**Proyecto**

En este módulo modernizamos nuestro caso de negocio transformando creando una arquitectura de procesamiento batch. Aquí comprenderás el uso de Orquestadores de pipelines dentro contexto práctico y real.

Realizarás un proyecto end to end utilizando todas las herramientas aprendidas en los módulos anteriores. Podrás aplicar las distintas fases de un proyecto:

* Definición de Arquitectura
* Identificación de fuentes
* Ingesta de Datos
* Modelamiento de Información
* Transformación y Procesamiento
* Orquestación de Data Pipelines
* Automatización

**Caso**

La empresa Retail.SA requiere centralizar la información del negocio y desea construir un repositorio de datos que le permita conocer más a sus clientes y tomar decisiones para mejorar su servicio y operación.

Requerimientos:

1. Desarrollar una Arquitectura básica para el desarrollo de la solución
2. Realizar ingesta de datos
3. Desarrollar los procesos ETL
   1. Top 5 clientes con mas ordenes
   2. Top 5 estados con más ordenes procesadas
   3. Cantidad de Ordenes por mes
   4. Cuáles son los Días de la semana de más facturación
   5. Top 5 tipos de ítems (Productos) más vendidos
   6. Top 5 clientes con mayor facturación
4. Automatizar los procesos
5. Desarrollar Data API’s
6. Publicar nuestro proyecto en GitHub