Process Book

Landen Kirkland | landenk@clemson.edu | C43786610
Joshua Pitts | jepitts@g.clemson.edu | C64307668
Nathan Frye | nfrye@g.clemson.edu | C10239405
Github Link: https://github.com/nathan-frye/nathan-frye.github.io.git
Website Link: https://nathan-frye.github.io/

Overview:

Formula 1 (F1) is arguably the highest level of motorsport in the world, with extreme levels of engineering and skill required to be successful. Oftentimes the difference that separates the greatest drivers from the others is just a fractions of a second. Each season of F1 sees the drivers competing in a series of races. Depending on where they finish in the race they will be awarded championship points which will be totaled at the end of the season. First place gets the most points and after that each driver will get less points, and the lowest positions get no points. In this project we will be looking at the period in F1 between 1990 and 2020 and showing how many points each driver got in the championship. This allows a user to see the best of the best, and the worst of the worst drivers, and you can even see how a driver changes over their career.

Questions:

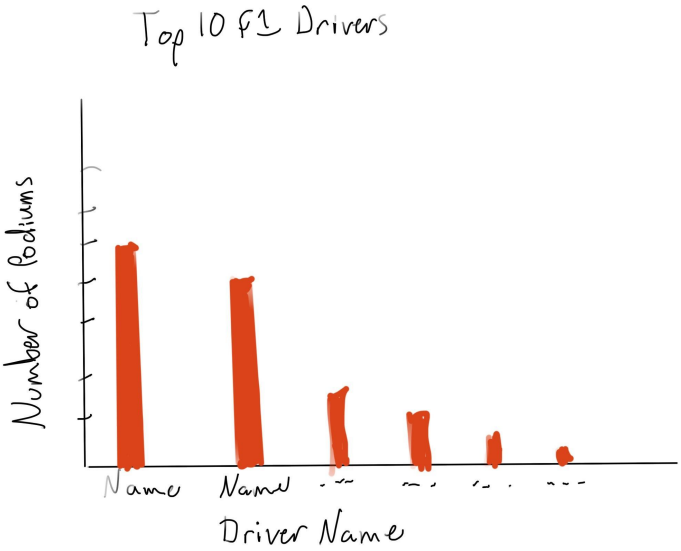
At the beginning of our project, our main question was “Who are the most successful F1 drivers?” We were originally going to only look at the ten best drivers based on the number of podium finishes they got, however it was decided that it would be better to look at all the drivers using points instead within the time frame, so that a user could make comparisons and decide for themselves who they thought the best are. In the end we could still have the same question, but now we can also answer “Who are the least successful F1 drivers?”

Data:

https://www.kaggle.com/rohanrao/formula-1-world-championship-1950-2020?select=results.csv

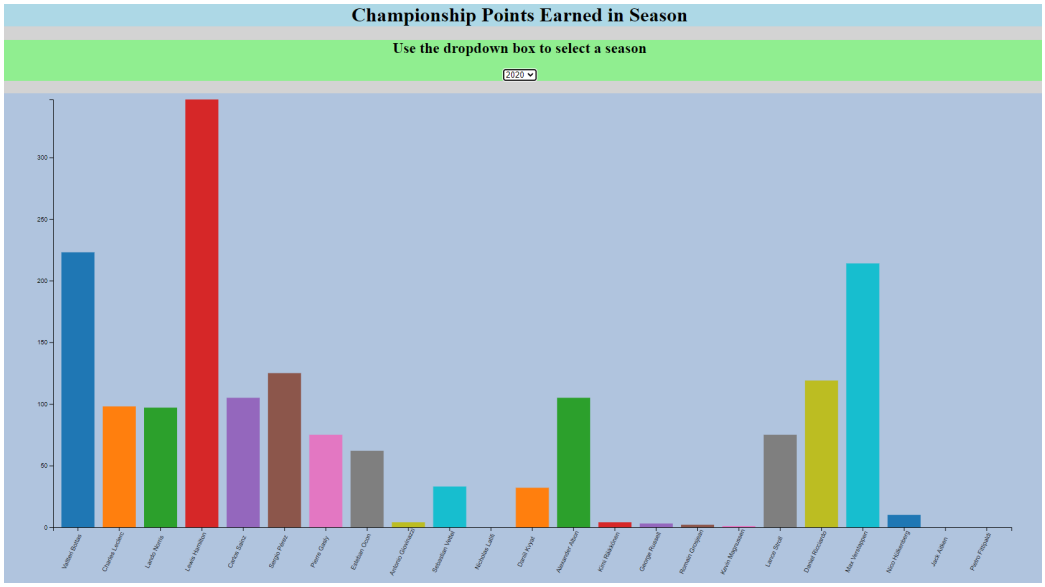
We used the dataset above for our project. It contains a huge amount of information about F1 from 1950, to 2020. If you wanted, you could look at every individual lap for every driver in every season, but we focused mainly on looking at one year at a time. There was no cleanup needed for this data, some unnecessary columns could be removed, but it wasn’t needed.

