Process Book

Overview and Motivation

A big part of the reason why we chose to visualize Formula 1 was due to our love for cars and passion for racing. Formula 1 has always been a sport of numbers, where milliseconds determine who's the champion. Visualizing these numbers in a graphical user interface will truly help with understanding the different trends in Formula 1. With this project we hope to spark interest in more Americans, and encourage them to start watching Formula 1.

Related Work

https://jasonjpaul.squarespace.com/formula-1-data-vis

https://www.f1trends.com/blog

https://kwjames.com/f1times.html

https://f1pace.com/p/2022-monaco-gp-lap-time-distribution/gp7_ridge.webp

Questions

- What are some trends seen throughout the history of Formula 1?
- How do drivers across multiple eras of Formula 1 compare?
- Which drivers and teams are scoring the majority of points in a season?
- How do drivers and teams perform as the season progresses?
- How do drivers perform in a given race?
- How do lap times compare in a given race between drivers?

## Data

The original API is the Ergast Developer API

API Link - Ergast Developer API

We will be using the dataset available on Kaggle.

Kaggle Dataset Link - Formula 1 World Championship (1950 - 2022) | Kaggle

Exploratory Data Analysis

We tried out a multi-line chart to show position changes in the race between the drivers. This turned out to be a pretty good design, so we decided to work on and improve on it for the final product.
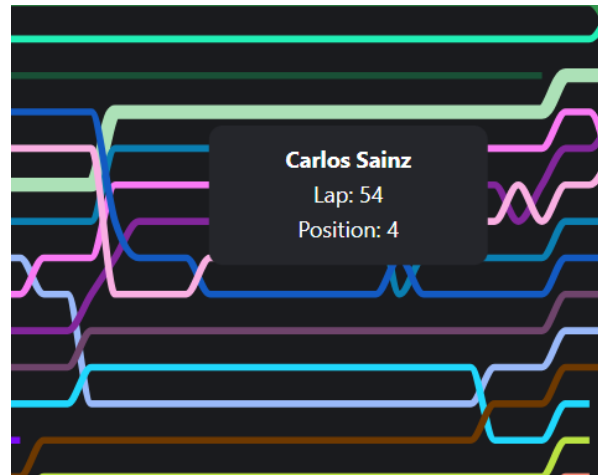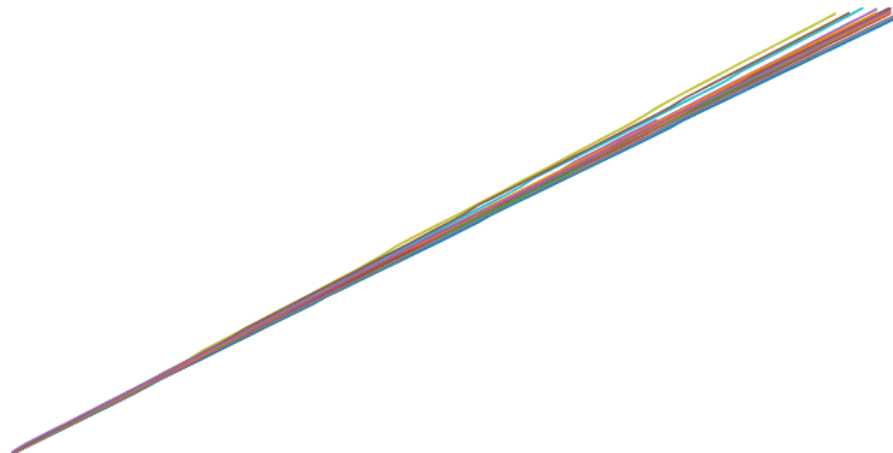

Race Position Multi-Line Chart

We decided that this visualization should be one of the main components of our project.

