Final Project – Pokémon Index

By Chen Peng (cp3rg)

Dataset

This dataset[1] is about the details of different Pokémon (Pokémon Game).

#	Name	Type 1	Type 2	Total	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation	Legendary
1	Bulbasaur	Grass	Poison	318	45	49	49	65	65	45	1	FALSE
2	Ivysaur	Grass	Poison	405	60	62	63	80	80	60	1	FALSE
3	Venusaur	Grass	Poison	525	80	82	83	100	100	80	1	FALSE
3	VenusaurMe	Grass	Poison	625	80	100	123	122	120	80	1	FALSE
4	Charmander	Fire		309	39	52	43	60	50	65	1	FALSE
5	Charmeleon	Fire		405	58	64	58	80	65	80	1	FALSE
6	Charizard	Fire	Flying	534	78	84	78	109	85	100	1	FALSE
6	CharizardMe	Fire	Dragon	634	78	130	111	130	85	100	1	FALSE
6	CharizardMe	Fire	Flying	634	78	104	78	159	115	100	1	FALSE
7	Squirtle	Water		314	44	48	65	50	64	43	1	FALSE
8	Wartortle	Water		405	59	63	80	65	80	58	1	FALSE
9	Blastoise	Water		530	79	83	100	85	105	78	1	FALSE
9	BlastoiseMe	Water		630	79	103	120	135	115	78	1	FALSE
10	Caterpie	Bug		195	45	30	35	20	20	45	1	FALSE

The features are described as bellows.

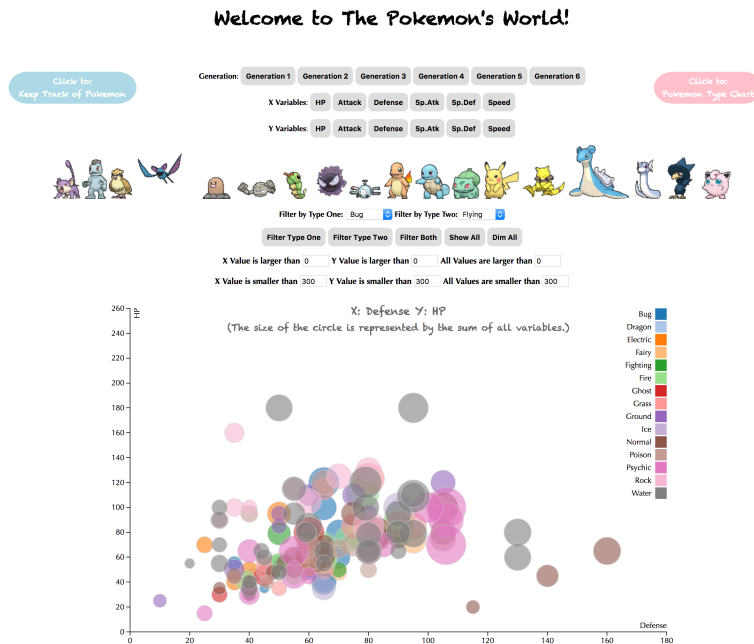| |
|---|
| #: ID for each pokemon |
| Name: Name of each pokemon |
| Type_1: Each pokemon has a type, this determines weakness/resistance to attacks |
| Type_2: Some pokemon are dual type and have 2 |
| Total: sum of all stats that come after this, a general guide to how strong a pokemon is |
| HP: hit points, or health, defines how much damage a pokemon can withstand before fainting |
| Attack: the base modifier for normal attacks (eg. Scratch, Punch) |
| Defense: the base damage resistance against normal attacks |
| SP_Atk: special attack, the base modifier for special attacks (e.g. fire blast, bubble beam) |
| SP_Def: the base damage resistance against special attacks |
| Speed: determines which pokemon attacks first each round |

There are 721 observations and 13 variables, but not all variables will be used. People used this dataset to discuss what features will most influence the pokemon, but my idea about the data visualization is to represent different skills for different pokemon, so people can clearly know about their pokemon to help them play the game. In other words, it is a Pokémon Pokédex, which is a mini-encyclopedia of Pokémon skills in different fields.

Data Visualization

The scatter plot was used to represented the relationship between each variable.

This is the initial page for my project.

---

1 HTTPS://WWW.KAGGLE.COM/ABCSDS/POKEMON

Generation: Generation 1 Generation 2 Generation 3 Generation 4 Generation 5 Generation 6

X Variables: HP Attack Defense Sp.Atk Sp.Def Speed

Y Variables: HP Attack Defense Sp.Atk Sp.Def Speed

Click to:
Keep Track of Pokemon

Click to:
Pokemon Type Chart

Filter by Type One: Bug Filter by Type Two: Flying

Filter Type One Filter Type Two Filter Both Show All Dim All

X Value is larger than 0   Y Value is larger than 0   All Values are larger than 0

X Value is smaller than 300   Y Value is smaller than 300   All Values are smaller than 300

X: Defense Y: HP
(The size of the circle is represented by the sum of all variables.)

