

UNIVERSIDAD NACIONAL DE ROSARIO

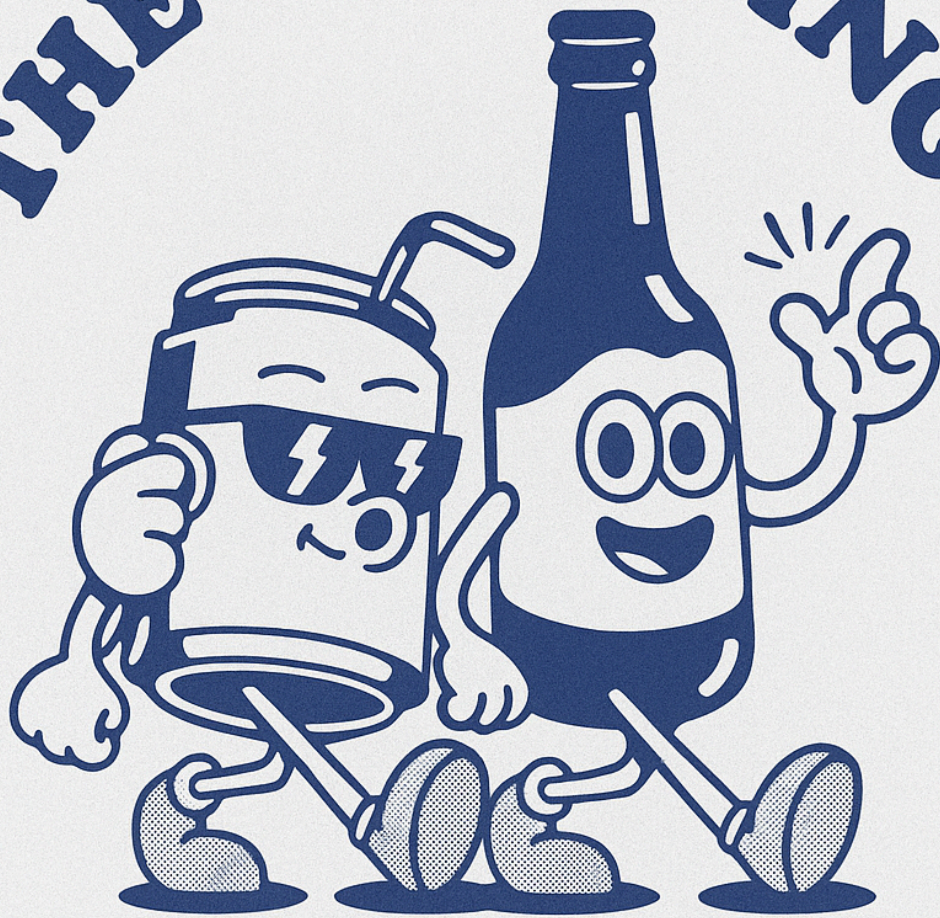
**Facultad de Ciencias Exactas,
Ingeniería y Agrimensura**

**Tecnicatura Universitaria en
Inteligencia Artificial**

**Trabajo práctico final Bases de datos 2
ETL “The drinking company”**

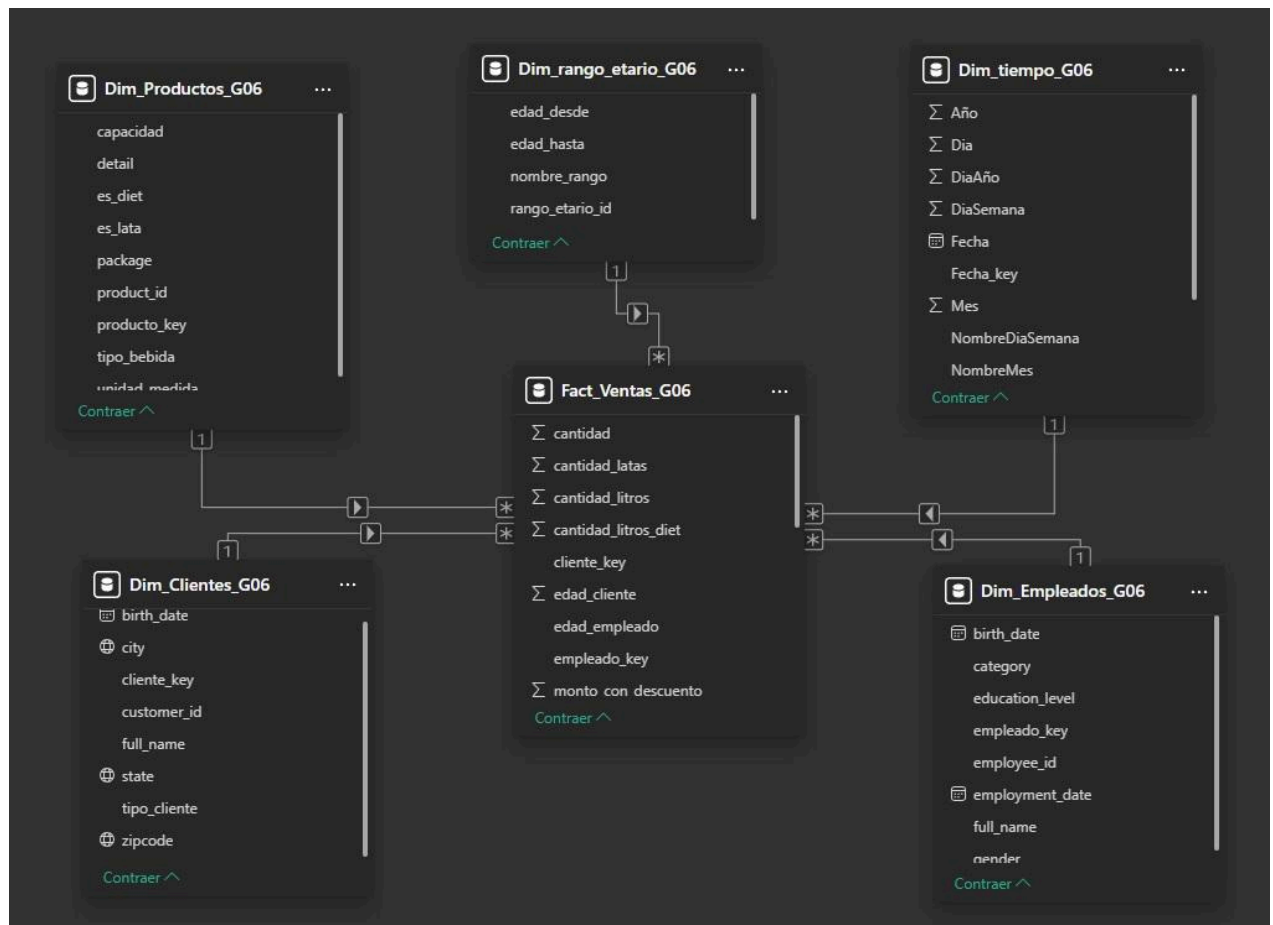
Integrantes: Manuela Almada, legajo: A-4692/2
Fabián Alvarez, legajo: A-4501/2

