

BDA - Project

Jan Nyberg, Carl-Victor Schauman

4/12/2022

Contents

Introduction	2
Description of the data	2
Description of models	3
Priors used	3
Rstan code	3
Running of stan model	3
Convergence diagnostics	3
Posterior predictive checks	3
Predictive performance assessment	3
Sensitivity analysis	3
Model comparison	3
Discussion of issues and potential improvements	3
Conclusion what was learned from the data analysis	3
Self-reflection	3

Introduction

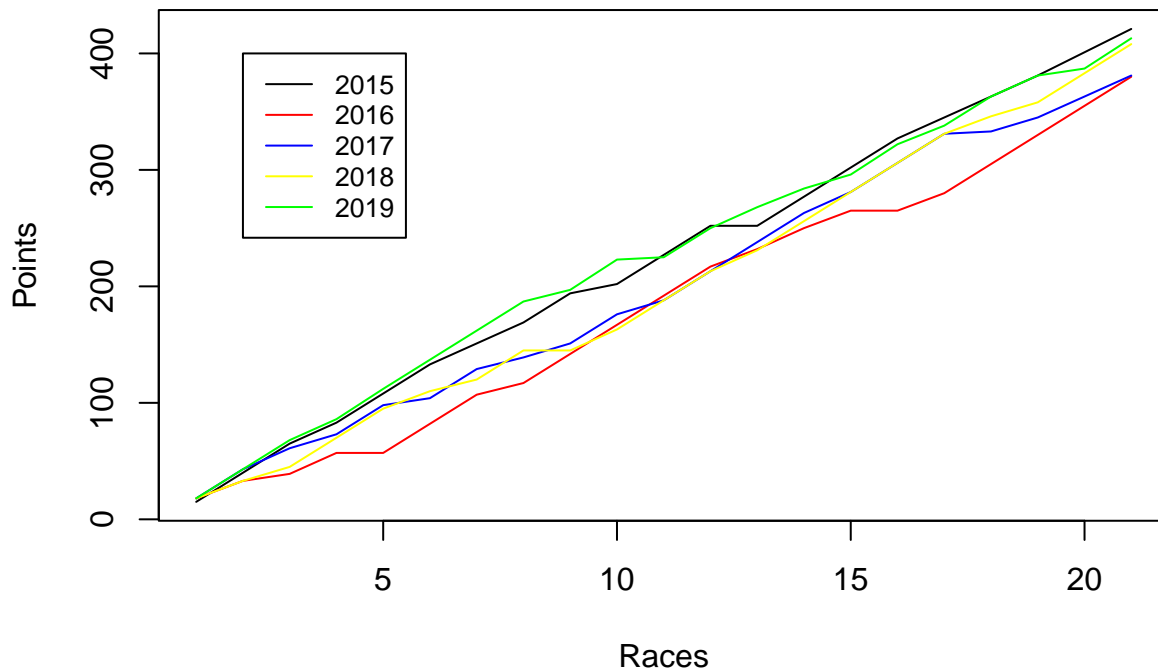
For this project, we are trying to predict Lewis Hamilton's average score in a season based on previous years scores. This is mostly just due to our curiosity if we are able to use this to somehow predict the score.

We are modeling his scores from five years, and trying to build a model using it. We want to see what kind of distribution the answer will be and how well it is able to estimate the following year. There are many factors we don't take into consideration, but we hope to see relatively good results and predictions.

Description of the data

The data we use is from Kaggle and can be found [here](#). We took Lewis Hamilton out of the data and chose the years 2015-2019. We selected his scores from all the races from those years. One thing to note with the data is that every year doesn't have an equal amount of races. To account for this we chose to fill in the missing races with the median for the year. This makes the data a bit inaccurate, however, it shouldn't have too big an effect on the data.

Hamilton cumulative points



Description of models

Priors used

Rstan code

Running of stan model

Convergence diagnostics

Posterior predictive checks

Predictive performance assessment

Sensitivity analysis

Model comparison

Discussion of issues and potential improvements

Conclusion what was learned from the data analysis

Self-reflection