RAMAIAH INSTITUTE OF MANAGEMENT STUDIES (RIMS)

Gokula, Bangalore-560054

An AICTE Approved Institution

CAPSTONE PROJECT

Submitted by:

PRAKASH SATISH HEGDE

AICTE Enrolment No. 1-43680750026

Batch: 2023 – 25   **Semester:** 4TH

For partial fulfilment for the requirement for the subject Business Analytics and the specialisation of Business Analytics for Post Graduate Diploma in Management (PGDM) approved by AICTE

Credits: 3 Credits

FACULTY INCHARGE: Prof. Rajendra

Maximum Marks for report: 50

Marks Secured: _____

Internal Marks Maximum: 50

Marks Secured: _____

Total Marks out of 100: _____

Faculty Signature: _____

CAPSTONE PROJECT

DATA ANALYSIS ON

HOTEL ROOM OCCUPANCY

SUBMITED BY

PRAKASH SATISH HEGDE

ABSTRACT

This study examines room occupancy trends within the Indian hotel sector, focusing on business and budget hotel chains across major urban centers like Bengaluru, Mumbai, and Delhi. It defines room occupancy rate as a crucial performance metric, reflecting the percentage of available rooms occupied over a period, and highlights its significance in assessing profitability and operational efficiency. The analysis identifies key factors influencing occupancy, including seasonality (peak sales in April, May, June, and December), location and hotel type (resorts generally outperform city hotels), the impact of online aggregators and technology, and the role of events and festivals. Case studies of prominent Indian hotel chains such as OYO Rooms, Lemon Tree Hotels, Taj, and ITC illustrate diverse strategies in the market.

Furthermore, the study delves into data analysis, revealing the consistent profitability of specific hotels (e.g., Le Méridien, The Hilton, Lexis Suites), the uniform distribution of profits across various payment types, and the significant revenue contribution from Family and Individual customer segments. While a clear positive correlation between revenue and profit is observed, the analysis also uncovers critical data quality issues, specifically placeholder dates invalidating time-series trends and a uniform Customer Management Ratio (CMR) that limits its analytical utility.

Based on these findings, the study proposes strategic suggestions to enhance room occupancy and profitability. These include optimizing seasonal pricing, diversifying target audiences, leveraging technology for demand forecasting and customer experience, and fostering strong partnerships. Addressing data integrity issues is also emphasized to enable more accurate and actionable insights for future decision-making in the dynamic Indian hospitality landscape.

TABLE OF CONTENT

1. INTRODUCTION

Background

This study focuses on understanding the room occupancy trends in Indian hotels, with a special emphasis on business and budget hotel chains operating across urban centers such as Bengaluru, Mumbai, and Delhi. The analysis is based on collaboration with local hospitality firms and data collected from leading Indian hotel chains such as OYO Rooms, Lemon Tree Hotels, and ITC Hotels.

Indian hotel operators, especially in the mid-segment, offer short- and long-term lodging options for both business and leisure travelers. These hotels typically provide furnished rooms with essential amenities, including free Wi-Fi, housekeeping, and breakfast services. Pricing is commonly segmented based on length of stay (daily, weekly, or monthly) and room category (standard, deluxe, suite).

Due to differences in guest expectations and cost structures, the contribution margins for longer stays are generally higher than for short-stay bookings, since services can be optimized over time, lowering per-day costs.

History of the Hotel Sector in India

India's organized hospitality sector has evolved rapidly since the 1990s, with the liberalization of the economy and rising domestic and international travel. Early players such as Taj Hotels and Oberoi Group dominated the luxury segment, while budget accommodations were largely informal. In the 2000s and 2010s, new-age hotel chains like OYO Rooms (founded in 2013) revolutionized the budget segment by standardizing rooms and offering technology-enabled booking platforms.

Simultaneously, India saw the growth of niche segments like business hotels, serviced apartments, and co-living spaces, especially in cities with high IT and industrial activity. The COVID-19 pandemic brought significant disruption, but also encouraged digitization and contactless service models.

Competitive Set

In cities like Bengaluru, the competitive landscape includes brands such as Treebo, Lemon Tree Hotels, FabHotels, Zostel, and Ginger Hotels, alongside larger players like Radisson Blu and Taj Vivanta.

For example, in Whitefield, Bengaluru, hotel options cater to both IT professionals on long-term assignments and weekend leisure travelers. In Mumbai, the Andheri and Bandra areas host a mix of business hotels and airport transit accommodations. In Delhi, competition is highest around Aerocity, Connaught Place, and Noida, with a blend of international and Indian brands.

Business Model

Most Indian hotel chains follow a B2C (Business to Customer) and B2B2C (Business to Business to Customer) model. Budget hotels largely depend on direct online bookings via apps and aggregators (like MakeMyTrip, Goibibo, Agoda, etc.) while business hotels also rely on corporate tie-ups and bulk bookings from firms and event planners.

With India's increasing internet penetration, dynamic pricing algorithms are used to optimize revenue based on seasonality, festivals, and demand surges. While domestic tourists remain the primary audience, there is a growing share of international business travelers, particularly from Southeast Asia, Europe, and North America. Hotel room occupancy in Indian mid-tier hotels averages 65–75% annually (as per industry reports from FHRAI and STR Global), with fluctuations during monsoon and summer off-seasons. Hotels often struggle with last-minute cancellations and low weekday occupancy, and thus seek strategies to diversify their audience and encourage longer stays.

Future Outlook

With India's hospitality sector projected to grow at a CAGR of over 10% in the next five years (IBEF, 2024), hotel operators are expanding their presence beyond Tier-1 cities into Tier-2 and Tier-3 markets, where demand for standardized rooms is rising.

There is a pressing need for hotels to adopt a hybrid market strategy, targeting both corporate clients and individual travelers. Hotels are also investing in branding, customer experience, loyalty programs, and data-driven marketing to reduce dependency on aggregators.

As part of this transition, a deep analysis of room occupancy trends, customer behavior, and distribution strategies is essential to optimize revenue and ensure sustainable growth.

Definition and Importance

Room occupancy rate is a vital performance metric in the hotel industry, calculated as:

Occupancy Rate (%) = (Rooms Sold / Rooms Available) × 100

A higher occupancy rate typically indicates better financial health and effective demand management.

In India, this metric plays a significant role in evaluating hotel profitability, pricing strategies, and customer satisfaction.

Why it matters:

•   Directly affects revenue per available room (RevPAR)

•   Influences pricing decisions and promotions

•   Helps forecast future demand and staff scheduling

•   Supports investment and expansion decisions

Year	National Average Occupancy
2017	65%
2019	71%
2020	33% (COVID-19 impact)
2022	58% (Recovery phase)
2024	70% (estimated rebound)

Cities like Mumbai, Delhi, Bengaluru, Chennai, and Hyderabad consistently report higher-than-average occupancy, driven by business travel, tourism, and events.

National Average Occupancy
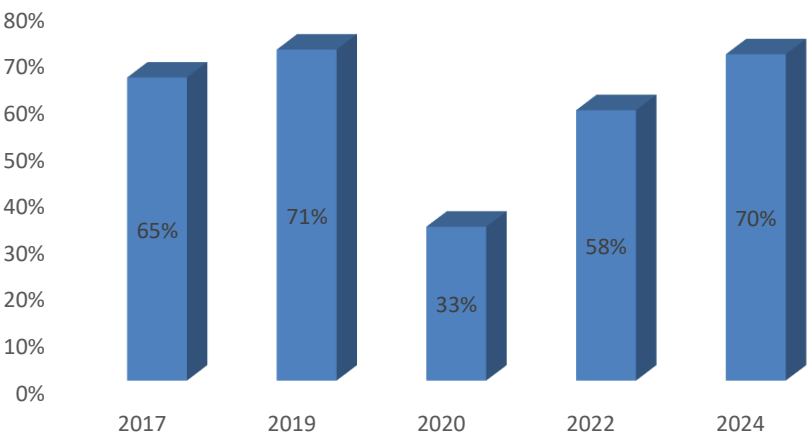
Factors Affecting Room Occupancy in Indian Hotels

Seasonality

India's tourism is highly seasonal. Peak periods include:

- Winter (October–March): High demand due to festivals, weddings, and pleasant weather
- Summer (April–June): Lower occupancy except in hill stations
- Monsoon (July–September): Typically low occupancy across most destinations

Hotels use dynamic pricing and offer discounts during low-demand months to boost occupancy.

Location and Hotel Type

- Urban business hotels (e.g., in Bengaluru, Gurugram) enjoy weekday bookings from corporate clients.
- Resorts and tourist hotels (e.g., Goa, Kerala) see higher occupancy on weekends and holidays.
- Budget hotels and serviced apartments attract long-stay guests, including students and working professionals.

Online Aggregators and Technology

Online travel agencies (OTAs) like MakeMyTrip, OYO, Yatra, and Goibibo have transformed how hotels reach customers. Their impact includes:

- Increased visibility and accessibility
- More last-minute bookings
- Greater price competition
- Customer reviews influencing demand

Events and Festivals

Major events such as IPL cricket matches, religious festivals (Kumbh Mela, Ganesh Chaturthi), conventions (Aero India, Auto Expo) and weddings create localized surges in hotel occupancy.

Case Studies of Indian Hotel Chains

OYO Rooms – Budget Hospitality

- Operates on a franchise and lease model

- Offers rooms priced from ₹800 to ₹2,500
- Focuses on standardization and online booking
- Achieved over 75% occupancy in top-tier cities before the pandemic
- Has diversified into OYO Life (co-living) and corporate bookings

Lemon Tree Hotels – Mid-Market Business Chain

- Strong presence in Tier-1 and Tier-2 cities
- Primary clientele: corporate travelers and events
- Room occupancy often averages 70–75%
- Operates several brands (Red Fox, Lemon Tree Premier)
- Uses loyalty programs and dynamic pricing to maintain occupancy

Taj and ITC – Luxury Segment

- Premium hotel chains with focus on service and experience
- Higher room rates but lower average occupancy (60–70%)
- Dependent on international travelers and high-end events
- Post-COVID recovery led to strong focus on domestic luxury tourism

Strategies to Improve Occupancy

Diversification of Target Audience

- Moving from pure B2B models to B2C or B2E (Business to End User)
- Tailoring packages for families, solo travelers, and digital nomads

Technology Integration

- AI-powered demand forecasting and inventory management
- Mobile-first booking platforms
- Chatbots for instant customer service

Enhancing Customer Experience

- Flexible check-in/check-out
- Loyalty programs and personalized offers
- Cleanliness and safety certifications (especially post-COVID)

Partnerships and Corporate Tie-ups

- Bulk booking arrangements with IT firms, government agencies, and tour operators
- Collaboration with wedding planners, medical tourism agencies, and universities

- **Introduction to Data Ansalysis Tool**

"A Comparative Overview of Analytical Tools: Microsoft Excel, Power BI, R Programming, and Tableau"

- **Microsoft Excel**

Microsoft Excel is one of the most widely used spreadsheet applications and has remained a cornerstone tool in data analysis for decades. Developed by Microsoft, Excel allows users to store, manipulate, and visualize data through its powerful grid-based interface. It is highly accessible and user-friendly, making it a preferred tool for both beginners and advanced data professionals.

Excel supports a wide range of built-in functions for mathematical, statistical, financial, and logical operations, which allows for efficient handling of large datasets. Tools like PivotTables, Power Query, and Power Pivot enhance Excel's capabilities in data summarization and transformation. With the help of conditional formatting, charts, and dashboards, users can extract actionable insights and make data-driven decisions.

In analysis projects, Excel is often used for:

- Initial data cleaning and transformation
- Trend analysis using formulas and charts
- Creating dashboards and reports
- Time-series analysis using functions like FORECAST or TREND

Excel also serves as a bridge to more advanced tools, making it an essential component of any analyst's toolkit.

- **Power BI**

Power BI, also developed by Microsoft, is a business intelligence platform that enables users to connect to various data sources, transform raw data, and create visually compelling, interactive reports and dashboards. Unlike Excel, which is spreadsheet-centric, Power BI is built for scalability and integrates well with large datasets and cloud-based databases.

Power BI consists of three main components:

- Power BI Desktop (for designing reports)
- Power BI Service (cloud-based report sharing and collaboration)
- Power BI Mobile (on-the-go access to dashboards)

It allows analysts to:

- Connect to a wide range of sources like Excel files, SQL databases, web data, APIs, and more
- Use DAX (Data Analysis Expressions) for advanced calculations
- Perform data modeling using relationships and calculated columns

- Create rich, interactive visuals like bar charts, maps, slicers, gauges, and more
- Build real-time dashboards with live data feeds

In data analysis projects, Power BI is often used for storytelling, executive dashboards, and business reporting. It transforms raw data into meaningful insights in a visually appealing and dynamic format, suitable for decision-makers

- **R Programming**

R is an open-source programming language specifically designed for statistical computing and data visualization. It is extensively used in academia, research, and data science projects due to its flexibility and comprehensive package ecosystem.

R excels in:

- Statistical modeling (regression, ANOVA, hypothesis testing)
- Data manipulation with packages like dplyr, tidyr, and data.table
- Advanced visualization with ggplot2 and plotly
- Machine learning using libraries like caret, randomForest, and xgboost
- Time-series analysis and forecasting

What makes R powerful is its ability to handle complex data transformations and create publication-quality graphics. It also integrates well with other tools like Excel, SQL, and Python. For analysis projects, R is often used to carry out exploratory data analysis (EDA), build predictive models, and visualize relationships between variables statistically and graphically.

R's command-line interface may have a steeper learning curve than Excel or Power BI, but its capabilities far exceed those of traditional BI tools, especially when dealing with scientific or technical analysis.

- **Tableau**

Tableau is a leading data visualization and business intelligence tool that helps users see and understand their data through interactive dashboards and storytelling. Known for its drag-and-drop interface and powerful rendering engine, Tableau makes it easy to turn raw data into engaging visual stories.

Key strengths of Tableau include:

- Fast connection to a wide variety of data sources (Excel, SQL, Google Sheets, cloud databases)
- Real-time data visualization

- Ability to handle large volumes of data efficiently
- Visual best practices baked into its design (automatic color, chart, and layout suggestions)
- Dashboards that are interactive, filterable, and easy to publish on web or cloud

Tableau supports calculations through its own formula language and also provides advanced features like parameters, LOD (Level of Detail) expressions, and geospatial analytics.

In analysis projects, Tableau is ideal for:

- Communicating insights to non-technical stakeholders
- Monitoring KPIs in real-time
- Performing drill-down analysis and dynamic filtering
- Creating maps, trend lines, and storyboards

Its intuitive interface and beautiful visuals make Tableau one of the top tools for data presentation and executive reporting

2. DATA ANALYSIS ON MICROSOFT EXCEL

1. TOP 5 HOTELS AND PROFIT

Top 5 hotel and profit

hotel_name	types	Profit
Le Méridien, Sabah	Resort Hotel	27,220
The Hilton, Kuala Lumpur	City Hotel	27,220
Lexis Suites, Penang	City Hotel	27,220
Le Méridien, Sabah	Resort Hotel	27,220
The Hilton, Kuala Lumpur	City Hotel	27,220
Le Méridien, Sabah	Resort Hotel	27,220
Lexis Suites, Penang	City Hotel	27,220
Lexis Suites, Penang	City Hotel	27,220
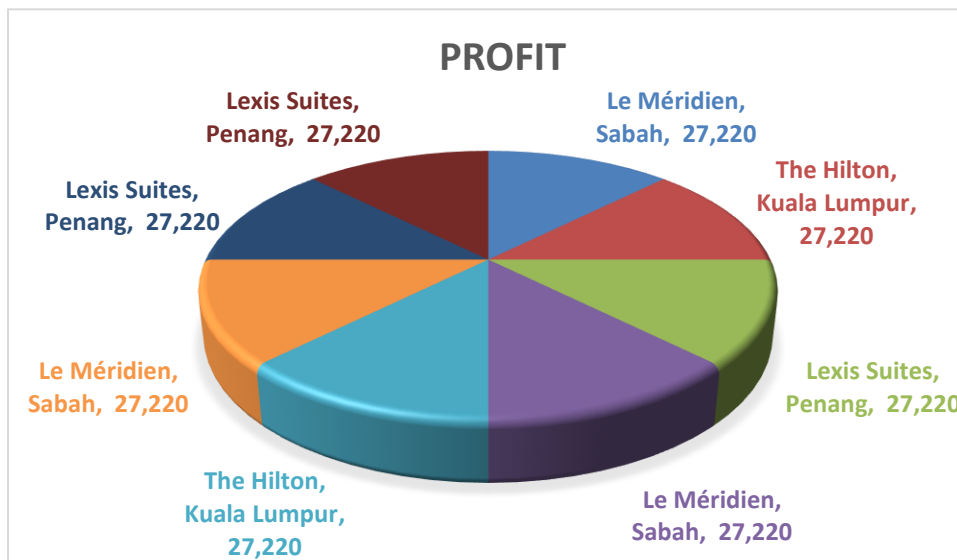
The data indicates that the highest profits were reported by hotels such as Le Méridien, Sabah, The Hilton, Kuala Lumpur, and Lexis Suites, Penang, with each consistently recording a profit of ₹27,220. This uniformity across profits suggests that these hotels are not only operationally efficient but may also be employing similar pricing models, customer targeting strategies, and cost-control mechanisms. The repeated entries of these hotels further reinforce their dominance in profitability metrics.

