



ITCS 453: Data Warehousing and Data Mining
Project 1 – Data Warehousing for Business Intelligence

Business Case Study: Agoda

(February 2025)

Prepared by

6588070 Nakarin Phoorahong

6588071 Thareerat Phothitham

6588087 Yaowapa Sabkasedkid

6588096 Panipak Sittiprasert

6588183 Achiraya Mankham

Presented To

Asst. Prof. Dr. Songsri Tangsripairoj

Faculty of Information and Communication Technology, Mahidol
University

Executive summary

This report is part of the Data Warehousing for Business Intelligence project for ITCS453: Data Warehousing and Data Mining. The report provides a comprehensive overview of the project, covering key topics such as project objectives, data sources, data warehouse design, the ETL (Extract, Transform, Load) process, data analysis, visualization, and conclusions. For this study, we have focused on Agoda, a leading online travel agency. The report explores the processes Agoda uses to collect and transform data into actionable insights, including the implementation of OLTP (Online Transaction Processing) and ETL processes. The project also uses a smaller dataset to demonstrate the application of these methodologies in a real-world context.

Table of contents

Project Overview	1
Business Case Study Overview (Agoda, n.d.) (Finspace, 2023)	1
Business Problems and Requirements for Building a Data Warehouse	1
Objectives and Scope of the Project	1
Importance of Data Warehouse and BI to Agoda	1
Data Sources	2
Overview of Operational Database (OLTP)	2
Data Dictionaries for Operational Database Tables	4
Explanation of Data Sources Used for OLTP Database	10
Data Warehouse Design	13
Overview of Data Warehouse Design	13
Star Schema Design for Agoda's Data Warehouse	13
ETL Process	17
Overview of the ETL (Extract, Transform, Load) Processes	17
Mapping Between OLTP and OLAP (Operational and Analytical Systems) and Tableau Prep	
Screenshots and Step-by-Step ETL Integration	17
Analysis and Visualization Reports	27
Overview of Tableau Desktop for Data Visualization	27
Report 1: Quarterly Revenue Report	28
Report 5: Average Amount and Booking Trends	34
Report 6: Hotel Performance Analysis	35
Discussion and Conclusion	37
VDO Clip demonstration	39
References	40

Project Overview

Business Case Study Overview (Agoda, n.d.) (Finspace, 2023)

Agoda is a digital travel platform that helps people find deals on hotel, flight, activities and many other things, with a global network of 4.5 Million hotels and holiday properties around the world to offer to customers. Agoda, both agoda website and applications is available in 39 languages and has customer support for 24/7, headquartered in Singapore. Agoda is part of Booking Holding, the well-known company in the Online Travel Agency industry that acquired Agoda in 2007. Typically, Travel Agencies have one important aspect is pricing, with the proper instruments and correct decision-making can create a huge difference in sales and profit when setting price for hotels and flights.

Business Problems and Requirements for Building a Data Warehouse

The need for data-driven strategy; personalization, dynamic pricing, and easier method access to queries to improve operational efficiency, and improve the outcome of decision-making processes. These will ensure the growth of revenue and the better outcome of decision making.

Example of Dynamic pricing and Demand Forecasting.

Dynamic price setting on hotel reservations and flight tickets based on seasonal, weather, holiday, competitor's price, and historical data with limited time inclusion such as 45 days or 90 days. Dynamic pricing can increase the sales which indirectly increase profits. Moreover, Demand Forecasting based on similar information can help gauge the dynamic pricing policy even better.

Objectives and Scope of the Project

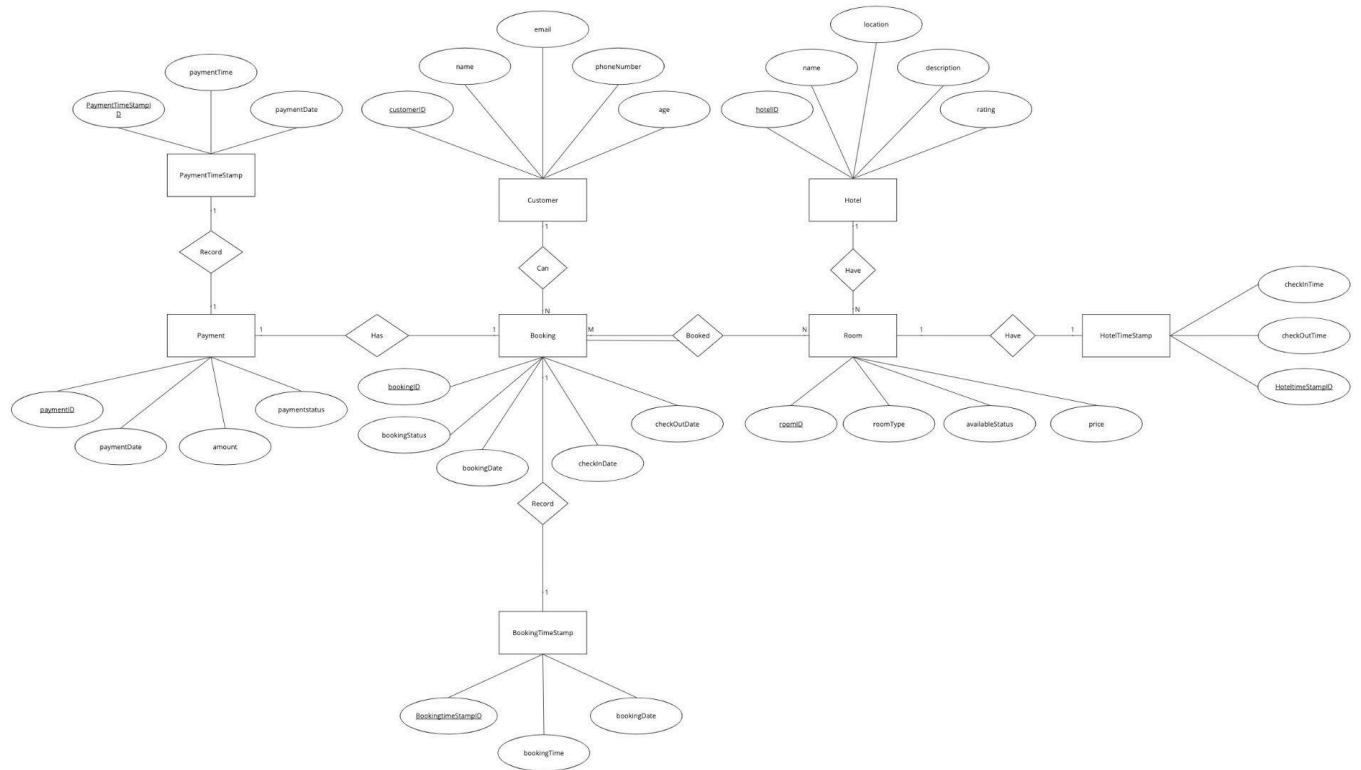
In this project we are focusing on only aspects of dynamic pricing, we focus on showing the difference that dynamic pricing can contribute to the rise of sales and profit, and represent the information that we can gather both from the internet that is already available publicly.

Importance of Data Warehouse and BI to Agoda

It is a crucial part of modern business processes in order to gain insightful knowledge, which leads to better decision-making in the end.

Data Sources

Overview of Operational Database (OLTP)



Detailed Entity-Relationship (ER) diagrams for Agoda's booking system are not publicly available, as they are proprietary designs. However, to understand the general structure of hotel reservation systems, resources such as the Hotel Booking System ER Diagram from academic project websites like FreeProjectz.com can provide useful insights [3]. These references offer a foundational understanding of booking systems similar to those used by companies like Agoda.

Customer-Booking Relationship (Make a Booking):



A "Make a Booking" relationship represents a transaction where a customer reserves a room, forming a one-to-many relationship where each customer can create multiple bookings, but each booking is linked to only one customer. This relationship is essential for managing booking history, tracking reservations, and enhancing personalized services such as tailored promotions, loyalty rewards, and customer support. Additionally, the system incorporates Booking Timestamp Data to record key time-related details for each booking. With a one-to-one relationship between Booking and Booking Timestamp Data, this structure ensures accurate event tracking, supports audit logs, and improves the management of payment processing and booking history.

Hotel-Room Relationship (Own):



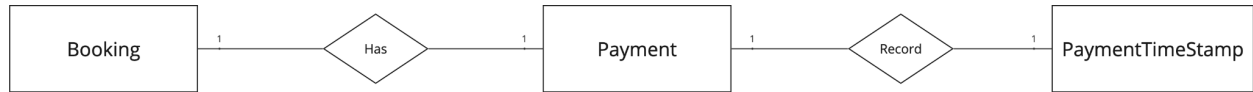
An "Own" relationship connects a hotel to the rooms it offers, forming a one-to-many relationship where each hotel can own multiple rooms, but each room belongs to a single hotel. This relationship is essential for managing room inventory, ensuring accurate linkage between rooms and their respective hotels, and providing users with reliable details such as pricing, availability, and amenities. It also enhances the search experience by enabling better filtering options. Additionally, the system includes Hotel Timestamp Data, which has a one-to-one relationship with each hotel. By maintaining precise timestamps, the system improves reservation management, enhances operational efficiency, and ensures a seamless booking experience.

Room-Booking Relationship (Book):

A "Book" relationship links rooms to bookings, representing the reservation of a room. This is a many-to-many relationship because a single room can be booked multiple times by different users on various dates, and a single booking can include multiple rooms (e.g., for group

stays). This relationship is essential for managing room availability, preventing double bookings, and providing accurate reservation data to both the users and hotel staff. It ensures the seamless operation of the reservation process.

Booking-Payment Relationship (Pay For):



A "Pay For" relationship represents the financial transaction associated with a booking, forming a one-to-one relationship where each booking corresponds to a single payment, and each payment is linked to one booking. This ensures that all reservations are accurately billed and settled, supporting the financial integrity of the system. It plays a crucial role in tracking transactions, managing refunds, and generating financial reports. Additionally, the system includes Payment Timestamp Data, which maintains a one-to-one relationship with each payment. By capturing precise payment timestamps, the system enhances financial accuracy, improves transaction monitoring, and ensures seamless payment processing.

Data Dictionaries for Operational Database Tables

1. Customer-Booking Relationship (Make a Booking)

- Customer Data

Field	Data Type	Description
customerID	INT(5)	Unique identifier for each customer
name	VARCHAR(50)	Customer name
email	VARCHAR(50)	Customer email address
phoneNumber	VARCHAR(50)	Customer phone number
age	INT(2)	Customer age

- Booking Data

Field	Data Type	Description
bookingID	INT(5)	Unique identifier for each booking
bookingDate	DATE	Date when the booking was created
checkinDate	DATE	Date of check-in
checkoutDate	DATE	Date of check-out
bookingStatus	ENUM	Status of the booking
customerID	INT(5)	Unique identifier for each customer
roomID	INT(5)	Unique identifier for each room

- Booking Timestamp Data

Field	Data Type	Description
bookingTimeStampID	INT(5)	Unique identifier for each booking timestamp
bookingTime	TIME	The specific time when a booking was made
bookingDate	DATE	The date on which a payment was made
bookingID	INT(5)	Unique identifier for each

		booking
--	--	---------

2. Hotel-Room Relationship (Own)

- Hotel Data

Field	Data Type	Description
hotelID	INT(5)	Unique identifier for hotel
name	VARCHAR(50)	Name of the hotel
location	VARCHAR(50)	Location of the hotel
description	TEXT	Description of hotel
rating	DECIMAL	Hotel rating

- Room Data

Field	Data Type	Description
roomID	INT(5)	Unique identifier for each room
roomType	ENUM	Type of room
price	DECIMAL	Room price (Dollar unit)
availableStatus	ENUM	Availability room
hotelID	INT(5)	Unique identifier for hotel

- Hotel Timestamp Data

Field	Data Type	Description
-------	-----------	-------------

HotelTimeStampID	INT(5)	Unique identifier for each hotel timestamp.
checkInTime	DATETIME	Records the date and time when a customer is check-in at the hotel.
checkOutTime	DATETIME	Records the date and time when a customer is check-out at the hotel.
roomID	INT(5)	Unique identifier for each room

3. Room-Booking Relationship (Book)

- Room Data

Field	Data Type	Description
roomID	INT(5)	Unique identifier for each room
roomType	ENUM	Type of room
price	DECIMAL	Room price (Dollar unit)
availableStatus	ENUM	Availability room
hotelID	INT(5)	Unique identifier for hotel

