Assignment 6
Exercise 1: Unix Text Analysis

1. Setting up the Virtual Machine on VMWare Fusion, with Centos7 operating system.

2. Completed downloading essential files including "The Raven.txt" file.

```
Resolving ia801509.us.archive.org (ia801509.us.archive.org)... 207.241.228.159
Connecting to ia801509.us.archive.org (ia801509.us.archive.org)|207.241.228.159|:443... connected.
HTTP request sent, awaiting response... 301 Moved Permanently
Location: https://archive.org/download/TheRaven_201601/The%20Raven.txt [following]
--2018-12-09 02:24:09--  https://archive.org/download/TheRaven_201601/The%20Raven.txt
Resolving archive.org (archive.org)... 207.241.224.2
Connecting to archive.org (archive.org)|207.241.224.2|:443... connected.
HTTP request sent, awaiting response... 302 Found
Location: https://ia801001.us.archive.org/18/items/TheRaven_201601/The%20Raven.txt [following]
--2018-12-09 02:24:10--  https://ia801001.us.archive.org/18/items/TheRaven_201601/The%20Raven.txt
Resolving ia801001.us.archive.org (ia801001.us.archive.org)... 207.241.228.81
Connecting to ia801001.us.archive.org (ia801001.us.archive.org)|207.241.228.81|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 6772 (6.6K) [text/plain]
Saving to: 'The Raven.txt'

100%[=================================================================>] 6,772       --.-K/s   in 0s

2018-12-09 02:24:10 (652 MB/s) - 'The Raven.txt' saved [6772/6772]

[root@localhost ~]# ls
anaconda-ks.cfg  pdftk-2.02-1.e16.x86_64.rpm  The Raven.txt  The Raven.txt.1
```

3. Using the *head* and *tail* commands to view the content of the text file

```
[root@localhost ~]# head 'The Raven.txt'

Once upon a midnight dreary, while I pondered, weak and weary,
Over many a quaint and curious volume of forgotten lore?
    While I nodded, nearly napping, suddenly there came a tapping,
As of some one gently rapping, rapping at my chamber door.
??Tis some visitor,? I muttered, ?tapping at my chamber door?
            Only this and nothing more.?

    Ah, distinctly I remember it was in the bleak December;
And each separate dying ember wrought its ghost upon the floor.
[root@localhost ~]# tail 'The Raven.txt'
    Leave my loneliness unbroken!?quit the bust above my door!
Take thy beak from out my heart, and take thy form from off my door!?
            Quoth the Raven ?Nevermore.?

    And the Raven, never flitting, still is sitting, still is sitting
On the pallid bust of Pallas just above my chamber door;
    And his eyes have all the seeming of a demon?s that is dreaming,
    And the lamp-light o?er him streaming throws his shadow on the floor;
And my soul from out that shadow that lies floating on the floor
            Shall be lifted?nevermore![root@localhost ~]#
```

4. Using the *less –N* command to view the content of the file in a pop-up-window like manner.

```
Once upon a midnight dreary, while I pondered, weak and weary,
Over many a quaint and curious volume of forgotten lore<97>
    While I nodded, nearly napping, suddenly there came a tapping,
As of some one gently rapping, rapping at my chamber door.
<93><92>Tis some visitor,<94> I muttered, <93>tapping at my chamber door<97>
        Only this and nothing more.<94>

    Ah, distinctly I remember it was in the bleak December;
And each separate dying ember wrought its ghost upon the floor.
    Eagerly I wished the morrow;<97>vainly I had sought to borrow
    From my books surcease of sorrow<97>sorrow for the lost Lenore<97>
For the rare and radiant maiden whom the angels name Lenore<97>
        Nameless here for evermore.

    And the silken, sad, uncertain rustling of each purple curtain
Thrilled me<97>filled me with fantastic terrors never felt before;
    So that now, to still the beating of my heart, I stood repeating
    <93><92>Tis some visitor entreating entrance at my chamber door<97>
Some late visitor entreating entrance at my chamber door;<97>
        This it is and nothing more.<94>

    Presently my soul grew stronger; hesitating then no longer,
<93>Sir,<94> said I, <93>or Madam, truly your forgiveness I implore;
    But the fact is I was napping, and so gently you came rapping,
    And so faintly you came tapping, tapping at my chamber door,
That I scarce was sure I heard you<94><97>here I opened wide the door;<97>
        Darkness there and nothing more.

    Deep into that darkness peering, long I stood there wondering, fearing,
Doubting, dreaming dreams no mortal ever dared to dream before;
The\ Raven.txt
```

5. "The Raven.txt" file was giving me difficulty due to its space. Therefore, I copied it to "theraven_orig.txt" and continued to work on this new file.
I performed some deleting of lines using the *sed* command as well.

