

# INSTITUTO POLITÉCNICO NACIONAL ESCUELA SUPERIOR DE CÓMPUTO



## ADMINISTRACIÓN DE SERVICIOS DE RED

---

### Práctica 5: Modelo en estrella de la base de datos de ingresos del metro y lattice de cubos.

---

Graciano Herrera Gabriel  
Meza Zamora Abraham Manuel

31 de mayo de 2022

# Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Desarrollo</b>	<b>3</b>
2.1. Rediseño . . . . .	3
2.2. Modificaciones a aplicar en las fases del ETL . . . . .	3
2.3. Visualización de cubos de datos . . . . .	4
<b>3. Conclusiones</b>	<b>10</b>

# 1. Introducción

En las bases de datos usadas para data warehousing, un esquema en estrella es un modelo de datos que tiene una tabla de hechos (o tabla fact) que contiene los datos para el análisis, rodeada de las tablas de dimensiones. Este aspecto, de tabla de hechos (o central) más grande rodeada de radios o tablas más pequeñas es lo que asemeja a una estrella, dándole nombre a este tipo de construcciones.

Las tablas de dimensiones tendrán siempre una clave primaria simple, mientras que en la tabla de hechos, la clave principal estará compuesta por las claves principales de las tablas dimensionales.

El esquema estrella separa los datos del proceso de negocios en: hechos y dimensiones. Los hechos contienen datos medibles, cuantitativos, relacionados con la transacción del negocio, y las dimensiones son atributos que describen los datos indicados en los hechos (una especie de meta-datos, o sea datos que describen otros datos).

Este esquema es ideal por su simplicidad y velocidad para ser usado en análisis multidimensionales (OLAP, Datamarts, EIS). Permite acceder tanto a datos agregados como de detalle.

El diseño de esquemas en estrella permite implementar la funcionalidad de una base de datos multidimensional utilizando una clásica base de datos relacional (más extendidas que las multidimensionales).

Otra razón para utilizar los esquemas en estrella es su simplicidad desde el punto de vista del usuario final. Las consultas no son complicadas, ya que las condiciones y las uniones (JOIN) necesarias solo involucran a la tabla de hechos y a las de dimensiones, no haciendo falta que se encadenen uniones y condiciones a dos o más niveles como ocurriría en un esquema en copo de nieve. En la mayoría de los casos son preferibles los de estrellas por su simplicidad respecto a los de copo de nieve por ser más fáciles de manejar.

Finalmente, es la opción con mejor rendimiento y velocidad pues permite indexar las dimensiones de forma individualizada sin que repercuta en el rendimiento de la base de datos en su conjunto.

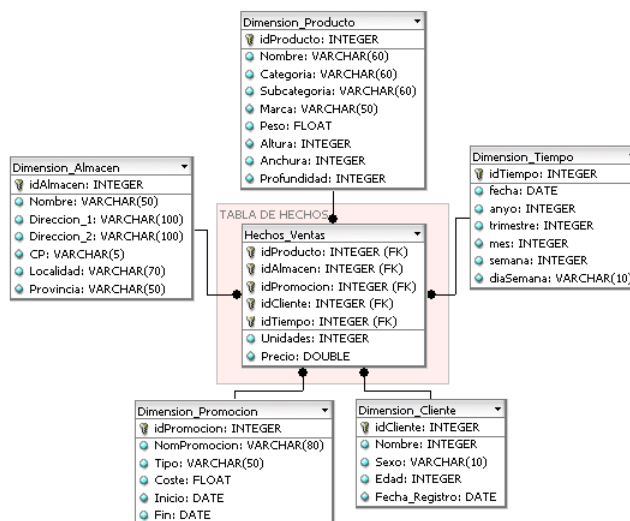


Figura 1: Modelo en estrella.

La solución propuesta para esta tarea consiste en hacer un análisis de como se va a armar el modelo en estrella y posteriormente con un script en python extraer los datos del dataset y escribir los queries necesarios para armar dicho modelo en sql.

## 2. Desarrollo

### 2.1. Rediseño

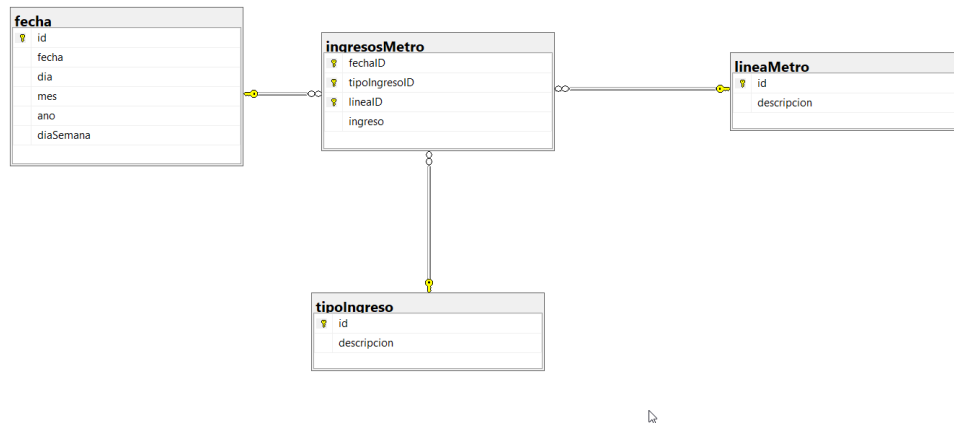


Figura 2: Diagrama ER.