- Booking Data

Field	Data Type	Description
bookingID	INT(5)	Unique identifier for each

		booking
bookingDate	DATE	Date when the booking was created
checkinDate	DATE	Date of check-in
checkoutDate	DATE	Date of check-out
bookingStatus	ENUM	Current status of the booking
customerID	INT(5)	Unique identifier for each customer
roomID	INT(5)	Unique identifier for hotel

4. Booking-Payment Relationship (Pay For)

- Payment Timestamp Data

Field	Data Type	Description
PaymentTimeStampID	INT(5)	Unique identifier for each timestamp record
paymentTime	TIME	The time when a payment transaction occurs
paymentDate	DATE	The date when a payment transaction was processed
paymentID	INT(5)	Unique identifier for each payment transaction

- Booking Data

Field	Data Type	Description
bookingID	INT(5)	Unique identifier for booking
bookingDate	DATE	Date when the booking was created
checkinDate	DATE	Date of check-in
checkoutDate	DATE	Date of check-out
bookingStatus	ENUM	Status of the booking
roomID	INT(5)	Unique identifier for each room
customerID	INT(5)	Unique identifier for each customer

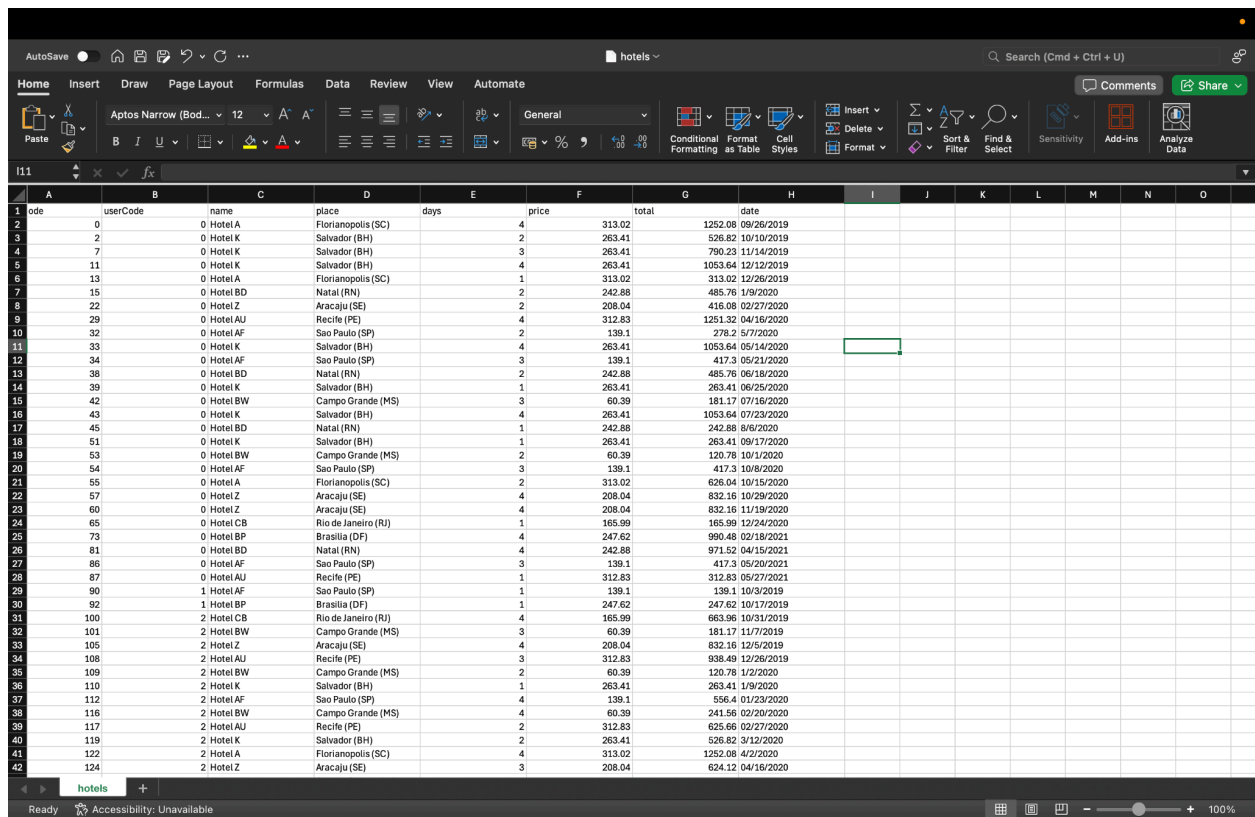
- Payment Data

Field	Data Type	Description
paymentID	INT(5)	Unique identifier for each payment
paymentDate	DATE	Date of the payment
amount	DECIMAL	Amount paid for the booking
paymentstatus	ENUM	Payment method used by the customer
bookingID	INT(5)	Unique identifier for booking

Explanation of Data Sources Used for OLTP Database

Due to privacy regulations and business confidentiality, real booking data from companies like Agoda is not accessible. This research utilizes a simulated dataset from Kaggle that represents typical patterns seen in online travel bookings. The dataset has been modified from its original Kaggle form to better suit the specific requirements of this study, while maintaining realistic patterns of booking platform behavior. No real customer or company information is used in this process.

1. Customer-Booking Relationship (Make a Booking)



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	code	userCode	name	place	days	price	total	date							
2	0	0	Hotel A	Florianopolis (SC)	4	313.02	1252.08	09/26/2019							
3	2	0	Hotel K	Salvador (BH)	2	263.41	526.82	10/10/2019							
4	7	0	Hotel K	Salvador (BH)	3	263.41	790.23	11/14/2019							
5	11	0	Hotel K	Salvador (BH)	4	263.41	1053.64	12/12/2019							
6	13	0	Hotel A	Florianopolis (SC)	1	313.02	313.02	12/26/2019							
7	15	0	Hotel BD	Natal (RN)	2	242.88	485.76	1/9/2020							
8	22	0	Hotel Z	Aracaju (SE)	2	208.04	416.08	02/27/2020							
9	29	0	Hotel AU	Recife (PE)	4	312.83	1251.32	04/16/2020							
10	32	0	Hotel AF	Sao Paulo (SP)	2	139.1	278.2	5/7/2020							
11	33	0	Hotel K	Salvador (BH)	4	263.41	1053.64	05/14/2020							
12	34	0	Hotel AF	Sao Paulo (SP)	3	139.1	417.3	05/21/2020							
13	38	0	Hotel BD	Natal (RN)	2	242.88	485.76	06/18/2020							
14	39	0	Hotel K	Salvador (BH)	1	263.41	263.41	06/25/2020							
15	42	0	Hotel BW	Campo Grande (MS)	3	60.39	181.17	07/16/2020							
16	43	0	Hotel K	Salvador (BH)	4	263.41	1053.64	07/23/2020							
17	45	0	Hotel BD	Natal (RN)	1	242.88	242.88	8/6/2020							
18	51	0	Hotel K	Salvador (BH)	1	263.41	263.41	09/17/2020							
19	53	0	Hotel BW	Campo Grande (MS)	2	60.39	120.78	10/1/2020							
20	54	0	Hotel AF	Sao Paulo (SP)	3	139.1	417.3	10/8/2020							
21	55	0	Hotel A	Florianopolis (SC)	2	313.02	626.04	10/15/2020							
22	57	0	Hotel Z	Aracaju (SE)	4	208.04	832.16	10/29/2020							
23	60	0	Hotel Z	Aracaju (SE)	4	208.04	832.16	11/19/2020							
24	65	0	Hotel CB	Rio de Janeiro (RJ)	1	165.99	165.99	12/24/2020							
25	73	0	Hotel BP	Brasilia (DF)	4	247.62	990.48	02/18/2021							
26	81	0	Hotel BD	Natal (RN)	4	242.88	971.52	04/15/2021							
27	86	0	Hotel AF	Sao Paulo (SP)	3	139.1	417.3	05/20/2021							
28	87	0	Hotel AU	Recife (PE)	1	312.83	312.83	05/27/2021							
29	90	1	Hotel AF	Sao Paulo (SP)	1	139.1	139.1	10/9/2019							
30	92	1	Hotel BP	Brasilia (DF)	1	247.62	247.62	10/17/2019							
31	100	2	Hotel CB	Rio de Janeiro (RJ)	4	165.99	663.96	10/31/2019							
32	101	2	Hotel BW	Campo Grande (MS)	3	60.39	181.17	11/7/2019							
33	105	2	Hotel Z	Aracaju (SE)	4	208.04	832.16	12/5/2019							
34	108	2	Hotel AU	Recife (PE)	3	312.83	938.49	12/26/2019							
35	109	2	Hotel BW	Campo Grande (MS)	2	60.39	120.78	1/2/2020							
36	110	2	Hotel K	Salvador (BH)	1	263.41	263.41	1/9/2020							
37	112	2	Hotel AF	Sao Paulo (SP)	4	139.1	556.4	01/23/2020							
38	116	2	Hotel BW	Campo Grande (MS)	4	60.39	241.56	02/20/2020							
39	117	2	Hotel AU	Recife (PE)	2	312.83	625.66	02/27/2020							
40	119	2	Hotel K	Salvador (BH)	2	263.41	526.82	3/12/2020							
41	122	2	Hotel A	Florianopolis (SC)	4	313.02	1252.08	4/2/2020							
42	124	2	Hotel Z	Aracaju (SE)	3	208.04	624.12	04/16/2020							

This dataset is a modified version of the "Agoda Hotels" dataset from Kaggle [4], with key additions to facilitate more detailed analysis. The dataset includes Customer Data, which contains fields such as customerID, name, email, phone number, and age. It also includes Booking Data, with fields like bookingID, bookingDate, checkinDate, checkoutDate, bookingStatus, customerID, and roomID. Additionally, Booking Timestamp Data is incorporated, including timeStampID, bookingDate, and bookingID, which provides crucial

time-related information. These additions enable a deeper understanding of customer behavior, booking patterns, and other insights not available in the original dataset.

2. Hotel-Room Relationship (Own)

	A	B	C	D	E	F	G	H	I	J	K
1		Hotel name	Location	Price	Discount	Rating					
2		0 Capital O 80961 Hotel Radisson Suite	India, Mumbai		2819 65% off	104					
3		1 OYO SilverKey Hotel Manas Residency	Chembur East, Mumbai		2702 65% off	410					
4		2 OYO Hotel Airport Metro Near Chhatra Andheri East, Mumbai			2289 64% off	879					
5		3 Collection O Hotel Kiwi International I Andheri East, Mumbai			2094 66% off	75					
6		4 OYO Druv's Taj Palace Near Chhatrapati Near Asapha Metro Station, Dhumal E			2656 65% off	686					
7		5 Collection O Hotel Jaysh Vira Parle East, Mumbai			3127 65% off	588					
8		6 OYO Hotel Blue Executive Near Chhat On main street, Near Saki Vihar compli			2772 65% off	1875					
9		7 OYO Blue Pearl Residency Near Chhat On main street, Near Asapha Metro St			1921 65% off	995					
10		8 OYO Hotel Blue Sapphire Residency N Near The Crescent Business Park, Saki			2220 65% off	1144					
11		9 OYO Hotel Blue Sea Near Chhatrapati On main street, Near Asapha Metro St			1921 65% off	2051					
12		10 Super OYO Hotel Divine Residency Ne Near Chirag Nagar Police Station, LBS			4069 65% off	644					
13		11 Capital O Hotel Adore Palace	Sakinaka Jharimari Link Road, Mumbai		2092 66% off	34					
14		12 OYO Hotel Greenforest Near Airport R Near Axis Bank, Customs Colony, Muri			2247 67% off	664					
15		13 Capital O Hotel Pearl Near US Embassy Near Wajan Kata Bus Stop, Indira Nagi			2896 65% off	119					
16		14 OYO Anjali Hotel Mumbai	Near Kanti Nagar, Andheri East, Mumb		2535 67% off	477					
17		15 OYO Rana Residency Near Western Ex Tolary Naka, Sher-e-Punjab Colony, M			1916 67% off	589					
18		16 Flagship The Mumbai Inn Near Chhatra The Mumbai Inn 90 Feet Road Next To I			1863 67% off	470					
19		17 OYO Royal Inn Near Versova Beach	Near Jankidevi Public School Road, Va		2740 67% off	465					
20		18 Capital O Hotel Zaid International NE Near Huntsman Building, Marol, Mumbai			3273 65% off	711					
21		19 OYO Flagship Om Residency Near Arp Near Ganesh nagar, Marol, Mumbai			2520 67% off	16					
22		20 SPOT ON Yatra Inn Near Lalbagh Botar Lalbagh, Bangalore			1040 64% off	2617					
23		21 OYO Townhouse 180 Jayanagar	Near Sagar Hospital, Tilak Nagar, Bang		1828 67% off	3524					
24		22 Super OYO Townhouse OAK Spruce Ho India, Bangalore			1423 75% off	253					
25		23 Super OYO Collection O Sabharwal Re Opposite Shanthinagar, Bangalore			1410 65% off	429					
26		24 OYO Townhouse Chic 364 Richmond R Near Baldwin Women College, Richm			2082 65% off	1457					
27		25 OYO Hotel Townhall	Near Dhankalakem Bank, JO Road, Ban		1070 65% off	1144					
28		26 Collection O Hotel Hemadri Residenc	Opp Saraswati Kalyana Mantapa, Bang		872 67% off	74					
29		27 Super OYO Capital O Boulevard Elite h B T M Layout, Bangalore			1365 66% off	2664					
30		28 OYO Sai Ram Residency	Cubbonpet Main Road, Bangalore		1092 67% off	207					
31		29 Super OYO Design Oyo Townhouse OA 4th Block, Rajaji Nagar, Bangalore			3053 65% off	1360					
32		30 SPOT ON Hotel Relax Inn Near Banash Banashankari 3rd Stage, Bangalore			645 77% off	69					
33		31 Super OYO Townhouse 385 Nayya Inte	Near Rammamurthy Nagar Police Stati		1274 67% off	2655					
34		32 OYO Sree Baiji Residency	Agrahara Dasarahalli, Rajaji Nagar, Ba		907 67% off	360					
35		33 Super OYO Townhouse 238 Tierra Woo	Halasuru, Bangalore		2276 65% off	1675					
36		34 Capital O Champion Lotus Near Mantri Near IFT, Vijayanagar, Bangalore			1339 65% off	1342					
37		35 Capital O Premium Hotel Victoria Heij	Near St. Philomena's Hospital, Xavier L		1508 67% off	957					
38		36 Townhouse OAK SRK White Inn Near Ru	Koramangala, Bangalore		1427 65% off	1603					
39		37 OYO Hotel Silver House Near Mantri S Arghar	Dasarahalli Near Madhu Hospitz		844 65% off	494					
40		38 Viceroy Emerald Near Raggedda Sri Pr	Near KEB Colony, BTM Layout 1, Bangl		879 65% off	1965					
41		39 Collection O Hotel Palm Tree	Near Halasuru Metro Station, Ulsoor, I		2566 65% off	1271					
42		40 Super OYO Townhouse 385 Nayya Inte	Near Rammamurthy Nagar Police Stati		1274 67% off	2655					

