Dataset Selection (10 pts)

You may pick any dataset for your final project, provided that the dataset has:

- 50+ cases/rows (1 pt)
- Numerical response (5 pts)
- Numerical predictor (2 pts)
- Categorical predictor (2 pts)

Exploratory Analysis (15 pts)

- Graphical displays for each variable
- Summary statistics for each variable
- One written paragraph per variable

Each variable will be worth 5 points (3 variables total: 1 response, two predictors).

SLR (50 pts)

- Plot and discussion of the variables (5 pts)
- Correlation between the variables (2 pts)
- Interpretation of R^2 (3 pts)
- Calculations and interpretation of slope and intercept for the model of the variables. When interpreting the intercept, clearly state if extrapolating (5 pts)
- Confidence intervals for slope and intercept (5 pts)
- Diagnostic plots with discussion (10 pts)
- Discussion of influential points (5 pts)
- Discussion of whether the linear model is appropriate (5 pts)
- Discussion of the relationship at hand–must be supported with statistical evidence. (10 pts)

Indicator Variable (15 pts)

- A new scatterplot–Does knowing the value of the categorical variable add any information to your prediction? (5 pts)
- Interpretation of adjusted R^2 value (2 pts)
- Calculation and interpretation of the coefficient on the categorical variable (3 pts)
- A confidence interval for the coefficient on the categorical variable (5 pts)

Report (10 pts)

- Formatted neatly with few errors
- Does not include unnecessary tables or graphs
- Does not include your code

Comments: