User Acceptance Test Document

Project: Data Warehouse Validation for Airbnb Dataset

Version: 1.0

Date: June 30, 2025

Prepared by: [Your Name]

# 1. Purpose

This document outlines the acceptance testing process to validate that the data pipeline correctly loads and transforms data from the Data Lake to the Data Warehouse. This ensures data integrity, accuracy, and completeness.

# 2. Scope

The scope of this test includes:  
- Verification that today's data in the Data Lake matches today's data in the Data Warehouse.  
- Batch ID consistency validation between the Data Lake and the Data Warehouse.  
- Basic row count checks for dimensional and fact tables.

# 3. Acceptance Criteria

|  |  |  |
| --- | --- | --- |
| Test ID | Test Description | Pass Criteria |
| TC001 | Row Count Check - Source vs Target | Row counts must match between source and target tables. |
| TC002 | Data Match Check - Source vs Target (by Surrogate Key) | No mismatches in key fields or row data. |
| TC003 | Batch ID Validation | Latest batch ID must match in both Data Lake and Warehouse. |

# 4. Test Cases

## Test Case TC001: Row Count Validation

Objective: Verify row count parity between source and target.  
Tables: RAW.REVIEWS → DEV.FCT\_REVIEWS (and similar for others)  
Expected Result: Row counts match exactly.

## Test Case TC002: Data Match on Key and Row Hash

Objective: Verify data accuracy at row level using surrogate key comparison.  
Logic:  
- Generate row hashes based on key columns (REVIEW\_SK or others).  
- Ensure each key exists in both tables.  
- Confirm that row hashes are identical for matching keys.  
Expected Result: Zero mismatches in keys and rows.

## Test Case TC003: Batch ID Match Validation

Objective: Confirm that the most recent batch ID loaded into the Data Warehouse matches the latest batch in the Data Lake.  
Validation: Query both layers for the max batch ID.  
Expected Result: MAX(batch\_id) matches in both source and target.

# 5. Test Execution and Sign-off

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test ID | Pass/Fail | Notes/Remarks | Tested By | Date |
| TC001 |  |  |  |  |
| TC002 |  |  |  |  |
| TC003 |  |  |  |  |

# 6. Approval

|  |  |  |  |
| --- | --- | --- | --- |
| Role | Name | Signature | Date |
| Data Engineer |  |  |  |
| Data Owner / SME |  |  |  |
| Project Manager |  |  |  |