present different use cases that handle different kind of questions or problems that occur or are of interest to our goal users. We want to make possible for them to look more detailed in the meetings sessions and to explore to only on the content of the meeting and the theme but also on the person who spoke during it, the number of members that were present and their membership in a particular political group etc.

Our project will consist of three different dashboards in Tableau (interactive of course) where the user will be able to filter throughout the different parts of the dashboard and answer for him, in the moment, most important question regarding a

politician, political groups, a specific theme or many other subjects. On hand of our project, he should be able to make rational and statistical decisions and investigate the data efficiently. The statistics that come as a result of the visualization can be used to predict some future features of political groups, to analyze a specific politician or a group etc.

RELATED WORK

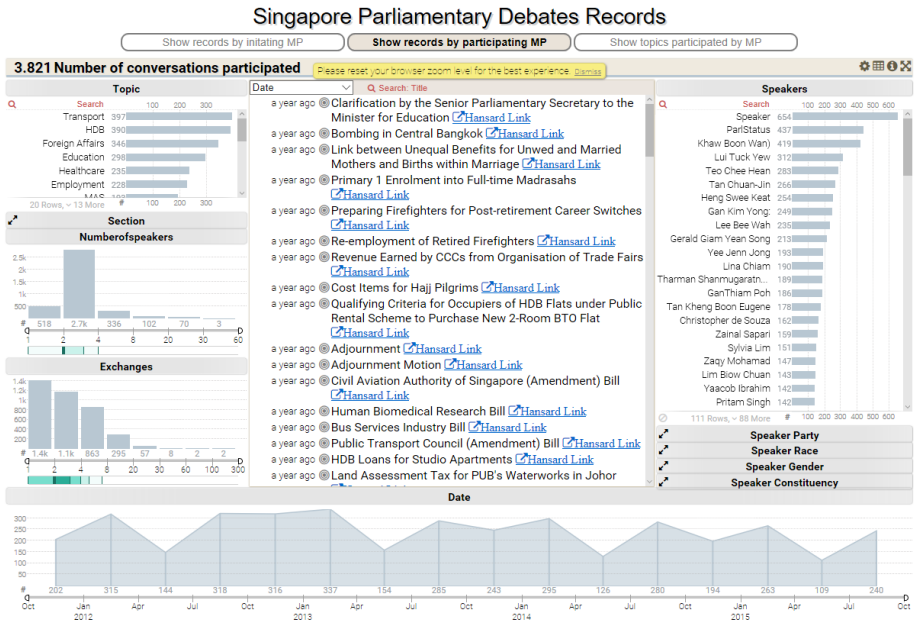
We found the Project integritywatch.eu. The people of this project want to give an overview of the Lobby meetings in the EU parliament. So they focus on a different problem than we have, but the problem of visualizing the data is quite similar.



http://www.integritywatch.eu/

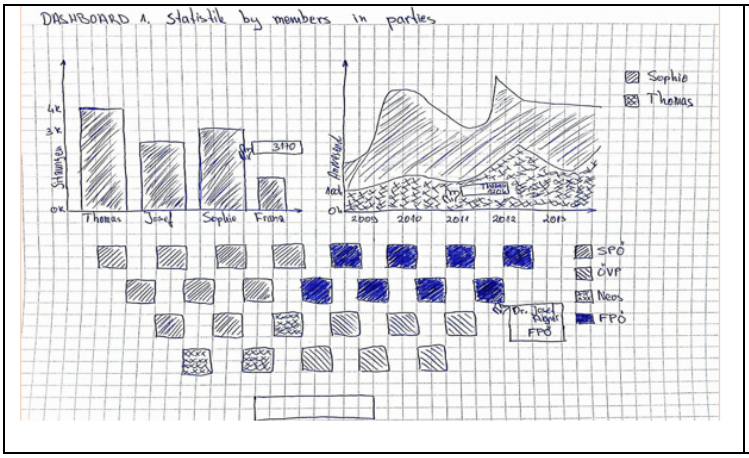
They also made a Word Cloud for visualizing the topics, just as we initially planned. Also, they made a bar chart for showing the top 10 hosts of a meeting. Since their focused problem was different than ours, they also included an overview of organizations and a list of hosts on the bottom of the Dashboard. We made a similar approach, since there are not many ways to show information about every person in the data.

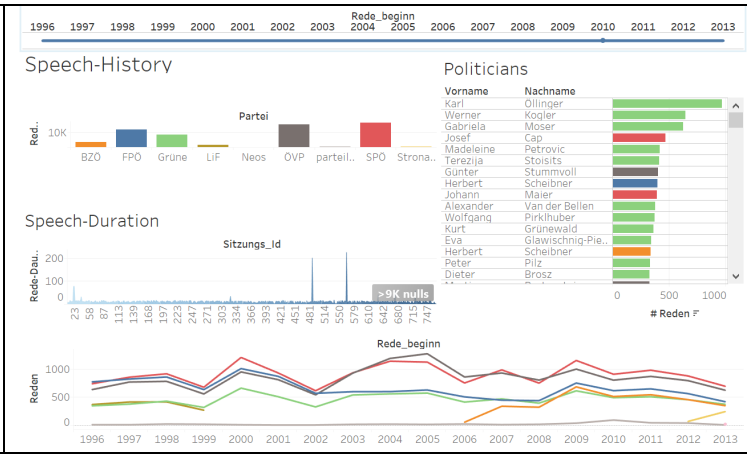
We also found a project, that visualizes the Debates at the parliament of Singapore.



https://s3-ap-southeast-1.amazonaws.com/hansardbrowser/index.html

This Project focusses mostly on what was talked about in the parliament. The graphs highlight information other graphs, when the mouse is hovered over an information. So you can see for example the participation of female speakers in the topics. We like this idea, because the whole dashboard changes according to the marked information. This helps to understand the data better and the user can filter the data for any problem they want to approach. On the bottom of the Dashboard is a timeline, that shows the amount of speeches held by time. We implemented a similar timeline, but in the end focused less on speech analysis. To show the participation of different speakers, they use many bar charts, which highlight when hovering over them. We decided to not just use bar charts, but also a bubble chart and a word cloud, to make it more appealing for users. Although we used a bubble chart for the politicians, to show how often they were participating, we changed our mind and used a similar approach to this problem, by visualizing the participation with bars. We also would have used a word cloud for the topics, instead of a bar chart. The advantage of using a bar chart here is, that they can show the participation of a specific person on a topic, not just overall, like we did. What we do not like about this visualization, is that the links in the middle of the dashboard use so much space which make it look very unappealing.





We included our whole first Dashboard concept, into our high fidelity prototype. The user can choose a politician, or a party and see the timeline for their participation. But since we already implemented an overview of the parties in our other Dashboard, we did not make another one for the second Dashboard.

APPROACH

Description of your visualization design

- **Bar Chart**

  In order to show the “gender statistics of the parliamentarians” in dashboard I, we decided to use a bar chart. The bar chart displays the general number of parliamentarians on the y axis and two bars on the x axis, which are filled by two colors, blue and red. The blue color represents the number of male parliamentarians und red the number of female parliamentarians. We kept the visualization in the bar chart dynamic, which means the gender visualization can be sorted by any year between 1997 and 2013.

