---

title: "Top 10 Boxers with Highest Net Worth"

author: "LaChandra Ash"

date: "February 26, 2020"

output: html\_document

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

**### Introduction**

Large and small businesses have been using Big Data analytics for years, and it plays a major role in sports. The NFL, premier league, and NBA have been utilizing big data, and boxing falls behind on the big data usage. Boxing is known to be the world's most ancient sport and is way behind on using big data. The average boxing fans are older than 50 years old. Unfortunately, boxing is losing fans and integrating analytics may help to increase the fan base again.

**### What can data science do to help increase profits and fan base for the boxer and trainer?**

Big data analytics may help to increase the boxing fan base. The first goal is to attract and maintain a younger fan base in boxing. The top eleven boxers in the industry with the highest net income include Mayweather, Foreman, Pacquiao, De La Hoya, Lewis, Sugar Ray, Vitaly, Ali, Hagler, and Hopkins. These eleven boxers were very successful in boxing, and their characteristics, boxing record, and net worth may help boxing trainers to choose boxers that have the potential of becoming a successful boxer and attract larger audiences to the boxing events. A larger fan base will increase the net worth of a boxer and increase the profits that will be received by the trainer.

**### Data that will be used to address the problem.**

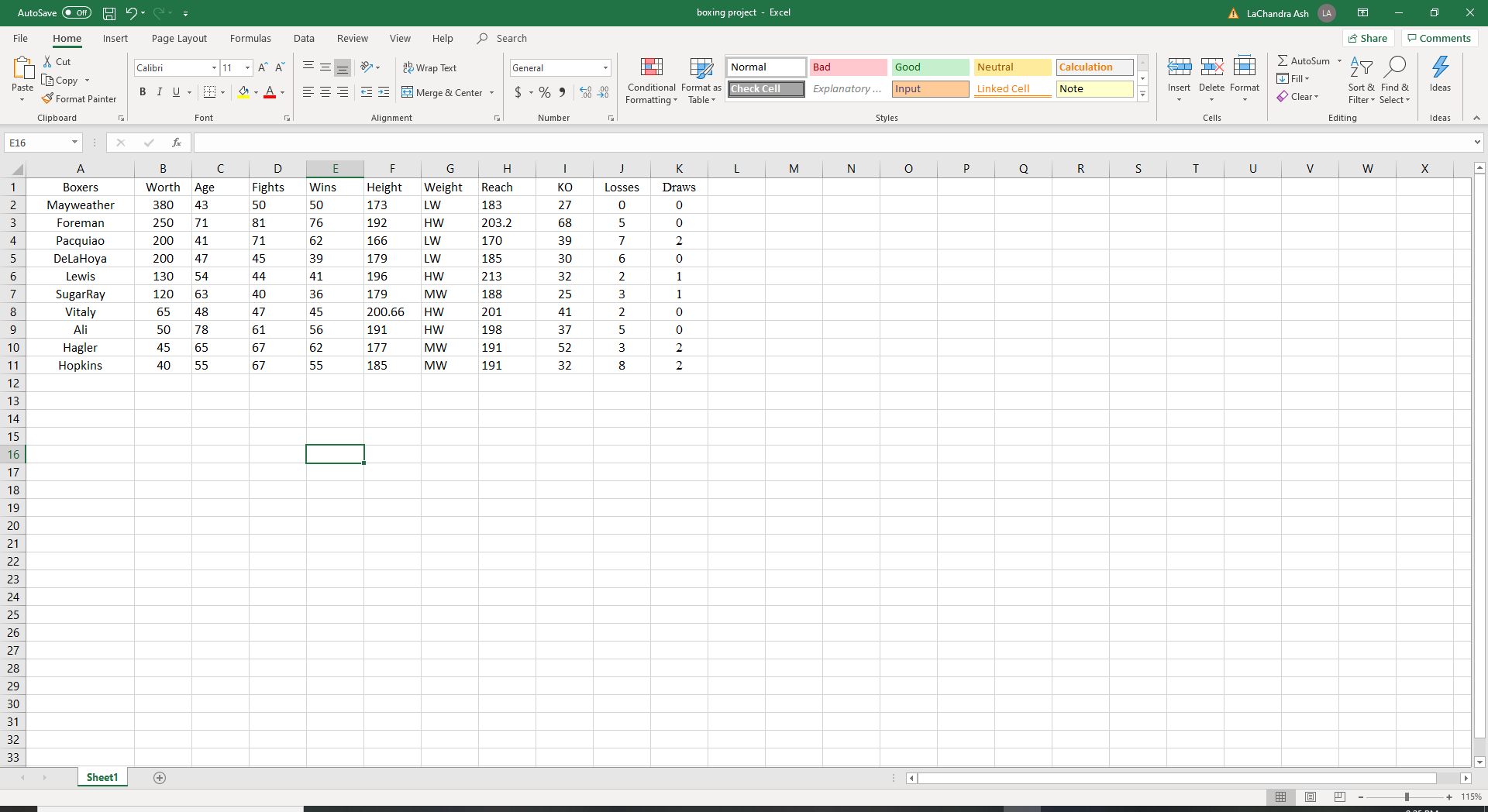
The data that I have used to address the problem include the boxer’s names, their net worth, age, total number of fights, total amount of wins, total amount of knock out wins, their losses, draws, weight class, height, and the length of their reach. I am going to compare the boxers’ physical characteristics and boxing records, to find relationships between boxers whom characteristics and boxing record made them the best boxing legends.

