Sorry, I'm not sorry

Let me take you back to 2021...

F1 2021 Championship



Lewis Hamilton (Mercedes) vs Max Verstappen (Red Bull)



Ferocious rivalry w/ incidents marking the season



Closest championship in F1 history



FORMULA ONE

Hamilton vs. Verstappen

Points over time. Select a circle for more information

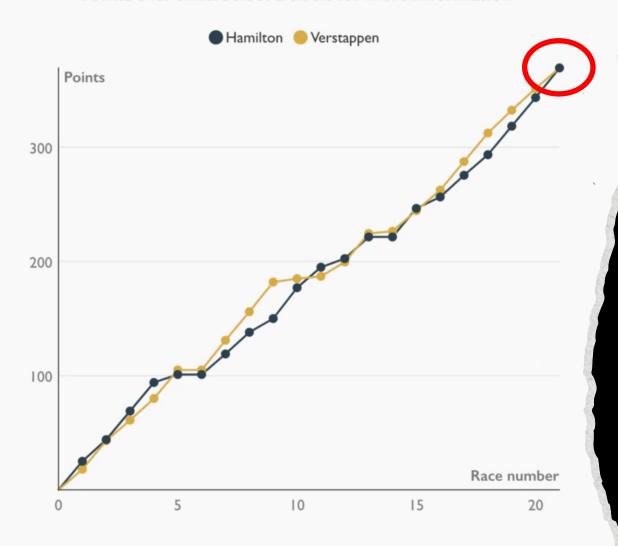


Chart: The Times and The Sunday Times

After **21 races**, Lewis Hamilton and Max Verstappen tied in points (**369.5 points each**)

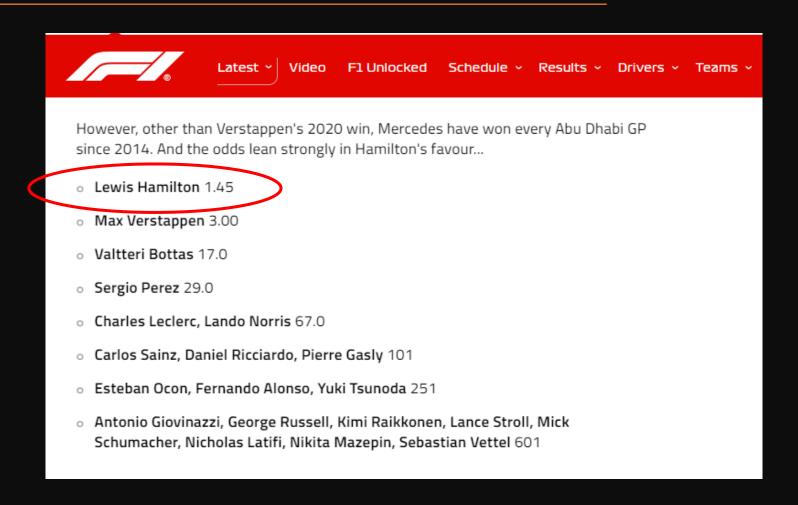
As data scientists

wathore adata....

visualisation

Last Race – Winner Takes It All...

The odds were in Lewis Hamilton's favour to win the last race...



"The clock is ticking. We are going to ned a MIRACLE in the final 10 laps."

Christian Horner—





Last lap...unexpected twist...



Max is the world champion...



Hamilton fans' petition DEMANDING Abu Dhabi 2021 result change reaches 100,000 signatures

B Lauren Sneath
■ Wednesday 19 April 2023 11:48 Updated: 12:00





F1 fans blast 'farcical' Abu Dhabi Grand Prix end: Max Verstappen win 'complete robbery'

MAX VERSTAPPEN sensationally won the F1 World Championship.

By SAM SMITH

00:47, Mon, Dec 13, 2021 | UPDATED: 07:19, Mon, Dec 13, 2021

















TV Fuming fans sent axed F1 race director Michael Masi death threats

For the first time, Masi has opened up on the level of horrific abuse he received from F1 fans on social media after the errors he made which cost Lewis Hamilton the 2021 title

Could we have predicted this outcome?

We will find out....

I THINK THIS IS VERY CLEVER

In reality the my BrainSt in

n project for me with my

Introduction to the Dataset & EDA Findings

- •Dataset Overview: Comprehensive F1 race data (on a driver level for each race) from 1995-2023.
- •Volume and Variety: Includes 11.2k rows and 37 columns (e.g. race winner, starting grid position, engine manufacturer, fastest lap time, etc.)
- •Data Quality Concerns: Identified missing values in 7 columns and inconsistencies (e.g. the races weren't in order).

•Preliminary EDA Insights:

- Dominance of certain teams (Mercedes & Red Bull) and drivers (Lewis Hamilton and Michael Schumacher) in historical data.
- The data is imbalanced (95.39% non-wins, 4.69% wins) which can challenge prediction models to avoid bias towards the majority class.

Next Steps: Data Processing, Feature Engineering, & Modeling

Data Cleaning: Refine dataset by handling missing values and duplicates

•Feature Selection: Identify predictive variables for race outcomes.

•Feature Engineering: Create new features from existing data for insights.

•Normalization: Scale data to prepare for modeling.

Baseline Model: Start with logistic regression for classification.

Thanks for your time