Process Book

Project Metadata

Project Title: F1Delta
Team Member: Srinanda Kishore Yallapragada
SPIRE ID: 32838512
Email: syallapragad@umass.edu
Project Repository: https://github.com/Srinanda-Yallapragada/F1Delta

Background and Motivation:

My main motivation to chose Formula 1 as a topic is simple. I am a fan! In recent years, F1 has witnessed a huge surge in popularity, particularly in the United States. The 2024-2025 season has already delivered some of the most shocking and historic moves in F1 history. The biggest among them is Lewis Hamilton's unexpected switch to Ferrari. This move seemed almost unimaginable given his world championships and legacy with Mercedes. Beyond Hamilton's transfer, the driver market has been extremely dynamic, with Esteban Ocon moving to Haas, Carlos Sainz heading to Williams, and Nico Hülkenberg moving to Sauber. While several new drivers are entering the grid, I aim to focus specifically on these driver shifts and the changes in their driving styles across the different cars.

Another reason for choosing this project is the immense amount of data available in F1. Being a highly technical sport, every aspect of performance (lap times, tire degradation, car telemetry data, throttle and break metrics, engine RPM and more) can be analyzed in depth. A lot of this data requires knowledge from different aspects of the sport, which may overwhelm new fans. I want to explore ways to visualize these technical aspects in a way that makes them more accessible to newer fans and helps them understand the technical aspects of the sport better.

Project Objectives:

Here are a few questions I want to explore

How do their results compare to their previous team's results? Are they finishing higher or lower than their old team's drivers? Is the move helping their career, or would they have been better off staying? Fans are always curious to know if a team switch was the right move or not, and this would answer that question.

How consistent are their race lap times? Answering this question could indicate how comfortable the drivers are in their new machinery.

How does switching teams impact a driver's racing style? Do drivers like Hamilton, Ocon, Sainz, and Hülkenberg change their braking points, racing lines, or cornering speeds when adapting to a new car? The answer to this question could indicate to us how much the driver has to adapt due to their team switch.

How have their qualifying push lap performances changed? Are their best qualifying lap times improving or worsening? Have their driving lines or braking points in qualifying changed with a new car? The answer to this question would indicate any strategy changes of the driver during qualification for the race.

An important thing to note about this comparison is that when the project is due, only 6 races of the new season will have taken place. This is necessary context to consider the visualized data. We also know that the drivers are likely to improve over time as they become more familiar with their new team and race engineers. However, that shouldn't stop us from asking interesting questions. These

drivers are paid millions of dollars to earn every point possible, and points scored early in the season are just as valuable as those scored later on, making this exploration worthwhile!

Another caveat to consider is Carlos Sainz's move from Ferrari, where he was racing near the front of the grid, to Williams, a team that typically finishes towards the back. His performance differences will likely be heavily influenced by the car's capabilities (or so I believe), which will make it particularly interesting to investigate.

Data:

I plan to use https://openf1.org/. OpenF1 is a free and open-source API that provides real-time and historical Formula 1 data. The entire website would fully be reliant on this data source and so I would have no local collection of any data. This source does update its dataset with about a 3 second delay from when an event happens on track, which may allow me to have a realtime visualization component, however this would be a reach goal of the project.

I will also occasionally reference https://www.statsf1.com/en/statistiques/pilote.aspx for static data.

Formula 1 cars get faster every single year. This means that there is a chance that the data I am analyzing simply favors the most recent car in all cases. In this case I will have to pivot my strategy of analysis. I would instead compare the current driver with the driver who filled their seat as a form of comparison.

Data Processing:

I do not expect significant data clean up as the OpenF1 documentation is quite clear on what type of data is being returned and in what format it is available. I will also be able to filter what data I pull from the API based on time which will make it easier to pull relevant data based on lap times.

I will need to do some calculations to derive new quantities of data such as minimum speed in a corner but these should not take too long and will be calculated realtime using the latest API data available.

Visualization Design:

Since the audience of this project is the general audience, my homepage includes overall statistics of the driver. These are generally very easy to understand with no background and consist of simple ideas like podium finishes and total number of fastest laps etc. The home page also has a list of all the races that have taken place. These are represented as "thumbnails" so that the user can pick each race that they want to investigate themselves. Since each race is unique, it is not possible to aggregate improvements across multiple races due to the differences in location and track layout.

On the left of the thumbnails, I will have the race results from 2024. On the right, I will have the results of 2025. If the race has not taken place yet, I will grey the results out.

On page two, I elected to pick an overall track view which labels all the existing corners. It also shades the part of the track where the previous and current teams are faster on the track. This gives a clear visual indication on which year's car was faster in what parts of the track. the page itself will have a separation down the middle indicating the left of the page is 2024 and the right is 2025. I will present some statistics in the form of bars which will quickly convey which statistic was better. The better statistic will be highlighted in the years particular shading, likely in the color of the team.

On a click of any particular turn, the website will take you to page three where we can see more detailed statistics. This allows users to see deeper statistics per turn. Here the user can see a lap by lap breakdown of each turns breaking points and other telemetry data like minimum speed and the line taken through the corner. A line on top of a section of the turn is the most intuitive way to

convey the path taken by a driver across the turn. The line data is not shown on page 2 because I belive it would clutter the visualization too much.

Must-Have Features:

List the features without which you would consider your project to have failed.

- Must have the quick stats page on page 1.
- Must have the track overview with completion statistics on page 1.
- Must have the faster sector visualization on page 2.
- Must have the bar graph visualization of race stats on page 2.

Optional Features:

- The turn statistics while nice to investigate are not needed to gain insight into how the driver is fitting into his new team.
- The line the driver takes is important to see a change in the driving style, but may prove to be a technical challenge to finish before the semester ends. I may have to visualize the entire line taken across a lap instead of breaking it into a corner by corner basis. This would lose the breakpoint data and the throttle gradient information.

Project Schedule:

remaining weeks 8

- week 1: Page 1 quick stats
- week 2: Track thumbnails + stats
- week 3: Track overview + faster sector highlighting.
- week 4: Race statistics + manual introduction of context like safety cars
- week 5: Turn statistics graphs
- week 6: Turn statistics lines
- week 7: Turn statistics lines
- week 8: Overall polish.

