

Universidad de San Carlos de Guatemala  
Facultad de Ingeniería  
Ingeniería en Ciencias y Sistemas  
Laboratorio Sistemas Organizacionales y Gerenciales 2  
Sección N+  
Ing. Mario José Bautista Fuentes  
Aux. Jimmy Daniel Larios Martínez

# **PRÁCTICA 1**

Guatemala, 31 agosto de 2024

## Objetivo:

Que el estudiante comprenda la importancia del análisis de datos dentro de una organización, reconociendo cómo los insights derivados del análisis pueden informar la toma de decisiones estratégicas, mejorar la eficiencia operativa y aumentar la ventaja competitiva en el mercado.

## Introducción:

Como analista de datos junior usted está trabajando para una tienda en línea de ropa y accesorios. La empresa quiere entender mejor sus ventas y el comportamiento de los clientes durante el último año.

Por lo cual se le ha proporcionado un archivo .csv con la información necesaria para realizar dicho análisis, el cual contiene los siguiente campos.

- ID de la orden
- Fecha de la compra
- ID del cliente
- Género del cliente
- Edad del cliente
- Categoría del producto (ropa, accesorios, calzado)
- Nombre del producto
- Precio del producto
- Cantidad comprada
- Total de la orden
- Método de pago
- Región de envío

## Tareas a realizar:

1. Preparación de datos:
  - Extraer los datos del archivo .csv.
  - Verificar si hay valores faltantes o duplicados y decidir cómo manejarlos.
  - Asegurarse de que los tipos de datos sean correctos para cada columna.
  - Cargar los datos a una base de datos SQL en la nube.
2. Análisis exploratorio:
  - Obtener los datos de la base de datos.
  - Calcular estadísticas básicas (media, mediana, moda) para las variables numéricas.
  - Crear visualizaciones para mostrar la distribución de ventas por categoría de producto y región.
3. Análisis de tendencias:
  - Determinar los meses con mayores y menores ventas.
  - Identificar los productos más vendidos y los menos populares.
4. Segmentación de clientes:
  - Agrupar a los clientes por edad y analizar sus patrones de compra.
  - Comparar el comportamiento de compra entre géneros.
5. Análisis de correlación:

- Investigar si existe una relación entre el total de la orden y la edad del cliente.
- Examinar si hay una correlación entre la categoría del producto y el método de pago preferido.
- 6. Visualización de datos:
  - Crear al menos siete gráficos diferentes (por ejemplo, gráfico de barras, gráfico de dispersión, gráfico de líneas) para representar los hallazgos más importantes.
- 7. Conclusiones y recomendaciones:
  - Basándose en el análisis realizado, proporcionar al menos cuatro conclusiones clave sobre las ventas y el comportamiento de los clientes.
  - Sugerir dos acciones concretas por estudiante que la empresa podría tomar para mejorar sus ventas o la satisfacción del cliente.
- 8. Responder a las preguntas:
  - ¿Cómo podrían los insights obtenidos ayudar a diferenciarse de la competencia?
  - ¿Qué decisiones estratégicas podrían tomarse basándose en este análisis para aumentar las ventas y la satisfacción del cliente?
  - ¿Cómo podría este análisis de datos ayudar a la empresa a ahorrar costos o mejorar la eficiencia operativa?
  - ¿Qué datos adicionales recomendarían recopilar para obtener insights aún más valiosos en el futuro?

## **Presentación:**

Se deberá realizar una presentación (Power Point, Canva, ...) en la cual se presentan los insights obtenidos así como las respuestas a las preguntas planteadas.

## **Documentación:**

La documentación debe contener:

- a) Planificación:
  - Cómo se dividieron las tareas entre los miembros del equipo.
  - Qué herramientas y tecnologías decidieron utilizar y por qué.
  - Cómo establecieron los plazos para cada fase del proyecto.
- b) Proceso de análisis:
  - Describa el enfoque paso a paso que siguieron para limpiar y preparar los datos.
  - Explique las decisiones tomadas durante el análisis exploratorio de datos.
  - Detalle los desafíos encontrados durante el análisis y cómo los superaron.
- c) Metodología:
  - Explique cómo seleccionaron las visualizaciones más apropiadas para sus hallazgos.
- d) Respuestas a las preguntas planteadas, se debe agregar una respuesta por estudiante a cada pregunta.

## Entregables

- Diagrama de la base de datos
- Código utilizado para la implementación
- Documentación
- Presentación

## Observaciones y restricciones

- La base de datos que se utilice puede ser cualquiera siempre y cuando sea relacional (MySQL, Oracle, Sql Server) y esté implementada en una plataforma en la nube (GCP, AWS, AZURE, ...)
- Se permite utilizar cualquier herramienta para realizar el análisis (Python, R, ...)
- Dudas o comentarios respecto a la práctica debe ser colocada en el foro de la semana correspondiente.
- La calificación se llevará a cabo por medio de Google Meet.
- Las copias completas/parciales serán merecedoras de una nota de 0 puntos y reportadas al catedrático de la sección y a la escuela de ciencias y sistemas.
- La práctica debe ser realizada en **parejas**.

**FECHA DE ENTREGA: 13 de septiembre hasta las 23:59:59 de 2024 en la plataforma UEDI.**