Data to drive

Personalized visualization in Formula One racing

den Hartog, Iris H.D.

Award date:
2022

Awarding institution:
Tilburg University

Data to drive: Personalized visualization in Formula One racing

I.H.D. den Hartog*
Student number: 1338277

Eindhoven University of Technology & Tilburg University

ABSTRACT

The fast paced world of Formula One depends on data to make quick decisions that could mean the difference in becoming world champion or falling of the podium. Personalized visualizations can be the helping hand in showing the data in an appropriate way to be able to make those decisions faster and easier. As personalized visualizations are an upcoming field, this paper contributes by creating a concept idea of a graphical user interface that distinguishes three target groups within a Formula One team; the driver, the data analyst and the team principal. Each have their own assets that make them valuable to the team and unique as a user. With this paper a model could be created that implements those differences and adapts the visualizations to the correct user group. So you have a better shot at becoming, Formula One world champion.

Index Terms: Personalized visualization, Formula One racing, Dash, GUI, Python

1 INTRODUCTION

Visualizations hold a lot of information, especially in the data driven field of racing. A Formula One team depends on data to make the correct decisions regarding set-up, strategy and driver choices instantly. For example, data about speed on different types of tyres is used to determine which tyre is best to start the race on. But during the race decisions need to be made about which tyres to switch to, these decisions need to be made in milliseconds. Nevertheless, not every team member is interested in the same data. An engineer of one driver might not be interested in the data of the other driver, while the data specialist would want to compare the two. Therefore, it is useful to have an interactive system that shows the people the data and the graphics they need to fulfill their respective function. The challenge of this project will be, how to distinguish between the different roles and adapt the visualizations to their respective needs. This is the main challenge as not many of these personalized systems already exist and none are focused on this specific field. The focus of this project will be on three different team members; the driver, the data analyst and the team principal. (*How many people work in an f1 team?*, 2020) Another challenge will be determining the specific needs of these different target groups.
"... user interests, intent, and visualization preferences are likely to be fundamentally different, yet vitally important." (Qian et al., 2021, p.1) This is the basis of personalized visualization. Visualizations are only useful if they appeal to the target audience. Showing a difficult plot to a toddler is almost as useless as showing a very simple plot to a data scientist. Thus, this field of research is crucial to get the right information across to the audience.
The field of Formula One racing was chosen as the main focus, in the first place because of personal interest in sports, and Formula One in particular. Secondly, Formula One is very open about the use of data. Every race weekend, the commentators but also the teams stun the audience with great visualizations and complex strategies created with all the data they gathered during practise sessions and on the simulator. Even though teams are hesitant to share their own data, they don't want to make the competition wiser than they already are, a great deal of information can be found on the internet although it be a lot simpler data. Data about race wins, pit stops and penalties are all up for grabs as everybody has been able to see it on TV already. This interest combined with the ability to easily obtain and use all this data lead to the decision to focus on this field for this project.
As stated before, there is an abundance of data to work with. However, this is mostly rather simple data; championship points, number of pit stops, and lap times. Therefor, the focus will be on more simple visualizations such as bar charts or line plots. Another reason for this choice is the lack of experts to check the research. There are no visualization experts or actual Formula One team members within reach to test or ask about the tool, thus the choice was made to keep it easy to give other people the chance to test and give their honest and valuable opinion.
At first the plan was to build a working Graphic User Interface (GUI) using a rule based model with plain rules. A minimum viable product was created but didn't meet the standards in terms of personalized visualizations, there was too much set in stone for it to be personalized. So, a new plan was thought out and because of a lack of time only a concept was made for the end product. In this concept an idea is introduced on how to automatically create personalized visualizations.
This paper is set up in the following way; in section 2 the research questions leading this paper will be discussed. Section 3 is about previous research done on this topic including research on personalized visualizations on its own and on Formula One data specifically. All the details on how and why the research was done and what the target audience entails is discussed in section 4. Furthermore, section 5 describes the data that has been used and how the different graphs were created. Visual design, section 6, explains the choices made regarding the GUI and the following concept design. The results of the evaluation will be given in section 7. Some methods used require a bit more explanation and the results need to be discussed, this is done in section 8. Lastly, section 9 draws a conclusion and recaps what has been done in this paper and what could be done in the future to improve the results.

