

— MEMORIA —

Informes y Consultas sobre un Data Mart PRÁCTICA 3

CURSO 2023/2024

Simón Alonso (821038)
Óscar Brizuela (820773)
Jan Carlos Franco (806955)

Índice

1. Consultas	3
1.1. Consultas básicas	3
1.1.1. Consulta básica 1	3
1.1.2. Consulta básica 2	4
1.2. Consultas con operadores de agrupación GROUP BY	4
1.2.1. Consulta con operadores de agrupación 1	4
1.2.2. Consulta con operadores de agrupación 2	5
1.2.3. Consulta con operadores de agrupación 3	6
1.2.4. Consulta con operadores de agrupación 4	7
1.2.5. Consulta con operadores de agrupación 5	8
1.3. Cuestiones sobre la sentencia CREATE DIMENSION y similares	8
1.4. Consultas MDX	9
2. Informes	9

1. Consultas

1.1. Consultas básicas

Es necesario apuntar que, a lo largo de toda la memoria, se ha considerado el concepto de “vuelos con retraso” como aquellos cuyo retraso (bien de salida o bien de llegada) sea positivo. De esta forma, las medidas que involucran medias de retrasos no se ven afectadas por aquellos vuelos cuyo “retraso”, en realidad, es un “adelanto”, puesto que el valor de dicho campo es negativo.

1.1.1. Consulta básica 1

Se ha planteado una solución con tablas temporales mediante la cláusula *WITH*, para hacer la consulta modular.

La tabla temporal *RETRASO_SALIDA_LLEGADA* devuelve el retraso de salida medio y el retraso de llegada medio por aeropuerto origen.

La tabla temporal *RETRASO_TENIENDO_EN_CUENTA_DESTINO* obtiene el retraso de salida medio en función del aeropuerto origen y la ciudad de destino.

```
WITH RETRASO_SALIDA_LLEGADA AS (  
    SELECT IATA_aeropuerto_origen, AVG(retraso_salida) AS  
    retraso_salidaT, AVG(retraso_llegada) AS retraso_llegadaT  
    FROM Vuelo  
    WHERE retraso_salida > 0 AND retraso_llegada > 0  
    GROUP BY IATA_aeropuerto_origen  
)  
RETRASO_TENIENDO_EN_CUENTA_DESTINO AS (  
    SELECT IATA_aeropuerto_origen, nombre_ciudad_destino, AVG(retraso_salida) AS  
    retraso_total  
    FROM Vuelo  
    WHERE retraso_salida > 0  
    GROUP BY IATA_aeropuerto_origen, nombre_ciudad_destino  
)  
  
SELECT  
    RETRASO_TENIENDO_EN_CUENTA_DESTINO.IATA_aeropuerto_origen, RETRASO_TENIENDO_EN_CUENTA_DESTINO.nombre_ciudad_destino, RETRASO_TENIENDO_EN_CUENTA_DESTINO.retraso_total AS  
    retraso_medio_dest, (RETRASO_SALIDA_LLEGADA.retraso_salidaT +  
    RETRASO_SALIDA_LLEGADA.retraso_llegadaT)/2 AS retraso_aeropuerto  
FROM RETRASO_SALIDA_LLEGADA, RETRASO_TENIENDO_EN_CUENTA_DESTINO  
WHERE RETRASO_SALIDA_LLEGADA.IATA_aeropuerto_origen =  
    RETRASO_TENIENDO_EN_CUENTA_DESTINO.IATA_aeropuerto_origen
```

1.1.2. Consulta básica 2

Esta consulta muy similar a la anterior, con la diferencia de que se añade una tabla temporal **RETRASO_DESTINO** que devuelve el retraso de salida y de llegada medios en función de la ciudad origen.

```
WITH RETRASO_SALIDA_LLEGADA AS (  
    SELECT IATA_aeropuerto_origen,nombre_ciudad_origen, AVG(retraso_salida) AS  
    retraso_salidaT,AVG(retraso_llegada) AS retraso_llegadaT  
    FROM Vuelo  
    WHERE retraso_salida > 0 AND retraso_llegada > 0  
    GROUP BY IATA_aeropuerto_origen,nombre_ciudad_origen  
),  
RETRASO_TENIENDO_EN_CUENTA_DESTINO AS (  
    SELECT IATA_aeropuerto_origen, nombre_ciudad_destino, AVG(retraso_salida) AS  
    retraso_total  
    FROM Vuelo  
    WHERE retraso_salida > 0  
    GROUP BY IATA_aeropuerto_origen,nombre_ciudad_destino  
),  
RETRASO_DESTINO AS (  
    SELECT nombre_ciudad_origen, AVG(retraso_salida) AS  
    retraso_salidaT,AVG(retraso_llegada) AS retraso_llegadaT  
    FROM Vuelo  
    WHERE retraso_salida > 0  
    GROUP BY nombre_ciudad_origen  
)  
  
SELECT  
RETRASO_TENIENDO_EN_CUENTA_DESTINO.IATA_aeropuerto_origen,RETRASO_TENIENDO_EN_CUENTA_DESTINO.nombre_ciudad_destino,RETRASO_TENIENDO_EN_CUENTA_DESTINO.retraso_total AS  
retraso_medio_dest,(R_AERO_OR.retraso_salidaT + R_AERO_OR.retraso_llegadaT)/2 AS  
retraso_aeropuerto,(R_AERO_DEST.retraso_salidaT + R_AERO_DEST.retraso_llegadaT)/2 AS  
retraso_aeropuerto_dest  
FROM RETRASO_SALIDA_LLEGADA AS  
R_AERO_OR,RETRASO_TENIENDO_EN_CUENTA_DESTINO,RETRASO_DESTINO AS R_AERO_DEST  
WHERE R_AERO_OR.IATA_aeropuerto_origen =  
RETRASO_TENIENDO_EN_CUENTA_DESTINO.IATA_aeropuerto_origen AND  
R_AERO_DEST.nombre_ciudad_origen =  
RETRASO_TENIENDO_EN_CUENTA_DESTINO.nombre_ciudad_destino  
ORDER BY RETRASO_TENIENDO_EN_CUENTA_DESTINO.IATA_aeropuerto_origen
```