**#Preprocessing and cleaning the data.**

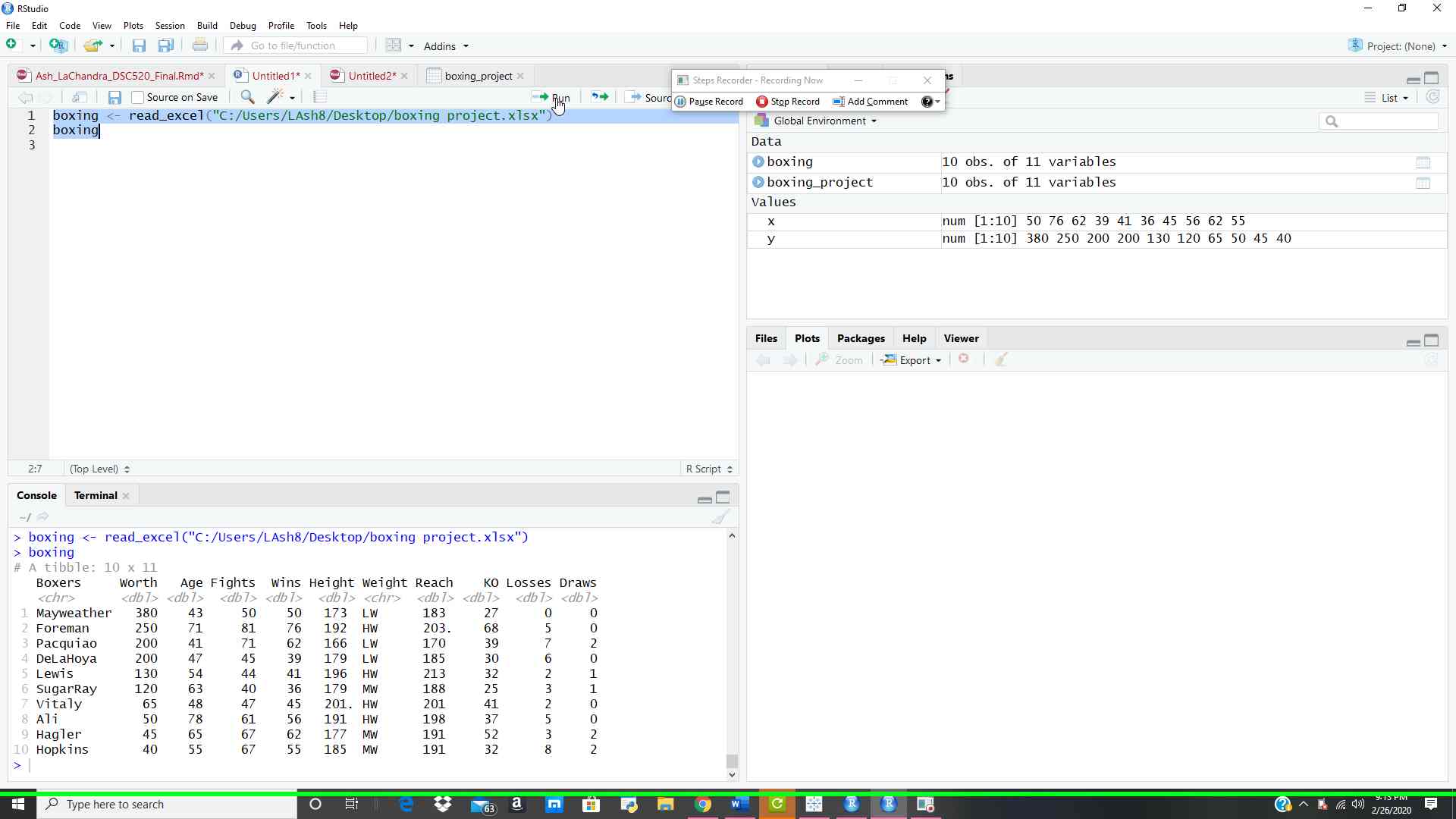
I started the preprocessing task by analyzing the data within the dataset in Excel. I searched for misspellings, missing information, abnormal variables, and typos. I center aligned all text in each row. I did not see any the raw data set looked good and I centered the variables within each cell so that the data will have a better looking appearance. I expanded the boxer column to allow their names to fit better within the Excel cell. The raw dataset was good and just had some minor issues that needed to be corrected.

