# Pet Box Subscription

## Background

PetMind is a retailer of products for pets. They are based in the United States.

PetMind sells products that are a mix of luxury items and everyday items. Luxury items include toys. Everyday items include food. The company wants to increase sales by selling more everyday products repeatedly. They have been testing this approach for the last year. They now want a report on how repeat purchases impact sales.

### Data

The dataset contains the sales records in the stores last year.

Find the attached dataset in the mail.



## Tasks

***Submit your answers directly to my email(segunumoru1@gmail.com) with the GitHub workspace link including a 7 slides PowerPoint presentation, do a live video presentation, and tag me on LinkedIn.***

1. For every column in the data:

a. State whether the values match the description given in the table above.

b. State the number of missing values in the column.

c. Describe what you did to make values match the description if they did not match.

2. Create a visualization that shows how many products are repeat purchases. Use the visualization to:

a. State which category of the variable repeat purchases has the most observations b. Explain whether the observations are balanced across categories of the variable repeat purchases

3. Describe the distribution of all of the sales. Your answer must include a visualization that shows the distribution.

4. Describe the relationship between repeat purchases and sales. Your answer must include a visualization to demonstrate the relationship.

5. Feel free to do create additional insight and visuals in line with the problem objective.

**N.B** – *To be submitted on 11:59pm (GMT +1), September 30, 2023.*

All the best!

Segun Umoru