

**Name:** Shruti More

### **Module 3: Customer Segmentation and Behavioural Analysis**

Module 3 focuses on analysing customer behaviour by creating meaningful customer segments using data-driven techniques. This module involves designing analytical attributes such as Guest Type, Customer Clusters, RFM Scores, and Demographic Segmentation to understand how different groups of customers behave.

Using Excel, Power BI, and DAX, this module enhances customer-level intelligence and helps hotels gain insights into guest loyalty, spending patterns, booking preferences, and travel purpose. By applying segmentation techniques, hotels can identify high-value guests, understand customer diversity, and design targeted strategies to improve guest satisfaction and revenue.

#### **1. Creating GuestType (Customer Persona Classification)**

To understand customer behaviour more effectively, a new analytical attribute called GuestType was created in Power BI using a DAX calculated column. This attribute classifies each guest into meaningful personas based on their travel purpose, booking channel, and stay duration.

##### **DAX Formula Used:**

```
GuestType =  
SWITCH(  
    TRUE(),  
  
        -- Business: Purpose = Business OR BookingChannel = Corporate  
        [Purpose] = "Business" || [BookingChannel] = "Corporate",  
        "Business",  
  
        -- Family: Vacation or Holiday with Long Stay  
        ([Purpose] = "Vacation" && [StayType] = "Long Stay") ||  
        ([Purpose] = "Holiday" && [StayType] = "Long Stay"),  
        "Family",  
  
        -- Corporate: Conference OR Medium Stay  
        [Purpose] = "Conference" || [StayType] = "Medium Stay",  
        "Corporate",  
  
        -- Default  
        "Solo"
```

)

The screenshot shows the Power BI Data Editor interface. On the left, the 'Structure' pane displays a DAX formula for 'GuestType':

```

1 GuestType =
2 SWITCH(
3   TRUE(),
4
5   -- Business: Purpose = Business OR BookingChannel = Corporate
6   [Purpose] = "Business" || [BookingChannel] = "Corporate",
7   "Business",
8
9   -- Family: Vacation or Holiday with long Stay
10  ([Purpose] = "Vacation" & [StayType] = "Long Stay") ||
11  ([Purpose] = "Holiday" & [StayType] = "Long Stay"),
12  "Family",
13
14  -- Corporate: Conference OR Medium Stay (adjust if needed)
15  ([Purpose] = "Conference" || [StayType] = "Medium Stay",
16  "Corporate",
17
18  -- Default
19  "Solo"
20 )
21
  
```

The 'Data' pane on the right lists various columns from the 'Booking' table, including BookingChannel, BookingID, BranchID, CancellationReason, CheckInDate, CheckOutDate, CustomerCluster, CustomerID, Duration, GuestType, LeadTime, MonthStart, MonthStartHierarchy, NumOfBookings, PaymentMethod, PromotionCodeUsed, Purpose, Revenue, RoomTypeID, and SourceType.

The main table view shows a sample of 25,000 rows from the 'Booking' table, with columns like DiscountApplied, BookingChannel, Purpose, StayType, PromotionCodeUsed, TaxAmount, SourceType, ADR, MonthStart, GuestType, NumOfBookings, and CustomerCluster.

This DAX formula categorizes guests into Business, Family, Corporate, or Solo based on their purpose of visit, stay duration, and booking channel. It helps classify guests for better analysis of customer behavior and hotel trends.

## 2. CustomerCluster

The CustomerCluster attribute categorizes customers based on their booking frequency and total revenue generated. This segmentation helps identify guest loyalty levels and spending patterns, allowing the hotel to focus on high-value customers and design personalized strategies for each group.

### DAX Formula:

CustomerCluster =

SWITCH(TRUE(),

[NumOfBookings] = 1 && [Revenue] < 5000, "First-Timer",

[NumOfBookings] >= 3 && [Revenue] < 15000, "Loyal Guest",

[Revenue] >= 15000, "High Spender",

"Other"

)

The screenshot shows the Power BI Column tools interface. In the top navigation bar, 'Shruti More\_Visualizations' is selected, and the status bar indicates 'Last saved: Today at 3:22 PM'. The 'Column tools' tab is active. On the left, the 'Structure' pane shows a formula:

```

1 CustomerCluster =
2 SNITC(TRUE()),
3 [NumOfBookings] = 1 && [Revenue] <
4 5000, "First-Timer",
5 [NumOfBookings] >= 3 && [Revenue] <
6 15000, "Loyal Guest",
7 [Revenue] >= 15000, "High Spender",
8 "Other"
9
10

