**Universidad de El Salvador**

****

**Facultad de Ingeniería y Arquitectura**

***Escuela de Ingeniería de sistemas informáticos***

*Asignatura: Curso de Especialización de Ingeniería de Datos*

*Catedrático: Ing. Fabricio Quintanilla*

***Tema: Análisis del proceso de vuelos para la empresa "Para Volar" que utiliza el software “Gamatron”***

*Kevin Eliseo Arevalo Beltran*

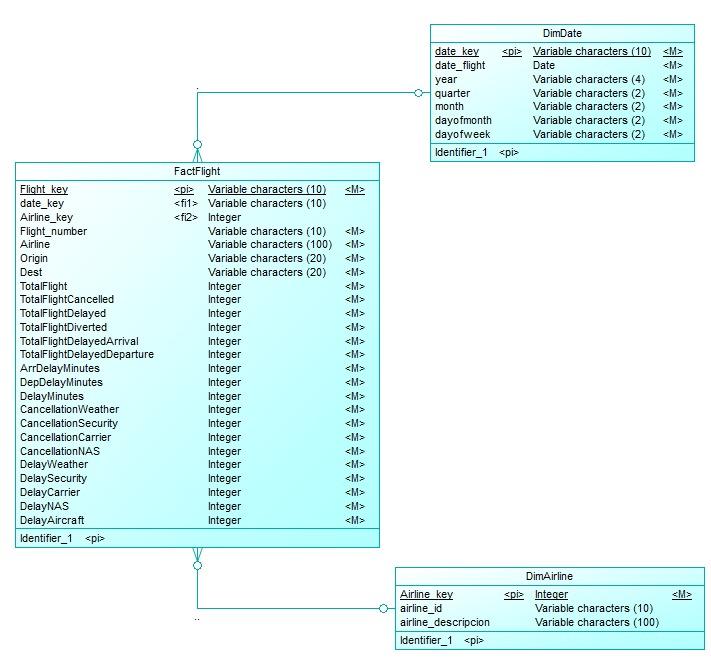
**Dimensiones**

* **DimAirline**
* **DimDate**
* **FactFlight**

**Métricas**

* Quiero conocer el número de vuelos totales
* Quiero conocer el número de vuelos retrasados, cancelados y diferidos
* Quiero conocer el promedio de vuelos retrasados
* Quiero conocer el promedio de minutos retrasados
* Quiero conocer la frecuencia de los motivos de cancelación
* Quiero conocer la frecuencia de los motivos de retraso
* Quiero conocer el número de vuelos retrasados en llegada
* Quiero conocer el número de vuelos retrasados en salida
* Quiero saber la proyección de vuelos cancelados, retrasados y diferidos

**Diseño de la Base de Datos**

****

**Script de la base**

***DimAirline***

CREATE TABLE "public"."DimAirline"(

"Airline\_key" INTEGER NOT NULL IDENTITY (1,1),

"AirlLine\_id" VARCHAR (10) NOT NULL,

"Airline\_descripcion" VARCHAR (100) NOT NULL,

UNIQUE ("Airline\_key"),

PRIMARY KEY("Airline\_key"))

DISTSTYLE KEY DISTKEY("AirlLine\_id")

SORTKEY ("AirlLine\_id")

ENCODE AUTO;

***DimDate***

CREATE TABLE "public"."DimDate"(

"Date\_key" INTEGER NOT NULL,

"Date\_flight" DATE NOT NULL,

"Year" VARCHAR (4) NOT NULL,

"Qarter" VARCHAR (1) NOT NULL,

"Month" VARCHAR (2) NOT NULL,

"DayOfMonth" VARCHAR (2) NOT NULL,

"DayOfWeek" VARCHAR (2) NOT NULL,

UNIQUE ("Date\_key"),

PRIMARY KEY("Date\_key"))

DISTSTYLE KEY DISTKEY("Date\_key")

SORTKEY ("Date\_key")

ENCODE AUTO;

