Statistical and Predictive Modeling II (DATA 2204)
Assignment #4 – Regularization (15% of Final Grade)
Professor: Ritwick Dutta

Mr. John Hughes is still trying to create an optimized model for his **EnergyUse-Cooling.csv** dataset. This time he would like you to use Regularization techniques.

If you recall the dataset has the following variables.

Independent Variables:

X1 - Relative Compactness
X2 - Surface Area
X3 - Wall Area
X4 - Roof Area
X5 - Overall Height
X6 - Orientation
X7 - Glazing Area
X8 - Glazing Area Distribution

Dependent Variable:

Y- Cooling Load

The Ask:

1. Create a PowerPoint (PPT) presentation that includes the following:
   a. Cover Page (Title, Name (1st and last) and Student Number)
   b. Rational Statement (summary of the problem or problems to be addressed by the PPT) – *2%*
   c. Present and explain **three (3) key insights** from the key metrics (Adj. R^2, MAE, RMSE) from each of the Optimized Regularization models (i.e. LASSO, Ridge, and Elastic Nets), but first **use Tukey to remove any outliers**. ***Note: nine (9) insights in total are required*** – *10%*
   d. State and explain **three (3) recommendations** for Mr. John Hughes for next steps. – *3%*

Attention: Please ensure that all key facts are in your slides and not in the notes section

Note: Please ignore the error for the Ridge output

Hint: Leverage the code from WK7b-Tutorial-RegTukey

Random State = 100 for all section

2. Provide a copy of your HTML Python Code

Please post your PowerPoint Document (.ppt) and HTML of Python Code via assignments under Assignment #4 by 11:59 p.m. on Thursday, March 24th, 2022

Grading Rubric

	Exemplary (14-15)	Proficient (10-13)	Incomplete (7-9)	Needs Improvement (0-6)
Analysis	Cover Page Complete Rational Statement is complete with supporting details Evaluation metrics and three (3) insights are presented and fully evaluated for each Optimized Regularization model	Cover Page Complete Rational Statement is complete with high-level supporting details Evaluation metrics and three (3) insights high-level evaluation for each Optimized Regularization model	Cover Page Incomplete Rational Statement is complete with missing supporting details Evaluation metrics and less than three (3) insights are presented and evaluated for each Optimized Regularization model	Cover Page missing Rational Statement missing Evaluation metrics and Insights are missing or incorrect.
Next Steps	Three (3) recommendations have been identified with detailed explanations.	Three (3) recommendations have been identified with only high-level explanations.	Less than Three (3) recommendations and incomplete explanations.	Recommendations are missing or incorrect.

Note: 50% Grade Penalty for missing Python HTML File