F1 DELTA

Data points
- Break pressed
- DRS pedal on
- Car gear
- Speed
- throttle

car data

- Driver details
- images
- team name
- team color

driver details

- time to leader
- fastest lap
- quali lap time
- lap times race
- Sector times

timing data

- track data
- car location on track
- Svg of track
- Weather

track details

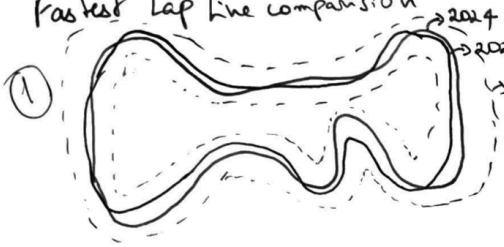
Fastest Sectors over Race distance

blue faster in corners
black faster on straights

Free Practice (5)

this would be cool.

Fastest Lap line comparison



(1)

↳ Appox shape of track.

(6)

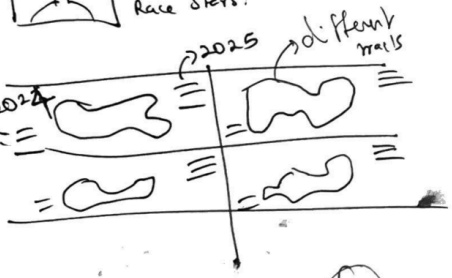
turnb Zoom in & show narrow line difference.

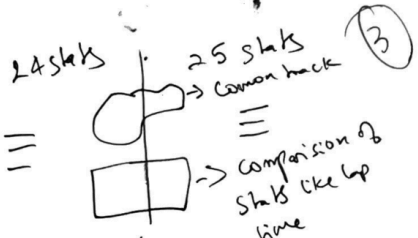
corner by corner comparision hard but good idea.

Home Page

Quick Stats
Podiums:
Wins:
Race Stats:
etc etc.

(2)





(3)

1 Page per track.
+ Notable events.

Filter:

good like ideas (2) (3) (6) (4).

would not do (5) as it is too confusing. The scales would not make sense.

Categorize:

(1) can be the overview of the track while

(6) happens when you click on a corner.

Each page would need contextualizing information.

Combine & Refine:

(2) can be the overview with multiple displays & interactable. when you click on a map, it brings up (1) with (3) below it contextualizing the info.

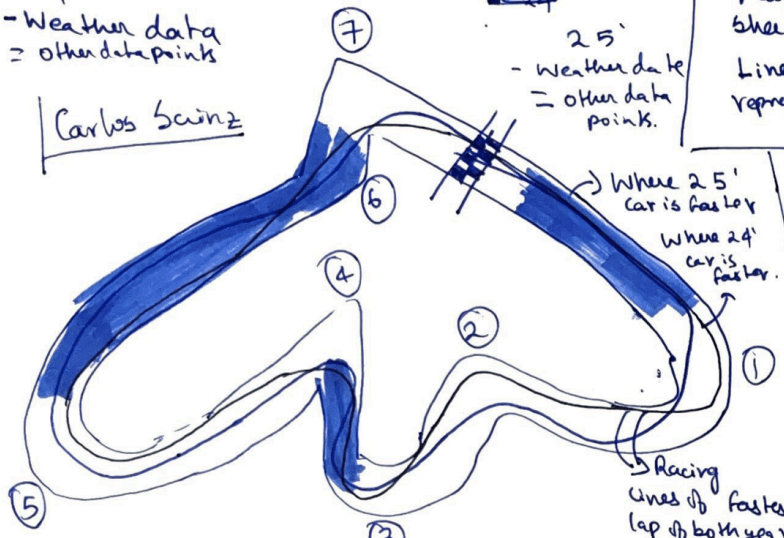
② MELBOURNE 24 vs 25

Srinanda Kishore Yallapragada
March 12, 2025
Sheet 2, F1 Delta
Line that driver takes representation on the track

24' ▷ quali | ▷ lap 3

- Weather data
= other data points

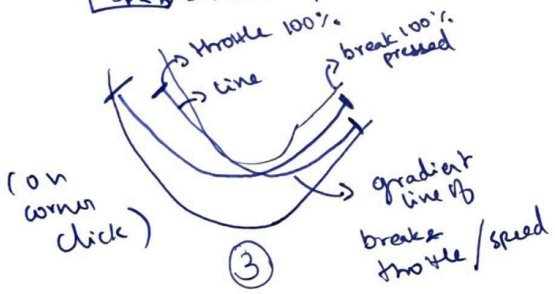
25'
- Weather data
= other data points.

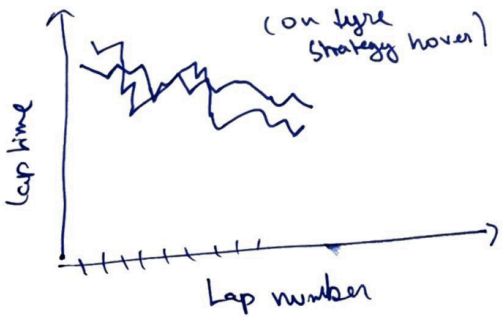


Soft 25	Med 45	Soft	57
~~MED~~ 2	Soft 33	~~MED~~	57

Tyre Strategy.

Operations
- Change the driver
- Select a corner to bring up breaking data & increase the size of the corner to ~~focus~~ focus on it.
- Can click on tyre strategy to see data on tyre stint as a graph.
- Select quali or lap number of race