Vis sketch used for prototype:



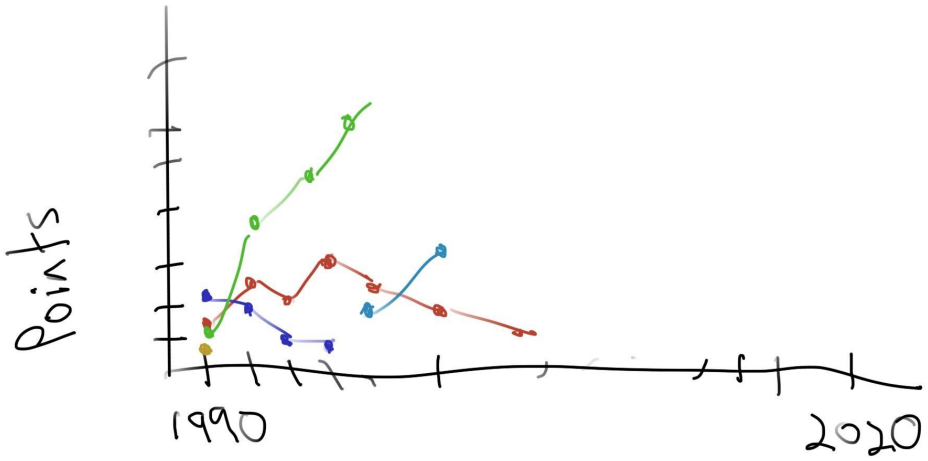
We decided to use points rather than podiums, as some points are awarded for more than just the top 3 positions.

First visualization prototype:



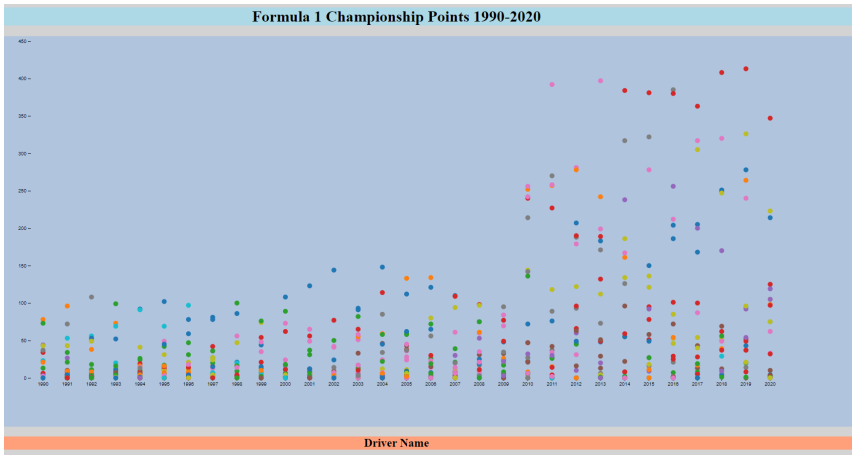
Years from 1950 to 2020 were selected using a dropdown box. Each driver that participated in any races for that year is shown on the graph along with how many total championship points they earned that year.

