

Rubric - Finding Good Books From Hooks

Ishan Korothe, 2023

Submission format: Python code (.py or .ipynb) file(s) and PDF explanation/review of case study

Individual Assignment

General Description: Develop a Python script/notebook or modify existing code to complete this project and explain how your changes impacted the results and what correlations you discovered. Your focus lies within the realm of popular books and what variables stand out in a hook.

Why am I doing this? This assignment focuses on applying what you learned in data analytics to expand upon a project that has been done to analyze good books from hooks. The analyses used here, in this case, are sentimental and parts-of-speech analysis, which you are free to modify as needed. This case study is intended to get you familiar with text-based data and how you can improve upon what has already been done. Think of this as an internship experience, for instance. Make this an opportunity for you to showcase your knowledge in various analytical methods and perhaps even make use of machine learning to help add more correlations to good hooks for good books.

What am I going to do? Your task is to modify existing code and/or do another form of data analytics based off the data gathered and either produce results that show better correlations or come up with a new factor that determines the hook of the book. Once you get something desirable or interesting, proceed to make a write up about how you derived your methodologies and what you did you get the results you got. For extra credit, find a way to collect books, their hooks, and other potential variables that could help you get great results and discover overlooked variables. You could also potentially change the objective of this project to focus on another variable rather than the hook itself for extra credit as well.

Tips for success:

- Dive deep and research different methods of text analyses people use.
- Learn about the text structure of hooks and what makes them good for a wide audience.
- Have fun!

How will I know I have Succeeded? You will meet expectations on CS2 Create Case Study when you follow the criteria in the rubric below.

Spec Category	Spec Details
Formatting	<ul style="list-style-type: none"> • Python scripts/notebooks – Files you created or modified from the GitHub repository. <ul style="list-style-type: none"> ○ Can be either .py or .ipynb ○ Specify your name, date, and file name at the top of the files you created/modified. ○ State explicitly whether the script/notebook was your creation or a modified version of a script/notebook from the GitHub repository below the line with your name, date, and file name. • Documentation – Explain your methodology/approach in what you created/modified and elaborate upon the results you got. <ul style="list-style-type: none"> ○ PDF format ○ Order <ul style="list-style-type: none"> ▪ Title ▪ Your name and date ▪ An explanation of your methodology/approach in what you did. ▪ An explanation of your results and what significance they have on this project. ▪ Length for the above two bullets should be around several sentences (3-5) to a couple paragraphs (2-3).
Code File (s)	<ul style="list-style-type: none"> • Follow the instructions stated in the formatting. • Ensure that the code is well-commented to help an interpreter understand what you did.
Documentation	<ul style="list-style-type: none"> • Goal: This should be used as a reference in your thought process on what you did for this case study. • Follow the instructions stated in the formatting. • Methodology/approach should elaborate upon how you brainstormed your idea and why you did it that way. Make sure to also talk about what exactly your method is trying to resolve. • Explanation for results must be detailed and discuss any improved/new correlations. If nothing was improved, state why this could be the case and how you would go about improving upon the project at hand.

Acknowledgements: Special thanks to Professor Alonzi and Harsh for this fun experience! 😊