

## Week 7 Deliverable

### Team Member's Details:

Name: Patrick Otieno  
Email: potieno578@gmail.com  
Country: Kenya  
College/Company: Umma University Specialization:  
Data Analyst

### Problem Description:

XYZ credit union in Latin America is performing very well in selling the Banking products (eg: Credit card, deposit account, retirement account, safe deposit box etc) but their existing customer is not buying more than 1 product which means bank is not performing good in cross selling (Bank is not able to sell their other offerings to existing customer). XYZ Credit Union decided to approach ABC analytics to solve their problem.

### Data Understanding

- **Type of Data:**
  - Customer demographic data
  - Transaction history
  - Product purchase behavior
  - Marketing engagement details
- **Problems Identified in Data:**
  - Missing values (NA values)
  - Presence of outliers
  - Skewed data distributions in key variables
  - Imbalanced classes (if applicable)

### Data Preprocessing & Approach

- **Addressing NA Values:**
  - Identify columns with missing data.
  - Strategies like mean/median imputation or predictive modeling to fill gaps.
- **Handling Outliers:**
  - Use statistical methods (e.g., IQR, Z-scores) to detect and cap/floor extreme values.
- **Fixing Skewness:**
  - Apply transformations like log or square root to normalize data.
- **Class Imbalance (if applicable):**
  - Use techniques like SMOTE or stratified sampling for balance.

**Why:** These steps ensure clean, reliable data for model training, improving analysis accuracy and predictive power.