

Pokemon Data Analysis

Tony Schmitt

aschmi5@clermson.edu

C36755261

<https://github.com/Tschmitt99/datavizgroup16>

Payton Beans

pbeans@clermson.edu

C21285765

Overview and Motivation

Pokemon is a game that focuses on the collection, training and battling of powerful creatures called Pokemon in a turn based team battle. Currently throughout the pokemon franchise there are 898 unique species of pokemon. In the game each Pokemon has a collection of stats and attribute data that play into the game's battle mechanics such as a single or double typing and a set of numerical stats.

We chose our data set because Pokemon is a complex game that was a part of many people's childhoods. When we were young we didn't explore the data behind Pokemon much at all and most choices in the game made by kids are more focused on what Pokemon we thought were interesting at the time. Now we would like to go back and find a better understanding of the game and it's wide array of creatures through data analysis.

Related Work

Across the internet you can find websites that can give you the stats for each pokemon and list their individual typings and weaknesses. This gave us a good point of reference for what information we would need to include. For individual pokemon visualizations many sites used bar graphs but we felt this wasn't effective for displaying the overall power of a pokemon so we wanted to look for a more effective method. As for visualizing all pokemon together there were very few examples aside from tables that could be sorted by a stat.

Questions to answer

Primarily our objective was to compare pokemon based on their stats and typings and help identify any pokemon with exceptional stats in some area that may go unnoticed due to shortcomings elsewhere. In the games this data is behind a set of menus and only gives the player the raw numerical data. This makes comparison between pokemon tedious ingame as the player must transcribe the stats of each pokemon to compare outside of the game or constantly change between menus. This inefficient menu design and poor information display leads to many players never taking the time to analyze the stats of their pokemon. While some players won't care, there is also a competitive scene for pokemon filled with players who want to understand every point of a pokemon's statistics. Creating a tool that allows for a rapid comparison of pokemon in a visual channel would help increase understanding and accessibility.

We also wanted to explore ways to overlap the stats of two or more pokemon so that a player can try to build a team of pokemon that covers all the stat metrics with strong pokemon.

This could also be helpful in identifying shortcomings in a player's existing team. Many websites can display multiple pokemon's stats in a table but we found this to be an inefficient design from a consumer standpoint. A row of values does a poor job of displaying power disparity when compared to many other visualizations and can show the scale between the values. We however did feel that there's a place for pure numerical data in pokemon so each pokemon selected does have its base stats and typing displayed below our visualizations.

Data

<https://data.world/steveinatx/pokemon-index/workspace/file?filename=pokemon.xlsx>

The source for our data comes from an international enterprise data catalog Data.World that allows people from across the world to upload datasets. Our data consists of the following 4 sections that we will be able to digest in various forms.

Pokemon:

This section of the data consists of all 898 unique species of Pokemon from across the generations of the games. This includes their name, their elemental type (one of 18 possible types), their health points or HP, attack stat value, defense stat value, speed stat value, as well as their special attack and special defense stat values.

Moves:

This section consists of all of the possible moves that Pokemon can perform. There are 608 total moves. Each value includes the move's name; its elemental type (much like the Pokemon themselves); its category (physical, special, or status); its power stat value; its accuracy from 1-100; its power points of PP (which is a numerical value representing how many times in a battle that a Pokemon can use that move); whether or not a technical machine or TM is required to gain access to the move; its 'effect' which is a description of what the move does; and the probability percentage that the effect it has (if applicable) hits or not.

Evolution:

This portion of the data includes which Pokemon evolve and what they can evolve into. It includes the names of both the evolving Pokemon and its evolution name, the level that it needs before it can evolve, any special required conditions for the evolution to occur, and the evolution type (level, stone, trade, happiness, other).

TypeChart:

The last section of the data is a very important part of the Pokemon battle mechanics; the TypeChart. In the Pokemon games, certain elements or "types" are considered to be more effective against others, which then provides multipliers to move stats when used in-game if the appropriate types are in play. This part of the data consists of the attacking Pokemon's type, the defending Pokemon's type, its effectiveness against the defender (normal, not very effective, no effect, or super effective), and the stat multiplier.

