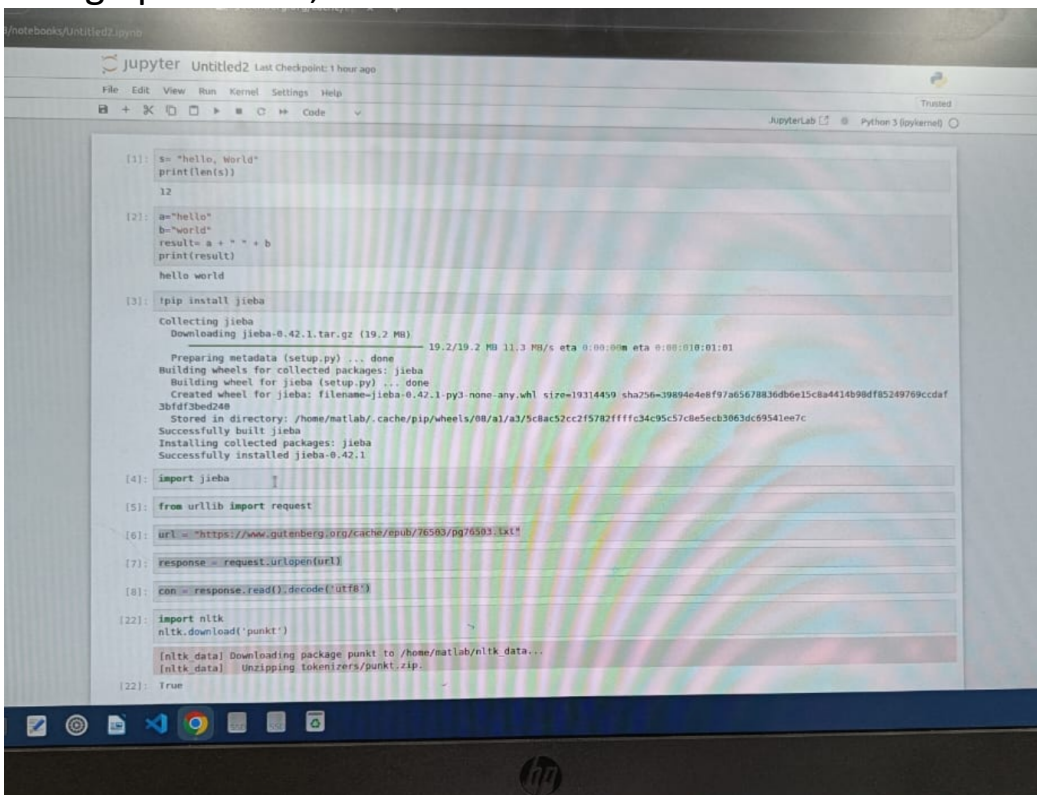
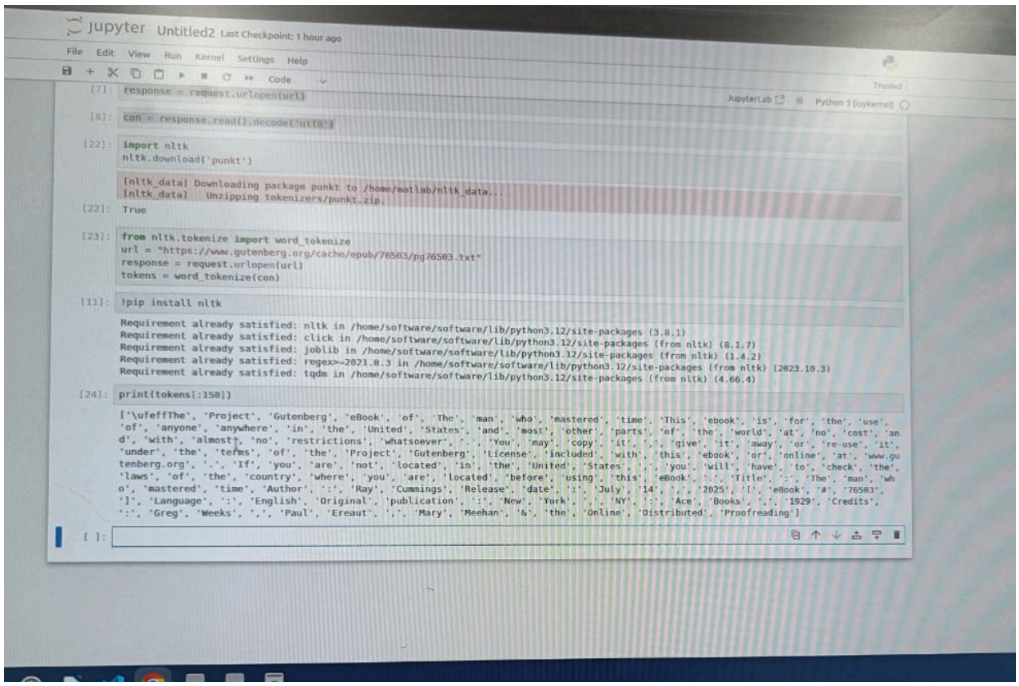
Lab 2

String operations ,text extraction





Documentation:-

Source of the Website

The data is taken from Project Gutenberg (https://www.gutenberg.org), a well-known website that provides free access to public domain books. Specifically, I used this link: https://www.gutenberg.org/cache/epub/76503/pg76503.txt,
which hosts the plain text version of the book *"The Man Who Mastered Time."*

What I'm Doing with the Book and Why

I'm taking the text from the book so I can practice programming. Basically, I want to:

- Learn how language is structured and how to break it down.
- Split the book into words/tokens
- See how computers read and work with text,

I'm not changing, copying, or selling the book—it's just for my own learning and understanding.

Copyrights and licenses

Most permissions not needed

Most permission requests we receive do not require a custom response. The vast majority of Project Gutenberg eBooks are in the public domain in the US. This means that nobody can grant, or withhold, permission to do with this item as you please.

"As you please" includes any commercial use, republishing in any format, making derivative works or performances, etc. Read more about the public domain in Wikipedia.

Linking

No permission is needed to link to www.gutenberg.org or to any page or address within it.

To link to a specific eBook, link to the landing page for that eBook, not to specific files. So, for example, you would link to www.gutenberg.org/ebooks/19033 rather than "deep linking" to the specific HTML, plain text, etc. www.gutenberg.org prohibits deep links automatically (based on the HTTP referrer), because from time to time the underlying structure of an eBook might change.