

Gestión de Datos Espaciales con DuckDB

Desde los Fundamentos de SQL hasta el
Análisis Geoespacial Avanzado



Qiusheng Wu

Gestión de Datos Espaciales con DuckDB

De los Fundamentos de SQL a la Analítica
Geoespacial Avanzada

Qiusheng Wu
2026

Contents

Prefacio	1
Introducción	3
Para Quién Es Este Libro	4
Qué Cubre Este Libro	5
Cómo Aprovechar al Máximo Este Libro	6
Convenciones Usadas en Este Libro	7
Descarga de los Ejemplos de Código	8
Tutoriales en Video y Recursos Complementarios	8
Comunidad y Retroalimentación	9
Agradecimientos	9
Acerca del Autor	9
Licencia y Derechos de Autor	10
I: Fundamentos de DuckDB	11
1. Primeros Pasos con DuckDB	13
1.1. Introducción	13
1.2. Objetivos de Aprendizaje	13
1.3. Qué Hace Diferente a DuckDB de las Bases de Datos Tradicionales	13
1.4. Cuándo (y Cuándo No) Usar DuckDB para Trabajo Espacial	15
1.5. Instalación del CLI de DuckDB y Ejecución de Tu Primera Consulta	17
1.6. Instalación del Cliente Python de DuckDB	18
1.7. Instalación de Visual Studio Code	22
1.8. Uso de la Interfaz de Usuario de DuckDB	25
1.9. Instalación del IDE SQL DBeaver	26
1.10. Puntos Clave	30
1.11. Ejercicios	31
2. SQL Esencial para Análisis Espacial	34
2.1. Introducción	34
2.2. Objetivos de Aprendizaje	34
2.3. Conjuntos de Datos de Ejemplo	35
2.4. Configuración del Entorno	35
2.5. Conectando a DuckDB	37
2.6. Instalar Extensiones	37
2.7. Lectura de Archivos CSV desde URLs	37
2.8. Creación de Tablas para Mejor Rendimiento	38
2.9. La Sentencia SELECT de SQL	40
2.10. Filtrando Datos con la Cláusula WHERE	47
2.11. Coincidencia de Patrones con LIKE	48
2.12. El Operador IN	50
2.13. El Operador BETWEEN	51
2.14. Combinando Datos con Uniones SQL	52
2.15. Planes de Consulta y Rendimiento	58
2.16. Agregando Datos para Estadísticas de Resumen	59
2.17. Sentencias Condicionales	62
2.18. Guardando Resultados	63
2.19. Características Avanzadas de SQL para Análisis Espacial	65
2.20. Comentarios y Documentación SQL	69

2.21. Conclusiones Clave	70
2.22. Ejercicios	71
3. Integración de DuckDB con Python	75
3.1. Introducción	75
3.2. Objetivos de Aprendizaje	75
3.3. Conjuntos de Datos de Ejemplo	76
3.4. Instalación y Configuración	77
3.5. Instalación y Carga de Extensiones	78
3.6. Lectura de Datos desde Múltiples Fuentes	79
3.7. Integración Perfecta con DataFrames de Pandas	82
3.8. Interoperabilidad con Polars	86
3.9. Conversión de Resultados y Formatos de Salida	86
3.10. Escribiendo Datos a Disco	89
3.11. Almacenamiento Persistente y Archivos de Base de Datos	89
3.12. Sentencias Preparadas y Parámetros	92
3.13. Puntos Clave	93
3.14. Ejercicios	94
II: Operaciones con Datos Espaciales	101
4. Carga de Formatos de Datos Espaciales	103
4.1. Introducción	103
4.2. Objetivos de Aprendizaje	104
4.3. Conjuntos de Datos de Ejemplo	105
4.4. Instalación y Configuración	105
4.5. Instalación y Carga de Extensiones	106
4.6. Descarga de Datos de Ejemplo	107
4.7. Carga de Archivos CSV con Coordenadas	107
4.8. Carga de Archivos JSON	111
4.9. Consultar DataFrames de Pandas Directamente	113
4.10. Carga de Archivos Parquet para Rendimiento	114
4.11. Carga de Archivos GeoJSON con Geometrías Espaciales	117
4.12. Carga de Shapefiles en Flujos de Trabajo Modernos	123
4.13. Carga de GeoParquet para Análisis Espacial Nativo de la Nube	125
4.14. Estrategias de Rendimiento para Carga de Datos	127
4.15. Solución de Problemas Comunes de Carga de Datos	128
4.16. Puntos Clave	129
4.17. Ejercicios	131
5. Exportación y Conversión de Datos Espaciales	134
5.1. Introduccion	134
5.2. Objetivos de Aprendizaje	135
5.3. Conjuntos de Datos de Muestra	136
5.4. Instalacion y Configuracion	137
5.5. Instalacion y Carga de Extensiones	137
5.6. Carga de Datos de Muestra	139
5.7. Exportacion a Pandas DataFrames	140
5.8. Exportacion a Archivos CSV	145
5.9. Exportacion a Archivos JSON	149
5.10. Exportacion a Archivos Excel	152
5.11. Exportacion a Archivos Parquet	155

