

```
In [28]: sentence = "Natural language processing makes it possible for computers to understand the human language. In natural language processing, human language is separated into fragments so that the grammatical structure of sentences and the meaning of words can be analysed and understood in context. This helps computers read and understand spoken or written text in the same way as humans. I am studying Natural Language Processing at Amrita University."
```

```
In [29]: print(sentence)
```

Natural language processing makes it possible for computers to understand the human language. In natural language processing, human language is separated into fragments so that the grammatical structure of sentences and the meaning of words can be analysed and understood in context. This helps computers read and understand spoken or written text in the same way as humans. I am studying Natural Language Processing at Amrita University.

```
In [30]: if(sentence.find("language") >= 0):
          print("Word language present in the sentence")
        else:
          print("Word language not present in the sentence")
```

Word language present in the sentence

```
In [31]: ind = sentence.index("human")
          print("Index value of the word human is:", ind)
```

Index value of the word human is: 78

```
In [32]: list_words = sentence.lower().split(" ")
          flag = False
          for i in range(len(list_words)):
              if(list_words[i] == "possible"):
                  print("Index value of the word possible is:", i)
                  flag = True
          if(not flag):
              print("Word possible not present in the sentence")
```

Index value of the word possible is: 5

```
In [33]: print("Third word of the given text is:", list_words[3-1])
```

Third word of the given text is: processing

```
In [34]: print("Number of Lines in the given text is:", len(sentence[:-1].split(".")))
          print("Number of Words in the given text is:", len(list_words))
          print("Number of Characters in the given text is:", len(sentence)-sentence.count(" "))
```

Number of Lines in the given text is: 4
Number of Words in the given text is: 67
Number of Characters in the given text is: 372

```
In [35]: vocabulary = set(list_words)
         print(len(vocabulary))
```

50

```
In [36]: freq = []
         lst_voc = list(vocabulary)
         for i in range(len(lst_voc)):
             c = list_words.count(lst_voc[i])
             freq.append(c)
         z=zip(lst_voc, freq)
         dict_voc = dict(z)
         from pprint import pprint
         pprint(dict_voc)
```

```
{'am': 1,
  'amrita': 1,
  'analysed': 1,
  'and': 3,
  'as': 1,
  'at': 1,
  'be': 1,
  'can': 1,
  'computers': 2,
  'context.': 1,
  'for': 1,
  'fragments': 1,
  'grammatical': 1,
  'helps': 1,
  'human': 2,
  'humans.': 1,
  'i': 1,
  'in': 3,
  'into': 1,
  'is': 1,
  'it': 1,
  'language': 4,
  'language.': 1,
  'makes': 1,
  'meaning': 1,
  'natural': 3,
  'of': 2,
  'or': 1,
  'possible': 1,
  'processing': 2,
  'processing,': 1,
  'read': 1,
  'same': 1,
  'sentences': 1,
  'separated': 1,
  'so': 1,
  'spoken': 1,
  'structure': 1,
  'studying': 1,
  'text': 1,
  'that': 1,
  'the': 4,
  'this': 1,
  'to': 1,
  'understand': 2,
  'understood': 1,
  'university.': 1,
  'way': 1,
  'words': 1,
  'written': 1}
```