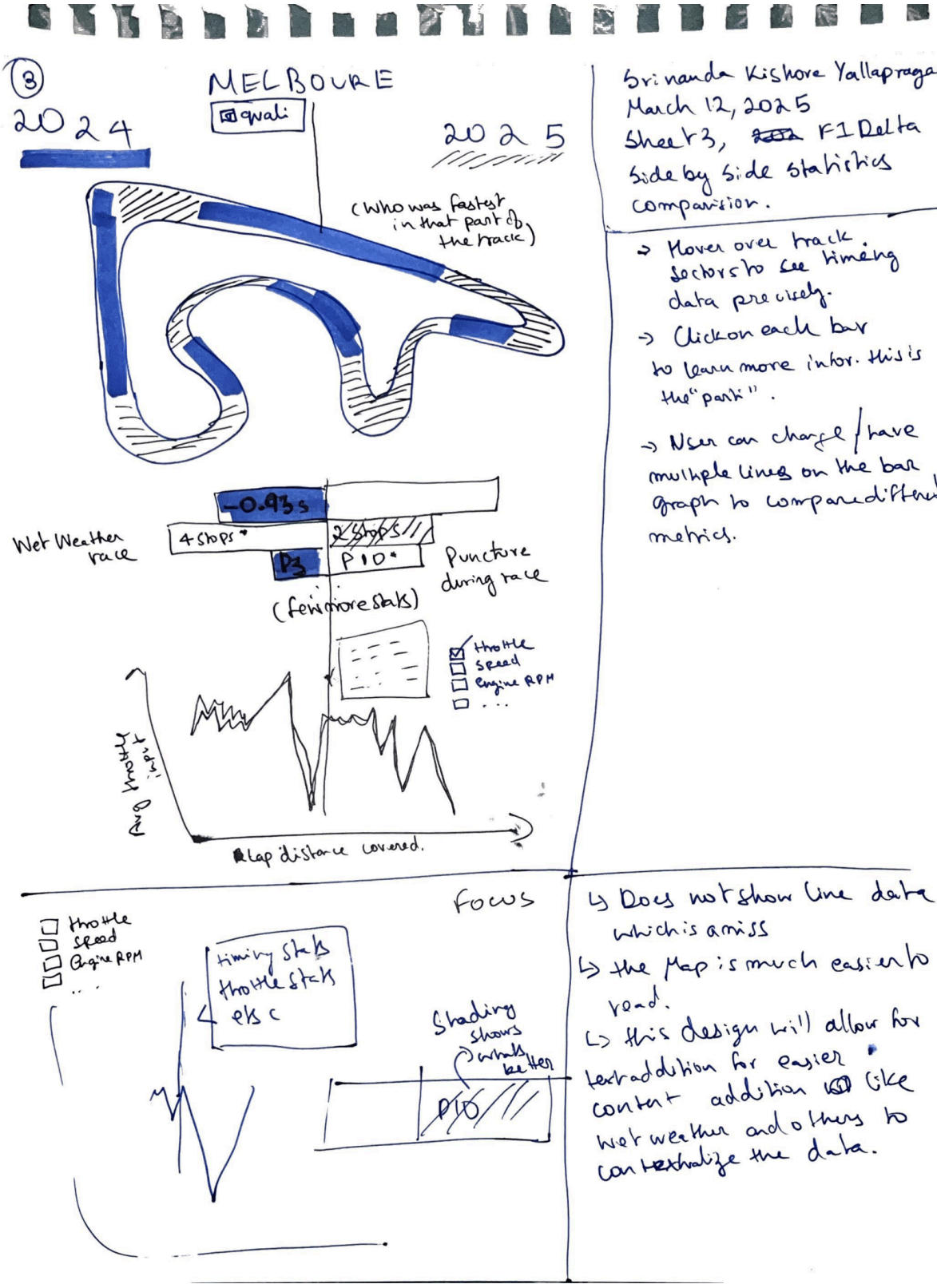




Flows

Discussion

- Tyre strategy may just be better as a graph
- What are the relevant stats
  - ↳ Safety cars, red flags
  - ↳ weather, wind
- Is gradient enough to show throttle inputs
- When in a corner?
  - ↳ I don't think so because from 100 break to 100 throttle is a wide margin.
- Would an aggregate lap line make sense?
  - ↳ Yes Avg the line because of overtakes skewing the data.



Srinanda Kishore Yallapragada
March 12, 2025
Sheet 3, ~~2222~~ F1 Delta
Side by Side Statistics
comparision.

→ Hover over track sectors to see timing data precisely.

→ Click on each bar to learn more infor. this is the "park".

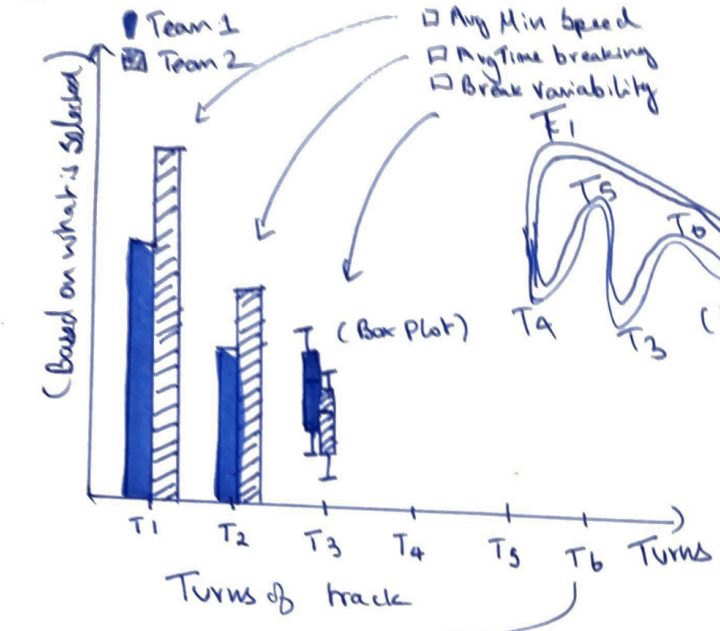
→ User can change / have multiple lines on the bar graph to compare different metrics.

↳ Does not show line data which is amiss

↳ the map is much easier to read.

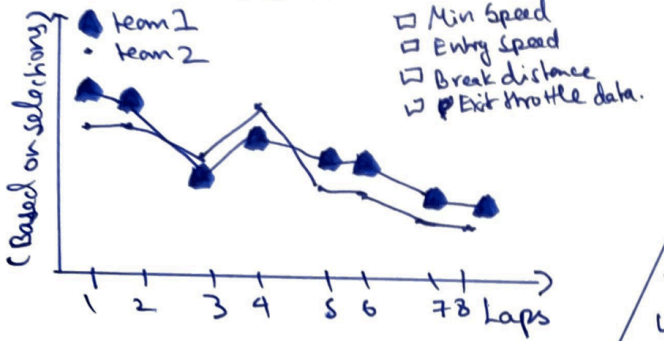
↳ this design will allow for text addition for easier content addition like wet weather and others to contextualize the data.

④ MELBOURNE RACE



(Select a particular turn)

Turn 6



Srinanda Kishore Yallapragada
March 13, 2025
Sheet 4, P1 Delta
Track turn based visualization of driving style.

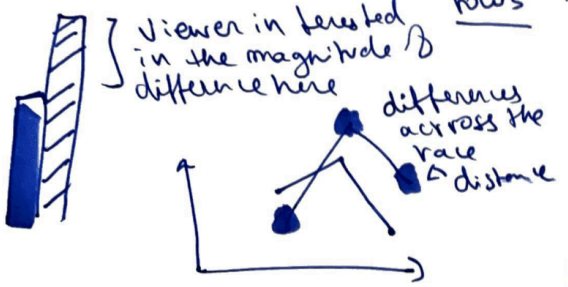
- Can interact with the multiple boxes to select what data to show on the Y axis. This would be aggregated across the full race
- Select a particular turn to bring up a lap by lap description of the available data. Similar to the other graph, we can pick what's on the y axis.

