
Awesome DuckDB

A curated list of awesome DuckDB libraries, tools and resources.

DuckDB is an in-process SQL OLAP database management system.

Chat with this page

You can chat with this page's content on HuggingChat.

Contents

- Resources
- Client APIs
- Tools Powered by DuckDB
 - Web Clients
- Libraries Powered by DuckDB
- SQL Clients and IDE that Support DuckDB
- Projects Powered by DuckDB
- Integrations
- Extensions
- Media
 - Talks
 - Podcasts
 - Blog Posts
- Contribute

Resources

- Official Documentation - Official DuckDB documentation.
 - Official Blog - Official DuckDB blog.
 - DuckDB Clients - Client APIs for DuckDB.
 - DuckDB Documentation PDF - The DuckDB documentation as a single PDF file.
- docker-duckdb - Docker image for DuckDB CLI.

-
- DuckDB setup - GitHub Action to install DuckDB in CI.
 - Serverless DuckDB over S3 - Running DuckDB over a data lake on S3 using lambda.
 - DuckDB snippets - Collection of snippets curated by MotherDuck.
 - DuckDB tldr page - DuckDB's entry in tldr pages, available in CLI via the `tldr duckdb` command.
 - DuckDB AWS Lambda layer - Run DuckDB in AWS Lambda functions.
 - Compatible DuckDB Extensions for AWS Lambda - Extensions specifically compiled for the AWS Lambda runtime (GLIBC 2.26).
 - Serverless DuckDB as API - Use DuckDB as API with Amazon API Gateway and AWS Lambda.
 - Serverless Parquet Repartitioner - Use DuckDB to repartition data in S3-based Data Lakes.
 - Observable notebooks - Notebooks using DuckDB on the Observable data visualization platform.
 - duckdb-nf - Example uses of DuckDB with Nextflow.

Client APIs

- C
- C++
- CLI
- Go
- Julia
- Node.js
- Python
- R
- Rust
- Swift
- TypeScript
- Wasm
- ADBC
- ODBC
- .NET
- Common Lisp

Tools Powered by DuckDB

- Rill Developer - Tool for effortlessly transforming data sets into powerful, opinionated dashboards using SQL.

-
- Ibis Project - A DataFrame API for interacting with DuckDB (and other compute engines).
 - MotherDuck - Serverless data warehouse powered by DuckDB.
 - Boiling Data - Serverless data analytics overlay on top of S3 Data Lakes.
 - Hex Dataframe SQL - Hex's Dataframe SQL cells are powered by DuckDB.
 - Mode - Mode uses DuckDB for their in-memory data engine.
 - VulcanSQL - DuckDB can be used as a caching layer or a data connector in VulcanSQL, a Data API framework for data folks to create REST APIs by writing SQL templates.
 - Tad - A fast, free, cross-platform tabular data viewer application powered by DuckDB.
 - Honeycomb Maps - A browser-based geospatial analysis tool leveraging DuckDB Wasm.
 - Bauplan - A serverless data transformation platform for data lakes.
 - Excalichart.com - A fast, free dashboard for exploring your data.
 - Malloy - Malloy is an experimental language for describing data relationships and transformations. Malloy connects to BigQuery and Postgres, and natively supports DuckDB.
 - Evidence - Generate reports using SQL and markdown. The DuckDB connector allows querying across DuckDB, csv, parquet and json.
 - Latitude - Latitude uses DuckDB to power data snapshots. Drop a CSV file and query it with SQL at the speed of light.
 - Census - Census's dataset diffing for incremental syncs is powered by DuckDB.
 - Huey - Blazing-fast & intuitive pivot tables on .parquet, .csv, .json files and .duckdb tables in the browser based on DuckDB WASM. Open source (MIT). Zero install!
 - Parquet Explorer - Visual Studio Code extension for exploring Parquet files with SQL, powered by DuckDB.
 - Iceburst - The real-time data lake for monitoring & security.
 - DQOps - Data quality platform for data engineers, data quality teams and data operations.
 - DatalakeStudio - Load, explore, transform your datasets and expose them via API. Integration with external APIs, S3, PostgreSQL and ChatGPT.
 - Spice.ai - A unified SQL query interface and portable runtime to locally materialize (using an embedded DuckDB), accelerate, and query datasets from any database, data warehouse, or data lake.