2 RESEARCH QUESTION

The following research questions and sub questions are leading in this process of building a GUI, and later on a concept, with a rule based model that presents and adapts personalized visualizations based on the user. The goal of this research is to build a tool that is effective in differentiating between the different target groups and automatically adapting the visualizations shown accordingly. The tool is supposed to learn to separate the multiple users and adapt the graphs to the needs of those users.

*e-mail: i.h.d.d.hartog@student.tue.nl

- How can a rule based model in combination with a visualization tool best distinguish different types of users?
- What is the optimal model threshold setting for showing extra (or less) detail to a user?
  - What differentiates the different users?
  - How can those differences be translated in a model?

3 Related work

People react slower and less accurate when presented with information that contains either too much or too little detail. (Oscar, Mejía, Metoyer, & Hooker, 2017) In a high velocity field like Formula One racing every second and millimeter counts so determining the amount of detail an individual needs is crucial. This can be done by using different view operations which hide or highlight parts that the user needs or does not. (Reichert, Kolb, Bobrik, & Bauer, 2012) Highlighting different parts of screen is a very effective way to direct the persons attention to that part. However, the system by Reichert et all (2012) makes it possible for the viewer themselves to build a personalized visualization, whereas the focus for this paper is on systems that make the visualizations personalized for the reader. An example of what could be interesting for a user to see is, for a driver it could be motivational to see a different driver going faster and the differences in their driving that lead to that. This example shows how visualizations could be used for personal learning outcomes as they can motivate and encourage users. (Hsiao et al., 2012)
Currently, there are studies that have focused on personalized visualization but they are sparse. Moreover, Oscar et al (2017) have shown the need for personalized visualization but have not yet proven how to accommodate that need. A similar situation is presented in the book by Ottley (2020), they go even a step further by presenting a way in which the user interaction can be detected and studied but as they say themselves "This work is a first step in learning about users live from their interactions... but an adaptive system could test the feasibility of applying this type of results in real time." (p.62) so a lot of work is still to be done. In terms of tracking the user interactions there are a few options; mouse clicks, mouse movements, the time between the mouse movements and clicks, and the speed of the mouse movements. Of course, there are far more complicated things to track such as eye movements but it is harder for a system to do that and adapt itself, so the focus for this study will be on the mouse interactions. The paper that comes the closest in terms of what this paper is trying to achieve is the one of Qian et al (2021) even though it is not focused on the niche market of Formula One. They created a framework that recommends visualizations based on past visualization interactions such as viewing, clicking on or creating a plot. It also adapts the next visualizations based on the graphs the previous interactions were with. The models that were used in this project could also be used in specific markets like this one after being retrained on the specific user needs and datasets in this field.
There is a paper that uses Formula One data in visualizations, Spenke & Beilken (2003) uses a tool by the name of InfoZoom to answer questions about the data. By clicking on buttons and text within this tool, different features of the data are highlighted to enable the answering of the questions. The tool shows database relations in a table format. A nice tool for displaying the data but not automatically personalized unfortunately.

4 Methodology

To answer the research questions and their sub questions first data has been found to start creating the visuals. Also, more literature research was done to better understand the research field and how interactive visualizations and GUI's work. Next the data and acquired knowledge have been used to create a GUI. The different user groups are described below, answering the sub question on how they can be differentiated. The data has been transformed into graphs using the different types of data, creating visuals that can be used for different purposes by different people. Within this set-up different model variables are tested to create personalized visual patterns. Different patterns are set up to belong to specific target groups, to translate the user differences into the model. The plan was to use simple rule based models first, for example a click on a specific plot means show more detail of that plot. After the first rules are implemented they need to be evaluated. This can be done by letting multiple people work with the GUI, keeping in mind that they represent a certain role. If the target group clicks on the plot that is designed for them and they are satisfied with the following plot then the intentions were right. This way it can be tested if the right visuals show up for the intended group. Along the way this proved not to be personalized enough. Also, it should be assessed if the research questions are answered. Then, depending on the outcomes of the different analyses, the process of designing and testing different model paths starts again. In this way the model is updated along the way, keeping the goals in mind and evaluating between every iteration to see what the optimal set-up is. This evaluation is done by interviewing different people while guiding them through the dashboard. Because of the current situation around covid-19 doing a user study is challenging as it is hard to show the dashboard through online resources and study people's interaction therefor the decision was made to keep the group small and personal. This unfortunately lead to the fact that the group is not very diverse in demographics; five out of six being men and two thirds being over sixty years old. Because of these issues no hard conclusions can be drawn from this evaluation but it is still used as advisory information and the comments made are taken into account. Furthermore, everyone in the user study was assumed to have the same basic knowledge of Formula One racing, the information they received beforehand was the same as well as the questions. The only difference was the role; driver, data analyst or team principal, they had to keep in mind when answering the questions. The questions and their answers are described in section 7.
The GUI is build using Dash because it is an easy framework to use with real time updates following from user interactions. (Plotly, 2017) User interactions will be tracked in the form of clicks. After clicking on a certain plot the user will be directed to another. In this way a specific user will be led into a sort of tunnel of information that is assumed to be of use to them. Of course, there is always the possibility of coming back to the homepage to be able to see other things.
After building this dashboard and evaluating it, it was deemed not personalized enough. Due to lack of time, the decision was made not to rebuild a whole dashboard but simply a concept of one. In terms of looks not much would change, the main difference is the way in which the user groups will be recognised by the system. The dashboard should be able to track the user's mouse movements and clicks. Below per user group will be described what kinds of user interaction would be assumed to be typical for that group and how that would result in different visualizations.

