****  
  
**Quality Control Form**

|  |  |
| --- | --- |
| **Team Name** | WHERE |
| **Team Members** | Isabel Benson, Lorcan Donnellan, Harvir Singh |
| **Client Name** | Mystic Manuscript |
| **Date of file received** | 22/09/2025 |
| **Name of file received** | Books Sales Records – V2 |
| **File Format received** | CSV (comma delimited) |
| **Size of file (KB) received** | 932 |
| **Encoding of the file** | UTF-8 |
| **Recorded Number of Columns** | 36 |
| **Recorded Number of Rows** | 3481 |
| **Name of Schema of Destination Table** | **all\_2509** |
| **Name of Destination Table** | **team\_where\_2\_book\_sales\_records\_ib** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID of SQL or Checked Rows** | **Description of Check** | **Result in SOURCE** | **Result in DESTINATION** | **OUTCOME** |
| *DD-P1-1 (Example Row – DELETE)* | *Count of rows* | *5* | *5* | *PASS* |
| *DD-P1-2 (Example Row – DELETE)* | *Count of columns* | *5* | *6* | *FAIL* |
| SQL-1 | **Count of rows** | 3480 | 3480 | *PASS* |
| SQL-2 | **Count of Distinct Rows** | 3480 | 3480 | PASS |
| SQL-3 | **Count of Columns** | 36 | 36 | PASS |
| SQL-4 | **Sum of column Sums** | 11680962.8 | 11680962.8 | PASS |
| SQL-5 | **Sum of Row Sums** | 11680962.8 | 11680962.8 | PASS |
| SQL-6 | **Date format Check**  *Compare 5 randomly chosen values from Source, check date matches in Destination.* | 1583  518  2885  3229  2341 | 1583  518  2885  3229  2341 | PASS |
| SQL-7 | **Eyeball Check**  *Inspect 5 randomly chosen records from Source* | 1350  952  751  3308  1532 | 1350  952  751  3308  1532 | PASS |
| SQL-8 | **Casing check**  *Viewed all columns to check formatting is the same* | Publication column has some in all caps and some not | Publication column has some in all caps and some not |  |
| SQL-9 | **Max, min and average price check**  *Completed for shipping\_price and item\_price* | |  | | --- | | 40 | | 6450 | | 377.73592 | |  | | 0 | | 850 | | 79.0284483 | | |  | | --- | | 40 | | 6450 | | 377.73592 | |  | | 0 | | 850 | | 79.0284483 | | PASS |
| SQL-10 | **Count of nulls** | 0 | 0 | PASS |

**Example for extra checks:** Count of Nulls - {COLUMN NAME} ; Aggregates (SUM/MIN/MAX) checks between Source and Destination

**[[ Your notes about your discoveries or quality checks. ]]**

* **Formula used for finding 5 random ids in our source (excel data sheet): =INDEX(SORTBY(A2:A3481, RANDARRAY(ROWS(A2:A3481))), SEQUENCE(5))**
* **Columns have had name updates:**
* **ID: id**
* **Purchase-date: purchase\_date**
* **Date: date**
* **Time: time**
* **Quarter: quarter**
* **Qtr: qtr**
* **Year: year**
* **Month: month**
* **Date\_1: day\_of\_month**
* **Special Day: special\_day**
* **Online Sale Offers: online\_sale\_offers**
* **Day: day**
* **Weekend: weekend**
* **Morning: morning**
* **Afternoon: afternoon**
* **Evening: evening**
* **Night: night**
* **Gender: gender**
* **Customer ID: customer\_id**
* **Gender F=1665 M=1815: gender2**
* **Product-Name: product\_name**
* **Item-Status: item\_status**
* **Quantity: quanity**
* **Currency: currency**
* **Item Price: item\_price**
* **Shipping-Price: shipping\_price**
* **Ship-City: ship\_city**
* **Ship-State: ship\_state**
* **Ship-Postal-Code: ship\_postal\_code**
* **Category: category**
* **Total\_amount: total\_amount**
* **Author: author**
* **Publication: publication**
* **Profit Percentage (%): profit\_percentage**
* **Profit (INR): profit\_inr**
* **Cost Price: cost\_price**
* **These have been updated on the excel file before creating the table in SQL for easier and quicker use**
* **Casing is different in “publication” column, has been kept the same as per instruction however can be changed in the future if needed**