### 2.2. Modificaciones a aplicar en las fases del ETL

a) Estrategia de extracción de datos

La estrategia de extracción de datos es bastante similar a las prácticas anteriores.

b) Estrategia de transformación

Se modela la base en función del modelo estrella.

c) Estrategia de carga de datos a su modelo del data lake

Se modificó el archivo para generar los queries.

### 2.3. Visualización de cubos de datos

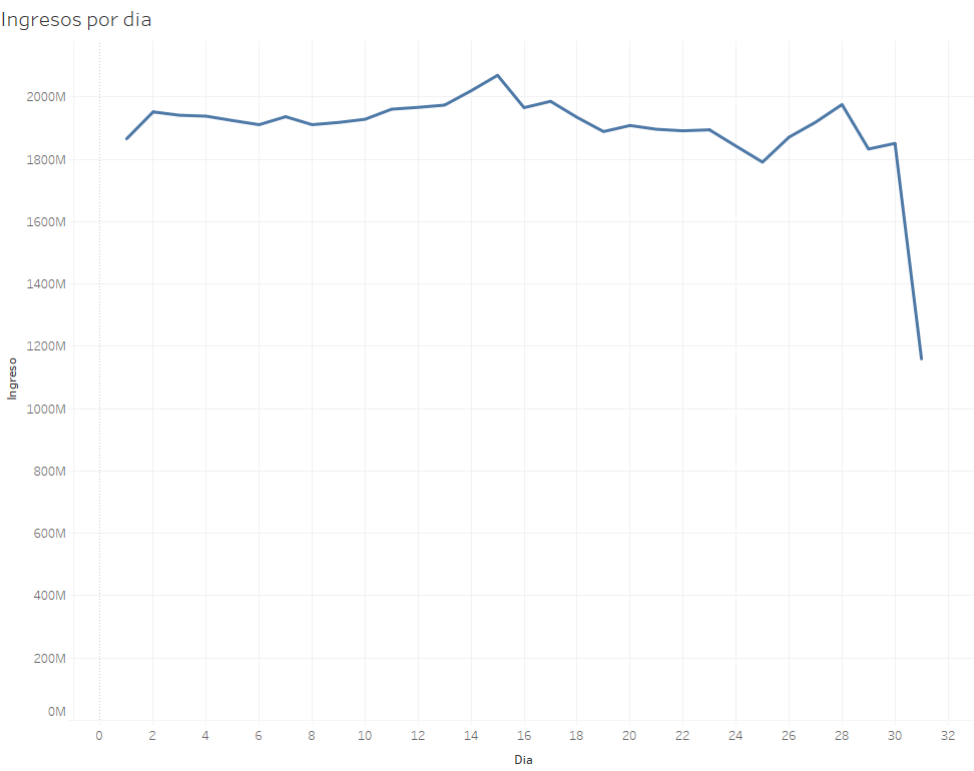


Figura 3: Ingresos por día.

