Investigating Relationships Between Financial News Sentiment and Stock Price Rubric

Why am I doing this? This assignment is your opportunity to apply the skills you've developed throughout your data science coursework. You'll reflect on and reproduce the work of a former classmate, allowing you to practice the full data science workflow — from data collection and aggregation to sentiment analysis, statistical testing, and interpreting the results.

Category	Requirements
Formatting	The project should be organized in a single, well-structured GitHub repository. It must include the following components: ● README file ● LICENSE file ● A scripts/ folder ● An outputs/ folder ● A data/ folder ● Any additional materials relevant to the project
README	The README should effectively guide the reader through the contents of the repository. It must: ● Provide context and background of the project ● Describe any software, tools, and platforms used ● Include clear instructions on how to reproduce the analysis using the repository contents ● Cite relevant sources such as articles, raw data sources, or any external references
LICENSE	The LICENSE file should inform users about the terms under which they may use and cite your repository. ● Choose an appropriate license when creating the repository (e.g., the MIT License is generally suitable)
SCRIPTS Folder	This folder should contain all relevant code written for the project. All scripts must be:

	● Well-commented and easy to understand ● Replicable and follow standard coding practices ● Written using consistent naming conventions (e.g., snake_case for Python, camelCase for Java)
DATA Folder	Include raw or processed data used in the project, provided the data size is within GitHub's upload limits. If not, then: ● Include a PDF in the data/ folder with: ○ A link to the data ○ Instructions for accessing or downloading it ○ A description of how the data was obtained
OUTPUTS Folder	Store any output files generated during the project, such as: ● CSV files ● Images ● Any other results relevant for analysis or interpretation
Results	By the end of the project, you should be able to: ● Use the Alpaca API to fetch stock and financial news data ● Aggregate and preprocess data effectively ● Perform sentiment analysis ● Apply moving averages and justify their use (especially for fluctuating sentiment and price data) ● Conduct and interpret time-lagged correlation analysis ● Perform and interpret Granger causality tests ● Communicate findings clearly