Most Dominant Formula 1 Drivers

Landen Kirkland | [landenk@clemson.edu](mailto:landenk@clemson.edu) | C43786610

Joshua Pitts | [jepitts@g.clemson.edu](mailto:jepitts@g.clemson.edu) | C64307668

Nathan Frye | [nfrye@g.clemson.edu](mailto:nfrye@g.clemson.edu) | C10239405

Github Link: https://github.com/nathan-frye/nathan-frye.github.io.git

**Background:**

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 few tenths of a second. Since it was first introduced to the modern era, there have been a few drivers and cars that were dominant in the sport, frequently finding themselves in the top 3 finishing positions, also called a podium finish. In this project we will be looking at the top 10 drivers to have competed in F1 since 1950 based on the number of podiums that they were able to get during their career and other accomplishments that came with that.

**Objectives:**

* Who are the most successful F1 drivers?
  + The primary objective with this project will be to provide an overview of the 10 most successful F1 drivers since 1950.
  + This will include finding a total number of podium finishes that each driver was able to achieve in their career. This will be our measure of what success means.
  + We will be able to compare the best drivers, and see perhaps if they have similar performances, or was one or two much more successful than the others.
  + Some additional information can be provided such as the number of championships won (having the most points at the end of the season) to see if perhaps some of the most winning drivers didn’t always win the championship. This could happen if they consistently got 3rd place while another got 1st, for example.
* How did the success of these drivers change over the time of their careers?
  + We will then further break this down for each driver, finding out specifically how many first, second, and third place finishes that they were able to get each season (one season each year).
  + This will allow us to see if the best drivers took some time to achieve their success by gaining experience, or finding a good team (each team develops their own car and each is a little different in performance).
  + For each season we will also include the team that they drove for and if that was a championship winning season.

**Data and Processing:**

We got our data from: <https://www.kaggle.com/rohanrao/formula-1-world-championship-1950-2020?select=results.csv>

This is a large dataset containing almost all of the available information from the 1950 season of F1 to the current 2021 season. It is split into multiple .csv’s containing information on different aspects of the sport. We will not have to do much cleanup with the data, but rather just pull out the data that we need to get just the info that we want. We will pull out the finishing position for every race that each driver participated in, their final position in the championship at the end of the season, the season (year), and the team that they drove for. This should be fairly easy to do with python, by just pulling all of the data from the .csv’s and then placing it into one or two smaller .csv’s.