  In dashboard II we have again used a bar chart in order to show the visualization of “held speeches by politicians”. On this bar chart we show you which political party held how many speeches per year. Further, this visualization can be sorted by any year between 1997 and 2013. The bar chart is linked with the table view in dashboard II. For further information, please read about the approach concerning the table view below.

- **Bubble Chart**

  In this data visualization we give you an overview of how many members of a political party are parliamentarians. Therefore, we have used a bubble chart. Each bubble represents a political party. The biggest bubble shows the political party

with the most members of parliament and the smallest bubble shows the political party with the least members of parliament. Further, if you move your mouse above the bubbles you will see the exact amount of the members of parliament per party.

- **Area Chart**

  For a general visualization of parliamentarians of each political party in Austria, we have used an area chart in dashboard I. The area chart gives you information about the political party name and number of parliamentarians of a political party of each year from 1997 till 2013. Moreover, the area chart is linked to the bubble chart. If you select any party in any year on the area chart, then you will see accordingly a bubble of the selected political party on the bubble chart and details on the politicians of the respected party on a table. The area chart is linked to the bubble chart and table view. For further information, please read the approach on the bubble chart and table view.

  The second area chart in dashboard II gives you the information on “speech duration”. It is linked to a bar chart and time line in dashboard II. Further, the user can click on any political party in the bar chart or / and select any year in the time line to see the information on the speeches accordingly.

- **Table View**

  In each dashboard we have provided a table. The table in dashboard I contains the details on the politicians, for example their title, first name, last name and party. To see these details you simply click on any political party in the area chart.

  The list in dashboard II gives you information on those politicians who held the speeches. The list is sorted by the most speeches holder. In order to see this information, you select any political party from the bar chart for speeches history.

- **Timeline**

  To show the time period from 1997 till 2013 as was provided in the data, we have used timeline. This enables a random selection.

Reasons for your design choices

- **Bar Chart**

  The reason why we have chosen bar charts is, that we wanted to keep the

visualization simple and friendly for our users. A bar chart is the best option, especially when the comparison visualization has to be shown. Further, it provides a quick and easily understandable view of a visualization.

- **Bubble Chart**

  The reason why we have chosen bubble bar for this visualization is, because on one side it takes less space on the screen and on the other side it is attractive for users. Especially bubble chart suits in our case because we have actual 9 political parties in the Austrian parliament. For the comparison of 9 different parties we have first tried other different charts, like bar chart and tree map but we couldn't get as good results as with the bubble chart.

- **Area Chart**

  The reason why we have chosen area chart for this visualization is, that the result of the comparison of the political parties with area chart was very neat and clean. The area chart is like the plot chart except that the area below the plotted line is filled in with color to indicate volume. The first political party (attribute) is plotted as a line with color filling followed by the second party, and so on.

- **Table View**

  With the table view you can easily show the information in a list form. The view is very common and it is uncomplicated to read the information provided and this is the reason why we have chosen it.

- **Timeline**

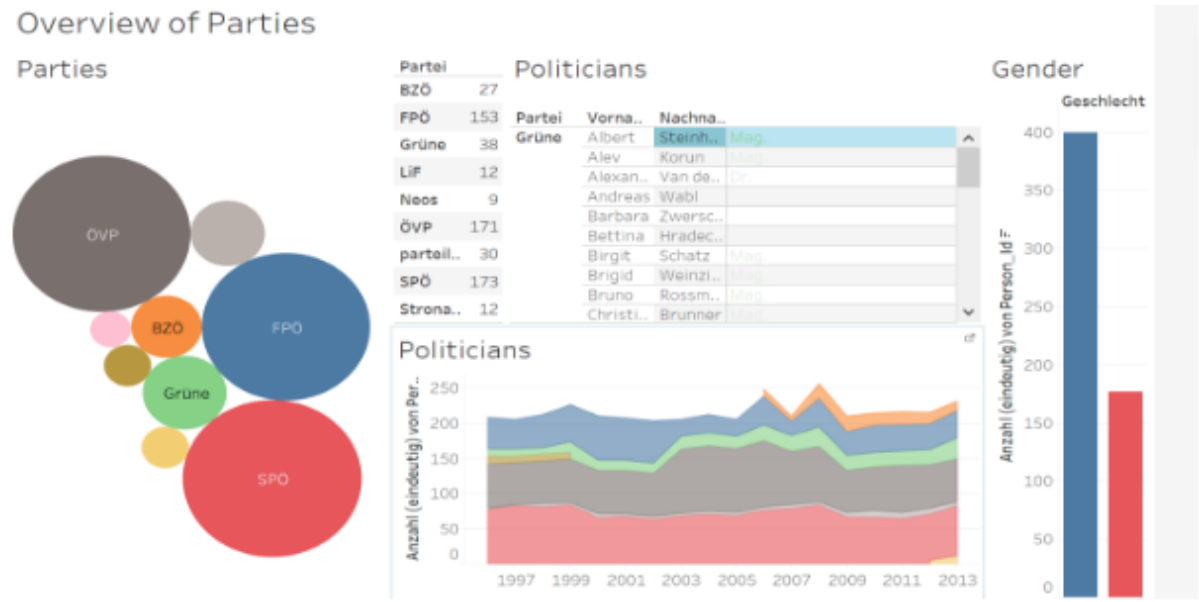
  The main reason why we have chosen the timeline is to provide the selection function to the users. With this function they can visualize the statistics of the selected year.

IMPLEMENTATION

Our entire project was implemented with Tableau. We've created two dashboards. The first dashboard with "Overview of parties" consists of five individual sheets (bubble chart, two tables, areachart and bar chart), which we also realized in tableau before. The second Speech - History dashboard consists of four sheets. Bar chart, area chart, line chart and a table.

