***Posgrado de actualización de Business Intelligence e Introducción al Data Mining***

Trabajo práctico cuatrimestral

*Integrantes:*

* *Ayala, Ariel*
* *Jacobacci, Florencia*
* *Rahman, Thomas*

**Modelo de negocio:**

En vista del auge de los vuelos de bajo costo que se está observando en la actualidad, un grupo empresario está evaluando la posibilidad de ingresar al mercado aeronáutico internacional con una nueva aerolínea low-cost de bandera argentina. Para esto, han acudido a nuestra consultora para obtener, mediante un análisis de mercado, una solución de inteligencia de negocio y una evaluación de la factibilidad y viabilidad del proyecto.

El mencionado análisis de mercado aeronáutico se plantea con el objetivo de identificar las rutas más explotadas (utilizadas con mayor frecuencia por las distintas aerolíneas) y, a su vez, la situación actual en la región latinoamericana en cuanto a aeropuertos disponibles y cantidad de vuelos que reciben o inician desde dicha ubicación.

Para realizar el análisis del mercado aeronáutico de forma efectiva, se ha comprado una base de datos que posee información acerca de las rutas internacionales vigentes, aerolíneas habilitadas para operar actualmente y los vuelos que operan actualmente en esas rutas.

El datamart generado a partir de esta base de datos es el de “Vuelos”. Cada vuelo corresponde a una aerolínea que opera en una ruta que conecta un aeropuerto de origen con un aeropuerto de destino. Esa ruta puede ser recorrida numerosas veces en distintos vuelos, con fechas y horarios de salida y de llegada estipulados. Por motivos climáticos, gremiales o de ocupación de las pistas en los aeropuertos, los vuelos pueden adelantarse o retrasarse, por lo que también se registra la hora de salida y de llegada real de cada vuelo.

A continuación las dimensiones y atributos (detalladas entre paréntesis) definidos en el Modelo Dimensional:

* **Avión** (*Marca, modelo, código*)

*Ejemplos:*

* Marca: Airbus. Modelo: a300. Código: AB3.
* Marca: Airbus. Modelo: a300-600. Código: ABY.
* Marca: Boeing. Modelo:747. Código:747.
* **Geografía\_origen** (*País, Ciudad, Aeropuerto*)

Cabe destacar que dentro del atributo “ciudad” se tiene la siguiente información: Nombre de ciudad, Timezone (Zona horaria). A su vez, dentro del atributo “aeropuerto” se tiene la siguiente información: ID, Nombre del aeropuerto, código IATA, código ICAO, Latitud, Longitud, Altitud (en pies) y tipo de aeropuerto (airport/aeropuerto, station/estacion, port/puerto, unknown/desconocido).

*Ejemplos:*

* País: Argentina. Ciudad: Buenos Aires. Aeropuerto: Ministro Pistarini International Airport. ID: 398., Código IATA: EZE. Código ICAO: SAEZ. Latitud: -34.8222. Longitud: -58.5358. Altitud: 67. Timezone: -3. Tipo de aeropuerto: airport/aeropuerto.
* País: USA. Ciudad: Elizabeth City. Aeropuerto: Elizabeth City Regional Airport & Coast Guard Air Station. ID: 3562. Código IATA: ECG. Código ICAO: KECG. Latitud: 36.26060104. Longitud: -76.17459869. Altitud: 12. Timezone: -5. Tipo de aeropuerto: station/estación.
* **Geografía\_destino** (*País, Ciudad, Aeropuerto*)
* **Fecha\_salida** (A*ño, Mes, Fecha*)
* **Fecha\_llegada** (A*ño, Mes, Fecha*)
* ***Hora\_salida\_estipulada*** *(Hora, Minuto)*
* **Hora\_llegada\_estipulada** *(Hora, Minuto)*
* **Hora\_salida\_real** *(Hora, Minuto)*
* **Hora\_llegada\_real** *(Hora, Minuto)*
* **Ruta** (*Ruta*)

El atributo ruta se encuentra compuesto por la siguiente información: ID, Día del servicio.

*Ejemplo:*

* ID: AR240. Día del servicio: lunes.
* **Aerolínea** (*Bandera, Aerolínea*)

Dentro del atributo aerolínea se encuentra la siguiente información: ID, Nombre de la

aerolínea, alias, IATA code, ICAO code y si está actualmente en actividad (“Y” para activa y “N” para inactiva).

*Ejemplo:*

* Bandera: Argentina. Nombre\_aerolínea: Aerolíneas Argentinas. ID: 412. IATA code: AR. ICAO code: ARG. En actividad: Y

A continuación se deja detalle de el set de datos utilizado para este trabajo practico:



Identificadas y detalladas las distintas dimensiones y atributos, se puede observar que el detalle de los vuelos sale de la combinación de las dimensiones y atributos mencionados.

Detalle de los hechos que se identificaron:

● Número máximo de pasajeros transportados por vuelo.

● Duración real del vuelo.

● Minutos de demora en la salida.

● Minutos de demora en la llegada.

Como resultado de la identificación de las dimensiones y atributos y, a su vez, de los hechos que se quieren obtener, se procede a realizar el gráfico de modelo dimensional. A continuación grafica del mismo:



A continuación damos un detalle de posibles métricas que se pueden obtener utilizando el datamart especificado anteriormente:

* Cantidad de aerolíneas por país, para poder identificar la competencia por los pasajeros en cada país.
* Cantidad de rutas que opera cada aerolínea.
* Cantidad de rutas cubiertas por aerolíneas de origen argentino.
* Cantidad de rutas con origen en Argentina (se puede ordenar por bandera de aerolínea para evaluar competencia).
* Cantidad de rutas con destino en Argentina (se puede ordenar por bandera de aerolínea para evaluar competencia).
* Cantidad de pasajeros que salen de Argentina (evaluar el tamaño del mercado nacional).
* Cantidad de pasajeros que llegan a Argentina (se puede separar por región para conocer el tamaño del mercado regional, ej: Latinoamérica).
* Puntualidad por aerolínea.
* Marcas y modelos predominantes en las flotas nacionales (para evaluar el estado de los potenciales competidores).

**DER:**



(Los rombos en los extremos de las flechas indica la multiplicidad de las relaciones)