

## Actividad: Estudio de Caso

### Extracción y Manipulación de Datos en un Entorno Laboral

#### Objetivo de la Actividad:

El objetivo de esta actividad es que los participantes comprendan la importancia de la obtención y manipulación de datos en el entorno laboral. A través de un estudio de caso realista, aplicarán técnicas de extracción de datos desde diferentes fuentes utilizando Python y sus librerías especializadas. Esto les permitirá optimizar procesos, mejorar la toma de decisiones y automatizar tareas en su trabajo diario.

#### Instrucciones:

1. Lee atentamente el estudio de caso presentado.
2. Analiza el problema y determina las estrategias adecuadas para la obtención y manipulación de los datos.
3. Responde las preguntas clave utilizando los conocimientos adquiridos en el módulo.
4. Entrega un informe detallado con tus respuestas y justificaciones.

#### Estudio de Caso:

María es analista de datos en una empresa de retail que vende productos en línea. Su equipo ha recibido la tarea de analizar las ventas del último trimestre para identificar tendencias y optimizar el inventario. Los datos se encuentran en distintos formatos:

- Un archivo CSV con el registro de ventas detallado.
- Un archivo Excel con reportes financieros generados por el área contable.
- Datos sobre precios de productos extraídos de una página web de la competencia.

#### Desafío:

María debe consolidar toda esta información en un solo informe que permita visualizar el comportamiento de las ventas y compararlas con los precios de la competencia. Para ello, debe realizar las siguientes tareas:

1. Cargar los datos desde el archivo CSV y analizar el total de ventas por categoría de producto.
2. Leer el archivo Excel y calcular el margen de ganancia promedio de cada categoría.
3. Extraer los precios de productos desde la web de la competencia utilizando herramientas de scraping.
4. Comparar los precios de la competencia con los datos de ventas y margen de ganancia para determinar si es necesario ajustar los precios de la empresa.
5. Generar un archivo Excel con los resultados obtenidos para compartirlo con el equipo de marketing y ventas.

#### Preguntas Clave:

1. ¿Qué librerías de Python utilizarías para realizar la extracción y manipulación de datos en este caso? Justifica tu respuesta.
2. ¿Cómo podrías leer los archivos CSV y Excel en Python? Explica el procedimiento con ejemplos de código.
3. ¿Qué ventajas ofrece la extracción de datos desde la web en comparación con otras fuentes? ¿Qué desafíos podrías encontrar al hacerlo?

4. ¿Cómo podrías automatizar el proceso para que María no tenga que repetir manualmente estos pasos cada trimestre?
5. ¿Cómo garantizarías la calidad y precisión de los datos extraídos antes de analizarlos?

## Rúbrica de Evaluación

Criterio	4 - Excelente	3 - Bueno	2 - Satisfactorio	1 - Insuficiente
<b>Identificación de herramientas</b>	Menciona todas las herramientas adecuadas y justifica su elección con ejemplos claros.	Menciona la mayoría de las herramientas necesarias con justificación parcial.	Menciona algunas herramientas, pero sin justificación suficiente.	No menciona herramientas relevantes o su justificación es incorrecta.
<b>Análisis del problema</b>	Explica detalladamente cómo resolver cada tarea con ejemplos de código.	Explica cómo resolver la mayoría de las tareas con ejemplos adecuados.	Explica soluciones de manera general sin ejemplos específicos.	No proporciona soluciones claras o no responde correctamente.
<b>Automatización y optimización</b>	Propone una estrategia eficiente para automatizar el proceso y mejorar el flujo de trabajo.	Propone una estrategia de automatización básica sin optimización avanzada.	Menciona la automatización, pero sin una implementación clara.	No menciona formas de automatización.
<b>Claridad y estructura del informe</b>	Presenta respuestas bien organizadas, con lenguaje claro y argumentos sólidos.	Presenta respuestas organizadas, pero con algunas fallas en la claridad.	Presenta respuestas con falta de estructura y poca claridad.	Presenta respuestas desordenadas y difíciles de entender.
<b>Evaluación de calidad de datos</b>	Identifica correctamente los criterios para validar la calidad de los datos y propone soluciones.	Identifica algunos criterios de calidad, pero con soluciones parciales.	Menciona criterios de calidad sin estrategias claras.	No menciona cómo evaluar la calidad de los datos.

**Total de Puntos:** 20 puntos

Este estudio de caso permitirá a los participantes aplicar conceptos clave en un contexto realista, fomentando el pensamiento analítico y la resolución de problemas mediante el uso de Python para la manipulación de datos.