# Programming Languages for Data Engineering

## Python Assignment

Name :Inam Abdul Ghafar

ID : 2495895

---

## About Python:

In terms of syntax, Python has always been known as a simple programming language. It prioritises readability and provides clutter - free, easy-to-learn syntax. Furthermore, the Python style guide, PEP 8, includes a set of rules to help with code formatting. Python can easily meet this requirement. Because it is a general programming language, it can generate CSV output for easy data interpretation in a spreadsheet. Python is not only multifunctional, but it is also lightweight and efficient when it comes to code execution. It can be used anywhere because it supports object-oriented, structural, and functional programming styles.

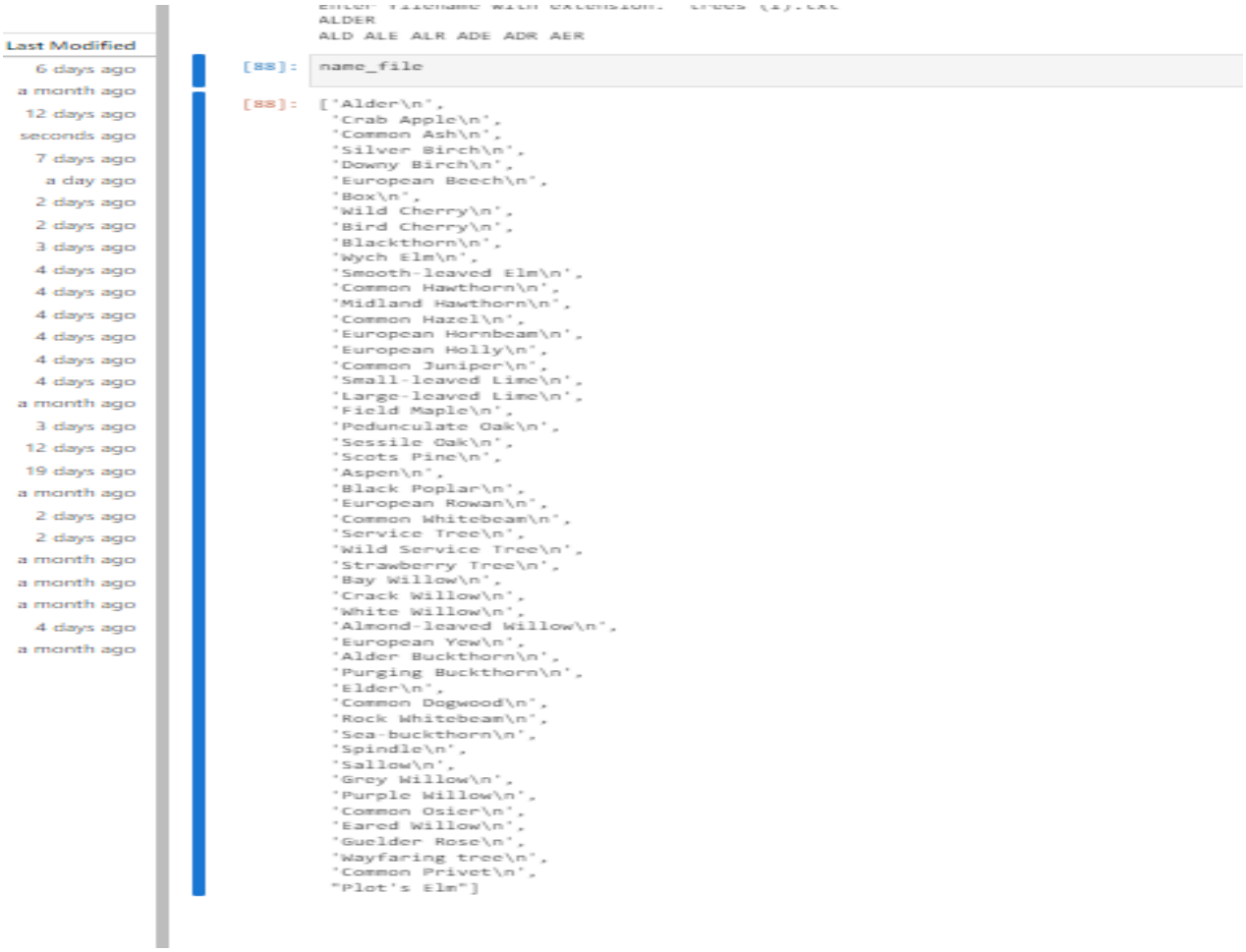
Therefore it provides very easy solutions for completing the assignment.

## Files:

As per asignment we have two files first one consisting of tree names and the second one was values.

