

Open in app ↗

Medium

 Search

PostgreSQL and DuckDB: Supercharging Ad-Hoc Data Analysis and ETL



Josef Machytka

1 min read · 11 hours ago



Listen



Share



More

On Wednesday, January 8th, 2025, I had the amazing opportunity to present my talk about the DuckDB database online at the “Postgres MeetUp for All” event. This fantastic and inspiring gathering was organized by Elizabeth Garrett Christensen, the Director of Content Marketing at Crunchy Data — “The Trusted Open Source Enterprise PostgreSQL Leader.” Thank you very much for inviting me!

My talk focused on the powerful ways DuckDB complements PostgreSQL, showcasing some of the most important use cases that can be especially beneficial for those working with PostgreSQL. Compared to the version of this talk I presented in October 2024 at the PostgreSQL MeetUp in Prague, this updated presentation dives deeper into new use cases and provides enriched examples, including Python integration.

For anyone already leveraging PostgreSQL for their data workflows, this session demonstrates how DuckDB can supercharge ad-hoc data analysis and streamline ETL processes, making it an indispensable tool in your data toolkit. If you're interested in exploring this further, you'll find a wealth of related insights in my Medium articles. Slides from the talk are [available on academia.edu](#) or [on my GitHub](#).

You can watch the recorded talk here:

Postgres & DuckDB with Josef Machytka



Original banner of this talk for social media:



Postgresql

Duckdb

Etl Process

Data Analysis

[Edit profile](#)

Written by Josef Machytka

53 Followers · 10 Following

I work as PostgreSQL specialist & database reliability engineer at NetApp Deutschland, Open Source Services division.

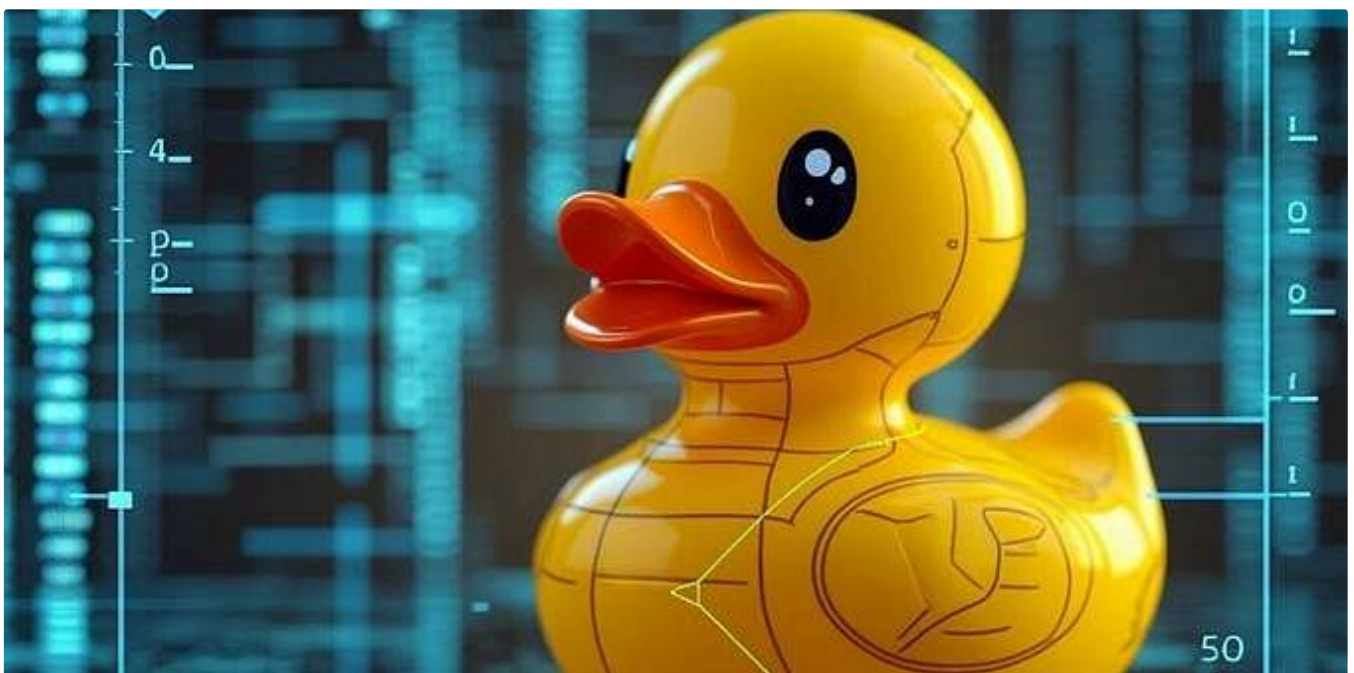
No responses yet



What are your thoughts?

