

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 believe 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.

①

F1 DELTA

Data points

- Break pressed
- DRS pedaled on
- Gear gear
- Speed
- Throttle
- Track data
- Car location on track
- Svg of track
- Weather

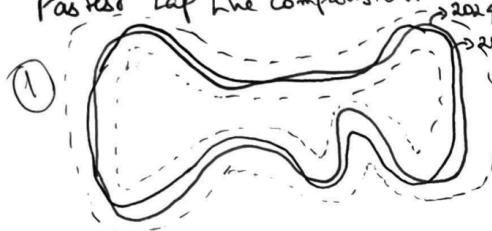
- Driver details
- Images
- Team name
- Team color

- Time to leader
- Fastest lap
- Quali lap time
- Lap times rate
- Sector times

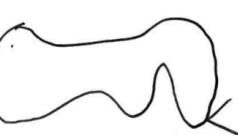
Timing
data

Fastest Lap line comparison

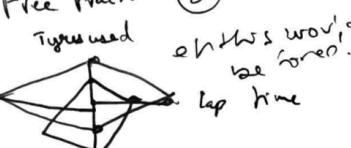
①



corner by corner
turn by turn zoom in & show
narrow line difference.



Free Practice ⑤



total num laps

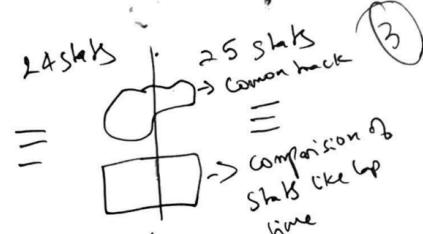
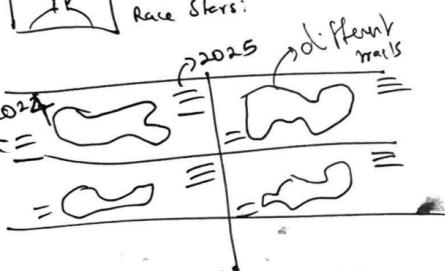
Home Page

②

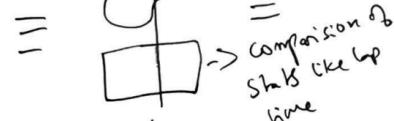


Quick stats
Podiums:
Wins:
Race Starts:

etc etc.



=
1 page per track.
+ Notable events.



Filter:

g like ideals ② ③ ⑥ ⑦.

g 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:

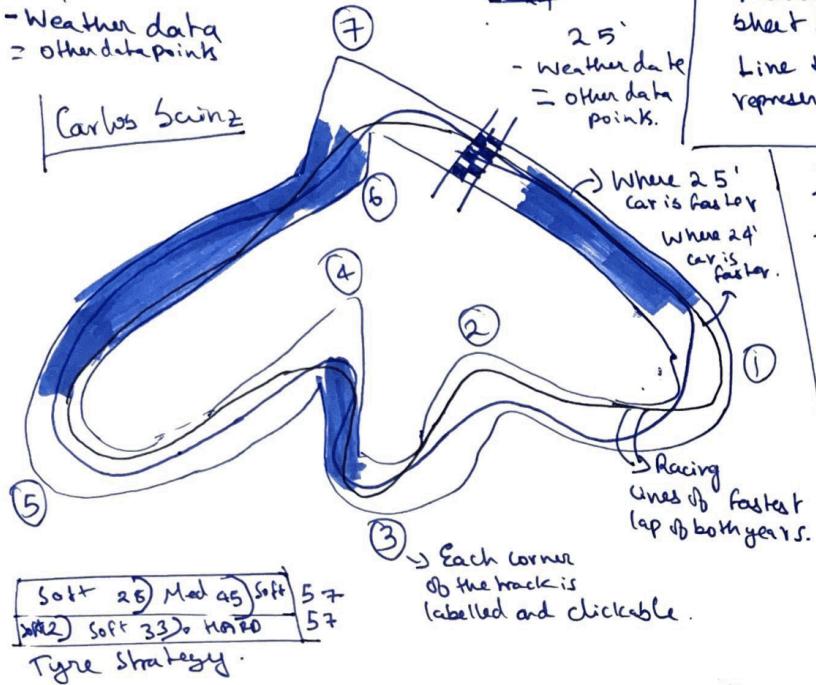
- ② can be the overview with multiple displays & interactive. When you click on a map, it brings up ① with ③ below it contextualizing the info.

