Exploring Lap Performance and Tyre Degradation in Formula 1 through Visual Analytics

Pisano Raffaele
pisano.1959863@studenti.uniroma1.it

1 Introduction

Formula 1 is an inherently **data-driven sport**, where data collection and visualization is crucial for making decisions on car setup, race strategy, tyre choice, and adjustments to environmental and track conditions. While fans often enjoy watching races visually, those with a deeper technical interest may wish to explore, analyze, and understand these data in a structured and effective way. Detailed race data can reveal patterns in team strengths, performance correlations under different conditions, tyre degradation trends, and even provide a foundation for anticipating future race outcomes.

However, despite their potential, this data is often poorly visualized and not easily accessible to enthusiasts. This makes it challenging for users to analyze lap performances, compare drivers or teams, and gain insights in a detailed and consistent way.

The goal of this project is to develop an **interactive visual analytics dashboard** that organizes and presents the 2025 Formula 1 race lap data in a clear and engaging way.

This application aims to provide F1 fans with a tool to explore lap performance patterns, understand tyre compound dynamics, and analyze both driver and team performances throughout the season. Through coordinated visualizations and interactive analytics, users can filter, highlight, and examine laps, gaining insights into both individual and team performance under varying conditions.

2 Related Work

2.1 Formula 1 Data Visualization

2.1.1 Online Systems

A number of online platforms provide access to Formula 1 data.

F1 Live Timing is the official Formula 1 platform that streams real-time race information. It offers basic features such as position tracking and driver comparisons, while advanced statistics like split times and tyre usage are only accessible with a subscription. Unlike my system, it does not allow for the exploration of past events and presents data mainly in tabular form, without advanced analysis or coordinated visual views.

F1 Tempo focuses on historical race data from 2018 onward, with graphical representations of lap times and telemetry. While it shares with my project the attention to lap-level performance, its visualizations are very simple and lack analytical tools such as dimensionality reduction or regression. The system mainly provides single charts per driver, track, or session.

Rapit also emphasizes race data visualization through interactive charts. It includes tyre strategy information and offers visually polished graphics, but its scope is limited to showing race progress rather than enabling deeper analysis of performance variables.

PitWall distinguishes itself by offering an extensive database, covering races as far back as 1950. It allows for lap comparisons and provides detailed driver reports, although its interface is heavily text and table based, with little attention to visual design.

F1 Dash is another live-oriented platform, displaying driver positions, tyre choices, lap times, sector data, and even radio communications in real time. Similar to F1 Live Timing, it is mainly designed for following ongoing races, making it suitable for casual fans rather than analytical exploration.

Formula1 Dashboard provides season-level statistics and visualizations related to drivers, teams, pit stops, and incidents. Although its charts are visually appealing, they are largely static, offer limited interactivity, and are not coordinated with each other.

In general, existing platforms either concentrate on live race tracking or provide season summaries through static or simple charts. The work proposed in this paper is positioned differently, aiming to support detailed and interactive analysis of lap data.

2.1.2 Academic Papers

The academic literature on Formula 1 data visualization highlights different approaches to understanding performance and supporting decision-making.

The study **"Visualization of Formula One Racing Results"**[1] presents an interactive tool for exploring driver results across multiple seasons. It combines a timeline, stacked bar charts, and statistics views to compare performances.

Hartog et al. (2021)[2], **"Personalized Visualization in Formula One Racing,"** propose tailoring data visualizations to the roles of team members, such as drivers, analysts, or principals. Their system adapts the displayed information depending on the user, highlighting the importance of personalization.

Latorre (2025)[3], **"F1 Data Analysis,"** developed an interactive tool to explore seasonal driver performance using bar charts, statistics panels, and trend lines.

Finally, the paper **"Automation of Data Analysis in Formula 1"**[4] addresses the increasing complexity of telemetry data collected by teams. The authors propose interactive dashboards built with Python/Dash to support decision-making in real time, highlighting the value of automation and scalability.

These works differ from the present study in that they primarily emphasize historical season-level analysis, role-specific visualization, or real-time decision support for teams, rather than offering a coordinated and interactive framework for the retrospective exploration of lap-level performance with integrated analytical techniques.

2.2 Formula 1 Data Visualization

Visual analytics and visualization techniques have been widely applied to competitive sports to help analysts and fans understand performance patterns, strategies, and outcomes. Several examples of these solutions are used in this work.

Scatter Plots

Scatter plots have been used to provide an overview of team or player performance in sports. For instance, Lage et al. (2016)[5] designed scatter plots for baseball data to compare team behaviors and overall performance trends. Similar to this system, scatter plots can visually reveal clusters of performance and highlight relationships between key variables.

