Individual Assignment: Data Visualization

Task

Create 1-2 multipanel figure(s) that demonstrate your mastery of data handling and visualization in R.

Your figures should showcase:

- Data wrangling using tidyr and dplyr
- Advanced ggplot2 visualization
- Spatial data handling and mapping (create a map in R or QGIS)
- Integration of various plot types into cohesive multipanel figure(s)

Technical Requirements

Your report must include:

1. Data transformations:

    - Show your use of tidyr and dplyr, for example:
        - Group_by() + summarize() operations
        - pivot_longer() or pivot_wider() operation

    _____

2. Visualizations including:

    - At least 3 different types of ggplot2 geoms, (including plots based on your data transformation)
    - One spatial visualization (map) using R or QGIS
    - Custom themes and thoughtful color choices
    - Proper handling of labels and legends

Written Report

Create a Quarto document that you can use for both GitHub submission and zoom presentation:

1. Data set Overview

   - Data source and context
   - Key variables
   - What your think of the data quality

   _____

2. Data Handling Section

   - Document and justify each major transformation
   - Include all code with comments
   - Explain how transformations support your visualization goals

   _____

3. Visualization Section

   - Individual figures with explanations
   - Color choice justification
   - Accessibility considerations
   - Final multipanel composition and layout

   _____

4. Reflection Section

   - New skills acquired
   - Challenges encountered and solutions
   - Areas for future improvement

Data Options

1. Your own data set:

   - It should contain a spatial component so you can create a map
   - Multiple variables suitable for different plot types (ex. categorical data, continuous data, interval data). Consult the teachers if uncertain.

2. TidyTuesday data set options:

   - [Birds observed at feeding stations in USA](#)
   - [Volcanic Eruptions](#)
   - Movies: [Monster movies](#) or [Summer movies](#)
     - For the movie data set, create a map based on 10 movies (Ex: the 5 worst and 5 best movies, production sites or another feature)

Github submission

- Upload Quarto document and rendered output
- Include README.md with project overview
- Submit your project GitHub page URL on the Canvas assignment page
- Include any extra data files you created (ex. coordinates for the movies)
- If you use your own data, don't make it public (consult teachers)

Zoom presentation (January 31)

- Reformat the quarto document to presentation mode for:
- 10-minute presentation
- 5 minutes for questions
- Focus on visualization choices and lessons learned

Contribution to TidyTuesday

Last year at least 30 courses used tidytuesday data sets.

If you used a TidyTuesday data set, the creators encourage you to contribute on social media. See the TidyTuesday GitHub page for details.

Let them know we used it too!

About

This assignment was created for the VS-GMAS course "P000085, Data handling and high-quality illustrations for publications", 3 credits, Spring semester 2025. Swedish University of Agricultural Sciences.