4.1 User groups

Only three different target groups are studied to answer the posed questions. This is to keep the project manageable in the first stage and adding target groups is always possible in the future. The chosen subjects are; the driver, the data analyst and the team principal. These groups are chosen because they make up a great deal of a Formula One team. Furthermore they have different goals, want to see different types of visualizations and they need to see different data.

4.1.1 Driver

Even though Formula One racing is often described as a team sport; one of the most, if not the most, important part of the team are the drivers. Without them no races are won, no points are gained and thus no money comes in. The drivers are assumed to have a fair amount of knowledge of the car and the data that results from driving it. They benefit from seeing specific data of their own driving so they can see what works for them and how they can improve themselves. For example, lap times over a race can show them how the car performs with different tyre types on full fuel in race conditions (Figure 1).
They want to analyse their own data thoroughly but in between practises there is not a lot of time so they need to move to the right plot quickly. Therefor, it is assumed that drivers do not click a lot on the dashboard but do move their mouse quickly to the things they want to see.

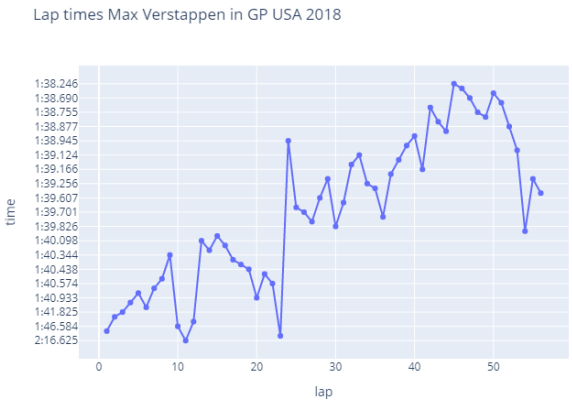

Figure 1: Lap times over a race

4.1.2 Data analyst

The data analyst of the team is the most specialised in reading and understanding the data. They are assumed to be capable of understanding all types of data and interpret them. Furthermore, they are able to compare data of different drivers and point out where one performed better than the other as well as make strategic decisions for the team. For example, comparing track positions over a race can improve strategic decisions regarding pit stops and tyre choices (Figure 2).
As the data analyst is the most at home in the data they know what they are looking for, so they know where to click and what to do. Therefor, the assumption is made that they move their mouse quickly, maybe even quicker than the drivers. Also the assumption is made that they make a lot of clicks as they are interested in multiple factors.

4.1.3 Team principal

Every Formula One racing team is led by a team principal. Ultimately they decide what happens within the team. They are assumed to have a bit of knowledge of the data but not as in depth as the drivers and the data analyst. As they are the team lead they are most interested in how the team is performing in comparison to the other teams and how they can improve the team's result. For example, they might be interested in how their team is doing in terms of overall wins compared to other teams (Figure 3).
The team principal is assumed not to be as used to working with the data as the other two groups are, so they take their time moving the

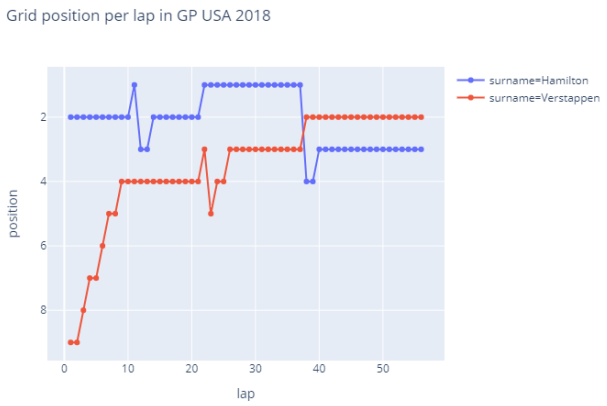

Figure 2: Grid positions in a race

mouse to what seems interesting. In terms of clicks it is assumed that they only click a couple of times or at least there is a lot of time in between their clicks.