Operations

Discussion

- Very mathematical as a visualization of data.
  ↳ Can be a live highlight across the track?
- User load is high due to selecting many interface elements
- It is effective as a quantative analysis. But the goal is for new comers to the sport.
  ↳ Might use some elements in the final design.

Flows

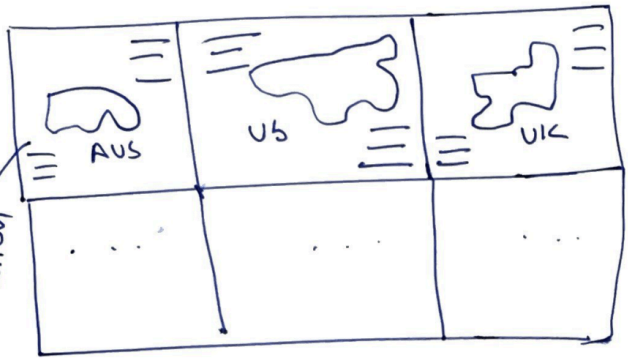


Srinanda Kithore Yallapragada
Sheet 5, March 13 2025
F1 Delta
Task: Selection of Visualizations and compiling ideas

⑤ Page 1 (Home Page)

Quick Stats
- Podiums
- Wins
- Poles
- Race stats
- Current team
- Previous team

List of tracks Overview



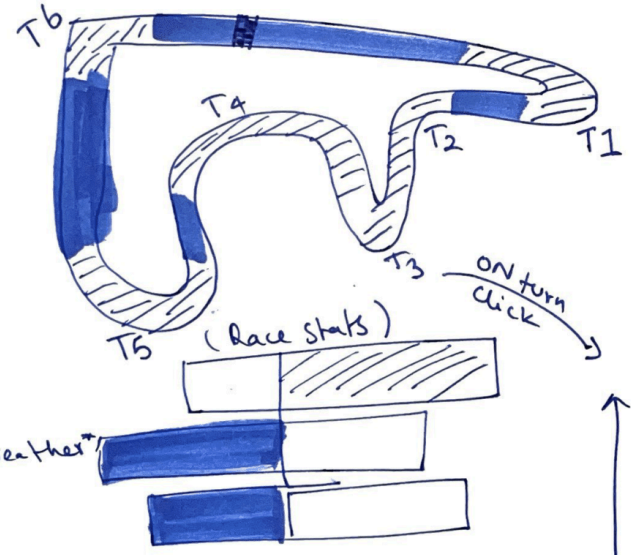
On track selection

Operations
- Select track to view
- Select driver to view
- Select turn of track
- Select turn data
- hover on track, bar & line graphs for numerical data
- Interaction with check boxes

Page 2 (Race Overview Comparision)

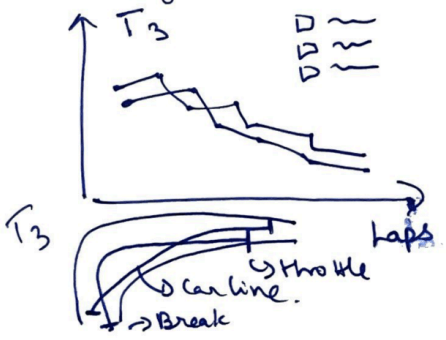
2024

2025



Details
- Requires data manipulation to calculate certain statistics
- Data needs manual inspection for Red flags
- Manually add context for race on Page 2
- Sector calculations need to learn
- Quick stats should update with API.
- Races line calc is hard
- detection of a turn is hard.

Page 3 (Turn Statistics)



Project Milestone additional information

The structure of the website has been established with this milestone. Since a lot of the data is static, I have decided not to use any backend services and have the front-end directly communicate with the api's themselves. This makes it easier for development, and running the server simply is a python command with the public folder so deploying this will also not be difficult. I will likely be using github pages itself to deploy this website.

I underestimated the amount of data processing I would need to do for this project. Unfortunately, I am finding it difficult to find an api that directly gives f1 statistics of race results. Aggregated statistics are not available and thus need to be calculated on my own. I have already sourced some of these statistics like number of wins, poles and podiums, but this will need additional manual code to gain all the statistics I would want to display in this project. I downloaded some of the results data directly as a json rather than trying to use the api so that I could directly pull out the race results data that I need. This data has come from https://github.com/jolpica/jolpica-f1 which is another service that provides f1 data. However, no free service has a direct statistics for each driver, they all mainly have the results.

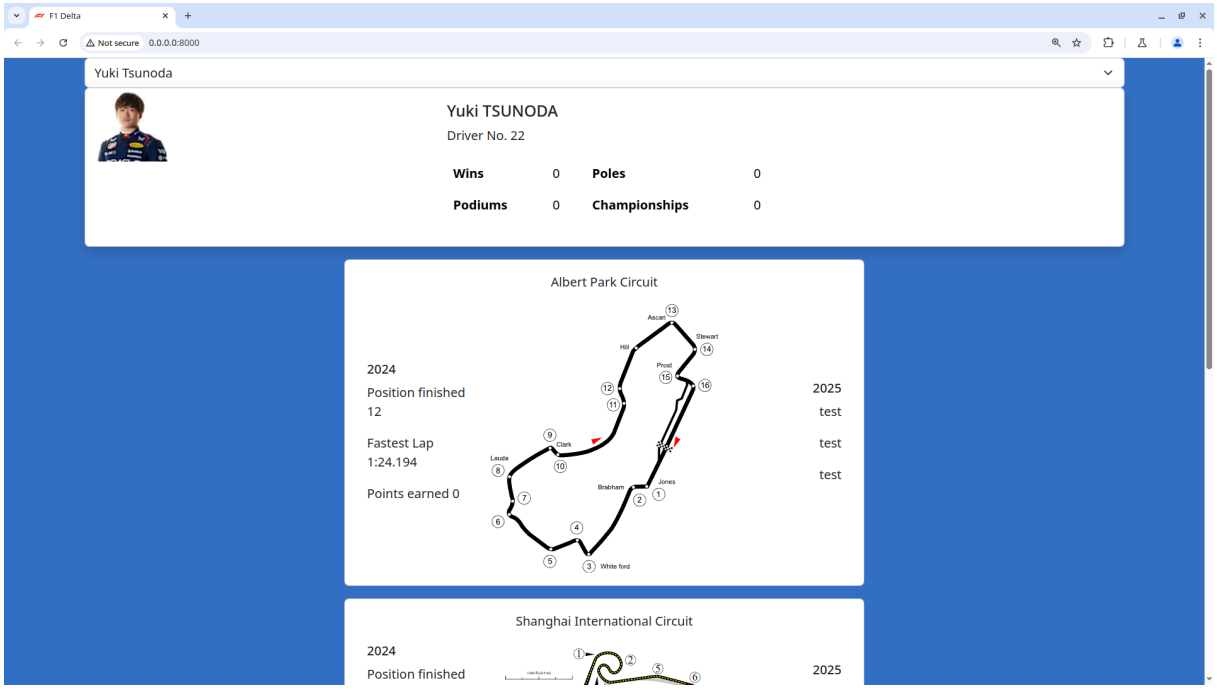
As of now the landing page provides a comparative overview of who the driver is in terms of their overall statistics like wins and poles, and then shows how much improvement is present across each race of the season that has occurred so far. The race data is pulled from an api, so as new races occur, the website automatically adds in additional race track cards. All geojson files have been acquired and loaded into the website. I will download the additional svg files as needed.

