****  
  
**Quality Control Form**

|  |  |
| --- | --- |
| **Team Name** | SELECT |
| **Team Members** | Katrina, Lucy, Ashan, Serena |
| **Client Name** | Mystic Manuscripts |
| **Date of file received** | 04/11/2024 |
| **Name of file received** | Books Sales Records - V2 |
| **File Format received** | CSV |
| **Size of file (KB) received** | 933 |
| **Encoding of the file** | UTF-8 |
| **Recorded Number of Columns** | 36 |
| **Recorded Number of Rows** | 3480 |
| **Name of Schema of Destination Table** | **mystic\_manuscripts** |
| **Name of Destination Table** | **select\_team** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID of SQL or Checked Rows** | **Description of Check** | **Result in SOURCE** | **Result in DESTINATION** | **OUTCOME** |
| 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 | 11680962.79 | PASS |
| SQL-5 | **Sum of Row Sums** | 11680962.8 | 11680962.79 | PASS |
| SQL-6 | **Date format Check**  *Compare 5 randomly chosen values from Source, check date matches in Destination.* | Checked rows:  10, 300, 2080, 2450, 3400 | Checked rows:  10, 300, 2080, 2450, 3400 | PASS |
| SQL-7 | **Eyeball Check**  *Inspect 5 randomly chosen records from Source* | Checked rows:  31, 3400, 180, 865, 467 | Checked rows:  31, 3400, 180, 865, 467 | PASS |
| SQL-8 | ***Count of nulls in gender\_char*** | 0 | 0 | PASS |
| SQL-9 | ***Count of nulls everywhere*** | 0 | 0 | PASS |
| SQL-10 | ***Sum of cost\_price*** | 1106422.61 | 1106422.61 | PASS |
| SQL-11 | ***Average profit\_percentage*** | 16.6612069 | 16.6612068965517241 | PASS |
| SQL-12 | ***Maximum item\_price*** | 6450 | 6450 | PASS |

**There seems to be lots of duplication in the columns, where information is repeated in different forms e.g ‘quarter’ stores the same information as ‘qtr’ and ‘year’.**

**There is inconsistent capitalisation in columns ‘city’, ‘ship-city’ and ‘ship-state’.**

**There are special characters in the ‘author’ and ‘publications’ columns which may want to be removed or looked at before beginning analysis.**