📝 Dataset Description_Wycliffe

The Sample Superstore dataset contains sales data from a fictional retail store. It includes customer orders, shipment details, profit/loss per sale, and regional segmentation. It's perfect for practicing:

- Sales & profit analysis
- Regional performance
- Category & sub-category trends
- Customer segmentation

Key Columns:

Order ID, Order Date, Ship Date, Customer Name, Segment, Region, City, State, Category, Sub-Category, Sales, Quantity, Discount, Profit

Analytic Questions to Explore in R using ggplot2

Here are five insightful questions you can visualize and manipulate:

1. Which product categories and sub-categories drive the most sales and profit?
   - ➢ Bar chart: Sub-Category vs Sales & Profit
   - ✓ Useful ggplot2 function: geom_bar(stat="identity")

2. How do sales and profits vary across different regions and states?
   - ➢ Heatmap or choropleth (if using maps), or bar plot of Region vs Profit
   - ✓ Explore high vs low-performing areas

3. Is there a relationship between discount given and profit margin?
   - ➢ Scatter plot: Discount vs Profit
   - ✓ Add regression line with geom_smooth(method="lm")

4. How do sales and profits trend over time (monthly or quarterly)?
   - ➢ Line chart using Order Date (aggregated by month/year)
   - ✓ Analyse seasonality in sales

5. Who are the most valuable customers (top 10 by profit)?
   - ➢ Bar chart: Customer Name vs Total Profit
   - ✓ Helps identify key customers to target