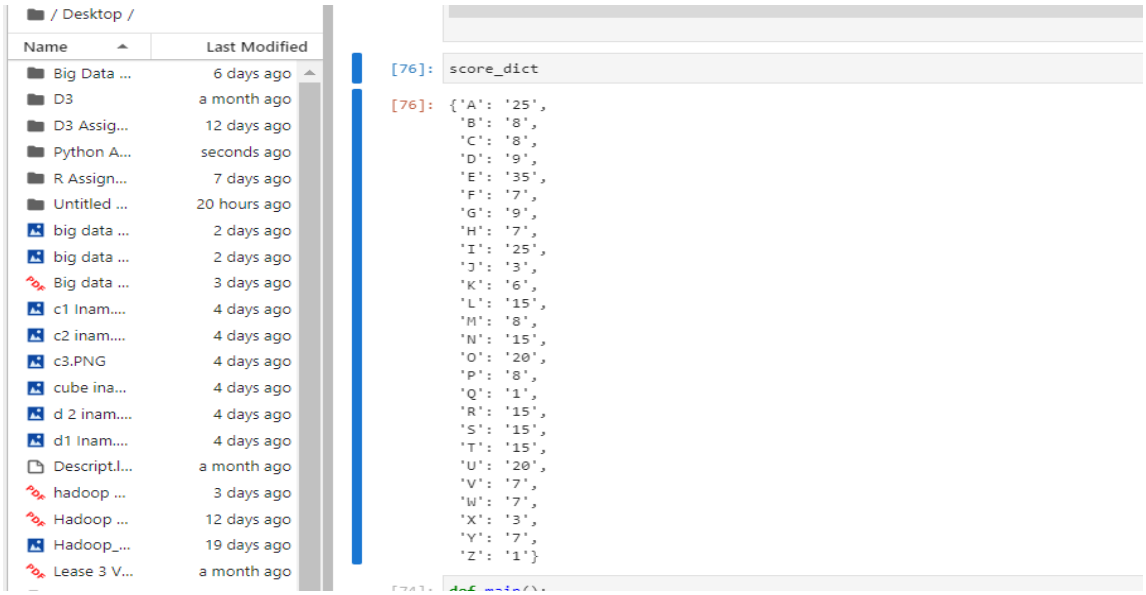
Sor first of all I have opened the files in a python3 tool .the code is given bellow.

```
with open("trees (1).txt","r") as name_f:
    name_file =name_f.readlines()
with open('values (1).txt', 'r') as val_f:
    file_ofValue = val_f.readlines()
```

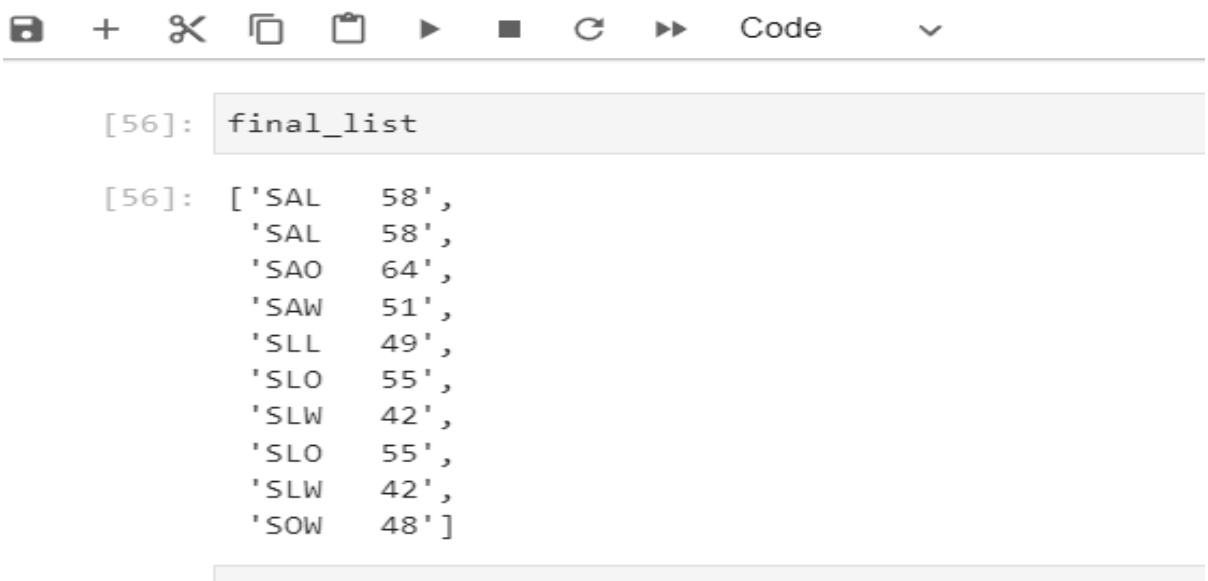



## Score :

Actually when I was starting to work assignment I use jupyter node book when I finished assignment then I use VS Code for python so here is the screen shot of the sore out put which is in dictionry .

