Unit 8 – Data Analysis and Visualisation

1. Title

Data Analysis and Visualisation

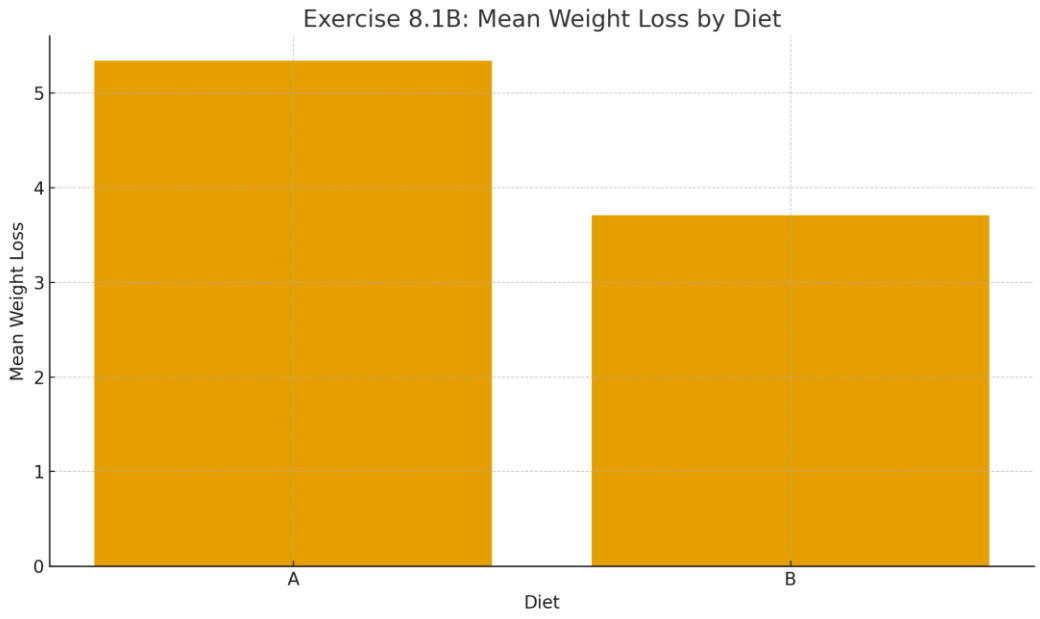
2. Summary of Activity

Completed the Charts worksheet using Excel datasets (Exe 8.1B, 8.2B, 8.3D, 8.4G, 8.6C). Produced appropriate charts (bar, line, scatter) and computed summary measures (mean, median, standard deviation) to interpret patterns and relationships in the data.

3. Results (Charts and Key Measures)

Exercise 8.1B – Diets (Bar Chart of Mean Weight Loss)

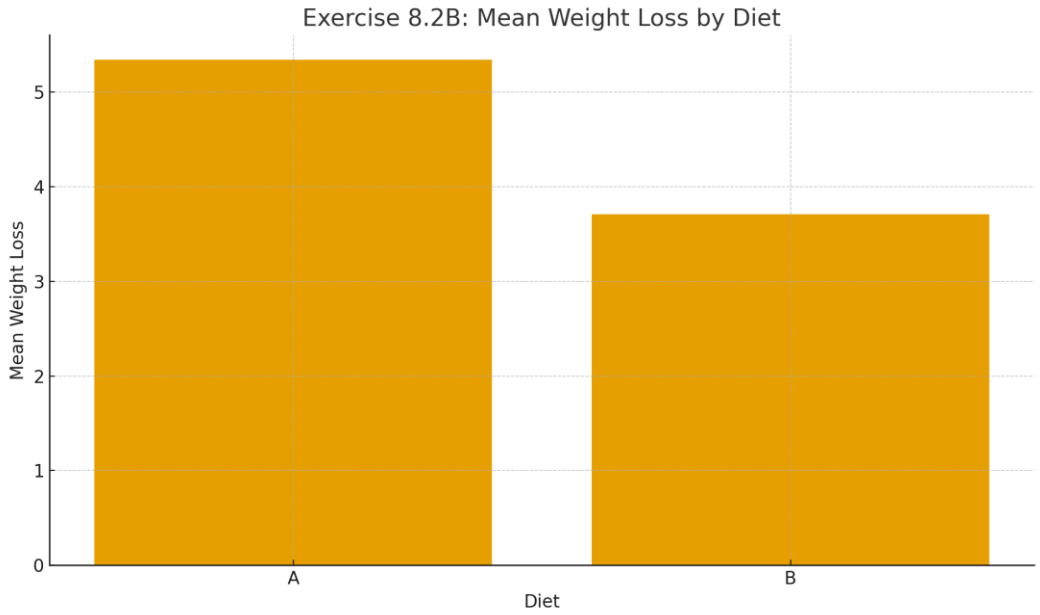
Mean (Diet A) = 5.341 ; Mean (Diet B) = 3.710



