

DS5001: Final Project Manifest

- Provenance: I got my source files from the Project Gutenberg website. Project Gutenberg is a library of over 60,000 free eBooks where they can be downloaded as raw text files or stored on kindles or other portable readers.
- Location:
 - The Adventures of Tom Sawyer: <https://www.gutenberg.org/ebooks/74>
 - Adventures of Huckleberry Finn: <https://www.gutenberg.org/ebooks/76>
 - The Innocents Abroad: <https://www.gutenberg.org/ebooks/3176>
 - A Tramp Abroad: <https://www.gutenberg.org/ebooks/119>
 - A Tale of Two Cities: <https://www.gutenberg.org/ebooks/98>
 - Great Expectations: <https://www.gutenberg.org/ebooks/1400>
 - Oliver Twist: <https://www.gutenberg.org/ebooks/730>
 - Hard Times: <https://www.gutenberg.org/ebooks/786>
- Description:
 - Corpus consists of texts from two of the greatest authors from the early 1800s, Mark Twain and Charles Dickens.
- ▪ Format:
 - All eight texts are in .txt format