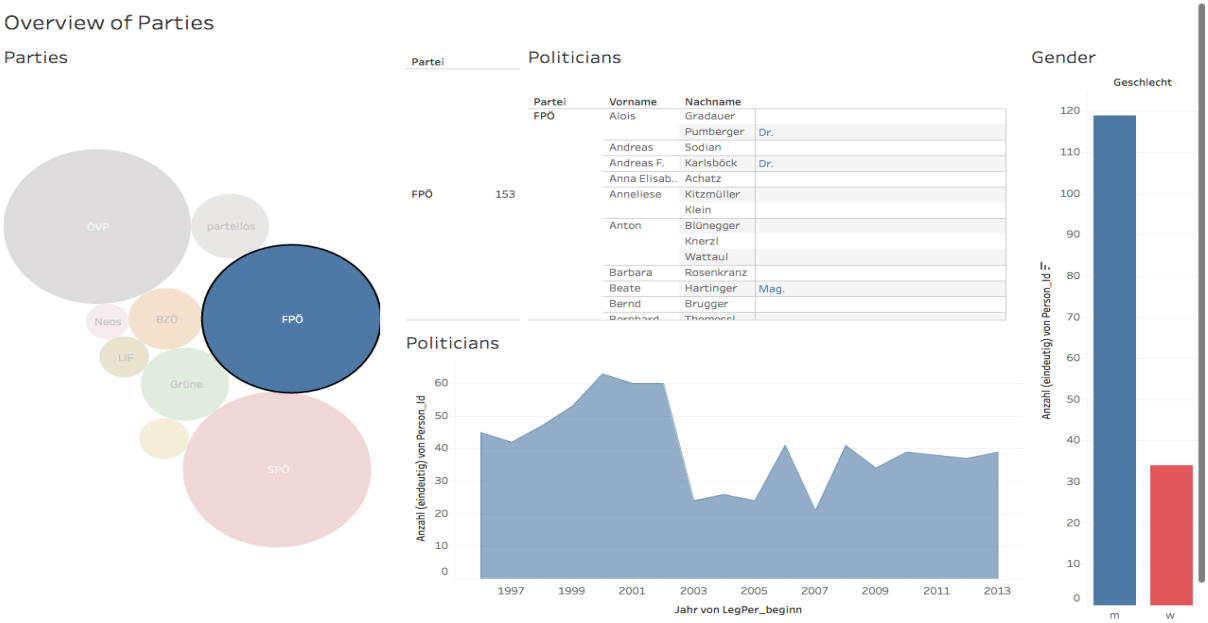
The biggest challenges were as first to get all the desired data from the texts. At the beginning of the project we had some great ideas but in the end we could not realize it because the data was missing. For example, to determine the seats of individual politicians and to visualize them just as in Parliament. Then go deeper into all speeches to determine which politician has talked about what topics etc. Unfortunately, that was not possible through NLP (Natural Language Processing). The second challenge was to create meaningful and well-fitting dashboards based on the sheets we created or could create. As first we created sheets so that all the data we had were used up. After that, we talked for a long time to define which sheet fits together and in what sense. We got the two dashboards. One with focus on statistics by politicians in parties and parties themselves and second with deeper overview of the speeches.

RESULTS

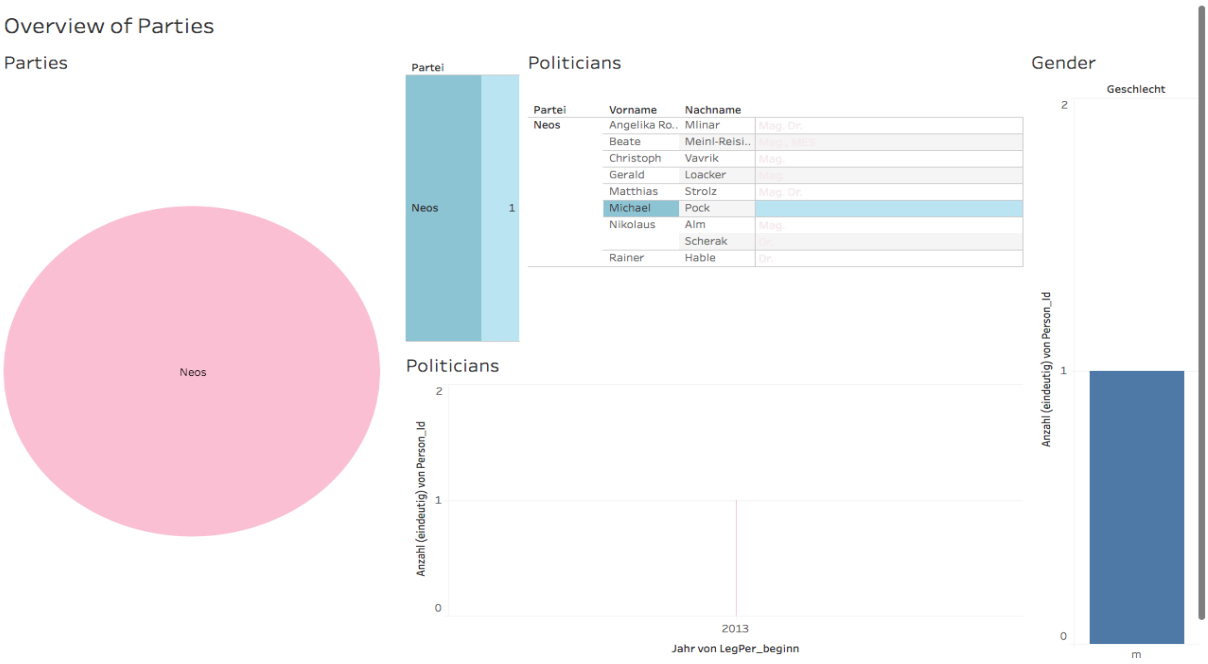


Picture1

As first you can see the Dashboard with overview of parties. On the main site as like on the *Picture1* you'll see the parties in bubble chart, Bigger bubbles mean bigger number of politicians in those party. In the tables you can see clear see the numbers and the names of politicians. On the bottom in the middle you see an area chart which represents the growth and decay of the numbers through the years.

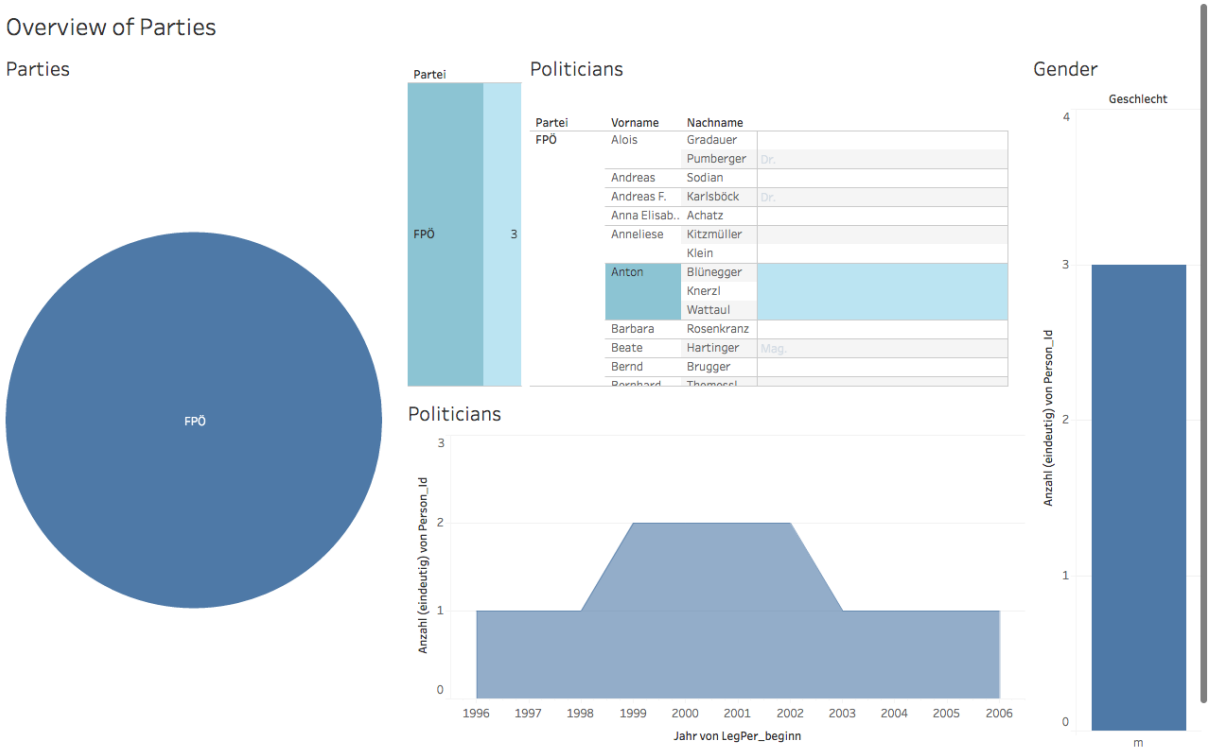


Picture2.



