DATA WAREHOUSE
SOLUTION FOR AN AIRLINE COMPANY

Diane Haiden

January 2025

1.0 EXECUTIVE SUMMARY

This project presents a business intelligence (BI) and data warehouse solution for an airline company, focusing on their reservation system. Some airlines use software that is limited in analyzing customer data and satisfaction metrics. To remain competitive, a new application and database system is needed. Southwest Airlines and Allegiant use modern platforms like AWS and Snowflake to their advantage.

Current Limitations: Existing flight operations solutions is specialized but limited in scope.

Need for Upgrade: A new system is necessary for better data management and competitiveness.

Proposed Solution: Adoption of AWS or Snowflake for a central data hub.

Benefits:

- Improved decision-making with actionable insights
- Enhanced customer experience through segmentation and analytics
- Better data quality
- Faster query performance
- Streamlined reporting
- Scalability

The project follows a structured four-step process for designing a data warehouse system, as proposed by Taniar & Rahayu (2022), ensuring a refined and efficient data model tailored for an airline company. This includes updating the star schema, populating fact and dimension tables, and executing SQL commands for implementation.

The execution steps first involve implementing the ETL (Extract, Transform, Load) workflow, where data is extracted from the source and loaded into the data warehouse, and accurate variable transfer is ensured. Following this, SQL commands are used to populate tables and capture both row counts and listings.

The report components also include updates to the business process, star schema, screenshots of row counts and table listings, and a comprehensive listing of SQL commands used. This approach aims to enhance services, improve customer satisfaction, support growth, and potentially increase profitability for an airline company.

Table of Contents

2.0 BUSINESS PROCESS

An airline OLTP system primarily manages transactions between customers and an airline company. An airline OLTP system can contain many systems that manage flights, reservations, customer service, customer information, etc. Our project focuses on flight and route analysis, aircraft utilization, passenger and customer insight, agent and service insight, transaction analysis, and operational insight.

- Flights – this business process manages crew, flight routes, and flight schedules.
- Aircraft – this business process manages aircraft-specific data
- Passenger – this business process manages passenger details, including the routes taken
- Agent/Customer Service – this process manages customer complaints and inquiries
- Reservations – this business process manages customer details, fares, schedules, seat availability, payment information, cancelations, and other changes.

3.0 BUSINESS QUESTIONS

Flights

1. What are the most frequent routes traveled by each customer?
2. Which flight routes generate the highest revenue?
3. Which flight routes are underperforming in passenger count or terms of revenue?
4. How many flights operate daily on each flight route?
5. Which flights experience the most delays?
6. How are flight schedules adjusted during unexpected events (i.e. bad weather)?
7. How does the frequency of flights change over the holidays or peak season?

Aircraft

1. What is the average number of flights operated by each aircraft per day or week?
2. How efficient is the usage of each aircraft in terms of flight occupancy and frequency

Passenger/Customer

1. How many passengers travel on each flight route on average?
2. Who are the most frequent passengers, and what routes do they prefer?
3. What are the common customer inquiries related to specific flights or routes?
4. How many passengers utilize particular services (e.g., meals, Wi-Fi, priority boarding)?

Customer Service

1. What are the most common complaints made by customers?
2. Are the complaints isolated by segment (i.e. no wifi on short flights)?
3. How is customer satisfaction measured?
4. What are the most requested services by customers?
5. What services contribute the most to customer satisfaction (i.e. free wifi)?

Reservations

1. How can we make seat availability more accurate and accessible for customers?
2. What patterns can we see in customers making reservations?
3. What preferences are consistently being selected for each customer?
4. How often do customers change their reservations?