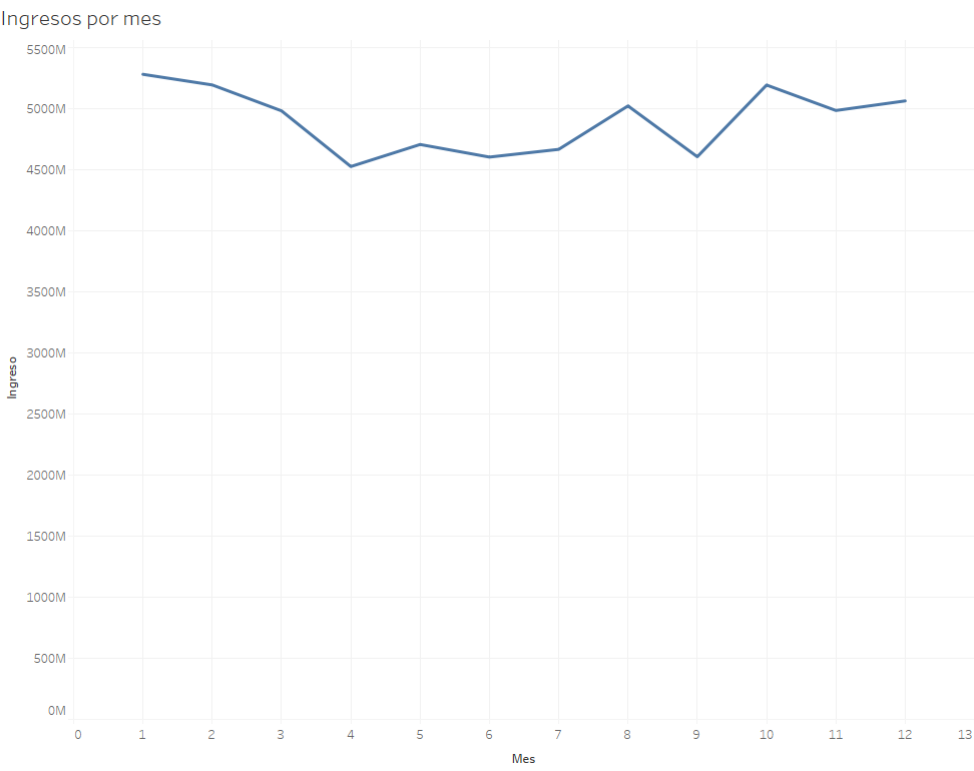


Figura 4: Ingresos por mes.

Ingresos por línea

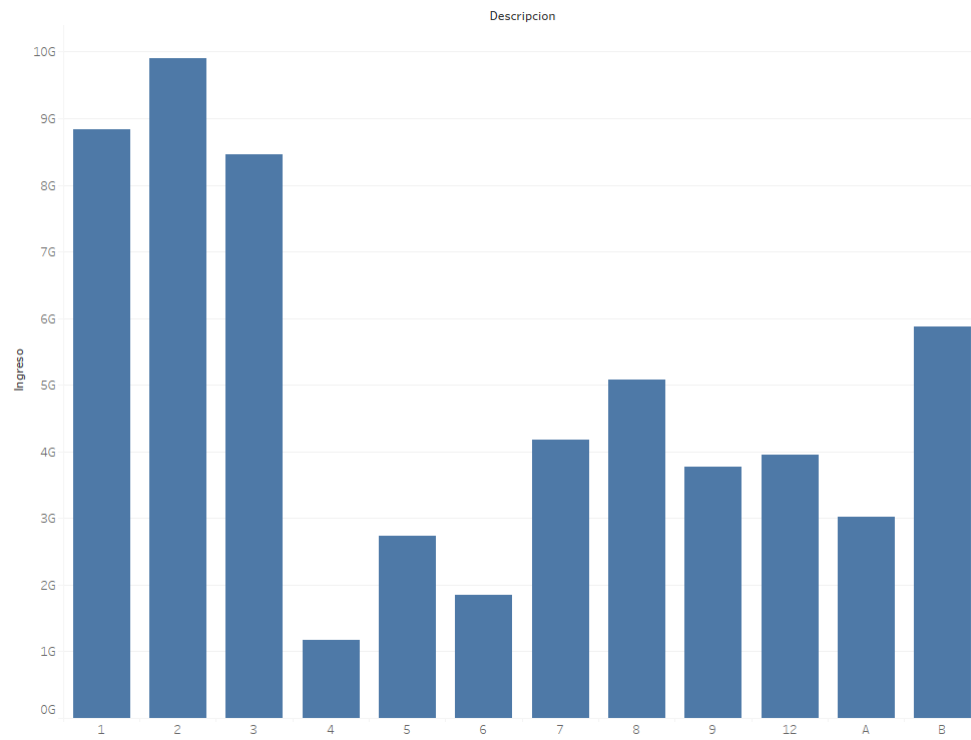


Figura 5: Ingresos por línea.