[Respond](#)

More from Josef Machytka





Josef Machytka

DuckDB Database File as a New Standard for Sharing Data?

This is not my original idea; I came across it in an excellent article titled “DuckDB Beyond the Hype” by Alireza Sadeghi. However, it...

Dec 30, 2024



17



2



ate the table

```
TABLE special_data_types (  
  INT AUTO_INCREMENT PRIMARY KEY,  
  me VARCHAR(50) NOT NULL,  
  atus ENUM('active', 'inactive', 'pending') NOT NULL,  
  rmissions SET('read', 'write', 'execute') NOT NULL,  
  all_number TINYINT NOT NULL,  
  dium_number MEDIUMINT NOT NULL,  
  scription TEXT,  
  ta BLOB,  
  eated_at DATE NOT NULL
```

ert 10 rows of data

```
INTO special_data_types (name, status, permissions, small_number, medium_number, description, data, created_at)  
, 'active', 'read,write', 5, 1000, 'Alice description', 'Alice data', '2023-01-01'),  
, 'inactive', 'read', 10, 2000, 'Bob description', 'Bob data', '2023-02-01'),  
, 'pending', 'write,execute', 15, 3000, 'Charlie description', 'Charlie data', '2023-03-01'),  
, 'active', 'read,write,execute', 20, 4000, 'David description', 'David data', '2023-04-01'),  
, 'inactive', 'execute', 25, 5000, 'Eve description', 'Eve data', '2023-05-01'),  
, 'pending', 'read,write', 30, 6000, 'Frank description', 'Frank data', '2023-06-01'),  
, 'active', 'read', 35, 7000, 'Grace description', 'Grace data', '2023-07-01'),  
, 'inactive', 'write,execute', 40, 8000, 'Hank description', 'Hank data', '2023-08-01'),  
, 'pending', 'read,write,execute', 45, 9000, 'Ivy description', 'Ivy data', '2023-09-01'),  
, 'active', 'execute', 50, 10000, 'Jack description', 'Jack data', '2023-10-01');
```



Josef Machytka

DuckDB as a Rudimentary Data Migration Tool

After exploring how to use DuckDB as an intelligent ETL tool for PostgreSQL, and how to extend its ETL capabilities with simple Python...

