Digital Lending Data Analyst Technical Test

Outline

This Digital Lending Data Analyst Technical Test outlines the assignment, from SQL assignment to extract and aggregate data from the SLIK dataset, including SLIK aggregated table, customer scorecard, and customer whitelist. Following this, it delves into data analysis and finally data visualization. The document details the dataset to be utilized and the specific deliverables expected from the test. Evaluation criteria are established to assess the quality and proficiency of the test submission, the review conducted will be based on this.

Context

As a Digital Lending Data Analyst, your role involves handling and analyzing various datasets to derive insights that drive business decisions. In this technical test, you will demonstrate your proficiency in SQL querying, data aggregation, analysis, and visualization using a specific dataset commonly utilized in our scenarios.

Assignment

A. Query

Your task involves creating SQL queries to extract and aggregate data from the SLIK dataset, focusing on key metrics and attributes relevant to digital lending analysis. In this assignment you should generate these tables. Use **Google BigQuery** to perform this task. SLIK = Sistem Layanan Informasi Keuangan (previously called SLIK while managed by BI, now called idebku managed by OJK), which is the financial regulator's database for loans.

1. SLIK aggregated table

   Aggregate the multiple tables to one output table that summarizes the loan information from individual customers. The loan application in this case is at **2023-11**

   Output Table

   a. **NIK:** Customer's unique identifier
   b. **flags_allcondition_count:** Facility count of loans in any condition.
   c. **flags_restructured_active_count:** Facility count of active loans that have been restructured
   d. **flags_chargewriteoff_count:** Facility count of loans that have been written off
   e. **plafond_credit_card_active_max:** Maximum plafond for active credit card facility
   f. **installment_active_max:** Maximum installment for active loans

g. **balance_personal_loan_active_sum:** Sum of balance for all active personal loans
h. **plafond_credit_card_active_sum:** Sum of plafond for all active credit card facility
i. **Installment_personal_loan_active_sum:** Sum of installment for active personal loans
j. **balance_credit_card_active_sum:** Sum of balance for all active credit card facility
k. **mob_active_avg:** Average months on book for all active loans.
l. **dpd_nonbank_allcondition_last_12months_max:** Maximum days past due for all nonbank loans in the 12 months prior to the loan application.
m. **collection_status_closed_unsecured_last_24months_max:** Maximum collection status for all closed, unsecured loans in the 24 months prior to the loan application.
n. **dpd_allcondition_last_3months_max:** Maximum days past due for all loans in the 3 months prior to the loan application.
o. **collection_status_allcondition_last_6months_max:** Maximum collection status for all loans in the 6 months prior to the loan application.
p. **collection_status_allcondition_last_12months_max:** Maximum collection status for all loans in the 6 months prior to the loan application.
q. **balance_unsecured_sum:** Sum of the balance for unsecured loans
r. **plafon_sum:** Sum of the plafond for all loans
s. **mob_allcondition_max:** Maximum months on book for all loans
t. **balance_sum:** Sum of the balance for all loans
u. **flags_active_dpd10plus_count:** Facility count of active loans that have been due more than 10 days
v. **flags_restructured_allcondition_unsecured_count:** Facility count of all unsecured loans that have been restructured

2. Customer scorecard

Create the customer scorecard which provides a comprehensive overview of key metrics and attributes associated with individual customers. This includes financial indicators, loan history, and product usage.

Output Table

a. NIK: Customer's unique identifier
b. SLIK Installment: Monthly CC outstanding *5% + Monthly Loan installment
c. SLIK Exposure: Monthly CC limit + Monthly Personal Loan outstanding
d. Max DPD (Days Past Due): Maximum number of days past due across all accounts
e. Max Col (Collections): Maximum collectability across all accounts
f. Number of CC (Active/Closed): Total number of active and closed credit card accounts

g. Number of Unsecured Personal Loans (Active/Closed): Total number of active and closed unsecured personal loan accounts
h. Number of Secured Personal Loans (Active/Closed): Total number of active and closed secured personal loan accounts

NIK	SLIK Installment	SLIK Exposure	Max DPD	Max Col	Number of CC Active	Number of CC Closed	Number of Unsecured Personal Loans Active	Number of Unsecured Personal Loans Closed	Number of Secured Personal Loans Active	Number of Secured Personal Loans Closed
123456	10,000	20,000	30	2	2	1	1	1	0	1
789012	15,000	30,000	60	3	1	1	1	0	1	1
...	...	...	...	...	...	...	...	...	...	...

3. Customer whitelist

Rules to be included in the whitelist are:

1. **No Active Col>2**
2. **No Active Restructure**
3. **No Write off**

Generate a table that lists whitelisted customers based on adherence to these rules. Additionally, if a customer is not whitelisted, the same table should indicate which rule each customer violated first.

Output Table

1. NIK: Customer's unique identifier
2. Result: Indicates whether the customer is whitelisted or not, along with the first rule they violated if applicable.

NIK	Result
123456	Whitelisted
789012	Active Collections > 2
345678	Active Restructure
901234	Write-off

B. Data Analysis

Once the dataset is aggregated based on the provided query, perform analysis based on the aggregated table you have created. These are the 3 topics you can explore (but not limited to):

- Segment the customer based on SLIK behavior
- Analyze characteristics by demographic factors.
- Evaluate customer scorecard metrics (from SLIK Installment and SLIK Exposure)

Candidates should proceed to **create the analysis with Looker Studio** (https://lookerstudio.google.com/) or **google sheets.**

C. Data Visualization

Using Looker Studio, create visualizations based on the analyses conducted. Visualize insights derived from the analysis.

Dataset

The dataset provided for this task is the SLIK dataset, which contains relevant information for digital lending analysis.

- **Dataset:**
  1. https://drive.google.com/file/d/1Hev21j-gQEOH3NyI3xS7ziN4QiwS3Hc4/view?usp=drive_link
  2. https://drive.google.com/file/d/1TarzLP47hawBdT9QKeG8SFkWNy2p13jr/view?usp=drive_link
  3. https://drive.google.com/file/d/1cTnQVPC4IFyBaYEW5BjW9PyHFI0ZvfMG/view?usp=drive_link
  4. https://drive.google.com/file/d/18NsTAWEgK3bYNZRC4qkgbJt_NPUesyDw/view?usp=drive_link

- **Additional Documents:**
  1. KODE REFERENSI (2).xlsx
  2. Pedoman Penyusunan Laporan SLIK - OJK

Deliverables

- Big Query SQL queries used for data extraction and aggregation. Queries should be pasted to **Google Docs** and attach the link when submitting this test.
- Aggregated datasets in Big Query as per the provided output requirements. Attach the **Table ID** (e.g `*project_name.dataset_name*.slik_aggregated_table`) of the output table when submitting the test. (Make sure share permission to yosira.jawata@jago.com, maria.sekar@jago.com, muhammad.darmawan@jago.com before hand)
- Analysis presented through **Looker Studio** or **Google Sheets** to report insights.

- Data visualizations created in **Looker Studio**

Evaluation Criteria

Candidates will be evaluated based on the following criteria:

- Accuracy and completeness of SQL queries
- Quality and depth of data analysis
- Clarity and relevance of insights derived
- Effectiveness of data visualization techniques
- Overall presentation and documentation of findings