

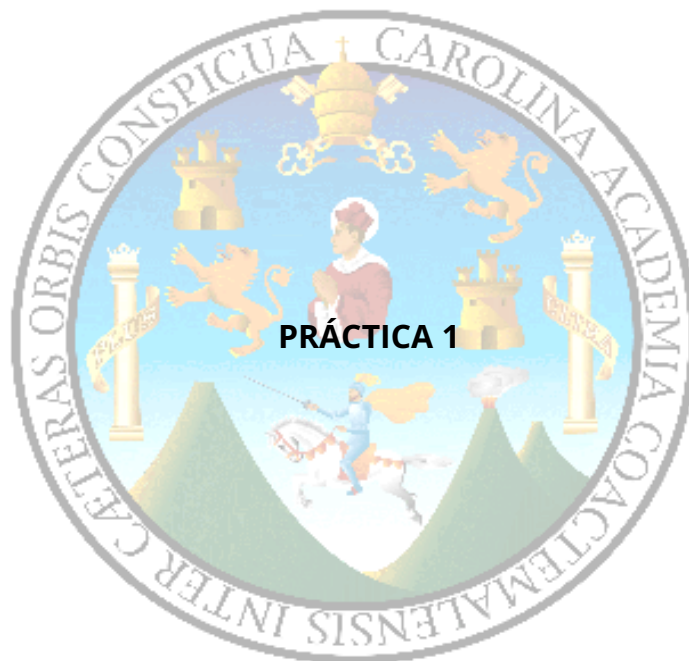
UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA

ESCUELA DE CIENCIAS Y SISTEMAS

SISTEMAS ORGANIZACIONALES Y GERENCIALES 2

SECCIÓN N°



202004725	Yeinny Melissa Catalán de León
202010223	Luis Angel Barrera Velásquez

Índice

División de Tareas.....	4
Herramientas y Tecnologías.....	4
Establecimiento de Fases del proyecto.....	4
Limpieza y preparación de Datos.....	4
Análisis Exploratorio de Datos.....	4
Desafíos Durante el Análisis.....	4
Selección de Visualizaciones.....	5

PLANIFICACIÓN

División de Tareas

Tarea	Luis Angel Barrera Velásquez	Yeinny Melissa Catalán de León
Preparación de Datos	✓	
Análisis Exploratorio	✓	
Análisis de Tendencias	✓	
Segmentación de Clientes		✓
Análisis de Correlación		✓
Visualización de Datos		✓
Conclusiones y Recomendaciones	✓	✓
Respuestas a las Preguntas	✓	✓

Herramientas y Tecnologías

Base de Datos

- **Base de Datos Utilizada:** MySQL
- **Plataforma en la Nube:** Google Cloud Platform (GCP)

Lenguaje de Programación

- **Lenguaje Principal:** Python

Entorno de Desarrollo

- **Entorno de Desarrollo:** Jupyter Notebook

Librerías y Paquetes Utilizados

- **Pandas:** Para la manipulación y análisis de datos.
- **Matplotlib:** Para la creación de gráficos y visualizaciones.
- **Seaborn:** Para la creación de gráficos estadísticos y visualizaciones más atractivas.
- **MySQL Connector/Python:** Para conectar y operar con la base de datos MySQL.

Herramientas de Visualización

- **Matplotlib:** Utilizado para gráficos básicos, como gráficos de barras y líneas.
- **Seaborn:** Utilizado para gráficos más avanzados, como gráficos de dispersión y distribuciones.

Justificación:

- **MySQL** fue elegido debido a su simplicidad y por su compatibilidad con python.
- **Google Cloud Platform (GCP)** se seleccionó por su escalabilidad e integración con otras herramientas, por lo cual se creó una máquina virtual con docker para almacenar la base de datos.
- **Python** es un lenguaje versátil y ampliamente utilizado en el análisis de datos por su variedad de librerías.
- **Jupyter Notebook** proporcionó un entorno interactivo y flexible para el desarrollo y documentación del análisis.
- **Pandas** facilitó el manejo del archivo csv.
- **Matplotlib** y **Seaborn** permitieron la creación de visualizaciones detalladas y comprensibles.

