



Introduction to Data Science
Supervised learning (Working with data)

Download data set form kaggel or Google datasets.

- **(10 pt)** Understand and describe it dimensions sparsity, and resolution
- **(10 pt)** Visualize the data using relevant visualization chart
- **(10 pt)** Aggregate and summarize the data

Create a list of 1000 random integers between 1 and 100000, then calculate the Z-Score to check for the outliers. Consider values Z-Score > 2 as outliers

- **(10 pt)** Solve the exercise using excel
- **(10 pt)** Solve the exercise using Python
- **(10 pt)** Comment on the results you find

Create a list of 10 random integers between 1 and 100.

- Standardize the numbers
- Normalize the numbers
- **(10 pt)** Solve the exercise using excel
- **(10 pt)** Solve the exercise using Python
- **(10 pt)** Comment on the results you find

Create a list of 100 random pair of integers (x,y) between 1 and 100. Draw visualization of the data using different

- **(10 pt)** Draw scatter plot of the numbers using Excel
- **(10 pt)** Draw scatter plot of the numbers using Python