1.2. Consultas con operadores de agrupación *GROUP BY*

1.2.1. Consulta con operadores de agrupación 1

Mediante el operador de agregación *ROLLUP*, obtenemos:

- Los retrasos medios por cada pareja <matrícula, IATA_aeropuerto_origen>
- Los retrasos medios por cada <IATA_aeropuerto_origen>
- El retraso medio total de todos los vuelos

Además, mediante el uso de la función de agregación *GROUPING*, se crea una máscara <noAeropuerto, noMatricula>, la cual muestra los valores nulos procedentes de las tuplas devueltas por el operador *ROLLUP* anterior. De esta forma, ordenando de manera descendente en función de los valores de esta máscara, obtenemos primero el retraso medio total de todos los vuelos, seguido de los retrasos medios de cada aeropuerto y, por último, los retrasos medios de cada pareja <matrícula, IATA_aeropuerto_origen>.

Se ha considerado como retraso solo aquellos vuelos cuyo valor de *retraso_salida* es positivo, ya que si es negativo es un adelanto.

```
SELECT matricula, IATA_aeropuerto_origen, AVG(retraso_salida) AS
retrasos_salidas_medio, GROUPING(IATA_aeropuerto_origen) AS noAeropuerto,
GROUPING(matricula) AS noMatricula
FROM Vuelo
where retraso_salida > 0
GROUP BY ROLLUP(IATA_aeropuerto_origen, matricula)
ORDER BY noMatricula DESC, noAeropuerto DESC;
```

1.2.2. Consulta con operadores de agrupación 2

Esta consulta está formada a su vez por dos subconsultas.

La primera de ellas, denominada como *tabla1*, es muy similar a la anterior, pero utilizando el operador de agrupación *CUBE*. Mediante este operador se obtiene la siguiente información:

- Los retrasos medios por cada pareja <matrícula, IATA_aeropuerto_origen>
- Los retrasos medios por cada <matrícula>
- Los retrasos medios por cada <IATA_aeropuerto_origen>
- El retraso medio total de todos los vuelos

Es importante destacar que estos resultados proceden sólo de aquellos vuelos cuya fecha está considerada como periodo vacacional. En este caso, a pesar de no ordenar en función del valor de la máscara, también se hace uso de la función de agregación *GROUPING*.

La siguiente subconsulta obtiene la misma información que la anterior pero para los vuelos cuya fecha está en periodo no vacacional.

Una vez tenemos ambas subconsultas, se realiza un *SELECT* que las engloba para obtener el retraso medio de salida de los vuelos en periodo vacacional junto al número de vuelos por

cada matrícula y aeropuerto origen. Se obtiene la misma información para los vuelos en periodo no vacacional.

De este modo podemos comparar el retraso medio en cada uno de los periodos por cada matrícula y aeropuerto origen.

Todos los resultados están ordenados de forma descendente dada la suma de los retrasos en periodo vacacional y no vacacional, es decir, se muestran primero aquellas matrículas y aeropuertos origen en los que la suma de los retrasos de ambos periodos es mayor.

```
SELECT tabla1.matricula, tabla1.IATA_aeropuerto_origen, tabla1.vuelos AS
vuelos_vacacion,
    tabla1.retrasos_salidas_medio AS retraso_salida_vacacion,tabla2.vuelos AS
vuelos_no_vacacion,tabla2.retrasos_salidas_medio AS retraso_salida_no_vacacion
FROM
    (SELECT matricula, IATA_aeropuerto_origen, COUNT(*) AS vuelos,
    AVG(retraso_salida) AS retrasos_salidas_medio,
    GROUPING(IATA_aeropuerto_origen) AS noAeropuerto, GROUPING(matricula) AS
noMatricula
    FROM Vuelo v,Fecha f
    WHERE retraso_salida > 0 AND periodo_vacacional = 'Periodo vacacional' AND
v.id_fecha = f.id_fecha
    GROUP BY CUBE(matricula, IATA_aeropuerto_origen)
    ORDER BY matricula, IATA_aeropuerto_origen) tabla1,

    (SELECT matricula, IATA_aeropuerto_origen, COUNT(*) AS vuelos,
    AVG(retraso_salida) AS retrasos_salidas_medio,
    GROUPING(IATA_aeropuerto_origen) AS noAeropuerto, GROUPING(matricula) AS
noMatricula
    FROM Vuelo v,Fecha f
    WHERE retraso_salida > 0 AND periodo_vacacional = 'Periodo no vacacional' AND
v.id_fecha = f.id_fecha
    GROUP BY CUBE(matricula, IATA_aeropuerto_origen)
    ORDER BY matricula, IATA_aeropuerto_origen) tabla2
WHERE tabla1.matricula = tabla2.matricula AND tabla1.IATA_aeropuerto_origen =
tabla2.IATA_aeropuerto_origen
ORDER BY (tabla1.retrasos_salidas_medio + tabla2.retrasos_salidas_medio) DESC;
```

1.2.3. Consulta con operadores de agrupación 3

Esta consulta obtiene la media de los retrasos de salida en la franja horaria especificada del aeropuerto de origen.

