



Session 29

Assignment 1 Question

Session 29: Assignment 1

Table of Contents

1. Introduction

2. Problem Statement

3. Output

1. Introduction

This assignment will help you to consolidate the concepts learnt in the session.

2. Problem Statement

In this assignment students have to find the frequency of words in a webpage. User can use urllib and BeautifulSoup to extract text from webpage.

Hint:

```
from bs4 import BeautifulSoup
```

```
import urllib.request
```

```
import nltk
```

```
response = urllib.request.urlopen('http://php.net/')
```

```
html = response.read()
```

```
soup = BeautifulSoup(html,"html5lib")
```

NOTE: The solution shared through Github should contain the source code used and the screenshot of the output.

3. Output

N/A