Proposal

Introduction

The 2021 Formula One season was one of the most exciting seasons in its history. Two rivals; Lewis Hamilton, and Max Verstappen, were neck and neck all season, with Verstappen eventually clinching the World Driver’s Championship. However, the driver that wins the championship is not always the best driver that season.

Questions

The question I will be answering is: Who was the best driver during the 2021 season; Lewis Hamilton, or Max Verstappen? I will be analyzing various data categories to determine the best performing driver over the past season.

Data

I will be using [this](https://www.kaggle.com/datasets/rohanrao/formula-1-world-championship-1950-2020/code?resource=download) dataset of the Formula 1 World Championship.

Approach

I will first import the data by uploading the CSV files to GitHub and then importing them to Python using the Pandas library. I will then clean it by removing any unnecessary data such as drivers and races not in the 2021 season.

I will analyze many data features between Hamilton and Verstappen such as Qualifying pace, average race pace, and driver mistakes. Various techniques will be used to analyze the data, and different data features will require different methods. Some visual aids will include line graphs, bar graphs, and box-and-whisker plots. These will be combined to illustrate who the ‘better’ driver was over the whole season.