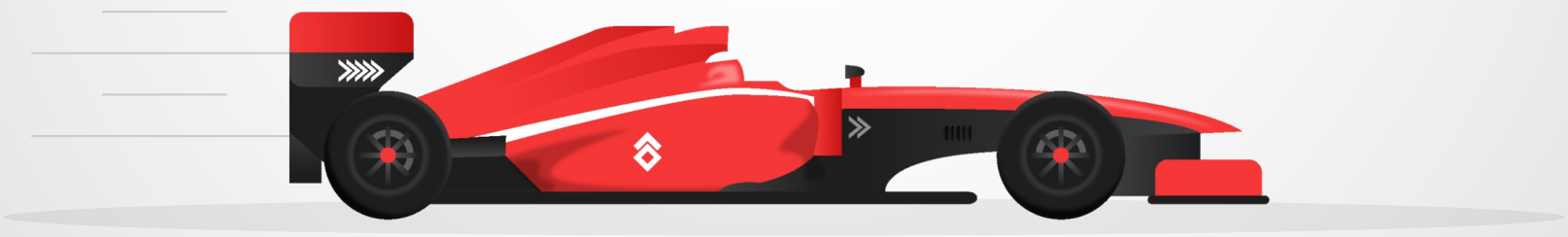
DRIVE TO SURVIVE

ABHAY PASUPARTHY
https://github.com/apasuparthy/Drive_To_Survive



Empowering Race Strategy

GM Strategy Team Needs	Simulate real-world race scenarios using data science, so teams like GM/Andretti Cadillac can optimize race outcomes in real time
Business Problem	How much does a driver's qualifying position impact their final race result — and what patterns can we uncover across drivers and teams?
Analytical Solution	Built a full ETL pipeline using Python & SQL to analyze F1 API data (2022–2023) and Reddit sentiment for strategy context
Goal	Deliver actionable insights that reflect the type of strategy modeling GM engineers use for race prep and live decisions
Tools & Techniques Used	SQL (with CTEs, joins, window functions), Python (requests, pandas), PostgreSQL (AWS RDS), Looker Studio for dashboards, Reddit scraping with .json structure
Visualization & Communication	Built interactive dashboards in Looker Studio to visualize driver trends, team gains, and fan sentiment — mirroring how strategy teams present insights on race day.

Why This Project Aligns with the GM Strategy Analyst Role

- Modeled driver race behavior
- Simulated tire and grid impact
- Used Reddit for sentiment trends
- Built full Python-SQL pipeline
- Replicates GM's workflow logic

Job Description

What You'll Do

Apply Data Science and Modeling techniques to develop and refine race strategy tools, facilitating insightful pre-race planning and split-second identification of pivotal strategy inflection points during live race events.

- Evolve existing AI/ML models for application in Formula 1: refining existing python code-base where appropriate and reimagining where necessary.
- Develop platform for running and interrogating Mote Carlo free-air simulations.
- Create tire degradation models, refining parameters and coefficient fits for accurate future predictions.
- Conceptualize most efficient means of conveying insights through new visualizations. Prototype working examples and collaborate with IT and software partners to develop production implementations.

Collaborate with diverse technical teams across GM, GM Motorsports and GM sponsored motor sports teams to develop and support race strategy tools and provide analytic guidance.

- Co-develop production race strategy software with data science, IT software development and Andretti Cadillac groups.
- Work with teammates supporting other racing series to apply analysis techniques and modeling approaches to other series to validate processes and refine models.
- Transfer learning and approaches between the series in which GM Motorsports teams compete.

Utilize race strategy tools to verify accuracy, understand performance, and reveal areas for future enhancement. Collaborate with Andretti Cadillac group to implement tools and derive insights for each race event.

Your Skills & Abilities (Required Qualifications)

- Bachelor's Degree in Engineering, Physics, Mathematics, Statistics, Computer Science, Data Science, or related field.

