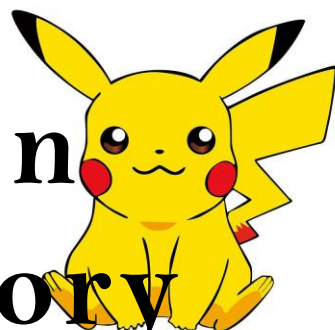LES DERNIERS
VAINQUEURS

Pokémon
Data Story

Process Book

Haolong Li, Yuheng Lu,
Xingyue ZHANG

Introduction

Problematic: An analysis into Pokémon's' abilities, attributes and types

We aim to optimize Pokémon recommendations for competitions. Key determining factors include most influential parameter, effective parameter sum, abilities, and type combinations.

We compare OU-tier Pokémon with others to pinpoint crucial parameters. The effective parameter sum, varying with roles and speeds, considers parameters relevant to a Pokémon's role rather than sum of all six. Pokémon abilities, especially those prevalent in higher tiers, could indicate potential strength. We also examine which dual-type combinations provide best defensive and offensive capabilities. These factors all contribute to competitive performance and can guide recommendations.

Data

We retrieved data from the following sources:
- [Pokemon All Status Data (Gen1 to 9)](#)
- [Complete Competitive Pokemon Dataset](#)
- [The Complete Pokemon Dataset](#)

We cleaned the data for missing values and derived statistics according to our needs.

Peer Assessment

- Haolong: Infrastructure of website, Home page, Abouts, Viz1
- Yuheng: Viz2
- Xingyue: Viz3

Viz1: Abilities

Deliverables

There are three visualizations in the Abilities Viz: a 3-d map showing all Pokémon's with either 1, 2 or 3 abilities; a word cloud indicating all abilities; and a bar-plot with adjustable quantity showing the best abilities.

…From Milestone 1?

All work are based on the data from milestone 1. I was able to get inspired to have more interesting plots than just a bar-plot from our analysis from the first milestone. To further illustrate notions about abilities, I added a word cloud to give the viewer a general feeling of what an ability is.
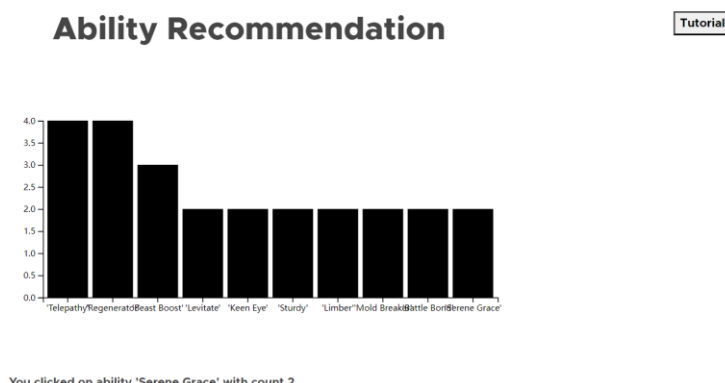
However I wasn't able to finish the initial design, which I planned to create a gallery-like exhibition of abilities, because the pictures of abilities aren't available.

Challenges

I wasn't quite content with my initial product, where I had just an ugly bar-plot.

The first design

I decided to add some coloring from [some palettes](#), also make the plot interactive, in that I can toggle something to change the plot, and when there are actions like onClick, onScrollTo, the plot can have some interesting reactions. They were not quite easy, because I had to deal with some back-end logics. For example, the non-linear toggling effect when the user first scroll to the plot is quite problematic, as I had to implement the whole thing from scratch.

The whole infrastructure of the project, on the other hand, is also quite a thing: GPT can't help! I had to understand all structures of the project, and do a lot of cleanups (the structure was far more messy before!)

Discoveries

- Occurrence of abilities follow a clear tailed distribution.
- Abilities that appear the most are not likely the best abilities, nor are the rarest abilities.
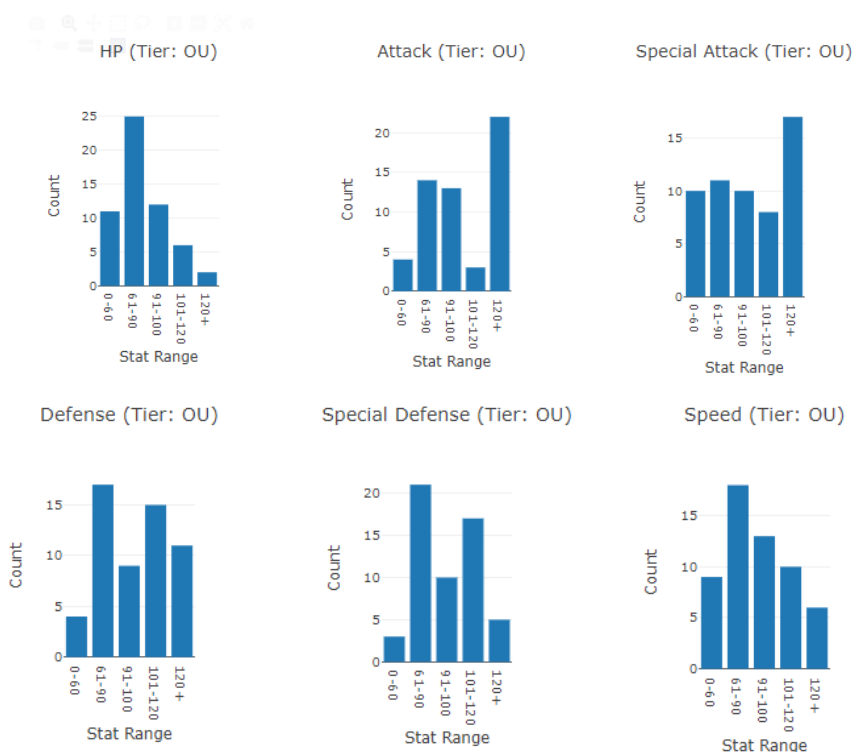
Viz2: Attributes