Basic sketch for next iteration



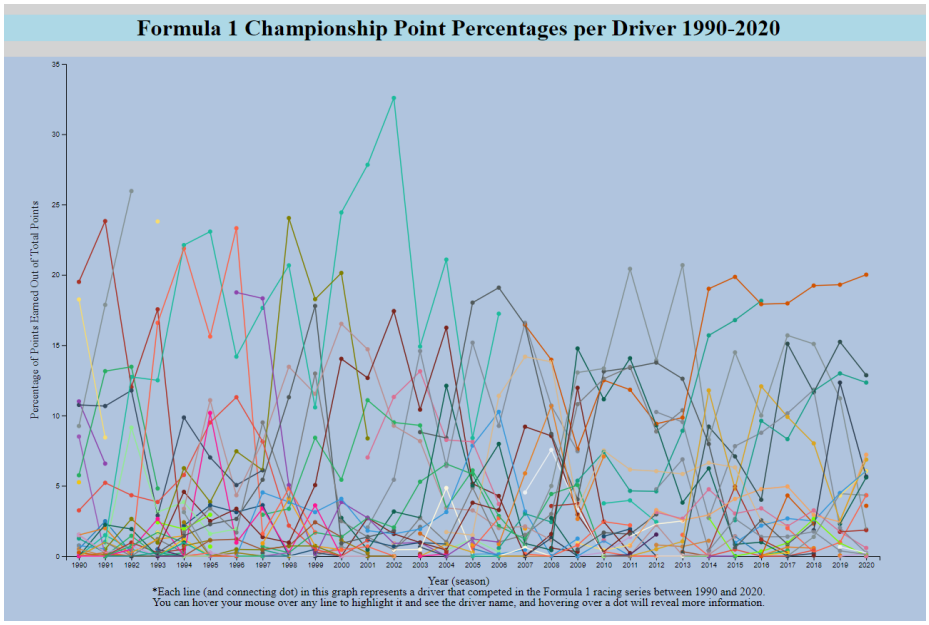
Changed to a more limited year range of 1990 to 2020 rather than starting at 1950. This has every driver represented for each year that they raced, with drivers participating in multiple years being connected by a line. Planned interaction will allow you to hover over a dot and see the driver's name, then if you click will bring up a secondary visualization for that driver and season so you can see their performance in each race that season. Going with this approach rather than a bar chart allows a user to much more easily compare between seasons.

Start of next iteration

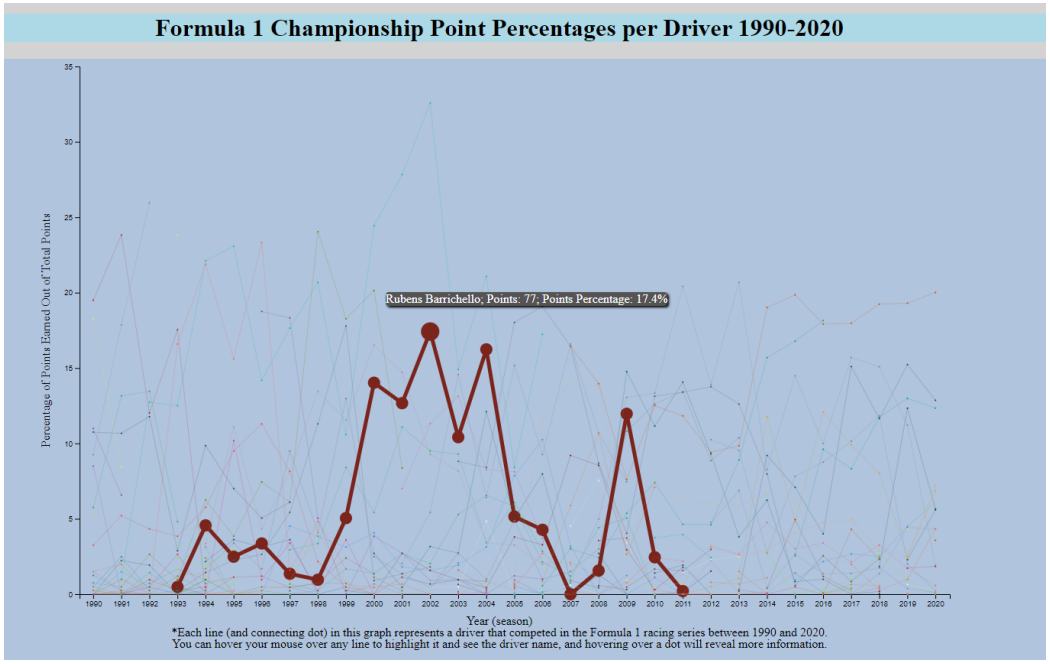


First iteration of the new prototype, missing connecting lines for drivers. Also will need to give drivers unique colors so that they aren't mistaken for others. In the future the y-axis will be based off of percentage of points earned rather than total, as you can see there are at least two different scoring systems that make the graph inconsistent.

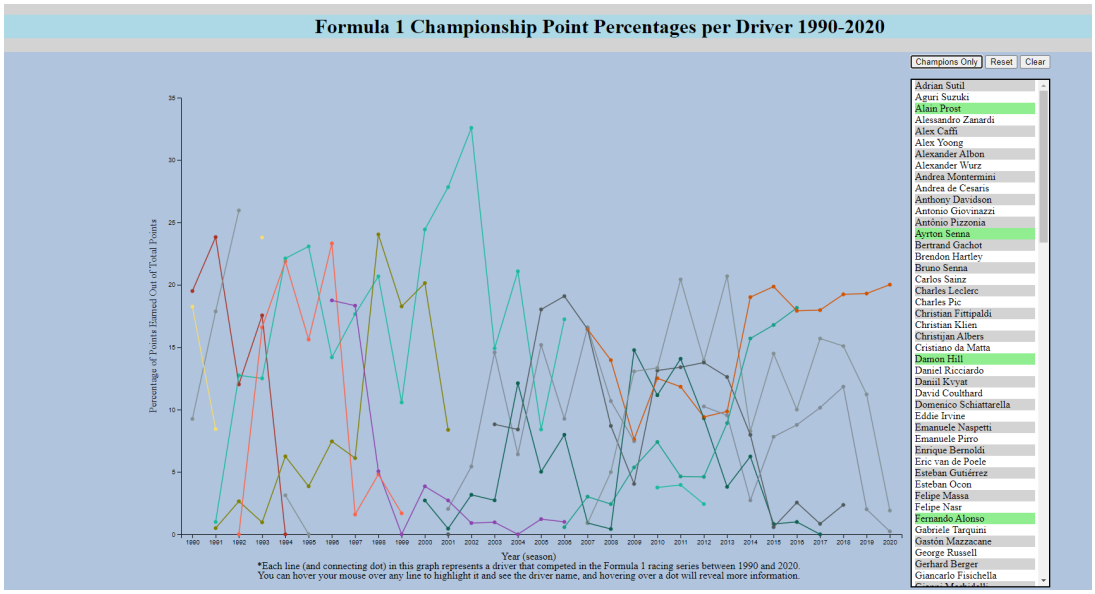
Main Visualization near completion



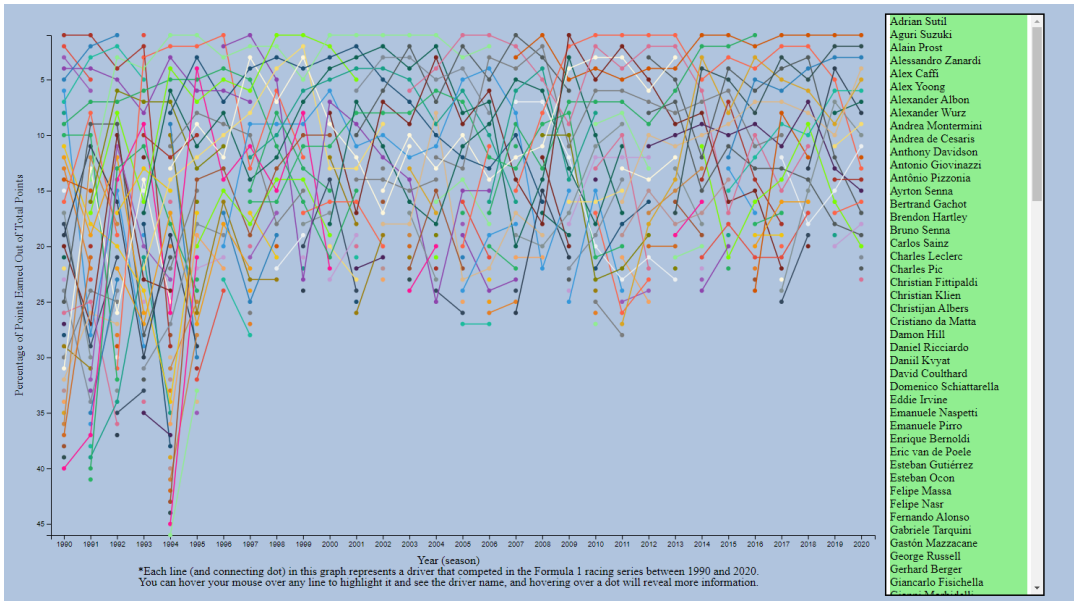
The graph looks fairly messy, but the interaction below makes it much easier to read.



