Gear Shifts and Data Rifts: Formula 1 Driver Prowess

Caleb King

Department of Mathematics and Statistics, Northern Arizona University, Flagstaff, AZ 86011

Office of Undergraduate Research and Creative Activity

Department of Mathematics and Statistics

Abstract

Deepened the understanding of the factors influencing driver performance in Formula 1 (F1) racing could provide meaningful insights to improve results on the racetrack. F1 data from the 2023 Singapore Grand Prix, including race results, lap times, and car telemetry, were collected using the FastF1 Python package, which accessed Formula 1's live timing service. This data was made available through the FastF1 plugin, which was hosted on GitHub. Data was available across multiple race circuits throughout the 2023 season, with diverse track layouts and weather conditions. The analysis focused on comparing the performance of the top 10 drivers in the race by examining various metrics such as lap times, sector times, speed, brake pressure, and DRS advantages. The goal was to identify patterns, strengths, and weaknesses for these drivers throughout the season by connecting driver statistics to the available lap and car data. This project employed statistical techniques, including one-way and two-way ANOVAs, to extract meaningful insights from the data. These insights could aid in optimizing driver strategies and performance by providing data-driven recommendations such as adjustments to driving techniques, setup preferences, or race strategies.

Introduction

F1 racing stands at the pinnacle of motorsport, where drivers push the boundaries of speed, skill, and technology. Understanding the intricate factors that influence driver performance in this highly competitive environment of the 2023 Singapore Grand Prix is crucial for teams aiming to gain a competitive edge on the racetrack. By delving deep into the data generated during this race, valuable insights could be unearthed, offering opportunities for optimizing driver strategies and performance.

In this study, comprehensive datasets from the 2023 F1 season were collected, comprising race results, lap times, and actual data on the car's performance. These datasets were sourced using the FastF1 Python package, which accessed Formula 1's live timing service and provided real-time data on the race circuit.

The primary objective of this study was to dissect the performance of F1 race drivers by analyzing a myriad of metrics such as lap times, sector times, speed, brake pressure, and DRS. By connecting driver statistics to the available telemetry and lap data, the aim was to identify patterns, strengths, and weaknesses exhibited by drivers throughout the 2023 season.

Methods

The data for this analysis was collected from the 2023 Singapore Grand Prix using the FastF1 plugin within the Python environment. The dataset included information on race results, lap times, telemetry data of the cars, and various race statistics. The initial step involved cleaning the collected data to ensure its suitability for analysis. This process included converting lap times from string format to seconds using a cushion lap_time_to_seconds function, as well as aggregating sector times for each driver on each lap using the dplyr package. Additionally, the data was filtered to include only the top 10 drivers based on their positions in the last lap of the race. Finally, the last step of the cleaning process was to remove all NA values from the data set.

Following data cleaning, various visualizations were created using the ggplot2 package to explore different aspects of driver performance . These visualizations included histograms to analyze lap time distributions, line charts to track lap time trends by driver and tire compound, box plots to examine lap time distributions by driver and tire compound, and faceted plots for sector analysis. To extract deeper insights, one-way and two-way ANOVAs were conducted using the lm function from the stats package. The one-way ANOVA examined the effect of driver on lap times, while the two-way ANOVA investigated the effects of driver, tire compound, and their interaction on lap times.

