

Computer Science project



Topic: Image extractor and text compiler

Class: XIth A2 & XIth C

Group member:

Paritosh (Roll no:25)

Shaarav Gore(Roll no: 49)

Jeremy Stanley (Roll no:13)

CERTIFICATE

This is to certify that _____,
a student of class XII, has successfully completed the
research project on the topic “ _____
_____, under the guidance of
_____.

References taken in making this project have been
declared at the end of the report.

Principal

Teacher In-charge

External Examiner

ACKNOWLEDGEMENT

I hereby acknowledge my deep sense of gratitude and indebtedness to the following personalities whose immense help, genius guidance, encouragement, necessary suggestions, initiations, enthusiasm and inspiration made this work a master art and a joint enterprise.

_____:- (Principal)

_____:- (PGT_____)

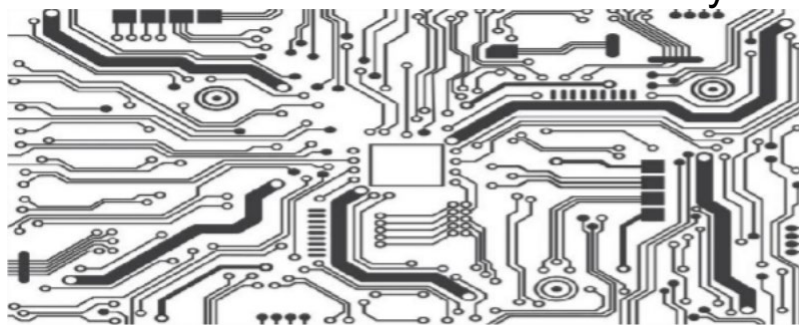
What is python?

Python is an interpreted high-level general-purpose programming language. Its design philosophy emphasizes code readability with its use of significant indentation. Its language constructs as well as its object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

How can python help us ?

Python isn't only for programmers and data scientists. Learning Python can open new possibilities for those in less data-heavy professions, like journalists, small business owners, or social media marketers. Python can also enable non-programmer to simplify certain tasks in their lives. Here are just a few of the tasks you could automate with Python:

- Keep track of stock market or crypto prices
- Send yourself a text reminder to carry an umbrella anytime it's raining
- Update your grocery shopping list
- Renaming large batches of files
- Converting text files to spreadsheets
- Randomly assign chores to family members
- Fill out online forms automatically



What is Python used for?

Python is commonly used for developing websites and software, task automation, data analysis, and data visualization. Since it's relatively easy to learn, Python has been adopted by many non-programmers such as accountants and scientists, for a variety of everyday tasks, like organizing finances.

What are the advantages of Python?

1. Easy to Read, Learn and Write
2. Improved Productivity
3. Interpreted Language
4. Dynamically Typed
5. Free and Open-Source
6. Vast Libraries Support
7. Portability



Why is python so popular?

When Guido van Rossum was creating python in the 1980s, he made sure to design it to be a general-purpose language. One of the main reasons for the popularity of python would be its simplicity in syntax so that it could be easily read and understood even by amateur developers also that's why Python language is incredibly easy to use and learn for new beginners and newcomers. The python language is one of the most accessible programming languages available because it has simplified syntax and not complicated, which gives more emphasis on natural language. Due to its ease of learning and usage, python codes can be easily written and executed much faster than other programming languages.

What are the disadvantages of Python?

1. Slow Speed
2. Not Memory Efficient
3. Weak in Mobile Computing
4. Database Access
5. Runtime Errors



Summary

Python is a **simple**, **versatile** and a complete **programming**

language. It is a great choice for beginners up to professionals. Although it has some disadvantages, we can observe that the advantages exceed the disadvantages. Even Google has made Python one of its primary programming languages.

Bibliography:

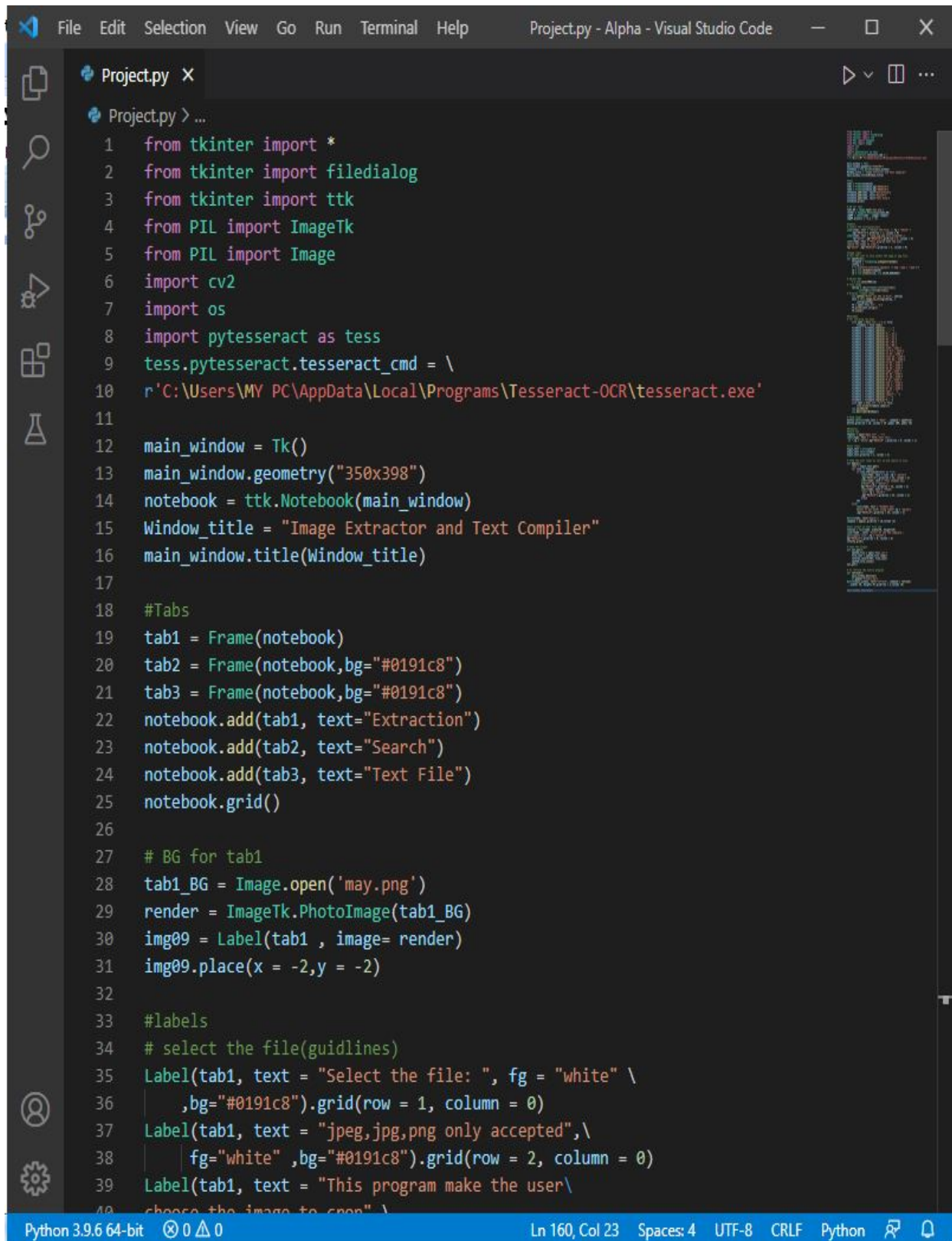
<https://techvidvan.com/tutorials/python-advantages-and-disadvantages/>

<https://www.coursera.org/articles/what-is-python-used-for-a-beginners-guide-to-using-python>