Data processing

In order to get our data into a consistent format and address an outlying information in the data set we did have to do some processing work. The pokemon games are released in generations and currently there are 8 unique generations of the game that each introduce new pokemon and game mechanics. In order to keep our data more consistent across generations we will remove Pokemon that only exist through the use of a game mechanic in later generations. An example of this is Mega Evolutions, these are upgraded versions of existing pokemon and, while they have unique stats and typings compared to the non Mega Evolved counterparts, not all Pokemon have a Mega Evolution.

We've also had to standardize the stats that each pokemon has. In the first generation of Pokemon each Pokemon had only one stat that contributed to attack or defence appropriately named "attack" and "defense", later generations divided each of these stats into a normal and special attack/defense stat. The first generation used the more simple stat spread for both cases of normal/special. To standardize the data we have added a special attack and special defence to each of the Pokemon from the first generation that is just a duplicate value of the original non-special stat. This wouldn't change any aspect of the game play and creates a more uniform data set for analysis.

Initial Visualization Design Prototypes:

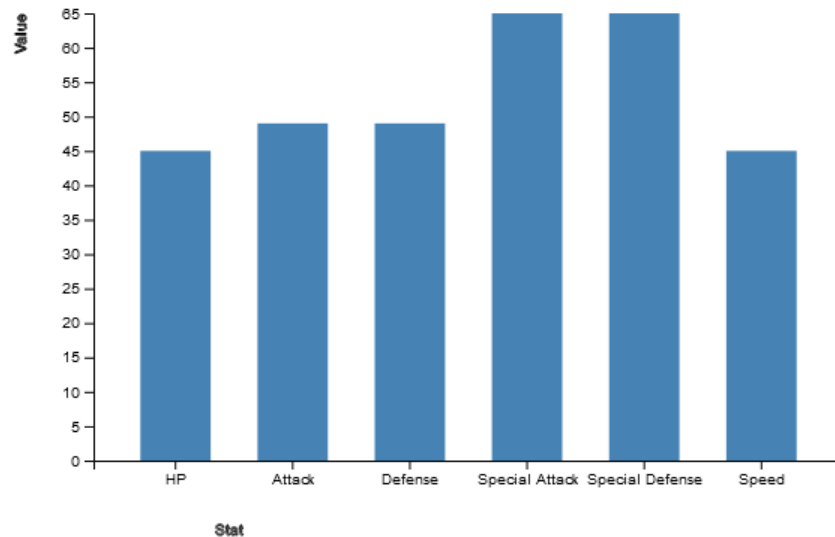
Below is a list of visualizations that we created when we were originally coming up with conceptual ideas for our project. We initially had some challenges coming up with ways to interpret the large amounts of data that came with our Pokemon dataset, so it was imperative for us to come up with prototype designs early so that we could decide which visualization ideas had value and which did not.

Early Visualization Concepts

1. Bar chart - individual Pokemon stats

Originally designed to also have a selector to pick a pokemon dropdown list or search bar to select. We moved away from this design as it became apparent that a bar chart would do a poor job for our comparisons. While comparing the speed of one pokemon to another would work well, there was no way to see a more holistic comparison. The best we could do would be a comparison of stat totals but we would then lose perspective on the individual stats.

Bulbasaur stats (GRASS)



- Concept art-

2. Normal Chart - type chart effectiveness

We were sure we wanted to include some way of displaying how the typing of a pokemon affects its stats in comparison to the base stats of a pokemon. It is very easy to look online and find charts like our example but these charts are always generalized. Ultimately cut as we couldn't come up with a way to make it interactive

- <https://codepen.io/FelixRilling/pen/ExdIC>

	Fire	Water	Grass	Electric	Ice
Fire	1/2	1/2	2	-	2
Water	2	1/2	1/2	-	-
Grass	1/2	2	1/2	-	-
Electric	-	2	1/2	1/2	-
Ice	1/2	1/2	2	-	1/2

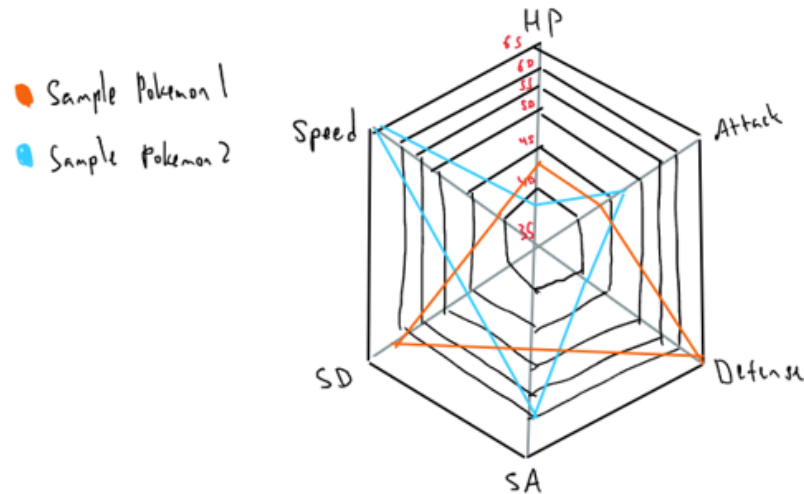
- Concept art-

3. Radar chart - Pokemon strengths & weaknesses

We ended up using the radar chart in our final design. We felt it kept an excellent balance of showing the magnitude of each individual stat while the area enclosed visualized how the stats all added up. This also allowed for easily overlapping the charts for multiple pokemon. The user would still have an

easy time seeing which pokemon has the higher HP stat and can identify which pokemon has more stats overall.

- Reference code - <https://www.anychart.com/blog/2020/07/07/radar-chart-js/>

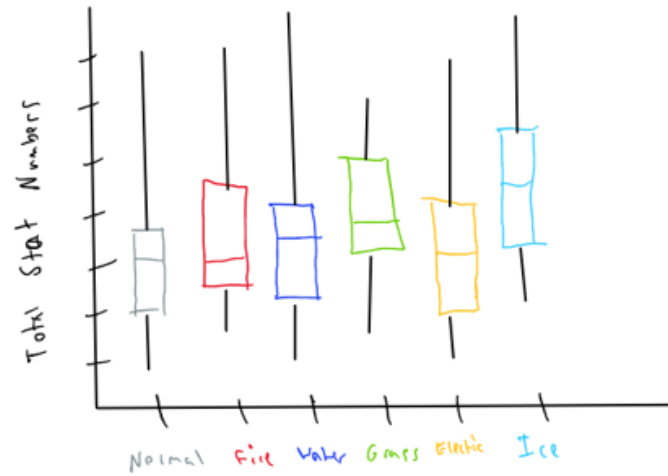


- Concept art -

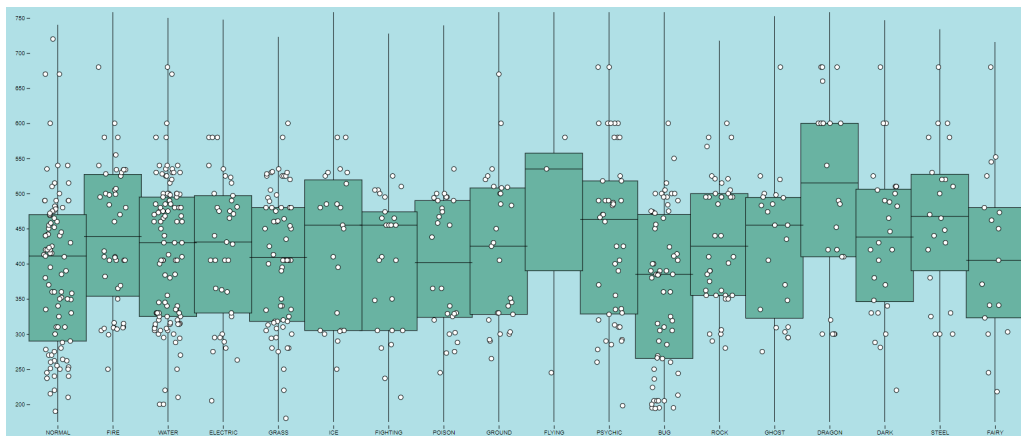
4. Box and Whisker Plot - Strongest/Weakest overall types

We developed this concept for a while, even having a working example shown below. We eventually scrapped the idea as the box and whiskers overall didn't make the data easier to parse. While it did help highlight which pokemon were more average and which stand out we felt that this was already apparent enough from just a scatter plot. We kept the origination of type groupings as typings are very important in the game and users will commonly want a pokemon with good stats of a certain type.

