

Automatización del flujo de ventas multicanal

Nombre: Jonathan Vásquez

Componentes del flujo de datos

Fuentes de datos:

- Plataformas de e-commerce propias (webshop).
- Puntos de venta físicos (POS).
- Aplicaciones móviles de la empresa.
- Marketplaces externos (Ej: MercadoLibre, Amazon, etc).

Procesador (Transformaciones necesarias):

- Limpieza de datos: eliminar duplicados, corregir formatos de fechas y montos.
- Normalización: unificar nombres de productos, categorías y códigos de clientes.
- Validación: verificar integridad de campos críticos (ID de transacción, fecha, cantidad).
- Agregaciones: calcular totales diarios, promedio de ventas, y métricas por canal.

Enrutador o decisor:

- Aplicar reglas para enviar ciertos registros a destinos específicos:
 - Ventas físicas → Data Warehouse general.
 - Ventas online y marketplace → Data Lake para análisis avanzado.
 - Transacciones fallidas o incompletas → cola de alertas para revisión manual.

Destino de los datos:

- **Data Warehouse (DWH):** consolidación de ventas diarias y generación de reportes.

- **Data Lake:** almacenamiento histórico y análisis avanzado (ej. machine learning o forecasting).
- **API interna:** exposición de datos en tiempo real para dashboards internos.

Plataforma de automatización sugerida

Plataforma: Apache NiFi

Justificación:

- Permite procesamiento de flujos en tiempo real y por lotes.
- Soporta gran variedad de conectores para bases de datos, APIs y archivos.
- Permite enrutamiento y reglas de decisión de manera visual.
- Fácil de monitorear y auditar, ideal para reducir errores humanos en la integración manual.

Beneficios esperados

- **Reducción de errores humanos:** automatización evita fallas en la consolidación manual.
- **Disponibilidad en tiempo real:** los datos de ventas se procesan y actualizan de manera inmediata.
- **Estandarización del proceso:** datos consistentes y uniformes desde distintas fuentes.
- **Monitoreo continuo:** alertas automáticas ante datos incompletos o inconsistentes.
- **Escalabilidad:** soporta incremento de volumen de datos sin necesidad de cambios significativos.

Propuesta de flujo paso a paso

1. Captura de datos:

- Conectar Apache NiFi a todas las fuentes: e-commerce, POS, apps móviles y marketplaces.

2. Ingesta inicial:

- Recepción de archivos CSV/JSON, API calls o streams de eventos de ventas.

3. Limpieza y normalización:

- Eliminar duplicados, corregir formatos de fechas y unificar nombres de productos y categorías.

4. Validación:

- Revisar integridad de campos críticos: ID de transacción, monto, fecha, canal de venta.

5. Transformaciones y agregaciones:

- Calcular métricas clave: ventas totales, promedio por canal, cantidad de productos vendidos.

6. Enrutamiento según reglas:

- Transacciones completas → Data Warehouse.
- Transacciones online/marketplace → Data Lake.
- Transacciones incompletas o fallidas → sistema de alertas para revisión manual.

7. Carga a destino:

- Enviar datos procesados al DWH y Data Lake.
- Exponer datos relevantes a dashboards internos vía API.

8. Monitoreo y auditoría:

- Configurar alertas automáticas para errores de procesamiento o retrasos.
- Generar logs para trazabilidad y control interno.