Final Project Checkpoint 4: Visualization with ggplot2

Objective: In this checkpoint, you will visualize your regression analysis results from Checkpoints 1-3 using the ggplot2 package in R. Visualizations help interpret the model performance, residuals, and the relationships among variables.

Reference the starter code in Github (Final_Project/Checkpoint4.R) for assistance on data preprocessing and the visualizations for tasks 1-3.

📌 Task 1: Visualize Predicted vs Actual Values (7 points)

1. Generate a scatter plot comparing the predicted values from your original regression model (Checkpoint 2) with the actual observed values.

2. Generate a scatter plot comparing the predicted values from your adjusted regression model (Checkpoint 3) with the actual observed values.

✍ Interpret the plots: Does the adjusted model more accurately predict the response variable?

📸 Attach a screenshot of your plots.

📌 Task 2: Residuals Distribution Visualization (7 points)

1. Create a histogram showing the distribution of residuals from your original model.

2. Create a histogram showing the distribution of residuals from your adjusted model.

✍Interpret the residual distributions: How does the distribution differ between models?

📸 Attach a screenshot of your plots.

📌 **Task 3: Residuals vs Predicted Values (6 points)**

1. Construct a scatter plot of residuals against the predicted values from the original model.

2. Construct a scatter plot of residuals against the predicted values from the adjusted model.

✍️ Interpret the residuals vs predicted values plot: Is there evidence of heteroscedasticity?

📸 Attach a screenshot of your plots.

After all tasks are complete:

📸 Attach the R Script from RStudio that contains your code from Tasks 1-3.

Submit answers on the provided Google Form link.