

**Juweek Adolphe**  
**Special Problems in Design Assi. #2**

**Step 1: Pokemon dataset (all 7 generations)**

Pokemon is a media franchise owned by The Pokemon Company. It has spawned video games, trading cards, and tv shows. It is centered on fictional creatures called Pokemon that fight each other at the direction of humans known as Pokemon Trainers. Each game is focused on a character who goes around and collects pokemon, train them for battle, and fight other trainers.

Pokemon was released in generations. The first few games, back in 1995, held generation I. Shortly after that came generation II. This has gone on ever since, with the Pokemon Company releasing new games up to generation VII.

**Step 2 - Pose an initial question you want to answer.**

I play the pokemon games on occasion and whenever I bring it up, there is always an argument about which generation is the best. There are many ways to compare them, but I decided on a simple one: the avg base total of each generation.

The object of the game is to win battles, and the higher a pokemon's base stats are, the more likely they are to win. I decided to choose this question to ask because it seemed pretty interesting.

Question: Which generation of Pokemon, statistically, are the strongest?

**Step 3 - Assess the fitness of your data for answering your question; wrangle if needed.**

*Look at the data, starting with the raw data. Does the data seem appropriate to the question? If not, do you need to find different data or a new question? Does the data need to be cleaned or formatted first? Perform any steps necessary to get the data into shape prior to visual analysis. (You might do this in Excel, for example).*

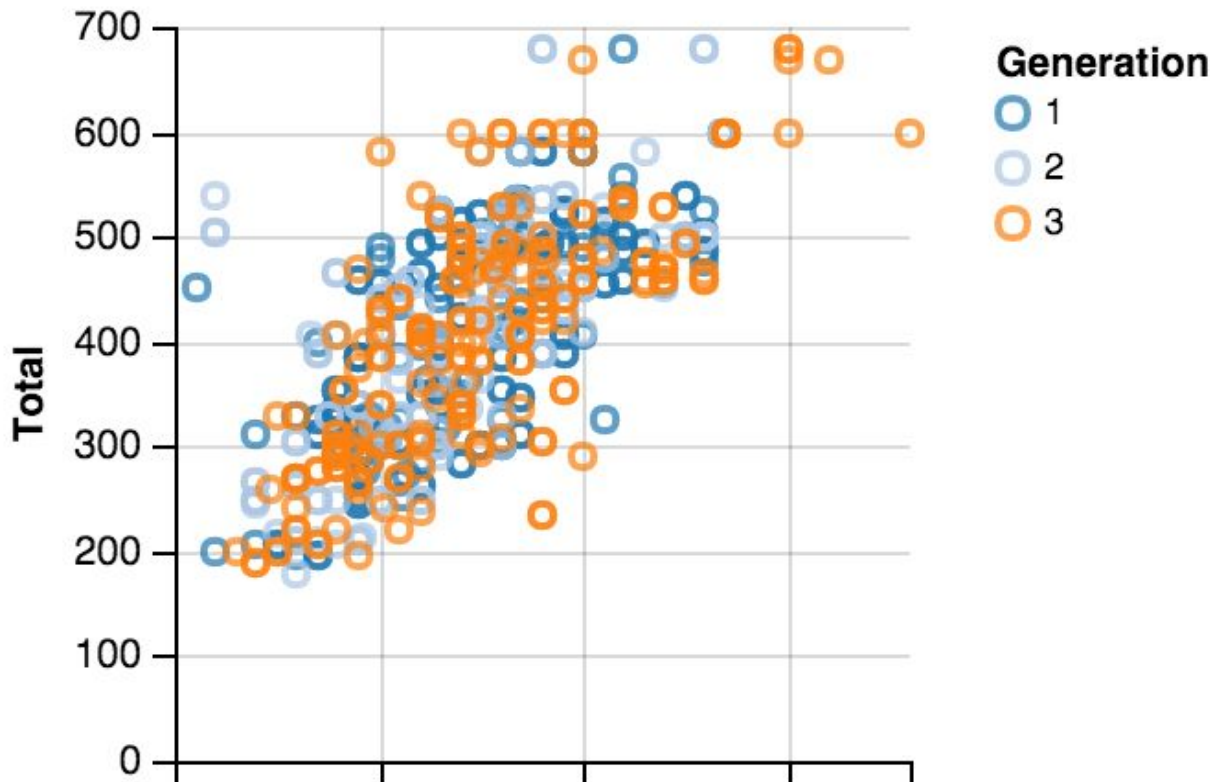
The data listed seemed mostly appropriate, but it had its problems. For one, Pokemon with multiple types were listed twice. When comparing stats of each generation, it was very likely that a pokemon and its stats could be posted multiple times, negatively

affecting the graphs. I removed the duplicates to prevent this problem. Most of the other information was fairly standardized, every stat was in integer format, and every stat had an approximate range of 40 - 60.

