**Jupyter File Code(20pts)**

2 Points : Code For Sourcing Data

2 : Clear and easy to follow, and the appropriate libraries are being used.

1 : Incorrect use of the libraries, and sourcing the data requires too many requests.

5 Points : Code For Database

3-5 : Code for creating the table(s) and the database(s) are clear and easy to follow. There is high quality SQL code being used, and we are following the rules for character escaping. We are following the rules for working with data, storing raw data in a table, and pulling from that table when needed. Any further changes to the data should be in a separate table(if needed), the raw data should never be altered, only accessed.

1-2 : There is little effort in making the database(s), and the table(s) themselves are hard to access, and hard to work with. Code was not made with extensibility in mind.

2 Points :

2:When inserting into the databases, we are using appropriately encapsulated helper functions to make the raw data easy to insert. Helper functions for accessing the database are extensible, allowing a very fine or broad selection of rows.

1: Helper functions do not encapsulate reusable extensible functions that will be needed over the lifetime of the project.

5 Points: Quality of Data Cleaning

5: All of the columns and rows are accounted for. There is uniformity to the columns and rows, and any Nulls or NaNs are either replaced, taken out, or filled in, with justification being given as to why. The Data is easy to traverse and to manipulate.

3-4 : The Data set has been cleaned, but there is still some variation of values between rows and columns. Nulls and NaNs were replaced or taken out, but no justification was given as to why. The Data can be worked, but it is difficult and time consuming.

1-2 : There was little to no effort made to clean the data or make it easier to work with. Working through the dataset was difficult because steps were not taken here to simplify it.

6 Points : Clean logical application of the Data Science Libraries.

5-6 : Pandas methods and functions were invoked with purpose, and redundant code was taken out of the project and replaced with more efficient code. Code that was going to be used again was encapsulated in functions, and called throughout the program. Use of Vectorization was included when the situation called for it, as well as mapping and applying. When applying regression we are using the correct loss function as well and tuning the hyperparameters for the optimal regression line. Matplotlib was used in a way to make regression lines easy to understand as well as other aspects of the dataset.

3-4 : The libraries were applied, but the application was difficult to read, and it was hard to justify the uses of the methods and functions being used. An attempt was made to refactor code, but improvements were still able to be made.

1-2 : No reasoning was used when calling methods and functions, and the libraries were misused.

**Jupyter File Code Structure(10pts)**

3 Points : Sense of flow within the Code.

3 : There was a clear flow to the program, and following through it was easy because logical steps were being followed from one cell to another, and the reader was being facilitated through the cells.

2: The flow was difficult to follow at times, but overall it was clear enough to follow alongside the project.

1: There was difficulty following through the project, and there was no attempt made to provide clarity and justification as to what was being done, when it was being done, and why it was being done.

3 Points : Segmented and Commented Code

3 : Lines of code were clearly commented on, and thoroughly explained what was happening in the cells. Cells were divided by logic, and each cell encapsulates a single change happening without in our project.

2: There was an attempt to comment on the code, but not enough clarity was provided, and there was ambiguity as to what was going on in a cell. Cells are overloaded and contain different applications of logic and code, and are not properly encapsulated.

1: There was little to no attempt to comment, and the cells are hard to follow along logically.

2 Points : Clear and labeled Visualizations

2: The visualizations are clearly labeled and sized, and it is clear as to what the visualization is meant to represent. A Write up is provided when needed.

1: The visualizations are not labeled and sized, and there is little to no understanding gained from the visualizations that were generated.

**Project Write Up(10pts)**

2 Points : Clarity

2: The document is easy to read and is clear and concise.

1: The document is submitted, but there are many spelling mistakes or grammatical mistakes that make it difficult to discern the meaning of the document.

0 : It is not possible to understand the document or read through it without being given clarification.

2 Points : Flow

2: There is a clear structure and flow to your write up.

1: There is confusion regarding the decision as to the order of things being explained.

0: It is difficult to understand or determine why the document flows in the way submitted.

6 Points - Details regarding the knowledge gained.

5-6 : You are able to convince the reader that you have gained either domain knowledge about the topic you were researching, or that you were able to acquire a new understanding of the Data Science Libraries. You are using the correct vocabulary and the correct terminology when talking about the methods and procedures that you took when working through the data. You are making clear and direct calls to specific methods or functions from your code and clearly explaining what you did, and why you undertook these tasks.

3-4 : You are making references to the code and visuals you generated, but it is not clear how you implemented them, and it’s not clear as to what you learned. You are often misnaming or incorrectly referencing aspects of the code.While it is clear that you learned something, you aren’t able to clearly explain what it is you learned.

1-2 : There is no indication that you learned anything from the project.

**Script(10pts)**

*Out of 10*

2 Points : Clarity

2: The document is easy to read and is clear and concise.

1: The document is submitted, but there are many spelling mistakes or grammatical mistakes that make it difficult to discern the meaning of the document.

0 : It is not possible to understand the document or read through it without being given clarification.

2 Points : Flow

2: There is a clear structure and flow to your write up.

1: There is confusion regarding the decision as to the order of things being explained.

0: It is difficult to understand or determine why the document flows in the way submitted.

6 Points : Content

5-6 : You highlighted the best parts of your project and your process, and it is clear as to why you choose to talk about certain topics. You explain thoroughly and clearly your thought processes and the knowledge you gained while working through sections of your project. It is clear that you have a high level understanding of Pandas and the DS Libraries, as well as your Data.

3-4 : You choose appropriate topics, but it is not clear why you choose these, or it is not clear what you learned about these topics. The script is well written, but had more time been spent preparing the script, it would be more convincing that a high level understanding of Pandas and the DS Libraries and your data exists.

1-2 : Your script does not showcase you or your talents, or what you learned about Pandas or your Data.

**Presentation (20pts)**

6 points : Over 5 minutes, under 10 minutes

5-6 : Your presentation falls between these two limits.

3-4 : Your presentation is 2-4 minutes too long or 1 minute too short.

1-2 : Your presentation is only 1-2 minutes long, or is way over time (5 minutes or more).

6 points : Clear and logical explanations of how the person went through the project

5-6 : Your presentation clearly shows what you learned, how you applied your knowledge, and is clear and concise.

3-4 : Your presentation shows growth, but does not convince the audience that you have taken a significant step in your Data Science journey.

1-2 : Your presentation is flat, and does not show any knowledge or mastery gained.

6 points : Confidence and clarity while talking about their project

5-6 : There is confidence in the presentation, and it is clear that the presentation was rehearsed and practiced.

3-4 : There are some hiccups during the presentation, but the presentation moves along, and overall is smooth and clear.

1-2: It is obvious that no practice was put into the presentation and that the script was not followed, or was not practiced.

2 points : Timeliness

2 Points : On Time

0 Points : Late