- Resource links
 - Source explained (go down to the box-and-whisker plot section) - <https://thelostlambda.github.io/pokestats/>
 - Helpful reference code for how to make these plots in d3 - https://www.d3-graph-gallery.com/graph/boxplot_horizontal.html

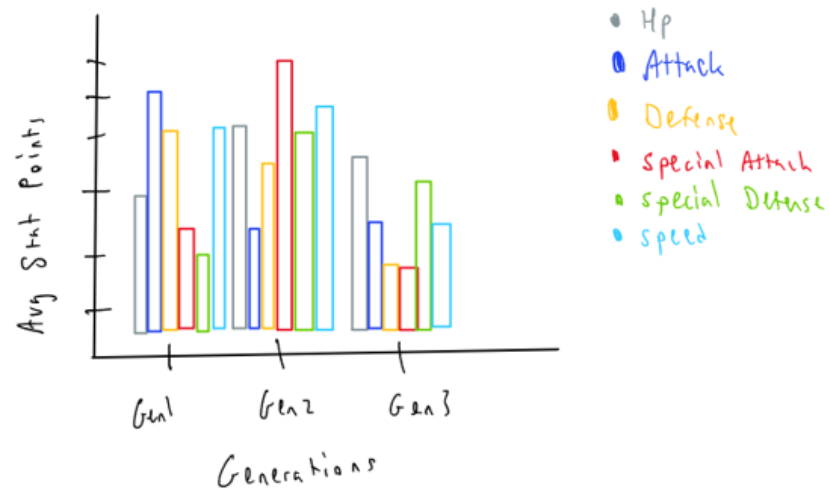


- Concept art -



5. Bar chart - average overall types by generation

Originally planned to be used in concert with the individual pokemon bar chart. This idea was quickly cut as it was uninteractive and we preferred to show each pokemon as individual points rather than in these larger groups.

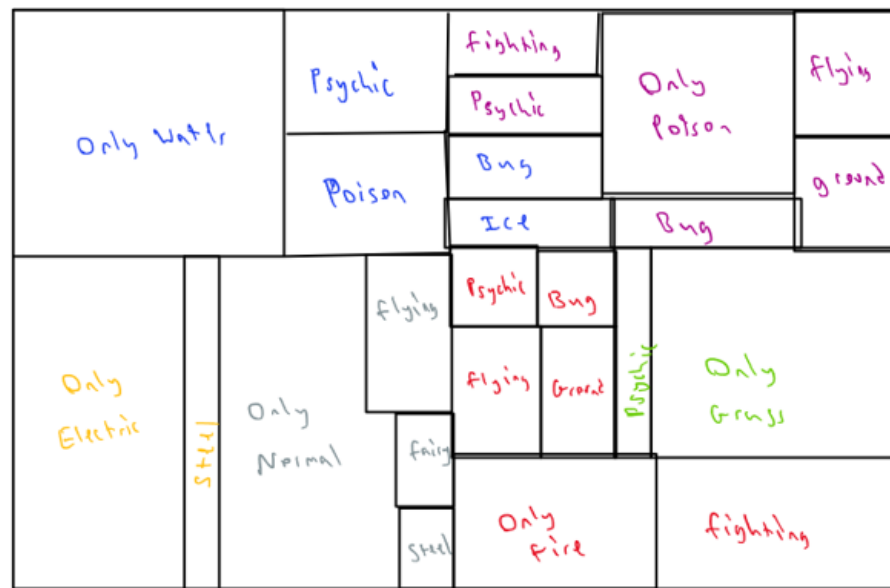


- Concept art -

6. Treemap - type representations

Although this would be a unique visualization to create, we ultimately cut this visualization because we found that it didn't have value to the end user.