The project now includes drivers who have remained with the same teams, providing a broader comparative analysis. Rookies have been excluded due to the lack of previous years racing data.

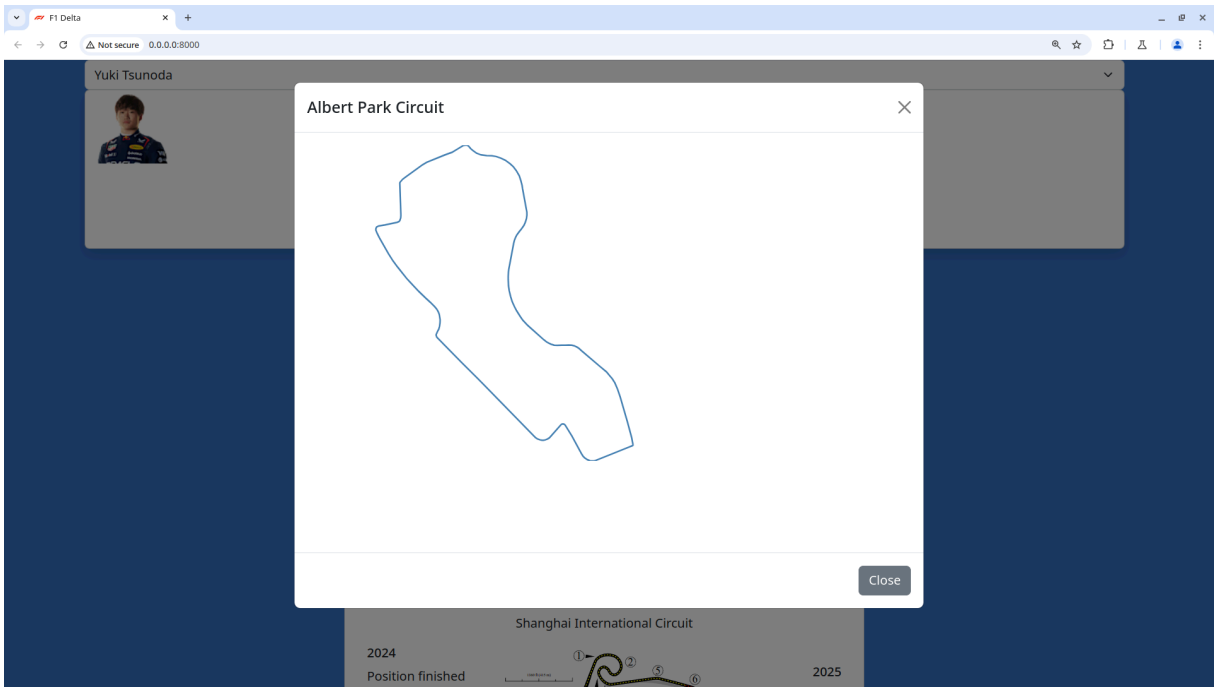
Unfortunately , during the initial design of the website, I did not account for the fact that formula 1 cars across the seasons improve drastically in their performance. I will need to change some of the visualization goals to focus on metrics that are not as dependant on absolute car performance as all the cars will generally be going faster than the previous years. this means that data like sector times will not provide additional value as the 2025 cars will simply be faster across the entire track. the visualization may provide some anomalies, but it does not answer general questions about the drivers improvements.

I added the team color of the racing driver in the background that changes as you change the driver which is a visual cue as to which team they are looking at. This would be an indicator for those who are familiar with the teams colors that we see on track.

Below are some screenshots from the work in progress website



below is the loaded geojson of the particular track



Final Submission

Project Metadata

Project Title: F1Delta
Team Member: Srinanda Kishore Yallapragada
SPIRE ID: 32838512
Email: syallapragad@umass.edu
Project Repository: https://github.com/Srinanda-Yallapragada/F1Delta

Overview and Motivation:

My main motivation to chose Formula 1 as a topic is simple. I am a fan! In recent years, F1 has witnessed a huge surge in popularity, particularly in the United States. The 2024-2025 season has already delivered some of the most shocking and historic moves in F1 history. The biggest among them is Lewis Hamilton's unexpected switch to Ferrari. This move seemed almost unimaginable given his world championships and legacy with Mercedes. Beyond Hamilton's transfer, the driver market has been extremely dynamic, with Esteban Ocon moving to Haas, Carlos Sainz heading to Williams, and Nico Hülkenberg moving to Sauber. While several new drivers are entering the grid, I aim to focus specifically on these driver shifts and the changes in their driving styles across the different cars.

Another reason for choosing this project is the immense amount of data available in F1. Being a highly technical sport, every aspect of performance (lap times, tire degradation, car telemetry data, throttle and break metrics, engine RPM and more) can be analyzed in depth. A lot of this data requires knowledge from different aspects of the sport, which may overwhelm new fans. I want to explore ways to visualize these technical aspects in a way that makes them more accessible to newer fans and helps them understand the technical aspects of the sport better.

Related Work:

Anything that inspired you, such as a paper, a website, visualizations we discussed in class, etc.

