

Stats 2600 Excerpt

Anna Sanders

12/22/2021

```
library(tidyverse)
library(dplyr)
library(readxl)

# The Whole Data
pokemon <- read.csv(file= "pokemon.csv")

# Easier to Read Data
pokemonlite <- pokemon%>% #an easier way to Look at the data
  select(pokedex_number, name, type1, type2, generation, is_legendary)
```

Note: This excerpt was taken from my STAT 2600 Final Project, where my group chose a topic and questions and completed exploratory analysis to help answer them. This excerpt includes the introduction section and my analysis section.

I WANT TO BE THE VERY BEST (like no one ever was)



Main Question and Why this is Important

Pokemon is a beloved franchise by many children and adults alike. As the franchise has grown from one video game in 1996, to now a multimedia monolith (including trading cards, movies, and a tv series), fans have been waiting in anticipation for the next game in the series, which was recently announced as Pokemon: Sword and Pokemon: Shield. Using data from the previous games (Red, Blue, Green, Yellow, Gold, Silver, Crystal, Ruby Sapphire, FireRed, LeafGreen, Emerald, Diamond, Pearl, Platinum, HeartGold, SoulSilver, Black, White, Black2,

White2, X, Y, Omega Ruby, Alpha Sapphire, Sun, Moon, Ultra Sun, Ultra Moon, Let's GO Eevee, Let's Go Pikachu), we wanted to create a list of recommendations to help become the very best (like no one ever was) in Sword and Shield, including which pokemon to bring with us from the previous generations.

Recommendations

Based off of our individual findings, we each have come up with four statements regarding various pokemon stats on how to be the very best (like no one ever was) in the new game.

1. From the analysis of the best generation and the best pokemon based off of their entry in the Pokedex, I would recommend using either Starters, Pseudo-legendaries, or Legendaries from generation 4 (Sinnoh region) on your team to become the very best.
2. Regarding legendary pokemon to put on your team, it's best to choose the pokemon Latias. If Latias isn't available, it's advised to choose a pokemon that is either a dragon or flying type, with great defensive stats.
3. Pokemon types that are more likely to be male have significantly better base totals, meaning that in essence, they are superior to the other Pokemon. Therefore, in order to try to be the best player out there, it is recommended that the player puts an emphasis on catching these Pokemon type.
4. Dragon and Steel Types are by far the most powerful on average and the new calculated epwr (an estimate for a pokemon's power based off of the highest attack stat and average Type effectiveness) can help estimate a pokemon's overall attack combat effectiveness.

List of References:

https://bulbapedia.bulbagarden.net/wiki/Base_stats (https://bulbapedia.bulbagarden.net/wiki/Base_stats)

https://bulbapedia.bulbagarden.net/wiki/Catch_rate (https://bulbapedia.bulbagarden.net/wiki/Catch_rate)

<https://bulbapedia.bulbagarden.net/wiki/Friendship> (<https://bulbapedia.bulbagarden.net/wiki/Friendship>)

<https://bulbapedia.bulbagarden.net/wiki/Damage> (<https://bulbapedia.bulbagarden.net/wiki/Damage>)

