



Information Technology Institute (ITI)
Smart Village

Analyze And Model Data Warehouse For Airline System Project

Prepared by:

Amr El-saeed Mohamed Ellakany

Donia Ayman Ahmed

Hafsa Ahmed Zahran

Mohab Mahmoud Amro Abd Elghafar

Shahd Mohamed Ali

Presented to:

Ebtehal El-Hossiny

Table of Contents

Contents

Introduction	3
Business Matrix.....	3
Dimensional Modeling Process.....	3
1- Business process	3
2- Granularity	3
3- Dimensions Tables.....	3
4- Facts Tables:	4
Logical Model.....	5
Tables Identifications	6
1- Passengers Dimension.....	6
2- Frequent Flyers Dimension.....	6
3- Flights Dimension.	6
4- Class Upgrade Dimension.....	6
5- Date Dimension.	6
6- Reservation Channels Dimension.....	6
7- Fare Basis Dimension.....	7
8- Hotels Dimension.....	7
9- Interactions Dimension.	7
10- Aircrafts Dimension.	7
11- Airports Dimension.....	7
12- Payment Methods Dimension.....	7
13- Redeem Dimension.	8
14- Staff Dimension	8
15- Flights Activity Fact.....	8
16- Customer Care Fact.	8
17- Frequent Flyers Fact.	9
Physical Model	9
Insert Dummy Data into The Physical Tables	23
Insert The Measured Data into Fact Tables.....	30

SQL Queries to Answer Business Questions:	33
Indexes For the Data Warehousing	38

Introduction

The project's purpose is to analyze the flight activities of some airline companies and their frequent flyers and model its data warehouse schema.

Business Matrix

Facts	Passengers	Frequent Flyers	Class Upgrade	Fare basis	Flights	Airports	Aircrafts	Channels	Payment method	Interactions	Redeem	Staff	Date	Hotels				
Flights Activity Fact	X	X	X	X	X	X	X				X	X	Sement number(DD)	Ticket Number (DD)	Schedualed Departure Time (DD)	Actual Departure Time (DD)	Scheduled Arrival Time (DD)	Actual Arrival Time (DD)
Customer care	X	X		X	X			X		X	X		Sement Number (DD)	Ticket Number (DD)				
Frequent Flyer		X		X	X				X		X			Ticket Number (DD)				

Dimensional Modeling Process

1- Business process

Analyze the Airlines flights' activities.

2- Granularity

Flights Activity Fact is per Leg level.

Customer Care Fact is per Ticket Number level.

Frequent Flyers Fact is per Ticket Number level.

3- Dimensions Tables

- 1- Passengers Dimension.
- 2- Frequent Flyers Dimension.
- 3- Flights Dimension.
- 4- Date Dimension.
- 5- Class Upgrade Dimension.
- 6- Channels Dimension.
- 7- Fare Basis Dimension.

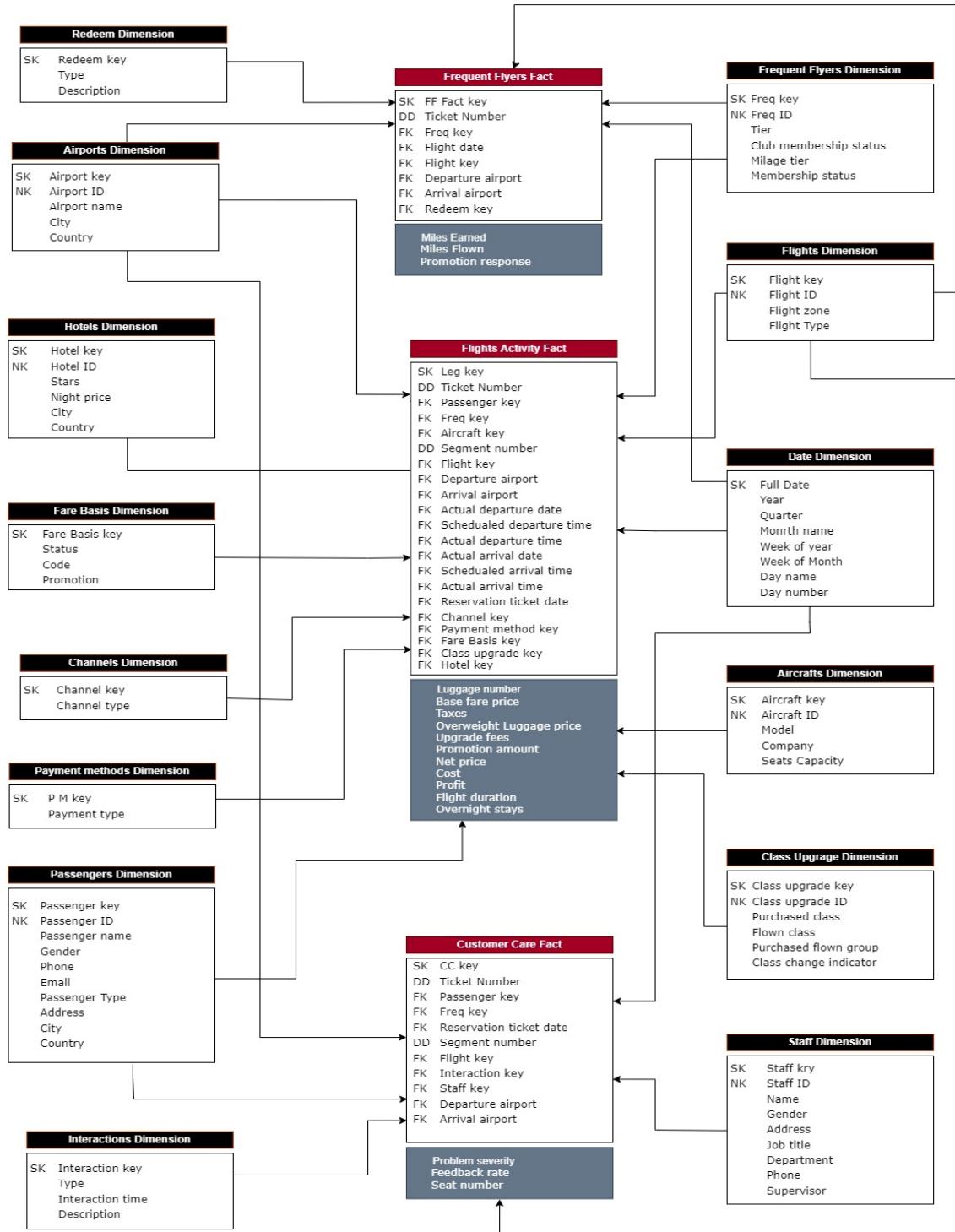
- 8- Hotels Dimension.
- 9- Interactions Dimension.
- 10- Airports Dimension
- 11- Aircrafts Dimension
- 12- Payment methods Dimension
- 13- Redeem Dimension
- 14- Staff Dimension

4- Facts Tables:

- 1- Flights Activity Fact.
- 2- Frequent Flyers Fact
- 3- Customer Care Fact.

Logical Model

Represents the facts and dimensions tables, and the relations between them.



Tables Identifications

1- Passengers Dimension

Represents data about passengers, like **Passenger ID, Name, Gender, Phone Number, email, address, city, country, Passenger Type: Regular or Frequent.**

2- Frequent Flyers Dimension

Represents data about Frequent Flyers, like **Their ID, Tier, Club Membership Status: Non Member, Club Member, Mileage Tier and Membership Status**

As follow:

Under 100,000 miles Silver

500,000-1,999,999 miles Platinum

100,000-499,999 miles Gold

3- Flights Dimension.

Represents data about the flight, like **flight ID, Flight Zone: Domestic or international and Flight type: Non-Stop, Direct, indirect.**

4- Class Upgrade Dimension.

Represents data about the flight, like **ID, Purchased class: Economy, Prem Economy, First, Business and Flown Class, purchased Flown Group Economy-Economy, Economy-Prem Economy,...etc, class change indicator: No change , Upgrade, Downgrade**

5- Date Dimension.

Represents the **Minute, Hour, Day, Month, Quarter, and Year** for each date stored for different business needs.

6- Reservation Channels Dimension.

Represents the channel that the ticket was purchased from. Like **channel Id, channel type** It can be **Website or the agency itself.**

7- Fare Basis Dimension.

Represents if the passenger fully paid the ticket price, or with a discount. It contains **id, code, status**.

This represented using code numbers which:

111 means Full Fare First Class.

112 means Discount Fare First Class.

221 means Full Fare Business Class.

222 means Discount Fare Business Class.

331 means Full Fare Economy Class.

332 means Discount Fare Economy Class.

8- Hotels Dimension.

Represents data about the hotels that the passengers stayed on during the transit flights, like hotel **ID, Stars, Night Price, city and country**.

9- Interactions Dimension.

Represents the passenger's different interactions with customer services, it contains the **Interaction Time** if it was **Before, Within, or After** the flight and the **Interaction Type** if it was **Feedback, Inquiry, or Complaint**.

10-Aircrafts Dimension.

Represents the data about the aircraft, it contains **ID, Model, Company and Seats Capacity**

11-Airports Dimension.

Represents the information about airport, it contains **ID, Name , City and Country**

12-Payment Methods Dimension.

Represents the information about the payment method, it contains **Payment method type: Credit or Cash**

13-Redeem Dimension.

Represents the information about the type of Redeem that customer will change his mile point with, it contains **type and description**.

As follow:

Flight awards	Get a discount on your ticket total price.
Upgrade awards	Get premium cabin upgrades.
Club and status awards	Complimentary amenities and services are available to make your travel more productive and relaxing.

14-Staff Dimension

Represents the information about staff like: **ID, Name, Gender, Address, job title, department, phone and his supervisor.**

15-Flights Activity Fact.

Represents the activity of the passenger during the flight and the data needed.

16-Customer Care Fact.

