

Proyecto Final – Estrategia, Procesos y Gobierno de Datos

Organización del trabajo y entregas

El trabajo se desarrollará en **grupos de 3 personas** y tendrá un esquema de entregas parciales obligatorias que permitirán dar seguimiento al avance del proyecto. Cada entrega será revisada y retroalimentada para que puedan mejorar su entrega final.

Cronograma de entregas

1. Primera entrega parcial – Octubre 20

- Definición del problema, necesidad del negocio y estrategia corporativa.
- Propuesta inicial del modelado de procesos (diagramas BPM).
- Búsqueda y selección de conjunto(s) de datos abiertos relevantes.
- Boceto de arquitectura de datos (puede estar en borrador).

2. Segunda entrega parcial – Noviembre 10

- Pipeline de procesamiento implementado (limpieza y normalización de datos).
- Análisis exploratorio inicial (EDA).
- Primer borrador del gobierno de datos: definición de roles, reglas y estándares.
- Avance documentado en GitHub (estructura de carpetas, README inicial).

3. Entrega final – Noviembre 24

- Proyecto completo en GitHub con todos los entregables definidos.
- Documentación final en README y carpeta de documentación.
- Dashboard o visualización funcional para toma de decisiones.
- Presentación del pitch final en clase virtual (10-12 minutos).

Objetivo del Proyecto

Desarrollar una solución de negocio basada en datos, que cubra desde la formulación estratégica, el diseño de procesos, la implementación técnica, el gobierno de datos y la visualización para la toma de decisiones, usando datos abiertos reales. Todo debe estar documentado y ejecutable en un repositorio colaborativo en GitHub.



Fuentes de Datos Abiertos (usar al menos una)

Algunas opciones sugeridas:

- Gobierno de Colombia – Datos Abiertos: <https://www.datos.gov.co/>
- Banco Mundial – Indicadores Globales: <https://data.worldbank.org/>
- Naciones Unidas – Datos de Desarrollo Humano: <https://data.un.org/>
- Kaggle (datasets públicos): <https://www.kaggle.com/datasets>
- Our World in Data: <https://ourworldindata.org/>
- Open Data Soft / portales municipales internacionales.

Entregables

Repositorio en GitHub con código, documentación y notebooks ejecutables.

- Carpeta `/notebooks/` : limpieza, EDA, visualización.
- Carpeta `/documentacion/` : estrategia, procesos, arquitectura, gobernanza.
- `README.md` : resumen del trabajo y cómo ejecutar el proyecto.
- Carpeta `/img/` o `/dashboard/` : visualizaciones o reportes.
- Pitch virtual (10-12 minutos): presentación del caso, solución y conclusiones.

Nota: El código que se desarrollo debe ser en Python, siguiendo las buenas prácticas de programación.

Rúbrica de Evaluación

Criterio	Descripción esperada	Porcentaje	Nivel Excelente (100%)
1. Problema y estrategia	Definir una necesidad real usando pensamiento crítico y herramientas estratégicas.	10%	Problema claro, conectado al entorno digital y con análisis estratégico profundo.
2. Procesos y automatización	Modelar flujos de proceso con BPMN u otra notación, identificar mejoras.	10%	Flujos completos y automatización propuesta con herramientas realistas.

3. Arquitectura de datos	Diseñar cómo se integrarán y transformarán los datos.	15%	Diagrama técnico completo con herramientas, pasos y justificación.
4. Limpieza y normalización	Procesamiento, corrección de errores, estandarización de datos.	10%	Código limpio y documentado que transforma los datos de forma reproducible.
5. Análisis exploratorio (EDA)	Identificar patrones, relaciones y estadísticas útiles.	10%	EDA visual, con insights relevantes conectados al problema
6. Gobierno de datos	Definir roles, políticas, estándares y trazabilidad.	20%	Documento técnico robusto con reglas, catálogo, seguridad y responsables.
7. Visualización final	Visualización clara que permita tomar decisiones estratégicas.	10%	Dashboard atractivo y útil con conexión a los objetivos del proyecto.
8. Trabajo colaborativo (GitHub)	Evidencia de trabajo en equipo, documentación clara en README.	15%	Repositorio completo, commits bien distribuidos, README completo, notebooks funcionales.

Recomendaciones Finales

- Justificar cada decisión tomada en el proceso.
- Usar para la visualización la herramienta que deseen (menos Excel)
- Sean claros y colaborativos: el trabajo en equipo será evaluado.