

Guía de resolución y retroalimentación – Actividad 2

Programación de visualizaciones automatizadas con análisis exploratorio

Asignatura: Visualización Avanzada y Automatización del Análisis de Datos
Dataset base: AdventureWorks Sales.xlsx
Trabajo en grupo

¿Qué pedía exactamente la Actividad 2?

- **Objetivos de la actividad:**
 - Diseñar un flujo de trabajo reproducible en Business Intelligence.
 - Implementar visualizaciones automatizadas (Power BI, Tableau, R o Python).
 - Aplicar análisis exploratorio avanzado (clustering, segmentación, PCA...).
 - Documentar el uso de IA y reflexionar críticamente.
 - Elaborar un vídeo explicativo del proyecto.
- **Entregables obligatorios:**
 - Documento principal (Word/PDF, máximo 8 páginas).
 - Archivo técnico de visualización: .pbix, .twb o scripts R/Python.
 - Vídeo MP4 (5–8 minutos).
 - Evidencias de uso de IA (anexo o enlace).
 - Hoja de control grupal cumplimentada.
- **Rúbrica (resumen):**
 - C1: Objetivos (2 puntos, 20%).
 - C2: Datos y modelización (2,5 puntos, 25%).
 - C3: Visualizaciones automatizadas (3 puntos, 30%).
 - C4: Vídeo (1,5 puntos, 15%).
 - C5: Reflexión sobre IA (1 punto, 10%).

Criterio 1 – Enfoque y objetivos (2 puntos)

- **Nivel excelente:**
 - Problema empresarial de Business Intelligence claramente definido.
 - 1 objetivo general bien formulado.
 - 3–5 objetivos específicos relevantes, coherentes y factibles con AdventureWorks.
- **Ejemplo de buena práctica: segmentación de clientes**
 - Tema: Segmentación de clientes de AdventureWorks en función de su valor y comportamiento de compra.
 - Objetivo general: Identificar segmentos de clientes con comportamientos de compra diferenciados para apoyar decisiones de marketing y fidelización.
 - Objetivos específicos:
 - Calcular el valor y frecuencia de compra de cada cliente en el periodo analizado.
 - Aplicar una técnica de clustering para agrupar clientes según su comportamiento (RFM u otra combinación de variables).
 - Caracterizar los segmentos obtenidos (ticket medio, margen, región, canal).
 - Diseñar visualizaciones dinámicas que permitan explorar los segmentos y su evolución.
 - Automatizar la actualización de datos y segmentos ante nuevos registros de ventas.

Criterio 2 (1) – Preparación de datos y calidad

- Uso de AdventureWorks Sales.xlsx como mínimo (opcionalmente enriquecido).
- Tareas de una entrega excelente:
 - Limpieza: tratamiento de nulos, outliers y errores de carga justificados.
 - Transformación: agregaciones por cliente/periodo y creación de tablas intermedias.
 - Creación de nuevas variables: VentasTotalesCliente, NumeroPedidos, Recencia, Margen estimado, etc.
- Ejemplo de redacción en el documento:
 - “A partir de la tabla de ventas se construyó una tabla de clientes agregados con las variables VentasTotalesCliente, NumeroPedidos, FechaUltimaCompra y MargenEstimado.”
 - “Se detectaron clientes sin región definida, agrupados como 'Desconocido'. Se excluyeron registros con fecha nula o importes negativos por considerarse errores.”
 - “Las variables para el clustering se normalizaron para evitar que la escala de las ventas dominara el análisis.”

Criterio 2 (2) – Modelización y análisis exploratorio avanzado

- **Se debe aplicar al menos una técnica de:**
 - Clustering (p. ej. k-means sobre variables de cliente o producto).
 - Segmentación (reglas de negocio, árbol de decisión, RFM...).
 - Reducción de dimensionalidad (PCA) para facilitar la visualización.
- **Documentación necesaria:**
 - Variables utilizadas y justificación de su selección.
 - Parámetros del modelo (nº de clusters, criterios) y pruebas realizadas.
 - Interpretación de cada grupo en términos de negocio.
- **Ejemplo orientativo (segmentación de clientes):**
 - Variables: VentasTotalesCliente, NumeroPedidos, RecenciaDias.
 - Técnica: k-means con $k = 4$ clusters (probando también $k = 3$ y $k = 5$).
 - Resultados: segmentos de alto valor, valor medio, clientes esporádicos y clientes reactivados.

Ejemplo de flujo de trabajo reproducible

- **Flujo de trabajo ideal (Power BI + R/Python):**
 - Importación de tablas de AdventureWorks (Sales, Customers, Products, Territory, Date) y creación del modelo de datos.
 - En Power Query: tabla agregada de clientes con variables para el modelo (RFM u otras).
 - Ejecución de script R/Python que realiza el clustering/segmentación y devuelve el segmento de cada cliente.
 - Integración del resultado (segmento) en el modelo de Power BI para usarlo en visualizaciones.
 - Guardado del script con comentarios para poder reutilizarlo (flujo reproducible).