Parallel Coordinates Plots

Parallel coordinates have been adopted in sports analytics to visualize multidimensional metrics. Legg et al.[6] applied them to rugby game data, while Chen et al. (2016)[7] represented score data using parallel coordinates to analyze player performance.

Line Charts

Line charts are frequently used to track changes in player or team metrics over time. Perin et al. (2016)[8] designed overlap-free line charts for soccer scores, while Goldman and Rao (2012)[9] used line charts in basketball to show estimated offensive rebound rates.

Relative Performance Visualization

In sports such as cycling, skiing, or golf, visualizations often emphasize relative differences from the top performer or winner, as it is more interesting than the general time (e.g., Wood 2015; CNW 2008)[10][11]. In this work, a similar approach is applied by visualizing lap times relative to the average lap time of each race, allowing fair comparisons across different tracks.

Use of Semantic Colors

Many sports visualizations use semantic colors to represent teams or players consistently (e.g., Won 2013)[12]. These semantic colors originate from commonly associated colors (like the blues of Chelsea in soccer) or as representative colors from the badges that are teams' symbols [13]. Similarly, in the system, teams are assigned distinct colors to maintain visual clarity. Where possible, these colors are chosen to reflect the traditional team colors, for example red for Ferrari.

3 Dataset

The dataset used in this project originates from the FastF1 API, which provides access to detailed Formula 1 data including timing, telemetry, and session results. For this study, all Formula 1 races of the 2025 season up to the Austrian Grand Prix in Spielberg were collected, with a focus exclusively on race laps.

Before analysis, a preprocessing step was performed to ensure data quality. Several types of records were discarded:

- **Incomplete laps**, where timing or metadata were missing.

- **Invalid laps**, such as those completed under Safety Car, Virtual Safety Car, yellow or red flag conditions.
- **Laps flagged** as inaccurate by the FastF1 API due to synchronization issues, even if the recorded lap time itself was available.
- **Track limits infringements**, where the car exited the track boundaries.

In addition, weather information was merged with lap data. Since FastF1 provides weather updates at one-minute intervals, each lap was matched with the nearest available weather record. This allowed each lap to be associated with a categorical indicator of rainfall (rain vs. dry).

For simplicity, some available details were excluded from the current dataset, such as individual sector times. These could be considered for future extensions of the system.

The final dataset consists of approximately 9,900 valid laps, described by 10 dimensions:

- **Driver**: the driver who completed the lap
- **LapTime**: the absolute time of the lap (in seconds)
- **LapNumber**: the sequential number of the lap in the race
- **Compound**: the tyre compound used (e.g., Soft, Medium, Hard)
- **TyreLife**: the number of laps completed on the current tyre set
- **Team**: the driver's constructor team
- **Track**: the circuit where the lap was recorded
- **Rainfall**: binary indicator of whether rain was present
- **TrackTemp**: track temperature in °C
- **LapTime_norm**: lap time normalized by the average lap time of the respective track

Normalization of lap times ensures comparability across different circuits, whose layouts and lengths can vary significantly. This way, performance analysis focuses on relative differences in lap execution rather than absolute timing, which would otherwise be biased by track characteristics.

4 System

4.1 Intended Users

The system is primarily designed for expert users with a strong background in Formula 1. These users are familiar with domain-specific concepts such as lap times, tyre strategies, race phases, and environmental conditions. Leveraging the dashboard, they can discover insights and analyze the data effectively and efficiently to understand, for example, how a race might unfold based on past performances, identify a team's strengths and weaknesses, explore correlations between different conditions and variables, and extract other valuable information for strategic decision-making.

4.2 System implementation

The implementation of the system consists of an interactive dashboard that provides advanced analytics and visualizations for Formula 1 lap data. The dashboard was designed to support both **selection** and **highlighting** mechanisms, enabling users to perform in-depth comparative analyses of driver performances, tyre behavior, and race conditions.

At the top of the interface, a **filter menu** allows users to define the subset of laps to be analyzed. Filtering can be applied along several dimensions, including **driver**, **track**, **team**, **race phase** (early, mid, late race), **temperature range**, **tyre compound**, and **weather conditions**. Once the filtering is applied, the system generates all analytics and visualizations dynamically. Users can then interact with the charts either by selecting laps, which constrains the dataset used in calculations, or by highlighting laps, which emphasizes them visually without altering the underlying computations.

