Case Study Rubric

Due: TBD
Submission format:
- Upload link to Github repository to Canvas
- Submit PDF of written analysis to Canvas

Why am I doing this? This is an opportunity to integrate the skills you have learned in this class to create a final deliverable that demonstrates the technical and conceptual skills you have developed. It will provide you with practice in synthesizing information and presenting it in a concise manner to inform an audience.

What am I going to do? You will be given a case study and asked to complete a specific task using specific models to draw conclusions. The case study will include a dataset, additional resources that provide information about the topic and technical skills, and the motivation for the case study. Your job is to create an analysis that answers the prompt provided in the case study and create a final deliverable that explains the conclusions drawn from the dataset. The deliverable will include:

- A written explanation of the graphs and the implications of the conclusions drawn
- A link to a github repository with all of the materials

How will I know I have Succeeded? You will meet expectations on this case study when you follow the criteria in the rubric below.

Spec Category	Spec Details
Formatting	- Written portion - Submit the written portion as a PDF file to canvas - Github Repository - Include scripts with all code/analysis, include folder with outputs, references folder
Written Portion	- Provide a brief overview of the topic - Include information about the variables and dataset - In a paragraph discuss the questions being answered in the case study and the implications of the analysis - In a paragraph discuss the modeling approach being used - Describe the analysis plan - Provide an explanation for why a particular model is being used - Include a short executive summary that outlines the structure of the document - Use short sentences

	● Include a section on EDA that highlights key information about the dataset
	  ○ Include graphs and visualizations that are representative of the EDA
	  ○ Highlight which questions were answered by the EDA and how it aided the analysis
	● Analysis
	  ○ Include graphs and visualizations demonstrating the relationship between variables in the dataset
	  ○ Include graphs of the models used
	    ■ Explained what is being shown in the graph and the implications
	  ○ Include graph of the predicted model used
	    ■ Highlight which variables most contribute to high crime rates based on the predictive model
	● Conclusion
	  ○ In at least 1 page, identify main conclusions drawn from analysis
	  ○ Suggest improvements for analysis and next steps
	● Implications
	  ○ Explain the relevance of the analysis
	  ○ Provide suggestions for policymakers based on the analysis

Github Repository	Analysis
	  ○ Include scripts of all data analysis
	    ■ Include detailed but brief comments about what each chunk of the code is doing
	  ○ EDA- create exploratory graphs to provide information about the dataset
	    ■ Such as: areas with highest crime rates, demographic information of victims, most common crimes, etc.
	  ○ Incorporate graphs using SARIMA and ARIMA models
	    ■ Compare models to

	identify influence of seasonal fluctuations
	    ■ Include paragraphs explaining what is being demonstrated by each model
	  ○ Use random forest to create a predictive model
	      ■ Include accuracy, precision, recall, F-1 score to evaluate the model
	Outputs Folder
	  ● Include any outputs as a PDF, even those that do not make it into the written portion
	  ● Provide a detailed title for each item in the output folder so they can be easily identified

References	● Both the github repository and the written portion should include a section with references in IEEE citation style