Se ha utilizado el operador de agrupación *CUBE*, de modo que selecciona los siguientes conjuntos:

- *<IATA_aeropuerto_origen franja_horaria>*, que representa la agrupación por aeropuerto de origen y franja horaria.
- *<IATA_aeropuerto_origen>*, que representa la agrupación por aeropuerto de origen independientemente de la franja horaria.
- *<franja_horaria>*, que representa la agrupación solo por franja horaria.
- *<>*, que representa la agrupación total sin restricciones adicionales.
-

```

SELECT
    IATA_aeropuerto_origen,
    franja_horaria,
    AVG(retraso_salida) AS retraso_salida_franja
FROM
    (SELECT
        id_fecha,
        IATA_aeropuerto_origen,
        CASE
            WHEN EXTRACT(HOUR FROM hora_estimada_salida) >= 6 AND
EXTRACT(HOUR FROM hora_estimada_salida) < 12 THEN 'Mañana'
            WHEN EXTRACT(HOUR FROM hora_estimada_salida) >= 12 AND
EXTRACT(HOUR FROM hora_estimada_salida) < 18 THEN 'Tarde'
            ELSE 'Noche'
        END AS franja_horaria,
        retraso_salida
    FROM
        Vuelo
    WHERE
        retraso_salida > 0) AS subquery
GROUP BY
    CUBE(IATA_aeropuerto_origen, franja_horaria)
ORDER BY
    IATA_aeropuerto_origen, franja_horaria;

```

1.2.4. Consulta con operadores de agrupación 4

Una función ventana es una función agregada aplicada a una partición ó subconjunto del resultado de una consulta

Una ventana está formada por una partición (*PARTITION*) y un marco (*FRAME*). Se define aplicando la cláusula *OVER* a la función agregada. Esta cláusula *OVER* define la partición ó subconjunto que forma la ventana y un marco (*FRAME*) dentro de la partición, que se define con la cláusula *PARTITION BY*.

Por tanto, utilizando la función de ventana *OVER (PARTITION BY IATA_aerolinea)*, se obtiene la suma total de vuelos por cada aerolínea para cada día de la semana. Esta suma total de vuelos de cada aerolínea es utilizada posteriormente para ordenar la salida de la consulta de manera decreciente, de forma que aquellas aerolíneas con más vuelos totales saldrán primero. Esto sería equivalente a realizar un *COUNT(*)* usando el *GROUP BY IATA_aerolinea*.

A su vez, se utiliza la función de agregación *ROLLUP*, para agrupar por:

- <IATA_aerolinea,nombre_dia_semana>
- <IATA_aerolinea>
- <>

De esta forma, se obtiene el número de vuelos totales (evidentemente, de todas las aerolíneas), el número de vuelos por cada pareja <IATA_aerolinea,nombre_dia_semana> y el número de vuelos por cada aerolínea.

```
SELECT
    IATA_aerolinea,
    nombre_dia_semana,
    COUNT(*) AS cantidad_vuelos,
    SUM(COUNT(*)) OVER (PARTITION BY IATA_aerolinea) AS total_vuelos_aerolinea
FROM
    Vuelo, Fecha
WHERE Vuelo.id_fecha = Fecha.id_fecha
GROUP BY
    ROLLUP(IATA_aerolinea, nombre_dia_semana)
ORDER BY
    IATA_aerolinea, nombre_dia_semana;
```

1.2.5. Consulta con operadores de agrupación 5

En esta consulta se utiliza la función de agregación *GROUPING SETS*, la cual nos permite especificar las agrupaciones más detalladamente.

En este caso se está realizando un producto cartesiano <id_fecha matricula> X <IATA_aeropuerto_origen IATA_aeropuerto_destino>. En total saldrán las siguientes agrupaciones:

- <id_fecha IATA_aeropuerto_origen>
- <id_fecha IATA_aeropuerto_destino>
- <matrícula IATA_aeropuerto_origen>
- <matrícula IATA_aeropuerto_destino>

```
SELECT id_fecha, matricula, IATA_aeropuerto_origen,IATA_aeropuerto_destino
    COUNT(*) AS num_vuelos_producto_fecha,
    AVG(retraso_salida) AS retrasos_salidas_medio
FROM Vuelo
WHERE retraso_salida > 0
GROUP BY GROUPING SETS(id_fecha, matricula),
    GROUPING SETS(IATA_aeropuerto_origen, IATA_aeropuerto_destino)
ORDER BY id_fecha, matricula, IATA_aeropuerto_origen;
```


1.3. Cuestiones sobre la sentencia *CREATE DIMENSION* y similares

La sentencia *CREATE DIMENSION* no forma parte del lenguaje SQL estándar. Sin embargo, cuando nos encontramos en el contexto de Oracle *OLAP*, sí que nos encontramos con dimensiones. A pesar de que no son directamente equivalentes a las dimensiones en bases de datos relacionales estándar, estas dimensiones permiten analizar y organizar datos de forma multidimensional.

La sentencia *CREATE DIMENSION* se utiliza para crear una de estas dimensiones, que define una relación de padre-hijo entre pares de conjuntos de columnas, donde todas las columnas de un mismo conjunto deben proceder de una misma tabla. Sin embargo, las columnas de otro conjunto pueden proceder de una tabla diferente.

El optimizador utiliza estas relaciones junto a vistas materializadas para realizar reescrituras de consultas.

Un ejemplo de la sentencia *CREATE DIMENSION*, procedente del manual de Oracle, podría ser:

```
CREATE DIMENSION customers_dim
  LEVEL customer    IS (customers.cust_id)
  LEVEL city        IS (customers.cust_city)
  LEVEL state       IS (customers.cust_state_province)
  LEVEL country     IS (countries.country_id)
  LEVEL subregion   IS (countries.country_subregion)
  LEVEL region      IS (countries.country_region)
  HIERARCHY geog_rollup (
    customer        CHILD OF
    city            CHILD OF
    state           CHILD OF
    country         CHILD OF
    subregion       CHILD OF
    region
  )
  JOIN KEY (customers.country_id) REFERENCES country
)
ATTRIBUTE customer DETERMINES
(cust_first_name, cust_last_name, cust_gender,
 cust_marital_status, cust_year_of_birth,
 cust_income_level, cust_credit_limit)
ATTRIBUTE country DETERMINES (countries.country_name)
;
```

Como se observa en la cláusula *HIERARCHY*, se establece una jerarquía entre los distintos niveles de la dimensión que va a ser creada. Por tanto, en base a los tipos de consultas que son solicitadas en esta práctica (utilizando los diversos operadores de agrupación), la sentencia *CREATE DIMENSION* no es necesaria, ni siquiera relevante. Quizá podría aportar algo de información para definir dimensiones específicas en el modelo en cuestión.

Sin embargo, al ser un esquema en estrella la implementación en una base de datos relacional un modelo de datos dimensional, podemos utilizar el parámetro de inicialización *STAR_TRANSFORMATION_ENABLED* con el objetivo de identificar las tablas de hechos y las restricciones en las tablas de dimensión de forma automática. Al asignar el valor de este parámetro a *TRUE*, la transformación se lleva a cabo solo si el coste del plan de transformación es menor que el coste del plan de no realizar la transformación. Por tanto, la transformación a modelo en estrella tendrá lugar cuando tengamos condiciones *WHERE* en dos o más tablas de dimensión. Esta transformación, a menudo, mejorará el rendimiento de las consultas que cuenten con restricciones muy selectivas, puesto que evita escanear la tabla de hechos completa.

2. Informes

Para la realización de los informes se ha utilizado la herramienta *Tableau Desktop*. Esta herramienta nos ha permitido visualizar el resultado de varias consultas que se han considerado importantes con el objetivo de conocer determinadas características de los datos.

Todas las consultas *SQL* correspondientes a estos informes, que se han ejecutado sobre la base de datos previamente cargada, se encuentran en el fichero *consultas_graficas.sql*, en el cual además se muestra una breve descripción del resultado de cada consulta.

2.1. Informe con información relevante sobre ciudades

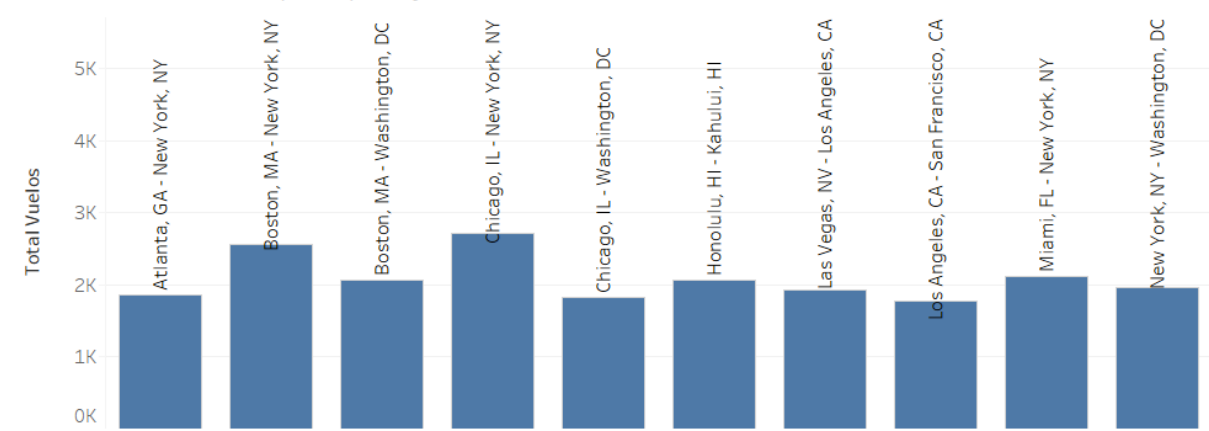
El primero de los gráficos en el informe muestra cuáles son las 10 parejas (A, B) de ciudades estadounidenses que tienen más vuelos entre sí. Por tanto, para cada columna del gráfico se considera la suma de los vuelos desde la ciudad A a la ciudad B y viceversa. Se ha decidido utilizar un gráfico de barras debido a que las diferencias entre las distintas parejas de ciudades son fácilmente visualizables en cuanto a las cantidades totales de vuelos.

El segundo de los gráficos en el informe muestra las 10 ciudades destino preferidas por los pasajeros para volar en lo que es considerado “periodo vacacional” en la base de datos (desde el 1 de enero de 2023 hasta el 15 de enero de 2023). Este gráfico hace uso de una de las funciones más interesantes que nos proporciona la herramienta: la función geográfica. Al obtener como resultado de la consulta la columna correspondiente al nombre de la ciudad (además del número de vuelos que tienen dicha ciudad como destino), *Tableau* ofrece automáticamente las latitudes y longitudes correspondientes a cada una de las ciudades cuando estas son importadas en una tabla. Sin embargo, debido a que los nombres de las ciudades devueltos por la consulta cuentan también con el nombre del estado (por ejemplo, “*Dallas, TX*”) debido a que pueden existir varias ciudades con el mismo nombre pero en distintos estados, hubo que eliminar el estado del nombre de la ciudad

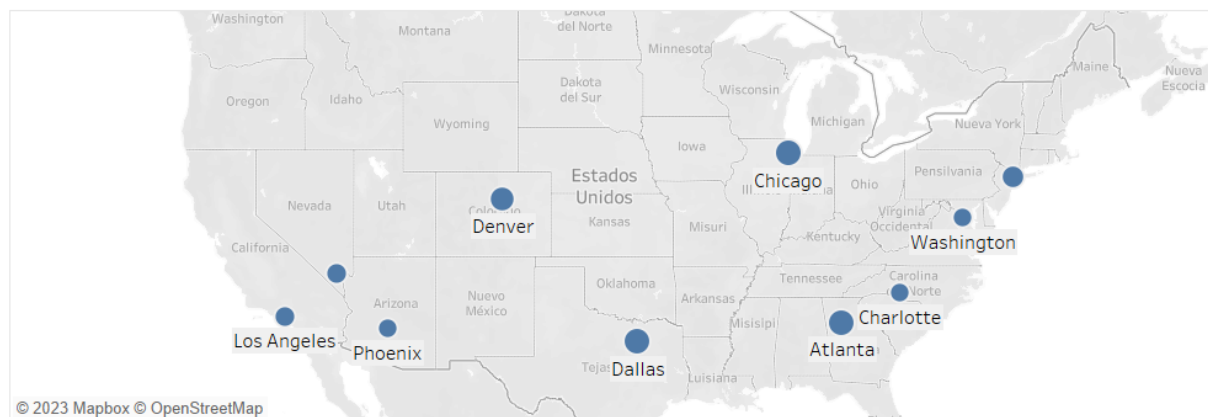
antes de importarlas como *Tabla*. Desafortunadamente, esta misma razón llevó a que *Tableau* detectara algunas de estas ciudades como ambiguas (puesto que, como hemos comentado antes, existen varias ciudades con el mismo nombre pero en distintos estados entre aquellas resultantes de la consulta). Para solventar este problema, se decidió introducir manualmente las coordenadas de cada una de estas “ciudades ambiguas” mediante la opción “Editar ubicaciones” de la pestaña “Mapa”, debido a que, al fin y al cabo, solo eran 6 ciudades.

Utilizar este gráfico permite visualizar muy fácilmente la localización de aquellas ciudades preferidas por los ciudadanos americanos en vacaciones, así como hacerse una idea del volumen de vuelos recibidos por cada una mostrado por el tamaño de los puntos azules. No obstante, al pasar el ratón por encima de cada punto se puede observar en la descripción emergente la cantidad exacta de vuelos recibidos en dicho periodo vacacional.

Top 10 parejas de ciudades con más vuelos entre sí



Top 10 ciudades a las que más se viaja en periodo vacacional



Informe 1

2.2. Informe con información relevante sobre estados

De la misma manera que en el segundo gráfico del informe anterior, se ha hecho uso de la función geográfica de *tableau* para identificar los distintos estados de forma automática. No obstante, para los dos gráficos de este informe se ha tenido que indicar que la columna “Estado”, importada como una tabla, hacía referencia a estados de EEUU, puesto que de no indicarlo no encontraba unas coordenadas (por defecto, la configuración de ubicación

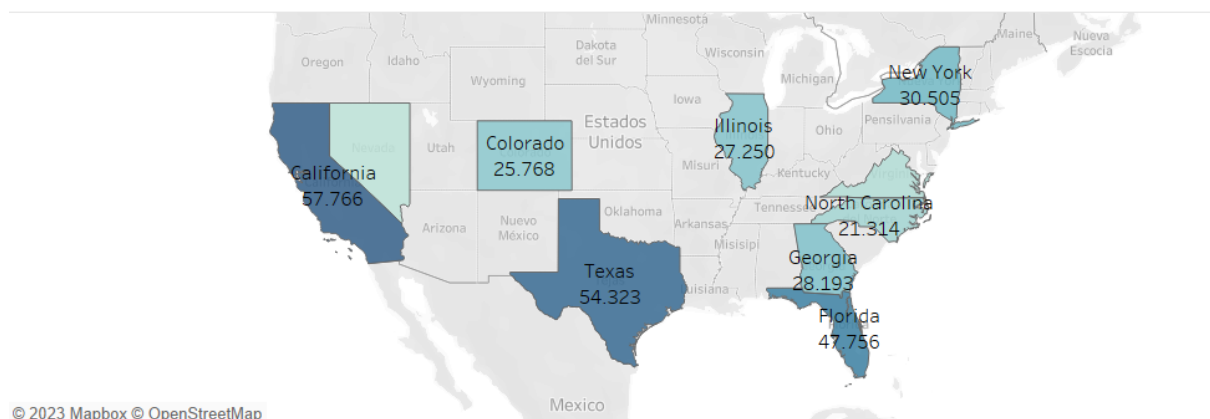
estaba en “España” y, como España no tiene estados, sino comunidades autónomas, Tableau no proporcionaba las coordenadas). Este paso se realizó también mediante la opción “Editar ubicaciones” de la pestaña “Mapa”.

El primero de los gráficos muestra los 10 estados preferidos por los pasajeros como destino, independientemente del periodo. Las etiquetas muestran tanto el nombre del estado como el número de vuelos recibidos. Además, la intensidad de la gama de colores utilizada muestra, de un vistazo rápido, cuáles son los estados que más vuelos reciben (y, por tanto, también los que menos) dentro de los 10 seleccionados.

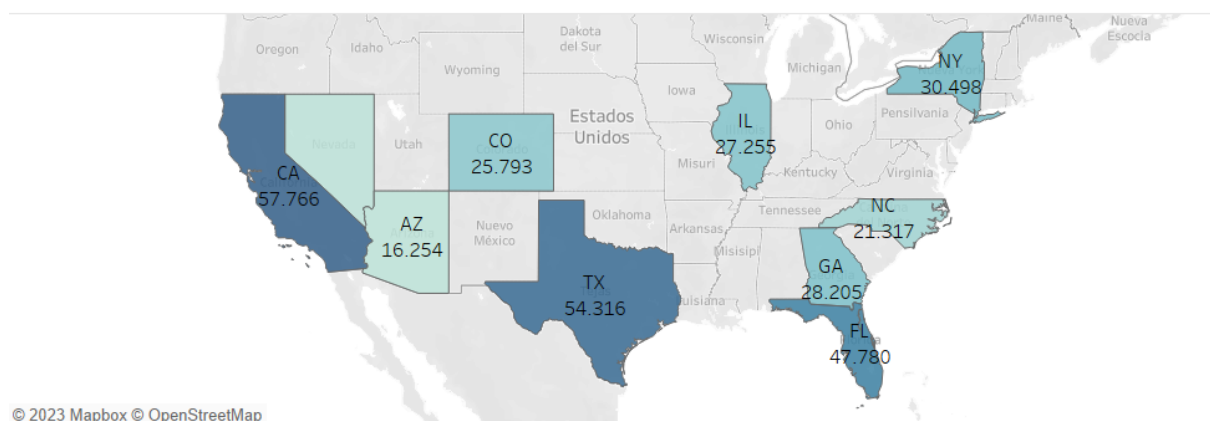
El segundo de los gráficos del informe muestra los 10 estados que más vuelos internos tienen, es decir, aquellos cuyo estado de origen y estado de destino del vuelo son el mismo. De la misma forma que en el anterior gráfico, cada etiqueta de estado muestra el nombre y la cantidad de vuelos recibidos, así como hacer uso de la intensidad de azul para mostrar rápidamente los estados con más vuelos internos (y los que menos) dentro del top 10.

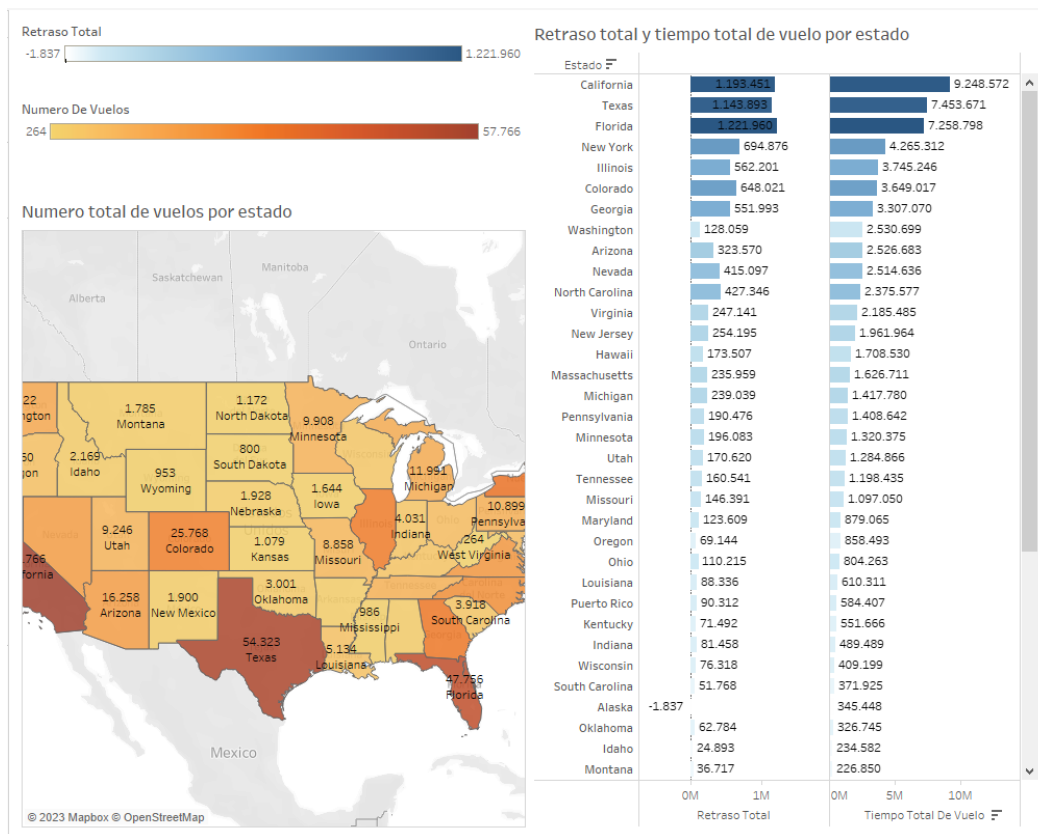
En el tercer gráfico del informe se muestra el número de vuelos de cada estado. Además, las gráficas laterales se han utilizado para que el usuario pueda comparar el retraso total acumulado de cada estado con el tiempo total de vuelo que se ha realizado en dicho estado.

