## 4. PROBLEM DESCRIPTION

The problem this project will address is that the relationship between the main survey result columns in the survey results public and survey results schema tables is not clearly laid out and thus insights about the data release cannot be obtained.

## 4.1. Techniques and tools:

We plan on using different data science techniques in R such as linear regression,Naïve Bayes classifier, KNN algorithm, decision trees, among others.

We also plan on using data tools in R such as data visualisation, random generation, advanced data visualisation.

## 4.2. Benefits

The following are some of the anticipated benefits at the end of this project;

* Obtaining insights about the relationship between salaries, machine learning and diversity in technology.
* Identifying the kind of users less likely to identify as part of the stack overflow community.
* Understanding the kind of users that are likely to participate in the stack overflow activities.
* Understanding those kinds of users that feel Kinship with fellow developers.

## 4.3. Project beneficiaries

The following are some off the beneficiaries of this project; Stack overflow developers, aspiring developers, applicants for jobs at stack overflow.

# 5. CONTACT INFORMATION

For more information,please contact our project chief executive on:

Mobile +256 705 395 400

Mobile +256 791 504 377

Email: faithnalumansi@gmail.com.