

**IDS 400 - PROGRAMMING FOR DATA SCIENCE IN
BUSINESS - FALL 2025**

LAB 1

dplyr Practice

1. Calculate the average Attack, Defense, and Speed for each primary type (Type 1). Which primary type has the highest average Speed?
2. How many missing values are there in each column? Which column has the most missing values?
3. Which non-Legendary Water-type Pokémon (Type 1 = "Water") have Attack > 90? List their Name, Type 1, Attack, and Generation?
4. Find which primary type (Type 1) appears most often in the dataset. How many times does it appear?
5. Which five Pokémon have the highest Speed? List their Name, Type 1, Type 2, Speed, and Legendary status.
6. For each (Type 1, Generation) pair, how many Pokémon are there? Within each generation, which pairs have the largest counts?
7. For each Type 1, which Pokémon have Sp. Def strictly above their type's average Sp. Def, and what are their Name, Type 1, and Sp. Def?
8. Compare Legendary and non-Legendary average Totals by Generation. Which Generation has the widest gap?
9. In each Generation, what number of Pokémon have a non-missing Type 2? Which generation has the highest share of dual-type Pokémon?
10. For each Type 1, return the Pokémon with the greatest Attack (include all ties), with columns Type 1, Name, Attack.
11. Which Type 1 has the biggest share of Legendary Pokémon?
12. Which Generation has the highest median Speed?
13. Which Type 1 shows the largest gap between average Sp. Atk and average Attack?

Submission

The due date for this assignment is Sunday, September 14, 2025 at 11:59 PM. It can be submitted in groups of up to 5 individuals. If you're stuck, please feel free to refer to the lab recording for solutions.