

## 1. Data Quality Checks

Bagian akhir kode menyediakan *data quality checks* seperti pengecekan completeness, orphan records, invalid values, dan duplikasi, serta menambahkan contoh query analitik untuk menguji performa dan melihat insight seperti trend enrollment, distribusi GPA, funnel admission, hingga hubungan antara absensi dan nilai.

## Data Quality Checks

Query:

```
-- Data Quality Checks
-----
-- 1. Completeness check
SELECT COUNT(*) AS NullName
FROM dbo.Dim_Student
WHERE FullName IS NULL;
GO

-- 2. Orphan check Fact <-> Dimension
SELECT COUNT(*) AS Orphan_Students
FROM dbo.Fact_Enrollment f
LEFT JOIN dbo.Dim_Student d ON f.StudentKey = d.StudentKey
WHERE d.StudentKey IS NULL;
GO

-- 3. Numeric grade validity
SELECT COUNT(*) AS InvalidGrades
FROM dbo.Fact_Enrollment
WHERE NumericGrade NOT BETWEEN 0 AND 4;
GO

-- 4. Duplicate fact combinations
SELECT StudentKey, CourseKey, SemesterKey, COUNT(*) AS Cnt
FROM dbo.Fact_Enrollment
GROUP BY StudentKey, CourseKey, SemesterKey
HAVING COUNT(*) > 1;
GO
```

The screenshot shows the SQL Server Management Studio interface. The Object Explorer on the left shows the database structure for 'akademik'. The central pane displays a T-SQL script named 'SQLQuery1.sql' with the following code:

```
-- 1. Completeness: cek NULL di Dim_Student
SELECT
    Dim_Student AS TableName,
    COUNT(*) AS TotalRows,
    SUM(CASE WHEN StudentNaturalID IS NULL THEN 1 ELSE 0 END) AS NullNIM,
    SUM(CASE WHEN FullName IS NULL THEN 1 ELSE 0 END) AS NullName,
    SUM(CASE WHEN Gender IS NULL THEN 1 ELSE 0 END) AS NullGender
FROM dbo.Dim_Student;
GO
```

The results pane shows a single row of data from the query:

TableName	TotalRows	NullNIM	NullName	NullGender
Dim_Student	0	NULL	NULL	NULL

At the bottom, a message indicates the query was executed successfully.

Output menunjukkan bahwa pada tabel Dim\_Student tidak ada nilai NULL.

The screenshot shows the SQL Server Management Studio interface. The Object Explorer on the left shows the database structure for 'akademik'. The central pane displays a T-SQL script named 'SQLQuery1...Pongo (76)\*' with the following code:

```
-- 2. Consistency: Orphan fact vs dim
SELECT
    COUNT(*) AS OrphanStudentFact
FROM dbo.Fact_Enrollment f
LEFT JOIN dbo.Dim_Student s ON f.StudentKey = s.StudentKey
WHERE s.StudentKey IS NULL;
GO
```

The results pane shows a single row of data from the query:

OrphanStudentFact
0

At the bottom, a message indicates the query was executed successfully.

Hasil pengecekan konsistensi data menunjukkan bahwa tidak terdapat *orphan record* pada tabel *Fact\_Enrollment*. Seluruh nilai *StudentKey* pada tabel fakta berhasil mencocokkan data yang ada pada tabel *Dim\_Student*, sehingga integritas referensial antara fakta dan dimensi berada dalam kondisi baik.

The screenshot shows the SQL Server Management Studio (SSMS) interface. On the left is the Object Explorer pane, which displays the database structure for 'BITLANCKA\MSSQLSERVER01' (SQL Server 16.0.1000.6). It includes nodes for Databases, Tables (containing System Tables, FileTables, External Tables, Graph Tables, and various fact and dimension tables), Views, and Security. The right side features a 'SQLQuery1....Pongo (76)\*' query window containing the following T-SQL code:

```
-- 3. Accuracy: numeric grade di luar range
SELECT COUNT(*) AS InvalidNumericGrade
FROM dbo.Fact_Enrollment
WHERE NumericGrade < 0 OR NumericGrade > 4;
GO
```

Below the query window is a 'Results' tab showing the output:

	InvalidNumericGrade
1	0

A status bar at the bottom indicates 'Query executed successfully.'

Tidak ditemukan nilai NumericGrade yang berada di luar rentang 0 sampai 4. Semua nilai NumericGrade pada tabel Fact\_Enrollment valid dan sesuai standar penilaian

The screenshot shows the SSMS interface with the following details:

- Object Explorer:** Shows the database structure for "BITLANCKA\MSSQLSERVER01 (SQL Server 16.0.1000.6)".
- SQLQuery1....Pongo (76)\*:** The current query window.
- Code:** A T-SQL query titled "4. Duplicate business key".

```
-- 4. Duplicate business key
SELECT
    StudentKey, CourseKey, SemesterKey,
    COUNT(*) AS DuplicateCount
FROM dbo.Fact_Enrollment
GROUP BY StudentKey, CourseKey, SemesterKey
HAVING COUNT(*) > 1;
GO
```
- Results:** The results pane is empty, showing the header "StudentKey CourseKey SemesterKey DuplicateCount".
- Status Bar:** "Query executed successfully." and connection information.

Pengecekan duplikasi pada tabel Fact\_Enrollment menunjukkan bahwa tidak terdapat kombinasi StudentKey, CourseKey, dan SemesterKey yang muncul lebih dari satu kali. Hal ini menandakan bahwa data transaksi enrollment tidak memiliki duplikasi dan sudah memenuhi aturan grain yang ditetapkan.

The screenshot shows the SQL Server Management Studio (SSMS) interface. On the left is the Object Explorer pane, which lists the database 'akademik' under 'BITLANCKA\MSSQLSERVER01 (SQL Server 16.0.1000.6)'. The 'Tables' node under 'akademik' is expanded, showing various tables like 'stg.Enrollment', 'dbo.Dim.Course', etc. The main area is a query editor titled 'SQLQuery1....Pongo (76)\*'. It contains the following T-SQL code:

```
-- 5. Rekonsiliasi jumlah record antara staging vs warehouse
SELECT 'stg.Enrollment' AS Source, COUNT(*) AS RecordCount
FROM stg.Enrollment
UNION ALL
SELECT 'Fact_Enrollment' AS Source, COUNT(*) AS RecordCount
FROM dbo.Fact_Enrollment;
GO
```

The results pane shows a single row of data:

Source	RecordCount
stg.Enrollment	0
Fact_Enrollment	0

At the bottom of the results pane, a message says 'Query executed successfully.'

Hasil rekonsiliasi antara tabel staging dan tabel warehouse menunjukkan bahwa kedua tabel, baik stg.Enrollment maupun Fact\_Enrollment, memiliki jumlah record 0. Hal ini menandakan bahwa belum terdapat data yang masuk ke staging, atau proses ETL untuk memuat data ke tabel Fact\_Enrollment belum dijalankan