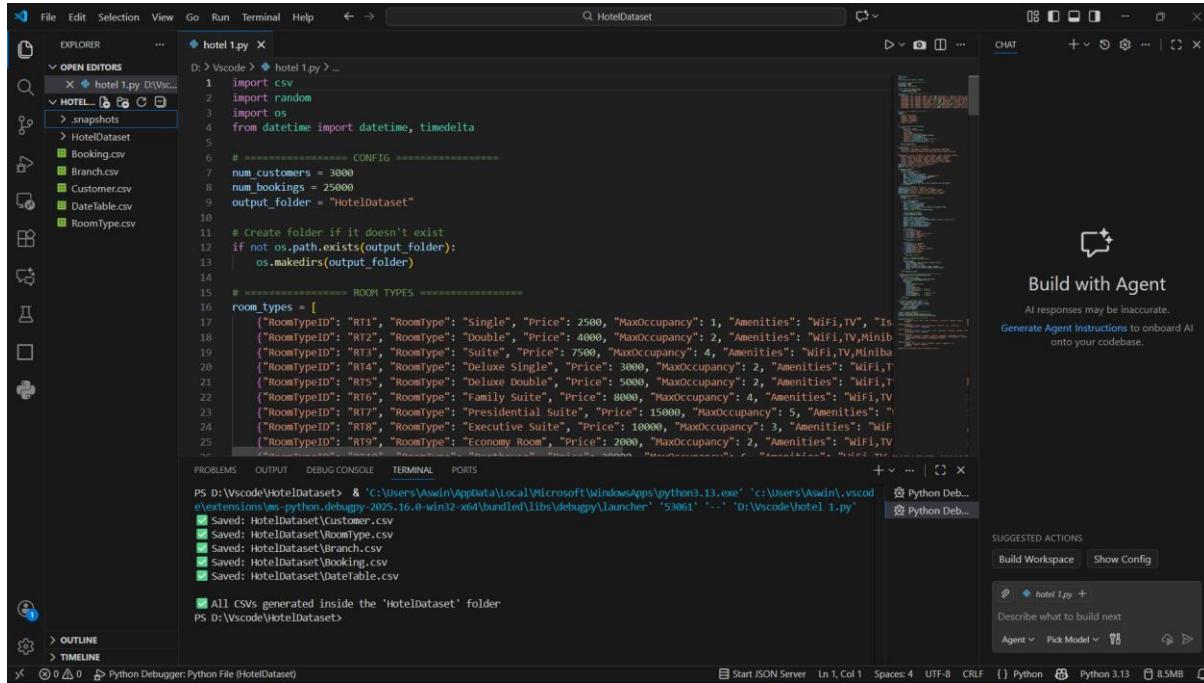


# Module 1: Data Modelling and Ingestion

## 1. Download the Dataset Using Python Code:



```
File Edit Selection View Go Run Terminal Help < > HotelDataset

OPEN EDITORS
HOTEL... > .snapshots
> HotelDataset
> Booking.csv
> Branch.csv
> Customer.csv
> DateTable.csv
> RoomType.csv

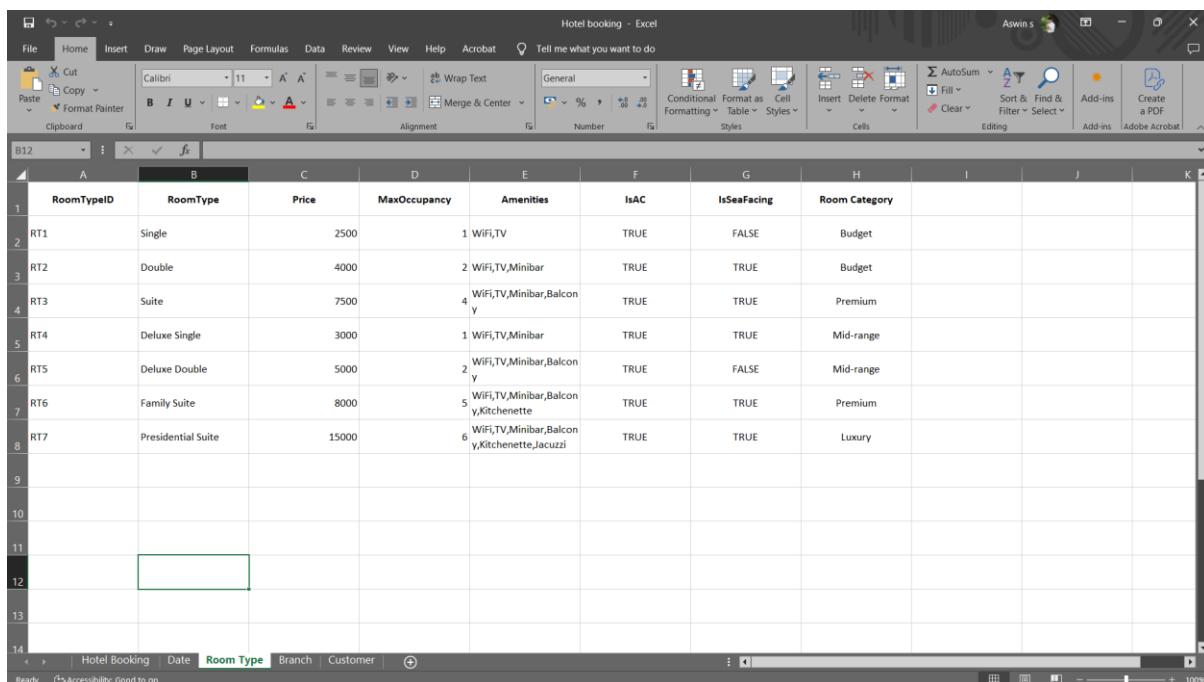
import csv
import random
import os
from datetime import datetime, timedelta

# ===== CONFIG =====
num_customers = 3000
num_bookings = 25000
output_folder = "HotelDataset"

# Create folder if it doesn't exist
if not os.path.exists(output_folder):
    os.makedirs(output_folder)

# ===== ROOM TYPES =====
room_types = [
    {"RoomTypeID": "RT1", "RoomType": "Single", "Price": 2500, "MaxOccupancy": 1, "Amenities": "WiFi,TV", "IsAC": true, "IsSeaFacing": false, "RoomCategory": "Budget"}, {"RoomTypeID": "RT2", "RoomType": "Double", "Price": 4000, "MaxOccupancy": 2, "Amenities": "WiFi,TV,Minibar", "IsAC": true, "IsSeaFacing": false, "RoomCategory": "Budget"}, {"RoomTypeID": "RT3", "RoomType": "Suite", "Price": 7500, "MaxOccupancy": 4, "Amenities": "WiFi,TV,Minibar,Balcony", "IsAC": true, "IsSeaFacing": false, "RoomCategory": "Premium"}, {"RoomTypeID": "RT4", "RoomType": "Deluxe