

# Informe Técnico: Análisis de la Eficiencia del Gasto Público en el Presupuesto de Hospitales del GCBA

## 1. Objetivo

El objetivo principal de este proyecto es analizar la eficiencia del gasto público ejecutado por los Hospitales del Gobierno de la Ciudad Autónoma de Buenos Aires (GCBA). A través de un proceso de análisis de datos estructurado, se busca identificar patrones de ejecución presupuestaria, comparar el desempeño entre distintas instituciones y comunas, y proporcionar indicadores clave (KPIs) que permitan una toma de decisiones informada para optimizar la asignación de recursos en el sistema de salud pública.

## 2. Proceso ETL (Python/Pandas)

El proceso de Extracción, Transformación y Carga (ETL) fue realizado utilizando el lenguaje Python, con la biblioteca Pandas como herramienta principal para la manipulación de datos.

### Extracción

Los datos fuente fueron extraídos de los registros detallados del Presupuesto del GCBA.

### Transformación

La fase de transformación se centró en asegurar la calidad y consistencia de los datos:

- **Limpieza de Datos:** Se identificaron y gestionaron valores nulos, registros duplicados y errores de formato. Se estandarizaron los nombres de columnas y se corrigieron inconsistencias en la tipificación de los datos.
- **Filtrado de Datos de Hospitales:** Se aplicaron filtros específicos para aislar únicamente los registros correspondientes a las unidades ejecutoras clasificadas como Hospitales, excluyendo otras dependencias de salud no hospitalarias.
- **Normalización:** Se normalizaron las cadenas de texto (ej. nombres de hospitales y actividades) a mayúsculas y se eliminaron caracteres especiales.

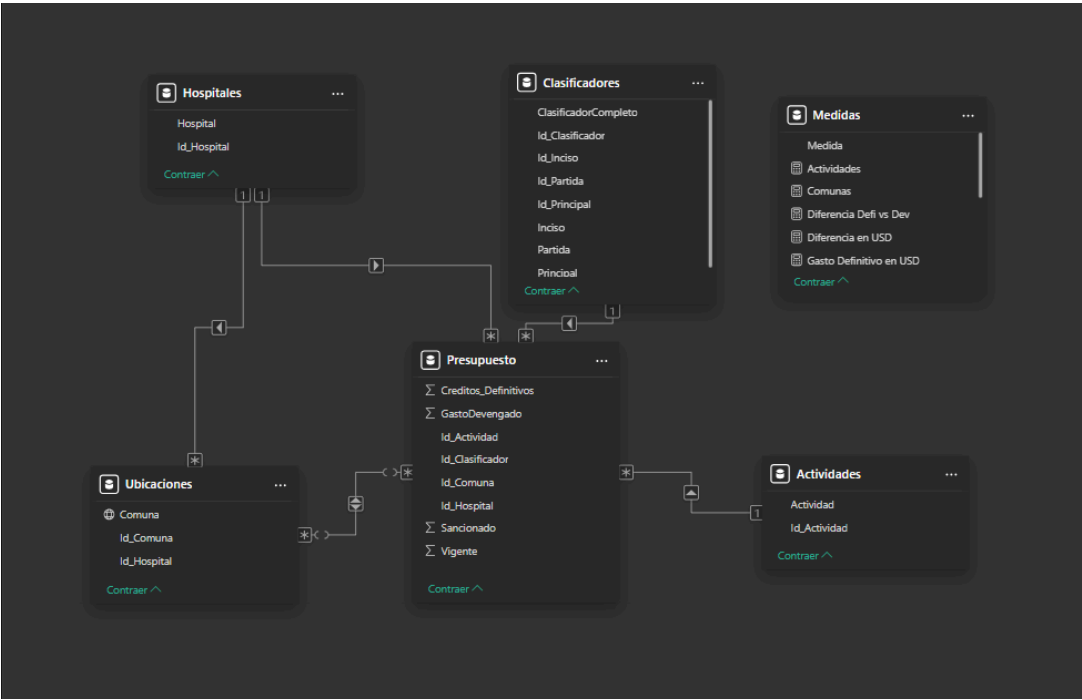
Esto fue crucial para asegurar la correcta correspondencia con las tablas de dimensiones durante el modelado.

A continuación, se detalla el flujo de transformación:

Etapa	Herramienta	Descripción de la Tarea
Carga Inicial	Python/Pandas	Lectura de archivos fuente (CSV/Excel).
Limpieza	Python/Pandas	Manejo de nulos, eliminación de duplicados, corrección de tipos de datos.
Filtrado	Python/Pandas	Selección de registros exclusivos de hospitales.
Normalización	Python/Pandas	Estandarización de nombres de columnas y valores categóricos.