Represents the passenger's different **interactions** and the customer service's employee who is responsible for it and in which flight this interaction happened and the passenger's seat number and his ticket's number. It measures the **Problem Severity** based on its **Type** and **Time** of the occurrence to work on different interactions based on their priorities:

- 6 for Complaint Within the flight.**
- 4 for Complaint Before the flight.**
- 2 for Complaint After the flight.**
- 6 for Inquiry Within the flight.**
- 4 for Inquiry Before the flight.**
- 2 for Inquiry After the flight.**

It also shows the passengers' **feedback rate [From 1 ~ 5]**

17-Frequent Flyers Fact.

Represents the frequent flyers process and the **Miles Earned** that each passenger gained based on the **Miles Flown** also shows the **redeem type** to change these miles and the **promotion response**.

Physical Model

The creation of tables using SQL queries. In this project, PL SQL is used for all queries.

1- Create Passenger Dimension

CREATE TABLE Passengers_Dim

(

Passenger_key NUMBER CONSTRAINT Passenger_pk_cons PRIMARY KEY,
Passenger_ID VARCHAR2(250) CONSTRAINT Passenger_ID_unq UNIQUE,
Name VARCHAR2(250),
Gender VARCHAR2(50),
Phone VARCHAR2(250),
Email VARCHAR2(250),
Passenger_Type VARCHAR2(100),
Address VARCHAR2(250),
City VARCHAR2(250),
Country VARCHAR2(250));

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
PASSENGER_KEY	1	1	N	NUMBER	None			<input type="checkbox"/>	<input type="checkbox"/>
PASSENGER_ID	2		Y	VARCHAR2 (250 Byte)	None			<input type="checkbox"/>	<input type="checkbox"/>
NAME	3		Y	VARCHAR2 (250 Byte)	None			<input type="checkbox"/>	<input type="checkbox"/>
GENDER	4		Y	VARCHAR2 (50 Byte)	None			<input type="checkbox"/>	<input type="checkbox"/>
PHONE	5		Y	VARCHAR2 (250 Byte)	None			<input type="checkbox"/>	<input type="checkbox"/>
EMAIL	6		Y	VARCHAR2 (250 Byte)	None			<input type="checkbox"/>	<input type="checkbox"/>
PASSENGER_TYPE	7		Y	VARCHAR2 (100 Byte)	None			<input type="checkbox"/>	<input type="checkbox"/>
ADDRESS	8		Y	VARCHAR2 (250 Byte)	None			<input type="checkbox"/>	<input type="checkbox"/>
CITY	9		Y	VARCHAR2 (250 Byte)	None			<input type="checkbox"/>	<input type="checkbox"/>
COUNTRY	10		Y	VARCHAR2 (250 Byte)	None			<input type="checkbox"/>	<input type="checkbox"/>

2- Create Frequent Flyers Dimension

```
CREATE TABLE Frequent_flyers_Dim  
(  
    Freq_key NUMBER CONSTRAINT Freq_pk_cons PRIMARY KEY,  
    Freq_ID VARCHAR2(250) CONSTRAINT Freq_ID_unq UNIQUE,  
    Tier VARCHAR2(250),  
    Club_Membership_Status VARCHAR2(250),  
    Milage_Tier VARCHAR2(250),  
    Mempership_Status VARCHAR2(250)  
);
```

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
FREQ_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
FREQ_ID	2		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
TIER	3		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
CLUB_MEMBERSHIP_STATUS	4		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
MILAGE_TIER	5		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
MEMBERSHIP_STATUS	6		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>

3- Create Flights Dimension

```
CREATE TABLE Flights_Dim  
(  
    Flight_key NUMBER CONSTRAINT flight_pk_cons PRIMARY KEY,  
    Flight_ID VARCHAR2(250) CONSTRAINT flight_unq_cons UNIQUE,  
    Flight_zone VARCHAR2(250),
```

Flight_type VARCHAR2(250));

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
► FLIGHT_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
FLIGHT_ID	2		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
FLIGHT_ZONE	3		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
FLIGHT_TYPE	4		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>

4- Create Class Upgrade Dimension

CREATE TABLE Class_Upgrade_Dim

(

Class_Upgrade_key NUMBER CONSTRAINT class_pk_cons PRIMARY KEY,
Class_Upgrade_ID VARCHAR2(250) CONSTRAINT classU_ID_unq UNIQUE,
Purchased_Class VARCHAR2(250),
Flown_Class VARCHAR2(250),
Purchased_Flown_Group VARCHAR2(250),
Class_Change_Indicator VARCHAR2(100)

);

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
► CLASS_UPGRADE_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
CLASS_UPGRADE_ID	2		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
PURCHASED_CLASS	3		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
FLOWN_CLASS	4		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
PURCHASED_FLOWN_GROUP	5		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
CLASS_CHANGE_INDICATOR	6		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>

5- Create Date Dimension

CREATE TABLE Date_Dim

```
(  
    Full_date DATE CONSTRAINT D_pk_cons PRIMARY KEY,  
    Year VARCHAR2(15),  
    Quarter VARCHAR2(15),  
    Month_Name VARCHAR2(15),  
    Week_of_Year NUMBER(6),  
    Week_of_Month NUMBER(6),  
    Day_Name VARCHAR2(20),  
    Day_Number NUMBER(6)  
);
```

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
FULL_DATE	1	1	N	DATE		None		<input type="checkbox"/>	<input type="checkbox"/>
YEAR	2		Y	VARCHAR2 (15 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
QUARTER	3		Y	VARCHAR2 (15 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
MONTH_NAME	4		Y	VARCHAR2 (15 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
WEEK_OF_YEAR	5		Y	NUMBER (6)		None		<input type="checkbox"/>	<input type="checkbox"/>
WEEK_OF_MONTH	6		Y	NUMBER (6)		None		<input type="checkbox"/>	<input type="checkbox"/>
DAY_NAME	7		Y	VARCHAR2 (20 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
DAY_NUMBER	8		Y	NUMBER (6)		None		<input type="checkbox"/>	<input type="checkbox"/>

6- Create Channel Dimension

CREATE TABLE Channels_Dim

```
(  
    Channel_key NUMBER CONSTRAINT channel_pk_cons PRIMARY KEY,  
    Channel_Type VARCHAR2(200)  
);
```

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
CHANNEL_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
CHANNEL_TYPE	2		Y	VARCHAR2 (200 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>

7- Create Fare Basis Dimension

CREATE TABLE Fare_Basis_Dim

(

Fare_Basis_key NUMBER CONSTRAINT fare_pk_cons PRIMARY KEY,
Status VARCHAR2(400),
Code VARCHAR2(100),
Promotion VARCHAR2(150)

);

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
►FARE BASIS_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
STATUS	2		Y	VARCHAR2 (400 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
CODE	3		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
PROMOTION	4		Y	VARCHAR2 (150 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>

8- Create Hotel Dimension

CREATE TABLE Hotels_DIM

(

Hotel_key NUMBER CONSTRAINT HT_pk_cons PRIMARY KEY,
Hotel_ID VARCHAR2(50) CONSTRAINT HT_unq_cons UNIQUE,
Stars NUMBER,
Night_Price NUMBER,
City VARCHAR2(20),
Country VARCHAR2(20));

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
►HOTEL_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
HOTEL_ID	2		Y	VARCHAR2 (50 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
STARS	3		Y	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
NIGHT_PRICE	4		Y	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
CITY	5		Y	VARCHAR2 (20 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
COUNTRY	6		Y	VARCHAR2 (20 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>

9- Create Interactions Dimension

```
CREATE TABLE Interactions_DIM  
(  
    Interaction_key NUMBER CONSTRAINT Int_pk_cons PRIMARY KEY,  
    TYPE VARCHAR2(500),  
    Interaction_time VARCHAR2(150),  
    Description VARCHAR2(600)  
);
```

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
►INTERACTION_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
TYPE	2		Y	VARCHAR2 (500 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
INTERACTION_TIME	3		Y	VARCHAR2 (150 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
DESCRIPTION	4		Y	VARCHAR2 (600 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>

10-Create Aircrafts Dimension

```
CREATE TABLE Aircrafts_Dim  
(  
    Aircraft_key NUMBER CONSTRAINT aircraft_pk_cons PRIMARY KEY,  
    Aircraft_ID VARCHAR2(100) CONSTRAINT Acraft_unq_cons UNIQUE,  
    Model VARCHAR2(100),  
    Company VARCHAR2(100),  
    Seats_Capacity NUMBER  
);
```

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
►AIRCRAFT_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
AIRCRAFT_ID	2		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
MODEL	3		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
COMPANY	4		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
SEATS_CAPACITY	5		Y	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>

11-Create Airports Dimension

CREATE TABLE Airports_Dim

```
(  
    Airport_key number CONSTRAINT airport_pk_cons PRIMARY KEY,  
    Airport_ID VARCHAR2(100) CONSTRAINT Aport_unq_cons UNIQUE,  
    Airport_name VARCHAR2(100),  
    City VARCHAR2(100),  
    Country VARCHAR2(100)  
);
```

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
AIRPORT_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
AIRPORT_ID	2		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
AIRPORT_NAME	3		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
CITY	4		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>
COUNTRY	5		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>

12-Create Payment methods Dimension

CREATE TABLE Payment_Methods_Dim

```
(  
    P_M_key NUMBER CONSTRAINT p_m_pk_cons PRIMARY KEY,  
    Payment_type VARCHAR2(100)  
);
```

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
P_M_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input type="checkbox"/>
PAYMENT_TYPE	2		Y	VARCHAR2 (100 Byte)		None		<input type="checkbox"/>	<input type="checkbox"/>

13-Create Redeem Dimension

CREATE TABLE Redeem_Dim

```
(  
    Redeem_key NUMBER CONSTRAINT Rd_pk_cons PRIMARY KEY,  
    TYPE VARCHAR2(250),  
    Description VARCHAR2(500)  
);
```

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
REDEEM_KEY	1	1	N	NUMBER		None		<input type="checkbox"/>	<input checked="" type="checkbox"/>
TYPE	2		Y	VARCHAR2 (250 Byte)		None		<input type="checkbox"/>	<input checked="" type="checkbox"/>
DESCRIPTION	3		Y	VARCHAR2 (500 Byte)		None		<input type="checkbox"/>	<input checked="" type="checkbox"/>

