

# PROYECTO 1 – CUBOS DE INFORMACION

## *Análisis de Datos*

Como parte del curso, trabajaremos un proyecto para efectuar un análisis sobre el set de datos de recepción de pagos de diversos productos en una transacción de agencias de banco a lo largo de casi 6 años conteniendo mas de 500,000 registros.

Los bancos como entidades financieras, captan pagos para clientes diversos por medio de sus cajas receptoras por ejemplo: pagos de servicio de luz, pagos de mensualidades de colegios, universidades, etc. Esto se efectúa por medio de una transacción específica de la caja, donde se registran datos relevantes para el cliente y la forma en que se efectuó el pago. El cliente en este caso, es la empresa o institución para la cual se esta efectuando el recaudo de dinero.



### **Set de datos**

El set de datos en realidad posee pocas columnas que denominaremos características. Existe dentro del set de datos como hemos visto, datos cuantitativos y datos cualitativos. Los campos se describen a continuación:

**Operación:** Forma parte de la Identificación única de la operación efectuada en una agencia, en conjunto con la columna agencia. Es decir, pueden existir varias operaciones con la misma numeración pero de diferente agencia.

**Recibo:** identificador del pago relevante para el cliente de la transacción, con ella el cliente identifica en su sistema de manera única lo que se esta cancelando.

**Feccobro:** Fecha en la que se efectuó el cobro, el formato de esta es DDMMYYYY.

Dato en el dataset	Fecha DDMMYYYY
1122023	01/12/2023
15112023	15/11/2023

**Horcobro:** Hora en que se efectuó el cobro. El formato en el cual se encuentra HHMMSS. Ejemplos de los datos:

Dato en el dataset	Hora HHMMSS
72312	07:23:12
111511	11:15:11

\*recordar que las agencias de banco atienden generalmente de 9:00 am a 21:00.

**Agencia:** código de la agencia donde se recibió el pago.

**Nombre:** Nombre de la agencia donde se recibió el pago.

**Tipo:** Tipo de agencia donde se recibió el pago. Datos que puede tomar esta columna:

Tipo	Descripción
NORMAL	Agencia tradicional.
CASETA	Colocada en empresas que tienen una gran cantidad de operaciones de banco y bajo las cuales se le colca una agencia dedicada para sus operaciones.
PREMIUM	Agencia dirigida al segmento Premium de clientes.

**Departamento:** Departamento donde se ubica la agencia.

**Municipio:** Municipio donde se ubica la agencia.

**Producto:** Descripción del producto del que se recibió el pago.

**Operador:** Código del cajero que recibió el pago. En conjunto el código de agencia identifica de manera única al cajero.

**ValorEfectivo:** Monto del pago recibido en efectivo.

**Valorcheque:** Monto del pago recibido en cheques del banco receptor.

**ValorOtrosMedios:** Monto del pago recibido en cheques de otros bancos que se reciben bajo la reserva usual.

**ValorTotal:** Monto total del pago recibido, esto sumara las tres columnas anteriores.

El proyecto consistirá en evaluar los datos presentados, y consistirá en dos fases:

### **Fase I – Proceso de ETL FECHA DE ENTREGA: 17/03/2023**

En este proceso usted pondrá en practica, diversos conocimientos obtenidos del proceso de extracción de datos, Transformación y carga de los datos. Esto incluye efectuar un diseño dimensional adecuado para el set de datos que le permitirá en la fase II efectuar visualizaciones adecuadas para el negocio.

#### **Entregables**

-Diseño Dimensional: detallado y explicado de manera adecuada, recuerde este es su blue print, los planos sobre los cuales construirá su cubo de información.

-Diseño Entidad Relación: El diseño a nivel de la base de datos indicando los campos y llaves creadas. Incluir también el DML de la creación de las tablas.

-Flujos de Tableau Prep: Archivo conteniendo el diseño del flujo de ETL y como se efectúan las cargas de las tablas dimensionales y de hechos.

-Manual Técnico: Explicando el proceso de carga y las operaciones sobre efectuadas a los datos, así como todo lo asumido al momento de efectuar este proceso.

## **FASE II – Visualización de datos Fecha de entrega: 30/03/2023**

En esta fase el objetivo es presentar los datos o los descubrimientos efectuados por medio de tableros adecuados y diseñados por usted para tal efecto. La visualización de datos debe de presentarse en dos tableros diferentes:

- Tablero del proceso del negocio: Diseñar un tablero que permita visualizar de manera simple el proceso de negocio analizado mostrando como se ha comportando la recepción de pagos de acuerdo a las diversas dimensiones identificadas.
- Tablero de descubrimiento y exploración de datos: explorar los datos de acuerdo a diversas visualizaciones de datos y tratar de identificar algún insight que pueda ayudar al negocio a mejorar la recepción de pagos o encontrar un dato que sea relevante para el funcionamiento del negocio.

Entregables:

Manual Técnico: detallando las decisiones tomadas y la forma en que se presentaron los datos.

Archivo de tableau Prep: Conteniendo el diseño de los tableros y hojas de trabajo efectuadas.

### **Notas adicionales**

Cualquier decisión de diseño es importante dejarla bien documentada.

Cualquier cosa que se asuma también es importante dejarla documentada.

Recuerde el diseño dimensional será la base del cubo de información, no solo para la visualización correcta de los datos, sino también para el rendimiento del mismo.