Data Overview

1. Project Context

This analysis supports the *Chip Category Review* project, aimed at uncovering purchasing patterns, customer segmentation, and behavioral trends to guide strategic decisions in product placement, pricing, and promotions.

We work with two primary datasets:

- **Transactions Dataset:** Captures detailed chip purchase activity.

- **Purchase Behavior Dataset:** Enriches transaction records with customer demographics and lifestyle segmentation.

---

2. Dataset Summary

Dataset Name	Source	Rows	Columns	Description
transaction.xlsx.	Point-of-Sale System (POS)	264,000+	7	Contains records of customer purchases, including date, customer ID, product type, and total spend.
purchase_behavior.csv	Internal Customer Database	5,000+	5	Links each customer ID to demographic and behavioral segments such as lifestage and premium/budget classification.

Key Features

📑 **transaction Dataset:**

- DATE: Date of transaction

- CUSTOMER_ID: Unique identifier for customer

- PRODUCT_ID: Purchased chip product

- QUANTITY: Number of units purchased

- PRICE: Unit price

- TOTAL_SPEND: QUANTITY × PRICE

- STORE_LOCATION: Store where purchase occurred

👤 **purchase_behavior Dataset:**

- CUSTOMER_ID: Unique identifier (foreign key to transaction)

- AGE_GROUP: Customer's age range (e.g., 25–34)

- LIFESTAGE: Demographic segment (e.g., "Young Singles/Couples")

- INCOME_LEVEL: Customer's income classification (e.g., Low, Medium, High)

- PREMIUM_CUSTOMER: Boolean flag or category (e.g., Premium vs. Budget)