Challenge encountered and Change in milestone3

I would like to define effective total for every pokemon and calculate it in milestone2. However, this seems not so meaningful after visualization. And there is a challenge that it is not able to use visualizing tools to explain such a technical and detailed questions in pokemon games. Therefore, in milestone 3, I focus on a broader view on the attributes of pokemons. However, the entity of this part is still based on what I have done in milestone2. Now I will still illustrate how pokemons' attributes variate corresponding to Tier and generation and then make recommendation from these aspects.

Reuse and expansion!

First, I want to illustrate the difference in 6 basic attributes between ordinary pokemons and pokemons that are used most frequently, i.e. Tier = 'OU'. In milestone2, I only use python matplotlib to draw the pictures and all of the images are static. In order to implement dynamic data visualization while make the analysis directly, I use plotly to draw the following bar plots and make them interactively in website.

6 Attributes' Distribution for Tier OU

…Ongoing

Second, I still look into the relation between the attributes and generation. Instead of just seeing the relationship between base(or effective) total and generation, I now see every attribute of a pokemon. Take HP as an example below. You can press a button on the website to change from 7 different options and see their relationship with generation

Finally, in order to do recommendation according to different gamers' requirement, I draw a radar chart for every pokemon in OU, and let the gamers select what they want directly by visualization.

Discoveries

For Tier, we can see it almost meets our expectations. For all 6 attributes of pokemons, we can see that all pokemons in OU have higher amount of 6 attributes.

For generation, we can see that although the 4th generation has a specially high number in mean figures, the trends of pokemon games obey the mentioned routine.
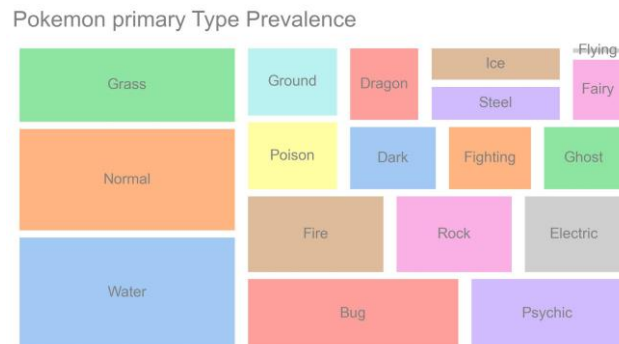
In radar chart, it is easy to visualize the features of different pokemons' attributes' distribution. Therefore, it is easier to select one pokemon by just seeing this radar chart.

Viz3: Types

Path and Challenges

My final design is mainly based on what I did in Milestone 2. For example, there was one visualization where I demonstrated Pokémon primary type prevalence:

I was wondering how I can make it more fun and interactive. Then I came up with the final design where there are bubbles in different colors and sizes. If there are more Pokémon of this primary type, the bubble will be larger. You can also click in the bubbles to see what Pokémon are in this type and check out their information.

In the beginning, I only had a bar chart graph for the type pair effectiveness. In the end, I made it more interactive and informative for the user by using a Sankey diagram.

One of the challenges I encountered was: How do I make sense of the data and tell a story? I didn't want to just have fancy graphs that provides useless information. I want my visualization to be meaningful and actually teach the reader something in a easy and fun way. I had to think about what information to include and which is the most effective way to present this information.

Another challenge is actually about how to calculate the type effectiveness value. I downloaded the 'game master' file from the actual Pokémon game and figured out how to obtain the type effectiveness value.

Care about Design

For my second diagram (Pokémon Type Effectiveness), I want to show which type dominates which other type. I chose to use the Chord diagram. A user can just hover over the chords and arcs to see which type is effective on another type and their type effectiveness value. You can also tell which type is better overall from the thickness of the arcs. I also chose colors that are pleasing on the eye.

For my third diagram, I wanted to explore the relationship between capture rate, primary type, and also generation. I used scatter plot and violin plot. Generation 1 to 7 are represented by 7 distinctive colors. The x-axis is primary type. From the violin plot, you can tell which type has higher or lower capture rate. Moreover, you can choose from 'All', 'legendary', and 'non-legendary'. It's cool when you compare the difference of capture rate between legendary and non-legendary Pokémon. The capture rate of legendary Pokémon is so much lower.

For my forth graph, I want users to easily see what the best pairings are for a a primary type. If you hover over the lines, you can easily see the effectiveness value for a type pair. This is very useful for Pokémon players.

For my final graph, I want to demonstrate the relationship between types and attributes. You can easily find out if a given type is good at attack or defense from the Radar chart. Also, the user can filter by base total value to find out which Pokémon has higher base total value. Moreover, you can choose whether to combine or separate the graphs. On the first slide, there are only legendary Pokémon. The second slide is about non-legendary Pokémon. The above elements offer a pretty comprehensive picture of types and attributes for users.