**Level 1 (Regular Data Science Questions):**

1. Calculate the average Speed across all Pokémon.

68.73

1. Find the median Base Special Attack for Pokémon in Evolution Stage 1.

50.0

1. What is the range (difference between maximum and minimum) of the Base HP across all Pokémon?

240

**Level 2 (Multiple Step Hard Data Science Questions):**

1. For each 'Hidden Ability', calculate the average Base Defense. Which Hidden Ability corresponds to the highest average Base Defense?

Overcoat

1. Group Pokémon by 'Evolution Stage' and calculate the median Length for each group. Which Evolution Stage has the longest median Length?

Evolution Stage 3

1. Identify the Pokémon with the highest Base Attack to Speed ratio. What is the name and ratio of this Pokémon?

Slowpoke with a ratio of 4.33

1. Calculate the variance of 'Speed' for Pokémon in each 'Hidden Ability' category. Which category has the highest variance?

Unnerve

1. For Pokémon with Base HP greater than 70, what is the correlation coefficient between 'Base Special Defense' and 'Length'?

0.329

1. Divide Pokémon into two groups based on median 'Base Attack' (above or below). Compare the average 'Base Special Attack' between these two groups. Which group has a higher average?

Pokemon above the median Base Attack have a higher average Base Special Attack

**Level 3 (Multistep Data Analysis Aspects of data science/ML):**

1. Using a clustering algorithm (e.g., K-Means), cluster the Pokémon based on Base HP, Base Attack, and Base Special Attack. How many clusters provide the most meaningful grouping?

7 clusters

1. Predict the 'Evolution Stage' of a Pokémon using a decision tree classifier based on all base stats. What is the accuracy of this model?

Accuracy: 68.89%

1. Identify outliers in the dataset with respect to Base HP, Base Attack, and Base Defense using an Isolation Forest algorithm. How many outliers are detected?

23 outliers in the dataset with respect to Base HP, Base Attack, and Base Defense

1. Implement Principal Component Analysis (PCA) on the base stats (HP, Attack, Defense, Special Attack, Special Defense, Speed) and determine how many components explain 90% of the variance.

5 Components

1. Using the features 'Base HP', 'Base Attack', 'Base Special Attack', 'Speed', train a Support Vector Machine (SVM) classifier to predict whether a Pokémon's 'Length' is above or below the median. What is the precision of this model?

0.9

1. Apply a time series forecasting model (e.g., ARIMA) to predict the next 5 values for 'Speed' if it were a time series. What are these 5 predicted values?

Next predicted values: [85.99, 74.08, 70.86, 69.99, 69.76]