

COMSATS UNIVERSITY



WAH CAMPUS

Title: Formula 1 Case Study

Submitted By

Name:

Shayan Babar	M.Arsal	Ali Raza	Saad Rashid
--------------	---------	----------	-------------

Registration No:

Fa19bcs051	Fa19-BCS-056	Fa19-BCS-047	Fa19-BCS-194
------------	--------------	--------------	--------------

Class/Section:

BSCS/6d

Submitted To

Teacher Name: *Ammara Zamir*

Date of Submission: *27/02/2022*

Formula 1 Case Study

Link# <https://aws.amazon.com/solutions/case-studies/formula-one/>

Speed is important whether you are loading large amounts of data or racing in a formula 1 car. With information being driving factors, how much downforce the car is caring, fuel loads, suspensions or how the aerodynamics are managed like any other sports sensors measure everything later analyst measure each and everything to analyze performance. This is where big data comes in view. The entire game is dependent on speed that is why experts have turned to data analytics to level up their performance.

Five V's

Volume: F1 teams handle loads amount of data. A single f1 car generates 1gb of data when its speed is at 200mph, and these cars reach up to 250mph (350 km/h). Big teams like Mercedes Amg have 200+ sensors on their cars and can collect more than 2gb of data on a single lap which on average takes 1 min 20 sec. And on a racing week these cars complete hundreds of these laps to simulate racing pace and collect data.

Variety: The diversity of data collected by f1 teams are different in range not only are they collecting data on how much speed is being generated by they collect as much as raw data as they can i.e how much response time does a driver have or how long does it take for the tyres to reach optimal temperature or g-force created at a single corner even data on drivers weight loss after a single race is also recorded. Constant video and audios are being recorded then analyzed to check what speed line a driver took to enter a certain corner and check if there is anything wrong with the car while running broken parts etc.

Velocity: This might be the most important v for the f1 teams as these races are held around the world and they have a large analytics teams and machineries for commuting therefore they cannot bring their whole analytics teams to the racing tracks. They are usually at the headquarters. Therefore, data needs to be transferred and results need to be received at great speeds there for they implement highly responsive architecture. Oracle Red bull f1 team have a huge \$100 million per year contract with Oracle who manage all the data engineering part of their team.

Veracity: Quality of data matters a lot for an f1 racing team. The quality is depended on how the data is collected or processed any fault in collection of analysis by slightest of margin can disturb the whole racing week. For example, analysis of how many liters of fuel needs to be filled to go through the whole race. If the data collected for this purpose was on a different engine mode or factors like yellow flags or rains were not analyzed will result in failure to complete race due to lack of fuel. Also adding more than required fuel will affect race pace as the car will be slower and other cars on less fuel will overtake easily. Hence precise calculation is needed.

Value: F1 teams use visual analysis capabilities to help drivers analyze on how to approach a circuit what corner speeds do what engine modes needed for different sectors of the racing track. Also race dynamics teams use this analysis to drill down different scenarios how the race will go on and build different strategies on that. Going back to 21 years there was no such reliance on data but now data is as basic as the fuel for the car for a formula 1 racing teams. The TIBCO Data science team has also developed custom “concertina” visualization by using “[Spotfire Mods](#)” for analyses on race day for Mercedes Amg f1 team .

Data Science Life Cycle

Business understanding: To build a proper model it is important to understand the problem the model will be helping to solve. Consider a f1 teams need a model for predicting how much the different tire air pressure effect the speed of the car. You first need to understand this problem. What the team is wanting to achieve from the prediction do they need to check how low pressure they can keep or how does this effect the dynamics of the car. In these situations, you need consultation with the experts in the domain and understand the problem.

Data acquisition and understanding: After understanding the problem statement clearly, we need to collect the data then break the problem into small components.

We need to understand what data is needed as a rough view taking about a tire point of view it might seem we only need to run the car on different air pressure then collect the highest records. But no there are 100 of other data points you also need to collect the compound of the tyres or will this increase the cars height etc.

All this different data points will create a complex and messy data set therefore data wrangling (process of cleaning data) needs to be applied.

Data needs to be explored to find patterns and get useful information.

Modeling: This is regarded as the core process. First, we select the right model that will give high accuracy. Depending on the data collected from different tyres compound and there their effect on different parts of the car then apply accurate machine learning models to the data. Generate results using different algorithms then test for accuracy.

Deployment: Last step is to deploy this model especially formula 1 being on the road all the time they need quick access to the data. The model can also be deployed so that it can give direct prediction to the drivers at real time on their steering wheel to change between different setups to support current air pressure of the tyres.

References:

[F1 Insights powered by AWS \(amazon.com\)](#)

[Data analysis in Formula 1: the difference between victory and defeat - Grupo MAPFRE Corporativo - Acerca de MAPFRE Data Valuation — What is Your Data Worth and How do You Value it? | by ODSC - Open Data Science | Medium](#)