```

The main area displays a table titled 'Booking' with 25,000 rows. The columns include: BookingChannel, Purpose, StayType, PromotionCodeUsed, TaxAmount, SourceType, ADR, MonthStart, GuestType, NumOfBookings, and CustomerCluster. The 'CustomerCluster' column contains four distinct values: High Spender, High Spender, High Spender, and High Spender. The right side of the interface shows the 'Data' pane with various columns listed.

This formula classifies guests into First-Timer, Loyal Guest, High Spender, or Other based on their booking count and total revenue. It helps the hotel identify guest value segments and understand their long-term contribution.

### 3. Creating Guest Country

The Guest Country field was created to simulate the geographical origin of guests by randomly assigning a country name to each record. This enhances the dataset for demographic analysis and enables country-wise segmentation, heatmaps, and trend insights in Power BI.

#### Excel Formula Used

```
=CHOOSE(RANDBETWEEN(1,46),
"India","United States","United Kingdom","Canada","Australia",
"Germany","France","Italy","Spain","Brazil","Mexico","China",
"Japan","South Korea","Russia","South Africa","Nigeria","Egypt",
"Kenya","Argentina","Chile","Colombia","Peru","Netherlands",
"Sweden","Norway","Denmark","Finland","Poland","Turkey",
"Saudi Arabia","United Arab Emirates","Israel","Thailand",
"Vietnam","Indonesia","Malaysia","Singapore","Philippines",
>New Zealand","Pakistan","Bangladesh","Sri Lanka","Nepal","Iran")
```

Shruti More\_Visualizations • Last saved: Today at 3:22 PM

File Home Help Table tools Column tools

Search

Name: Nationality Data type: Text

Format: Text Summarization: Don't summarize

Structure: Data category: Uncategorized

Formatting: Sort by column: Sort

Properties: Data groups: Groups

Sort: Manage relationships: Relationships

New column: Calculations

**Data**

CustomerID Name Email Phone City State LoyaltyTier AgeGroup Gender Nationality

C666 Rohan Sethi rohan.sethi@example.com 9520214537 Chennai Delhi Platinum 51+ Male Romania

C1184 Simran Verma simran.verma@example.com 9413830145 Chennai Delhi Platinum 18-25 Male Chile

C1312 Aditya Kapoor aditya.kapoor@example.com 9931244618 Chennai Delhi Platinum 51+ Male Australia

C1528 Rohan Bhatia rohan.bhatia@example.com 9409101211 Chennai Delhi Platinum 36-50 Male Malaysia

C1944 Ishan Kapoor ishan.kapoor@example.com 9273785688 Chennai Delhi Platinum 26-35 Male Thailand

C2096 Aarav Kapoor aarav.kapoor@example.com 9335834188 Chennai Delhi Platinum 36-50 Male Canada

C2538 Ishan Sharma ishan.sharma@example.com 9896212628 Chennai Delhi Platinum 18-25 Male Austria

C3071 Aditya Kapoor aditya.kapoor@example.com 9109576886 Chennai Delhi Platinum 36-50 Male Qatar

C3424 Aditya Sharma aditya.sharma@example.com 9146636222 Chennai Delhi Platinum 51+ Male China

C3513 Neha Verma neha.verma@example.com 9622890013 Chennai Delhi Platinum 36-50 Male Italy

C3524 Kavya Kapoor kavya.kapoor@example.com 9595256537 Chennai Delhi Platinum 51+ Male Mexico

C4443 Ishan Bhatia ishan.bhatia@example.com 9997171695 Chennai Delhi Platinum 36-50 Male Egypt

C4508 Tanya Kapoor tanya.kapoor@example.com 9192118650 Chennai Delhi Platinum 51+ Male Saudi Arabia

C4660 Rajat Verma rajat.verma@example.com 9509580064 Chennai Delhi Platinum 36-50 Male Thailand

C5067 Neha Kapoor neha.kapoor@example.com 9193749947 Chennai Delhi Platinum 36-50 Male Nepal

C5150 Rajat Bhatia rajat.bhatia@example.com 9715393478 Chennai Delhi Platinum 26-35 Male Oman

C5298 Simran Verma simran.verma@example.com 9188075044 Chennai Delhi Platinum 26-35 Male Sri Lanka

C5359 Aarav Sharma aarav.sharma@example.com 9569458176 Chennai Delhi Platinum 51+ Male Saudi Arabia

C5901 Rajat Sharma rajat.sharma@example.com 9915807582 Chennai Delhi Platinum 36-50 Male Iraq

C6092 Neha Sharma neha.sharma@example.com 9462523218 Chennai Delhi Platinum 18-25 Male Iran

C6310 Aarav Bhatia aarav.bhatia@example.com 9117201084 Chennai Delhi Platinum 36-50 Male Switzerland

C6465 Kavya Sharma kavya.sharma@example.com 9244533229 Chennai Delhi Platinum 36-50 Male United Arab Emirates

C6928 Aditya Sethi aditya.sethi@example.com 9375107746 Chennai Delhi Platinum 26-35 Male Australia

C7136 Neha Verma neha.verma@example.com 9800204239 Chennai Delhi Platinum 26-35 Male Nigeria

C7166 Tanya Kapoor tanya.kapoor@example.com 9425549935 Chennai Delhi Platinum 36-50 Male South Africa

C7335 Ishan Sharma ishan.sharma@example.com 9137734384 Chennai Delhi Platinum 26-35 Male Peru

C7520 Aarav Sethi aarav.sethi@example.com 9844892461 Chennai Delhi Platinum 18-25 Male Saudi Arabia

Table: Customer (15,000 rows) Column: Nationality (59 distinct values)

Update available (click to download)

This formula randomly assigns a country from a predefined list using RANDBETWEEN and CHOOSE. It helps create realistic guest demographics for country-wise analysis and visualizations.

#### 4. Guest Summary Table:

The Guest Summary Table aggregates guest-level stay information by summarizing each customer's earliest check-in date and most recent check-out date. This table is useful for understanding guest history, identifying long-term customers, and analysing patterns in first and last visits.

#### DAX Formula Used

```
Tbl_Guests =  
SUMMARIZE(  
    'TblBookings',  
    'TblBookings'[CustomerID],  
    "FirstStay", MIN('TblBookings'[CheckInDate]),  
    "LastStay", MAX('TblBookings'[CheckOutDate])  
)
```

The screenshot shows the Power BI Visualizations interface. At the top, there's a navigation bar with File, Home, Help, and Table tools. The Table tools tab is selected. Below the navigation bar, there's a search bar and a share button. The main area displays a table titled 'Tbl\_Guests'. The table has six columns: CustomerID, FirstStay, LastStay, RecencyScore, FrequencyScore, and SpendScore. The data shows various guest IDs with their first and last stay dates, and corresponding scores. To the right of the table is a 'Data' pane containing a tree view of data models and tables, with 'Tbl\_Guests' selected. The bottom right corner of the interface has a note: 'Update available (click to download)'.

This DAX formula creates a summary table showing each customer's first and last stay dates using MIN and MAX. It helps track guest visit history and analyse long-term customer relationships.

## 5. Scoring (RFM Analysis):

To evaluate customer value and behaviour, an RFM scoring model was created using three key metrics:

- Recency – how recently a customer stayed
- Frequency – how often they booked
- Monetary (Spend) – total revenue generated

These scores help identify loyal customers, high-value spenders, and inactive guests, improving segmentation and targeted decision-making.

### a) Recency Score

#### DAX Formula

RecencyScore =

DATEDIFF(

CALCULATE(

MAX('TblBooking'[CheckOutDate]),

FILTER('TblBooking', 'TblBooking'[CustomerID] = 'Tbl\_Guests'[CustomerID])

),

TODAY(),

DAY

)

Screenshot of Power BI Column Tools pane showing the DAX formula for RecencyScore:

```

1 RecencyScore =
2 DATEDIFF(
3   CALCULATE(
4     MAX('TblBooking'[CheckOutDate]),
5     FILTER('TblBooking', 'TblBooking'[CustomerID] = 'Tbl_Guests'[CustomerID])
6   ),
7   TODAY(),
8   DAY
9 )
10

```

The formula calculates the number of days since a customer's most recent stay. Lower values indicate recently active guests, while higher values show less recent or inactive customers.

Table: Tbl\_Guests (12,201 rows) Column: RecencyScore (1,800 distinct values)

Update available (click to download)

This formula calculates how many days have passed since a customer's most recent stay. Lower values indicate recently active guests, while higher values show less recent or inactive customers.

## b) Frequency Score

### DAX Formula

FrequencyScore =

CALCULATE(

COUNTROWS('TblBooking'),

FILTER('TblBooking', 'TblBooking'[CustomerID] = 'Tbl\_Guests'[CustomerID])

)

Screenshot of Power BI Column Tools pane showing the DAX formula for FrequencyScore:

```

1 FrequencyScore =
2 CALCULATE(
3   COUNTROWS('TblBooking'),
4   FILTER('TblBooking', 'TblBooking'[CustomerID] = 'Tbl_Guests'[CustomerID])
5 )
6

```

The formula counts the number of bookings for each customer. Lower values indicate more frequent customers, while higher values show less frequent or inactive customers.

Table: Tbl\_Guests (12,201 rows) Column: FrequencyScore (36 distinct values)

Update available (click to download)

This formula counts how many bookings each customer has made. It helps identify highly frequent guests versus one-time visitors.

### c) Spend Score

#### DAX Formula

SpendScore =

CALCULATE(

SUM('TblBooking'[DayRevenue]),

FILTER('TblBooking', 'TblBooking'[CustomerID] = 'Tbl\_Guests'[CustomerID])

)

The screenshot shows the Power BI Column Tools pane. At the top, there are tabs for Name, Data type, Structure, Format, Summarization, Sort by column, Data groups, Manage relationships, and New column. The Name tab is selected, showing the formula:

```

1 SpendScore =
2 CALCULATE(
3     SUM('TblBooking'[DayRevenue]),
4     FILTER('TblBooking', 'TblBooking'[CustomerID] = 'Tbl_Guests'[CustomerID])
5 )
6
  
```

Below the formula, there is a table view showing data for 12,201 rows. The columns are CustomerID, FirstStay, LastStay, RecencyScore, FrequencyScore, and SpendScore. The SpendScore column contains values like 947352, 586456, etc. On the right side of the table, there is a Data pane listing various tables and their fields, with 'SpendScore' highlighted under 'Tbl\_Guests'.

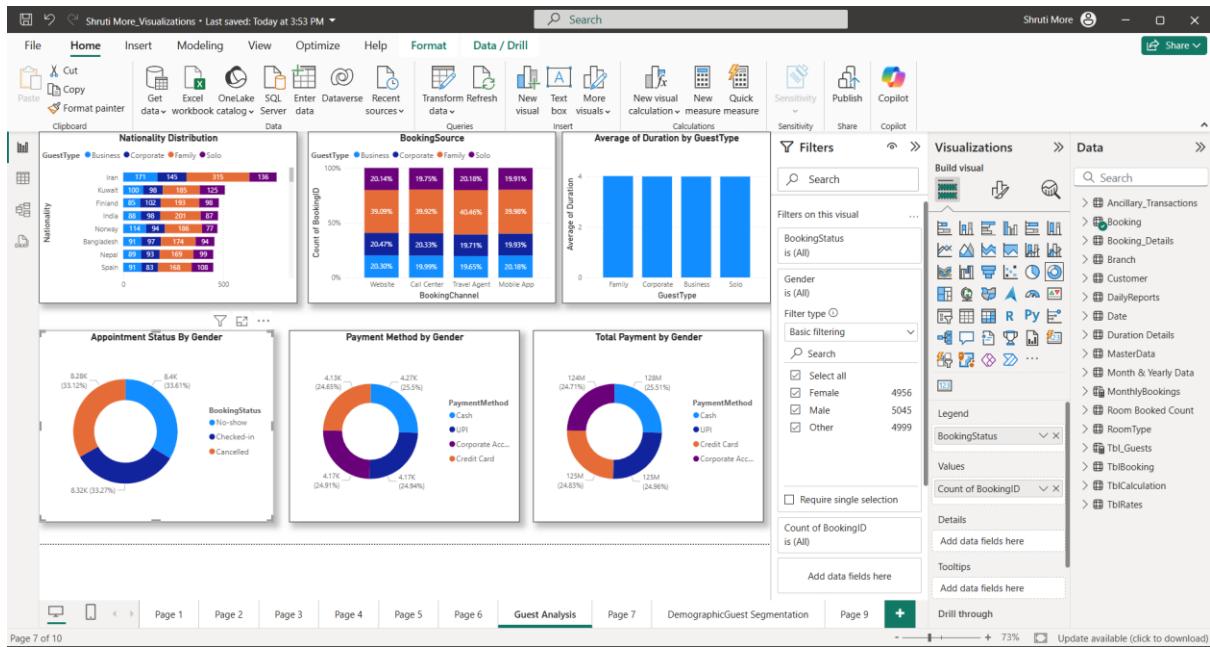
This formula calculates the total revenue generated by each customer across all stays. It highlights high spenders and helps analyse revenue contribution per guest.