Web Clients

- Online DuckDB Shell - Online DuckDB shell powered by WebAssembly.
- SQL Workbench - DuckDB-WASM based SQL Workbench for running queries on local or remote data, being able to show data as tables or visually as graphs, and sharing queries via URLs.
- Sekuel Playground - Query your local parquet, csv, json. Your data will not be sent out of the device you are using.

-
- CSVFiddle - Free tool to explore and share insights from CSV files using SQL. Import data, write SQL, then instantly share it with anyone.
 - Codapi - Embed executable code snippets directly into your product documentation, online course or blog post.
 - QuackDB - Open-source online DuckDB SQL playground and editor.
 - WhatTheDuck - WhatTheDuck is an open-source web application built on DuckDB. It allows users to upload CSV files, store them in tables, and perform SQL queries on the data.

Libraries Powered by DuckDB

- Mosaic - An extensible framework for linking databases and interactive views.
- Fugue - A unified interface for distributed computing. Fugue executes SQL, Python, Pandas, and Polars code on Spark, Dask and Ray without any rewrites.
- Splink - A free Python library for fast, accurate data deduplication and record linkage.
- Simple-data-analysis - Easy-to-use and high-performance JavaScript library for data analysis.
- duckdb_fdw - DuckDB Foreign Data Wrapper for PostgreSQL.
- @jetblack/duckdb-react - A context manager for React and duckdb-wasm.
- QuackOSM - A Python library for downloading and transforming raw OpenStreetMap data into GeoParquet files.

SQL Clients and IDE that Support DuckDB

- Harlequin - The DuckDB IDE for your terminal. (GitHub).
- qStudio - A free SQL tool specialized for data analysts. It runs on every operating system and allows easy browsing of tables and charting of results.
- DuckDB SQL Tools - Free DuckDB SQL Tools for VS Code IDE. Premium version available with advanced features.
- VSCode SQLTools (Free) - Free open-source VSCode extension to query and explore your DuckDB databases with latest DuckDB support.
- DBeaver - DBeaver is a universal database access and development tool that can be used to connect almost any type of database.
- DataGrip - Paid SQL IDE by JetBrains that supports many different database technologies, including DuckDB.
- Duckling - A fast viewer for CSV/Parquet files and DuckDB/SQLite, based on Tauri.
- rsql - CLI for DuckDB, LibSQL, MariaDB, MySQL, PostgreSQL, SQLite3 and SQL Server.

Projects Powered by DuckDB

- NBA Monte Carlo - Monte Carlo simulation of the NBA season, leveraging Meltano, dbt, DuckDB and Evidence.
- Datadex - Open source and local friendly data platform to collaborate on Open Data using DuckDB, Dagster, dbt, and Quarto.
- [endoflife.date](#) database - Daily dumps of endoflife.date data.
- [transfermarkt-datasets](#) - Curated football datasets from Transfermarkt.
- nodbi - NoSQL Database Connector for R, providing a common API across Elasticsearch, CouchDB, MongoDB, SQLite, PostgreSQL, and DuckDB.

Integrations

- dbt-duckdb - DuckDB dbt adapter.
- data load tool - DuckDB destination - Extract and load data from APIs to DuckDB using dlt.
- target-duckdb - Load data to DuckDB based on Singer spec.
- Airbyte DuckDB destination - Load data to DuckDB with Airbyte.
- Kestra DuckDB plugin - Run queries with DuckDB to schedule data transformations and process automations, and run event-driven anomaly detection pipelines.
- Arrow Flight SQL Server - An example implementation of the Arrow Flight SQL protocol what runs in a client-server setup with DuckDB or SQLite as backends.
- SQLFlow - Enables SQL-based stream-processing, powered by DuckDB.
- nf-sqldb - This plugin provides support for interacting with SQL databases in Nextflow scripts.

Extensions

- Official Extensions - Official DuckDB extensions.
 - [arrow](#) - A zero-copy data integration between Apache Arrow and DuckDB.
 - [duckdb_iceberg](#) - For reading Iceberg tables.
 - [duckdb_aws](#) - For handling AWS credentials.
 - [duckdb_azure](#) - For using the Azure Blob storage.
 - [mysql](#) - To read from and write to MySQL databases.
 - [postgres](#) - To read from and write to PostgreSQL databases.
 - [spatial](#) - Enables geospatial processing.
 - [sqlite](#) - To read from and write to SQLite databases.
- DuckDB Extension Radar - Repository that contains DuckDB extensions on GitHub. Refreshed daily.