**#Imported the Data**

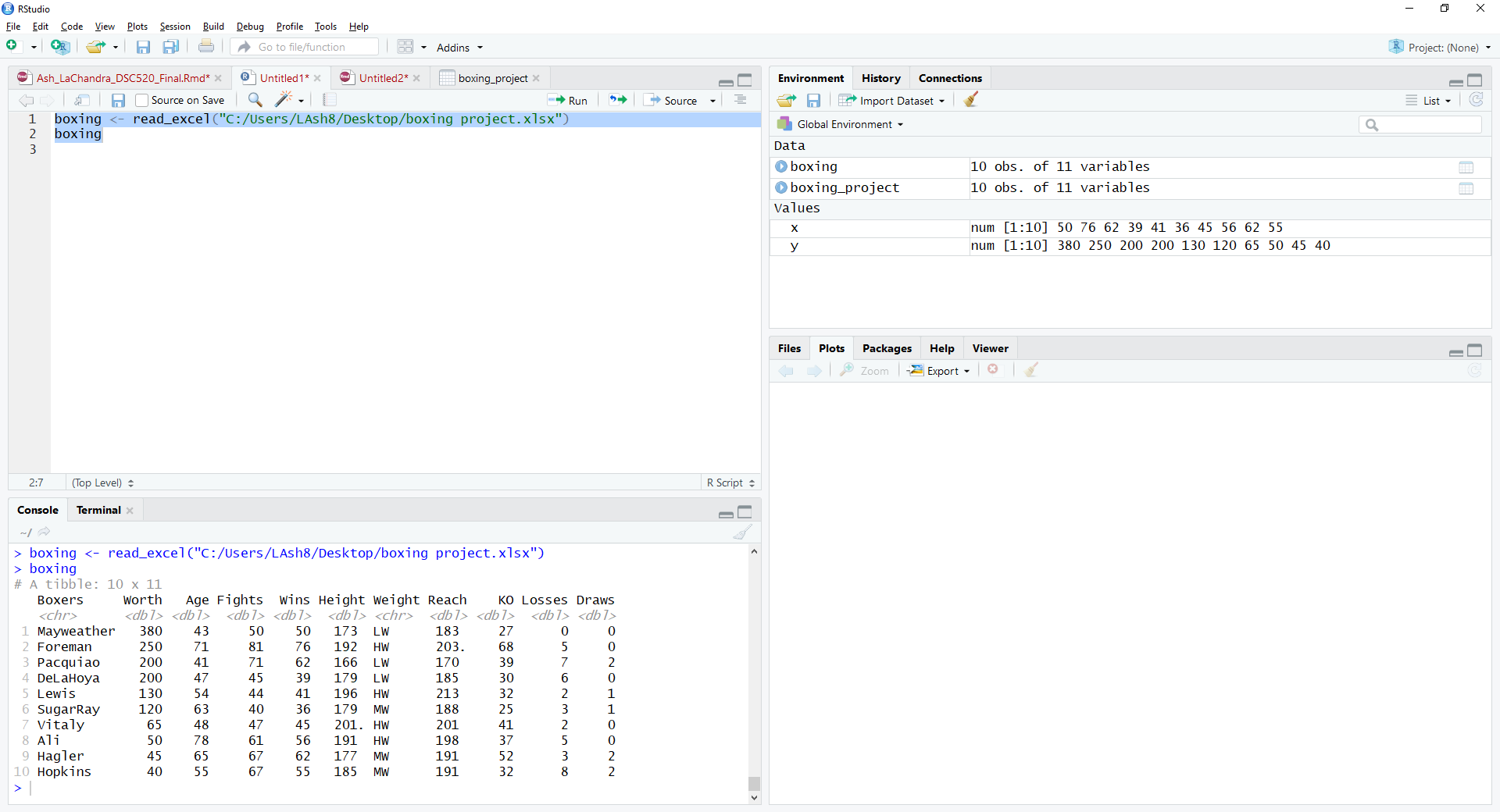
**1. The raw boxing dataset was located within Excel**.