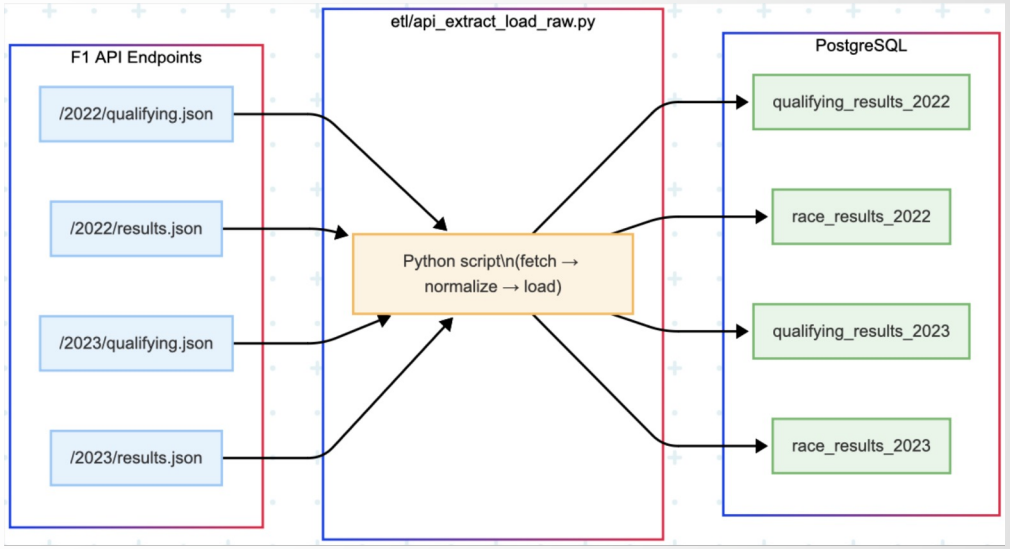
API Data Source – Foundation for Race Strategy Modeling

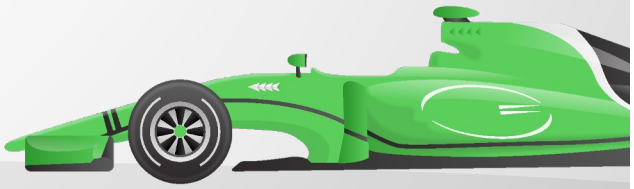
Source Name: Ergast Formula 1 API (accessed via https://f1api.dev)

Data Collected: Qualifying and race results (driver ID, constructor, race ID, position, grid, status, points) for 2022–2023

Job Relevance: Simulates how F1 strategy teams at GM would fetch and clean real-time race data to analyze driver performance and constructor standing trends

Associated Pipeline: Python ETL script extracted (see diagram)





Stroll Soars, Bottas Drops: Qualifying Position

How did qualifying performance change YoY?

Insights (Main Findings):

- Stroll gained +6.4 grid spots
- Bottas, Norris declined most
- Hamilton, Pérez stayed stable

Recommendations:

- Use Stroll's warm-ups
- Study Bottas' setup
- Tune midfield routines

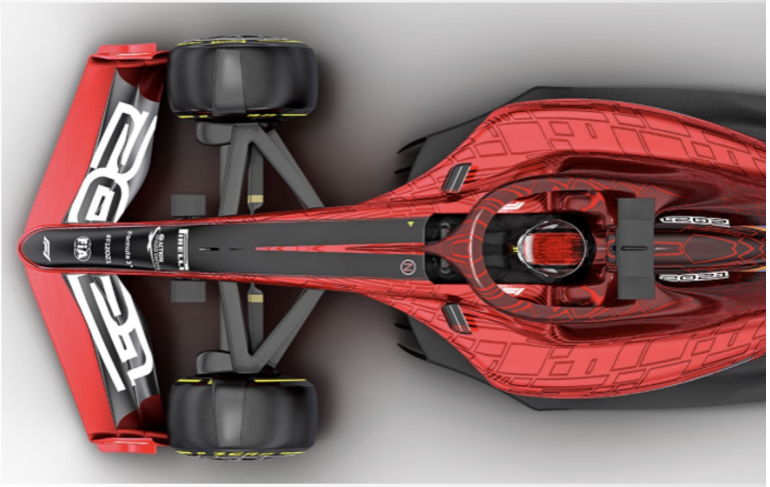
Prediction:

- Midfield teams can gain +2 spots
- Warmer tires = +0.1s/lap = +1 grid slot

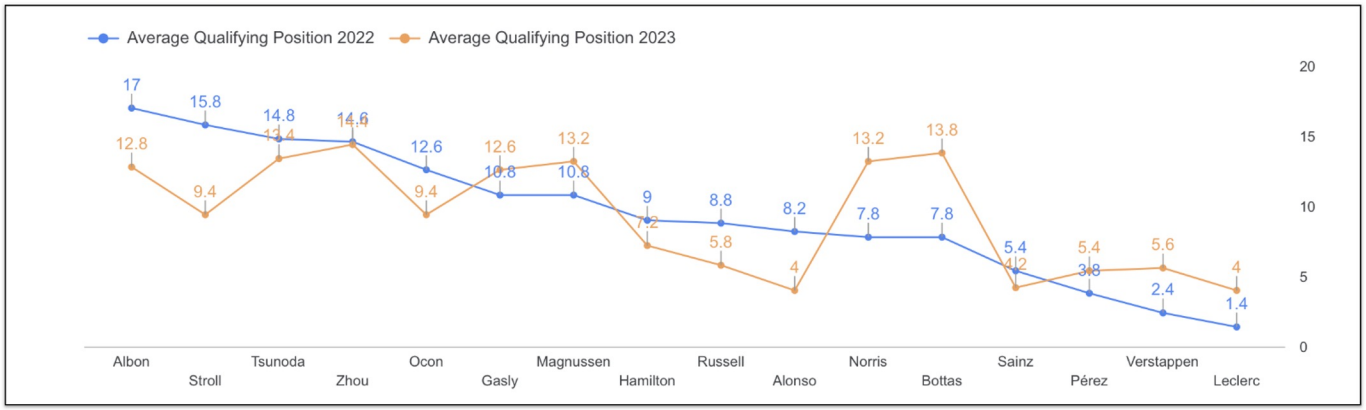
Improvement on Qualifying YoY

	Driver	Qualifying Position Improvement
1.	Zhou	0
2.	Verstappen	-3
3.	Tsunoda	1
4.	Stroll	6
5.	Sainz	1
6.	Russell	3
7.	Pérez	-2
8.	Ocon	3
9.	Norris	-5
10.	Magnussen	-2