***FactFlight***

CREATE TABLE "public"."FactFlight"(

"Flight\_key" INTEGER NOT NULL,

"Flight\_number" Varchar (10) NOT NULL,

"Airline\_key" INTEGER NOT NULL,

"Date\_key" INTEGER NOT NULL,

"Airline" Varchar (100) NOT NULL,

"Origin" Varchar(20) NOT NULL,

"Dest" Varchar(20) NOT NULL,

"TotalFlight" INTEGER DEFAULT 1 NOT NULL,

"TotalFlightCancelled" INTEGER DEFAULT 0 NOT NULL,

"TotalFlightDelayed" INTEGER DEFAULT 0 NOT NULL,

"TotalFlightDiverted" INTEGER DEFAULT 0 NOT NULL,

"TotalFlightDelayedArrival" INTEGER DEFAULT 0 NOT NULL,

"TotalFlightDelayedDeparture" INTEGER DEFAULT 0 NOT NULL,

"ArrDelayMinutes" INTEGER DEFAULT 0 NOT NULL,

"DepDelayMinutes" INTEGER DEFAULT 0 NOT NULL,

"DelayMinutes" INTEGER DEFAULT 0 NOT NULL,

"CancellationWeather" INTEGER DEFAULT 0 NOT NULL,

"CancellationSecurity" INTEGER DEFAULT 0 NOT NULL,

"CancellationCarrier" INTEGER DEFAULT 0 NOT NULL,

"CancellationNAS" INTEGER DEFAULT 0 NOT NULL,

"DelayWeather" INTEGER DEFAULT 0 NOT NULL,

"DelaySecurity" INTEGER DEFAULT 0 NOT NULL,

"DelayCarrier" INTEGER DEFAULT 0 NOT NULL,

"DelayNAS" INTEGER DEFAULT 0 NOT NULL,

"DelayAircraft" INTEGER DEFAULT 0 NOT NULL,

UNIQUE ("Flight\_key", "Date\_key", "Airline\_key"),

PRIMARY KEY("Flight\_key", "Date\_key", "Airline\_key"),

Foreign key("Airline\_key") references "DimAirline"("Airline\_key"),

Foreign key("Date\_key") references "DimDate"("Date\_key"))

DISTSTYLE KEY DISTKEY("Flight\_key")

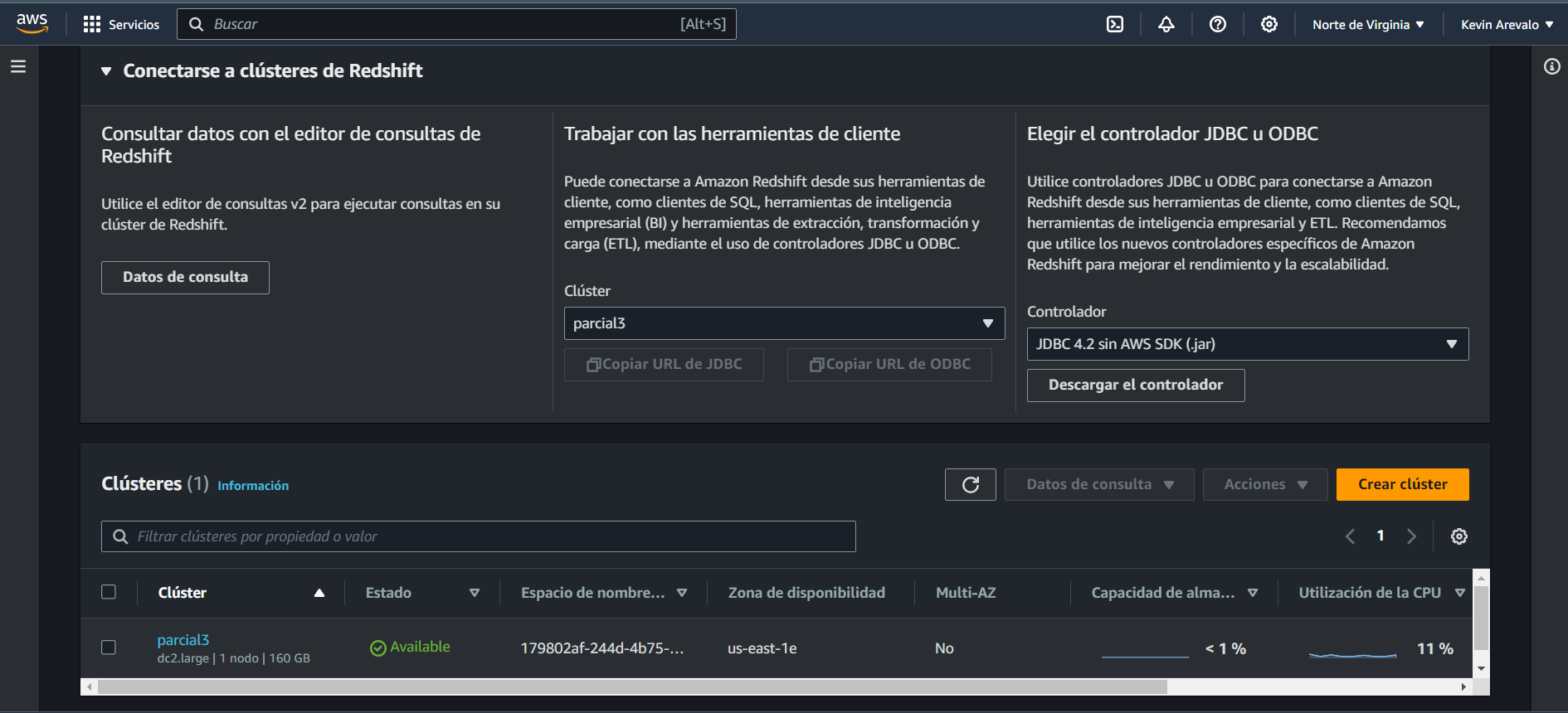
SORTKEY ("Flight\_key", "Date\_key", "Airline\_key")