HP

Defense

Bug
Dragon
Electric
Fairy
Fighting
Fire
Ghost
Grass
Ground
Ice
Normal
Poison
Psychic
Rock
Water

There are some buttons, and let me introduce them.

Generation: Generation 1 Generation 2 Generation 3 Generation 4 Generation 5 Generation 6

X Variables: HP Attack Defense Sp.Atk Sp.Def Speed

**Y Variables**: HP Attack Defense Sp.Atk Sp.Def Speed

Firstly, since there are six different generations, I added the button to reach different generations. People could click the button to get the different dataset from different generations. Secondly, people can use the X Variables button and Y variables button to change the variable of x-axis and y-axis, so people could explore the variables that they want.

As for the scatter plot, every circle represents one pokemon. The size of the circle was represented by the sum of all variables, and the color was represented by the type one. Thus, if the circle gets closer to the right, the value of X variable is larger, and if the circle gets closer to the top, the value of Y variable is larger.
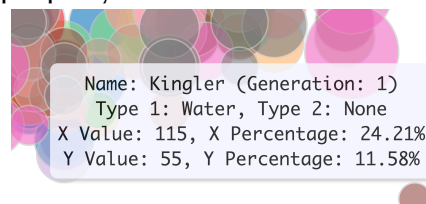
Filter by Type One: Bug Filter by Type Two: Flying

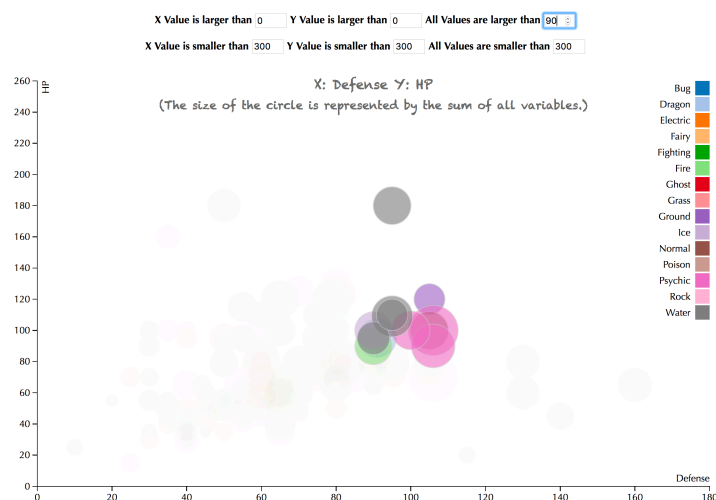Filter Type One Filter Type Two Filter Both Show All Dim All

Then, we can come to the types of pokemon. If your mouse is on the gif of different pokemon, you could know about their types. I just choose the classic pokemon to represent their types. As for the filter button, people could filter the data by type one, type two or the both. People could choose the kind they want and click the button "Filter Type One", "Filter Type Two" or "Filter Both", so they can get the conclusion in the scatter plot.

Let us look at an example.
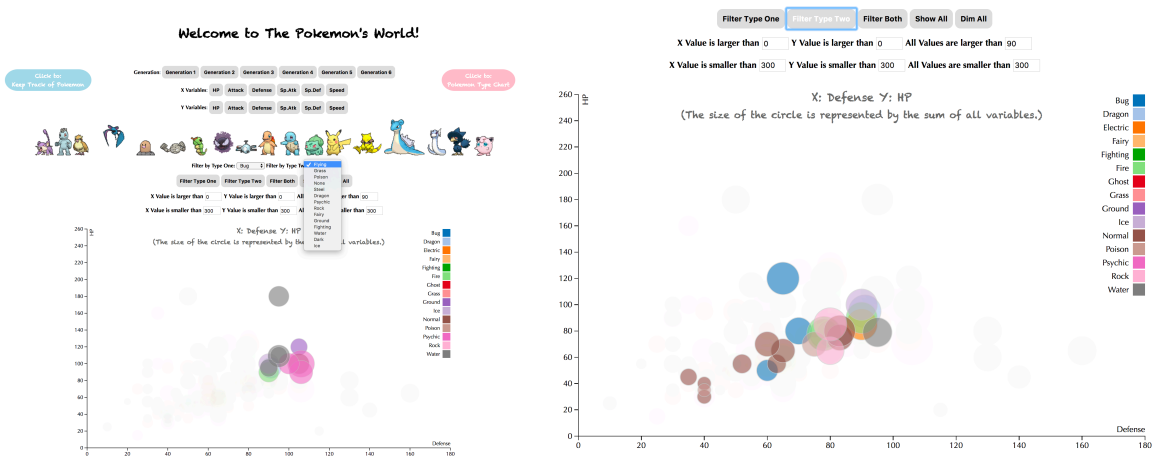
I also added tooltips to help people learn about the circle in the scatterplot, which contains the name, generation, type 1, type 2, the value and the percentage of the pokemon. The percentage is represented by the value divided by the total variable, which will help people to know how important this property is.

```
Name: Kingler (Generation: 1)
   Type 1: Water, Type 2: None
X Value: 115, X Percentage: 24.21%
Y Value: 55, Y Percentage: 11.58%
```

Furthermore, people could use the limit button to limit the range of the value.