The project drew inspiration and utilized data from several sources:

1. OpenF1 (https://openf1.org/): Initially considered as the primary data source for real-time and historical F1 data.
2. StatsF1 (https://www.statsf1.com/en/statistiques/pilote.aspx): Referenced for static driver statistics.
3. jolpica-f1 (https://github.com/jolpica/jolpica-f1): Used as a source for downloaded race results data in JSON format, particularly when aggregated statistics were not readily available via APIs.
4. F1TV Tech Talk: The show's telemetry data presentations served as an inspiration for the types of visualizations to explore.
5. FastF1 Python API: This became the primary tool for accessing and processing F1 data after initial challenges with other sources.(https://docs.fastf1.dev/gen_modules/examples_gallery/index.html)
6. https://rapit.com.br/ which is an f1 visualization I found on the internet

Questions:

The central question driving this project was: How does switching teams impact a Formula 1 driver's performance and racing style?

Initially, the project aimed to explore:

1. How do their results compare to their previous team's results?

2. How consistent are their race lap times in the new car?
3. How does their racing style (braking points, racing lines, cornering speeds) change?
4. How have their qualifying push lap performances changed?

Over the course of the project, these questions evolved:

1. The initial focus on direct lap time comparisons between seasons (e.g., 2024 vs. 2025) was re-evaluated. It became apparent that year-on-year car development in F1 leads to significantly faster cars, making direct time comparisons less indicative of driver adaptation and more a reflection of car improvement. This meant that visualizations based purely on sector times or lap times showing 2025 cars as faster across the board wouldn't provide deep insights into driver-specific changes.
2. A new question was introduced: How do drivers perform relative to their teammates in both their old and new teams? As I was going through the project I realized that performance relative to the drivers teammate is also judged quite a lot in f1.A lot of drivers have often said that "In F1 your teammate is also your worst enemy".

Ultimately, the visualizations aim to answer:

1. How do the drivers' qualifying lap performances compare to their previous team/season?
2. How do drivers manage the different engine characteristics and car dynamics of their new teams, as reflected in telemetry data like speed and RPM?
3. In which parts of the track do drivers carry more or less speed compared to their previous car/setup?
4. How did the drivers fare in races relative to their teammates from both the current and previous year?

Data:

The project utilized data from several F1 data sources:

The project initially intended to heavily rely on OpenF1 (https://openf1.org/) for its free and open-source API, aiming for real-time and historical data. StatsF1 (https://www.statsf1.com/en/statistiques/pilote.aspx) was used as a supplementary source for static driver statistics like total wins, poles, and podiums.

I found that it was not trivial to work with the data from OpenF1 so I switched the project to use the FastF1 Python API. This library proved to be far more convenient for accessing and processing detailed F1 telemetry and session data.

Data was collected primarily via API calls using the FastF1 library. This included session information, lap timings, and car telemetry data (speed, RPM, positions, etc.).

Data cleaning was minimal after switching to FastF1, as the library provides relatively clean and structured data.

A significant data processing step was data extrapolation for telemetry readings. Telemetry data from different cars (and even for the same car across different laps or sessions) are often not perfectly synchronized in time. To make meaningful comparisons it was necessary to extrapolate values to align them onto a common time or distance axis.

Exploratory Data Analysis:

What visualizations did you use to initially look at your data? What insights did you gain? How did these insights inform your design?

I reviewed existing Formula 1 data visualization websites to understand common ways F1 data is presented and to identify potential visualization types. A lot of these I found on reddit made by other users. a few are listed earlier in the project.

I also am a fan of F1TV's "Tech Talk" segments, which often break down telemetry data (like speed traces and gear shifts) to explain driver performance and car differences. This served as an inspiration for the types of telemetry visualizations that ended up in my project

I found while watching this show that they often use a touch screen and zoom in on different parts of the telemetry so I wanted that to be a feature of my graphs as well.

After the semester midpoint deadline, I ended up switching the entire tech stack as while d3js is a really good visualization tool, I found that working in python especially given the accessibility of tools for data visualization like matplotlib plotly streamlit and pandas and more made it much easier and faster to build websites for interactive visualizations. However the website performance is terrible so please be patient as the website loads.

Design Evolution:
What are the different visualizations you considered? Justify the design decisions you made using the perceptual and design principles you learned in the course. Did you deviate from your proposal?

I deviated significantly. I really wanted to have some kind of driver racing line visualization incorporated here as I find that to the the most interesting information but I could not make that work using the geojson files and the telemetry data. One of the issues I faced was that I could not accurately map the geojson data and the telemetry data effectively enough that I could trust the data that I was visualizing.

I tried to overcome this by only visualizing the driver laps without the track underneath it, but then I ran into the issues of not even being able to see the differences in the driver lines taken. At that point I gave up on trying to visualize the driver line taken.

I initially in the proposal considered having timing based data, but that ended up being a problem. I overlooked a crucial detail which was that f1 is not a static sport. The cars get so much faster each year to the point that the slowest team on the grid in 2025 will be driving close to or better than the fasters teams of 2024. This means that a lot of this timing data visualization would not provide any useful information into the change of driving.

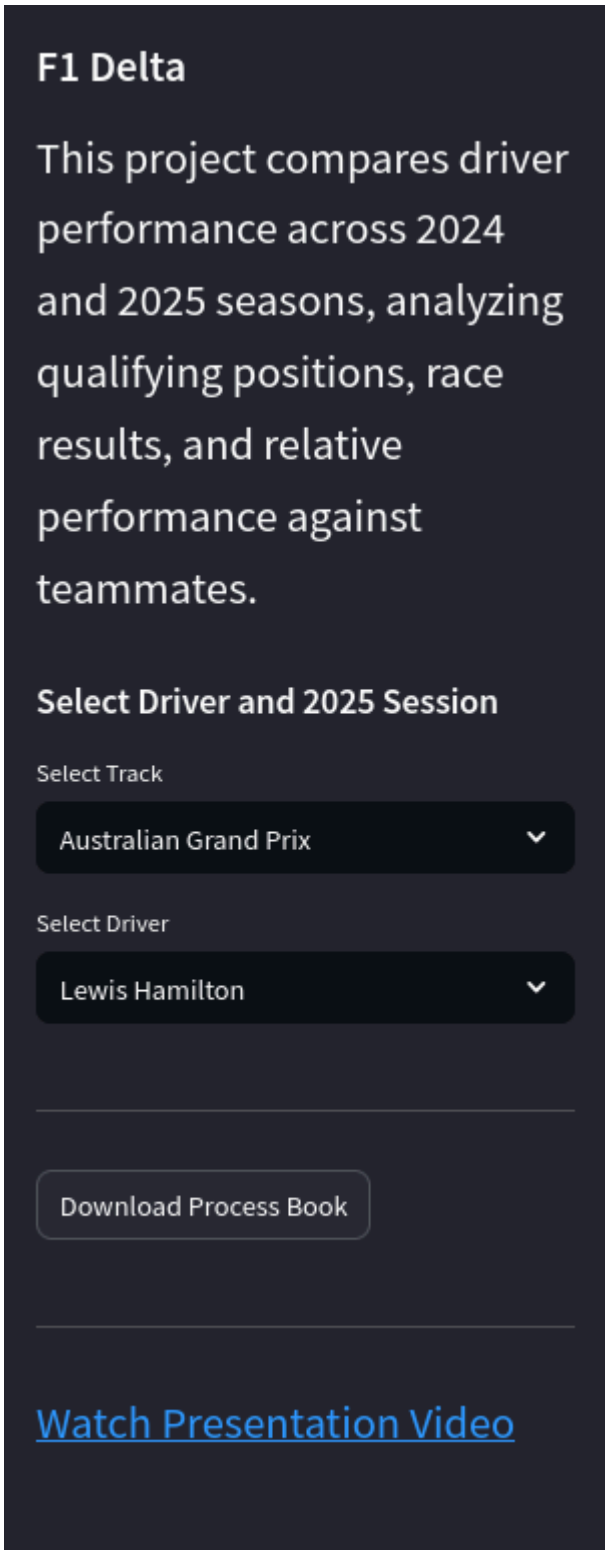
After the mid semester deadline. I found it quite frustrating to work with the open f1 api. I should have done a better job at picking the tools and started on python from the start. I switched the entire tech stack over to python to use fastf1 which meant a major rewrite and that significantly reduced the scope that I initially wanted to achieve. This also meant a lot of wasted hours that did not count towards the final project which is someth8ing I would avoid if I could do this project again. This was also due to some personal health issues which prevented my from realizing the full potential of this visualization.

I knew after the rewrite that I did not want to try experimental visualizations so I stuck with the tried and true line graphs for the speed and engine RPM visualization data. There are not a lot of other way to visualization that data effectively especially considering that I wanted to be able to zoom into the graph and have it be interactive.

I also found it really effective to color code important data with the team colors of the driver. It creates a lot of contrast and easily allows the viewer to at a glance know which year the data is referring to when they take a look at it.

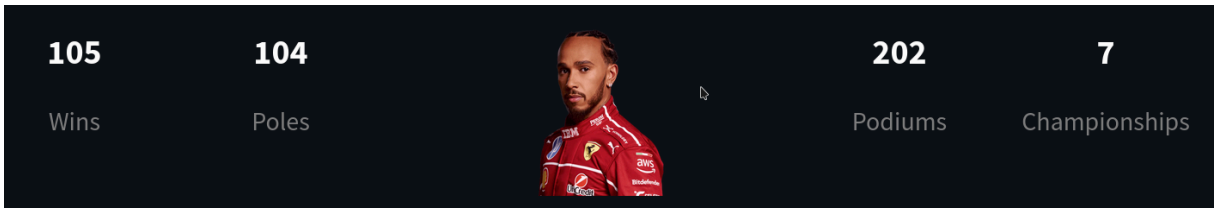
Implementation:
Describe the intent and functionality of the interactive visualizations you implemented. Provide clear and well-referenced images showing the key design and interaction elements.

Driver and Session selector



- This is a straightforward drop down toggle to select the different drivers and sessions to load the visualizations for.

Driver overview



- Loads the lifetime stats and an image of the drivers. It gives the data a face to associate with which I found to be helpful.

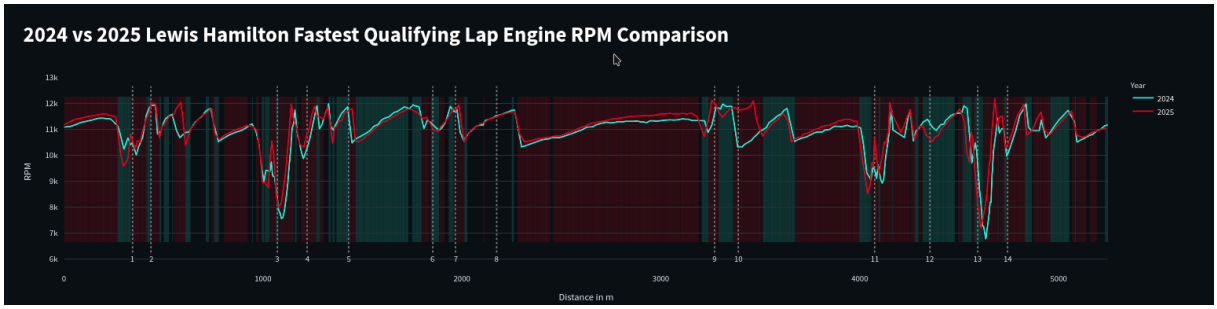
Track SVG

- The track svg gives the viewer information on where the corners are with the different corners being labeled.
- These labels are useful for the future graphs because they map the numeric data to points on the track which is useful in translating the data we are looking at into a tangible thing.

Written overview

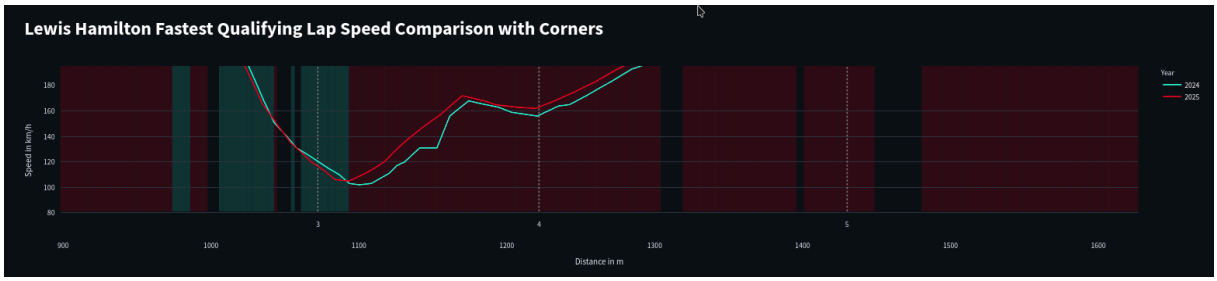
- Simple text generation based on results. it will say better or worse based on the results.
- color coded to the teams colors which makes it easier to distinguish the results through the rest of the page and the visualizations.

