

Análisis Estratégico y Predictivo: Caso AdventureWorks

1. Introducción y Marco Metodológico: El objetivo del presente informe es desarrollar una estrategia de negocio basada en datos (*Data-Driven Strategy*). Se ha seguido metodología **OSEMN**, finalizando con una propuesta estratégica y predictiva para la dirección.

2. O – Obtain & Scrub (Obtención y Depuración): el dataset incluye variables transaccionales y sociodemográficas para la segmentación de la cartera.

Se ejecutó una de limpieza para asegurar la integridad del análisis (*Data Quality*):

- **Integridad de Entidad:** Verificación cruzada de identificadores únicos (**CustomerID** y **PersonID**) para confirmar la unicidad de registros y evitar duplicidades.
- **Limpieza de Texto:** Eliminación de espacios residuales en variables categóricas para asegurar la consistencia en los filtros y agrupaciones.
- **Normalización Geográfica:** Corrección de la fragmentación en "United States" (dividido por regiones), agrupándolo bajo una etiqueta nacional única.
- **Consolidación de Categorías:** Simplificación de la variable *Occupation*, unificando "Skilled Manual" y "Manual" bajo la etiqueta única "**Manual**", optimizando así la segmentación.

Nota: Se utilizó IA como herramienta de consulta técnica para validar funciones de limpieza y lógica de agrupación.

3. E – Explore (Exploración Visual): Se realizó un Análisis Exploratorio de Datos (EDA) para detectar patrones. Se generaron visualizaciones clave, como gráficos para la distribución geográfica y diagramas de sectores para el perfil laboral y de estudios.

El análisis visual revela en las Ventas por País una fuerte concentración en Norte América; el desglose por Ocupación sitúa al comprador ideal en una clase media activa (Manual y Professional); las familias con 2 hijos o menos concentran la mayoría de las ventas.

4. M – Model (Modelado Estadístico): Para fundamentar las decisiones, se profundizó en el análisis visual aplicando técnicas estadísticas que cuantifican la relación entre variables.

Se calculó la matriz de correlación (**Pearson**) para identificar las palancas de compra, detectando una **relación inversa significativa** entre la variable *TotalChildren* (Número de hijos) y la decisión de compra (*BikePurchase*). Esto confirma que, a mayor carga familiar, menor es la propensión al consumo del producto.

Regresión Lineal Simple: Mediante un modelo de regresión, se calculó la sensibilidad de la demanda ante la natalidad. El coeficiente resultante (**-0,039**) indica que, por cada hijo adicional en el promedio familiar, la probabilidad de compra se reduce casi un **4%**.

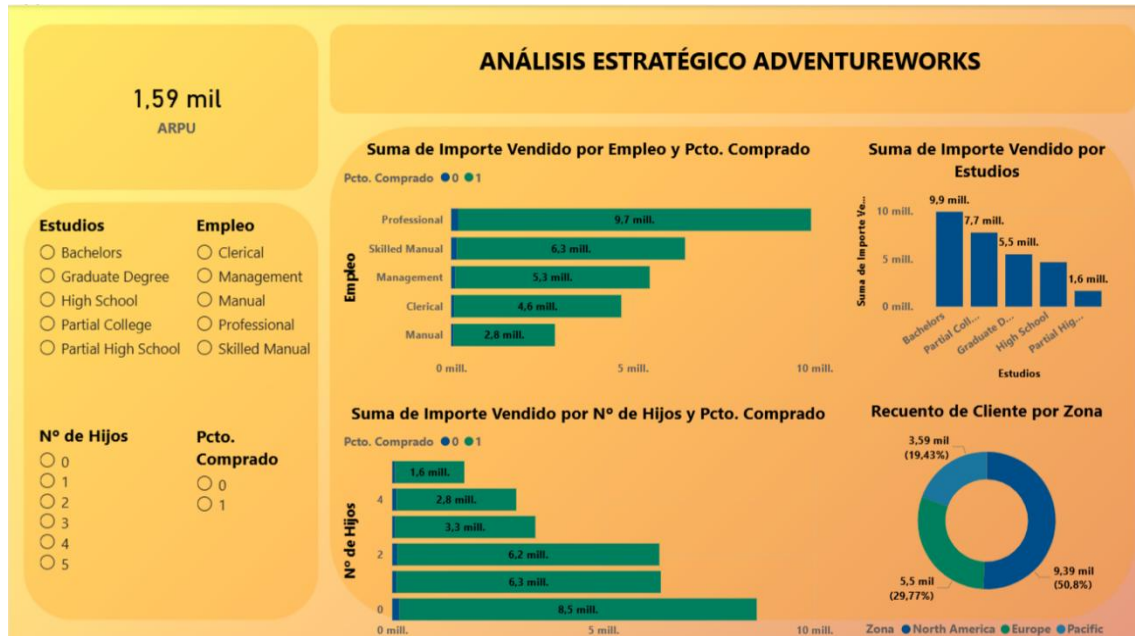
Ticket Medio Real: Se modeló el gasto promedio aislando exclusivamente a los compradores de bicicletas, obteniendo una media global ajustada de **3.145,5**, dato clave para las proyecciones financieras.

