

Case Study Rubric – Replication of Project

DS 4002

Individual Assignment

General Description: This rubric describes the steps needed to complete the case replication regarding hurricane activity. This includes the steps to open the data, complete the analysis, and finalize with visualizations. Submit to canvas your GitHub repository with Jupyter Notebook and PDF file.

Why am I doing this? Replication is an important part of studying data science. This project tests you with the ability to replicate the project of another student. This is invaluable as when you enter the workforce you will need to be able to reproduce the work of others. You will also be able to potentially learn a new topic and skills, to broaden your toolkit as a data scientist. Gaining exposure to different approaches when looking at a data science problem will be invaluable for your academic and professional career.

What am I going to do? Access the GitHub repository provided in the Hook Document. Follow the Rubric steps below to replicate the project and conclude with the RMSE result. Reflect on whether you think the model could be used by climatologists to increase awareness on climate change factors and their influence on hurricane activity.

Tips for Success:

- Follow instructions carefully and in the order they are listed
- Don't be afraid to make mistakes, you might be learning something new for the first time
- Reach out to course instructors, peers and online resources if you need assistance

How will I know I succeeded? You will meet expectations for the Case Study Replication Rubric if you follow the steps below:

Spec Category	Spec Details
Formatting	<ul style="list-style-type: none">● <u>Goal:</u> produce one GitHub repository with the information below<ul style="list-style-type: none">○ Your repository should contain:<ul style="list-style-type: none">■ HurricaneAnalysis.ipynb<ul style="list-style-type: none">● Follow the steps in the “Analysis” and “Results and Visualization” subheadings below and upload the finished .ipynb file.■ Conclusion.pdf<ul style="list-style-type: none">● A summary of your findings as detailed in the “Conclusion” section
Data Acquisition	<ul style="list-style-type: none">● <u>Goal:</u> download and familiarize yourself with the dataset

	<p>provided in the GitHub to replicate the project.</p> <ul style="list-style-type: none"> ● Navigate to the GitHub repository provided in the Hook Doc ● From the “DATA” folder download the dataset “Hurricane_Dataset.csv”. ● View “DataDict.pdf” also in the “DATA” folder to clear any confusion on the variables.
Analysis	<ul style="list-style-type: none"> ● <u>Goal</u>: visualize the data, conduct Granger’s Causality Test, Cointegration Test, and finalize Vector Autoregressive (VAR) model. ● In the “CODE” folder, make a copy of the “HurricaneAnalysis.ipynb” file. Some steps are completed for you, and the required packages are imported. ● Run each cell at the beginning of the notebook, up until Granger’s Causality Test. You will then see visualizations presented. Familiarize yourself with these visualizations and the data, and make comments on your Jupyter Notebook with any trends you observe. ● Conduct Granger’s Causality test using the guidelines provided in the notebook, ensuring a significant p-value. ● Conduct Cointegration Test to prove significant independent variables. ● Run the provided code that separates the data into train-test sets and conducts the Augmented-Dickey Fuller test. ● Create the VAR model <ul style="list-style-type: none"> ○ Suggested: utilize AIC scores and the Durbin-Watson statistic.
Results and Visualization	<ul style="list-style-type: none"> ● <u>Goal</u>: produce statistics and visualizations that explain the results from your VAR model. ● Calculate the RMSE and interpret the result. ● Plot the 2018-2021 forecasted data alongside the 2018-2021 actual data. Interpret the result. How strong is the model’s predictive power?
Conclusion	<ul style="list-style-type: none"> ● <u>Goal</u>: summarize your results and make a reflection in a document titled “Conclusion.pdf”. ● Answer the following questions: <ul style="list-style-type: none"> ○ What are your key findings? ○ Are your results strong enough to be applied in the real world? ○ What would you change about this case study? ○ What were some challenges you faced?