ENCODE AUTO;

***Resumen de herramientas utilizadas***

**RedShift Cluster: parcial3**

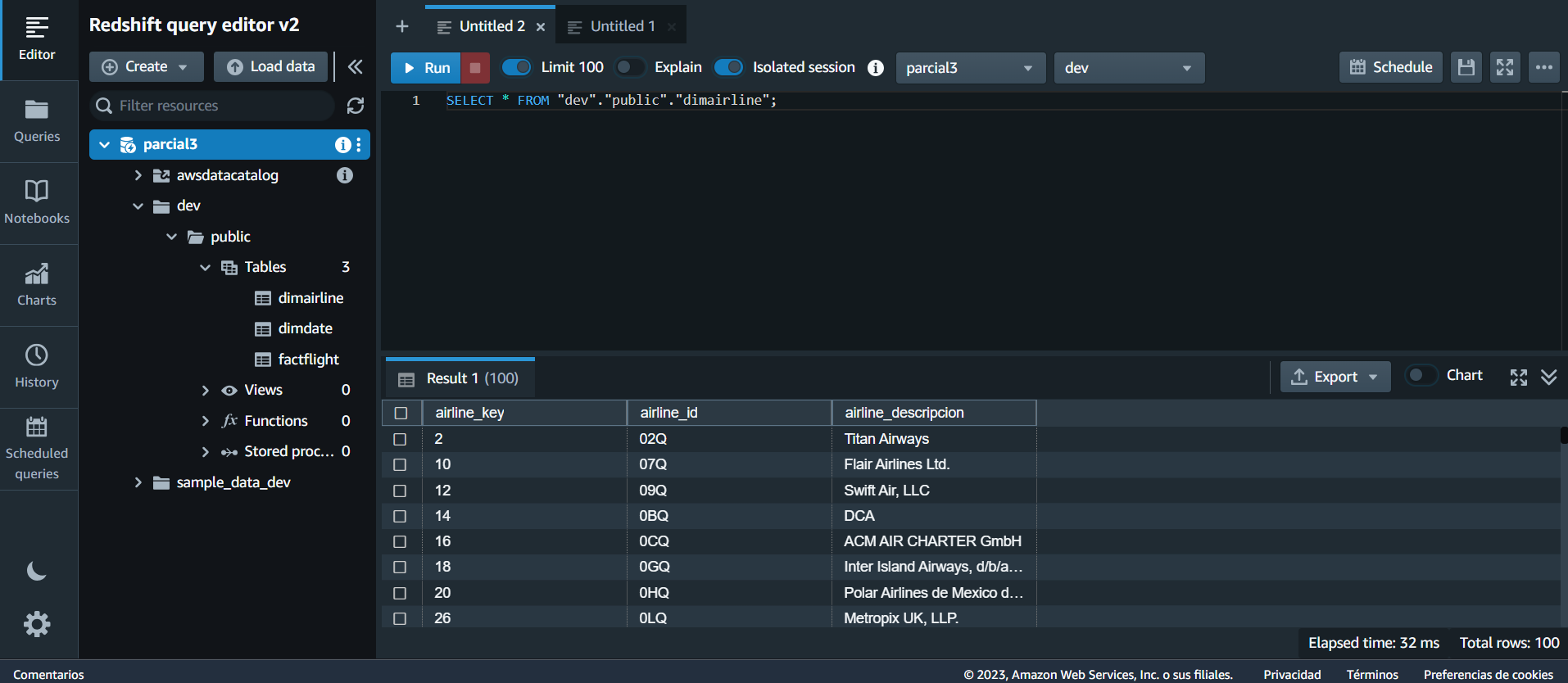
**Cluster habilitado**

****

