

009 MANEJO DE INFORMACIÓN GEOGRÁFICA



Geodata

Los datos geoespaciales se han vuelto cada vez más importantes y frecuentes en las aplicaciones modernas y los flujos de trabajo de ingeniería de datos, con casos de uso que van desde servicios basados en la ubicación hasta el monitoreo ambiental.

La información geográfica sirve para representar, analizar y visualizar datos relacionados con ubicaciones en la Tierra. Es fundamental para entender patrones espaciales, tomar decisiones basadas en el lugar, y gestionar recursos naturales, planificación urbana, transporte, y respuesta a emergencias, entre otros. Ayuda a conectar datos con su ubicación, lo que permite un análisis más profundo y contextualizado de fenómenos físicos, sociales y económicos.

CONCEPTOS ELEMENTALES

Algunos conceptos elementales que deben aprender para hablar el lenguaje de mapas son:

1. **GIS (Sistema de Información Geográfica):** Tecnología para capturar, almacenar, analizar y visualizar datos geográficos.
2. **Layer (Capa):** Conjunto de datos geográficos superpuestos que representan diferentes tipos de información (como carreteras, ríos, ciudades).
3. **Polígono:** Forma geométrica cerrada que representa áreas en un mapa, como terrenos o lagos.
4. **Punto:** Coordenada específica en un mapa, representada por una latitud y una longitud, usada para indicar ubicaciones precisas como una tienda o un punto de interés.
5. **Línea:** Conjunto de puntos conectados que representa caminos, ríos, o fronteras.
6. **Buffer:** Área de influencia alrededor de un punto, línea o polígono, utilizada para análisis espacial.

7. **Proyección Cartográfica:** Método para representar la superficie curva de la Tierra en un mapa plano.
8. **Raster:** Imagen compuesta de celdas o píxeles, cada uno con un valor, utilizada para representar datos continuos como elevación o temperatura.
9. **Vector:** Representación de datos geográficos usando puntos, líneas, y polígonos.
10. **Geocodificación:** Proceso de convertir direcciones u otras descripciones de lugares en coordenadas geográficas.

FORMATOS DE ARCHIVO

A continuación, se presentan algunos de los formatos más utilizados para almacenar y compartir datos geoespaciales:

1. **Shapefile (.shp):** Formato vectorial común en SIG, compuesto por varios archivos auxiliares.
2. **GeoJSON (.geojson):** Formato basado en JSON para representar datos geoespaciales en aplicaciones web.
3. **KML (.kml):** Formato XML utilizado para datos geoespaciales en Google Earth y SIG.
4. **GPX (.gpx):** Formato XML para el intercambio de datos de rastreo y rutas desde dispositivos GPS.
5. **WKT (Well-Known Text) (.wkt):** Formato de texto para representar geometrías espaciales de manera legible.
6. **GML (Geography Markup Language) (.gml):** Formato XML compatible con estándares OGC para datos geoespaciales.
7. **Raster (.tif, .png, .jpg):** Formatos de imagen para datos ráster en SIG, como imágenes satelitales.
8. **DXF (Drawing Exchange Format) (.dxf):** Formato CAD para dibujos y planos, integrado con SIG.
9. **HDF5 (.h5, .hdf5):** Formato para almacenar grandes volúmenes de datos multidimensionales.
10. **NetCDF (.nc):** Formato para datos multidimensionales, común en modelado climático y oceánico.
11. **PostGIS (.sql):** Formato para datos espaciales en bases de datos PostgreSQL con extensión PostGIS.
12. **SDE (Spatial Data Engine):** Formato para almacenar datos espaciales en bases de datos Esri.
13. **Parquet (.parquet):** Formato de almacenamiento en columna eficiente para grandes volúmenes de datos geoespaciales.

DUDCKDB Y DATOS GEOGRÁFICOS

Para manejar información geográfica en DuckDB, lo primero que debes hacer es instalar y activar el componente geoespacial.

```
INSTALL spatial;
LOAD spatial;
```

Documentación oficial: <https://duckdb.org/docs/extensions/spatial/overview.html>

MÉTODOS COMUNES

DuckDB ofrece una amplia gama de funcionalidades y formatos de archivos para el procesamiento de datos geográficos (como se muestra en el siguiente [link](#)). Sin embargo, para aprovechar al máximo estas capacidades, se requieren conocimientos técnicos más profundos. A continuación, listaremos las funcionalidades más elementales que puedes utilizar para empezar.

ST_READ : es una función que se utiliza para leer datos geoespaciales desde archivos en formato WKT (Well-Known Text) o WKB (Well-Known Binary). Ejemplo:

```
SELECT *
FROM ST_Read("ruta/archivo.geojson")
```

ST_GEOMFROMWKB : que se utiliza para crear un objeto geoespacial a partir de una representación en formato WKB (Well-Known Binary) en un objeto geométrico que puede ser utilizado para operaciones espaciales. Ejemplo:

```
SELECT ST_GEOMFROMWKB( '<binary_data>' );
```

ST_GEOMFROMWKT : es una función que se utiliza para convertir una representación textual de una geometría en formato WKT (Well-Known Text) en un objeto geométrico que puede ser procesado para análisis espaciales. Ejemplo:

```
SELECT ST_GEOMFROMWKT( 'POINT(1 1)');
```

ST_GeomFromText --- ajustar

ST_WITHIN : es una función que se utiliza para determinar si una geometría está completamente contenida dentro de otra geometría. Ejemplo:

```
-- Verificar si un punto está dentro de un polígono
SELECT
ST_WITHIN(
    ST_GEOMFROMWKT( 'POINT(1 1)' ),
    ST_GEOMFROMWKT( 'POLYGON((0 0, 0 2, 2 2, 2 0, 0 0))' )
);
```

ST_POINT : es una función que se utiliza para crear un objeto de punto geográfico a partir de coordenadas de longitud y latitud. Ejemplo:

```
SELECT ST_POINT(6.209004649498719, -75.56798270165233);
```

ST_TRANSFORM : es una función que se utiliza para convertir geometrías de un sistema de coordenadas a otro. Esto es fundamental cuando se trabaja con datos geoespaciales que pueden estar en diferentes sistemas de referencia o proyecciones. Ejemplo:

```
SELECT ST_TRANSFORM(
    ST_POINT(6.209004649498719, -75.56798270165233),
    'EPSG:4326', 'EPSG:3857'
);
```

SP_DISTANCE : es una función que se utiliza para calcular la distancia entre dos geometrías. Ejemplo:

```
SELECT ST_Distance(
    ST_TRANSFORM(ST_POINT(6.209004649498719, -75.56798270165233), 'EPSG:4326',
    'EPSG:3857'),
    ST_TRANSFORM(ST_POINT(6.271055144475173, -75.57311620567457), 'EPSG:4326',
    'EPSG:3857')
) / 1000 distancia
```