This dataset is based on a modified version of the "Oyo Hotel Rooms" dataset from Kaggle [5], with several key additions for enhanced analysis. The dataset now includes Hotel Data, featuring fields such as hotelID, name, location, description, and rating. It also includes Room Data, with details like roomID, roomType, price, availableStatus, and hotelID. Additionally, Hotel Timestamp Data has been added, which includes HoteltimeStampID, checkInTime, and checkOutTime. These enhancements allow for a more comprehensive analysis of hotel operations, room availability, and time-based data, providing richer insights than the original dataset.

3. Room-Booking Relationship (Book)

This dataset uses a merged dataset. It combines enhanced versions of the "Agoda Hotels" and "Oyo Hotel Rooms" datasets from Kaggle. The merge, based on the shared roomID, links

booking details (from Agoda, including bookingID, bookingDate, checkinDate, checkoutDate, and bookingStatus) with room information (from Oyo, including roomID, roomType, price, and availableStatus). This combined dataset allows for analysis of how room characteristics and meal options influence booking behavior.

4. Booking-Payment Relationship (Pay For)

This dataset is a further enhanced version of the "Agoda Hotels" dataset from Kaggle. In addition to the original booking data, we have added fields related to payments, including paymentID, paymentDate, paymentAmount, and paymentStatus. Furthermore, Payment Timestamp Data has been incorporated, with fields such as timeStampID, paymentTime, and paymentDate. These additions, when linked to the existing booking data (such as bookingID, bookingDate, checkinDate, checkoutDate, bookingStatus, roomID, and customerID), enable a more in-depth analysis of payment trends, success rates, and the relationship between booking details and payment behavior.

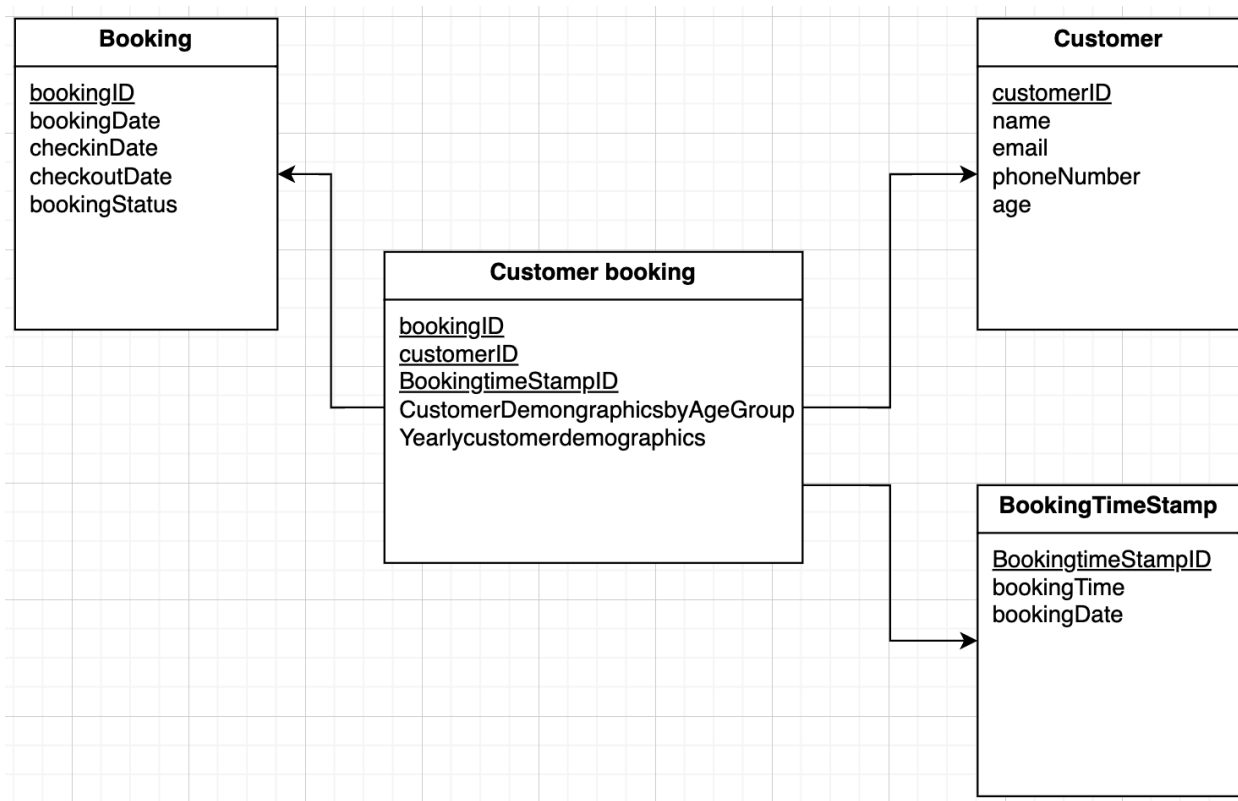
Data Warehouse Design

Overview of Data Warehouse Design

The Star Schema in data warehouse design organizes data into a central fact table for numerical metrics, surrounded by dimension tables for context. This structure simplifies queries, enhances performance, and supports efficient analysis. In a hotel booking system like Agoda, it helps track bookings, payments, and customer behavior for better decision-making.

Star Schema Design for Agoda's Data Warehouse

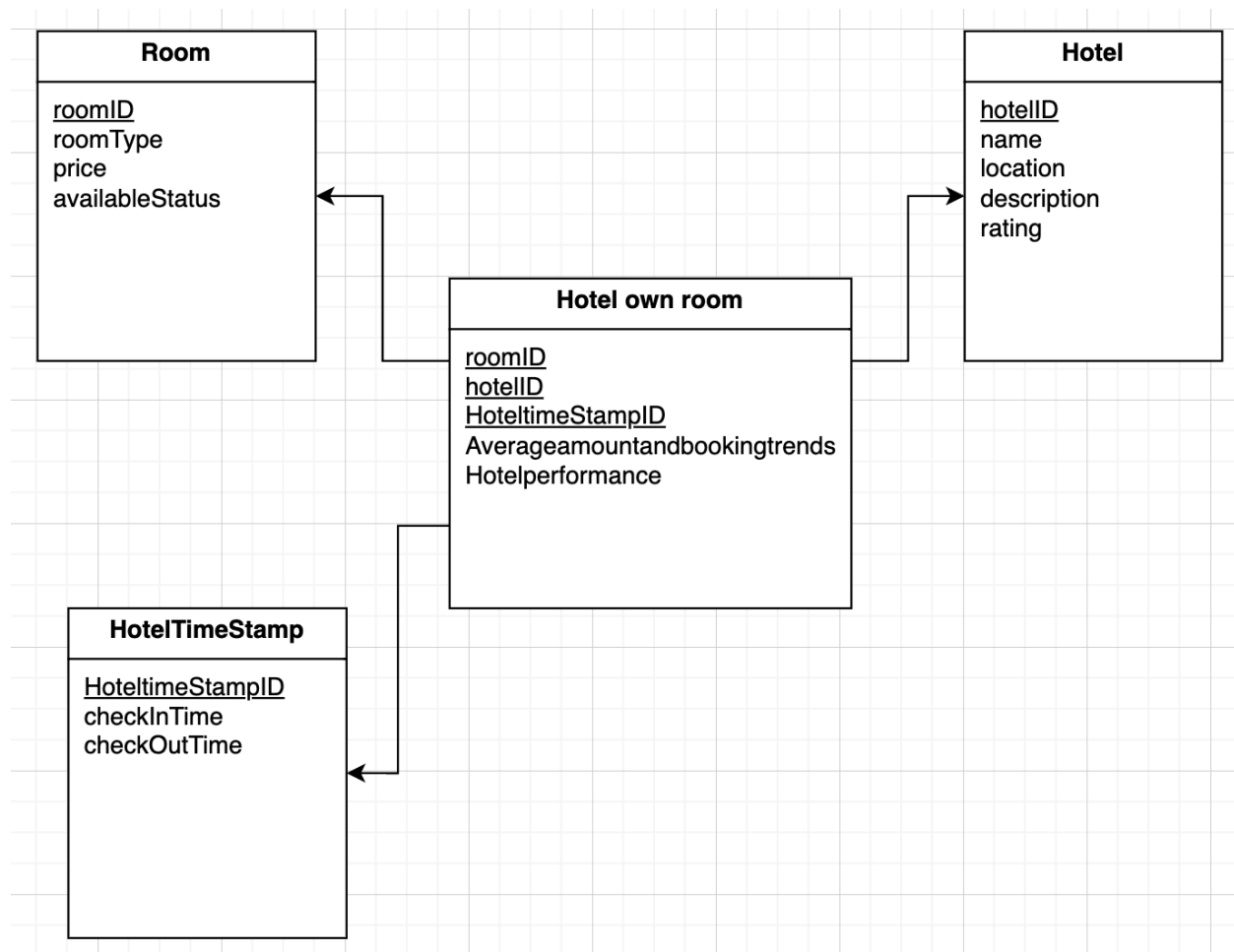
Customer booking



This star schema represents a booking system database with a central fact table, Customer Booking, which records transactional data and connects to multiple dimension tables. The fact table includes bookingID, customerID, and BookingtimeStampID as foreign keys linking to their respective dimension tables, along with demographic attributes like CustomerDemographicsbyAgeGroup and YearlycustomerDemographics. The Booking table contains bookingID as the primary key and attributes such as bookingDate, checkinDate, checkoutDate, and bookingStatus, representing details of each booking. The Customer table stores customer-related information

with customerID as the primary key and attributes like name, email, phoneNumber, and age. The BookingTimeStamp table manages timestamp-related details, with BookingtimeStampID as the primary key and attributes bookingTime and bookingDate. This schema follows a star structure, with Customer Booking at the center and the surrounding dimension tables providing descriptive context, enabling efficient analytical queries on customer behavior, booking trends, and demographic insights.