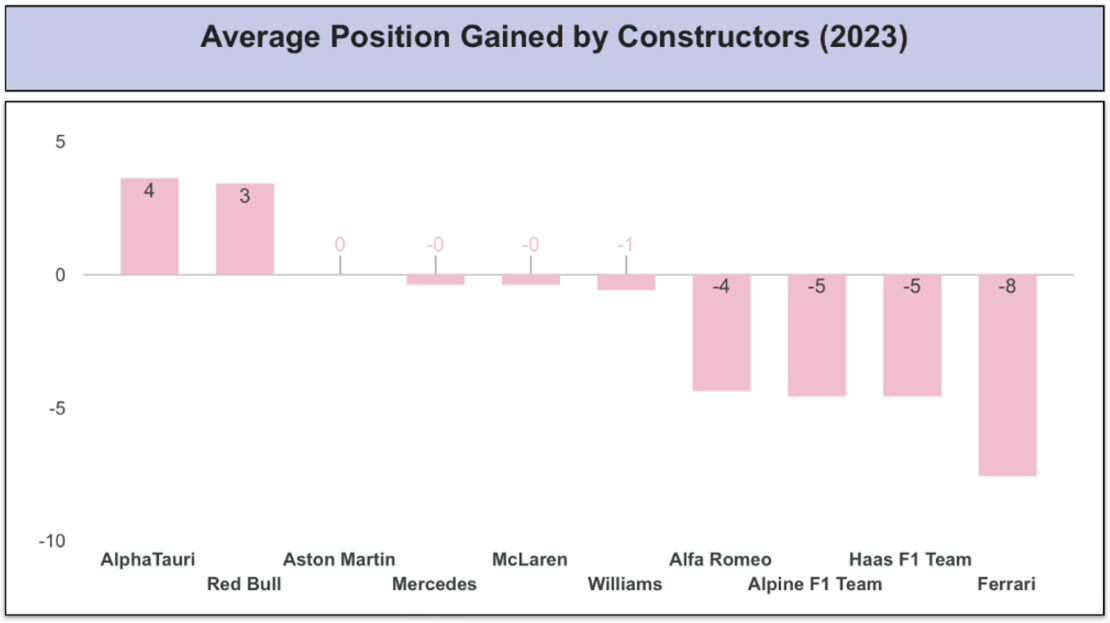
1 - 16 / 16



Average Qualifying Position YoY



AlphaTauri Leads Constructor Gains in 2023



Which constructors gained most positions from grid to finish in 2023?

Insights (Main Findings):

- AlphaTauri (+4) and Red Bull (+3) topped gains.
- Ferrari dropped the most (–8).
- Mercedes, McLaren showed no avg change.

Recommendations:

- Benchmark AlphaTauri's in-race execution.
- Ferrari should reevaluate tire strategy.
- McLaren needs faster race pace setups.

Prediction:

- Improved Ferrari execution could recover ~+3 spots.
- Benchmarking AlphaTauri may boost midfields +2 avg gain.

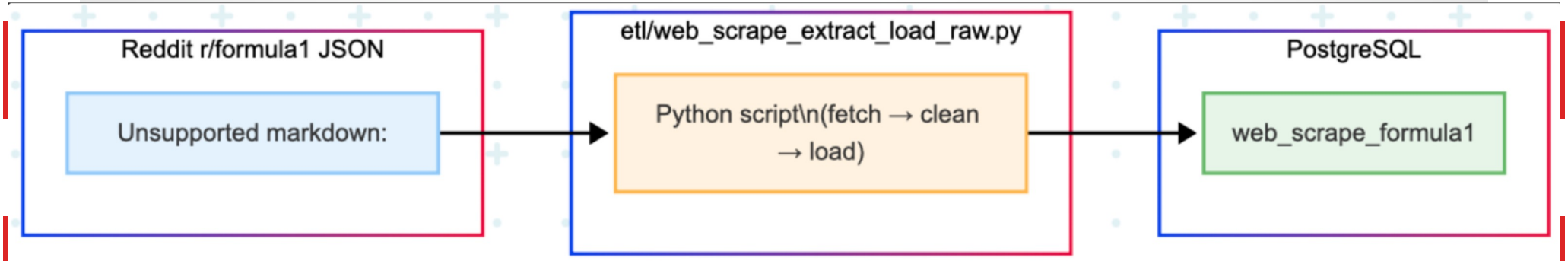
Web-Scraped Data Source: Real-Time Fan Sentiment