5.12.	Exportacion a Formato GeoJSON	157
5.13.	Exportacion a Formato Shapefile	161
5.14.	Exportacion a Formato GeoPackage	162
5.15.	Conclusiones Clave	164
5.16.	Ejercicios	166
6.	<i>Operaciones y Funciones de Geometría</i>	171
6.1.	Introducción	171
6.2.	Objetivos de Aprendizaje	171
6.3.	Conjuntos de Datos de Ejemplo	172
6.4.	Instalación y Configuración	173
6.5.	Conectando a DuckDB y Cargando Extensiones	174
6.6.	Comprendiendo los Tipos de Geometría	176
6.7.	Trabajando con Points	179
6.8.	Trabajando con Linestrings	182
6.9.	Trabajando con Polygons	184
6.10.	Trabajando con Colecciones	186
6.11.	Visualizando Datos Espaciales de NYC	187
6.12.	Procesamiento de Geometría	190
6.13.	Validez y Robustez de Geometría	194
6.14.	Conclusiones Clave	196
6.15.	Ejercicios	197
7.	<i>Consultas y Relaciones Espaciales</i>	202
7.1.	Introducción	202
7.2.	Objetivos de Aprendizaje	203
7.3.	Conjuntos de Datos de Ejemplo	203
7.4.	Instalación y Configuración	204
7.5.	Conexión a DuckDB y Carga de Extensiones	204
7.6.	Entendiendo las Relaciones Espaciales	206
7.7.	Probando Identidad Geométrica	206
7.8.	Relaciones Topológicas	209
7.9.	Relaciones Basadas en Distancia	217
7.10.	Filtrando por Umbral de Proximidad	219
7.11.	Consultas de Vecino Más Cercano	220
7.12.	Puntos Clave	220
7.13.	Ejercicios	222
8.	<i>Uniones Espaciales Avanzadas</i>	225
8.1.	Introducción	225
8.2.	Objetivos de Aprendizaje	226
8.3.	Conjuntos de Datos de Ejemplo	227
8.4.	Instalación	227
8.5.	Importación de Bibliotecas y Configuración	227
8.6.	Conectando a DuckDB	227
8.7.	Uniones de Intersección	228
8.8.	Uniones Dentro de Distancia	234
8.9.	Unión Avanzada	238
8.10.	Transformaciones de Sistemas de Coordenadas	239
8.11.	Referencia de Funciones de Relaciones Espaciales	242
8.12.	Conclusiones Clave	243