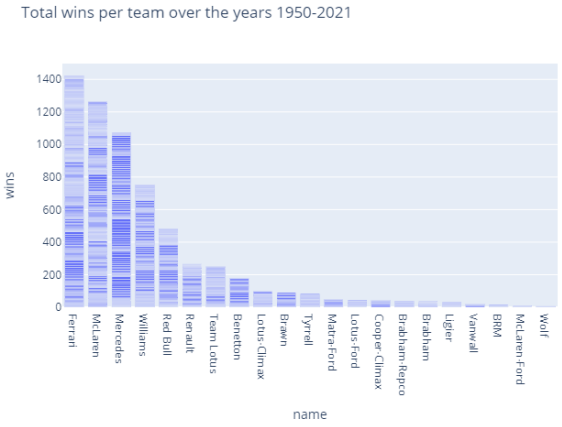

Figure 3: Total wins per team

4.1.4 Comparison

To answer the sub question "What differentiates the different users?", they are compared. Firstly, their assumed knowledge of data differs. Sorting them from least to most knowledge: the team principal, the driver and the data analyst. Secondly, their interests regarding different types of data are different. A driver might not be interested in how the team is doing in comparison to other teams, they care about their own performance and progress. Similarly, the team principal is not interested in specific sector times of one driver, they want to see the bigger picture. This shows that the types of data they are interested in and the level of detail they want and can understand differentiates the multiple users.
In terms of clicks and mouse movements the users can be separated by amount of clicks and the speed with which they move their mouse. The driver and team principal are assumed to make a small amount of clicks, the driver because they need to make fast decisions and know what they want to see. The team principal for opposite reasons, they do not know what they can expect with the exception that if they do click multiple times it is with a lot of time in between as

they study what they click first. So, the data analyst is the only user to be assumed to make a lot of clicks initially as they want to see as much of the data as possible. Then looking at mouse movement speed, the driver and data analyst know where to go so they move their mouses quickly across the dashboard while the team principal again takes time.

5 Data & Analysis

The data used consists of multiple datasets found on Kaggle. (*Kaggle: Your Machine Learning and Data Science Community*, n.d.) These datasets are all open for public use so there is no issue downloading and storing them, the only personal information they contain are the names and birth dates of the drivers. This does not pose an issue as it is open data that is already well known to the world. These datasets contain different variables regarding circuits, drivers, teams, results and timestamps. Using these different types of data creates opportunities to give insights useful to multiple users. For example, the team principal would want to compare how their team is doing in comparison to other teams in terms of results. The driver is more interested in how they are doing themselves in specific sectors and on similar circuits. The data analyst can compare the in-depth data of the two drivers to get the most out of the team. In this way the different roles within a team are interested in different parts of the data and contrasting levels of detail. The main dataset that is used contains data about the Formula One world championships from 1950 up to and including the latest (2021) season. (Rao, 2020) It consists of 13 tables with in total 75 different variables. The dataset has variables that are not of use for this particular case such as; URL links to different driver and circuit pages and latitude and longitude coordinates of the circuits. Therefor these are filtered out. To create the multiple graphs, different tables of the dataset are combined using the same variables present in multiple tables (Figure 8). The tables concerning the same subject are combined; constructor, constructorResults and constructorStandings. As well as; lapTimes and drivers. With the multiple combinations, the graphs used in the GUI have been created.

