**A blue and white logo

AI-generated content may be incorrect.SQL ASSESSMENT WORKBOOK V2 Step-1: SYSTEMIC ISSUE DETECTION**

**EduFin SQL Skill Assessment Workbook: A Beginner-Friendly Guide to SQL Queries, Business Logic, and Output Interpretation**

**Program: Skill AI Path – Data Analyst Pretraining Track  
Module: EduFin Risk Analytics Simulation  
Assessment Type: Skill Validation – SQL Query Writing + Business Understanding  
Prepared For: EduFin Data Analyst Cohort  
Organization: Krishnav Tech | Skill AI Path**

**Objective**

To assess and validate your foundational SQL skills, including query construction, systemic error detection, reconciliation analysis, and interpretation of institutional data quality, preparing you for **root cause investigation** in financial data analytics.

# SQL Skill Check Assessment – Pretraining Workbook

## PART A: Query Writing (60 points)

**Question 1 (10 points)**  
Write a query to count how many loans are missing an institution\_id, and show the percentage of such loans out of the total portfolio.

**Your Answer:**

**Question 2 (10 points)**  
Group institutions by partnership\_tier and show:

* total loan count,
* total loan amount,
* discrepancy (SUM(loan\_amount) – SUM(reported\_amount)).

**Your Answer:**

**Question 3 (10 points)**  
Calculate error rate (%) for each institution, where invalid loans are defined as:

* missing disbursement\_date,
* or loan\_amount <= 0.

**Your Answer:**

**Question 4 (10 points)**  
Write a reconciliation query showing, for each institution:

* actual\_disbursed (from loans),
* reported\_disbursed (from reported\_amount),
* discrepancy value.  
  Only include institutions with discrepancies above ₹1 crore.

**Your Answer:**

**Question 5 (10 points)**  
List all loans where payments exist in the payments table but no disbursement\_date is recorded in the loans table. Include institution\_name, loan\_id, and payment\_amount.

**Your Answer:**

**Question 6 (10 points)**  
Create a CASE-based classification of institutions:

* “Clean” (<2% errors),
* “Moderate Risk” (2–10%),
* “High Risk” (>10%).  
  Return institution\_name, error\_rate\_percent, and risk\_category.

**Your Answer:**

## PART B: Multiple Choice Questions (40 points)

**Question 7:**  
Which clause is used to filter grouped data?  
A) WHERE  
B) HAVING  
C) GROUP BY  
D) ORDER BY

**Answer:**

**Question 8:**  
In reconciliation queries, why calculate SUM(actual) – SUM(reported) instead of comparing row by row?  
A) Reduces NULL errors  
B) Captures total discrepancies at institution level  
C) Improves performance  
D) Required by SQL standard

**Answer:**

**Question 9:**  
What does this query return?

SELECT COUNT(\*) - COUNT(institution\_id) FROM loans;

A) Number of loans with NULL institution\_id  
B) Number of unique institutions  
C) Number of total loans  
D) Number of duplicate loans

**Answer:**

**Question 10:**  
Which function would you use to replace NULL loan\_amounts with 0?  
A) NULLIF()  
B) COALESCE(loan\_amount, 0)  
C) ROUND()  
D) FORMAT()

**Answer:**

**Question 11:**  
If a loan has loan\_amount = -5000, it is considered:  
A) Clean record  
B) Invalid record  
C) Missing record  
D) Duplicate record

**Answer:**

**Question 12:**  
What is the purpose of a CASE statement in systemic validation queries?  
A) To sort results  
B) To apply conditional classifications  
C) To calculate averages  
D) To join two tables

**Answer:**

**Question 13:**  
Why join payments with loans when checking reconciliation?  
A) To improve formatting  
B) To validate that every payment maps to a valid loan  
C) To increase row counts  
D) To reduce missing data

**Answer:**

**Question 14:**  
Which SQL keyword ensures institutions with >₹1 crore discrepancies can be filtered after aggregation?  
A) WHERE  
B) HAVING  
C) GROUP BY  
D) CASE

**Answer:**

**Question 15:**  
What does ROUND(value, 2) do in an error rate calculation?  
A) Rounds to 2 decimal places  
B) Returns 2 records only  
C) Converts to integer  
D) Multiplies value by 2

**Answer:**

**Question 16:**  
Which of the following best describes systemic issue detection using SQL?  
A) Checking only customer demographics  
B) Identifying and quantifying process gaps via queries  
C) Formatting results for executives  
D) Writing SELECT queries without JOINs

**Answer:**

## PASSING CRITERIA

* Minimum Score Required: **80 out of 100**
* Query Writing: **At least 48/60**
* MCQ Section: **At least 32/40**
* Time Limit: **90 minutes**
* Retakes Allowed: Unlimited until 80% is achieved

## SKILLS YOU VALIDATE

By completing this assessment, you demonstrate:

* Systemic issue detection using SQL queries
* Reconciliation of reported vs actual financial data
* Error rate measurement and classification
* NULL handling and data completeness checks
* Use of CASE, GROUP BY, and HAVING in real-world reporting

© 2025 Skill AI Path. All Rights Reserved.  
Created by: Viresh Gendle | [tech@skillaipath.com](mailto:tech@skillaipath.com) | [skillaipath.com](http://skillaipath.com/)  
Licensed for educational and internal business use.  
Attribution required when sharing. Commercial redistribution prohibited.