Tony Schmitt

Professor Federico

CPSC 4030

9/29/2021

Dataset for Milestone 1

As a child I used to play a lot of the Pokémon games and I thought it would be an interesting study to see if there is a bias amongst them and how it might affect the balance of the game. In doing so in this research, I have re-kindled an interest in this subject and gotten to delve quite deep into the balancing and mechanics involved in making a game like this work. This quest led me to the online “Smogon” tier system. Smogon's tier system is used to rank Pokémon into several groups based on their perceived power and usage in competitive play. These tiers dictate which Pokémon can be used in different situations in competitive play. Each situation, or metagame as they refer to it, encompasses different Pokémon, and therefore each one is unique in its style of play. The standards set by Smogon's tier system seek to balance competitive battling, ensuring no Pokémon is "too powerful" or over-centralizes the metagame it appears in. This allows players the flexibility to show "flair" in their team design rather than being restricted to using a particular Pokémon to win.

The big question that arose from this to me came in the form of the tiers. I would like to know the decisions used for certain Pokémon to be sorted as “superior” to others even though the games should in theory be balanced to allow all Pokémon in each game to be usable and effective at winning against certain others.

<https://data.world/steveinatx/pokemon-index>

Table

Description automatically generated

Figure 1: Pokémon complete dataset sample

The data included in this dataset includes many specific information about all of the Pokémon that can be used to help answer the issue of understanding superiority and value. The list contains all 718 Pokémon and the columns contain the Name of each, their Type (1 of 18 elements that exist within the game), the number of them that exist in the world, how much health they have, as well as their attack and defense stats (with special attacks included). There are 10 columns used to present this information.