Ingresos por tipo

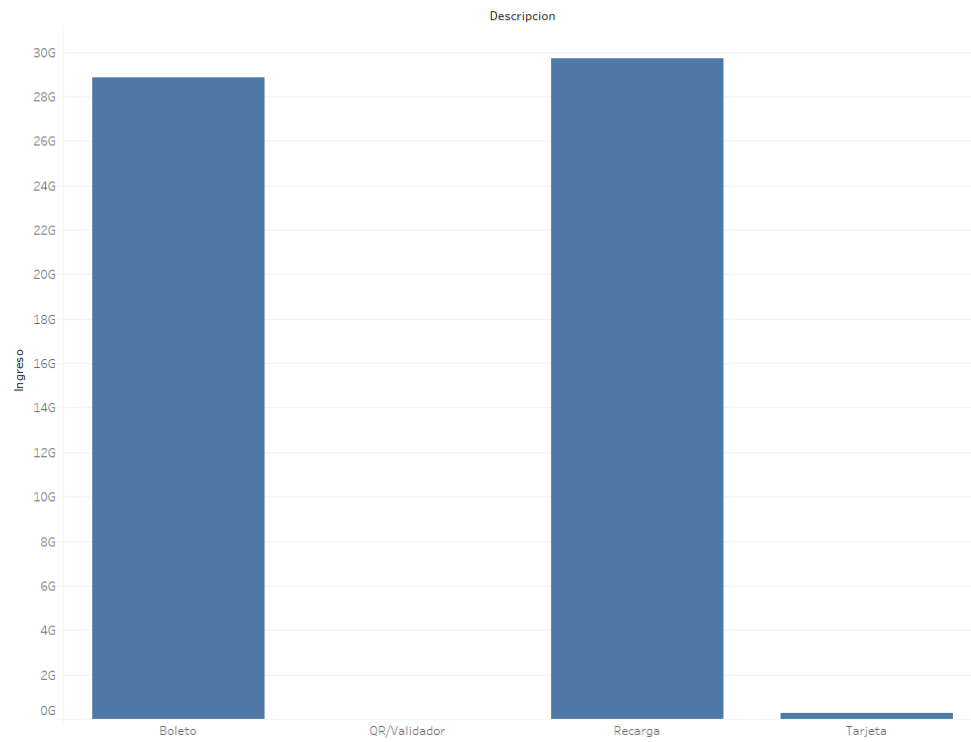


Figura 6: Ingresos por tipo.

Ingresos por línea y día de la semana

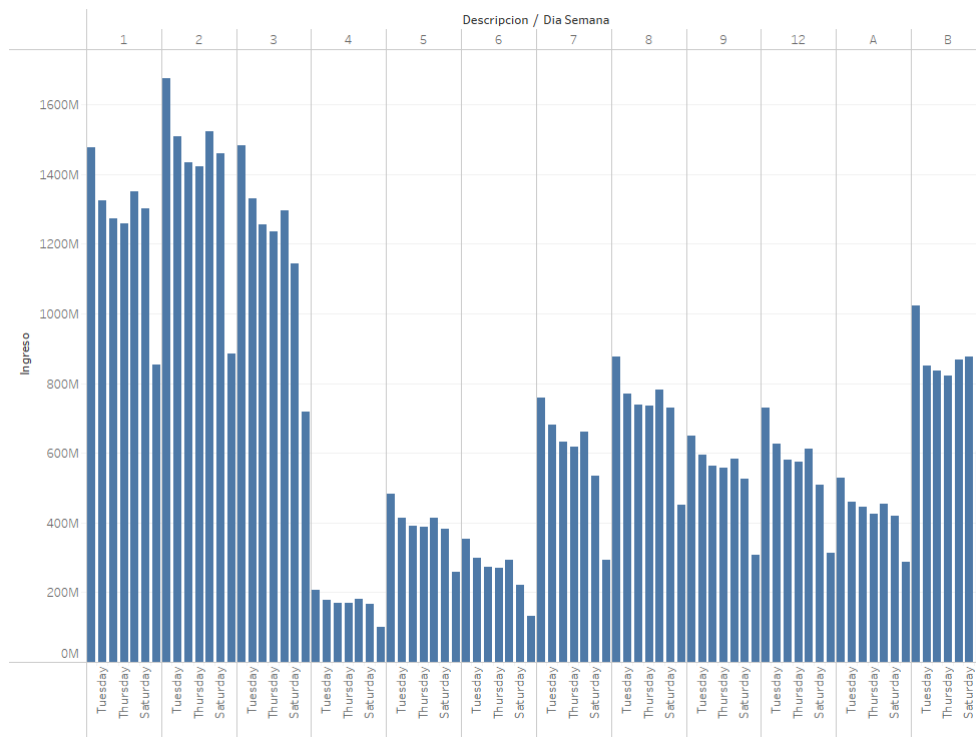


Figura 7: Ingresos por línea y por día de la semana.

Ingresos por línea y mes

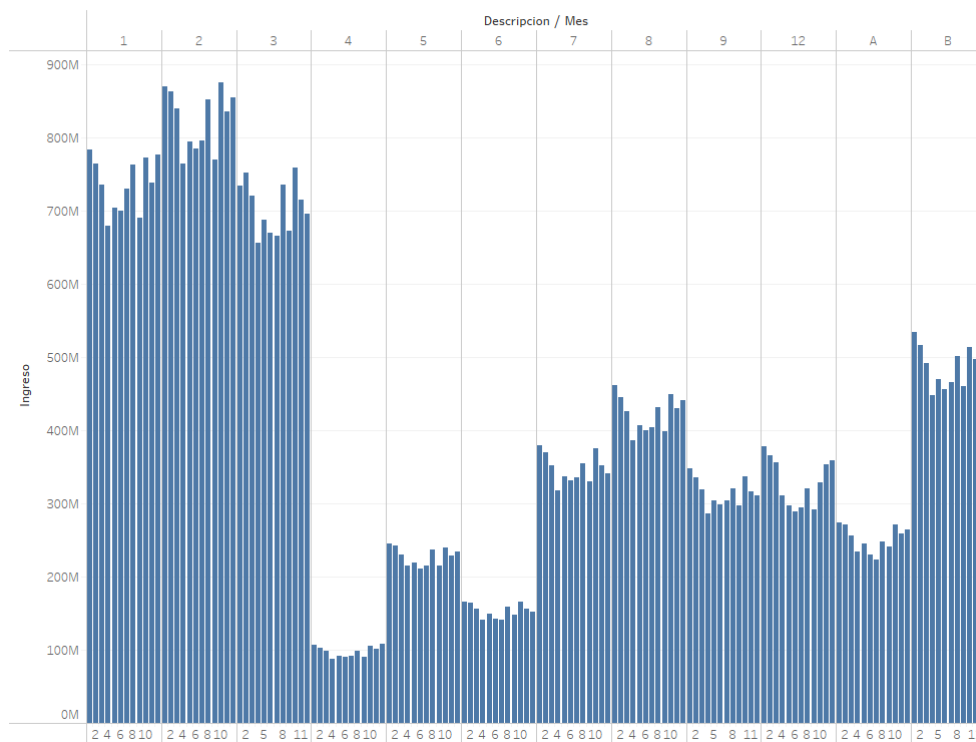


Figura 8: Ingresos por línea y por mes.

Ingreso por línea y año

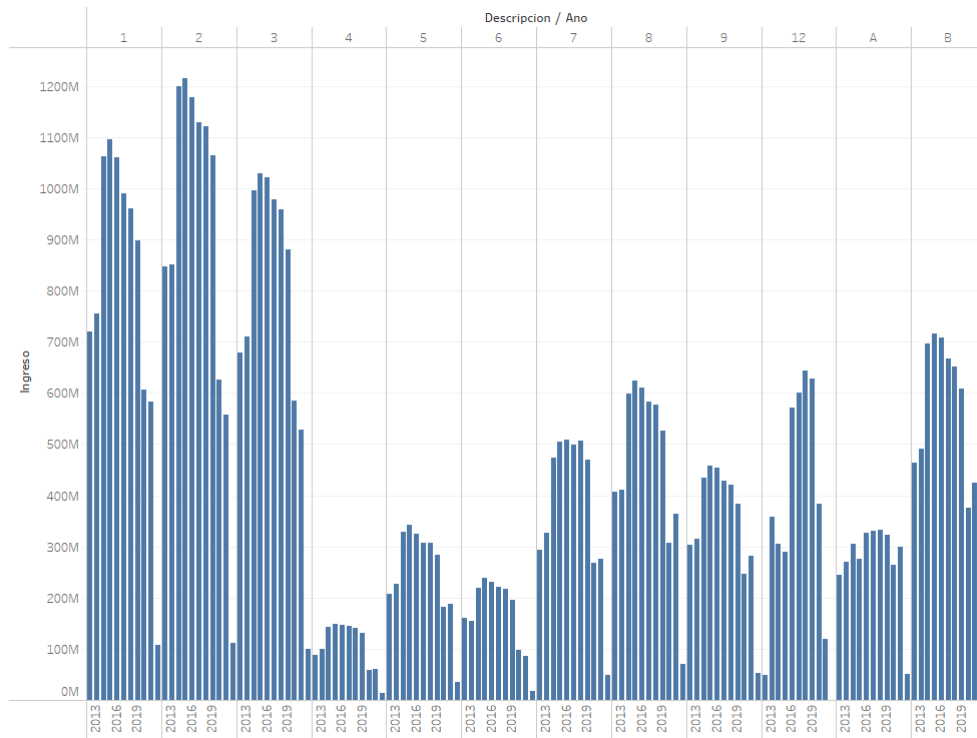


Figura 9: Ingresos por línea y por año.

