

[Get started](#)[Open in app](#)

towards
data science

[Follow](#)

569K Followers



You have **2** free member-only stories left this month. [Sign up for Medium and get an extra one](#)

Top 10 Databases to Use in 2021

MySQL, Oracle, PostgreSQL, Microsoft SQL Server, MongoDB, Redis, Elasticsearch, Cassandra, MariaDB, IBM Db2



Md Kamaruzzaman Jan 20 · 21 min read ★

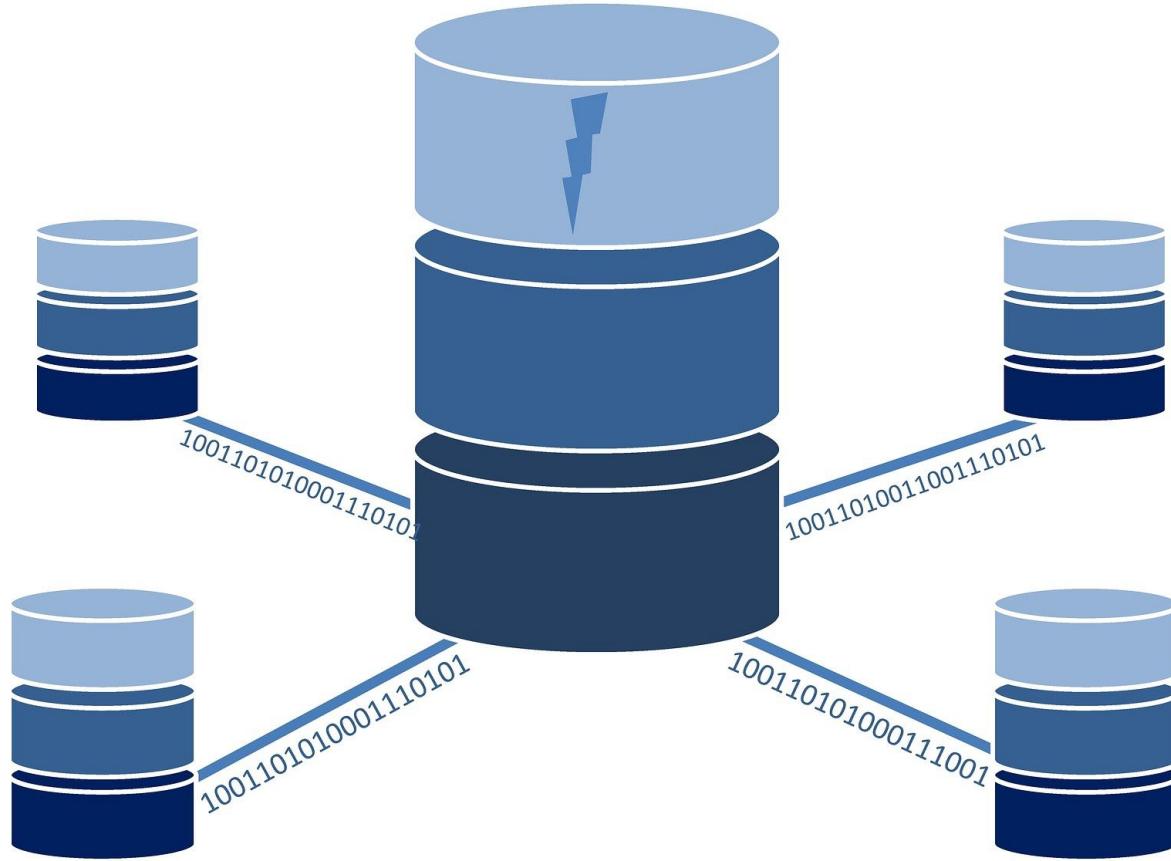


Image by [Tumisu](#) from [Pixabay](#)

Databases are the cornerstone of any Software Applications. You will need one or more databases to develop almost all kind of Software Applications: Web, Enterprise, Embedded Systems, Real-Time Systems, AI, ML, HPC, Blockchain, IoT, and many other applications.