8.13. Ejercicios	245
9. Visualización Interactiva de Datos	250
9.1. Introducción	250
9.2. Objetivos de Aprendizaje	251
9.3. Conjuntos de Datos de Muestra	251
9.4. Instalación y Configuración	252
9.5. Descarga de Datos de Muestra	252
9.6. Conexión a DuckDB y Carga de Extensiones	252
9.7. Visualización de Datos de Puntos	253
9.8. Visualización de Datos de Líneas	257
9.9. Visualización de Datos de Polígonos	259
9.10. Conclusiones Clave	264
9.11. Ejercicios	266
10. Trabajo con Teselas Vectoriales y PMTiles	270
10.1. Introducción	270
10.2. Objetivos de Aprendizaje	271
10.3. Conjuntos de Datos de Ejemplo	272
10.4. Instalación y Configuración	273
10.5. Visualización de Teselas Vectoriales Directamente desde Archivos	273
10.6. Conversión de Datos Vectoriales a PMTiles	278
10.7. Visualización de PMTiles	280
10.8. Puntos Clave	289
10.9. Ejercicios	291
III: Analítica Geoespacial del Mundo Real	297
11. Análisis del Inventario Nacional de Humedales de EE.UU.	299
11.1. Introducción	299
11.2. Objetivos de Aprendizaje	300
11.3. Conjuntos de Datos Utilizados en Este Capítulo	301
11.4. Comprendiendo la Fuente del Conjunto de Datos	303
11.5. Accediendo a Datos de Humedales con DuckDB	305
11.6. Visualizando Distribuciones de Humedales	309
11.7. Análisis de Humedales a Escala Nacional	313
11.8. Puntos Clave	326
11.9. Ejercicios	328
12. Análisis de Huellas de Edificios Globales	331
12.1. Introducción	331
12.2. Objetivos de Aprendizaje	332
12.3. Acerca del Conjunto de Datos	333
12.4. Instalación y Configuración	335
12.5. Instalación y Carga de Extensiones	336
12.6. Explorando los Datos Disponibles	336
12.7. Análisis Regional con Filtrado de Caja Delimitadora	340
12.8. Comparación Multi-Regional	349
12.9. Filtrado por Fuente de Datos y Evaluación de Calidad	355
12.10. Agregación Espacial Basada en Grilla	360
12.11. Agregando Datos de Edificios con Grillas Hexagonales H3	365
12.12. Conclusiones Clave	376
12.13. Ejercicios	378

13. <i>Análisis de Datos de Taxis de NYC</i>	381
13.1. Introducción	381
13.2. Objetivos de Aprendizaje	382
13.3. Acerca del Conjunto de Datos	383
13.4. Instalación	385
13.5. Importación de Bibliotecas	385
13.6. Instalación y Carga de Extensiones	385
13.7. Carga de Datos de Taxis	386
13.8. Análisis Temporal	390
13.9. Carga de Datos de Zonas de Taxis	397
13.10. Análisis Espacial	400
13.11. Análisis de Flujo de Viajes	408
13.12. Análisis de Pagos y Económico	412
13.13. Análisis del Comportamiento de Pasajeros	416
13.14. Análisis Multi-Mensual	418
13.15. Visualización	420
13.16. Consejos de Optimización de Rendimiento	426
13.17. Puntos Clave	427
13.18. Ejercicios	429
14. <i>Desarrollo de Dashboards Interactivos con Voilà y Solara</i>	432
14.1. Introducción	432
14.2. Objetivos de Aprendizaje	433
14.3. Instalando Voilà y Solara	434
14.4. Introducción a Hugging Face Spaces	434
14.5. Creando una Aplicación Básica de Voilà	436
14.6. Creando una Aplicación Web Avanzada con Solara	446
14.7. Puntos Clave	454
14.8. Ejercicios	455

Prefacio

Introducción

En un mundo cada vez más impulsado por los datos, la capacidad de gestionar y analizar eficazmente la información espacial se ha vuelto esencial. Desde la planificación urbana y el monitoreo ambiental hasta la logística y los servicios personalizados basados en ubicación, los datos geoespaciales forman la columna vertebral de numerosas aplicaciones que influyen en nuestra vida diaria. Sin embargo, trabajar con datos espaciales a menudo se ha considerado un dominio especializado y complejo, que requiere herramientas intrincadas y una curva de aprendizaje pronunciada.

Entra DuckDB.

DuckDB es una base de datos analítica innovadora diseñada para la eficiencia y la facilidad de uso. Como base de datos OLAP (Procesamiento Analítico en Línea) en proceso, opera directamente dentro de tu aplicación, eliminando la necesidad de despliegues de servidores separados y configuraciones engorrosas. Esta naturaleza integrada, junto con su arquitectura orientada a columnas y motor de ejecución vectorizada, hace que DuckDB sea excepcionalmente rápido para consultas analíticas, incluso con grandes conjuntos de datos. Inicialmente ganando popularidad por sus capacidades de procesamiento de datos de propósito general, el ecosistema en rápida expansión de DuckDB (especialmente sus extensiones para datos espaciales) representa un cambio transformador.

Este libro, “***Gestión de Datos Espaciales con DuckDB: Desde Conceptos Básicos de SQL hasta Análisis Geoespacial Avanzado***,” tiene como objetivo desmitificar los datos geoespaciales y mostrar cómo DuckDB empodera a todos (desde analistas de datos y científicos hasta desarrolladores y profesionales de SIG) para aprovechar sus capacidades con una simplicidad y velocidad sin precedentes. Creemos que el análisis espacial robusto debe ser accesible, no confinado a software especializado costoso o lenguajes de programación complejos. Con DuckDB, esa accesibilidad se convierte en realidad.

