

Data Warehouse Project

Airline Data Warehouse

Supervised by:
Eng. Nada Fathallah
Eng. Youssef Etman



Implementation done by:

Abdullah Ahmed Salah
Hady Ahmed Zahran
Omar Ashraf Hamdy

Introduction.....	(3)
Full Logical Model.....	(5)
Flight Activity Fact.....	(6)
Reservation Fact.....	(7)
Hotel Fact.....	(8)
Frequent Flyer Fact.....	(9)
Customer Care Fact.....	(10)
Random Queries Sample.....	(11)

Introduction

In an ever-evolving aviation landscape, staying ahead requires a deep understanding of customer behavior, operational efficiency, and strategic foresight. As a trusted advisor to the executive management of a leading airline company, our mandate is clear: to delve into the intricacies of current business processes, identify untapped opportunities, and pave the way for expansion.

Our inaugural focus lies on the heart of our operations: flight activity. Understanding the patterns and preferences of our frequent flyers is paramount. Through meticulous analysis, we aim to decipher the nuances of their travel habits: from preferred routes and fare basis to upgrade frequency and loyalty program engagement. Additionally, we'll explore their response to promotional offers, the duration of their stays, and the prevalence of elite status within our customer base.

Integral to this examination is a comprehensive review of our reservation processes. We recognize the imperative of aligning operational efficiency with financial viability. Whether bookings are made through traditional channels or emerging platforms, our aim is to unearth insights that optimize profitability while ensuring seamless customer experiences.

Furthermore, we acknowledge the critical role of customer care interactions. From pre-trip inquiries to post-travel feedback, each interaction presents an opportunity for refinement and enhancement. By harnessing customer insights and addressing concerns proactively, we reinforce our commitment to excellence and bolster brand loyalty.

Using a Kimball-style dimensional modeling approach for your airline's data analysis provides several benefits that align well with your objectives of understanding customer behavior, operational efficiency, and strategic foresight.

The below logical model serves as the bedrock upon which our exploration unfolds, encapsulating the intricacies of flight activity, reservation processes, and customer care interactions. By encapsulating the entirety of our business operations, it ensures a holistic understanding and facilitates informed decision-making by the executive management.

Through this model, we aim to not only uncover existing patterns and inefficiencies but also to illuminate pathways for growth and innovation. By delving deep into the core of our business, we pave the way for strategic enhancements that resonate across all facets of our operations.

As we navigate through the following sections, it becomes evident that our analysis transcends isolated components; rather, it encompasses the interconnectedness of our business processes, offering a comprehensive view that fosters synergy and optimization.

- **Business Process**

- Reservation
- Flights Activity
- Frequent Flyers
- Customer Care
- Hotel

- **Grain Level**

- Reservation Fact: Segment Level
- Flight Activity Fact: Segment Level
- Hotel Fact: Reservation Level
- Frequent Flyer Fact: Trip Level
- Customer Care Fact: Interaction Level

