QUALITY CONTROL FORM

OFFUTURE

The future of office supplies



Version 1.0.0

02/05/2024

# Quality Control

# Quality Control Objectives

Data Cleaning and Validation:

* Remove or correct inaccuracies, inconsistencies, and errors.
* Validate data against predefined rules, handling missing or null values.

Quality Control Checks:

* Perform thorough checks to verify data integrity.
* Maintain integrity throughout cleaning and transformation.

Documentation and Reporting:

* Document all procedures and provide relevant comments.

Marking Data Passed QC:

* Implement markers to indicate data that has passed quality control checks.
* Clearly label data that meets QC standards for easy identification.

# 

# Quality Control ID’s

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID of SQL or Checked Rows** | **Description of Check** | **Result in SOURCE (Excel)** | **Result in DESTINATION (SQL)** | **Outcome** | **Comments** |
| QC1 | Count of rows | 51290 | 51290 | PASS |  |
| QC2 | Count of Distinct rows | 51290 | 51290 | PASS |  |
| QC3 | Count of Columns | 24 | 24 | PASS |  |
| QC4 | Sum of column Sums | 1882578769 | 1882578768.6 | PASS | Slight discrepancy due to rounding in Excel |
| QC5 | Sum of Row Sums | 1882578769 | 1882578768.6 | PASS | Slight discrepancy due to rounding in Excel |
| QC6 | Ship Date format Check  5 randomly chosen values from Source checked against Destination | ***24296*** *– 11/11/2014*  ***26165*** *– 18/02/2013*  ***8510*** *– 06/09/2014*  ***16325*** *– 08/01/2013*  ***17284*** *– 13/06/2011* | ***24296*** *– 11/11/2014*  ***26165*** *– 18/02/2013*  ***8510*** *– 06/09/2014*  ***16325*** *– 08/01/2013*  ***17284*** *– 13/06/2011* | PASS | Dates match in same format |
| QC7 | Order Date format Check  5 randomly chosen values from Source checked against Destination | ***47241*** *– 03/03/2011*  ***12879*** *– 04/12/2012*  ***36939*** *– 02/02/2013*  ***26444*** *– 02/03/2012*  ***549*** *–25/11/2013* | ***47241*** *– 03/03/2011*  ***12879*** *– 04/12/2012*  ***36939*** *– 02/02/2013*  ***26444*** *– 02/03/2012*  ***549*** *–25/11/2013* | PASS | Dates match in same format |
| QC8 | Eyeball Check  Compare 5 randomly chosen rows from Source against Destination | 40035  13007  10281  31775  47441 | 40035  13007  10281  31775  47441 | PASS- | All values matching |

# Null Value ID’s

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID of SQL or Checked Rows** | **Description of Check** | **Result in SOURCE (Excel)** | **Result in DESTINATION (SQL)** | **Outcome** | **Comments** |
| **QC8.1** | Count of Nulls - Row ID | 0 | 0 | PASS |  |
| **QC8.2** | Count of Nulls - Order ID | 0 | 0 | PASS |  |
| **QC8.3** | Count of Nulls - Order Date | 0 | 0 | PASS |  |
| **QC8.4** | Count of Nulls - Ship Date | 0 | 0 | PASS |  |
| **QC8.5** | Count of Nulls - Ship Mode | 0 | 0 | PASS |  |
| **QC8,6** | Count of Nulls - Customer ID | 0 | 0 | PASS |  |
| **QC8.7** | Count of Nulls - Customer Name | 0 | 0 | PASS |  |
| **QC8.8** | Count of Nulls - Segment | 0 | 0 | PASS |  |
| **QC8.9** | Count of Nulls - City | 0 | 0 | PASS |  |
| **QC8.10** | Count of Nulls - State | 0 | 0 | PASS |  |
| **QC8.11** | Count of Nulls - Country | 0 | 0 | PASS |  |
| **QC8.12** | Count of Nulls - Postal Code | 41296 | 41296 | PASS |  |
| **QC8.13** | Count of Nulls - Market | 0 | 0 | PASS |  |
| **QC8.14** | Count of Nulls - Region | 0 | 0 | PASS |  |
| **QC8.15** | Count of Nulls - Product ID | 0 | 0 | PASS |  |
| **QC8.16** | Count of Nulls - Category | 0 | 0 | PASS |  |
| **QC8.17** | Count of Nulls - Sub-Category | 0 | 0 | PASS |  |
| **QC8.18** | Count of Nulls - Product Name | 0 | 0 | PASS |  |
| **QC8.19** | Count of Nulls - Sales | 0 | 0 | PASS |  |
| **QC8.20** | Count of Nulls - Quantity | 0 | 0 | PASS |  |
| **QC8.21** | Count of Nulls - Discount | 0 | 0 | PASS |  |
| **QC8.22** | Count of Nulls - Profit | 0 | 0 | PASS |  |
| **QC8.23** | Count of Nulls - Shipping Cost | 0 | 0 | PASS |  |
| **QC8.24** | Count of Nulls - Order Priority | 0 | 0 | PASS |  |