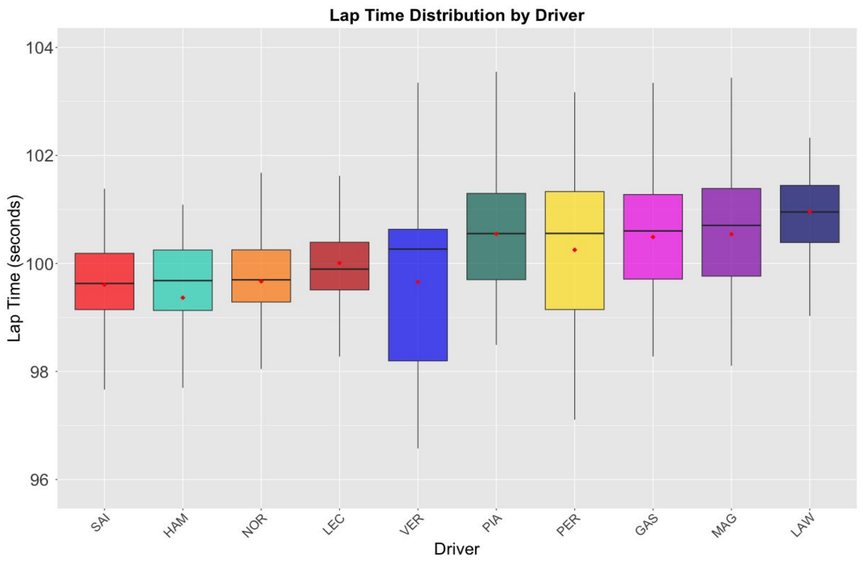
In the next phase of the analysis, advanced statistical techniques such as regression were employed to extract deeper insights from the dataset. These techniques would help uncover patterns, correlations, and trends in driver performance, ultimately contributing to a comprehensive understanding of the factors influencing the outcomes in the 2023 Singapore Grand Prix. Through this multi-faceted approach to data analysis, this study aimed to provide valuable insights into F1 driver performance during the 2023 season, which could inform strategic decisions and contribute to performance optimization on the racetrack.

Singapore Lap Data

Driver	Lap Number	Stint	PB	Tire Compound	Team	Position	Lap Time	Sector 1	Sector 2	Sector 3
VER	61	2	True	MEDIUM	Red Bull Racing	6	96.575	28.434	41.086	27.055
GAS	46	2	True	HARD	Alpine	6	98.277	28.779	41.942	27.556
PER	61	2	True	MEDIUM	Red Bull Racing	9	97.108	28.694	41.344	27.070
LEC	46	2	True	HARD	Ferrari	3	98.275	28.631	42.111	27.533
MAG	48	3	True	SOFT	Haas F1 Team	14	98.107	28.655	41.834	27.618
NOR	46	2	True	HARD	McLaren	2	98.046	28.630	41.946	27.470
LAW	47	2	True	HARD	AlphaTauri	8	99.028	28.975	42.289	27.764
HAM	47	3	True	MEDIUM	Mercedes	5	95.867	28.210	40.658	26.999
SAI	47	2	True	HARD	Ferrari	1	97.666	28.676	41.569	27.421
PIA	46	2	True	HARD	McLaren	7	98.492	28.878	42.070	27.544

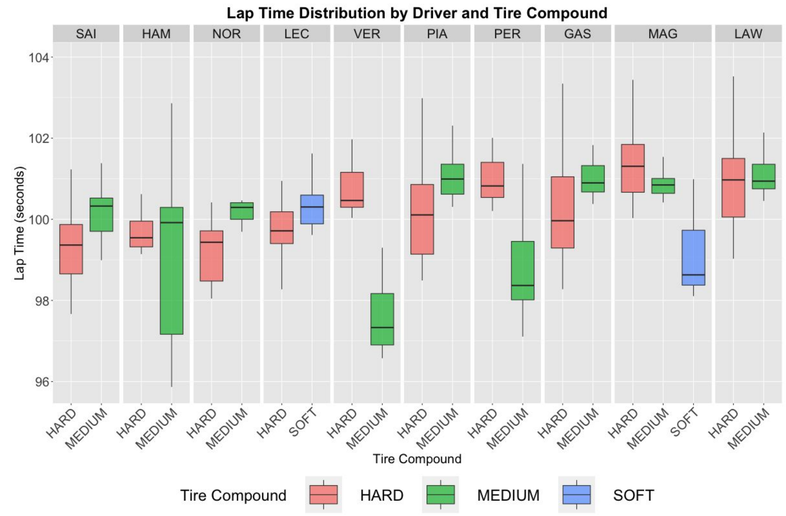
- The data shows the fastest lap times and other variables for each driver.
- MAG is the only driver who set his fastest lap on the soft tire compound, which is generally the fastest but least durable tire. He also set his fastest lap the latest in the race (lap 14) compared to other drivers. This could suggest MAG's team employed a different tire strategy.
- VER and PER (teammates at Red Bull Racing) have very similar fastest laps, with both on medium tires. Their sector times are also very close. This indicates the Red Bull car was performing well in this race.
- Sector 2 seems to be the longest sector, with lap times around 41-42 seconds, while Sector 3 is the shortest at around 27 seconds. Sector 2 is also around the same length of sector 1s. This information could be useful for teams in planning race strategies.

