

Week 2 Workshop

Objective: Data Visualisation and Exploration I

Requirements:

- Attend Week 2 Lecture
- Complete all pre-workshop activities

Optional:

- Read the case study: Prasanna et al. (2019) Happy Cow Ice Cream: Data-Driven Sales Forecasting, *Harvard Business School Press*

In the pre-workshop activities, you are required to:

- Set up your R environment
 - Install R and RStudio. Please refer to the video instructions provided on Ed
- Data visualisation using line charts
 - Import data from a CSV file ("weekly_sales_staff_students.csv") into R. The dataset includes weekly sales of ice cream to two consumer groups (students and staff).
 - Explore the imported data (rows and columns).
 - Create simple time series plots of weekly sales for staff only to figure out how sales for staff change over time.
 - Create multiple line charts of weekly sales for students and staff to figure out how sales change over time for both staff and students and compare the two consumer groups
- Data visualisation using bar plots, pie charts and box plots
 - Import data from a CSV file ("weekly_sales_flavours.csv") into R. The dataset includes weekly sales of ice cream to two consumer groups (students and staff) for 16 flavours: Apricot, Banana, Chai Tea, Cherry Almond, Chocolate, Coffee, Ginger, Green Tea, Hazelnut, Lime Coconut, Mango, Mint Choco, Pistachio, Pure Coconut, Red Bean, and S Caramel.
 - Create a subset of the data for the staff category only.
 - Aggregate the data per flavour for staff – total sales per ice cream flavour for staff.
 - Create bar charts, pie charts and box plots to compare ice cream sales per flavour.

Case study

Mary Schroeder, the CEO of Hong Kong ice cream maker Happy Cow, was thinking about how best to increase her sales reach. During her three years at the company, Mary had already grown sales from low

levels when she bought the company. Founded by two vegans who wanted to produce healthy ice cream, Happy Cow ice cream was now found in supermarkets and health stores around the region. However, its customer profile had not changed significantly, and the company is still depending on expatriates and overseas-educated Chinese for most of its sales.

To expand further, Happy Cow must increase its market share among local Chinese people. To boost its brand and obtain valuable customer data, Mary opens a Happy Cow retail store at the University of Hong Kong campus, where 80% of customers are local. Mary was looking at the raw store sales data, seven months since the store opened. She believed that she would be able to use these data to decide on an optimal strategy. She gave you the responsibility to analyse the data and develop a plan.

Question 1

Problem: How to increase sales of Happy Cow Ice Cream?

Data: You are given the weekly_sales_staff_students.csv dataset from the Happy Cow store located at the University of Hong Kong (HKU). The dataset captures the weekly sales of ice cream over six months across two types of consumers (students and staff).

In the pre-workshop activities, you explored the weekly sales performance of the two consumer groups (students and staff), visualising time series. The graph should look like the following:

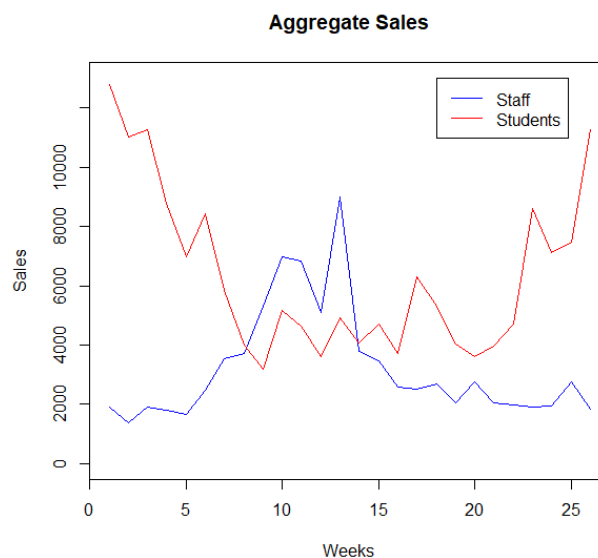


Figure 1. weekly sales performance of the two consumer groups (students and staff)

In your groups discuss the following:

- What does the time series tell you about sales? Sales are inverted proportional for the staff and students
- Based on the graph, what do you think are potential influences driving the observed sales patterns?
- After discussing the findings with your group, share your **insights** with the class.

B. During the teaching term, student would have higher sales since there is a constant flow of student, where as during non-teaching term, the staff would took a break from grading exams to have a snake, which results in the inverse proportional relationship of sales. (Also, the company employees different price strategy fro staff and students)

Question 2

Problem: Both Mary and Prem, the senior sales assistants, believe that different flavours sell better. Does the weekly data back this up?

Data: You are given the weekly_sales_flavours.csv dataset from the Happy Cow store located at the University of Hong Kong (HKU). The dataset captures the weekly sales of ice cream to two consumer groups (students and staff) for 16 flavours over six months.

In the pre-workshop activities, you explored the weekly sales to staff per flavour as follows:

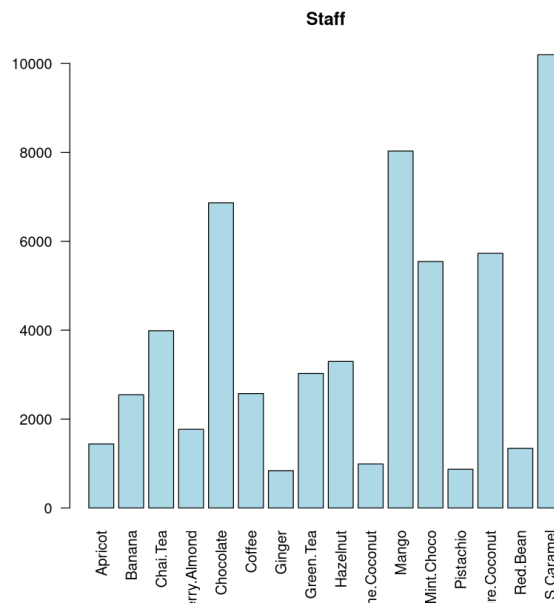


Figure 2. Weekly sales per flavour for staff

As you notice from Figure 2, there are too many ice cream flavours, many of which have similarities (e.g. fruity flavours, chocolaty flavours, etc.). One way to simplify the data and create a bar plot that is easier to interpret, is to group ice cream flavours into categories, e.g., as follows:

- fruit: Apricot, Banana, Cherry Almond (Cherry.Almond), Ginger, Lime Coconut (Lime.Coconut), Mango, Pure Coconut (Pure.Coconut), Red Bean (Red.Bean)
- caramel: Salted Caramel (S.Caramel)
- chocolate: Chocolate, Mint Chocolate (Mint.Choco)
- tea/coffee: Chai Tea (Chai.Tea), Green Tea (Green.Tea), Coffee
- nut: Hazelnut, Pistachio

You are requested to compare sales to staff and students across flavour categories:

- Visualise the total sales per flavour category for staff and students to directly compare sales for staff and students.
- After discussing the findings with your group, share your **insights** with the class
- What flavours sell better for students and staff?
- What would you **recommend** Mary to do to increase ice cream sales based on the insights derived from the findings.