There are six buttons for people to choose. They could find the pokemon whose x value is larger than 100, they could find the pokemon whose y value is smaller than 200, they also could find the pokemon whose both values are larger than 100, and they could input any numbers they want.

Finally, I also added the button to reach two websites. One is the type chart, and the other one is to help people use the name to locate the specific pokemon.
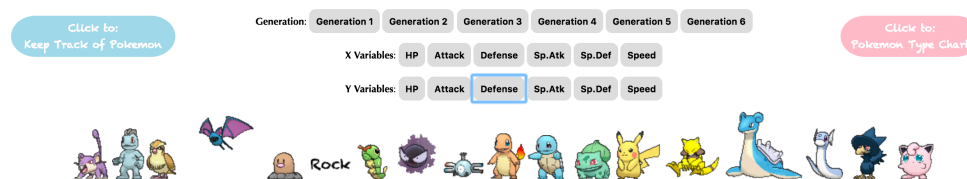

Discussion
We can explore the relationship between two variables, so people could use it to decide what pokemon they want to choose. On the one hand, they could use different variables to decide the pokemon. On the other hand, they also can choose pokemon by types.

For example, when people need to fight with someone else, and they will choose several pokemon they have. They will consider about different properties of the pokemon, such as HP, Attack, Defense, and types. Thus, they can use the index to find the most appropriate pokemon for themselves. If the person focuses on attack, but he also cares about the defense, so he could choose X and Y variables to attack and defense, and limit both values to 100. He could get the plot like this in generation 1.
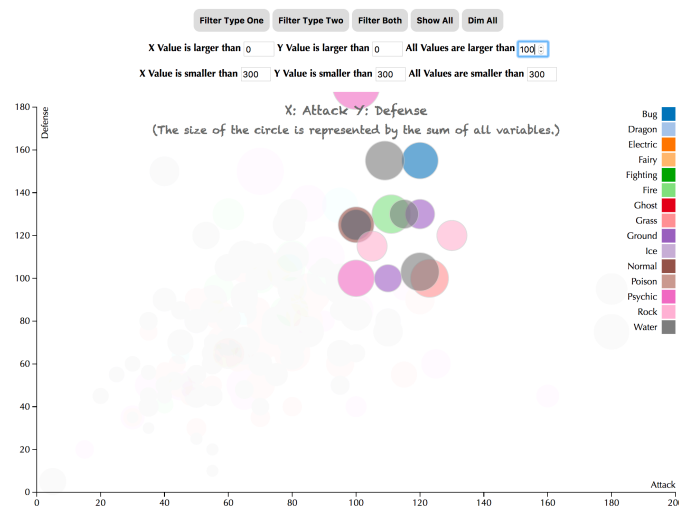
There are different types here, so he can click to the type chart and decide which types he would like to choose. Finally, he could choose several pokemon in this plot. In case, if he

can't remember what the pokemon is, he also could click to keep track of pokemon using their names.

I also find some patterns for this dataset.
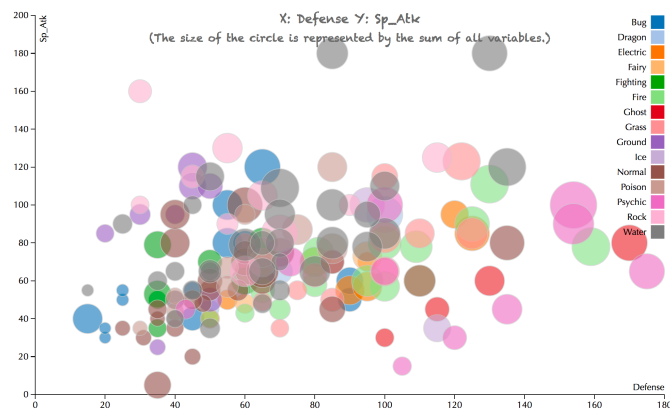
This is the generation one. As we can see, the pokemon whose type one is ghost have a high defense, so people could choose them if people focus on defense.

Besides, we couldn't only judge a type of pokemon by only one or two properties, since there is no absolute evidence to decide the types of pokemon.

However, it will be better if some parts of it could be improved. Firstly, people could only compare two variables here, which is a little inconvenient, so in the future, I will consider about making a plot to contain all the properties the pokemon have. Furthermore, if I can accomplish zooming in on the picture, it will be clearer for people to observe.

In conclusion, if people want to win a pokemon game, they firstly need to have enough pokemon, and then, they need to design a good strategy to let their pokemon play their strengths. Besides, the evolution of pokemon is also a good way to improve your pokemon.