Data Solutions Analyst Case Study

Scenario: Nithio is working with a distributed energy company that wants to enter the solar lighting market in Tanzania. We need to analyze the data and identify relevant features and patterns for our own understanding, in addition to answering specific questions for the client about the target market.

The client wants to know the following:

1. What does the energy and electricity market look like in Tanzania?
   - Think about major sources of lighting and energy, as well as variation in sources and consumption
2. Who or what type of households are most likely to purchase/need solar lighting systems?
   - Think about people who are not connected to the electricity grid and can afford to pay for solar lightning
3. Where are these households/people geographically concentrated?
   - Think about analyzing the data at different levels of geography
4. What's the total market size for their solar lighting products and which should they deploy first in Tanzania: (1) HomePro, $50; (2) PowerLite, $100; (3) SunMaster, $250
   - Think about total addressable market (TAM) in monetary terms or by the number of people/households that can potentially purchase the product.

Expectations:
1. Perform analysis on the prescribed data set and prepare a brief presentation of your results (6-8 slides. No more than 8 slides). During the case interview, you will present your slides and lead the rest of the team in a discussion. This will be a collaborative brainstorming activity. It will last roughly 60 minutes. Please send us the code that you used for the analysis on the day of the interview.
2. Write a 1-2 page (no more than 2 pages) brief on the problem statement, your analysis, your findings and recommendations. This brief should be complementary to your presentation.

Data: Tanzania National Panel Survey 2020-21, Wave 5

You will need to create an account with the World Bank and download the micro data for this survey. You may also find the questionnaires and other materials to be helpful. You will also only need to use the household modules and the consumption aggregate data.

Additional things to keep in mind:

- You are welcome to draw upon other research and data; however, the data from the link above should be the primary source for your analysis. As a hint, you may find population data from the World Bank useful.
- Please complete your analysis using R or Python. Since you will need to share your code via email, please include commented text to describe your assumptions.
- Think about what you are presenting and writing from the client perspective: ask yourself, is this useful for the client and why/how.
- Please understand the limitations of the analysis, but also push the envelope to really create "actionable" insights for the client.

What we value:

- Subnational/regional analysis
- Analysis and the presentation of the results in a clear and concise manner
- Be thoughtful about presenting information (charts, prose, bullets, figures) in the slides. You can also include charts and figures in the brief.
- Well structured brief that clearly communicates analysis objectives and finding with good overall flow.
- As much as we are interested in how you do your analysis, we are interested in how you interact with the team during the case interview.

You can address all questions to Njeri Kara (n.kara@nithio.com).