6 Visual design

The GUI consists of a main dashboard including three figures (Figure 4). These figures differ in their intended audience, figure 3, top one on the page, is designed to inform the team principal of the wins of their team compared to other teams. Figure 2, in the middle, compares grid positions of 2 rival drivers, a source of information to the data analyst. The third figure, number 1, is supposed to inform the respective driver of their own lap times. Due to these different target groups, depending on what plot is clicked a next one appears. What plot appears depends on the targeted group with the initial plot.
In creating this concept idea to make the dashboard more personalized, the same layout with the three plots is used simply because the different groups are interested in different things. The difference will be in the way in which the next plots are shown. The first user tests showed that not every user group would click on the graph that was thought to be of interest to them. So instead of only depending on what plot they click also the speed of the mouse movements and the amount of clicks will be taken into account. The next change will be in how the next plots are shown. Instead of having a pop up of the next plot, a new page will be loaded with more plots depending on the different measured factors. Still the most important variable that is taken into account is the first plot they clicked on followed by how fast and how many times they click on the next plots. This adapts the visualizations that are shown.

Figure 4: GUI dashboard

6.1 Team principal

Figure 3 contains the total wins per team over the years 1950 up to and including 2021. When a specific team is clicked another plot appears, for example figure 5 appears when Red Bull is clicked. This new graph shows the wins per year and on what track they are achieved by hovering over the bars.
For the team principal they are not only expected to be interested in the first plot as they might be interested in the lap times as well. However, they are assumed to move the mouse at a slower speed and only click a couple of times on the exact plot they want to see. So after slowly clicking the initial plot, a next page would load with more interesting plots for them.

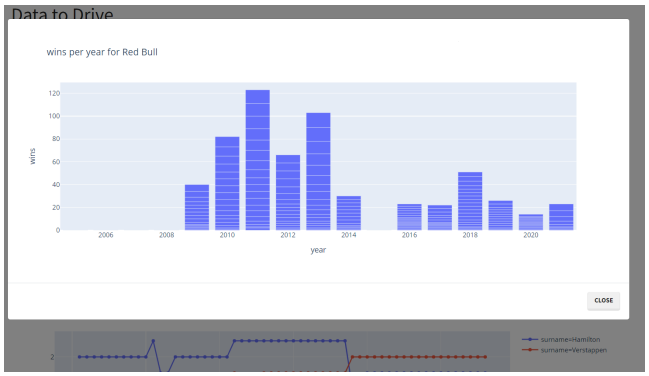

Figure 5: Click first plot on Red Bull

6.2 Data analyst

When a user clicks on plot 2, a more detailed graph appears containing the time it took to race a lap over a whole race for 2 rival drivers (Figure 6).
The data analysts in the user study indeed all clicked on the second plot but for them also speed and clicks should be taken into account. As assumed they move their mouses rather quickly as they know what to click on and in the next pages they click on loads of different things as they are interested in multiple factors.

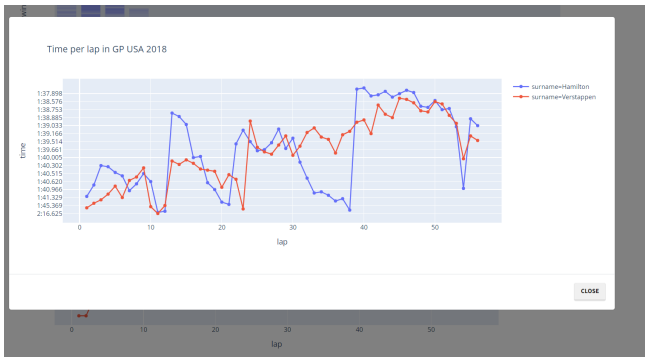

Figure 6: Laptimes over a race

6.3 Driver

The last plot (1) showing the lap times leads to a graph that shows the best sector times, figure 7. Each Formula One track is divided into three sectors. If the three separate best sector times together are faster than the fastest lap of that driver, it means that there is room left for improvement.
The drivers met the expectations in the small user study as well. Still the other variables weigh in, in the decision on how the next page with visualizations is going to look. Assuming their mouse speed is quicker than that of the team principal but not as fast as the data scientist's. The clicks they perform are expected not to be of a high quantity on the plots that are of the most interest to them. So the next page that shows up, assuming the driver is the user, contains multiple plots of that specific driver's statistics and other interests.