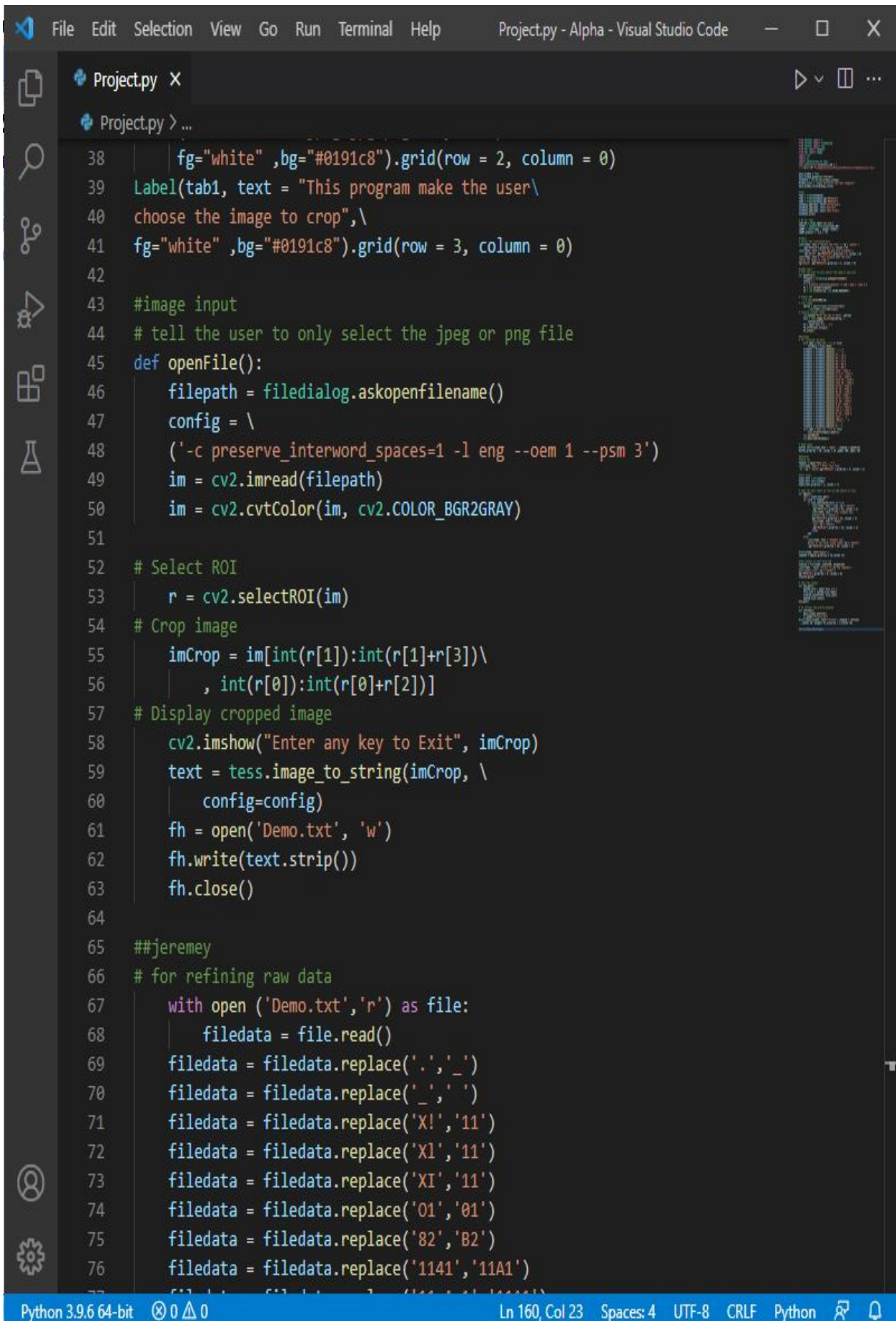
<https://www.python.org/doc/essays/blurb>

