**Data Quality Assessment**

|  |  |  |
| --- | --- | --- |
| Data Quality dimensions | Quality issues in the table | Action |
| Correct values | Date column  Invalid date 7,9,11,12/2020 | Recorrected data values |
| Records that are duplicates | 9 duplicate values in columns  (code, hotel name, code in Area table) | Removed duplicates values |
| values Formats | Room-nights is in string format | Reformat column  Room-nights from string to int |
| Missing values | Country column missing 10 values | Delete NA rows that represent 0.004 from the whole data |

**Data Edit:**

* Remove word(code) in column code to sort the code using just numbers in both table
* Change the name of the code column to country\_code