



# PostgreSQL

*powst · gres · kyoo · el*

Neea Rusch      24 August 2022

# Outline

1. Introduction
2. Technical demo
3. Research uses

# Quick introduction

- Object-relational database management system (ORDBMS)
- Free and open-source software (FOSS)
- Developed by PostgreSQL Global Development Group
- Mature, widely-used and popular

## Mature – many decades of history

- 1970s Project **Ingres** developed at UC Berkeley
- 1980s **Postgres** – the improved version of Ingres
- 1993 Development ended at UC Berkeley
- 1994 Open source community adopts **Postgres95**
- 1996 Final renaming to **PostgreSQL**

# Widely-used and popular

Connection libraries available for many programming languages:  
C++, Java, Julia, Python, Node, etc.

Over 158,000 questions on Stack Overflow,  
70% have accepted answers.

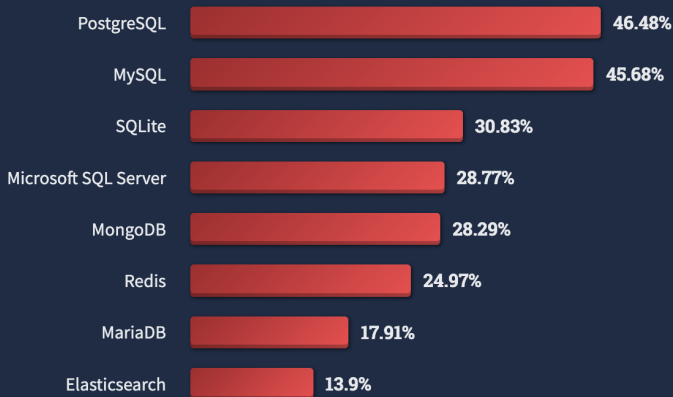
Some notable users:



# Stack Overflow Developer Survey 2022

**Q:** Which database environments have you done extensive development work in over the past year, and which do you want to work in over the next year?

**48,788 professional developers answered:**



# Features

All perks of relations databases

- schemas
- foreign keys
- flexible queries
- ACID transactions
- indices
- view tables
- stored procedures
- triggers, etc.

# Features

Rich native data type support, including:

arbitrary-precision numerics	arrays	boolean
character (text, varchar, char)	bit strings	binary
IPv4 and IPv6 addresses	composite	enum
date/time	JSON	money



# Features

- Extensible by design e.g., GIS, text and image processing, ...
- Possible to process complex data types e.g., geographic data
- Memory-efficient for performance-critical tasks
- Object-oriented features e.g., tables inheritance
- Highly ANSI SQL-compliant
- Flexible full text search

# Suitability assessment

Potential use cases:

- Structured data (can mix different kinds of data)
- Scales up to multiple terabytes of data
- No fee, even for commercial software products

Not ideal for: non-relational data, highly-distributed systems, offline/real-time sync, and excessive for small amounts of data.

# Technical Demo

# Setup basics

- Available for all major operating systems  
Linux, Windows, macOS, BSD, Solaris . . .
- Many front-end interfaces to interact with database  
`psql` is the primary official front-end
- Default user is `postgres` and a password is not required
- Default port 5432

Demo notebook on Google Colab 

# Research uses of PostgreSQL

```
SELECT *  
FROM articles  
WHERE body LIKE '%PostgreSQL%'  
ORDER BY citations  
LIMIT 3
```

# Research uses

The #1 cited (56) paper<sup>1</sup>

- Researchers integrated decision tree data mining techniques into PostgreSQL
- Compared to an alternative tool (Weka), observed approx. 2-4x speedup in analysis times, for 100K - 1M records

---

<sup>1</sup>Viloria, Amelec, et al. *"Integration of data mining techniques to PostgreSQL database manager system."* Procedia Computer Science 155 (2019): 575-580.

# Research uses

The #2 cited (45) paper<sup>2</sup>

- Premise: interest in content-based image retrieval is increasing
- Propose an image-handling extension: PostgreSQL-IE
- Introduce a new image data type that can associate images
- Enable new feature extraction procedures, similarity queries

---

<sup>2</sup>Guliatto, Denise, et al. "PostgreSQL-IE: An image-handling extension for PostgreSQL." *Journal of digital imaging* 22.2 (2009): 149-165.

# Research uses

The #3 cited (33) paper<sup>3</sup>

- Report describes PostgreSQL development team
- Uncovered by analyzing repository history, software artifacts
- Found there is a stable small team of core members
- User momjian (**Bruce Momjian**) is a key contributor

---

<sup>3</sup>German, Daniel M. "A study of the contributors of PostgreSQL." Proceedings of the 2006 international workshop on Mining software repositories. 2006.



# Final remarks

PostgreSQL – flexible, expandable, open source ORDBMS



## References

Perkins, Luc, et al. "Seven Databases in Seven Weeks", 2nd Ed., 2018.

PostgreSQL documentation: <https://www.postgresql.org>

SO developer survey: <https://survey.stackoverflow.co/2022/>

Stackshare.io: <https://stackshare.io/postgresql>

Wikipedia <https://en.wikipedia.org/wiki/PostgreSQL>