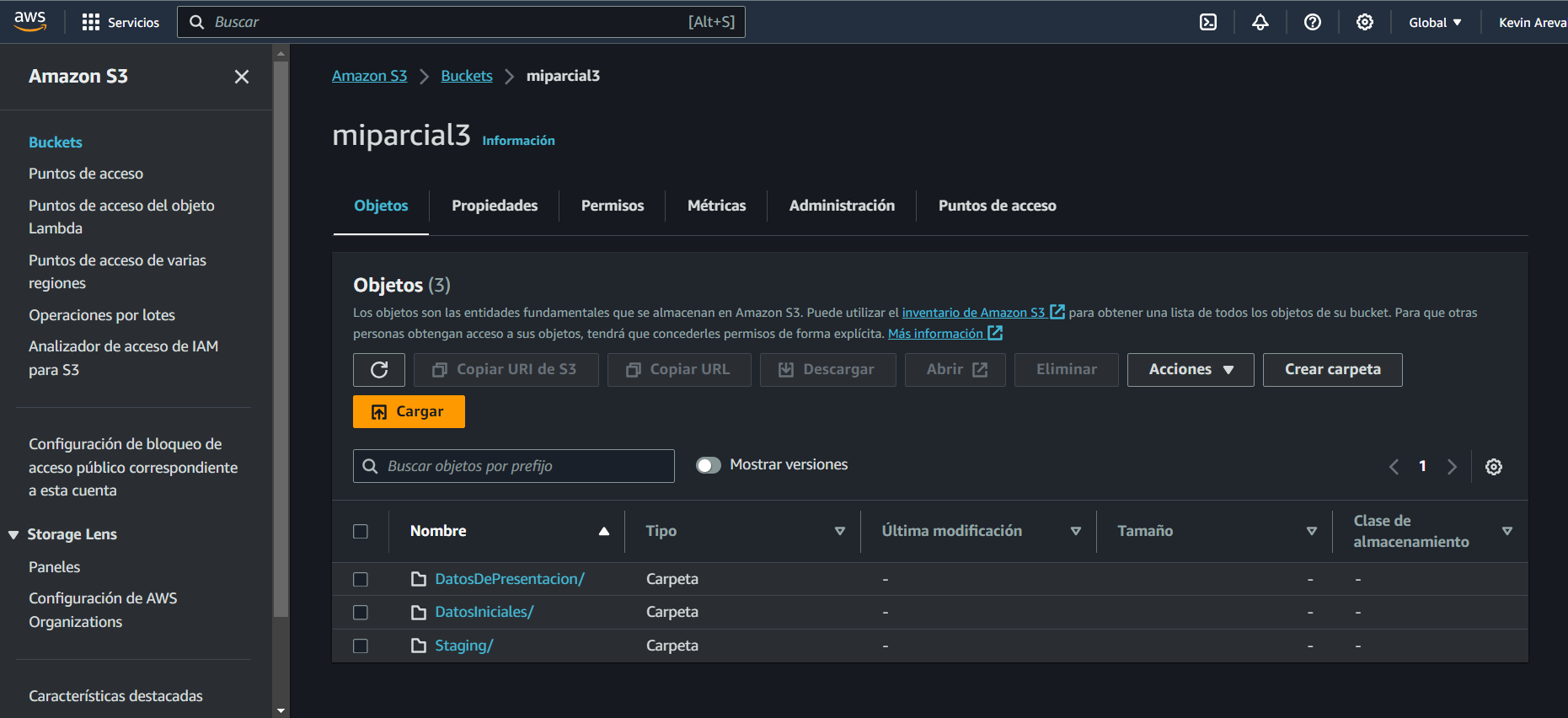
**Detalle del cluster**

****

**Ejemplo de DimAirline con los datos cargados desde S3**

****

**S3 Bucket: miparcial3**

****

**Talend**

