Project 2 Instructions

Goal

To understand the complexities of gathering data to answer a question that might not necessarily have data readily available, designing and validating a regression model to predict defined outcomes and presenting model results effectively.

Instructions

Come up with a question independent of data gathering, gather data, conduct EDA and develop a regression model that answers your question.

Using Rmarkdown develop a report of your project and publish to Rdocs so all students can see your work. Make sure your report answers at least the following questions:

1. How did you develop your question and what relevant research has already been completed on this topic?
2. How did you gather and prepare the data for analysis?
3. How did you select and determine the correct regression model to answer your question?
4. How reliable are your results?
5. What predictions can you make with your model? Examples
6. What additional information or analysis might improve your model results or work to control limitations?

Develop a presentation that provides an overview of your results, inclusive of the limitations and be prepared to demonstrate your knowledge in class. **15 slides max.**

Deliverables

Please turn in your final copy of each of these items in Blackboard:

- Rmarkdown report answering the questions detailed organized in alignment with the Data Lifecycle.
- Any slides or resources used for the presentation

Grading (See Rubric)

1. 33% Summary Report
2. 33% Model Development, Selection and Usage
3. 33% Presentation