14-Create Staff Dimension:

CREATE TABLE Staff_Dim

```
(  
    Staff_key NUMBER CONSTRAINT ST_pk_cons PRIMARY KEY,  
    Staff_ID VARCHAR2(500) CONSTRAINT ST_unq_cons UNIQUE,  
    NAME VARCHAR2(500),  
    Gender VARCHAR2(20),  
    Address VARCHAR2(250),  
    Job_Title VARCHAR2(250),  
    Department VARCHAR2(250),  
    Phone VARCHAR2(15),
```

Supervisor VARCHAR2(80));

Column Name	ID	Pk	Null?	Data Type	Default	Histogram	Encryption Alg	Salt	Seq/Trigger
STAFF_KEY	1	1	N	NUMBER				<input type="checkbox"/>	<input type="checkbox"/>
STAFF_ID	2		Y	VARCHAR2 (500 Byte)				<input type="checkbox"/>	<input type="checkbox"/>
NAME	3		Y	VARCHAR2 (500 Byte)				<input type="checkbox"/>	<input type="checkbox"/>
GENDER	4		Y	VARCHAR2 (20 Byte)				<input type="checkbox"/>	<input type="checkbox"/>
ADDRESS	5		Y	VARCHAR2 (250 Byte)				<input type="checkbox"/>	<input type="checkbox"/>
JOB_TITLE	6		Y	VARCHAR2 (250 Byte)				<input type="checkbox"/>	<input type="checkbox"/>
DEPARTMENT	7		Y	VARCHAR2 (250 Byte)				<input type="checkbox"/>	<input type="checkbox"/>
PHONE	8		Y	VARCHAR2 (15 Byte)				<input type="checkbox"/>	<input type="checkbox"/>
SUPERVISOR	9		Y	VARCHAR2 (80 Byte)				<input type="checkbox"/>	<input type="checkbox"/>

15-Create Flights Activity Fact

```
CREATE TABLE flights_activity_Fact
(
    Leg_key NUMBER CONSTRAINT Leg_pk_cons PRIMARY KEY,
    Ticket_number NUMBER CONSTRAINT TN_unq_cons UNIQUE,
    Passenger_key NUMBER,
    Freq_key NUMBER,
    Aircraft_key NUMBER,
    Segment_number NUMBER ,
    Flight_key NUMBER,
    Departure_airport NUMBER,
    Arrival_airport NUMBER,
    Actual_departure_date DATE,
    Schedualed_departure_time TIMESTAMP,
    Actual_departure_time TIMESTAMP,
    Actual_arrival_date DATE,
    Schedualed_arrival_time TIMESTAMP,
    Actual_arrival_time TIMESTAMP,
    Flight_duration NUMBER,
    Reservation_Ticket_Date DATE,
```

```

Channel_key NUMBER,
Payment_method_key NUMBER,
Fare_basis_key NUMBER,
Class_upgrade_key NUMBER,
Luggage_Number NUMBER,
Base_fare_price NUMBER,
Taxes NUMBER,
Overweight_Luggage_price NUMBER,
Upgrage_fees NUMBER,
Transaction_fees NUMBER,
Promotion_amount NUMBER,
Net_price NUMBER,
Cost NUMBER,
Profit NUMBER,
Hotel_key NUMBER,
overnight_stays NUMBER
);

ALTER TABLE flights_activity_Fact
ADD(
    CONSTRAINT pass_fk_cons FOREIGN KEY(Passenger_key) REFERENCES
Passengers_Dim(Passenger_key),
    CONSTRAINT FF_fk_cons FOREIGN KEY(Freq_key) REFERENCES
Frequent_flyers_Dim(Freq_key),
    CONSTRAINT aircraft_fk_cons FOREIGN KEY(aircraft_key) REFERENCES
Aircrafts_Dim(Aircraft_key),
    CONSTRAINT flight_fk_cons FOREIGN KEY(Flight_key) REFERENCES
Flights_Dim(Flight_key),
);

```

CONSTRAINT airport_fk_cons **FOREIGN KEY**(Departure_airport) **REFERENCES**
Airports_Dim(Airport_key),

CONSTRAINT airport_fk_cons **FOREIGN KEY**(Arrival_airport) **REFERENCES**
Airports_Dim(Airport_key),

CONSTRAINT actual_date_departure_fk_cons **FOREIGN KEY**(Actual_departure_date)
REFERENCES Date_Dim(Full_date),

CONSTRAINT actual_date_arrival_fk_cons **FOREIGN KEY**(Actual_arrival_date)
REFERENCES Date_Dim(Full_date),

CONSTRAINT resv_t_date_fk_cons **FOREIGN KEY**(Reservation_Ticket_Date)
REFERENCES Date_Dim(Full_date),

CONSTRAINT ch_fk_cons **FOREIGN KEY**(Channel_key) **REFERENCES**
Channels_Dim(Channel_key),

CONSTRAINT pm_fk_cons **FOREIGN KEY**(Payment_method_key) **REFERENCES**
Payment_Methods_Dim(P_M_key),

CONSTRAINT fb_fk_cons **FOREIGN KEY**(Fare_basis_key) **REFERENCES**
Fare_Basis_Dim(Fare_Basis_key),

CONSTRAINT cu_fk_cons **FOREIGN KEY**(Class_upgrade_key) **REFERENCES**
Class_Upgrade_Dim(Class_Upgrade_key),

CONSTRAINT H_fk_cons **FOREIGN KEY**(Hotel_key) **REFERENCES** Hotels_DIM(Hotel_key)
);

LUGGAGE_NUMBER	BASE_FARE_PRICE	TAXES	OVERWEIGHT_LUGGAGE_PRICE	UPGRADE_FEES	TRANSACTION_FEES	PROMOTION_AMOUNT	NET_PRICE	COST	PROFIT	HOTEL_KEY	OVERNIGHT_STAYS	
0	10.5	10.7		32.57	17.83	19.23	0.3	90.53	128.66	-38.13	41	0
4	14.97	19.11		41.42	11.52	19.17	0.32	105.87	122.46	-16.59	27	1
1	19.19	14.02		38.74	19.21	19.56	0.44	110.28	86.18	24.1	54	1
1	16.14	18.33		19.45	13.91	16.7	99	-14.47	94.24	-108.71	17	0
3	11.7	16.57		32.76	13.83	17.93	0.24	92.55	81.75	10.8	94	1
0	16.06	10.23		22.25	12.36	15.24	99	-22.86	93.3	-116.16	94	1
1	14.62	15.13		21.83	16.02	12.45	99	-18.95	114.59	-133.54	59	0
0	17.51	13.93		23.39	19.1	19.54	99	-5.53	145.48	-151.01	76	0
2	14.68	18.85		16.36	10.49	11.26	99	-27.36	115.52	-142.88	85	1
2	16.82	12.65		14.54	12.99	16.16	0.3	72.86	137.21	-64.35	40	1
0	16.55	17.44		17.27	13.77	19.47	0.41	84.09	138.69	-54.6	72	1
4	12.61	12.43		26.68	18.14	13.1	99	-16.04	125.47	-141.51	53	0
2	18.49	13.07		30.67	19.3	15.47	99	-2	147.28	-149.28	10	1
1	11.43	10.88		43.5	10.92	17.52	99	-4.75	140.24	-144.99	83	0
0	14.65	13.34		38.58	15	16.29	99	-1.14	124.36	-125.5	23	0
0	15.13	15.91		27.87	11.93	18.17	0.42	88.59	116.1	-27.51	60	1
1	18.52	12.61		15.46	16.79	12.29	99	-23.33	130.64	-153.97	25	1
2	19.77	18.92		28.45	16.36	16.46	0.11	99.85	96.94	2.91	52	0
3	19.56	10.5		48.64	18.73	15.5	99	13.93	108.58	-94.65	43	1

16-Create Customer Care Fact

```
CREATE TABLE Customer_Care_fact
```

```
(
```

```
    CC_key NUMBER CONSTRAINT cc_pk_cons PRIMARY KEY ,  
    Ticket_number NUMBER CONSTRAINT TN_unq_consss UNIQUE,  
    Passenger_key NUMBER ,  
    Freq_key NUMBER ,  
    Reservation_Ticket_Date DATE ,  
    Segment_number NUMBER ,  
    Flight_key NUMBER ,  
    Interaction_key NUMBER ,  
    Staff_key NUMBER ,  
    Problem_severity NUMBER ,  
    Feedback_Rate NUMBER ,  
    Deprature_airport NUMBER ,  
    Arrival_airport NUMBER ,  
    Seat_Number NUMBER
```

```
);
```

```
ALTER TABLE Customer_Care_fact
```

```
ADD(
```

```
    CONSTRAINT pass_fk_consss FOREIGN KEY(Passenger_key) REFERENCES  
    Passengers_Dim(Passenger_key),
```

```
    CONSTRAINT FF_fk_consss FOREIGN KEY(Freq_key) REFERENCES  
    Frequent_flyers_Dim(Freq_key),
```

```
    CONSTRAINT resv_t_date_fk_consss FOREIGN KEY(Reservation_Ticket_Date) REFERENCES  
    Date_Dim(Full_date),
```

```
    CONSTRAINT flight_fk_consss FOREIGN KEY(Flight_key) REFERENCES Flights_Dim(Flight_key),
```

```

CONSTRAINT int_fk_consss FOREIGN KEY(Interaction_key) REFERENCES
Interactions_DIM(Interaction_key),

CONSTRAINT st_fk_consss FOREIGN KEY(Staff_key) REFERENCES Staff_Dim(Staff_key),

CONSTRAINT airport_fk_conssecs FOREIGN KEY(Deprature_airport) REFERENCES
Airports_Dim(Airport_key),

CONSTRAINT airport_fk_conssecs FOREIGN KEY(Arrival_airport) REFERENCES
Airports_Dim(Airport_key)

);

