

## 1.1 Antes de iniciar

Todo el proyecto más el video se encuentra en el siguiente repositorio: [github.com/JSPRM/proyectomineria](https://github.com/JSPRM/proyectomineria)

## Proceso de selección y de negocio

El tópico seleccionado y a tratar en este documento "Deuda vigente de la Ciudad de México", que entre los conjuntos de datos a elegir ([datos.cdmx.gob.mx](https://datos.cdmx.gob.mx)) nos pareció ser la mejor opción al ser recientemente actualizado y contar con suficientes campos cuantitativos, un buen punto de partida para empezar a poner en práctica las aptitudes obtenidas a lo largo del curso.

Para darnos una visión general del concepto de deuda y sus tipos tomaremos un vistazo a distintas definiciones.

Deuda estatatal: se refiere únicamente a la deuda de los gobiernos estatales y sus entes públicos, excluyendo municipios, alcaldías o sus organismos paramunicipales.

Deuda subnacional: se refiere a todas las obligaciones de pago o pasivos, contratadas directa o indirectamente, derivadas de crédito o financiamiento a cargo de los estados, municipios, entes estatales o municipales, la Ciudad de México y sus alcaldías. La deuda subnacional incluye la amortización, los intereses, gastos y comisiones que genera la contratación de la misma.

Asimismo, para comprender el conjunto de datos en cuestión, hay que darle un vistazo al diccionario de datos proporcionado.

Nombre completo	Definición	Variable
Año	Año fiscal vigente, del 1 de enero al 31 de diciembre.	ANIO
Trimestre	Enero-marzo, abril-junio, julio-septiembre y octubre-diciembre.	TRIMESTRE
Mes	Mes de cierre de cada trimestre.	MES
No. de Registro	Número que asigna la Unidad de Coordinación con Entidades Federativas de la Secretaría de Administración y Finanzas.	NO_REGISTRO
Tipo deuda	Plazo de la deuda.	TIPO_DEUDA
Detalle de tipo deuda	Tipo de banco que realiza el crédito.	DETALLE_TIPO_DEUDA
Acreedor	Banco que otorga el crédito.	ACREEDOR
Concepto	Nombre del crédito.	CONCEPTO
Inicio del crédito	Fecha en la cual inicia el crédito.	INICIO_CREDITO
Fin del crédito	Fecha en la cual finaliza el crédito.	FIN_CREDITO
Días de contrato	Días de duración del crédito.	DIAS_CONTRATO
Días restantes	Días restantes para la terminación del crédito.	DIAS_RESTANTES_CONTRATO
Disposición inicial del crédito	Recurso utilizado del monto del crédito contratado.	DISPOSICION_INICIAL_CREDITO
Colocación al periodo	Proceso mediante el cual el gobierno recibe créditos que se normalizan a través de contratos, que amparan obligaciones derivadas del ejercicio presupuestal, los cuales representan medios de financiamiento para el Sector Gobierno.	COLOCACION_PERIODO
Pago de amortizaciones al periodo	Pago al capital de un crédito en un periodo.	AMORTIZACIONES_PERIODO
Pago de intereses al periodo	Pago al interés de un crédito en un periodo.	INTERESES_PERIODO
Pago por servicio de la deuda	Pago de amortizaciones más el costo financiero en un periodo.	PAGO_SERVICIO_DEUDA
Saldo al periodo	Monto restante por pagar del principal de un crédito al cierre del Trimestre.	SALDO_PERIODO
Endeudamiento al periodo	Saldo del crédito al trimestre menos saldo del crédito del trimestre anterior.	ENDEUDAMIENTO_PERIODO
Tasa	Tasa de interés fija o variable (TIIE) que se paga por el crédito. La TIIE Es una tasa representativa de las operaciones de crédito entre bancos. Es calculada diariamente (para plazos 28 y 91 días) por el Banco de México con base en cotizaciones presentadas por las instituciones bancarias mediante un mecanismo diseñado para reflejar las condiciones del mercado de dinero en moneda nacional. La TIIE se utiliza como referencia para diversos instrumentos y productos financieros, tales como tarjetas de crédito.	TASA
Sobretasa	Costo adicional a la tasa de interés variable que se paga por un crédito.	SOBRETASA
Tasa Final	Tasa más la sobretasa de un crédito.	TASA_FINAL

A continuación procedimos a analizar el conjunto de datos con apoyo de la herramienta Dataiku DSS.

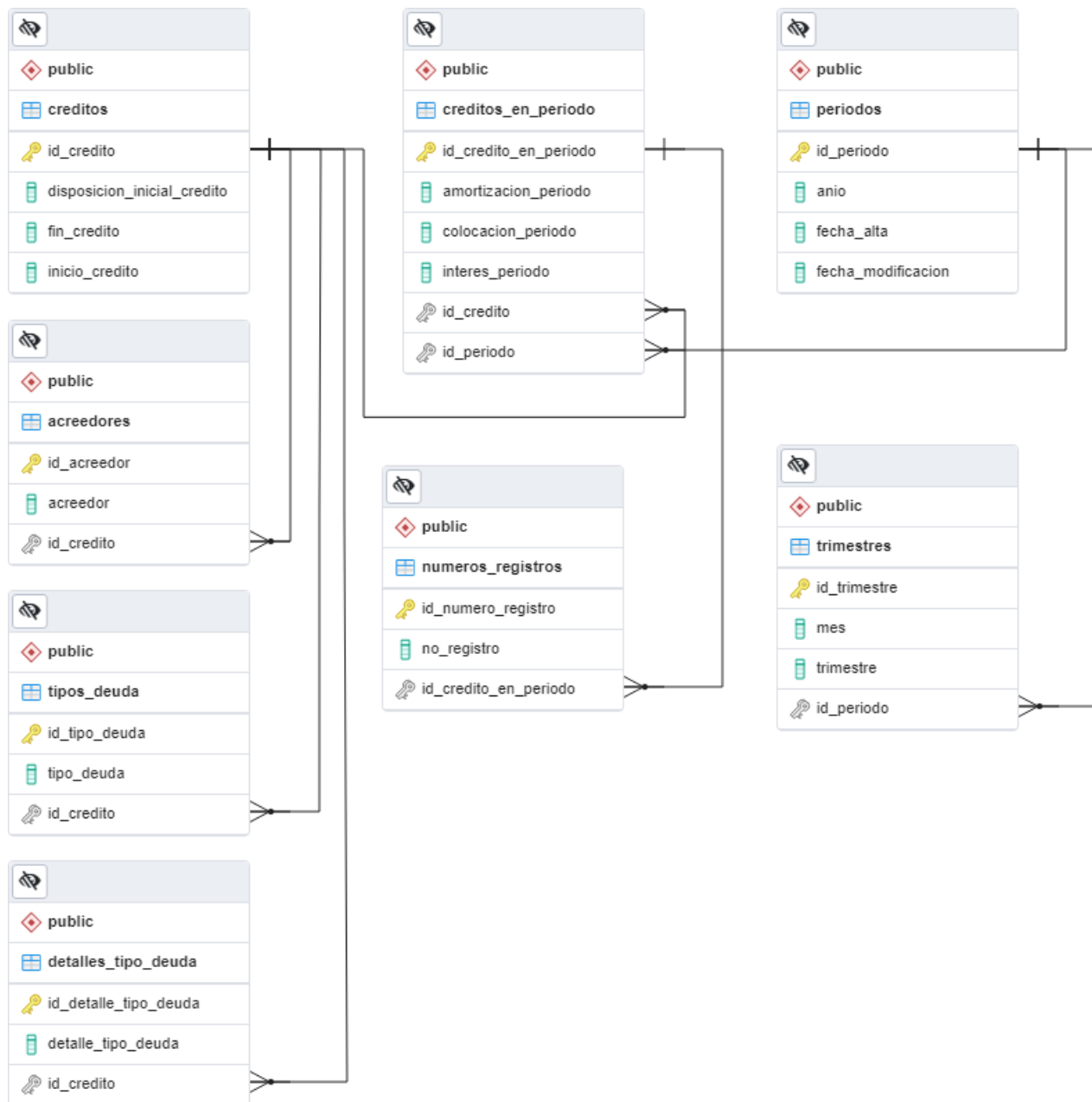
Una vez analizado al detalle tanto el diccionario de datos como los propios datos, nos dimos la libertad de eliminar algunos campos, estos fueron: tasa, sobretasa y tasa final. La razón principal fue que, a nuestro parecer, estos campos no tienen un formato adecuado, debido a que en ellos se trata de sumar cadenas de

