Question-2

- **Assumptions:**
  - No spelling errors in the query.
  - No punctuation marks will be present in the query
  - No single characters present in the query.

- **Preprocessing of Data:**
  - Text of the file converted to lower case.
  - Metadata removal.
  - Removal of punctuation marks.
  - Tokenization of text stream.

- **Preprocessing of Query:**
  - Text converted to lower case.
  - Tokenization of query text.

- **Methodology:**
  - After the preprocessing of data, positional index was created using the standard steps.
    - It has been stored as a dictionary.
  - The query text was then tokenized and the documents containing all the words of the query were then retrieved.
  - Positions of the words were searched in these documents then according to the position in the query through the maintained index.