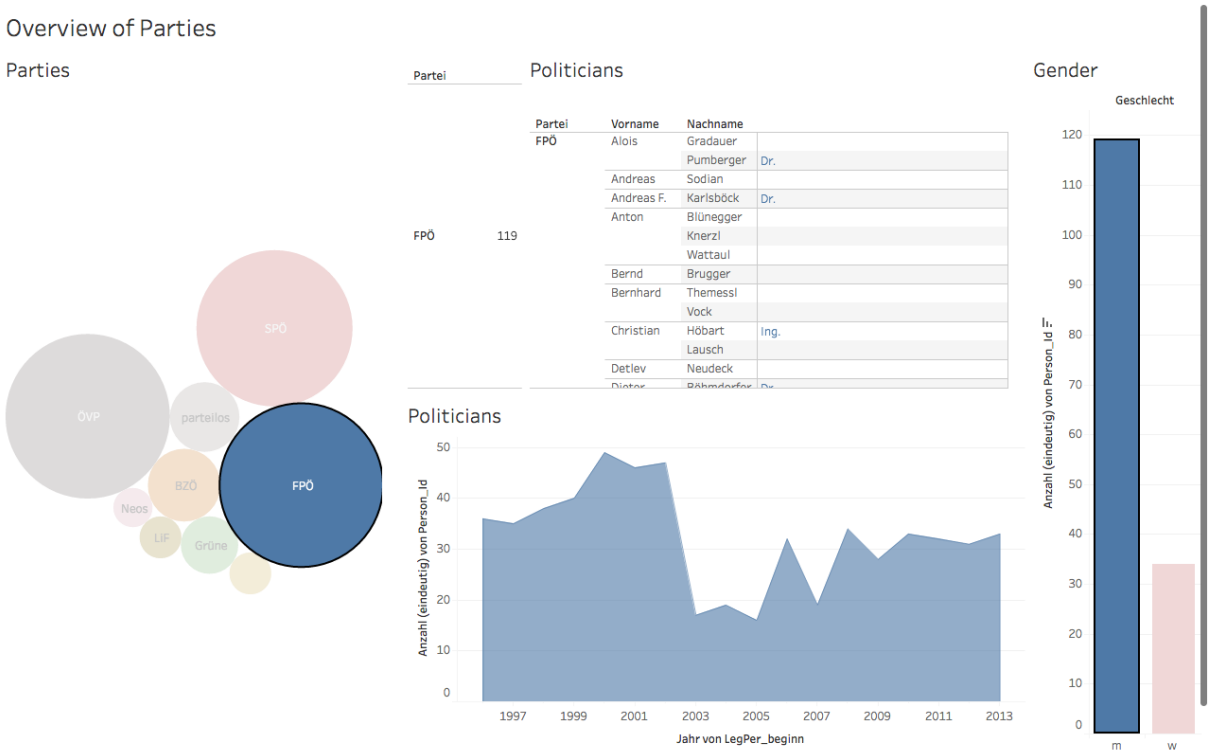
Picture3.

If you now click on the bubble/party or the party name in the table you want to research in detail, the visualisation will change. You'll only see the numbers, names, changing area and gender for the selected party (*Picture2 and Picture 3*).



Picture4.

If you want to know more about some politician you can choose him/her in the table and see the statistics as on *Picture 4*.

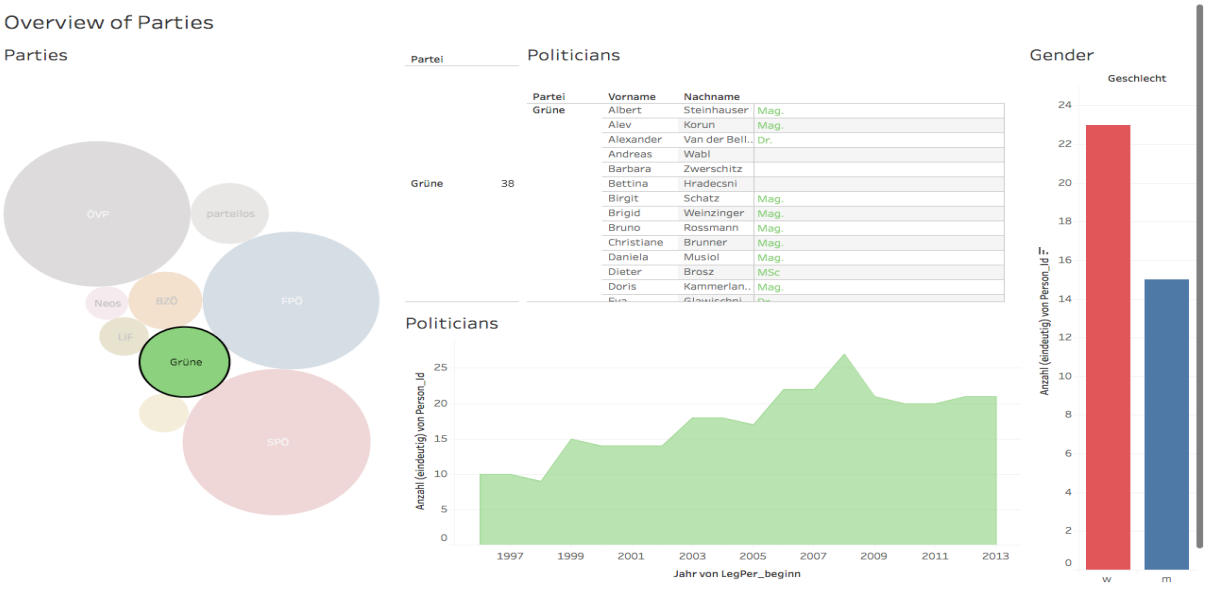


Picture5.