**Visualization ideas:**

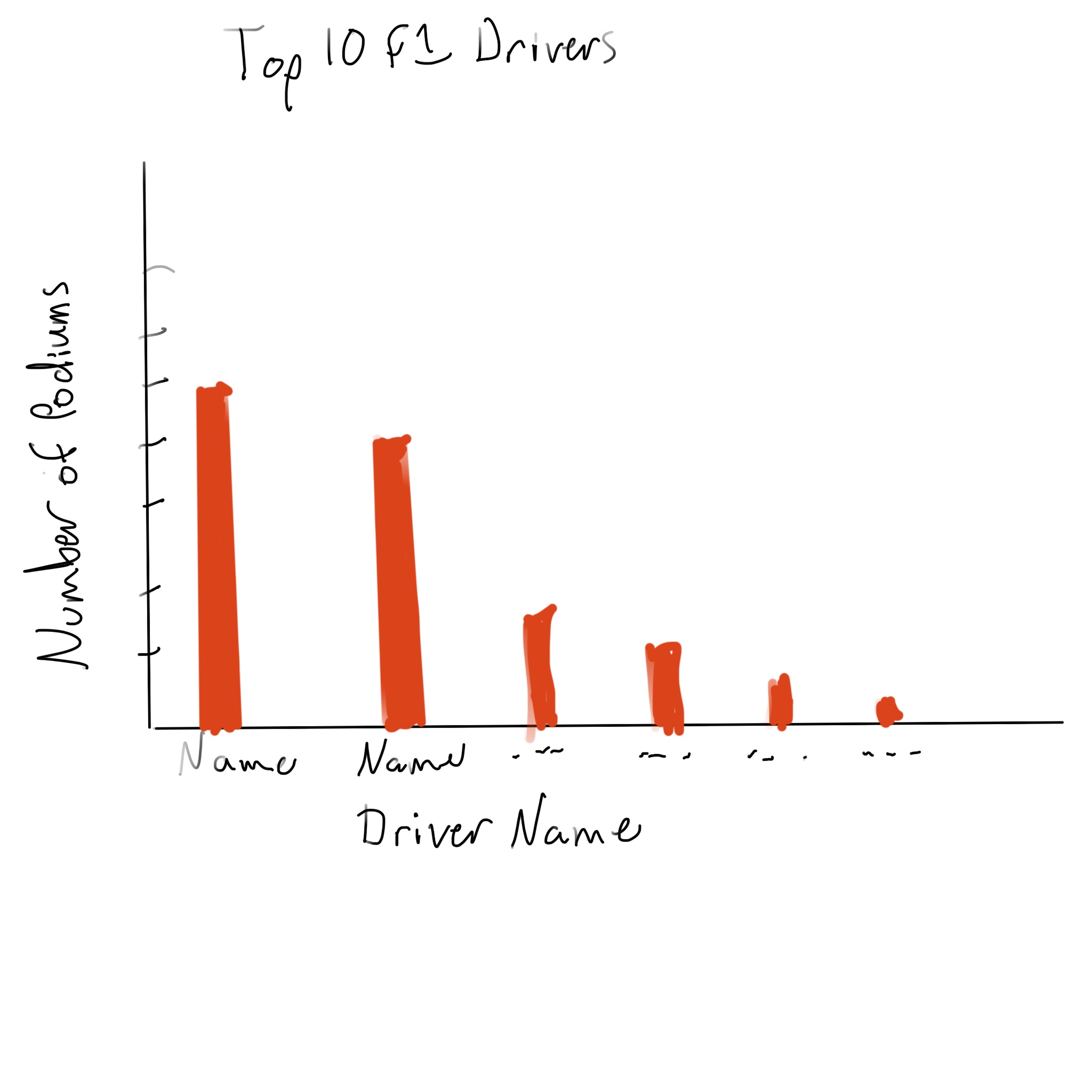
* Top 10 drivers to directly compare the top 10
  + Must Have Features
    - Total number of podiums each driver was able to get.
    - Easy side by side comparison.
  + Optional Features
    - Further breakdown to see specific 1st, 2nd, and 3rd place totals.
    - Show year range that each driver was in.
* Individual driver visualizations to get specific data about the career of each driver
  + Must Have Features
    - Number of podiums achieved each year that they raced.
    - Total number of 1st, 2nd, and 3rd place finishes they got.
  + Optional Features
    - Championship position (or maybe just highlight if they got 1st here)
      * Championship position of their team mate? (Could be a good comparison to another driver in a mostly identical car)
    - Most successful track to see if there was a track that they won a lot of races in particular.
    - Team they drove for each season.
  + There will be 10 of these, one for each driver.