THE DRINKING



COMPANY

1. Diagrama modelo estrella



2. E.T.L. The drinking company

El proceso comienza desde Management Studio SQL Server ejecutando el archivo 01_crear_bases_datos.sql para crear las bases de datos:

- bd_staging_2025_G06
- db_intermedia_2025_G06
- datawarehouse_2025_G06

Luego en bd_staging_G06 se crean las tablas donde se cargará la información de las fuentes de origen de los distintos sectores de la empresa. Estas tablas se generan automáticamente desde SSIS en Visual Studio.

- stg_empleados_G06
- stg_clientes_min_G06
- stg_clientes_may_G06
- stg_productos_G06
- stg_regiones_G06
- stg_feriados_G06
- stg_stock_G06
- stg_ventas_historicas_G06
- stg_billing_G06
- stg_billing_detail_G06
- stg_prices_G06
- stg_discounts_G06

En bd_intermedia_G06 desde Management Studio ejecutar el archivo 02_crear_int_RangoEtario.sql para crear la tabla:

- int_rango_etario_G06

En datawarehouse_2025_G06 desde Management Studio se crean las tablas de dimensiones ejecutando: 03_crear_dimensiones.sql.

- Dim_Clientes_G06
- Dim_Empleados_G06
- Dim_Productos_G06
- Dim_rango_etario_G06
- Dim_tiempo_G06

Luego se cargan las fechas a la dimensión tiempo ejecutando 04_cargar_fechas.sql.

En datawarehouse_2025_G06 se ejecuta 05_crear_Fact_ventas.sql para crear la tabla de hechos

Ahora si desde Visual Studio con SSIS se ejecuta la solución ETL_TDC.sln para ir cargando las tablas desde los archivos de origen hasta llenar las dimensiones en el datawarehouse, pasando previamente por algunas transformaciones o limpieza de datos.

Se recomienda ejecutar esas cargas de a una dimensión a la vez, deshabilitando el resto de bloques hasta llegar al último que es la carga de las ventas en la Fact_Ventas (todas las dimensiones ya deben existir y estar cargadas para que se pueda generar).

Adjuntamos el archivo 06_cargar_FactVentas.sql como backup, pero el mismo se ejecuta automáticamente en el último bloque del flujo de datos en SSIS.
Finalmente agregamos al datawarehouse una tabla log de errores para capturar registros con datos faltantes: 07_log_errores.sql.

3. Dashboard en Power BI

Cantidad de litros consumidos y productos adquiridos por cliente en el tiempo.

Compra promedio por cliente en el tiempo.

Clientes minoristas más valiosos según el principio de Pareto.



Litros vendidos por rango etario y por tipo de bebida.
Cantidad de productos y litros vendidos a lo largo del tiempo.



Monto de ventas por empleado según género y edad.
Monto de ventas por empleado según antigüedad en la empresa.
Ranking de empleados con mayor monto de ventas.



Ranking de productos vendidos por región.

Productos menos vendidos.

Litros de bebidas diet vendidas por año.

Cantidad de latas vendidas por año.



Clientes más valiosos para la empresa.

Monto de compra a través del tiempo segmentado por mes, trimestre y año.