```

CC_KEY	TICKET_NUMBER	PASSENGER_KEY	FREQ_KEY	RESERVATION_TICKET_DATE	SEGMENT_NUMBER	FLIGHT_KEY	INTERACTION_KEY	STAFF_KEY	PROBLEM_SEVERITY	FEEDBACK_RATE
1	10001	1	1	12/12/2021	14	1	4	1	2	
2	10002	2	3	12/12/2021	12	2	5	2	6	
3	10003	3	5	12/12/2021	14	5	9	3		2
4	10004	4	9	12/12/2021	14	4	8	50		3
5	10005	5	10	9/19/2020	14	3	5	30	6	
6	10006	6	13	9/19/2020	14	2	4	10	2	
7	10007	7	50	2/6/2021	14	1	2	40	6	
8	10008	8	17	2/4/2021	14	4	4	10	2	
9	10009	9	30	2/4/2021	14	2	1	50	2	
10	10010	10	20	9/16/2021	14	5	4	11	2	
11	10011	11	40	9/16/2021	14	6	4	13	2	
12	10012	12	40	9/9/2021	14	1	3	17	4	
13	10013	13	8	8/31/2021	14	3	4	22	2	
14	10014	14	50	8/31/2021	14	4	8	25		3
15	10015	15	14	8/17/2021	14	1	4	32	2	
16	10016	16	12	8/17/2021	14	2	1	35	2	
17	10017	17	11	12/12/2021	12	3	2	16	6	
18	10018	18	22	8/11/2021	12	4	5	1	6	
19	10019	19	25	8/1/2021	12	6	3	7	4	

17-CREATE Frequent Flyers Fact

```

CREATE TABLE Frequent_Flyers_fact
(
    FF_Fact_key Number CONSTRAINT freqf_fact_pk_cons PRIMARY KEY ,
    Ticket_number NUMBER CONSTRAINT TTN UNIQUE,
    Freq_key NUMBER,
    Flight_date DATE,
    Flight_key NUMBER,
    Deprature_airport NUMBER,
    Arrival_airport NUMBER,
    Miles_flown NUMBER,
    Miles_earned NUMBER,
)

```

```

promotion_response VARCHAR2(20),
Redeem_key NUMBER
);

```

```

ALTER TABLE Frequent_Flyers_fact
ADD(
    CONSTRAINT FF_fk_cons FOREIGN KEY(Freq_key) REFERENCES
Frequent_flyers_Dim(Freq_key),
    CONSTRAINT f_date_fk_cons FOREIGN KEY(Flight_date) REFERENCES Date_Dim(full_date),
    CONSTRAINT flight_fk_cons FOREIGN KEY(Flight_key) REFERENCES
Flights_Dim(Flight_key),
    CONSTRAINT airport_fk_cons FOREIGN KEY(Deprature_airport) REFERENCES
Airports_Dim(Airport_key),
    CONSTRAINT airport_fk_cons FOREIGN KEY(Arrival_airport) REFERENCES
Airports_Dim(Airport_key),
    CONSTRAINT rd_fk_cons FOREIGN KEY(Redeem_key) REFERENCES
Redeem_Dim(Redeem_key)
);

```

FF_FACT_KEY	TICKET_NUMBER	FREQ_KEY	FLIGHT_DATE	FLIGHT_KEY	DEPRATURE_AIRPORT	ARRIVAL_AIRPORT	MILES_FLOWN	MILES_EARNED	PROMOTION_RESPONSE	REDEEM_KEY
1	10001	1	12/15/2021	1	1	3	2500	250	yes	1
2	10002	2	1/15/2021	2	2	5	2500	250	yes	2
3	10003	3	1/15/2021	5	3	7	2500	250	no	3
4	10004	4	1/15/2021	4	5	9	2500	250	no	2
5	10005	5	10/25/2020	3	4	8	2000	200	yes	3
6	10006	6	10/25/2020	2	7	11	3500	350	yes	2
7	10007	7	3/10/2021	1	9	20	3000	300	no	1
8	10008	8	3/2/2021	4	10	30	3500	350	no	3
9	10009	9	3/2/2021	2	13	14	3500	350	yes	2
10	10010	10	10/10/2021	5	1	7	4600	460	no	1
11	10011	11	3/10/2021	6	5	8	3500	350	yes	1
12	10012	12	3/10/2021	1	15	13	2000	200	no	2
13	10013	13	9/9/2021	3	20	30	3500	350	yes	1
14	10014	14	9/9/2021	4	22	32	3500	350	no	1
15	10015	15	9/9/2021	1	17	78	4500	450	no	3
16	10016	16	9/1/2021	2	33	55	4000	400	yes	2
17	10017	17	12/18/2021	3	18	16	4000	400	yes	2
18	10018	18	8/20/2021	4	19	11	2000	200	no	3
19	10019	19	8/15/2021	6	20	55	1500	150	no	1

Insert Dummy Data into The Physical Tables

The data inserted was using excel files that have dummy data and PLSQL blocks generating random data

1- Passengers Dimension Data

PASSENGER_KEY	PASSENGER_ID	NAME	GENDER	PHONE	EMAIL	PASSENGER_TYPE	ADDRESS	CITY	COUNTRY
1 P-1	Rutter Prestney	Female	917-808-1103	rprestney0@moonfruit.com	Regular	78511 School Lane	Bartica	Guyana	
2 P-2	Esmaria Uzzell	Female	979-499-6498	euzzell1@wunderground.com	Frequent	070 Mitchell Plaza	Xiangshui	China	
3 P-3	Heinrich Siaskowski	Male	640-434-5592	hsiaskowski2@hao123.com	Regular	4472 Susan Plaza	Nunbaundelha	Indonesia	
4 P-4	Dean Scannell	Agender	378-843-4956	dscannell3@toplist.cz	Frequent	597 Sherman Circle	Nanshan	China	
5 P-5	Michail Gifkins	Male	542-649-9636	mgifkins4@instagram.com	Regular	32 Oakridge Hill	Shtip	Macedonia	
6 P-6	Emilee Bunce	Female	604-227-5606	ebunce5@ehow.com	Regular	2 Westend Center	Bafoulabé	Mali	
7 P-7	Yuma Pitway	Male	241-580-6213	ypitway6@patch.com	Frequent	03228 Merrick Avenue	Riihimäki	Finland	
8 P-8	Mycah Giffaut	Male	752-598-1590	mgiffaut7@barnesandnoble.com	Frequent	6020 Shoshone Junction	Nangka	Indonesia	
9 P-9	Carine Wannan	Female	209-135-5521	cwannan8@bbb.org	Regular	9 Farmco Pass	Bonga	Philippines	
10 P-10	Molli Roughey	Female	539-876-1055	mroughey9@youku.com	Frequent	549 Anhal Alley	Pakel	Indonesia	
11 P-11	Linus Gothrup	Male	414-876-2266	lgothrupa.imgur.com	Regular	822 Lerdahl Terrace	Vallegrande	Bolivia	
12 P-12	Rori Ravel	Female	670-475-6607	rravelb@marketwatch.com	Frequent	89 Springs Court	Chhaglnaiya	Bangladesh	
13 P-13	Genia Verheyden	Female	541-840-8413	gverheyenc@washingtongpost.com	Regular	170 Susan Pass	Zhugang	China	
14 P-14	Terrijo Summerson	Female	696-596-8896	tsummersond@time.com	Regular	276 Clove Circle	Suphan Buri	Thailand	
15 P-15	Patten Learmouth	Agender	922-602-4842	plearmouth@friendfeed.com	Frequent	258 Leroy Lane	Vallenar	Chile	
16 P-16	Ginevra Jillins	Female	416-718-0449	gjillinsf@goo.gl	Frequent	23 Buhler Point	Nogata	Japan	

2- Frequent Flyers Dimension Data

FREQ_KEY	FREQ_ID	TIER	CLUB_MEMBERSHIP_STATUS	MILEAGE_TIER	MEMBERSHIP_STATUS
1 FF-1	FF-1	Basic	Non-Member	Under 100,000 miles	Silver
2 FF-2	FF-2	MidTier	Club Member	100,000-499,999 miles	Gold
3 FF-3	FF-3	WarriorTier	Non-Member	500,000-1,999,999 miles	Platinum
4 FF-4	FF-4	MidTier	Non-Member	100,000-499,999 miles	Gold
5 FF-5	FF-5	Basic	Non-Member	Under 100,000 miles	Silver
6 FF-6	FF-6	WarriorTier	Club Member	500,000-1,999,999 miles	Platinum
7 FF-7	FF-7	WarriorTier	Non-Member	500,000-1,999,999 miles	Platinum
8 FF-8	FF-8	MidTier	Non-Member	100,000-499,999 miles	Gold
9 FF-9	FF-9	Basic	Non-Member	Under 100,000 miles	Silver
10 FF-10	FF-10	Basic	Club Member	Under 100,000 miles	Silver

3- Class Upgrade Dimension Data

	CLASS_UPGRADE_KEY	CLASS_UPGRADE_ID	PURCHASED_CLASS	FLOWN_CLASS	PURCHASED_FLOWN_GROUP	CLASS_CHANGE_INDICATOR
▶	1	CU-1	Economy	Economy	Economy-Economy	No Class Change
	2	CU-2	Economy	Prem Economy	Economy-Prem Economy	Upgrade
	3	CU-3	Economy	Business	Economy-Business	Upgrade
	4	CU-4	Economy	First	Economy-First	Upgrade
	5	CU-5	Prem Economy	Economy	Prem Economy-Economy	Downgrade
	6	CU-6	Prem Economy	Prem Economy	Prem Economy-Prem Economy	No Class Change
	7	CU-7	Prem Economy	Business	Prem Economy-Business	Upgrade
	8	CU-8	Prem Economy	First	Prem Economy-First	Upgrade
	9	CU-9	Business	Economy	Business-Economy	Downgrade
	10	CU-10	Business	Prem Economy	Business-Prem Economy	Downgrade
	11	CU-11	Business	Business	Business-Business	No Class Change
	12	CU-12	Business	First	Business-First	Upgrade
	13	CU-13	First	Economy	First-Economy	Downgrade
	14	CU-14	First	Prem Economy	First-Prem Economy	Downgrade
	15	CU-15	First	Business	First-Business	Downgrade
	16	CU-16	First	First	First-First	No Class Change

4- Flights Dimension Data

	FLIGHT_KEY	FLIGHT_ID	FLIGHT_ZONE	FLIGHT_TYPE
▶	1	F-ID-1	Domestic	Non-stop
	2	F-ID-2	Domestic	Direct
	3	F-ID-3	Domestic	Indirect
	4	F-ID-4	International	Non-stop
	5	F-ID-5	International	Direct
	6	F-ID-6	International	Indirect

5- Fare Basis Dimension Data

	FARE_BASIS_KEY	STATUS	CODE	PROMOTION
▶	1	111	Full Fare First Class	No
	2	112	Discount Fare First Class	Yes
	3	221	Full Fare Business Class	No
	4	222	Discount Fare Business Class	Yes
	5	331	Full Fare Economy Class	No
	6	332	Discount Fare Economy Class	Yes

6- Aircrafts Dimension Data

AIRCRAFT_KEY	AIRCRAFT_ID	MODEL	COMPANY	SEATS_CAPACITY
1	ARC-1	7AC Champ, 7ECA Citabria	Champion Aircraft Corporation	2
2	ARC-2	7GCAA, 7GCBC, 7KCAB Citabria	Champion Aircraft Corporation	2
3	ARC-3	Beechcraft Baron / 55 Baron	Beechcraft	6
4	ARC-4	Beechcraft 1900	Beechcraft	19
5	ARC-5	Beechcraft Baron / 58 Baron	Beechcraft	6
6	ARC-6	Bell 429	Bell	7
7	ARC-7	Bell 212	Bell	15
8	ARC-8	Bell 412	Bell	13
9	ARC-9	Avro RJ70	Avro International Aerospace	70
10	ARC-10	Hawker Siddeley HS 748	Avro International Aerospace	36
11	ARC-11	Avro RJ85	Avro International Aerospace	90
12	ARC-12	Avro RJ100	Avro International Aerospace	111
13	ARC-13	Airbus A319	Airbus	156
14	ARC-14	Airbus A330-900neo	Airbus	460
15	ARC-15	Airbus A319 (sharklets)	Airbus	156
16	ARC-16	Airbus A318	Airbus	110
17	ARC-17	Airbus A330-200	Airbus	268
18	ARC-18	Airbus A320	Airbus	180
19	ARC-19	Airbus A380-800	Airbus	520
20	ARC-20	Airbus A310-200	Airbus	230
21	ARC-21	Airbus A320neo	Airbus	146
22	ARC-22	Airbus A340-300	Airbus	290

7- Airports Dimension Data

AIRPORT_KEY	AIRPORT_ID	AIRPORT_NAME	CITY	COUNTRY
1	AP-1	Guyuan Liupanshan Airport	Villa Nueva	Guatemala
2	AP-2	Andorra la Vella Heliport	San Francisco	United States
3	AP-3	Millicent Airport	Veshnyaki	Russia
4	AP-4	Mackinac Island Airport	Luodong	China
5	AP-5	Stauning Airport	Baixiang	China
6	AP-6	Bauerfield International Airport	Sirāhā	Nepal
7	AP-7	Xangongo Airport	Sambonggede	Indonesia
8	AP-8	Hokitika Airfield	Xiaping	China
9	AP-9	Selfridge Air National Guard Base Airport	Golubinci	Serbia
10	AP-10	Sanfebagar Airport	Yuyao	China
11	AP-11	Berz-Macomb Airport	Karsin	Poland
12	AP-12	Clovis Municipal Airport	Mszana Dolna	Poland
13	AP-13	Chigorodó Airport	Fahraj	Iran
14	AP-14	Rongelap Island Airport	Zaragoza	Colombia
15	AP-15	Cowarie Airport	Ustupo	Panama
16	AP-16	Nea Anchialos Airport	Paradyż	Poland
17	AP-17	Mercedita Airport	Changyōn	North Korea

8- Channel Dimension Data

	CHANNEL_KEY	CHANNEL_TYPE
►	1	Airline Website
	2	Travel Agency

9- Payment methods Dimension Data

	P_M_KEY	PAYMENT_TYPE
►	1	Cash
	2	Credit

10- Interactions Dimension Data

	INTERACTION_KEY	TYPE	INTERACTION_TIME	DESCRIPTION
►	1	Complaints	Before Flight	The customer writes down any problem he encountered before the trip
	2	Complaints	Within Flight	The customer writes down any problem he encountered during the trip
	3	Complaints	After Flight	The customer writes down any problem he encountered after the trip
	4	Inquire	Before Flight	The customer inquires about anything before the trip
	5	Inquire	Within Flight	The customer inquires about anything during the trip
	6	Inquire	After Flight	The customer inquires about anything after the trip
	7	Feedback	Before Flight	customer feedback about the trip, services and offers
	8	Feedback	Within Flight	customer feedback about the trip, services and offers
	9	Feedback	After Flight	customer feedback about the trip, services and offers

11- Redeem Dimension Data

	REDEEM_KEY	TYPE	DESCRIPTION
►	1	Flight awards	Get discount on your ticket total price
	2	Upgrade awards	Get premium cabin upgrades
	3	Club and status awards	Complimentary amenities and services are available to make your travel more productive and relaxing.

12- Staff Dimension Data

STAFF_KEY	STAFF_ID	NAME	GENDER	ADDRESS	JOB_TITLE	DEPARTMENT	PHONE	SUPERVISOR
1	ST-1	Armand Burkill	Male	Suite 43	Professor	Maintenance	385-642-0533	Tommie Laugheran
2	ST-2	Evangelina Brende	Female	13th Floor	Duty Officer	Services	599-365-6831	Steffane Trevena
3	ST-3	Woody Cleugh	Male	Apt 1960	Research Nurse	Operations	816-595-5080	Maximilianus Cousin
4	ST-4	Boote Barosch	Male	Suite 5	Payment Adjustment Coordinator	Finance divisions	289-110-9790	Ozzie Fillary
5	ST-5	Blake Bleesing	Female	Suite 14	Help Desk Technician	Maintenance	220-217-7423	Fina Salman
6	ST-6	Kellyann Setchfield	Female	Suite 44	Clinical Specialist	Operations	654-376-6995	Ali Butland
7	ST-7	Drusi Witty	Female	Room 342	Electrical Engineer	Marketing	286-751-5110	Vallie Kelsall
8	ST-8	Timothy Bolstridge	Male	10th Floor	Mechanical Systems Engineer	Training	974-152-3728	Ellary Edgson
9	ST-9	Allsun Tregoning	Female	Apt 670	Biostatistician II	Operations	356-117-6198	Jourdan Kleinstern
10	ST-10	Euell Bussel	Male	PO Box 80183	Financial Advisor	Human Resources	136-929-5138	Brit Tempest
11	ST-11	Lorraine Boarleyson	Female	Apt 914	Media Manager III	Services	366-256-4943	Devin Grzegorzewicz
12	ST-12	Catharine Whitehair	Female	9th Floor	Speech Pathologist	Services	354-389-6399	Ernesta Videler
13	ST-13	Adriaens Blaske	Female	5th Floor	Junior Executive	Operations	389-339-5142	Berry Godain
14	ST-14	Ainsley Foxon	Female	Room 330	Research Associate	Operations	396-780-8574	Odelia Kurton
15	ST-15	Steffen Streeton	Male	Room 367	Junior Executive	Training	668-622-6781	Angelico Ankers
16	ST-16	Jasun Yesichev	Male	Suite 5	Assistant Professor	Maintenance	143-314-4440	Fairfax McGuiness
17	ST-17	Scottie Basini-Gazzi	Male	Apt 433	Marketing Manager	Finance divisions	208-485-4214	Angelo de la Tremoille
18	ST-18	Xyline Rushman	Female	PO Box 75816	Budget/Operations Analyst IV	Operations	735-645-5330	Anna-diane Lesaunier

13- Date Dimension Data

FULL_DATE	YEAR	QUARTER	MONTH_NAME	WEEK_OF_YEAR	WEEK_OF_MONTH	DAY_NAME	DAY_NUMBER
12/12/2021	2021	Q4	December	51	3	Sunday	12
12/13/2021	2021	Q4	December	51	3	Monday	13
12/14/2021	2021	Q4	December	51	3	Tuesday	14
12/15/2021	2021	Q4	December	51	3	Wednesday	15
12/16/2021	2021	Q4	December	51	3	Thursday	16
12/17/2021	2021	Q4	December	51	3	Friday	17
12/18/2021	2021	Q4	December	51	3	Saturday	18
12/19/2021	2021	Q4	December	52	4	Sunday	19
12/20/2021	2021	Q4	December	52	4	Monday	20
12/21/2021	2021	Q4	December	52	4	Tuesday	21
12/22/2021	2021	Q4	December	52	4	Wednesday	22
12/23/2021	2021	Q4	December	52	4	Thursday	23
12/24/2021	2021	Q4	December	52	4	Friday	24
12/25/2021	2021	Q4	December	52	4	Saturday	25
12/26/2021	2021	Q4	December	53	5	Sunday	26
12/27/2021	2021	Q4	December	53	5	Monday	27
12/28/2021	2021	Q4	December	53	5	Tuesday	28

14- Hotels Dimension Data

	HOTEL_KEY	HOTEL_ID	STARS	NIGHT_PRICE	CITY	COUNTRY
▶	1	H-1	5	680	Manolo Fortich	Philippines
	2	H-2	5	670	Nanzhen	China
	3	H-3	4	400	Durham	United States
	4	H-4	5	680	El Carmen	Peru
	5	H-5	3	398	Olsztyn	Poland
	6	H-6	1	133	Powassan	Canada
	7	H-7	4	410	Bayaguana	Dominican Republic
	8	H-8	3	362	Aparecida do Taboado	Brazil
	9	H-9	3	350	Hammam-Lif	Tunisia
	10	H-10	2	200	Tayturka	Russia
	11	H-11	5	688	Buy	Russia
	12	H-12	5	689	Whitchurch	United Kingdom
	13	H-13	4	420	Kasempa	Zambia
	14	H-14	2	220	Jetaral	China
	15	H-15	4	430	Gaoyu	China