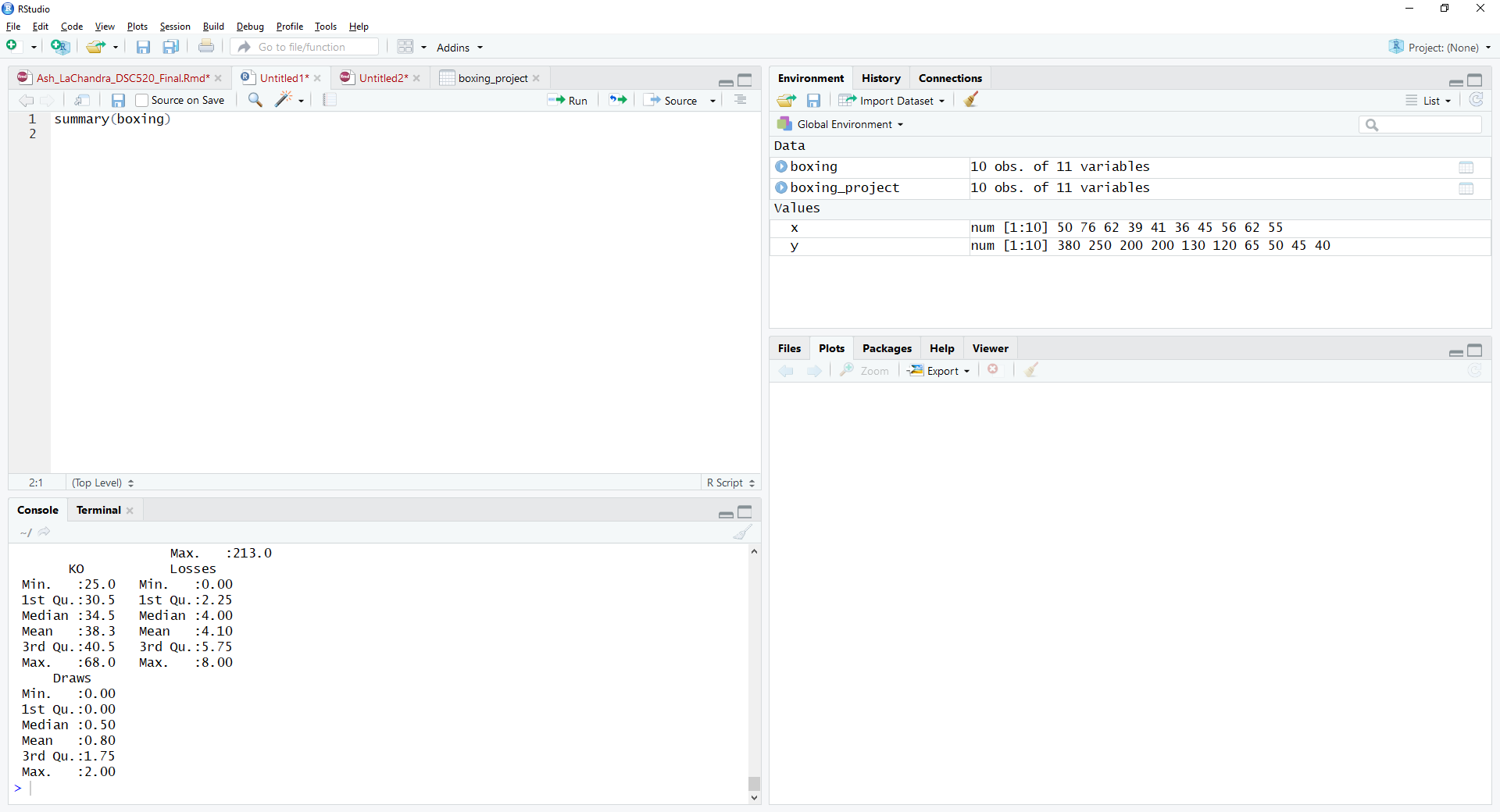
**2. The clean boxing dataset was imported into RStudio.**