Design Evolution

One thing we thought might've been a good idea at the time was to use a graph that displays the lap times cumulatively. However, after implementing it this is what it looked like:


As shown, the data is very difficult to read compared to the individual lap data (Shown below)
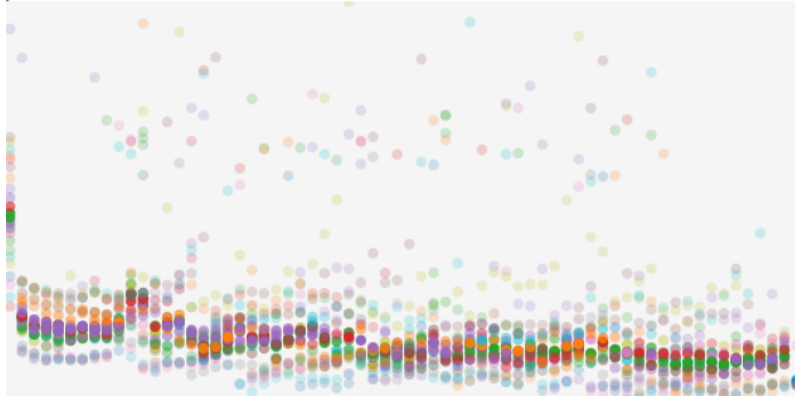
Overall, we believe this is the case since it not only highlights how competitive Formula 1 races are, but it also highlights how the spread between racers is fairly close. Also, the cumulative graph gives the false impression that the top line is the best time, which it clearly isn't since the lowest time is always the best. Since it is misleading to display the data this way, we decided to do away with using a cumulative graph and use only the lap-time graph above.

Implementation
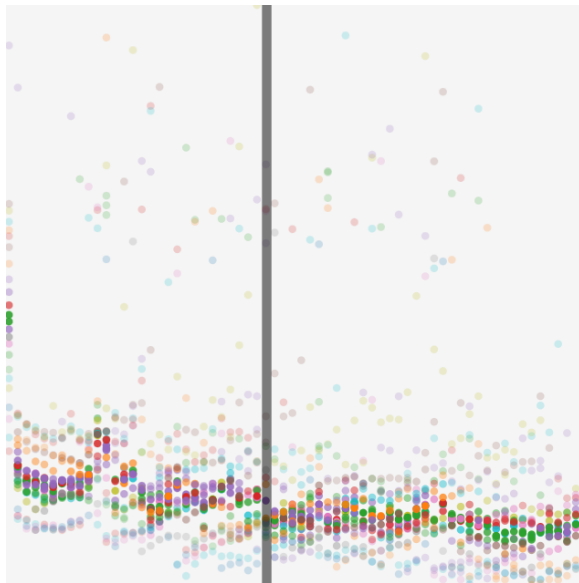
In this interaction, we wanted to use a rectangular bar to highlight all the racers (i.e. circles) in each lap. In this implementation, you can easily move your mouse across this particular svg to highlight different laps.
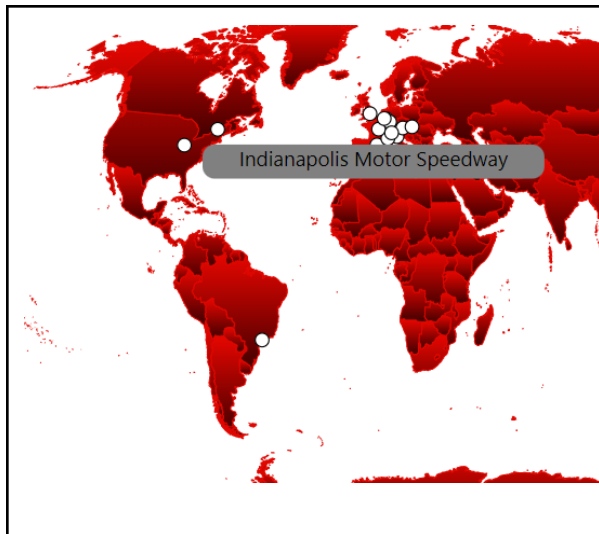
In the world map view, we wanted to display the race track's name when the user hovered over a particular circle. The purpose of this was to make it easier for users to find the name and location of particular races.
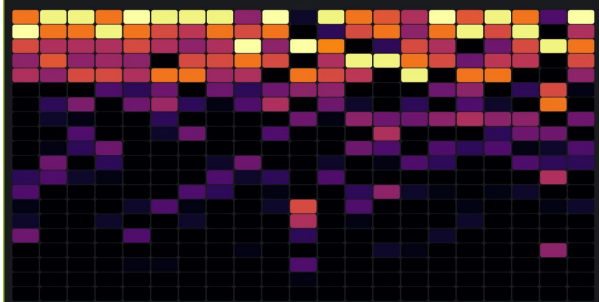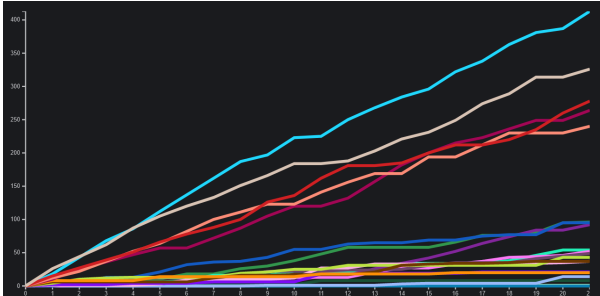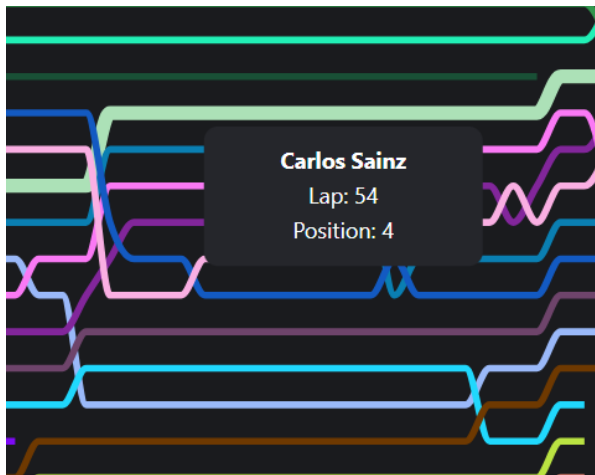

