Converting File Data in to Lists:

In [1]:
```
# Function to read a file into a lost of lines
# Each element in the list is one line in the file

def readfiletolist(filepath):
    with open(filepath,'r') as f:
        filedata=f.read()
        lines=filedata.split('\n')
        #lines=[]
        #for line in f:
            #lines.append(lines)
    return lines
filepath='DataFiles/data.txt'
readfiletolist(filepath)
```

Out[1]:
```
['new data',
 'line2\\line3line2',
 'line3Line2',
 'Line3line4line 5Line3line4line 5line4',
 'line 5',
 'Line4',
 'Line 5',
 '']
```

In [6]:
```
# Read the entire data from a file

def readfile(filepath):
    with open(filepath,'r') as f:
        #print(type(f))
        filedata=f.read()         # reads the entire data
    return filedata
filepath='DataFiles/data.txt'
print(readfile(filepath))
```

```
new data
line2\line3line2
line3Line2
Line3line4line 5Line3line4line 5line4
line 5
Line4
Line 5
```

In [9]:
```
# Function to print the char count in a files

def charcountlinesinfile(filepath):
    count=len(readfile(filepath))
    return count
charcountlinesinfile(filepath)
```

Out[9]: 95

In [7]:
```
# Function to print the word count in files

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

Out[7]: 8

In [32]:
```
# Function to print unique word count

def uniquecount(filepath):
    filedata=readfile(filepath).split()
    x=[]
    for i in filedata:
        if i not in x:
            x.append(i)
    return x
filepath='DataFiles/data.txt'
uniquecount(filepath)
```

Out[32]:
```
['new',
 'data',
 'line2\\line3line',
 'line3Line2',
 'Line3line4line',
 '5Line3line4line',
 '5line4',
 'line',
 '5',
 'Line4',
 'Line']
```

In [47]:

```
# Function to print the frequency count of all words in a file
# Frequency Distribution


# 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

def frequencycount(filepath):
    x=uniquecount(filepath)
    filedata=readfile(filepath).split()
    for i in x:
        c=0
        for j in filedata:
            if i==j:
                c=c+1
        print(i,c,'\n')

filepath='DataFiles/data.txt'
frequencycount(filepath)
```

```
new 1

data 1

line2\line3line 1

line3Line2 1

Line3line4line 1

5Line3line4line 1

5line4 1

line 1

5 2

Line4 1

Line 1
```

In [20]:
```
# fUNCTION TO GET UNIQUE ELEMENTS IN A LIST
# [1,2,3,3,2,1] --->[1,2,3]
# First create empty unique list[1,2,3]

    # create an empty unique list
    # check if 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 exists,move on to the next element in the main list
def uniquedata(li):

    unique=[]
    for element in li:


        if element not in unique:

            unique.append(element)
    return unique
li=[1,2,3,3,2,1]
uniquedata(li)
```

Out[20]: [1, 2, 3]

In [ ]:
```
```