Top 10 estados con más vuelos como destino



Top 10 estados con más vuelos internos





Informe 2

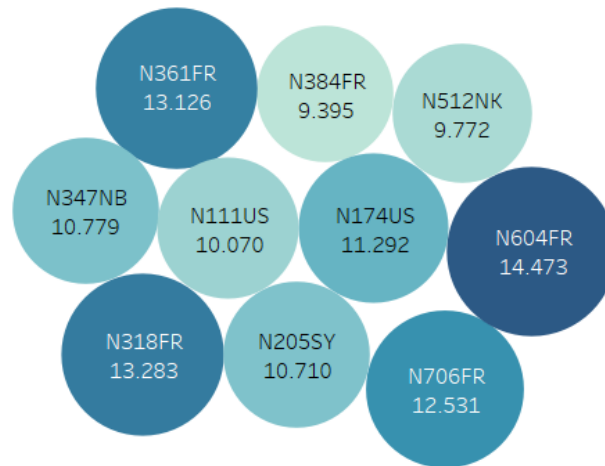
2.3. Informe con información relevante sobre aviones

Este informe se ha centrado en las matrículas de ciertos aviones, con el fin de visualizar información relevante.

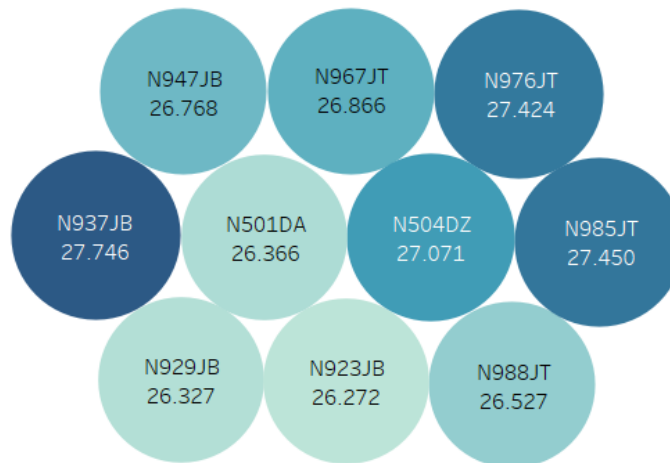
El primero de los gráficos muestra los 10 aviones que cuentan con más minutos de retraso totales. Las etiquetas muestran dicha cantidad de minutos, y tanto el diámetro de los círculos como la intensidad del color permiten hacerse una idea rápida de los resultados. Vemos, por ejemplo, que estos aviones cuentan en su matrícula con el sufijo *FR*.

El segundo gráfico de este informe permite visualizar los aviones cuyo tiempo de vuelo total es mayor. De la misma forma que antes, tanto el color como el diámetro de los círculos permite hacerse una idea rápida de los datos, además de mostrar los minutos totales de vuelo. Podemos comprobar que existen varios aviones en cuya matrícula se encuentra el sufijo *JB*, además de comprobar que todos se encuentran entre los 26000 y los 28000 minutos de vuelo.

Top 10 aviones con más minutos totales de retraso



Top 10 aviones con mayor tiempo de vuelo



Informe 3

2.4. Informe con información relevante sobre aeropuertos

Debido a que tableau no solo proporciona información geográfica sobre ciudades, estados y países, sino también sobre aeropuertos, se ha vuelto a utilizar la función geográfica, esta vez en función del *ATA* del aeropuerto.

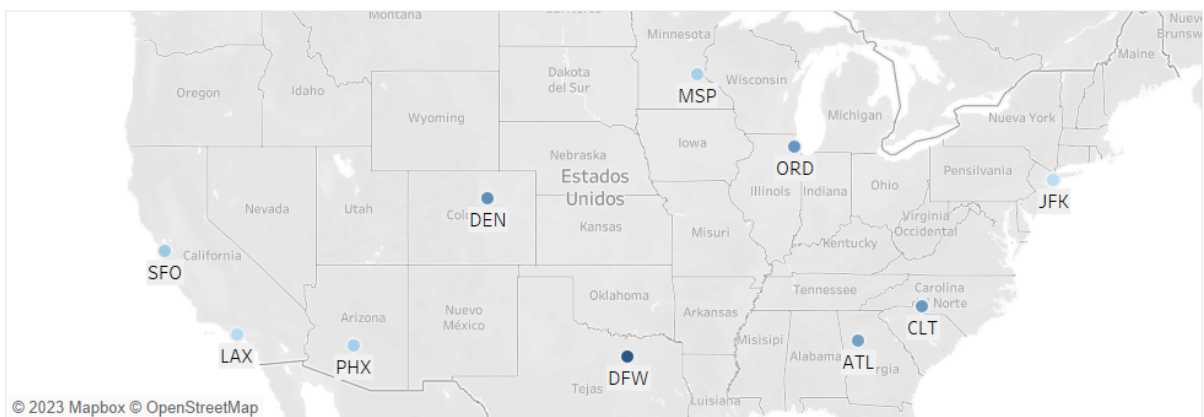
El primer gráfico del informe muestra los 10 aeropuertos con mayor cantidad de vuelos de salida. El mapa muestra los códigos *ATA* y la intensidad del color permite tener una percepción del volumen de vuelos de salida. No obstante, en la descripción emergente al deslizar el cursor por encima se puede observar el número de vuelos de salida exacto correspondiente.

