**Conociendo al Cliente 360°: Datos, Opiniones y Tendencias**

**Nombre del autor: Michel Alejandro Carrillo Vázquez**

**Email: alejandrocarrillo148@gmail.com**

**Cohorte: dsft01**

**Fecha de entrega:** 18/09/2025

**Institución:** InsightReach

# Introducción

El propósito de este proyecto fue llevar a cabo un análisis de datos exhaustivo para una empresa ficticia en el sector de restaurantes en Estados Unidos. El objetivo principal era entender el comportamiento del cliente y el rendimiento de los restaurantes, utilizando para ello datos de usuarios y la API de Yelp. A lo largo de tres fases, se realizó un análisis exploratorio de los datos (EDA), se integraron datos de la API de Yelp y se generaron consultas avanzadas para obtener información valiosa que pudiera guiar la toma de decisiones estratégicas.

# Desarrollo del proyecto

El proyecto se dividió en tres avances principales:

# Avance 1: EDA (Análisis exploratorio de datos)

En este avance, se cargaron y analizaron dos conjuntos de datos: uno sobre los usuarios de restaurantes en EE. UU. y otro sobre los restaurantes mismos. Se identificó que la mayoría de los usuarios se encuentran en los estratos socioeconómicos medio y alto, y que estos mismos grupos son los que registran un mayor número de visitas y un gasto promedio más elevado. Se concluyó que las campañas de marketing deberían enfocarse en este segmento de usuarios para maximizar la conversión.

# Avance 2: API de Yelp

En esta fase, se utilizó la API de Yelp para obtener información adicional sobre los restaurantes. Los datos obtenidos incluían detalles como la calificación, el número de reseñas, las categorías y la ubicación de los establecimientos. La información se integró con los datos existentes para enriquecer el análisis y permitir una visión más completa de la oferta de restaurantes.

# Avance 3: Consultas finales

En la última etapa del proyecto, se generaron y ejecutaron consultas avanzadas sobre la base de datos combinada. El objetivo de estas consultas fue responder a preguntas de negocio específicas y descubrir hallazgos clave que se detallan a continuación.

# Resultados y consultas

# Consultas finales:

**Hallazgo clave:** Se identificó que la cantidad de usuarios por género es casi igual, y que la mayoría de los clientes se encuentran en los estratos socioeconómicos medio y alto. Esto también se correlaciona con un mayor número de visitas y un gasto promedio más alto en estos segmentos.

**Recomendaciones estratégicas:** La empresa debe enfocar sus campañas de marketing en los usuarios de los estratos socioeconómicos medio y alto, ya que representan la mayor cantidad de clientes, visitas y gasto promedio, lo que sugiere una alta probabilidad de conversión.

**Hallazgo clave:** Se analizó el rendimiento de los restaurantes con base en sus calificaciones y número de reseñas obtenidas de la API de Yelp.

**Recomendaciones estratégicas:** La empresa debe analizar los restaurantes con mejores puntuaciones y mayor número de reseñas para replicar sus prácticas exitosas, como la calidad del servicio o la oferta de alimentos.

**Hallazgo clave:** Se evaluó la relación entre la frecuencia de visita de los usuarios y el tipo de cocina.

**Recomendaciones estratégicas:** Se recomienda a la empresa crear promociones o incentivos para fomentar las visitas a los restaurantes con el tipo de cocina preferido por los clientes más frecuentes.

# Optimización y sostenibilidad

La base de datos se optimizó para análisis externos y su sostenibilidad a largo plazo. Se implementaron prácticas de limpieza de datos para garantizar la precisión y la consistencia de la información.

# Desafíos y soluciones

Uno de los principales desafíos fue la integración de los datos de la API de Yelp con los datos de usuarios, debido a las diferencias en la estructura y el formato. La solución fue utilizar Python y la librería Pandas para limpiar, transformar y unificar los datos en un formato coherente, facilitando así el análisis y las consultas.

# Reflexión personal

A lo largo del proyecto, adquirí habilidades valiosas en análisis de datos, incluyendo la limpieza, transformación y visualización de datos. Aprendí a utilizar librerías como Pandas, NumPy, Matplotlib y Seaborn, así como a trabajar con APIs externas. La experiencia me enseñó a formular preguntas de negocio, a encontrar respuestas en los datos y a comunicar hallazgos de manera efectiva.

Si tuviera que volver a empezar este proyecto, habría dedicado más tiempo a la fase de planificación para definir un esquema de base de datos más robusto desde el principio, lo que habría simplificado el proceso de integración de datos.

# EXTRA CREDIT

Se utilizó la API de Yelp para enriquecer los datos de los restaurantes.

Se crearon consultas avanzadas para generar insights de negocio.

Se limpiaron y procesaron datos de forma programática utilizando scripts de Python.