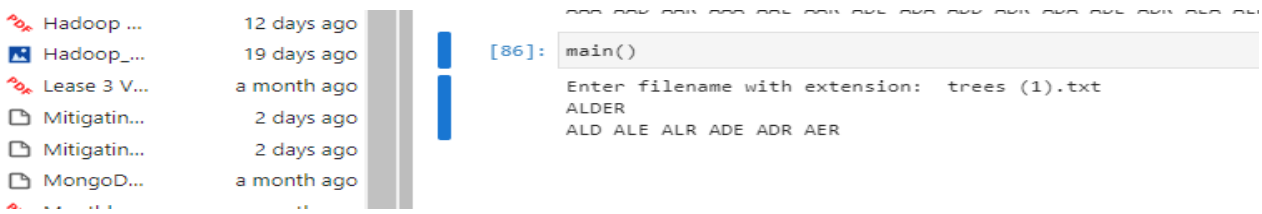



Here is the abbreviation of the out put bellow.



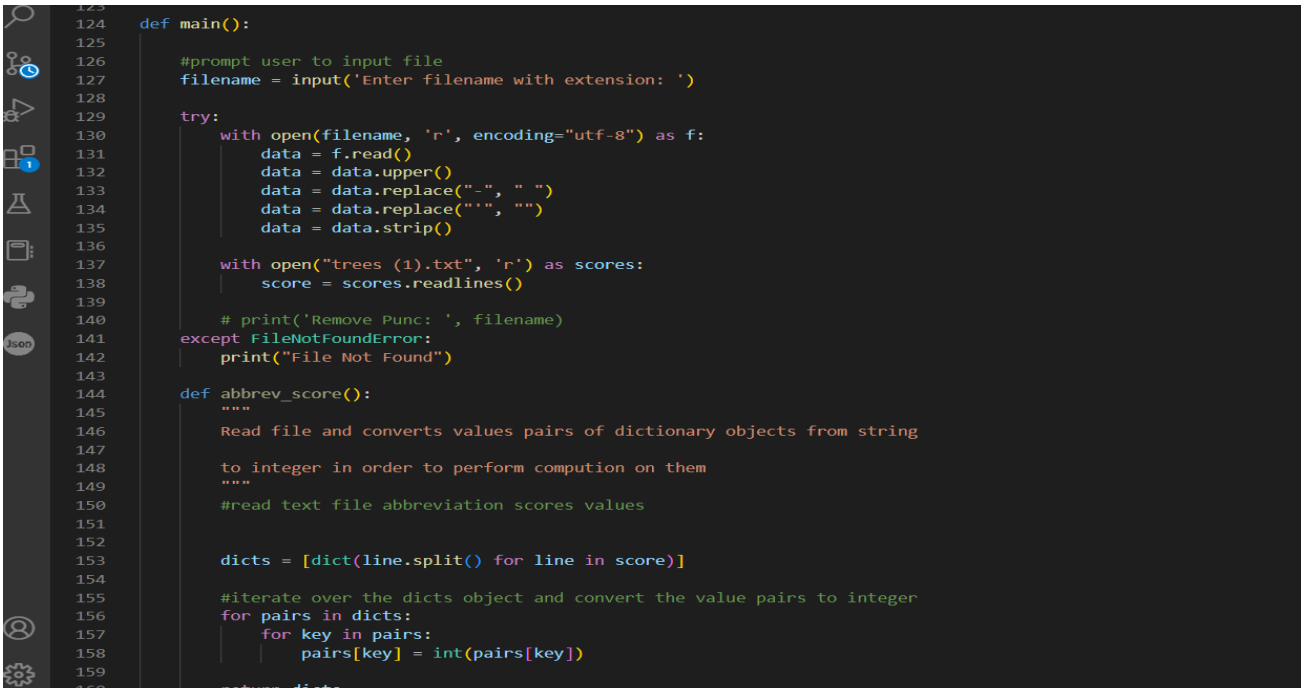

Here is the out put of just one word Alder which is the first word in tree text file you can see bellow more than one abbreviation has been created.similarly I can apply so on in the text file. Each abbreviation is given a score which indicates how good it is.

Enter the your file then it will the out of abbreviation like as bellow.