### Visualization and Analysis: (Output)

## 6. Visualization & Analysis (Output)

After creating analytical fields like GuestType, CustomerCluster, RFM Scores, and demographic attributes, several Power BI visualizations were created to understand customer behavior and segmentation.

These visuals help interpret patterns in nationality, booking behavior, spending habits, and customer demographics.



### a) Nationality Distribution by Guest Type

A clustered bar chart was created to show how different guest personas (Business, Corporate, Family, Solo) are distributed across countries.

This helps identify key international markets contributing to hotel bookings.

### b) Booking Source Analysis

A 100% stacked bar chart displays the proportion of bookings made via Website, Call Center, Travel Agent, and Mobile App.

This visualization highlights which channels are most preferred by different guest types.

### c) Average Duration by Guest Type

A bar chart shows how long each customer group stays on average.

Family guests typically have longer stays, while Business and Solo travelers have shorter visits.

### d) Appointment Status by Gender

A donut chart illustrates the distribution of Checked-in, No-show, and Cancelled bookings for Male and Female customers.

This helps understand reliability trends based on gender.

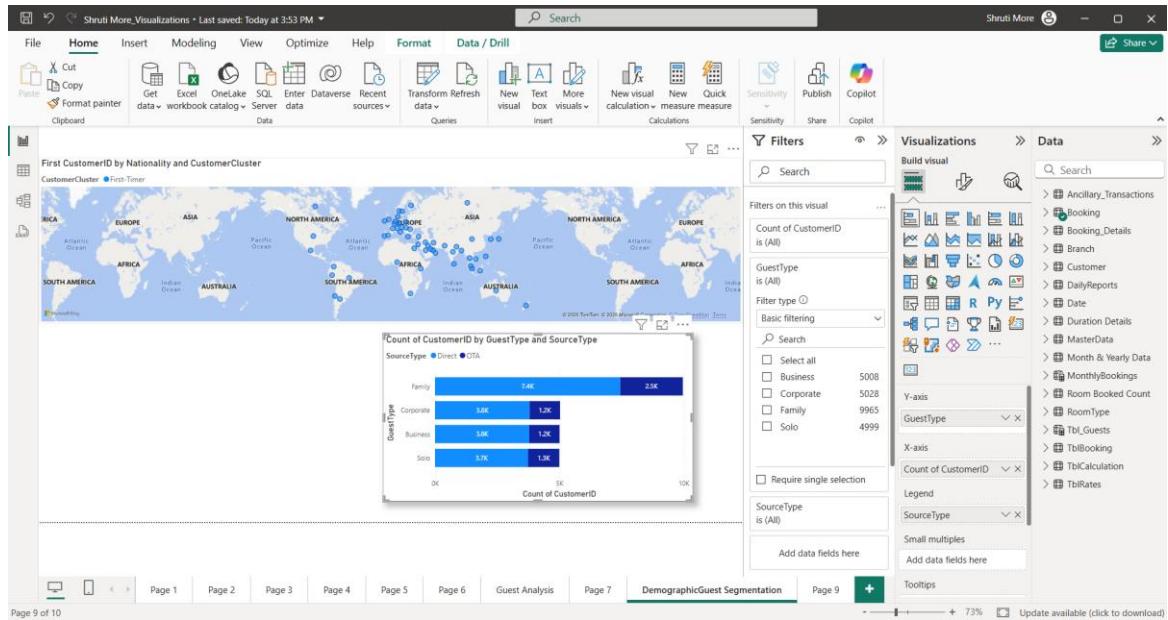
### e) Payment Method by Gender

Another donut chart shows which payment method (Cash, UPI, Credit Card, Corporate Account) is used more frequently by each gender.

This supports decisions related to billing methods and promotions.

## f) Total Payment by Gender

This chart shows total revenue contribution from each gender by payment mode. It helps identify spending behavior and possible high-value segments.

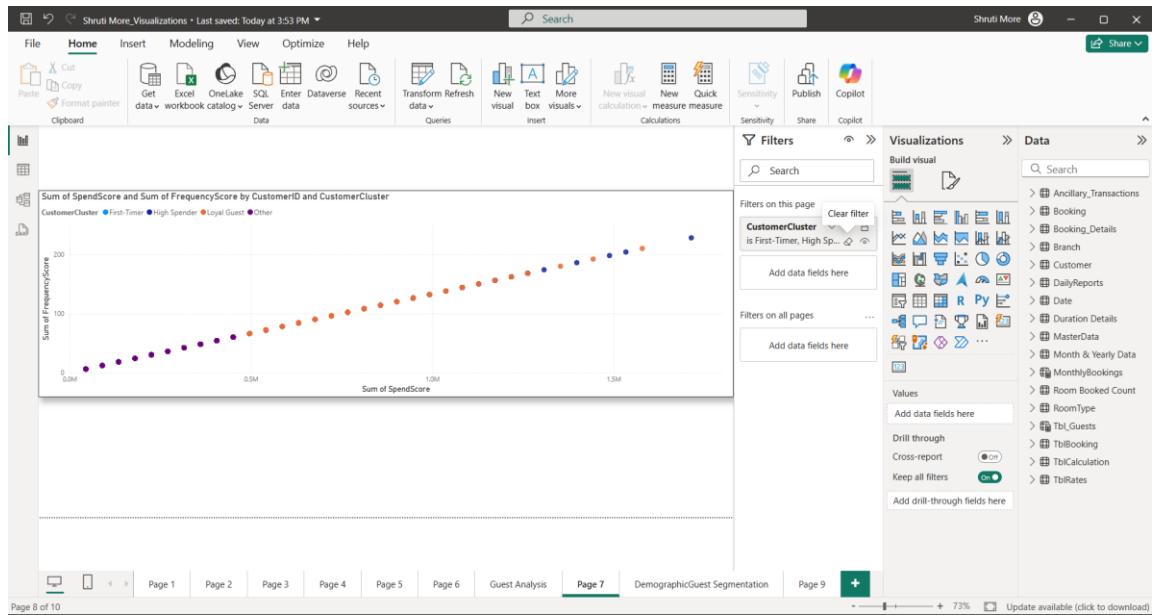


## g) Geographic Distribution of First-Time Customers

A map visualization highlights the locations of first-time guests around the world. This helps identify potential markets for targeted hotel promotions.