The distinction is crucial:

- **Selection** means that only the chosen laps are retained for calculations and visualization, while all other laps are excluded from the charts.
- **Highlighting**, instead, emphasizes laps visually (for example through opacity or size changes) but keeps them part of the dataset and calculations, with the exception of the **Boxplot**, which is recomputed using the highlighted laps as if they were selected.

4.2.1 Charts

The dashboard integrates six main types of charts, each with specific purposes and interaction capabilities. The **Boxplot** displays the distribution of normalized lap times (LapTime*) for each driver based on the highlighted laps. For each driver, boxplots are recomputed with every selection. Individual laps are represented as points in front of each driver's boxplot, enabling detailed performance comparisons.

The **Parallel Coordinates** visualization provides a multidimensional view of correlations between parameters such as lap number, tyre compound, normalized lap time (LapTime*), tyre life, track, weather, and temperature. The axes can be rearranged interactively. Laps can be highlighted directly within this visualization, and highlighted laps from other charts are also reflected here through changes in opacity and sizing.

The **Table** presents lap data in a tabular format. Users can highlight or clear highlights for specific laps, or highlight all laps at once. Sorting by different criteria (such as lap number, driver, or tyre) is also supported.

The **Tyre Wear Line Plot** illustrates the relationship between tyre life and LapTime*. This visualization recomputes a regression line based on the highlighted laps, showing the average degradation of tyres over time. For each selected compound, a trend line is generated with slope (in seconds per lap) and R^2 (coefficient of determination) as quality metrics, providing insight into tyre performance across stints.

The **PCA Scatter Plot**, based on Principal Component Analysis, is computed only on the selected laps. Users can highlight laps within the PCA view, and these highlights propagate across all other linked visualizations. Additionally, clustering can be applied using k-means, with cluster quality evaluated via silhouette scores. Once clustering is performed, laps across the dashboard can be colored by cluster, team, or track.

Finally, the **Scatterplot Matrix (SPLOM)** displays pairwise correlations between key lap variables such as LapTime*, lap number, tyre life, track temperature, and weather. Each subplot can be selected and zoomed individually. Laps highlighted in a zoomed view are automatically highlighted across the rest of the dashboard.

4.2.2 Coordination

The core of the system lies in the **coordination between visualizations**. All charts display only the laps selected via the top-level filter. The **Parallel Coordinates**, **SPLOM**, **Table**, and **PCA Scatter Plot** are **bidirectionally linked**, allowing highlights to be initiated in any of them and reflected across the others. On the other hand, the **Boxplot** and the **Tyre Wear Line Plot** are **one-way linked**: they reflect highlighted laps by recomputing the entire chart, but they do not serve as sources of new highlights.

This coordinated approach ensures consistency across the dashboard and provides users with a seamless experience when exploring and analyzing race data, while also offering the opportunity to **discover actionable insights** on driver performance, strategy, and race dynamics.

4.3 Insights

During the exploration of the dataset through the visualization system, several non-trivial insights emerged that provide a deeper understanding of the data.

4.3.1 Tire life and performance

A counterintuitive pattern was observed: the average lap time tends to improve as tire life increases. In other words, worn tires appear to deliver better performance. This trend, visible in the scatterplot of TireLife vs. LapTime*, shows reduced variance and an overall improvement in lap times as the stint progresses. However, the line plots reveal that this effect is primarily associated with the hard compound. Regression analysis indicates an average gain of about 0.022s per lap, with a cumulative improvement of around -0.924s in LapTime*. This suggests that, for the hard tire, the benefit of reduced fuel load outweighs the negative impact of tire wear.

4.3.2 Hard vs. Medium tire performance

Across the full dataset, hard tires surprisingly show better overall performance than the mediums, along with the expected lower degradation rate. Yet, when isolating the first quarter of the race, the situation reverses: mediums outperform hards. A possible explanation is that hard tires suffer more from heavy fuel loads and the close racing dynamics at the beginning of a race (dirty air, slipstream, battles). From the second quarter onward, the hards regain their advantage, gradually surpassing the mediums in both pace and degradation.

4.3.3 Team performance and temperature sensitivity

Temperature filtering revealed a strong team-specific trend:

- Red Bull tends to perform better in colder conditions and struggles more as temperatures rise.
- Ferrari, on the other hand, appears to benefit from higher track temperatures while underperforming in the cold.

This difference was clearly visible in the boxplot distribution of lap times across temperature ranges.