Single", "Price": 3000, "MaxOccupancy": 2, "Amenities": "WiFi,TV,Minibar", "IsAC": true, "IsSeaFacing": false, "RoomCategory": "Mid-range"}, {"RoomTypeID": "RT5", "RoomType": "Deluxe Double", "Price": 5000, "MaxOccupancy": 4, "Amenities": "WiFi,TV,Minibar,Balcony", "IsAC": true, "IsSeaFacing": false, "RoomCategory": "Mid-range"}, {"RoomTypeID": "RT6", "RoomType": "Family Suite", "Price": 8000, "MaxOccupancy": 5, "Amenities": "WiFi,TV,Minibar,Balcony,Kitchenette", "IsAC": true, "IsSeaFacing": false, "RoomCategory": "Premium"}, {"RoomTypeID": "RT7", "RoomType": "Presidential Suite", "Price": 15000, "MaxOccupancy": 5, "Amenities": "WiFi,TV,Minibar,Balcony,Kitchenette,Jacuzzi", "IsAC": true, "IsSeaFacing": false, "RoomCategory": "Luxury"}, {"RoomTypeID": "RT8", "RoomType": "Executive Suite", "Price": 10000, "MaxOccupancy": 3, "Amenities": "WiFi,TV,Minibar,Balcony", "IsAC": true, "IsSeaFacing": false, "RoomCategory": "Premium"}, {"RoomTypeID": "RT9", "RoomType": "Economy Room", "Price": 2000, "MaxOccupancy": 2, "Amenities": "WiFi,TV", "IsAC": true, "IsSeaFacing": false, "RoomCategory": "Budget"}]
```