Ingresos por tipo y día

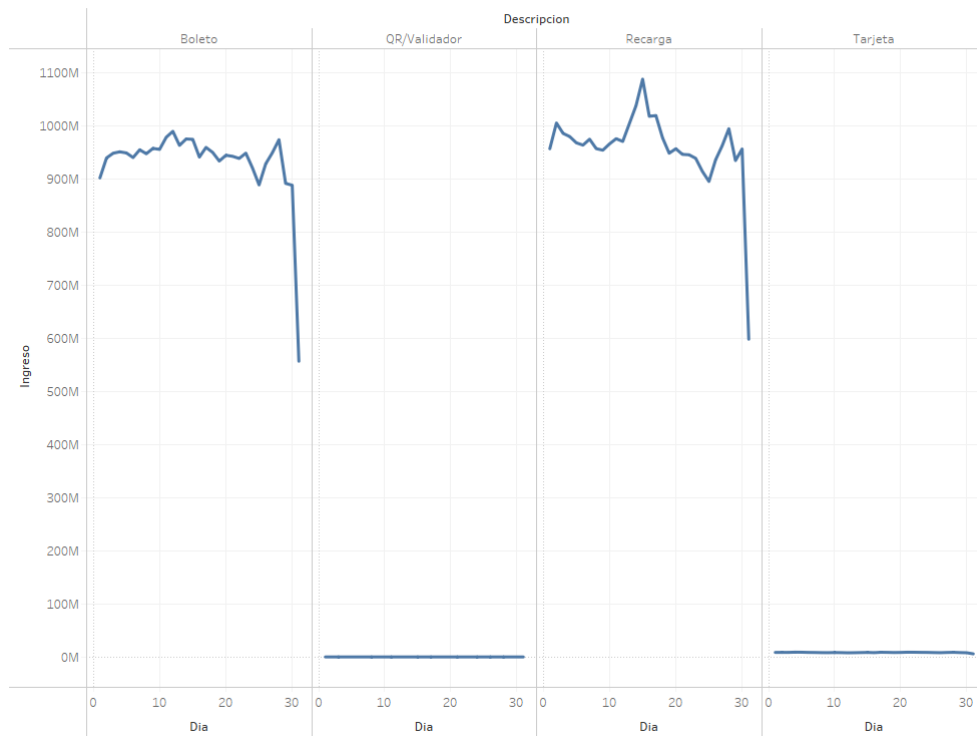


Figura 10: Ingresos por tipo y por día.



Ingresos por tipo y mes

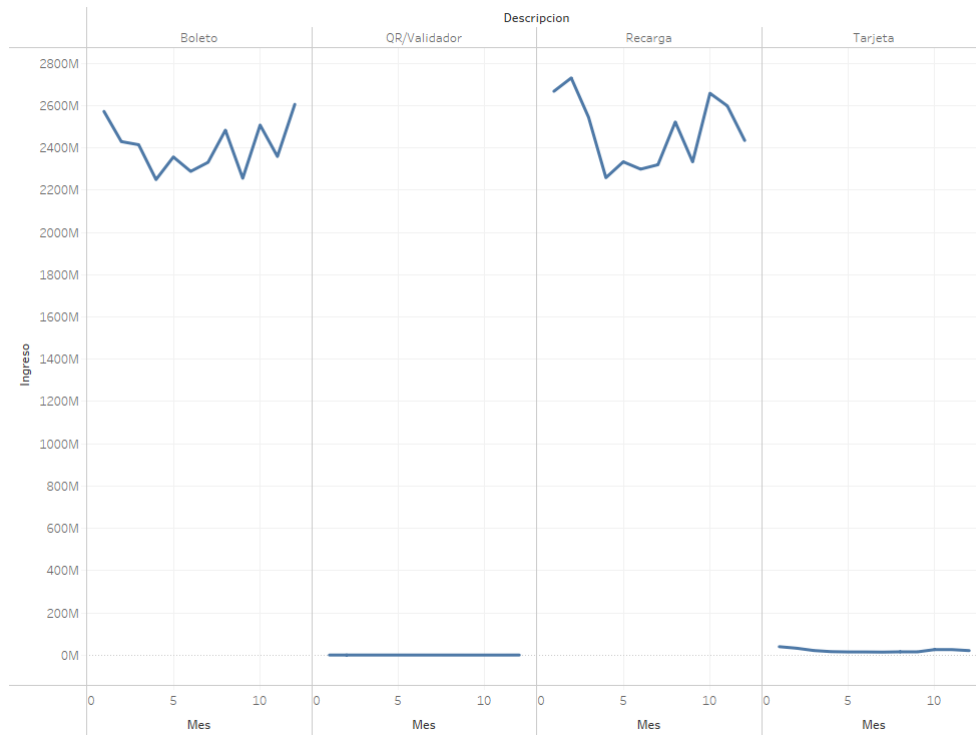


Figura 11: Ingresos por tipo y por mes.

Ingreso por tipo y día de la semana

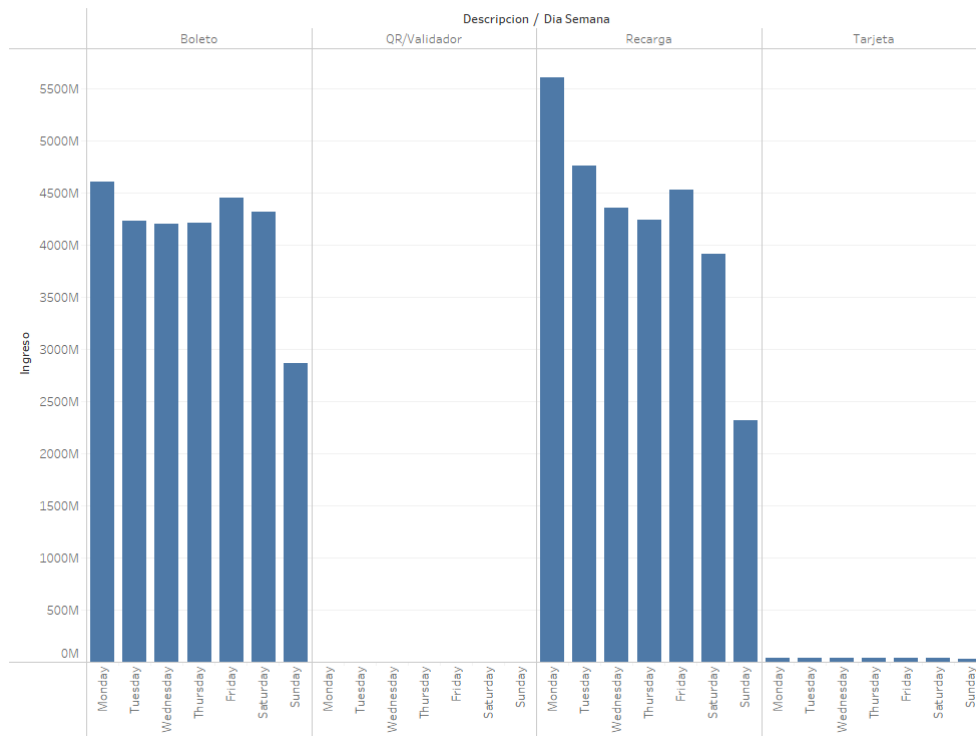


Figura 12: Ingresos por tipo y por día de la semana.

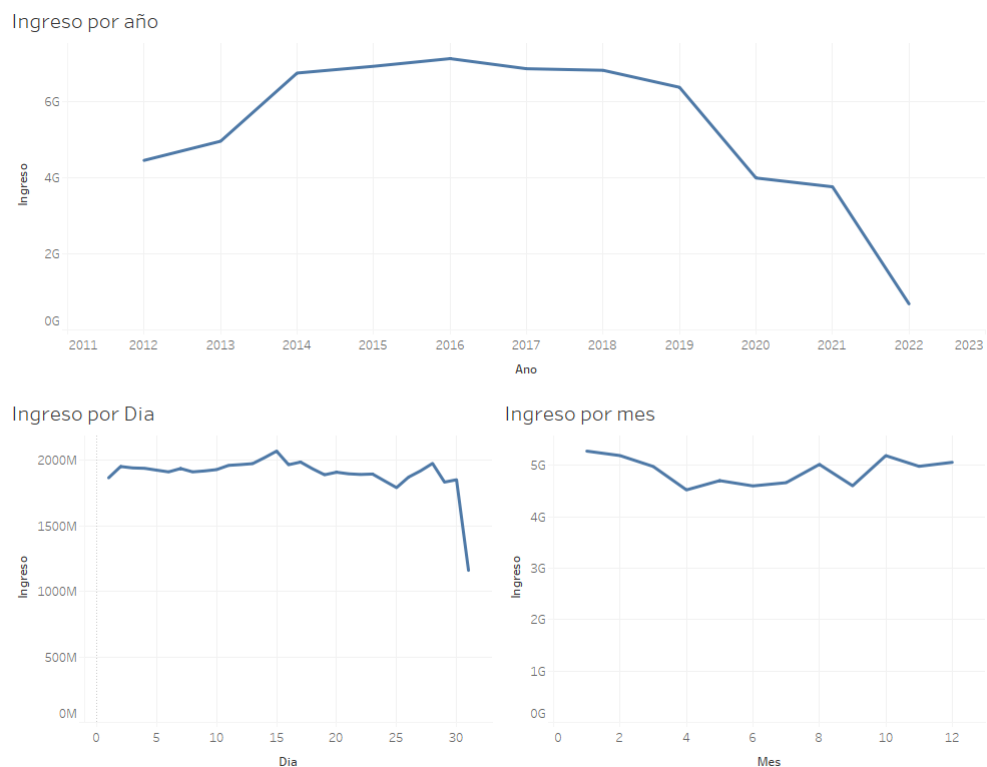


Figura 13: Vista 1.