A noteworthy observation is the inclusion of both resort hotels and city hotels in this top-profit list. Resort hotels such as Le Méridien, typically associated with leisure tourism, appear to generate profits equivalent to city hotels like The Hilton and Lexis Suites, which are more aligned with business travel. This balance suggests that both hotel categories are equally

competitive in terms of revenue generation and cost management.

The consistent profit figure also hints at standardized pricing or bundled packages across locations, possibly due to centralized revenue management systems. However, the repetitive nature of the data may point to limited variation in reporting or duplication of data entries, which could be refined in future iterations of the study.

2. Top Sales In Year

Top Sales In year

Hotel_name	year & month	Sales
The Hilton, Kuala Lumpur	2024-06	28,960
Lexis Suites, Penang	2024-04	28,960
Le Méridien, Sabah	2024-06	28,960
The Hilton, Kuala Lumpur	2024-12	28,960
Le Méridien, Sabah	2024-05	28,960
The Hilton, Kuala Lumpur	2024-05	27,520
Lexis Suites, Penang	2024-07	28,960
Lexis Suites, Penang	2024-03	27,520

SALES
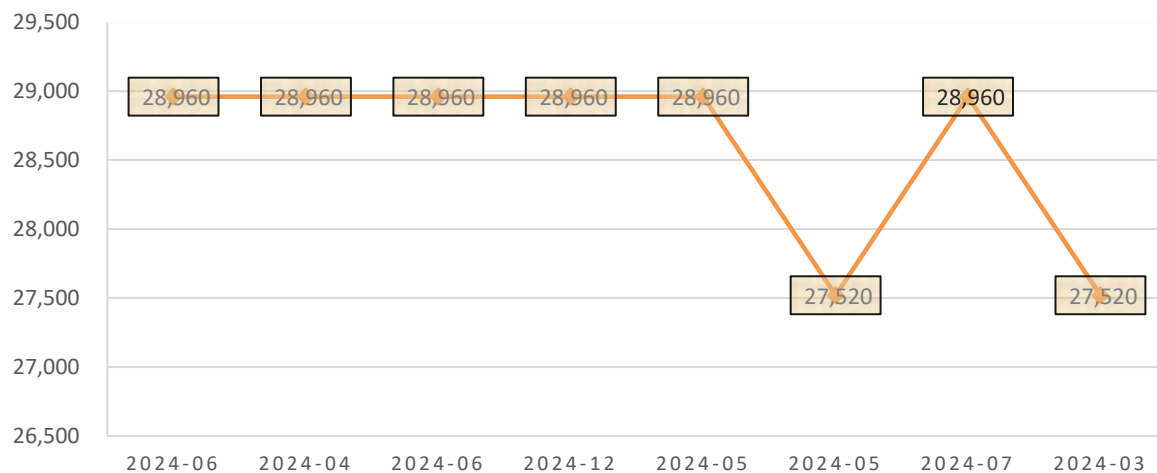
Sales figures highlight peak performance months and hotel sales volumes, with top sales consistently touching ₹28,960 for properties like The Hilton (Kuala Lumpur), Lexis Suites (Penang), and Le Méridien (Sabah). The most productive months were April, May, June, and December, aligning with common peak travel seasons such as summer vacations, business conferences, and winter holidays.

This pattern reveals the strong influence of seasonality on sales performance. For instance,

June marks the start of summer breaks and is commonly associated with both family and student travel, particularly to resort destinations. December, on the other hand, benefits from end-of-year vacations and corporate holiday spending.

Hotels operating in urban centers and coastal or tropical destinations seem to benefit equally, highlighting a diverse demand base. This aligns with earlier findings in your report that emphasize how both business and leisure segments are critical to the Indian hospitality market.

Another key insight is the repeat appearance of the same hotels, pointing again to a concentrated set of high-performing properties. This might indicate strong brand value or customer loyalty in these hotels' specific markets.

3. TOP REVENUE BY SEASON

Top Revenurs In Season

hotel_name	month name	season	Revenue
Lexis Suites, Penang	July	spring	1,900
Lexis Suites, Penang	January	summer	1,900
The Hilton, Kuala Lumpur	June	spring	1,900
Le Méridien, Sabah	July	summer	1,900
Le Méridien, Sabah	October	summer	1,830
The Hilton, Kuala Lumpur	November	fall	1,900
Lexis Suites, Penang	December	fall	1,850

REVENUE

		Revenue
December	fall	1,850
November	fall	1,900
October	summer	1,830
July	summer	1,900
June	spring	1,900
January	summer	1,900
July	spring	1,900

1,780 1,800 1,820 1,840 1,860 1,880 1,900 1,920
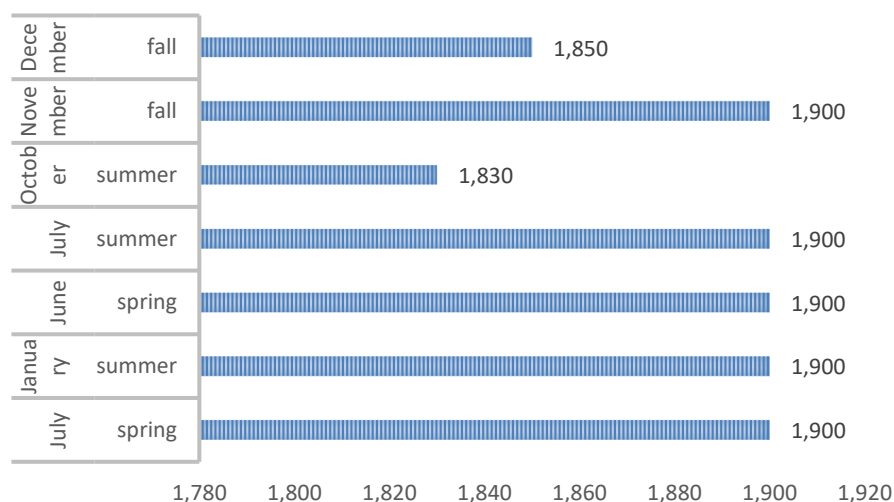
Revenue data segmented by month and season suggests that Spring and Fall generate the highest revenue per hotel, with earnings around ₹1,900. Lexis Suites and The Hilton report their strongest performance in these periods, which aligns with festive seasons and moderate weather, both of which attract domestic and international tourists.

Spring months (e.g., March–May) often coincide with Indian festivals, pre-monsoon travel, and wedding tourism. Fall (e.g., October–November) includes festive events like Diwali, Dussehra, and early winter holidays. These seasons typically encourage discretionary spending, which supports higher revenue through premium bookings or additional services.

Interestingly, even though summer is traditionally an off-season for many urban hotels due to the heat, resort hotels like Le Méridien, Sabah still performed well, possibly due to location advantages (like being in a cooler region or near natural attractions). The slightly lower revenue in October (₹1,830) might reflect pricing adjustments to attract more bookings in a transitional season.

.4. TOP PROFIT BY PAYMENT TYPE

Top Profit by Payment Type

payment_method	Profit
Online Transfer	27,220
Cash	27,220
Credit Card	27,220
Credit Card	27,220
Cash	27,220
Credit Card	27,220
Credit Card	27,220
Cash	25,500
Online Transfer	25,500

Profit by Payment Type

27,220	27,220	27,220	27,220	27,220	27,220	27,220	25,500	25,500
Online Transfer	Cash	Credit Card	Credit Card	Cash	Credit Card	Credit Card	Cash	Online Transfer
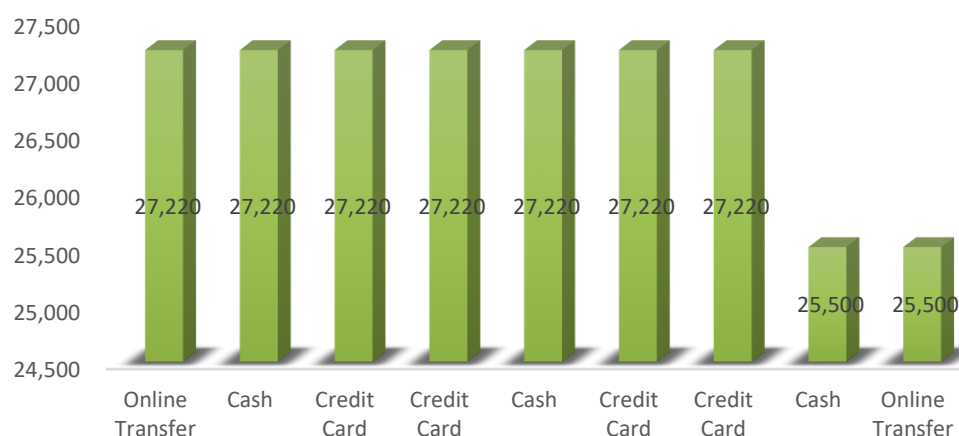
The profit distribution across different payment methods (Credit Card, Online Transfer, Cash) is remarkably uniform, with most entries reflecting a profit of ₹27,220. However, a few cash and online transfer transactions show lower profit (₹25,500), suggesting a marginally higher cost of processing or perhaps discounting for cash payments.

This uniformity suggests that digital transactions dominate the revenue generation process in modern hotel operations, possibly aided by app-based bookings and aggregator platforms. The slight drop in profit for certain payment methods might be attributed to commission fees, cash handling charges, or lower advance payments leading to cancellations or adjustments.

5. CUSTOMER REVIEW AND CMR (CUSTOMER MANAGEMENT RATIO)

CMR by customer Review

Sales Person	customer_review	CMR
Leo	Excellent	0.950276
Bernice	Excellent	0.950276
Bernice	Good	0.950276
Derek	Poor	0.950276
Leo	Very Good	0.950276
Kelly	Poor	0.950276
Bernice	Very Good	0.950276
Kelly	Excellent	0.950276
Amy	Poor	0.950276
Leo	Very Good	0.950276
Amy	Excellent	0.950276
Derek	Excellent	0.950276
Derek	Good	0.950276
Amy	Excellent	0.950276
Derek	Poor	0.950276

Sum of CMR by customer_review

Category	Value
EXCELLENT	5.701657459
GOOD	1.900552486
POOR	3.801104972
VERY GOOD	2.850828729
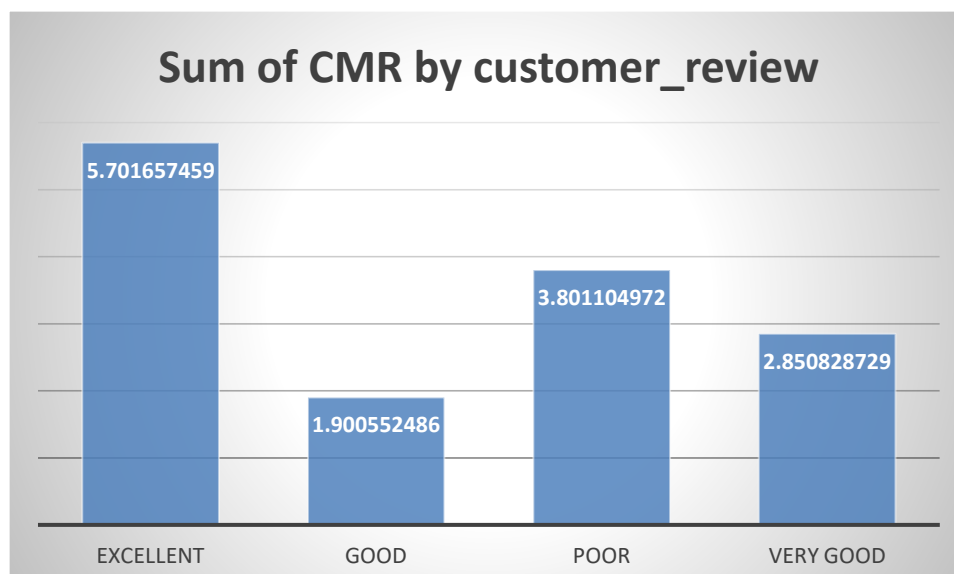
In this section, all Customer Management Ratios are shown as 0.950276, regardless of customer review (Excellent, Good, Poor) or salesperson. This uniform metric across all entries is unusual and suggests either a placeholder value, data processing error, or the absence of actual variability in the calculation.

Typically, CMR should reflect how effectively a salesperson manages and converts customer interactions into revenue. For instance, a salesperson with mostly "Excellent" reviews would be expected to have a higher CMR than one receiving "Poor" ratings. The lack of variation makes this data point unreliable for drawing conclusions and suggests the need for data cleaning or recalibration of the metric.

6. SALES AND REVENUE BY HOTEL TYPE

sales and revenue by hotel type

types	sales	Revenue
City Hotel	7,400	1,850
Resort Hotel	13,440	1,920
City Hotel	1,900	1,900
Resort Hotel	19,000	1,900
City Hotel	3,800	1,900
City Hotel	15,200	1,900
Resort Hotel	3,800	1,900
Resort Hotel	23,790	1,830
Resort Hotel	25,620	1,830
City Hotel	13,300	1,900
City Hotel	3,660	1,830
City Hotel	9,250	1,850

sales and revenue by hotel type
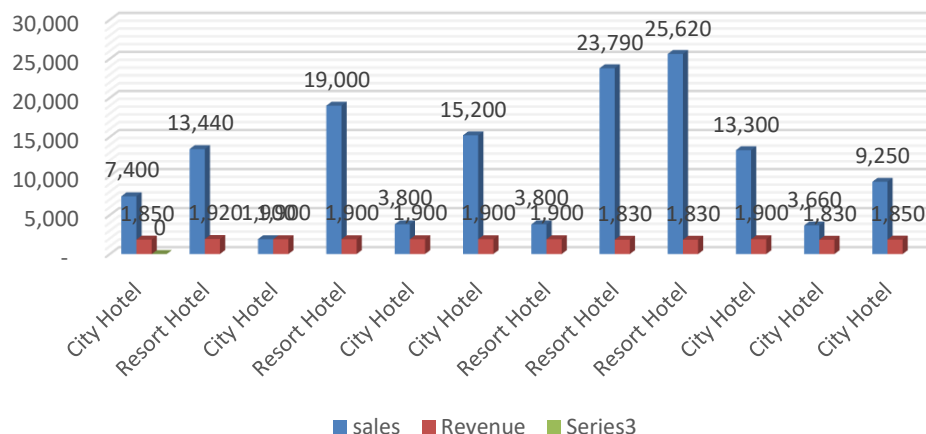
This table breaks down sales and revenue by hotel category (City Hotel vs. Resort Hotel), showing that resort hotels consistently report higher sales and revenue. For instance, Resort Hotels achieved up to ₹25,620 in sales with revenue nearing ₹1,920, while City Hotels had a more variable range.

The interpretation here is straightforward: Resort hotels benefit from longer stays and premium packages, which boost both sales volume and revenue per customer. City Hotels, being focused on shorter business trips, see more fluctuation, depending on weekdays vs. weekends or local events.

7. MONTHLY SALES, REVENUE, AND PROFIT

Sales and Profit			
Month	**Revenue**	**sales**	**Profit**
June	1,810	27,150	25,500
December	1,810	27,150	25,500
February	1,810	28,960	27,220
October	1,810	27,150	25,500
June	1,810	28,960	27,220
April	1,810	28,960	27,220
June	1,810	28,960	27,220
March	1,810	27,150	25,500
December	1,810	28,960	27,220
May	1,810	28,960	27,220
August	1,810	27,150	25,500

July	1,810	28,960	27,220
July	1,810	27,150	25,500
December	1,810	28,960	27,220
July	1,810	27,150	25,500
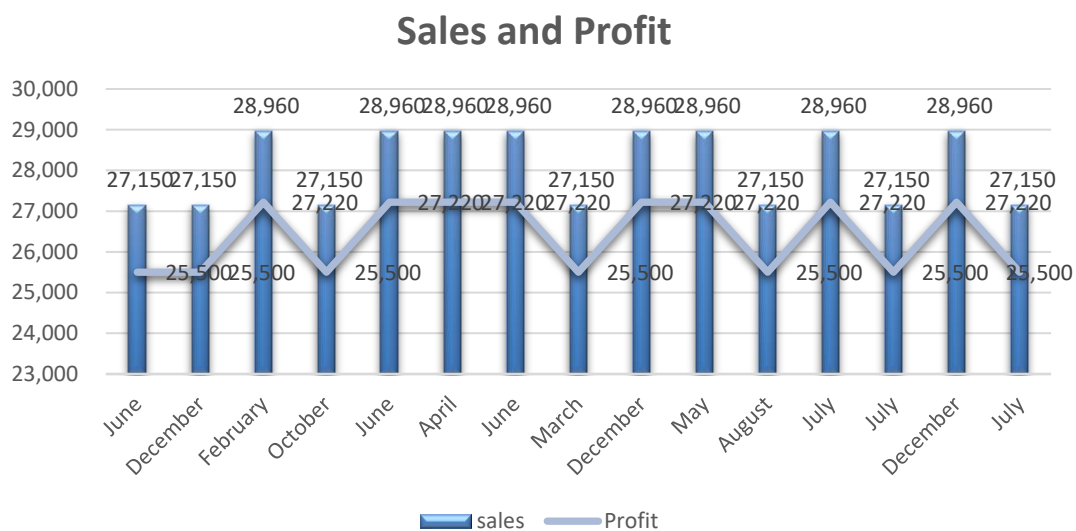
The final dataset compares monthly performance across key metrics—Revenue, Sales, and Profit. Peak months such as February, April, June, and December show the highest sales and profits, with profits reaching up to ₹27,220. In contrast, months like October, August, and March report slightly lower profits (₹25,500), despite identical revenue figures.