## 2. Insert All Dataset into Excel:



	RoomTypeID	RoomType	Price	MaxOccupancy	Amenities	IsAC	IsSeaFacing	Room Category
2	RT1	Single	2500	1	WiFi,TV	TRUE	FALSE	Budget
3	RT2	Double	4000	2	WiFi,TV,Minibar	TRUE	TRUE	Budget
4	RT3	Suite	7500	4	WiFi,TV,Minibar,Balcony	TRUE	TRUE	Premium
5	RT4	Deluxe Single	3000	1	WiFi,TV,Minibar	TRUE	TRUE	Mid-range
6	RT5	Deluxe Double	5000	2	WiFi,TV,Minibar,Balcony	TRUE	FALSE	Mid-range
7	RT6	Family Suite	8000	5	WiFi,TV,Minibar,Balcony,Kitchenette	TRUE	TRUE	Premium
8	RT7	Presidential Suite	15000	6	WiFi,TV,Minibar,Balcony,Kitchenette,Jacuzzi	TRUE	TRUE	Luxury

### 3.Creating Duration Calculation using Formula

Duration = MAX(0, F2 - E2)

	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	RoomTypeID	BranchID	CheckInDate	CheckOutDate	Duration	Revenue	BookingStatus	CancellationReason	PaymentMethod	DiscountApplied	BookingChannel	Purpose	Stay Type	DateID
2	RT5	B3	03-07-2025	08-07-2025	5	25000	No-show	Corporate Accou	0 Call Center	Business	Short Stay	20200101		
3	RT5	B6	07-05-2025	08-05-2025	1	4250	Cancelled	Weather	Cash	15 Mobile App	Conference	Medium Stay	20200102	
4	RT2	B6	17-01-2024	22-01-2024	5	18000	No-show	Credit Card	10 Call Center	Holiday	Long Stay	20200103		
5	RT4	B2	18-08-2023	20-08-2023	2	6000	No-show	Credit Card	0 Website	Holiday	Long Stay	20200104		
6	RT2	B6	04-10-2023	08-10-2023	4	16000	Cancelled	Change of plans	Cash	0 Call Center	Other	Other	20200105	
7	RT7	B7	11-09-2024	14-09-2024	3	38250	No-show	Cash	15 Travel Agent	Conference	Medium Stay	20200106		
8	RT5	B3	30-06-2025	04-07-2025	4	18000	Checked-in	UPI	10 Website	Holiday	Long Stay	20200107		
9	RT3	B2	07-07-2023	08-07-2023	1	6000	Checked-in	Corporate Accou	20 Mobile App	Holiday	Long Stay	20200108		
10	RT6	B3	12-05-2025	15-05-2025	3	19200	No-show	Credit Card	20 Mobile App	Vacation	Long Stay	20200109		
11	RT2	B4	31-08-2023	04-09-2023	4	16000	Cancelled	Change of plans	Corporate Accou	0 Call Center	Other	Other	20200110	
12	RT7	B7	06-09-2024	11-09-2024	5	60000	No-show	UPI	20 Mobile App	Vacation	Long Stay	20200111		