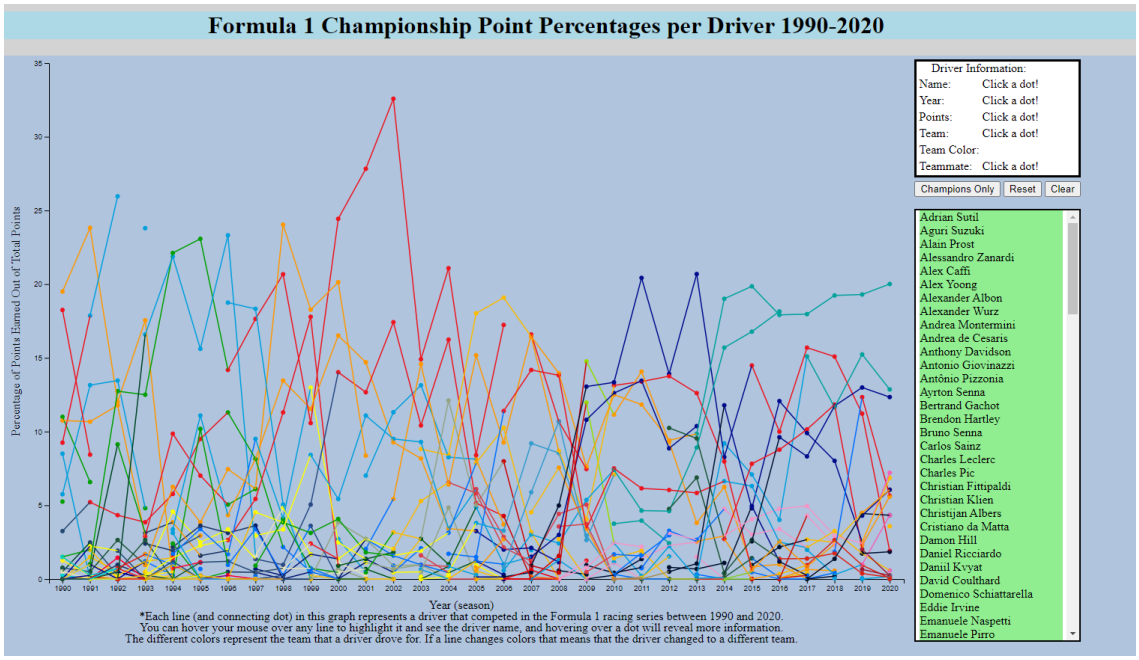
Hovering over elements highlights the selected one, and dims the others, allowing the user to clearly see the driver that they are interested in. There is also a popup showing the name, and if hovering over a dot, the points and points percentage for that year for the driver. A difficult part of getting to this point is the way in which the points are graphed, using a loop, makes it so that only the most recent year graphed is easily accessible. We fixed this by storing the dot information through each iteration of the loop, and using it later to add the lines one year at a time.



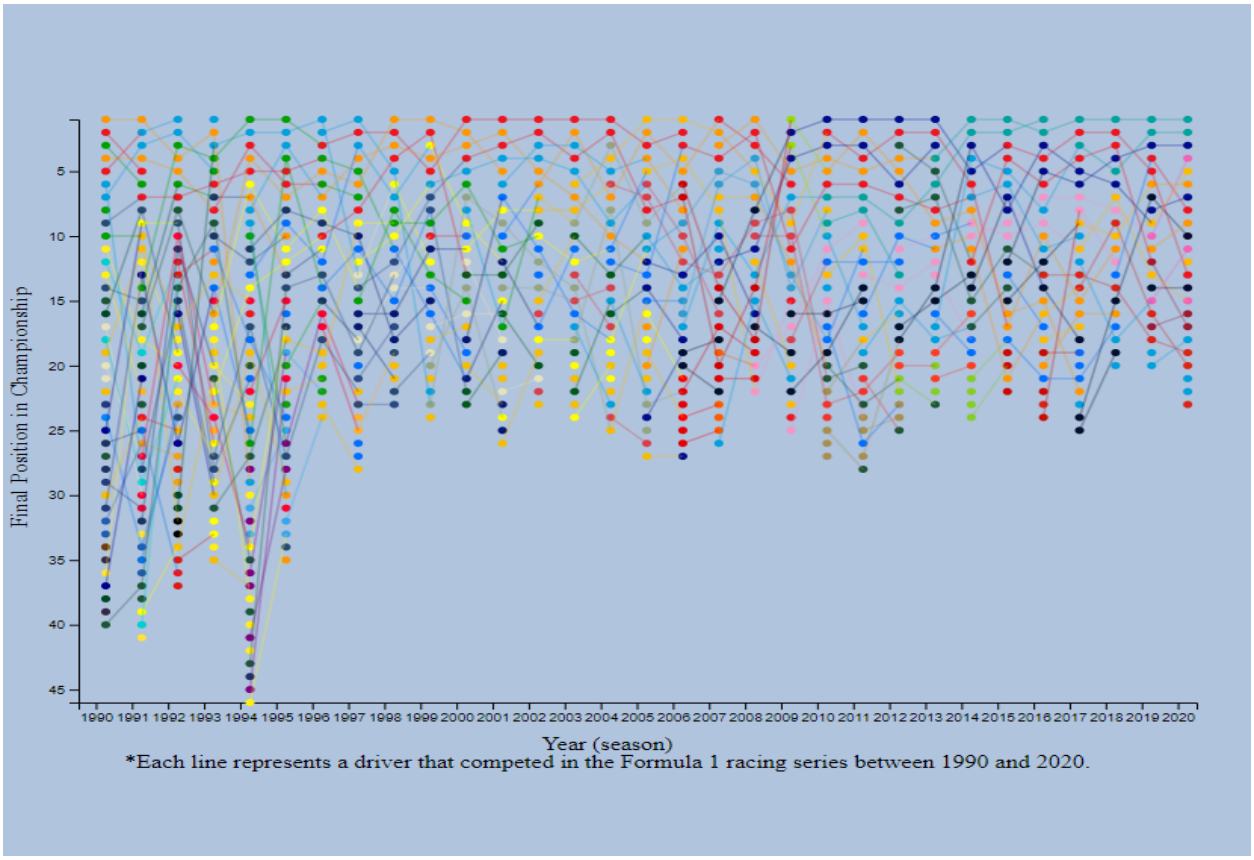
Added interaction elements at the top allow you to clear, reset, or display only drivers which have won a championship. There is also a box on the right with the name of every driver. Clicking a name will display in on the graph, so a user can select just the drivers that they want.



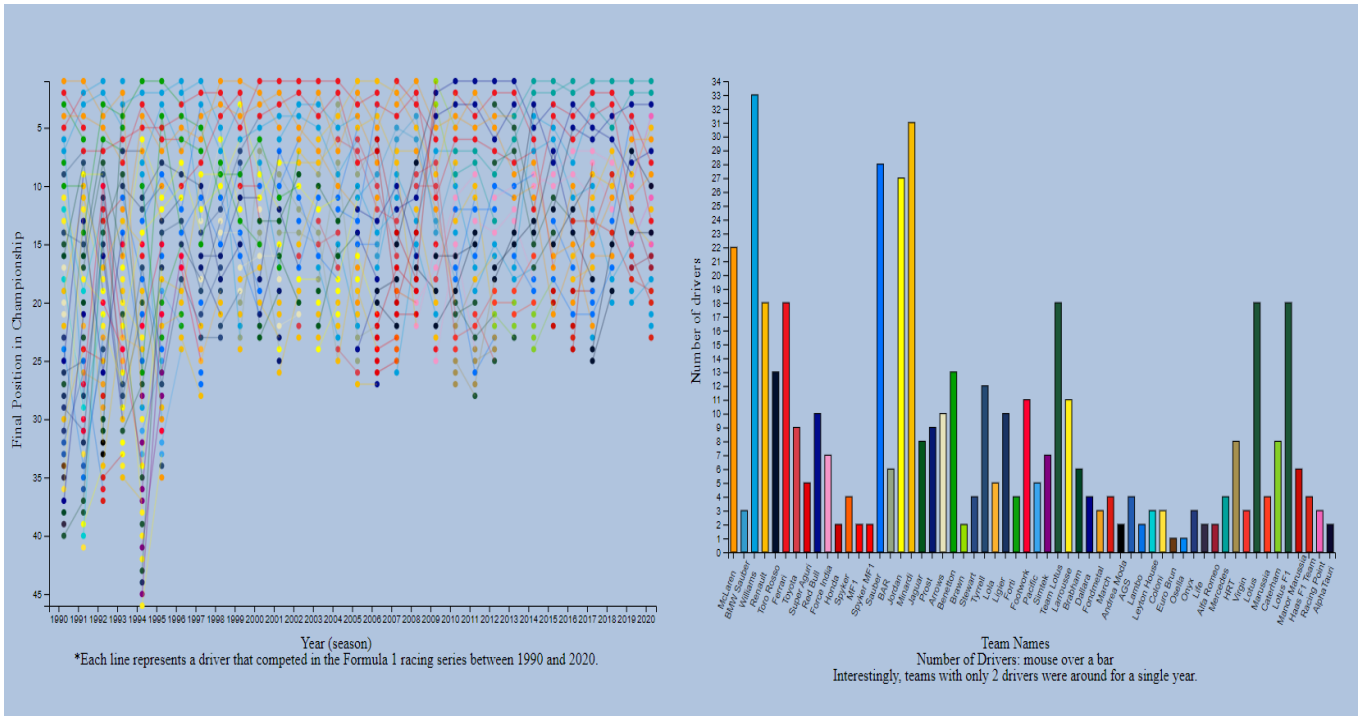
An alternative design which uses final championship position rather than point percentages. The left side is very messy here.