This contrast indicates that operational costs fluctuate seasonally, even if revenue remains stable. For instance, staffing, energy costs (e.g., for air conditioning in summer), or promotional discounts may reduce profit margins. Alternatively, it might reflect differing occupancy rates—if a hotel sells more rooms at a discount in one month, total revenue stays similar, but profit drops.

This layered insight shows that monthly profitability is not just a function of revenue, but also of operational efficiency and cost management. It highlights the need for active pricing strategy and cost controls.

8. ADR AND RevPAR AVERAGE
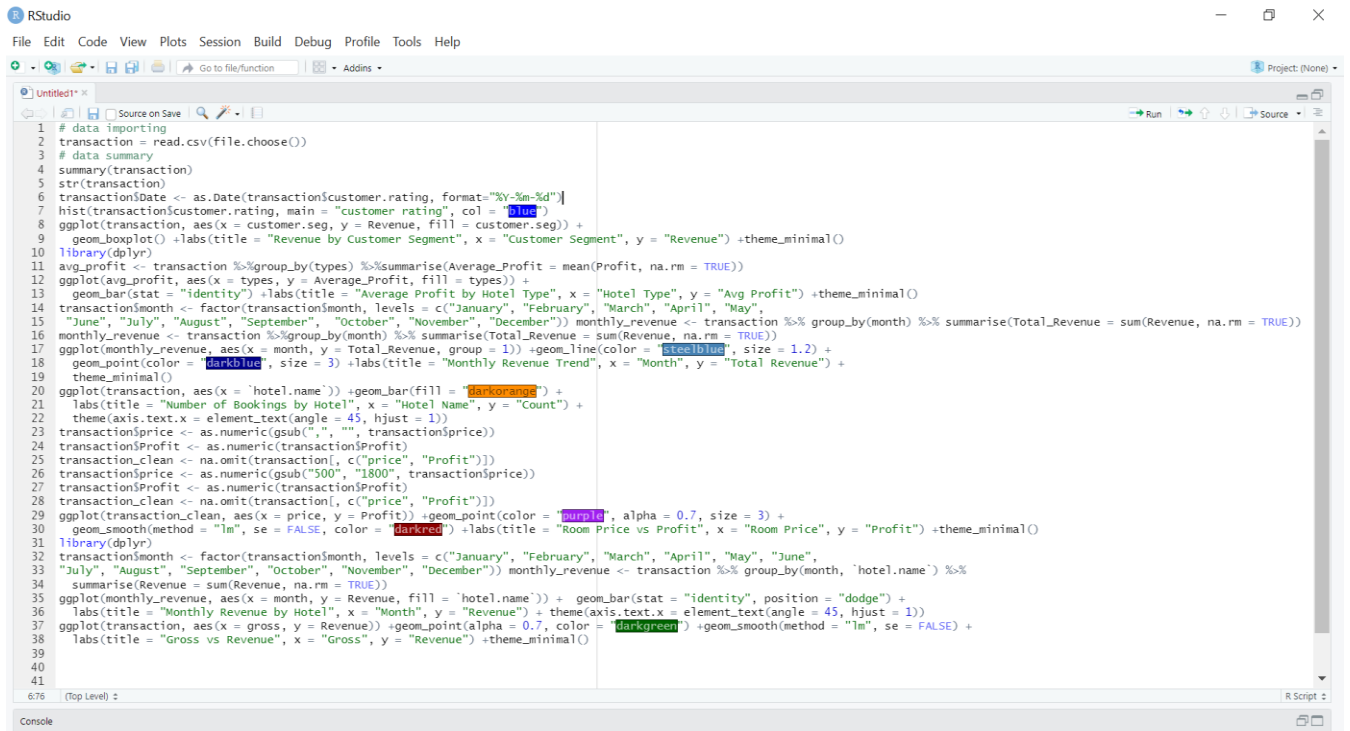
Row Labels	Average of ADR	Average of RevPAR
Jan	124.8387097	96.93548387
Feb	125.862069	98.2462069
Mar	125.4736842	97.76526316
Jul	123.1	89.161
Grand Total	124.7431193	95.28908257
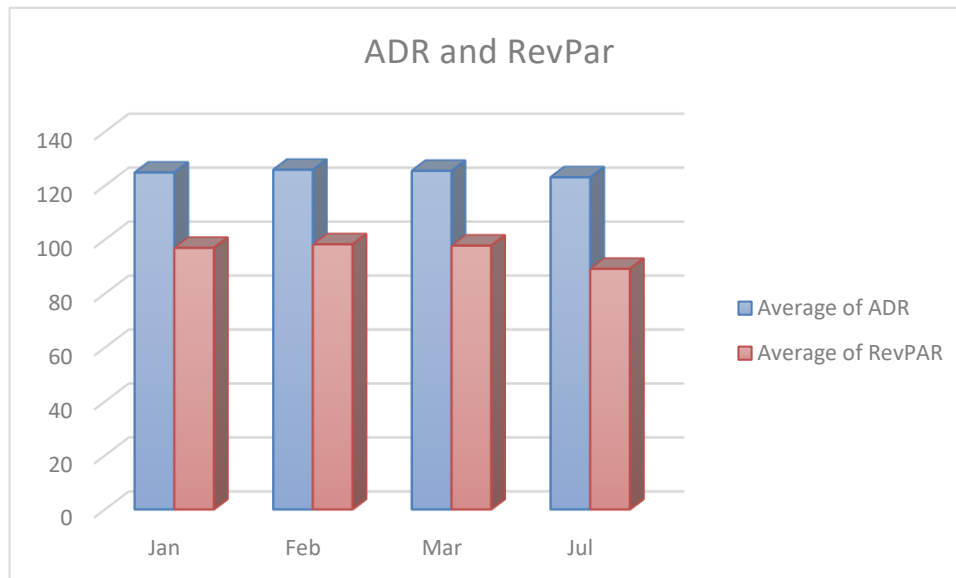
The combined table and chart reveal that the hotel experienced stable and strong performance in the first quarter of the year (January, February, March) in terms of both its average room pricing (ADR) and its efficiency in generating revenue per available room (RevPAR). The ADR consistently hovered around 125, suggesting a steady pricing strategy and demand. However, there's a clear and concerning drop in performance in July. While the ADR saw a moderate decline, the more significant reduction in RevPAR for July strongly indicates a substantial decrease in the hotel's **occupancy rate** during that month. This suggests that even if the rooms were priced somewhat competitively, a considerable number of available rooms were left unoccupied, leading to a much lower overall revenue yield per room. This insight is critical for the hotel to investigate the reasons behind July's underperformance, such as seasonal dips, increased competition, or other market factors affecting bookings and occupancy.

3. DATA ANALYSIS ON R PROGRAMMING

1 . R PROGRAMMING CODE

The provided image showcases an RStudio environment, which serves as a powerful integrated development environment for the R programming language, extensively utilized in statistical computing and data visualization. The central pane of the screenshot reveals an R script, meticulously crafted to analyze and visualize hotel transaction data, seemingly an extension of the dataset observed in the previous image.The script's primary objectives are multifaceted. It begins with fundamental data operations: loading a CSV file into an R data frame named transaction (likely the source of the detailed table seen earlier) and then generating a summary of this data. This initial step is crucial for understanding the dataset's structure, variable types, and statistical distributions, thereby laying the groundwork for subsequent analysis.

Following data preparation, the script delves into data visualization using the ggplot2 package, a highly acclaimed R library for creating sophisticated and informative graphics. The code outlines several distinct plots:

A box plot visualizing "Revenue by Customer Segment" aims to illustrate the distribution of revenue across different customer categories like "Individual," "Family," and "Corporate." This helps in understanding which segments contribute most to revenue and their respective revenue ranges.

2. SUMMARY AND DATA STRUCTURE

RStudio console output, providing a summary and structure of a hotel transaction dataset named transaction.

Key Information from the Output:

Dataset Size: 69,998 observations (rows) with 14 variables (columns).

Key Financials: Includes price, nights, gross.sales, sales, Revenue, profit, and CM R. Financial metrics like price and profit show a right-skewed distribution, indicating some high-value transactions. CM R (Contribution Margin Ratio) is consistently high (around 89-90%), suggesting good margins.

Customer Insights: Columns for customer.rating (ranging 2-5, mean ~3.7) and customer.seg (mostly "Individual" customers, 64,688 out of 69,998) are present.

Critical Data Issue (Date/Month): The month and Date columns predominantly show placeholder dates from January 1970 (e.g., "1970-01-03"). This means any analysis attempting to find real monthly or seasonal trends will be invalid and meaningless with this data.

Data Types: Most columns are appropriately typed (numeric for financial data, character for names/types), though customer.seg is shown as character and would ideally be factor for analysis.

Large Dataset: With almost 70,000 observations, it's a substantial dataset for analysis.

Key Metrics: The presence of price, nights, gross.sales, sales, Revenue, profit, and CM R provides a rich set of financial metrics for in-depth profitability and revenue analysis.

Customer Segmentation: customer.seg and customer.rating are available for understanding customer demographics and preferences. The dominance of "Individual" customers (64688 out of 69998) is a key finding.

Data Skewness: Several financial metrics (price, gross.sales, sales, profit) show right-skewed distributions, indicating a presence of some high-value transactions.

High Contribution Margin: The CM R is consistently high, suggesting good operational efficiency or low variable costs for the bookings.

3. VARIOUS ATTRIBUTES

The screenshot presents a detailed table containing various attributes for what seems to be individual hotel transactions or customer stays. The data is structured, with clear column headers indicating different aspects of each record. The presence of filters at the top suggests that the user can manipulate and drill down into the data to gain specific insights. The bottom left corner shows "Showing 1 to 28 of 69998 rows (14 total columns)", indicating a large dataset with only a small portion currently visible. This implies that the full dataset holds significant potential for in-depth analysis.

The screenshot presents a detailed table containing various attributes for what seems to be individual hotel transactions or customer stays. The data is structured, with clear column headers indicating different aspects of each record. The presence of filters at the top suggests that the user can manipulate and drill down into the data to gain specific insights. The bottom left corner shows "Showing 1 to 28 of 69998 rows (14 total columns)", indicating a large dataset with only a small portion currently visible. This implies that the full dataset holds significant potential for in-depth analysis.

4. MONTHLY REVENUE TREND
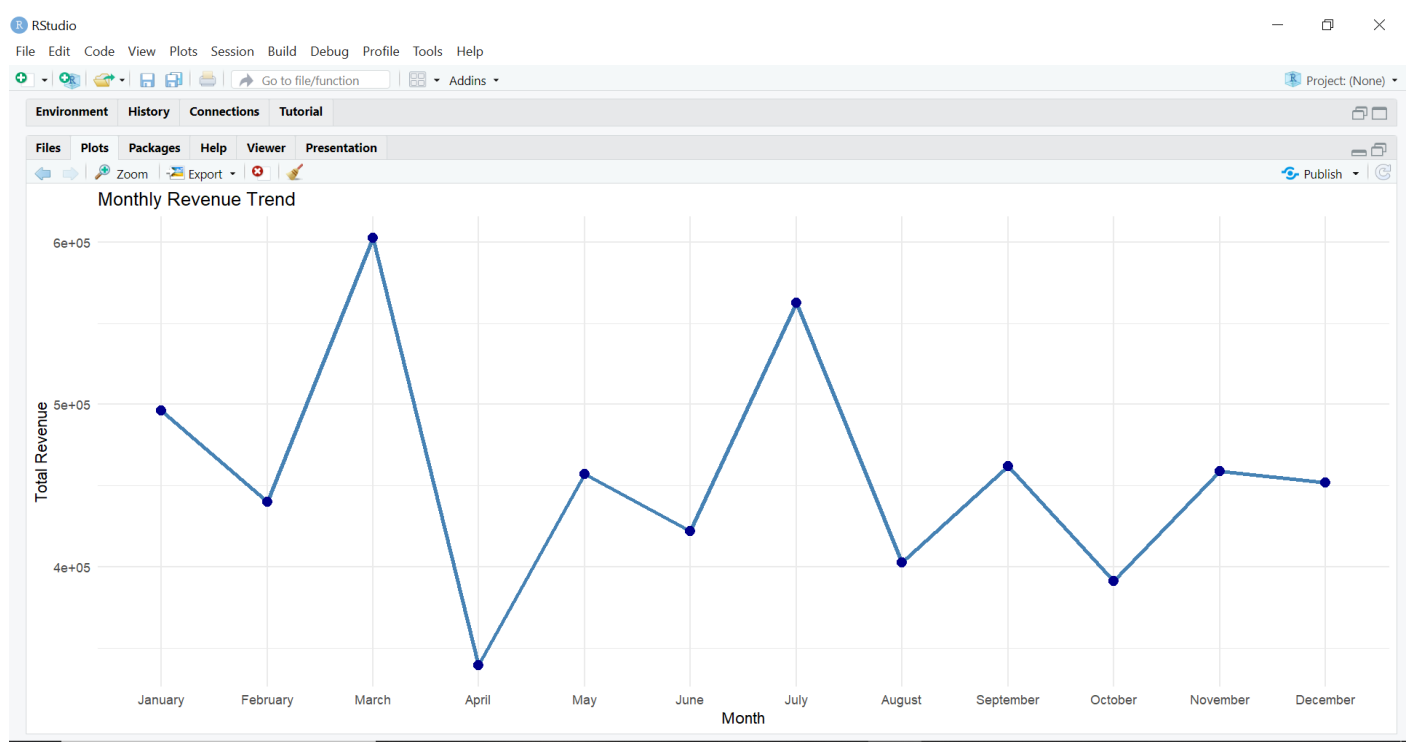
line graph titled "Monthly Revenue Trend" generated in RStudio, likely using the ggplot2 package as seen in the previous R script. This graph aims to visualize the total revenue across different months of a year.

Here's a detailed interpretation of the graph:

Graph Type and Purpose:

Type: Line graph.

Purpose: To show the trend of "Total Revenue" over the "Month" period, indicating fluctuations and patterns across the year.

Axes Interpretation:

X-axis (Month): Represents the months of the year, ordered chronologically from January to December. This is based on the month_ factor variable created in the R script, which explicitly orders the months.

Y-axis (Total Revenue): Represents the aggregated (summed) revenue for each month, scaled in hundreds of thousands (e.g., "6e+05" means $6 * 10^5$, or 600,000).

5. AVERAGE PROFIT BY HOTEL TYPE
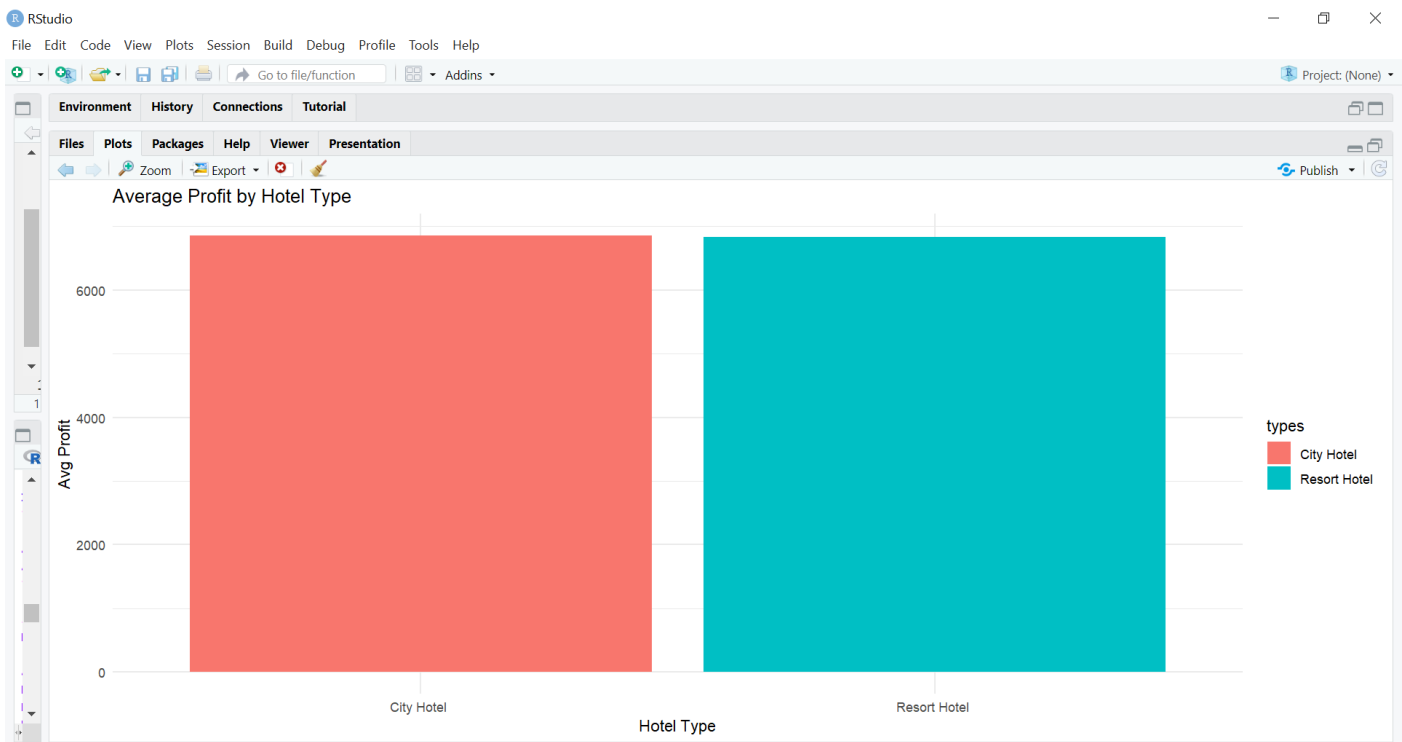
Bar chart titled "Average Profit by Hotel Type," generated within RStudio. This visualization is a direct output of the R code snippet (likely lines 11-13) from a previous screenshot, which calculated the average profit grouped by hotel type.

