
Proyecto No. 3 Implementación OLAP (Mismos grupos que el proyecto 2)

Modalidad y fecha de entrega

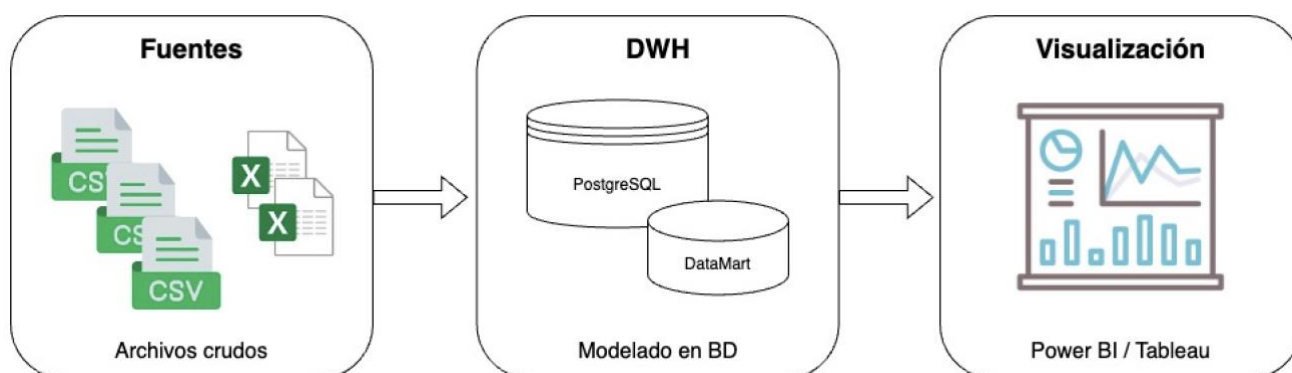
- El proyecto se hará en **grupos de 4 y un trío** según los grupos del proyecto anterior (un único grupo de 3 permitido) y debe de ser enviado antes de la fecha límite de entrega.
- El proyecto se entregará en la semana de exámenes finales (fecha y horario pendiente de definir), entre el 17 y 21 de Noviembre de 2025.
- El proyecto se presentará en clase ante el resto de la audiencia el mismo día, es decir, los tres grupos que presentarán a sus mismos compañeros de clase. Se deberá entregar previo a las fechas de presentación, aunque el orden de presentación será sorteado más adelante.
- No se permitirá la entrega o envío de proyecto más allá de la fecha límite, todos los grupos entregan el mismo día y en la misma hora.

Descripción general del proyecto

El objetivo principal del proyecto es construir un *pipeline* de datos bajo un escenario productivo y que refleje las diferentes etapas de los datos, desde los datos fuente que podemos encontrar o extraer de alguna parte, pasando por una etapa de transformación y procesamiento, y finalmente explotando estos datos en análisis que puedan brindarle un **valor agregado a los negocios** a través de **analítica prescriptiva**. La información que utilice en su proyecto estará almacenada en un repositorio centralizado que permita la consulta de estos según las arquitecturas de datos que hemos visto durante el curso.

El proyecto contempla el uso de tecnologías de bases de datos para la creación y carga de modelos de datos, con el objetivo de utilizar lenguaje SQL para investigación, desarrollo y presentación de resultados sobre preguntas de negocio para apoyo de toma de decisiones. Deberá aplicar todos los conceptos y buenas prácticas que ha podido aprender durante todo el curso.

La idea de este último proyecto es implementar la parte analítica y de visualización, construyendo **un caso de uso a su elección**. Podrán, incluso, utilizar el mismo caso de uso que utilizaron para su proyecto 2, pero no será obligatorio puesto que quedará a discreción de cada grupo. Cada grupo, deberá implementar la siguiente arquitectura a partir de un caso de uso a su elección:



A continuación, se explica cada punto de la arquitectura:

- Fuentes:** deberán de ser archivos **crudos** en cualquier formato: .csv, .txt, .xlsx, etc. Las fuentes pueden ser *datasets* de ejemplo que encuentre en cualquier repositorio o lugar de internet o bien, datos que el grupo decida generar de manera aleatoria. Para generar datos de manera aleatoria puede utilizar [este recurso](#). Para buscar datasets en internet, puede hacerlo de cualquier lugar, como por ejemplo repositorios robustos y variados como [kaggle](#). Recuerde que estas fuentes deberán de modelar un caso de uso que el grupo considere apropiado, y del cual deberán generar las visualizaciones más adelante. Los casos de uso a seleccionar **no podrán repetirse** entre los distintos grupos de proyecto. Será importante que cada grupo pueda seleccionar su caso de uso con antelación suficiente. Si el caso de uso lo desean cambiar más adelante lo podrán hacer siempre y cuando no repitan otro ya escogido y que el grupo acepte las repercusiones en cuanto al tiempo disponible para realizar el resto de actividades.
- Modelado en BD:** La idea con este punto es que pueda modelar las fuentes previas basándose en un **modelado estrella**, implementado en PostgreSQL. Sobre este RDBMS, deberá analizar qué datos le interesan visualizar y sobre estos, crear un **datamart** con diferentes tablas que construyan el análisis posterior. Dichas vistas serán la entrada de datos para sus visualizaciones.
- Visualización:** Por último, sobre alguna herramienta de visualización, deberá hacer una conexión hacia PostgreSQL y conectarse a las fuentes de su DataMart (es decir, la base de datos que tendrá almacenadas todas las tablas que le permitan construir su tablero de reportes). Dentro de las herramientas de visualización podrá escoger una entre las siguientes: [PowerBI Desktop](#) o [Tableau Desktop](#). Ambas son gratuitas, la diferencia es que Power BI no está disponible para descargarse sobre sistemas operativos Mac (a menos que se utilice PowerBI Server) , y Tableau sí se puede utilizar tanto en Windows como en Mac. Esta selección quedará a

criterio del grupo. Para poder conectar su herramienta de visualización a su instancia de PostgreSQL deberá tener en mente los siguientes datos de PostgreSQL:

- i. **Server:** el nombre de su servidor. Al estar instalado de manera local, generalmente es “localhost”.
- ii. **User:** el usuario que configuraron para postgres. El usuario default es “postgres”.
- iii. **Password:** la contraseña asociada a ese usuario. Si la olvidó, lo más fácil es volver a instalar PostgreSQL.
- iv. **Puerto:** No siempre se pide, pero el puerto es el 5432 por default.

Consideraciones básicas:

- 1) Está modelando un caso de uso. El enfoque que se le debe de dar a este proyecto es un **enfoque de negocio**, tal cual lo indican los fundamentos de inteligencia de negocio (BI).
- 2) No se complique buscando fuentes, puede utilizar cualquiera, incluso fuentes que pudimos utilizar a lo largo del curso y/o que no tengan formatos complicados de procesar. De hecho, podrá utilizar el mismo dataset construido para el proyecto anterior (Proyecto 02), bajo la consideración estipulada previamente en donde no se pueden repetir casos de uso. Recuerde que para cargar las fuentes hacia la base de datos en PostgreSQL, será necesario que aplique lo que utilizó previamente en los dos proyectos anteriores. Lo importante es que modelen un caso de uso sobre el cual sea fácil y/o intuitivo realizar los análisis y visualizaciones posteriores.
- 3) El DataMart se resume en crear una base de datos con algunas tablas ya procesadas o que contengan alguna **vista** o resumen de varias tablas. Lo importante, es que en su base de datos construyan un modelo estrella conformado por al menos una tabla de hechos (fact) y el resto de dimensiones (tablas catálogo).
- 4) La visualización deberá ir orientada a negocio. Es decir, los gráficos presentados deberán de ser bastante explicativos, fáciles de entender y que den una respuesta o análisis útil según el caso de negocio utilizado. Deberá de utilizar como mínimo **5 tipos de gráficos distintos** en su tablero de visualización. Se recomienda complementar este tablero con algunos [KPIs](#) (Key Performance Indicator), métricas, etc., siempre y cuando sean **representativos** al caso de uso. La cantidad de gráficos, KPIs, métricas y demás utilizadas en el tablero no tienen ningún límite y quedan a completa decisión del grupo. Eso sí, mientras más robusto, completo, interesante y atractivo visualmente quede este tablero, podrán optar a 10 puntos extras del proyecto.

Por favor, tomen en cuenta aspectos como: tipos de gráfico que hagan sentido, paleta de colores correcta, nombres y títulos de gráfico con su leyenda de ejes claramente establecida, etc.

- 5) Cada grupo tendrá que hacer una presentación **orientada a negocio**. Explicarán el caso de uso, explicarán las fuentes y el DataMart construido pero sobretodo, harán especial énfasis en la presentación de su tablero de visualización. Sobre esta presentación habrán otros 5 puntos extra para el mejor tablero de visualización elegido por la clase.

- 6) Podrá contactar a su catedrático para cualquier duda del proyecto en cualquier momento. No lo deje a última hora. Se utilizará al menos un día de cada semana de clases dedicado completamente al trabajo del proyecto durante el horario asignado. Será vital que aprovechen este tiempo para avanzar y resolver dudas.

Los pasos sugeridos para el inicio correcto del proyecto son (en este mismo orden):

1. Instalación del manejador PostgreSQL.
2. Selección de fuentes según el caso de negocio utilizado. **Fecha máxima de aprobación:** Viernes 07 de Noviembre de 2025.
3. Procesamiento, limpieza e ingesta de datos hacia PostgreSQL. Fecha *recomendada* de finalización: 11 de Noviembre de 2025.
4. Modelado estrella y construcción de DataMart. Fecha *recomendada* de finalización: 14 de Noviembre de 2025.
5. Instalación de herramienta de visualización.
6. Configuración y conexión entre herramienta de visualización y PostgreSQL. Fecha *recomendada* de finalización: 17 de Noviembre de 2025.
7. Diseño previo de las métricas, KPIs y gráficos a implementar en el tablero.
8. Creación de tablero de visualización.

Temas a reforzar

- Diseño de bases de datos
- Vistas
- Construcción de modelo estrella para Datamarts
- Uso de herramientas de inteligencia de negocios

Documentos a entregar

- Documentación de modelo de datos correspondiente a la Base de Datos utilizada (diagrama estrella a alto nivel detallando tablas, campos y relaciones).
- Código fuente de la lógica de procesamiento, limpieza e ingesta desarrollada para las fuentes.

-
- Exportar tablero de visualización como imagen o PDF.
 - Dump de su base de datos.

VI. Evaluación

- Modelado estrella, e implementación y construcción del DataMart: 10%
- Herramientas de BI y tablero de visualización: 70%
- Calidad de la presentación (20 minutos máximo de presentación, califican estudiantes y profesor): 20%
 - **Total: 100% (15 puntos netos)**

Puntos extras (15 pts máximo):

- Creatividad, análisis y diseño del tablero de visualización construido, siendo representativo para el negocio: **10 puntos máximo**
- Mejor tablero y presentación, a criterio de los participantes: **5 puntos extra.**