**Visualization Sketches:**

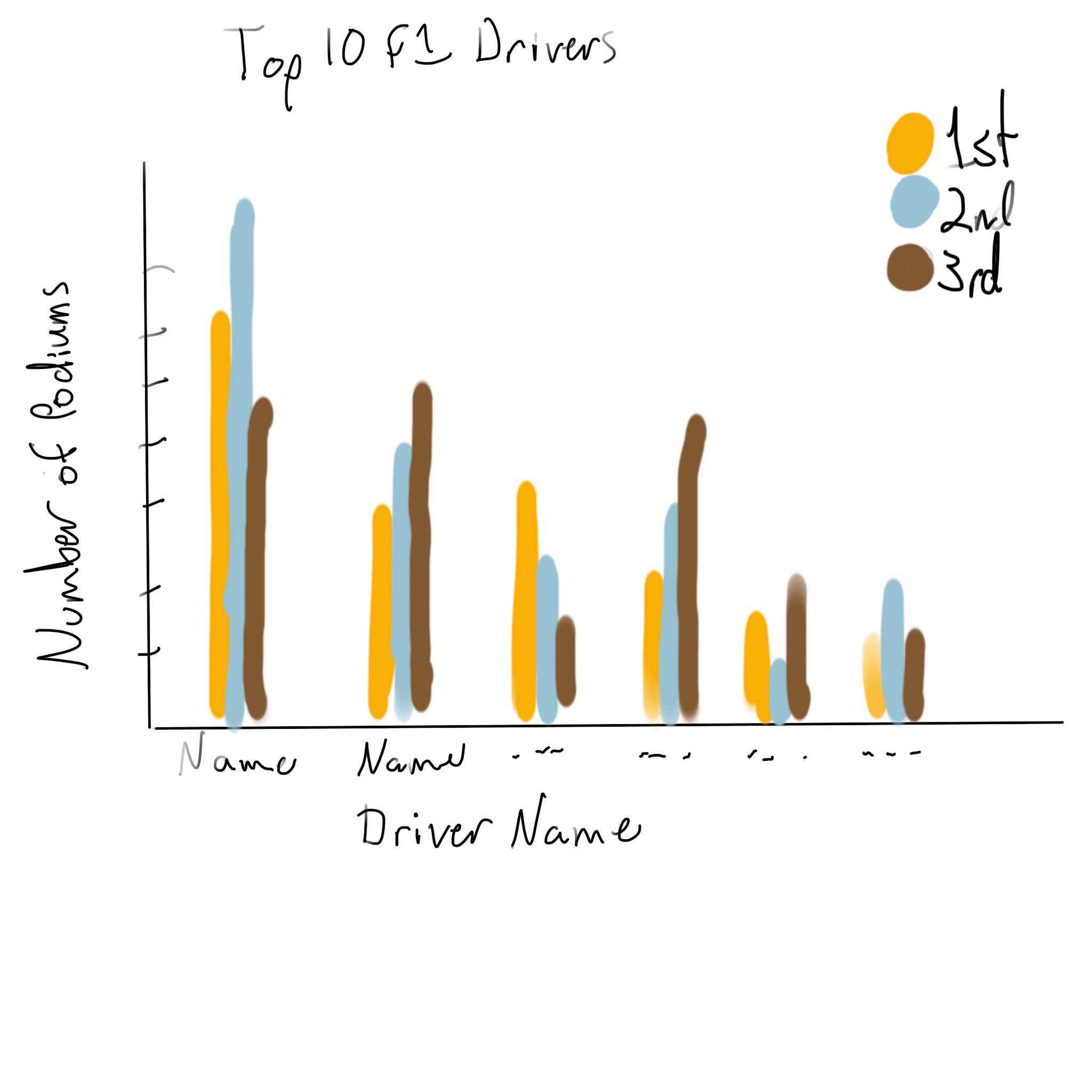
Top 10 drivers sketches:



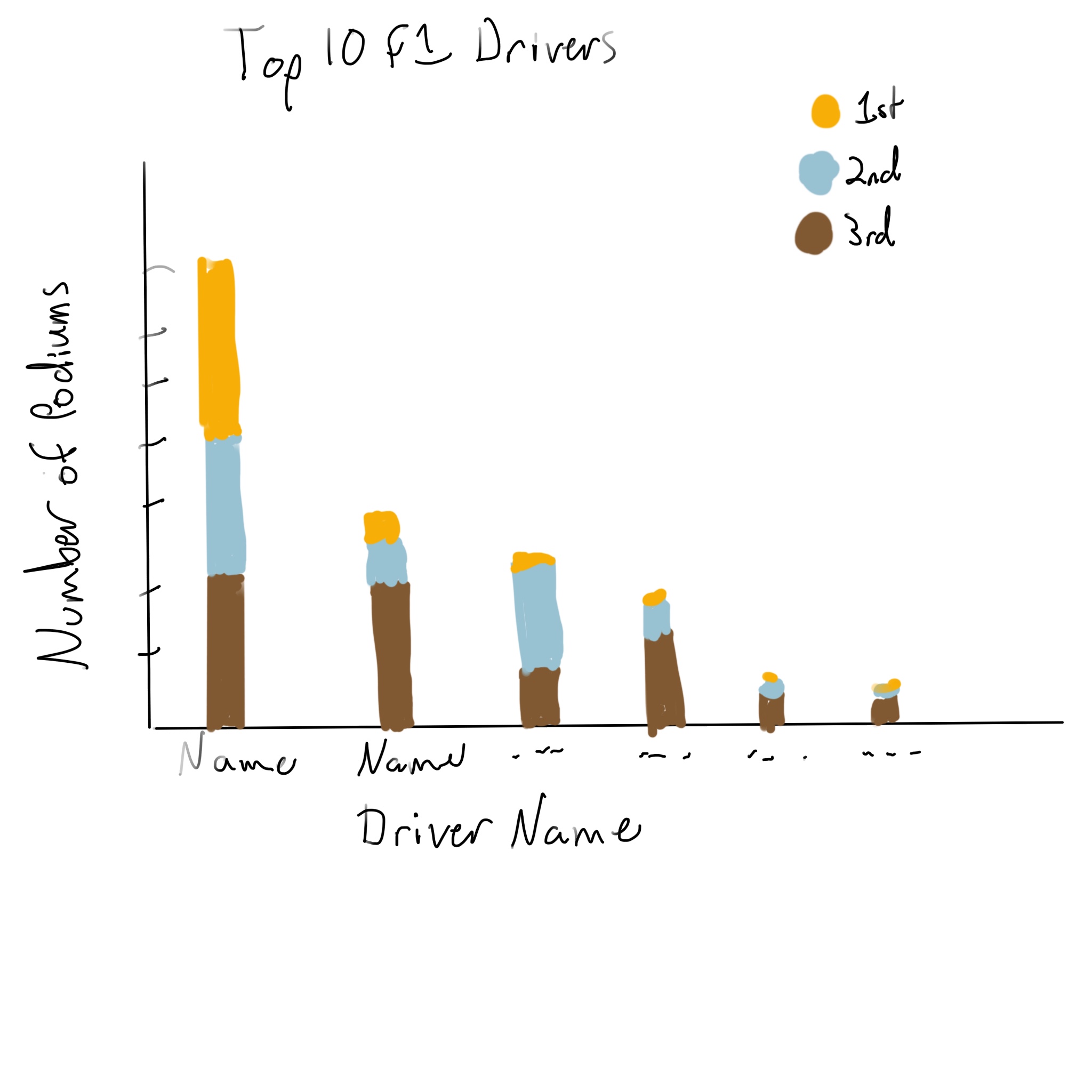
For this sketch, the size of the circle represents the number of podiums for each driver. It is a fairly simple visualization and would work okay as long as the drivers’ circles aren’t too close in size. We could also split these into pie charts to show the number of 1st, 2nd, and 3rd finishes but that may be too confusing.



This visualization will be better, especially if any of the bars are closer in height. It is also fairly simple and the only information shown will be the number of podium finishes. This may be good to use just a total number of podiums without breaking it down further, as each driver will have more information shown for them in the other visualizations.

