

SPATIAL DATA MANAGEMENT WITH DUCKDB

From SQL Basics to Advanced
Geospatial Analytics



Qiusheng Wu

Spatial Data Management with DuckDB

From SQL Basics to Advanced Geospatial
Analytics

Qiusheng Wu
2025

Contents

Preface	1
Introduction	3
Who This Book Is For	3
What This Book Covers	5
Getting the Most Out of This Book	6
Conventions Used in This Book	7
Downloading the Code Examples	7
Video Tutorials and Supplementary Resources	8
Community and Feedback	8
Acknowledgments	9
About the Author	9
Licensing and Copyright	10
I: DuckDB Foundations	11
1. Getting Started with DuckDB	13
1.1. Introduction	13
1.2. Learning Objectives	13
1.3. What Makes DuckDB Different from Traditional Databases	13
1.4. When (and When Not) to Use DuckDB for Spatial Work	15
1.5. Installing DuckDB CLI and Running Your First Query	16
1.6. Installing the DuckDB Python Client	18
1.7. Installing Visual Studio Code	21
1.8. Using the DuckDB UI	24
1.9. Installing DBeaver SQL IDE	25
1.10. Key Takeaways	29
1.11. Exercises	30
2. Essential SQL for Spatial Analysis	33
2.1. Introduction	33
2.2. Learning Objectives	33
2.3. Sample Datasets	34
2.4. Setting Up the Environment	34
2.5. Connecting To DuckDB	35
2.6. Install Extensions	35
2.7. Reading CSV Files from URLs	36
2.8. Creating Tables for Better Performance	37
2.9. The SQL SELECT Statement	38
2.10. Filtering Data with the WHERE Clause	45
2.11. Pattern Matching with LIKE	47
2.12. The IN Operator	48
2.13. The BETWEEN Operator	49
2.14. Combining Data with SQL Joins	50
2.15. Query Plans and Performance	56
2.16. Aggregating Data for Summary Statistics	56
2.17. Conditional Statements	60
2.18. Saving Results	61
2.19. Advanced SQL Features for Spatial Analysis	62
2.20. SQL Comments and Documentation	66

2.21. Key Takeaways	67
2.22. Exercises	68
3. DuckDB Python Integration	72
3.1. Introduction	72
3.2. Learning Objectives	72
3.3. Sample Datasets	73
3.4. Installation and Setup	74
3.5. Installing and Loading Extensions	75
3.6. Reading Data from Multiple Sources	76
3.7. Seamless Integration with Pandas DataFrames	79
3.8. Polars Interoperability	82
3.9. Result Conversion and Output Formats	82
3.10. Writing Data To Disk	85
3.11. Persistent Storage and Database Files	85
3.12. Prepared Statements and Parameters	88
3.13. Key Takeaways	88
3.14. Exercises	90
II: Spatial Data Operations	95
4. Loading Spatial Data Formats	97
4.1. Introduction	97
4.2. Learning Objectives	98
4.3. Sample Datasets	98
4.4. Installation and Setup	99
4.5. Installing and Loading Extensions	100
4.6. Downloading Sample Data	101
4.7. Loading CSV Files with Coordinates	101
4.8. Loading JSON Files	104
4.9. Querying Pandas DataFrames Directly	106
4.10. Loading Parquet Files for Performance	108
4.11. Loading GeoJSON Files with Spatial Geometries	110
4.12. Loading Shapefiles into Modern Workflows	116
4.13. Loading GeoParquet for Cloud-Native Spatial Analysis	118
4.14. Data Loading Performance Strategies	120
4.15. Troubleshooting Common Data Loading Issues	121
4.16. Key Takeaways	121
4.17. Exercises	123
5. Exporting and Converting Spatial Data	126
5.1. Introduction	126
5.2. Learning Objectives	127
5.3. Sample Datasets	128
5.4. Installation and Setup	128
5.5. Installing and Loading Extensions	129
5.6. Loading Sample Data	130
5.7. Exporting to Pandas DataFrames	132
5.8. Exporting to CSV Files	136
5.9. Exporting to JSON Files	140
5.10. Exporting to Excel Files	142
5.11. Exporting to Parquet Files	145