Nuestro viaje comienza con los conceptos fundamentales de datos espaciales y su representación, estableciendo una base sólida en SQL para trabajar con puntos, líneas y polígonos. A medida que avanzamos, descubrirás cómo las características geoespaciales nativas de DuckDB (mejoradas por su extensión compatible con PostGIS) permiten operaciones sofisticadas como uniones espaciales, buffering y búsquedas de vecinos más cercanos a través de elegantes consultas SQL. Exploraremos varias aplicaciones del mundo real, demostrando cómo cargar, transformar, analizar y visualizar conjuntos de datos espaciales, empoderándote para extraer información significativa de la información geográfica.

Ya sea que estés buscando integrar el análisis espacial en tus pipelines de datos, realizar consultas geoespaciales ad-hoc rápidas o desarrollar aplicaciones interactivas conscientes de la ubicación, este libro servirá como tu guía completa. Cubriremos temas que van desde configurar tu entorno DuckDB e importar diversos formatos de archivos espaciales (como Shapefiles, GeoJSON y GeoParquet) hasta ejecutar tareas analíticas complejas e integrarse con herramientas de visualización.

Nuestro objetivo no es solo enseñarte sintaxis, sino cultivar una comprensión de por qué estas herramientas y técnicas son poderosas. Al final de este libro, serás competente en el uso de DuckDB como tu motor predilecto para la gestión y análisis de datos espaciales, desbloqueando nuevas posibilidades para tus proyectos y empoderándote para tomar decisiones informadas y espacialmente conscientes.

Únete a nosotros mientras profundizamos en la emocionante intersección de las capacidades analíticas de DuckDB y el rico mundo de los datos geoespaciales. El futuro del análisis espacial accesible está aquí, y funciona con DuckDB.

Para Quién Es Este Libro

Este libro está diseñado para cualquier persona que enfrenta las complejidades del análisis moderno de datos espaciales. Si alguna vez has pasado horas esperando que termine una unión espacial, has luchado por cargar grandes conjuntos de datos geográficos en memoria, o has deseado una forma más sencilla de combinar el poder de SQL con operaciones espaciales, este libro es para ti.

Encontrarás el Mayor Valor Si Eres

Un Profesional de SIG frustrado por las limitaciones del software de escritorio al manejar grandes conjuntos de datos. Estás familiarizado con QGIS o ArcGIS, pero necesitas analizar millones de características, procesar extensas pistas GPS o integrar el análisis espacial en flujos de trabajo automatizados.

Un Científico de Datos o Analista que frecuentemente encuentra datos de ubicación. Te sientes cómodo con Python y pandas, pero los datos espaciales a menudo se sienten como un misterio. Quieres incorporar dimensiones geográficas en tus análisis sin sumergirte en software SIG complejo.

Un Desarrollador de Software que construye aplicaciones que incorporan características espaciales. Necesitas consultas espaciales rápidas, deseas evitar infraestructura pesada de bases de datos y prefieres trabajar con SQL familiar en lugar de bibliotecas espaciales especializadas.

Un Investigador o Académico en campos como geografía, ciencias ambientales o planificación urbana. Tu investigación involucra grandes conjuntos de datos espaciales, y requieres métodos de análisis reproducibles y escalables que puedan adaptarse a volúmenes de datos crecientes.

Un Profesional de Inteligencia de Negocios que trabaja con datos empresariales basados en ubicación. Ya sean ubicaciones de tiendas, rutas de entrega, territorios de clientes o carteras inmobiliarias, necesitas fusionar métricas empresariales con información espacial.

