

# Project Title:

The Data Behind the Perfect Cup: Exploring Coffee Shop Trends and Customer Preferences

## Team Members:

1. Amy Salton
2. Christian Molina
3. Jade Russell
4. Lavenya Ammu Mohanasundaram
5. Saniya Kulkarni

## Project Description/Outline:

The goal of this project is to conduct exploratory data analysis (EDA) on coffee shop-related data to uncover trends in customer preferences, popular menu items, peak hours, and regional patterns in coffee consumption. Analyzing public datasets, we aim to identify key factors influencing customer behavior, product popularity, and pricing strategies within the coffee shop industry.

## Research Questions to Answer:

1. Based on sales and customer reviews, what are the most popular beverages or snacks?
2. When are the peak hours for coffee shop visits, and how do seasonal trends affect customer preferences?
3. How do product preferences vary across different regions or cities?
4. Does product pricing correlate with customer ratings or reviews?
5. Can we detect any anomalies in product sales, such as sudden spikes in specific items?

## Datasets to Be Used:

1. **Yelp Open Dataset:** Reviews and ratings from coffee shop customers across different regions.
2. **Starbucks Sales and Financial Data:** Sales trends, product performance, and seasonal variations.
3. **Coffee Shop Sales Data (Kaggle):** Detailed sales and customer information from local coffee shops.
4. **National Coffee Association (NCA) Data:** Reports on coffee consumption trends and consumer preferences.

## Rough Breakdown of Tasks:

1. **Data Collection:** Gather datasets from the sources mentioned above.
2. **Data Cleaning:** Prepare the datasets by cleaning missing or inconsistent data points.
3. **Exploratory Data Analysis (EDA):**
  - Analyze customer preferences based on product popularity, pricing, and ratings.
  - Investigate peak hours and seasonal spikes using time-of-purchase data.

- Identify regional patterns by comparing product preferences across different locations.

4. **Visualization:**

- Create bar plots for top-selling items and customer ratings.
- Develop heatmaps to visualize peak hours for coffee shop visits.
- Use word clouds to display popular drink names from reviews.