caracteres con números decimales.

Dado por concluidas las modificaciones y limpieza del conjunto de datos original pasamos a plantearnos el modelo OLTP.

## Diseño base de datos: Modelo OLTP

El objetivo en nuestra mente al realizar este modelo fue el de proporcionar facilidad al momento de crear nuevos registros de créditos, pero asimismo tener siempre la posibilidad de redefinir parte del registro, ya que es muy posible que durante el ciclo de vida de la base de datos algunas terminologías/categorías sean modificadas.



Se hizo uso de PostgreSQL y pgAdmin para llevar a cabo tanto la base de datos como el diagrama entidad relación presentado.

El backup de la base de datos con todos sus registros se encuentra en el repositorio indicado al inicio de este documento.

## Consultas de negocios: Modelo OLTP

¿Cuál es el total de amortizaciones en cada trimestre?

```
SELECT SUM(amortizacion_periodo), anio, trimestre
from creditos_en_periodo c
INNER JOIN periodos p on c.id_periodo = p.id_periodo
GROUP BY anio, trimestre
ORDER BY anio asc;
```

¿Cuál es el total de intereses pagados en cada trimestre?

```
SELECT SUM(interres_periodo), anio, trimestre
from creditos_en_periodo c
INNER JOIN periodos p on c.id_periodo = p.id_periodo
GROUP BY anio, trimestre
ORDER BY anio asc;
```

¿Cuánto se ha pagado de interés a cada banco/acreedor?

```
SELECT SUM(interres_periodo), acreedor
from creditos_en_periodo c
INNER JOIN creditos c2 on c2.id_credito = c.id_credito
GROUP BY acreedor;
```