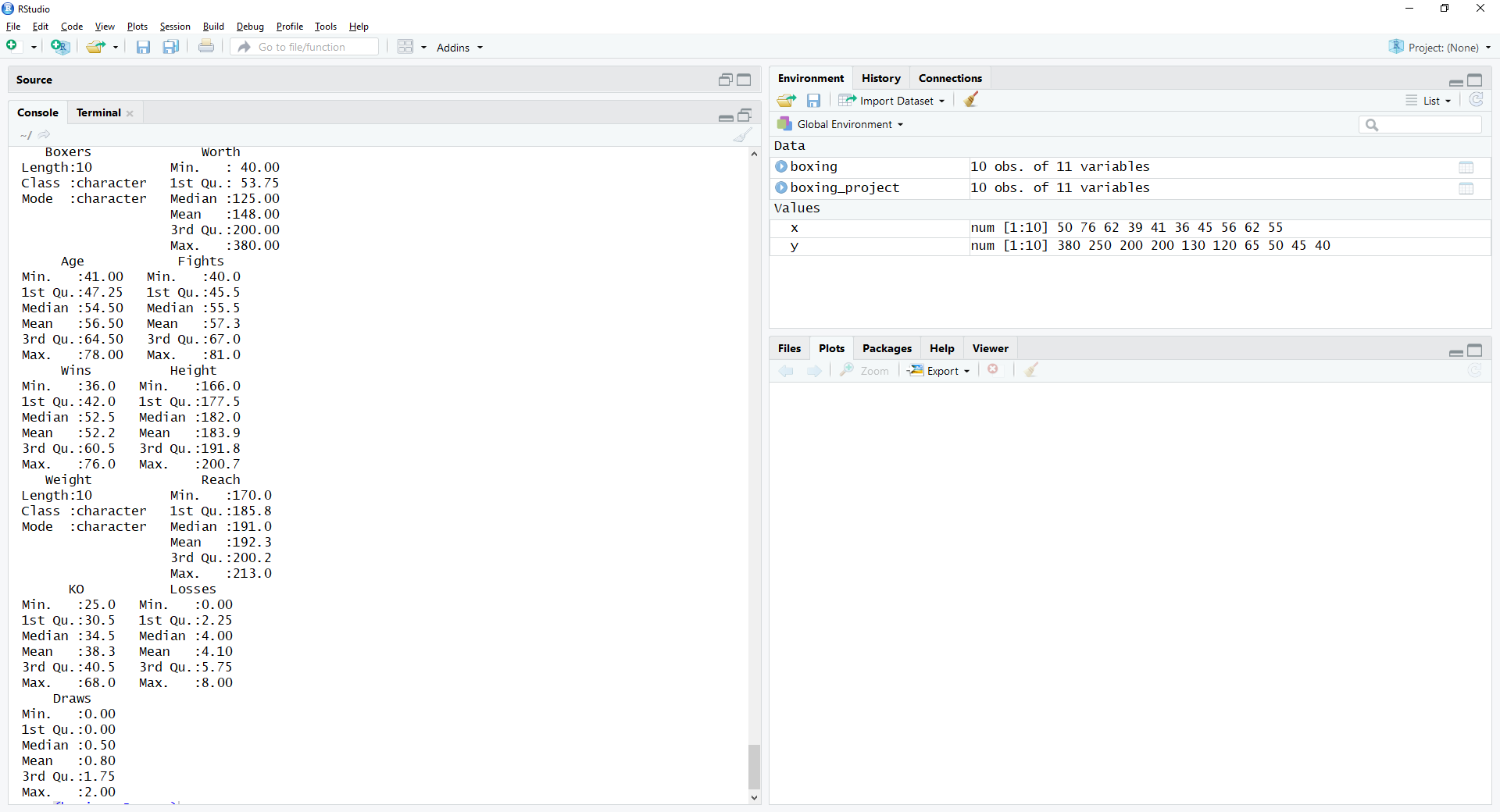
**3. The output of the printed dataset.**



