Documentación Técnica del Proyecto GeoSalud

Durante el desarrollo del proyecto GeoSalud, se integraron, limpiaron, transformaron y analizaron más de 186 millones de registros provenientes de fuentes gubernamentales abiertas. Debido al volumen y complejidad de los datos, se aplicaron principios de ingeniería y arquitectura de datos, utilizando herramientas tipo ETL (Extract, Transform, Load) como KNIME. Esta operación a gran escala se considera dentro del contexto de Big Data, ya que implicó procesamiento masivo, normalización de columnas, creación de claves únicas, unión de múltiples fuentes heterogéneas y generación de variables derivadas para análisis geoespacial y proyecciones de salud pública a nivel localidad.

# 1. Introducción

GeoSalud es un proyecto de análisis de datos públicos del sector salud en México, con enfoque en zonas marginadas. Fue desarrollado utilizando herramientas de análisis y visualización de datos como KNIME, Power BI, Python y Flutter. La finalidad es identificar localidades con cobertura médica insuficiente y proponer soluciones basadas en datos, como unidades móviles, brigadas médicas o clínicas nuevas.

# 2. Fuentes de Datos Utilizadas

- RENUE (Registro Nacional de Unidades Económicas) del INEGI  
- Catálogo de localidades y población por localidad (INEGI)  
- Infraestructura médica disponible (Secretaría de Salud / Datos Abiertos)  
- Coordenadas geográficas de localidades  
- Datos abiertos en formato CSV y Excel

# 3. Procesamiento de Datos en KNIME

El flujo de trabajo en KNIME se diseñó para depurar, transformar y unir los datos poblacionales con la base de establecimientos médicos. Se realizaron los siguientes pasos:

a) Lectura de archivos: Excel Reader y CSV Reader  
b) Filtrado por entidad (Guanajuato) y limpieza de columnas  
c) Creación de columna 'llave\_union' combinando clave entidad, municipio y localidad  
d) Unión de RENUE con establecimientos médicos mediante 'llave\_union'  
e) Cálculo de población futura a 5 años con tasa de crecimiento anual (ejemplo: 2.2%)  
f) Clasificación de cobertura médica actual y futura mediante semáforo basado en la regla 1 unidad médica por cada 100 personas

# 4. Exportación de Resultados

Se exportaron los datos procesados en formato Excel y JSON para integrarlos en la aplicación móvil desarrollada con Flutter. También se preparó una versión para visualización en Power BI con tarjetas resumen, filtros y vistas municipales/locales.

# 5. Aplicación Móvil (GeoSalud App)

La aplicación fue desarrollada con Flutter y Dart. Se incluyeron los siguientes elementos:  
- Pantalla de inicio con filtros por municipio y localidad  
- Visualización de tarjetas por localidad con: población actual, clínicas, semáforo actual, población futura, recomendación futura  
- Iconos visuales y recomendaciones en colores (verde, amarillo, rojo)  
- Botón de mapa con pines para mostrar ubicación de localidades

# 6. Consideraciones Técnicas

- Proyecto desarrollado en Mac con chip M2 y Windows (Power BI)  
- Archivos finales exportados desde KNIME a JSON  
- Flutter utilizado para compilar aplicación como APK Android  
- Uso de lógica condicional para semáforos y banderas  
- Se respetó el principio de no modificar visualmente la app una vez estabilizada

# 7. Conclusión

GeoSalud demuestra el uso estratégico de datos públicos para detectar desigualdades en servicios de salud y proponer soluciones. El proyecto es escalable, visual, y está preparado para presentaciones profesionales o propuestas institucionales.