CS Rubric – Climate Prediction

DS 4002 – Spring 2024 - Nishita Cheekatamarla
Submission format: Upload GitHub repository link and PDF to canvas
Individual Assignment

General Description: This assignment will allow you to practice predictive modeling using climate data containing many different parameters.

Prerequisites - Familiarity with coding in Python and/or R.

Why am I doing this? The goal of this assignment is to work on unfamiliar models and create a deliverable presenting your work. Data science work typically involves working with problems that can have messy, complicated solutions and this project can provide a way for you to experience that environment on a smaller scale. This is a chance for you to use your technical skills as an upcoming data scientist to find a potential solution to a problem and make further improvements to it.

- Course Learning Objective: Build a predictive model based on given data.
- Course Learning Objective: Prepare and document findings in a case study format.

What am I going to do? Begin by reading the hook document to understand the main context for the problem. Then go through the provided resources (articles, dataset, etc) in order to help you decide how you want to create your model. Outside research is not necessary but is encouraged! Other resources can provide you with further context and information regarding the topic of interest. Then use the provided datasets to create the predictive model that can be used to predict different climate parameters for the future. Submit a written document detailing the procedures and findings of your project as a whole.

Deliverables include:

- Multipage PDF document explaining your findings and the process used to obtain them, including a hypothesis, data cleaning procedures, analysis plan, conclusions, and next steps.
- Online github repository including all resources, figures, and used code.

Tips for success:

- Focus, focus, focus – Avoid distractions
- Be bold - This is your chance to try something new, creative, and fun.
- Don't overthink it - This could be a difficult problem to solve that may take many tries. Try your best.
- Communicate with the instructors. This is a creative assignment, they may be able to assist you.
- Do your research on the topic at hand and the technical information needed for the project.

How will I know I have succeeded? You will meet expectations on the case study when you follow the specific criteria in the rubric below:

Formatting	● One GitHub repository (submitted via link on canvas)
	  o The top level page should contain the following:
	  ▪ SRC folder
	  ▪ DATA folder
	  ▪ OUTPUT folder
	● Multipage PDF document including the following sections:
	  o Introduction/Hypothesis
	  o Data Cleaning/Preprocessing/Methodology
	  o Analysis Plan
	  o Tricky Analysis Decision
	  o Findings and Results
	  o Conclusions
	  o Next Steps and Future Work
	  o References

GitHub	● Goal: This repository should contain all source code, data files, and all outputs and figures generated throughout the project ● SRC folder: all source code used in the project ● DATA folder: includes the given data, along with any supplemental data used throughout the project ● OUTPUT folder: all figures and/or outputs generated during the project, labeled with specific and detailed titles/axis labels or explained in a README.md file.
PDF Document	● Goal: Multipage PDF document explaining all the steps taken to complete this case study along with the results, conclusions, and analysis written using professional vocabulary. ● Introduction/Hypothesis: Give context to the problem and describe what you learned from the explanation article given. Include your hypothesis and the goals for your project. ● Data Cleaning/Preprocessing/Methodology: Describe the steps you took to modify the data and why it was necessary to do so. ● Analysis Plan: Beginning of the main section of the document. Include details regarding the type of model used and how you plan to evaluate the results produced by it. Include a quantifiable goal. ● Tricky Analysis Decision: Identify and describe a trivial decision made during the process and why it impacts the overall product. ● Findings and Results: Address the hypothesis and present the results of the model you built. Interpret the results in context of the problem. ● Conclusions: Discuss the overarching, general conclusions drawn from the results of the project. Discuss main takeaways of the overall project.

	● Next Steps and Future Work: Discuss how these findings open new avenues for future work and things you would change moving forward. ● References: Use IEEE documentation style to cite your sources

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