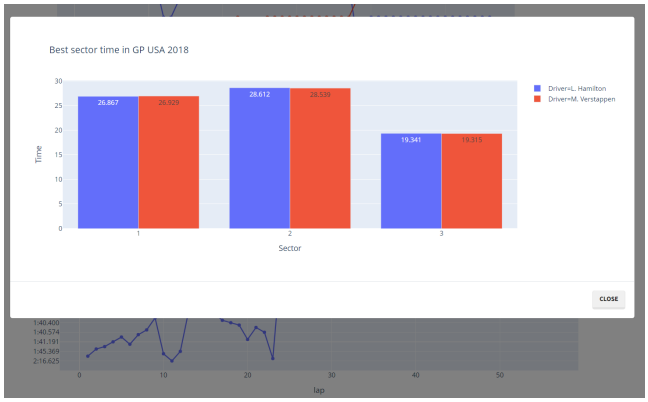

Figure 7: Best sector times

7 Results

Within this GUI three paths are created for the three different user groups. These paths are created manually as there was too little time to build and train a machine learning model. Different people were asked to portray a certain user group. They were informed on what role they fulfilled and what the purpose is of that user within a Formula One team. The questions that were asked and the answers that were given are stated below.

7.1 User study

- Question 1: A dashboard is being created on which the different team members can see all kinds of data and statistics regarding the races and the car. What do you want to see on such a dashboard, in terms of data?
  - Driver 1: I would mostly like to see my own speed and performance, but also the comparison to other drivers to see if they are faster in certain sectors, high- and low speed corners.
  - Driver 2: Your own performance with respect to your rivals, to see where you can improve. How the engine is working and how it reacts to your driving style, the same for the tyres and the steering wheel. To see if your reaction capacity is good enough.
  - Data analyst 1: I would want to see the lap and sector times, so you can see who is faster where. Maybe add how hard they brake and if they brake earlier to see if that can improve.
  - Data analyst 2: I would like to compare races, see the weather forecast and how it affects tyres. Compare how other teams perform on different tracks.
  - Team principal 1: What I would like to see is grid positions, championship results and how the team is doing. What the best pit stop strategy is and how to get an advantage out of that.
  - Team principal 2: Top speed, if everything about the car works as it should and how the competition is doing.
- Question 2: What do you think of this dashboard?
  - Driver 1: It is a clear overview.
  - Driver 2: At first you only see your own lap times and no comparison. Maybe a comparison with last year so you see your own improvements would be good.
  - Data analyst 1: Not really impressive, just three plots below one another.
  - Data analyst 2: It is hard to understand the first plot, others are good.
  - Team principal 1: It could use more color to highlight the differences. Needs a larger font. Would be nice to immediately see the differences between driver times.
  - Team principal 2: Looks good, clear.
- Question 3: Which plot would you click on first?
  - Driver 1: Figure 1
  - Driver 2: Figure 1 as that is the most relevant for me.
  - Data analyst 1: Figure 2
  - Data analyst 2: Figure 2
  - Team principal 1: Figure 1
  - Team principal 2: Figure 3

- Question 4: What do you think of the plot that follows?
  - Driver 1: There are very small differences between the two drivers and they are hard to see, except when you look at the actual numbers you see the difference. So they should be put into perspective, by changing the range of the axis for example.
  - Driver 2: Good to see the difference between you and your rival, so you can see where there is room for improvement.
  - Data analyst 1: Very clear, you see exactly the times per lap.
  - Data analyst 2: It makes you look at the outliers first and then at which points the difference is very small and why. So that is very good.
  - Team principal 1: (After clicking Mercedes in the first plot). Nice for the books but we don't live in the past so not necessary. Lap times are more important to me than wins.
  - Team principal 2: (After clicking Mercedes in the first plot). Very clear, important for the team principal to see at what circuits they performed well.
- Question 5: Any other comments?
  - Driver 1: It would be nice to see at which point in the round the difference is made between drivers. So at what corner one performs better than the other.
  - Driver 2: You should make it easier to see in one quick look if you are faster than your opponents and where exactly on track.
  - Data analyst 1: No.
  - Data analyst 2: Nice that there are lines in this plot (figure 6) to make it easier to compare. Also good that it gives more information when you hover.
  - Team principal 1: It would be nice to have all the drivers of the top four teams in one graph. Very busy but as it depends on small differences, you can draw conclusions fast.
  - Team principal 2: No, it all looks good.