Establecimiento de Fases del proyecto

1. Preparación y Planificación (31 de agosto de 2024)

- **Objetivo:** Definir el alcance del proyecto, asignar roles y responsabilidades, y establecer un cronograma de trabajo.
- **Actividades:**
 - Revisión del archivo CSV y definición de tareas.
 - Planificación de la estructura de la base de datos y herramientas necesarias.
 - Establecimiento de plazos para cada fase del proyecto.

2. Preparación de Datos (1 - 3 de septiembre de 2024)

- **Objetivo:** Limpiar y preparar los datos para el análisis.
- **Actividades:**
 - Extracción de datos del archivo CSV.
 - Verificación y manejo de valores faltantes y duplicados.
 - Ajuste de tipos de datos y carga de datos en la base de datos MySQL en GCP.

3. Análisis Exploratorio (4 - 6 de septiembre de 2024)

- **Objetivo:** Realizar un análisis inicial de los datos para identificar patrones y obtener estadísticas básicas.
- **Actividades:**
 - Cálculo de estadísticas básicas (media, mediana, moda) para las variables numéricas.
 - Creación de visualizaciones para mostrar la distribución de ventas por categoría de producto y región.

4. Análisis de Tendencias (7 - 8 de septiembre de 2024)

- **Objetivo:** Identificar las tendencias en las ventas y el comportamiento de los clientes.
- **Actividades:**
 - Determinación de los meses con mayores y menores ventas.
 - Identificación de los productos más vendidos y los menos populares.

5. Segmentación de Clientes (9 - 10 de septiembre de 2024)

- **Objetivo:** Agrupar y analizar el comportamiento de los clientes según características demográficas.
- **Actividades:**
 - Agrupación de clientes por edad y análisis de patrones de compra.
 - Comparación del comportamiento de compra entre géneros.

6. Análisis de Correlación (11 de septiembre de 2024)

- **Objetivo:** Investigar relaciones entre diferentes variables en los datos.
- **Actividades:**
 - Investigación de la relación entre el total de la orden y la edad del cliente.
 - Análisis de la correlación entre la categoría del producto y el método de pago preferido.

7. Visualización de Datos (12 de septiembre de 2024)

- **Objetivo:** Crear visualizaciones para representar los hallazgos más importantes del análisis.
- **Actividades:**
 - Creación de gráficos diversos para ilustrar los resultados del análisis.

8. Conclusiones y Recomendaciones (13 de septiembre de 2024)

- **Objetivo:** Sintetizar los hallazgos del análisis y proporcionar recomendaciones para la empresa.
- **Actividades:**

- Redacción de conclusiones clave sobre las ventas y el comportamiento de los clientes.
- Sugerencia de acciones concretas para mejorar las ventas y la satisfacción del cliente.

PROCESO DE ANÁLISIS

Limpieza y preparación de Datos

Mejorar y Optimizar los procesos de Ecommerce de la empresa, desde la gestión de productos hasta las relaciones con proveedores y clientes

Análisis Exploratorio de Datos

Nuestra empresa está orientada hacia el crecimiento sostenido y la excelencia operativa, con un enfoq.

Desafíos Durante el Análisis

Nuestra empresa está orientada hacia el crecimiento sostenido y la excelencia operativa, con un enfoq.

METODOLOGÍA

Selección de Visualizaciones

Mejorar y Optimizar los procesos de Ecommerce de la empresa, desde la gestión de productos hasta las relaciones con proveedores y clientes

PREGUNTAS

1. ¿Cómo podrían los insights obtenidos ayudar a diferenciarse de la competencia?

2. ¿Qué decisiones estratégicas podrían tomarse basándose en este análisis para aumentar las ventas y la satisfacción del cliente?
3. ¿Cómo podría este análisis de datos ayudar a la empresa a ahorrar costos o mejorar la eficiencia operativa?
4. ¿Qué datos adicionales recomendarían recopilar para obtener insights aún más valiosos en el futuro?