

# Proyecto Prueba:

El Proyecto prueba que sugerimos para la Competencia de la WorldSkill 2025 es sobre un **Sistema de Ventas** una solución informática para la empresa GSW General Supplies Warehouse es una empresa creada como proyecto prueba. Esta empresa se dedica a abastecer pequeñas empresas y consumidores de alimentos comestibles, pero para sus actividades se desarrollará un sistema de ventas como solución informática que manejará de una forma más eficiente sus actividades. En el proceso de realizar esta solución informática se deberá cumplir con unas series de módulos que también forman parte de todo lo sugerido para la competencia. Estos módulos deben ser elaborados en los tiempos y días establecidos.

Son siete módulos y en cada uno se realizarán sus actividades.

## Modulo Primero:

Este paquete de sesión contiene los siguientes materiales:

1. Descripción del proyecto y las tareas
2. Instrucciones para los concursantes
3. Limpieza y transformación de datos
4. Análisis de las tendencias de venta
5. Análisis del rendimiento de los productos
6. Calculo del ciclo de vida del cliente

# Descripción del proyecto y las tareas

En esta sesión, se analizará los datos de GSW General Supplies Warehouse para obtener datos limpios para ser utilizados en las tomas de decisiones.

## Directrices

1. **Facilidad de uso:** Presente datos en un formato claro y comprensible.
2. **Buen aspecto:** para todas las visualizaciones e informes.
3. **Buen funcionamiento:** Asegúrese de que todos los análisis y cálculos son precisos y no contienen errores.
4. **Seguridad:** Trate los datos de los clientes de forma confidencial y cumpla la normativa sobre privacidad de datos.
5. **Puntualidad:** Complete todas las tareas dentro del plazo especificado.

## Consideraciones técnicas

1. **Limpieza de datos:** Abordar los valores que faltan, las incoherencias y los problemas de formato en los conjuntos de datos proporcionados.
2. **Análisis de datos:** Aplicar técnicas estadísticas adecuadas para el análisis de los datos.
3. **Visualización de datos:** Crear gráficos y tablas claros e informativos para presentar los resultados.

## Consideraciones adicionales

- El análisis debe ser reproducible y estar bien documentado.
- Utilice etiquetas y explicaciones claras para todas las visualizaciones y tablas.
- Organizar la información de forma lógica para facilitar su comprensión por las partes interesadas.

# Instrucciones para los concursantes

## 1. Carga y consulta de datos

### Objetivo

Demuestre su capacidad para cargar, inspeccionar y comprender los conjuntos de datos proporcionados, identificando posibles problemas de calidad de los datos y preparándolos para su posterior análisis.

### Tareas

- 1.1. Cree una base de datos llamada Modulo1\_DB.
- 1.2. En la base de datos Modulo1\_DB, importe los archivos base1 y base3 con extensión xlsx a SQL Server.
- 1.3. Muestre de la columna CódigoCiudad del archivo base1 los que comienzan con S y terminan con M.
- 1.4. Determine los tipos de datos de cada columna.
- 1.5. Determina los registros con fecha no válidas. Cuantos son y cuales son.
- 1.6. Mostrar los registros cuyas fechas estén entre 2023 y 2024.
- 1.7. Determine el número de filas con **DocumentoCliente** negativos.
- 1.8. Numero de filas con atributos vacíos o nulos.
- 1.9. Realice un **dashboard** filtrado por: cantidad vendida por producto y por mes. Para el Dashboard utilice el archivo **dashboard.xlsx**. Realice una Tabla Dinámica para este filtro. Luego convierta estos archivos en PDF para su entrega.

### Entregables:

- I. Archivo .sql del SSMS con las sentencias SQL que resuelven las tareas del 1 al 8
- II. Archivo PDF y xlsx con el DashBoard solicitado en la tarea 9.

## 2. Limpieza y transformación de datos

### Objetivo

Utilizando el archivo **base3.xlsx** realice las tareas de transformación que se solicitan a continuación:

### *Problemas de calidad de datos y soluciones*

Este conjunto de datos presenta varios problemas comunes que requieren limpieza y transformación para ser analizados de manera efectiva.

- 2.1 Formato de fechas inconsistente: Las fechas en las columnas Fecha Nacimiento y Fecha Venta tienen formatos diferentes (ej. DD/MM/YYYY y lunes, 15 de septiembre de 2025). Para estandarizar, se debe convertir todo a un formato único, como YYYY-MM-DD.
- 2.2 Variabilidad en la notación de precios: La columna Precio Venta tiene diferentes monedas y símbolos (, \$, EUR, usd) y algunos valores no tienen símbolo. La solución es estandarizar la moneda (por ejemplo, a USD) y convertir todos los valores al mismo formato de moneda.
- 2.3 Inconsistencia en los nombres de ciudades: La ciudad puerto Plata aparece con la primera letra en minúscula. Es fundamental normalizar los nombres de las ciudades para evitar errores de conteo o clasificación.
- 2.4 La fila que tiene un valor null en la columna. Se debe decidir cómo manejar estos datos faltantes: se pueden imputar con más común o simplemente dejar el valor como 'Desconocido' si no se puede inferir.

### *Transformación de datos*

- 2.5 División de columnas: La columna Nombre Completo contiene el nombre y el apellido en una sola celda. Para un análisis más granular, como agrupar por apellido o buscar por nombre, esta columna debe dividirse en dos nuevas columnas: Nombre y Apellido.
- 2.6 Creación de nuevas variables: Se puede crear una columna Edad a partir de la columna Fecha Nacimiento y la fecha actual. Esto facilitaría el análisis demográfico del cliente sin tener que calcular la edad cada vez.

### **Entregables:**

- I. Archivo de Excel con el conjunto de datos transformados.

### 3. Análisis de las tendencias de datos

#### Objetivo

Calcule y visualice las tendencias de ventas de GSW general supplies warehouse a lo largo del tiempo. Para esto utilice el archivo de Excel que ha transformado en el punto anterior (2).

#### Tareas

3.1 Calcule los ingresos totales por ventas, el número de transacciones y el valor promedio de los pedidos al mes.

3.2 Cree gráficos de líneas para cada una de las tres métricas a lo largo del tiempo (mensualmente).

3.3 Cree una tabla que muestre los 3 ~~productos~~ más ~~ventas~~ por cantidad vendida, incluyendo sus nombres, cantidad total y ~~ingresos~~ totales.

3.4 Gráfico de barras: ~~Ingresos totales por categoría de productos.~~

**Entregables:**

I. Archivo de Excel con el dashboard creado.