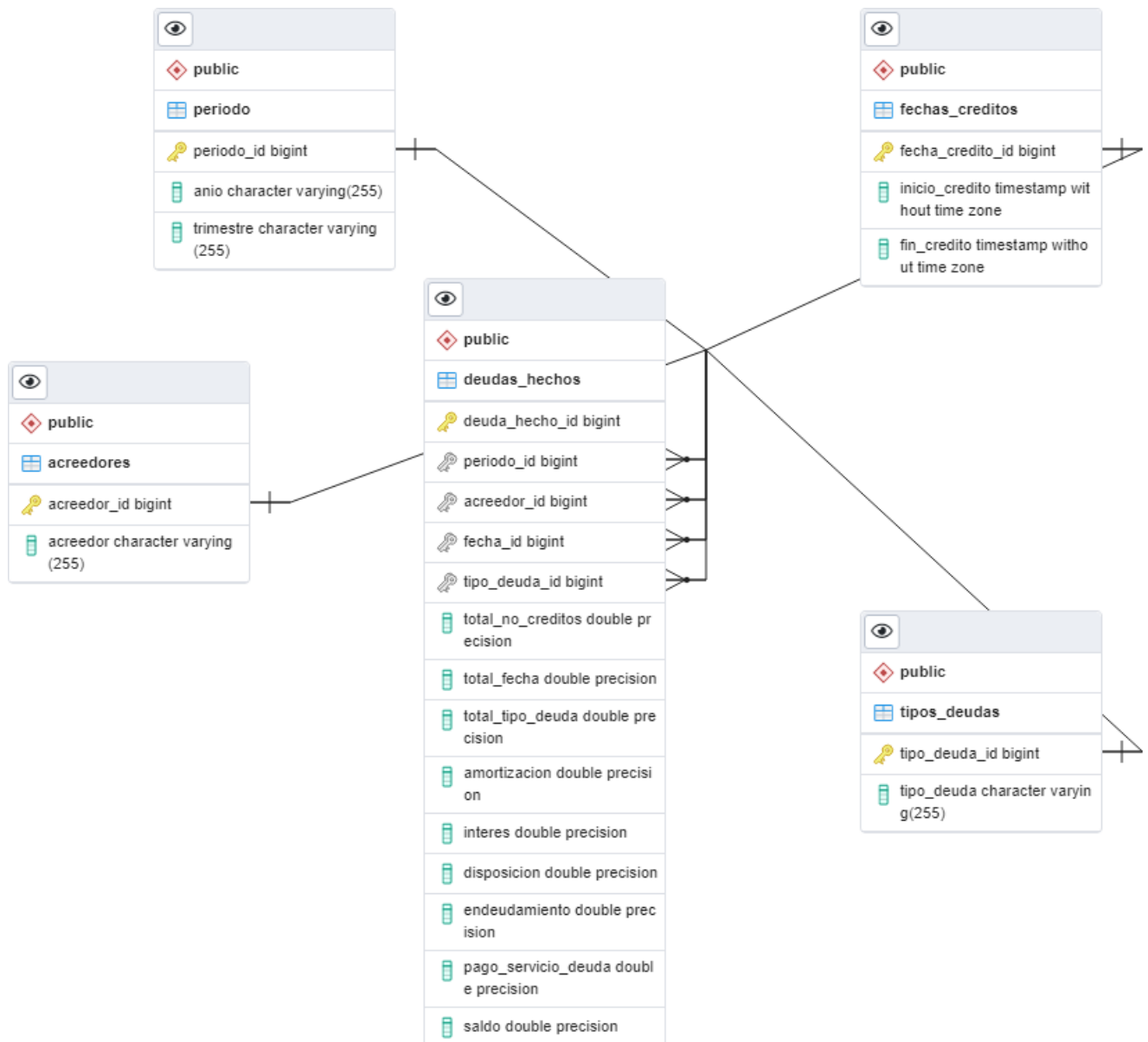
¿Cuánto me ha prestado cada banco?

```
SELECT SUM(disposicion_inicial_credito), acreedor
from creditos
GROUP BY acreedor;
```

¿Cuánto se debe por crédito?

```
SELECT tmp.sum, tmp.disposicion_inicial_credito, tmp.disposicion_inicial_credito - tmp.sum as
    resta_por_pagar
FROM (
    SELECT SUM(amortizacion_periodo), inicio_credito, fin_credito, disposicion_inicial_credito
    FROM creditos_en_periodo
        INNER JOIN creditos c on c.id_credito = creditos_en_periodo.id_credito
    GROUP BY inicio_credito, fin_credito, disposicion_inicial_credito
) as tmp;
```

## Diseño Data Warehouse: Modelo OLAP



Se planteó un modelo de tipo estrella como fue requerido, contando con cuatro tablas de dimensiones, una de ellas dedicada a las fechas de inicio y finalización de los créditos con el objetivo de tener un registro histórico.

Usando Dataiku realizamos las transformaciones necesarias a los datos dentro del OLTP para posteriormente ser guardados en nuestra base de datos OLAP.

El flujo de ETL lo podemos observar a continuación



Recomendaciones: Como ya hemos aprendido, la base de datos OLAP la ocuparemos para datos que no se vayan a modificar y que nos sirvan de registro histórico, también con estos datos haremos análisis y posibles modelos de aprendizaje automático.

Por otra parte, la base OLTP la usaremos para subir los nuevos registros que se realicen con regularidad, en el caso de este negocio los registros son semestre, y hacer consultas que sea necesario.

En general este trabajo fue un ejercicio interesante para nosotros, aprendimos mucho y nos hemos dado cuenta de áreas en las que podemos mejorar. Nos ha sido de mucha ayuda para conocer nuestras fortalezas y gustos, que en este momento es crítico para nuestro desarrollo en el ámbito laboral que ya es una realidad.