****

**Quality Control Form**

|  |  |
| --- | --- |
| **Team Name** | Call of Data |
| **Team Members** | Ben, Riya, Mary |
| **Client Name** | Mystic Manuscripts |
| **Date of file received** | 15.01.2024 |
| **Name of file received** | Books Sales Records V2 |
| **File Format received** | CSV |
| **Size of file (KB) received** | 1.53 MB |
| **Encoding of the file** | UTF-8 |
| **Recorded Number of Columns** | 36 |
| **Recorded Number of Rows** | 3,481 (*including header*) |
| **Name of Schema of Destination Table** | **student** |
| **Name of Destination Table** | **call\_of\_data** |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ID of SQL or Checked Rows** | **Description of Check** | **Result in SOURCE** | | | **Result in DESTINATION** | | | **OUTCOME** |
| **1.** | **Count of rows** | 3480 | | | 3480 | | | **PASS** |
| **2.** | **Count of Distinct Rows** | 3480 | | | 3480 | | | **PASS** |
| **3.** | **Count of Columns** | 36 | | | 36 | | | **PASS** |
| **4.** | **Sum of column Sums** | 1481390811.79 | | | 1481390811.79 | | | **PASS** |
| **5.** | **Sum of Row Sums** | 1,481,390,811.79 | | | 1481390811.79 | | | **PASS** |
| **6.** | **Date format Check** | ***Format  ROW ID – Data Example:*** *100 – 13/12/2024 (DD/MM/YYYY)* | | | Row ID 100 Date 2024-12-13 | | | **PASS** |
|  | **1.** | ROW ID 5  Date Purchase Date  04-01-2018 01-04-2018 20:12 | | | ROW ID 5  Date Purchase Date  04-01-2018 01-04-2018 20:12 | | |  |
|  | **2.** | ROW ID 200  Date Purchase Date  20-02-2018 20-02-2018 16:53 | | | ROW ID 200  Date Purchase Date  20-02-2018 20-02-2018 16:53 | | |  |
|  | **3.** | ROW ID 300  Date Purchase Date  07-03-2018 03-07-2018 08:59 | | | ROW ID 300  Date Purchase Date  07-03-2018 03-07-2018 08:59 | | |  |
|  | **4.** | ROW ID 600  Date Purchase Date  28-04-2018 28-04-2018 06:14 | | | ROW ID 600  Date Purchase Date  28-04-2018 28-04-2018 06:14 | | |  |
|  | **5.** | ROW ID 750  Date Purchase Date  13-05-2018 13-05-2018 10:23 | | | ROW ID 750  Date Purchase Date  13-05-2018 13-05-2018 10:23 | | |  |
| **7.** | **Eyeball Check** | Row ID: 71 ,742, 1494 ,2072, 3256 | | |  | | | **PASS** |
|  |  |  | | |  | | |  |
|  |  |  | | |  | | |  |
| **8.** | **Count of Nulls -** |  | | |  | | | **PASS** |
|  | ID | 0 | | | 0 | | |  |
|  | Purchase-date | 0 | | | 0 | | |  |
|  | Date | 0 | | | 0 | | |  |
|  | Time | 0 | | | 0 | | |  |
|  | Quarter | 0 | | | 0 | | |  |
|  | Qtr | 0 | | | 0 | | |  |
|  | Year | 0 | | | 0 | | |  |
|  | Month | 0 | | | 0 | | |  |
|  | Date | 0 | | | 0 | | |  |
|  | Special Day | 0 | | | 0 | | |  |
|  | Online Sale Offers | 0 | | | 0 | | |  |
|  | Day | 0 | | | 0 | | |  |
|  | Weekend | 0 | | | 0 | | |  |
|  | Morning | 0 | | | 0 | | |  |
|  | Afternoon | 0 | | | 0 | | |  |
|  | Evening | 0 | | | 0 | | |  |
|  | Night | 0 | | | 0 | | |  |
|  | Gender | 0 | | | 0 | | |  |
|  | Customer ID | 0 | | | 0 | | |  |
|  | GenderFM | 0 | | | 0 | | |  |
