

Data Systems Analyst Technical Assessment- 19th July 2023 2.00-4.00pm

Kidogo is a social enterprise that improves access to high-quality, affordable early childhood care & education in East Africa's low-income communities. We partner with local women ("Mamapreneurs") to start or grow their own childcare centers through a social franchising approach. We provide 'Mamapreneurs' with training, mentorship and support tools so that they can give children in their care the best start to life, while their mothers go to work for the day. Kidogo currently operates in 30+ communities across eight counties and has plans to scale across Kenya.

Question.

You have been provided with an Excel file named 'kidogo_data_systems_analyst_work_assignment' that contains four tabs:

- 1. qip_data:** This tab includes information about daycares that have enrolled in Kidogo's Quality Improvement Program (QIP). The QIP is a three-month program designed to introduce new daycares to Kidogo's model before they sign a contract and become members. Each row represents a daycare with its details.
- 2. joining_session_data:** This tab contains data on the daycares that have completed the QIP program and signed a contract to become Kidogo members. Each row represents a community along with its daycares that have signed the contract. The daycares that have signed the contract are listed in the 'kidogo_code_signed_contract' column.
- 3. qip_metadata:** This tab serves as a data dictionary providing a description of the 'qip_data' tab.
- 4. joining_session_metadata:** This tab serves as a data dictionary providing a description of the 'joining_session_data' tab.

Task: Your task is to provide a summary report on the following metrics:

1. Conversion rate: Calculate the percentage of QIP daycares that transitioned into members per community.
2. Collection rate: Determine the average collection rate of the QIP fee per community.

Procedure:

1. Preprocessing: Utilize Python, preferably Google Colab, to preprocess the 'joining_session_data' and 'qip_data'. During this step, make note of any data quality issues encountered during the cleaning process. These issues should be included in your summary report.
2. Reporting: Present your findings on the key metrics mentioned above, including the Conversion rate and Collection rate.
3. Presentation: Create a PowerPoint or Google Slides presentation that you will deliver to the business stakeholders during a one-on-one interview. Identify and communicate any key insights and recommendations derived from your analysis.

Submission

1. Prepare a single .zip file that includes the following:

Nairobi, Kenya



- a. Cleaned data Excel file containing two tabs: 'cleaned_qip_data' and 'cleaned_joining_session_data'.
 - b. Jupyter Notebook file named 'data_preprocessing.ipynb' that documents the data cleaning process. Ensure your code includes useful comments for the reproducibility of your analysis during the assessment scoring.
 - c. A PowerPoint presentation file (.pptx) or a Google Slides presentation file (.ppt) showcasing key insights and recommendations from your analysis.
2. Name your file submission .zip file as "[Your Name]-[Data Systems Analyst Assignment]".
 3. You have 2 hours (from 2.00pm to 4.00pm) to complete this assignment and submit the zip file to hr@kidogo.co anything shared after 4.00pm will NOT be considered for grading.

Nairobi, Kenya