Nov 30, 2024



26



2



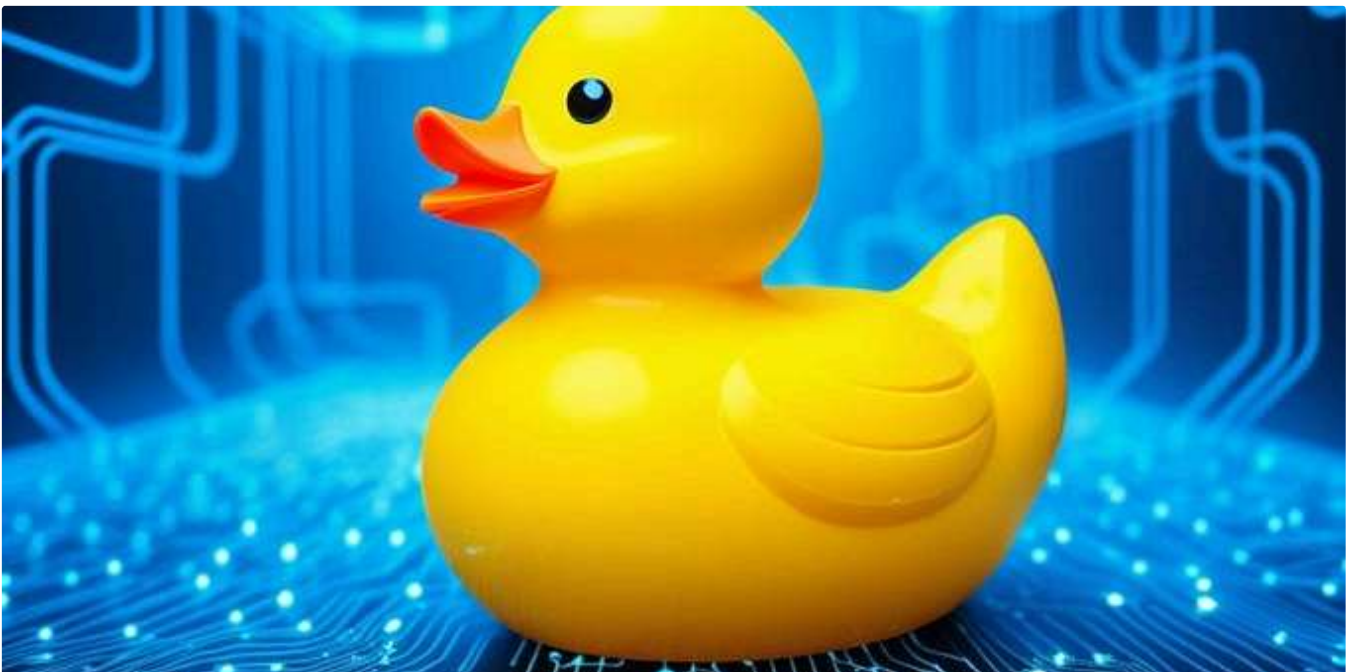


 Josef Machytka

Quick and Easy Data Exports to Parquet Format Using DuckDB

The Parquet format has become almost an industry standard for Data Lakes and Data Lakehouses, thanks to its efficiency and compact storage...

Dec 5, 2024  11



 Josef Machytka

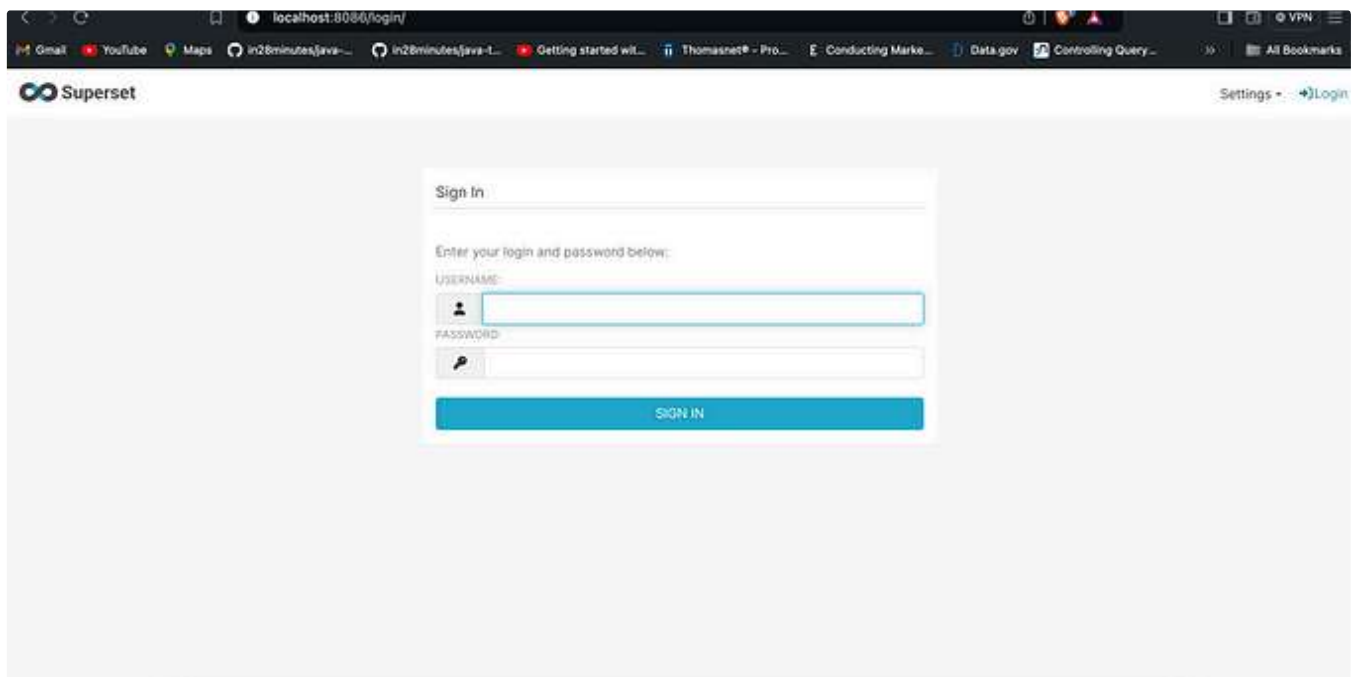
Extending DuckDB ETL Capabilities with Python

