

# Airline

# Data warehouse

Project by :

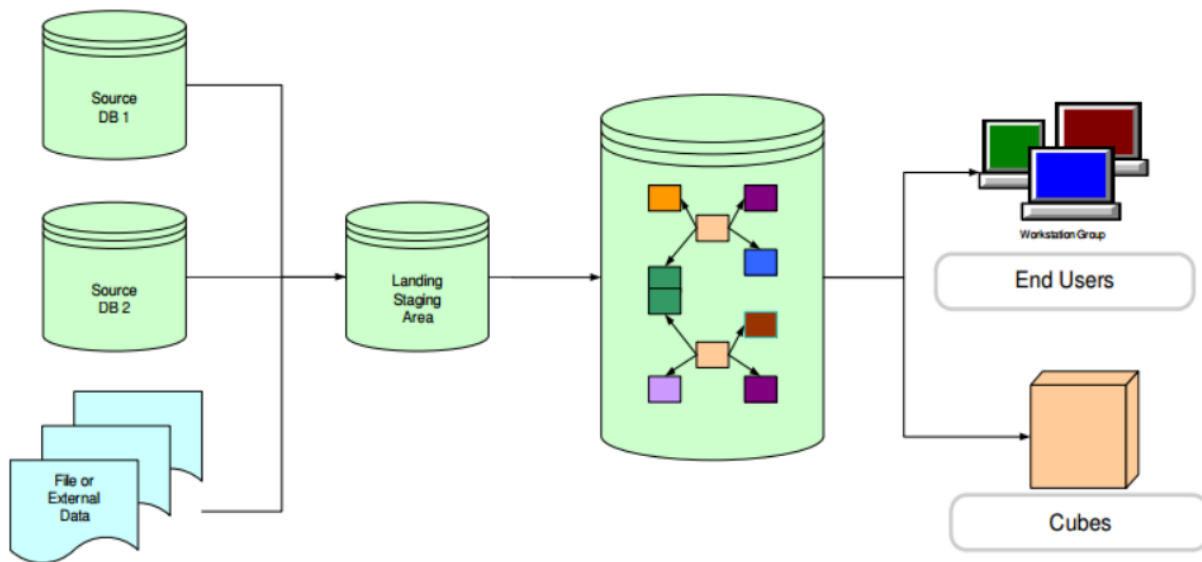
- Yousra Mohamed
- Abdelaziz amin
- Ibrahim elsayed

## Airline data warehouse model

For entire the business analysis system of our case we considered that we have a little changes in requirements of different departments so we decided to choose Kimball approach (2-layer) in our data warehouse model

Why?

- It will give us easier understanding in case the business needs to read the entire process and relations of data elements
- Achieve the business needs, we already know the requirements so we developed the system over these reqs
- Speed , no much effort to understand everything in a unified place like(enterprise DWH) to analyze new business process, we just need to identify how sources will interact.



### Sources used in business understanding process :

- [Airline Pricing Strategies Explained: Types, Examples & Tips | Pricefx](#)
- [Taxes and Fees on Flights: What You Need to Know | Prince of Travel](#)

# Executive's data mart

## Identification steps

### 1- Business process: **flights monitoring**

- For executives we assume that they would like to see how flights is going and what is the rate of flights per day, week or month, also they may be interested in total revenue of the flight and if the flight has landed well or faced diversion due to weather conditions.
- They may be interested in how many flights has been cancelled over a period of time
- Also Occupancy rate of the flight is an important aspect to be measured

### 2- Granularity: **one row per flight**

- Level of granularity choosing depended on the measures the executives would like to see, these measures could be result from level of granularity we have identified.

