

Assignment - I

Hometake Lab Assignment

Instructions:

- This is the first assignment
- It is a programming assignment
- You should upload a python script to the google drive before deadline, with the following format:
`FullName_ID_XXXX_XX_Assignment_1.py`
- Any written answer other than the code, if there is any, should be answered on the script with a comment
- All the answer to the solution should be shown(Outputted) upon running
- Do not turn in **ChatGPT** , **Gemini** and **Bing** solutions as your own.

Questions:

1. Write regular expressions for the following languages.

- The set of all alphabetic strings;
- The set of all lower case alphabetic strings ending in a b;
- The set of all strings from the alphabet a, b such that each a is immediately preceded by and immediately followed by a b

2. Extracting and Replacing Email address:

- Given the doc:
doc = "For more information details please mail us at: infodesk@aau.edu.et
You can also reach us for inquiries related to registration at:
registration@aau.edu.et"
- Regex expression** should contain a proper format which technically will have email address and a domain name separated by "@"
 - Email: Email Name (Local Part)+ "@" + domain name [Top Level , Second Level , Third Level]
- Write a proper **regular expression**
- Write a **function** that validates an email with the given regex expression. It should return a bool that validates emails in general with the given rule of regex expression of emails.
- Find all addresses** given in the document. Hint: Use **re.findall()** function
- Replace** the given information desk email with an "**information_desk@aau.et**".
Hint: Use **re.sub()**

3. Extract all dates from a text in a format 'dd/mm/yyyy'. Hint: Use **re.findall()**

- Write the **regex expression** for the date format.
- Given the texts:
text_1 = "Today's date is 13/04/2024 and tomorrow will be 14/04/2024. Last year, on this day, it was 13/04/2023."

text_2 = "We have several dates noted here: the project starts on 01/05/2024, the review is on 15/08/2024, and the end date is expected to be 30/11/2024. Please mark these in your calendar."

text_3 = "In the dataset provided, we noticed some entries are incorrectly dated, such as entry 05/13/2024, which doesn't conform to the dd/mm/yyyy format. Correct dates are 12/12/2023, 25/10/2023, and possibly 31/12/2023 which might need verification due to its year-end placement."

Write a **function** that returns dates from the texts and the regex expression (pattern) from the previous question.

- 4. Username validation system:** Let's consider a simple username validation system , that allows users to accept a username that is 8-20 characters long, a combination of letters, numbers and underscores. **Hint:** Use **re.match()**
- Write a regex that accepts this username
 - Write a function that returns "Username Valid!" if it satisfies the rule "Username Valid, please try another one" if not.
 - [Optional Question / Bonus] - Identify the particular error and return the error.

Deadline: 18/04/24