#### DS5610 EDA: Data Presentation Grading Rubric

**Assignment Overview**: In this assignment, you will showcase your proficiency in exploratory data analysis (EDA) by preparing a 5-minute in-class presentation. You will choose a dataset of your preference that involves panel data, spatial data, text data, or network data. Your presentation should encompass both descriptive statistics and exploratory visualizations. The assignment aims to evaluate your ability to effectively communicate your insights using high-quality tables, figures, and the underlying code.

**Presentation (10 Points)**

Introduction/Motivation (1 Point)

(0) The introduction is missing or inadequately covers the dataset and the motivating questions.

(1) A clear and concise introduction is provided, presenting the dataset and open- ended motivating questions effectively.

Data Description (1 Point)

(0) The description of the original data is missing or lacks clarity.

(1) A clear and detailed description of the original data is provided, explaining what parts of the dataset are necessary to answer the motivating questions.

Analysis (2-3 Slides) (6 Points)

(0) Fewer than 2 slides for analysis are presented.

(1) Analysis slides are present but are incomplete or lack clarity.

(2) Analysis slides are present, clear, and well-organized.

(4) Analysis slides are clear, well-organized, and effectively highlight aspects of the data related to the motivation.

(6) Analysis slides are clear, well-organized, effectively highlight aspects of the data, and showcase insightful exploratory visualizations supported by well- commented code snippets. At least one analysis slide is specific to the assigned data type (panel data, spatial data, text data, or network data).

Takeaway (1 Point)

(0) The takeaway is missing or lacks interpretation.

(1) A clear interpretation of the results of the analysis is provided, including the identification of trends and speculations about the data's appearance.

Clarity and Communication (2 Points)

(0) The presentation is disorganized and difficult to follow.

(1) The presentation is somewhat organized, but key points are unclear.

(2) The presentation is well-organized and effectively communicates key points.

**Code (10 Points)**

Code Quality (6 Points)

(0) Code is missing, incomplete, or of poor quality.

(2) Code is present but lacks proper organization or clarity. It may run with problems that require modification before the intended result can be replicated.

(4) Code is well-organized and clear, with some comments, but may have minor issues. There may be some unnecessary code/comments, but the code should run without any problems.

(6) Code is well-organized, clear, and includes comprehensive comments, making it highly reproducible and informative. Any data wrangling is done efficiently and there is no unnecessary code or commenting.

Code Correspondence (4 Points)

(0) Code is not provided for most of the tables and figures in the presentation.

(1) Some code is missing or incomplete for the tables and figures in the presentation.

(2) Code is provided for most tables and figures but lacks clarity or organization. It is difficult to tell which code sections produce which presentation outputs.

(4) Code is provided for all tables and figures, is well-organized, and clearly corresponds to the visuals in the presentation. The code could be easily altered by someone other than the original author to make modifications to the visualizations