



UTN.BA

UNIVERSIDAD TECNOLÓGICA NACIONAL
FACULTAD REGIONAL BUENOS AIRES

T.P. FRBA - VIAJES

Estrategia

GESTIÓN DE DATOS

1º Cuatrimestre - 2020

GRUPO 29 - LOS_BORBOTONES

Coen, Ezequiel

Conde, Nicolás

Lizarraga Sanchez, Ever Rolando

Stiberman, Alan

ÍNDICE

SCRIPT CREACIÓN INICIAL.....	3
ENTIDADES.....	3
MIGRACIÓN.....	5
STORED PROCEDURES.....	6
FUNCIONES.....	6
SCRIPT MODELO BI.....	8
ENTIDADES.....	8
MIGRACIÓN.....	9
STORED PROCEDURES.....	9
FUNCIONES.....	10

SCRIPT CREACIÓN INICIAL

ENTIDADES

CLIENTE

Cuenta con el campo *cliente_id* como clave primaria, y los campos *cliente_apellido*, *cliente_nombre*, *cliente_dni*, *cliente_fecha_nac*, *cliente_mail* y *cliente_telefono*. El campo *cliente_id* se determinó como una PK autonumérica, para identificar a cada cliente unívocamente, ya que existían DNI repetidos.

SUCURSAL

El campo *sucursal_id* es la clave primaria y es autonumérica. Posee también los campos *sucursal_direccion*, *sucursal_mail* y *sucursal_telefono*. El id de la sucursal comienza en 2, para que coincida el ID de cada sucursal que figura en su mail.

AEROLÍNEA

aerolínea_codigo es la clave primaria y dentro de esta tabla también se encuentra el campo *aerolínea_razon_social*. Se utiliza para identificar a aquellas empresas que trabajan con pasajes. Se utiliza una PK autonumérica.

AVIÓN

Esta tabla está compuesta por una PK *avión_id*, y el campo *avión_modelo*.

VUELO

Está compuesta por el campo *vuelo_codigo*, determinada como clave primaria, y los campos *vuelo_fecha_salida*, *vuelo_fecha_llegada*, *vuelo_avion_id* (que hace referencia al ID del respectivo avión), *vuelo_ruta_aerea_id*, (que hace referencia a la ruta aérea correspondiente) y *vuelo_aerolinea_codigo* (que representa a la empresa de aerolínea dueña del vuelo).

CIUDAD

Se compone de un campo autonumérico *ciudad_codigo*, determinado como clave primaria, para identificar a cada ciudad, y el campo *ciudad_detalle*.

RUTA_AÉREA

Está compuesta por los campos *ruta_aerea_id*, *ruta_aerea_codigo*, *ruta_aerea_ciu_origen* y *ruta_aerea_ciu_destino*. Se determinó al campo *ruta_aerea_id* como clave primaria, para identificar unívocamente cada ruta aérea ya que existían rutas aéreas con el mismo código de ruta aérea y distintas ciudades de origen y/o destino por lo que no se podía identificar unívocamente a la ruta aérea. Los últimos dos campos son los códigos de las ciudades pertenecientes a la ruta aérea.

TIPO_BUTACA

Se compone de un campo autonumérico *tipo_butaca_codigo* como clave primaria, y otro campo *tipo_butaca_detalle*. Esta tabla está directamente relacionada con la tabla butaca.

BUTACA

Tiene el campo *butaca_id* como clave primaria de tipo autonumérico, y además la tabla contiene los campos *butaca_numero*, *butaca_tipo_butaca_codigo* y *butaca_avion_id*. Esta tabla se relaciona con el tipo de butaca, el avión y el pasaje.

GRUPO_HOTELARIO

Posee un campo *grupo_hotelario_codigo* como clave primaria, y el campo *grupo_hotelario_razon_social*. Se emplea para identificar a aquellas empresas que trabajan con pasajes. Se utiliza una PK autonumérica.

HOTEL

La tabla posee la información de cada hotel perteneciente a los distintos grupos hotelarios. Se compone de los campos *hotel_codigo* (clave primaria autonumérica), *hotel_calle*, *hotel_nro_calle*, *hotel_cantidad_estrellas* y *hotel_grupo_hotelario_codigo*.

TIPO_HABITACIÓN

Está compuesta por dos campos: *tipo_habitacion_codigo* (clave primaria) de tipo decimal y *tipo_habitacion_detalle* de tipo nvarchar.

HABITACIÓN

La tabla posee cada habitación de un determinado hotel con su respectivo precio de compra y de venta. También se refleja en qué piso del hotel se encuentra, si está en el frente o no y por último, su tipo de habitación correspondiente. Está compuesta por los campos *habitacion_numero* y *habitacion_codigo_hotel* como clave primaria compuesta para identificar unívocamente una habitación, y también por los campos *habitacion_piso*, *habitacion_frente*, *habitacion_costo*, *habitacion_precio* y *habitacion_tipo_habitacion_codigo*.

FACTURA

Se compone de los campos *factura_numero* (clave primaria), *factura_fecha* y también se agregan las referencias al cliente que ha comprado y a la sucursal que realizó la venta mediante los campos *factura_cliente_id* y *factura_sucursal_id*.

COMPRA_EMPRESA_TURISMO

La tabla está compuesta por los campos *compra_empr_numero* (clave primaria) y *compra_empr_fecha*, que hacen referencia a los datos relacionados a las compras de pasajes y estadías que realiza la empresa de turismo a las aerolíneas y grupos hoteleros.

PASAJE

Posee la información de cada pasaje comprado a las aerolíneas, indicando también a qué número de compra y número de factura (en caso de que hayan sido vendidos) se vinculan. Se compone de los campos *pasaje_codigo* (clave primaria), *pasaje_costo*, *pasaje_precio*, *pasaje_vuelo_codigo*, *pasaje_butaca_id*, *pasaje_factura_numero*, *pasaje_compra_numero* y por último, *pasaje_valido*. Este último campo permite visualizar, mediante un bit, qué pasajes son válidos según la estrategia tomada en la migración.

ESTADIA

Posee la información de cada estadía comprada a los grupos hoteleros, indicando también a qué número de compra y número de factura (en caso de que hayan sido vendidos) se vinculan. Se compone de los campos *estadia_codigo* (clave primaria), *estadia_fecha_inicial*, *estadia_fecha_fin*, *estadia_cantidad_noches*, *estadia_hotel_codigo*, *estadia_habitacion_numero*, *estadia_costo*, *estadia_precio*, *estadia_factura_numero*, *estadia_compra_numero*.

MIGRACIÓN

La migración se realizó en base a la tabla Maestra provista por la cátedra. Para la misma, se fueron identificando los diferentes campos a migrar del sistema anterior y se lo integró a las tablas creadas para el nuevo sistema.

Algunos aspectos que se tuvieron en cuenta para realizarla fueron:

- En la tabla CLIENTE se agregó un campo *cliente_id*, ya que en los datos provistos se encontraron registros en los cuales había clientes con mismo DNI, por lo que se creó otro campo que los identifique.
- Se ha separado a las EMPRESAS según el producto que nos venden en las tablas AEROLÍNEAS y GRUPOS_HOTELARIOS.
- Se creó una tabla CIUDAD para persistir aquellas ciudades que pertenecen a alguna RUTA AÉREA.
- En la tabla RUTA AÉREA se agregó un campo *ruta_aerea_id*, ya que en los datos provistos se encontraron registros en los cuales existían códigos de rutas aéreas iguales, que tenían distintas ciudades de origen y destino, por lo que se tomó esta decisión para identificar cada una unívocamente.
- Se ha creado una tabla TIPO BUTACA para persistir la información de su tipo aparte de la propia butaca, y cada tipo con su respectivo código.
- En la tabla HABITACIÓN se decide utilizar los campos *habitacion_numero* y *habitacion_hotel_codigo* como clave primaria compuesta para identificar unívocamente a cada habitación de su correspondiente hotel.
- Se decidió referenciar tanto desde la tabla PASAJE como la tabla ESTADÍA a los respectivos registros relacionados a su compra en la tabla COMPRA y a su venta en la tabla FACTURA (en caso de que el producto haya sido vendido). De esta forma un registro comparte ambos datos, y no existen repeticiones.
- En la tabla PASAJE se decide agregar un campo denominado *pasaje_valido*, con tipo de dato bit. Este campo permite visualizar si un pasaje es válido, ya

que en la tabla Maestra existían distintos pasajes que estaban vinculados a una misma butaca de un vuelo, a realizarse en la misma fecha y hora.

STORED PROCEDURES

- Para llevar a cabo la migración de los datos de la tabla maestra a cada una de las tablas de nuestro modelo utilizamos un stored procedure por cada tabla. El nombre de cada stored procedure figura con nombre "*migracion_insert_[nombre de la tabla]*".
- Además creamos un stored procedure específico (*migracion_validar_pasajes*) que se encarga de marcar los pasajes que no sean válidos, cambiando el valor de su campo *pasaje_valido* a 0. Este proceso se ha dividido en tres updates que verifican la validez de los pasajes cargados. En primer lugar se actualizan aquellos pasajes repetidos que ambos hayan sido vendidos. En segundo lugar, aquellos pasajes repetidos en los cuales uno fue vendido y el otro no. Y por último, aquellos pasajes repetidos, de los cuales ninguno fue vendido.

FUNCIONES CREADAS

get_codigo_ciudad_by_detalle_ciudad

Obtiene el valor de la PK (ciudad_codigo) de una ciudad a través su nombre (ciudad_detalle).

get_hotel_codigo_by_calle_y_nro_calle

Obtiene el valor de la PK (hotel_codigo) de un hotel a través de su calle y su número de la calle.

get_grupo_hotelario_codigo_by_empresa_razon_social

Obtiene el valor de la PK (grupo_hotelario_codigo) de un grupo hotelero a través de su razón social (empresa_razon_social).

get_aerolinea_codigo_by_empresa_razon_social

Obtiene el valor de la PK (aerolinea_codigo) de una aerolínea a través de su razón social (empresa_razon_social).

get_tipo_butaca_codigo_by_tipo_butaca_detalle

Obtiene el valor de la PK (tipo_butaca_codigo) de un tipo de butaca a través de su detalle de tipo de butaca (tipo_butaca_detalle).

get_butaca_id_by_avion_numero_butaca_tipo_butaca

Obtiene el valor de la PK (butaca_id) de una butaca a través de su avión (butaca_avion_id), su número de butaca (butaca_numero) y su tipo de butaca (tipo_butaca_codigo).

get_ruta_aerea_id_by_ruta_aerea_codigo_ciudad_origen_ciudad_destino

Obtiene el valor de la PK (`ruta_aerea_id`) de una ruta aérea a través de su código (`ruta_aerea_codigo`), su código de ciudad de origen (`ruta_aerea_ciu_origen`) y su código de ciudad de destino (`ruta_aerea_ciu_destino`). Allí, con esos datos, busca los nombres de las respectivas ciudades utilizando los códigos de las mismas.

SCRIPT MODELO BI

ENTIDADES

DIMENSION_SUCURSAL

El campo *dim_sucursal_id* es la clave primaria y es autonumérica. Posee también los campos *dim_sucursal_direccion*, *dim_sucursal_mail* y *dim_sucursal_telefono*. El id de la sucursal comienza en 2, para que coincida el ID de cada sucursal que figura en su mail.

DIMENSION_TIPO_HABITACION

Está compuesta por dos campos: *dim_tipo_habitacion_codigo* (clave primaria) de tipo decimal y *dim_tipo_habitacion_detalle* de tipo nvarchar.

DIMENSION_TIPO_BUTACA

Se compone de un campo autonumérico *dim_tipo_butaca_codigo* como clave primaria, y otro campo *dim_tipo_butaca_detalle*. Esta tabla está directamente relacionada con la tabla butaca.

DIMENSION_AVION

Esta tabla está compuesta por una PK *dim_avión_id*, y el campo *dim_avión_modelo*.

DIMENSION_RUTA_AÉREA

Está compuesta por los campos *dim_ruta_aerea_id*, *dim_ruta_aerea_codigo*, *dim_ruta_aerea_ciu_origen* y *dim_ruta_aerea_ciu_destino*. Se determinó al campo *dim_ruta_aerea_id* como clave primaria, para identificar unívocamente cada ruta aérea ya que existían rutas aéreas con el mismo código de ruta aérea y distintas ciudades de origen y/o destino por lo que no se podía identificar unívocamente a la ruta aérea. Los últimos dos campos son los códigos de las ciudades pertenecientes a la ruta aérea.

DIMENSION_AEROLÍNEA

dim_aerolínea_codigo es la clave primaria y dentro de esta tabla también se encuentra el campo *dim_aerolínea_razon_social*. Se utiliza para identificar a aquellas empresas que trabajan con pasajes. Se utiliza una PK autonumérica.

DIMENSION_GRUPO_HOTELARIO

Posee un campo *dim_grupo_hotelario_codigo* como clave primaria, y el campo *dim_grupo_hotelario_razon_social*. Se emplea para identificar a aquellas empresas que trabajan con pasajes. Se utiliza una PK autonumérica.

HECHO_PASAJE

La tabla posee una primary key compuesta, la cual está formada por los campos: *hecho_pasaje_mes_anio*, *hecho_pasaje_aerolinea_codigo*, *hecho_pasaje_edad_cliente*, *hecho_pasaje_sucursal_id*, *hecho_pasaje_avion_id*, *hecho_pasaje_tipo_butaca_codigo* y *hecho_pasaje_ruta_aerea_id*. Además, posee los

campos *hecho_pasaje_costo_compra_promedio*,
hecho_pasaje_precio_venta_promedio, *hecho_pasaje_cantidad_pasajes_vendidos* y
hecho_pasaje_ganancia.

HECHO_ESTADIA

La tabla posee una primary key compuesta, la cual está formada por los campos: *hecho_estadia_mes_anio*, *hecho_estadia_grupo_hotelario_codigo*, *hecho_estadia_edad_cliente*, *hecho_estadia_sucursal_id* y *hecho_estadia_tipo_habitacion_codigo*. Además, posee los campos: *hecho_estadia_cant_estrellas_promedio*, *hecho_estadia_cantidad_camas*, *hecho_estadia_costo_compra_promedio*, *hecho_estadia_precio_venta_promedio*, *hecho_estadia_cantidad_habitaciones_vendidas* y *hecho_estadia_ganancias*.

MIGRACIÓN

Algunos aspectos que se tuvieron en cuenta para realizarla fueron:

- Decidimos crear dos tablas de hechos, HECHO_PASAJE y HECHO_ESTADIA. Nos pareció que esta es la mejor forma de dividir la información en el modelo OLAP ya que todos los datos considerados en el modelo transaccional giran en torno a las acciones (ya sea de compra o venta) que se realizan sobre pasajes y estadías. Además, tomamos esta decisión hay que nos dimos cuenta de que los hechos relacionados a pasajes y estadías no poseen las mismas dimensiones de interés, por lo tanto no podríamos considerarlos en una única tabla de hechos ya que de ser así figurarían muchas columnas de dimensiones en NULL (lo cual no debería pasar ya que las PKs de una tabla no pueden tener valor nulo).
- Para las dimensiones de tiempo y edad del cliente no nos pareció necesario crear una tabla aparte ya que de cada dimensión no encontramos ningún dato más de los que nos fue otorgado por la catedra en la tabla maestra que sea relevante para el análisis que se desea hacer sobre los datos.

STORED PROCEDURES

- Para llevar a cabo la migración de los datos desde el modelo OLTP a cada una de las tablas de dimensiones del modelo BI utilizamos un stored procedure por cada tabla. El nombre de cada stored procedure figura con nombre "*migracion_[nombre de la tabla]*".
- Para llevar a cabo la migración de los registros que componen las tablas de hechos (HECHOS_PASAJE y HECHO_ESTADIA) utilizamos un stored procedure para cada uno en el cual migramos los datos de las tablas PASAJE y ESTADIA del modelo OLTP (haciendo un insert-select). En dicho select ya devolvemos precalculados los valores de las columnas (de las tablas de hechos) que corresponden a información que obtenemos del análisis de los datos del modelo OLTP (como la cantidad de camas, la cantidad de estrellas promedio de estadías, precios de venta promedios, etc.)

FUNCIONES

`cantidad_cameras_vendidas`

Obtiene el valor de la cantidad de camas que posee una habitación según su tipo (`tipo_habitacion_codigo`).