DuckDB has recently become my go-to solution for small ETL tasks. It is an exceptional database created by people who quite obviously...

Nov 24, 2024 🖱 26

[See all from Josef Machytka](#)

Recommended from Medium

 Vivek Suresh

DuckDB and Superset Combination to Power Analytics

Since DuckDB has been all the rage in the past 1 year in data engineering circles and also because we have been recommending it for clients...

Sep 17, 2024 🖱 13





 In Dev Genius by Aleksei Aleinikov 

10 Ways to Work with Large Files in Python: Effortlessly Handle Gigabytes of Data!

Handling large text files in Python can feel overwhelming. When files grow into gigabytes, attempting to load them into memory all at once...

Dec 1, 2024  304  3



Lists



Practical Guides to Machine Learning

10 stories · 2138 saves



ChatGPT prompts

51 stories · 2445 saves



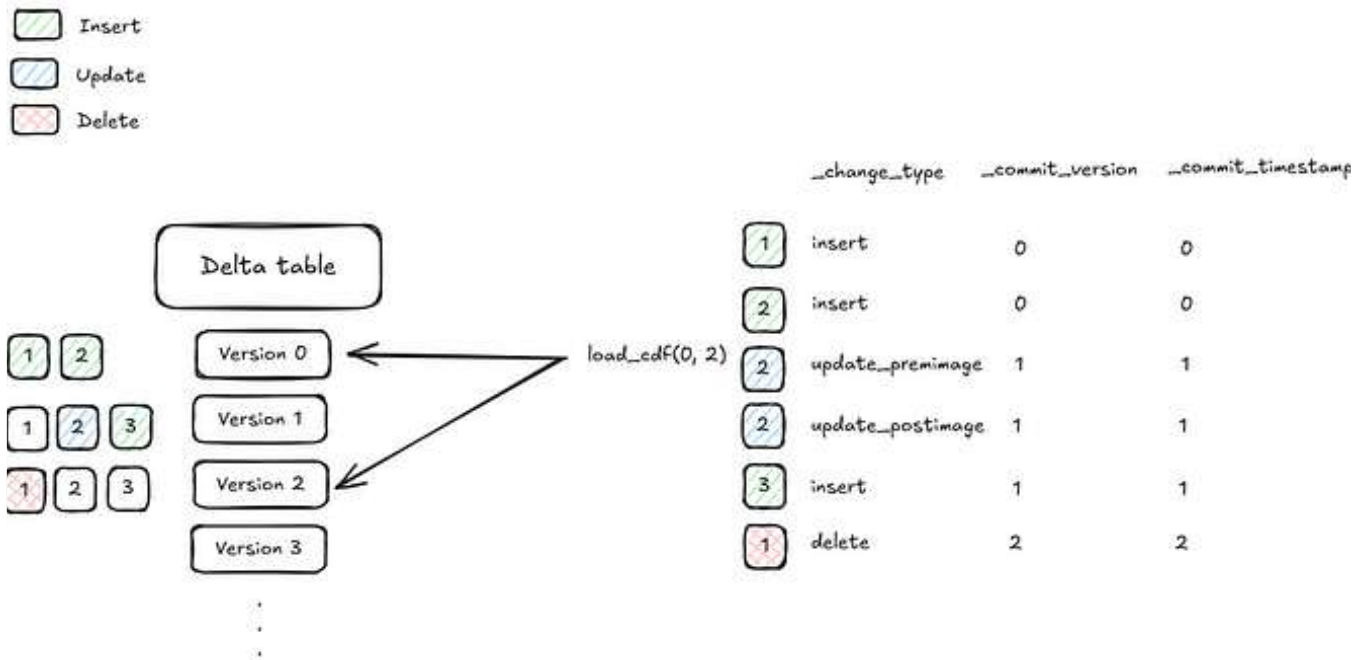
Staff picks


