**R Training Project 2**

The data was collected on selected households in Kenya from different regions on financial services usage. You have been asked to study the data and answer the following questions. You should perform all the statistical analysis required and write up the key findings. You should remember that the target audiences do not have a statistical background.

* Use the SPSS file: **“FinAccess\_SPSS.sav**

We want to address the following

1. Import the data to R in csv format (this is SPSS data) and name it financial\_data.csv
2. Clean the data before undertaking any analysis
3. Conduct exploratory analysis of the data and write a few bullet points on any descriptive statistics (summary statistics, tables and graphs) you find interesting and why you find them interesting.
4. Create a Cross Table of age\_coded and cluster\_type
5. Select the variables, age and monthly\_income. Test for normality. If possible, does monthly income depend on the age of the respondent?
6. Select the categorical variables on age and monthly income. Perform a Chi-square test. State what you are testing and your findings
7. Perform a one way ANOVA between Monthly income and highest education of the respondent. Does monthly income depend on the level of education?
8. Test for statistical significance using one sample t test to find out if the average income is less than Kshs. 15,000?
9. Using the generalized linear model, perform the following
10. Does mobile usage depend on gender, religion, education and credit usage?
11. Does bank usage depend on financial literacy, numeracy, savings usage and credit usage

**Remark:**

Save the work using Rmackdown for assessment