```
In [1]:  text = "My name is harshita khangarot, I m asst. Prof. in JIET-jodhpur. I have 7 Ye
```

```
In [2]:  # Splits at space
         print(text.split())
```

```
['My', 'name', 'is', 'harshita', 'khangarot,', 'I', 'm', 'asst.', 'Prof.', 'in',
'JIET-jodhpur.', 'I', 'have', '7', 'Years', 'of', 'Exp,I', 'love', 'Jodhpur', 'an
d', "Jodhpur's", 'sweets.']
```

```
In [3]:  # Splits at '.'
         print(text.split('.'))
```

```
['My name is harshita khangarot, I m asst', ' Prof', ' in JIET-jodhpur', " I have
7 Years of Exp,I love Jodhpur and Jodhpur's sweets", '']
```

```
In [4]:  import nltk
         nltk.download('all')
```

```
[nltk_data] Downloading collection 'all'
[nltk_data]    |
[nltk_data]    | Downloading package abc to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package abc is already up-to-date!
[nltk_data]    | Downloading package alpino to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package alpino is already up-to-date!
[nltk_data]    | Downloading package averaged_perceptron_tagger to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package averaged_perceptron_tagger is already up-
[nltk_data]    |     to-date!
[nltk_data]    | Downloading package averaged_perceptron_tagger_ru to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package averaged_perceptron_tagger_ru is already
[nltk_data]    |     up-to-date!
[nltk_data]    | Downloading package basque_grammars to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package basque_grammars is already up-to-date!
[nltk_data]    | Downloading package bcp47 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package bcp47 is already up-to-date!
[nltk_data]    | Downloading package biocreative_ppi to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package biocreative_ppi is already up-to-date!
[nltk_data]    | Downloading package bllip_wsj_no_aux to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package bllip_wsj_no_aux is already up-to-date!
[nltk_data]    | Downloading package book_grammars to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package book_grammars is already up-to-date!
[nltk_data]    | Downloading package brown to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package brown is already up-to-date!
[nltk_data]    | Downloading package brown_tei to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package brown_tei is already up-to-date!
[nltk_data]    | Downloading package cess_cat to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package cess_cat is already up-to-date!
[nltk_data]    | Downloading package cess_esp to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package cess_esp is already up-to-date!
[nltk_data]    | Downloading package chat80 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package chat80 is already up-to-date!
[nltk_data]    | Downloading package city_database to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package city_database is already up-to-date!
[nltk_data]    | Downloading package cmudict to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package cmudict is already up-to-date!
[nltk_data]    | Downloading package comparative_sentences to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package comparative_sentences is already up-to-
[nltk_data]    |     date!
[nltk_data]    | Downloading package comtrans to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
```

```
[nltk_data]    |    Package comtrans is already up-to-date!
[nltk_data]    | Downloading package conll2000 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package conll2000 is already up-to-date!
[nltk_data]    | Downloading package conll2002 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package conll2002 is already up-to-date!
[nltk_data]    | Downloading package conll2007 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package conll2007 is already up-to-date!
[nltk_data]    | Downloading package crubadan to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package crubadan is already up-to-date!
[nltk_data]    | Downloading package dependency_treebank to
[nltk_data]    |      C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package dependency_treebank is already up-to-date!
[nltk_data]    | Downloading package dolch to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package dolch is already up-to-date!
[nltk_data]    | Downloading package europarl_raw to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package europarl_raw is already up-to-date!
[nltk_data]    | Downloading package extended_omw to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package extended_omw is already up-to-date!
[nltk_data]    | Downloading package floresta to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package floresta is already up-to-date!
[nltk_data]    | Downloading package framenet_v15 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package framenet_v15 is already up-to-date!
[nltk_data]    | Downloading package framenet_v17 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package framenet_v17 is already up-to-date!
[nltk_data]    | Downloading package gazetteers to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package gazetteers is already up-to-date!
[nltk_data]    | Downloading package genesis to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package genesis is already up-to-date!
[nltk_data]    | Downloading package gutenberg to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package gutenberg is already up-to-date!
[nltk_data]    | Downloading package ieer to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package ieer is already up-to-date!
[nltk_data]    | Downloading package inaugural to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package inaugural is already up-to-date!
[nltk_data]    | Downloading package indian to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package indian is already up-to-date!
