

### 1. PROJECT TITLE

-- Project\_Gutenberg\_using\_R

### 2. PROJECT DESCRIPTION

-- R script scraping corpora from Project Gutenberg for text analysis and comparison.

-- Uses two science authors and two science fiction authors:

- "Experiments with Alternate Currents of High Potential and High Frequency" by Nikola Tesla, Gutenberg ID: 13476.
- "Relativity: The Special and General Theory" by Albert Einstein, Gutenberg ID: 30155.
- "A Journey to the Centre of the Earth" by Jules Verne, Gutenberg ID: 18857.
- "The Time Machine" by H. G. Wells, Gutenberg ID: 35.

### 3. PROJECT FOLDER CONTENTS

-- This README.pdf file.

-- Link\_to\_Youtube\_Presentation.pdf file with link to video presentation, "CAP\_6640\_Presentation\_#2" on YouTube (<https://youtu.be/nMOEhUmcs1s>).

-- R script, "NLP\_tool\_demo.R"

-- Expected output files (4): Rplot\_\*.png files showing top twenty words in each corpora as bar plots.

-- Expected output files (4): Rplot\_\*\_cloud.png files showing top twenty words in each corpora as word clouds.