Prerrequisitos Esenciales

Deberías sentirte cómodo con:

- **Programación en Python:** entender variables, funciones y cómo importar bibliotecas (no se requiere experiencia avanzada)
- **Conceptos de análisis de datos:** filtrar registros, agregar datos y unir tablas
- **Fundamentos de SQL:** familiaridad con cláusulas SELECT, WHERE y GROUP BY (cubriremos los aspectos espaciales)
- **Conceptos básicos de datos espaciales:** entender que los datos tienen una ubicación (latitud/longitud, proyecciones)

Conocimientos Útiles (Pero No Requeridos)

- Experiencia con pandas, GeoPandas o Jupyter notebooks
- Exposición previa a bases de datos o data warehouses
- Familiaridad con software SIG (QGIS, ArcGIS, PostGIS)
- Conocimiento de formatos de archivos espaciales (GeoJSON, Shapefiles, Parquet)

Si Eres Nuevo en la Programación Python

Si eres nuevo en la programación geoespacial con Python, el siguiente libro proporciona una excelente introducción tanto a los conceptos fundamentales de SIG como a la programación en Python:

Wu, Q. (2025). *Introduction to GIS Programming: A Practical Python Guide to Open Source Geospatial Tools*. Independently published. ISBN 979-8286979455. <https://www.amazon.com/dp/B0FFW34LL3>

Qué Cubre Este Libro

Este libro ofrece un viaje estructurado desde los conceptos básicos de SQL hasta el análisis geoespacial avanzado, equipándote con habilidades prácticas a través de ejemplos del mundo real. Cada capítulo progresa desde consultas simples hasta análisis espaciales complejos, desarrollando tu experiencia en la gestión moderna de datos geoespaciales.

Parte I: Fundamentos de DuckDB (*Capítulos 1-3*)

Domina los conceptos esenciales que sustentan todo el contenido posterior:

- **Capítulo 1: Primeros Pasos con DuckDB:** Instalación, consultas iniciales e información sobre cómo DuckDB revoluciona el análisis espacial.
- **Capítulo 2: SQL Esencial para Análisis Espacial:** Patrones SQL clave para filtrar, agregar, unir y optimizar consultas adaptadas para datos espaciales.
- **Capítulo 3: Integración de DuckDB con Python:** Combina el poder de SQL con la flexibilidad de pandas para crear un flujo de trabajo de análisis espacial sin interrupciones.

Al final de la Parte I, consultarás conjuntos de datos espaciales con confianza e integrarás DuckDB en cualquier pipeline de análisis basado en Python.

Parte II: Operaciones con Datos Espaciales (*Capítulos 4-10*)

Sumérgete en el kit de herramientas espaciales principal, cubriendo todo desde la carga de datos hasta el análisis avanzado:

- **Capítulo 4: Carga de Formatos de Datos Espaciales:** Importa varios formatos incluyendo coordenadas CSV, GeoJSON desde APIs, Shapefiles masivos y GeoParquet alojado en la nube.
- **Capítulo 5: Exportación y Conversión de Datos Espaciales:** Transforma tus resultados en cualquier formato que tus interesados requieran.
- **Capítulo 6: Operaciones y Funciones de Geometría:** Crea, mide y transforma características espaciales utilizando funciones SQL.
- **Capítulo 7: Consultas y Relaciones Espaciales:** Domina los equivalentes espaciales de las uniones de tablas: contención, intersección y proximidad.
- **Capítulo 8: Uniones Espaciales Avanzadas:** Combina conjuntos de datos por ubicación en lugar de IDs, que es la esencia del análisis espacial.
- **Capítulo 9: Visualización Interactiva de Datos:** Genera mapas y gráficos convincentes que comuniquen efectivamente las narrativas espaciales de tus datos.
- **Capítulo 10: Trabajo con Teselas Vectoriales y PMTiles:** Despliega mapas interactivos capaces de manejar millones de características sin problemas.

Al final de la Parte II, serás experto en gestionar cualquier formato de datos espaciales, ejecutar operaciones complejas y crear visualizaciones profesionales.

Parte III: Análisis Geoespacial del Mundo Real (*Capítulos 11-14*)

Explora cuatro estudios de caso completos usando conjuntos de datos reales a gran escala:

- **Capítulo 11: Análisis del Inventario Nacional de Humedales de EE.UU.:** Realiza análisis ambientales en los 50 estados, procesando millones de polígonos de humedales.
- **Capítulo 12: Análisis de Huellas de Edificios Globales:** Analiza datos urbanos usando huellas de edificios globales de Overture Maps.
- **Capítulo 13: Análisis de Datos de Viajes en Taxi de NYC:** Descubre patrones espaciotemporales de cientos de millones de viajes en taxi, revelando información sobre la movilidad urbana.
- **Capítulo 14: Desarrollo de Dashboards Interactivos con Voilà y Solara:** Construye y despliega aplicaciones web que hacen tus análisis accesibles para los interesados.

Al final de la Parte III, tendrás proyectos dignos de portafolio que demuestran tus capacidades avanzadas de análisis espacial.

Temas Transversales a lo Largo del Libro

- **Rendimiento a Escala:** Técnicas que son efectivas ya sea manejando miles o millones de características espaciales.
- **Flujos de Trabajo Nativos de la Nube:** Procesa datos directamente desde S3 e intégrate sin problemas con pilas de datos modernas.
- **Análisis Reproducible:** Código y métodos compartibles que pueden ser versionados y desplegados en producción.
- **Desafíos de Datos del Mundo Real:** Aborda problemas de proyección, valores faltantes y preocupaciones de calidad de datos.
- **Patrones de Integración:** Combina DuckDB con el ecosistema geoespacial más amplio de Python para una funcionalidad mejorada.

Qué Hace Diferente a Este Libro

A diferencia de las discusiones teóricas o tutoriales específicos de herramientas, este libro enfatiza *resolver problemas reales*. Cada técnica está arraigada en desafíos analíticos reales, demostrada con conjuntos de datos reales, y explicada en términos claros de cuándo y por qué usarla.

Cómo Aprovechar al Máximo Este Libro

Para maximizar tu experiencia de aprendizaje con este libro, considera las siguientes recomendaciones:

Configura un Entorno de Desarrollo Adecuado: Instala Python y las bibliotecas requeridas como se describe en el Capítulo 1. Un entorno bien configurado te ahorrará tiempo y frustración a lo largo de tu viaje de aprendizaje. Considera usar conda o uv para gestionar tus paquetes de Python, ya que esto simplifica la instalación de bibliotecas geoespaciales.

Sigue los Ejemplos de Código: Este libro está diseñado para ser interactivo. No solo leas el código; escríbelo, ejecútalo y experimenta con modificaciones. La comprensión viene a través de la práctica, y cada ejemplo desarrolla habilidades que necesitarás más adelante.

Trabaja los Ejercicios: Cada capítulo incluye ejercicios diseñados para reforzar los conceptos que has aprendido. Estos no son extras opcionales; son una parte integral del proceso de aprendizaje. Comienza con los ejercicios guiados, luego desafíate con tus propios proyectos.

Usa Datos Reales: Aunque el libro proporciona conjuntos de datos para ejemplos y ejercicios, intenta aplicar las técnicas a datos de tu propio campo o intereses. Esto te ayudará a entender cómo los conceptos se aplican a escenarios del mundo real y desarrollar confianza en tus habilidades.

Construye Proyectos: A medida que progresas a través del libro, considera trabajar en un proyecto personal que te interese. Esto podría ser analizar datos de tu investigación, crear mapas para tu comunidad, o resolver un problema que has encontrado en tu trabajo.

Sé Paciente Contigo Mismo: La programación puede ser frustrante, especialmente cuando estás aprendiendo. Espera encontrar errores, pasar tiempo depurando y ocasionalmente sentirte atascado. Esto es normal y parte del proceso de aprendizaje. Toma descansos cuando sea necesario, y recuerda que la experiencia se desarrolla gradualmente a través de la práctica constante. Si te quedas atascado, no dudes en pedir ayuda en el repositorio de GitHub del libro.

Sigue Practicando: Las habilidades en este libro requieren práctica regular para mantenerse y desarrollarse. Reserva tiempo regularmente para trabajar en proyectos de programación geoespacial, aunque sean pequeños.

Convenciones Usadas en Este Libro

Este libro usa varias convenciones para ayudarte a navegar el contenido y entender los ejemplos de código:

Formato de Código: Todo el código Python aparece en fuente monoespaciada dentro de bloques de código. Cuando el código aparece dentro del texto regular, está formateado `así`. Los nombres de archivos y directorios también están formateados en fuente monoespaciada.

Ejemplos de Código: La mayoría de los ejemplos de código son completos y ejecutables. Incluyen comentarios explicando los conceptos clave y técnicas que se demuestran. Los números de línea pueden incluirse para referencia en el texto acompañante.

```
# Este es un ejemplo de un bloque de código
import leafmap
m = leafmap.Map()
m.add_basemap("OpenTopoMap") # agregar un mapa base al mapa
m
```

Guía de Estilo SQL: Para consistencia y legibilidad, los ejemplos SQL siguen estos patrones:

- **Palabras clave en MAYÚSCULAS:** `SELECT`, `FROM`, `WHERE`, `JOIN`
- **Los nombres de funciones preservan mayúsculas y minúsculas:** `ST_Area()`, `read_csv_auto()`
- **Nombres de tablas y columnas en minúsculas:** `cities`, `population`
- **Indentación para legibilidad:** Las consultas multilínea están formateadas para claridad

```
SELECT name, ST_Area(geometry) as area
FROM neighborhoods
WHERE borough = 'Manhattan'
ORDER BY area DESC;
```

Instrucciones de Línea de Comandos: Los comandos a ingresar en la línea de comandos o terminal se muestran con un símbolo `$` (no escribas el símbolo `$` mismo):

```
$ pip install leafmap
$ python script.py
```

Descarga de los Ejemplos de Código

Todos los ejemplos de código, conjuntos de datos y materiales complementarios para este libro están disponibles gratuitamente en GitHub:

<https://github.com/giswqs/duckdb-spatial>

Para descargar los materiales, puedes usar uno de los siguientes métodos:

- **Clonar el repositorio** (si tienes Git instalado):

```
$ git clone https://github.com/giswqs/duckdb-spatial.git
```

- **Descargar como ZIP** (si prefieres no usar Git):
 - Visita la página del repositorio en GitHub
 - Haz clic en el botón verde **Code**
 - Selecciona **Download ZIP**
 - Extrae los archivos a tu ubicación preferida
- **Navegar archivos individuales** en línea a través de la interfaz de GitHub si solo necesitas ejemplos específicos

El repositorio se actualiza regularmente con correcciones, mejoras y ejemplos adicionales. Revisa periódicamente para actualizaciones, o **sigue** el repositorio en GitHub para ser notificado de cambios.

Si encuentras errores en el código o tienes sugerencias para mejoras, por favor abre un issue o envía un pull request en GitHub. Las contribuciones de la comunidad ayudan a hacer este recurso mejor para todos.

Tutoriales en Video y Recursos Complementarios

Complementando el contenido escrito, este libro está respaldado por una serie completa de tutoriales en video que recorren conceptos clave y proporcionan ejemplos adicionales:

<https://tinyurl.com/duckdb-spatial-videos>

Los videos están diseñados para complementar, no reemplazar, el material escrito. Son particularmente útiles para:

- Aprendizajes visuales que se benefician de ver el código siendo escrito y ejecutado
- Entender conceptos complejos a través de múltiples explicaciones
- Aprender sobre el flujo de trabajo de desarrollo y mejores prácticas
- Ver cómo abordar problemas y depurar issues

La lista de reproducción está organizada para seguir la estructura del libro. Puedes verlos en orden a medida que progresas a través del libro, o saltar a temas específicos según sea necesario.

Los videos fueron creados en el otoño de 2023 cuando estaba enseñando el curso [Spatial Data Management](https://geog-414.gishub.org)¹ en la Universidad de Tennessee. Aunque el curso ha concluido, los videos siguen siendo relevantes y pueden usarse como referencias para el libro. Se agregarán videos adicionales en el futuro.

¹<https://geog-414.gishub.org>

Comunidad y Retroalimentación

Agradezco la retroalimentación, preguntas y sugerencias de los lectores. Tu aporte ayuda a mejorar el libro y lo hace más útil para la comunidad de programación geoespacial.

Para preguntas y discusiones relacionadas con el libro:

- GitHub Issues: <https://github.com/giswqs/duckdb-spatial/issues>
- GitHub Discussions: <https://github.com/giswqs/duckdb-spatial/discussions>

Tipos de retroalimentación que son particularmente útiles:

- Errores o explicaciones poco claras en el texto o código
- Sugerencias para ejemplos adicionales o casos de uso
- Ideas para nuevos temas o capítulos
- Reportes de problemas de compatibilidad con diferentes sistemas operativos o versiones de bibliotecas
- Historias de éxito de cómo has aplicado las técnicas del libro

Agradecimientos

Este libro existe gracias a las contribuciones de muchas personas y la comunidad más amplia de código abierto geoespacial.

Primero, quiero agradecer al **equipo de desarrollo de DuckDB** por crear un sistema de base de datos tan excepcional. Su compromiso con el rendimiento, la simplicidad y los principios de código abierto ha hecho que el análisis espacial sea accesible para una audiencia mucho más amplia. Un reconocimiento especial a los miembros del equipo que han contribuido a las capacidades espaciales de DuckDB.