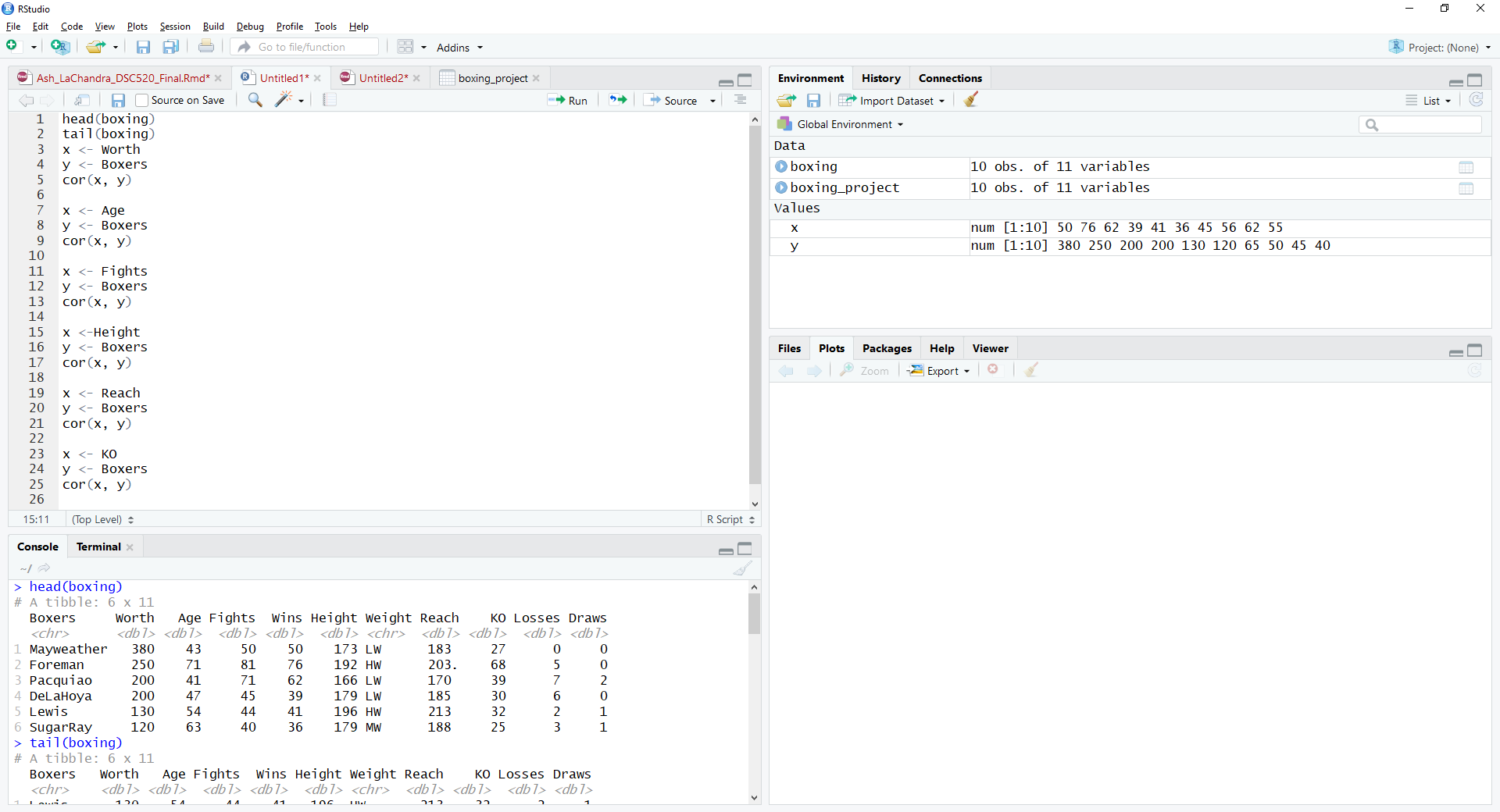
**4**. **Summary of the boxing data.**



**5. The output of the boxing data.**



**6. Correlation of the dataset.**



|  |
| --- |
| Correlation output for all: [1] 0.1057331 |
|  |
| |  | | --- | |  | |

**###Interesting insights that the data provided.**

The interesting insights were Mayweather and Foreman had the highest net worth versus the other boxers. Hopkins, Ali, Vitaly, and Hagler net worth’s were lower than the other six boxers. Mayweather continues to have the highest net worth versus the other nine boxers. Some boxers competed in more fights than the other boxers. Foreman participated in more fights versus the other boxers, and Sugar Ray had the least amount of fights underneath his belt.

The boxers with the highest net worth may have more sponsors and fought more fights versus the other boxers. Mayweather was the only boxer whom never was defeated in the boxing ring. Hopkins had the least amount of net worth versus the other boxers.

**###Limitations of analysis**

I did not see limitations in the analysis, and I believe that the boxers not listed in this project can increase their net worth if they possess similar physical characteristics and outstanding boxing records like Mayweather and Foreman. Boxing trainers can use this analysis to find the trends in the to get a good idea of which boxers they train may become successful and increase the growth of profits for the trainer.