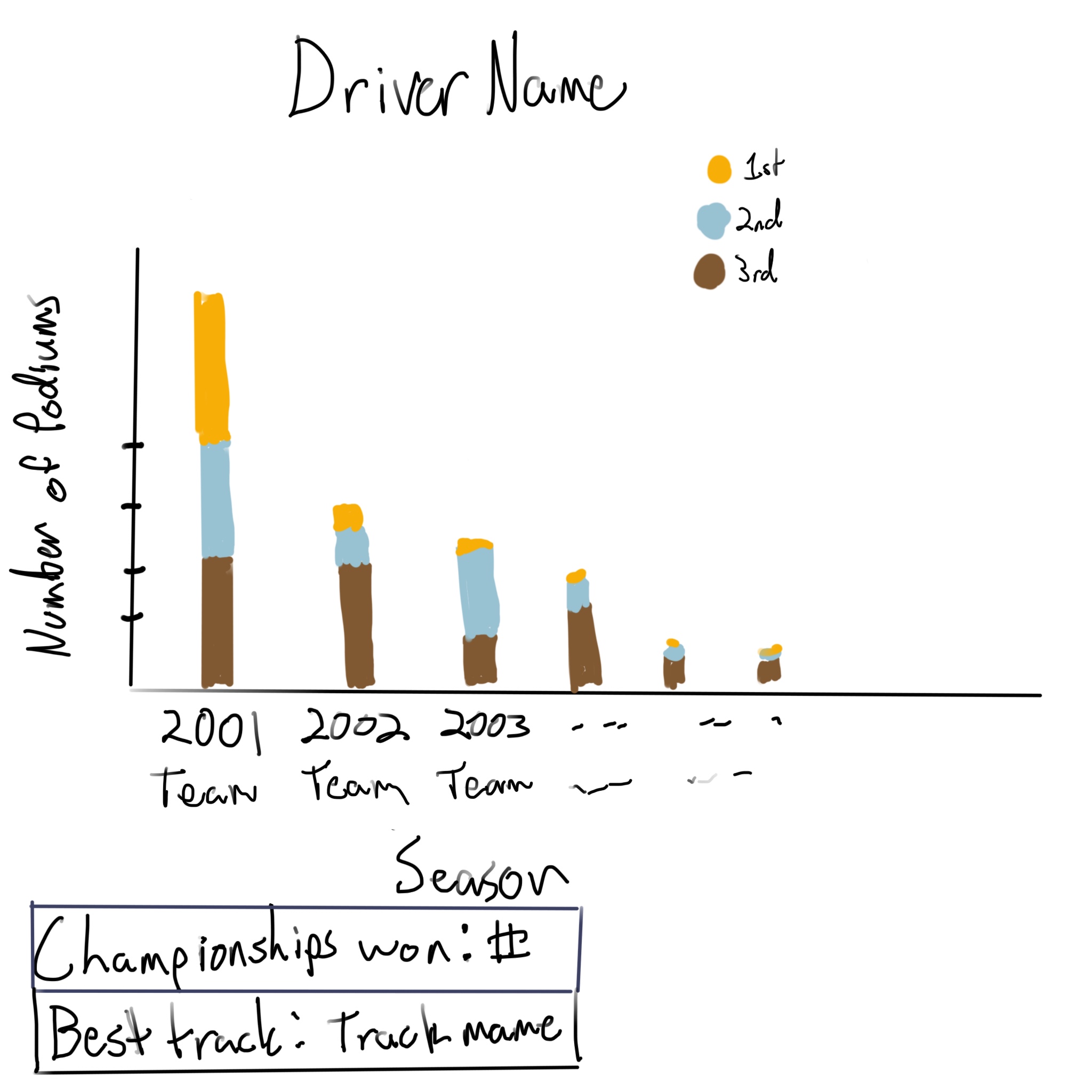


This visualization is similar to the one before, but breaks the bars down by numbers of 1st, 2nd, and 3rd place finishes. This one will be good if you are wanting to compare the number of each type of podium finish between drivers.

