

# Pokemon

## Tidy Tuesday Week 13

Nicole Bonge

### Table of contents

<b>1</b>	<b>Motivation</b>	<b>1</b>
<b>2</b>	<b>Research Question</b>	<b>1</b>
2.1	Data Visualization . . . . .	2
2.2	Linear Regression Results . . . . .	2
2.3	Conclusion . . . . .	3

## 1 Motivation

Pokemon is a global icon. It is a TV series that expanded into video games, Pokemon cards/collectibles, and merchandise that made it into a multimillion dollar franchise. Growing up, Pokemon was one of my favorite shows and to this day, it remains one of the fondest memories from my childhood.

## 2 Research Question

Before I begin my project, I should know what I am doing and how to do it. This quarter, I am in my first *Introduction to R* course so my project directly applies the course material. I utilized my lecture notes to learn the technicalities of coding it and the mathematics behind each line I wrote.

After doing my research, I took this approach to solve my own problem: **Could I predict battle outcomes?**

My report is organized into the following section:

1. Loading the Data Set

2. Data Visualization
3. Linear Regression.

## 2.1 Data Visualization

Data visualization is a powerful tool. Figure 1 shows the characteristics by their Types.

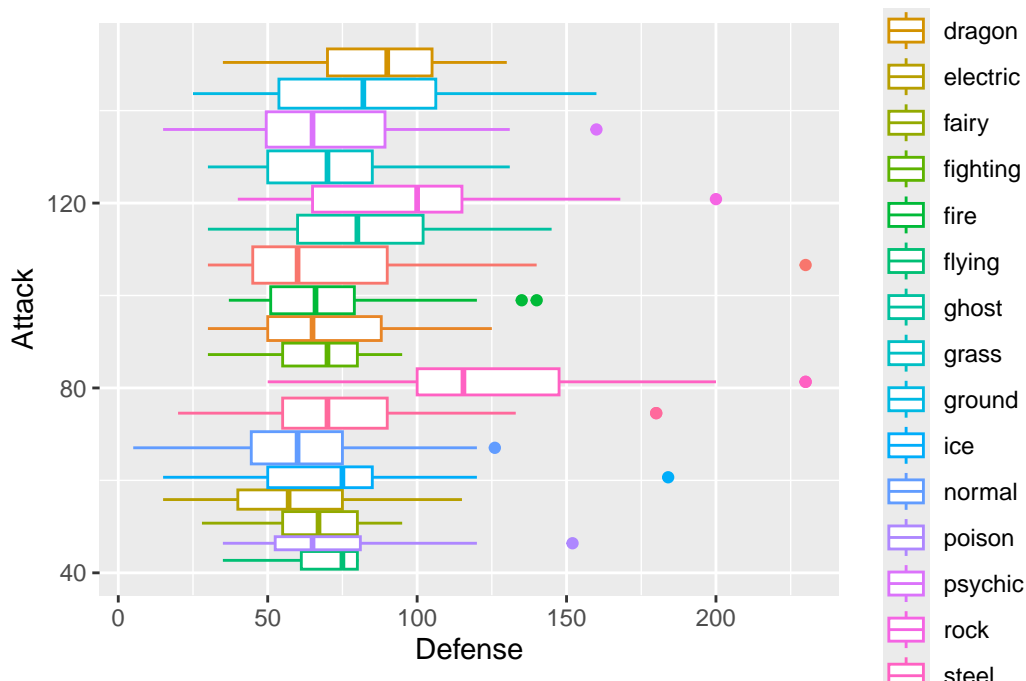


Figure 1: Boxplot of denfense and attack by their types

Second, I check out the relationships among six dimensions of information of Pokemon using pairwise scatterplot. As shown in Figure 2, I did not see clear relationships.

## 2.2 Linear Regression Results

Table 1: Linear regression of six indices on HP

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	29.9779308	2.6931666	11.131109	0.0000000
Speed	-0.1134998	0.0292901	-3.875018	0.0001140
Attack	0.3055534	0.0281869	10.840257	0.0000000
Defense	-0.0721483	0.0313883	-2.298572	0.0217472

Table 1: Linear regression of six indices on HP

	Estimate	Std. Error	t value	Pr(> t )
SpAttack	0.1405615	0.0282497	4.975673	0.0000008
SpDefense	0.2442290	0.0348479	7.008426	0.0000000

As Table 1 shows, *Speed* and *Defense* have negative effects on *HP*, while *Attack* has positive effect on Pokemon's HP. *Defense* has no significant relationship with *HP*.

## 2.3 Conclusion

With this project, I achieved a lot of firsts. It was the first time that I worked with such a large dataset. Not only were there 800 observations, but there were also 12 different variables. This meant I had a lot to play around with and explore, but also there was room for a lot of error.

It was also the first time I have worked with multiple linear regression. I am accustomed to performing linear regression analysis on the relationship between two variables, so working with eight was a new challenge.

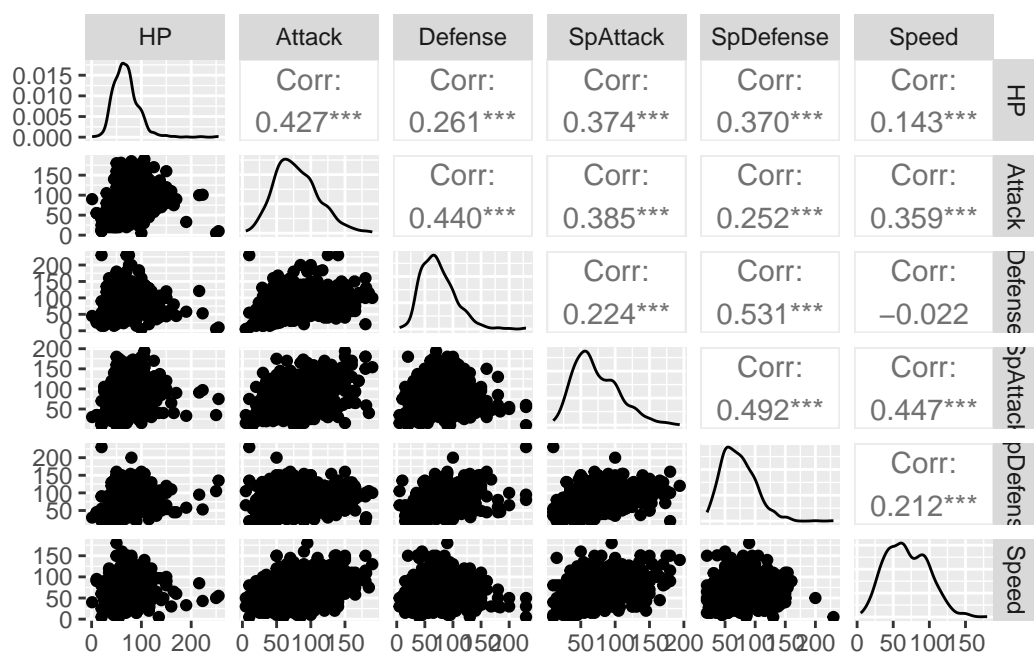


Figure 2: Pairwise scatterplot