4.3.4 Behavior of intermediate tires

By filtering laps under cold conditions (<25°C) and projecting them into the PCA space, three distinct groups emerged: intermediate laps in rain, intermediate laps in sunny conditions, other laps. Analyzing the first two groups, despite similar average performance, degradation analysis shows a clear advantage for intermediates used in drying conditions rather than in wet ones. This is likely because, as the track dries, intermediate tires gradually lose their grooves and become smoother, increasingly resembling slicks. This transformation improves their long-run performance on a drying track.

4.3.5 Temperature effects on soft tires

Soft (red) tires demonstrate better performance and lower degradation in colder conditions (<25°C). In contrast, when temperatures are higher, they degrade faster and lose performance, making them less effective in hot races.

4.3.6 Driver-specific patterns

- Lewis Hamilton showed exceptional performance on hard tires in the second half of races, even outperforming McLaren over a wide range of laps. However, in the first half of races, his performance dropped significantly, often falling outside the top 10.
- Charles Leclerc displayed the opposite pattern: he was among the fastest on hard tires in the first half of races (ranked 2nd out of 20), but slowed down in the second half, dropping to 7th.

This suggests possible setup differences or strategic choices affecting long-run tire management.

5 Conclusion

This project introduced an interactive visual analytics dashboard for Formula 1 lap data, enabling expert users to explore tyre dynamics, team strategies, and driver performance. Through coordinated visualizations, the system revealed non-trivial insights such as counterintuitive tire degradation patterns, compound-specific behaviors, team sensitivity to temperature, and driver-specific trends. Overall, the dashboard proves effective in transforming raw race data into actionable knowledge, supporting deeper understanding of race dynamics.

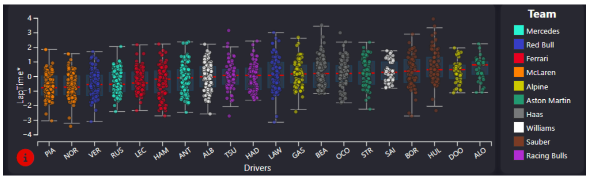
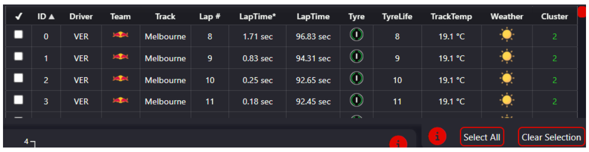


Figure 1: Boxplot



✓	ID ▲	Driver	Team	Track	Lap #	LapTime*	LapTime	Tyre	TyreLife	TrackTemp	Weather	Cluster
	0	VER		Melbourne	8	1.71 sec	96.83 sec		8	19.1 °C		2
	1	VER		Melbourne	9	0.83 sec	94.31 sec		9	19.1 °C		2
	2	VER		Melbourne	10	0.25 sec	92.65 sec		10	19.1 °C		2
	3	VER		Melbourne	11	0.18 sec	92.45 sec		11	19.1 °C		2

Select All
Clear Selection

Figure 2: Table

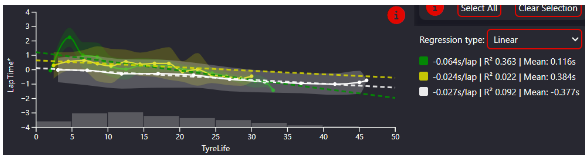


Figure 3: Tyre Deg Graph

References

[1] M. Vincken, *Visualisation of formula one racing results*, 2016. [Online]. Available: https://www.academia.edu/19974065/Visualisation_of_Formula_One_Racing_Results_ENGLISH_.

[2] I. H. d. Hartog, *Data to drive: Personalized visualization in formula one racing*, 2022. [Online]. Available: https://research.tue.nl/files/306910238/Thesis_BDS_Iris_den_Hartog.pdf.

[3] A. La Torre Castelltort, *F1 data analysis*, 2025. [Online]. Available: https://diposit.ub.edu/dspace/handle/2445/1201/simple-search?etal=0&filter_field_1=subject&filter_type_1=equals&filter_value_1=Treballs+de+fi+de+grau&order=asc&query=&rpp=35&sort_by=0&start=6405.

