

This data set includes 721 Pokemon, including their number, name, first and second type, and basic stats: HP, Attack, Defense, Special Attack, Special Defense, and Speed, etc. These are the raw attributes that are used for calculating how much damage an attack will do in the games. The dataset contains:

#: 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

Please address the following questions using proper statistical methods

- (1) Which variable (or variables) can define the type of a pokemon, if any? Also interpret your final model.
- (2) predict whether a pokemon is legendary or not.

For each question, please justify your method (e.g, why you prefer model A over model B).

Format: Report should be submitted in a pdf file by December 1 at 6pm to blackboard. The same grade will be given to all group members. When submitting the report, I recommend you to upload the file only through one student's account. This makes me easier to know which is the right version for grading. (Last year, some students in the same group submitted different versions.) Please make sure the names of the all group members and the NetID are printed clearly on the first page of the project.

Report that is late receives 20% off within 24 hours late and 0 after that. No more than 8 pages in normal format with font size 12. All the analysis should be done in R and all programs used for data analysis should be turned in together with the report (programs not counted into the 8 page).