**One Acre Fund - Data Specialist role**

**Written assessment**

<< Scope and guidelines >>

* Relevant data can be found in the below zipper folder:

<https://www.dropbox.com/sh/piajywfsn96n9p6/AABGcjQ2lm0MIDqARM5wxoCba?dl=0>

* You can do the task using any software, but we prefer you demonstrate your knowledge in R/ Python at least for some part of the analysis.
* We are not looking for perfection in your assignment, but rather at the approach you have taken and the logic and considerations behind it.
* For the detailed method and procedure, please share your notes, codes or documents that explain how you came to the results.
* Please submit your written materials, code, and notes no later than 48 hours (2 days) after receiving this assignment.

**Assignment 1:**

One Acre Fund is a data-driven organisation, recognising the value and importance of data to support informed decision-making. As part of this position, you will be expected to independently lead data analysis workflows. This will require you to creatively address complex research and organisational questions. As part of this, you will have to rely on your ability to proactively review analytical options that could address these problems and test them. We want you to demonstrate this here.

In many cases, analytical questions will likely see you collating data from external sources and combining them with data from either One Acre Fund or the CGIAR to provide recommendations or outputs to support decision-making. One such required output is the spatial distribution of crops across Rwanda.

Imagine you have access to a ground-truthed dataset with the GPS locations of some farms producing banana throughout the country. You have been asked to use these data to *predict geospatially* the likelihood of the presence of banana across Rwanda. Provide an outline of an analysis you would perform to predict the distribution of banana across Rwanda. Provide a list of other data that would be needed to make these spatial predictions. Feel free to also provide references of other works that have implemented similar approaches.

**Deliverable 1:** Word Document that outlines your approach (3 pages maximum, including any figures and/or reference)

**Assignment 2:**

In One Acre Fund, it is key that the organisation understands their clients and to improve the products and interventions that are available to them. To do this, One Acre Fund collects extensive data from their clients through digital surveys. In the zipped folder you will find an example of such survey data- ‘Survey\_Data’. It represents one survey where more than 200 households were interviewed across a number of Rwandan districts. The survey contains data relating to general household details (e.g. age, education, household members), agronomic details (yields, fertiliser use), and economic/market details (income, percentages of crops sold).

Explore the dataset to analyse the effects of market orientation (percentage of crops cultivated that are sold, rather than consumed) on household dynamics, specifically looking at dietary diversity (number of food groups consumed) and income (off farm and on-farm). Also consider the effects of regionality on these factors. A rough and quick analysis is sufficient. Provide a script that shows:

1. how basic data quality control was done, including some visualisation of the data;
2. a statistical model used to quantify the effects of market orientation on total household income and dietary diversity;
3. suggest household characteristics which might explain market orientation.

In addition, provide a concise list of bullets with how you would propose next steps to investigate the specific effects that market orientation might have upon nutrient security in farming households of Rwanda and how One Acre Fund could use this information.

**Deliverable 2:** R/ Python script (.R/Py file) with your code and comments on next steps

**Good luck.**