Topic: Image extractor and text compiler

Code:



```
1  from tkinter import *
2  from tkinter import filedialog
3  from tkinter import ttk
4  from PIL import ImageTk
5  from PIL import Image
6  import cv2
7  import os
8  import pytesseract as tess
9  tess.pytesseract.tesseract_cmd = \
10 r'C:\Users\MY PC\AppData\Local\Programs\Tesseract-OCR\tesseract.exe'
11
12 main_window = Tk()
13 main_window.geometry("350x398")
14 notebook = ttk.Notebook(main_window)
15 Window_title = "Image Extractor and Text Compiler"
16 main_window.title(Window_title)
17
18 #Tabs
19 tab1 = Frame(notebook)
20 tab2 = Frame(notebook,bg="#0191c8")
21 tab3 = Frame(notebook,bg="#0191c8")
22 notebook.add(tab1, text="Extraction")
23 notebook.add(tab2, text="Search")
24 notebook.add(tab3, text="Text File")
25 notebook.grid()
26
27 # BG for tab1
28 tab1_BG = Image.open('may.png')
29 render = ImageTk.PhotoImage(tab1_BG)
30 img09 = Label(tab1 , image= render)
31 img09.place(x = -2,y = -2)
32
33 #labels
34 # select the file(guidelines)
35 Label(tab1, text = "Select the file: ", fg = "white" \
36         ,bg="#0191c8").grid(row = 1, column = 0)
37 Label(tab1, text = ".jpeg,jpg,png only accepted",\
38         fg="white" ,bg="#0191c8").grid(row = 2, column = 0)
39 Label(tab1, text = "This program make the user\
40         choose the image to open" )
```

```
File Edit Selection View Go Run Terminal Help Project.py - Alpha - Visual Studio Code
Project.py X
Project.py > ...
38 fg="white",bg="#0191c8").grid(row = 2, column = 0)
39 Label(tab1, text = "This program make the user\
40 choose the image to crop",\
41 fg="white",bg="#0191c8").grid(row = 3, column = 0)
42
43 #image input
44 # tell the user to only select the jpeg or png file
45 def openFile():
46     filepath = filedialog.askopenfilename()
47     config = \
48     ('-c preserve_interword_spaces=1 -l eng --oem 1 --psm 3')
49     im = cv2.imread(filepath)
50     im = cv2.cvtColor(im, cv2.COLOR_BGR2GRAY)
51
52     # Select ROI
53     r = cv2.selectROI(im)
54     # Crop image
55     imCrop = im[int(r[1]):int(r[1]+r[3])\
56                , int(r[0]):int(r[0]+r[2])]
57     # Display cropped image
58     cv2.imshow("Enter any key to Exit", imCrop)
59     text = tess.image_to_string(imCrop, \
60                               config=config)
61     fh = open('Demo.txt', 'w')
62     fh.write(text.strip())
63     fh.close()
64
65 ##jeremey
66 # for refining raw data
67 with open ('Demo.txt','r') as file:
68     filedata = file.read()
69     filedata = filedata.replace('.', '_')
70     filedata = filedata.replace('_', ' ')
71     filedata = filedata.replace('X!', '11')
72     filedata = filedata.replace('X1', '11')
73     filedata = filedata.replace('XI', '11')
74     filedata = filedata.replace('01', '01')
75     filedata = filedata.replace('82', 'B2')
76     filedata = filedata.replace('1141', '11A1')
77     filedata = filedata.replace('1111', '111111')
```

Python 3.9.6 64-bit 0 0 Ln 160, Col 23 Spaces: 4 UTF-8 CRLF Python

File Edit Selection View Go Run Terminal HelpProject.py - Alpha - Visual Studio Code

Project.py X

Project.py

```
64
65  ##jeremey
66  # for refining raw data
67      with open ('Demo.txt','r') as file:
68          filedata = file.read()
69          filedata = filedata.replace('.', '_')
70          filedata = filedata.replace('_', ' ')
71          filedata = filedata.replace('X!', '11')
72          filedata = filedata.replace('Xl', '11')
73          filedata = filedata.replace('XI', '11')
74          filedata = filedata.replace('01', '01')
75          filedata = filedata.replace('82', 'B2')
76          filedata = filedata.replace('1141', '11A1')
77          filedata = filedata.replace('11 A-1', '11A1')
78          filedata = filedata.replace('11 A2', '11A2')
79          filedata = filedata.replace('11th A2', '11A2')
80          filedata = filedata.replace('11th AZ', '11A2')
81          filedata = filedata.replace('11-A2', '11A2')
82          filedata = filedata.replace('Xl A2', '11A2')
83          filedata = filedata.replace('XI A2', '11A2')
84          filedata = filedata.replace('1142', '11A2')
85          filedata = filedata.replace('XI A2', '11A2')
86          filedata = filedata.replace('11 AZ', '11A2')
87          filedata = filedata.replace('X A2', '11A2')
88          filedata = filedata.replace('11thA2', '11A2')
89          filedata = filedata.replace('11A 2', '11A2')
90          filedata = filedata.replace('114 2', '11A2')
91          filedata = filedata.replace('X A2', '11A2')
92          filedata = filedata.replace('X\ A2', '11A2')
93          filedata = filedata.replace('(me)', ' ')
94          filedata = filedata.replace('(Host)', ' ')
95          filedata = filedata.replace('|', '')
96          filedata = filedata.replace(' ', '')
97      with open ('Demo.txt','w') as file:
98          file.write(filedata.lower())
99      cv2.waitKey(0)
100     cv2.destroyAllWindows()
101
102  # Read image
103  button =Button(tah1.text = "Open" , command = openFile)
```