- Source for inspiration:
<https://blog.revolutionanalytics.com/2016/07/an-analysis-of-pokemon-go-types-created-with-r.html>
- Helpful reference for how to make a treemap in d3:
<https://www.d3-graph-gallery.com/treemap.html>

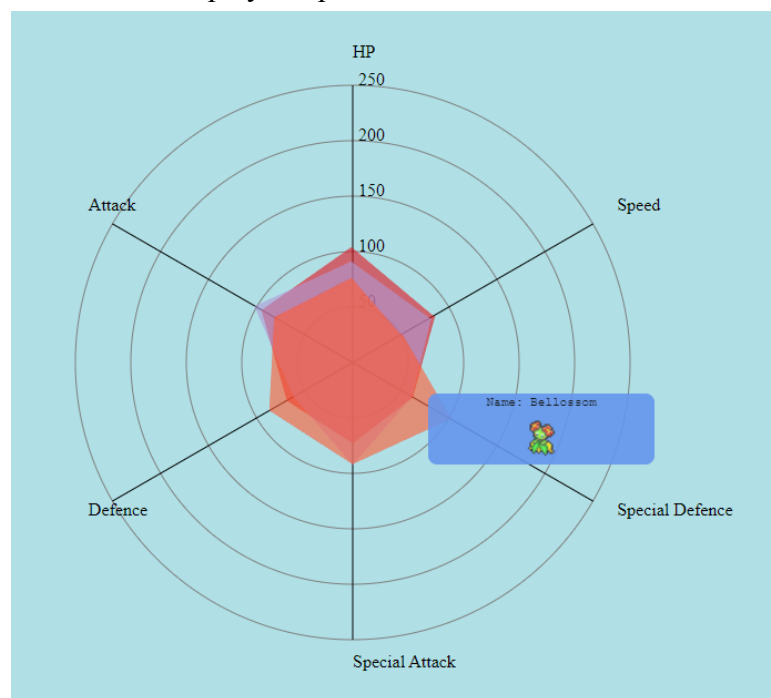


- Concept art -

Current Features: Implementation

The intent of our visualizations was to provide the user with an adequate amount of data from user-selected comparison options to be able to benefit in game from using our visualizations. The features we have consist of:




- A comparison of pokemon by user selected stats to compare the overall pokemon data set.
- A ranking of the strongest Pokemon in each stat category with type groupings
- Comparing multiple user-specified Pokemon to each other in an overlapped radar chart. Hovering over an area will display the pokemon's icon and name



- A self updating table detailing showing adjusted stats for typing matchups

Types	Attacking with.....	Type	Attack Stat	Sp. Attack Stat	When Defending	Effective Defence	Effective Sp. Defence
BUG	Normal	POISON	61	61	50% Damage Reduction	80	80
DARK	100% Bonus Damage	FIGHT	122	122	Normal	40	40
DRAGON	Normal	POISON	61	61	Normal	40	40
ELECTRIC	Normal	POISON	61	61	Normal	40	40
FIGHT	100% Bonus Damage	POISON	122	122	50% Damage Reduction	80	80
FIGHT	Normal	POISON	61	61	50% Damage Reduction	80	80
FIRE	Normal	POISON	61	61	Normal	40	40
FLYING	Normal	POISON	61	61	Normal	40	40
GHOST	50% Damage Reduction	POISON	30.5	30.5	Normal	40	40
GRASS	100% Bonus Damage	POISON	122	122	50% Damage Reduction	80	80
GROUND	Normal	FIGHT	61	61	100% Bonus Damage	20	20
ICE	100% Bonus Damage	FIGHT	122	122	Normal	40	40
NORMAL	100% Bonus Damage	FIGHT	122	122	Normal	40	40
POISON	50% Damage Reduction	POISON	30.5	30.5	50% Damage Reduction	80	80
PSYCHIC	Normal	POISON	61	61	100% Bonus Damage	20	20
ROCK	100% Bonus Damage	FIGHT	122	122	Normal	40	40
STEEL	100% Bonus Damage	FIGHT	122	122	Normal	40	40
WATER	Normal	POISON	61	61	Normal	40	40

- A table to display the base stats of selected pokemon

			
Name	Kricketot	Duskull	Croagunk
Index	401	355	453
Primary Type	BUG	GHOST	POISON
Secondary Type			FIGHT
HP	37	20	48
Attack	25	40	61
Defence	41	90	40
Special Attack	25	30	61
Special Defence	41	90	40
Speed	25	25	50
Total	194	295	300

Evaluation:

By providing interactive visualizations that allow users to interact with the data and the various Pokemon, we have found