MATH158 Project Part 1

Lathan Liou and Alex Gui

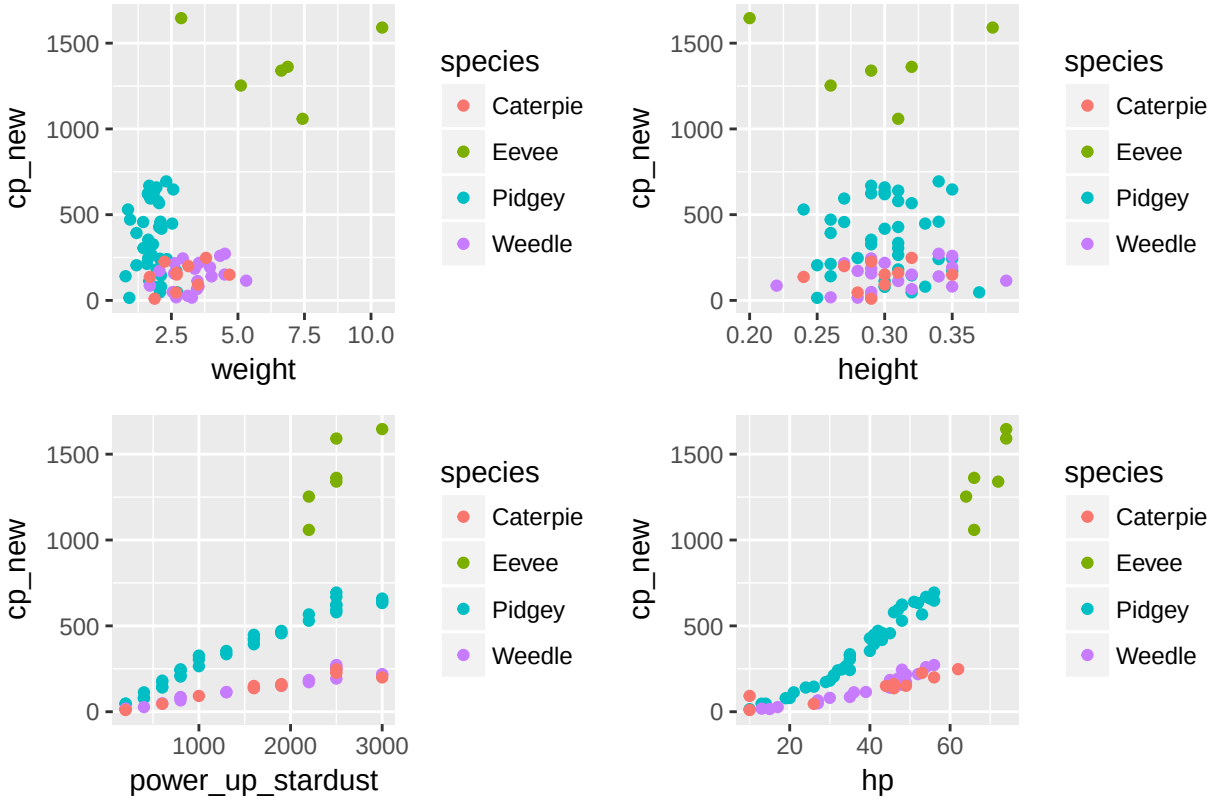
Data Description

This data comes from OpenIntro, and the dataset contains 26 variables of interest, as listed below. The observational units are different Pokemon.

1. name: A unique name given to the Pokémon
2. species: The Pokémon's type, e.g. Pidgey.
3. cp: Pre-evolution Combat Power, which is a summary of the Pokémon's strength for battling prior to the evolution of the Pokémon.
4. hp: Pre-evolution Hit Points, which is a summary of how difficult it is to weaken the Pokémon in a battle.
5. weight: Pre-evolution weight, in kilograms.
6. height: Pre-evolution height, in meters.
7. power_up_stardust: Pre-evolution stardust required to power up the Pokémon.
8. power_up_candy: Pre-evolution candy required to power up the Pokémon.
9. attack_weak: The name of the pre-evolution weaker attack of the Pokémon.
10. attack_weak_type: The type of the pre-evolution weaker attack.
11. attack_weak_value: The damage done by the pre-evolution weaker attack.
12. attack_strong: The name of the pre-evolution stronger attack.
13. attack_strong_type: The type of the pre-evolution stronger attack.
14. attack_strong_value: The damage done by the pre-evolution stronger attack.
15. cp_new: Post-evolution Combat Power.
16. hp_new: Post-evolution Hit Points.
17. weight_new: Post-evolution weight, in kilograms.
18. height_new: Post-evolution height, in meters.
19. power_up_stardust_new: Post-evolution stardust required to power up the Pokémon.
20. power_up_candy_new: Post-evolution candy required to power up the Pokémon.
21. attack_weak_new: The name of the post-evolution weaker attack.
22. attack_weak_type_new: The type of the post-evolution weaker attack.
23. attack_weak_value_new: The damage done by the post-evolution weaker attack.
24. attack_strong_new: The name of the post-evolution stronger attack.
25. attack_strong_type_new: The type of the post-evolution stronger attack.
26. attack_strong_value_new: The damage done by the post-evolution stronger attack.

Problem 2

Figure 1: Exploratory Plots



Note: Skimr output is attached in separate paper. Had problems knitting skimr output.

In this project, we will mainly focus on: how can we predict a pokemon's combat power after evolution based on their previous combat features?

We first have a quick look at the old combat power in our data set: it has a median of 169, min of 10 and max of 619. The distribution is skewed to the right. The new combat power after evolution roughly doubles and the distribution is more skewed.

As for the continuous variables, weight distribution is also skewed to the right whereas height is more homogenously distributed. As for the categorical variables, attach strong and attack weak are not evenly distributed. It seems like there are more pokemon in some certain attack strong values. Interestingly, power up stardust does not change before and after evolution.The distribution of cp_new is skewed right. Other distributions of note are those of height and hp, which are fairly symmetrical. Some relationships that we decided to highlight were weight vs. cp, height vs. cp, power_up_stardust vs. cp, and hp vs. cp, with the data points colored by species.

Sampling is from a random process, based on our understanding of the game mechanics, in which a randomly generated pokemon with randomly generated stats will appear, available for capture. In other words, if everyone in our class captured 100 Pidgeys, we would all have similar distributions of Pidgeys's stats. This data could be representative of all Pidgey, Caterpie, Eevee, and Weedle in the mobile game, Pokemon Go.