**Week 9**

Data Science Intern at Data Glacier
Project: Bank Marketing (Campaign)

Name: Adrian Baysa
Email: adrianbaysa2@gmail.com
Country: Philippines
Batch Code: LISUM16

1. **Problem Description**

    ABC Bank wants to sell its term deposit product to customers and before launching the product they want to develop a model which help them in understanding whether a particular customer will buy their product or not (based on customer's past interaction with bank or other Financial Institution).

2. **Github Repolink:** https://github.com/agbaysa/dataglacier_week9

3. **Data Transformation**

    The following data transformation was done on the data:

    a. Drop the 'duration' column as required.
    b. Dropping of duplicate rows
    c. Profiling of the dataset using pandas-profiling to highlight the following:
        i. Number of Variables
        ii. Number of observations
        iii. Missing Data
        iv. Duplicate Rows
        v. Data size
        vi. Data Types
        vii. Distribution of Continuous Features
        viii. Cardinality of Categorical Features
        ix. Correlations of Continuous Features
    d. Nominal labelling of the following columns:
        i. job: nominal labelling and grouping of categories to address high cardinality
        ii. marital: nominal labelling
        iii. education: nominal labelling
        iv. default: nominal labelling
        v. housing: nominal labelling
        vi. contact: nominal labelling
        vii. campaign: nominal labelling and grouping of categories to address high cardinality
        viii. poutcome: nominal labelling
        ix. y: nominal labelling
        x. pdays: binning of past due days by bucket (e.g. 1, 30, 60 days, etc.) in order to reduce cardinality and for EDA purposes