MODELLING AIRLINE COMPANY

Data Warehouse Project

Datawarehouse Fundamentals Course

ITI Intake 43 | Team 1 (Mariam – Khalid – Amira)

Table of Contents

I. Introduction

A. Problem Description

A major airline company decided to hire you in order to assist the executive management to **analyze their current business processes and expand the company by discovering new opportunities.**

- Executives decided that the first deliverable should focus on the flight activity in order to ensure good ongoing business process.
- The marketing department wants to analyze what flights the company's frequent flyers take, what fare basis they pay, how often they upgrade, how they earn and redeem their frequent flyer miles, whether they respond to special fare promotions, how long their overnight stays are, and what proportion of these frequent flyers have gold, platinum or titanium status.
- Your analysis shall also include the reservation process where finance team will be interested in analyzing the company profit. Note that reservation processes can take place through multiple channels.
- Airline company also provides customer care interaction before, within and after your trip in order to handle customer inquiries, complaints and keep their feedback for business enhancements. Analysis should include interaction type and problem severity (if issue exists)

State the processes for your company then apply your modeling technique/s with demonstrative layering methodology to **design the logical and physical design** to support such kind of analysis for decision support.

B. Modelling Technique

- **Data Vault** was chosen as a way to store the Airline System Data, then Dimensional Modeling was built in order to report and query and aggregate the data from the system.

We choose Data vault for two main reasons:

1. The data of the passenger in every change in one entity, the whole tables is written again.
2. The parallel loading increases the performance and the optimization

C. The Modelling Process

1. The aim is to analyze the Airline business processes and expand the company by discovering new opportunities
2. The Business has three main processes, the **grain** will be measured as:
   - Reservation fact, it will be measure it per transaction.
   - Costumer Care fact, it will be measured per transaction.
   - Flight activity fact, it will be measured per day.

3. There are sixteen-dimension table in three different processes:

Booking_Channel	Represent the way that passenger use to book, if online or from an agent.
Airport	Represents an information about the airport like its name, ID and Code.
Flight	Represents the information about the Flight, its ID, the Depature_time, Arrival_Time, duration and if the flight is cancelled.
Location	Represents the information about the airport location, its ID, the city, state, country, Latitude and Longitude.
Route	Represents the information about route of the flight like the origin airport, Destination_Airport and the distance.
Airplane	Represents the information about route airplanes like the airplane name and the model number, Manufacture_code, Capacity, and the Weight.
Flight_Type	Represents the information about the flight if its non-stop flight or connected flight and no of stops of the connected flights with the number of layover hours.
d_date	Represent the date, date of week, date of month, date of year, month number, quarter, year.
Interaction	Represent the ID of the interaction and its type of time like, before or after or within the flight.
Survey	Represent the way that customer gives his/her feedback, the type whether its online by email or offline as a paper, and the category of the survey like the type of services that the customer give it a score.
Fare	Represents the fare of the ticket, how the fare is calculated like Base_fare and taxes and if there is an Extra Fees.
Booking_Class	Represents the booking class ID and its type whether it's an economy with its types and how much the price with increase by percentage or Business class with its type and how much the price with increase by percentage or first class with its type and how much the price with increase by percentage by calculating Class_Fare_ Per%.
Ticket	Represents the Ticket ID, status(Paid or unpaid) and its type (e-ticket or paper) and the class of his ticket.
Passenger	Represent the information about each Passenger, such as Passenger ID, first name, last name, Gender, date of birth, phone number, address, and the status as if this passenger is a frequent flyer or not.
Frequent Flyer	Represent the information about each Frequent Flyer as we see the Frequent Flyer as a service that the passenger must apply for. It includes the Frequent Flyer ID , the day he started to be a Frequent Flyer, how much distance of flights he/she traveled, the amount he paid to buy miles, Total number of miles he/she has, if he/she accepts a promotion (no of promotion) and his/her Home_Airport.

Frequent Flyer category	Represent the information about Frequent Flyer category such ID and the Frequent Flyer type (Silver, Gold, Blue, Elite, Titanium) and min number of miles to earn this type and max number to be in this type.