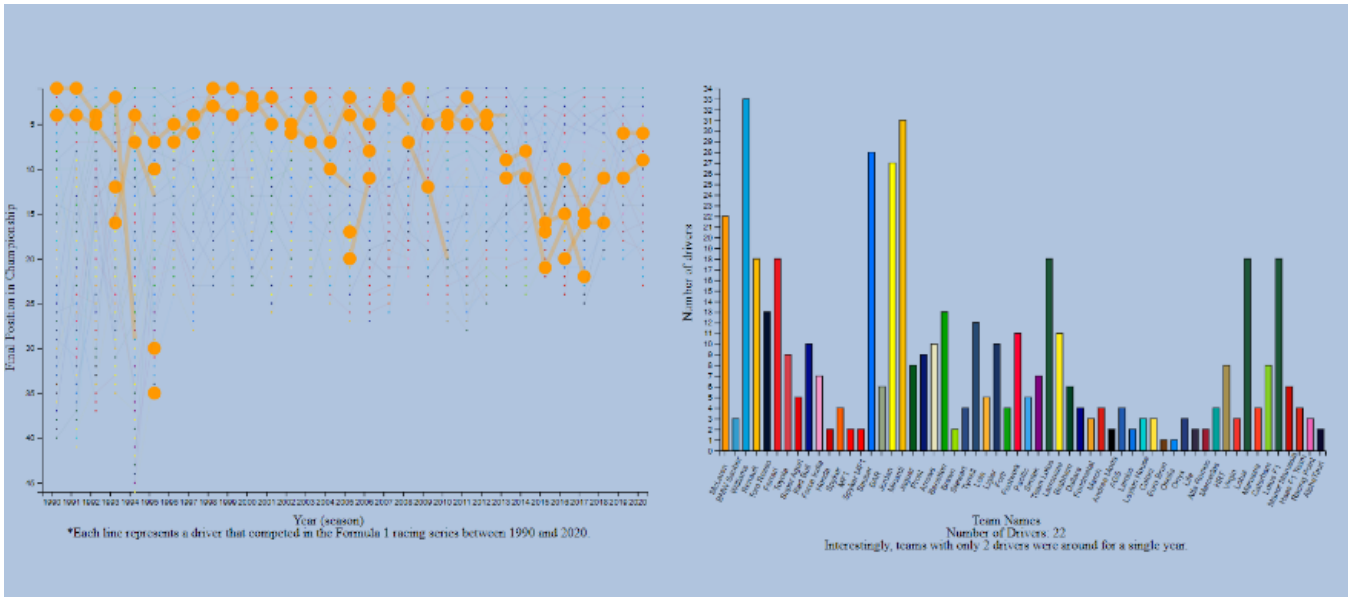
Here is the final iteration. A text box in the upper right corner was added so that clicking a driver will show more information for that driver and year, like team and teammate. With this design, you can answer our main questions by seeing who the best and worst drivers are, and everyone in between. If a user wanted, they could select a single driver, and see how their career evolved over the years that they drove which leads to a possible expansion/improvement. It would be interesting to be able to select a driver, and have more information to better see their career development, but that is beyond what we were just trying to show here. The bottom of the graph, however, was too “busy” and it was too hard to locate unique dots (drivers) in the graph.



