### ****1.1 Most Improved Manufacturer (Best Description Quality)****

SELECT [Manufacturer name], COUNT(\*) AS total\_products,

SUM(CASE WHEN [Description Quality] = 'good' THEN 1 ELSE 0 END) AS good\_quality\_count,

ROUND(SUM(CASE WHEN [Description Quality] = 'good' THEN 1 ELSE 0 END)\*100.0/COUNT(\*), 2) AS good\_quality\_percentage

FROM product\_data

GROUP BY [Manufacturer name]

ORDER BY good\_quality\_percentage DESC, good\_quality\_count DESC

LIMIT 1

**Explanation:**  
Identifies the manufacturer with the highest percentage of good product descriptions.  
Ranks by both quality % and total number of good-quality entries.  
Highlights which manufacturer currently leads in product content quality.  
Useful for recognition, benchmarking, or identifying best practices.

**1.2 Manufacturers with Bad Quality Descriptions**

SELECT [Manufacturer name], COUNT(\*) AS total\_products,

SUM(CASE WHEN [Description Quality] = 'bad' THEN 1 ELSE 0 END) AS bad\_quality\_count,

ROUND(SUM(CASE WHEN [Description Quality] = 'bad' THEN 1 ELSE 0 END)\*100.0/COUNT(\*), 2) AS bad\_quality\_percentage

FROM product\_data

GROUP BY [Manufacturer name]

ORDER BY bad\_quality\_count DESC

**Explanation:**  
This query lists all manufacturers and calculates how many of their products have bad descriptions.  
It also gives the percentage of bad quality descriptions out of total products.  
Helps identify manufacturers with the most content quality issues.  
Prioritizes improvement efforts based on actual data volume.

**2. Field Completion Rates by Manufacturer**

SELECT [Manufacturer name], 'Field Name' AS field\_name,

ROUND(SUM(CASE WHEN [Field Name] IS NOT NULL THEN 1 ELSE 0 END)\*100.0/COUNT(\*), 2) AS completion\_rate

FROM product\_data

GROUP BY [Manufacturer name]

**Explanation:**  
This is repeated per field to calculate how well each manufacturer fills key fields.  
It returns a % showing field completeness (e.g., for EAN, image, technical details).  
Highlights gaps in data entry by field and manufacturer.  
Supports targeted actions for improving data completeness.

**3.1 EAN vs Description Quality Correlation (Additional insights)**

SELECT CASE WHEN [EAN] IS NULL THEN 'Missing EAN' ELSE 'Has EAN' END AS ean\_status,

ROUND(SUM(CASE WHEN [Description Quality] = 'bad' THEN 1 ELSE 0 END)\*100.0/COUNT(\*), 2) AS bad\_description\_percentage

FROM product\_data

GROUP BY ean\_status

**Explanation:**  
Compares product quality based on whether EAN is present or missing.  
Shows % of bad descriptions in both groups.  
Validates if missing EAN is a strong indicator of poor product content.  
Helps enforce EAN compliance for better overall quality.

**3. 2 Missing Field Combinations (Additional Insights)**

SELECT SUM(CASE WHEN [Field] IS NULL THEN 1 ELSE 0 END) AS missing\_field, COUNT(\*) AS total\_records

FROM product\_data

✅ **Explanation:**  
Counts how many records are missing each important product field.  
Helps quantify the overall scale of missing information.  
Useful for spotting patterns in poor data uploads or integrations.  
Forms the basis for quality control and cleanup strategy.