796 stories · 1558 saves



Natural Language Processing

1884 stories · 1530 saves

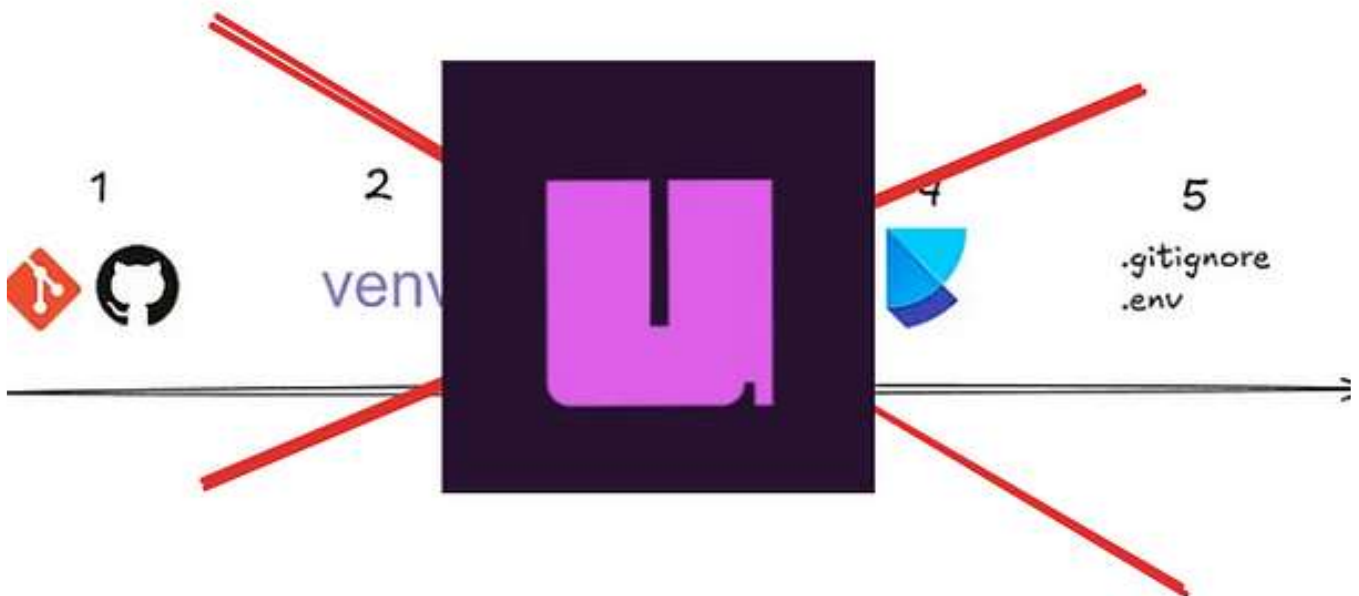



 Pablo Cabeza

Incremental pipelines with Delta and Polars

How to apply watermark pipelines at a small scale with Change Data Feed

Nov 29, 2024  2



 Lorenzo Uriel

In Love With uv

Forget about everything that I teach you in this article!

Dec 4, 2024  115  2





PY In Python in Plain English by Raphael Schols

Frustrated by Manual Reporting? Build a Self-Service Portal with Python

Build Streamlit Portals for Excel and PowerPoint

★ Dec 16, 2024 🖱 255 💬 2



DE In Towards Data Engineering by Sandun Lakshan

10 Best Python Text User Interface (TUI) Libraries for 2025

Text-based user interfaces (TUIs) are a simple way to create interactive applications that run in a terminal. They let users control...



Dec 11, 2024



36



1



See more recommendations