

PROJECT SPECIFICATION

Write a Data Science Blog Post

Code Functionality and Readability

CRITERIA	MEETS SPECIFICATIONS
Code is readable (uses good coding practices - PEP8)	Code has easy-to-follow logical structure. The code uses comments effectively and/or Notebook Markdown cells correctly. The steps of the data science process (gather, assess, clean, analyze, model, visualize) are clearly identified with comments or Markdown cells, as well. The naming for variables and functions should be according to PEP8 style guide.
Code is functional.	All the project code is contained in a Jupyter notebook, which demonstrates successful execution and output of the code.
Write code that is well documented and uses functions and classes as necessary.	Code is well documented and uses functions and classes as necessary. All functions include document strings. DRY principles are implemented.

Data

CRITERIA	MEETS SPECIFICATIONS
Project follows the CRISP-DM Process while analyzing their data.	Project follows the CRISP-DM process outlined for questions through communication. This can be done in the README or the notebook. If a question does not require machine learning, descriptive or inferential statistics should be used to create a compelling answer to a particular question.
Proper handling of categorical and missing values in the dataset.	Categorical variables are handled appropriately for machine learning models (if models are created). Missing values are also handled appropriately for both descriptive and ML techniques. Document why a particular approach was used, and why it was appropriate for a particular situation.

Analysis, Modeling, Visualization

CRITERIA	MEETS SPECIFICATIONS
There are 3-5 business questions answered.	There are between 3-5 questions asked, related to the business or real-world context of the data. Each question is answered with an appropriate visualization, table, or statistic.

Github Repository

CRITERIA	MEETS SPECIFICATIONS
Student must publish their code in a public Github repository.	Student must have a Github repository of their project. The repository must have a README.md file that communicates the libraries used, the motivation for the project, the files in the repository with a small description of each, a summary of the results of the analysis, and necessary acknowledgements. Students should not use another student's code to complete the project, but they may use other references on the web including StackOverflow and Kaggle to complete the project.

Blog Post

CRITERIA	MEETS SPECIFICATIONS
Communicate their findings with stakeholders.	Student must have a blog post on a platform of their own choice (can be on their website, a Medium post or Github blog post). Student must communicate their results clearly. The post should not dive into technical details or difficulties of the analysis - this should be saved for Github. The post should be understandable for non-technical people from many fields.
There should be an intriguing title and image related to the project.	Student must have a title and image to draw readers to their post.
The body of the post has paragraphs that are broken up by appropriate white space and images.	There are no long, ongoing blocks of text without line breaks or images for separation anywhere in the post.
Each question has a clearly communicated solution.	Each question is answered with a clear visual, table, or statistic that provides how the data supports or disagrees with some hypothesis that could be formed by each question of interest.