El segundo de los gráficos muestra los 10 aeropuertos que tienen un retraso de salida mayor en la franja horaria matutina, considerándose esta entre las 6:00 am y las 11:59 am. Los puntos muestran la información de la misma forma que otro gráfico de este informe.

Top 10 aeropuertos con mayor cantidad de vuelos



Top 10 aeropuertos con más retraso por la mañana



Informe 4

2.5. Informe con información relevante sobre días de la semana

El primer gráfico del informe muestra el total de vuelos de cada día de la semana. A diferencia de los gráficos anteriores, este gráfico de burbujas no tiene en cuenta la intensidad del color para hacer referencia al volumen de vuelos, sino tan solo el diámetro del círculo.

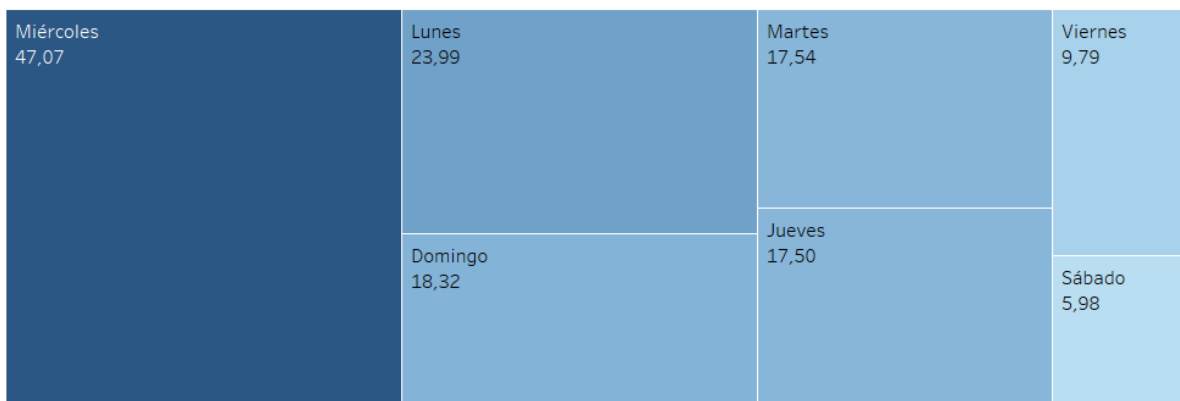
No obstante, cabe destacar que, al tratarse de datos absolutos y no relativos (o promedios), este gráfico puede llevar a malas interpretaciones. Debido a que enero de 2023 empezó un domingo y terminó un martes, hay un miércoles, un jueves, un viernes y un sábado menos con respecto a domingo, lunes y martes. Por tanto, este gráfico no pretende indicar que los lunes, por ejemplo, se tiende a viajar más que los sábados: simplemente muestra los datos de vuelos totales.

El segundo gráfico muestra, en un gráfico en forma de mapa en árbol, los retrasos medios para cada día de la semana. La dimensión del rectángulo y la intensidad del color hacen las veces del diámetro del círculo y la intensidad del color en los gráficos anteriores. En este caso, por tanto, sí podríamos concluir que los miércoles existe una mayor cantidad de retrasos respecto a los viernes y los sábados.

Total de vuelos por cada día de la semana



Retraso promedio cada día de la semana



Informe 5

2.6. Informe con información relevante sobre aerolíneas

En este informe ambos gráficos son gráficos de burbujas.

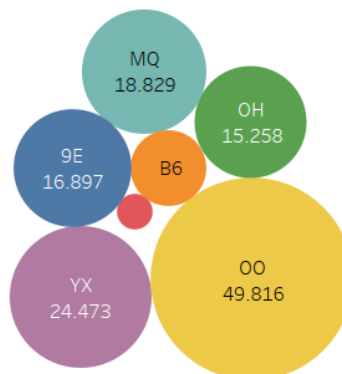
El primero muestra el promedio de tiempo de vuelo de cada aerolínea. Cada color identifica a una aerolínea, y el diámetro de los círculos permiten hacerse una idea de la cantidad de minutos promedios para cada empresa.

El segundo muestra el número de vuelos que cada aerolínea ha operado sin utilizar un avión de la marca *Airbus* o *Boeing*. Evidentemente, aquellas que solo operan vuelos con aviones de dichas marcas no aparecen en el gráfico. Además de aparecer el número de vuelos totales de cada aerolínea que cumplen estas características, el diámetro también proporciona una idea de las empresas que menos utilizan *Airbus* o *Boeing* en sus vuelos.

Duración promedio del tiempo de vuelo de cada aerolínea



Número de vuelos operados por cada aerolínea que no han sido volados por un Airbus o un Boeing



3. Conclusiones

Hemos sido capaces de realizar consultas complejas, tanto utilizando los operadores de agrupación comunes como aquellos que no son de primer nivel.

También hemos extraído información relevante de la base de datos previamente poblada, con el objetivo de visualizarla de la forma más sencilla y clara posible.

Tarea	Jan	Simón	Óscar
Punto 1: Consultas básicas	3		
Punto 2: Consultas con operadores <i>GROUP BY</i>	5		
Punto 3: Informes	5		
Realización de la memoria	3		
Total: 48	16		