

Questions for Session 4

- Q 1. Write Python programmes to answer the following questions, respectively
- (a) Create a pandas data frame *marks* to express the following table

Name	Business statistics	Simulation
John	65	60
Anna	72	65
Emma	56	64

- (b) Select column *Simulation*;
- (c) Selection the first row;
- (d) Delete column *Simulation*;
- (e) Delete the last row
- (f) Add a new column called 'Grades' to marks: Grade = ['Merit', 'Merit', 'Merit']

- Q 2. Suppose

$$A = \begin{pmatrix} 1 & 2 & 6 & 2 \\ 0 & 7 & 1 & 3 \\ 4 & 0 & 2 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 5 & 2 & 3 & 2 \\ 2 & 1 & 4 & 4 \\ 2 & 5 & 7 & 1 \end{pmatrix}$$

Write a Python programme to answer the following questions

- (a) Calculate $A + B$, $A - B$, $A \circ B$, $A \oslash B$, respectively, and
- (b) Calculate the dot product between A and B^T .

- Q 3. Download the dataset [Adult Exercise.csv](#) from Moodle, write a Python programme to answer each of the following questions.

- (a) Read the file into a dataframe called AE_df;
- (b) Show the first 5 rows and the last 5 rows of AE_df, respectively;
- (c) Find out the data structure of AE_df;
- (d) Find out the number of observations, mean, minimum value, maximum value of AE_df and then assign the data to a data frame called *basicStatistics*;
- (e) Is the variable name *hours-per-week* valid? If not, rename it to be *hours_per_week* and then find mean, minimum value, maximum value, respectively;
- (f) Write *basicStatistics* to a csv-formatted file called bS.csv.