Assignment: 2
Course: Automation
Date: 24th April 2022
Due Date: 14th May 2022
Submission on: Blackboard
Weightage: 15%
Total marks: 30

Description:

This assignment is related to automation concepts that we have discussed in the lecture or we will discuss in the coming lectures. This assignment aims to push you towards knowing the usage of some of those concepts. We have discussed the basics of these concepts during lectures, and now it's time for you to dig more, research, and puts some effort into implementing those concepts on your own. An assignment does not always contain the things, which you have already done exactly during Lecture/Practical. The assignment aims to make you put some effort on your own to understand and implement the things.

Mainly, there are three parts in the assignment as follows,

- To write a script to extract data from a file and plot it using any tool.
- To build a tool for web scrapping.
- To create a tool to read excel sheet and send emails automatically.

The description and explanation of each part are mentioned in its particular section.

Requirements:

- Provide a detailed report on all of these three tasks.
- What do you understand from the explanation given as a description?
- Provide a step by step configuration steps for each task.
- Provide requirements for each task in its section.
- Provide files and screenshots for proof that you have completed these tasks yourself.
- Provide the files for each task separately with the report. So that, I can run files easily.
- Each student is supposed to work individually on their assignment. If two assignments will be found similar. Both will get ZERO marks. So, please do not share your assignment answers and files with others and do not copy from the others.
- There is no word limit for the report. But, do not unnecessarily extend the description. If you can explain anything in four lines, do not occupy one full page for the explanation.
- Submit your report in PDF format, except those code files.

Report organization:

It's up to you, how you organize your report. You have a free hand. Following are the suggestions in terms of what I am looking for.

1. Summary
2. Section 1 or Task 1
   a. Description/explanation, requirements (HW/SW), Brief Methodology, Implementation (Steps with screenshots), observations (Easiness/Difficulties/New-things),
3. Section 2 or Task 2
   a. Description/explanation, requirements (HW/SW), Brief Methodology, Implementation (Steps with screenshots), observations (Easiness/Difficulties/New-things),
4. Section 3 or Task 3
   a. Description/explanation, requirements (HW/SW), Brief Methodology, Implementation (Steps with screenshots), observations (Easiness/Difficulties/New-things),
5. Conclusion
6. References

Task 1:

This task aims to extract specific information from a trace file. A sample file is given with this assignment. You can find different column data in that sample file. Your task is to extract the following and plot it using the file given in the previous practical for GNU Plot. The parameter which you need to show is Packet delivery ratio which can be defined as, 'the ratio of number of packets received at the destination to the number of packets sent from the source'. The performance is better when packet delivery ratio is high.

• Column number 5 'Total Packets at Tx node'
• Column number 7 'Packets received at Rx node'
• Column number 3 'Total number of packets'
• Perform a simple arithmetic operation of Col7/Col3 and multiply it by 100 to get a Packet Delivery Ratio (PDR) of Rx node.

PDR = Tot packet received at Rx node/Tot packet transmitted x 100
= Column 7/ Column 3 *100

This will be the PDR for the Rx nodes.

**In the similar way you can plot also PDR for Tx node. C5/C3*100, and plot both the lines in the same graph.

• This third column data you must plot.
• Write a script to extract the two columns and a calculated third column and save it in a new file.
• Use that new file as an input to plot the graph. Change the X-axis, Y-axis, and legends names appropriately with new ranges on the x and y-axis. Write this second script for the graph plotting. Now your Y-axis will be PDR (%) and X-axis TCP streams (1, 2, 3, ….. X)
• You can write a third script to control the above scripts OR you can do it individually and finally run a command for Gnuplot to plot the new graph.

The task is to plot the PDR of TX and RX nodes on a single graph using scripts. Means, the script will be written to read the values from csv or excel sheet and save it in a particular format in a new file from which the required PDR column can be read by a script again to plot it using GNUPLOT syntax script or some other tool.

****

- For the submission, provide these script files, in the report and separately.
- You do not have to plot, col3 or col5, you only have to plot the third column as PDR.
- You can either use the awk script. AWK uses the same linux command syntax. Or you can use simply
- the Linux commands to extract the information and save it in a different file.
- You can directly use the col information to plot a graph. But at this stage, this will be too difficult for
- you. What will be easy for you? is that work on the script to extract data first and save it in a different file, and then use this file as an input to plot.
- The main aim is to automate the extraction and plotting of the graph, so try to automate this process as much as you can. You can mention in your report, the level of automation. For example, you can use just a single script to perform all the above options. You can write separate scripts for each task.
- You can use R as well for plotting if you are comfortable with R software. But remember, automation
- is aimed here, so try to automate the stuff using scripts and files for plotting.

TASK 2:

Write a program in python to extract data from any job search websites.
From that website, you must capture the following for the python developer-related jobs.

- Job title
- Company
- Date posted
- Post link
- Last date to apply (if available)
- Description

Provide those in xlsx file.
For this task, you should include a code file and an output file in your report. Also, upload your code file separately.

Task 3:

You are working in a financial firm. Your boss has given you the job to track the people who have missed their monthly installment of May and send them an email for the reminder to clear their dues. The list could be very long and for you to send emails to each person individually, by copying the email address from excel sheets and send. It will take days. It will be a boring job as well, maintaining the spreadsheets and emailing people. That can be a repeated job every month.
This task can be automated in two parts, first, take the input from financial spreadsheets to update the dues excel sheet automatically, to update the sheet for those who have paid their dues. Second, a code can be used to read the dues-related excel sheet and email automatically those only who have missed their dues.
For this assignment, you should focus on the second part ONLY, where you already have an excel sheet with updated details on people who have missed their dues. Your job here is to write a code in python to read from the excel sheet and send emails automatically to only those who have missed their dues.
At a high level, your task will include the following parts,
• Read data from an Excel spreadsheet.
• Find all members who have not paid their dues for the month of May.
• Find their email addresses and send them personalized reminders.

Weightage:

	Marks	Distribution
Task 1	10	Description (description, methodology (way to solve and its clarity)) – 4 Implementation (code, running, output including screenshots) - 6
Task 2	10	Description (description, methodology (way to solve and its clarity)) – 4 Implementation (code, running, output including screenshots) - 6
Task 3	10	Description (description, methodology (way to solve and its clarity)) – 4 Implementation (code, running, output including screenshots) - 6