15- Flights Activity Fact

LUGGAGE_NUMBER	BASE_FARE_PRICE	TAXES	OVERWEIGHT_LUGGAGE_PRICE	UPGRAGE_FEES	TRANSACTION_FEES	PROMOTION_AMOUNT	NET_PRICE	COST	PROFIT	HOTEL_KEY	OVERNIGHT_STAYS
0	10.5	10.7	32.57	17.83	19.23	0.3	90.53	128.66	-38.13	41	0
4	14.97	19.11	41.42	11.52	19.17	0.32	105.87	122.46	-16.59	27	1
1	19.19	14.02	38.74	19.21	19.56	0.44	110.28	86.18	24.1	54	1
1	16.14	18.33	19.45	13.91	16.7	99	-14.47	94.24	-108.71	17	0
3	11.7	16.57	32.76	13.83	17.93	0.24	92.55	81.75	10.8	94	1
0	16.06	10.23	22.25	12.36	15.24	99	-22.86	93.3	-116.16	94	1
1	14.62	15.13	21.83	16.02	12.45	99	-18.95	114.59	-133.54	59	0
0	17.51	13.93	23.39	19.1	19.54	99	-5.53	145.48	-151.01	76	0
2	14.68	18.85	16.36	10.49	11.26	99	-27.36	115.52	-142.88	85	1
2	16.82	12.65	14.54	12.99	16.16	0.3	72.86	137.21	-64.35	40	1
0	16.55	17.44	17.27	13.77	19.47	0.41	84.09	138.69	-54.6	72	1
4	12.61	12.43	26.68	18.14	13.1	99	-16.04	125.47	-141.51	53	0
2	18.49	13.07	30.67	19.3	15.47	99	-2	147.28	-149.28	10	1
1	11.43	10.88	43.5	10.92	17.52	99	-4.75	140.24	-144.99	83	0
0	14.65	13.34	38.58	15	16.29	99	-1.14	124.36	-125.5	23	0
0	15.13	15.91	27.87	11.93	18.17	0.42	88.59	116.1	-27.51	60	1
1	18.52	12.61	15.46	16.79	12.29	99	-23.33	130.64	-153.97	25	1
2	19.77	18.92	28.45	16.36	16.46	0.11	99.85	96.94	2.91	52	0
3	19.56	10.5	48.64	18.73	15.5	99	13.93	108.58	-94.65	43	1

16- Frequent Flyers Fact

FF_FACT_KEY	TICKET_NUMBER	FREQ_KEY	FLIGHT_DATE	FLIGHT_KEY	DEPRATURE_AIRPORT	ARRIVAL_AIRPORT	MILES_FLOWN	MILES_EARNED	PROMOTION_RESPONSE	REDEEM_KEY
1	10001	1	12/15/2021	1	1	3	2500	250	yes	1
2	10002	2	1/15/2021	2	2	5	2500	250	yes	2
3	10003	3	1/15/2021	5	3	7	2500	250	no	3
4	10004	4	1/15/2021	4	5	9	2500	250	no	2
5	10005	5	10/25/2020	3	4	8	2000	200	yes	3
6	10006	6	10/25/2020	2	7	11	3500	350	yes	2
7	10007	7	3/10/2021	1	9	20	3000	300	no	1
8	10008	8	3/2/2021	4	10	30	3500	350	no	3
9	10009	9	3/2/2021	2	13	14	3500	350	yes	2
10	10010	10	10/10/2021	5	1	7	4600	460	no	1
11	10011	11	3/10/2021	6	5	8	3500	350	yes	1
12	10012	12	3/10/2021	1	15	13	2000	200	no	2
13	10013	13	9/9/2021	3	20	30	3500	350	yes	1
14	10014	14	9/9/2021	4	22	32	3500	350	no	1
15	10015	15	9/9/2021	1	17	78	4500	450	no	3
16	10016	16	9/1/2021	2	33	55	4000	400	yes	2
17	10017	17	12/18/2021	3	18	16	4000	400	yes	2
18	10018	18	8/20/2021	4	19	11	2000	200	no	3
19	10019	19	8/15/2021	6	20	55	1500	150	no	1

17- Customer Care Fact Data

ICER_KEY	FREQ_KEY	RESERVATION_TICKET_DATE	SEGMENT_NUMBER	FLIGHT_KEY	INTERACTION_KEY	STAFF_KEY	PROBLEM_SEVERITY	FEEDBACK_RATE	DEPRATURE_AIRPORT	ARRIVAL_AIRPORT	SEAT_NUMBER
1	1	580	14	1	1	1	6	5	1	3	167
2	3	580	12	2	2	2	1	5	2	5	291
3	5	580	14	5	3	3	6	4	3	7	35
4	9	580	14	4	4	50	2	1	5	9	160
5	10	580	14	3	5	30	6	2	4	8	206
6	13	580	14	2	6	10	4	4	2	5	98
7	50	583	14	1	7	40	7	1	1	3	268
8	17	583	14	4	8	10	5	2	5	9	176
9	30	583	14	2	9	50	8	2	2	5	72
10	20	583	14	5	8	11	5	4	3	7	42
11	40	583	14	6	7	13	7	3	5	8	234
12	66	583	14	1	6	17	4	5	1	3	299
13	88	583	14	3	5	22	6	5	4	8	196
14	77	583	14	4	4	25	2	1	5	9	51
15	14	586	14	1	3	32	6	2	1	3	43
16	12	586	14	2	2	35	6	3	2	5	19
17	11	586	12	3	1	16	4	2	4	8	124
18	22	586	12	4	3	1	5	1	5	9	54
19	25	586	12	6	1	7	7	1	1	3	219

Insert The Measured Data into Fact Tables

1- Insert into Flights Activity Fact [Net price & Profit]

```
DECLARE
  CURSOR Cur IS
    SELECT Leg_key, Net_price, Base_fare_price, Taxes, Overweight_Luggage_price, Upgrage_fees,
Transaction_fees, Promotion_amount, Cost, Profit
    from Flights_activity_fact
    FOR UPDATE OF Net_price ;

BEGIN
  FOR Rec IN Cur
    Loop
      UPDATE Flights_activity_fact
      SET Net_price = (Base_fare_price+
Taxes+nvl(Overweight_Luggage_price,0)+nvl(Upgrage_fees,0)+Transaction_fees)-
nvl(Promotion_amount ,0)
      WHERE CURRENT OF Cur;
      UPDATE Flights_activity_fact
      SET Profit =Net_price-Cost
      WHERE CURRENT OF Cur;

    END LOOP;
  END;
```

2- Insert into Customer Care Fact [Seat number]

```
DECLARE
  CURSOR Cur IS
    SELECT Flight_key ,Seat_Number
    from Customer_Care_fact
    FOR UPDATE OF Seat_Number;

BEGIN
  FOR Rec IN Cur
    LOOP
      IF (Rec.Flight_key in (1,2,3,4,5,6))
      THEN
        UPDATE Customer_Care_fact
        SET Seat_Number = TRUNC(dbms_random.value(1,300),0)
```

```

        WHERE CURRENT OF Cur;

    END IF;

END LOOP;

END;

```

3- Insert into Customer Care Fact [Problem severity & Feedback Rate]

```

DECLARE

v_type VARCHAR(250);
v_time VARCHAR(250);

CURSOR Cur IS
    SELECT Interaction_key ,Problem_severity
    from Customer_Care_fact
    FOR UPDATE OF Problem_severity;

BEGIN

FOR Rec IN Cur

LOOP

SELECT Interaction_time, TYPE
INTO v_time, v_type
FROM Interactions_DIM
WHERE Interaction_key = Rec.INTERACTION_KEY;

IF (v_time = 'Within Flight' AND v_type = 'Complaints')
THEN

UPDATE Customer_Care_fact
SET Problem_severity = 6
WHERE CURRENT OF Cur;

ELSIF (v_time = 'Before Flight' AND v_type = 'Complaints')
THEN
    UPDATE Customer_Care_fact
    SET Problem_severity = 2
    WHERE CURRENT OF Cur;

ELSIF (v_time = 'After Flight' AND v_type = 'Complaints')
THEN
    UPDATE Customer_Care_fact
    SET Problem_severity = 4
    WHERE CURRENT OF Cur;

ELSIF (v_time = 'Within Flight' AND v_type = 'Inquire')

```

```

THEN
  UPDATE Customer_Care_fact
  SET Problem_severity= 6
  WHERE CURRENT OF Cur;

ELSIF (v_time = 'Within Flight' AND v_type = 'Feedback')
THEN
  UPDATE Customer_Care_fact
  SET Feedback_Rate = TRUNC(dbms_random.value(1,5),0)
  WHERE CURRENT OF Cur;

ELSIF (v_time = 'Before Flight' AND v_type = 'Inquire')
THEN
  UPDATE Customer_Care_fact
  SET Problem_severity = 2
  WHERE CURRENT OF Cur;

ELSIF (v_time = 'Before Flight' AND v_type = 'Feedback')
THEN
  UPDATE Customer_Care_fact
  SET Feedback_Rate= TRUNC(dbms_random.value(1,5),0)
  WHERE CURRENT OF Cur;
ELSIF (v_time = 'After Flight' AND v_type = 'Feedback')
THEN
  UPDATE Customer_Care_fact
  SET Feedback_Rate = TRUNC(dbms_random.value(1,5),0)
  WHERE CURRENT OF Cur;

ELSIF (v_time = 'After Flight' AND v_type = 'Inquire')
THEN
  UPDATE Customer_Care_fact
  SET Problem_severity = 4
  WHERE CURRENT OF Cur;
END IF;

END LOOP;
END;

```

SQL Queries to Answer Business Questions:

1- Which Flight types do frequent flyers take more depending on their tier?

```
SELECT fd.Flight_type, ff.Tier, COUNT(fa.Freq_key) AS Frequent_Flyers_number  
      FROM flights_activity_Fact fa  
      JOIN Flights_DIM fd  
        ON fa.Flight_key = fd.Flight_key  
      JOIN Frequent_flyers_Dim ff  
        ON ff.Freq_key = fa.Freq_key  
      GROUP BY fd.Flight_type, ff.Tier ;
```

FLIGHT_TYPE	TIER	FREQUENT_FLYERS_NUMBER
Non-stop	MidTier	2
Non-stop	Basic	4
Indirect	Basic	4
Indirect	WarriorTier	1
Direct	WarriorTier	4
Direct	Basic	2
Non-stop	WarriorTier	2

2- Where are the most used hotels for overnights?

```
SELECT hd.Country, hd.City, hd.Hotel_ID , count(fa.passenger_key)  
      FROM Hotels_DIM hd  
      JOIN flights_activity_Fact fa  
        ON hd.Hotel_key = fa.Hotel_key  
      GROUP BY hd.Country, hd.City, hd.Hotel_ID  
      ORDER BY hd.Country ;
```

COUNTRY	CITY	HOTEL_ID	COUNT(FA.PASSENGER_KEY)
New Zealand	Pleasant Point	H-60	1
Nicaragua	San José de Bocay	H-25	1
Nigeria	Kaltungo	H-94	2
Peru	Ferreñafe	H-41	1
Philippines	Villanueva	H-72	1
Russia	Khatanga	H-85	1
Russia	Tayturka	H-10	1

3- What is the average number of nights passengers stay in each hotel?

```
SELECT h.Hotel_ID, h.Stars, TRUNC(AVG(fa.overnight_stays),0)
AVG_number_of_nights
FROM Hotels_DIM h
JOIN flights_activity_Fact fa
ON fa.Hotel_key = h.Hotel_key
GROUP BY h.Hotel_ID, h.Stars
ORDER BY AVG_number_of_nights DESC;
```

HOTEL_ID	STARS	AVG_NUMBER_OF_NIGHTS
H-72	1	1
H-60	4	1
H-54	4	1
H-27	3	1

4- What is the percentage of passengers to aircraft seats Capacity for every flight in each aircraft If there were stop stations?

```
SELECT fa.Flight_key, fa.Segment_number, ac.Aircraft_key,
Round( (COUNT(fa.Passenger_key) / ac.Seats_Capacity)* 100,2) AS
Capacity_Percentage
FROM flights_activity_Fact fa
JOIN Aircrafts_Dim ac
ON fa.Aircraft_key = ac.Aircraft_key
GROUP BY fa.Flight_key, fa.Segment_number, ac.Aircraft_key, ac.Seats_Capacity
ORDER BY Capacity_Percentage DESC ;
```

FLIGHT_KEY	SEGMENT_NUMBER	AIRCRAFT_KEY	CAPACITY_PERCENTAGE
	2	12	0.9
	3	12	0.9
	4	12	0.9
	6	12	0.9
	1	14	0.87
	2	14	0.65
	4	14	0.65

5- What are the numbers of complaints based on their Problem severity per year?

```
SELECT d.Year, i.TYPE, i.Interaction_time, f.Problem_severity,  
       COUNT(f.Problem_Severity) AS "No. of Complaints", i.Description  
  FROM Interactions_DIM i  
 JOIN Customer_Care_fact f  
    ON i.Interaction_key = f.Interaction_key  
 JOIN Date_Dim d  
   ON d.Full_date = f.Reservation_Ticket_Date  
 WHERE i.TYPE = 'Complaints'  
 GROUP BY d.YEAR, i.TYPE, i.Interaction_time, f.Problem_severity, i.Description ;
```

	YEAR	TYPE	INTERACTION_TIME	PROBLEM_SEVERITY	No. of Complaints	DESCRIPTION
▶	2021	Complaints	After Flight	4	2	The customer writes down any problem he encountered after the trip
	2021	Complaints	Before Flight	2	2	The customer writes down any problem he encountered before the trip
	2021	Complaints	Within Flight	6	2	The customer writes down any problem he encountered during the trip

6- Which is the most severity interaction complaint for each year?

```
SELECT d.Year, i.TYPE, MAX(f.Problem_Severity)  
  FROM Interactions_DIM i  
 JOIN Customer_Care_fact f  
   ON i.Interaction_key = f.Interaction_key  
 JOIN Date_Dim d  
  ON d.Full_date = f.Reservation_Ticket_Date  
 WHERE i.TYPE = 'Complaints'  
 GROUP BY d.Year, i.TYPE ;
```

	YEAR	TYPE	MAX(F.PROBLEM_SEVERITY)
▶	2021	Complaints	6

7- Find the interaction type for inquiries and complaints and the employee who worked on it

```
SELECT i.TYPE, i.Interaction_time, f.Problem_severity, st.NAME AS Customer_Services_name  
FROM Interactions_DIM i  
JOIN Customer_Care_fact f  
ON i.Interaction_key = f.Interaction_key  
JOIN Staff_Dim st  
ON f.Staff_key = st.Staff_key  
WHERE i.TYPE = 'Complaints' OR i.Type = 'Inquire'  
ORDER BY i.TYPE;
```

TYPE	INTERACTION_TIME	PROBLEM_SEVERITY	CUSTOMER_SERVICES_NAME
Complaints	Within Flight	6	Jasun Yesichev
Complaints	After Flight	4	Drusi Witty
Complaints	Before Flight	2	Ike Bown
Complaints	After Flight	4	Scottie Basini-Gazzi
Inquire	Before Flight	2	Mareah Baudic
Inquire	Within Flight	6	Kirsteni Jenson
Inquire	Before Flight	2	Cloris Milington

8- What are the numbers of Feedback Ratings based on their degree of rate per year?

```
SELECT d.Year, i.TYPE, f.Feedback_Rate,  
COUNT(f.Feedback_Rate) AS "No. of Feedback Ratings", i.Description  
FROM Interactions_DIM i  
JOIN Customer_Care_fact f  
ON i.Interaction_key = f.Interaction_key  
JOIN Date_Dim d  
ON d.Full_date = f.Reservation_Ticket_Date  
WHERE i.TYPE = 'Feedback'
```

```

GROUP BY d.Year, i.TYPE, f.Feedback_Rate, i.Description
ORDER BY COUNT(f.Feedback_Rate) DESC ;

```

YEAR	TYPE	FEEDBACK_RATE	No. of Feedback Ratings	DESCRIPTION
► 2021	Feedback	3	2	customer feedback about the trip, services and offers
2021	Feedback	2	1	customer feedback about the trip, services and offers

9- Which Reservation Channel contributes more to the sales of the company?

```

SELECT ch.Channel_Type AS "Reservation Channel", SUM(fa.Net_price) AS
Total_sales
FROM flights_activity_Fact fa
JOIN Channels_Dim ch
ON fa.Channel_key = ch.channel_key
GROUP BY ch.Channel_Type
ORDER BY Total_sales DESC;

```

Reservation Channel	TOTAL_SALES
► Airline Website	357.12
Travel Agency	265

10- What are the profits of the company per Quarter?

```

SELECT d.Year, d.Quarter, SUM(fa.Profit) AS "Profit"
FROM flights_activity_Fact fa
JOIN Date_Dim d
ON fa.Reservation_Ticket_Date = d.full_date
GROUP BY d.Year, d.Quarter ;

```

YEAR	QUARTER	Profit
► 2021	Q3	-799.48
2021	Q4	-293.3
2020	Q3	-105.36
2021	Q1	-427.43

Indexes For the Data Warehousing

1. Introduction

Data warehouse solutions have recently grown in significance for decision-makers. Most of the queries against a large data warehouse are complex and iterative. The ability to answer these queries efficiently is a critical issue in the data warehouse environment. Indexing is a way to optimize the performance by minimizing the number of disk accesses required when a query is processed. If the right index structures are built on columns, the performance of queries, especially ad hoc queries, will be greatly enhanced. In this report, we discuss indexing techniques being used. Indexing in a data warehouse may be tricky; too few indexes will result in quick data loads but delayed query responses, while too many indexes will affect data loads slow down and storage requirements go up, but query response is good.

2. Factors used to determine which indexing technique should be built on a Column Characteristics of indexed column

- A. A column has its own characteristics which we can use to choose a proper index. These characteristics are given below:
 - Cardinality data: The cardinality data of a column is the number of distinct values in the column. It is better to know that the cardinality of an indexed column is low or high since an indexing technique may work efficiently only with either low cardinality or high cardinality.
 - Distribution: The distribution of a column is the occurrence frequency of each distinct value of the column. The column distribution guides us to determine which index type we should take.
 - Value range: The range of values of an indexed column guides us to select an appropriate index type. For example, if the range of a high cardinality column is small, an indexing technique based on bitmap should be used. Without knowing this information, we might use a B-Tree resulting in a degradation of system performance.

- B. Understanding the data and the usage in the SQL language. Knowing the columns that will be queried helps us choose appropriate index types for them. For example, which columns will likely be a part of the selection list, join constraints, application constraints, the ORDER BY clause, or the GROUP BY clause?

3. Indexing Techniques

3.1 B+-Tree indexes

The most used data structures to improve query processing time in operational databases. A B+-Tree indexes on an attribute C of a relation R is a balanced tree providing associative access to the tuples of R on the basis of the values of the C key. The leaves of this tree are linked to each other and store the pointers to the disk blocks containing tuples. Internal nodes create a sort of map to find key values quickly. The top level of the index is called the root. The lowest level is called the leaf node. All other levels in between are called branches. Both the root and branch contain entries that point to the next level in the index. Leaf nodes consisting of the index key and pointers pointing to the physical location in which the corresponding records are stored.

3.2 Bitmap Index

The bitmap representation is an alternate method of the row ids representation. It is simple to represent and uses less space and CPU efficient than row ids. when the number of distinct values of the indexed column is low. The indexes improve complex query performance by applying low-cost Boolean operations such as OR, AND, and NOT in the selection predicate on multiple indexes at one time to reduce search space before going to the primary source data. Many variations of the Bitmap Index (Pure Bitmap Index, Encoded Bitmap, etc.) aiming to reduce space requirement as well as improve query performance.