② MELBOURNE 24 vs 25

D qual 1 D lap 3

- Weather data
= other data points

Carlos Sainz



Srinanda Kishore Yallapragada

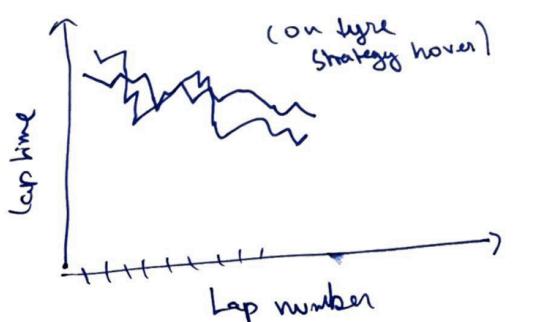
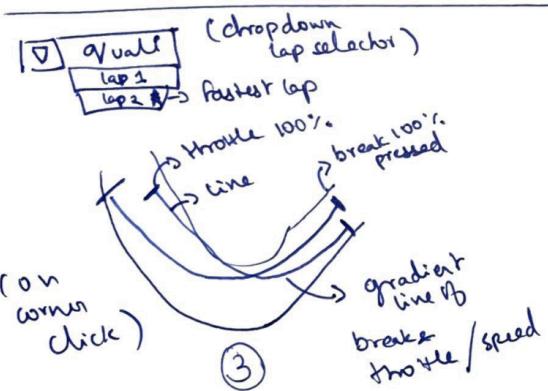
March 12, 2025

Sheet 2, F1 Delta

Line that driver takes
representation on the track

Operations

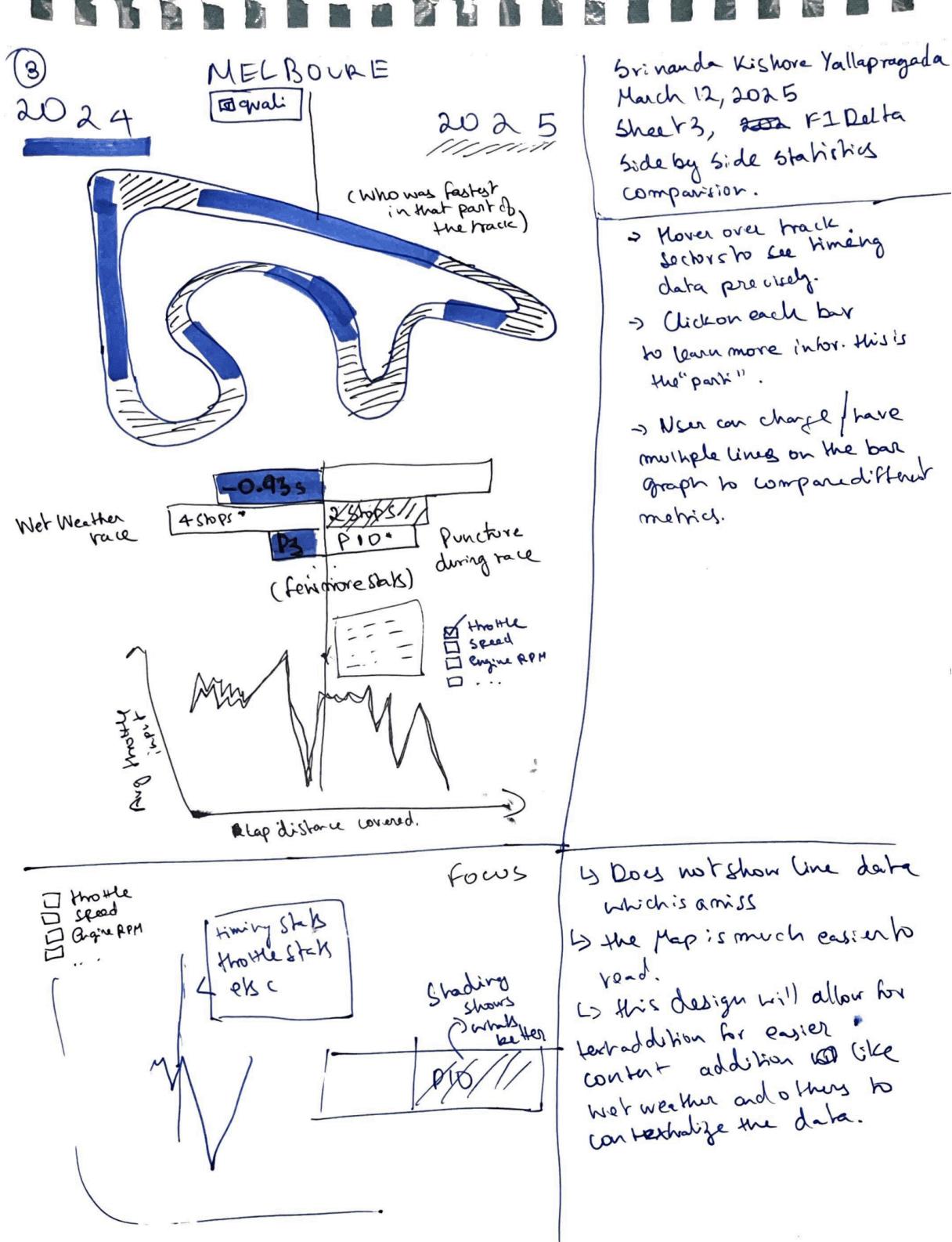
- Change the driver
- Select a corner to bring up breaking data & increase the size of the corner to focus on it.
- Can click on tyre strategy to see data on tyre stint as a graph.
- Select qual or lap number of race