The multi-line chart to visualize position changes between drivers in a race.
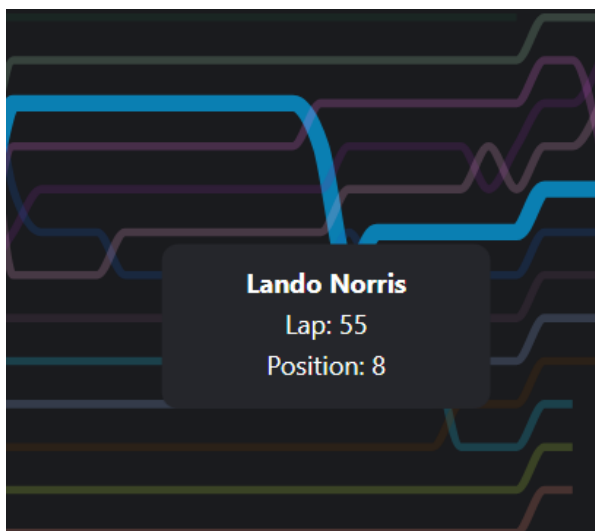
When the user clicks on a line, we lower the opacity of the other lines to draw attention to the selected line.

The user can select multiple lines to compare drivers.

	A heatmap to show which drivers are scoring the majority of the points at each race in a season.
	A kernel density chart to show where the majority of the position changes happened during a race. We also use a line chart on top of it to emphasize the fact.
	We show a line chart to show how teams and drivers progress as the season goes on.

Evaluation

We believe what's true for the race visualizations is how close each driver's times are to one another and how competitive each race is. Unlike sports such as College Football where it's usually two teams competing at a time in a game, the outcome of Formula 1 races are much more variable as more than ten teams and twenty drivers are competing in the same race at the same time. With the exception of top racers such as Lewis Hamilton and Michael Schumacher, Formula 1 races are hard to predict.