-
- Deltatable - DuckDB extension for reading Delta tables.
 - Lance - Integrate Lance (modern columnar data format for ML implemented in Rust) with DuckDB.
 - PRQL - Run PRQL commands directly within DuckDB.
 - duckdb-engine - SQLAlchemy driver for DuckDB.
 - Scrooge-McDuck - DuckDB extension to support a set of aggregation functions and data scanners on financial data.
 - duckdb-pytables - DuckDB extension to allow running SQL on arbitrary data sources.
 - ODBC Scanner DuckDB Extension - DuckDB extension to read data directly from databases supporting the ODBC interface.
 - duckdb-jfr-extension - DuckDB extension to read JFR (Java Flight Recorder) files directly.
 - QDuckDB - Plugin for reading DuckDB spatial tables in QGIS software.
 - ERPL - DuckDB SAP connector using RFC, ODP, or BICS.
 - duckdb-extension-template-zig - A Zig & Nix toolkit template for building extensions against multiple versions of DuckDB using Zig, C or C++.
 - DuckPGQ - DuckDB extension for graph workloads that supports the SQL/PGQ standard.
 - h3-duckdb - Adds support for the H3 discrete global grid system.

Media

Talks

- DuckCon #4 playlist
- State of the Duck @ DuckCon #4 - Hannes Mühleisen and Mark Raasveldt.
- In-Process Analytical Data Management with DuckDB @ PyData Amsterdam - Hannes Mühleisen.
- DuckDB: The Power of a Data Warehouse in your Python Process @ PyData Yerevan - Gábor Szárnyas.
- DuckDB: Bringing analytical SQL directly to your Python shell @ EuroPython - Pedro Holanda.
- DuckDB keynote @ Data + AI Summit 2023 - Hannes Mühleisen.
- State of the Duck @ DuckCon #3 - Hannes Mühleisen and Mark Raasveldt.
- DuckCon #3 playlist
- DuckDB: Bringing Analytical SQL Directly To Your Python Shell @ FOSDEM - Pedro Holanda.
- State of the Duck @ DuckCon #2 - Hannes Mühleisen & Mark Raasveldt.
- DuckDB Extensions @ DuckCon - Pedro Holanda & Sam Ansmink.
- Developing Systems in Academia: The Good, the Bad, and the not-so-Ugly Duckling @ CIDR - Hannes Mühleisen.
- DuckDB An Embeddable Analytical Database @ FOSDEM - Hannes Mühleisen.

-
- DuckDB tutorials playlist by Learn Data with Mark - Mark Needham.
 - DuckDB tutorials playlist by MotherDuck - Mehdi Ouazza (MotherDuck).
 - Nextflow and database uses: powering data engineering, exploring DuckDB, and beyond - Edmund Miller.
 - Why should you care about DuckDB? @ Dublin DuckDB meetup - Mihai Bojin.
 - Exploring Monte Carlo Simulations With DuckDB @ Dublin DuckDB meetup - James McNeill.
 - DuckDB and recommenders : a lightning fast synergy @ Dublin DuckDB meetup - Khalil Muhammad.

Podcasts

- Developer Voices: Implementing Hardware-Friendly Databases - Hannes Mühleisen.
- The Geek Narrator: DuckDB Internals - Mark Raasveldt.
- Software Engineering Daily: DuckDB - Hannes Mühleisen.
- Data Engineering Podcast: Move Your Database To The Data And Speed Up Your Analytics With DuckDB - Hannes Mühleisen.
- The Analytics Engineering Podcast: The Personal Data Warehouse - Jordan Tigani.

Blog Posts

- Modern Data Stack in a Box - Fast, free, and open-source Modern Data Stack deployed on a laptop using the combination of DuckDB, Meltano, dbt, and Apache Superset.
- How to use DuckDB, Motherduck and Kestra for ETL - How DuckDB can transform data, mask sensitive PII information, detect anomalies in event-driven workflows, and streamline reporting use cases.
- DuckDB vs. MotherDuck — how do they compare - What are key differences between them, and when to choose each of these options.
- Building DuckDB Extensions with Zig and Nix - For Nix users and Zig developers familiar with DuckDB looking to extend its capabilities with custom extensions.
- Exploring StarCraft 2 data with Airflow, DuckDB and Streamlit - Example project using DuckDB to persist API data, but also explains how to use DuckDB as a versatile data manipulation tool in data wrangling scripts.
- DuckDB: The Rising Star in the Big Data Landscape
- How to Make a DuckDB Extension for a Table Function? - How to make a DuckDB extension to fetch data from external sources.

Contribute

Contributions welcome! Read the contribution guidelines first.