

Integración del Conjunto de Datos de Límites Municipales PDET

Análisis Geoespacial basado en MGN2024 del DANE

Entrega 2 — Proyecto Final: Administración de Bases de Datos

Pontificia Universidad Javeriana Bogotá

Facultad de Ingeniería — Ingeniería de Sistemas

Autores:

Natalia Ávila
Juan Diego Arias
Nicolás Camacho
Santiago Mesa

3 de noviembre de 2025

Índice

1. Resumen Ejecutivo	2
2. Adquisición y Verificación de Datos	2
2.1. Fuentes Utilizadas	2
2.2. Verificación	2
3. Integridad y Formato de los Datos	2
3.1. Normalización de Códigos DANE	2
3.2. Verificación de Correspondencias	3
4. Integración Espacial en NoSQL	3
4.1. Conversión a Formato GeoJSON	3
4.2. Estructura para MongoDB	3
5. Documentación del Proceso	3
5.1. Herramientas Utilizadas	3
5.2. Flujo de Trabajo	4
6. Conclusiones	4
7. Referencias	4

1. Resumen Ejecutivo

Este documento describe el proceso completo de integración de los límites administrativos oficiales de los municipios de Colombia, obtenidos del *Marco Geoestadístico Nacional 2024 (MGN2024)* del DANE, con el conjunto de municipios priorizados por los *Programas de Desarrollo con Enfoque Territorial (PDET)*.

El objetivo principal fue asegurar que el proyecto se enfoque únicamente en los territorios PDET, garantizando la integridad, consistencia y compatibilidad geoespacial de los datos.

2. Adquisición y Verificación de Datos

2.1. Fuentes Utilizadas

- **Límites Administrativos MGN2024 (DANE):** versión nacional con todos los niveles geográficos. Se filtró únicamente el nivel municipal (`MGN_ADM_MPIO_GRAFICO.shp`).
- **Listado de Municipios PDET:** archivo CSV oficial con 170 registros, incluyendo subregión PDET, código DANE, departamento y nombre del municipio.

2.2. Verificación

Ambos conjuntos fueron verificados antes de su integración:

- Los shapefiles fueron cargados correctamente con `GeoPandas`.
- Se ajustó la lectura del CSV usando codificación `latin1` y separador `;`.
- Se verificó la existencia de 170 municipios PDET válidos.

3. Integridad y Formato de los Datos

3.1. Normalización de Códigos DANE

Para permitir la unión entre ambos conjuntos:

- Se creó un campo `cod_dane_completo` combinando `dpto_ccdgo` y `mpio_ccdgo`.
- Se estandarizó el formato a cinco dígitos con ceros a la izquierda.

3.2. Verificación de Correspondencias

Se realizó la unión entre los municipios del DANE y los del archivo PDET mediante:

```
gdf_pdet = gdf.merge(  
    pdet,  
    left_on="cod_dane_completo",  
    right_on="Código DANE Municipio",  
    how="inner"  
)
```

El resultado fue exitoso: **170 municipios coincidentes**.

4. Integración Espacial en NoSQL

4.1. Conversión a Formato GeoJSON

El conjunto resultante se exportó en formato GeoJSON para permitir su integración con bases NoSQL:

```
gdf_pdet.to_file(  
    "/content/MGN2024_MUNICIPIOS_PDET.geojson",  
    driver="GeoJSON"  
)
```

4.2. Estructura para MongoDB

Cada documento contiene:

- Código DANE del municipio.
- Nombre del departamento y municipio.
- Subregión PDET.
- Geometría en formato Polygon o MultiPolygon.

5. Documentación del Proceso

5.1. Herramientas Utilizadas

- Python 3.12, Google Colab.

- Librerías: `pandas`, `geopandas`, `matplotlib`.
- Fuentes de datos: DANE (MGN2024), ART (PDET).

5.2. Flujo de Trabajo

1. Descarga y descompresión del paquete MGN2024.
2. Identificación del shapefile municipal.
3. Lectura y limpieza del CSV PDET.
4. Unión por código DANE.
5. Validación de los 170 municipios.
6. Exportación a formatos `.shp` y `.geojson`.

6. Conclusiones

- Se integraron exitosamente los límites municipales oficiales con los municipios PDET.
- El resultado mantiene coherencia geográfica y semántica.
- El archivo GeoJSON está listo para ser cargado en MongoDB u otra base NoSQL.

7. Referencias

- Departamento Administrativo Nacional de Estadística (DANE). *Marco Geoestadístico Nacional 2024*.
- Agencia de Renovación del Territorio (ART). *Listado de Municipios PDET, 2024*.
- Open Geospatial Consortium (OGC). *GeoJSON Specification*, 2016.