



PROYECTO INTEGRADOR

"Descubriendo la Base de Datos de Fast Food"

Carrera: Data Analytics

Cohorte: DA-FT18

Estudiante: Melisa Rossi

Email: melirossi.mr@gmail.com

Fecha de entrega: 04 de Septiembre de 2025

Introducción

El desarrollo de este proyecto representó una valiosa inmersión en el mundo de la gestión y el análisis de datos, asumiendo el rol de Data Analyst para la marca en crecimiento **Fast Food**, dentro del competitivo sector de comida rápida.

Aunque la empresa había logrado un rápido crecimiento con la apertura de nuevas sucursales, enfrentaba un desafío importante: la información se encontraba fragmentada en múltiples archivos de Google Sheets y documentos manuales, lo que dificulta la eficiencia operativa y la toma de decisiones estratégicas.

Mi intervención tuvo como objetivo liderar la transición hacia una base de datos relacional que permita centralizar la información de órdenes, productos, empleados, clientes, sucursales y métodos de pago. Asimismo, se buscó generar consultas avanzadas en SQL Server que aportaran respuestas a preguntas de negocio críticas para la gestión estratégica de la empresa.

Módulo 2

Desarrollo del proyecto

La tarea consistió en cuatro avances progresivos:

Avance 1: Diseño y creación de la base de datos

En esta primera etapa se construyó la base de datos “FastFood” y se definió su modelo relacional.

- Se crearon las tablas mediante la sentencia CREATE TABLE. Se crearon las tablas: Ordenes, Productos, Clientes, Empleados, Sucursales, Mensajeros, Categorías, TiposPago, OrigenesOrden, y, la tabla intermedia “DetalleOrdenes” como relación entre Ordenes y Productos.
- Se establecieron claves primarias y foráneas para garantizar la integridad referencial y modelar adecuadamente las relaciones.

Esta normalización permite reducir redundancia y preparar la base para consultas eficientes.

Avance 2: Población y primeras consultas exploratorias

En esta fase se realizaron las primeras cargas de datos y la aplicación de sentencias DML, lo que permitió validar que las relaciones entre tablas funcionaran correctamente.

- Se implementó la sentencia INSERT INTO para insertar nuevos registros a cada tabla de la base de datos.
- Se aplicaron sentencias como UPDATE y DELETE para poner a prueba la base de datos.
- Se ejecutaron consultas simples de exploración, cantidad total de registros únicos, empleados por departamento, productos por categorías, etc., mediante funciones agregadas COUNT, AVG.

Módulo 2

Este avance permitió comprender en profundidad la lógica del modelo de datos y familiarizarse con la sintaxis de SQL para su correcta implementación.

Avance 3: Consultas de negocio avanzadas

Aquí se formularon las primeras consultas de análisis avanzado para responder preguntas clave de negocio, mediante funciones de agregación:

- Análisis de ventas globales y ventas por tipo de pago para medir facturación total y distribución de medios de pago, mediante la función agregada SUM .
- Obtención de promedios, como promedio de precios de productos por categoría y promedio de cantidad de productos por orden, mediante la función AVG.
- Orden mínima y máxima por sucursal para identificar dispersiones en tickets, aplicando las funciones MIN y MAX respectivamente; y, mayor número de kilómetros recorridos para entrega, nuevamente mediante la función agregada MAX.
- Sucursal con la venta promedio más alta y sucursales con ventas superiores a \$1000, indicadores de desempeño comparativo. La primera aplicando la función agregada AVG y trayendo el valor solicitado mediante TOP 1, y la segunda, aplicando SUM dentro de un filtro HAVING.
- Comparación de ventas antes y después del 1 de julio de 2023, útil para analizar tendencias temporales, .
- Actividad de ventas por horario (mañana, tarde, noche) para entender patrones de consumo, mediante las funciones de agregación COUNT, AVG y MAX de distintas columnas de la tabla "Ordenes".

Este avance permitió obtener métricas concretas que reflejan el funcionamiento de la empresa y brindan insumos estratégicos.