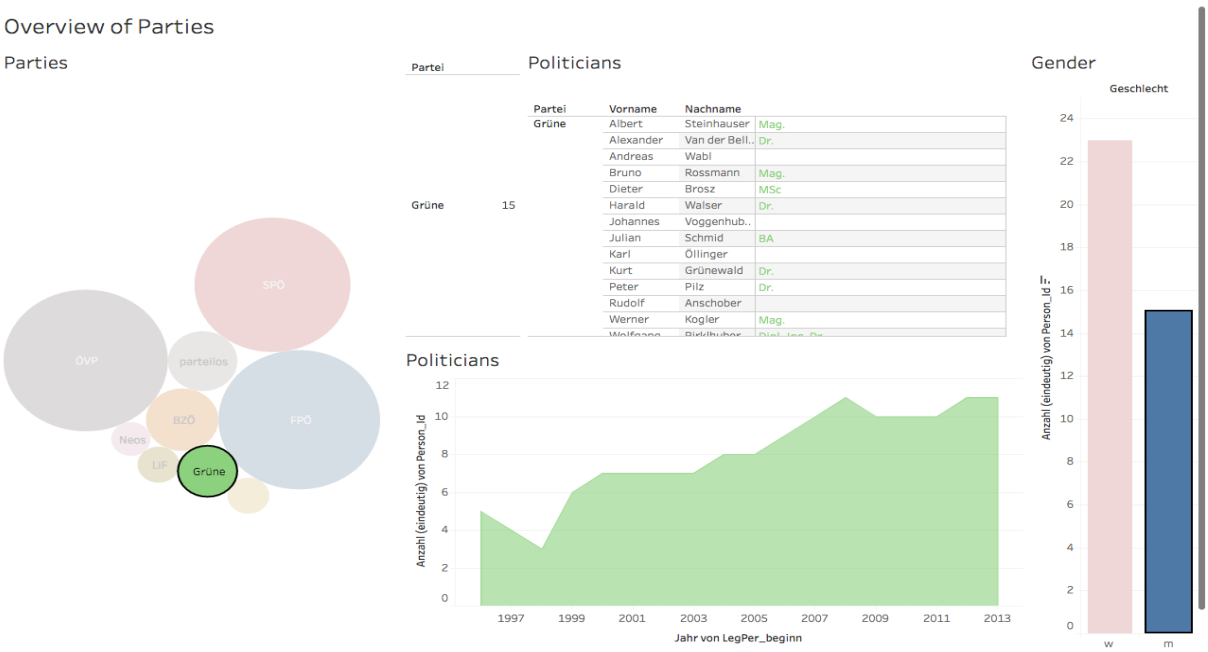
Next possibility is to choose the gender and go deeper into the statistics. As showed on the *Picture5* when you click on male gender the statistics will change and be

visualised.

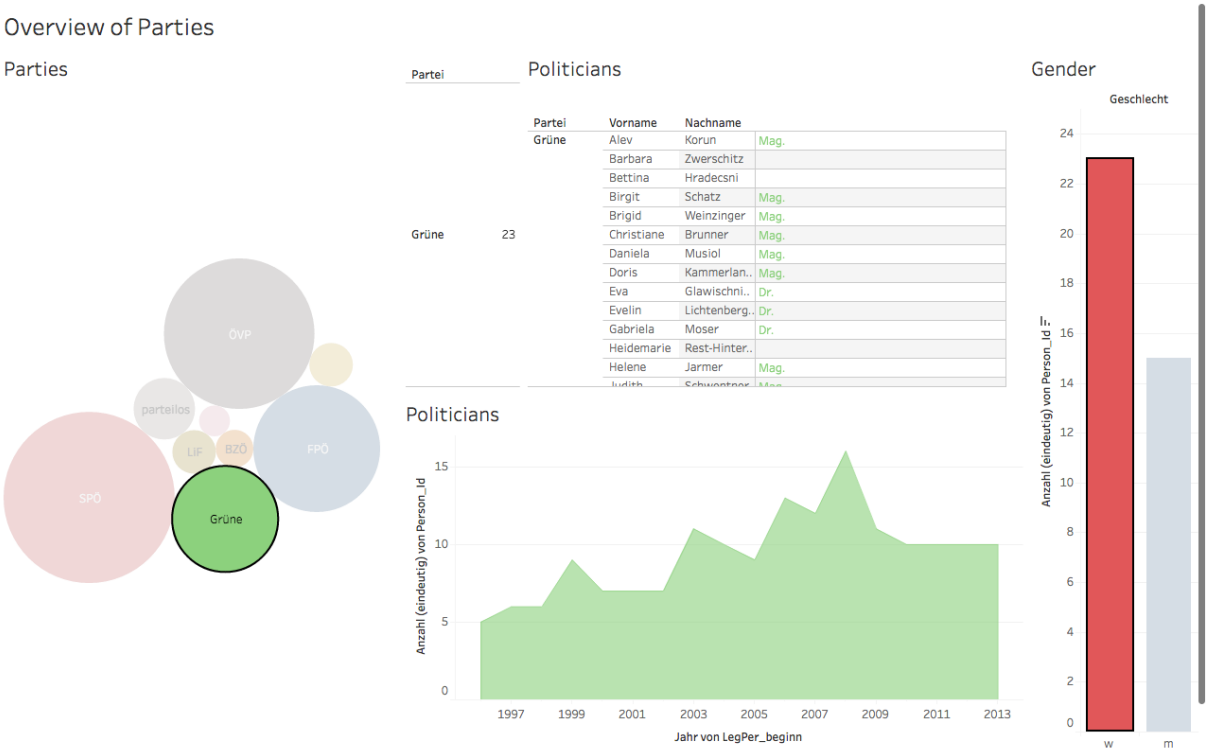
More examples for other party on *Picture6,Picture7,Picture8* and *Picture 9.*



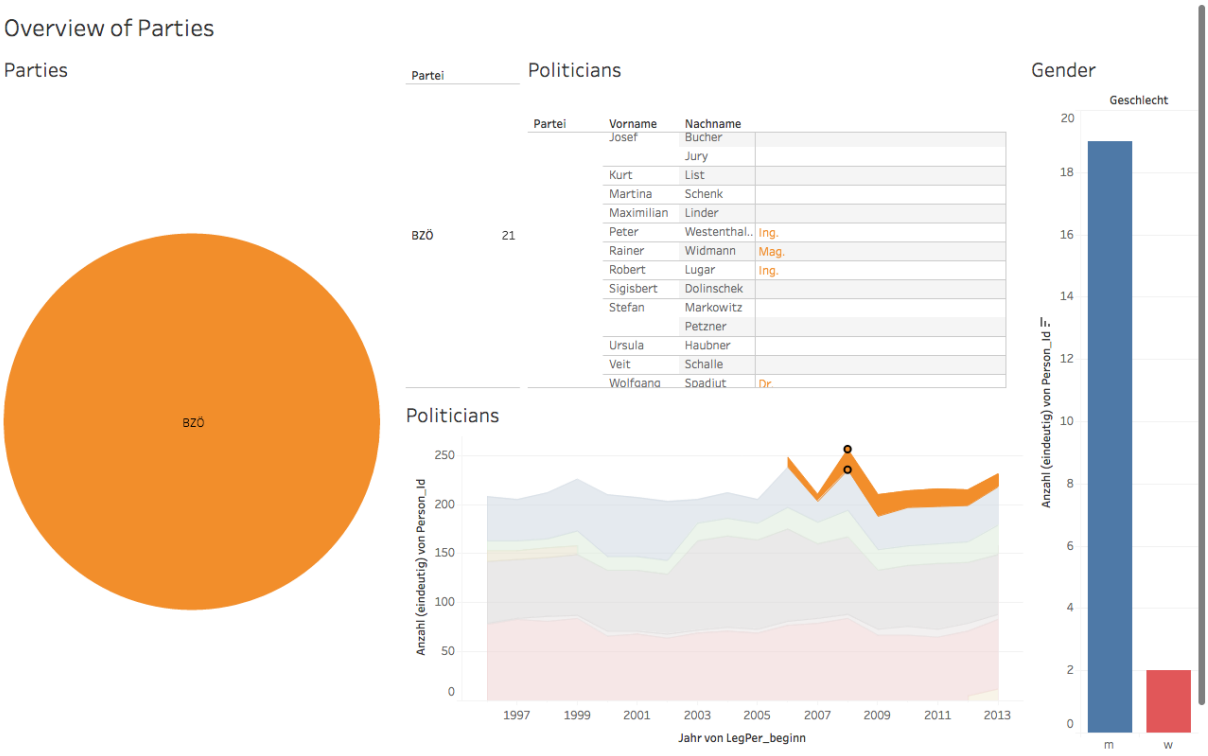
Picture6.