|  | Product-Name | 0 | | | 0 | | |  |
|  | Item-Status | 0 | | | 0 | | |  |
|  | Quantity | 0 | | | 0 | | |  |
|  | Currency | 0 | | | 0 | | |  |
|  | Item Price | 0 | | | 0 | | |  |
|  | Shipping-Price | 0 | | | 0 | | |  |
|  | Ship-City | 0 | | | 0 | | |  |
|  | Ship-State | 0 | | | 0 | | |  |
|  | Ship-Postal-Code | 0 | | | 0 | | |  |
|  | Category | 0 | | | 0 | | |  |
|  | Total\_amount | 0 | | | 0 | | |  |
|  | Author | 0 | | | 0 | | |  |
|  | Publication | 0 | | | 0 | | |  |
|  | Profit\_Percentage | 0 | | | 0 | | |  |
|  | Profit (INR) | 0 | | | 0 | | |  |
|  | Cost Price | 0 | | | 0 | | |  |
|  |  |  |  |  |  |  |  |  |
| **9a, 9b, 9c** | **Sum| Min| Max** | **Sum** | **Min** | **Max** | **Sum** | **Min** | **Max** | **PASS** |
|  | Cost Price | 1,106,422.61 | 34.80 | 5,611.50 | 1,106,422.61 | 34.80 | 5,611.50 |  |
|  | Profit (INR) | 208,139.18 | 5.20 | 838.50 | 208,139.18 | 5.20 | 838.50 |  |
|  | Profit\_Percentage | 57,981 | 13 | 18 | 57,981 | 13 | 18 |  |
|  | Total\_amount | 1,589,599 | 80 | 7,300 | 1,589,599 | 80 | 7,300 |  |
|  | Ship-Postal-Code | 1,463,652,909 | 110,001 | 855,107 | 1,463,652,909 | 110,001 | 855,107 |  |
|  | Shipping-Price | 275,019 | 0 | 850 | 275,019 | 0 | 850 |  |
|  | Item Price | 1,314,521 | 40 | 6,450 | 1,314,521 | 40 | 6,450 |  |
|  | Quantity | 3,655 | 1 | 10 | 3,655 | 1 | 10 |  |
|  | Date\_1 | 51,551 | 1 | 31 | 51,551 | 1 | 31 |  |
|  | Month | 20,701 | 1 | 12 | 20,701 | 1 | 12 |  |
|  | Year | 7,024,924 | 2,018 | 2,019 | 7,024,924 | 2,018 | 2,019 |  |
|  | Qtr | 8,194 | 1 | 4 | 8,194 | 1 | 4 |  |
|  | ID | 6,056,940 | 1 | 3,480 | 6,056,940 | 1 | 3,480 |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | **ADDITIONAL CHECKS** |  |  |  |  |  |  |  |
|  | There were discrepancies for few column names:  (The issues persisted for file version -2 as well) | --Profit Percentage  (%)  --Total\_  amount  --Gender  F=1665  M=1815 |  |  |  |  |  | **FAIL (\*\*)** |
|  | Special character check | Â – were encountered in data file. |  |  |  |  |  | **PASS** |
|  | Checked the data types for each column according to data. |  |  |  |  |  |  | **PASS** |
| **10.** | Checked for the duplicate rows in the data. | 0 |  |  | 0 |  |  | **PASS** |
| **11.** | Mean of the columns sums | 67,335,946 |  |  | 67,335,946 |  |  | **PASS** |
| **12.** | Mean of the row sums | 425,687 |  |  | 425,687 |  |  | **PASS** |

**NOTE**:

1. If the data entered into the column “purchase date” is followed by the timestamp standard, it could be interpreted as the format and used in further calculations.
2. (\*\*) The following columns names were not appropriate in version 1 and version 2. The client granted permission to change it. We also recommend to avoid including a newline symbol in the column name for a more correct reflection in the database in future.

* Profit Percentage -> Profit\_Percentage

(%)

* Total\_ -> Total\_amount

amount

* Gender -> GenderFM

F=1665

M=1815