Avance 4: Consultas exploratorias y enriquecimiento de la información

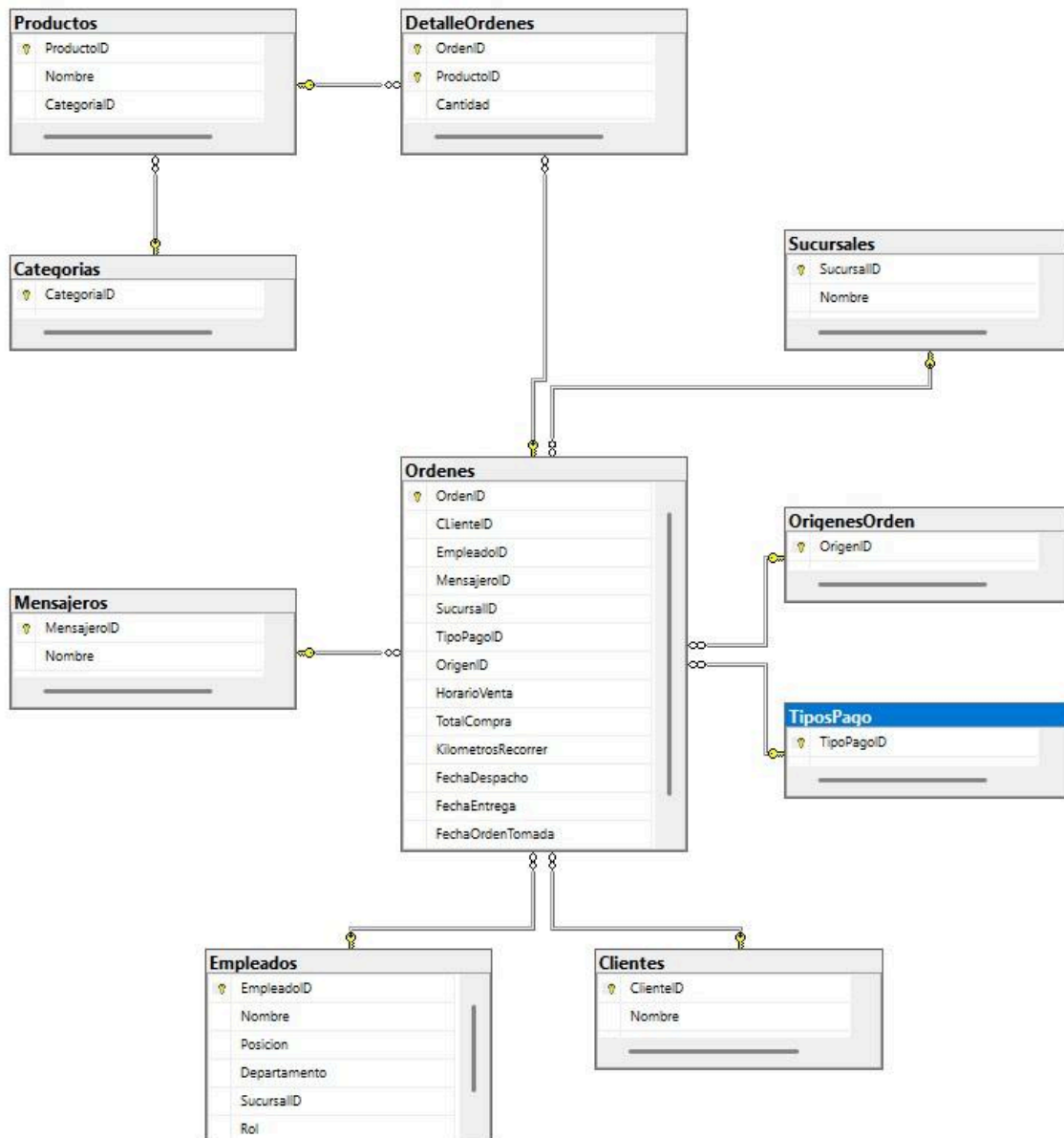
En la última etapa se desarrollaron consultas más complejas que implican la sentencia JOIN entre múltiples tablas:

- Listar productos con su categoría y detección de productos sin categoría asignada, asegurando calidad de datos. Donde se realiza un JOIN entre las tablas Productos y Categorías.
- Empleados por sucursal, para analizar la asignación de personal. A través de un JOIN entre las tablas Empleados y Sucursales.
- Detalle completo de órdenes, incluyendo los nombres de clientes, empleados y mensajeros, para obtener una visión amplia de cada transacción. Para esto se utiliza la cláusula JOIN repetidas veces para obtener el valor solicitado de las tablas Ordenes, Clientes, Empleados y Mensajeros.
- Productos vendidos por sucursal y categoría, un cruce que aporta insights sobre el mix de productos y su desempeño por ubicación. Nuevamente, mediante el uso JOIN, obtenemos los valores necesarios de las tablas Ordenes, DetalleOrdenes, Productos, Categorías y Sucursales.

Este avance consolidó el proyecto, mostrando la capacidad de la base para responder tanto a consultas estratégicas como a análisis operativos y de calidad de datos.

Módulo 2**Resultados y consultas**

Para una mejor comprensión de la base de datos, se presenta el esquema relacional:



Respecto al análisis realizado en los avances 3 y 4 encontramos que:

Módulo 2

1. Del **total de ventas globales**, se obtiene la facturación acumulada de \$9123.51.

Resulta importante usar esta métrica como base para fijar metas de crecimiento y compararla con períodos anteriores, en el caso correspondiente.