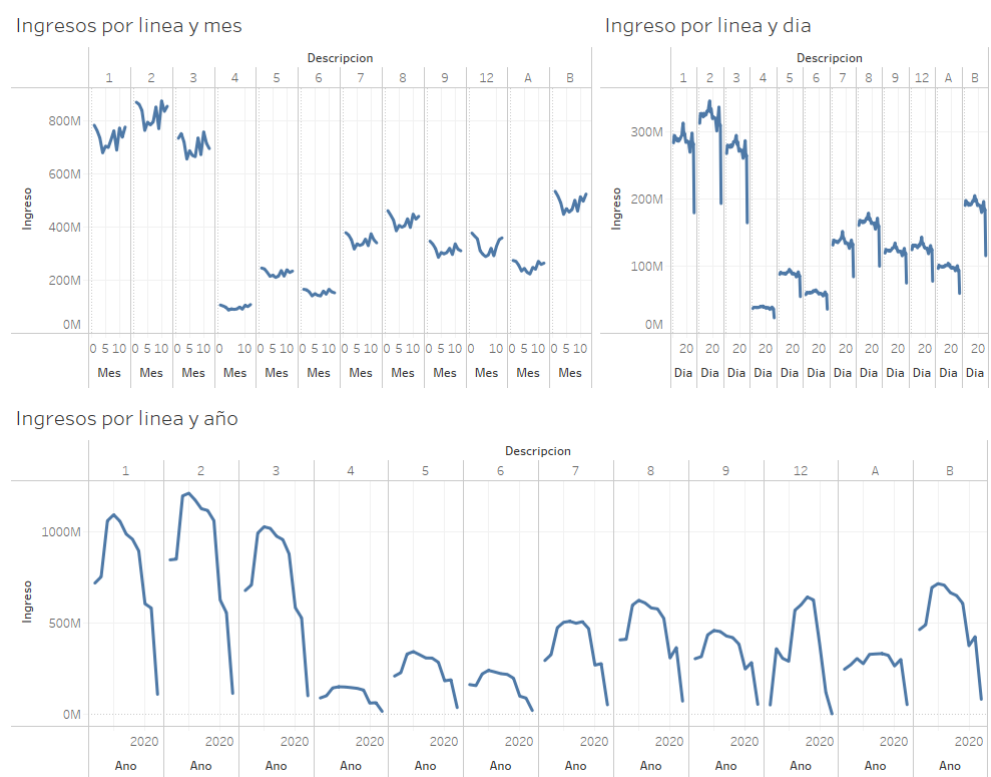


Figura 14: Vista 2.

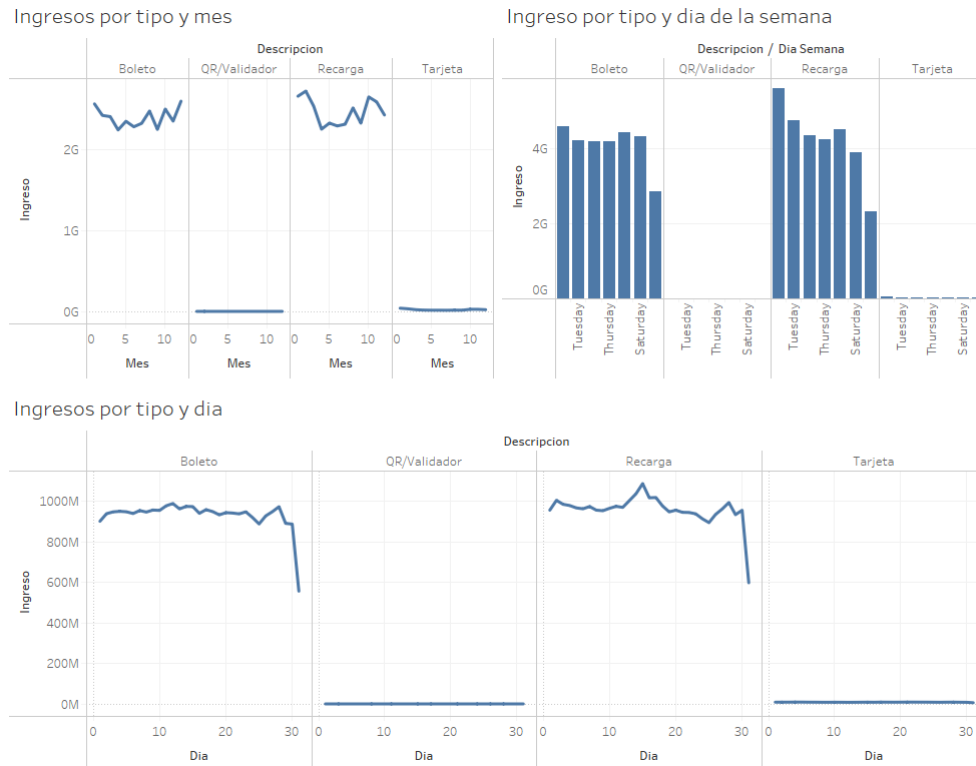


Figura 15: Vista 3.

### 3. Conclusiones

- Graciano Herrera Gabriel

El desarrollo de esta práctica nos enseñó las bondades de usar cubos de datos ya que con este ejemplo podemos ver cómo es que funcionan y la importancia que tienen para la minería de datos, crear cubos de datos adecuados y hacer sus gráficas respectivas nos ayuda a visualizar de mejor manera nuestra información a través de las distintas dimensiones de esta.

- Meza Zamora Abraham Manuel

La implementación de cubos de datos nos facilita el análisis de información con medidas y parámetros específicos según se necesite. Así mismo, facilitan la creación de gráficas al filtrar los datos que son más relevantes.