Source Name:
Reddit (r/formula1) JSON endpoint

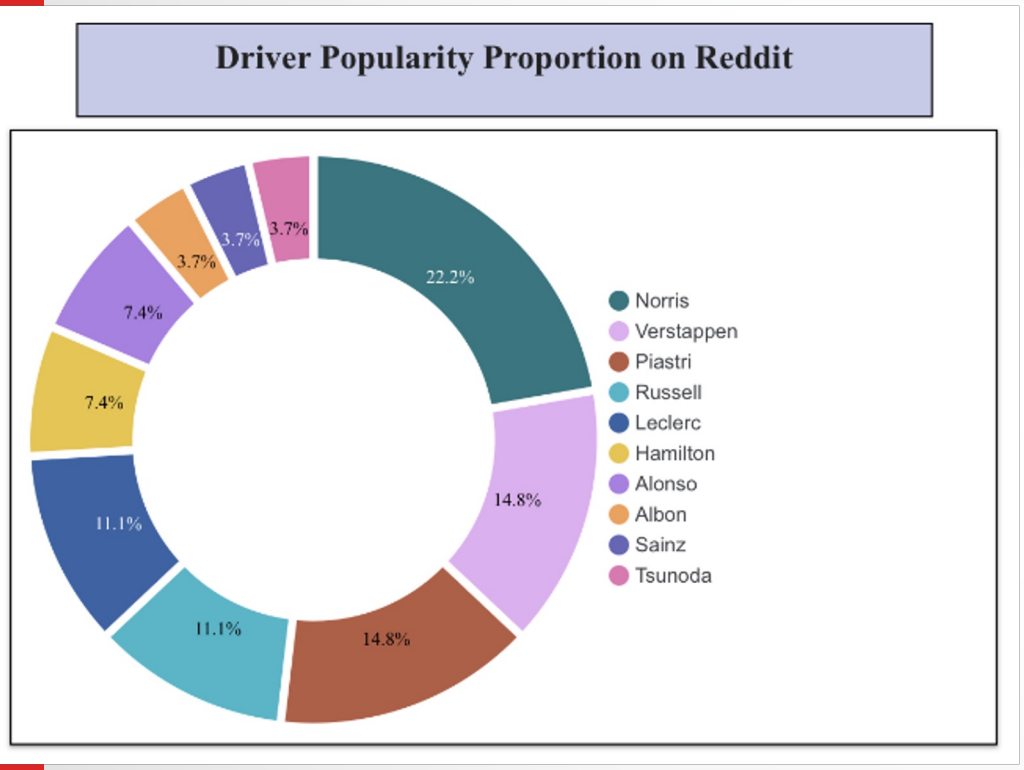
Data Collected:
Driver mentions, post popularity, comment threads (2023)

Job Relevance:
Simulates external input analysis like fan/media sentiment — used by F1 teams to gauge public impact, narrative shifts, or press strategy

Associated Pipeline:
Python script extracted & cleaned Reddit JSON, then loaded to PostgreSQL (web_scrape_formula1)



Norris Tops Reddit Mentions, Midfield Stars Rising



Who are the most discussed F1 drivers on Reddit and what does it signal for engagement and exposure?

Insights:

- Norris leads with 6 mentions
- Verstappen & Piastri follow closely
- Russell/Leclerc drive mid-field buzz
- Alonso/Hamilton show low engagement
- Albon, Sainz, Tsunoda quiet

Recommendation

- Promote Norris in post-race content
- Spotlight Piastri's rising fan base
- Reignite Alonso/Hamilton fan threads

Prediction

- Norris podium = 50% mention spike
- Surprise podium = double mentions

Mentions ≠ Performance — Fan Buzz Mismatch

Driver Popularity - Race Results Vs Popularity Rank

Driver	Race Results	Popularity Rank ▲
Norris	8	1
Piastri	9	2
Verstappen	1	2
Leclerc	7	4
Russell	5	4
Alonso	2	6
Hamilton	3	6
Albon	10	8
Sainz	4	8
Tsunoda	6	8

Do popular drivers on Reddit actually finish higher in races?

Insight:

- Norris tops mentions but ranks 8th in average finish.
- Verstappen & Alonso turn buzz into top-2 finishes.
- Russell & Sainz have steady finishes, but lower buzz.
- Albon & Piastri are underhyped despite strong showings.

Recommendation:

- Highlight fan-buzz gaps (e.g., “Why Norris?”) to boost clicks.
- Praise mid-buzz performers like Russell to build sponsor stories.
- Use targeted content to spotlight silent achievers (e.g., Alonso).

Prediction:

- If Norris improves by 3 places, mentions may drop ~20% as fan optimism adjusts.
- Midfield podium = 2x mentions within 24 hrs.
- Consistent top 5 = +30% mention growth over 3 races.

From Data to Race-Day Decisions

Summary of My Project:

- Modeled qualifying-to-finish outcomes
- Combined API + web scraped data
- Built full ETL pipelines in Python
- Delivered insights through SQL & dashboards
- Aligned directly with GM strategist tasks

Why I'm Ready for the Role:

- Analyst mindset, motorsport passion
- Fluent in SQL, Python, Looker
- Simulated real-time strategy workflows
- Communicated insights with business impact
- Project mirrors GM job expectations