2. **Promedio de precios de productos por categoría**, muestra la distribución de precios de las distintas categorías o tipos de productos. Encontramos que: La categoría 1 tiene un precio promedio de \$9.99; la categoría 2 un promedio de \$3.49; la categoría 3 de \$6.49; la categoría 7 un precio promedio de 2.99 y la categoría 10 de 12.49.

Al identificar, por ejemplo, que las categorías 10, 1 y 7 representan pizzas, hamburguesas y helados respectivamente, se confirma que los precios promedio obtenidos resultan consistentes con el tipo de producto.

A partir de este análisis, se puede considerar si los precios impactan en el volumen de ventas para ajustar la estrategia de mercado.

3. Orden mínima y máxima por sucursal. En este caso arroja el mismo resultado para ambas funciones, ya que solo hay 1 valor cargado por cada sucursal.
4. **Mayor número de kilómetros recorridos para entrega**, encontramos que es de 12.5 Km.
5. El **promedio de cantidad de productos por orden** es de 3.
6. Del **total de ventas por tipo de pago**, encontramos que el tipo de pago 9, 8 y 2 son los que tienen ventas más altas, siendo de \$1095, \$1085 y \$1075 respectivamente. Sin embargo, no hay mayores diferencias entre los distintos tipos de pago, ya que el valor mínimo es de \$920 correspondiente al tipo de pago 3.

Módulo 2

7. Respecto a la **sucursal con la venta promedio más alta**, encontramos que la sucursal 9, Sucursal Lago, tiene el mayor promedio de ventas, correspondiente a \$1095.
8. Respecto a las sucursales con ventas mayores a \$1000 encontramos a las sucursales 9, 8, 2, 7, 1 ordenadas de forma descendente.

A partir de esta consulta, se podría pensar prácticas comerciales de la sucursal con mayor promedio de ventas y replicarlas en otras.

9. **Comparación de ventas antes y después del 1 de julio 2023**, encontramos que las ventas en el primer semestre de 2023 fue de \$979.75 y en el segundo semestre \$1081.66. Esto permite evaluar si hubo crecimiento o caída en la facturación promedio después de mitad de año, y en este caso encontramos un leve incremento.

Respecto a esto, podemos estudiar la posibilidad de reforzar campañas de marketing en períodos de menor rendimiento.

10. **Análisis de actividad de ventas por horario**, dentro de los 3 momentos u horarios (mañana, tarde y noche) encontramos que en todos se realizaron 3 ventas. El turno con promedio de compras más alto es el de la tarde con \$1038.33, le sigue el de la mañana con \$1016.17 y por último el horario de la noche con \$986.66
11. Se consultó a qué sucursal asignada se encuentra cada empleado. Todos los empleados registrados pertenecen a la sucursal 1, Sucursal Central.
12. Se buscó, además, si existen productos sin categoría asignada, lo que busca detectar problemas de calidad en los datos (productos mal categorizados), pero en este caso todos los productos pertenecen a una categoría.
13. **Productos vendidos por sucursal y categoría**, muestra qué productos (por categoría) se venden más en cada sucursal. Todos los registros de artículos vendidos pertenecen a la "Sucursal Central". Las categorías con más artículos vendidos son las de comida rápida y pizzas con 8 artículos cada

Módulo 2

una; postres, con 7 artículos; helados, con 5 artículos vendidos y 4 ensaladas.

Consultas finales

1. **Eficiencia de los mensajeros:** ¿Cuál es el tiempo promedio desde el despacho hasta la entrega de los pedidos gestionados por todo el equipo de mensajería?

```
SELECT AVG(DATEDIFF(MINUTE, FechaDespacho, FechaEntrega)) AS  
TiempoPromedioEntrega  
FROM Ordenes  
WHERE MensajeroID IS NOT NULL;
```

2. **Análisis de Ventas por Origen de Orden:** ¿Qué canal de ventas genera más ingresos?

```
SELECT TOP 1  
    ori.Descripcion AS CanalVenta,  
    SUM(o.TotalCompra) AS TotalIngresos  
FROM Ordenes AS o  
INNER JOIN OrigenesOrden AS ori  
    ON o.OrigenID = ori.OrigenID  
GROUP BY ori.Descripcion  
ORDER BY TotalIngresos DESC;
```

Módulo 2

3. **Productividad de los Empleados:** ¿Cuál es el nivel de ingreso generado por Empleado?

```
SELECT
    e.Nombre AS Empleado,
    SUM(o.TotalCompra) AS IngresosGenerados
FROM Ordenes AS o
INNER JOIN Empleados e
    ON o.EmpleadoID = e.EmpleadoID
GROUP BY e.Nombre
ORDER BY IngresosGenerados DESC;
```

4. **Análisis de Demanda por Horario y Día:** ¿Cómo varía la demanda de productos a lo largo del día?