[nltk_data]    | Downloading package jeita to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package jeita is already up-to-date!
[nltk_data]    | Downloading package kimmo to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package kimmo is already up-to-date!
[nltk_data]    | Downloading package knbc to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |    Package knbc is already up-to-date!
[nltk_data]    | Downloading package large_grammars to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
```

```
[nltk_data]    |   Package large_grammars is already up-to-date!
[nltk_data]    | Downloading package lin_thesaurus to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package lin_thesaurus is already up-to-date!
[nltk_data]    | Downloading package mac_morpho to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package mac_morpho is already up-to-date!
[nltk_data]    | Downloading package machado to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package machado is already up-to-date!
[nltk_data]    | Downloading package masc_tagged to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package masc_tagged is already up-to-date!
[nltk_data]    | Downloading package maxent_ne_chunker to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package maxent_ne_chunker is already up-to-date!
[nltk_data]    | Downloading package maxent_treebank_pos_tagger to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package maxent_treebank_pos_tagger is already up-
[nltk_data]    |       to-date!
[nltk_data]    | Downloading package moses_sample to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package moses_sample is already up-to-date!
[nltk_data]    | Downloading package movie_reviews to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package movie_reviews is already up-to-date!
[nltk_data]    | Downloading package mte_teip5 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package mte_teip5 is already up-to-date!
[nltk_data]    | Downloading package mwa_ppdb to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package mwa_ppdb is already up-to-date!
[nltk_data]    | Downloading package names to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package names is already up-to-date!
[nltk_data]    | Downloading package nombank.1.0 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package nombank.1.0 is already up-to-date!
[nltk_data]    | Downloading package nonbreaking_prefixes to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package nonbreaking_prefixes is already up-to-date!
[nltk_data]    | Downloading package nps_chat to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package nps_chat is already up-to-date!
[nltk_data]    | Downloading package omw to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package omw is already up-to-date!
[nltk_data]    | Downloading package omw-1.4 to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package omw-1.4 is already up-to-date!
[nltk_data]    | Downloading package opinion_lexicon to
[nltk_data]    |     C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package opinion_lexicon is already up-to-date!
[nltk_data]    | Downloading package panlex_swadesh to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package panlex_swadesh is already up-to-date!
[nltk_data]    | Downloading package paradigms to C:\Users\ansuya
[nltk_data]    |     bohra\AppData\Roaming\nltk_data...
[nltk_data]    |   Package paradigms is already up-to-date!
[nltk_data]    | Downloading package pe08 to C:\Users\ansuya
```

```
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package pe08 is already up-to-date!
[nltk_data]    | Downloading package perluniprops to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package perluniprops is already up-to-date!
[nltk_data]    | Downloading package pil to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package pil is already up-to-date!
[nltk_data]    | Downloading package pl196x to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package pl196x is already up-to-date!
[nltk_data]    | Downloading package porter_test to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package porter_test is already up-to-date!
[nltk_data]    | Downloading package ppattach to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package ppattach is already up-to-date!
[nltk_data]    | Downloading package problem_reports to
[nltk_data]    |       C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package problem_reports is already up-to-date!
[nltk_data]    | Downloading package product_reviews_1 to
[nltk_data]    |       C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package product_reviews_1 is already up-to-date!
[nltk_data]    | Downloading package product_reviews_2 to
[nltk_data]    |       C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package product_reviews_2 is already up-to-date!
[nltk_data]    | Downloading package propbank to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package propbank is already up-to-date!
[nltk_data]    | Downloading package pros_cons to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package pros_cons is already up-to-date!
[nltk_data]    | Downloading package ptb to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package ptb is already up-to-date!
[nltk_data]    | Downloading package punkt to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package punkt is already up-to-date!
[nltk_data]    | Downloading package qc to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package qc is already up-to-date!
[nltk_data]    | Downloading package reuters to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package reuters is already up-to-date!
[nltk_data]    | Downloading package rslp to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package rslp is already up-to-date!
[nltk_data]    | Downloading package rte to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package rte is already up-to-date!
