



Excel Project : Advanced Power Query Transformations

Data Cleaning, Transformation,
and Insights

Bharadwaj Dubbaka

Project Overview



Objective: This project focuses on applying advanced Power Query techniques to clean, transform, and extract insights from real-world datasets. Using two CSV files—`bookings_data.csv` and `rooms_data.csv`—we explore practical data handling challenges.



Significance: The project demonstrates essential skills for data analysts, including cleaning messy data, integrating datasets, and uncovering actionable insights. These are vital for making data-driven decisions in business scenarios.



Goal: By the end of the project, we aim to deliver a clean, structured dataset that can support meaningful analysis and reporting.

Dataset Details

bookings_data.csv:

- Includes date, property details, city codes, room IDs, bookings, and capacity.

rooms_data.csv:

- Links room IDs with room classes.

Step 1 - Data Cleaning

Corrected inconsistent values in property_name (e.g., 'Atliq bay' to 'Atliq Bay').

A screenshot of a data filter dialog for the 'property_name' column. The dialog has a 'Search' field and a list of values with checkboxes. The values are: (Select All), Atliq bay, Atliq Bay, and Atliq Palace. The 'Atliq bay' and 'Atliq Bay' entries are circled in blue, indicating they are the focus of the correction.

A screenshot of the same data filter dialog for 'property_name' after the correction. The 'Atliq bay' entry has been replaced with 'Atliq Bay', and it is now circled in black.

Trimmed whitespace from property_type.

A screenshot of a data table for the 'property_type' column. The table has a single column with the following values: Luxury, Business, Luxury, Luxury, Business, Business, and Business. The first 'Luxury' and the last 'Business' are circled in black, indicating they are the focus of the trimming operation.

A screenshot of the same data table for 'property_type' after trimming whitespace. The values are now: Luxury, Business, Luxury, Luxury, Business, Business, and Business. The first 'Luxury' and the last 'Business' are no longer circled.

Split city|city_code column into City and City Code.

A screenshot of a data table for the 'city|city_code' column. The table has a single column with the following values: Mumbai|+91 22, Delhi|+91 11, Mumbai|+91 22, Mumbai|+91 22, and Delhi|+91 11. The first row is circled in black, indicating it is the focus of the splitting operation.

A screenshot of the same data table after splitting the 'city|city_code' column into two columns: 'property_type' and 'city'. The values are now: Mumbai, Delhi, Mumbai, Mumbai, and Delhi.

Step 2 - Data Transformation

Created new column:
'Availability Status'.
Logic: 'Sold out' if bookings =
capacity, otherwise 'Vacant'.



ABC 123 Availability Status
Sold Out
Sold Out
Sold Out
Vacant
Vacant
Vacant
Vacant

Added column: 'occ%'
(occupancy percentage).
Formula: bookings / capacity
(formatted as percentage).



% occ%
100.00%
100.00%
100.00%
14.81%
35.00%
16.67%
23.68%

Merged bookings_data and
rooms_data on room_id to
add room_class.



% occ%	ABC room_class
100.00%	Elite
100.00%	Standard
100.00%	Standard
14.81%	Premium
35.00%	Premium
16.67%	Presidential
23.68%	Elite

Reordered columns for
better readability.



ABC room_id	ABC room_class
RT2	Elite
RT1	Standard
RT1	Standard
RT3	Premium
RT3	Premium
RT4	Presidential

Step 3 - Extracting Insights

Extracted month names from the date column for trend analysis.



date	Month
01/06/2022	June
01/06/2022	June
01/06/2022	June
01/06/2022	June
01/06/2022	June
01/06/2022	June

Analyzed properties with high occupancy rates.



	successful_bookings	total capacity	sold out	vacant	total occ%
Atliq Bay					
Elite	164	240	150	14	68.33%
Premium	146	216	135	11	67.59%
Presidential	17	36	12	5	47.22%
Standard	46	160	20	26	28.75%
Atliq Bay Total	373	652	317	56	57.21%
Atliq Palace					
Elite	162	304	114	48	53.29%
Premium	77	160	40	37	48.13%
Presidential	66	144	36	30	45.83%
Standard	196	328	123	73	59.76%
Atliq Palace Total	501	936	313	188	53.53%

property_type
Business
Luxury

city
Delhi
Mumbai

Categorized properties as 'Sold Out' or 'Vacant' for availability status.

Project Outcomes

Enhanced knowledge of Power Query for cleaning and transformations.

Gained insights into property bookings and room classifications.

Prepared dataset for advanced analytics and reporting.