

INFORME DEL ANALISIS

1. Introducción

El presente proyecto tiene como objetivo realizar un análisis exhaustivo de datos de transacciones de comercio electrónico correspondientes al periodo 2009–2010.

Para ello se han utilizado dos conjuntos de datos en bruto procedentes de distintas fuentes, los cuales han sido integrados, transformados y analizados mediante herramientas de análisis de datos en Python y visualización en Power BI.

El propósito principal del análisis es:

- Evaluar el rendimiento comercial del negocio.
- Analizar la evolución temporal de las ventas.
- Identificar patrones por región y país.
- Examinar el impacto de los descuentos en la facturación.
- Detectar los productos con mayor contribución al ingreso.

2. Descripción de los conjuntos de datos

Se utilizaron dos datasets en bruto:

1. Online Retail Transactions (Excel)

Contiene información detallada de transacciones:

- Invoice
- StockCode
- Description
- Quantity
- InvoiceDate
- Price
- Customer ID
- Country

2. Sales Data (CSV)

Contiene información adicional relacionada con descuentos y métricas comerciales.

Ambos datasets fueron integrados para generar un conjunto final unificado.

3. Limpieza y transformación de datos

La fase de preparación de datos se realizó en Python utilizando Pandas.

◆ **Principales acciones realizadas:**

- Eliminación de valores nulos.
- Corrección de tipos de datos.
- Eliminación de registros inconsistentes.
- Verificación de duplicados.
- Validación de coherencia matemática.

◆ **Creación de nuevas variables:**

- `Total_Amount = Quantity × Price`
(Ventas brutas por transacción)
- `Discounted_Revenue`
(Ventas netas tras aplicar descuentos)
- Variables temporales:
 - `Year`
 - `Month`
 - `YearMonth`
 - `DayOfWeek`
- `Variable Region`
(Agrupación geográfica de países)

El dataset final resultante contiene:

- **400.916 registros**
- Más de 20 variables analíticas

Cumpliendo los requisitos mínimos establecidos.

4. Análisis descriptivo

Se realizó un análisis estadístico básico del dataset final:

- Número total de transacciones: 400.916
- Media de ventas por transacción: 21,95 €
- Desviación estándar: 77,76
- Venta máxima registrada: 15.818,40
- El ticket promedio es bajo, lo que indica un modelo basado en volumen.
- Existe dispersión significativa en los importes de venta, con presencia de valores extremos.
- La mayoría de transacciones se concentran en rangos bajos de importe.

5. Análisis temporal

El análisis mensual muestra:

- Tendencia creciente durante el año.
- Fuerte incremento en el último trimestre (Q4).
- Pico máximo de ventas en noviembre.
- Ligera caída en diciembre.

Se observa un comportamiento estacional claro, probablemente asociado a campañas promocionales y periodo navideño.

6. Análisis por región

El análisis geográfico revela:

- La región **West** concentra aproximadamente el 93% del volumen total de ventas.
- Las regiones North, South y East presentan participación significativamente menor.

Existe una fuerte concentración geográfica del negocio, lo que implica:

- Alta dependencia de una región dominante.
- Oportunidad estratégica de expansión en regiones menos desarrolladas.

7. Ventas brutas vs ventas netas

Se realizó una comparativa entre:

- Ventas Totales (brutas)
- Ventas Netas (tras descuento)

Resultados:

- Ventas Totales: 8,80 millones €
- Ventas Netas: 7,49 millones €
- Impacto medio del descuento: ~15%

Los descuentos tienen un impacto relevante en la facturación final, reduciendo aproximadamente un 15% del ingreso bruto.

8. Análisis de productos

El análisis de los 10 productos con mayor volumen de ingresos muestra:

- Concentración significativa en un conjunto reducido de referencias.
- Existencia de productos “estrella” que impulsan el negocio.

El negocio depende parcialmente de productos clave, lo que puede representar riesgo si la demanda de estos disminuye.

9. Conclusiones generales

El análisis realizado permite concluir que:

- El negocio presenta un modelo basado en alto volumen y ticket medio reducido.
- Existe una marcada estacionalidad en el último trimestre del año.
- La región West domina ampliamente el mercado.
- Los descuentos impactan significativamente en la facturación final.
- La cartera de productos muestra concentración en referencias clave.

10. Valor del dashboard operativo

El dashboard desarrollado en Power BI permite:

- Monitorizar KPIs clave en tiempo real.
- Analizar evolución temporal.
- Comparar ventas brutas vs netas.
- Filtrar por año, región y país.
- Identificar productos de mayor rendimiento.

Esto facilita la toma de decisiones estratégicas basadas en datos.

Conclusión Final

El proyecto demuestra competencias en:

- Limpieza y transformación avanzada de datos.
- Análisis descriptivo y estadístico.
- Modelado y creación de métricas de negocio.
- Desarrollo de dashboard interactivo.
- Interpretación estratégica de resultados.