### 3. Modelado de Datos (Esquema de Estrella)

Para facilitar el análisis dimensional y optimizar el rendimiento de las consultas, se implementó un Esquema de Estrella. Este esquema consta de una Tabla de Hechos central rodeada por Tablas de Dimensiones. Se crea una tabla “Medidas” para que sean de facil acceso.



### Tabla de Hechos

La Tabla de Hechos (Ejecución Presupuestaria) contiene las métricas cuantitativas clave y las claves foráneas que la relacionan con las dimensiones.

- **Claves Foráneas:** ID\_Hospital, ID\_Ubicacion, ID\_Actividad, ID\_Clasificadores.
- **Métricas (Hechos):** Presupuesto Sancionado (Monto), Presupuesto Vigente (Monto), Gasto Devengado (Monto), Gasto Definitivo(Monto).

### Tablas de Dimensiones (Dimension Tables)

Dimensión	Atributos Clave	Propósito
Hospitales	ID_Hospital, Hospitall.	Permite filtrar y analizar el gasto por institución.
Ubicaciones	ID_Comuna, ID_Hospital, Comuna.	Facilita el análisis geográfico del gasto y la eficiencia por comuna.
Actividades	ID_Actividad, Actividades.	Permite segmentar el gasto por las actividades aprobadas por el GBCA.
Clasificadores	ID_Clasificador, ClasificadorCompleto,	Permite segmentar el gasto por destino funcional (ej. insumos, personal, infraestructura).

Dimensión	Atributos Clave	Propósito
	ID_Inciso, Inciso, ID_Principal, Principal, ID_Partida, Partida.	

## 4. Visualización (Power BI)

La visualización de los datos modelados se llevó a cabo utilizando Power BI, permitiendo una exploración interactiva y la presentación clara de los hallazgos. El enfoque se centró en la creación de paneles que respondan a la pregunta de eficiencia del gasto.

### Indicadores Clave de Desempeño (KPIs)

Se diseñaron los siguientes KPIs de ejecución presupuestaria para medir la eficiencia del gasto en los hospitales del GCBA:

- **Porcentaje de Ejecución (Devengado / Definitivo):** Este indicador mide el avance real del gasto frente al presupuesto final aprobado. Según el tablero, la ejecución actual es del **83,20%**, lo que refleja la capacidad de la institución para transformar el **Monto Definitivo (\$307,03M)** en **Monto Devengado (\$255,43M)**.
- **Distribución del Gasto por Hospitales y Comunas:** Mide la asignación y utilización de recursos segmentada por centro de salud y ubicación geográfica. Permite visualizar, a través del mapa y el gráfico de barras, que el Htal. Argerich lidera el gasto con **\$23M**, y que la **Comuna 4** es la zona con mayor concentración presupuestaria.
- **Variación Presupuestaria (Sancionado vs. Definitivo):** Permite identificar la diferencia entre el **Monto Sancionado (\$1.418,80M)** y el **Monto Definitivo (\$307,03M)**. Este KPI es crucial para detectar reasignaciones significativas o ajustes en la planificación original durante el ejercicio fiscal.
- **Composición Económica del Gasto:** Evalúa en qué conceptos se invierten los recursos, destacando que el rubro de **Gastos en Personal (\$199,12M)** representa la mayor parte del presupuesto ejecutado, seguido por Bienes de Consumo.

### Mapas Geográficos por Comuna

Se utilizó la dimensión **Ubicaciones** para generar mapas que visualizan la ejecución presupuestaria agregada a nivel de Comuna.

- **Análisis de Densidad de Gasto:** Los mapas muestran, mediante sombreados de color, la concentración de gasto (total devengado) por comuna.

- **Eficiencia por Localización:** Al cruzar los KPIs de ejecución con el mapa, se pueden identificar rápidamente las comunas con hospitales que presentan la menor o mayor eficiencia, permitiendo dirigir inspecciones o ajustes de política pública a ubicaciones específicas.



Este trabajo valida la importancia de una arquitectura de datos robusta previa a la visualización. Al desacoplar la lógica de transformación de la capa de presentación, se ha optimizado el rendimiento del modelo, garantizando tiempos de respuesta ágiles incluso al filtrar por múltiples dimensiones como ubicación geográfica o clasificadores económicos.

El pipeline diseñado es modular y escalable. La estructura de tablas de hechos y dimensiones independientes permite integrar fácilmente nuevos períodos fiscales o ampliar el alcance a otras áreas de salud sin necesidad de reestructurar el modelo central. Este enfoque sienta las bases para futuras automatizaciones, donde la ingesta de datos podría programarse para actualizar el tablero de control en tiempo real ante cada nueva publicación oficial del GCBA.