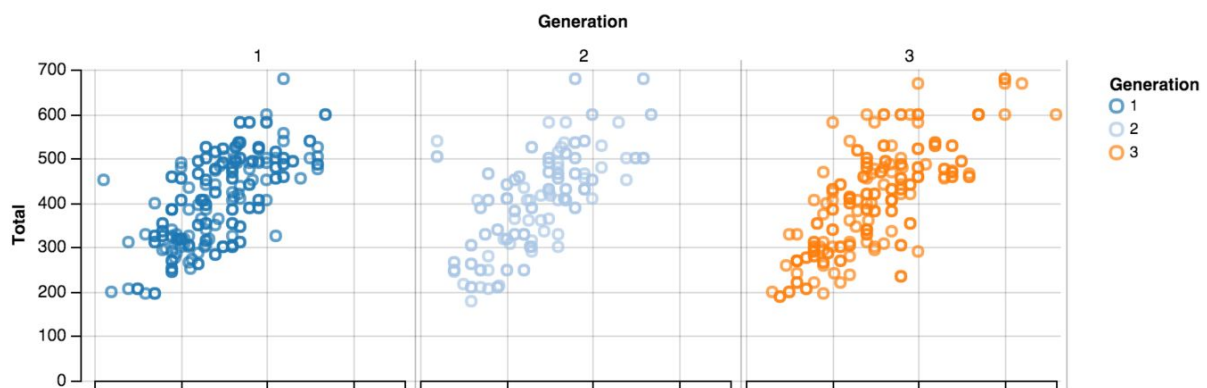
#### **Step 4 - Explore**

*Build a visualization that answers your initial question. As you build the visualization, you may find your question changes. It may become more specific, may seem like a wrong question. The visualization may suggest new questions. Track the history of your questions and how you answered them. Fully designed visualizations are not needed at this point. This is a notebook of your exploratory process.*

The first question I asked was comparing generations. I planned to plot every pokemon's total base stats in each generation on a graph and see if there was a distribution of points that led to any conclusions. However, that turned out to be inconclusive, simply because there was just too much data. At first, I started with all 7 generations, but making out any plot point was impossible. Even when I dwindled it down to 3 generations, it still felt cluttered. This is how it looked when each graph was shown on the graph.



Even when the graphs were separated into their own visualizations, it still didn't do much good. It was hard to compare them with one another, and although differences could be made out, it was almost impossible to tell what they were.



Thus, the question I had posed initially proved to not be the right one; I had to restructure it. However, I still wanted to use the Pokemon dataset because it was

formatted fairly nicely, and the data was organized in a clean fashion that allowed manipulation if I wanted to.

Instead of a quantitative nature, I looked at the data from a nominal point of view. As mentioned earlier, each pokemon has one or more types associated with it. There were about 16 in total, and each of them had their advantages and disadvantages (for instance, fighting types usually have high attack stats). However, I wondered if it were possible to compare them on an objective standard. Thus, I rephrased my question:

Step 2: Is a pokemon type generally stronger than another?

Step 3: To find this out, I started at first with the original dataset. I took each pokemon and I charted their type with their base total stat. This turned out to be generally ok, but Pokemon is a fairly balanced game; throughout the generations, there were a total of 700+ pokemon. There was a wide range of pokemon and stats, and a specific type of data was inconclusive from the graph.

To fix this problem, I narrowed the dataset to just the first game. This included only around 150 pokemon. I listed the dataset as intended, keeping in the duplicates that were removed from the first generation. This data set proved to be a bit more varied in terms of stats, and it was still ample enough for me to draw significant conclusions.

## **Step 5 - Final Visualization**

*After answering all the questions you have to your satisfaction, think of a way to word the final or most important answer(s)—your findings. Design a way to present the data and your findings as clearly as possible. Use visualization software tools (and an illustration program if needed) to produce a final image. Write a caption for the final graphic, as well as a paragraph describing the graphic and your findings.*

After looking at the data, one finds out a lot about the pokemon, and the types, of Generation 1. It is shown clearly that psychic types seem to be the overall strongest, for their ending range is the highest out of all the pokemon. Ice types seem to have the least standard deviation, for its line is quite short; this means that all ice types are pretty comparable to each other and closer in strength to one another than another more varied type (like bug or flying). This chart shows that the strongest pokemon in the game is a psychic type pokemon, while the absolute weakest pokemon are both poison and bug.

It is fairly easy to compare types due to length and position, and it is easy to look up a type due to color. This graph, therefore, does a great job of making perceiving the information fairly simple and easy.