Rows

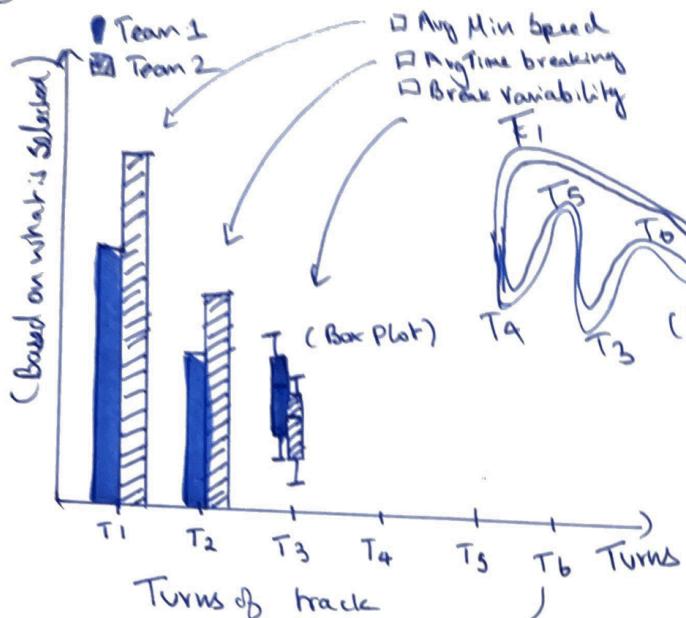
Discussion

- 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 0% to 100% throttle is a wide margin.
- Would an aggregate lap line make sense?
 - ↳ Yes Avg the line because of overtaking skewing the data.

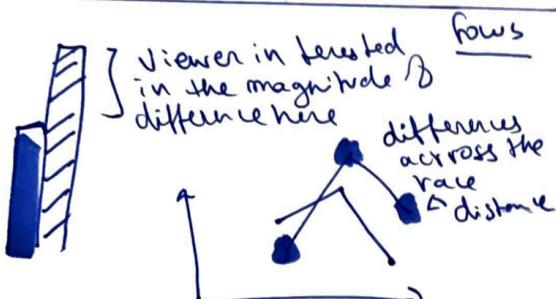
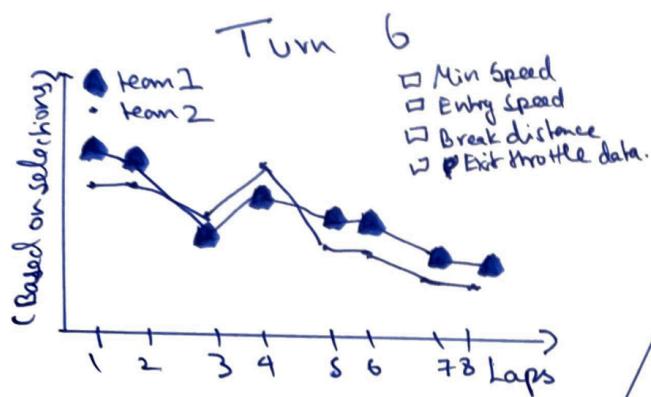


④

MELBOURNE RACE



(Select a particular turn)



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 graphs, 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
- GT is effective as a quantitative analysis. But the goal is for new comers to the sport.
 - ↳ might use some elements in the final design.



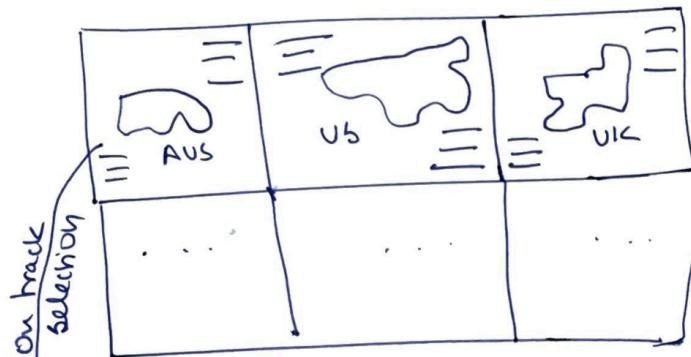
⑤ Page 1 (Home Page)



Quick Stats

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

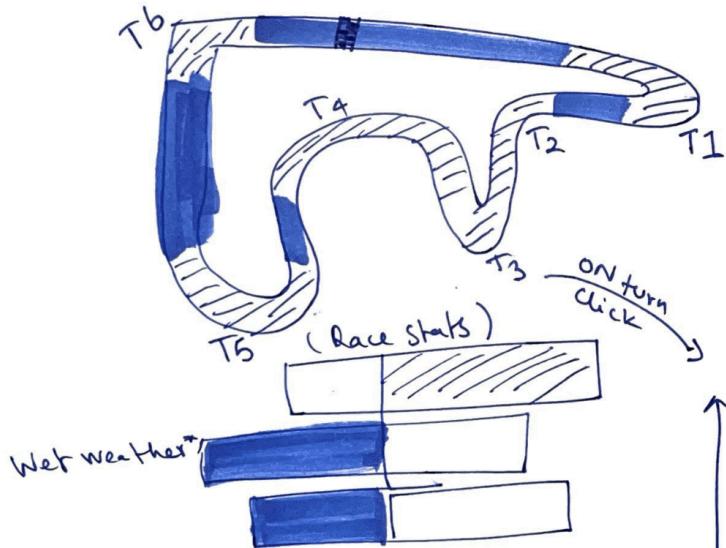
List of tracks Overview



Page 2 (Race Overview Comparison)

2024

2025
|||||



Srinanda Kishore Yallapragada
Sheet 5, March 13 2025

F1 Delta

Task: Selection of Visualizations and computing ideas

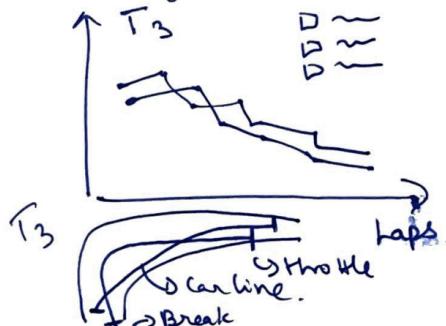
Operations

- 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

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
- determining 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.

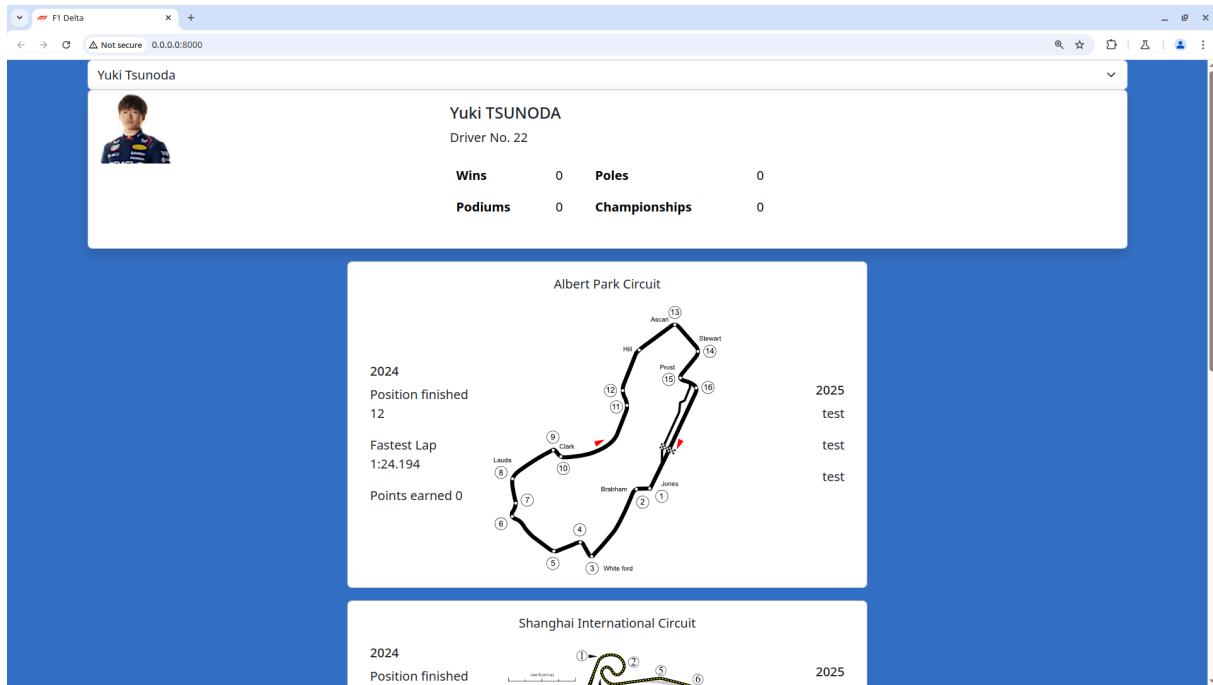
As of now the landing page provides a comparative overview of who the driver is in terms of their overall statistics like wins and poles, and then shows how much improvement is present across each race of the season that has occurred so far. The race data is pulled from an api, so as new races occur, 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.