4.0 GRAIN OF THE STAR MODEL

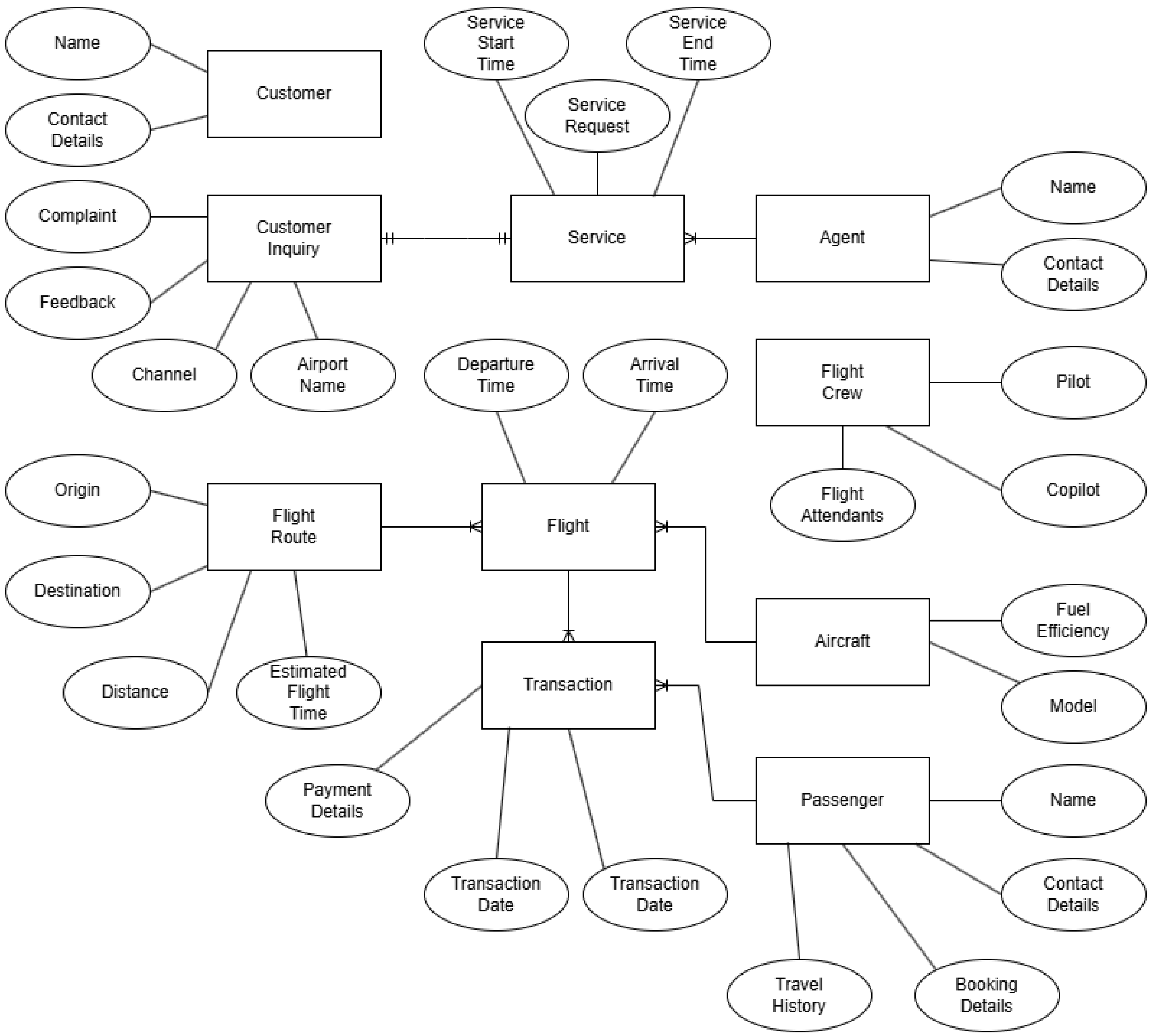
Flight Granularity
To optimize the flight business process, the most detailed grain is the combination of flight routes (e.g. LAX to JFK), aircraft, fuel consumption per route, departure time, arrival time, flight attendants, pilot, and copilot.

Reservations Granularity
To optimize the reservations business process, granularity can be broken down into three groups: flight, passenger, and transaction. At the flight level, the most detailed grain is the combination of a flight number, seat availability, and departure and arrival times. At the passenger level, the most detailed grain is the combination of a passenger's personal information, their booking details, and their travel history. At the transaction level, the most detailed grain is the combination of payment details, transaction dates, and ticket prices.

Customer Services Granularity
To optimize the customer services business process, the most detailed grain is the combination of customer inquiries, complaints, feedback, channels (phone, online chat, and in-person), agent, customer details, service request, service time and date from start to finish, and airport name.

By defining the granularity at these levels, airlines can gain more insights and control over each business process. This will help the airline company make more informed business decisions.

