

Project Milestone 2: Data Visualization

I306 Statistics for Informatics

Overview

Building on your Milestone 1 exploratory analysis, this milestone focuses on creating effective visualizations that reveal patterns, relationships, and distributions in your data.

Due: End of Week 8 **Points:** 40

Requirements

Create a minimum of 5 visualizations that explore different aspects of your dataset. Your visualizations should include at least:

- One visualization showing the distribution of a numeric variable
- One visualization showing the distribution of a categorical variable
- One visualization showing the relationship between two numeric variables
- One visualization showing the relationship between a numeric and categorical variable
- One visualization of your choice that reveals something interesting about your data

Deliverables

Submit a Quarto document (.qmd) and its rendered PDF containing:

1. Visualizations (25 points)

For each visualization:

- Include a clear, descriptive title
- Label axes appropriately (with units if applicable)
- Use color purposefully (not just decoratively)
- Include a figure caption

Use `ggplot2` for all visualizations.

2. Narrative Interpretation (15 points)

For each visualization, write 2-3 sentences explaining:

- Explain why you chose to present this visualization in this assignment
- What pattern or relationship the visualization reveals
- Why this is interesting or important for understanding your data
- Any limitations or caveats (e.g., outliers, missing data)

Visualization Guidelines

Apply the principles we've discussed in class:

- Choose the right plot type for your data and question
- Avoid chartjunk—maximize data-ink ratio
- Use informative labels, not variable names
- Consider your audience

Suggested Plot Types

Data Type	Recommended Plots
One numeric	Histogram, density plot, boxplot
One categorical	Bar chart
Two numeric	Scatterplot, line plot
Numeric + Categorical	Side-by-side boxplots, faceted histograms
Two categorical	Stacked/grouped bar chart, mosaic plot

Submission

Submit your `.qmd` file and rendered rendered output (PDF or HTML) to Canvas by the due date.

Grading Rubric

Component	Points	Criteria
Visualization 1	5	Correct type, properly formatted
Visualization 2	5	Correct type, properly formatted
Visualization 3	5	Correct type, properly formatted
Visualization 4	5	Correct type, properly formatted
Visualization 5	5	Correct type, properly formatted
Narrative	15	Insightful interpretations for each visualization

Tips

- Don't just make plots—tell a story with your data

- Each visualization should answer a specific question
- Consider what you learned in Milestone 1 when deciding what to visualize
- If you notice something unexpected, explore it further