# CS4 Rubric – Gas Price Modelling

**DS 4002 – Spring 2024**
**Submission format:**

- **Upload link to github repo and presentation pdf to canvas**

**Group Assignment**

**General Description:** Submit to canvas a link to your project repository and a pdf of you 3-minute presentation

Preparatory Assignments – None

**Why am I doing this?** This is your opportunity to apply your modeling skills to a real-world scenario and compete with the class to earn extra credit. This project will test your ability to work as a group and come up with ways to improve the predictive accuracy of your model.

- Course Learning Objective: use analytical skills to create and refine a time-series model.
- Course Learning Objective: perform well in a group setting.
- Course Learning Objective: submit and communicate results to the class.

**What am I going to do?** Your group will explore the factors that contribute to fluctuations in US gas prices (national average). After reading the hook document, check out the github contents. Your group has been given data on national average gas prices from 1999 to 2020. You will need to build a model to test this data using the best practices of machine learning. You will submit a github link with the R code that creates your model and performs the analysis. You will also give a 3-minute presentation to the class conveying the results of your analysis, including RMSE test metrics. Finally, each group's model will be tested on the 2021-2023 data, with the models with the lowest RMSE values being awarded extra credit for the class.

**Tips for success:**

- Think outside the box – what variables are the best indicators of gas price? You might need to add ones that are not already available
- Carefully consider different options for a model during the selection phase
- Make sure to reach out to your Professor and TA for assistance

**How will I know I have Succeeded?** You will meet expectations on CS3 Gas Price Modelling when you follow the criteria in the rubric below

| Spec Category | Spec Details |
|---|---|
| Formatting | ● One github Repository (submitted via link on canvas)<br>○ A README.md file (which auto displays)<br>○ A LICENSE.md file (use MIT as default)<br>○ A DATA folder |

| | |
|---|---|
| | ○ A SCRIPTS folder<br>○ AN OUTPUT folder<br>● One pdf of slides for a 3-minute presentation |
| README.md | ● Goal: This file serves as an orientation to everyone who comes to your repository, it should enable them to get their bearings.<br>● Provide a brief description of the project, model used, a guide to your repository, and a brief summary of results<br>● Include any references in IEEE format |
| LICENSE.md | ● Goal: This file explains to a visitor the terms under which they may use and cite your repository.<br>● Select an appropriate license from the GitHub options list on repository creation.<br>● Usually, the MIT license is appropriate. |
| DATA folder | ● Goal: Contains all the data used for analysis<br>● Upload data in .csv format<br>● Make sure to update the data if you added any variables for analysis and make note of this<br>● Include sources of where the data came from |
| SCRIPTS folder | ● Goal: Document the model creation and analysis process<br>● Include an R markdown file that processes the data and creates your model. This file should also have a training and testing set and output RMSE values on both sets<br>● Any type of model is acceptable, make sure to include any data cleaning as well |
| OUTPUT folder | ● Goal: Show any figures or results<br>● You should add an appendix which contains the summary statistics on your model, and the final results for analysis<br>● You should also include an. RData file with your model that can be used for prediction purposes on the 2021-2023 dataset for the competition |
| Presentation (pdf) | ● Goal: Demonstrate an outline of the presentation you will give to the class<br>● Submit the presentation pdf on canvas<br>● This needs to show methodology, model selection process, refinements, and results<br>● Aim for a 3-minute presentation |
| References | ● All references should be listed at the end of the document<br>● Use IEEE Documentation style (link) |

Acknowledgements: Special thanks to Jess Taggart from UVA CTE for coaching on making this rubric. This structure is pulled from Streifer & Palmer (2020).