NOTA: Esta consulta no puede ser implementada sin una definición clara del horario (mañana, tarde, noche) en la base de datos existente. Asumiremos que HorarioVenta refleja esta información correctamente.

```
SELECT o.HorarioVenta, SUM(do.Cantidad) AS ProductosVendidos
FROM Ordenes AS o
INNER JOIN DetalleOrdenes AS do
    ON o.OrdenID = do.OrdenID
GROUP BY o.HorarioVenta;
```

Módulo 2

5. **Comparación de Ventas Mensuales:** ¿Cuál es la tendencia de los ingresos generados en cada periodo mensual?

```
SELECT YEAR(FechaOrdenTomada) AS Anio,  
        DATENAME(MONTH, FechaOrdenTomada) AS Mes,  
        SUM(TotalCompra) AS IngresoTotal  
FROM Ordenes  
GROUP BY YEAR(FechaOrdenTomada), MONTH(FechaOrdenTomada),  
          DATENAME(MONTH, FechaOrdenTomada)  
ORDER BY MONTH(FechaOrdenTomada);
```

Módulo 2

Hallazgos clave

1. Encontramos que el promedio del tiempo de entrega de las órdenes desde su despacho es de 30 minutos.
2. Del análisis de las ventas por canal de ventas (presencial, drive thru, teléfono, en línea, app móvil), hallamos que el canal de ventas con mayor ingreso es el presencial, con un total de ingresos de \$2140.
3. Del análisis de la productividad de los empleados, encontramos que de los 9 empleados, Ethan Martinez tiene el total de ingresos más alto con \$1095; en segundo lugar encontramos a Olicia García con un ingreso total de \$1085 y en tercer puesto a Jane Smith con \$1075 de ingreso total.
4. En este caso, no podemos realizar la comparación de la demanda a lo largo del día, ya que tenemos solo datos del horario “mañana”, donde la venta fue de 32 productos.
5. Del análisis de tendencia de ingresos, concluimos que, si bien los ingresos de Enero y Febrero fueron buenos, en Marzo se registra una marcada baja siendo el mes con el menor ingreso \$920, sin embargo, a partir de ahí se muestra una tendencia estable con un leve crecimiento hacia el final del período 2023, alcanzando el valor máximo en septiembre (\$1095).

Para ampliar los resultados, se podrían implementar éstas mismas consultas sobre la base de datos actualizada en el Extra Credit.

Módulo 2

Desafíos y soluciones

Uno de los desafíos que encontré en el transcurso del proyecto integrador, fue que en el avance 3 pregunta 9, dónde al realizar consultas comparando fechas (por ejemplo, calcular el promedio de ventas antes del 1 de julio de 2023), el resultado fue NULL al utilizar un filtro como `< '2023-07-01'`, a pesar de que los datos existían en la base.

Para buscar una solución, recurrí a ChatGPT, para buscar explicación de por qué podría estar ocurriendo esto y encontrar ayuda para solucionarlo. Se descubrió que el problema podría ser un tema de formato regional, y para evitar problemas (que solo aparecían con ésta consulta) la IA me sugirió usar el formato ISO, quitando los guiones medio que separan el año, mes y día.

Reflexión personal

Realizar este proyecto integrador fue una gran oportunidad para poner en práctica todo lo aprendido a lo largo del módulo 2. Asumir el rol de *Data Analyst* para la empresa Fast Food me permitió experimentar de primera mano los desafíos de crear una base de datos relacional centralizada y de elaborar consultas en SQL para extraer información valiosa.

El diseño de tablas, la carga de datos y el uso de claves primarias y foráneas me ayudaron a comprender la importancia de la normalización y la integridad referencial. Además, al construir consultas con funciones de agregación como SUM, MAX, AVG y otras sentencias, pude generar métricas que respondieron a preguntas de negocio reales, dándome una visión más clara de cómo los datos impulsan la toma de decisiones.

En conclusión, este proyecto no solo reforzó mi comprensión teórica de SQL, sino que también me permitió aplicarla de forma práctica.

Módulo 2

Extra Credit

- En primer lugar, se realizó la carga de nuevos datos ficticios a las tablas “DetalleOrdenes”, “Empleados” y “Ordenes”.
- Posteriormente se realizaron nuevas consultas, obteniendo así nuevas insights y resultados.

Archivo:

<https://drive.google.com/file/d/1nrEdWBZCXylmecAhssFE6HN1IBtH72We/view?usp=sharing>