

ADVANCE PYTHON – LAB ASSIGNMENT

STRING OPERATIONS IN PYTHON

22MID0136

DHEVISHREE GN

15-07-25

The screenshot shows a JupyterLab window titled 'Untitled2' running on a local host. The notebook contains a text snippet about Hawaiian history, followed by a code cell that prints the length of the text (177843). The next cell shows the installation of the 'nltk' library. The final cell attempts to import 'word_tokenize' from 'nltk.tokenize' and fetch text from a URL, but it results in a 'SyntaxError: invalid syntax' due to a missing closing quote in the URL string.

```
who have studied the matter will be able to appreciate. Here it will suffice to say that of his race there was none like him before, there has been none like him since. In all that shadowy time from the dawn of Hawaiian history to the establishment of intercourse with the western world, the time of heroes eight or nine feet high, who wielded spears ten yards long; heroes who fought with gods and received aid from gods, as the Greek warriors at Troy from Minerva and Apollo--heroes like Kiha of the magic conch, like Liloa and Umi and Iono, there was none who accomplished what Kanehameha did by the patient toil and dauntless courage of forty years.
```

```
[14]: print(len(raw))
177843
```

```
[15]: !pip install nltk

Requirement already satisfied: nltk in /home/software/software/lib/python3.12/site-packages (3.8.1)
Requirement already satisfied: click in /home/software/software/lib/python3.12/site-packages (from nltk) (8.1.7)
Requirement already satisfied: joblib in /home/software/software/lib/python3.12/site-packages (from nltk) (1.4.2)
Requirement already satisfied: regex==2021.8.3 in /home/software/software/lib/python3.12/site-packages (from nltk) (2023.10.3)
Requirement already satisfied: tqdm in /home/software/software/lib/python3.12/site-packages (from nltk) (4.66.4)
```

```
[16]: import nltk
from nltk.tokenize import word_tokenize url="https://www.gutenberg.org/cache/epub/76499/pg76499.txt"
url="https://www.gutenberg.org/cache/epub/76499/pg76499.txt"
response=request.urlopen(url)
response=request.urlopen(url)
tokens=word_tokenize(raw)

Cell [16], line 2
from nltk.tokenize import word_tokenize url="https://www.gutenberg.org/cache/epub/76499/pg76499.txt"
SyntaxError: invalid syntax
```