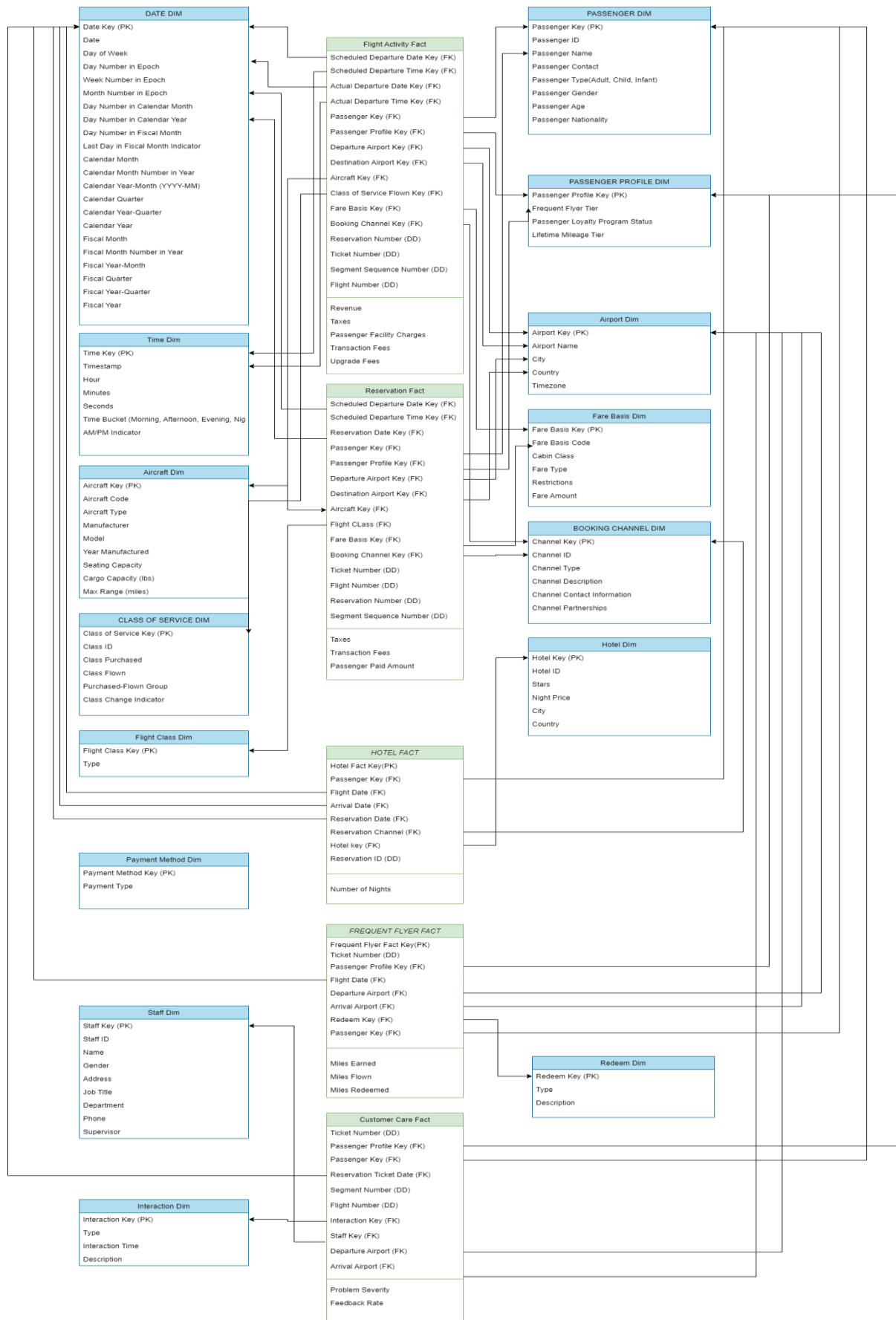
- **Dimensions Tables**

- Date Dimension
- Time Dimension
- Aircraft Dimension
- Airport Dimension
- Booking Channel Dimension
- Class of Service Flown Dimension
- Fair Basis Dimension
- Flight Class Dimension
- Hotel Dimension
- Interaction Dimension
- Passenger Dimension
- Passenger Profile Dimension
- Payment Method Dimension
- Redeem Dimension
- Staff Dimension

- **Fact Tables**

- Reservation Fact
- Flight Activity Fact
- Hotel Fact
- Frequent Flyer Fact
- Customer Care Fact