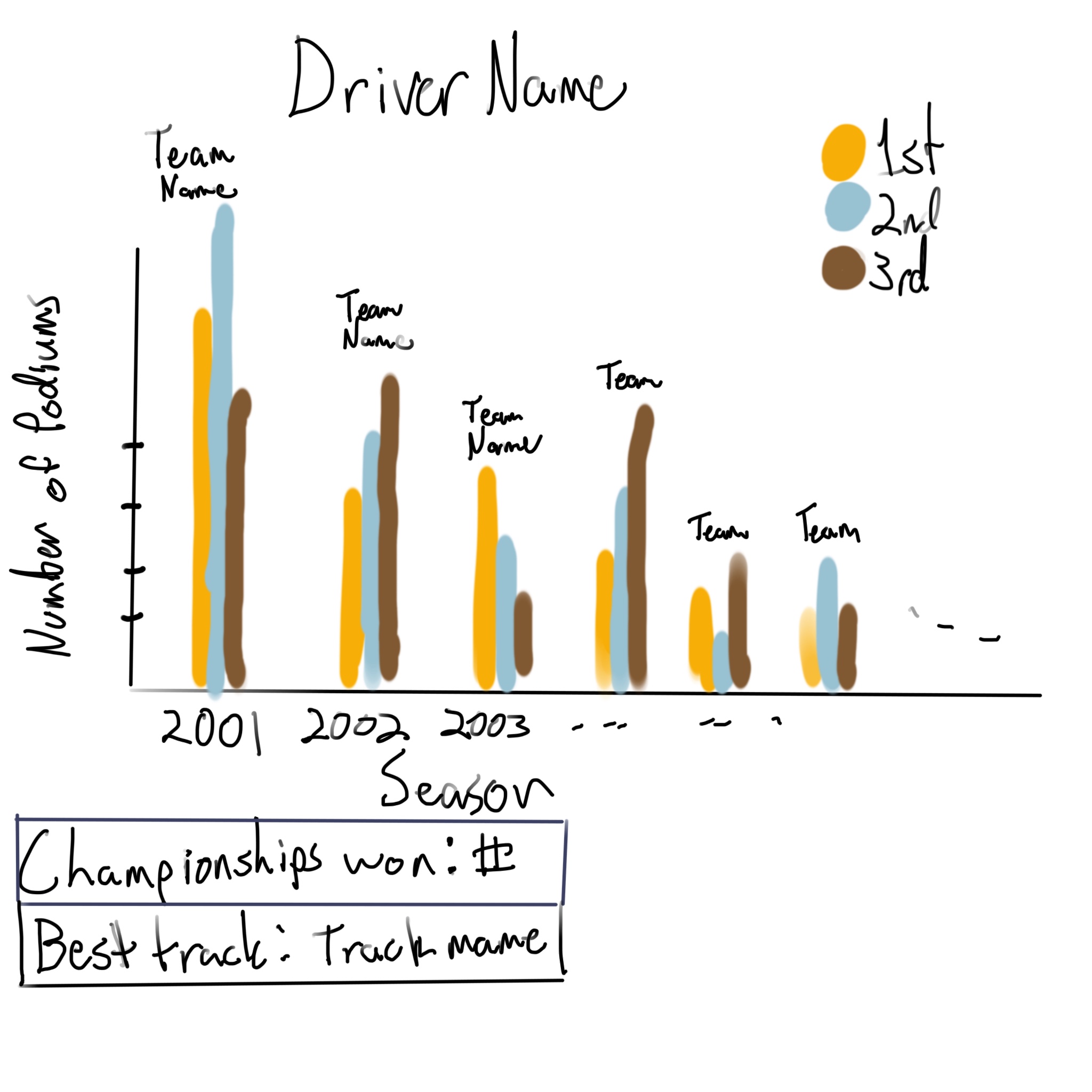


This is the best design we have for showing the most data possible, while still retaining the ability to best compare the total number of podiums between drivers.