Criterio 3 – Visualizaciones dinámicas y automatizadas (3 puntos)

- **Características de las visualizaciones en nivel excelente:**
 - Claridad: gráficos legibles, sin saturación de información.
 - Coherencia con los objetivos: cada gráfico responde a una pregunta de negocio.
 - Interactividad real: filtros, segmentadores, navegación por segmentos, drill-down.
 - Automatización: el dashboard se actualiza con nuevos datos sin rehacer el trabajo manual.
- **Elementos recomendables en un dashboard de segmentación:**
 - KPIs por segmento: N° de clientes, ventas totales, margen, ticket medio.
 - Gráfico de dispersión o burbujas: variables del clustering en los ejes, color por segmento y tamaño por n° de clientes o margen.
 - Tabla o matriz: distribución de segmentos por país/región, canal y categoría de producto.
 - Gráfico de evolución temporal: ventas por segmento a lo largo del tiempo.

Ejemplo de automatización de las visualizaciones

- En Power BI / Tableau:
 - Fuente de datos centralizada (AdventureWorks en OneDrive, SharePoint o base de datos).
 - Configuración de actualización programada en el servicio (Power BI Service o Tableau Cloud/Server).
 - El modelo de datos y las medidas se recalculan automáticamente con cada actualización.
 - Scripts de R/Python (si se usan) se ejecutan de nuevo para reasignar segmentos.
 - El dashboard se mantiene siempre actualizado para el usuario final.
- En R / Python puros:
 - Script que lee AdventureWorks, limpia, transforma, segmenta clientes y genera gráficos o informes.
 - Exportación de resultados a informes HTML, Markdown o notebooks reutilizables.
 - Programación de la ejecución (cron, tareas programadas, etc.).

Criterio 4 – El vídeo explicativo (1,5 puntos)

- **Requisitos del vídeo:**
 - Duración entre 5 y 8 minutos.
 - Explicación clara y estructurada del proyecto.
- **Estructura orientativa:**
 - Min 0–1,5: Presentación del problema y de los objetivos.
 - Min 1,5–3,5: Explicación de los datos, limpieza y variables creadas.
 - Min 3,5–6: Recorrido por el dashboard y principales insights.
 - Min 6–8: Comentario sobre el rol de la IA, limitaciones y aprendizajes.
- **Claves de calidad:**
 - Lenguaje comprensible, evitando tecnicismos innecesarios.
 - Uso de una narración lógica y no un simple 'pase de pantallas' del dashboard.

Criterio 5 – Reflexión crítica sobre el uso de IA (1 punto)

- **Contenido esperado en la reflexión (máx. 200 palabras):**
 - Utilidad concreta de la IA en el proyecto (qué aportó realmente).
 - Limitaciones encontradas (respuestas genéricas, desconocimiento del dataset real, propuestas poco adecuadas).
 - Influencia real en el resultado final (qué ideas se adoptaron, cuáles se descartaron y por qué).
- **Esquema de respuesta excelente:**
 - Aportes: cómo ayudó la IA a elegir variables, técnicas o tipos de visualización.
 - Limitaciones: por qué sus sugerencias no siempre eran aplicables a AdventureWorks.
 - Impacto: cómo se combinaron las propuestas de la IA con el EDA propio y la teoría de la asignatura.
- **Lo que se evalúa no es 'hablar bien' de la IA, sino mostrar pensamiento crítico sobre su uso.**

La entrega “de 10 puntos”

- **Checklist de una entrega excelente (Documento):**
 - Tema y problema empresarial bien definidos.
 - Objetivo general + 3–5 objetivos específicos claros.
 - Proceso de datos: limpieza, transformaciones, creación de variables, calidad y limitaciones.
 - Modelización avanzada (clustering, segmentación, PCA) explicada paso a paso e interpretada.
 - Descripción del flujo automatizado de visualización.
 - Reflexión crítica sobre IA estructurada y concreta.
- **Checklist de una entrega excelente (Archivo técnico y vídeo):**
 - Archivo .pbix / .twb / scripts con flujo reproducible y visualizaciones alineadas con los objetivos.
 - Dashboard dinámico, con filtros útiles y buena jerarquía visual.
 - Vídeo de 5–8 minutos con relato claro: problema, datos, modelo, dashboard e IA.
- **Anexos:**
 - Evidencias de IA (capturas, enlaces).
 - Hoja de control grupal completa y coherente.

Hoja de control y trabajo en equipo

- **Hoja de control grupal:**
 - Cada estudiante valora la contribución de los demás (sí/no).
 - Un 'no' puede implicar una penalización de hasta 3 puntos personales.
 - Si el grupo marca 'no' a todos los miembros, la actividad se suspende.
 - Se entrega una única hoja de control por grupo.
- **Trabajo en equipo y calificación individual:**
 - La hoja de control permite diferenciar la nota cuando no todos han participado por igual.
 - Es una herramienta de corresponsabilidad, no un mero trámite administrativo.
 - Un buen resultado del grupo requiere coordinación, reparto de tareas y comunicación continua.

**Muchas gracias por
vuestra atención**

unir
LA UNIVERSIDAD
EN INTERNET

unir LA UNIVERSIDAD
EN INTERNET

www.unir.net