Full Logical Model

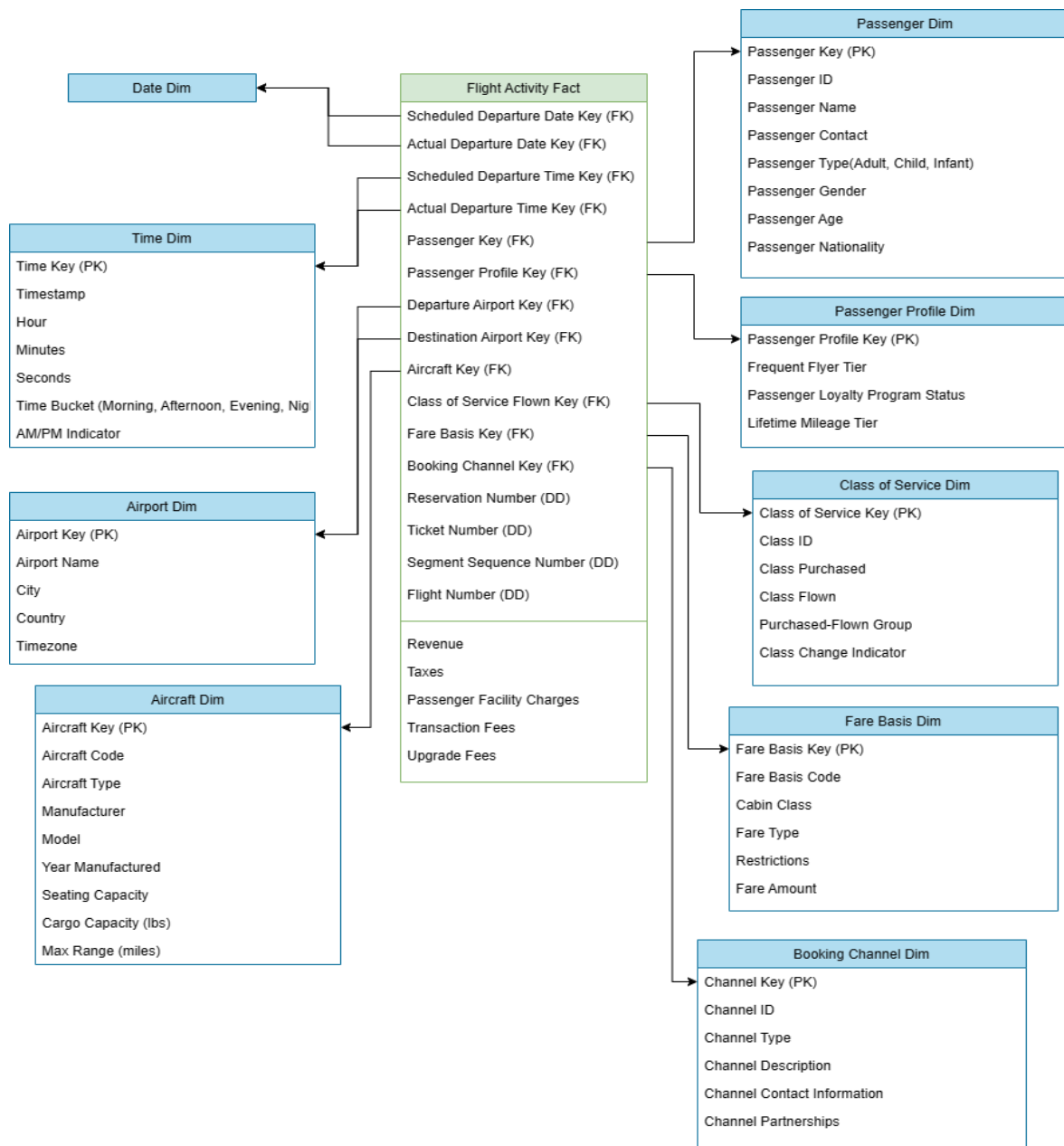


Flight Activity Fact

The Flight Activity fact table records details of flight activities, encompassing scheduled and actual departure dates and times, passenger information, departure and destination airports, aircraft, class of service flown, booking channel, fare basis, ticket number, flight number, reservation number, and segment sequence number. Additional measures such as taxes, transaction fees, revenue, passenger facility charges, and upgrade fees are also captured. Analysis of this data enables airlines to monitor flight operations, revenue generation, and passenger activity. Insights from analysis guide decisions on scheduling, revenue management, and service enhancements to optimize operational efficiency and passenger satisfaction while maximizing profitability.

Logical data model in details: Using Star schema to represent the model.

Grain: Segment



Reservation Fact

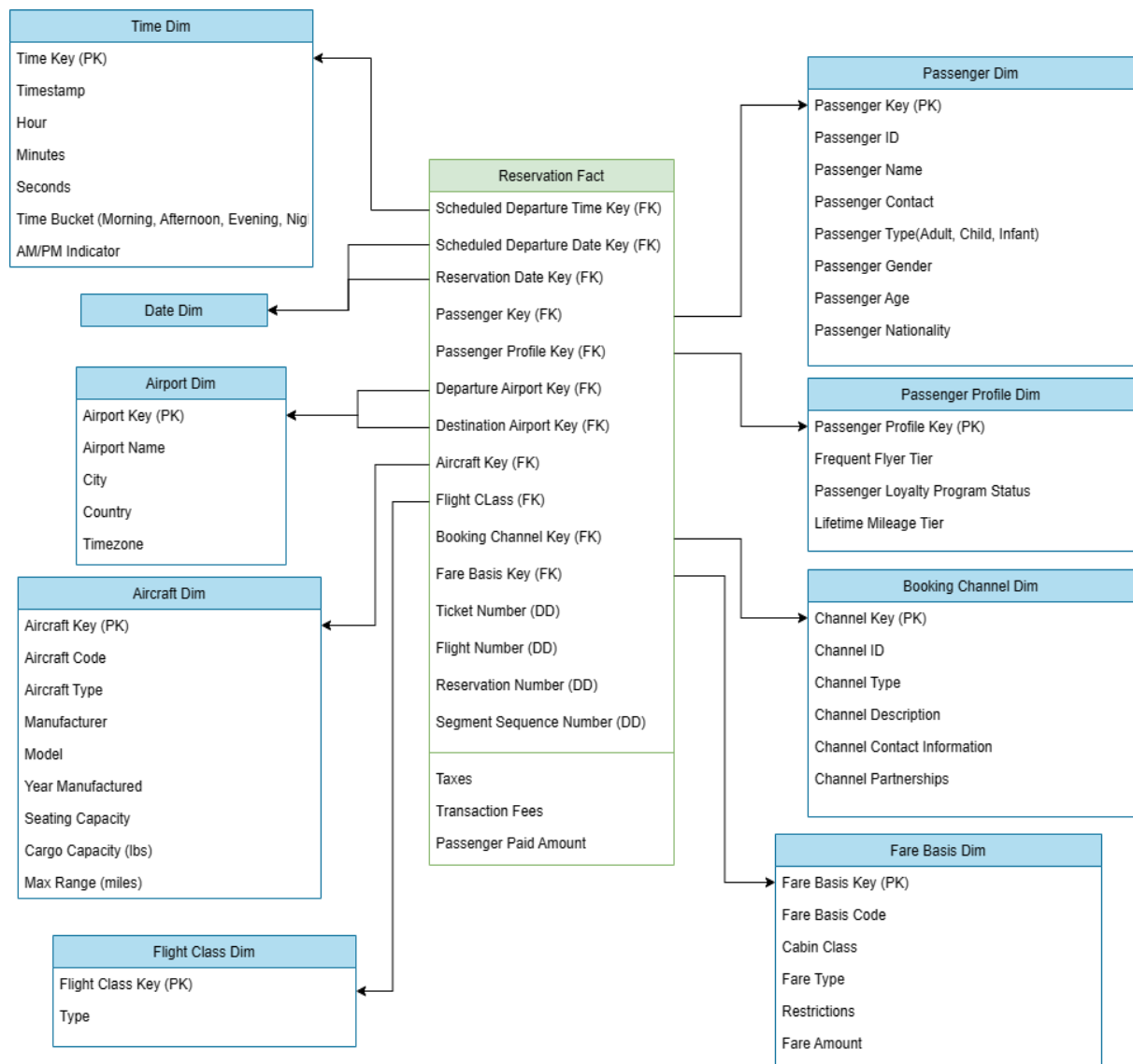
Passengers make flight reservations through various channels. The Reservation fact table captures details of these reservations.

Recorded information includes scheduled departure time and date, reservation date, passenger details, departure and destination airports, aircraft, flight class, booking channel, fare basis, ticket number, flight number, reservation number, and segment sequence number.

Additional measures such as taxes, transaction fees and passenger-paid amount are also recorded. Analysis of this data allows airlines to understand booking patterns, track revenue, and optimize operations. Insights gained from analysis inform decisions related to pricing strategies, capacity management, and customer service enhancements.

Logical data model in details: Using Star schema to represent the model.

Grain: Segment



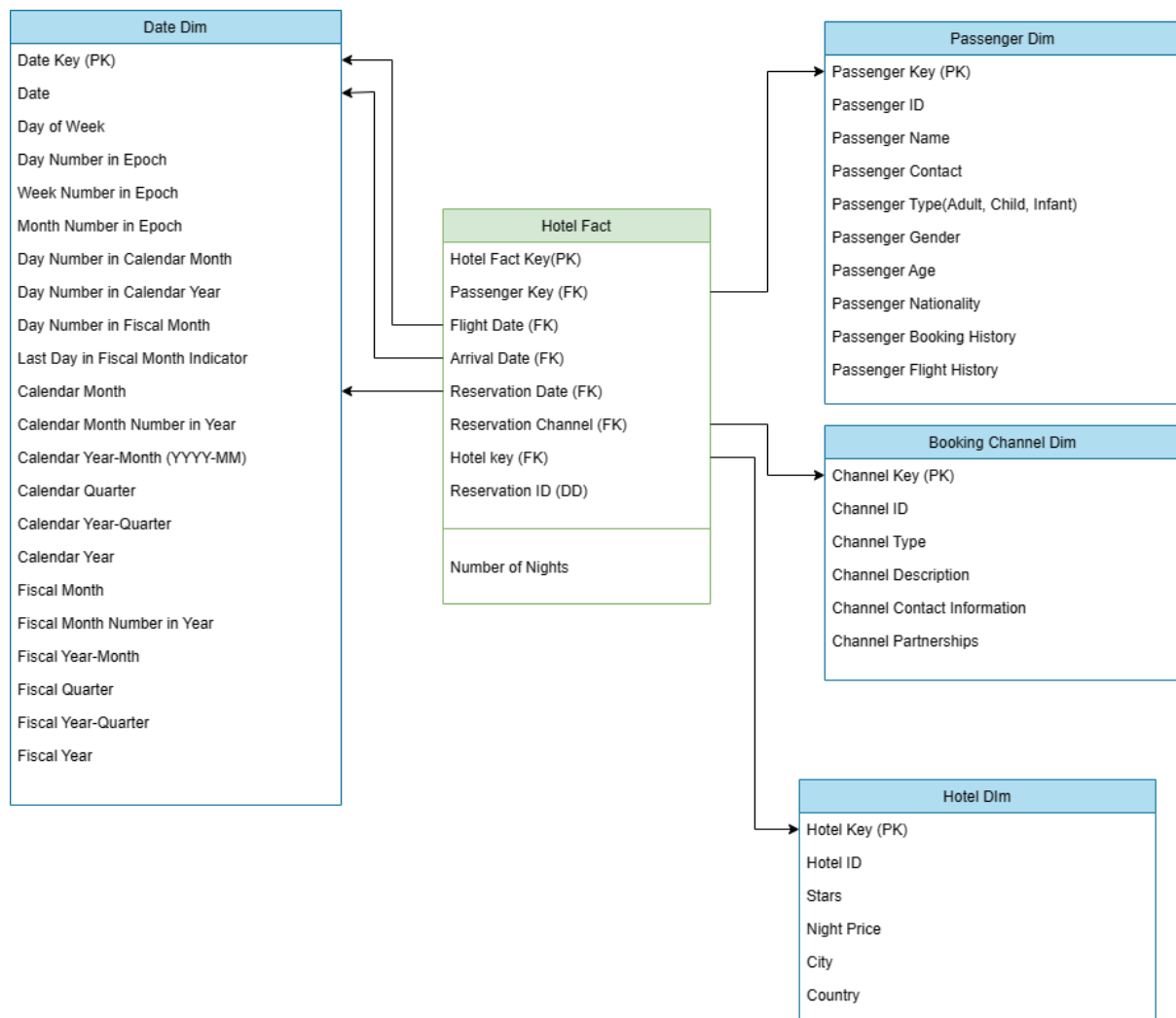
Hotel Fact

Passengers make reservations for hotels in advance of their travel dates. The reservation process involves selecting a hotel, specifying the dates of stay, and choosing a reservation channel.

Reservation details are stored, including the passenger, flight date, reservation date, reservation channel, hotel information, and the duration of the stay. This information allows analysis of hotel booking patterns, such as popular booking channels, average length of stay, and booking trends over time.

Logical data model in details: Using Star schema to represent the model.

Grain: Hotel Reservation



Frequent Flyer Fact

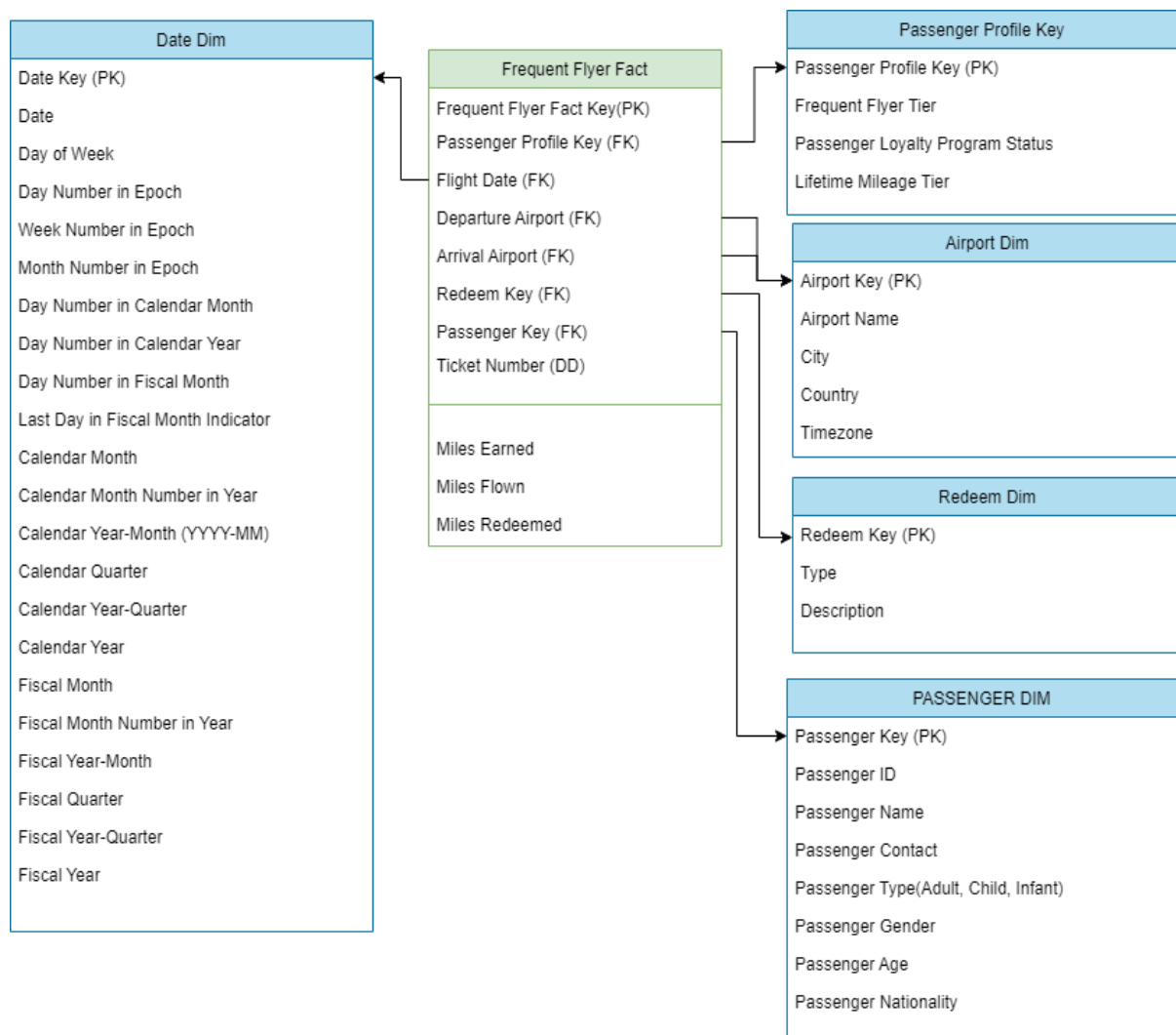
Passengers participate in frequent flyer programs offered by airlines to earn rewards based on their travel activity. The Frequent Flyer fact table captures key metrics related to a passenger's participation in the frequent flyer program.

This includes tracking information such as flight dates, departure and arrival airports, miles earned, miles flown, and miles redeemed. Passengers accumulate miles or points for each flight they take, which can be redeemed for rewards such as free flights, upgrades, or other benefits.

The fact table serves as a central repository for this data, enabling airlines to analyse passenger behaviour, track program effectiveness, and tailor rewards to meet customer preferences. Insights gained from analysing this data can inform marketing strategies, customer retention efforts, and the overall management of the frequent flyer program to enhance customer satisfaction and loyalty.

Logical data model in details: Using Star schema to represent the model.

Grain: Flight



Customer Care Fact

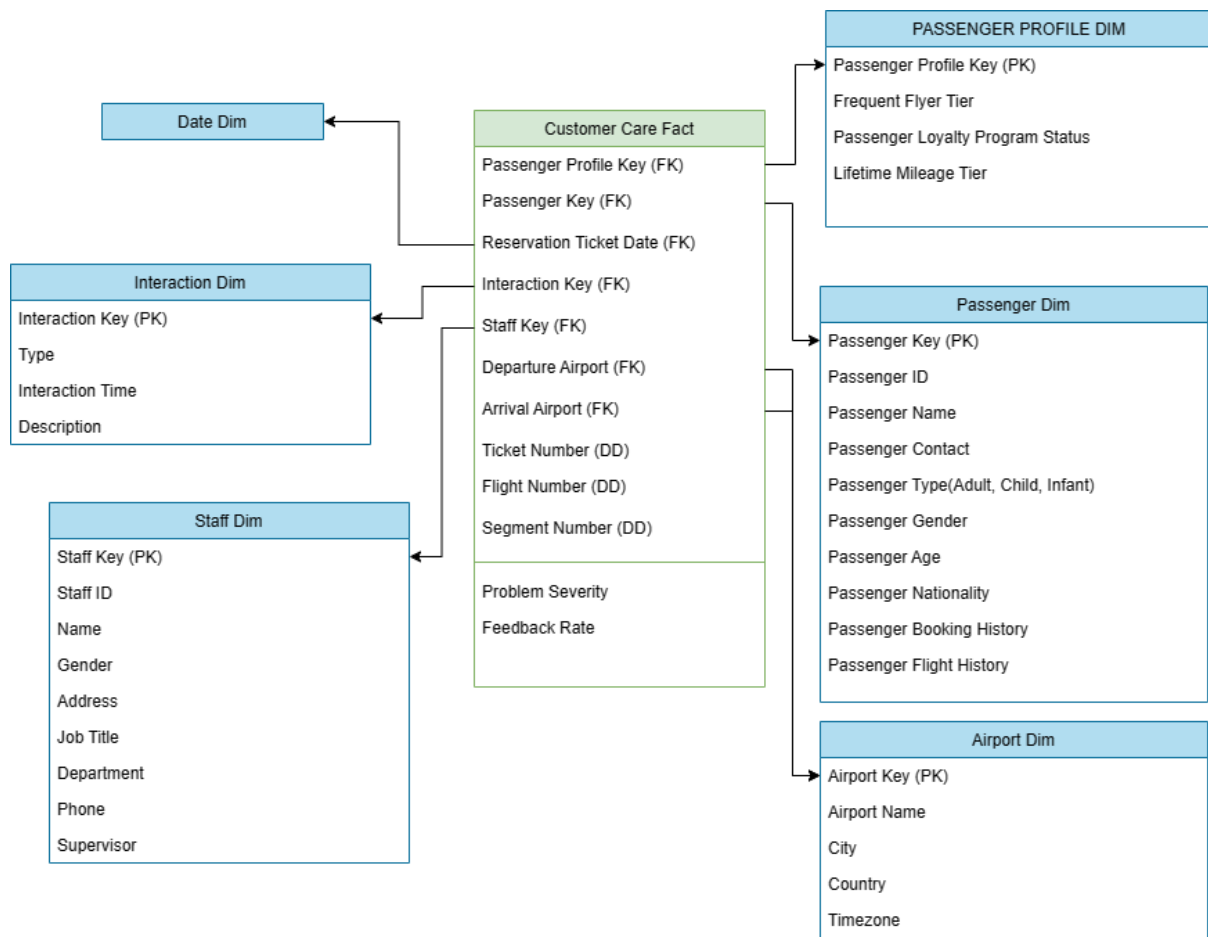
Passengers engage with customer care services for assistance with travel-related issues. The Customer Care fact table captures details of these interactions.

Recorded information includes ticket number, passenger details, reservation ticket date, airport details, flight segment, interaction type, staff involved, problem severity, and feedback rating.

Data analysis identifies recurring issues, measures problem resolution effectiveness, and informs improvements to enhance customer satisfaction.

Logical data model in details: Using Star schema to represent the model.

Grain: Interaction



Sample Business Queries

This query retrieves detailed information about passengers including their lifetime mileage tier, total miles earned, total miles flown, and their rank based on miles flown. It joins the Frequent_Flyer_Fact table with the Passenger_Dim table and the Passenger_Profile_Dim table to get the necessary information and aggregates the data accordingly.

```

1 SELECT F.PASSENGER_KEY, P.PASSENGER_FIRSTNAME, P.PASSENGER_LASTNAME, P.PASSENGER_NATIONALITY, PP.LIFETIME_MILEAGE_TIER, SUM(f.Miles_Earned) AS Total_Miles_Earned, SUM(f.Miles_Flown) AS Total_Miles_Flown,
2 RANK() over (ORDER BY SUM(f.Miles_Flown) DESC) AS Rank_Flyer
3 FROM Frequent_Flyer_Fact f
4 JOIN Passenger_Dim p on F.PASSENGER_KEY = P.PASSENGER_KEY
5 JOIN Passenger_Profile_Dim pp on F.PASSENGER_PROFILE_KEY = PP.PASSENGER_PROFILE_KEY
6 GROUP BY F.PASSENGER_KEY, P.PASSENGER_FIRSTNAME, P.PASSENGER_LASTNAME, P.PASSENGER_NATIONALITY, PP.LIFETIME_MILEAGE_TIER;

```

PASSENGER_KEY	PASSENGER_FIRSTNAME	PASSENGER_LASTNAME	PASSENGER_NATIONALITY	LIFETIME_MILEAGE_TIER	TOTAL_MILES_EARNED	TOTAL_MILES_FLOWN	RANK_FLYER
1831	Ella	Carter	Scotland	1,000,000-1,999,999 miles	440	18895	1
976	Oliver	Ross	England	2,000,000-2,999,999 miles	728	18265	2
753	Dominic	Bates	Australia	100,000-499,999 miles	307	18162	3
1789	Alexa	Garcia	Scotland	100,000-499,999 miles	1283	15943	4
1428	Ariana	Cruz	England	1,000,000-1,999,999 miles	625	15803	5
1158	William	Thompson	Germany	2,000,000-2,999,999 miles	1241	14897	6
1736	Allison	Johnson	Scotland	Under 100,000 miles	611	14850	7
1905	Benjamin	Parker	Northern Ireland	1,000,000-1,999,999 miles	785	14569	8
46	Benjamin	Jackson	England	100,000-499,999 miles	898	14479	9
835	Leila	Grant	France	100,000-499,999 miles	1313	13907	10

On this query, we show the total number of nights spent by each passenger through different booking channels. It joins the Hotel_Fact table with the Booking_Channel_Dim table to get the booking channel descriptions and aggregates the data accordingly.

```

9 SELECT H.Passenger_Key, BC.CHANNEL_DESCRIPTION,
10 SUM(Number_of_Nights) AS Total_Nights_Spent
11 FROM Hotel_Fact H
12 JOIN Booking_Channel_Dim BC ON H.RESERVATION_CHANNEL_KEY = BC.CHANNEL_KEY
13 GROUP BY H.Passenger_Key, BC.CHANNEL_DESCRIPTION;

```

PASSENGER_KEY	CHANNEL_DESCRIPTION	TOTAL_NIGHTS_SPENT
1957	Limited Booking	8
1744	Conference Reservations	1
947	Media Booking	5
66	Magazine Reservations	15
580	Event Booking	1
856	Review Booking	8
1434	Website	5
602	Web Booking	16
1701	Business Website	10
232	Social Booking	2
1843	Lab Reservations	5
1015	Mobile App	15

This query provides insights into the count of reservations made through each booking channel, ordered by the number of reservations in descending order. It joins the Hotel_Fact table with the Booking_Channel_Dim table to get the booking channel descriptions and counts the occurrences of each channel description.

17	
18	▶ <code>SELECT BC.CHANNEL_DESCRIPTION, Count(BC.CHANNEL_DESCRIPTION) Reservation_Count</code>
19	<code>FROM Hotel_Fact H</code>
20	<code>JOIN Booking_Channel_Dim BC ON H.RESERVATION_CHANNEL_KEY = BC.CHANNEL_KEY</code>
21	<code>GROUP BY BC.CHANNEL_DESCRIPTION</code>
22	<code>order by Reservation_Count desc;</code>
23	
24	

Data Grid	
<div> Call Stack DBMS Output (disabled) CodeXpert Breakpoints Watches Profiler REF CURSOR Results Script </div>	
<div> Cancel </div>	
CHANNEL_DESCRIPTION	RESERVATION_COUNT
▶ Mobile Application	2543
Web Booking	1728
Travel Agent	1716
Military Reservations	908

This query provides insights into the count of interactions categorized by problem severity level and interaction type. It joins the Customer_Care_Fact table with the Interaction_Dim table to get the interaction types and counts the occurrences of each combination.