[nltk_data]    | Downloading package sample_grammars to
[nltk_data]    |       C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package sample_grammars is already up-to-date!
[nltk_data]    | Downloading package semcor to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package semcor is already up-to-date!
[nltk_data]    | Downloading package senseval to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package senseval is already up-to-date!
[nltk_data]    | Downloading package sentence_polarity to
```

```
[nltk_data]    |        C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package sentence_polarity is already up-to-date!
[nltk_data]    | Downloading package sentiwordnet to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package sentiwordnet is already up-to-date!
[nltk_data]    | Downloading package shakespeare to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package shakespeare is already up-to-date!
[nltk_data]    | Downloading package sinica_treebank to
[nltk_data]    |        C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package sinica_treebank is already up-to-date!
[nltk_data]    | Downloading package smultron to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package smultron is already up-to-date!
[nltk_data]    | Downloading package snowball_data to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package snowball_data is already up-to-date!
[nltk_data]    | Downloading package spanish_grammars to
[nltk_data]    |        C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package spanish_grammars is already up-to-date!
[nltk_data]    | Downloading package state_union to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package state_union is already up-to-date!
[nltk_data]    | Downloading package stopwords to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package stopwords is already up-to-date!
[nltk_data]    | Downloading package subjectivity to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package subjectivity is already up-to-date!
[nltk_data]    | Downloading package swadesh to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package swadesh is already up-to-date!
[nltk_data]    | Downloading package switchboard to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package switchboard is already up-to-date!
[nltk_data]    | Downloading package tagsets to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package tagsets is already up-to-date!
[nltk_data]    | Downloading package timit to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package timit is already up-to-date!
[nltk_data]    | Downloading package toolbox to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package toolbox is already up-to-date!
[nltk_data]    | Downloading package treebank to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package treebank is already up-to-date!
[nltk_data]    | Downloading package twitter_samples to
[nltk_data]    |        C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package twitter_samples is already up-to-date!
[nltk_data]    | Downloading package udhr to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package udhr is already up-to-date!
[nltk_data]    | Downloading package udhr2 to C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package udhr2 is already up-to-date!
[nltk_data]    | Downloading package unicode_samples to
[nltk_data]    |        C:\Users\ansuya
[nltk_data]    |        bohra\AppData\Roaming\nltk_data...
[nltk_data]    |      Package unicode_samples is already up-to-date!
```

```
[nltk_data]    |   Downloading package universal_tagset to
[nltk_data]    |       C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package universal_tagset is already up-to-date!
[nltk_data]    |   Downloading package universal_treebanks_v20 to
[nltk_data]    |       C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package universal_treebanks_v20 is already up-to-
[nltk_data]    |         date!
[nltk_data]    |   Downloading package vader_lexicon to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package vader_lexicon is already up-to-date!
[nltk_data]    |   Downloading package verbnet to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package verbnet is already up-to-date!
[nltk_data]    |   Downloading package verbnet3 to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package verbnet3 is already up-to-date!
[nltk_data]    |   Downloading package webtext to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package webtext is already up-to-date!
[nltk_data]    |   Downloading package wmt15_eval to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package wmt15_eval is already up-to-date!
[nltk_data]    |   Downloading package word2vec_sample to
[nltk_data]    |       C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package word2vec_sample is already up-to-date!
[nltk_data]    |   Downloading package wordnet to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package wordnet is already up-to-date!
[nltk_data]    |   Downloading package wordnet2021 to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package wordnet2021 is already up-to-date!
[nltk_data]    |   Downloading package wordnet2022 to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package wordnet2022 is already up-to-date!
[nltk_data]    |   Downloading package wordnet31 to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package wordnet31 is already up-to-date!
[nltk_data]    |   Downloading package wordnet_ic to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package wordnet_ic is already up-to-date!
[nltk_data]    |   Downloading package words to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package words is already up-to-date!
[nltk_data]    |   Downloading package ycoe to C:\Users\ansuya
[nltk_data]    |       bohra\AppData\Roaming\nltk_data...
[nltk_data]    |     Package ycoe is already up-to-date!
[nltk_data]    |
[nltk_data]  Done downloading collection all
```