Estoy agradecido con mis **colegas y estudiantes** de la Universidad de Tennessee que han proporcionado retroalimentación, probado ejemplos y ayudado a refinar el contenido a través del curso **Spatial Data Management**. Sus preguntas y perspectivas han hecho este libro mucho más sólido.

La **comunidad de código abierto geoespacial** merece un reconocimiento particular. Proyectos como GDAL, GeoPandas, Shapely y muchos otros forman la base que hace posible el análisis espacial moderno. El espíritu colaborativo de esta comunidad continúa inspirando mi trabajo.

Gracias a los **primeros lectores** que proporcionaron retroalimentación sobre los borradores de los capítulos y ayudaron a identificar áreas que necesitaban aclaración o mejora. Sus diversas perspectivas (desde profesionales experimentados de SIG hasta recién llegados a la ciencia de datos) ayudaron a asegurar que este libro sirva a su audiencia prevista.

Quiero reconocer a **mi familia** por su paciencia y apoyo durante las muchas noches y fines de semana dedicados a escribir. Su comprensión y aliento hicieron posible este proyecto.

Finalmente, gracias a **tú**, el lector, por tu interés en el análisis de datos espaciales y las herramientas de código abierto. Son profesionales como tú quienes impulsan la innovación y hacen del campo geoespacial un lugar tan emocionante para trabajar.

Si este libro te ayuda a resolver problemas espaciales más efectivamente, entonces todo el esfuerzo ha valido la pena.

Acerca del Autor

El Dr. Qiusheng Wu es Profesor Asociado en el Departamento de Geografía y Sostenibilidad de la Universidad de Tennessee, Knoxville. También es Amazon Scholar. La investigación del Dr. Wu se centra

en avanzar el análisis geoespacial de código abierto a través de la computación en la nube y GeoAI. Es el creador y mantenedor de varios paquetes de Python de código abierto ampliamente utilizados, incluyendo [Geemap](https://geemap.org)², [Leafmap](https://leafmap.org)³, [SAMGeo](https://samgeo.gishub.org)⁴, y [GeoAI](https://opengeoai.org)⁵, que integran plataformas geoespaciales basadas en la nube con análisis y visualización impulsados por IA. El trabajo del Dr. Wu conecta la teledetección, la observación de la Tierra y la inteligencia artificial para hacer que los datos geoespaciales a gran escala sean más accesibles, reproducibles e inteligentes para investigadores, educadores y profesionales en todo el mundo. Sus proyectos de código abierto se pueden encontrar en GitHub en <https://github.com/opengeos>.

Licencia y Derechos de Autor

Este libro abraza los principios de la ciencia abierta y la educación abierta. Para apoyar la transparencia, el aprendizaje y la reutilización, los **ejemplos de código** en este libro se publican bajo una licencia [Creative Commons Atribución 4.0 Internacional \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/). Esto significa que eres libre de copiar, modificar y distribuir el código, incluso para propósitos comerciales, siempre que se otorgue el crédito apropiado.

Por favor atribuye el uso del código citando el libro o enlazando al repositorio de GitHub:

Wu, Q. (2025). *Spatial Data Management with DuckDB: From SQL Basics to Advanced Geospatial Analytics*. Independently published. PDF edition ISBN 979-8993859705; Print edition ISBN 979-8274710572. <https://duckdb.gishub.org>

Mientras que el código está disponible libremente, el **texto, figuras e imágenes** en este libro están **protegidos por derechos de autor** por el autor y no pueden ser reproducidos, redistribuidos o modificados sin permiso explícito. Esto incluye todo el contenido escrito, diagramas personalizados y visualizaciones incrustadas a menos que se indique lo contrario.

Si deseas reutilizar o adaptar cualquier material que no sea código del libro (por ejemplo, para enseñanza, presentaciones o publicaciones), por favor contacta al autor para solicitar permiso.

Este enfoque de licencia dual ayuda a equilibrar el acceso abierto a los materiales de aprendizaje con la protección del trabajo creativo original. Gracias por respetar estos términos y apoyar a la comunidad geoespacial de código abierto.

²<https://geemap.org>

³<https://leafmap.org>

⁴<https://samgeo.gishub.org>

⁵<https://opengeoai.org>