RPM engine visualization



- the background color of higher value is highlighted. The threshold for it to be highlight is 1% threshold. This makes it easier to see which car was higher on the graph without having to zoom in immediately or use the tooltip.
- We can click on 2024 and 2025 to disable and enable the lines incase the user wants to only see one line.
- We see the dotted lines that show where across the lap the driver is. These dotted lines correspond to the corner numbers which are shown in the svg of the track above.
- Hovering over the lines brings up a tooltip for more precise information of the numerical data.

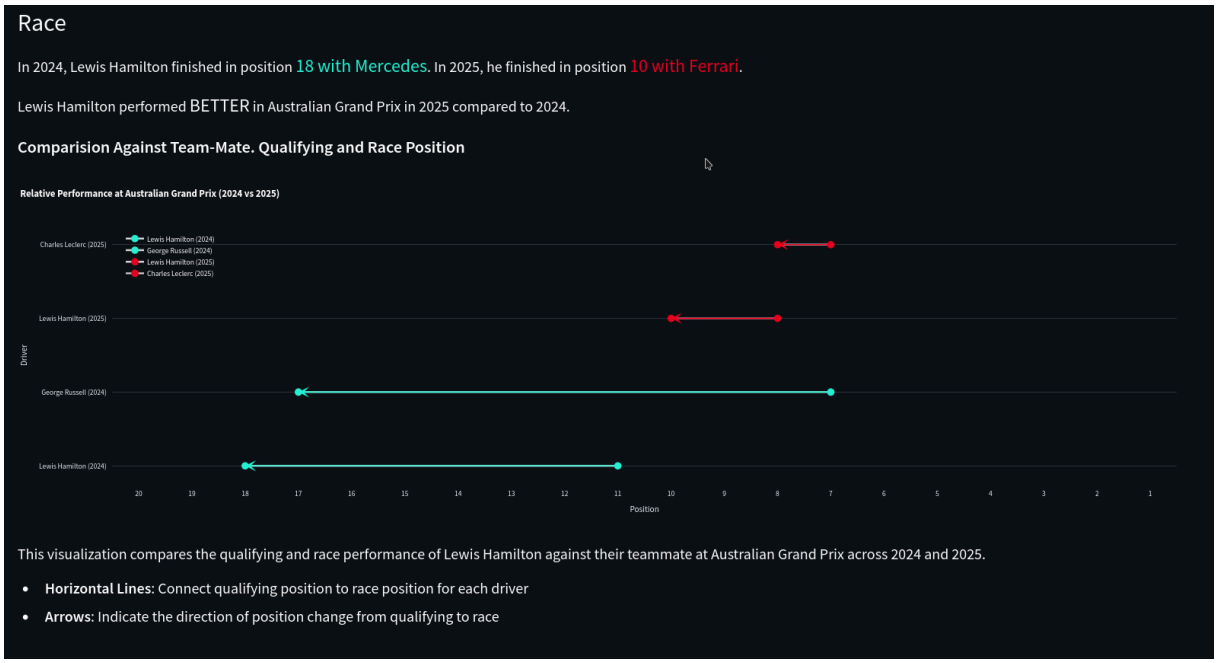
Speed visualization



- It is an iterative graph as you can zoom into the graph to view different parts of the data in closer detail

- options to pan, auto scale, full screen, zoom in zoom out and reset on all the visualizations.

race visualization



- There is an arrow that indicates if the driver went forwards or backwards during the race
- Comparison with teammates easily visible on the same axis

Evaluation:

What did you learn about the data by using your visualizations? How did you answer your research questions? How well does your visualization work, and how could you further improve it? (Be honest here. Limitations are a part of any project, and they will be noticeable during the grading process. Acknowledging them in this section indicates thoughtfulness in your design process, and, as such, will only help your grade.)

What did you learn about the data by using your visualizations?

1. The visualizations revealed specific driving pattern changes. For instance, it was observed that Lewis Hamilton and Carlos Sainz exhibited a trend of entering racing corners at a slower speed in some comparisons, a pattern also noted for a few other drivers. This was in the Australian grand prix which was the first race of the season. I am choosing to interpret this as that the drivers are not confident in their new cars. So while the new cars are overall faster, the drivers have some kind of hesitation and are not fully committing to later breaking points. This shows that they are not fully confident in their new cars which makes sense as this is the first race with their new team.
2. I found that the position graph which compared a driver's race progress to their teammate's to be highly effective. It clearly conveyed relative performance, showing whether a driver moved forwards or backwards from their qualifying position and how their teammate fared, which is a strong indicator of performance within the F1 context.

How did you answer your research questions?

1. How do the drivers' qualifying lap performances compare to their previous team/season?

   By looking at the drivers overall lap time, position qualified and telemetry data. It gives us information about how they drove the car through the telemetry. It provides an overall idea of performance by comparing the qualifying positions which is relative to the entire grid. This gives us context on how the season in going.

2. How do drivers manage the different engine characteristics and car dynamics of their new teams, as reflected in telemetry data like speed and RPM?

   Using the telemetry graphs, we can see in which areas which cars were faster and how the engine responds to the drivers. We can see spikes in RPM and differences in entry and exit speed from the graphs.

3. In which parts of the track do drivers carry more or less speed compared to their previous car/ setup?

   Generally, we see that the drivers who have changed speeds are carrying less speed into the entry of corner in the track while they are hitting higher speeds through the straights.

4. How did the drivers fare in races relative to their teammates from both the current and previous year?

   Using the position graph, we can see the relative performance of the drivers to their teammates. It shows forwards or backwards progress through the race and relative finishing positions to give a visual indication of the performance of the drivers.

Some of the limitations are

1. The current visualizations, while informative at a surface level, do not fully capture the deep nuances of driver adaptation, such as precise changes in braking points or detailed racing line alterations. This was an initial goal that could not be realized due to technical challenges in accurately mapping telemetry to track visuals and the scope reduction from the tech stack change. The visualization also do not provide an overview of enough context. There are plenty more telemetry data and exceptions to the telemetry data such as safety cars, accidents and more that are ignored for the sake of this visualization. Unfortunately it would be too much effort to incorporate these especially after the rewrite.
2. For more complex plots like RPM traces, adding annotations or brief explanations of what to look for could help users derive more meaning. Currently there is a general set of words for every graph but having graphics that point out what to look for in every graph would be interesting to have. This can be auto generated (for example highlight in a bright color when there is a difference of over 10 % let say) and would make it easier to parse the technical graphs data.