**Lab 5: Learning Python for Data Science**

**Objectives**

* To learn Python basics.
* To learn the basic operations in data science using Python.

**Equipment**

* A Windows 10 PC with Internet connection

**Procedures**

**Part 1. Learning Python for Data Science**

* 1. Learn Python basics by working through the page [Learn Python 3 with Jupyter Notebook] at:

<https://gist.github.com/kenjyco/69eeb503125035f21a9d>

The Jupyter Notebook file ‘**learning-python3.ipynb**’ is needed.

* 1. Work through an example data analysis at the page [Jupyter Notebook for Beginners Tutorial]:

<https://www.dataquest.io/blog/jupyter-notebook-tutorial/>

The CSV file of the data is ‘**fortune500.csv**’.

* 1. Another tutorial on data science using Python is the page [A Complete Python Tutorial to Learn Data Science from Scratch] at:

<https://www.analyticsvidhya.com/blog/2016/01/complete-tutorial-learn-data-science-python-scratch-2/>

* 1. A more comprehensive treatment on the topic can be found at the page [Python for Data Science]:

<https://courses.analyticsvidhya.com/courses/introduction-to-data-science>

To access the page, you need to register an account at:

<https://id.analyticsvidhya.com/auth/signup/>

**Additional Resources**

* 成為python數據分析達人的第一課

<http://moocs.nccu.edu.tw/course/123/intro>

* Introduction to Computer Science and Programming in Python

<https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-0001-introduction-to-computer-science-and-programming-in-python-fall-2016/>

* Jupyter Notebook: An Introduction

<https://realpython.com/jupyter-notebook-introduction/>

* Mastering Markdown

<https://guides.github.com/features/mastering-markdown/>

* Top Python IDEs for 2019

<https://www.datacamp.com/community/tutorials/top-python-ides-for-2019>

**Requirements**

**Program 1 (P1) – Printing stars**

1. Two questions are asked:

* How many stars are needed?
* 1-sided or 2-sided?

1. If the answers are 5 [stars] and 2 [-sided], the output is:

|  |  |
| --- | --- |
|  | \*  \*\*  \*\*\*  \*\*\*\*  \*\*\*\*\*  \*\*\*\*  \*\*\*  \*\*  \* |

1. If the answers are 8 [stars] and 1 [-sided], the output is:

|  |  |
| --- | --- |
|  | \*  \*\*  \*\*\*  \*\*\*\*  \*\*\*\*\*  \*\*\*\*\*\*  \*\*\*\*\*\*\*  \*\*\*\*\*\*\*\* |

1. You’re encouraged to handle exceptions like non-numeric input.

**Hint**

* Use the built-in function input() to read a string from the input.

**Program 2 (P2) – Solving quadratic equations**

1. The coefficients of several quadratic equations () are given in the CSV file ‘**abc\_lab5\_p2\_input.csv**’.
2. The content of the file is:

|  |  |
| --- | --- |
|  | a,b,c  1,5,3  2,-4,2  3,2,-1  3,2,1 |

1. The program reads in the CSV file and outputs the following content on the screen:

**Basic** requirement:

|  |  |
| --- | --- |
|  | a b c D p q  1 5 3 13 -4.303 -0.697 (Real roots: -4.303 and -0.697)  2 -4 2 0 1 1 (Repeated roots: 1)  3 2 -1 16 -1 0.333 (Real roots: -1 and 0.333)  3 2 1 -8 NaN NaN (Complex roots are not calculated) |

**The alignment should be more nice-looking.**

**The precision or the number of decimal places shown is for reference only**. We are not taking a Mathematics course.

**Advanced** requirement:

|  |  |
| --- | --- |
|  | a b c D p q  1 5 3 13 -4.303 -0.697 (Real roots: -4.303 and -0.697)  2 -4 2 0 1 1 (Repeated roots: 1)  3 2 -1 16 -1 0.333 (Real roots: -1 and 0.333)  3 2 1 -8 -0.333 0.471 (Complex roots: -0.333 + 0.471j and -0.333 – 0.471j) |

1. The program also outputs the following content to the CSV file ‘**abc\_lab5\_p2\_output.csv**’:

**Basic** requirement:

|  |  |
| --- | --- |
|  | a,b,c,D,p,q  1,5,3,13,-4.303,-0.697 (Real roots: -4.303 and -0.697)  2,-4,2,0,1,1 (Repeated roots: 1)  3,2,-1,16,-1,0.333 (Real roots: -1 and 0.333)  3,2,1,-8,, (Complex roots are not saved) |

**Advanced** requirement:

|  |  |
| --- | --- |
|  | a,b,c,D,p,q  1,5,3,13,-4.303,-0.697 (Real roots: -4.303 and -0.697)  2,-4,2,0,1,1 (Repeated roots: 1)  3,2,-1,16,-1,0.333 (Real roots: -1 and 0.333)  3,2,1,-8,-0.333,0.471 (Complex roots: -0.333 + 0.471j and -0.333 – 0.471j) |

1. You’re encouraged to handle exceptions like linear equation, non-numeric input.

**Hint**

* How to create new columns derived from existing columns?

<https://pandas.pydata.org/docs/getting_started/intro_tutorials/05_add_columns.html>

* Square root of the column in pandas python

<http://www.datasciencemadesimple.com/square-root-of-the-column-in-pandas-python-2/>

**Submission**

* Present your programs in Jupyter Notebook files with code/ description/ comments
* ‘**Chan\_Tai\_Ming\_P1.ipynb**’ and ‘**Chan\_Tai\_Ming\_P2.ipynb**’ are the filenames of student Chan Tai Ming’s notebook files
* Upload the notebook files to Moodle before the due date as informed by the lecturer

30/3/2022