### 3- Dims:

#### a. **Airport:** **role playing dimension** (participate with different views)

- airport\_id : surrogate key for identifying every airport
- name : name of the airport
- country
- city
- time zone
- num\_of\_gates : number of gates or terminals of the airport

#### b. **Plane:**

- plane\_id : surrogate key for identifying every plane
- engine\_num : a natural key like engine number in cars
- model : manufacturer model

#### c. **Cancellation reasons:**

- reason\_id : surrogate key for identifying every cancellation reason
- category : like(technical issue – weather conditions –reservations not enough )
- desc

#### d. **flight:** **degenerated dimension**

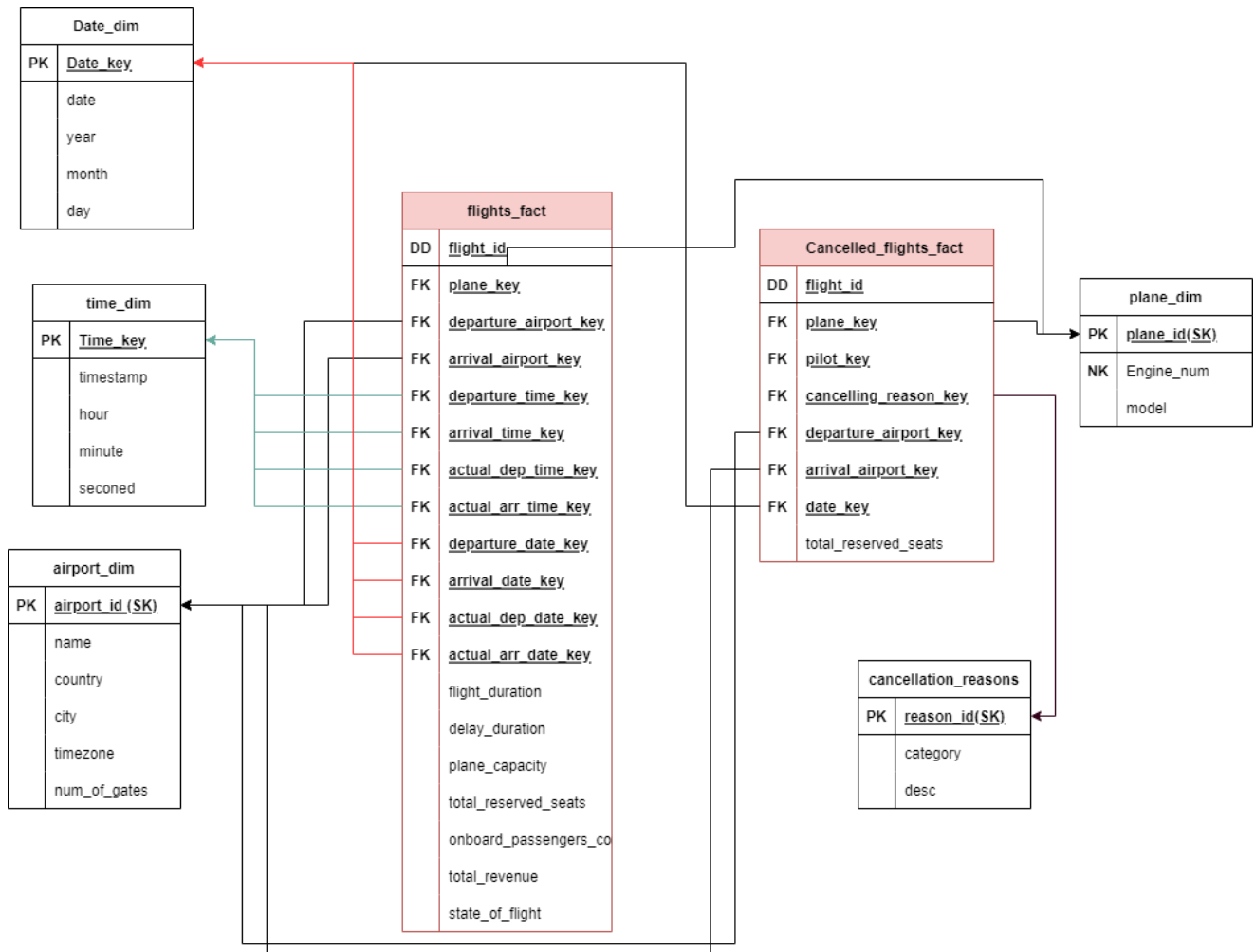
#### e. **date :** **role playing dimension** (participate with different views)

#### f. **time :** **role playing dimension** (participate with different views)

### 4- Measures:

- **Total\_reserved\_seats** : number of reserved seats in the plane which is indicator to the rate of occupancy of the plane capacity
- **Flight duration** : time of the flight takes from departure to arrival
- **Delay duration** : time from the scheduled departure time to the actual time
- **Plane capacity** : full capacity of the plane
- **Total reserved seats** : the number of seats reserved by clients
- **Onboard passengers count** : number of actual passengers in the flight
- **Total revenue** : revenue of the flight = sum of all payments by clients in the flight
- **State of flight** : indicator if flight landed well or faced some issues in the sky or during landing

## Model:



# Marketing data mart

## Identification steps

### 1- **Business process:** client reservation process

- For the marketing team they are interested to follow the actions of the client during the reservation process from the first step till onboarding
- They would like to know how the client will reserve his ticket and if he is a frequent flyer, would he use his points or miles in reservation, how he will interact with discounts during specific periods of the year

### 2- **Granularity:** one row per client reservation

- Level of granularity choosing depended on the measures the marketing would like to see so they have interest in client behavior.

### 3- **Dims:**

#### a. **Airport:** role playing dimension (participate with different views)

- airport\_id : surrogate key for identifying every airport
- name : name of the airport
- country
- city
- time zone
- num\_of\_gates : number of gates or terminals of the airport

#### b. **flight:** it will serve need of analyze client behavior according to specific flights with respect to a specific airport

- flight\_id : surrogate key for identifying every flight
- departure airport key : foreign key reference to airport dim
- arrival airport key : : foreign key reference to airport dim
- departure, arrival, actual dep and actual arrival time

#### c. **fare base dim:** fare base is the adjusted price of the ticket that client will pay during reservation and it has some aspects that it depends on

- fare\_base\_id : surrogate key for identifying every fare base
- type : like(business-A , business-B , economy-J, economy-K )
- base miles : describe the miles that
- class bonus :
- total miles earned perc : percentage of total flight miles that client would get in frequent flyer system miles
- qualification dollars percentage : percentage of paid dollars that client will be rewarded in the frequent flyer system charge

#### d. **Booking channel :**

- Booking\_channel\_key : surrogate key for identifying every booking channel
- type : like (agency – website – call center )

e. **payment method** :

- payment\_method\_key : surrogate key for identifying every payment method
- type : like (cash – visa )

f. **ticket**: **degenerated dimension**

g. **date** : **role playing dimension** (participate with different views)

h. **time** : **role playing dimension** (participate with different views)

i. **Class**: classes of the flight

- class\_id : surrogate key for identifying every class
- type : like (first class – business – economy )
- desc : full description of the class

j. **Tiers**: it describes the different tiers of the frequent flyers of the airline

- tier\_id : surrogate key for identifying every tier
- type : like (gold – silver – platinum )
- miles\_perc: earned miles percentage of flight miles
- upgrade miles perc : percentage of miles earned when upgrade
- additional luggage allowance : kg allowance for tier

k. **Frequent flyer** : **slowly changing dimension**

Tracks changing in frequent flyers tiers and accumulated score of miles and dollars

- Frequent flyer id : surrogate key for identifying every tier
- Current\_tier\_id : a foreign key to tiers dimension
- Current\_miles\_num : total earned miles score
- Current qualification dollars amount : total earned dollars score
- Start date : current transaction inserting date
- End date : transaction end date , if null means it is the latest update for the flyer

l. **client** :

- client\_id : surrogate key for identifying every client
- passport id : natural key for the client
- name
- gender
- date of birth
- nationality
- contact number

m. **client mapping dim** : **mapping dimension**

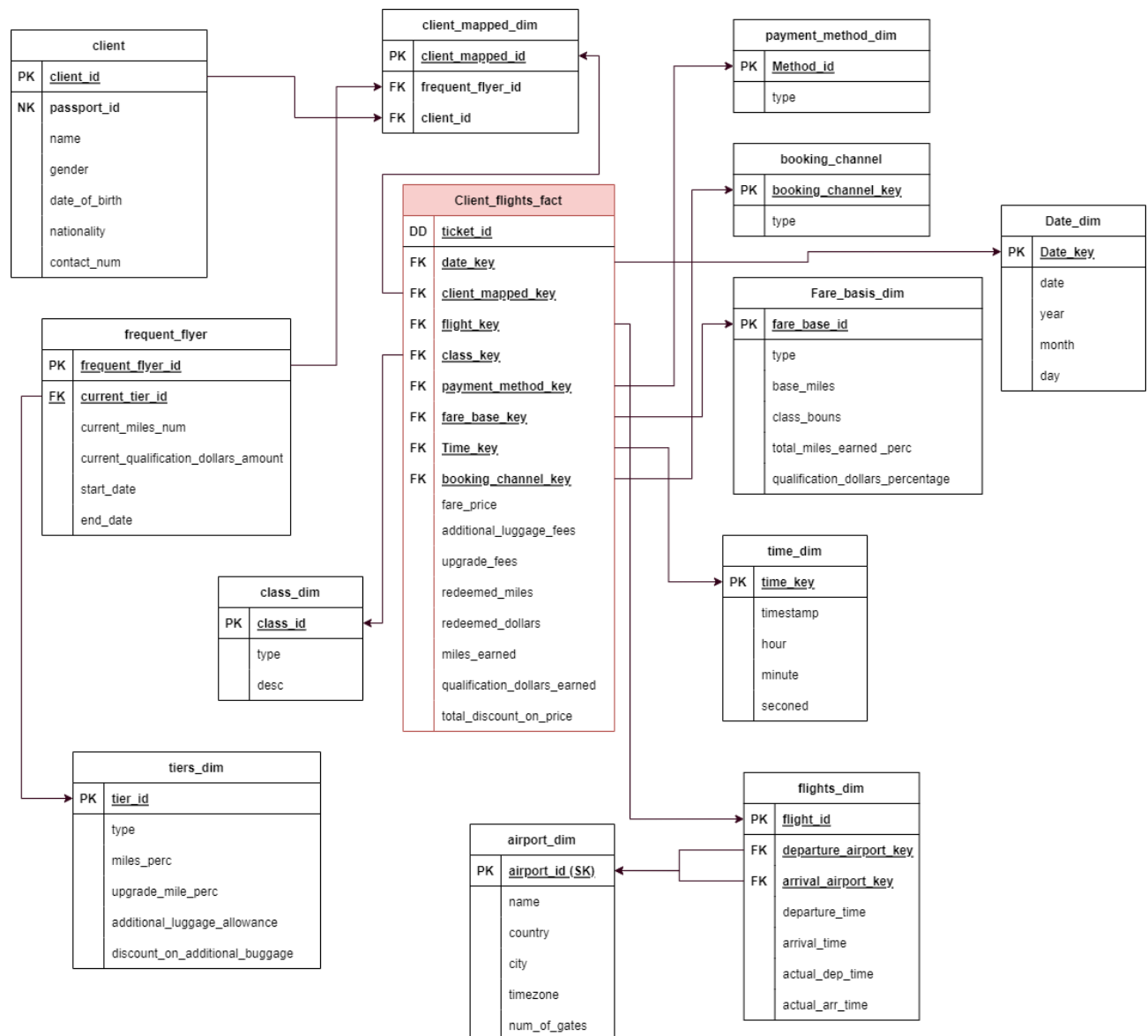
if the normal client become a frequent flyer (SCD) so we will have the same client in two different representation key in the same model so we used mapping dim to link these two ids of the same client

- client mapped id : surrogate key for identifying every client (this will be involved in fact table)
- Frequent flyer id : foreign key references frequent flyer dim
- client id : foreign key references clients dim

#### 4- Measures:

- **Fare price** : the price that client will pay for basic package of the fare base type
- **Additional luggage fees** : cost of additional luggage
- **Upgrade fess** :
- **Redeemed miles** : in case of client redeemed form his score in miles
- **Redeemed dollars** : in case of client redeemed form his dollars charge
- **Miles earned** : miles earned according to tier and rewarding program
- **Qualification dollars earned** : dollars earned according to tier and rewarding program
- **Total discount on price** : the deduction from the fare base price

#### Data Model:



## Finance data mart

### Identification steps

#### 1- **Business process:** client reservation

- For the finance team they would be interested in analyzing financial elements in the process of ticket reservation for the client
- Taxes – additional costs – airport fees – facilities fess – actual price paid after discounts

#### 2- **Granularity:** one row per client reservation

- Level of granularity choosing depended on the measures the finance would like to see so they have interest in client reservation cost analysis which nclude regular cost and additional costs

#### 3- **Dims:**

##### a. **Airport:** role playing dimension (participate with different views)

- airport\_id : surrogate key for identifying every airport
- name : name of the airport
- country
- city
- time zone
- num\_of\_gates : number of gates or terminals of the airport

##### b. **flight:** it will serve need of analyze client behavior according to specific flights with respect to a specific airport

- flight\_id : surrogate key for identifying every flight
- departure airport key : foreign key reference to airport dim
- arrival airport key : : foreign key reference to airport dim
- departure, arrival, actual dep and actual arrival time

##### c. **fare base dim:** fare base is the adjusted price of the ticket that client will pay during reservation and it has some aspects that it depends on

- fare\_base\_id : surrogate key for identifying every fare base
- type : like(business-A , business-B , economy-J, economy-K )
- base miles : describe the miles that
- class bonus :
- total miles earned perc : percentage of total flight miles that client would get in frequent flyer system miles
- qualification dollars percentage : percentage of paid dollars that client will be rewarded in the frequent flyer system charge

##### d. **Booking channel :**

- Booking\_channel\_key : surrogate key for identifying every booking channel
- type : like (agency – website – call center )



e. payment method :

- payment\_method\_key : surrogate key for identifying every payment method
- type : like (cash – visa )

f. ticket: **degenerated dimension**

g. date : **role playing dimension** (participate with different views)

h. time : **role playing dimension** (participate with different views)

i. Class: classes of the flight

- class\_id : surrogate key for identifying every class
- type : like (first class – business – economy )
- desc : full description of the class

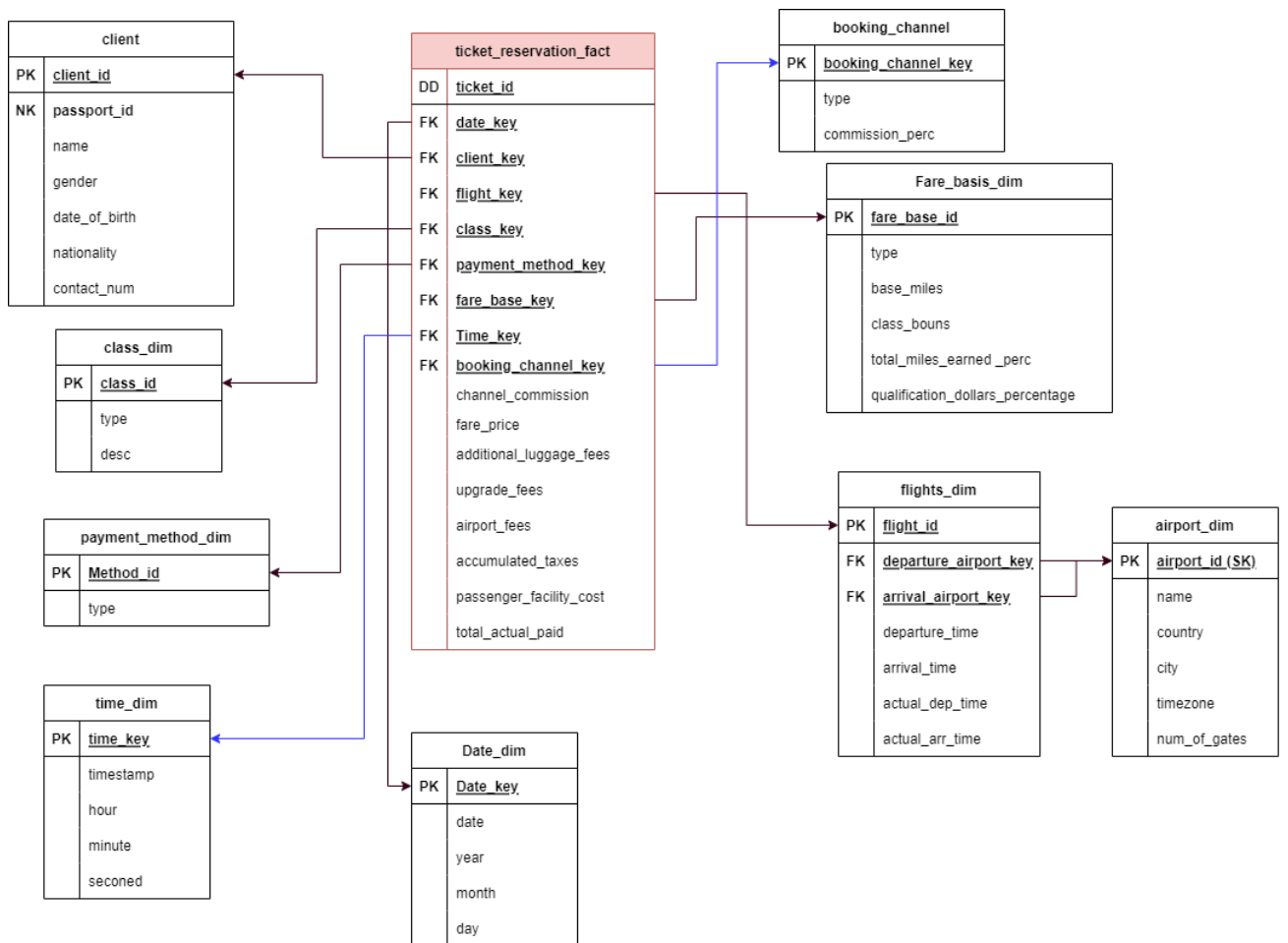
j. client :