Nota: Se utilizó IA para validar la significancia estadística de las correlaciones observadas.

5. N – Interpretación: Estrategia y Predicción: El Dashboard monitoriza cinco indicadores clave. Su interpretación estratégica es la siguiente:

- **Ingreso Medio por Cliente (ARPU):** Para medir el valor medio por cliente. Calcula el ingreso promedio que genera cada cliente único. Permite evaluar la rentabilidad individual.
- **Ventas por Ocupación:** Para identificar los nichos laborales más rentables. Mide el volumen de ingresos acumulado según el perfil profesional del cliente (Occupation). Identifica qué nichos laborales (ej. *Professional* vs *Manual*) tienen mayor peso en la facturación para adaptar el lenguaje publicitario.
- **Índice de Gasto por Educación (HEI):** Relaciona el nivel de estudios alcanzado con el gasto total realizado, buscando identificar si la formación académica es un predictor de consumo. Ayuda a decidir en qué canales mediáticos invertir (por ejemplo, redes profesionales si el segmento con postgrado es el más rentable).
- **Tasa de Conversión Familiar:** Para entender el impacto de tener hijos en la compra de bicis. Mide la probabilidad de compra de una bicicleta en función de si el cliente tiene hijos (*TotalChildren*). Determinar si la bicicleta se vende como un artículo de ocio familiar, permitiendo reorientar la oferta hacia productos infantiles o de adultos.

- **Distribución Regional (Group Sales Share):** Analiza el rendimiento comercial agrupando los países en sus grandes regiones estratégicas (Group): North America, Europe y Pacific. Permite una visión macro del negocio para la toma de decisiones sobre expansión internacional y asignación de presupuestos logísticos por continente.



5.2. Interpretación Predictiva (Escenarios a 3-5 años)

La Hipótesis: *La bajada de la natalidad impulsará las ventas. (Nota: Hipótesis propia tras descartar las sugerencias de la IA por falta de relevancia)."*

Análisis de Escenarios: Aplicando el coeficiente de regresión (-0,039), el modelo cuantifica 3 escenarios: **Pesimista (-5%):** Aporta +203.668 €, un impacto positivo pero residual. **Moderado (-10%):** Genera +407.335 € de beneficio extra sin cambios estratégicos. **Optimista (-20%):** El escenario más lucrativo (tendencia Child-Free), sumando +814.670 € a la facturación.

| Proyección Escenarios Bajada de la Natalidad | | | | | | |
|--|----------------|----------------|----------------------|-----------|----------|-----------|
| Country | Total Clientes | Promedio Hijos | *Bajada de Natalidad | 20% | 10% | 5% |
| | | | Ticket Medio País | Optimista | Moderada | Pesimista |
| Australia | 3.591 | 1,63 | 4.154,06 | 192.054 | 96.027 | 48.013 |
| United States | 7.819 | 2,00 | 2.562,45 | 316.692 | 158.346 | 79.173 |
| Canada | 1.571 | 2,12 | 2.814,66 | 73.865 | 36.933 | 18.466 |
| France | 1.810 | 1,74 | 3.178,22 | 78.995 | 39.498 | 19.749 |
| United Kingdom | 1.913 | 1,63 | 3.232,17 | 79.291 | 39.646 | 19.823 |
| Germany | 1.780 | 1,67 | 3.156,37 | 73.773 | 36.886 | 18.443 |
| Total | | | | 814.670 | 407.335 | 203.668 |