# **Proyecto: Diseño de arquitectura Cloud para un negocio de datos**

## **Objetivo**

Diseñar una **arquitectura en la nube** que soporte un negocio basado en datos. El proyecto combina:

* Diseño de un **modelo de negocio**.
* Creación de una **arquitectura cloud** completa (ingesta → procesamiento → almacenamiento → análisis → visualización → opcional ML).
* Uso de **web scraping o APIs públicas** para ilustrar con un ejemplo de datos reales.

## **Fases del proyecto**

### **1. Definición del negocio**

* Elijan un **sector** (ej.: e-commerce, turismo, deportes, salud, transporte, streaming, educación, etc.).
* Definan un **modelo de negocio basado en datos**, respondiendo:
  1. ¿Qué problema resuelven?
  2. ¿Qué datos necesitan recolectar?
  3. ¿Qué valor ofrecen a clientes/usuarios gracias a los datos?

👉 Entregable: ppt mostrando el modelo de negocio.

### **2. Arquitectura Cloud**

* Elegir un **proveedor cloud**: AWS, Azure o Google Cloud (solo de forma conceptual, no necesitan acceso real).
* Diseñar la **arquitectura de datos** incluyendo:
  1. **Ingesta de datos**: APIs, web scraping, formularios, sensores, logs, etc.
  2. **Procesamiento (ETL/ELT)**: transformación de datos crudos → datos limpios.
  3. **Almacenamiento**: decidir qué usar para datos estructurados (SQL) y no estructurados (NoSQL).
  4. **Análisis y visualización**: dashboards, reporting, BI.
  5. **Opcional**: módulos de **Machine Learning / AI** para predicciones, segmentaciones o recomendaciones.

👉 Entregable:

* **Diagrama de arquitectura** (puede ser en PowerPoint, Draw.io, Miro, Figma…).
* Documento explicando las decisiones tomadas: por qué eligieron cada servicio (ej.: S3 vs BigQuery vs Cosmos DB).

### **3. Ejemplo de datos (Web Scraping o API)**

* Implementar una **pequeña demo en Python** que obtenga datos de una fuente real:
  + **API pública** (ej.: PokeAPI, OpenWeather, Quotes API, etc.).
  + **Web scraping** sencillo (ej.: Books to Scrape, Wikipedia).
* Convertir esos datos a un **DataFrame de pandas**.
* Explicar **cómo se integrarían** en su arquitectura (ej.: ingesta → base de datos → análisis).

👉 Entregable:

* Jupyter Notebook con el código de scraping/API + DataFrame.
* Breve explicación de qué representan esos datos en el contexto de su negocio.

## **Presentación final**

Cada grupo deberá presentar en **10 minutos**:

1. El **modelo de negocio**.
2. El **diagrama de arquitectura cloud**.
3. La **demo con datos reales** y su integración en la arquitectura.

## **Criterios de evaluación**

* Creatividad y claridad en el modelo de negocio (20%).
* Completitud y coherencia en la arquitectura (30%)
* Uso correcto de scraping/API y pandas (20%).
* Claridad en la documentación y presentación (20%).
* Opcional: uso de ML/IA (10% bonus).