Twenty years ago, choosing a database was much easier. In most cases, you could have chosen one of the Relational Database Systems and you were done. But in modern Software Development, choosing the right database is one of the most challenging tasks.

With the rise of Microservices, Cloud, Distributed Applications, Global Scaling, Semi-Structured Data, Big Data, Fast Data, Low Latency Data: the traditional **SQL** databases are now joined by various **NoSQL**, **NewSQL**, and **Cloud** databases.

There are a whopping **343** databases at present. Here I will shortlist ten databases from them according to the following criteria:

- Key Features
- Popularity
- Trending
- Mainstream uses
- Bright Future

I will discuss the following for each Database:

- An overview, including a brief history.
- 5 Key Features.
- When to use them.
- When not to use them.
- Popularity.
- Trending.
- Database As a Service (DBaaS).
- Alternatives.

... . . .

1. MySQL

In 1995, two Software Engineers, Michael Widenius and David Axmark, created the Open Source Relational Database Management System (RDBMS) MySQL. Since its inception, MySQL quickly became popular in the industry and community for its enterprise-grade features and free, flexible (GPL) community license, and upgraded commercial license. Among the open-source Databases, while PostgreSQL focuses on innovation and advanced features, MySQL focuses on robustness, stability, and maturity.

Today, MySQL is one of the most popular and widely used SQL databases. It is also one of the most used databases in Web Applications. Some of the world's largest Web-Scale

applications (e.g., Facebook, Uber) uses MySQL.

5 Key Features

- Open source RDBMS with two licensing models: free Community Server and proprietary Enterprise Server.
- Offers ACID transactional guarantee (with InnoDB engine). In terms of CAP, it offers immediate Consistency.
- Offers horizontal partitioning (sharding) via its **Shared Nothing MySQL Cluster**. As a result, it offers high availability and high throughput with low latency and near-linear Scaling.
- With its MySQL Cluster, it offers multi-master ACID transactions.
- Multi-model database and supports both structured data (SQL) and semi-structured data (JSON).

When to Use MySQL

- Structured Data (SQL) with an ACID transaction guarantee.
- Horizontal Scalability is a key requirement, especially with Write heavy data.
- Multi-Master ACID transaction is a fundamental requirement.
- Data Security is a key feature.
- A Converged database is required, i.e., one database for both OLTP and OLAP workload.

When not to Use MySQL

- “**Distributed SQL**” is required where millions of transactions should be handled in a globally distributed database.
- Data is extremely relational (e.g., Social Media), i.e., Graph like data.
- Data is Semi-structured, i.e., JSON data with advanced query features.
- Advanced-data Protection (e.g., Masking, Obfuscating, Throttling) is a must.

MySQL As a Service

- [MySQL Database Service](#) (Multi-Cloud, OLTP, and OLAP)
- [ScaleGrid](#) (Horizontal Scaling)
- [Vitess](#) (Horizontal Scaling)
- [Aiven for MySQL](#) (Multi-Cloud)
- [Amazon RDS for MySQL](#)
- [Oracle MySQL Cloud Service](#) (Horizontal Scaling)
- [Azure MySQL Database](#)
- [Google Cloud SQL for MySQL](#)

Alternatives

- [Oracle](#)
- [PostgreSQL](#)
- [Microsoft SQL Server](#)
- [MariaDB](#)
- [IBM Db2](#)
- [Amazon Aurora](#)
- [Google Cloud Spanner](#)

Popularity:

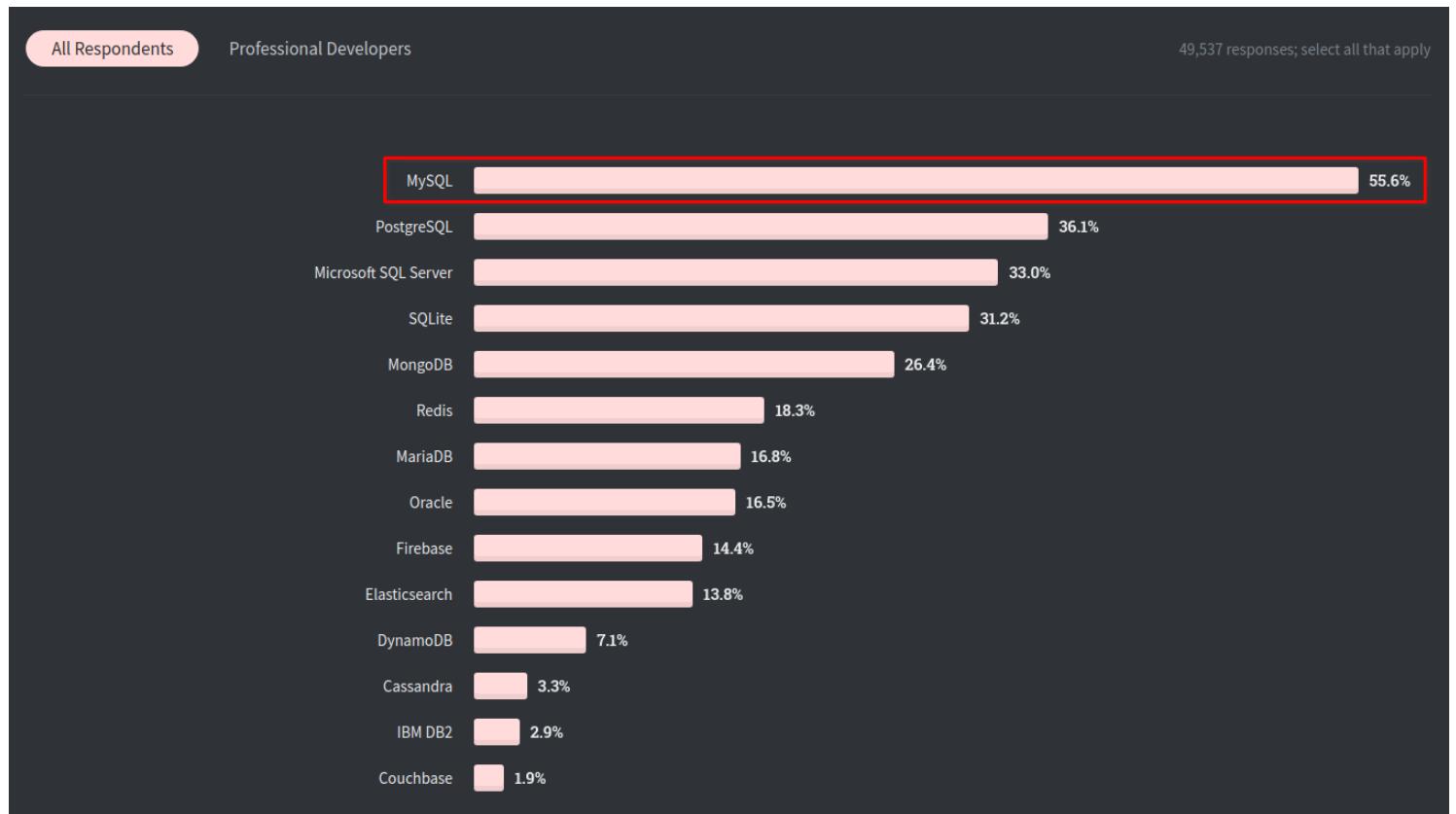
MySQL Community Edition is the most widely used free database in the industry. Also, its commercial version is used extensively in the industry.

The most prominent database ranking site DB-Engines ranked MySQL as the second-highest ranked database, just behind Oracle:



Image by Author (data: [DB-Engines](#))

The Stack Overflow Developer Survey 2020 has placed MySQL as the most popular database for 2020:



Source: [Stack Overflow Developer Survey, 2020](#)

Trending

With 25 years of history, MySQL is not the trendiest database at present. But unlike other older databases, it is still maintaining high traction, as shown below:



Also, Google Trends has shown a comparatively stable graph in the last few years:



2. Oracle

When Edgar F. Codd's published his revolutionary paper "A Relational Model of Data for Large Shared Data Banks" (1970) on the Relational Database Management System (RDBMS), it has completely changed the landscape of database Systems. The paper particularly inspired a young Software Engineer **Larry Ellison** (current CTO of Oracle Corporation). He later created the world's first commercially available RDBMS system Oracle in 1979. Since then, Oracle remained the leading commercial RDMBS System and dominated the Unix and Linux Systems. Over the last 41 years, Oracle has evolved with time and contributed to the RDBMS and the overall database Systems innovations.