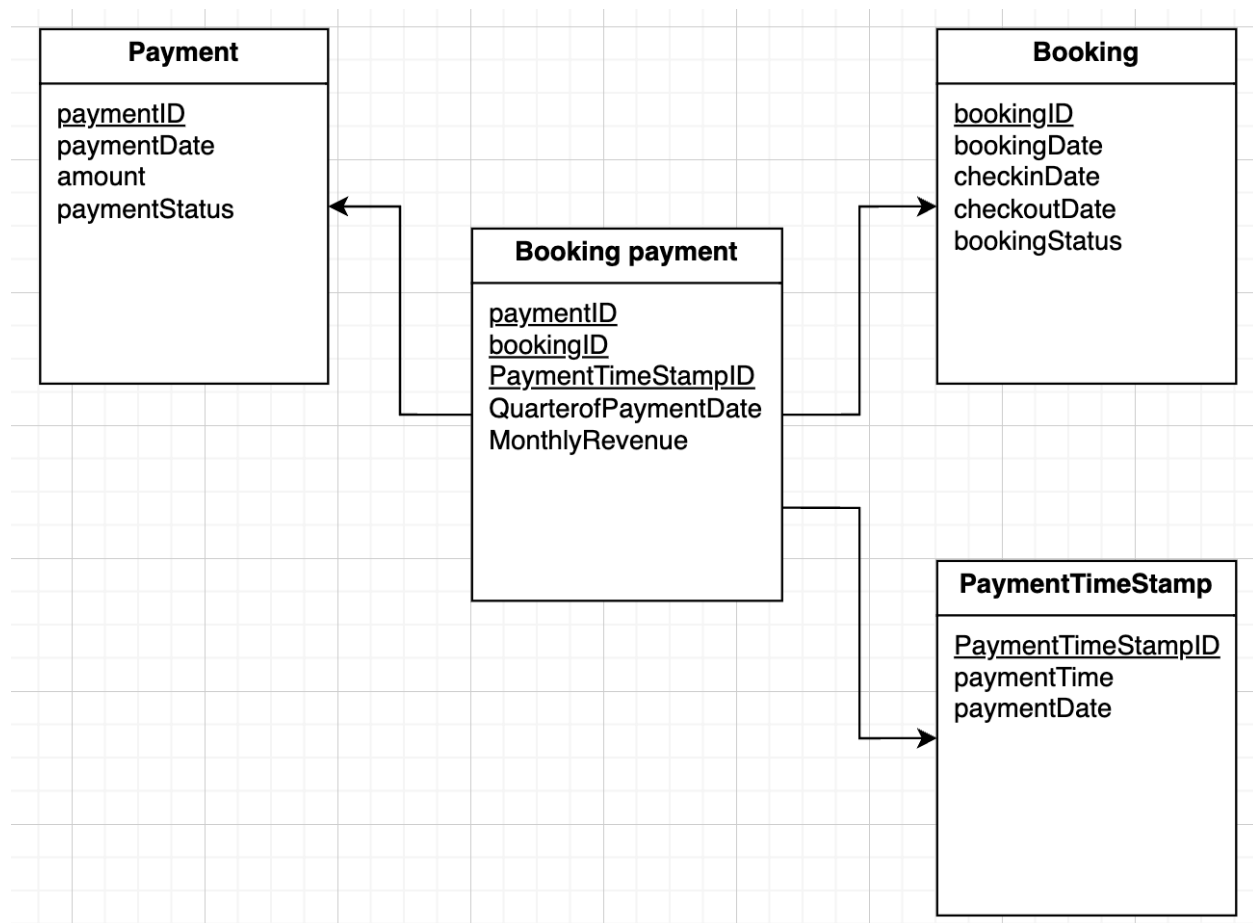
Hotel own room



This star schema represents a hotel room booking system with a central fact table, Hotel own room, which tracks room ownership and booking trends while connecting to multiple dimension tables. The fact table includes roomID, hotelID, and HoteltimeStampID as foreign keys linking to their respective dimension tables, along with attributes like Averageamountandbookingtrends and Hotelperformance for analytical insights. The Room table contains roomID as the primary

key and attributes such as roomType, price, and availableStatus, describing room details. The Hotel table stores information about hotels, with hotelID as the primary key and attributes like name, location, description, and rating. The HotelTimeStamp table captures time-related details with HotelTimeStampID as the primary key and attributes checkInTime and checkOutTime. This schema follows a star structure, with Hotel own room as the central fact table and dimension tables providing contextual information, enabling efficient queries on room availability, hotel performance, and booking trends.

Booking payment



This star schema represents a payment and booking system, with the central fact table, Booking payment, recording payment transactions linked to bookings. The fact table contains paymentID, bookingID, and PaymentTimeStampID as foreign keys referencing their respective dimension tables, along with attributes like QuarterofPaymentDate and MonthlyRevenue for financial analysis. The Payment table, with paymentID as the primary key, stores details about

transactions, including paymentDate, amount, and paymentStatus. The Booking table, with bookingID as the primary key, holds information such as bookingDate, checkinDate, checkoutDate, and bookingStatus, representing reservation details. The PaymentTimeStamp table, identified by PaymentTimeStampID, captures the temporal aspect of payments with paymentTime and paymentDate. This star schema structure is optimized for financial reporting and revenue analysis by enabling efficient queries on booking payments, revenue trends, and payment timing pattern.

ETL Process

Overview of the ETL (Extract, Transform, Load) Processes

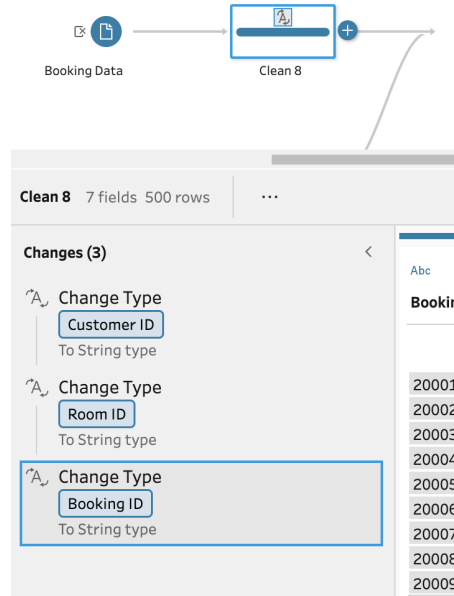
ELT is short from Extract, Transformation, and Loading. It is the processes used for collecting, filtering, integrating, and aggregating data.

Mapping Between OLTP and OLAP (Operational and Analytical Systems) and Tableau Prep Screenshots and Step-by-Step ETL Integration

This is an overview of the flow.

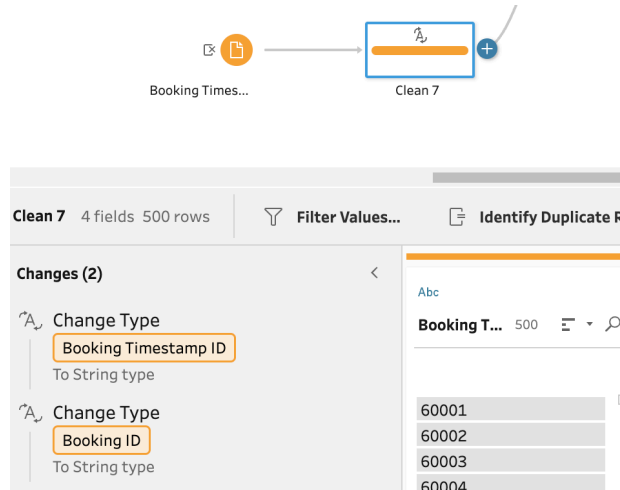


Booking Data



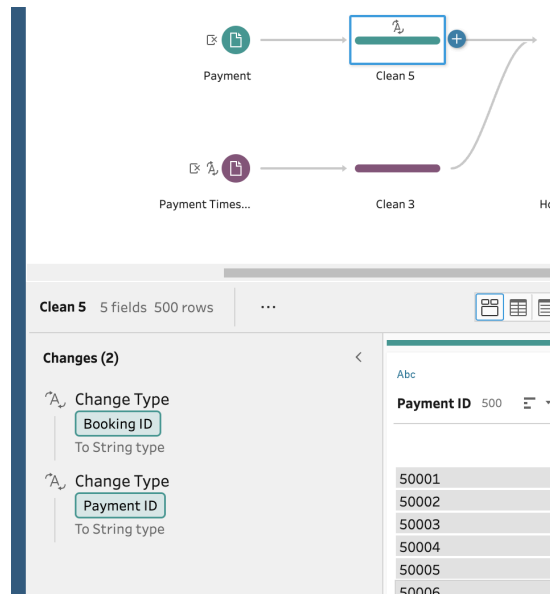
In the booking data, Data (CustomerID, RoomID, BookingID) are cleaned by changing the type of data from the whole number which has “,” in it to String Type.

Booking TimeStamp



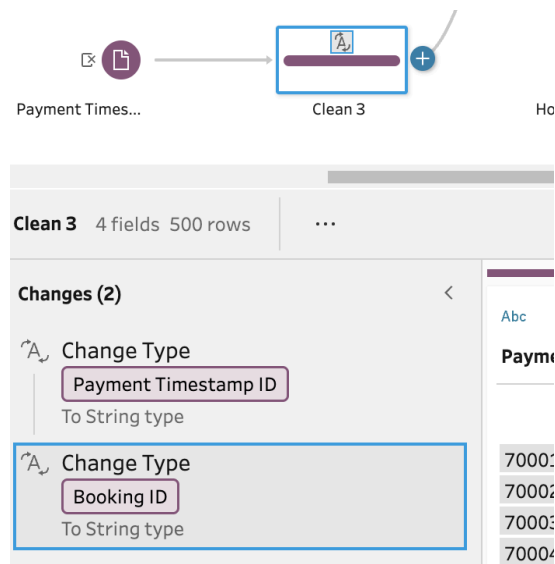
In the Booking TimeStamp data, booking timeStampID and BookingID are changed from whole number to stringtype.

Payment



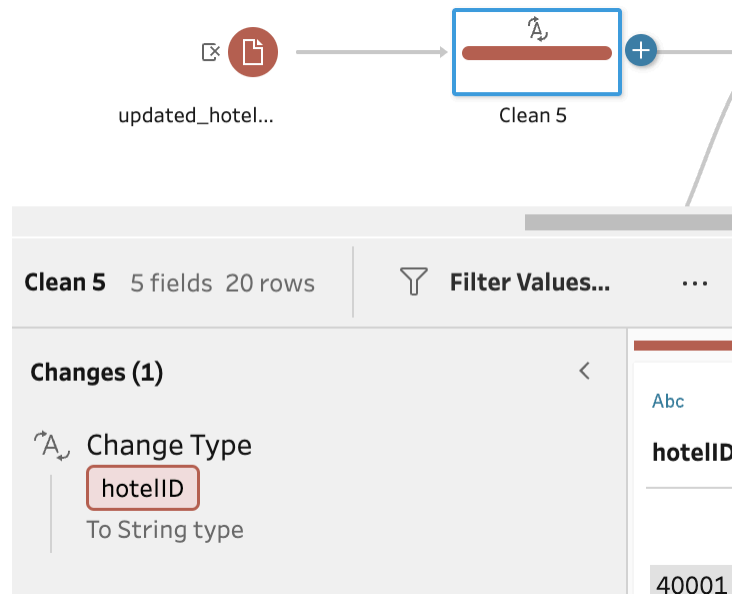
In the payment data, “Booking ID” and “Payment ID” data type are changed from whole number to string.

Payment TimeStamp



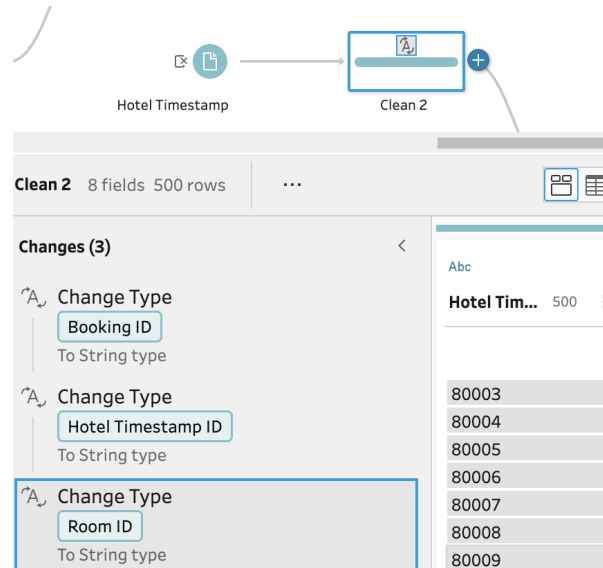
In the payment TimeStamp, “Payment TimeStampID” and “BookingID” data type are changed from whole number to string.