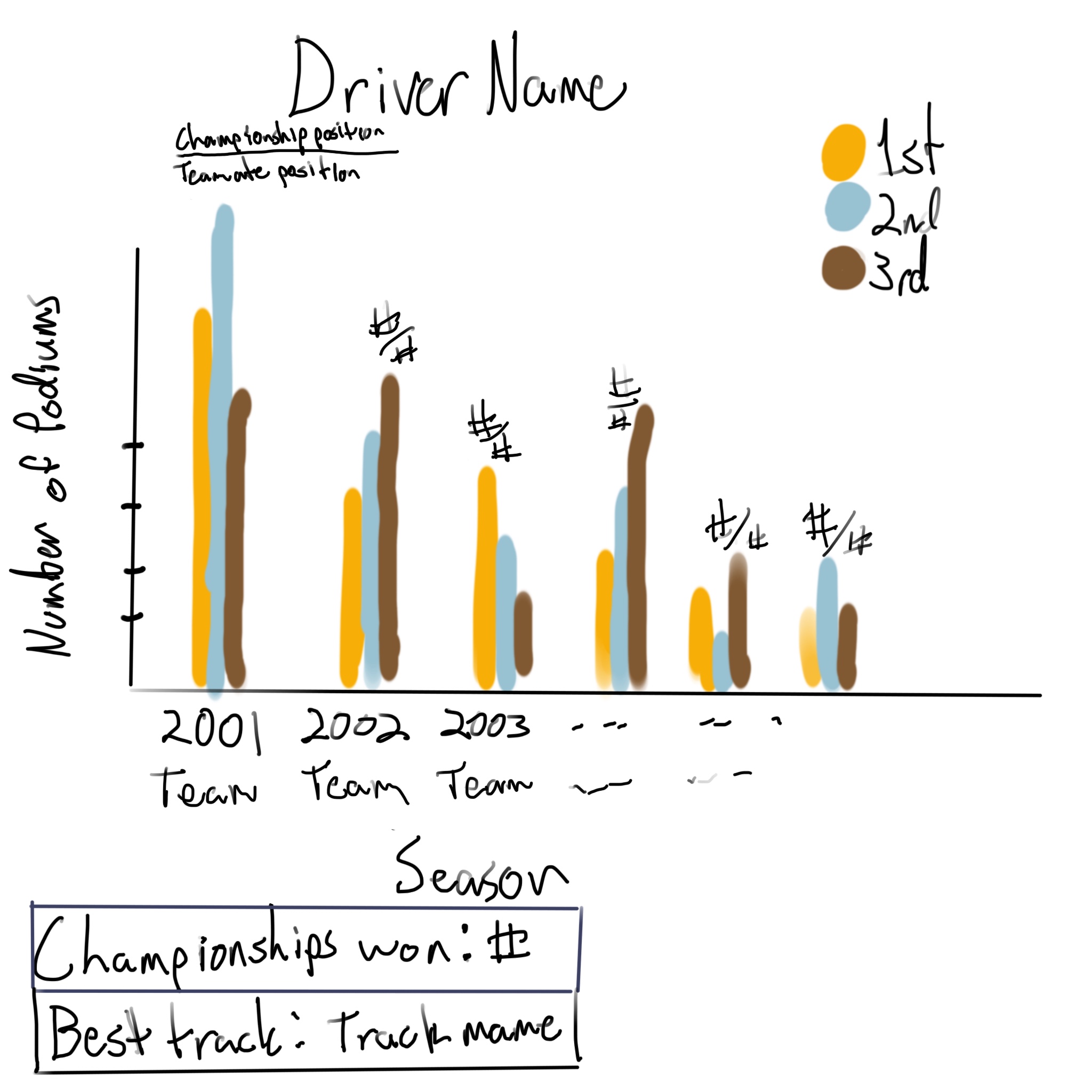
Individual driver sketches:



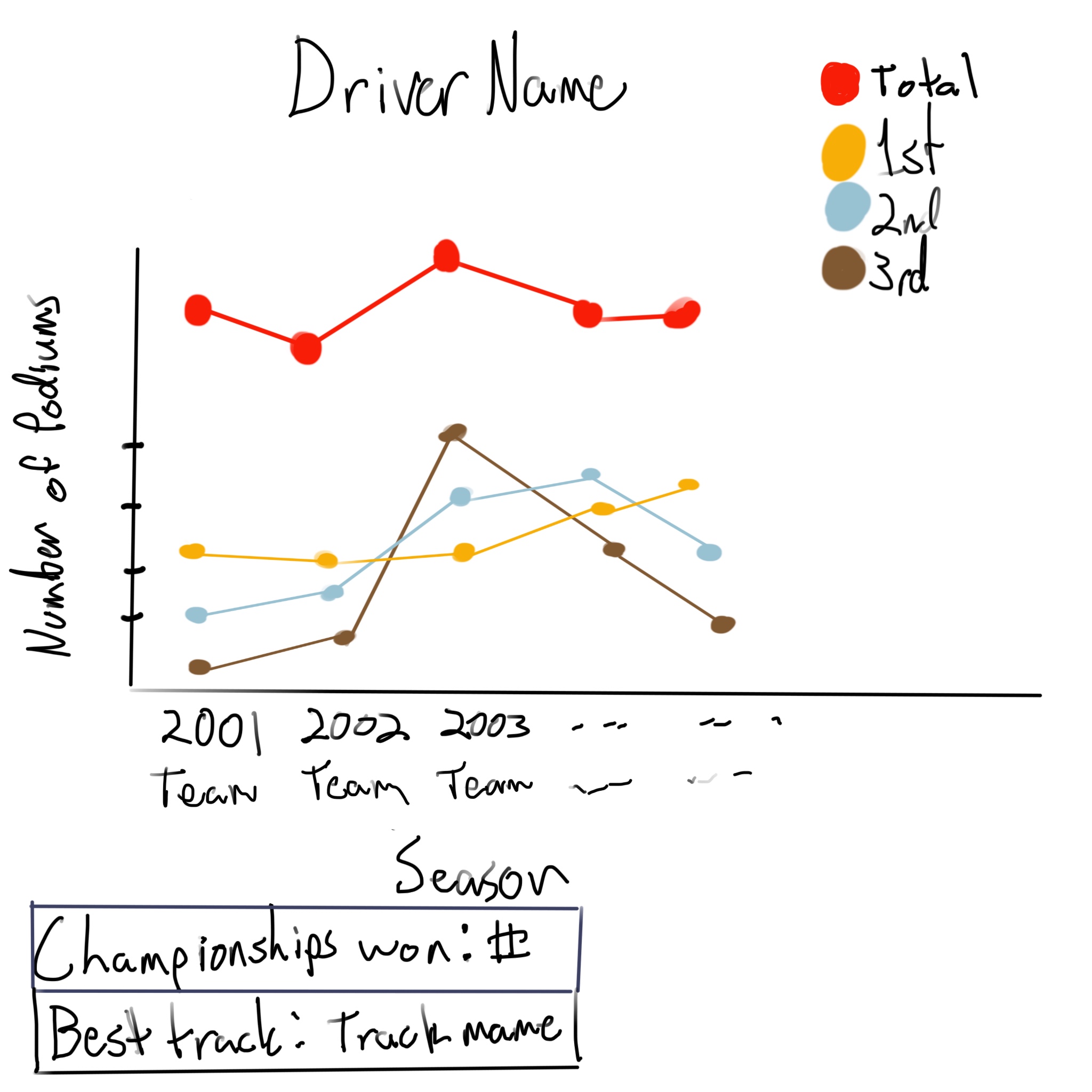
Very similar to the best for the top 10 visualization, this adds by giving info for the whole career of a driver.



This One is a little better, by dividing the bars you can better see how many of each type of podium the driver got. For this we are not as worried about comparing the totals for each as we were before, but a breakdown of the different types.

****

Similar to the previous, but with team name added under the year, and above the bars you have the final championship position over the position that their teammate got. The added features here are not as important though and only give extra information.



This is the best visualization for the individual drivers that we have. It uses a plot rather than bars, better than showing the change over the course of the driver’s career. It also has a line showing the total, but this can be removed if it would cause issues with scaling to make the graph look good. Like the rest, it has a simple box showing the number of championships won, and the track which the driver had the most success. The track is also optional and can be removed if needed.

**Schedule:**

October 17-23 - Pull relevant information out of the dataset.

October 24-30 - Rough creation of visualizations.

October 31-Nov 6 - Refinement of visualizations.

Nov 7 - Prototype Due

Nov 14 - Peer evaluation

Nov 30- Dec 2 - Oral Presentation

Dec 5 - Final Delivery