|  |
| --- |
|  |
| DW Project DocumentationTeam MembersAbdulrahman HatemAbdullah GalalNajat KhalafNada Nady |
|  |
| First of all, we will define The**Business Process** |
| We are modeling a DW model for a major airline to assist the executive management to analyze their current business processes and expand the company by discovering new opportunities, and this will be approached through some points we take care of: |
| * Flight Activity      * In this process will analyze flight activity including data on routes, passenger numbers. The report would also include an analysis of trends and patterns in flight activity, as well as recommendations for optimizing the airline's operations. such as number of passenger, empty seats, number of crew , flight ID, aircraft ID, airport ID , captain id, flight attendant. |

* Marketing Analysis
* 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.
* Reservations
* Mainly focus on Profit for particular (Customers ID, Channel ID, Flight ID, FBC ID, Payment Method , Date ID )
* Customer Care

This process focus on

* Handling customer inquiries, complaints and feedbacks
* Interaction type and problem severity

### Grains

## Flight Fact

Per flight (Aircraft ID, route, source airport, destination airport, departure time, arrival time, captain, and flight)

## Marketing Analysis

Per each customer , from the perspective of the marketing team , they are focusing on customer "lowest level", is he upgrade , uses the promotions and discounts , how often he travels, at the end the marketing team can take action.

## Reservation

Here we use it for each Ticket, which contain (customer ID, flight ID, aircraft ID , route ID , from a certain airport at a given departure time, to an arrival airport at a specific arrival time, fare class type, channel ID , and a payment method)

## Customer Care

For the level of each complain we can track the action, feedback, problem severity also for specific customer, on a specific flight, channel, payment method, with a specific crew members, and exact date.

### Facts & Measurements

1.Flights Fact

Using this fact, we can keep track of each flight and gather information that will enable us to identify which flights are most popular, also which (individual aircraft, route, source airport, destination airport, departure time, arrival time, captain, and flight attendant)

* 1. Measurements:
     + Number of passengers.
     + Number of empty seats.
     + Number of the crew.
  2. Dimensions:
* Crew Member
* Aircraft
* Airport
* Flight

### Facts & Measurements

2. Marketing Analysis Fact

We can say that it is factless fact, no measurements have been calculated, we only use it to help Marketing team to analyze the frequent customers, are they using promotions, miles bonus?

1.Measurements:

* + - No measurements

2.Dimensions:

* Flight date
* Promotion
* Flyer miles
* Upgrade
* Customer
* Flight

### Facts & Measurements

3.Reservation Fact

Using this fact, we can extract which customers brings the most value, which date "season"/flight/airplane/airport makes the best profit, which reservation channel and which payment method delivers the most money.

1.Measurements:

* + - Price.
    - Discounts.
    - Overnight stand.

2.Dimensions:

* Customer
* Payment method
* Channel
* Flight date
* Fare class
* Flight

### Facts & Measurements

4.Customer Care Fact

Using this fact, we keep up with information about customer satisfaction and suggests solutions to enhance business success by attending to customers' requirements and categorize the problems.

1.Measurements:

* + - Severity type.

2.Dimensions:

* Customer
* Flight
* Payment method
* Channel
* Flight date
* Action
* Crew members

### Dimensions

1. Crew Member

* Data about all the employees in the company.

1. Customer

* Data about the passengers.

1. Date

* Calendar dimension for any DW Model.

1. Channel

* The channels are the source for tickets’ reservation.

1. Payment Method

* The different methods of payment.

1. Fare Class

* The tickets classes (Economy, Business, etc).

1. Flyer Miles

* Holding miles' number of the flight.

### Dimensions

1. Promotion

* Data about the discounts that the company offer.

1. Upgrade

* Data about types of upgrades the company offer.

1. Aircraft

* Data about the airplanes the company owns.

1. Airport

* Data about the airports that the aircrafts use.

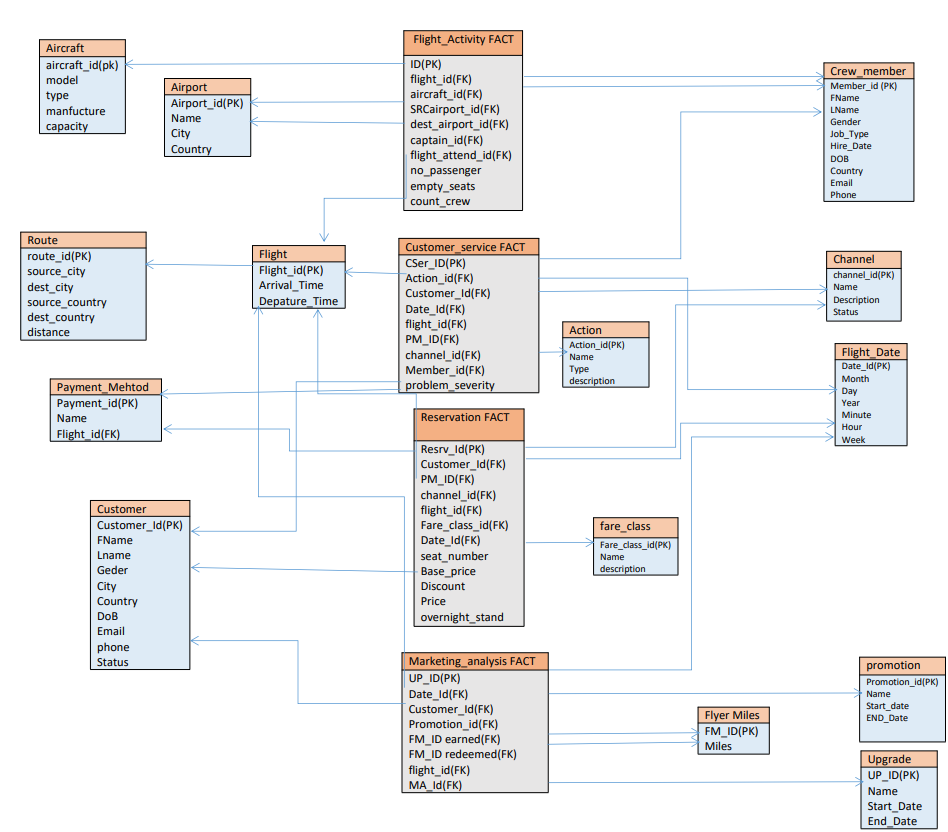
1. Flight

* Data about the specific times for each flight.

1. Route

* Data about the details for each flight.

### Schema



* Why we used Star schema?

We find that the star schema is suitable for our case study and perform better in implementation, after searching also we find that star schema is better for many reasons like:

* Streamlined query execution: Star schema modeling enables us to streamline and improve query execution. Queries may be run effectively and rapidly without the use of intricate joins or subqueries since the fact table serves as the foundation of the schema and is linked to the dimension tables by foreign keys.
* Star schema modelling is very scalable and has a great data handling capacity. The schema can handle massive volumes of data without affecting query performance or data analysis by dividing the data into smaller, more manageable tables.
* Better data analysis: Star schema modelling offers a clear and simple method for analyzing data from several aspects. By traversing through the dimension tables, analysts can quickly drill down into the data, enabling more intricate analysis and insights.

### Indexes

For Indexes, We used

1. Cluster index on each PK columns " By default on Toad".
2. B-tree Index on FK " idx\_flight\_activity\_flight\_id ON hr.flight\_activity(flight\_id);"
3. Unique Index on E-mail column.

### Types of Indexes in Oracle

Oracle defines two types of indexes: the B-Tree (Balanced Tree) Index and the Bitmap Index.

|  |  |
| --- | --- |
| B-tree | Bitmap |
| * Index-organized tables   An index-organized table differs from a heap-organized because the data is itself the index. | * Function-based indexes   This type of index includes columns that are either transformed by a function, such as the UPPER function, or included in an expression. B-tree or bitmap indexes can be function-based. |
| * Reverse key indexes   In this type of index, the bytes of the index key are reversed, for example, 103 is stored as 301. The reversal of bytes spreads out inserts into the index over many blocks. | * Application domain indexes   A user creates this type of index for data in an application-specific domain. The physical index need not use a traditional index structure and can be stored either in the Oracle database as tables or externally as a file. |
| * UNIQUE Descending indexes   This type of index stores data on a particular column or columns in descending order. |  |
| * B-tree cluster indexes |

### Sample of SQL queries

## 1-Which customers use our services more frequent ?

select \* from(

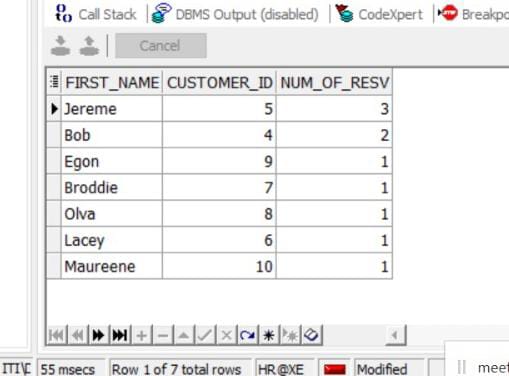
select first\_name ,r.customer\_id, count(r.customer\_id) num\_of\_resv

from hr.reservation r,hr.customer c

where r.customer\_id=c.customer\_id

group by first\_name,r.customer\_id

)order by num\_of\_resv desc;



## 2-What are the most popular routes/aircrafts ?

select \* from (

select model,air.aircraft\_id ,count(air.aircraft\_id) aircraft\_freq

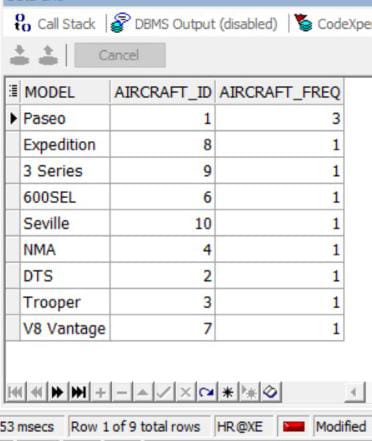
from hr.flight\_activity f,hr.aircraft air

where air.aircraft\_id = f.aircraft\_id

group by air.aircraft\_id,air.model

)

order by aircraft\_freq desc;



## 3-What are the most popular destcountry?

select \* from (

select r.DEST\_COUNTRY ,r.route\_id ,count(f.route\_id) as freq\_country

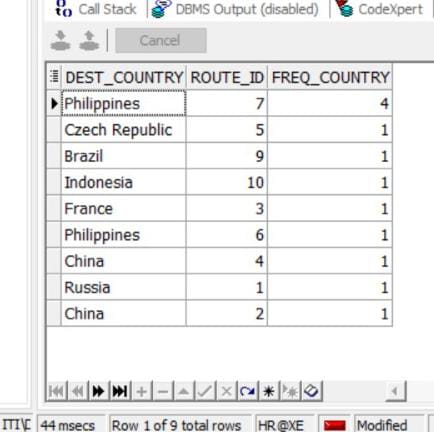
from hr.flight f,hr.route r

where f.route\_id=r.route\_id

group by r.route\_id ,r.DEST\_COUNTRY

)

order by freq\_country desc;



## 4-Which booking channel is most rewarding?

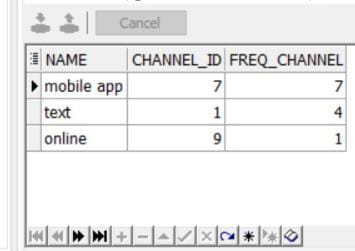
select ch.name,ch.channel\_id ,count(r.channel\_id) freq\_channel

from hr.reservation r ,hr.channel ch

where r.channel\_id=ch.channel\_id

group by ch.CHANNEL\_ID ,ch.name

order by freq\_channel desc;



## 5-Which year have more reservations?

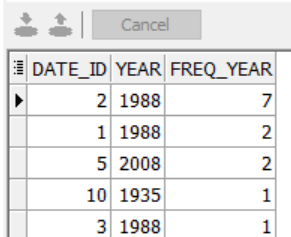
select fd.date\_id,year,count(fd.date\_id) freq\_year

from hr.flight\_date fd,  hr.reservation r

where r.date\_id=fd.date\_id

group by fd.date\_id ,fd.year

order by freq\_year desc;



## 6-Which payment method most popular ?

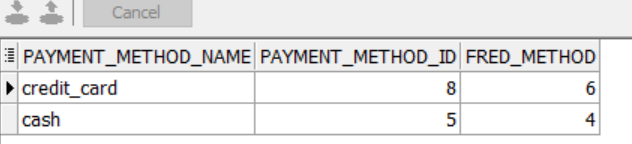
select pm. PAYMENT\_METHOD\_NAME,pm.PAYMENT\_METHOD\_id,count(r.PM\_ID) fred\_method

from hr.reservation r,hr.payment\_method pm

where r.pm\_id=pm.PAYMENT\_METHOD\_id

group by pm.PAYMENT\_METHOD\_id,pm. PAYMENT\_METHOD\_NAME

order by fred\_method desc;



## 7-How much miles they earn ?

select first\_name, m.customer\_id,sum(miles) earn\_miles

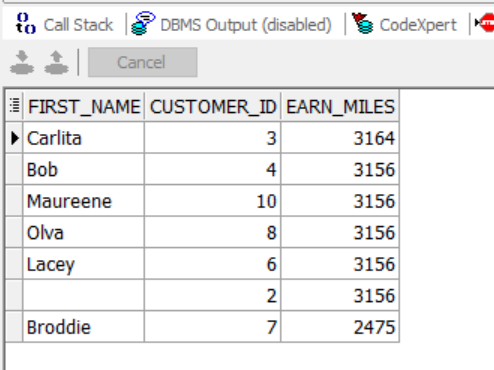
from hr.marketing\_analysis M ,hr.Flyer\_Miles F, hr.customer c

where M.Flyer\_miles\_earned\_id = f.Flyer\_miles\_id and

m.customer\_id =c.customer\_id

group by m.customer\_id,first\_name

order by earn\_miles desc;

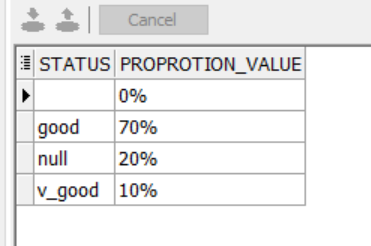


## 8-Proprotion of membership status passengers

select status, count(status) /(select count(status) from hr.customer )\*100||'%' proprotion\_value

from hr.customer

group by status;



## 9-Profit per year

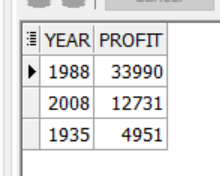
select year, sum(price) as profit

from hr.reservation r ,hr.flight\_date  d

where r.date\_id =d.date\_id

group by year

order by profit desc;



## 10-Profit from every reservation channel

select name,sum(price) as profit

from hr.reservation r ,hr.channel ch

where r.channel\_id=ch.channel\_id

group by name

order by profit desc;

