

What Makes the Best Tennis Player?

Steven Herrera and Ethan Shen

11/09/2018

Introduction

Data

We used the following data set: `atp_matches_2017`, created by JeffSackMann on GitHub.

Here is a link to the dataset: https://github.com/JeffSackmann/tennis_atp/blob/master/atp_matches_2017.csv.

The dataset that we will be using for our analysis is not the same as the one we initially received, as multiple dataset manipulation techniques were used in Excel and in R. Look at **Additional Work**.

Here is a glimpse of our dataset:

Analysis

- 1) Description of the response variable and its variable type
- 2) Description of the explanatory variables and the population coefficients you wish to understand using statistical inference
- 3) Hypothesis/hypotheses regarding your question of interest
- 4) Proposed methods to use for the analysis

Reference

Additional Work

- 1) In order to obtain the dataset in our RMD, we will first need the data to fit nicely into a CSV file. In excel, this is what the dataset originally looked like before and after using the **Text to Columns** function:

```
png("before.png", 490, 350)
```