Here are the components you should aim to cover in your \*\*paper\*\*:

\* Problem statement and hypothesis

\* Description of your data set and how it was obtained

\* Description of any pre-processing steps you took

\* What you learned from exploring the data, including visualizations

\* How you chose which features to use in your analysis

\* Details of your modeling process, including how you selected your models and validated them

\* Your challenges and successes

\* Possible extensions or business applications of your project

\* Conclusions and key learnings

Your \*\*presentation\*\* should cover these components with less breadth and less depth. Focus on creating an engaging, clear, and informative presentation that tells the story of your project.

You should create a GitHub repository for your project that contains the following:

\* \*\*Project paper:\*\* any format (PDF, Markdown, etc.)

\* \*\*Presentation slides:\*\* any format (PDF, PowerPoint, Google Slides, etc.)

\* \*\*Code:\*\* commented Python scripts, and any other code you used in the project

\* \*\*Data:\*\* data files in "raw" or "processed" format

\* \*\*Data dictionary (aka "code book"):\*\* description of each variable, including units