Picture7.

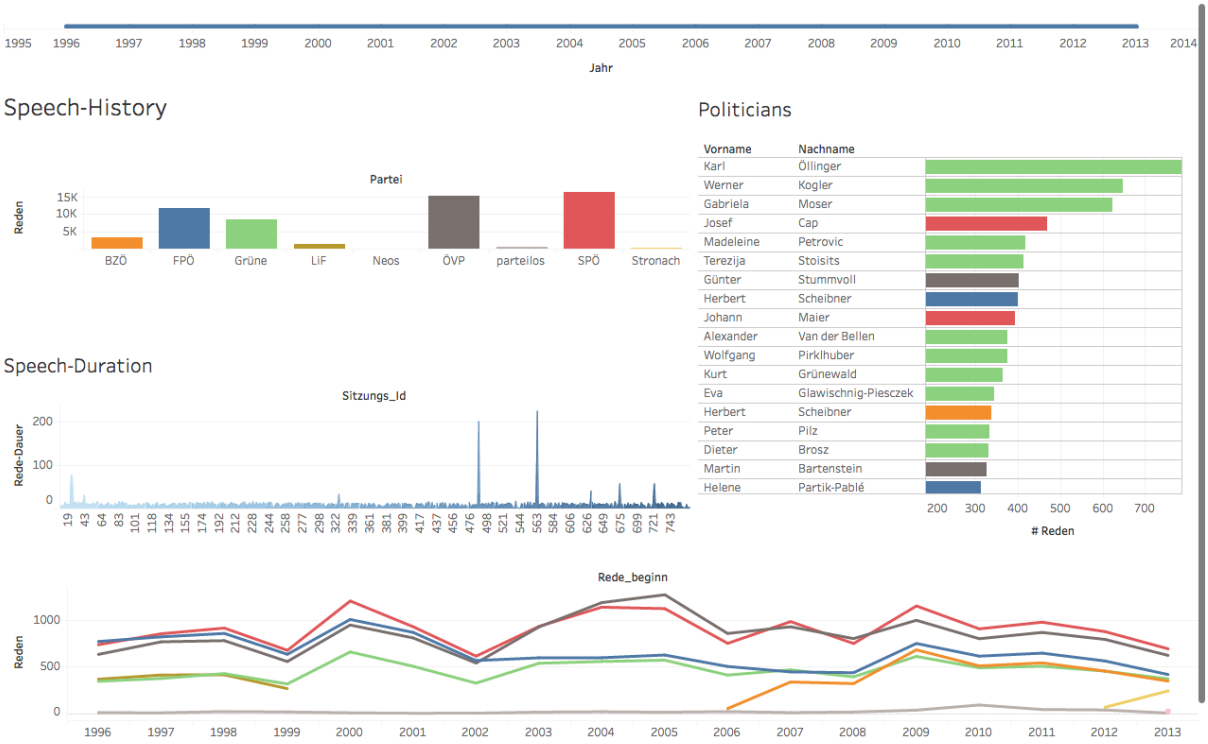


Picture8.



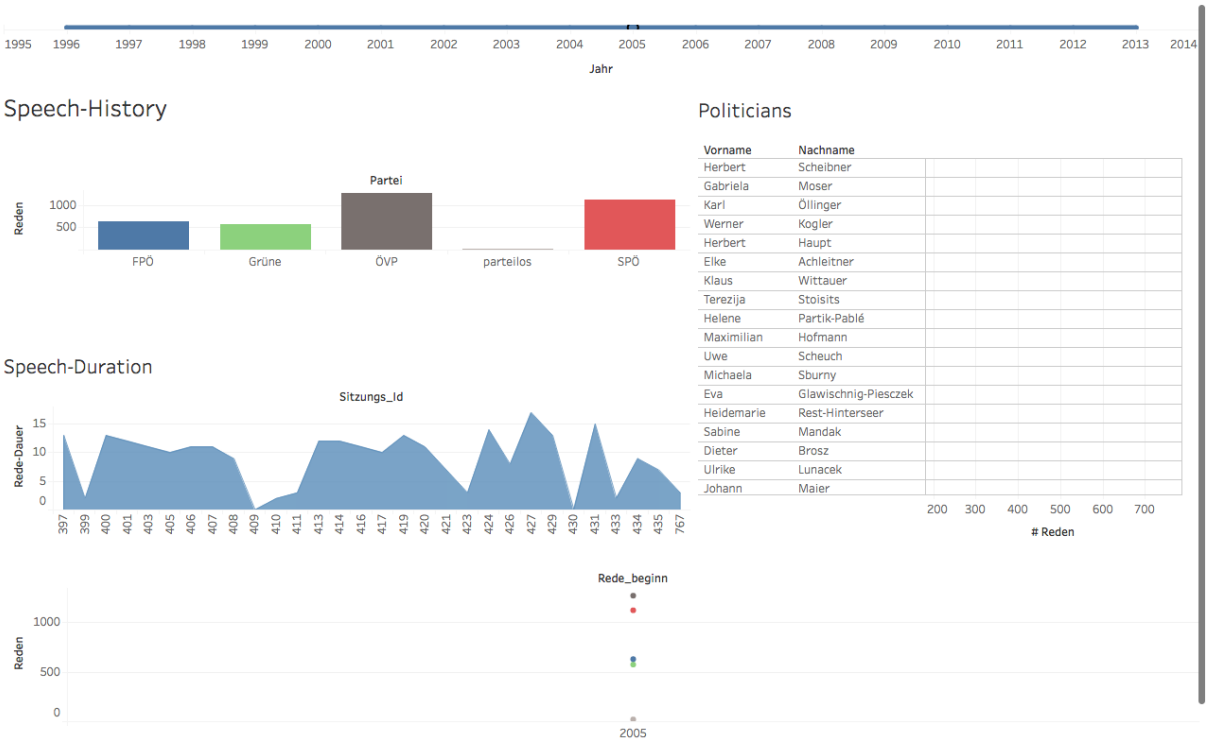
Picture9.