<pre>SELECT Customers FROM Base Table WHERE Region = W AND Rating = L</pre>	<p>Base Table</p> <table border="1" style="display: inline-table; vertical-align: top;"> <thead> <tr><th>Cust</th><th>Region</th><th>Rating</th><th>...</th></tr> </thead> <tbody> <tr><td>C1</td><td>N</td><td>H</td><td>...</td></tr> <tr><td>C2</td><td>S</td><td>M</td><td>...</td></tr> <tr><td>C3</td><td>W</td><td>L</td><td>...</td></tr> <tr><td>C4</td><td>W</td><td>H</td><td>...</td></tr> <tr><td>C5</td><td>S</td><td>L</td><td>...</td></tr> <tr><td>C6</td><td>W</td><td>L</td><td>...</td></tr> <tr><td>C7</td><td>W</td><td>H</td><td>...</td></tr> </tbody> </table>	Cust	Region	Rating	...	C1	N	H	...	C2	S	M	...	C3	W	L	...	C4	W	H	...	C5	S	L	...	C6	W	L	...	C7	W	H	...	<p>Region Index</p> <table border="1" style="display: inline-table; vertical-align: top;"> <thead> <tr><th>RowId</th><th>N</th><th>S</th><th>E</th><th>W</th></tr> </thead> <tbody> <tr><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>2</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>3</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>4</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>5</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>6</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>7</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> </tbody> </table>	RowId	N	S	E	W	1	1	0	0	0	2	0	1	0	0	3	0	0	0	1	4	0	0	0	1	5	0	1	0	0	6	0	0	0	1	7	0	0	0	1	<p>Rating Index</p> <table border="1" style="display: inline-table; vertical-align: top;"> <thead> <tr><th>RowId</th><th>H</th><th>M</th><th>L</th></tr> </thead> <tbody> <tr><td>1</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>2</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>3</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>4</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>5</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>6</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>7</td><td>1</td><td>0</td><td>0</td></tr> </tbody> </table>	RowId	H	M	L	1	1	0	0	2	0	1	0	3	0	0	1	4	1	0	0	5	0	0	1	6	0	0	1	7	1	0	0
Cust	Region	Rating	...																																																																																																								
C1	N	H	...																																																																																																								
C2	S	M	...																																																																																																								
C3	W	L	...																																																																																																								
C4	W	H	...																																																																																																								
C5	S	L	...																																																																																																								
C6	W	L	...																																																																																																								
C7	W	H	...																																																																																																								
RowId	N	S	E	W																																																																																																							
1	1	0	0	0																																																																																																							
2	0	1	0	0																																																																																																							
3	0	0	0	1																																																																																																							
4	0	0	0	1																																																																																																							
5	0	1	0	0																																																																																																							
6	0	0	0	1																																																																																																							
7	0	0	0	1																																																																																																							
RowId	H	M	L																																																																																																								
1	1	0	0																																																																																																								
2	0	1	0																																																																																																								
3	0	0	1																																																																																																								
4	1	0	0																																																																																																								
5	0	0	1																																																																																																								
6	0	0	1																																																																																																								
7	1	0	0																																																																																																								
	<p>Base Table</p> <table border="1" style="display: inline-table; vertical-align: top;"> <thead> <tr><th>Cust</th><th>Region</th><th>Rating</th></tr> </thead> <tbody> <tr><td>C1</td><td>N</td><td>H</td></tr> <tr><td>C2</td><td>S</td><td>M</td></tr> <tr><td>C3</td><td>W</td><td>L</td></tr> <tr><td>C4</td><td>W</td><td>H</td></tr> <tr><td>C5</td><td>S</td><td>L</td></tr> <tr><td>C6</td><td>W</td><td>L</td></tr> <tr><td>C7</td><td>W</td><td>H</td></tr> </tbody> </table>	Cust	Region	Rating	C1	N	H	C2	S	M	C3	W	L	C4	W	H	C5	S	L	C6	W	L	C7	W	H	<p>Region Index</p> <table border="1" style="display: inline-table; vertical-align: top;"> <thead> <tr><th>RowId</th><th>N</th><th>S</th><th>E</th><th>W</th></tr> </thead> <tbody> <tr><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>2</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>3</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>4</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>5</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>6</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>7</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> </tbody> </table>	RowId	N	S	E	W	1	1	0	0	0	2	0	1	0	0	3	0	0	0	1	4	0	0	0	1	5	0	1	0	0	6	0	0	0	1	7	0	0	0	1	<p>Rating Index</p> <table border="1" style="display: inline-table; vertical-align: top;"> <thead> <tr><th>RowId</th><th>H</th><th>M</th><th>L</th></tr> </thead> <tbody> <tr><td>1</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>2</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>3</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>4</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>5</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>6</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>7</td><td>1</td><td>0</td><td>0</td></tr> </tbody> </table>	RowId	H	M	L	1	1	0	0	2	0	1	0	3	0	0	1	4	1	0	0	5	0	0	1	6	0	0	1	7	1	0	0								
Cust	Region	Rating																																																																																																									
C1	N	H																																																																																																									
C2	S	M																																																																																																									
C3	W	L																																																																																																									
C4	W	H																																																																																																									
C5	S	L																																																																																																									
C6	W	L																																																																																																									
C7	W	H																																																																																																									
RowId	N	S	E	W																																																																																																							
1	1	0	0	0																																																																																																							
2	0	1	0	0																																																																																																							
3	0	0	0	1																																																																																																							
4	0	0	0	1																																																																																																							
5	0	1	0	0																																																																																																							
6	0	0	0	1																																																																																																							
7	0	0	0	1																																																																																																							
RowId	H	M	L																																																																																																								
1	1	0	0																																																																																																								
2	0	1	0																																																																																																								
3	0	0	1																																																																																																								
4	1	0	0																																																																																																								
5	0	0	1																																																																																																								
6	0	0	1																																																																																																								
7	1	0	0																																																																																																								
	<p>Base Table</p> <table border="1" style="display: inline-table; vertical-align: top;"> <thead> <tr><th>Cust</th><th>Region</th><th>Rating</th></tr> </thead> <tbody> <tr><td>C1</td><td>N</td><td>H</td></tr> <tr><td>C2</td><td>S</td><td>M</td></tr> <tr><td>C3</td><td>W</td><td>L</td></tr> <tr><td>C4</td><td>W</td><td>H</td></tr> <tr><td>C5</td><td>S</td><td>L</td></tr> <tr><td>C6</td><td>W</td><td>L</td></tr> <tr><td>C7</td><td>W</td><td>H</td></tr> </tbody> </table>	Cust	Region	Rating	C1	N	H	C2	S	M	C3	W	L	C4	W	H	C5	S	L	C6	W	L	C7	W	H	<p>Region Index</p> <table border="1" style="display: inline-table; vertical-align: top;"> <thead> <tr><th>RowId</th><th>N</th><th>S</th><th>E</th><th>W</th></tr> </thead> <tbody> <tr><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>2</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>3</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>4</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>5</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>6</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>7</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> </tbody> </table>	RowId	N	S	E	W	1	1	0	0	0	2	0	1	0	0	3	0	0	0	1	4	0	0	0	1	5	0	1	0	0	6	0	0	0	1	7	0	0	0	1	<p>Rating Index</p> <table border="1" style="display: inline-table; vertical-align: top;"> <thead> <tr><th>RowId</th><th>H</th><th>M</th><th>L</th></tr> </thead> <tbody> <tr><td>1</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>2</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>3</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>4</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>5</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>6</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>7</td><td>1</td><td>0</td><td>0</td></tr> </tbody> </table>	RowId	H	M	L	1	1	0	0	2	0	1	0	3	0	0	1	4	1	0	0	5	0	0	1	6	0	0	1	7	1	0	0								
Cust	Region	Rating																																																																																																									
C1	N	H																																																																																																									
C2	S	M																																																																																																									
C3	W	L																																																																																																									
C4	W	H																																																																																																									
C5	S	L																																																																																																									
C6	W	L																																																																																																									
C7	W	H																																																																																																									
RowId	N	S	E	W																																																																																																							
1	1	0	0	0																																																																																																							
2	0	1	0	0																																																																																																							
3	0	0	0	1																																																																																																							
4	0	0	0	1																																																																																																							
5	0	1	0	0																																																																																																							
6	0	0	0	1																																																																																																							
7	0	0	0	1																																																																																																							
RowId	H	M	L																																																																																																								
1	1	0	0																																																																																																								
2	0	1	0																																																																																																								
3	0	0	1																																																																																																								
4	1	0	0																																																																																																								
5	0	0	1																																																																																																								
6	0	0	1																																																																																																								
7	1	0	0																																																																																																								
		Region = W	AND	Rating = L																																																																																																							

3.3 Join and Star Indexes

Even though star schemata explicitly aim at reducing the number of joins necessary to retrieve data, joins still mainly determine the cost of OLAP queries. Researchers have been working for a long time to reduce the join execution time. To do this, they have developed more efficient join algorithms.

A join index between two relations R_1 and R_2 , stores couples of tuples RIDs that fulfill the $C_1 <\text{operator}> C_2$: join predicate between two attributes $C_1 \in R_1$ and $C_2 \in R_2$.

Join indexes strongly improve performance because they preprocess and explicitly store RIDs of the matching tuples in two relations. To check for the tuples to fulfill your join predicate, you no longer need to check all the combinations, but you need to scan your index.

4. References

Book:

Matteo Golfarelli, *Data Warehouse Design, Modern Principles and Methodologies*

Paper:

Indexing Techniques for Data Warehouses' Queries

A Comparative Study of Indexing Techniques in Data Warehouse

Indexing Techniques to Enhance the Performance of Data Warehouses

The impact of indexes on data warehouse performance

Website:

https://docs.oracle.com/cd/B12037_01/server.101/b10736/indexes.htm#i1006549