Currently, Oracle is the number one commercially supported database and one of the widely used RDBMS overall. Its latest release (21.c) has added many innovative features that will make it an attractive option in the coming years.

5 Key Features

- Proprietary RDBMS.
- Offers ACID transactional guarantee. In terms of CAP, it offers immediate Consistency as a single Server.
- Advanced Multi-Model databases supporting Structured Data (SQL), Semi-Structured Data(JSON, XML), Spatial Data, and RDF Store. Offers multiple access pattern depending on the specific Data Model
- Offers Blockchain Tables.
- Supports both OLTP and OLAP workload.

When to Use Oracle

- If a company wants to have a Converged database or Master Database (One database for OLTP and OLAP).
- Traditional transactional workloads with structured (SQL) data, and when ACID transaction guarantee is a key criterion.
- Blockchain Table is required.

- For Data Warehousing.
- A multi-model database including Geospatial Data type is an essential requirement.

When not to Use Oracle

- If a company wants to save money on a database.
- Multi-Master ACID transaction is a must-have feature.
- Data is Semi-structured, i.e., JSON data with advanced query functions.
- Data is extremely relational (e.g., Social Media), i.e., Graph like data.

Oracle As a Service

- [Oracle Converged Database](#)
- [Amazon RDS for Oracle](#)

Alternatives

- [MySQL](#)
- [PostgreSQL](#)
- [Microsoft SQL Server](#)
- [MariaDB](#)
- [IBM Db2](#)
- [Amazon Aurora](#)
- [Google Cloud Spanner](#)

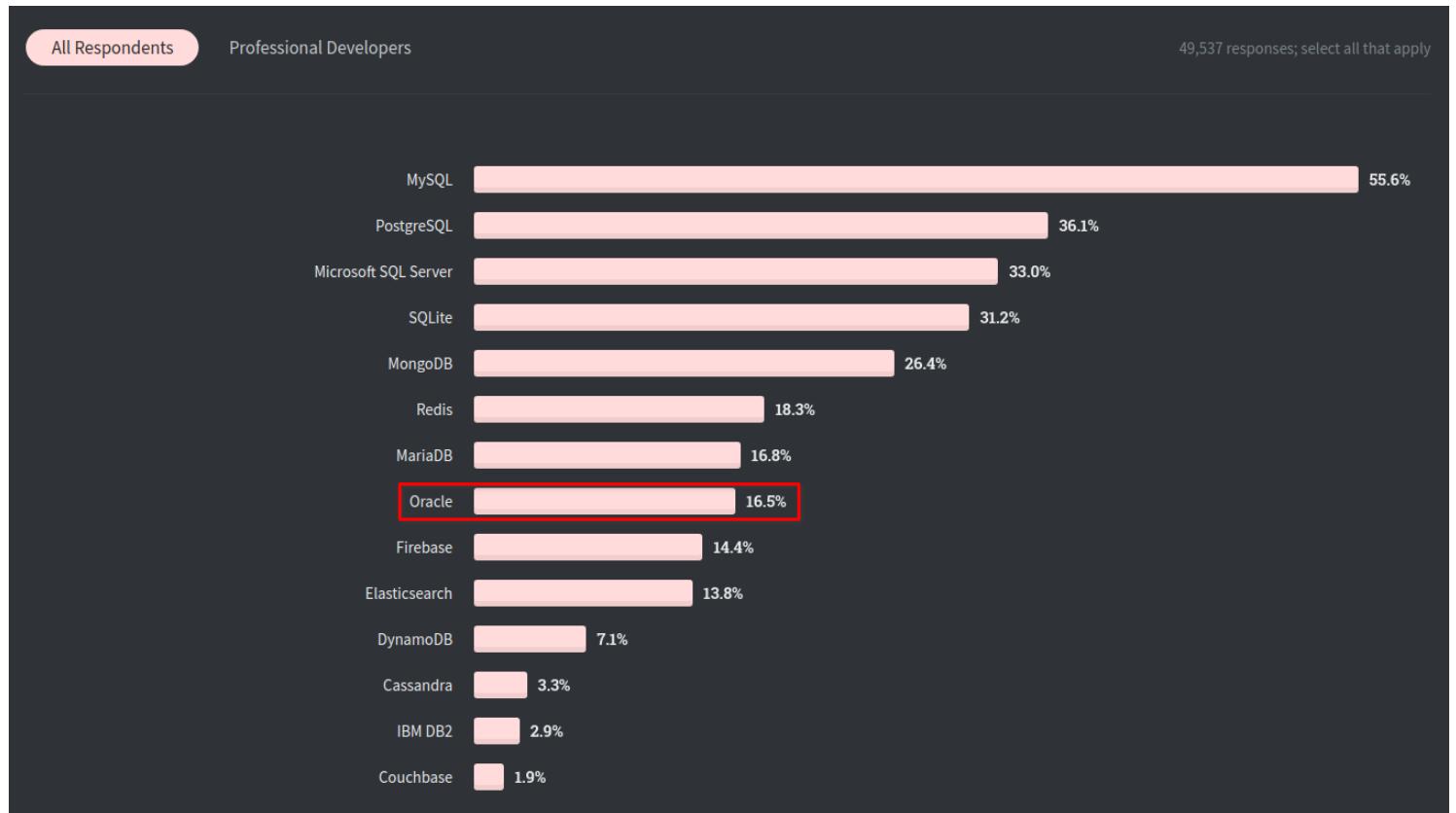
Popularity

For four decades, Oracle is providing a stable, mature, enterprise-grade database to its customer. Although it is facing stiff competition from the open-source SQL databases and NoSQL databases, it is still the most popular database System as ranked by DB-Engines:



Image by Author (Data: [DB-Engines](#))

However, in the Stack Overflow Developer Survey, it was ranked in 8th position as the list is dominated by open-source databases:



Source: [Stack Overflow Developer Survey, 2020](#)

Trending

Oracle probably hit a Jackpot with its newest release. As a result, it has gained lots of traction in the last couple of years. Considering the market is moving to open-source databases, it is quite a feat:



Also, Google trends shows a curiously flat curve for Oracle, which is rare for a commercial database:

Source: [Google Trends](#)

3. PostgreSQL

Since the early 1970s, [UC Berkeley](#) is working to shape the modern Database Management Systems via its ground-breaking database project [Ingres](#). In 1986, the legendary [Michael Stonebraker](#) led the POSTGRES (Post-Ingres) project to tackle the existing database Projects' problems. [PostgreSQL](#) was born as part of the POSTGRES project, which is a Relational Database Management System. Over the past 30 years, PostgreSQL leads the way in Modern database Development, contributing many innovations, and Michael Stonebraker got a [Turing Award](#) in 2014 primarily for his work in PostgreSQL.

Today, PostgreSQL is one of the most used databases. It is also the most Advanced Open Source Relational Database.

5 Key Features

- Open Source RDMBS with a very permissible PostgreSQL license. There are many extensions of PostgreSQL with various advanced features.
- Offers ACID transactional guarantee. In terms of CAP, it offers immediate Consistency as a single Server.
- **Citus Data**, a PostgreSQL extension, offers much-coveted **Distributed SQL** features, i.e., a database with the transactional guarantee of an SQL database and horizontal scaling of NoSQL database.
- It offers more Advanced indexes liked partial Index, Bloom Filters. Also, creating an Index in PostgreSQL is non-blocking.
- It is a feature-rich multi-model database supporting Structured Data (SQL), Semi-Structured Data (JSON, XML), Key-Value, Spatial Data.

