

UNIVERSIDAD SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA

ESCUELA DE CIENCIAS Y SISTEMAS

LABORATORIO: SISTEMAS

ORGANIZACIONALES Y GERENCIALES 2

SECCIÓN: N

AUXILIAR: JOSÉ LUIS REYNOSO TIJ

GRUPO: 20



PRACTICA 1

Nombre

Kelly Mischel Herrera Espino
Alberto Gabriel Reyes Ning

Carné:

201900716
201612174

Indice

Documentación.....	2
a) Planificación.....	2
b) Proceso de análisis.....	3
c) Metodología.....	3
d) Respuestas a las preguntas.....	3
¿Cómo podrían los insights obtenidos ayudar a diferenciarse de la competencia?.....	4
¿Qué decisiones estratégicas podrían tomarse basándose en este análisis?.....	4
¿Cómo podría este análisis ayudar a ahorrar costos o mejorar la eficiencia operativa?.....	4
¿Qué datos adicionales recomendarían recopilar para obtener insights más valiosos?.....	4

Documentación

a) Planificación

¿Cómo se dividieron las tareas entre los miembros del equipo?

El equipo se conformó por: .

- **Alberto Gabriel Reyes Ninn** se enfocó en la parte técnica, incluyendo la implementación del código Python, conexión a la base de datos PostgreSQL, carga de datos y parte del desarrollo del script de análisis.
- **Kelly Mischel Herrera Espino** brindó orientación y apoyo en la parte de documentación, análisis de resultados y definición de fórmulas estadísticas, así como en la redacción de los insights interpretativos.

¿Qué herramientas y tecnologías decidieron utilizar y por qué?

- **Python** fue utilizado como lenguaje de programación por su amplia compatibilidad con análisis de datos.
- **pandas y matplotlib** fueron elegidos para la manipulación de datos y la generación de gráficos.
- **PostgreSQL en AWS RDS** como base de datos relacional en la nube debido a su robustez y escalabilidad.
- **SQLAlchemy y psycopg2** para manejar las conexiones entre Python y PostgreSQL.
- **DBeaver** para verificar el estado de las bases de datos y tablas visualmente.
- **Markdown** fue utilizado para documentar el análisis en un archivo .

¿Cómo establecieron los plazos para cada fase del proyecto?

El proyecto se dividió en tres fases principales:

1. **Preparación del entorno y limpieza de datos** – 2 días
2. **Análisis y visualización** – 3 días

3. Documentación y revisión final – 2 días

Cada fase fue asignada y revisada colaborativamente al finalizar.

b) Proceso de análisis

¿Qué enfoque paso a paso siguieron para limpiar y preparar los datos?

1. Se cargó el archivo CSV que se brindó previamente para el post análisis.
2. Se realizó la eliminación de duplicados y valores nulos.
3. Se realizó una transformación de las columnas de fecha al formato adecuado.
5. Se dividieron los datos en tablas de dimensiones y hechos bajo un esquema estrella.
6. Se insertaron los datos limpios en PostgreSQL.

¿Qué decisiones se tomaron durante el análisis exploratorio?

Se centró en variables clave como edad, categoría del producto, región de envío, y total de orden. Se calcularon medidas estadísticas básicas como media, mediana, moda y se priorizó la exploración visual de las ventas y patrones de compra.

¿Qué desafíos encontraron y cómo los resolvieron?

Uno de los desafíos que surgieron fue manejar errores de codificación (años vs anio) y evitar duplicados al insertar en la base de datos. Estos se resolvieron cambiando a psycopg2 y aplicando filtrado previo a la carga con drop_duplicates. También se detectaron problemas relacionados con las claves primarias duplicadas al no reiniciar las tablas, lo que se soluciona mediante un script que elimina y recrea las tablas correctamente.

c) Metodología

¿Cómo seleccionaron las visualizaciones más apropiadas?

Se eligieron gráficos de barras para comparar totales por categoría o región, gráfica de líneas para tendencias mensuales, dispersión para relaciones entre edad y total de orden, boxplots para comparar géneros, y un heatmap para correlaciones.

La decisión se basó en la claridad de interpretación y la capacidad de cada tipo de gráfico para representar correctamente la relación entre variables.

d) Respuestas a las preguntas

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

- **Alberto Gabriel Reyes Ning:**
Los insights permiten tomar decisiones basadas en datos reales del comportamiento del cliente, como identificar qué región genera más ingresos o qué productos se venden mejor. Esto ayuda a diseñar estrategias enfocadas que la competencia podría no tener.
- **Kelly Mischel Herrera Espino:**
Los insights permiten un refuerzo en la personalización, es decir que se pueden lanzar campañas específicas, aumentando la fidelización de los clientes tomando en cuenta que segmento.

¿Qué decisiones estratégicas podrían tomarse basándose en este análisis?

- **Alberto Gabriel Reyes Ning:**
Un refuerzo en el tema del inventario en productos más vendidos como suéteres y realizar un aumento en el marketing en la región.
- **Kelly Mischel Herrera Espino:**
Aplicar promociones en las categoría de ropa, adaptando el contenido de marketing por género y así tener un enfoque personalizado para estos grupos

¿Cómo podría este análisis ayudar a ahorrar costos o mejorar la eficiencia operativa?

- **Alberto Gabriel Reyes Ning:**
Al obtener las regiones que registraron menos ventas se puede realizar un ajuste en las rutas de envío y reducir costos de logística innecesarios.
- **Kelly Mischel Herrera Espino:**
Se puede evitar un sobre stock en productos que tienen una baja rotación, optimizando así los recursos de bodeg

¿Qué datos adicionales recomendarían recopilar para obtener insights más valiosos?

- **Alberto Gabriel Reyes Ning:**
Un historial de navegación para identificar intereses no convertidos.
- **Kelly Mischel Herrera Espino:**
Datos sobre devoluciones y también reseñas de productos, esto para poder entender la satisfacción del cliente.