**This assignment is worth 5% of the overall grade.**

In this assignment, you will practice simple modeling of customer churn management on small set of pre-processed data with small number of attributes by applying visual, statistical and learning techniques.

**Data:** use the attached file: [Telecom-Usage-Details.csvPreview the document](https://canvas.instructure.com/courses/1846650/files/88328012/download?wrap=1)

**Columns:**the first line of the above file contains columns names for your reference. You might need to investigate the business meaning of each in more details.

**Tasks:**

1- Examine the distribution and importance of key variables including visual and statistical analysis. Discuss your results. (Code: 5 points - Explanation: 5 points)

2- Find out the best way to segment customers using K-means based on the Tenure and Total Charges variables in the dataset. Can we reflect the results on customer churn behavior? How? What are the Interpretations? (Code: 15 points - Explanation: 15 points)

3- Build simple models using Logistic Regression to predict customer churn behavior based on the most important variables in the provided dataset. You should build 3 models by randomly selecting the training and test sets used each time. Use 70% training and 30% testing. Compare and discuss the 3 resulting models and their performance in the test session. (Code: 20 points - Explanation: 20 points)

4- Plot the ROC curve of the 3  models overlay them on same visual with the associated AUC result. (Code: 10 points - Explanation: 10 points)

**Deliverables:** Jupiter Notebook containing the Python code of the above 4 tasks and the explanation for each task/step in the task in the same sequence. The explanation part is your report with half the grade so please make it clear, objective and to the point. Code will not be graded without associated explanation.