Python 3.9.6 64-bit 0 0 Ln 143, Col 34 Spaces: 4 UTF-8 CRLF Python

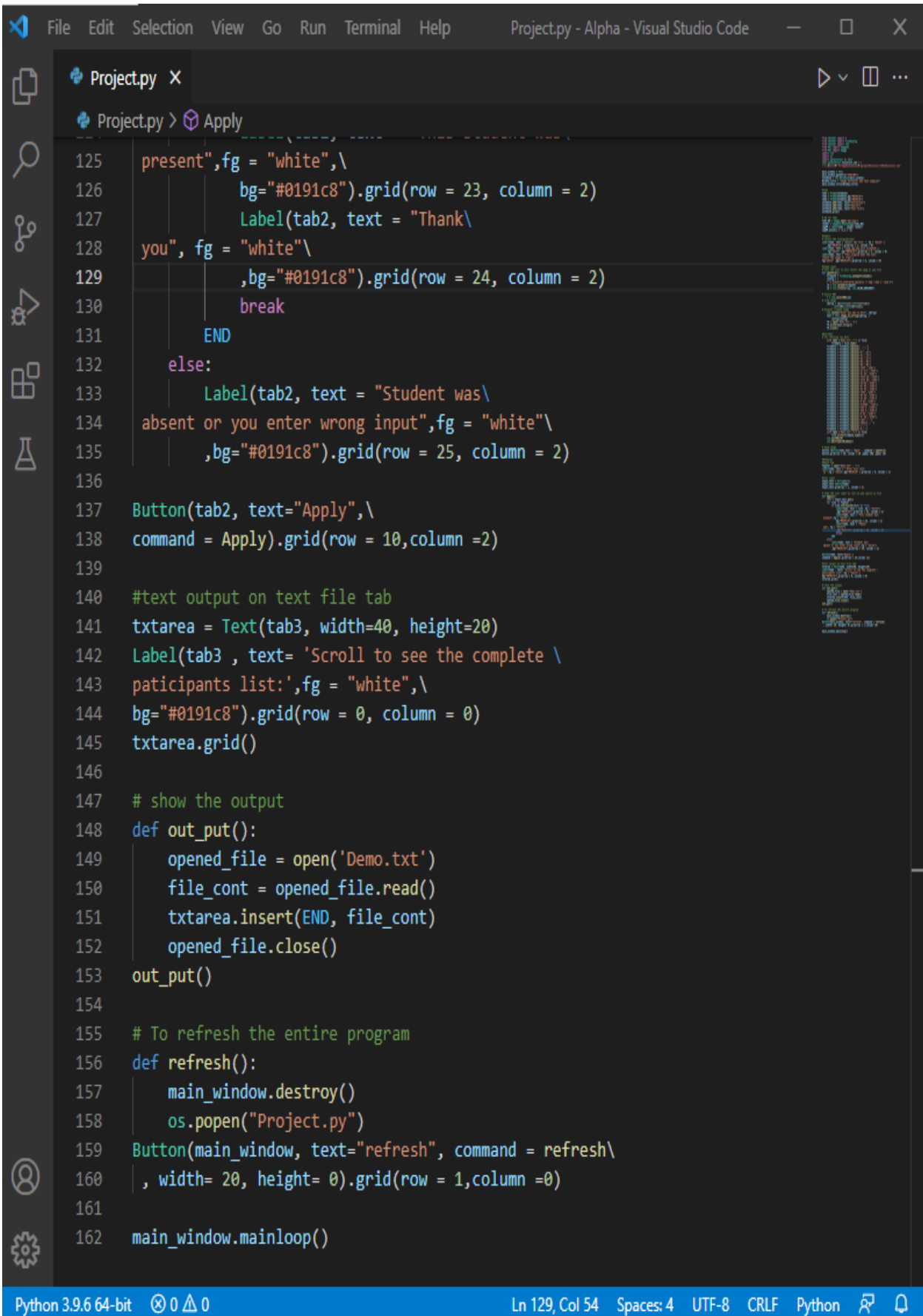
File Edit Selection View Go Run Terminal Help Project.py - Alpha - Visual Studio Code