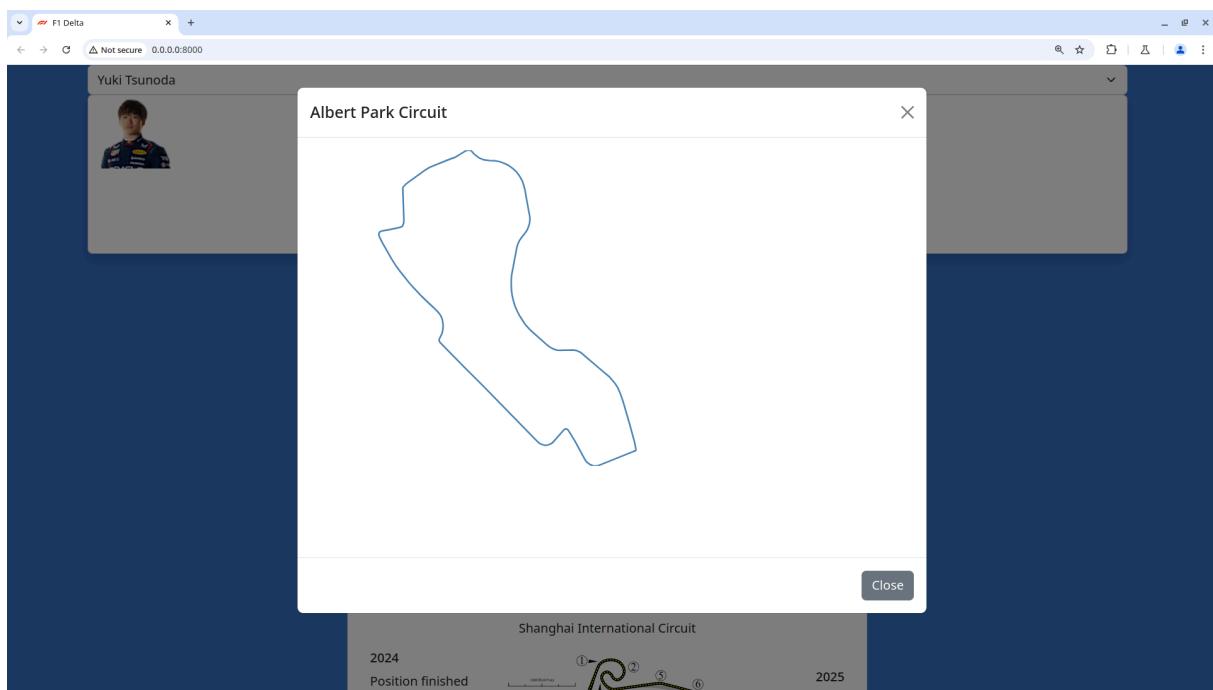
Unfortunately , during the initial design of the website, I did not account for the fact that formula 1 cars across the seasons improve drastically in their performance. I will need to change some of the visualization goals to focus on metrics that are not as dependant on absolute car performance as 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



Project Final Submission

Overview and Motivation:

Provide an overview of the project goals and the motivation for it. Consider that this will be read by people who did not see your project proposal.

Related Work:

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

Questions:

What questions are you trying to answer with your visualization(s)? How did these questions evolve/change over the course of the project? What new questions did you consider as the project progressed?

The most general question I wanted to answer was whether the driver was performing better or worse at the new team.

In my initial proposal I did not consider adding any visualization that had the teammate of the drivers. 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”. So to reflect this I added a question of how the drivers performed relative to their teammates.

At the end, the visualization answers the following questions

- How do the drivers on their qualifying lap perform compared to their previous team?
- How do the drivers race the different engines from the different teams?
- Where do the drivers carry more speed compared to their previous team
- how did the drivers do in the race relative to their teammates from both the years.

Data:

Include information about the source, how you collected it (e.g., web scraping), cleaning methods, etc.

- initially wanted to use open f1 to stick with d3js, but soon realized what a pain that was and switched to fastf1 python api which was far more convenient to work with
- data cleaning was not necessary after the switch to fastf1
- data extrapolation was needed because the telemetry readings are desynced between the different cars

meaning one car has its speed marked at time 1 second and the other car has it at time 1.1. So we need to extrapolate to fill in the gaps at common times to make the visualization.

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?

- websites online
- f1tv has a tech talk show which sometimes goes over the telemetry data which is an inspiration for this project

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 be 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 fastest 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 something I would avoid if I could do this project again. This was also due to some personal health issues which prevented me from realizing the full potential of this visualization.

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.

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.)

I found that Lewis Hamilton is going slower on entering into the racing corners. This trend was repeated across a few other drivers as well.

The engine RPM graph while it looks cool, it is hard to derive actual useful information from it.

The position graph I feel like conveys the relative performance to teammates well because you can clearly see if the driver moved forwards or backwards from their initial position and you can see how their teammate performance compares which I found to be a strong point..

I feel like these visualizations that I have so far haven't done a good job of covering how the drivers are actually adapting to the cars at a deep level. These visualizations provide a surface level story, but do not give any insight into breaking points or driver line changes which is something I wanted to achieve but could not make happen.