## h) Customer Count by Guest Type and Source Type

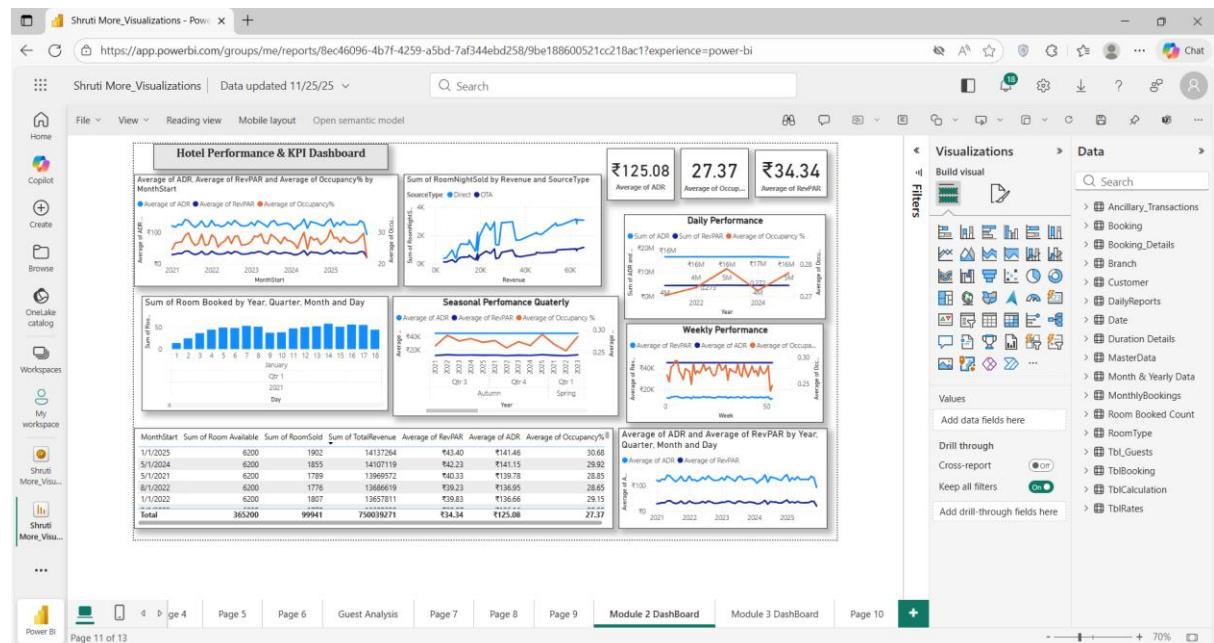
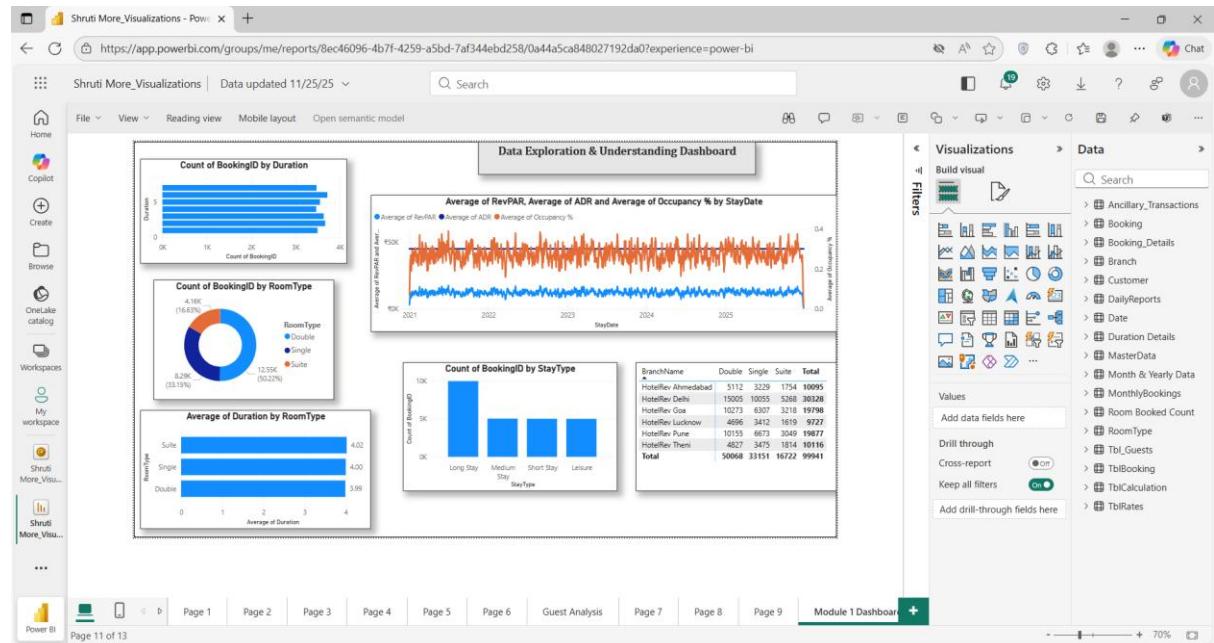
A horizontal bar chart compares how many customers (Direct vs OTA) belong to each guest category. This helps understand channel preference for different customer groups.