6. Word counts using *wc –l* and *wc –m* commands.

```
[root@localhost ~]# wc -l theraven_orig.txt
125 theraven_orig.txt
[root@localhost ~]# wc -m theraven_orig.txt
6657 theraven_orig.txt
```

7. Searing for words/phrases using *grep* and using *–n* commands to display results with line number.

```
[root@localhost ~]# grep -n "raven" theraven_orig.txt
53:?Though thy crest be shorn and shaven, thou,? I said, ?art sure no craven,
```

8. Searching for words/phrases without case sensitivity using the *grep –E –n* command.

```
[root@localhost ~]# grep -E -n "(R|r)aven" theraven_orig.txt
45:In there stepped a stately Raven of the saintly days of yore;
53:?Though thy crest be shorn and shaven, thou,? I said, ?art sure no craven,
54:Ghastly grim and ancient Raven wandering from the Nightly shore?
56:        Quoth the Raven ?Nevermore.?
65:    But the Raven, sitting lonely on the placid bust, spoke only
79:    But the Raven still beguiling all my fancy into smiling,
98:        Quoth the Raven ?Nevermore.?
105:        Quoth the Raven ?Nevermore.?
112:        Quoth the Raven ?Nevermore.?
119:        Quoth the Raven ?Nevermore.?
121:    And the Raven, never flitting, still is sitting, still is sitting
```

9. Deleting, finding and replacing of characters from the txt file using *tr –d [:punct:]* to delete punctuations, *tr [:upper:] [:lower:]* commands for convert all text to lowercase.

```
[root@localhost ~]# tr -d [:punct:] < theraven_orig.txt > theraven_nopunct.txt
[root@localhost ~]# less theraven_nopunct.txt

[root@localhost ~]# tr [:upper:] [:lower:] < theraven_nopunct.txt > theraven_lower.txt
[root@localhost ~]# grep -n "raven" theraven_lower.txt
45:in there stepped a stately raven of the saintly days of yore
53:?though thy crest be shorn and shaven thou? i said ?art sure no craven
54:ghastly grim and ancient raven wandering from the nightly shore?
56:          quoth the raven ?nevermore?
65:   but the raven sitting lonely on the placid bust spoke only
79:   but the raven still beguiling all my fancy into smiling
98:          quoth the raven ?nevermore?
105:          quoth the raven ?nevermore?
112:          quoth the raven ?nevermore?
119:          quoth the raven ?nevermore?
121:   and the raven never flitting still is sitting still is sitting
```

10. Using *tr –d '\r'* command to remove false carriage return characters;
    Using *tr ' ' '\n'* to delete blank spaces
    Using *sort* command to sort the words in the text file.
    Using *uniq –c* command to count numbers of repetitions.

    Finally, using the *head* and *tail* commands, the final file (theraven_wordfreq.txt) displays the results of the counts.

```
[root@localhost ~]# tr -d '\r' < theraven_lower.txt > theraven_lf.txt
[root@localhost ~]# tr ' ' '\n' < theraven_lf.txt > theraven_oneword.txt
[root@localhost ~]# sort theraven_oneword.txt > theraven_sorted.txt
[root@localhost ~]# less theraven_sorted.txt
[root@localhost ~]# uniq -c theraven_sorted.txt > theraven_wordfreq.txt
[root@localhost ~]# head theraven_wordfreq.txt
    432
     15 a
      7 above
      1 adore?
      1 again
      1 agreeing
      2 ah
      1 aidenn
      1 air
      4 all
[root@localhost ~]# tail theraven_wordfreq.txt
      6 word
      1 wore
      1 ?wretch?
      1 wrought
      2 yet
      2 yore
      1 yore?
      2 you
      1 you??here
      1 your
```