- client\_id : surrogate key for identifying every client
- passport id : natural key for the client
- name
- gender
- date of birth
- nationality
- contact number

**4- Measures:**

- **Channel Commission** : in case of client reserve a ticket through outsource agency
- **Fare price** : the price that client will pay for basic package of the fare base type
- **Additional luggage fees** : cost of additional luggage
- **Upgrade fess**
- **Airport fees** : fees of airport the where flight will land
- **Accumulated taxes**: total governmental taxes over the client ticket
- **Passenger facility cost** : cost of airport facility for the client
- **Total actual paid** : the final price that client has paid for the ticket

## Data Model:



## Customer care data mart

### Identification steps

- 1- **Business process:** monitor actions (complaints – inquiries - rating )
  - For the customer care team, the business process will be any kind of communication type between customer(client) and the airline
- 2- **Granularity:** one row per client communication event
  - Communication event refers to complaint, inquiry or rating transaction
- 3- **Dims:**
  - a. **Airport:** role playing dimension (participate with different views)
    - airport\_id : surrogate key for identifying every airport
    - name : name of the airport
    - country
    - city
    - time zone
    - num\_of\_gates : number of gates or terminals of the airport
  - b. **flight:** it will serve need of analyze client complaints according to specific flights with respect to a specific airport
    - flight\_id : surrogate key for identifying every flight
    - departure airport key : foreign key reference to airport dim
    - arrival airport key : : foreign key reference to airport dim
    - departure, arrival, actual dep and actual arrival time
  - c. **channel of interaction :**
    - channel id : surrogate key for identifying every interaction channel
    - type : like (website – email – call center - app )
  - d. **interaction type:**
    - interaction type id : surrogate key for identifying every interaction type
    - type : like (complaint – inquiry )
    - severity : how urgent this type is (urgent – normal )
  - e. **case state dim:** describe state of the transaction
    - case state id : surrogate key for identifying every state type
    - state : like (finished – waiting – cancelled )
    - priority : (high – medium - low )
  - f. **issue ticket id :** degenerated dimension
    - it identifies every transaction in the fact and like in call center business
  - g. **date :** role playing dimension (participate with different views)
  - h. **time :** role playing dimension (participate with different views)

i. **Class**: classes of the flight

- class\_id : surrogate key for identifying every class
- type : like(first class – business – economy )
- desc : full description of the class

j. **client** :

- client\_id : surrogate key for identifying every client
- passport id : natural key for the client
- name
- gender
- date of birth
- nationality
- contact number

k. **agent** :

- agent\_id : surrogate key for identifying every agent
- national id : natural key for the agent
- name
- gender
- date of birth
- nationality
- contact number
- title : position of the agent in the customer care department

**4- Measures:**

- **Overall feedback** : rating of 5 from the client on the service of customer care
- **Flight rating** : rating of 5 from the client on the flight experience
- **Service rating** : rating of 5 from the client on the airline service
- **Overall rating**: overall rating of 5 from the client

## Data Model:

