

[« Prev](#)[Next »](#)

## Database Weekly

Issue 108 — June 9, 2016

### FEATURED

Be warned, there's a heavy dose of Apache Spark in this week's issue, thanks to this week's Spark Summit bringing lots of news out of the woodwork :-)

**Why Spark Is On Fire: A Conversation with Creator Matei Zaharia** — Apache Spark, the open source framework for distributing computing across clusters of machines, has caught on like wildfire among companies looking for insights from their growing masses of data.

ROBERT HOF

**MongoDB Queries Don't Always Return All Matching Documents** — It's essential to understand some of MongoDB's characteristics before scaling it up significantly in production to get the results you expect.

DAVID GLASSER

**Why Some of The Fastest Growing DBs Are Also The Most Experimental** —

Everyone has heard about MongoDB and Cassandra, but what other databases are making big gains against Oracle and Microsoft?

MATT ASAY

**Try Compose Elasticsearch, Get a Shirt** — Have you been wanting to try Elasticsearch? This month only, get a limited edition t-shirt when you deploy Elasticsearch on Compose. All new accounts are completely free for 30-days.

COMPOSE **SPONSORED**



**Google BigQuery Gets Beta Standard SQL Support** — Till now, you've had to query BigQuery with the SQL-like BQL (BigQuery Query Language) but now full SQL support is coming to the popular data warehouse service.

JORDAN NOVET

**Correlated Subqueries Are Evil and Slow. Or Are They?** — A common myth in SQL is the idea that correlated subqueries are slow, but here's a look at why you should always test for your specific circumstances.

LUKAS EDER

**Planning A Simple Online Store Data Model** — The latest in Vertabelo's long line of common database use case visualizations. This time, how you'd lay out a simple online store's data model.

VERTABELO

**5 Mistakes Beginners Make When Working with Databases** — “There’s a number of bad practices that tend to get picked up when working with databases, here’s a rundown of just a few.”

CRAIG KERSTIENS

**5 Reasons You Should Use SQLite in 2016** — Reposting in case you missed this earlier in the year - SQLite is frequently underestimated, but there are many compelling reasons to use it in various contexts.

CHARLES LEIFER

**Using Spark for Anomaly (Fraud) Detection** — Anomaly detection is a method used to detect outliers in a dataset and take some action, e.g. fraudulent transactions.

MICHAEL VOGIATZIS

## **JOBS**

**Work on a Better Stack!** — On Hired, engineers typically get 5+ job offers in 1 week. Find that new opportunity you've been craving and get access to 3,500+ companies instantly.

HIRED.COM

## **IN BRIEF**

[Microsoft Announces Major Commitment to Apache Spark](#) — Both Spark and R are making their way into most of Microsoft’s big data and analytics offerings.

MICROSOFT **NEWS**

[Apache Spark Adoption by the Numbers](#) — It’s the top open source big data project of the moment.

DATANAM **NEWS**

[Hadoop 3 Poised to Boost Storage Capacity, Resilience with Erasure Coding](#)

DATANAM **NEWS**

[Corporate Giving: EDB’s Contributions to PostgreSQL](#) — Robert Hass outlines some of EDB’s recent contributions to the PostgreSQL ecosystem. Like any other company that supports PostgreSQL, EDB makes technical contributions to the PostgreSQL community in a variety of ways: code; patch review, testing, and commit; external tools; and packaging.

ENTERPRISEDB **NEWS** **SPONSORED**

[Proposing a 'Stream' Data Type for Redis](#)

SALVATORE SANFILIPPO **NEWS**

[Apache Cassandra 3.6 Released](#)

APACHE.ORG **NEWS**

[Spark Makes Inroads into NoSQL Ecosystem](#) — “The primary use case for deploying Spark and NoSQL databases together involves bridging the transactional and analytic divide.”

DATANAM **NEWS**

[Calculating Distance Between Points on a Globe with SQL](#)

PERISCOPE

[Calculating Common Salesforce Reports in SQL](#)

PERISCOPE

## [Packet Capturing MySQL with Rust](#)

PAUL LACROSSE

## [19 Resources for Learning About Database Design Errors](#)

JEFFREY EDISON

## [Postgres 9.6 New Features with Examples \(PDF\)](#) — A thorough 50 page (PDF) guide to the new features.

HEWLETT PACKARD ENTERPRISE

## [Beverly Hills 90210 and ZIP+4: Handling Addresses in Data Models](#)

JEFFREY EDISON

## [Find Redis Memory Leaks with RedisGreen](#) — RedisGreen's new memory map models your database, helping you track down problems and gain new insights.

REDISGREEN **TOOLS** **SPONSORED**

## [miniDBA Desktop: A Free Dashboard Tool for SQL Server](#)

MINIDBA **TOOLS**

## [Introducing the Neo4j 3.0 Apache Spark Connector](#)

MICHAEL HUNGER **TOOLS**

## [rpgffi: Create Postgres Extensions in Rust](#) — A Rust Postgres FFI for building Postgres extensions in the modern, safe systems programming language.

ALEX NEWMAN **CODE**

[« Prev](#)

[Next »](#)

Want to subscribe? Enter your address here

**Subscribe now »**

Easy to unsubscribe at any time. Your e-mail address [is safe](#) — here's [our privacy policy](#).