Date : 22 June 2019

Day Objectives

- File Handling
  - Basic File Data Processing -Accessing and Modifying File Data
  - Character Count
  - Line Count
  - File Size
  - Word Count
  - Unique Word Count

In [27]:
```
# Read a File - File should exist(read mode)
# Write to a File - Existing(append mode) or New File(write mode)

def readFile(filepath):
    with open(filepath,'r') as f:
        filedata = f.read()
    return filedata
filepath = 'Data-files/data.txt'
#filedata = readFile(filepath)

#for line in filedata.split():
  #  print(line)
print(readFile(filepath))
```

```
new Gautham
```

In [25]:
```
# write into a file
def writeFile(fname,fdata):
    with open(fname,"w")as f:
        f.write(fdata)
    return
fname='Data-files/data.txt'
writeFile(fname,"new Gautham")
print(readFile(fname))
```

```
new Gautham
```

In [29]:
```
#append into a file
def writeFile(fname,fdata):
    with open(fname,"a")as f:
        f.write(fdata)
    return
fname='Data-files/data.txt'
writeFile(fname,"new Gautham")
print(readFile(fname))

# spliting the file data
filedata = readFile(filepath)
for line in filedata.split():
    print(line)
```

```
new Gauthamnew Gauthamnew Gautham
new
Gauthamnew
Gauthamnew
Gautham
```

In [42]:
```
#finding lines,characters and words in given file

with open('Data-files/data.txt') as infile:
    lines=0
    words=0
    characters=0
    for line in infile:
        wordslist=line.split()
        lines=lines+1
        words=words+len(wordslist)
        characters += sum(len(word) for word in wordslist)
print(lines)
print(words)
print(characters)
```

```
2
8
50
```

In [47]:
```
#unique word count
def uniquecount(file_name):
    count={}
    with open(file_name,'r')as f:
        l=f.read().split()
        for i in l:
            if i in count:
                count[i]+=1
            else:
                count[i]=1
        print(count)
file_name = 'Data-files/data.txt'
uniquecount(file_name)
```

```
{'new': 3, 'Gautham': 3, 'Sai_ram_Gautham': 1, 'uytrw': 1, 'ramarao': 1, 'sai':
1, 'ram': 1}
```

In [38]:
```
#function to read a file into a list of lines
# each element in the list is one line in the file
def readFileIntoList(filepath):
    with open(filepath,'r')as f:
        filedata = f.read()
        lines = filedata.split("\n")
        #lines = []
       # for line in f:
        #    lines.append(line)
    return lines
filepath = 'Data-files/data.txt'
readFileIntoList(filepath)
```

Out[38]:
```
['new Gautham new Gautham new Gautham Sai_ram_Gautham', 'uytrw']
```

In [39]:
```
# function to count the number of characters in a file

def charCountFile(filepath):
    count = len(readFile(filepath))
    return count
charCountFile(filepath)
```

Out[39]:
```
57
```

In [44]:
```
#funetion to count the number of words in a file

import re

def wordCountFile(filepath):
    pattern = '[ \n]'
    filedata = readFile(filepath)
    count = len(re.split(pattern,filedata))
    return count
wordCountFile(filepath)
```

Out[44]:
```
11
```

In [52]:
```
# function to get unique elements in a list

#[1,2,3,3,2,1]->>[1,2,3]
# create a empty unique list [1,2,3]

def uniqueData(list):
    # create an empty unique list
    unique = []
    #for every element in the main list,check if it exists in the unique lis
    #If it does not exist,add it to unique list
    #else if it already exist,move on to the next

    for element in list:
        if element not in unique:
            unique.append(element)
    return unique

list=[1,2,3,3,2,1]

uniqueData(list)
```

Out[52]:
```
[1, 2, 3]
```

In [53]:
```
#function to print the frequency count of a

#Data in Line 1
#Data in Line 2
#Data in Line 3
#o/p
#Data : 3
#in : 3
#Line : 3
#1 : 1
#2 : 1
#3 : 1

import re

def Frequencycount(file_name):
    count={}
    with open(file_name,'r')as f:
        l=f.read().split()
        for i in l:
            if i in count:
                count[i]+=1
            else:
                count[i]=1
        print(count)
file_name = 'Data-files/data.txt'
Frequencycount(file_name)
```

```
{'new': 3, 'Gautham': 3, 'Sai_ram_Gautham': 1, 'uytrw': 1, 'ramarao': 1, 'sai':
1, 'ram': 1}
```

In [ ]:
```
1
```