Hotel



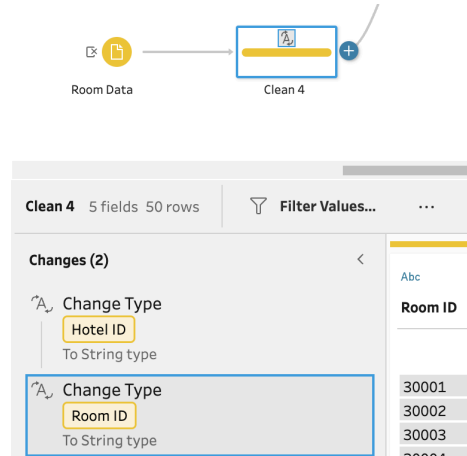
In the hotel data, “hotelID” data type is changed from whole number to string.

Hotel TimeStamp



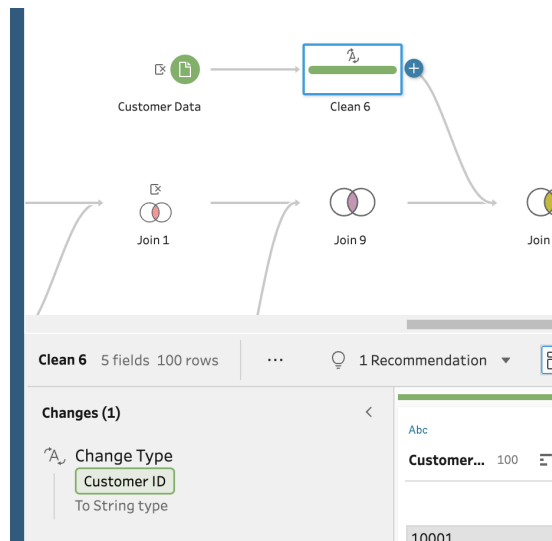
In the hotel timestamp data, “BookingID”, “Hotel TimeStampID” and “RoomID” data type are changed to string from whole number.

RoomData



In the Room data, “HotelID” and “RoomID” data types are changed from whole number to string.

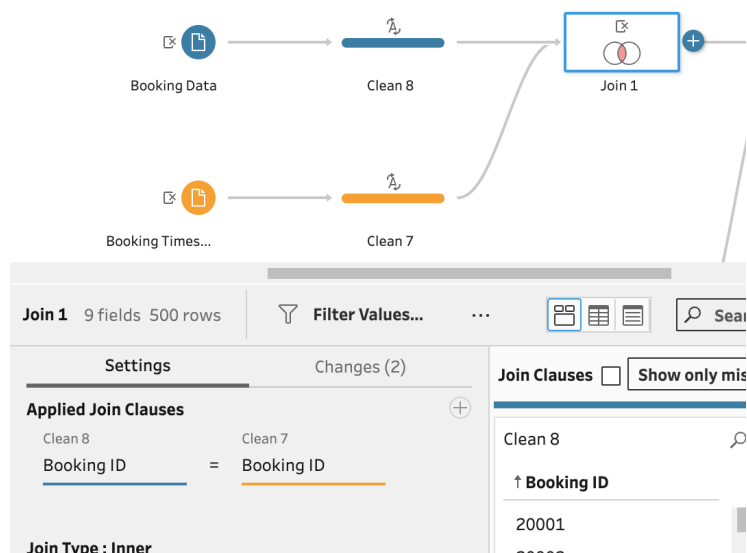
Customer Data



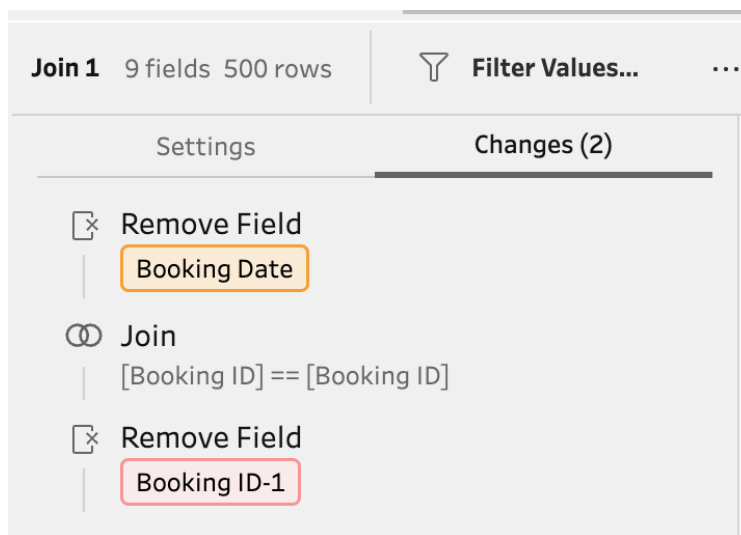
In the customer data, “CustomerID” data types are changed from whole number to string.

Integrating data together

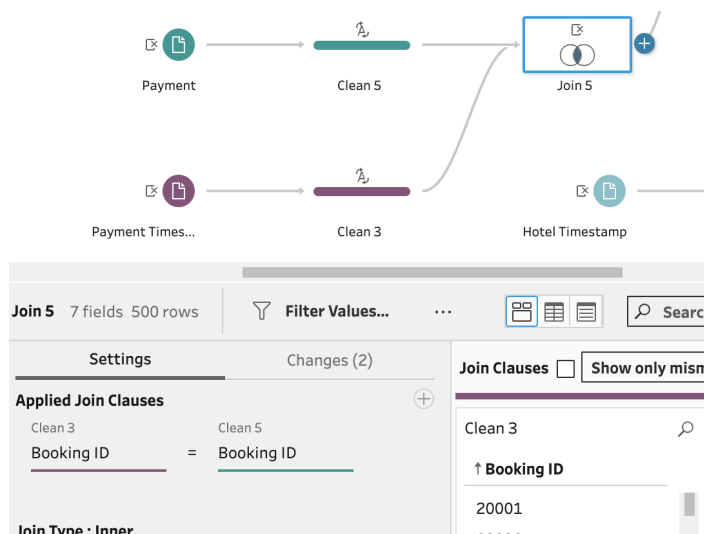
1. Booking Data and Booking Timestamp



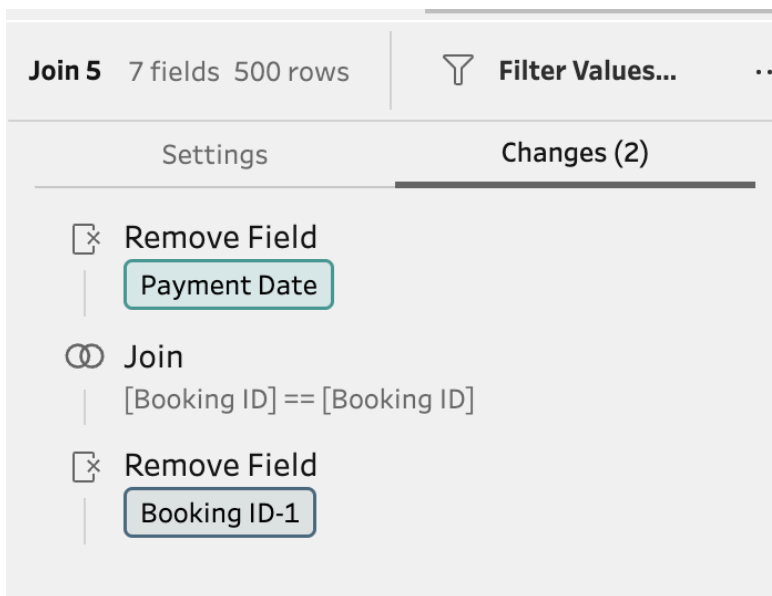
Join Booking Data and Booking TimeStamp together using BookingID and then clean data by removing duplicates columns of data.

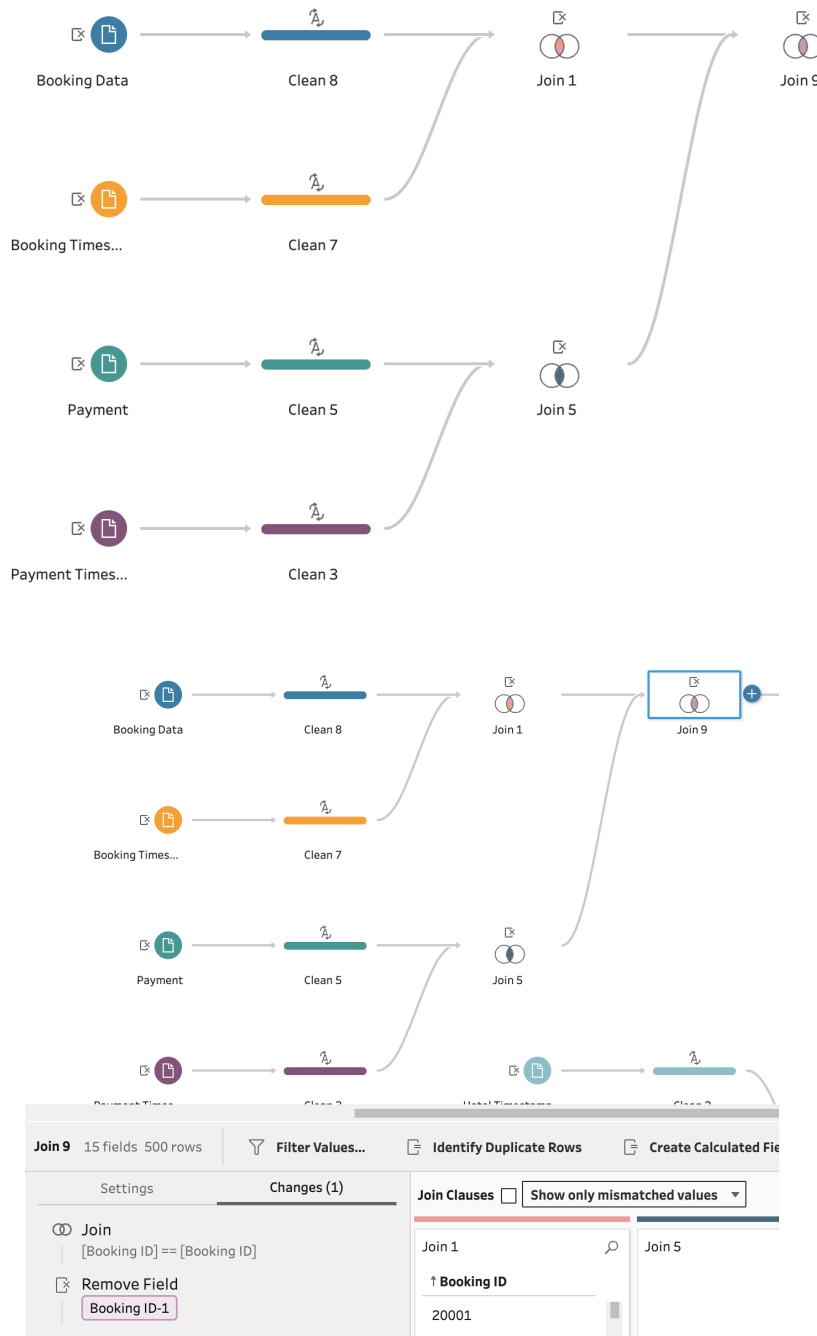


2. Payment and Payment TimeStamp



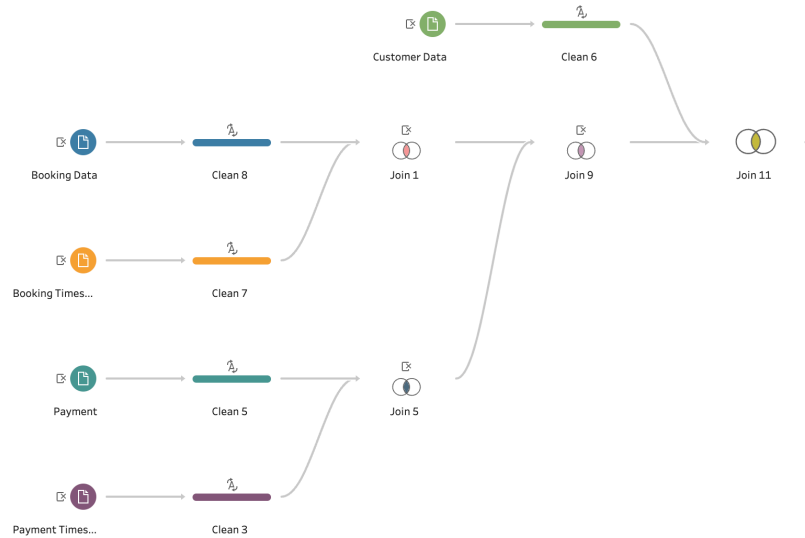
Join Payment and Payment TimeStamp together by using BookingID and then remove duplicate columns out.



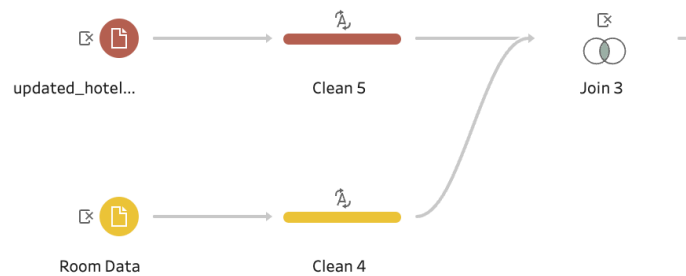


3. Join the already joined data from booking data and booking time stamp with the already joined data between payment and payment timestamp and then clean the duplicate columns out.

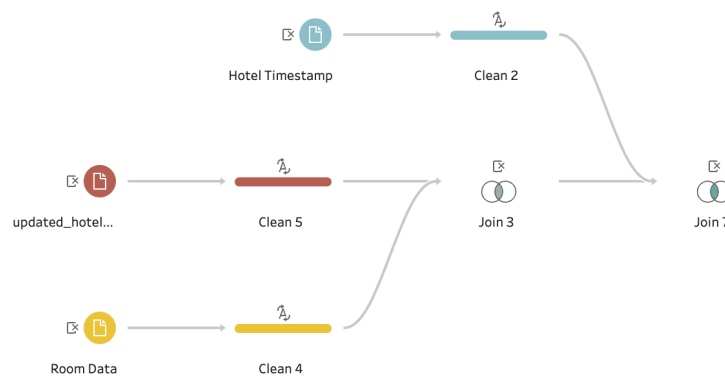
4. Join the joined data previous step with customer data.



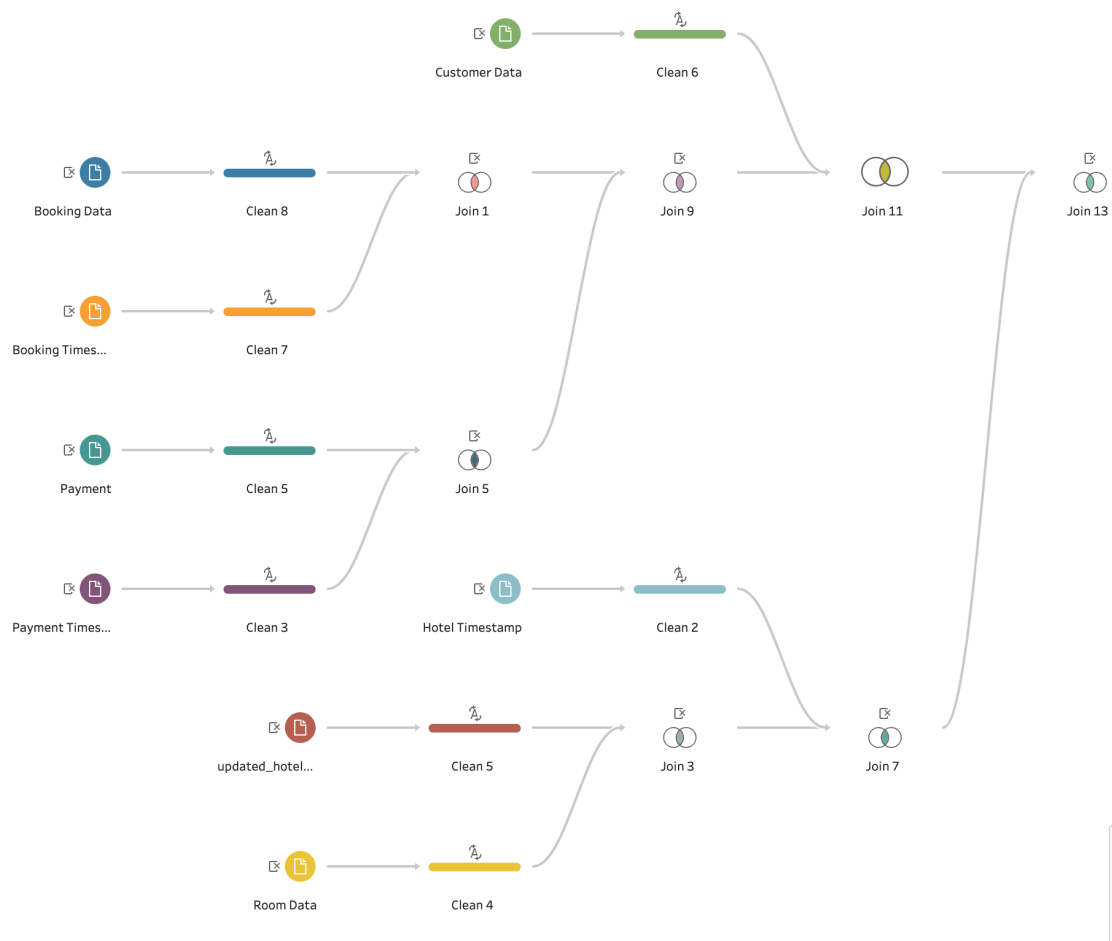
5. Join hotel data with room data and cut duplicate columns out.



6. Join the previous step with hotel timestamp data and cut duplicate columns out.



7. Join all previous steps together.

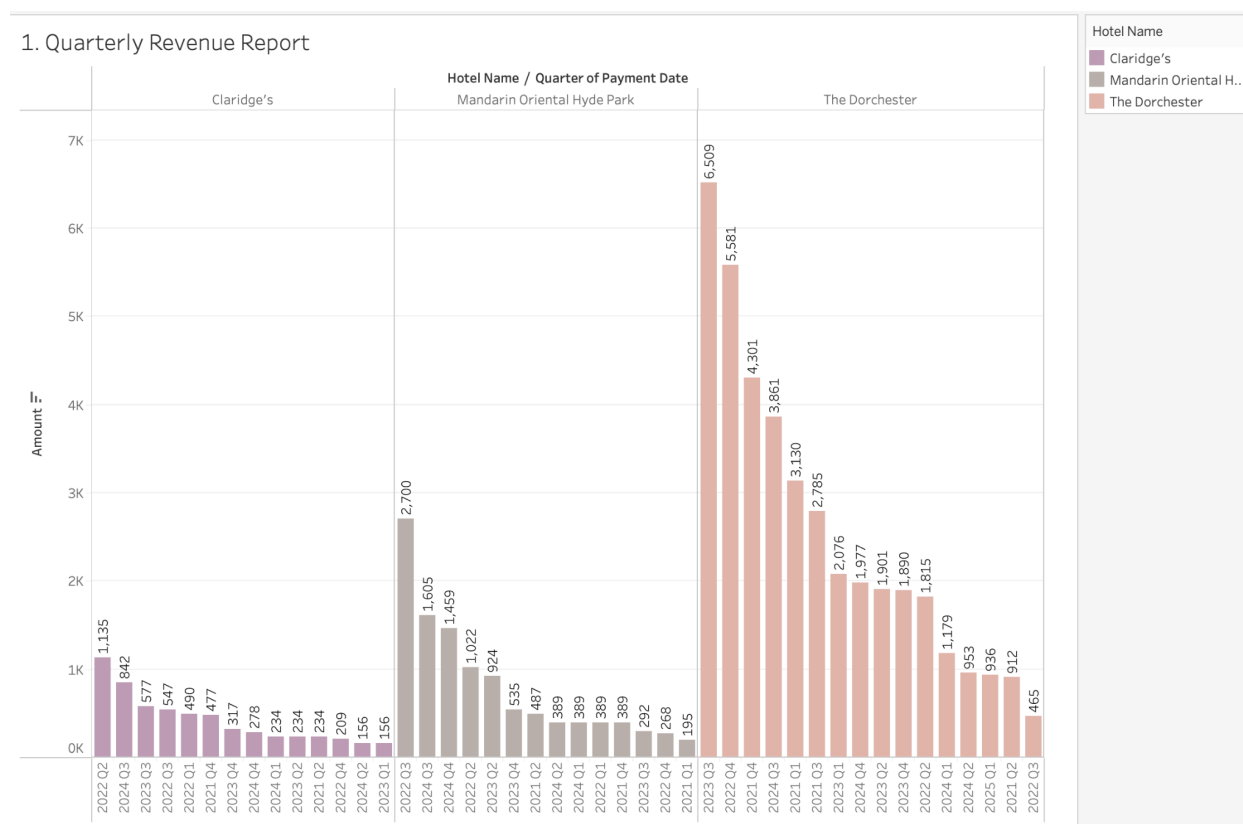


Analysis and Visualization Reports

Overview of Tableau Desktop for Data Visualization

Tableau Desktop is a powerful data visualization tool that allows users to create interactive and shareable dashboards. It helps in simplifying raw data into easily understandable formats through visualizations like graphs, charts, and maps. For Agoda, Tableau Desktop provides a platform to analyze large volumes of booking, customer, and pricing data, facilitating better decision-making and operational efficiency.

Report 1: Quarterly Revenue Report



This report visualizes the quarterly revenue trends of three hotels—Claridge's, Mandarin Oriental Hyde Park, and The Dorchester—located in London, UK. It provides insights into revenue fluctuations over time, enabling businesses to analyze performance, forecast future revenue, and refine pricing strategies.

Key Insights from the Visualization

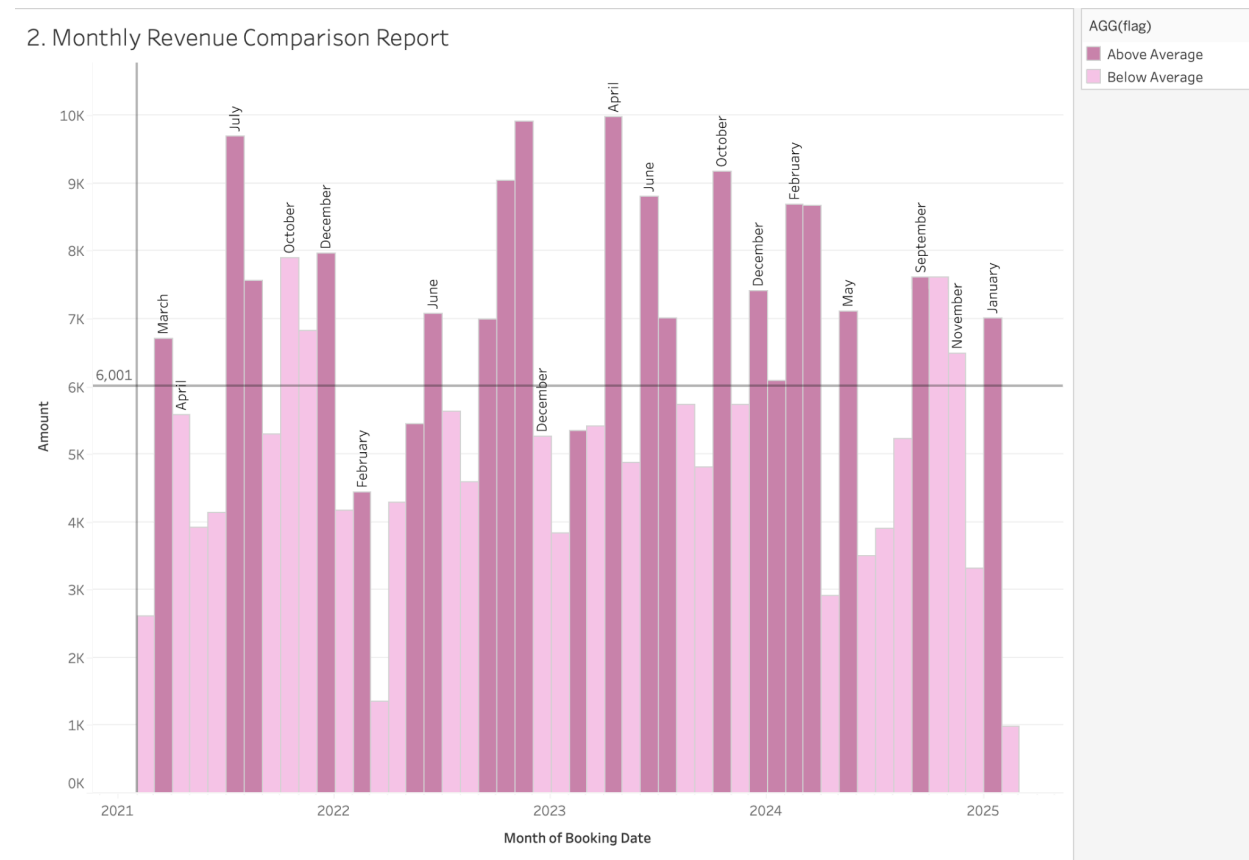
- **X-Axis:** Quarter of Payment Date (2021 Q1 to 2024 Q3)
- **Y-Axis:** Revenue amount in \$ per quarter

- **Visualization Type:** Bar chart segmented by hotel

How This Report Meets Business Requirements

- **Hotel Performance Analysis:**
 - Identifies revenue trends for each hotel, highlighting peak and low-performing quarters.
 - It helps management understand seasonality and demand shifts.
- **Financial Forecasting:**
 - Enables data-driven revenue predictions based on historical performance.
 - Supports pricing, promotional, and operational adjustments to maximize revenue.
- **Competitive Benchmarking:**
 - Compares revenue performance across different hotels, offering insights into market positioning.
 - Identifies potential performance gaps and areas for improvement in pricing strategies.



Report 2: Monthly Revenue Comparison Report



This report analyzes monthly revenue performance over different years, comparing each month's revenue to the average performance. It categorizes months into above-average and

below-average revenue periods, helping businesses identify seasonal trends and optimize marketing and pricing strategies accordingly.

Key Insights from the Visualization

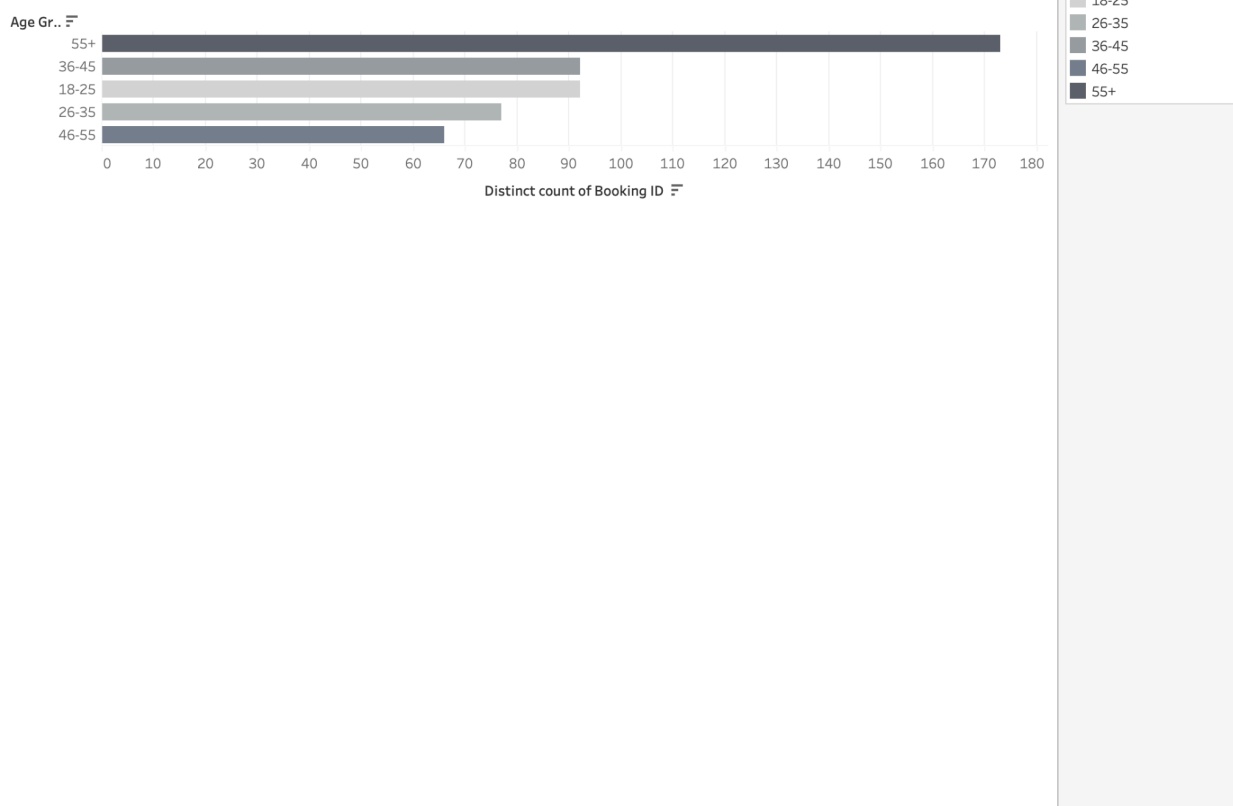
- **X-Axis:** Month of Booking Date (2021 to 2025)
- **Y-Axis:** Revenue amount in \$ per month
- **Visualization Type:** Bar chart highlighting above-average months  and below-average months 

How This Report Meets Business Requirements

- **Seasonality Analysis:**
 - Identifies peak and slow months, allowing businesses to adjust pricing and marketing strategies accordingly.
 - Helps allocate marketing budgets more effectively to capitalize on high-demand periods.
- **Revenue Optimization:**
 - Hotels can increase promotional efforts during historically below-average months to boost bookings.
 - Supports data-driven pricing adjustments to maximize revenue.
- **Budget Planning:**
 - Aids in strategic financial planning by aligning operational resources with expected demand fluctuations.
 - Helps hotels prepare for seasonal shifts in occupancy and revenue.

Report 3: Customer Demographics by Age Group

3. Customer Demographics by Age Group



This report analyzes customer distribution by age group, offering insights into the dominant demographics. By understanding the age composition of their customer base, hotels can tailor marketing strategies, customize services, and enhance guest experiences.

Key Insights from the Visualization

- **X-Axis:** Distinct count of Booking ID
- **Y-Axis:** Age groups (18-25, 26-35, 36-45, 46-55, 55+)
- **Visualization Type:** Horizontal bar chart representing the number of bookings per age group

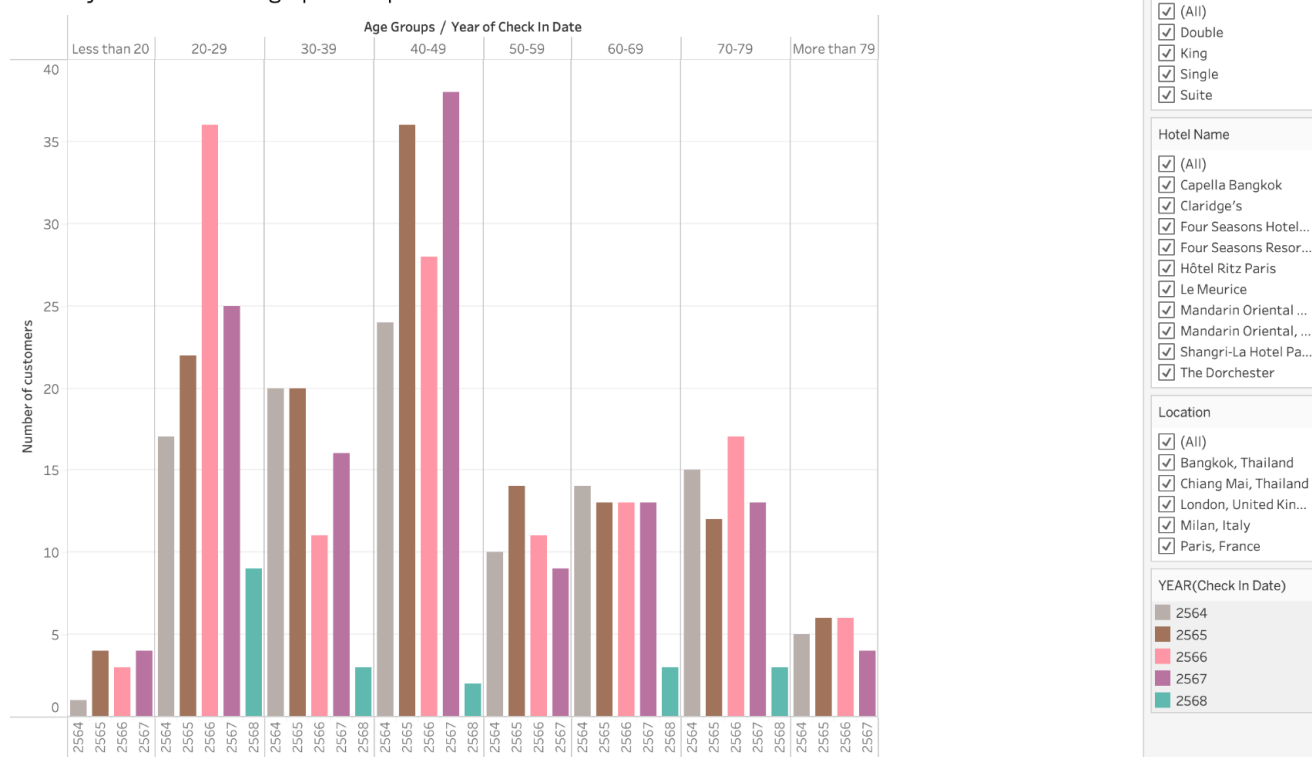
How This Report Meets Business Requirements

- **Personalized Marketing:**
 - Helps create age-specific promotions and loyalty programs to attract and retain different customer segments.
 - Enables targeted advertising on platforms popular among specific age groups.
- **Service Customization:**
 - If a significant portion of bookings come from older customers (55+), hotels can enhance accessibility features, wellness services, and senior-friendly amenities.

- Younger demographics (18-35) may benefit from tech-savvy features, nightlife recommendations, and digital concierge services.
- **Target Audience Insights:**
 - Identifying the most frequent age groups helps refine advertising strategies and partnerships with travel agencies and online platforms catering to those demographics.
 - Supports long-term customer retention strategies by adapting offerings to evolving age-based trends.

Report 4: Yearly Customer Demographics Report

4. Yearly Customer Demographics Report



This report visualizes customer demographics by age groups over the years of Check-in Date. It shows trends in the distribution of customers based on age and year, allowing Agoda to understand how different age groups are engaging with the platform over time.

Key Insights from the Visualization

- X-Axis:
 - Age Group: Less than 20, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, More than 79
 - Years Check-in Date (2564 , 2565 , 2566 , 2567 , 2568)

- Y-Axis: Number of customers in each age group per year.
- Visualization Type: A bar chart to show the number of customers for each age group over multiple years.

How This Report Meets Business Requirements

- Data-Driven Strategy for Personalization:

This report segments customers by age group and tracks their behavior over time, enabling Agoda to create personalized campaigns. For example, if the 40-49 age group is booking frequently, Agoda can target this group with exclusive promotions, tailored ads, or loyalty rewards.

- Dynamic Pricing Optimization:

By identifying age groups with higher booking rates, Agoda can adjust dynamic pricing. For instance, if younger travelers (20-29) are booking less often, Agoda could implement student discounts or special offers to increase bookings from this group.

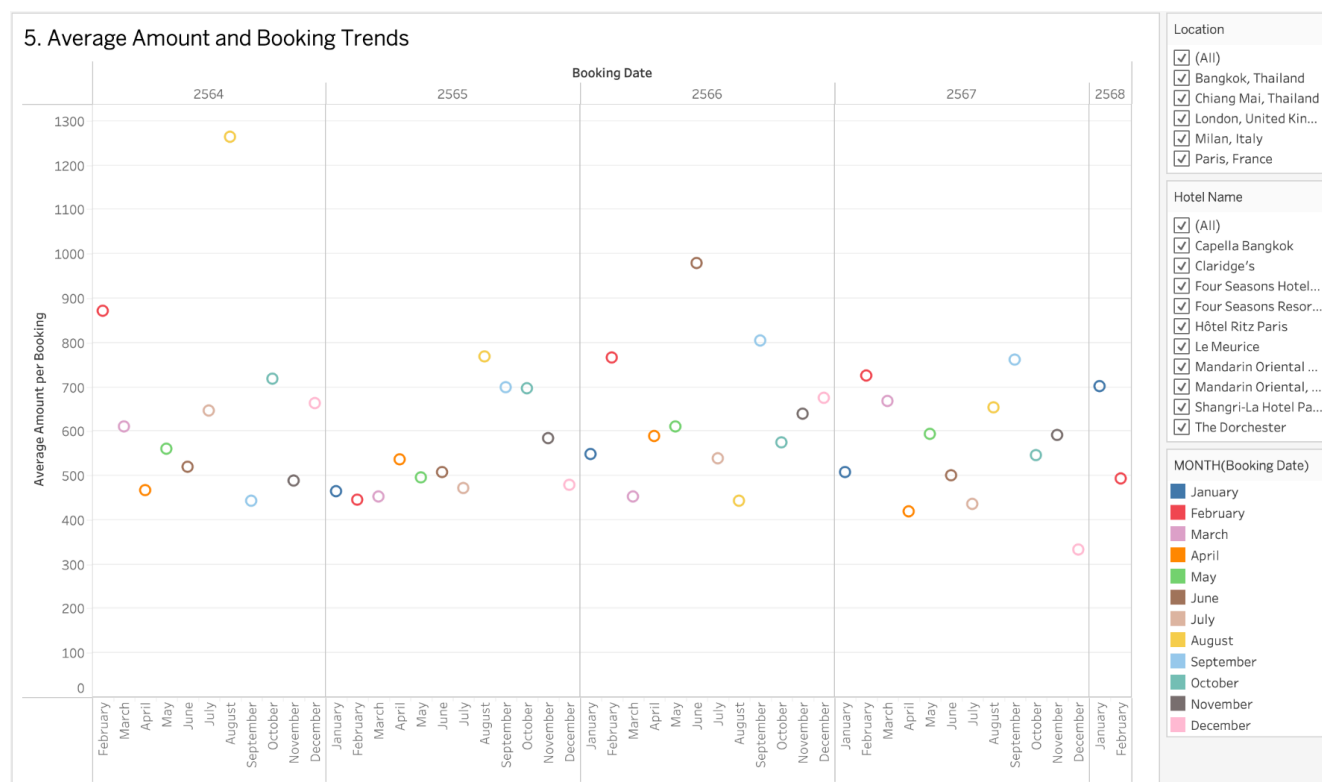
- Operational Efficiency & Resource Allocation:

This report allows hotels to optimize room availability based on the dominant age group of travelers. For instance, if older travelers (70-80) show a rise in bookings, Agoda can suggest senior-friendly rooms, improved accessibility services, and additional medical assistance.

- Forecasting Demand & Business Growth Strategy:

Tracking yearly changes in demographics helps forecast future booking trends, enabling long-term strategic planning. For example, a decline in younger travelers might require a shift in marketing strategies or a re-evaluation of services targeted at this segment.

Report 5: Average Amount and Booking Trends



This report will visualize how the average booking amount varies over time and highlight trends in customer bookings. It helps Agoda understand how pricing, seasonality, and booking patterns change over different time periods.

Key Insights from the Visualization

- X-Axis:
 - Booking Date:
 - Month (January ■, February ■, March ■, April ■, May ■, June ■, July ■, August ■, September ■, October ■, November ■, December ■)
 - Years (2564, 2565, 2566, 2567, 2568)
- Y-Axis: Average amount per booking.
- Visualization Type: A scatter plot with color-coded points indicating booking months.

How This Report Meets Business Requirements

- Price Optimization:

The visualization helps Agoda identify patterns of high and low booking amounts based on seasonality. For instance, during holidays, prices can be increased using dynamic pricing, and discounts can be introduced during off-peak months to encourage bookings.

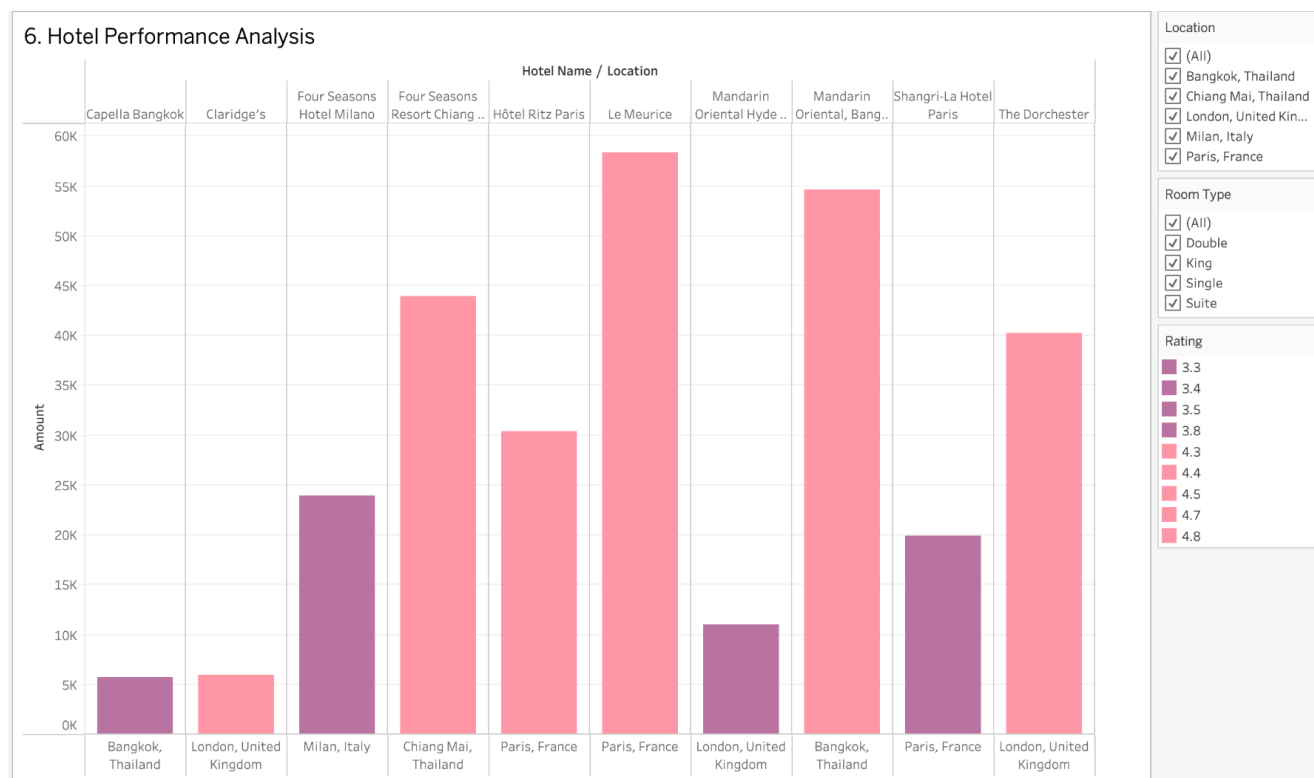
- Demand Forecasting:

Tracking how average booking amounts fluctuate over time helps Agoda forecast future demand and adjust pricing strategies. For example, the system can anticipate high-demand months and prepare pricing adjustments accordingly.

- Revenue Growth:



By understanding when customers are likely to spend more, Agoda can optimize revenue during peak months and avoid underpricing during slow periods. Targeted promotions and dynamic pricing can drive increased revenue during high-demand periods.

Report 6: Hotel Performance Analysis



This report offers an analysis of how each hotel performs in terms of revenue generation, and how customer ratings correlate with total revenue. This enables Agoda to identify top-performing hotels and correlate revenue with customer satisfaction.

Key Insights from the Visualization

- X-Axis:
 - Booking Date: Hotel Name, Hotel Location
- Y-Axis: Total revenue for each hotel
- Visualization Type: Bar chart, with color-coding indicating ratings (3.0 - 3.9 , 4.0 - 4.9 )

How This Report Meets Business Requirements

- **Revenue Optimization:**

The performance insights allow Agoda to focus on high-rated hotels in marketing campaigns and special offers. For example, hotels with ratings above 4.0 could be featured in premium campaigns, potentially increasing overall profitability.
- **Customer Satisfaction:**

By correlating customer ratings with revenue, this report provides insights into which hotels offer the best experiences and generate the highest returns. This helps Agoda prioritize marketing and customer service improvements.
- **Hotel Performance Monitoring:**

This report tracks which hotels are driving the most revenue, enabling Agoda to monitor performance and identify top performers. This data can also be used to address any underperforming properties, either through enhanced marketing or service improvements.

Discussion and Conclusion

In conclusion, the study, design, and development of the data warehouse for Agoda have significantly enhanced the company's ability to harness the power of data. Data is integral to Agoda's operations, as it informs every aspect of the business from customer bookings to pricing strategies, hotel management, and overall operational efficiency.

Key Benefits of the Data Warehouse and BI Implementation:

1. **Centralized Data Management:** The creation of a centralized data warehouse enables Agoda to consolidate large volumes of data from various sources such as bookings, payments, customer demographics, and hotel information. This single source of truth eliminates data silos, ensuring consistent, accurate, and accessible data for decision-making.
2. **Informed Decision-Making:** With Business Intelligence tools like Tableau integrated with the data warehouse, Agoda can analyze complex data and generate actionable insights. BI helps the company make data-driven decisions regarding pricing strategies, customer targeting, and inventory management, which are critical for improving profitability and staying competitive in the travel industry.
3. **Revenue Optimization:** BI allows Agoda to perform advanced analytics, identifying trends and patterns in customer behavior, seasonal demand, and booking preferences. This enables dynamic pricing, promotional strategies, and yield management techniques that maximize revenue and optimize hotel room availability. By forecasting demand more accurately, Agoda can adjust strategies proactively to capture higher-value bookings during peak periods and adjust during slower seasons.
4. **Enhanced Customer Insights:** The data warehouse empowers Agoda to segment and analyze customer data based on various attributes like demographics, booking history, and preferences. This level of customer insight allows Agoda to create personalized marketing campaigns, deliver tailored promotions, and improve customer experiences, fostering loyalty and repeat business.
5. **Operational Efficiency:** Through BI reporting and real-time data analysis, Agoda can streamline its operations, from forecasting room availability to optimizing resource

allocation. This leads to cost reductions, faster decision-making, and the ability to focus on high-impact areas, thus improving overall operational efficiency.

6. **Strategic Growth:** The data warehouse not only supports day-to-day operations but also lays the foundation for long-term strategic growth. By continuously analyzing historical trends and market conditions, Agoda can identify emerging opportunities, potential risks, and areas for expansion, ensuring that the company remains agile in a rapidly evolving industry.
7. **Competitive Advantage:** With the insights provided by BI tools, Agoda can stay ahead of competitors by anticipating market shifts, adjusting strategies in real time, and responding faster to customer demands. The ability to leverage data gives Agoda a distinct edge, allowing the company to position itself as a leader in the online travel industry.
8. **Compliance and Data Security:** As customer data becomes more valuable, protecting sensitive information is paramount. The data warehouse centralizes security measures, ensuring compliance with data privacy regulations, and maintaining customer trust. With robust data governance in place, Agoda can safeguard personal and financial data while adhering to global standards.

Conclusion

In conclusion, the integration of a data warehouse and BI system empowers Agoda to unlock the full potential of its data. It enhances decision-making, optimizes revenue, improves customer service, and ensures operational efficiency. This transformation not only strengthens Agoda's position in the market but also enables the company to evolve and adapt to changing customer needs, market conditions, and technological advancements. By leveraging data-driven insights, Agoda is well-positioned to drive growth, improve business outcomes, and maintain a competitive edge in the online travel industry.

VDO Clip demonstration

<https://drive.google.com/file/d/10-1kyIYHtMZlD5zFRGbEJs8VxBr7iaX9/view?usp=sharing>

References

- [1] Agoda. (n.d.). *About Agoda*. Agoda. Retrieved February 7, 2025, from <https://www.agoda.com/about-agoda/>
- [2] Finspace. (2023, August 8). *Do you know Booking and Agoda are own by the same person*. Finnomena. Retrieved February 7, 2025, from <https://www.finnomena.com/finspace/agoda-booking/>
- [3] “Hotel Booking System ER Diagram | Academic Projects,” Freeprojectz.com, 2017. https://www.freeprojectz.com/entity-relationship/hotel-booking-system-er-diagram#google_vignette (accessed Feb. 17, 2025).
- [4] A. Mittal, “Dataset of Agoda Hotels,” Kaggle.com, 2021. <https://www.kaggle.com/datasets/atulmittal199174/dataset-of-agoda-hotels> (accessed Feb. 17, 2025).
- [5] S. maheshwari, “Oyo hotel rooms,” Kaggle.com, 2023. <https://www.kaggle.com/datasets/sonu1maheshwari/oyo-hotel-rooms> (accessed Feb. 17, 2025).