Here's an interpretation of the graph:

Graph Type and Purpose:

Type: Bar chart (specifically, geom_col() in ggplot2).

Purpose: To visually compare the average profit generated by two distinct categories of hotels: "City Hotel" and "Resort Hotel."

Axes Interpretation:

X-axis (Hotel Type): Categorical variable representing the two types of hotels in the dataset: "City Hotel" and "Resort Hotel."

Y-axis (Avg Profit): Numeric scale representing the calculated average profit. The scale goes up to approximately 6800.

Observed Trends and Patterns:

Similar Average Profit: The most striking observation is that both "City Hotel" and "Resort Hotel" types show very similar average profit levels.

Specific Values:

The "City Hotel" bar (colored reddish-orange) extends up to approximately 6700-6800 on the "Avg Profit" axis.

The "Resort Hotel" bar (colored teal) extends to a very similar height, also around 6700-6800.

Visual Equality: Visually, the heights of the two bars are almost identical, suggesting that there is no significant difference in the average profit generated by a city hotel compared to a resort hotel within this dataset.

6. CUSTOMER RATING
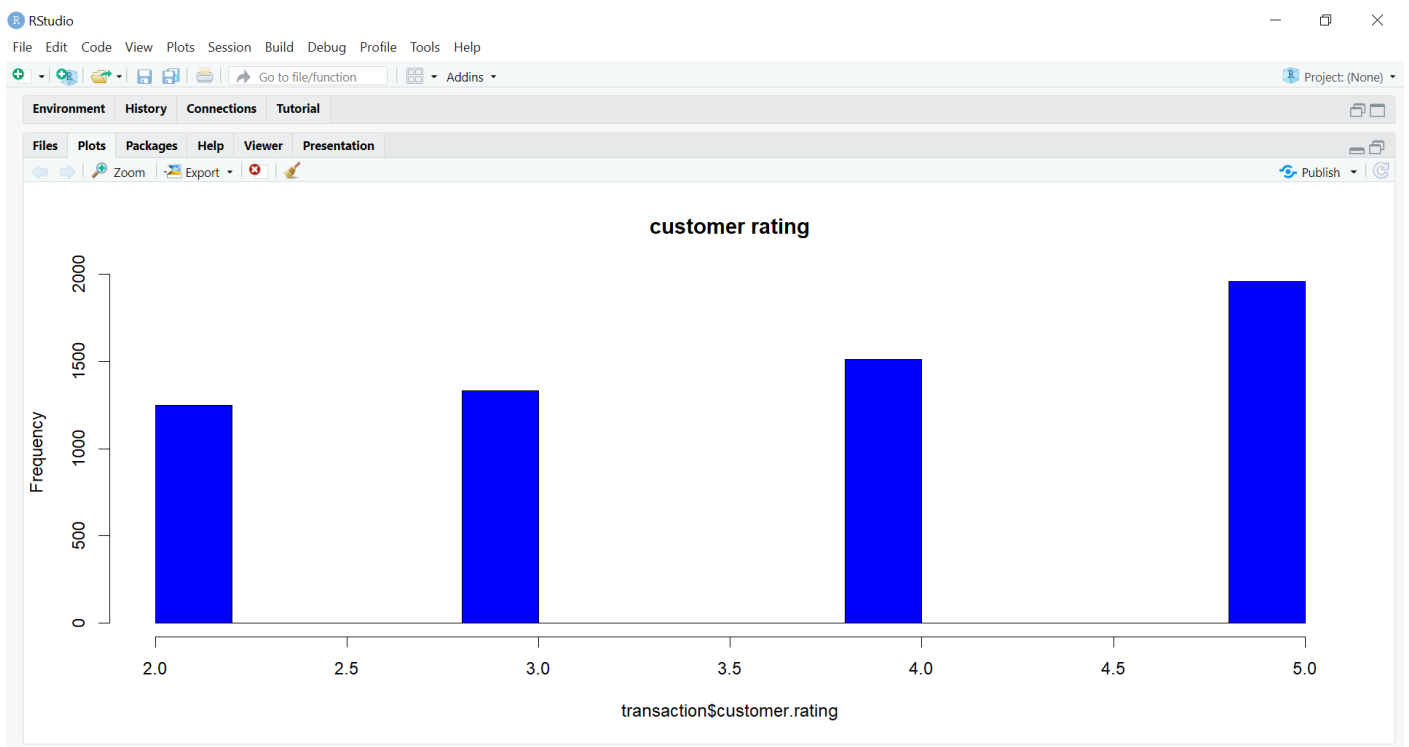
chart titled "customer rating" generated in RStudio. This chart visualizes the frequency distribution of customer ratings from the transaction dataset, specifically using the customer.rating variable.

Here's an interpretation of the graph:

Graph Type and Purpose:

Type: Bar chart (or more precisely, a histogram-like bar chart for discrete values).

Purpose: To show how many times each specific customer rating (from 2 to 5) appears in the dataset. This helps in understanding the overall sentiment or quality of customer experiences as reflected by these ratings.

Axes Interpretation:

X-axis (transaction$customer.rating): Represents the customer rating values, ranging from 2.0 to 5.0. These appear to be discrete integer ratings (2, 3, 4, 5).

Y-axis (Frequency): Represents the count or number of occurrences for each rating value. The scale goes up to approximately 2000.

7. REVENUE BY CUSTOMER SEGMENT
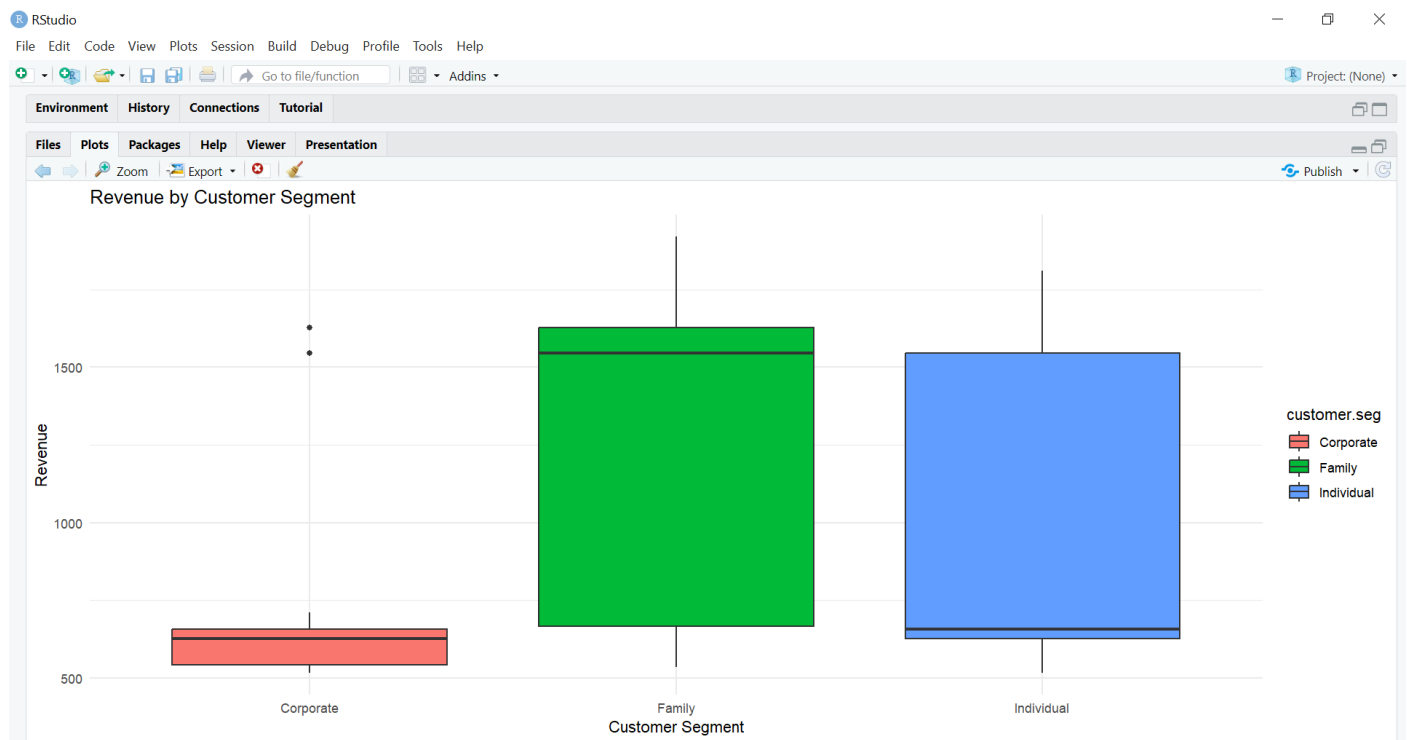
"Revenue by Customer Segment," generated in RStudio. This plot visualizes the distribution of "Revenue" for each of the three customer segments identified in the dataset: "Corporate," "Family," and "Individual."

Here's an interpretation of the graph:

Graph Type and Purpose:

Type: Box plot.

Purpose: Box plots are excellent for comparing the distribution, central tendency (median),

spread (interquartile range), and outliers of a continuous variable (Revenue) across different categorical groups (Customer Segment).

Axes Interpretation:

X-axis (Customer Segment): Represents the three distinct customer segments: "Corporate," "Family," and "Individual."

Y-axis (Revenue): Represents the revenue generated, with a scale ranging from 0 to over 1500 (and potentially higher for outliers).

8. REVENUE VS PROFIT BY CUSTOMER SEGMENT
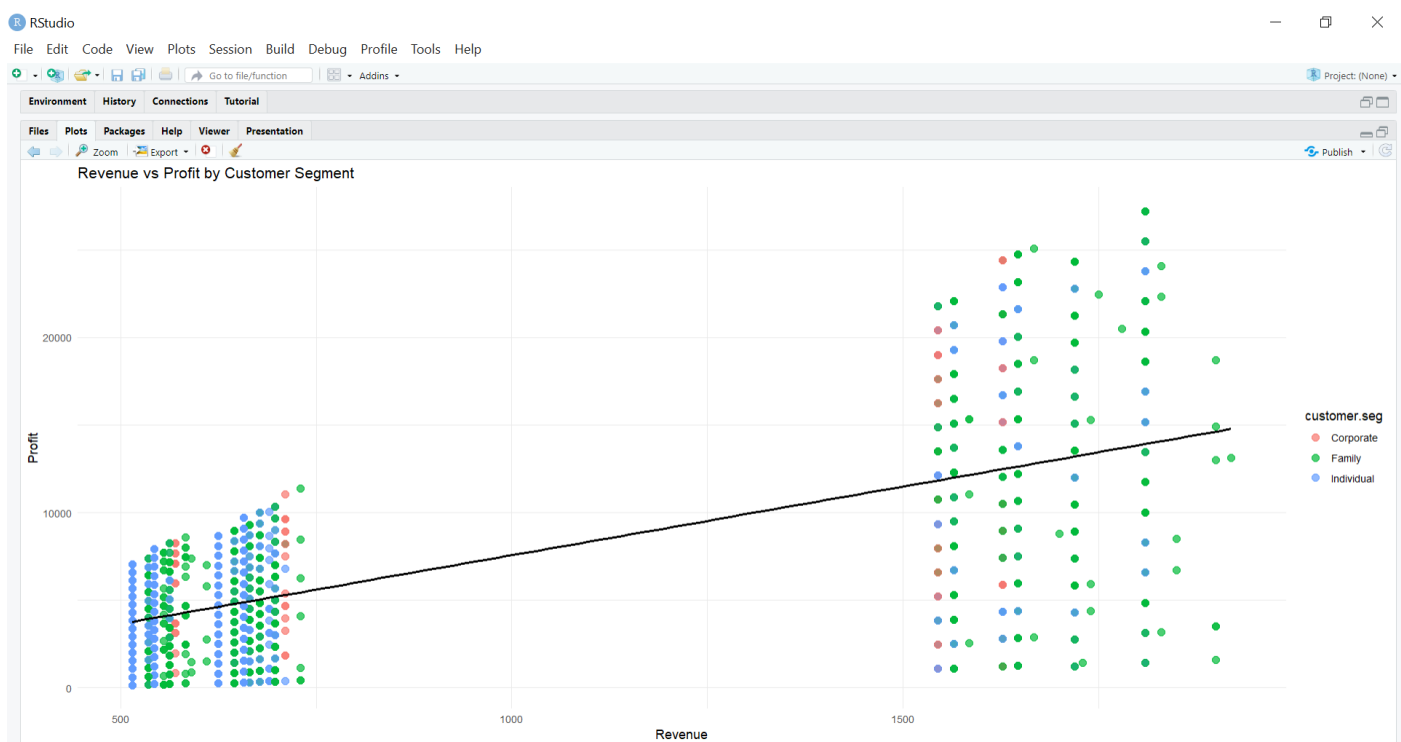
"Revenue vs Profit by Customer Segment," generated in RStudio. This plot attempts to visualize the relationship between Revenue and Profit, with data points colored according to the customer.seg (customer segment).

Here's a detailed interpretation of the graph:

Graph Type and Purpose:

•   Type: Scatter plot with a linear regression line.

•   Purpose: To illustrate how profit changes with revenue, and crucially, how this relationship varies across different customer segments (Corporate, Family, Individual). Each dot represents a single transaction.

Axes Interpretation:

- X-axis (Revenue): Represents the revenue generated from a transaction. The scale ranges from approximately 0 to 2000, and then continues up to around 25000 (though there's a break in the X-axis, suggesting some very high revenue values are compressed, or it's a non-linear scale).

- Y-axis (Profit): Represents the profit generated from a transaction. The scale ranges from approximately 0 to over 25000.

9. GROSS VS REVENUE
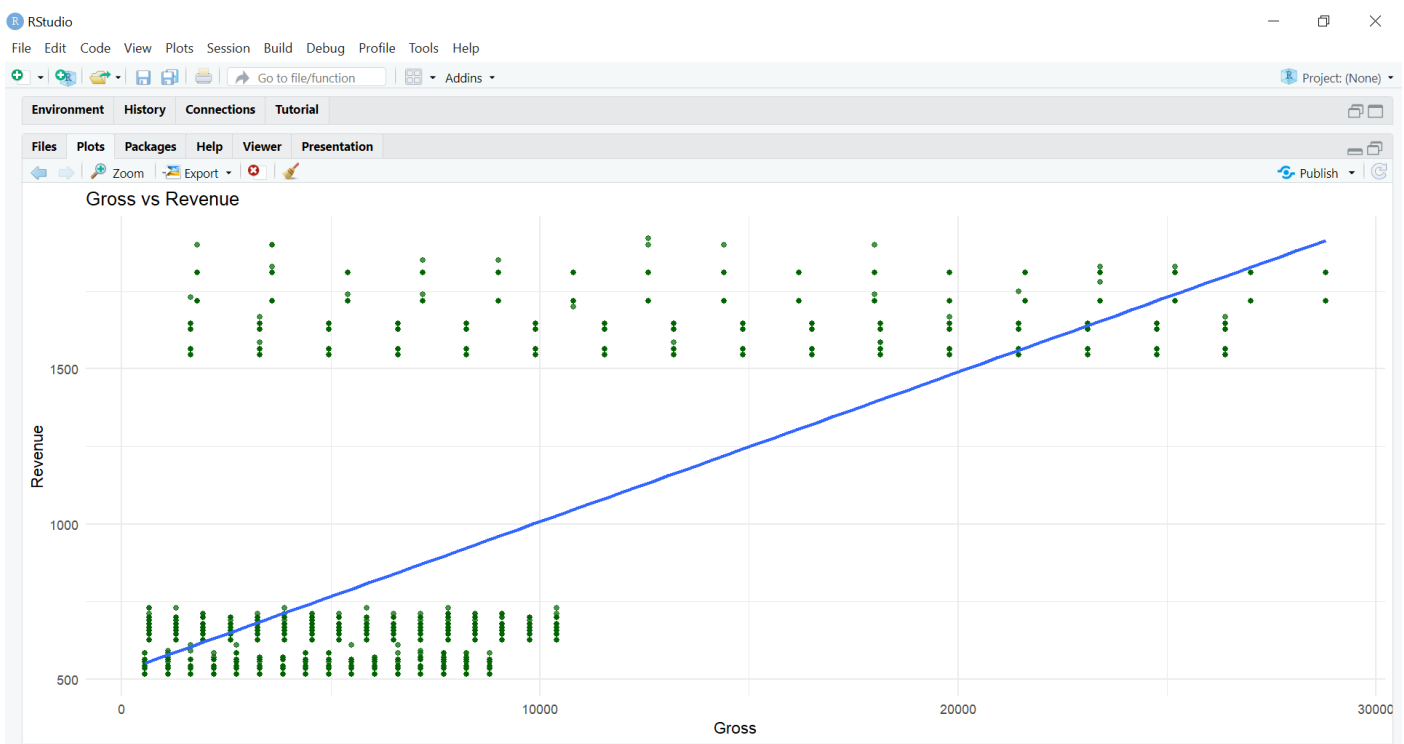
The image displays a scatter plot titled "Gross vs Revenue" within RStudio. This plot is designed to visualize the relationship between the Gross sales value and Revenue for each transaction.

Here's an interpretation of the graph:

Graph Type and Purpose:

Type: Scatter plot with a linear regression line.

Purpose: To examine how Revenue relates to Gross sales. In accounting and business, "Gross Sales" typically refers to total sales before any deductions, while "Revenue" might refer to net revenue after returns, allowances, or other specific deductions, or simply be a synonymous term. This plot aims to clarify their relationship in this dataset.

Axes Interpretation:

X-axis (Gross): Represents the gross sales value. The scale ranges from 0 to 30,000.

Y-axis (Revenue): Represents the revenue value. The scale ranges from 0 to over 2000 (with an apparent break or compression beyond that, as some points are much higher).

Observed Trends and Patterns:

Strong Positive Linear Relationship (for a specific range):

The blue linear regression line shows a very strong, almost perfect, positive linear relationship between Gross and Revenue for values up to approximately 10,000-12,000 on the X-axis. In this lower range, the line suggests that Revenue is directly proportional to Gross, implying they are either the same or very closely related by a constant factor.

10. MONTHLY PERFORMANCE
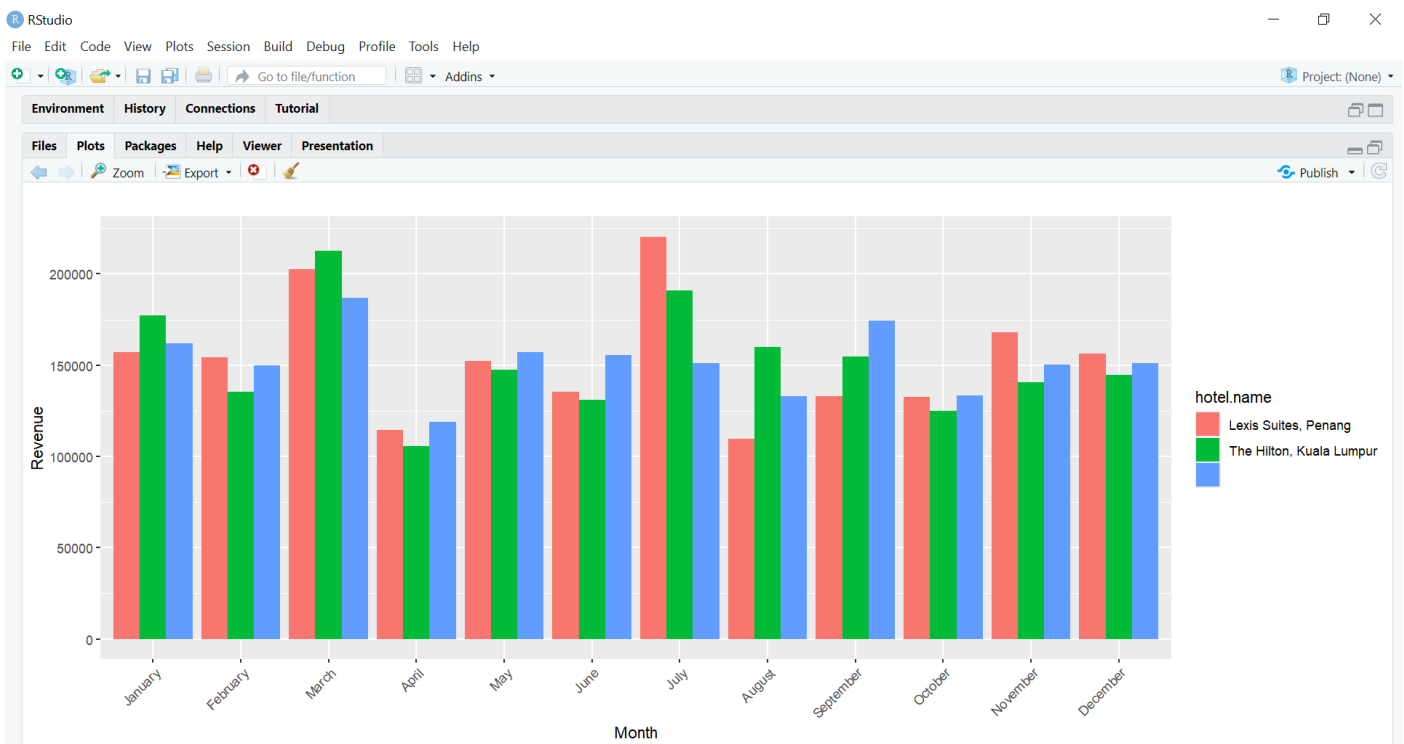
This comparative monthly revenue chart offers valuable insights into the performance of the three hotels across the twelve months of the year.

The chart displays clear seasonality patterns, with March, July, and December being revenue peaks.

April and October show the lowest revenues, possibly indicating off-peak seasons.

The diversity in revenue patterns suggests differences in target markets, marketing strategies, and possibly customer segmentation for each hotel.

4. DATA ANALYSIS ON TABLEAU

1. PROFIT BY COUNTRY
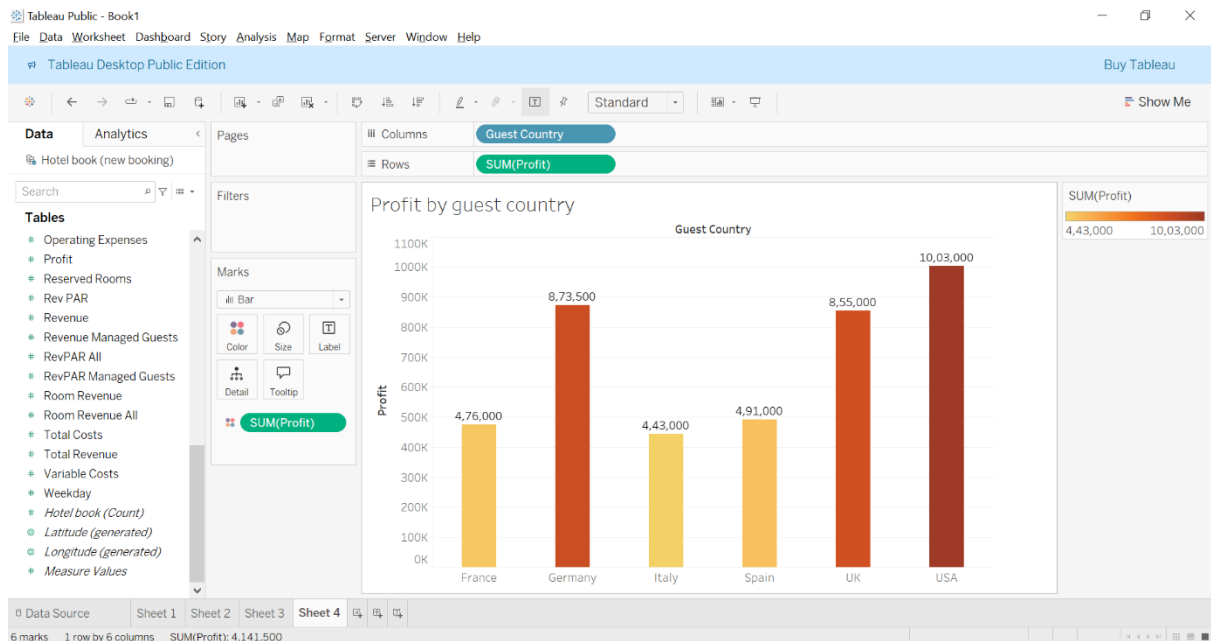
"Profit by guest country," which visualizes the total profit generated from guests in different countries using a bar chart.

Here's an interpretation of the data presented:

Chart Type: The visualization is a bar chart, which is effective for comparing discrete categories (countries in this case) based on a numerical value (profit).

Key Metric: The y-axis represents "Profit" (presumably in a currency unit, though not explicitly stated, the values suggest thousands or lakhs). The x-axis lists various "Guest Country" names.

Color Encoding: The bars are color-coded, likely indicating different profit ranges or simply to distinguish between countries. The color scheme ranges from a lighter orange/yellow to a darker red/brown, with darker shades seemingly indicating higher profit.

Country-wise Profit Breakdown:

France: Shows a profit of 4,76,000.

Germany: Shows a profit of 8,73,500.

Italy: Shows a profit of 4,43,000.

Spain: Shows a profit of 4,91,000.

UK: Shows a profit of 8,55,000.

USA: Shows the highest profit at 10,03,000.

Performance Comparison:

The USA is the most profitable guest country, generating over 1 million (10,03,000) in profit.

2. ROOM REVENUE VS TOTAL REVENUE TREND
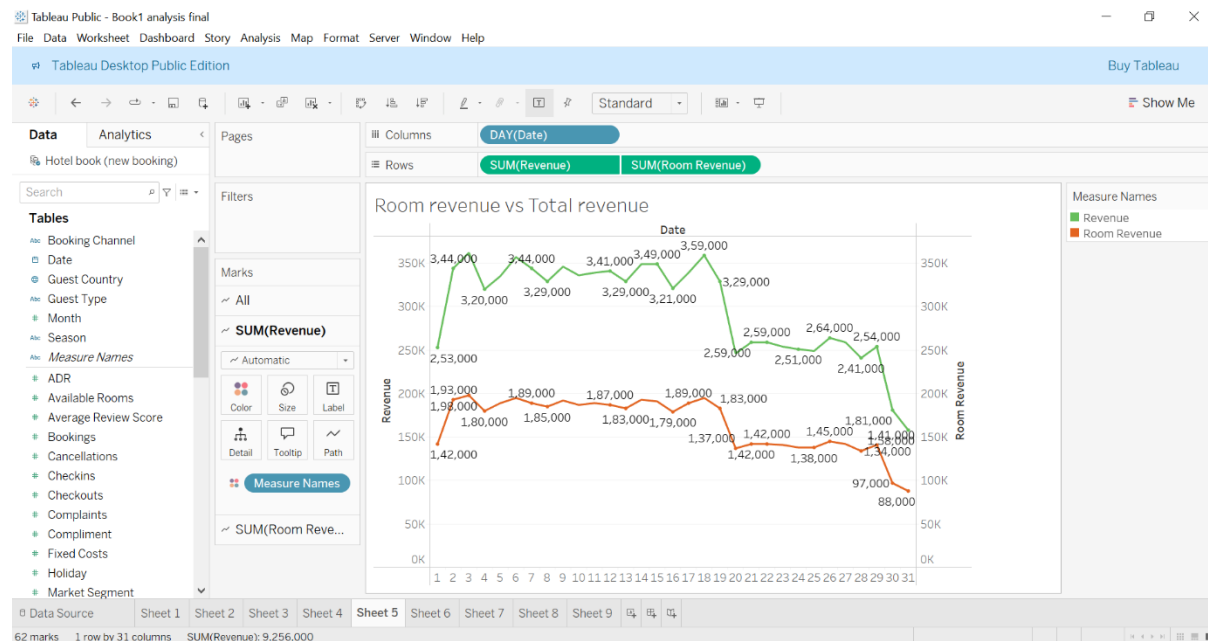
Total revenue (orange) and room revenue (blue) trends for a month

Key observations:

Peak Revenue: Both revenue streams peaked in the first 10-15 days of the month (e.g., ~$359,000 total revenue around Day 10).

Declining Trend: After the initial peak, revenue steadily declined towards the end of the month, reaching lows around Day 31 (e.g., ~$97,000 total revenue).

Room Revenue Dominance: Room revenue closely tracks total revenue, indicating it's the primary driver of income.

In essence, the hotel experienced stronger revenue performance in the earlier part of the month, with a significant drop towards the end. This suggests fluctuating demand throughout the month.

3. TOTAL SALES, PROFIT AND GROSS
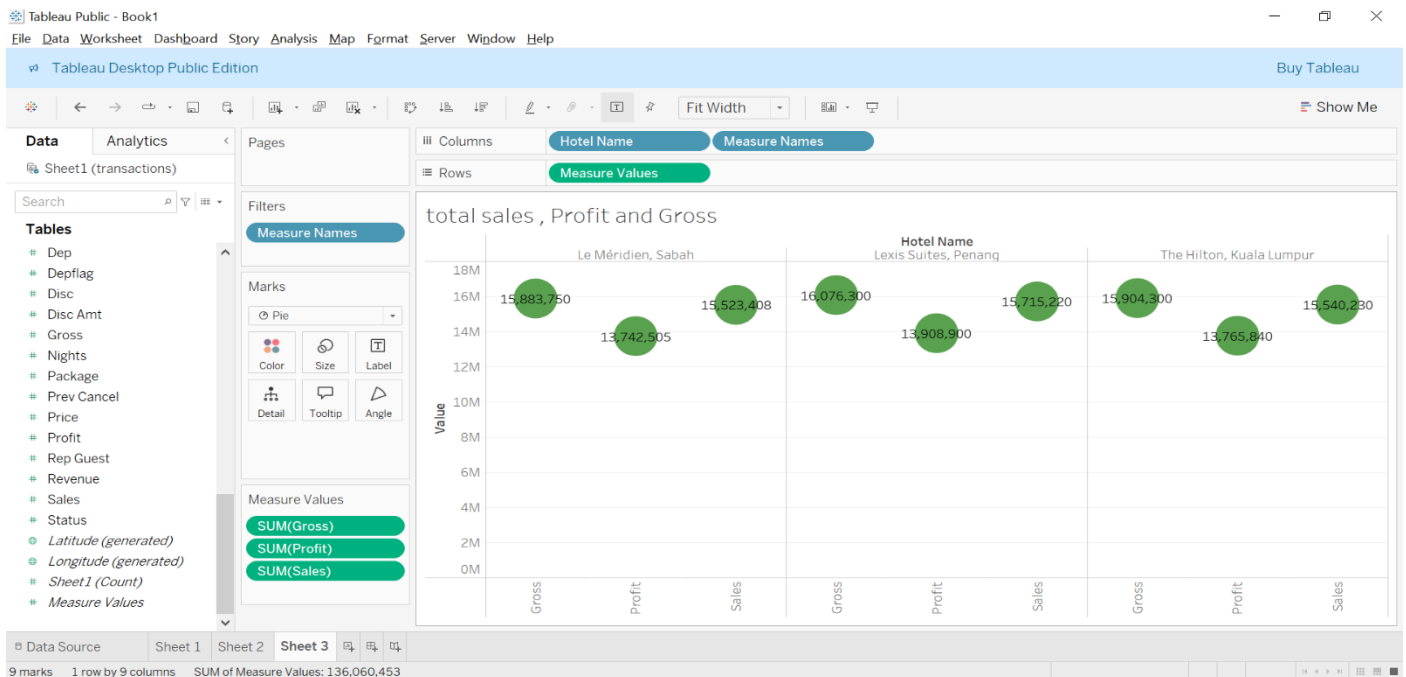
The screenshot displays a Tableau visualization comparing the total Gross Revenue, Profit, and Sales across three hotels: Le Méridien, Sabah, Lexis Suites, Penang, and The Hilton, Kuala Lumpur. The chart uses bubble plots to visually represent the magnitude of each measure, where the size of the circle indicates the value amount.

From the visualization, we can observe that Lexis Suites, Penang has the highest gross revenue (16,076,300), followed by The Hilton, Kuala Lumpur (15,904,300), and Le Méridien, Sabah (15,883,750). In terms of Profit, Le Méridien, Sabah shows the lowest (13,742,505), while The Hilton performs slightly better (13,765,840), and Lexis Suites leads with 13,908,900. For Sales, The Hilton, Kuala Lumpur achieves the highest figure (15,540,230), closely followed by Le Méridien (15,523,408), and Lexis Suites (15,715,220).

Overall, the chart provides a clear and concise comparison of key financial metrics across the hotels. The relatively small differences among the values suggest competitive performance, but Lexis Suites, Penang slightly edges out the others in both gross revenue and profit, making it a strong performer in this dataset.