5.0 ENTITY RELATIONSHIP DIAGRAM OF SOURCE OLTP



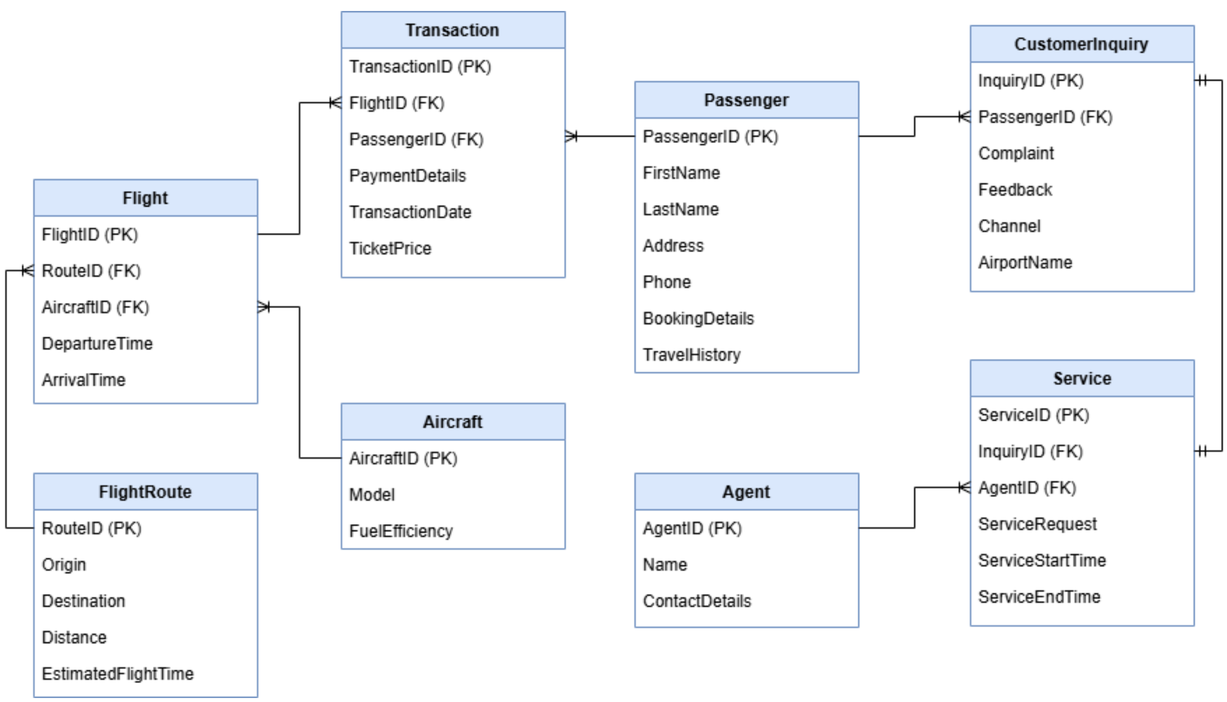
6.0 DIMENSION TABLES, VARIABLES, AND ATTRIBUTES

The airline data model has 5 dimensions which include Flight, FlightRoute, Aircraft, Passenger, and CustomerInquiry. The Flight dimension contains a FlightID (primary key), RouteID (foreign key to the FlightRoute dimension), AircraftID (foreign key to the Aircraft dimension), DepartureTime, and ArrivalTime. The FlightRoute dimension contains a RouteID (primary key), origin, destination, distance, and the EstimatedFlightTime. The Aircraft dimension contains an AircraftID (primary key), Model, and FuelEfficiency. The Passenger dimension contains the PassengerID (primary key), Name, ContactDetails, BookingDetails, and TravelHistory. The CustomerInquiry dimension contains the InquiryID (primary key), Complaint, Feedback, Channel, and the AirportName.