8 Discussion

After first building a dashboard with three different paths for the different target groups and then letting some users test it, the decision was made to go in a slightly different direction. The answers that were given in the user study contributed to the concept idea that was created afterwards.
For the driver the choice was made to show their lap times and best sector times, because the assumption was made that drivers are mostly interested in their own race. After the user study it became clear that drivers are indeed mostly interested in their own results but also appreciate the comparison to other teams and drivers especially seeing where on the track they gain the most. So for the concept it is important that the driver group gets to see more comparisons and a clear overview of the speed in all places on the track.
The data analyst got to see the grid positions of two drivers over a whole race and the lap times of those two drivers over the same race. This was done because the data analyst is expected to compare the two drivers in a team and explain the differences. Overall the data analysts were satisfied with the plots they got to see but they wanted more detail. In the concept more detail needs to be shown in terms of specific times in every stage of the lap, more information on braking and the weather forecast. These details could not have been implemented in this dashboard as the necessary information was not present in the dataset currently used.
Up next, the team principal had the total wins of all the teams over the seasons and the wins per season per track to their disposal. The reason for this choice was the assumption that the team principal is most interested in how the team is performing overall in terms of wins and scoring points. This group needed the most adjustments after the user study, the team principals want to see more detail of the championship results and more information on how the driver is doing in terms of speed and pit stops. One disagreed with the showing of the team wins per circuit in the past as they were deemed nice for the books but not important. Still this should be kept in as the data from the past might be important in the future. Sure the cars change but a lot can be learned from the knowledge at which type of circuits the car performed well and the other way around.
To answer the first research question "How can a rule based model in combination with a visualization tool best distinguish different types of users?" the rules described in the concept idea are kept in mind. At this moment the best way to distinguish the users is through what plot they click on, how fast they move the mouse and how many times they click. Of course these rules need to be tested and perfected but for now this seems the best way.
The second research question "What is the optimal model threshold setting for showing extra (or less) detail to a user?" cannot be fully answered right now. In the concept it is described which user groups are assumed to move their mouses slow or fast but the exact threshold settings need to be determined after extensive testing for which the capacity is lacking right now.
"What differentiates the different users?", sub question one, is answered extensively in section 4.1.4. The user groups are in the concept idea separated by the aforementioned variables and their interests.
Finally, the last sub question needs to be answered "How can those differences be translated in a model?". For now a rule based model on the different variables would help in the gathering of more data. After, to really get the niche differences it would be good to train a machine learning model as that might uncover differences that have not been spotted before.

9 Conclusion

To conclude, a concept has been created to implement a model into a graphic user interface that is able to separate three different target groups of a Formula One team; the driver, the data analyst and the team principal. Each of these target groups can be distinguished using the plots in which they are interested, the speed they move their mouse and the amount of times they click on the screen. To set an optimal threshold for such a model, more research and testing is needed on a larger scale. Due to the current situation and time restrictions only a limited user study was conducted, this was not optimal but still gave some useful insights.

9.1 Future work

In the future, more research needs to be done amongst specialists on the topic of Formula One racing and data, a lot of decisions right now are based on assumptions as there is little knowledge on what these people actually want. Moreover, better rule based models or even machine learning models could be trained to better adjust the visualizations to the target groups. When a good model is in place, the target groups could even be extended to include more types of team members.

Acknowledgments

This bachelor end project is supported by Eindhoven university of technology and Tilburg university. I would like to thank my supervisor Stef van den Elzen for his support, creative freedom and great ideas.

REFERENCES

How many people work in an f1 team? (2020, December). Retrieved from https://flowracers.com/blog/how-manypeople-work-in-an-f1-team/

Hsiao, I., Guerra, J., Parra, D., Bakalov, F., König-Ries, B., & Brusilovsky, P. (2012). Comparative social visualization for personalized e-learning. In *In proceedings of the international working conference on advanced visual interfaces* (pp. 303–307).

Kaggle: Your machine learning and data science community. (n.d.). Retrieved from https://www.kaggle.com/

Oscar, N., Mejía, S., Metoyer, R., & Hooker, K. (2017). Towards personalized visualization: Information granularity, situation, and personality. In *In proceedings of the 2017 conference on designing interactive systems (dis '17)* (pp. 811–819). doi: https://doi.org/10.1145/3064663.3064704

Ottley, A. (2020). *Adaptive and personalized visualization.* Retrieved from https://www.morganclaypool.com/doi/abs/10.2200/S00973ED1V01Y201912VIS01

Plotly. (2017, Nov). *introducing dash* . Author. Retrieved from https://medium.com/plotly/introducing-dash-5ecf7191b503

Qian, X., Rossi, R., Du, F., Kim, S., Koh, E., Malik, S., … Ahmed, N. (2021, Feb). *Personalized visualization recommendation.*

Rao, R. (2020, Jan). *Formula 1 world championship (1950 - 2021).* Retrieved from https://www.kaggle.com/rohanrao/formula-1-world-championship-1950-2020/version/13

Reichert, M., Kolb, J., Bobrik, R., & Bauer, T. (2012). Enabling personalized visualization of large business processes through parameterizable views. In *In proceedings of the 27th annual acm symposium on applied computing* (pp. 1653–1660).

Spenke, M., & Beilken, C. (2003). Visualization and analysis of formula one racing results with infozoom-the demo. In *Interact.* Retrieved from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.493.475&rep=rep1&type=pdf

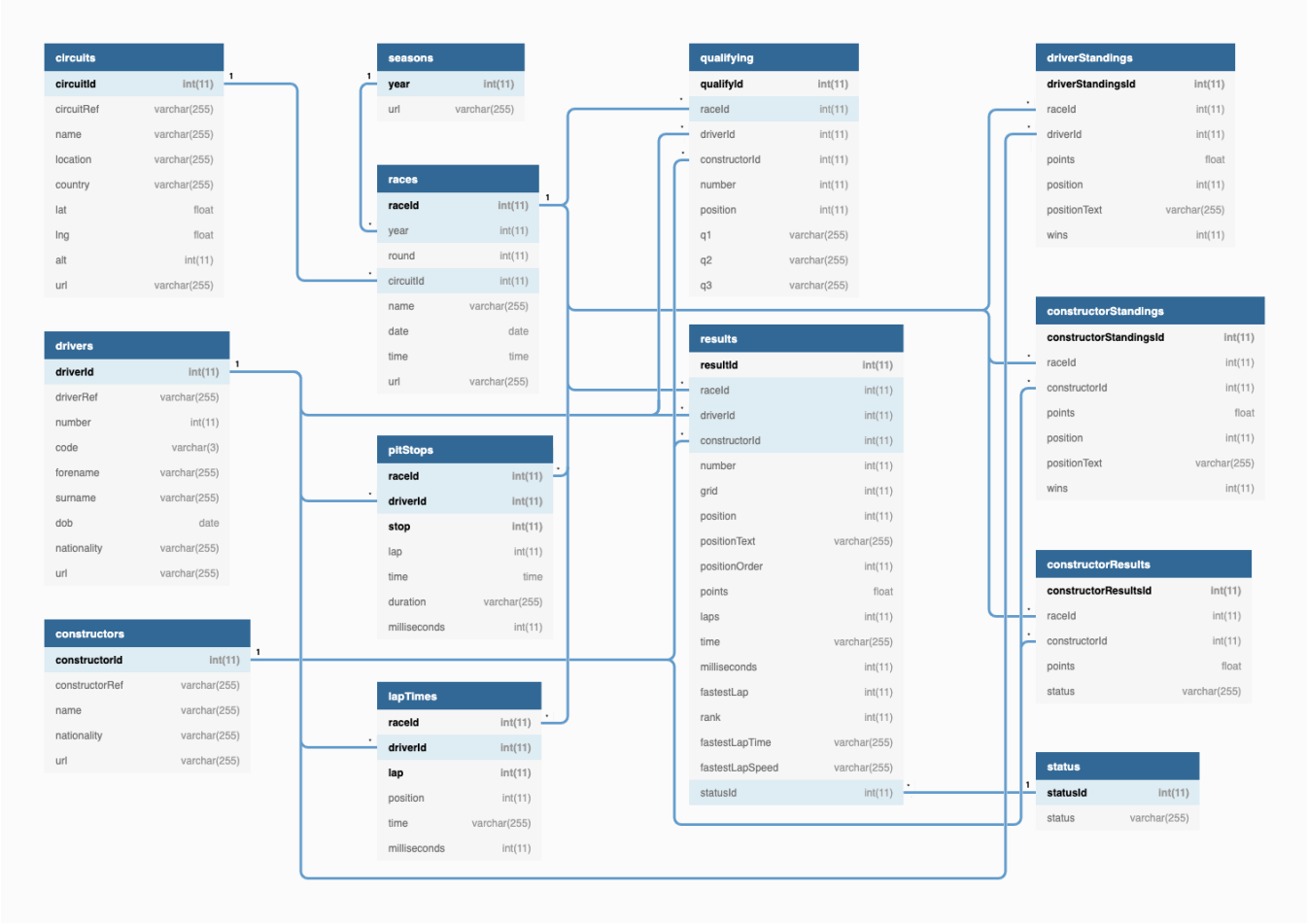

Figure 8: Scheme of the dataset