5.12.	Exporting to GeoJSON Format	147
5.13.	Exporting to Shapefile Format	151
5.14.	Exporting to GeoPackage Format	152
5.15.	Key Takeaways	154
5.16.	Exercises	155
6.	<i>Geometry Operations and Functions</i>	160
6.1.	Introduction	160
6.2.	Learning Objectives	160
6.3.	Sample Datasets	161
6.4.	Installation and Setup	162
6.5.	Connecting to DuckDB and Loading Extensions	163
6.6.	Understanding Geometry Types	165
6.7.	Working with Points	167
6.8.	Working with Linestrings	171
6.9.	Working with Polygons	172
6.10.	Working with Collections	174
6.11.	Visualizing NYC Spatial Data	175
6.12.	Geometry Processing	178
6.13.	Geometry Validity and Robustness	181
6.14.	Key Takeaways	183
6.15.	Exercises	185
7.	<i>Spatial Queries and Relationships</i>	189
7.1.	Introduction	189
7.2.	Learning Objectives	190
7.3.	Sample Datasets	190
7.4.	Installation and Setup	191
7.5.	Connecting to DuckDB and Loading Extensions	191
7.6.	Understanding Spatial Relationships	192
7.7.	Testing Geometric Identity	193
7.8.	Topological Relationships	195
7.9.	Distance-Based Relationships	203
7.10.	Filtering by Proximity Threshold	205
7.11.	Nearest Neighbor Queries	206
7.12.	Key Takeaways	207
7.13.	Exercises	208
8.	<i>Advanced Spatial Joins</i>	211
8.1.	Introduction	211
8.2.	Learning Objectives	211
8.3.	Sample Datasets	212
8.4.	Installation	213
8.5.	Library Import and Configuration	213
8.6.	Connecting to DuckDB	213
8.7.	Intersection Joins	214
8.8.	Distance Within Joins	219
8.9.	Advanced Join	223
8.10.	Coordinate System Transformations	225
8.11.	Spatial Relationship Functions Reference	227
8.12.	Key Takeaways	228

8.13. Exercises	230
9. <i>Interactive Data Visualization</i>	234
9.1. Introduction	234
9.2. Learning Objectives	234
9.3. Sample Datasets	235
9.4. Installation and Setup	235
9.5. Downloading Sample Data	236
9.6. Connecting to DuckDB and Loading Extensions	236
9.7. Visualizing Point Data	237
9.8. Visualizing Line Data	240
9.9. Visualizing Polygon Data	243
9.10. Key Takeaways	248
9.11. Exercises	250
10. <i>Working with Vector Tiles and PMTiles</i>	254
10.1. Introduction	254
10.2. Learning Objectives	255
10.3. Sample Datasets	256
10.4. Installation and Setup	256
10.5. Visualizing Vector Tiles Directly from Files	257
10.6. Converting Vector Data to PMTiles	261
10.7. Visualizing PMTiles	263
10.8. Key Takeaways	272
10.9. Exercises	274
III: Real-World Geospatial Analytics	279
11. <i>Analyzing US National Wetlands Inventory</i>	281
11.1. Introduction	281
11.2. Learning Objectives	282
11.3. Datasets Used in This Chapter	282
11.4. Understanding the Dataset Source	285
11.5. Accessing Wetland Data with DuckDB	287
11.6. Visualizing Wetland Distributions	290
11.7. National-Scale Wetland Analysis	294
11.8. Key Takeaways	306
11.9. Exercises	308
12. <i>Analyzing Global Building Footprints</i>	311
12.1. Introduction	311
12.2. Learning Objectives	312
12.3. About the Dataset	312
12.4. Installation and Setup	315
12.5. Installing and Loading Extensions	315
12.6. Exploring Available Data	316
12.7. Regional Analysis with Bounding Box Filtering	319
12.8. Multi-Region Comparison	328
12.9. Data Source Filtering and Quality Assessment	333
12.10. Grid-Based Spatial Aggregation	338
12.11. Aggregating Building Data with H3 Hexagonal Grids	343
12.12. Key Takeaways	353
12.13. Exercises	355

13. Analyzing NYC Taxi Data	358
13.1. Introduction	358
13.2. Learning Objectives	359
13.3. About the Dataset	359
13.4. Installation	361
13.5. Library Import	362
13.6. Installing and Loading Extensions	362
13.7. Loading Taxi Data	362
13.8. Temporal Analysis	366
13.9. Loading Taxi Zone Lookup Data	373
13.10. Spatial Analysis	376
13.11. Trip Flow Analysis	383
13.12. Payment and Economic Analysis	387
13.13. Passenger Behavior Analysis	391
13.14. Multi-Month Analysis	392
13.15. Visualization	395
13.16. Performance Optimization Tips	401
13.17. Key Takeaways	402
13.18. Exercises	404
14. Developing Interactive Dashboards with Voilà and Solara	406
14.1. Introduction	406
14.2. Learning Objectives	407
14.3. Installing Voilà and Solara	408
14.4. Introduction to Hugging Face Spaces	408
14.5. Creating a Basic Voilà Application	409
14.6. Creating an Advanced Web Application with Solara	419
14.7. Key Takeaways	427
14.8. Exercises	427

Preface

Introduction

In an increasingly data-driven world, the ability to effectively manage and analyze spatial information has become essential. From urban planning and environmental monitoring to logistics and personalized location-based services, geospatial data forms the backbone of numerous applications that influence our daily lives. However, working with spatial data has often been viewed as a specialized and complex domain, requiring intricate tools and a steep learning curve.

Enter DuckDB.

DuckDB is an innovative analytical database designed for efficiency and user-friendliness. As an in-process OLAP (Online Analytical Processing) database, it operates directly within your application, eliminating the need for separate server deployments and cumbersome configurations. This embedded nature, coupled with its column-oriented architecture and vectorized execution engine, makes DuckDB exceptionally fast for analytical queries, even with large datasets. Initially gaining popularity for its general-purpose data processing capabilities, DuckDB's rapidly expanding ecosystem (especially its extensions for spatial data) represents a transformative shift.

This book, ***"Spatial Data Management with DuckDB: From SQL Basics to Advanced Geospatial Analytics,"*** aims to demystify geospatial data and showcase how DuckDB empowers everyone (from data analysts and scientists to developers and GIS professionals) to leverage its capabilities with unprecedented simplicity and speed. We believe that robust spatial analytics should be accessible, not confined to costly specialized software or complex programming languages. With DuckDB, that accessibility becomes a reality.

Our journey begins with the fundamental concepts of spatial data and its representation, establishing a solid foundation in SQL for working with points, lines, and polygons. As we advance, you will discover how DuckDB's native geospatial features (enhanced by its PostGIS-compatible extension) enable sophisticated operations like spatial joins, buffering, and nearest neighbor searches through elegant SQL queries. We will explore various real-world applications, demonstrating how to load, transform, analyze, and visualize spatial datasets, empowering you to extract meaningful insights from geographic information.

Whether you're looking to integrate spatial analysis into your data pipelines, perform quick ad-hoc geospatial queries, or develop interactive location-aware applications, this book will serve as your comprehensive guide. We will cover topics ranging from setting up your DuckDB environment and importing diverse spatial file formats (like Shapefiles, GeoJSON, and GeoParquet) to executing complex analytical tasks and integrating with visualization tools.

Our aim is not just to teach you syntax but to cultivate an understanding of why these tools and techniques are powerful. By the end of this book, you will be proficient in using DuckDB as your go-to engine for spatial data management and analysis, unlocking new possibilities for your projects and empowering you to make informed, spatially-aware decisions.

Join us as we delve into the exciting intersection of DuckDB's analytical capabilities and the rich world of geospatial data. The future of accessible spatial analytics is here, and it runs on DuckDB.

Who This Book Is For

This book is designed for anyone grappling with the complexities of modern spatial data analysis. If you've ever spent hours waiting for a spatial join to finish, struggled to load large geographic datasets into memory, or wished for a more straightforward way to combine SQL's power with spatial operations, this book is for you.

You'll Find the Most Value If You Are

A GIS Professional frustrated by the limitations of desktop software when handling large datasets. You're familiar with QGIS or ArcGIS, but you need to analyze millions of features, process extensive GPS tracks, or integrate spatial analysis into automated workflows.

A Data Scientist or Analyst who frequently encounters location data. You're comfortable with Python and pandas, but spatial data often feels like a mystery. You want to incorporate geographic dimensions into your analyses without diving into complex GIS software.

A Software Developer building applications that incorporate spatial features. You need fast spatial queries, wish to avoid heavy database infrastructure, and prefer working with familiar SQL over specialized spatial libraries.

A Researcher or Academic in fields like geography, environmental science, or urban planning. Your research involves large spatial datasets, and you require reproducible, scalable analysis methods that can adapt to growing data volumes.

A Business Intelligence Professional dealing with location-based business data. Whether it's store locations, delivery routes, customer territories, or real estate portfolios, you need to merge business metrics with spatial insights.

Essential Prerequisites

You should be comfortable with:

- **Python programming:** understanding variables, functions, and how to import libraries (expertise is not required)
- **Data analysis concepts:** filtering records, aggregating data, and joining tables
- **SQL fundamentals:** familiarity with SELECT, WHERE, and GROUP BY clauses (we'll cover the spatial aspects)
- **Basics of spatial data:** understanding that data has a location (latitude/longitude, projections)

Helpful Background (But Not Required)

- Experience with pandas, GeoPandas, or Jupyter notebooks
- Prior exposure to databases or data warehouses
- Familiarity with GIS software (QGIS, ArcGIS, PostGIS)
- Knowledge of spatial file formats (GeoJSON, Shapefiles, Parquet)

If You're New to Python Programming

If you're new to geospatial Python programming, the following book provides an excellent introduction to both foundational GIS concepts and Python programming:

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>

What This Book Covers

This book offers a structured journey from SQL basics to advanced geospatial analytics, equipping you with practical skills through real-world examples. Each chapter progresses from simple queries to complex spatial analyses, building your expertise in modern geospatial data management.

Part I: DuckDB Foundations (*Chapters 1-3*)

Master the essential concepts that underpin all subsequent content:

- **Chapter 1: Getting Started with DuckDB:** Installation, initial queries, and insights into how DuckDB revolutionizes spatial analytics.
- **Chapter 2: Essential SQL for Spatial Analysis:** Key SQL patterns for filtering, aggregating, joining, and optimizing queries tailored for spatial data.
- **Chapter 3: DuckDB Python Integration:** Combine SQL's power with pandas' flexibility to craft a seamless spatial analysis workflow.

By the end of Part I, you'll confidently query spatial datasets and integrate DuckDB into any Python-based analysis pipeline.

Part II: Spatial Data Operations (*Chapters 4-10*)

Dive into the core spatial toolkit, covering everything from data loading to advanced analytics:

- **Chapter 4: Loading Spatial Data Formats:** Import various formats including CSV coordinates, GeoJSON from APIs, massive Shapefiles, and cloud-hosted GeoParquet.
- **Chapter 5: Exporting and Converting Spatial Data:** Transform your results into any format your stakeholders require.
- **Chapter 6: Geometry Operations and Functions:** Create, measure, and transform spatial features utilizing SQL functions.
- **Chapter 7: Spatial Queries and Relationships:** Master spatial equivalents of table joins: containment, intersection, and proximity.
- **Chapter 8: Advanced Spatial Joins:** Combine datasets by location rather than IDs, which is the essence of spatial analysis.
- **Chapter 9: Interactive Data Visualization:** Generate compelling maps and charts that effectively communicate your data's spatial narratives.
- **Chapter 10: Working with Vector Tiles and PMTiles:** Deploy interactive maps capable of handling millions of features smoothly.

By the end of Part II, you'll be adept at managing any spatial data format, executing complex operations, and creating professional visualizations.

Part III: Real-World Geospatial Analytics (*Chapters 11-14*)

Explore four comprehensive case studies using large-scale, real datasets:

- **Chapter 11: Analyzing the US National Wetlands Inventory:** Conduct environmental analyses across all 50 states, processing millions of wetland polygons.
- **Chapter 12: Analyzing Global Building Footprints:** Analyze urban data using global building footprints from Overture Maps.
- **Chapter 13: Analyzing NYC Taxi Trip Data:** Uncover spatiotemporal patterns from hundreds of millions of taxi trips, revealing insights into urban mobility.

- **Chapter 14: Developing Interactive Dashboards with Voilà and Solara:** Build and deploy web applications that make your analyses accessible to stakeholders.

By the end of Part III, you'll have portfolio-worthy projects showcasing your advanced spatial analysis capabilities.

Cross-Cutting Themes Throughout

- **Performance at Scale:** Techniques that are effective whether handling thousands or millions of spatial features.
- **Cloud-Native Workflows:** Process data directly from S3 and integrate seamlessly with modern data stacks.
- **Reproducible Analysis:** Shareable code and methods that can be version-controlled and deployed in production.
- **Real-World Data Challenges:** Tackle projection issues, missing values, and data quality concerns.
- **Integration Patterns:** Combine DuckDB with the broader Python geospatial ecosystem for enhanced functionality.

What Makes This Book Different

Unlike theoretical discussions or tool-specific tutorials, this book emphasizes *solving real problems*. Each technique is rooted in actual analytical challenges, demonstrated with real datasets, and explained in clear terms of when and why to use it.

Getting the Most Out of This Book

To maximize your learning experience with this book, consider the following recommendations:

Set Up a Proper Development Environment: Install Python and the required libraries as described in Chapter 1. A well-configured environment will save you time and frustration throughout your learning journey. Consider using conda or uv to manage your Python packages, as this simplifies the installation of geospatial libraries.

Follow Along with Code Examples: This book is designed to be interactive. Don't just read the code; type it out, run it, and experiment with modifications. Understanding comes through practice, and each example builds skills you'll need later.

Work Through the Exercises: Each chapter includes exercises designed to reinforce the concepts you've learned. These are not optional extras; they are an integral part of the learning process. Start with the guided exercises, then challenge yourself with your own projects.

Use Real Data: While the book provides datasets for examples and exercises, try applying the techniques to data from your own field or interests. This will help you understand how the concepts apply to real-world scenarios and build confidence in your abilities.

Build Projects: As you progress through the book, consider working on a personal project that interests you. This could be analyzing data from your research, creating maps for your community, or solving a problem you've encountered in your work.

Be Patient with Yourself: Programming can be frustrating, especially when you're learning. Expect to encounter errors, spend time debugging, and occasionally feel stuck. This is normal and part of the

learning process. Take breaks when needed, and remember that expertise develops gradually through consistent practice. If you get stuck, don't hesitate to ask for help on the book's GitHub repository.

Keep Practicing: The skills in this book require regular practice to maintain and develop. Set aside time regularly to work on geospatial programming projects, even if they're small ones.

Conventions Used in This Book

This book uses several conventions to help you navigate the content and understand the code examples:

Code Formatting: All Python code appears in monospaced font within code blocks. When code appears within regular text, it is formatted like this. File and directory names are also formatted in monospaced font.

Code Examples: Most code examples are complete and runnable. They include comments explaining the key concepts and techniques being demonstrated. Line numbers may be included for reference in the accompanying text.

```
# This is an example of a code block
import leafmap
m = leafmap.Map()
m.add_basemap("OpenTopoMap") # add a basemap to the map
m
```

SQL Style Guide: For consistency and readability, SQL examples follow these patterns:

- **Keywords in UPPERCASE:** SELECT, FROM, WHERE, JOIN
- **Function names preserve case:** ST_Area(), read_csv_auto()
- **Table and column names in lowercase:** cities, population
- **Indentation for readability:** Multi-line queries are formatted for clarity

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

Command Line Instructions: Commands to be entered at the command line or terminal are shown with a \$ prompt (don't type the \$ symbol itself):

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

Downloading the Code Examples

All code examples, datasets, and supplementary materials for this book are freely available on GitHub:

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

To download the materials, you can use one of the following methods:

- **Clone the repository** (if you have Git installed):

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

- **Download as ZIP** (if you prefer not to use Git):
 - Visit the GitHub repository page
 - Click the green **Code** button
 - Select **Download ZIP**
 - Extract the files to your preferred location
- **Browse individual files** online through the GitHub interface if you only need specific examples

The repository is regularly updated with corrections, improvements, and additional examples. Check back periodically for updates, or **watch** the repository on GitHub to be notified of changes.

If you find errors in the code or have suggestions for improvements, please open an issue or submit a pull request on GitHub. Community contributions help make this resource better for everyone.

Video Tutorials and Supplementary Resources

Complementing the written content, this book is supported by a comprehensive series of video tutorials that walk through key concepts and provide additional examples:

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

The videos are designed to complement, not replace, the written material. They're particularly helpful for:

- Visual learners who benefit from seeing code being written and executed
- Understanding complex concepts through multiple explanations
- Learning about the development workflow and best practices
- Seeing how to approach problems and debug issues

The playlist is organized to follow the book's structure. You can watch them in order as you progress through the book, or jump to specific topics as needed.

The videos were created in Fall 2023 when I was teaching the [Spatial Data Management](#)¹ course at the University of Tennessee. Although the course has concluded, the videos remain relevant and can be used as references for the book. Additional videos will be added in the future.

Community and Feedback

I welcome feedback, questions, and suggestions from readers. Your input helps improve the book and makes it more useful for the geospatial programming community.

For book-related questions and discussions:

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

Types of feedback that are particularly helpful:

- Errors or unclear explanations in the text or code
- Suggestions for additional examples or use cases

¹<https://geog-414.gishub.org>

- Ideas for new topics or chapters
- Reports of compatibility issues with different operating systems or library versions
- Success stories of how you’ve applied the techniques from the book

Acknowledgments

This book exists thanks to the contributions of many individuals and the broader open-source geospatial community.

First, I want to thank the **DuckDB development team** for creating such an exceptional database system. Their commitment to performance, simplicity, and open-source principles has made spatial analytics accessible to a much broader audience. Special recognition goes to the team members who’ve contributed to DuckDB’s spatial capabilities.

I’m grateful to my **colleagues and students** at the University of Tennessee who have provided feedback, tested examples, and helped refine the content through the **Spatial Data Management** course. Their questions and insights have made this book much stronger.

The **open-source geospatial community** deserves particular recognition. Projects like GDAL, GeoPandas, Shapely, and countless others form the foundation that makes modern spatial analysis possible. The collaborative spirit of this community continues to inspire my work.

Thanks to the **early readers** who provided feedback on draft chapters and helped identify areas that needed clarification or improvement. Their diverse perspectives (from seasoned GIS professionals to data science newcomers) helped ensure this book serves its intended audience.

I want to acknowledge **my family** for their patience and support during the many evenings and weekends spent writing. Their understanding and encouragement made this project possible.

Finally, thanks to **you**, the reader, for your interest in spatial data analysis and open-source tools. It’s practitioners like you who drive innovation and make the geospatial field such an exciting place to work.

If this book helps you solve spatial problems more effectively, then all the effort has been worthwhile.

About the Author

Dr. Qiusheng Wu is an Associate Professor in the Department of Geography & Sustainability at the University of Tennessee, Knoxville. He is also an Amazon Scholar. Dr. Wu’s research focuses on advancing open-source geospatial analytics through cloud computing and GeoAI. He is the creator and maintainer of several widely used open-source Python packages, including [Geemap](https://geemap.org)², [Leafmap](https://leafmap.org)³, [SAMGeo](https://samgeo.gishub.org)⁴, and [GeoAI](https://opengeoai.org)⁵, which integrate cloud-based geospatial platforms with AI-powered analysis and visualization. Dr. Wu’s work bridges remote sensing, Earth observation, and artificial intelligence to make large-scale geospatial data more accessible, reproducible, and intelligent for researchers, educators, and practitioners worldwide. His open-source projects can be found on GitHub at <https://github.com/opengeos>.

²<https://geemap.org>

³<https://leafmap.org>

⁴<https://samgeo.gishub.org>

⁵<https://opengeoai.org>

Licensing and Copyright

This book embraces the principles of open science and open education. To support transparency, learning, and reuse, the **code examples** in this book are released under a [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/) license. This means you are free to copy, modify, and distribute the code, even for commercial purposes, as long as appropriate credit is given.

Please attribute code usage by citing the book or linking to the GitHub repository:

Wu, Q. (2025). *Spatial Data Management with DuckDB: From SQL Basics to Advanced Geospatial Analytics*. <https://duckdb.gishub.org>

While the code is freely available, the **text, figures, and images** in this book are **copyrighted** by the author and may not be reproduced, redistributed, or modified without explicit permission. This includes all written content, custom diagrams, and embedded visualizations unless otherwise noted.

If you wish to reuse or adapt any non-code material from the book (for example, for teaching, presentations, or publications), please contact the author to request permission.

This dual licensing approach helps balance open access to learning materials with the protection of original creative work. Thank you for respecting these terms and supporting the open-source geospatial community.

Revolutionize Your Geospatial Data Analysis with DuckDB

Transform how you work with geospatial data using **DuckDB**, the revolutionary analytical database that makes spatial analysis faster, simpler, and more accessible than ever before. This comprehensive guide takes you from SQL fundamentals to advanced geospatial analytics, empowering you to tackle real-world spatial challenges and create scalable analytics pipelines.

What You Will Learn

- Master DuckDB fundamentals
- Work with diverse data formats
- Perform spatial operations
- Conduct advanced spatial analysis
- Create interactive visualizations
- Handle large-scale datasets
- Integrate with the Python ecosystem
- Apply real-world solutions
- Build production workflows



Scan the QR code to visit the
book website and download
code examples from
<https://duckdb.gishub.org>

