Assignment 1 Outputs

PFA the Screenshots in the zipped file also.

Task 1

```
Enter number from 1 to 9 for the corresponding question :
1) Print the number of words and sentences contained in the file given as input.
2) Print the number of words starting with consonants and vowels
3) List all the email ids in the file given as input.
4) Print the sentences and number of sentences starting with a given word in an input file.
5) Print the sentences and number of sentences ending with a given word in an input file.
6) Given a word and a file as input, print the count of that word and sentences containing that word in the input
file.
7) Given an input file, print the questions present, if any, in that file.
8) List the minutes and seconds mentioned in the date present in the file given as input.
9) List the abbreviations present in a file given as input.
10) Exit

 Enter the task number :
1

 Please Enter the File Name :
58055
Lines in Metadata:  8
Number of sentences in text(without metadata):  9
Number of words in text(without metadata):  132
```

Task 2

```
Enter number from 1 to 9 for the corresponding question :
1) Print the number of words and sentences contained in the file given as input.
2) Print the number of words starting with consonants and vowels
3) List all the email ids in the file given as input.
4) Print the sentences and number of sentences starting with a given word in an input file.
5) Print the sentences and number of sentences ending with a given word in an input file.
6) Given a word and a file as input, print the count of that word and sentences containing that word
file.
7) Given an input file, print the questions present, if any, in that file.
8) List the minutes and seconds mentioned in the date present in the file given as input.
9) List the abbreviations present in a file given as input.
10) Exit

 Enter the task number :
2

 Please Enter the File Name :
58045
Number of words starting with vowels :220
Number of words starting with consonants :495
```

Task 3

```
    Enter the task number :
    3

    Please Enter the File Name :
    58045

    All the emails present in the document are:

    fulk@cs.rochester.edu
    1993Apr5.184453.8394@cs.rochester.edu
    TIGGER.93Apr2110251@satyr.Sylvan.COM
    1993Apr3.161757.19612@cs.rochester.edu
    1993Apr5.151818.27409@trentu.ca
    1993Apr5.151818.27409@trentu.ca
    xtkmg@trentu.ca
    1993Apr3.161757.19612@cs.rochester.edu
    fulk@cs.rochester.edu
    fulk@cs.rochester.edu

    All the unique emails present in the document are:

    1993Apr3.161757.19612@cs.rochester.edu
    TIGGER.93Apr2110251@satyr.Sylvan.COM
    1993Apr5.151818.27409@trentu.ca
    xtkmg@trentu.ca
    1993Apr5.184453.8394@cs.rochester.edu
    fulk@cs.rochester.edu
```

Task 4

```
    Enter the task number :
    4

    Please Enter the File Name :
    58045

    Please enter the word you want to search :
    IN
    Input word does not represent a number

    Number of sentences starting with in :2
    The sentence are as follows:


    1: In article <1993Apr5.151818.27409@trentu.ca> xtkmg@trentu.ca (Kate Gregory) writes:
    >In article <1993Apr3.161757.19612@cs.rochester.edu> fulk@cs.rochester.edu (Mark Fulk) writes:
    >>
    >>Another uncommon problem is maternal hemorrhage.

    2: In view of the intensity of the birth experience, I doubt surroundings
    have much importance anyway.
```

Task 5

```
 Enter the task number :
5

 Please Enter the File Name :
58045

 Please enter the word you want to search :
time
Input word does not represent a number

 Number of sentences starting with time :3
 The sentence are as follows:


1: It is hard
>>to see how you could handle it at home, and you wouldn't have very much time.

2: Domestic horses and cows typically have a worse time.

3: I'm planning a hospital birth
>this time.
```

Task 6

Finding a word that does not represent a number

```
 Enter the task number :
6

 Please Enter the File Name :
105125

 Please enter the word you want to search :
the
Note : Input word does not represent a number

****    Output When Stemming was not used       ***


 Number of occurence of input word :13
Number of sentence which contain input word :7
 The sentence are as follows:
```

```
    1: The package includes:

            Frame, swingarm and Earles fork
            Final drive(27/8)
            Rear fender, front fender, large tank
            Steel wheels laced w/ ss spokes, one new rim, one EC used rim
            All brackets, brake hdware, footrests etc.

    2: The colour scheme used is the original BMW optional Bristol grey
            Items not painted are: Earles rigid section(needs straightening,
                    this rigid section was bent when I received it, not from
                    being mounted to the R60US frame),Earles swingarm and the
                    headlight shell.

    3: I have the paintthat was used and can
                    finish the non-painted items if desired.
```

```
    4: The striping
                 was done by Holt Design.

    5: The project was terminated due to lack of
            time and the R69(35hp) engine that was to be used was transferred
            to another project.

    6: I would prefer to sell the package as whole.

    7: I am negotiable as to what can be included in the package, I will
            not separate any of the painted items.
```

The below part will also be printed if comments are removed from the stemmer part output in ques6( )
[Currently, in code it won't be printed as it is commented out]

```
****     Output when Applying porter stemmer on both input file in preproceesing and word      ***


 Number of occurence of input word(after stemming) :13
Number of sentence which contain input word :7
 The sentence are as follows:


1: The package includes:

        Frame, swingarm and Earles fork
        Final drive(27/8)
        Rear fender, front fender, large tank
        Steel wheels laced w/ ss spokes, one new rim, one EC used rim
        All brackets, brake hdware, footrests etc.

2: The colour scheme used is the original BMW optional Bristol grey
        Items not painted are: Earles rigid section(needs straightening,
               this rigid section was bent when I received it, not from
               being mounted to the R60US frame) Earles swingarm and the
```

```
                 being mounted to the R60US frame),Earles swingarm and the
                 headlight shell.

    3: I have the paintthat was used and can
                 finish the non-painted items if desired.

    4: The striping
                 was done by Holt Design.

    5: The project was terminated due to lack of
            time and the R69(35hp) engine that was to be used was transferred
            to another project.

    6: I would prefer to sell the package as whole.

    7: I am negotiable as to what can be included in the package, I will
            not separate any of the painted items.
```

Finding a word that represents a number (Implemented in Task 4 and 5 also)

```
 Enter the task number :
 6

 Please Enter the File Name :
 58913

 Please enter the word you want to search :
 fiVe
Note : Input word represents the number :5

****    Output Without using Stemming    ***


 Number of occurence of input word :2
Number of sentence which contain input word :2
 The sentence are as follows:


1: 2)  About 5 or 10 years ago, I remember hearing that carmel color
       (obtained by caramelizing sugar -- a common coloring and flavoring
       agent) had been found to be carcinogenic.

2: 3)  About 5 or 10 years ago, I remember Earl Grey tea being implicated
       as carcinogenic, because it contains oil of bergamot (an extract
       from the skin of a type of citrus fruit).
```

Task 7

```
 Enter the task number :
 7

 Please Enter the File Name :
 105117

 Number of questions in input file: 2
 The questions are as follows:


1: But did the Hurt
> study make any distinction between an *ability* to swerve and a *failure*
> to swerve?

2: Hard braking and swerving tend to be mutually exclusive
> manouvres - did Hurt draw any conclusions on which one is generally preferable?
```

Task 8

```
 Enter the task number :
8

 Please Enter the File Name :
105117

 The outputs are as follows:
50 Minutes 05 seconds
```

Task 9

```
 Enter the task number :
9

 Please Enter the File Name :
58915
The abbreviations are:
Dr.
SLMR
```