#Anna My main question deals with which pokemon are the strongest in terms of total stats (base_total column) and how to predict if a certain Type (or Types) is likely to have a higher chance of having larger total stats.

```

pokemonpwr <- pokemon%>%
  select(pokedex_number, type1, type2, name, hp, attack, defense, speed, sp_attack, sp_defense,
base_total,is_legendary)

poketype1 <- pokemonpwr%>%
  group_by(type1)%>%
  summarise(mean= mean(base_total), n=n(), sum=sum(base_total))

poketype2 <- pokemonpwr%>%
  group_by(type2)%>%
  summarise(mean= mean(base_total), n=n(), sum=sum(base_total))

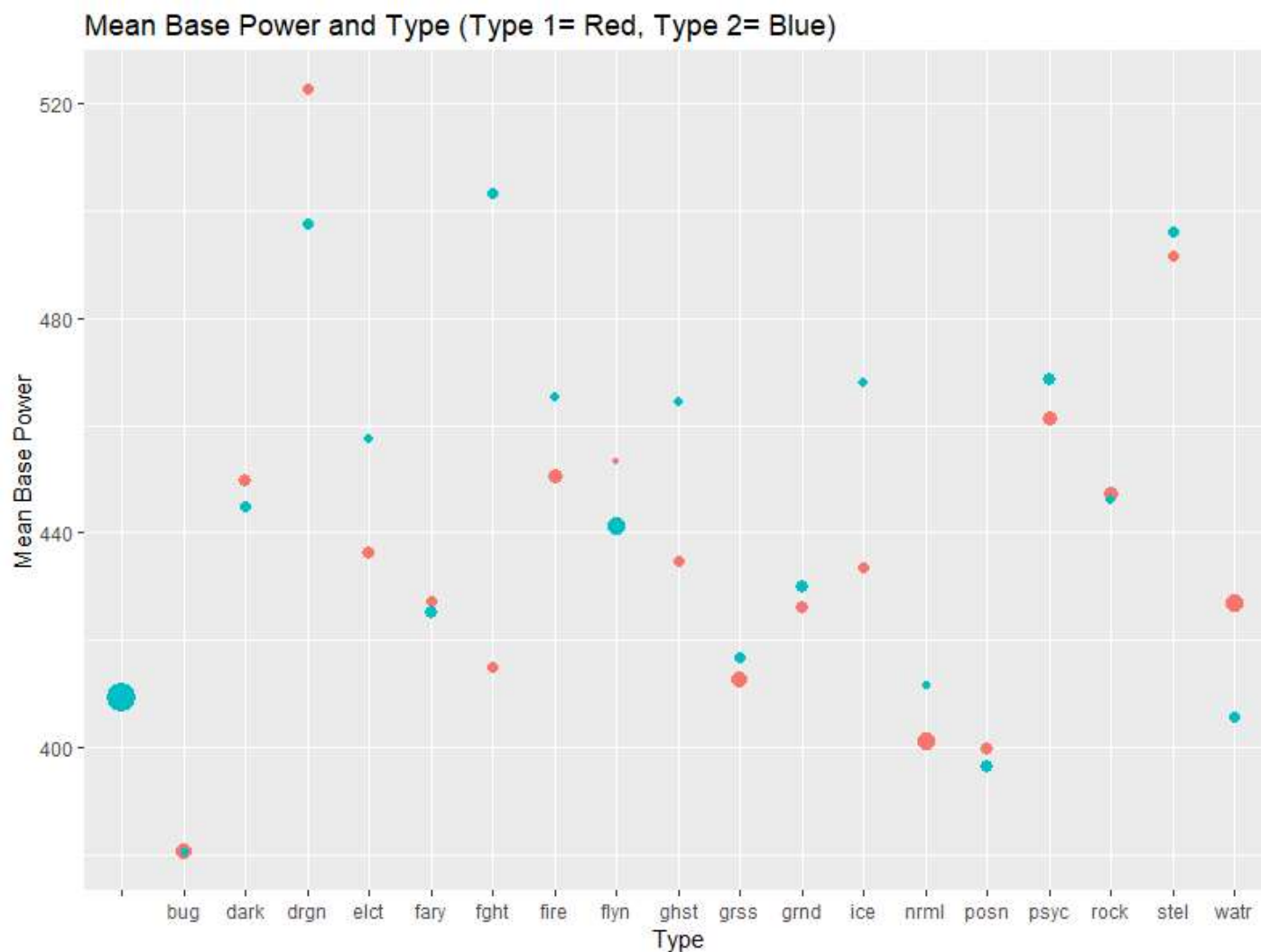
pokeeff <- pokemon%>%
  select(name, attack, sp_attack, against_bug, against_dark, against_dragon, against_electric, a
against_fairy, against_fight, against_fire, against_flying, against_ghost, against_grass, against
_ground, against_ice, against_normal, against_poison, against_psychic, against_rock, against_ste
el, against_water)%>%
  mutate(eff= (against_bug+ against_dark+ against_dragon+ against_electric+ against_fairy+ again
st_fight+ against_fire+ against_flying+ against_ghost+ against_grass+ against_ground+ against_ic
e+ against_normal+ against_poison+ against_psychic+ against_rock+ against_steel+ against_water)/
(18))

pokeeff2 <- pokeeff%>%      # I used a recently learned function apply() to pick the highest value i
n a row
  mutate(attackh= apply(pokeeff[,2:22], 1, max))

pokeeff3 <- pokeeff2%>%
  mutate(epwr= attackh*eff) #Expected Power (for comparison, not acutall calculation)

ggplot()+
  geom_point(poketype1, mapping=aes(x=type1, y=`mean`, size=n, color="Type 1"))+
  geom_point(poketype2, mapping=aes(x=type2, y=`mean`, size=n, color="Type 2"))+
  labs(x= "Type", y= "Mean Base Power", title= "Mean Base Power and Type (Type 1= Red, Type 2= B
lue)")+
  theme(legend.position= "none")+
  scale_x_discrete(labels= abbreviate)
  scale_x_discrete(breaks=c("", "bug", "dark", "dragon", "electric", "fairy", "fighting", "fire"
, "flying", "ghost", "grass", "ground", "ice", "normal", "poison", "psychic", "rock", "steel",
"water"))