Our second Dashboard represents the Speech - History.



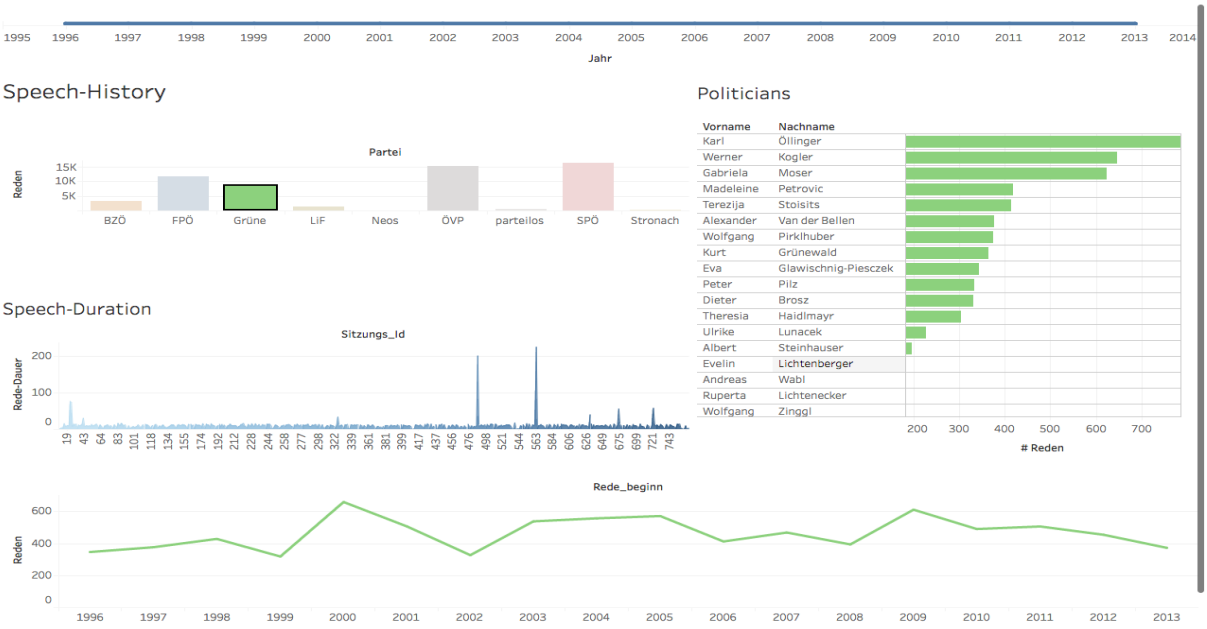
Picture10.

On the *Picture10* you can see the complete statistics made from the given speech data. On the barchart you can see and later choose the party, on the line chart the growth and decay of the number form speeches, in the area chart the statistics made by duration of speeches and in the table the names of politicians who held the speech and how many times. You can also navigate with a timeline on the top.



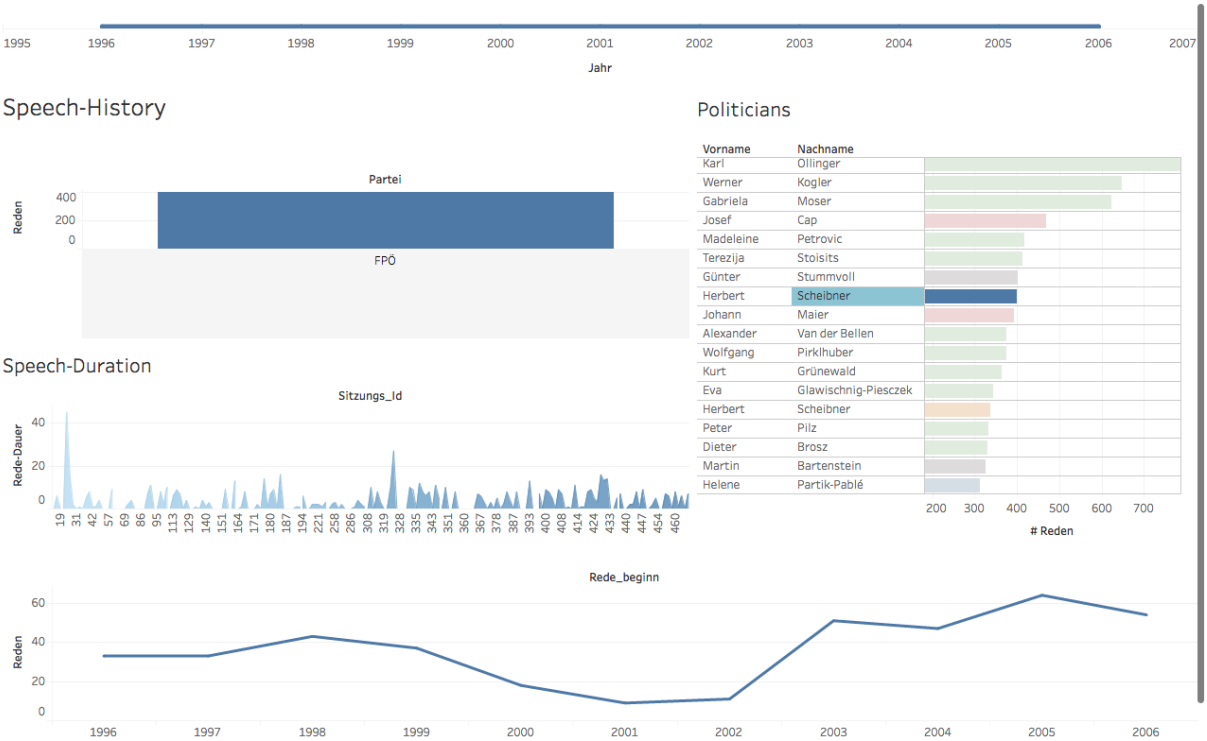
Picture11.

When you select the year the visualisation will change as on the *Picture11.* You'll only be able to see the already mentioned statistics for the selected year.