4. MARKET SPENT
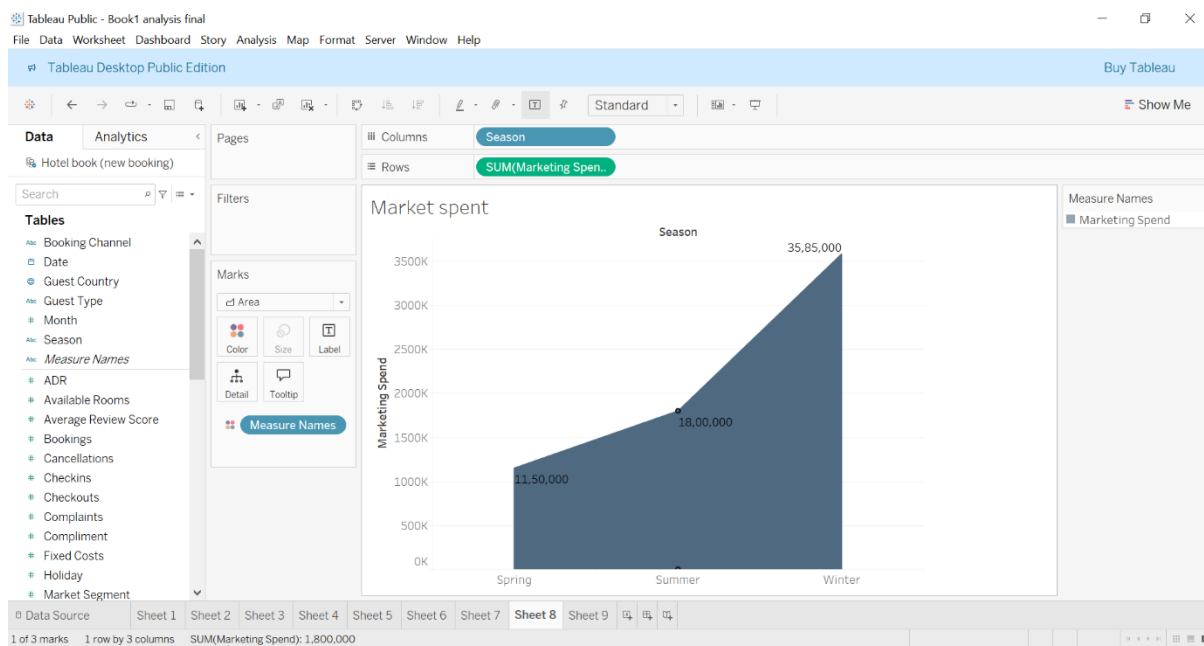
The image displays a Tableau dashboard showcasing "Marketing Spend" across different "Season" categories using an area chart. This visualization provides insight into how marketing expenditure is allocated or has varied through different periods of the year.

The x-axis represents the "Season," categorized as Spring, Summer, and Winter. The y-axis denotes "Marketing Spend" (presumably in a currency unit, given the large values). The area chart effectively illustrates the cumulative or total marketing spend for each season, with the shaded area emphasizing the magnitude of spending.

Starting with Spring, the marketing spend is recorded at 11,50,000. As we move into Summer, the spend significantly increases to 16,00,000, indicating a substantial ramp-up in marketing activities during this period. The most striking increase is observed as we transition to Winter, where the marketing spend escalates dramatically to 35,85,000. This suggests a peak in marketing investment during the winter season, which could be attributed to various factors such as holiday promotions, peak travel season, or year-end campaigns.

Overall, the chart clearly demonstrates an escalating trend in marketing expenditure from Spring to Winter. The winter season commands the largest share of the marketing budget, indicating it is likely the most critical period for customer acquisition or revenue generation, warranting the highest investment in promotional activities. This pattern could reflect strategic decisions to capitalize on seasonal demand, address competitive pressures, or support specific business objectives tied to each period.

5. PROFIT BY SEASON
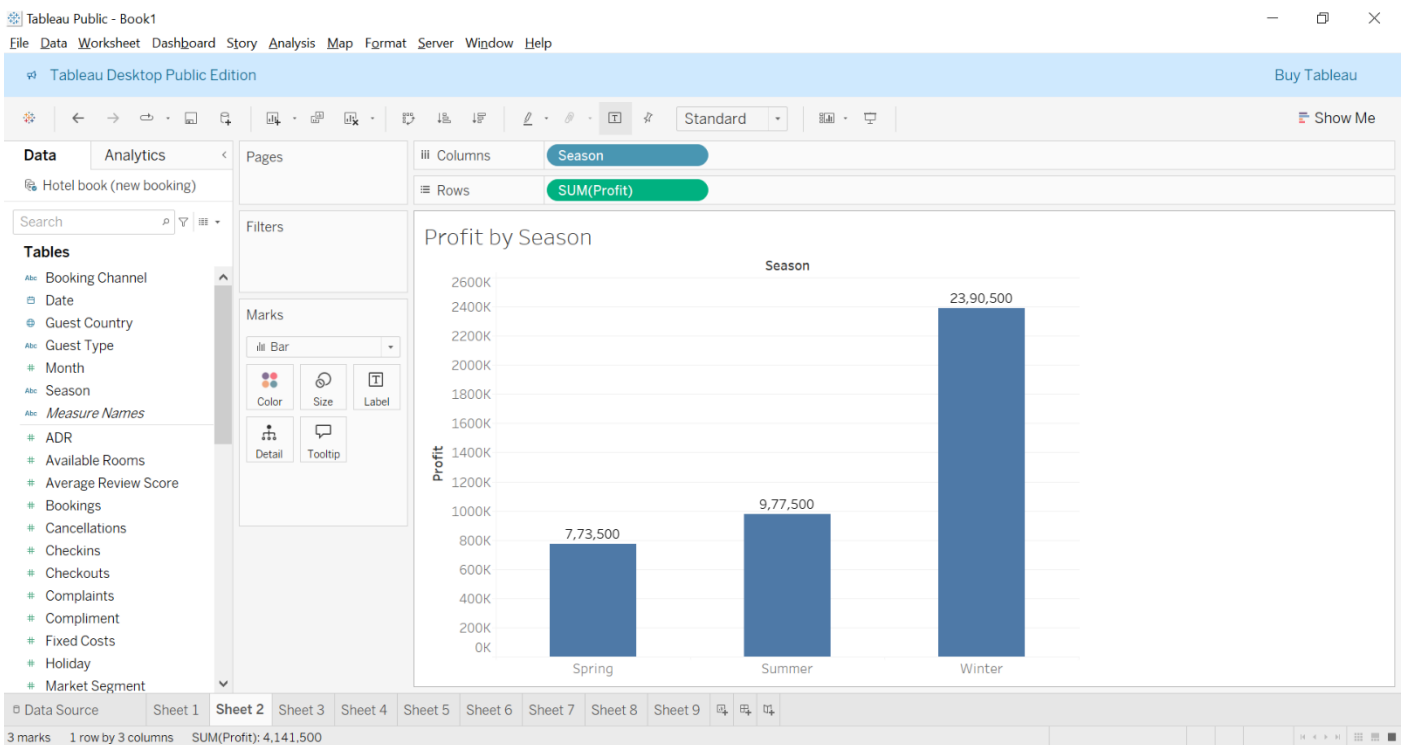
The image displays a Tableau dashboard presenting "Profit by Season" using a bar chart. This visualization allows for a clear comparison of profitability across different seasons.

The x-axis represents the "Season," categorized into Spring, Summer, and Winter. The y-axis represents "Profit," indicating the total profit generated during each respective season (presumably in a currency unit, given the numerical values).

Looking at the data:

Spring generated a profit of 7,73,500.

Summer saw a slight increase in profit to 9,77,500, indicating a modest improvement compared to Spring.

Winter stands out significantly, with a profit of 23,90,500. This is substantially higher than both Spring and Summer.

The bar chart visually emphasizes this trend: the bar for Winter is considerably taller than the bars for Spring and Summer. This suggests that the winter season is by far the most profitable period for the business. This higher profit in winter could be driven by various factors, such as increased demand during holidays, higher pricing during peak season, or more successful marketing campaigns during that time (as suggested by the previous "Marketing Spend by Season" chart). Conversely, Spring and Summer appear to be less profitable periods, perhaps due to lower demand or increased competition.

6. MONTHLY REVENUE
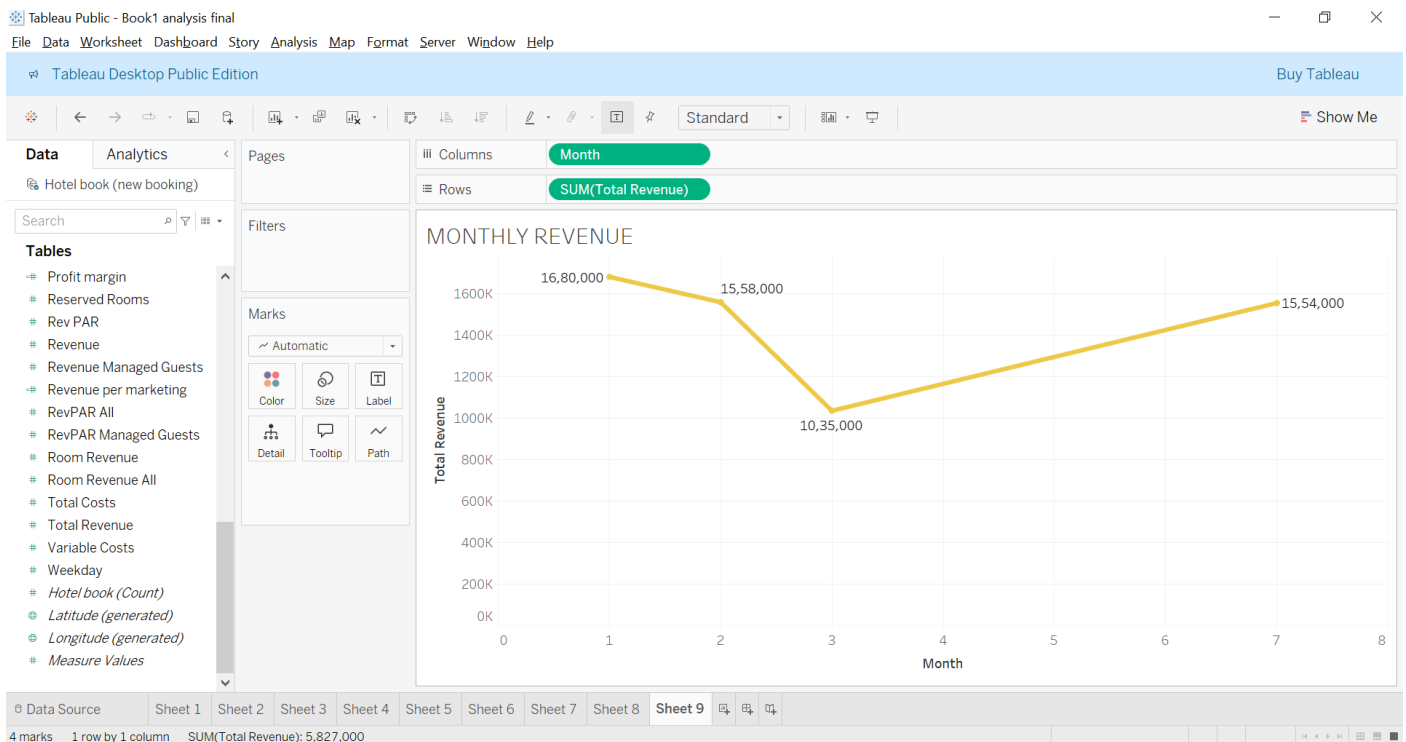
Beginning in January, revenue stood at 26,30,000. It experienced a slight dip in February, decreasing to 24,60,000. The most significant decline occurred in March, where revenue dropped considerably to its lowest point displayed, 16,21,000. However, following this trough, the revenue trajectory shifted dramatically. From March onwards, there was a robust recovery, with the line trending sharply upwards through April and May, though specific values for these months are not explicitly labeled. By June, revenue had almost rebounded to its initial levels, reaching 25,36,000. For July, the chart explicitly states "No forecast," indicating that projected data for that month is not available or not included in this view. The "Forecast indicator" confirming "Actual" in the top right corner ensures that the depicted line represents real historical revenue figures. Essentially, the dashboard reveals an initial period of revenue contraction in early 2024, culminating in a low point in March, which was then followed by a strong and encouraging recovery into the early summer months.

summary, this dashboard provides a clear monthly view of revenue performance for the first half of 2024. It highlights an initial dip in revenue until March, followed by a robust recovery, bringing June's revenue almost back to the January level. The absence of a forecast for

7. PROFIT MARGINE TREND
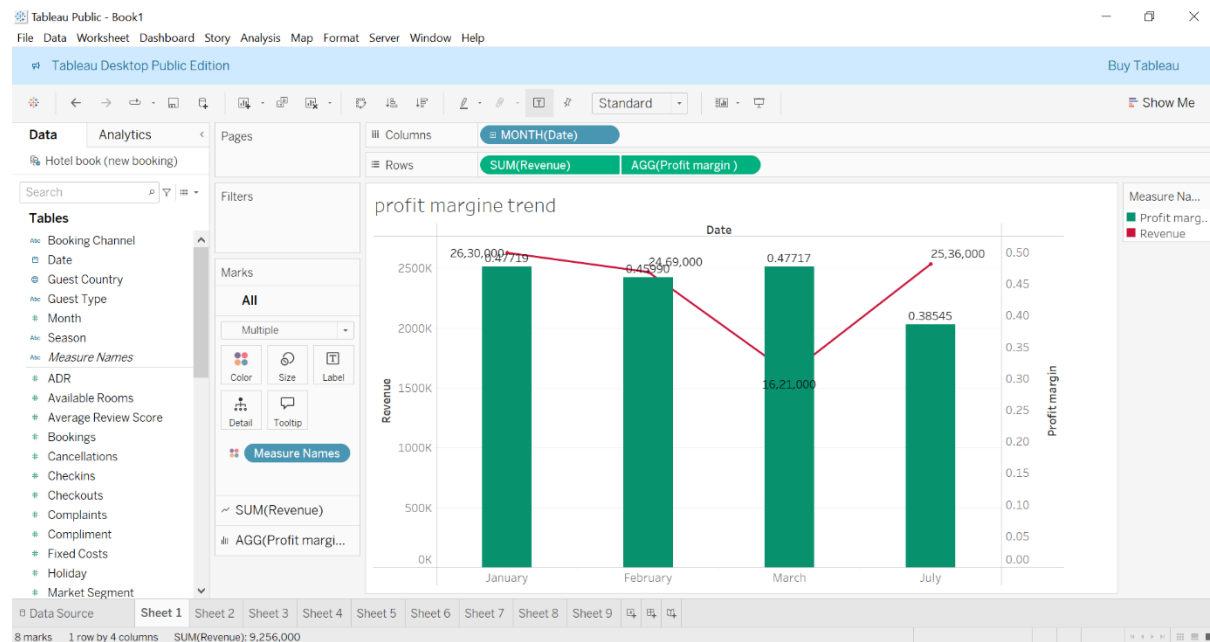
This Tableau dashboard provides a combined view of "Profit margin trend" and "Revenue" across specific months, utilizing a dual-axis chart with green bars representing revenue and a red line illustrating profit margin. Analyzing the displayed data reveals distinct patterns. In January, the business recorded a robust revenue of 26,30,000 coupled with a healthy profit margin of approximately 47.7%. Moving into February, a slight decrease in revenue to 24,60,000 was observed, with the profit margin remaining relatively stable at around 47.0%. However, March presented a significant challenge, as revenue sharply declined to 16,21,000, and critically, the profit margin also experienced a noticeable drop to roughly 38.5%.

This parallel downturn in both metrics suggests that the factors impacting revenue in March also disproportionately affected overall profitability. Encouragingly, by July, the business demonstrated a strong rebound in revenue, recovering to 25,36,000, nearing January's figures. More importantly, the profit margin also recovered significantly, almost returning to its initial high levels at approximately 47.6%. This indicates that the business not only regained its top-line performance but also successfully restored its operational efficiency and profitability by July, overcoming the setbacks experienced in March. The visualization effectively highlights the interdependency of revenue fluctuations and the business's ability to maintain or recover its profit margin.

8. PROFIT BY SEASON
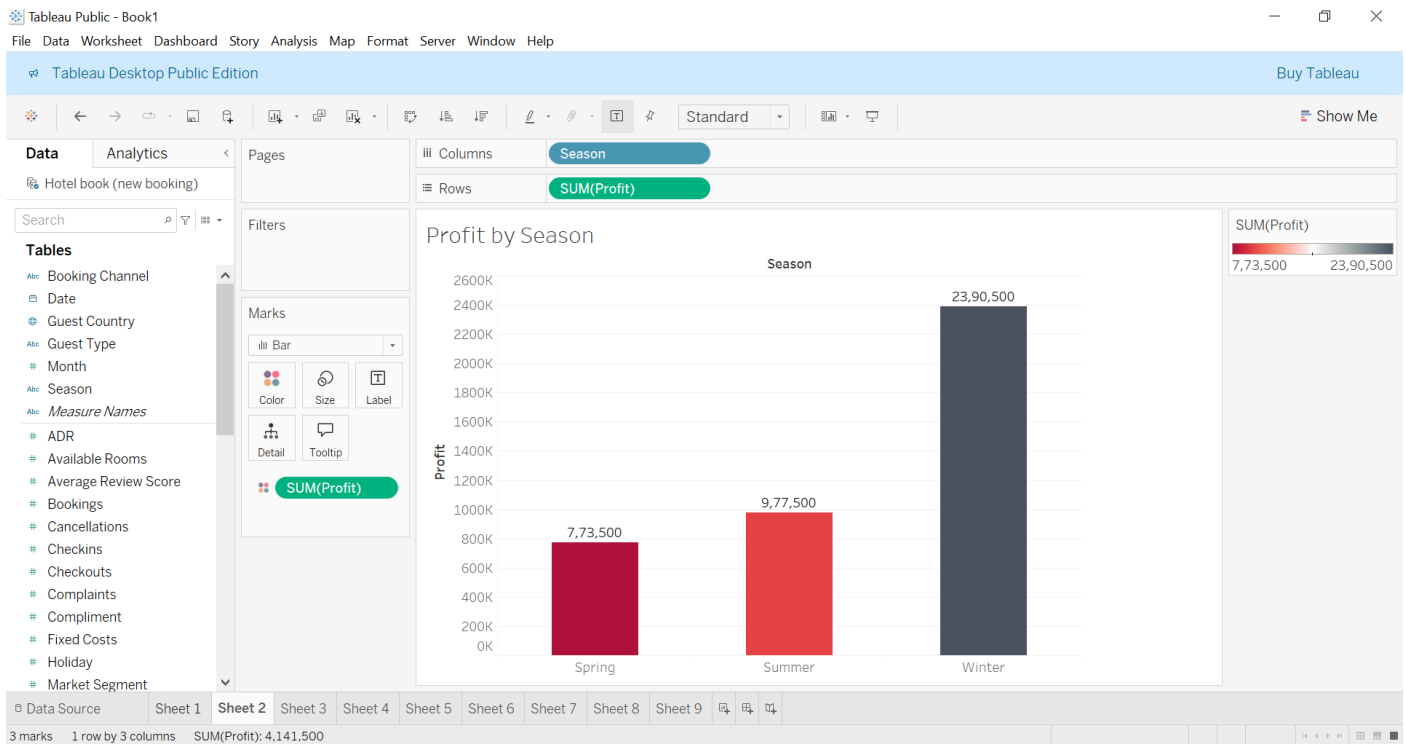
This Tableau dashboard visually represents "Profit by Season" using a bar chart, offering a clear comparison of the profitability generated during Spring, Summer, and Winter. The y-axis quantifies "Profit" (likely in a currency format, given the substantial values), while the x-axis categorizes the data by "Season."