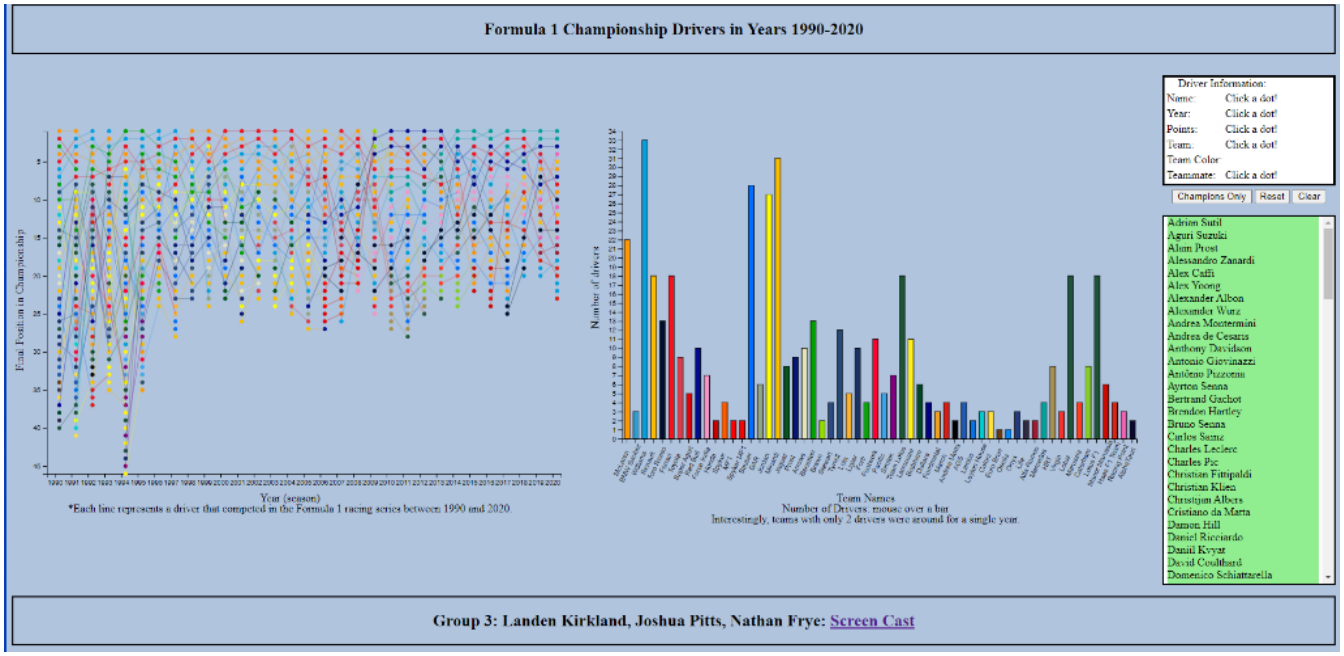
Reverting back to a previous idea, we displayed the final position of each driver for each year. This time we reduced the strength of the lines to alleviate some of the visual clutter in the left-most portion of the graph. Dots are now clearly distinguishable and shifts in final position are still visible through the line connections. We still have a problem with what the colors represent.



We created a bar graph that displays the total number of drivers for each unique team that is displayed in the line graph. The bar graph also serves as a legend for our color scheme by indicating what color each team is represented by in both graphs.



Additionally, we added consistent mouse-over interaction with the bar graph. Mousing over a bar will highlight all dots and lines corresponding to the drivers of that team on our first visualization, allowing users to see how that team's drivers placed over the years. Users can also see when lines protrude from dots, but do not connect to another dot easily with this display - indicating that a driver had changed to a different team for that year.



Our final visualization combines our first visualization, our second visualization, and our selection list/table. We kept the bar chart next to the line chart so that colors could be seen at once and directly compared between the two graphs. Additionally, all elements of our visualization are displayed in a single view. We felt that having to scroll through a webpage, or visit different pages to view individual elements would place a memory burden on the user and interactions would suffer if changes were not immediately noticeable.

Evaluation:

One thing that we learned by looking at the data is that many of the drivers in F1 are very inconsistent. There are a few that seem to do really well and are mostly at the top of the graph, but many midfield drivers are all over, one year they may get 0 points, and the next they are near the top. We also learned that the way that points are given to drivers changes a lot year to year, which is why we moved away from using only points fairly quickly, but luckily it is not an issue in our final visualization. Our main questions for this project was “Who are the most and least successful F1 drivers in our time frame?” The way we decided to answer this question was to allow the user to actually see every driver in one of the best measures of success in F1, their championship position at the end of every season. Successful drivers won lots of races and points, and finished higher than the unsuccessful drivers. Not only this, but the visualization allows you to filter to see only drivers that you want to, and you are able to see the change in their success over their whole career as some drivers were very successful in one year, only to fall into mediocrity the next year. We also wanted a user to be able to compare teams as well, so we have each dot and line the color of the team that they represent, and the graph on the right shows how many drivers drove for each particular team. A user can mouse over a bar for a team, and highlight all the drivers and lines for that team, so they can also see if a particular team contributed to the success of a driver or not as well, as each team makes their own car and not all are equal.

We think that our visualization works quite well in order to compare different drivers and teams. The filtering tools we added make it much easier to show only drivers that you want and see how they are different, or if you want to compare one driver to all the others, you can mouse over their name and easily see them outlined in the graph. There are also some areas that the visualization can be improved. To start, the graph with the drivers is quite cluttered and can be hard to read sometimes, we tried to alleviate this by lowering the line opacity, and allowing a user to highlight a driver to see them easily, but when all drivers are there it can be hard to follow a specific driver without filtering. Another improvement would be to implement better interaction between the two graphs. As it is, the team graph can be used to highlight all drivers for that team, but they do not influence each other more than that. Additionally, color use could have been improved, a lot of colors are similar and can cause some strain on the visualization. Color was a consistent issue throughout the project, but we struggled to find a balance between the number of unique colors to the amount of unique teams.