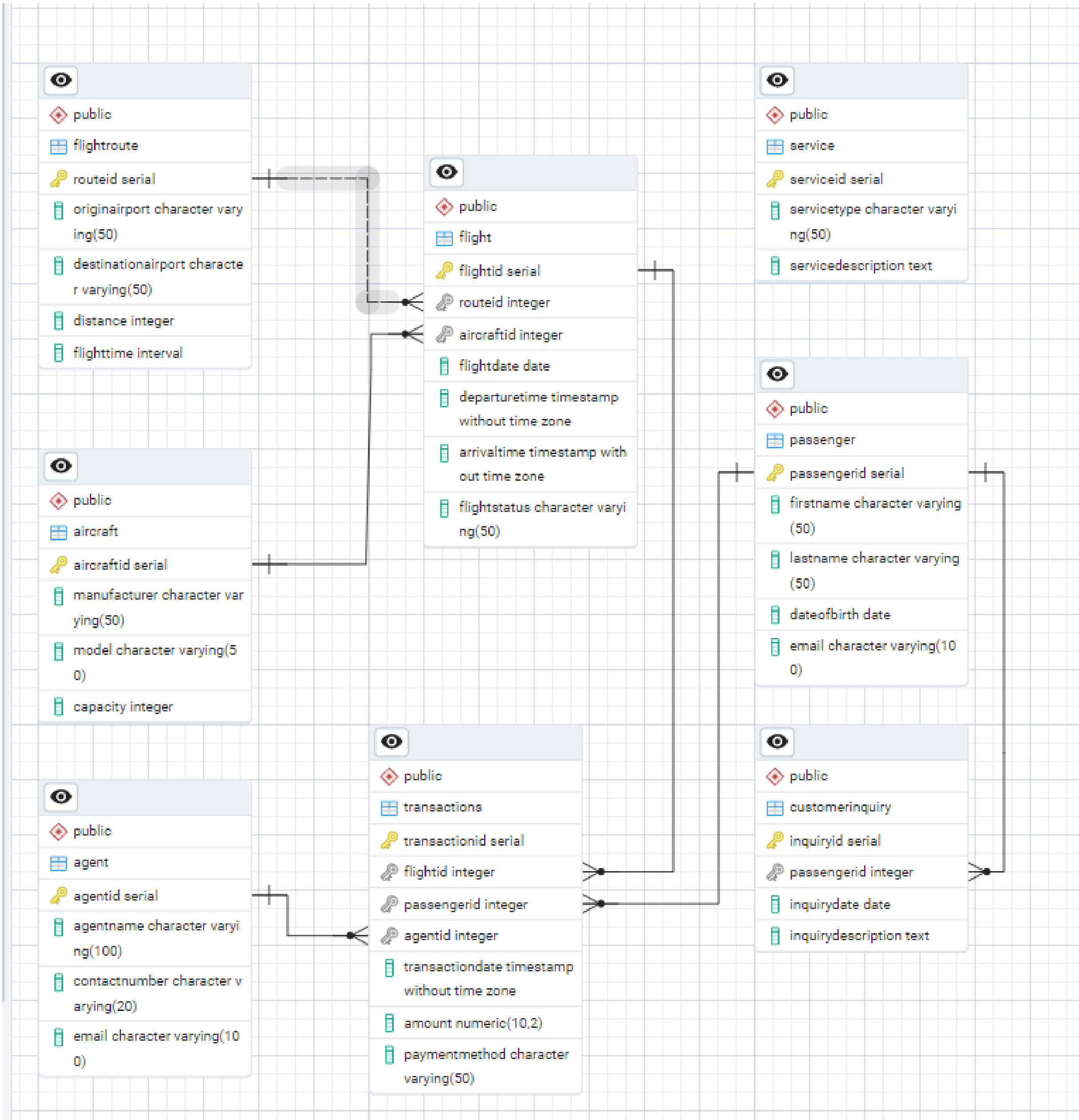
7.0 FACTS TABLE, VARIABLES, AND ATTRIBUTES

The airline data model contains a Transaction Fact table that captures measurable data. Its attributes include a TransactionID (primary key), FlightID (foreign key to the Flight dimension), PassengerID (foreign key to the Passenger dimension), PaymentDetails, TransactionDate, and TicketPrice.

8.0 STAR MODEL - CONCEPTUAL



9.0 FINAL STAR MODEL - PHYSICAL



10.0 CONCLUSION

This project is intended to deliver a simple data warehouse solution, addressing the critical need for improved data management and analytics capabilities across the airline industry. By adopting advanced platforms, airlines can access a centralized data hub that facilitates better decision-making, enhances customer satisfaction, and streamlines operations. This project also demonstrated a clear methodology for building similar systems in the future, making the process easier and less risky. Overall, this initiative sets the foundation for airlines to remain competitive in an industry where the smart use of data is paramount.

11.0 REFERENCE

Taniar, D., & Rahayu, W. (2022). *Data warehousing and analytics: Fueling the data engine.* Springer Nature. ISBN: 9783030819781