22 June 2019

Day Objectives:

- Basic File Data Processing
  - Accessing and Modifying file data
- character count
- Line Count
- File Size
- Word Count
- Unique Word count

Read a File - File should exists

Write to a file - Existing(appending) or New File

withh open(filePath,'mode') as f: print(type(f))

- here f is a simple name of the file or a variable/object
- open is the function to open the file

In [1]:

```
#1.Char count

def countCharOfFile(filename):

    with open(filename,'r') as f:
        print(len(f.read()))

countCharOfFile('DataFiles/data.txt')
```

```
14
```

In [3]:

```
# 2.File size
def fileSize(filename):

    with open(filename,'r') as f:
        print(len(f.read()),"bytes")

fileSize('DataFiles/data.txt')
```

```
14 bytes
```

In [6]:

```
# 3.Lines Counts
def lineCount(filename):

    with open(filename,'r') as f:
        count = len(f.readlines())
    print(count)
lineCount('DataFiles/data.txt')
```

```
3
```

In [13]:

```
# 4.Word Count

def countWord(filename):

    with open(filename,'r') as f:
        filedata = f.read().split()
        print(len(filedata))
countWord('DataFiles/data.txt')
```

```
8
```

In [17]:

```
# 5.Search word count
def searchForWordcount(filename,word):
    count = 0
    with open (filename,'r') as f:
        filedata = f.read().split()
        for i in range(0,len(filedata)):
            if (word == filedata[i]):
                count+=1
    print(count)
uniqueWordcount('DataFiles/data.txt','Hello')
```

```
2
```

In [54]:

```
# 6.Unique word count using regular expressions
def uniqueWordCount(filename):
    count = 0
    with open(filename,'r') as f:
        filedata = f.read().split()
        ds = {}
        ls = []
        print("Word","     :","Count")
        print("------------------------")
        for i in range(0,len(filedata)):
            count = 0
            temp = []
            if filedata[i] in filedata:
                for j in range(0,len(filedata)):
                    if(filedata[i]==filedata[j]):
                        count +=1
                        temp.append(filedata[i])
            ds[filedata[i]] = temp

        for key in ds:
            print(key,"     :",len(ds[key]))

uniqueWordCount('DataFiles/data.txt')
```

```
Word      : Count
------------------------
Hai      : 2
Hello      : 3
hia      : 1
is      : 1
seven      : 1
six      : 1
```

In [50]:

```
# 6.Unique word count sir's logic ,Optimised logic
def uniqueWordCount(filename):
    count = 0
    with open(filename,'r') as f:
        filedata = f.read().split()
        ls = []
        for i in filedata:
            if i not in ls:
                ls.append(i)
        print(ls,"Length is :",len(ls))
uniqueWordCount('DataFiles/data.txt')
```

```
  File "<ipython-input-50-758966a99408>", line 15

    ^
SyntaxError: unexpected EOF while parsing
```

In [ ]:

```
```