Assessment details

General

The **Python programming lab practical**is split into five separate lab submissions that relate to the instructional videos contained in the weekly materials. For each submission, you will solve a number of Python programming problems using Jupyter Notebook.

Your Jupyter notebook should include at least one markdown cell and at least one code cell for each task contained within the separate labs. The markdown cell briefly introduces the task and the code cell includes the Python code that you implement the task. Also note that your Python program must contain meaningful comments and sensible variable and function names. Otherwise, up to half of your original mark can be deducted at the marker’s discretion. Please submit your lab work before the submission deadline.

Specific

Please follow the specific instructions to complete this lab. You may use this template to complete and submit these tasks: [ITEC610\_Assessment 1 (IPYNB 131KB)](https://canvas.online.acu.edu.au/courses/535/files/126451?wrap=1)[Download ITEC610\_Assessment 1 (IPYNB 131KB)](https://canvas.online.acu.edu.au/courses/535/files/126451/download?download_frd=1)

Task 1 (3 points)

Let df be a pandas DataFrame constructed with the following code:

data = np.array([0, 7, 3, 6, 2, 8, 5, 9, 4]).reshape(3, -1)  
df = pd.DataFrame(data, index=['One', 'Two', 'Three'], columns=['a', 'b', 'c'])

**What is the output of the following code (answer questions A-T)? Please append with detailed comments about why.**

**A.** print(df)  
**B.** df[‘a’]  
**C.** df[‘One’]  
**D.** df.loc[‘Two’]  
**E.** df[:2]  
**F.** df.iloc[:,:2]  
**G.** list(df.columns)  
**H.** list(df.index)  
**I.** df[‘b’][‘Two’]  
**J.** list(df.iloc[2, :])

**K.** df.drop('a', axis=1)  
**L.** df[df.a !=5]  
**M.**list(df.sum(axis=0))  
**N.** df.iloc[:, list(df.sum(axis=0) < 17)]  
**O.** df.sort\_values(by='c')  
**P.**df.sort\_values(by='Two', axis=1)

**Q.** df.T  
**R.** (df<=2).any(axis=0)  
**S.** df.applymap(lambda x: x\*2-1)  
**T.**df.apply(lambda x: max(x), axis=1)

**Note:**Don’t forget to use text cell in Google Colab to explain what the code does. If you only provide the outputs, you will not be awarded the full mark. Questions R, S and T require some language features that you may not have learned in this unit. We ask that you to learn this language feature yourself: 'any' and lambda expression.

Task 2 (1 point)

**Step 1:**Copy and paste the data below and save the file as ‘datainput.csv’.

id, emp\_name, emp\_salary, start\_date, dept

1, Vishavjeet, 999.3, 2018-12-04, IT  
2, Nimesh, 878.4, 2017-09-24, IT  
3, Sara, 611, 2014-11-15, HR  
4, Gulfam, 729, 2014-05-11, IT  
5, Jack, 845.25, 2016-07-27, Finance  
6, Jessica, 878, 2014-05-21, Marketing  
7, Saurav, 632.8, 2018-07-30, Operations  
8, Sachin, 722.5, 2017-06-17, SEO

**Step 2:**Read data into a pandas DataFrame with variable name ‘data’.

**Step 3:**Complete the following:

* 1. Print the first 5 rows, and all of the columns for those rows.
  2. Print the rows from position 5 onwards, and columns from position 2 onwards.
  3. Print the first two columns, and all of the rows for the column.
  4. Print from 6th row, and all of the columns for that row.

Task 3 (1 point)

**Step 1:** Copy and paste the above data in the sheet 1 of an excel file ‘input.xlsx’.

**Step 2:** Copy and paste the following data as the sheet 2 data of the same excel file:

# Data in Sheet2  
id name zipcode  
1, Vishavjeet, 778888  
2, Nimesh, 553322  
3, Sara, 665544  
4, Gulfam, 663355  
5, Jack, 110078  
6, Jessica, 110059  
7, Saurav, 676770  
8, Sachin, 221144

**Step 3:**Complete the following:

* 1. Read the sheet 1 data into a pandas DataFrame with variable name ‘df1’ and then read the sheet 2 data into a pandas DataFrame with variable name ‘df2’.
  2. Print the first five rows of the ‘salary’ column of sheet 1 and Print the first five rows of the ‘zipcode’ column of sheet 2.
  3. Merge df1 and df2 as df3 and then write df3 to the sheet 3 of the excel file.

Task 4 (1 point)

**Step 1:** Copy and paste the data below and save the file as ‘input.json’.

{  
"ID":["1","2","3","4","5","6","7","8" ],  
"Name":["Vishavjeet","Amit","Nimesh","Sara","Saurav","Gulfam","Anupam","Guru" ],  
"Salary":["623.3","515.2","611","729","843.25","578","632.8","722.5" ],  
"StartDate":[ "1/1/2013","9/23/2014","11/15/2013","5/11/2014","3/27/2015","5/21/2013",  
"7/30/2013","6/17/2014"],  
"Dept":[ "IT","Operations","IT","HR","Finance","IT","Operations","Finance"]  
}

**Step 2:**Complete the following:

1. Read the json file into a pandas DataFrame with variable name ‘data’
2. Print the data and then dump the data into a json string and print the dumped string again with indent of 4 for an easy read.
3. Remove the ‘Salary’ column of data and then writes the new data to another json file ‘output.json’.
4. Load another json file from the Web with URL: https://api.exchangerate-api.com/v4/latest/USD into a data frame ‘df’ and then print this data frame.

Some helpful websites and resources

* [Google CoLabLinks to an external site.](https://colab.research.google.com/notebooks/) (Google, n.d.) environment.
* [PythonLinks to an external site.](https://www.python.org/) (Python Software Foundation, 2022) official website.