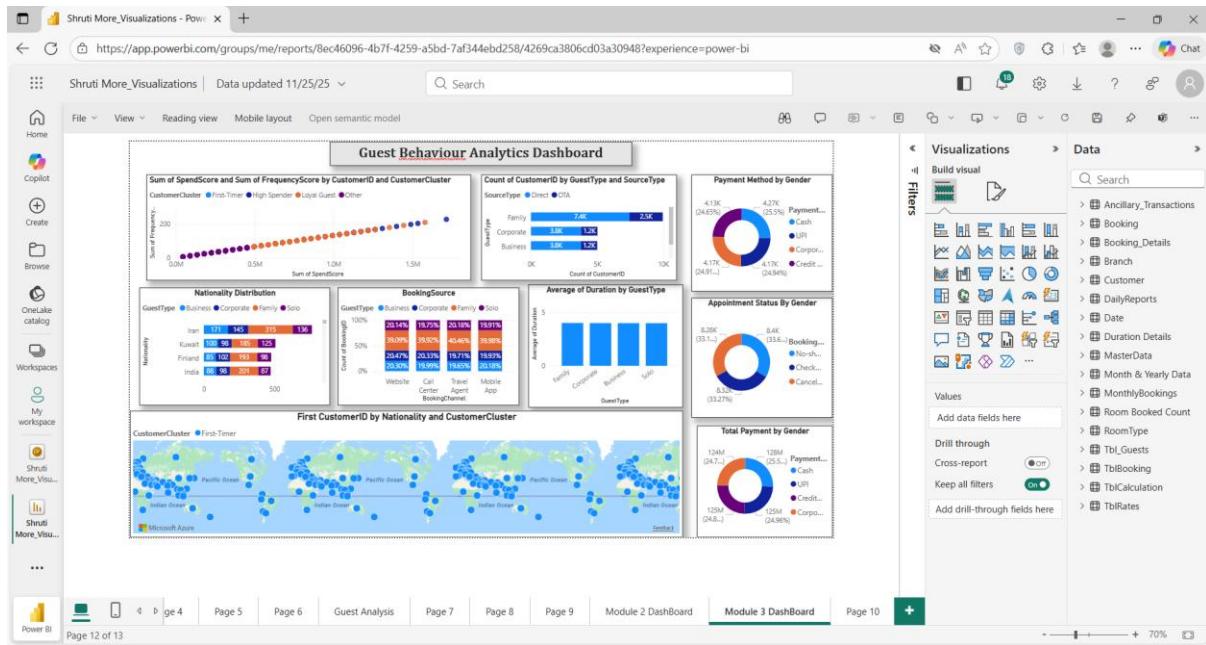


## I) Spend Score vs Frequency Score (Scatter Plot)

This scatter plot compares customers based on how much they spend (SpendScore) and how often they visit (FrequencyScore). Each dot represents a customer and colors indicate their CustomerCluster.

## DASHBOARD





## Insights and Learning from Module 3

### 1. Clear Customer Segmentation

Using calculated fields like GuestType and CustomerCluster, customers were grouped into Business, Family, Corporate, Solo, First-Timer, Loyal Guest, and High Spender. This helped understand each group's unique behavior.

### 2. Behaviour Patterns Identified

RFM scoring (Recency, Frequency, Spend) revealed how often customers visit, how recently they stayed, and how much revenue they contribute. This highlighted loyal vs. high-value customers.

### 3. Booking and Spending Trends

Visualizations showed differences in booking channels, payment methods, nationality distribution, and stay duration across guest types. These trends help identify which customer segments are most profitable.

### 4. Geographic Insights

Map visuals showed the spread of customers globally, helping identify strong markets and potential areas for growth.

### 5. Better Customer Understanding

By analysing segmentation, stay patterns, and RFM scores, the hotel can improve marketing, personalize service, and increase customer retention.

### 6. Strong Foundation for Predictive Analysis

The structured segmentation and scoring system prepares the data for deeper analysis like churn prediction and targeted promotions.

## Conclusion

Module 3 successfully enhanced customer understanding by segmenting guests based on their behavior, booking patterns, and spending levels. Through analytical fields like GuestType, CustomerCluster, and

RFM scoring, we identified distinct customer groups and their preferences. The visualizations provided meaningful insights into nationality distribution, booking sources, stay duration, and payment trends.

Overall, this module helps the hotel make data-driven decisions, improve customer engagement, personalize services, and strengthen marketing strategies. It provides a solid foundation for targeting high-value customers, improving retention, and planning future business growth.