Airline case study: multidimensional implementation

José Luis Puente Bodoque

January, 2023

Abstract

In this report we show the script content that implements the data warehouse for the airline case study on Oracle SQL Developer. Then we generate the star schema on Oracle Data Modeler.

1 The script

The `airlinedw.sql` script is a slight modification of the script provided as style guide. Next we show for each table the `CREATE TABLE` sentences and talk about the most remarkable aspects of their attributes.

1.1 Airplane

`airplane` is the dimension table for *airplane* dimension in our design.

```
CREATE TABLE AIRPLANE (
  id                NUMBER(10) NOT NULL,
  airplane_type     VARCHAR2(45) DEFAULT NULL NULL,
  departure_hour    VARCHAR2(5) DEFAULT NULL NULL,
  arrival_hour      VARCHAR2(5) DEFAULT NULL NULL,
  capacity          NUMBER(10) DEFAULT NULL NULL,
  first_capacity    NUMBER(10) DEFAULT NULL NULL,
  business_capacity NUMBER(10) DEFAULT NULL NULL,
  coach_capacity    NUMBER(10) DEFAULT NULL NULL,
  PRIMARY KEY (id))
;
```

`id` is the primary key and stores an integer number of 10 digits. It cannot be null. `departure_hour` and `arrival_hour` store both the departure and arrival hours, respectively, in 'HH:MM' format. The maximum length is 5 bytes and cannot be greater because we will never consider seconds. `capacity` attribute stores the number of available seats in the airplane. It also represents the top additive dimension of a hierarchy composed of `first_capacity`, `business_capacity` and `coach_capacity` dimensions. All the attributes except for the primary key may be null by default if we do not specify a value when loading the data.

1.2 Airport

`airport` is the dimension table for *airport* dimension in our design.

```
CREATE TABLE AIRPORT (
  id             NUMBER(10) NOT NULL,
  airport_type   VARCHAR2(45) DEFAULT NULL NULL,
  radar_type     VARCHAR2(45) DEFAULT NULL NULL,
  city           VARCHAR2(45) DEFAULT NULL NULL,
  state          VARCHAR2(45) DEFAULT NULL NULL,
  name           VARCHAR2(45) DEFAULT NULL NULL,
  PRIMARY KEY (id))
;
```

`id` is the primary key and stores an integer number of 10 digits. It cannot be null. `state` attribute stores the state where the airplane is located. It also represents the top dimension of a hierarchy composed of its lower `city` dimension. All the attributes except for the primary key may be null by default if we do not specify a value when loading the data.

1.3 Payment Channel

`payment_channel` is the dimension table for *payment_channel* dimension in our design.

```
CREATE TABLE PAYMENT_CHANNEL (
  id             NUMBER(10) NOT NULL,
  payment_type   VARCHAR2(45) DEFAULT NULL NULL,
  PRIMARY KEY (id))
;
```

`id` is the primary key and stores an integer number of 10 digits. It cannot be null. `payment_type` stores the payment channel (cash,

card or PayPal) and it may be null by default if we do not specify any value when loading the data.

1.4 Fare

`fare` is the dimension table for *fare* dimension in our design.

```
CREATE TABLE FARE (
  id              NUMBER(10) NOT NULL,
  seat_type       VARCHAR2(45) DEFAULT NULL NULL,
  description     CLOB DEFAULT NULL NULL,
  restriction     VARCHAR2(45) DEFAULT NULL NULL,
  PRIMARY KEY (id))
;
```

`id` is the primary key and stores an integer number of 10 digits. It cannot be null. All the attributes except the primary key may be null by default if we do not specify a value for them when loading the data.

1.5 Hour

`hour` is the dimension table for *hour* dimension in our design.

```
CREATE TABLE HOUR (
  id              NUMBER(10) NOT NULL,
  hour            VARCHAR2(10) DEFAULT NULL NULL,
  only_hour       VARCHAR2(10) DEFAULT NULL NULL,
  time_of_day     VARCHAR2(40) DEFAULT NULL NULL,
  PRIMARY KEY (id))
;
```

`id` is the primary key and stores an integer number of 10 digits. It cannot be null. `time_of_day` is the top dimension of a hierarchy composed of its lower dimension `hour`, which contains in turn the lower dimension `only_hour`. All the attributes except the primary key may be null by default if we do not specify a value for them when loading the data.

1.6 Passenger

`passenger` is the dimension table for *passenger* dimension in our design.

```
CREATE TABLE PASSENGER (
  id                NUMBER(10) NOT NULL,
  client_type       VARCHAR2(45) DEFAULT NULL NULL,
  city              VARCHAR2(45) DEFAULT NULL NULL,
  state             VARCHAR2(45) DEFAULT NULL NULL,
  name              VARCHAR2(45) DEFAULT NULL NULL,
  address           VARCHAR2(45) DEFAULT NULL NULL,
  zip               VARCHAR2(45) DEFAULT NULL NULL,
  income            VARCHAR2(45) DEFAULT NULL NULL,
  birth_date        VARCHAR2(45) DEFAULT NULL NULL,
  marital_status    VARCHAR2(45) DEFAULT NULL NULL,
  sex               VARCHAR2(45) DEFAULT NULL NULL,
  PRIMARY KEY (id))
;
```

`id` is the primary key and stores an integer number of 10 digits. It cannot be null. `state` attribute stores the passenger's state. It also represents the top dimension of a hierarchy composed of its lower `city` dimension. `birth_date` is a string in format "mmmm dddd, yyyy". All the attributes except the primary key may be null by default if we do not specify a value for them when loading the data.

1.7 Boarding Pass

`h_boarding_pass` is the fact table of the airline data warehouse.

```
CREATE TABLE H_BOARDING_PASS (
  id_boarding_pass              NUMBER(10) NOT NULL,
  id_trip                       NUMBER(10) NOT NULL,
  id_leg_origin                 NUMBER(10) NOT NULL,
  id_leg_destination            NUMBER(10) NOT NULL,
  id_trip_origin                NUMBER(10) NOT NULL,
  id_trip_destination           NUMBER(10) NOT NULL,
  id_departure_hour             NUMBER(10) NOT NULL,
  id_arrival_hour               NUMBER(10) NOT NULL,
  id_airplane                   NUMBER(10) NOT NULL,
  id_payment_channel            NUMBER(10) NOT NULL,
  id_fare                       NUMBER(10) NOT NULL,
  id_passenger                  NUMBER(10) NOT NULL,
  base_fare          BINARY_DOUBLE DEFAULT NULL NULL,
  miles              NUMBER(10) DEFAULT NULL NULL,
  delay              NUMBER(10) DEFAULT NULL NULL,
  ticket_number      NUMBER(10) DEFAULT NULL NULL,
  PRIMARY KEY (id_leg_origin, id_leg_destination,
     id_trip_origin, id_trip_destination,
     id_departure_hour, id_arrival_hour, id_airplane,
```

```
    id_payment_channel, id_fare, id_passenger)
,
  CONSTRAINT sk_H
    UNIQUE (id_boarding_pass)
   ,
  CONSTRAINT fk_H_AIRPORT1
    FOREIGN KEY (id_leg_origin)
    REFERENCES AIRPORT (id)
   ,
  CONSTRAINT fk_H_AIRPORT2
    FOREIGN KEY (id_leg_destination)
    REFERENCES AIRPORT (id)
   ,
  CONSTRAINT fk_H_AIRPORT3
    FOREIGN KEY (id_trip_origin)
    REFERENCES AIRPORT (id)
   ,
  CONSTRAINT fk_H_AIRPORT4
    FOREIGN KEY (id_trip_destination)
    REFERENCES AIRPORT (id)
   ,
  CONSTRAINT fk_H_HOUR1
    FOREIGN KEY (id_departure_hour)
    REFERENCES HOUR (id)
   ,
  CONSTRAINT fk_H_HOUR2
    FOREIGN KEY (id_arrival_hour)
    REFERENCES HOUR (id)
   ,
  CONSTRAINT fk_H_AIRPLANE
    FOREIGN KEY (id_airplane)
    REFERENCES AIRPLANE (id)
   ,
  CONSTRAINT fk_H_PAYMENT
    FOREIGN KEY (id_payment_channel)
    REFERENCES PAYMENT_CHANNEL (id)
   ,
  CONSTRAINT fk_H_FARE
    FOREIGN KEY (id_fare)
    REFERENCES FARE (id)
   ,
  CONSTRAINT fk_H_PASSENGER
    FOREIGN KEY (id_passenger)
    REFERENCES PASSENGER (id)
   )
;
```

The primary key of the fact table is the composite key of the foreign

keys that point to its origin dimension table. `id_boarding_pass` is the surrogate key, has uniqueness constraint but is not part of the primary key. The fact table also contains the measures. At the time of setting the integrity constraints and references related to the keys, we must distinguish between the key name in and out of the fact table scope.

2 Star schema

Once we have inserted the tables, we generate the star schema on Oracle SQL Data Modeler.

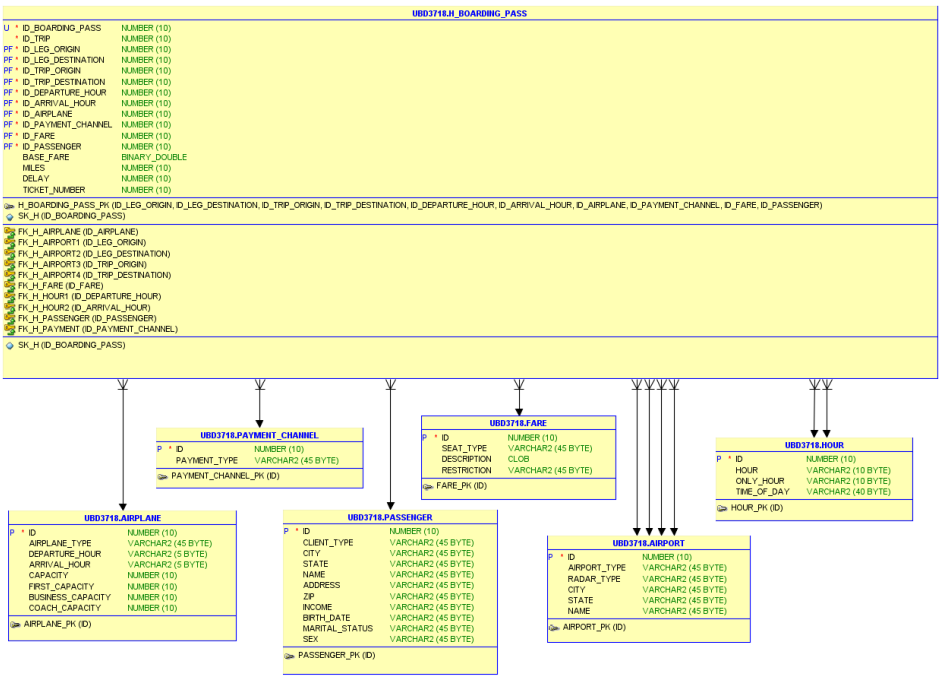

Figure 1: Star schema on Oracle Data Modeler

(*This report ends here.*)