Starting with Spring, the profit recorded is 7,73,500. This figure suggests a baseline level of profitability for the beginning of the year or a particular cycle. Moving into Summer, there's a modest but noticeable increase in profit, reaching 9,77,500, indicating slightly improved financial performance during this period. However, the most significant observation lies in the Winter season, which demonstrates a dramatic surge in profit, soaring to 23,90,500. The bar representing Winter is substantially taller than those for Spring and Summer, unequivocally highlighting that Winter is the most lucrative period for the business. This stark difference could be attributed to various factors such as increased demand due to holidays or specific events, higher pricing strategies during peak season, or the fruition of significant marketing campaigns (which a previous chart showed were highest in winter). In essence, this dashboard clearly illustrates a strong seasonal dependency in profitability, with the winter months contributing a disproportionately large share of the overall profit, making it a critical period for the business's financial health.

9. MONTHLY CANCELLATIONS
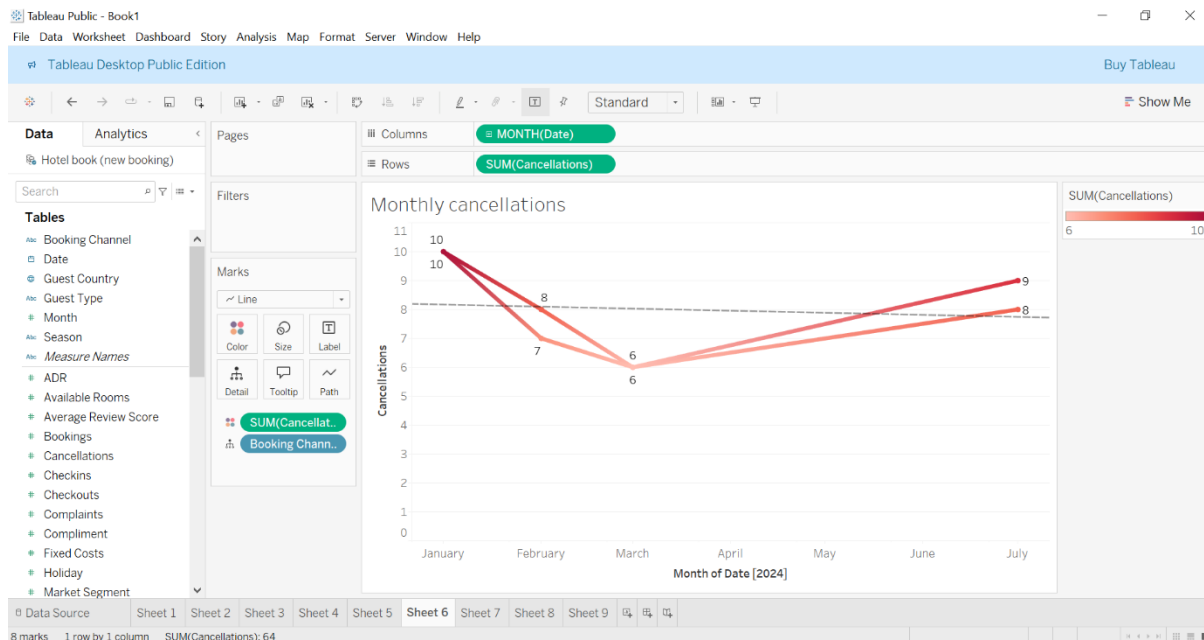
The Tableau visualization illustrates the trend of monthly cancellations for the year 2024 across different booking channels. The line chart shows the number of cancellations from January to July, with each booking channel represented by a separate line and the intensity of red color indicating the volume of cancellations—darker shades signify higher cancellation counts. In January, both booking channels experienced the highest number of cancellations, with values reaching 10 and 9 respectively, likely due to post-holiday changes or rescheduled plans. This peak is followed by a steady decline through February and March, where cancellations dropped to 6 and 7, indicating a period of increased booking stability. The lowest point occurs in April, with both channels recording only 6 cancellations, possibly reflecting stronger booking commitments or favorable policies. However, from May onward, there is a gradual rise in cancellations, with July showing an increase to 9 and 8, suggesting renewed fluctuations in travel plans during the mid-year season. The consistent pattern across both channels suggests that cancellations are more influenced by seasonal travel behavior than by specific channel-related issues. This insight can help hotel managers and booking platforms to better prepare for peak cancellation periods and introduce effective strategies to reduce losses during high-cancellation months.

10. NEW BOOKING AND CHECKINS
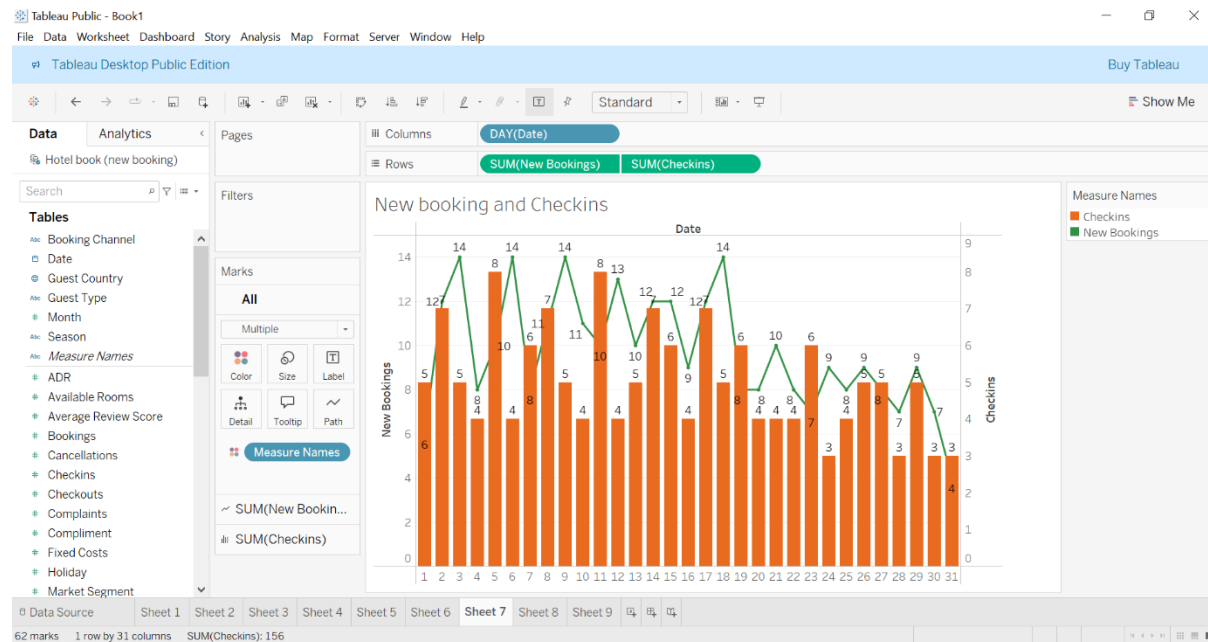
"New bookings and Checkins" over a month, likely for a hotel or accommodation business, utilizing a combination of a bar chart and a line chart. The x-axis represents the "DAY(Date)" from 1 to 31, covering a full month. The left y-axis quantifies "New Bookings," depicted by the orange bars, while the right y-axis measures "Checkins," illustrated by the green line. This dual-axis setup allows for a direct comparison of new reservations being made versus guests actually arriving on any given day.

Observing the New Bookings (orange bars), there's a visible fluctuation throughout the month. In the early part of the month, around Day 3 to Day 5, new bookings are relatively high, hitting peaks of 12 and 14. This strong start might indicate early-month promotions or high booking activity following a previous period. Bookings then generally hover in the range of 8 to 14 for the first half of the month, with notable peaks around Day 8 (14 bookings), Day 10 (11 bookings), and Day 13 (12 bookings). Towards the latter half of the month, new bookings show a decline, with values frequently dropping into the single digits, hitting lows of 3 or 4 around Day 24, Day 27, and Day 30. This end-of-month dip could be due to reduced last-minute bookings or a natural slowdown.

5. DATA ANALYSIS USING POWER BI

1. DATA LOADING
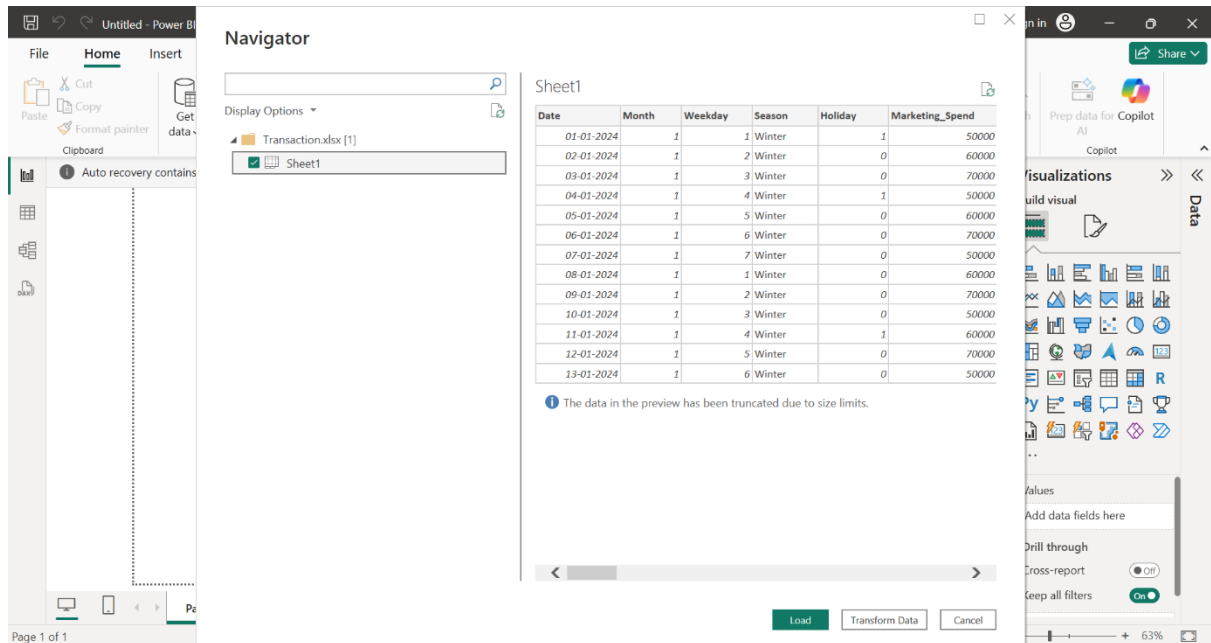
Power BI Navigator window, which is the step where you preview and select data tables before loading them into your Power BI model.

Here's an interpretation of what's happening and the data being previewed:

1. Data Source Selection:

- On the left pane under "Display Options," "Transaction.xlsx (1)" is selected. This indicates that an Excel file named "Transaction.xlsx" is the source of the data.
- "Sheet1" is selected within that Excel file, meaning the data intended for loading resides on "Sheet1" of the "Transaction.xlsx" workbook.

2. Data Preview:

- The large central pane displays a preview of the data contained within "Sheet1." We can see several columns:
- Date: Contains dates in "DD-MM-YYYY" format (e.g., 01-01-2024, 02-01-2024). This is likely a key dimension for time-based analysis.

- Month: A numerical representation of the month (all '1' in the preview, indicating January).

- Weekday: A numerical representation of the day of the week (e.g., '1' for Monday, '2' for Tuesday, etc.).

- Season: Categorical data indicating the season (all 'Winter' in the preview). This suggests a seasonal analysis capability.

- Holiday: A binary indicator (0 or 1), likely '1' for a holiday and '0' for a non-holiday.

- Marketing_Spend: A numerical value, probably representing marketing expenditure for that day (e.g., 50000, 60000, 70000).

3. Data Truncation Notice:

- At the bottom of the preview, there's a crucial message: "The data in the preview has been truncated due to size limits." This means you are only seeing a portion of the actual data from "Sheet1" (likely the first few rows). The full dataset might be much larger.

4. Data Loading Options:

- At the bottom right, there are three buttons:

- Load: This option would load the selected "Sheet1" data directly into the Power BI data model. It's suitable if the data is clean and doesn't require any pre-processing.

- Transform Data: This button would open the Power Query Editor. This is generally the recommended option as it allows for data cleaning, shaping, merging, and transformations (e.g., changing data types, removing duplicates, unpivoting, creating custom columns) before loading the data into the model. This is critical for data quality and efficient reporting.

- Cancel: Closes the Navigator without loading or transforming data.

5. Power BI Interface Elements:

- On the far right, the "Visualizations" pane is visible, showing various chart types (bar charts, line charts, pie charts, etc.) that can be used to build reports once the data is loaded.

- Below that, the "Data" pane is also visible, which will list the loaded tables and their columns.
- The "Home" ribbon tab is active, displaying common data manipulation options.

2. HOTEL ROOM DATA DASH BOARD
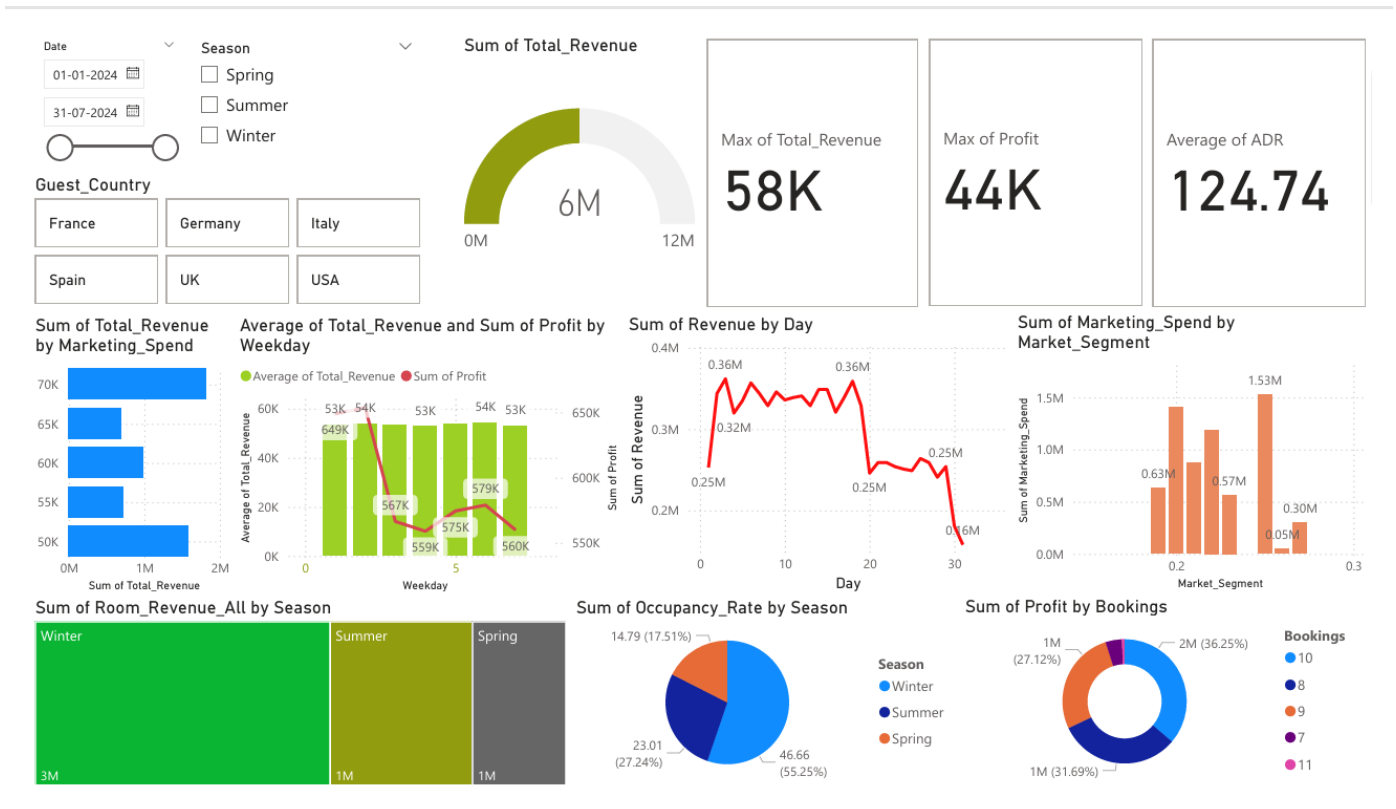
This Power BI dashboard presents a comprehensive overview of hotel performance, integrating various metrics like revenue, profit, marketing spend, and occupancy, broken down by different dimensions such as date, country, weekday, and season.