Interpretation: Diet A shows a higher mean weight loss than Diet B.

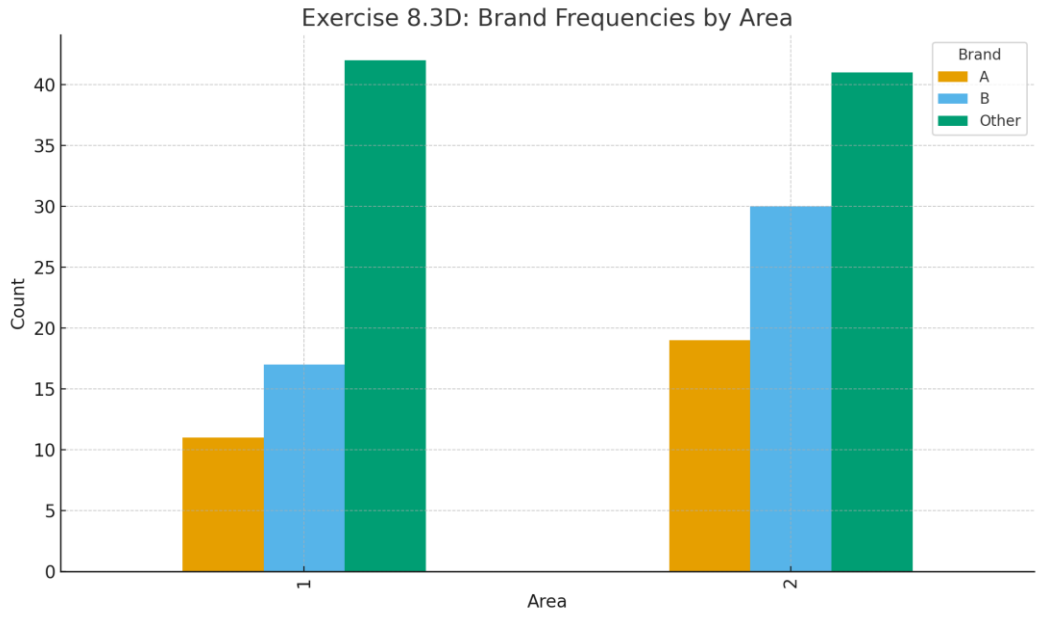
Exercise 8.2B – Diets (Replication Check)

Replicated means: A = 5.341 ; B = 3.710



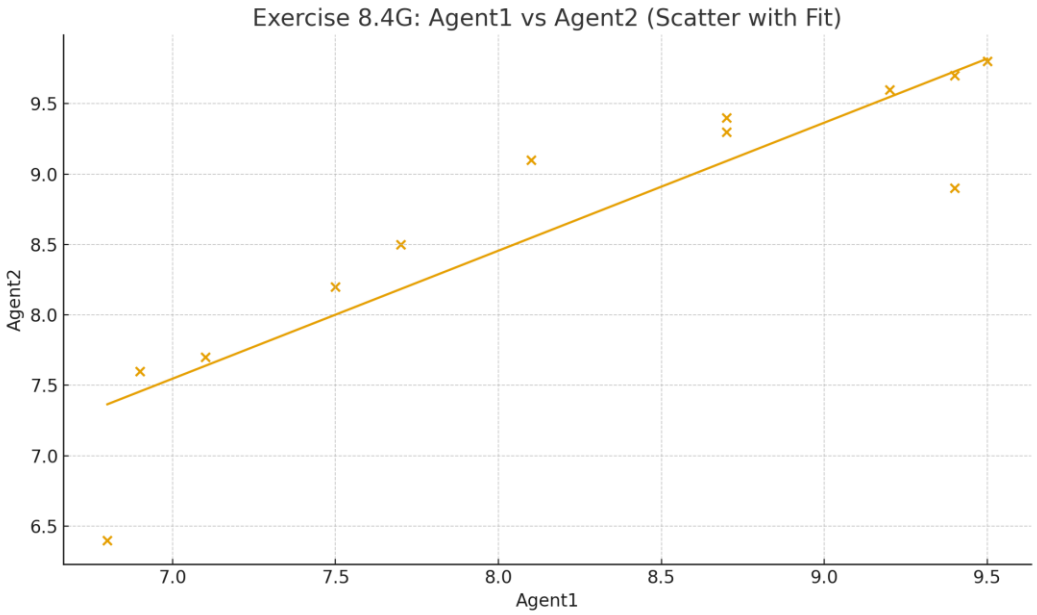
Interpretation: Replication confirms Diet A > Diet B on mean loss, consistent with Unit 7 inference.

