This project aims to support the Category Manager, Julia, in making data-driven decisions for an upcoming **category review** of chip products. The primary goal is to analyze transaction and customer data to understand purchasing **trends and behaviors**, especially across different **customer segments**.

The client is particularly interested in understanding:

- **Who buys chips** (e.g., families, young singles, seniors)

- **How they shop** (e.g., frequency, quantity, preferred brands)

- **What drives chip spending** (e.g., pack size, brand loyalty, product pricing)

To address this, we will analyze two datasets:

1. **QVI Transaction Data** – contains individual purchase records

2. **QVI Purchase Behavior Data** – includes customer attributes such as LIFESTAGE and PREMIUM_CUSTOMER

We are using Python as an optional tool for this analysis, due to its flexibility with large datasets and strong data visualization capabilities. While Excel could be used, it is less suitable for the volume and complexity of this project.

The analysis process involves:

- Data quality checks (e.g., missing values, data formatting)

- Feature engineering (e.g., extracting pack size and brand names)

- Creating metrics that capture chip purchasing behavior

- Segmenting customers to reveal actionable insights

Deliverables include a structured report summarizing findings and strategic recommendations, along with a PDF export of the working code (.Rmd file).