

Project 00 (Problem Set 4) Handout

Project Overview

In this project, you will apply the D³M process to analyze supermarket sales data and develop actionable insights. Your goal is to choose a business question, analyze the dataset, create compelling visualizations, and present your findings in a structured format.

Project Steps

1. Choose a Question

Select one of the questions below or propose your own (must be answerable with the available dataset).

Example Questions:

Sales & Revenue Analysis

- How do invoice values and branch contributions to revenue vary over time?

Customer Behavior

- What are the most popular product lines, and how frequently do customers purchase them?

Customer Satisfaction

- How do factors such as product line, time of purchase, payment method, and store location impact customer ratings?

Payment Methods

- Which channels of payment methods contribute most to sales, and how should payment methods be adjusted?

2. Develop a Hypothesis

What do you expect to find? Why?

Example: “We hypothesize that electronic accessories have the highest total sales, but health and beauty products receive the highest customer ratings.”

3. Explore & Prepare Data

- Load the `supermarket_sales.csv` dataset in R.
- Identify relevant variables that relate to your question.

- Clean and prepare the data for analysis (e.g., filter, group, aggregate).

4. Analyze the Data

What patterns and trends emerge?

- Use appropriate statistical measures (averages, totals, percentages) to support your findings.
- Perform comparisons across different factors (e.g., time, product line, location).

5. Visualize the Data

Tell a clear story with the visuals. Select the most effective visuals to communicate your insights (e.g., bar charts, line graphs, heat maps).

Follow the six guidelines for effective visualizations covered in class:

1. Identify your objective
2. Show the data
3. Reduce the clutter
 - Label axes and use appropriate scales
4. Integrate the graphics and text
 - Use active titles (e.g., “Weekend Sales Surpass Weekday Sales” instead of “Sales by Day”)
 - Use whitespace effectively
5. Avoid the spaghetti chart
6. Start with gray

6. Interpret Your Results

What do your findings reveal?

- Do the results support or contradict your hypothesis?
- What are the limitations of your analysis?

7. Make Recommendations

Based on your findings, what actionable steps would you suggest?

Example: “Since members purchase more health and beauty products on weekends, the store should offer weekend promotions for this category.”