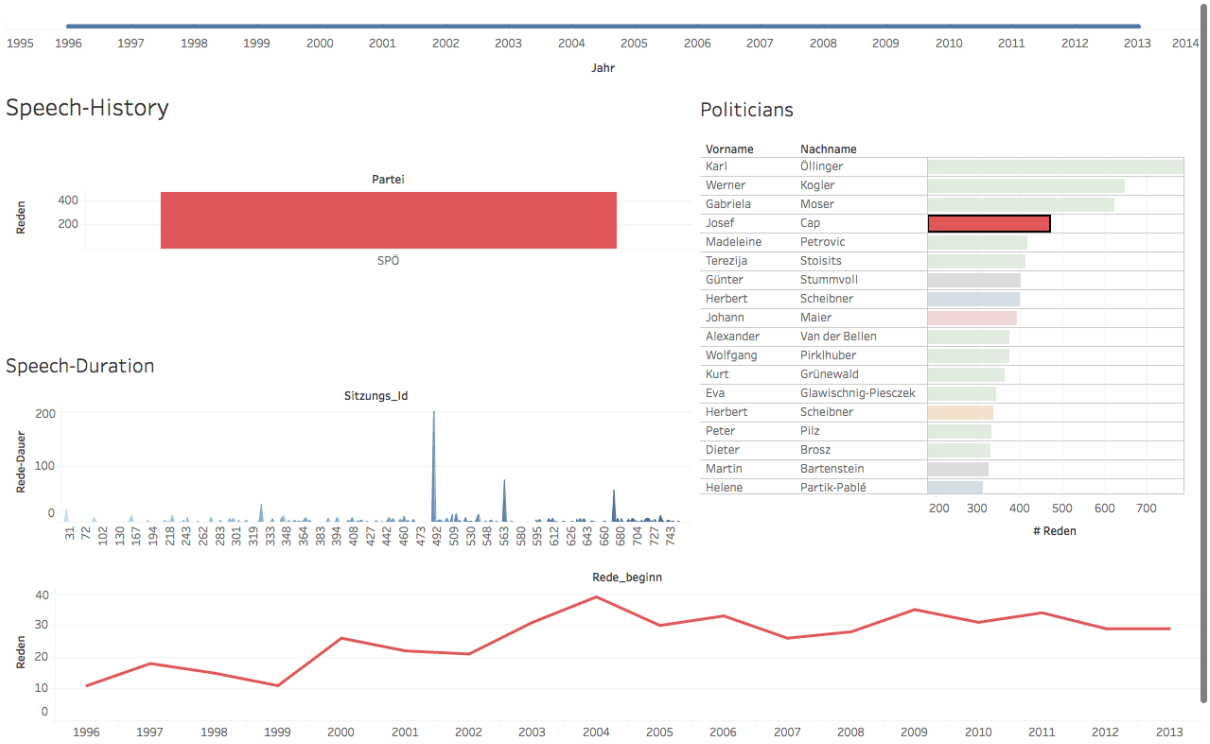
Picture12.

Next you can see what happens if you choose one party (*Picture12*).



Picture13.

On *Picture13* is a politician selected and his statistics are shown.



Picture14.

Picture14 shows what happens if you click on the speeches held by one politician. You see when he held the speeches and when.

Until now we have spoken about the functionality of the program and how we can use it. In order to understand its meaning a little bit better we will look now at a real life scenario of use. Maria hat a seminar that she has to write for the school about a specific political group (she is a political science student). She has sat down and wants to learn more about the SPO political group.

- Maria can click on the party at the bubble chart. With this action every other chart, including the table, will adjust itself and show the statistics for the clicked political group.
- Maria can see how big the number of members in a political groups is. She can also roughly compare the political groups on the bubble chart by their sizes and make decisions on which is smaller, bigger or roughly equal in comparison to the other ones.
- She can also see how the number of the members has changed throughout the years. The line graph shows how drastically the number has changed and which groups are very near to each other.
- On the right side she can see the the number of female and male politicians in a political group.
- On the table in the middle she can see names of all politicians that are have membership in the group that she clicked.

If she removes the filter by the political groups, the can see the names of all politicians, the overall number of females and males.

Performance of the system
We are very satisfied with the performance of our first Dashboard “Overview of parties”. The scrolling works on all of our computers fluently. Filtering the data and rendering the new Visualizations takes a few seconds, depending on the chosen filter. But this part of performance is bound to Tableau’s server, since they are computing the data before it gets rendered.

Our second Dashboard “Speech History” on the other hand, did not perform as good as the other Dashboard. This was definitely due to the high numbers of politicians in our bubble chart. Filtering the data took significantly longer than on the other Dashboard. We thought about limiting the amount of bubbles in the chart to make it faster and easier to read. At the same time, we did not want to hide non - redundant data. One way would be to exclude politicians, that were not that often participating, but then we lose some information, because every politician has to be selectable so their timeline can be analyzed. Another method would have been to just limit the numbers of bubbles. However, we did not find a way to do this exactly as we want, and at the same time we would also lose some information. In the end, we decided to not exclude some data and show the whole bubble chart, although it was a bit overblown and limited the system’s performance. Now we changed the bubble chart to a table, which shows the number of speeches held, due to professor Sedlmair’s feedback on our bubble chart. While this change not only made the Dashboard easier on the eyes, the system is now also faster at filtering the data and rendering the visualization.

As already mentioned above, we got some good, constructive feedback by Professor Sedlmair on our visualization. Especially our second Dashboard was criticized for causing confusion. So we made some changes to our implementation and removed the “pages” feature and also made a new table to show the highest number of speeches for each politician. To make it more readable, we sorted the table after the amount of held speeches. So the most important people are on top. We think this table is better readable than the bubble chart and we do not lose information.

DISCUSSION

- **Strengths and weaknesses of our approach and implementation**

  One of the strengths of our implementation is the simplicity of it. Every person (also without any previous knowledge about the field) can sit down and work with it. The interaction is simple to get to know with. The diagrams are charts that can be seen every day and which show the data in a non complicated

way. The data shown on the diagrams can be easily manipulated with the filters on the dashboards. The diagrams are responsive and you can click on a diagram in order to filter the other ones.This was easily achievable with Tableau, since we could focus on the design of our visualization without having to care about technical details. We did not have to know how it is rendered, or when which graph is drawn. Also Tableau offers a public server, which can easily process all the data we have. With Tableau we also could not make unconventional visualizations that no one understands. Unfortunately, this also is one of our implementations weaknesses.

A weakness of our approach is, that we are bound to the limits of Tableau. We have a limited influence on the system's performance and its general design philosophy. Without tableau we could have made more optimizations on performance and make the system our own. At the same time it would have been more prone to errors, which results in a higher investment on the technical perspective of a system. Also we are not able to use unconventional designs for our graphs. Technical details were not our focus, so we decided to use Tableau for the implementation, so we could focus on the design.

- **Lessons we learned**

  By an implementation of a visualisation is the clarity and readability of the program very important. In order for the users to use a product (efficiently), they how to understand and see clear the functions of it. The more the users give time into learning the functionality of an product, the less efficient and popular it becomes.

  The data should be well prepared for the use and also the redundant information should be removed in order to process the data faster and better. Concentrating on needless information costs time as not preparing the data.

  All prototypes and models of the product should be discussed with some future users from the goal group in order to get valuable feedback. The users and the software engineers often do not look at things the same way. Therefore feedback from the users can help improve the product, make customers more satisfied etc.

  By a software creation, in the focus should be the needs of the customers not only to create a solution of a problem. Sometimes the solution of the problem and the needs of users are not the same thing. The solution can be way more complicated than a customer needs and therefore more inefficient. Concentration on the customer needs and modifying it so that they support them, not make thing more difficult, will result in a creation of a useful future product.

Milestone Plan - tasks separation between the group members

TASK	INES BACIC	CHRISTIAN LASCSAK	TARIK MESANOVIC	UMAIR ALTAF
Motivation				
Related Work				
Approach				
Implementation				
Results				
Discussion				
M4 Dokument				
M4 Präsentation				
Homepage				