Group Assignment 2 (G2): Project Prototype

Interactive-Visual Data Analysis
Fall 2025

DUE: November 25, 23:59

General Task Description

This group assignment represents the core coding period of your project. You will now prepare a prototype of the tool you've characterized in G1. In this work phase, please continuously refer to the project characterization, visual encoding, and data modeling decisions you've already described in G1.

Your ultimate submission will be a video demonstration of your prototype, done similarly to your A1 assignment submission (the IVDA Tutorial). Prototype videos will be shown to the entire class on Thursday, Nov 27, so that everyone gets timely and useful feedback for the project refinement phase afterwards.

This assignment is worth 15 points in total for each student. Answers should relate back to your group's G1 submission. Generally, the points are allocated as follows, with a more detailed breakdown in the **Video Submission Instructions** and the **Coding Response Instructions**:

- **Computational Support:** Advanced data preprocessing/enrichment and integration of machine learning methods to suit your visualization requirements
- **Visual Analytics Development:** Supporting all of your identified tasks in your tool's prototype using an IVDA approach

Submission

You will submit 2 items to your OLAT group's folder: a video response and a PDF of your Jupyter notebook. Please submit all of your work **in your OLAT group as a single PDF and a single MP4 file**. Late submissions face the usual late policy, as outlined in the course syllabus, and will apply to all group members.

Your video response should be **an MP4, of no more than 4 minutes in length**, structured similarly to that of A1 (though you may also add a few slides, as needed). Videos will face point deductions when containing an unnatural, overly fast voiceover that is difficult to understand. Your coding response should be **a PDF submission of your team's Jupyter Notebook**. More details on each in the next couple of sections.

Video Submission Instructions (10 Points)

Your video response should be **an MP4 of no more than 4 minutes**, and include an update on your IVDA tool's development progress:

- Your team name, and the names of all of your group members
- A Nested Model summary characterization of your project, as presented in lecture and exercise (1 point)
- A reminder (or update) of your tool's task list (0.5 points)
- A demonstration of your tool's functionalities, with voiceover description (2 points)
- A clear characterization of the visual encoding, interactions, and view composition that your visualization(s) employ, with demonstration and voiceover description that clearly ties these decisions back to your project's task, domain, and user characterization (5 points)
- A reflection on how well your interaction and encoding decisions support your tasks, with discussion on what you could refine to support them better (1.5 points)

- Proper referencing of the sources used in your development, excluding IVDA lecture and exercise material

Video submissions with accelerated interaction and/or voice speed that impede the video understanding will receive a point deduction, depending on the severity of the issue.

Coding Response Instructions (5 Points)

Your PDF submission of your team's Jupyter Notebook should cover the entire data preprocessing, wrangling, and modeling process, including:

- Your team name, and the names of all of your group members
- A link to your group's GitLab or GitHub private repository* (mandatory)
- A brief (1-2 sentence) statement of how each team member contributed to your progress in completing G2 (0.5 points)
- A reminder of your tool's task list. Please ensure this matches the list you will provide in your video!
- A clear and complete characterization of your data, including:
  - Information on each (relevant) attribute/feature's data type and distribution of values (0.5 points)
  - Information of each (relevant) attribute/feature's data quality: amount of missing values, presence of extreme values/outliers, consistency, validity (see L11 slides – Data Wrangling and Visual Preprocessing) (0.5 points)
  - Information on your goals: what do you need to visualize and model with this data? What will it take to prepare your data for those goals? (100 words) (0.5 points)
- Clear and successful implementation of your data preparation goals, including:
  - Visualizations to appropriately and clearly present the data wrangling progress and results (1 point)
  - Visualizations to appropriately and clearly present the data modeling results, with a brief description of significant insights (100 words) (1 point)
- A short reflection on how well your data wrangling and modeling decisions support your tasks, with discussion on what you could refine to support them better (100 words) (1 point)

* Please invite mlacayo or mlacayoemery for GitLab or GitHub, respectively.