

## Ejercicio N° 1: Automatización del flujo de ventas multicanal

### Introducción

Una empresa con operaciones online y físicas recibe diariamente grandes volúmenes de información desde diferentes canales de venta: e-commerce, puntos físicos, aplicaciones móviles y marketplaces. Actualmente, el equipo de datos integra manualmente los reportes, lo que genera cuellos de botella y errores humanos. Se requiere automatizar este flujo para que los datos lleguen estandarizados y en tiempo real al data warehouse de la compañía.

### 1. Componentes del flujo

- Fuentes de datos: e-commerce, puntos de venta físicos, aplicaciones móviles y marketplaces externos.
- Procesadores: limpieza de datos, conversión de formatos, normalización de campos, validación de calidad.
- Enrutador o decisor: reglas para enviar datos a distintos destinos (Data Lake, Data Warehouse, API de reportes).
- Destinos: Data Warehouse centralizado, Data Lake para históricos y API de análisis en tiempo real.

### 2. Plataforma de automatización seleccionada

Se selecciona Apache NiFi como herramienta principal, ya que permite diseñar flujos de integración de datos de forma visual, conectando múltiples orígenes (archivos, APIs, bases de datos, servicios cloud). Además, facilita el monitoreo y la trazabilidad del proceso. Alternativamente, Talend puede utilizarse en contextos con ETL más complejos.

### 3. Beneficios esperados

- Reducción de errores humanos gracias a la automatización.
- Disponibilidad en tiempo real de los datos.
- Estandarización en los formatos de ventas de todos los canales.
- Monitoreo continuo con alertas en caso de fallos.
- Visión única y consolidada de las ventas multicanal.

#### **4. Flujo propuesto (paso a paso)**

1. Recolección automática de datos desde e-commerce, POS, apps móviles y marketplaces.
2. Validación de estructura y limpieza de duplicados.
3. Transformación de los formatos a un estándar común (ejemplo: Parquet/CSV unificado).
4. Normalización de campos: códigos de producto, divisas, IDs de clientes.
5. Enrutamiento de datos según reglas de negocio: Data Lake, Data Warehouse o API de reportes.
6. Aplicación de buenas prácticas: trazabilidad, validación automática y alertas.
7. Monitoreo en tiempo real desde Apache NiFi.

#### **Conclusión**

Con la implementación de este flujo automatizado, la empresa logra integrar y estandarizar la información de ventas multicanal, garantizando disponibilidad en tiempo real, reducción de errores y consolidación de datos para la toma de decisiones estratégicas.