Analysis Arena #001

Participant name: T S Akhilesh

Problem statement: Explore the Pokémon dataset to uncover interesting insights and solve the following optimization problem

Dataset used: The dataset used is the Pokémon dataset, which includes various attributes for each Pokémon, such as their name, types, base stats, abilities, and more.

Data Exploration: .head(), .shape, .info(), and .isnull().sum() methods are used to get an overview of the dataset, its structure, and missing values.

Handling Missing Data

- Numerical columns ('height_m', 'weight_kg', 'percentage_male') are filled with mean values.
- Categorical column ('type2') is filled with 'None'.

Exploratory Data Analysis (EDA)

Various EDA tasks are performed:

- Counting unique Pokémon names.
- Distribution of legendary vs. normal Pokémon.
- Pokémon count by generation and type.
- Top Pokémon classifications and abilities.

Optimization Process

- Calculates and ranks Pokémon by total base stats and number of abilities.
- Visualizes type coverage of Pokémon.

Synergy and Team Optimization

- Defines functions to calculate synergy scores among Pokémon teams and optimize a team for battle effectiveness.

Predictive Modeling

- Prepares data for machine learning:
    - Encodes categorical features using LabelEncoder.
    - Splits data into training and testing sets.
- Trains a RandomForestClassifier model and evaluates its performance using metrics like accuracy, precision, recall, and F1-score.