Out[4]:   True

In [5]:
```python
from nltk.tokenize import word_tokenize
word_tokenize(text)
```

Out[5]:
```
['My',
 'name',
 'is',
 'harshita',
 'khangarot',
 ',',
 'I',
 'm',
 'asst',
 '.',
 'Prof.',
 'in',
 'JIET-jodhpur',
 '.',
 'I',
 'have',
 '7',
 'Years',
 'of',
 'Exp',
 ',',
 'I',
 'love',
 'Jodhpur',
 'and',
 'Jodhpur',
 "'s",
 'sweets',
 '.']
```

In [6]:
```python
from nltk.tokenize import sent_tokenize
sent_tokenize(text)
```

Out[6]:
```
['My name is harshita khangarot, I m asst.',
 'Prof. in JIET-jodhpur.',
 "I have 7 Years of Exp,I love Jodhpur and Jodhpur's sweets."]
```

In [7]:
```python
from nltk.tokenize import TreebankWordTokenizer
tokenizer = TreebankWordTokenizer()
tokenizer.tokenize(text)
```

```
Out[7]:  ['My',
          'name',
          'is',
          'harshita',
          'khangarot',
          ',',
          'I',
          'm',
          'asst.',
          'Prof.',
          'in',
          'JIET-jodhpur.',
          'I',
          'have',
          '7',
          'Years',
          'of',
          'Exp',
          ',',
          'I',
          'love',
          'Jodhpur',
          'and',
          'Jodhpur',
          "'s",
          'sweets',
          '.']
```

```
In [8]:  from nltk.tokenize import WordPunctTokenizer
         tokenizer = WordPunctTokenizer()
         tokenizer.tokenize(text)
```

Out[8]:
```
['My',
 'name',
 'is',
 'harshita',
 'khangarot',
 ',',
 'I',
 'm',
 'asst',
 '.',
 'Prof',
 '.',
 'in',
 'JIET',
 '-',
 'jodhpur',
 '.',
 'I',
 'have',
 '7',
 'Years',
 'of',
 'Exp',
 ',',
 'I',
 'love',
 'Jodhpur',
 'and',
 'Jodhpur',
 "'",
 's',
 'sweets',
 '.']
```

In [10]:
```python
import nltk
file_content = open("abc.txt").read()
tokens = nltk.word_tokenize(file_content)
print(tokens)
```

```
['My', 'name', 'is', 'harshita', 'khangarot', ',', 'I', 'm', 'asst', '.', 'Prof.',
'in', 'JIET-jodhpur', '.', 'I', 'have', '7', 'Years', 'of', 'Exp', ',', 'I', 'lov
e', 'Jodhpur', 'and', 'Jodhpur', "'s", 'sweets', '.']
```

In [11]:
```python
#Using Regular Expression
from nltk.tokenize import RegexpTokenizer
```

In [12]:
```python
tokenizer = RegexpTokenizer("[\w']+")
text = "Let's see how it's working."
tokenizer.tokenize(text)
```

Out[12]:
```
["Let's", 'see', 'how', "it's", 'working']
```

In [13]:
```python
import re
my_string = "Let's write RegEx!"
PATTERN = r"\w+"
re.findall(PATTERN, my_string)
```

Out[13]:
```
['Let', 's', 'write', 'RegEx']
```

In [14]:
```python
my_string = "Let's write RegEx! Won't that be fun? I sure think so. Can you find 4
```

In [15]:
```python
# Write a pattern to match sentence endings: sentence_endings
sentence_endings = r"[.?!]"
```

```python
# Split my_string on sentence endings and print the result
print(re.split(sentence_endings, my_string))
```

```
["Let's write RegEx", " Won't that be fun", ' I sure think so', ' Can you find 4 s
entences']
```

In [16]:
```python
# Find all capicalized words in my_string and print the result
capitalized_words = r"[A-Z]\w+"
print(re.findall(capitalized_words, my_string))
```

```
['Let', 'RegEx', 'Won', 'Can']
```

In [17]:
```python
# Split my_string on spaces and print the result
spaces = r"\s+"
print(re.split(spaces, my_string))
```

```
["Let's", 'write', 'RegEx!', "Won't", 'that', 'be', 'fun?', 'I', 'sure', 'think',
'so.', 'Can', 'you', 'find', '4', 'sentences']
```

In [18]:
```python
# Find all digits in my_string and print the result
digits = r"\d+"
print(re.findall(digits, my_string))
```

```
['4']
```

In [ ]:

# Split my_string on sentence endings and print the result
print(re.split(sentence_endings, my_string))

# Find all capicalized words in my_string and print the result
capitalized_words = r"[A-Z]\w+"
print(re.findall(capitalized_words, my_string))