|  |  |  |
| --- | --- | --- |
| Checkpoint I | Checkpoint I: Project Proposal | |
| Group: | G11 |
| Date: | 2022/10/03 |
|  |  |

# Domain

Pokémon is a series of video games in which you can build teams of animal-like creatures (Pokémon) and pit them against other people’s teams in turn-based battles. Pokémon players can battle their teams against each other, online, competitively.

Each team can have up to six Pokémon.

Each Pokémon has the following main attributes:

* A set of Base Stats (Health Points, Attack, Defense, Special Attack, Special Defense, Speed);
* An ability, which may give them an advantage in battle (abilities may be split into groups according to function);
* An item which they may carry and which may have an in-battle effect;
* A nature, which influences their reaction to certain types of items;
* A set of up to four moves (associated with a type and damage class – special, physical, or status –, and which have the attributes “accuracy” and “power”);
* A generation of Pokémon games that they were first created in.

In this project, we will be analyzing common patterns in Pokémon attributes – both in the games, and in Pokémon competitive playing.

# Dataset

## Dataset description:

## There are 19 joint datasets available, all regarding information related to the competitive Pokémon video game, with particular reference to the period February - August 2022. Due to time constraints, we will initially be using 5 out of those 19 datasets as basis for our visualization (over 35 attributes total, over 70.000 items total), creating derived datasets as necessary.

## The data selected includes information about Pokémon (e.g.: df\_pokemon.csv), Pokémon prevalence in teams (e.g.: bridge\_pokemon\_pokemon\_USED\_IN\_TEAMS\_WITH.csv), Pokémon attributes (e.g.: df\_moves.csv, bridge\_pokemon\_moves\_MAY\_LERN.csv), and attributes’ prevalence (e.g.: bridge\_pokemon\_move\_USED\_WITH\_MOVE.csv).

## The data used is available at <https://www.kaggle.com/datasets/giorgiocarbone/complete-competitive-pokmon-datasets-may-2022>. We will build 5+ database tables from the selected datasets and use them to interpret the information as necessary.

# Example Questions

Competitive Pokémon battling generally occurs between two teams of six Pokémon, and there is no disadvantage to having a full team of six, so competitive teams with less than six members are rare, and, for this analysis, it is considered that a Pokémon in a team always has at least one teammate.

Pokémon types work in a rock-paper-scissors mechanic (e.g.: the Fire-Type is strong against the Grass-Type; the Grass-Type is strong against the Water-Type; the Water-Type is strong against the Fire-Type). Pokémon in a team that are of different Types may help cover for each other’s weaknesses in battle.

Pokémon attributes influence the result of a battle. For example:

* A move will be more powerful if its type corresponds to that of the Pokémon using it, and less powerful if it is a move the opponent is strong against;
* A move’s power may be influenced by the Speed Stat of the Pokémon it is used by/against, or by an item in use;
* An attack move of the Special damage class will do more damage if a Pokémon has a higher Special Attack
* A Pokémon with a higher Speed Stat will usually act first in a turn;
* Moves have power (how hard they hit), and Power Points (PP) – using a move reduces its PP, and when PP runs out, you cannot use that move for the rest of the battle.

A few relevant questions (all regarding competitive play):

* For two Pokémon who are teammates, which are the most often used Pokémon type combinations between each teammate?
* How has the HP Stat of Poison-Type Pokémon evolved throughout generations?
* How does a Pokémon’s Type (and number of Types) influence its Stats, and the Stats prioritized by competitive players?
* How do a Pokémon’s Stats influence the damage class of moves (Special, Physical, Status) chosen by the players?
* Do competitive players prioritize move availability (PP) or move power, when choosing moves for their Pokémon?

# Data Sample

Ex:

(from “df\_pokemon.csv”)

id; name;[...];generation;[...];Type1;Type2;[...]; HP; [...]; Usage Percentage; [...]

890; Eternatus; [...]; 8; [...]; Poison; Dragon; [...]; 140; [...]; 20; [...]

(from “bridge\_pokemon\_pokemon\_USED\_IN\_TEAMS\_WITH.csv”)

Use\_Percentage(%); Pokemon; Teammate

25.902%; Pikachu; Incineroar

(from “bridge\_pokemon\_move\_USED\_WITH\_MOVE.csv”)

Move; Use\_Percentage(%); Pokemon; Name

Protect; 98.672%; Incineroar; Protect

(from “df\_moves.csv”)

Name; [...]; Power; [...]; PP; Damage\_class

Double Slap; [...]; 15; [...]; 10 Physical

(from “bridge\_pokemon\_pokemon\_USED\_IN\_TEAMS\_WITH.csv”)

Move; Use\_Percentage(%); Pokemon; Name

Protect; 98.672%; Zacian Crowned Sword; Protect