

M-STAY: Homestay and Rental Service

Remove Duplicates

- Booking Table : Removed 2 duplicate *booking_id* entries.
- Host Table : Removed 4 duplicate *host_id* entries.

Remove Error Date

- Host Table : Identified invalid *host_since* date (16/05/9999) , set to NULL, and then delete the invalid date.

Remove Inaccurate Value

- Listing Table : Identified negative *listing_price* value, update to NULL and remove them.

Remove Invalid Relationships

- Identified mismatched foreign keys in Review, Host Verification, Property Amenity Tables.
- Set invalid foreign keys to NULL and removed them.

Delete Null Values

- Delete NULL in *review_comment* (Review Table), *amm_description* and *amm_id* (Amenity Table).

Data Cleaning

42.26K

AVERAGE Booking Cost

13M

TOTAL Number of Bookings

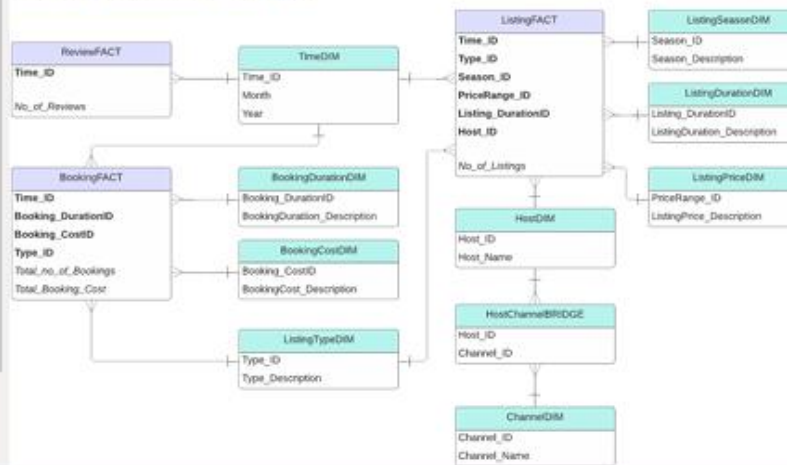
4867

Number of Reviews

4935

Number of Listings

Star Schema Design



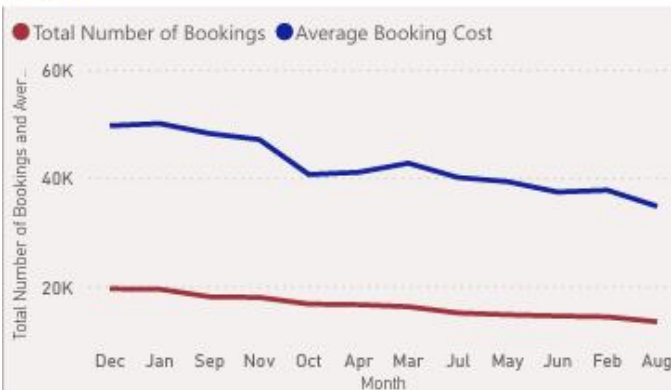
Analysis Result Summary

1. In 2021, **April** had the **highest number of bookings** (140,515), while October had the fewest (7,544).
2. **Medium-term bookings** (30-90 nights) consistently remained the **most popular** choice across all years.
3. The **highest number of listings** was observed during **Summer 2016**.
4. The **number of reviews** showed steady growth from 2010, **peaking in 2017**.

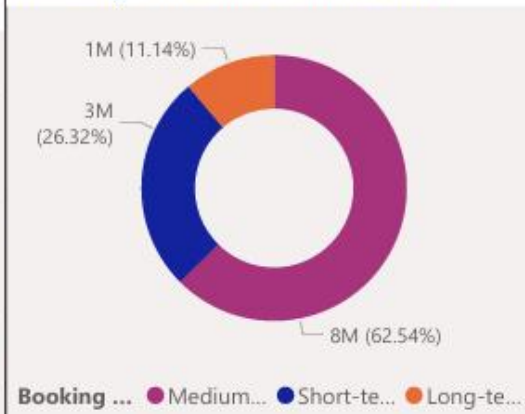
YEAR

All

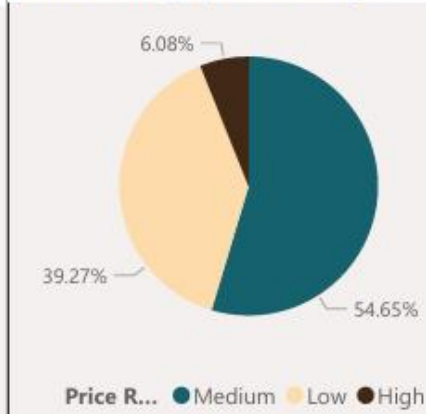
Total Number of Bookings and Average Booking Cost by Month



Booking Duration Distribution



Total Listing by Price Range



Number of Listing by Season