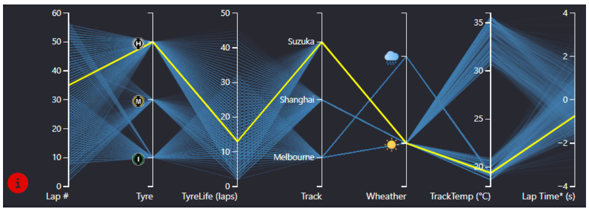


Figure 4: Parallel Coordinates

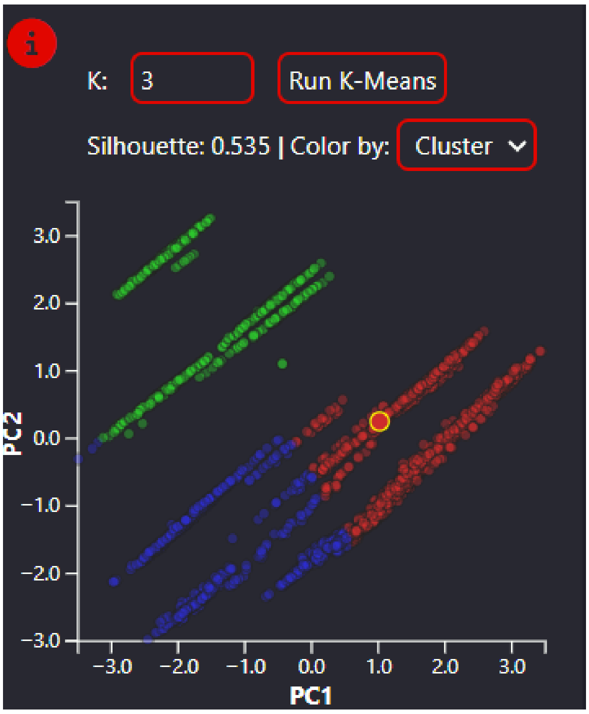


Figure 5: Pca

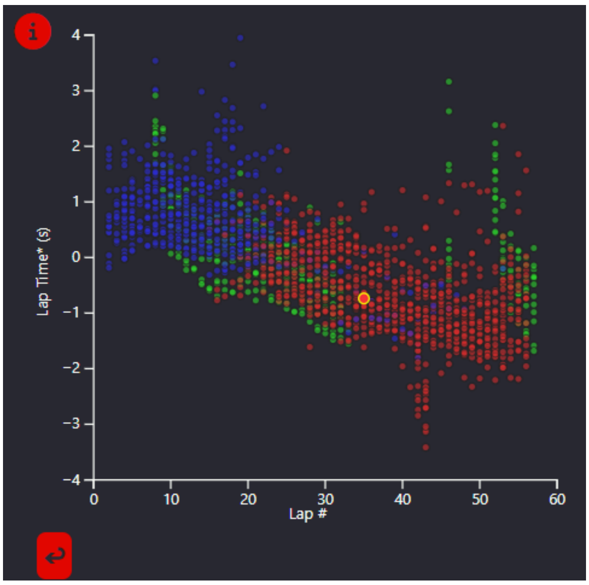


Figure 6: SPLOM - opend one

[4] A. M. Prescott Delzell Patrick McCabe, *Automation of data analysis in formula 1*, 2019. [Online]. Available: https://digitalcommons.calpoly.edu/imesp/254/.

[5] M. Lage, J. P. Ono, D. Cervone, J. Chiang, C. Dietrich, and C. T. Silva, "Statcast dashboard: Exploration of spatiotemporal baseball data," *IEEE Computer Graphics and Applications*, 2016.

[6] P. A. Legg, D. H. Chung, M. L. Parry, *et al.*, "Transformation of an uncertain video search pipeline to a sketch-based visual analytics loop," *IEEE Transactions on Visualization and Computer Graphics*, 2013.

[7] W. Chen, T. Lao, J. Xia, *et al.*, "Gameflow: Narrative visualization of nba basketball games," *IEEE Transactions on Multimedia*, 2016.

[8] C. Perin, J. Boy, and F. Vernier, "Using gap charts to visualize the temporal evolution of ranks and scores," *IEEE Computer Graphics and Applications*, 2016.

[9] M. Goldman and J. M. Rao, "Effort vs. concentration: The asymmetric impact of pressure on nba performance," in *Proceedings of the MIT Sloan Sports Analytics Conference*, 2012.

[10] J. Wood, "Visualizing personal progress in participatory sports cycling events," *IEEE Computer Graphics and Applications*, 2015.

[11] A. Cox, V. Nguyen, and J. Ward, *How woods wins a major*, 2008. [Online]. Available: https://www.nytimes.com/2008/05/01/sports/golf/01woods.html.

[12] K. Wongsuphasawat, "A narrative display for sports tournament recap," in *1st Workshop on Sports Data Visualization, IEEE VIS*, 2013.

[13] V. Setlur and J. D. Mackinlay, "Automatic generation of semantic icon encodings for visualizations," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014.