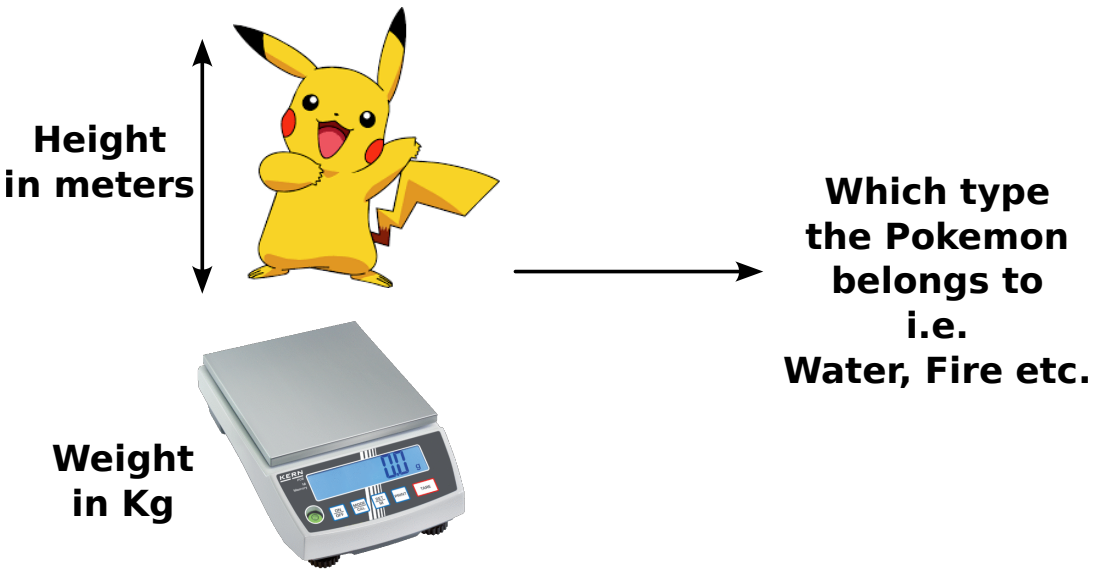
TD1

Introduction to Data Science en Python - ENSISA CPB2

Ali El Hadi ISMAIL FAWAZ

November 14, 2024

Exercice I: Pokemon Types



Our goal in this exercise is to try and understand any relationship between the information of a Pokemon and its main type. These information are the height, weight, attack value and defense value of each Pokemon.

Question 1: Reading the data

Read the 'pokemon.csv' file provided to you using the following command in *pandas* library:

```
import pandas as pd

df = pd.read_csv('pokemon.csv')
```

Question 2: Statistical Information

Extract the mean value of the height over all pokemons, as well as the min value, the max value, and the standard deviation. Use numpy for this question

```
import numpy as np
```

Question 3: BMI

The Body Mass Index (BMI) of each Pokemon is calculated as:

$$BMI = \frac{\text{weight in kg}}{\text{height in m}^2}$$

Given that the two columns, height and weight are already given in m and kg respectively, add the new column "bmi" that contains the BMI for each Pokemon.

```
df["bmi"] = ... # your code here
```

Question 4: Indexing

How many Pokemon are shorter than 0.9 meters ?

Question 5: Indexing

Print the rows where the Pokemon is of type "Water".

Question 6: Statistics on Indexed Rows

Get the mean and standard deviation of the BMI column only for the Water Pokemon.

Question 7: Intepreting Statistics on Types

Do the same as in question 6 for Fire and Grass Pokemon. Which type between Water, Fire and Grass seems to have a more stable BMI in your opinion ?

Question 8: Plotting Histograms

Plot a histogram of the attack values of all Pokemon. Use *matplotlib*, read docs on this package to know more, here is how to import it and plt a histogram [1]:

```
import matplotlib.pyplot as plt

plt.hist(..., bins=30)
```

Change the value of bins and observe what happens.

Question 9: Plotting Pie Charts

Plot a pie chart of the Pokemon types counts, label the parts of the pie chart with the pokemon type name.
Hint: use `plt.pie()` and `np.unique()` to do this.

Question 10: Box Plots

Use a box plot to show the median's placement for the *defense* column of Water, Fire and Dark Pokemon. Index the x-axis with the Pokemon type.
Hint: Use `plt.boxplot`

Question 11: Dispersion Plot

For this question, you will have to install Seaborn [2] using `pip install seaborn`. Import this package as follows:

```
import seaborn as sns
```

Plot a heatmap using Seaborn of the correlation matrix between the height, weight, attack, defense values of all Pokemon. What can you conclude ? Change the color map of the heat map to "jet" and set the min value to 0 and max value to 1. Read documentation to know how to do this.

[1] https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.hist.html
[2] https://seaborn.pydata.org/