# sum/min/max/avg ID’s

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID of SQL or Checked Rows** | **Description of Check** | **Result in SOURCE (Excel)** | **Result in DESTINATION (SQL)** | **Outcome** | **Comments** |
| **QC9.1** | Sum | Min | Max | AVG – Row ID | 1315357695  1  51290  25645.5 | 1315357695  1  51290  25645.5 | PASS |  |
| **QC9.2** | Sum | Min | Max | AVG – Sales | 12642501.91  0.444  22638.48  246.4905812 | 12642501.91  0.444  22638.48  246.4905812 | PASS |  |
| **QC9.3** | Sum | Min | Max | AVG – Quantity | 178312  1  14  3.47655 | 178312  1  14  3.47655 | PASS |  |
| **QC9.4** | Sum | Min | Max | AVG – Discount | 7329.728  0  0.85  0.142908 | 7329.728  0  0.85  0.142908 | PASS |  |
| **QC9.5** | Sum | Min | Max | AVG – Profit | 1467457.291  -6599.978  8399.976  28.61098248 | 1467457.291  -6599.978  8399.976  28.61098248 | PASS |  |
| **QC9.6** | Sum | Min | Max | AVG – Shipping Cost | 1352820.69  0  933.57  26.37591519 | 1352820.69  0  933.57  26.37591519 | PASS |  |
| **QC9.7** | Sum | Min | Max | AVG – Postal Code | 551572652  1040  99301  55190.3794 | 551572652  1040  99301  55190.3794 | PASS |  |
| **QC10.1** | Sum | Min | Max | AVG – LENGTH OF Order ID | 695987  9  15  13.56964321 | 695987  9  15  13.56964321 | PASS |  |
| **QC10.2** | Sum | Min | Max | AVG – LENGTH OF Order Date | 512900  10  10  10 | 512900  10  10  10 | PASS |  |
| **QC10.3** | Sum | Min | Max | AVG – LENGTH OF Ship Date | 512900  10  10  10 | 512900  10  10  10 | PASS |  |
| **QC10.4** | Sum | Min | Max | AVG – LENGTH OF Ship Mode | 658721  8  14  12.84306882 | 658721  8  14  12.84306882 | PASS |  |
| **QC10.5** | Sum | Min | Max | AVG – LENGTH OF Customer ID | 400812  5  8  7.81462273 | 400812  5  8  7.81462273 | PASS |  |
| **QC10.6** | Sum | Min | Max | AVG – LENGTH OF Customer Name | 664012  7  22  12.94622733 | 664012  7  22  12.94622733 | PASS |  |
| **QC10.7** | Sum | Min | Max | AVG – LENGTH OF Segment | 453778  8  11  8.84729967 | 453778  8  11  8.84729967 | PASS |  |
| **QC10.8** | Sum | Min | Max | AVG – LENGTH OF City | 431826  2  35  8.419302008 | 431826  2  35  8.419302008 | PASS |  |
| **QC10.9** | Sum | Min | Max | AVG – LENGTH OF State | 494486  3  36  9.640982648 | 494486  3  36  9.640982648 | PASS |  |
| **QC10.10** | Sum | Min | Max | AVG – LENGTH OF Country | 453233  4  32  8.836673816 | 453233  4  32  8.836673816 | PASS |  |
| **QC10.11** | Sum | Min | Max | AVG – LENGTH OF Market | 18540  2  6  3.6149 | 18540  2  6  3.6149 | PASS |  |
| **QC10.12** | Sum | Min | Max | AVG – LENGTH OF Region | 340496  4  14  6.63864301 | 340496  4  14  6.63864301 | PASS |  |
| **QC10.13** | Sum | Min | Max | AVG – LENGTH OF Product ID | 779352  15  16  15.19500877 | 779352  15  16  15.19500877 | PASS |  |
| **QC10.14** | Sum | Min | Max | AVG – LENGTH OF Category | 659389  9  15  12.85609281 | 659389  9  15  12.85609281 | PASS |  |
| **QC10.15** | Sum | Min | Max | AVG – LENGTH OF Sub-Category | 370852  3  11  7.230493274 | 370852  3  11  7.230493274 | PASS |  |
| **QC10.16** | Sum | Min | Max | AVG – LENGTH OF Product Name | 1582652  5  127  30.85693118 | 1582652  5  127  30.85693118 | PASS |  |
| **QC10.17** | Sum | Min | Max | AVG – LENGTH OF Order Priority | 277330  3  8  5.4070969 | 277330  3  8  5.4070969 | PASS |  |

# Comments

Excel (Source):

* Well organized dataset. Labels columned well. “Sales” could be renamed to “Revenue” for better clarification.
* A lot of missing values for postal code. May need revision for next excel sheet.
* Could have included specificities such as currency of profit/sales.
* Market category segmentation could be improved for better for analysis usage.
* Values of sales/profit could be standardized to 2 d.p. for more accurate usage when executing queries.

SQL:

* No issues importing
* No issues in manipulating data from excel for queries

For any further questions/details. Please contact our team at digital futures for assistance.