```



The plot above helps visualize the mean (or average) base power for each type (the names have been shortened). As you can see, the Dragon (drgn) and Steel (stel) Types have the highest mean base power for pokemon with either as their first or second type. We can also see that the worst types on average for base power are Bug (by a large margin) and Poison.

The top four pokemon with the highest total stats:

| Name | Total Stat |
|----------|------------|
| Mewtwo | 780 |
| Rayquaza | 780 |
| Kyogre | 760 |
| Groudon | 760 |

Top three non-legendary pokemon with the highest total stats:

| Name | Total Stat |
|-----------|------------|
| Tyranitar | 700 |
| Salamence | 700 |

| Name | Total Stat |
|-----------|------------|
| Metagross | 700 |

Top three first-type with the highest average total stats:

| Type | Average Total Stat |
|---------|--------------------|
| Dragon | 522.7778 |
| Steel | 491.5833 |
| Psychic | 461.2642 |

Top three second-type with the highest average total stats:

| Type | Average Total Stat |
|----------|--------------------|
| Fighting | 522.7778 |
| Dragon | 497.4706 |
| Steel | 495.9545 |

Top pokemon for epwr:

| Pokemon | Epwr |
|-----------|----------|
| Heracross | 220.9722 |
| Mewtwo | 215.5556 |
| Tyranitar | 209.5556 |
| Kyurem | 203.0556 |
| Rayquaza | 202.5000 |

In pokeff3, I created a new column that multiplied the highest attack value (attack or sp_attack) with an average calculation of the attack multiplier for type. This new value epwr only takes into account attack and type multipliers, but it still is a good base measurement of power vs. any pokemon of any type on average. Higher numbers for this statistic relate to overall attack power as attack is calculated by multiplying both attack/sp_attack (based on what type of move the pokemon uses) and the type modifier. I thought of multiplying the type effectiveness against by the amount of pokemon of that type to get a weighted average but decided since we do not yet know the distribution of new pokemon and their types, it would be better to not (as the ratio might change).

From this data, we can conclude that one should catch any pokemon that are Dragon and Steel type, regardless if it is the pokemon's first or second type. We can also conclude that legendary pokemon are in fact more powerful than any non-legendary pokemon. Lastly, the new column can help predict how well an older generation pokemon will do in battle, but is not an exact estimate due to other factors (e.g. level, defense, move power).