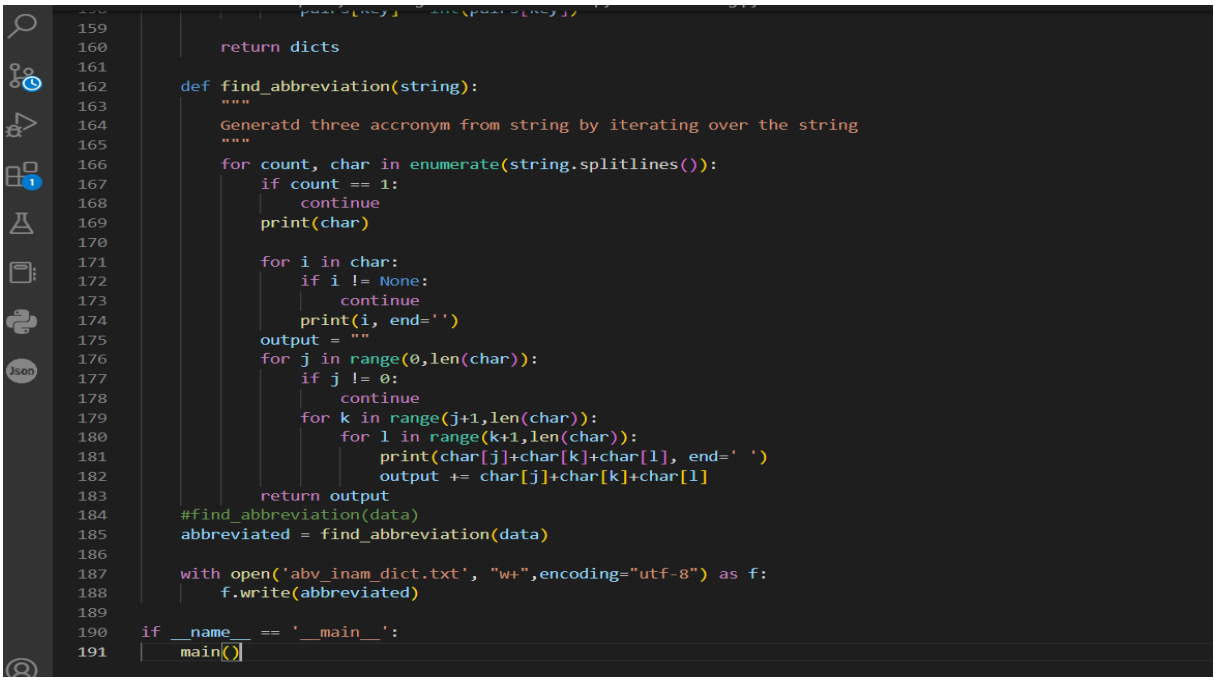



## Code of Main file:

I have created many functions and also I have created two main function this code is for taken the file then it will create the Upper level of the abbreviation. This code I use in VS Code as well as in Jupiter notebook .I have attached the python code file you can see easily.



```
def main():

    #prompt user to input file
    filename = input('Enter filename with extension: ')

    try:
        with open(filename, 'r', encoding="utf-8") as f:
            data = f.read()
            data = data.upper()
            data = data.replace("-", " ")
            data = data.replace("'", "")
            data = data.strip()

        with open("trees (1).txt", 'r') as scores:
            score = scores.readlines()

        # print('Remove Punc: ', filename)
    except FileNotFoundError:
        print("File Not Found")

    def abbrev_score():
        """
        Read file and converts values pairs of dictionary objects from string

        to integer in order to perform compution on them
        """
        #read text file abbreviation scores values


        dicts = [dict(line.split() for line in score)]

        #iterate over the dicts object and convert the value pairs to integer
        for pairs in dicts:
            for key in pairs:
                pairs[key] = int(pairs[key])
```



```
        return dicts

    def find_abbreviation(string):
        """
        Generatd three accronym from string by iterating over the string
        """
        for count, char in enumerate(string.splitlines()):
            if count == 1:
                continue
            print(char)

            for i in char:
                if i != None:
                    continue
                print(i, end='')
            output = ""
            for j in range(0,len(char)):
                if j != 0:
                    continue
                for k in range(j+1,len(char)):
                    for l in range(k+1,len(char)):
                        print(char[j]+char[k]+char[l], end=' ')
                        output += char[j]+char[k]+char[l]
            return output
    #find_abbreviation(data)
    abbreviated = find_abbreviation(data)

    with open('abv_inam_dict.txt', "w+",encoding="utf-8") as f:
        f.write(abbreviated)

if __name__ == '__main__':
    main()
```

I have more output but it will mad long file so I have attached code you can check all the output yourself.

Thank you