Project.py X

Project.py > openFile

```
101
102 # Read image
103 button =Button(tab1,text = "Open" , command = openFile)
104 button.grid(row = 10, column = 10 ,pady= 100, padx= 10)
105
106 ##shaarav
107 #Seach tab
108 fopener = open("Demo.txt" , 'r')
109 Label(tab2, text = " Enter full roll\
110 | no ",fg = "white",bg="#0191c8" ).grid(row = 0, column = 1)
111
112 #text input
113 Input_text = StringVar()
114 Input_text =Entry(tab2)
115 Input_text.grid(row = 1, column = 2)
116
117 # take the user input as roll no and seacrh in file
118 def Apply():
119     roll = Input_text.get()
120     for line in fopener :
121         if line.startswith(roll) == True:
122             Label(tab2, text = line ,fg = "white"\
123             ,bg="#0191c8").grid(row = 22, column = 2)
124             Label(tab2, text = "This student was\
125 present",fg = "white",\
126             bg="#0191c8").grid(row = 23, column = 2)
127             Label(tab2, text = "Thank\
128 you", fg = "white"\
129             ,bg="#0191c8").grid(row = 24, column = 2)
130             break
131         END
132     else:
133         Label(tab2, text = "Student was\
134 absent or you enter wrong input",fg = "white"\
135         ,bg="#0191c8").grid(row = 25, column = 2)
136
137 Button(tab2, text="Apply",\
138 command = Apply).grid(row = 10,column =2)
139
140 #text output on text file tab
```

Python 3.9.6 64-bit 0 0 Ln 90, Col 48 Spaces: 4 UTF-8 CRLF Python



```
Project.py X
Project.py > Apply
125 present",fg = "white",\
126         bg="#0191c8").grid(row = 23, column = 2)
127 Label(tab2, text = "Thank\
128 you", fg = "white"\
129         ,bg="#0191c8").grid(row = 24, column = 2)
130 break
131 END
132 else:
133     Label(tab2, text = "Student was\
134 absent or you enter wrong input",fg = "white"\
135         ,bg="#0191c8").grid(row = 25, column = 2)
136
137 Button(tab2, text="Apply",\
138 command = Apply).grid(row = 10,column =2)
139
140 #text output on text file tab
141 txtarea = Text(tab3, width=40, height=20)
142 Label(tab3 , text= 'Scroll to see the complete \
143 participants list:',fg = "white",\
144         bg="#0191c8").grid(row = 0, column = 0)
145 txtarea.grid()
146
147 # show the output
148 def out_put():
149     opened_file = open('Demo.txt')
150     file_cont = opened_file.read()
151     txtarea.insert(END, file_cont)
152     opened_file.close()
153 out_put()
154
155 # To refresh the entire program
156 def refresh():
157     main_window.destroy()
158     os.popen("Project.py")
159 Button(main_window, text="refresh", command = refresh\
160     , width= 20, height= 0).grid(row = 1,column =0)
161
162 main_window.mainloop()
```

Python 3.9.6 64-bit 0 0 0 Ln 129, Col 54 Spaces: 4 UTF-8 CRLF Python

Output:

