# Taylor Swift Lyrics Topic, Sentiment, and Emotional analysis Rubric

**Purpose:**

This project is designed to help you develop foundational skills in text analysis, focusing on extracting meaningful insights from text data using R and Python. By working through this project, you will:

- ★ Understand and apply topic modeling, sentiment analysis, and emotion analysis using R.
- ★ Learn how to interact with a large language model (LLM) through Python API prompts.
- ★ Gain hands-on experience in cleaning, preprocessing, analyzing, and visualizing text data.
- ★ Build confidence and knowledge to design and execute your own text analysis project on a dataset of your choice.

You will start with a guided case study, using provided scripts, tools, and resources. The project will also introduce you to best practices for organizing your work, documenting your process, and sharing your insights using GitHub.

| Spec Category | Spec Details |
|---|---|
| Formatting | Goal: Organize your code<br>Your final github should have a:<br>● README.md file: A clear overview of your project, including methodology, findings, and references in IEEE style<br>● LICENSE.md file: Use the MIT license as the default<br>● Processing Script: For data cleaning and preprocessing<br>● Analysis Script: For sentiment, emotion, and topic modeling analysis<br>● Master Script: compiled code<br>● Annotated Taylor Swift Master Script |
| Engage with Case Study | Goal: Engage with the Taylor Swift code and learn from it<br>In your project repository, upload a script containing annotations of the Taylor Swift master script. In your annotations write about:<br>● The changes you observed in significance or topics tweaking hyperparameters like sparsity in document term matrix and the value of k in the topic modeling section<br>● The changes observed in transformer emotion confidence scores when hyper parameters were tweaked |
| Create your own project | Goal: Apply knowledge gained from the Taylor swift case study.<br>Dataset sources are provided for you below. Select text data from any of these sources and put into practice what you have learnt. Your project must contain:<br>● Sentiment Analysis(using SentimentAnalysis package in R)<br>● Emotional Analysis(using transforEmotion package in R)<br>● Topic Modeling(using topicmodels package in R)<br>You do not need to include significance testing, but you can if you want. The primary goal of this study is to get you comfortable with using data science to derive insight from the emotions and topics available in vast amounts of text data. |

| | |
|---|---|
| Materials | Data<br>https://archive.ics.uci.edu/<br>https://www.kaggle.com/<br><br>Learn more about Packages<br>Stefan Feuerriegel, “SentimentAnalysis Vignette,” cran.r-project.org.<br>“Topic Modeling with R,” ladal.edu.au.<br>Data Centric Inc., “Tutorial on topic modelling in r tutorial”<br>AWS, “What is Sentiment Analysis? - Sentiment Analysis Explained - AWS”<br>[9]<br>N. Proellochs, “SentimentAnalysis package - RDocumentation,”<br><br>Citation<br>https://www.mybib.com/tools/ieee-citation-generator<br><br>GitHub: https://github.com/cann-emma/TSwiftLyricAnalysis |
| References | All references should be listed at the end of the README.md document<br>Use IEEE Documentation style |

**Acknowledgements:** Special thanks toProfessor Alonzi, Layla Ranjbar, and Jess Taggart from UVA CTE for coaching on making this rubric.This structure is pulled from Streifer & Palmer (2020).