Results - Lap Time



- The narrow boxes for drivers like SAI, LEC, and VER imply that their lap times were consistently close to their respective medians, showcasing strong and steady performance.
- The wider boxes and longer whiskers for drivers like LAW and MAG suggest greater variability in lap times, which could be due to factors such as traffic, tire management, or strategic decisions.
- The drivers with the most consistent lap times (SAI, LEC, VER) are also among the fastest, suggesting that consistency and speed often go hand in hand in Formula 1.
- The similar distribution shapes for drivers from the same team (e.g., VER and PER, NOR and PIA) suggest that car performance and team strategy play a significant role in determining lap time consistency and overall performance.
- Overall, there seems to be a notable difference in performance between the top drivers (VER, LEC, PER) and the rest of the field, as indicated by their lower median and mean lap times.
- The one-way ANOVA revealed a statistically significant difference in lap times among the drivers, indicating that driver performance had a significant impact on lap times during the 2023 Singapore Grand Prix.
- The box plot illustrates that drivers from the same team (e.g., VER and PER, SAI and LEC) had similar median lap times and box sizes, indicating the importance of car performance and team strategy in maintaining consistent lap times.

Results - Tire Compound



- The significant interaction effect between driver and tire compound suggests that the impact of tire compound on lap times varied across drivers. This implies that the optimal tire strategy may differ for each driver and team, depending on their specific car setup and driving style.
- The hard tire compound appears to be a popular choice, with 6 out of the 10 drivers opting for it. This suggests that the track conditions and race strategy favored the use of harder, more durable tires.
- GAS, despite being on the hard tire compound, has a higher median lap time compared to other drivers on the same compound. This could indicate challenges specific to the driver or car, rather than the tire choice itself.
- The fastest lap times were set on the medium tire compound, as evidenced by the lower median and mean lap times for drivers using mediums (VER, PER, HAM).
- LEC, the only driver on soft tires, demonstrated competitive lap times compared to drivers on medium and hard tires. The two-way ANOVA results showed a significant positive interaction effect for LEC with soft tires, indicating that the soft tire compound was particularly advantageous for LEC's performance.

Conclusion

In conclusion, the analysis of lap time data from the 2023 Singapore Grand Prix provided valuable insights into the factors influencing driver performance. By employing a combination of statistical techniques, including one-way ANOVA, two-way ANOVA, pairwise comparisons, and visualizations such as box plots, we were able to uncover the complex interplay between driver skill, tire compound, and team strategy.

- The one-way ANOVA revealed significant differences in lap times among drivers ($F(9, 540) = 9.621$, $p < 0.001$), highlighting the crucial role of individual driver performance.
- The two-way ANOVA showed significant main effects of both driver ($F(9, 529) = 19.35$, $p < 0.001$) and tire compound ($F(2, 529) = 19.35$, $p < 0.001$), as well as a significant interaction effect ($F(18, 529) = 19.35$, $p < 0.001$), indicating that the optimal tire strategy varied across drivers and teams.
- Drivers on the same team and tire compound exhibited similar lap time distributions, suggesting the importance of team strategy and car setup in achieving consistent performance.

References

Philipp Schaefer. (2024). *FastF1,* https://github.com/theOehrly/Fast-F1

Acknowledgements

Dr. Robert Buscaglia, NAU Data Science Capstone Professor, Department of Mathematics and Statistics
Northern Arizona, Department of Mathematics and Statistics
Northern Arizona University, Office of Undergraduate Research and Creative Activity