

## 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: <ul style="list-style-type: none"><li>• README file</li><li>• LICENSE file</li><li>• A scripts/ folder</li><li>• An outputs/ folder</li><li>• A data/ folder</li><li>• Any additional materials relevant to the project</li></ul>
README	The README should effectively guide the reader through the contents of the repository. It must: <ul style="list-style-type: none"><li>• Provide context and background of the project</li><li>• Describe any software, tools, and platforms used</li><li>• Include clear instructions on how to reproduce the analysis using the repository contents</li><li>• Cite relevant sources such as articles, raw data sources, or any external references</li></ul>
LICENSE	The LICENSE file should inform users about the terms under which they may use and cite your repository. <ul style="list-style-type: none"><li>• Choose an appropriate license when creating the repository (e.g., the MIT License is generally suitable)</li></ul>
SCRIPTS Folder	This folder should contain all relevant code written for the project. All scripts must be:

	<ul style="list-style-type: none"> <li>● Well-commented and easy to understand</li> <li>● Replicable and follow standard coding practices</li> <li>● Written using consistent naming conventions (e.g., snake_case for Python, camelCase for Java)</li> </ul>
DATA Folder	<p>Include raw or processed data used in the project, provided the data size is within GitHub's upload limits. If not, then:</p> <ul style="list-style-type: none"> <li>● Include a PDF in the data/ folder with: <ul style="list-style-type: none"> <li>○ A link to the data</li> <li>○ Instructions for accessing or downloading it</li> <li>○ A description of how the data was obtained</li> </ul> </li> </ul>
OUTPUTS Folder	<p>Store any output files generated during the project, such as:</p> <ul style="list-style-type: none"> <li>● CSV files</li> <li>● Images</li> <li>● Any other results relevant for analysis or interpretation</li> </ul>
Results	<p>By the end of the project, you should be able to:</p> <ul style="list-style-type: none"> <li>● Use the Alpaca API to fetch stock and financial news data</li> <li>● Aggregate and preprocess data effectively</li> <li>● Perform sentiment analysis</li> <li>● Apply moving averages and justify their use (especially for fluctuating sentiment and price data)</li> <li>● Conduct and interpret time-lagged correlation analysis</li> <li>● Perform and interpret Granger causality tests</li> <li>● Communicate findings clearly</li> </ul>