Exercise 8.3D – Brand Frequencies by Area (Column Chart)



Interpretation: Area 2 shows higher counts for brands A and B than Area 1, while "Other" is similar across areas. This suggests brand preference differences by area.

Exercise 8.4G – Agent Scores by Batch (Line Chart + Scatter)

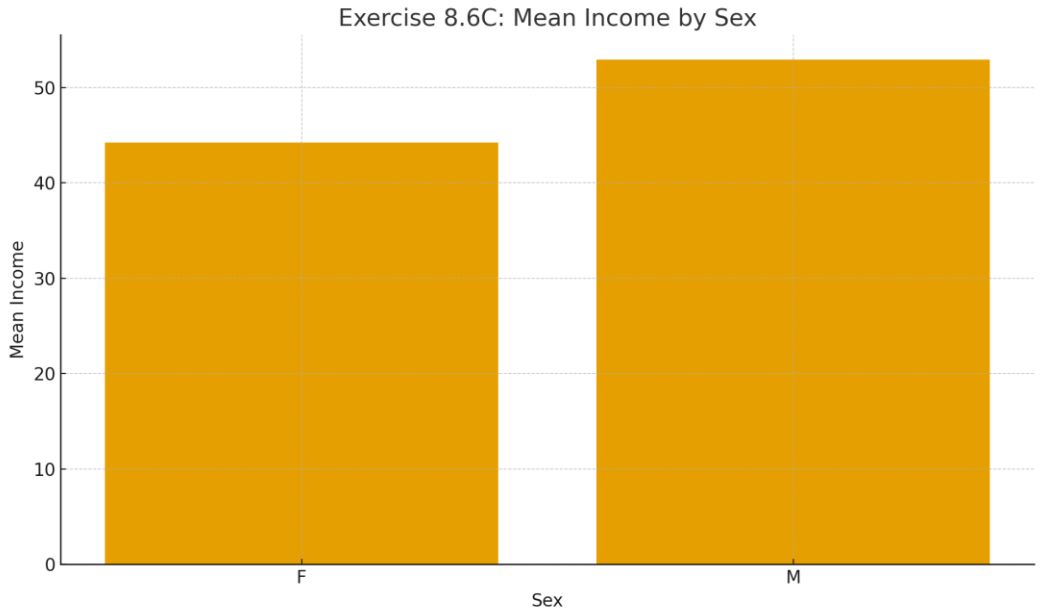


Key measures: Mean Agent1 = 8.25, SD = 1.03; Mean Agent2 = 8.68, SD = 1.04; Correlation(Agent1, Agent2) = 0.901.

Interpretation: Agent2 averages slightly higher than Agent1; the strong positive correlation (~0.90) indicates consistency of batch evaluation across agents.

Exercise 8.6C – Mean Income by Sex (Bar Chart)

Mean Income – Female = 44.23; Male = 52.91.



Interpretation: The sample indicates higher mean income for males than females. Further inferential testing (e.g., independent t-test) would be needed to assess statistical significance.

4. Reflection

Visualisation clarified underlying patterns from Unit 7's inferential results. Bar and column charts highlighted clear group differences; scatterplots and trendlines revealed relationships and supported forecasting; and summary measures contextualised the scale of effects. These techniques translate directly to cybersecurity reporting (incident frequencies, response-time trends, and correlation between controls and outcomes).