26	
27	• <code>SELECT C.Problem_Severity, I.TYPE,</code>
28	<code>COUNT(*) AS Interaction_Count</code>
29	<code>FROM Customer_Care_Fact C</code>
30	<code>JOIN Interaction_Dim I on C.INTERACTION_KEY = I.INTERACTION_KEY</code>
31	<code>GROUP BY C.Problem_Severity, I.TYPE</code>
32	<code>ORDER BY Interaction_Count DESC;</code>
33	
34	

Data Grid		
<div> Call Stack DBMS Output (disabled) CodeXpert Breakpoints Watches Profiler REF CURSOR Results Script Output Data Grid </div>		
<div> Cancel </div>		
PROBLEM_SEVERITY	TYPE	INTERACTION_COUNT
▶ 4	In Person	62
1	Support Ticket	62
4	Support Ticket	59
2	In Person	58
5	Support Ticket	54
4	Phone Call	52
2	Phone Call	51
5	In Person	51
3	Phone Call	50
5	Phone Call	50

This query provides insights into the count of flights for each combination of departure and destination airports (flights the company's flyers mostly take). It joins the Flight_Activity_Fact table with the Airport_Dim table twice to retrieve the names of the departure and destination airports and counts the occurrences of each combination. Finally, it orders the result set by the flight count in descending order.

35	
36	SELECT count(FA.PASSENGER_KEY) Flight_Count , AR.AIRPORT_NAME AS Departure_Airport, A.AIRPORT_NAME AS Destination_Airport
37	FROM Flight_Activity_Fact FA
38	JOIN Airport_Dim AR on FA.DEPARTURE_AIRPORT_KEY = AR.AIRPORT_KEY
39	JOIN Airport_Dim A on FA.DESTINATION_AIRPORT_KEY = A.AIRPORT_KEY
40	GROUP BY AR.AIRPORT_NAME, A.AIRPORT_NAME
41	ORDER BY Flight_Count DESC;
42	
43	
44	

FLIGHT_COUNT	DEPARTURE_AIRPORT	DESTINATION_AIRPORT
9	Sharm El Sheikh International Airport	Abu Simbel Airport
7	Abu Simbel Airport	Borg El Arab Airport
7	Hong Kong International Airport	Sydney Airport
6	Assiut Airport	Paris Charles de Gaulle Airport
6	Tokyo Haneda Airport	Mersa Matruh Airport
6	Abu Simbel Airport	Marsa Alam International Airport
6	Toronto Pearson International Airport	Sydney Airport
6	Abu Simbel Airport	Sohag International Airport
5	Dubai International Airport	Abu Simbel Airport
5	Assiut Airport	Asyut International Airport
5	Asyut International Airport	San Francisco International Airport
5	Paris Charles de Gaulle Airport	John F. Kennedy International Airport
5	Zurich Airport	El Arish International Airport

This query provides insights into the total revenue generated from flights, categorized by frequent flyer tier, aircraft type, and destination country, and city. It joins multiple dimension tables with the fact table to gather relevant information and aggregates the data accordingly. Finally, it orders the result set by total revenue in descending order.

45	
46	SELECT SUM(FA.BASE_FARE_REVENUE) as Total_Revenue, PP.FREQUENT_FLYER_TIER, AR.AIRCRAFT_TYPE, AP.COUNTRY, AP.CITY
47	FROM Flight_Activity_Fact FA
48	JOIN Passenger_Profile_Dim PP on FA.PASSENGER_PROFILE_KEY = PP.PASSENGER_PROFILE_KEY
49	JOIN Aircraft_Dim AR on FA.AIRCRAFT_KEY = AR.AIRCRAFT_KEY
50	JOIN Airport_Dim AP on FA.DESTINATION_AIRPORT_KEY = AP.AIRPORT_KEY
51	GROUP BY PP.FREQUENT_FLYER_TIER, AR.AIRCRAFT_TYPE, AP.COUNTRY, AP.CITY
52	ORDER BY Total_Revenue DESC;
53	
54	

TOTAL_REVENUE	FREQUENT_FLYER_TIER	AIRCRAFT_TYPE	COUNTRY	CITY
4142.12	WarriorTier	Wide-body	Egypt	Cairo
3529.76	WarriorTier	Regional jet	United Arab Emirates	Dubai
3509.69	Basic	Regional jet	Switzerland	Zurich
3506.56	Basic	Regional jet	Canada	Toronto
3350.77	WarriorTier	Regional jet	France	Paris
3103.61	WarriorTier	Turboprop	Switzerland	Zurich
3092.93	MidTier	Turboprop	Egypt	St. Catherine
3006.63	Basic	Regional jet	Egypt	Mersa Matruh
2986.83	Basic	Turboprop	Singapore	Singapore
2985.25	MidTier	Turboprop	Egypt	Borg El Arab
2975.26	WarriorTier	Turboprop	Egypt	Assiut
2920.69	Basic	Turboprop	United States	New York City
2904.35	Basic	Regional jet	Australia	Sydney
2850.24	WarriorTier	Regional jet	Australia	Sydney