13	RT2	B1	02-01-2024	09-01-2024	7	22400	Cancelled	Other	Credit Card	20 Call Center	Vacation	Long Stay	20200112	
14	RT6	B5	05-05-2025	10-05-2025	5	38000	Checked-in	Credit Card	5 Call Center	Business	Short Stay	20200113		
15	RT6	B1	16-11-2023	21-11-2023	5	34000	Checked-in	Cash	15 Travel Agent	Conference	Medium Stay	20200114		
16	RT2	B4	07-04-2025	14-04-2025	7	28000	Cancelled	Other	Cash	0 Travel Agent	Other	Other	20200115	
17	RT1	B4	08-08-2025	15-08-2025	7	15750	Checked-in	Cash	10 Website	Business	Short Stay	20200116		
18	RT2	B3	28-05-2023	30-05-2023	2	8000	No-show	UPI	0 Mobile App	Conference	Medium Stay	20200117		
19	RT3	B5	08-06-2024	09-06-2024	1	6750	Cancelled	Price	Cash	10 Mobile App	Business	Short Stay	20200118	
20	RT6	B1	26-05-2024	27-05-2024	1	8000	No-show	Credit Card	0 Mobile App	Business	Short Stay	20200119		
21	RT4	B7	10-07-2024	15-07-2024	5	12000	No-show	Corporate Accou	20 Mobile App	Other	Other	20200120		
22	RT7	B5	10-10-2023	15-10-2023	5	67500	Cancelled	Price	Credit Card	10 Travel Agent	Other	Other	20200121	
23	RT5	B6	19-12-2024	22-12-2024	3	14250	No-show	Corporate Accou	5 Travel Agent	Holiday	Long Stay	20200122		
24	RT5	B6	09-09-2025	12-09-2025	3	12750	Checked-in	Cash	15 Website	Conference	Medium Stay	20200123		
25	RT6	B5	27-11-2023	30-11-2023	3	20400	No-show	Corporate Accou	15 Call Center	Vacation	Long Stay	20200124		
26	RT3	B3	14-09-2023	18-09-2023	4	28500	No-show	Credit Card	5 Website	Conference	Medium Stay	20200125		
27	RT1	B4	30-09-2023	04-10-2023	4	9500	Checked-in	Cash	5 Travel Agent	Conference	Medium Stay	20200126		
28	RT6	B3	29-06-2025	01-07-2025	2	12800	No-show	Cash	20 Mobile App	Other	Other	20200127		
29	RT1	B1	29-09-2023	04-10-2023	5	10000	Checked-in	Credit Card	20 Website	Other	Other	20200128		
30	RT1	B4	01-01-2025	02-01-2025	1	2250	Cancelled	Other	Corporate Accou	10 Mobile App	Other	Other	20200129	
31	RT3	B6	08-03-2025	15-03-2025	7	44625	Cancelled	Other	Corporate Accou	15 Mobile App	Business	Short Stay	20200130	

### 4.Creating Revenue Calculation Using Formula:

Revenue =VLOOKUP(C2, '[RoomType]Room Type'!A:C, 3, FALSE) \* G2 \* (1 - (L2 / 100))

	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	RoomTypeID	BranchID	CheckInDate	CheckOutDate	Duration	Revenue	BookingStatus	CancellationReason	PaymentMethod	DiscountApplied	BookingChannel	Purpose	Stay Type	DateID
2	RT5	B3	03-07-2025	08-07-2025	5	25000	No-show	Corporate Accou	0 Call Center	Business	Short Stay	20200101		
3	RT5	B6	07-05-2025	08-05-2025	1	4250	Cancelled	Weather	Cash	15 Mobile App	Conference	Medium Stay	20200102	
4	RT2	B6	17-01-2024	22-01-2024	5	18000	No-show	Credit Card	10 Call Center	Holiday	Long Stay	20200103		
5	RT4	B2	18-08-2023	20-08-2023	2	6000	No-show	Credit Card	0 Website	Holiday	Long Stay	20200104		
6	RT2	B6	04-10-2023	08-10-2023	4	16000	Cancelled	Change of plans	Cash	0 Call Center	Other	Other	20200105	
7	RT7	B7	11-09-2024	14-09-2024	3	38250	No-show	Cash	15 Travel Agent	Conference	Medium Stay	20200106		
8	RT5	B3	30-06-2025	04-07-2025	4	18000	Checked-in	UPI	10 Website	Holiday	Long Stay	20200107		
9	RT3	B2	07-07-2023	08-07-2023	1	6000	Checked-in	Corporate Accou	20 Mobile App	Vacation	Long Stay	20200108		
10	RT6	B3	12-05-2025	15-05-2025	3	19200	No-show	Credit Card	20 Mobile App	Other	Other	20200109		
11	RT2	B4	31-08-2023	04-09-2023	4	16000	Cancelled	Change of plans	Corporate Accou	0 Call Center	Other	Other	20200110	
12	RT7	B7	06-09-2024	11-09-2024	5	60000	No-show	UPI	20 Mobile App	Vacation	Long Stay	20200111		
13	RT2	B1	02-01-2024	09-01-2024	7	22400	Cancelled	Other	Credit Card	20 Call Center	Vacation	Long Stay	20200112	
14	RT6	B5	05-05-2025	10-05-2025	5	38000	Checked-in	Credit Card	5 Call Center	Business	Short Stay	20200113		
15	RT6	B1	16-11-2023	21-11-2023	5	34000	Checked-in	Cash	15 Travel Agent	Conference	Medium Stay	20200114		
16	RT2	B4	07-04-2025	14-04-2025	7	28000	Cancelled	Other	Cash	0 Travel Agent	Other	Other	20200115	
17	RT1	B4	08-08-2025	15-08-2025	7	15750	Checked-in	Cash	10 Website	Business	Short Stay	20200116		
18	RT2	B3	28-05-2023	30-05-2023	2	8000	No-show	UPI	0 Mobile App	Conference	Medium Stay	20200117		
19	RT3	B5	08-06-2024	09-06-2024	1	6750	Cancelled	Price	Cash	10 Mobile App	Business	Short Stay	20200118	
20	RT6	B1	26-05-2024	27-05-2024	1	8000	No-show	Credit Card	0 Mobile App	Business	Short Stay	20200119		
21	RT4	B7	10-07-2024	15-07-2024	5	12000	No-show	Corporate Accou	20 Mobile App	Other	Other	20200120		
22	RT7	B5	10-10-2023	15-10-2023	5	67500	Cancelled	Price	Credit Card	10 Travel Agent	Other	Other	20200121	
23	RT5	B6	19-12-2024	22-12-2024	3	14250	No-show	Corporate Accou	5 Travel Agent	Holiday	Long Stay	20200122		
24	RT5	B6	09-09-2025	12-09-2025	3	12750	Checked-in	Cash	15 Website	Conference	Medium Stay	20200123		
25	RT6	B5	27-11-2023	30-11-2023	3	20400	No-show	Corporate Accou	15 Call Center	Vacation	Long Stay	20200124		
26	RT3	B3	14-09-2023	18-09-2023	4	28500	No-show	Credit Card	5 Website	Conference	Medium Stay	20200125		
27	RT1	B4	30-09-2023	04-10-2023	4	9500	Checked-in	Cash	5 Travel Agent	Conference	Medium Stay	20200126		
28	RT6	B3	29-06-2025	01-07-2025	2	12800	No-show	Cash	20 Mobile App	Other	Other	20200127		
29	RT1	B1	29-09-2023	04-10-2023	5	10000	Checked-in	Credit Card	20 Website	Other	Other	20200128		
30	RT1	B4	01-01-2025	02-01-2025	1	2250	Cancelled	Other	Corporate Accou	10 Mobile App	Other	Other	20200129	

## 5.Creating StayType Calculation Using Formula:

StayType =SWITCH(TRIM(UPPER(N2)),

"BUSINESS", "Short Stay",  
 "CONFERENCE", "Medium Stay",  
 "VACATION", "Long Stay",  
 "HOLIDAY", "Long Stay",  
 "FAMILY", "Leisure",  
 "LEISURE", "Leisure",  
 "GROUP", "Group Stay",  
 "SOLO", "Solo Stay",  
 "UNKNOWN", "Unknown",  
 "Other"

)

Hotel booking - Excel

The screenshot shows an Excel spreadsheet titled "Hotel booking - Excel". The formula in cell O2 is:

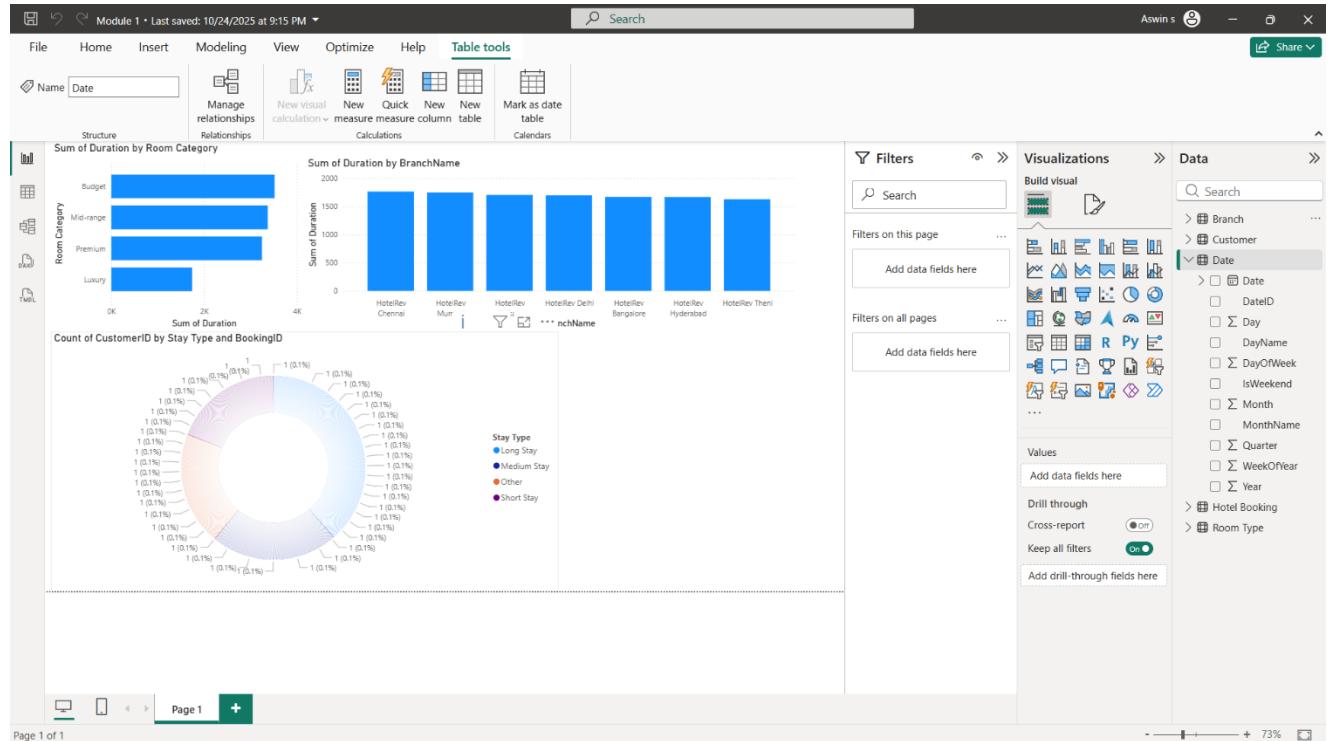
```
=SWITCH(TRIM(UPPER(N2)),
"BUSINESS", "Short Stay",
"CONFERENCE", "Medium Stay",
"VACATION", "Long Stay",
"HOLIDAY", "Long Stay",
"FAMILY", "Leisure",
"LEISURE", "Leisure",
"GROUP", "Group Stay",
"SOLO", "Solo Stay",
"UNKNOWN", "Unknown",
"Other"
)
```

Below the formula, there is a table with the following data:

	C	D	E	F	G	H	I	J	K	L	M	N	O	P
	RoomTypeID	BranchID	CheckInDate	CheckOutDate	Duration	Revenue	BookingStatus	CancellationReason	PaymentMethod	DiscountApplied	BookingChannel	Purpose	Stay Type	DateID
2	RT5	B3	03-07-2025	08-07-2025	5	25000	No-show	Corporate Accou	0 Call Center	Business	Short Stay		20200101	
3	RT5	B6	07-05-2025	08-05-2025	1	4250	Cancelled	Weather	Cash	15 Mobile App	Conference	Medium Stay		20200102
4	RT2	B6	17-01-2024	22-01-2024	5	18000	No-show		Credit Card	10 Call Center	Holiday	Long Stay		20200103
5	RT4	B2	18-08-2023	20-08-2023	2	6000	No-show		Credit Card	0 Website	Holiday	Long Stay		20200104
6	RT2	B6	04-10-2023	08-10-2023	4	16000	Cancelled	Change of plans	Cash	0 Call Center	Other	Other		20200105
7	RT7	B7	11-09-2024	14-09-2024	3	38250	No-show		Cash	15 Travel Agent	Conference	Medium Stay		20200106
8	RT5	B3	30-06-2025	04-07-2025	4	18000	Checked-in		UPI	10 Website	Holiday	Long Stay		20200107
9	RT3	B2	07-07-2023	08-07-2023	1	6000	Checked-in	Corporate Accou	20 Mobile App	Holiday	Long Stay			20200108
10	RT6	B3	12-05-2025	15-05-2025	3	19200	No-show		Credit Card	20 Mobile App	Vacation	Long Stay		20200109
11	RT2	B4	31-08-2023	04-09-2023	4	16000	Cancelled	Change of plans	Corporate Accou	0 Call Center	Other	Other		20200110
12	RT7	B7	06-09-2024	11-09-2024	5	60000	No-show		UPI	20 Mobile App	Vacation	Long Stay		20200111
13	RT2	B1	02-01-2024	09-01-2024	7	22400	Cancelled	Other	Credit Card	20 Call Center	Vacation	Long Stay		20200112
14	RT6	B5	05-05-2025	10-05-2025	5	38000	Checked-in		Credit Card	5 Call Center	Business	Short Stay		20200113
15	RT6	B1	16-11-2023	21-11-2023	5	34000	Checked-in		Cash	15 Travel Agent	Conference	Medium Stay		20200114
16	RT2	B4	07-04-2025	14-04-2025	7	28000	Cancelled	Other	Cash	0 Travel Agent	Other	Other		20200115
17	RT1	B4	08-08-2025	15-08-2025	7	15750	Checked-in		Cash	10 Website	Business	Short Stay		20200116
18	RT2	B3	28-05-2023	30-05-2023	2	8000	No-show		UPI	0 Mobile App	Conference	Medium Stay		20200117
19	RT3	B5	08-06-2024	09-06-2024	1	6750	Cancelled	Price	Cash	10 Mobile App	Business	Short Stay		20200118

## 6. Import the Excel into PowerBI and Create the Visualization:

- Duration by Room Category
- Duration by Branch Name
- Count of Customer ID by Stay Type and Booking ID



**Module 1 documentation has been successfully completed.**