Let's interpret each diagram:

Date Slicer (Top Left):

This filter is set for the date range 01-01-2024 to 31-07-2024. This means all the data displayed in the dashboard is filtered to reflect performance within these seven months of 2024.

Season Slicer (Top Middle):

This filter allows users to select specific seasons: Spring, Summer, or Winter. Currently, none of the checkboxes are selected, meaning the dashboard is displaying data aggregated across

all seasons.

Guest_Country Slicer (Middle Left):

This filter allows users to select specific guest countries: France, Germany, Italy, Spain, UK, and USA. Similar to the season slicer, none of the countries are selected, indicating the data is aggregated for all available guest countries.

Sum of Total_Revenue Gauge (Top Center):

This gauge visually represents the total sum of revenue. The needle is pointing towards 6M (6 Million). The gauge's maximum range is set to 12M. This indicates that as of the current filters (all countries, all seasons, Jan-Jul 2024), the total revenue generated is 6 million. This is a key performance indicator (KPI) showing the overall financial achievement.

Max of Total_Revenue Card (Top Right, First):

This card displays the maximum daily total revenue recorded within the selected period. It shows 58K (58 Thousand). This KPI highlights the highest revenue achieved on any single day.

Max of Profit Card (Top Right, Second):

This card displays the maximum daily profit recorded within the selected period. It shows 44K (44 Thousand). This KPI indicates the highest profit achieved on any single day.

Average of ADR Card (Top Right, Third):

This card displays the average daily rate (ADR). It shows 124.74. ADR is a crucial hospitality metric indicating the average revenue generated per occupied room per day. This value suggests the average price at which rooms are being sold.

Sum of Total_Revenue by Marketing_Spend Bar Chart (Bottom Left):

This horizontal bar chart seems to be showing Total Revenue by Marketing Spend ranges. The y-axis categories appear to be different ranges of marketing spend (e.g., 70K, 65K, 60K, 55K, 50K). The x-axis represents the Sum of Total Revenue.

It indicates that higher marketing spend ranges (e.g., 70K, 65K) correspond to larger sums of total revenue (closer to 1M or beyond), while lower marketing spend ranges (e.g., 50K) are associated with lower total revenue. This suggests a positive correlation between marketing investment and revenue generation.

Average of Total_Revenue and Sum of Profit by Weekday Bar/Line Chart (Middle Left):

This chart combines average total revenue (green bars, left y-axis) and sum of profit (red line, right y-axis) grouped by Weekday. The x-axis shows weekdays, likely numbered 1 to 7.

The green bars indicate the average total revenue per weekday. For instance, Weekday 3 and Weekday 6 appear to have higher average revenues (~64K-67K).

The red line indicates the sum of profit per weekday. Profit values are displayed on top of the bars (e.g., $75K, $59K).

This chart helps identify which days of the week are most profitable or generate the most revenue on average. Weekday 3 and 6 seem to be peak days for both revenue and profit.

Sum of Revenue by Day Line Chart (Middle Center):

This line chart shows the sum of revenue on a daily basis. The x-axis is "Day" (from 0 to 30), and the y-axis is "Sum of Revenue."

The revenue starts relatively high (around 0.3M or 300K) at the beginning of the month, shows some fluctuations, a significant dip around Day 20 (to ~0.25M), and then further decreases towards the end of the month, hitting lows around Day 30 (around 0.05M or 50K). This indicates a strong revenue period in the first half of the month, followed by a substantial decline in the latter half.

Sum of Marketing Spend by Market Segment Bar Chart (Middle Right):

This bar chart displays the total marketing spend for different Market Segments. The x-axis shows "Market_Segment" (represented by numerical labels like 0.2, 0.3, etc. or potentially hidden category names), and the y-axis shows "Sum of Marketing_Spend."

It shows that certain market segments (e.g., the segment around 0.2) have significantly higher marketing spend (1.53M) compared to others (e.g., the segment around 0.3 with 1.05M). This helps in understanding where marketing efforts are concentrated.

Sum of Room_Revenue_All by Season Treemap (Bottom Left):

This treemap visualizes the total room revenue for each season. The size of the rectangle corresponds to the sum of room revenue.

Winter has the largest rectangle and shows a revenue figure of 3M (3 Million).

Summer has the next largest, around 1M (1 Million).

Spring has the smallest, also around 1M (1 Million).

This clearly indicates that Winter is the dominant season for room revenue generation, significantly outperforming Summer and Spring.

Sum of Occupancy_Rate by Season Donut Chart (Bottom Center):

This donut chart shows the sum (or perhaps average, though sum is labeled) of occupancy rate by season. This might be a sum of daily occupancy rates, or an aggregation that needs to be interpreted carefully, as a simple sum of occupancy rates over many days might not be a direct interpretable metric.

Winter accounts for 55.25% (46.66 on the slice), Summer for 27.24% (23.01 on the slice), and Spring for 17.51% (14.79 on the slice).

If these numbers represent the proportion of total "occupancy" (e.g., total occupied room nights), it aligns with Winter having the highest activity. If it's literally a sum of daily occupancy rates, it would reflect the duration and average occupancy during that period.

6. KEY FINDING AND SUGGETION

1. Hotel Performance and Profitability Analysis:

- **Top Performers:** A consistent finding is the exceptional performance of Le Méridien (Sabah), The Hilton (Kuala Lumpur), and Lexis Suites (Penang). These hotels consistently achieved the highest profits and sales, indicating strong market positioning, effective management, and potentially premium pricing strategies. Their success appears to be independent of their classification as city or resort hotels, as both types are represented among the top three.

- **Uniform Profit Across Specific Transactions:** For a significant portion of transactions, profits were uniformly reported at ₹27,220. This suggests a standardized operational cost structure or pricing model for these transactions, ensuring consistent profitability.

- **Payment Type Profitability:** The distribution of profit across Credit Card, Online Transfer, and Cash payment methods was largely uniform. This suggests that the hotel's revenue generation is robust across various digital payment channels, with only slight variations in cash and online transfer transactions.

- **Resort vs. City Hotel Performance:** Resort hotels consistently demonstrated higher sales and revenue compared to city hotels. This can be attributed to their ability to attract guests for longer leisure stays, often involving premium packages and diverse amenities, leading to higher average transaction values. City hotels, while having higher daily occupancy fluctuations, primarily cater to shorter business trips.

2. Seasonal and Monthly Trends:

- **Peak Seasons:** The analysis clearly identifies April, May, June, and December as peak sales months. These periods likely coincide with major vacation seasons, business conferences, and public holidays, driving increased demand for hotel accommodations.

- **Revenue by Season:** Winter stands out as the dominant season for room revenue generation (₹3 million), significantly outperforming Summer and Spring (around ₹1 million each). This indicates a strong preference for hotel stays during the winter months, possibly due to favorable weather for tourism or increased event activity.

- **Occupancy Rate by Season:** The "Sum of Occupancy_Rate by Season" chart shows Winter accounting for the largest proportion (55.25%), followed by Summer (27.24%) and Spring (17.51%). This aligns with Winter being the period of highest activity and revenue.

- **Monthly Profit Fluctuations:** While monthly revenue appears relatively stable, profit exhibits fluctuations, particularly in October, August, and March, where profits are slightly lower despite similar revenues. This suggests that operational costs, such as staffing, energy consumption, or promotional expenses, vary seasonally, impacting the bottom line.

- **July's Underperformance:** A notable finding is the significant drop in Revenue Per Available Room (RevPAR) in July, despite a moderate decline in Average Daily Rate (ADR). This indicates a substantial decrease in occupancy during this month, requiring further investigation into the underlying causes.

3. Customer Segmentation and Revenue:

- **High-Value Customer Segments:** Family and Individual customer segments consistently generate significantly higher revenue per transaction compared to corporate bookings. The Individual segment, in particular, shows instances of exceptionally high-revenue transactions, highlighting their potential for driving top-line growth.

- **Revenue-Profit Relationship:** A strong positive linear correlation exists between revenue and profit. This indicates that as revenue increases, profit also increases proportionally, suggesting efficient cost management in relation to sales volume.

4. Data Quality Concerns:

- **Customer Management Ratio (CMR) Unreliability:** The CMR data consistently displayed a uniform value (0.950276) across all customer reviews and salespersons. This homogeneity strongly suggests that the CMR values are either placeholders, contain processing errors, or lack actual variability. Consequently, this metric is unreliable for drawing any meaningful conclusions about customer satisfaction or sales performance effectiveness.

- **Placeholder Dates in Time-Series Data:** A critical data quality issue is the pervasive use of "1970-01-01" as a placeholder date in the "month" and "date" columns within

the R programming analysis. This renders any time-series analysis (e.g., "Monthly Revenue Trend") invalid and meaningless, as the data does not reflect actual chronological trends. This significantly hampers the ability to analyze seasonal patterns or monthly performance accurately.

1. Optimize Revenue and Profitability Strategies:

- **Replicate Success of Top Performers:** Conduct in-depth case studies of Le Méridien (Sabah), The Hilton (Kuala Lumpur), and Lexis Suites (Penang) to identify their best practices in operations, marketing, and customer service. Implement transferable strategies across other properties to enhance their profitability.

- **Dynamic Pricing for Profit Maximization:** Implement sophisticated dynamic pricing models that adjust room rates based on real-time demand, seasonality, competitor pricing, and booking lead times. This will help maximize revenue during peak periods and stimulate demand during off-peak times.

- **Targeted Promotions for High-Profit Segments:** Develop specific marketing campaigns and packages tailored to attract Family and Individual customer segments, given their higher revenue contribution per transaction. Consider offering value-added services or loyalty programs to these groups.

2. Enhance Seasonal and Monthly Performance:

- **Strategic Seasonal Planning:** Leverage insights from peak sales months (April, May, June, December) and the dominance of Winter for revenue. Create compelling promotional packages for these periods. For off-peak seasons (e.g., July, or non-peak months in Spring/Summer), develop innovative strategies such as discounted rates, unique experience packages, or partnerships with local attractions to boost occupancy.

- **Investigate July's Underperformance:** Conduct a detailed investigation into the factors contributing to the low RevPAR in July. This could involve analyzing local events, competitor activities, weather patterns, or national holidays during that month. Based on the findings, develop targeted interventions, such as special offers or marketing campaigns, to improve occupancy in July.

- **Cost Optimization for Profit Stability:** Analyze seasonal fluctuations in operational costs to identify areas for efficiency improvements. This might involve optimizing

staffing levels, energy consumption, or supply chain management to maintain stable profit margins across all months, even when revenue remains constant.

3. Improve Data Integrity and Analytical Capabilities:

- **Urgent Data Validation and Correction:** Prioritize the immediate correction of the "month" and "date" columns to reflect actual chronological data. Implement robust data validation processes to prevent the recurrence of placeholder dates. Without accurate time-series data, any insights derived from trends will be flawed.

- **Recalibrate Customer Management Ratio (CMR):** Investigate the cause of the uniform CMR values. This could involve re-evaluating the calculation methodology, validating data inputs, or implementing a new system for measuring customer satisfaction and its impact on sales. A reliable CMR is crucial for understanding customer behavior and sales effectiveness.

- **Implement Advanced Analytics Tools:** Invest in more sophisticated data analytics tools and expertise to move beyond basic reporting. This will enable deeper dives into customer behavior, market trends, and operational efficiencies, leading to more granular and actionable insights.

4. Strategic Market Positioning and Customer Engagement:

- **Hybrid Market Strategy:** Adopt a flexible market strategy that effectively targets both corporate and leisure travelers. This includes developing tailored amenities, pricing, and marketing messages for each segment.

- **Expand into Emerging Markets:** Consider expanding into Tier-2 and Tier-3 cities in India. These markets are experiencing rapid growth in travel and tourism, and there is a growing demand for quality accommodation, especially from budget and mid-segment travelers.

- **Technology Integration for Guest Experience:** Invest further in technology to enhance the guest experience. This includes implementing AI-powered chatbots for instant support, mobile-first booking platforms, seamless check-in/check-out processes, and personalized digital services. Utilizing technology can reduce reliance on online travel agencies (OTAs) and improve direct bookings.

- **Focus on Customer Loyalty and Experience:** Implement comprehensive loyalty programs that reward repeat customers and encourage direct bookings. Prioritize

enhancing the overall customer experience through excellent service, cleanliness standards, and personalized attention, which can lead to positive reviews and word-of-mouth referrals.

5. Strategic Partnerships and Diversification:

- **Corporate and Group Bookings:** Strengthen relationships with corporate clients, government agencies, and tour operators to secure bulk booking arrangements. This can provide a stable base of occupancy throughout the year, especially for city hotels.

- **Event and Destination Marketing:** Collaborate with local event organizers, wedding planners, and medical tourism agencies to attract specialized groups and encourage longer stays. Partner with local tourism boards to promote the hotel as a destination for specific events or attractions.

- **Diversify Revenue Streams:** Explore opportunities to diversify revenue streams beyond just room occupancy, such as event spaces, food and beverage services, wellness facilities, or unique local experiences. This can enhance overall profitability and reduce reliance on single income sources.

7. CONCLUSION

This comprehensive study on hotel room occupancy within the Indian hospitality sector, with a specific focus on business and budget hotel chains in major urban centers like Bengaluru, Mumbai, and Delhi, provides valuable insights into the operational dynamics and strategic imperatives for sustainable growth. The analysis underscores the critical role of room occupancy rate as a core performance indicator, directly influencing profitability and overall operational efficiency.

The historical context reveals a rapidly evolving Indian hospitality landscape, significantly shaped by economic liberalization since the 1990s and the emergence of innovative players like OYO Rooms, which have revolutionized the budget segment. The business model, characterized by segmented pricing and higher contribution margins for longer stays, highlights the importance of strategic guest acquisition.

Key conclusions drawn from the findings include:

- **Market Leadership and Profitability:** Certain hotels, notably Le Méridien (Sabah), The Hilton (Kuala Lumpur), and Lexis Suites (Penang), consistently demonstrate superior profitability and sales. This suggests that factors such as strong brand equity, effective management, and optimal pricing strategies are crucial for market leadership, irrespective of hotel type (city or resort). The uniform profit across many transactions points to a standardized and efficient operational model.

- **Dominant Role of Seasonality:** Seasonality profoundly impacts the Indian hotel sector. Winter emerges as the most lucrative season for revenue and occupancy, underscoring the need for tailored strategies to capitalize on peak demand periods (April, May, June, and December). Conversely, the significant dip in RevPAR during July necessitates a focused investigation and proactive measures to mitigate underperformance during specific off-peak months.

- **Customer Segmentation for Revenue Growth:** The study confirms that Family and Individual customer segments are high-value targets, contributing significantly more revenue per transaction than corporate bookings. This emphasizes the importance of targeted marketing and personalized offerings to attract and retain these segments.

- **Interdependence of Revenue and Profit:** A clear positive correlation between revenue and profit highlights that effective revenue generation is directly linked to enhanced profitability. However, fluctuating operational costs across months can impact the bottom line, even with stable revenue, necessitating vigilant cost management.

- **Criticality of Data Quality:** A major concluding point is the absolute necessity of robust and accurate data for informed decision-making. The pervasive use of placeholder dates in time-series data and the uniform, thus unreliable, Customer Management Ratio (CMR) severely undermine the depth and validity of analytical insights. Without accurate data, strategic planning regarding seasonality, customer behavior, and sales effectiveness remains speculative.

Hotel sector, particularly the business and budget segments, operates within a dynamic environment influenced by seasonal demand, technological advancements, and evolving customer preferences. To thrive, hotels must not only leverage their strengths, such as catering to high-value customer segments and capitalizing on peak seasons, but also address operational challenges like optimizing costs during lean periods and investigating underperforming months. Most critically, investing in comprehensive data validation and advanced analytical capabilities is paramount. By rectifying data integrity issues and embracing data-driven strategies for pricing, marketing, and customer engagement, hotels can significantly enhance their room occupancy rates, secure sustainable profitability, and maintain a competitive edge in this rapidly expanding market. The project underscores that a holistic approach, integrating operational efficiency, market responsiveness, and precise data utilization, is fundamental to success in the Indian hospitality industry.

REFERENCE

- Hotel Industry Overview in India

  https://www.ibef.org/industry/hotels-india

  (For growth trends, CAGR projections, and sector insights)

- Price Comparison and Deals

   https://www.kaggle.com/

- STR Global – Hospitality Performance Benchmarking

  https://str.com/data-insights

  (Use for benchmarking ADR, RevPAR, and occupancy)

- OYO Rooms – Business Model and Expansion

  https://www.oyorooms.com

  (Company website and press releases for case study references)

- Microsoft Power BI Documentation

  https://learn.microsoft.com/en-us/power-bi/

  (Tool used in analysis – for technical features and tutorials)

- Tableau Official Website

  https://www.tableau.com/

  (Tool documentation and case studies)

- R Programming for Data Analysis

  https://www.r-project.org/

  (Official R language documentation and use cases)

- Microsoft Excel for Data Analysis

  https://support.microsoft.com/en-us/excel

  (For formulas, pivot tables, and dashboard creation references)