4. The Fact tables with its measures will be represented below:

Reservation fact	The reservation fact connects the passenger with the flight, the booking Channel, Booking Class, Ticket, Fare, and date. - The fact will measure the total price (fare + booking class percentage) - The fact will measure the Miles earned if the ID is connected to a frequent flyer (number of miles is calculated according to the flight distance+booking class) - The fact will measure the Miles_Redeem (total miles per passenger – the miles used to upgrade) - The fact will measure the NoOfFlightsLayOver (count number of layover flights from the flight type that is connected to the flight table) - The fact will measure the Class_Change_Indicator (if there is a difference between the ticket class and the Booking class, then there is an upgrade)
Costumer Care fact	The Costumer Care fact connects the passenger with the interaction and the survey. - The fact table will measure the score of the survey (by collecting the average of every category in the survey) - The fact table will measure the Problem_severity (if the score of the survey is less then 2)
Flight activity fact	The Flight activity fact connects the flight with the date. - The fact table will measure the Total flights canceled from the flight table (Aggregate the fact by calculating the sum). - The fact table will measure the Total flights Duration from the flight table (Aggregate the fact by calculating the sum). - The fact table will measure the NoOfDelayedFlights from the flight table (Aggregate the fact by calculating the difference in the Schedualed_depature_time and Depature_time or the difference Schedualed_arrival_time and arrival time).

II. Data Vault Model

The following Diagram shows the Data Vault Model for the Airline System:

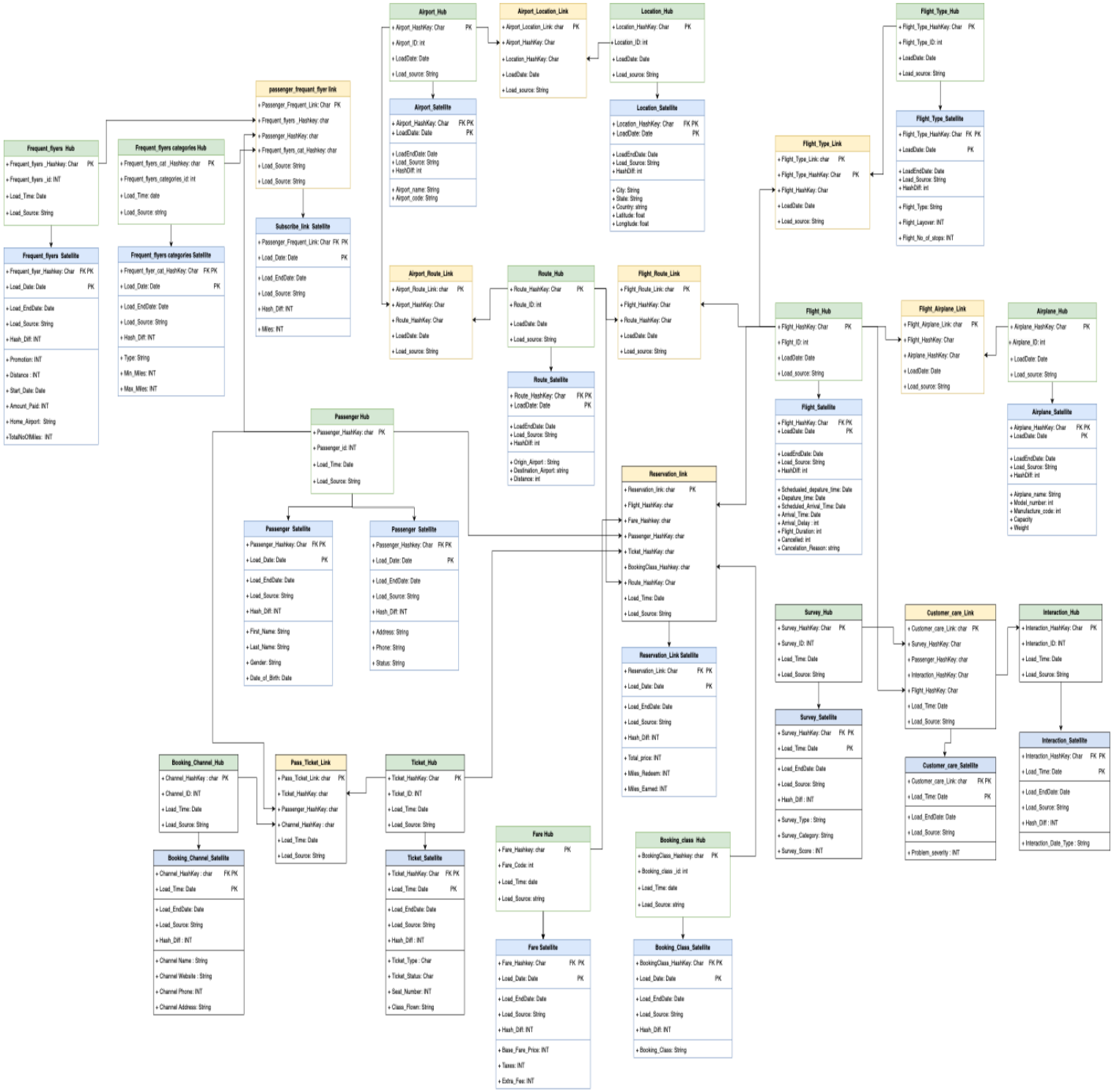

Figure 1: Data Vault Model

III. Dimensional Modeling Model

The following Diagram shows the Dimensional Modelling for the Airline System:

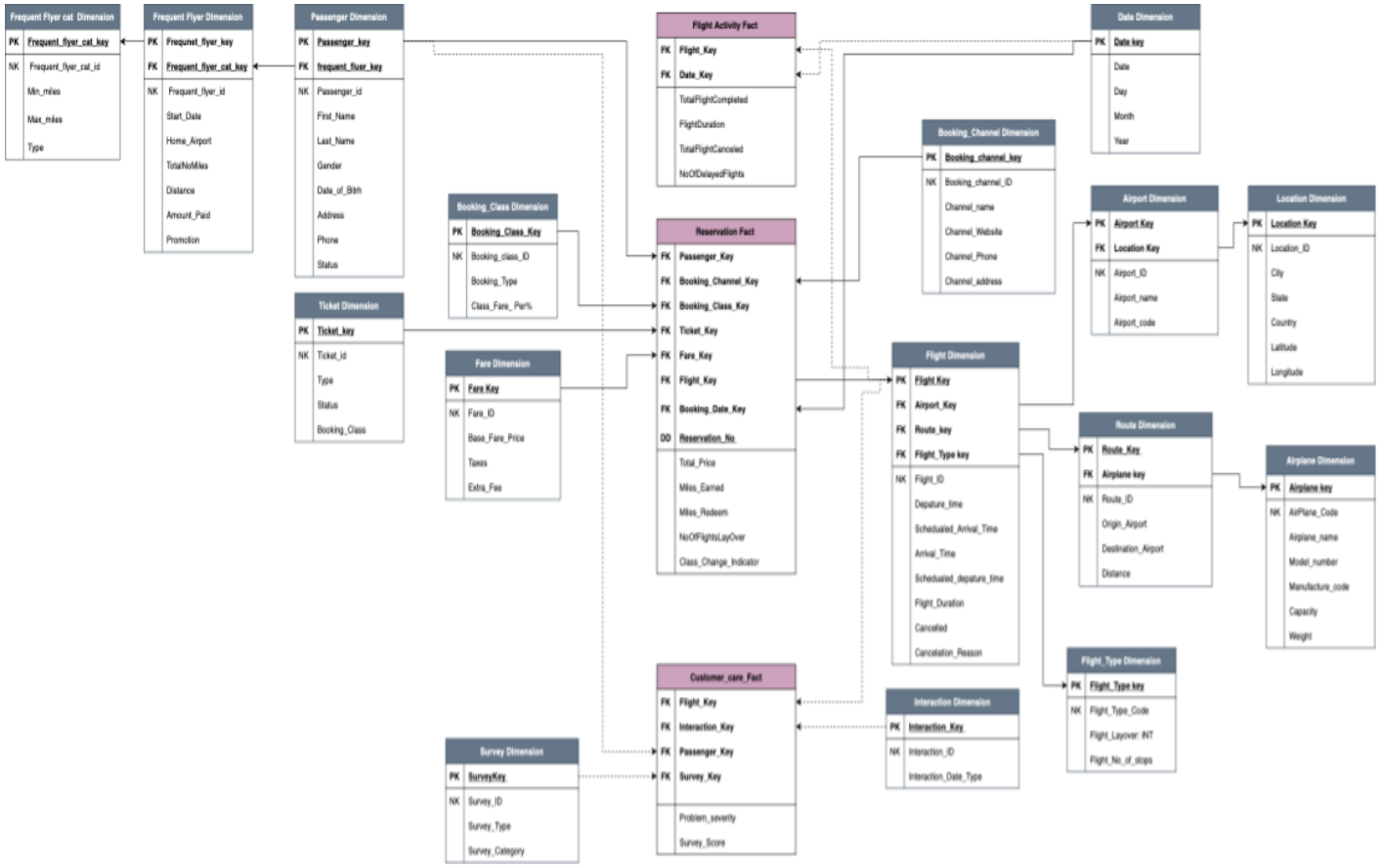

Figure 2: Dimensional Modelling