When to Use PostgreSQL

- Data is structured (SQL), and ACID transaction guarantee is a key criterion with strict Data integrity.
- When “**Distributed SQL**” is required, where millions of transactions should be handled in a globally distributed database.
- A sophisticated query planner is wanted.
- Advanced Reliability and Disaster Recovery is desired, e.g., PITR, Active Standbys.
- A multi-model database including Geospatial Data type is expected.

When not to Use PostgreSQL

- Multi-Master ACID transaction is a must-have feature.
- Data is Semi-structured, i.e., JSON data with advanced query planers.
- A Converged database (OLTP and OLAP) is required.
- Data is extremely relational (e.g., Social Media), i.e., Graph like data.

PostgreSQL As a Service

- [EDB for PostgreSQL](#) (Horizontal Scaling)
- [CitusData](#) (Distributed SQL)
- [ScaleGrid](#) (Horizontal Scaling)
- [Aiven for PostgreSQL](#) (Multi-Cloud)
- [Amazon RDS for PostgreSQL](#)
- [Azure PostgreSQL Database](#)
- [Google Cloud SQL for PostgreSQL](#)

Alternatives

- [MySQL](#)
- [Oracle](#)
- [Microsoft SQL Server](#)

- [MariaDB](#)
- [IBM Db2](#)
- [Amazon Aurora](#)
- [Google Cloud Spanner](#)

Popularity

PostgreSQL is one of the oldest open-source and free databases. It has tremendous community support and is used in the Community as well as in the industry.

It is one of the most popular databases according to the DB-Engines:



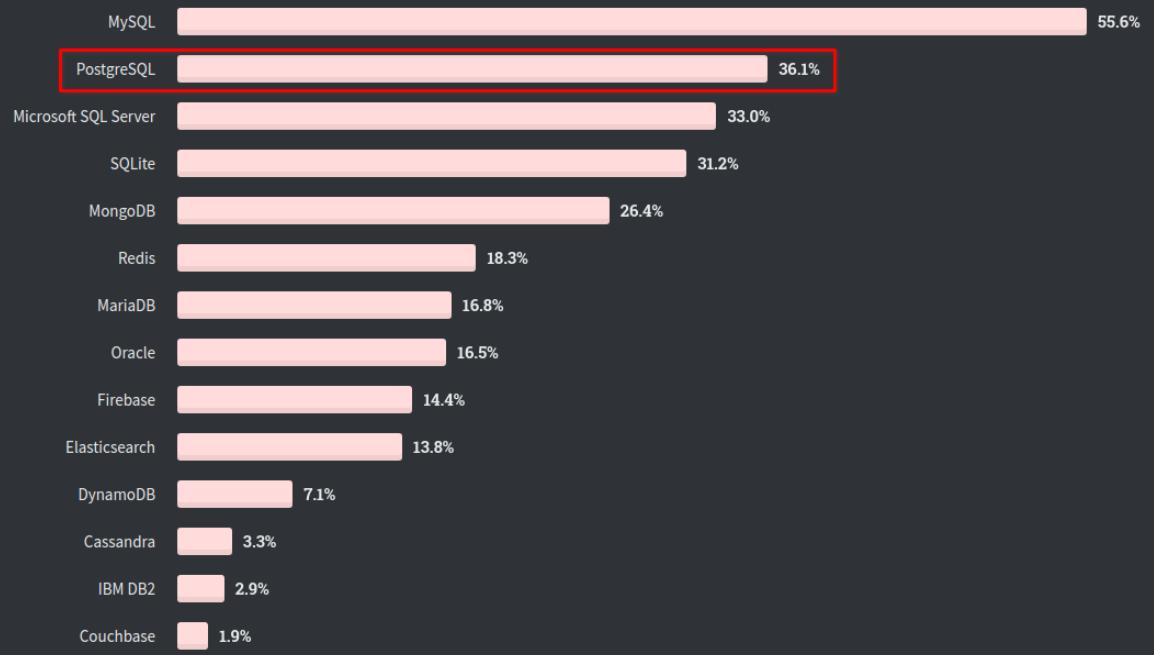
Image by Author (Data: [DB-Engines](#))

According to the Stack Overflow Developer Survey, it was the second most popular database in 2020:

All Respondents

Professional Developers

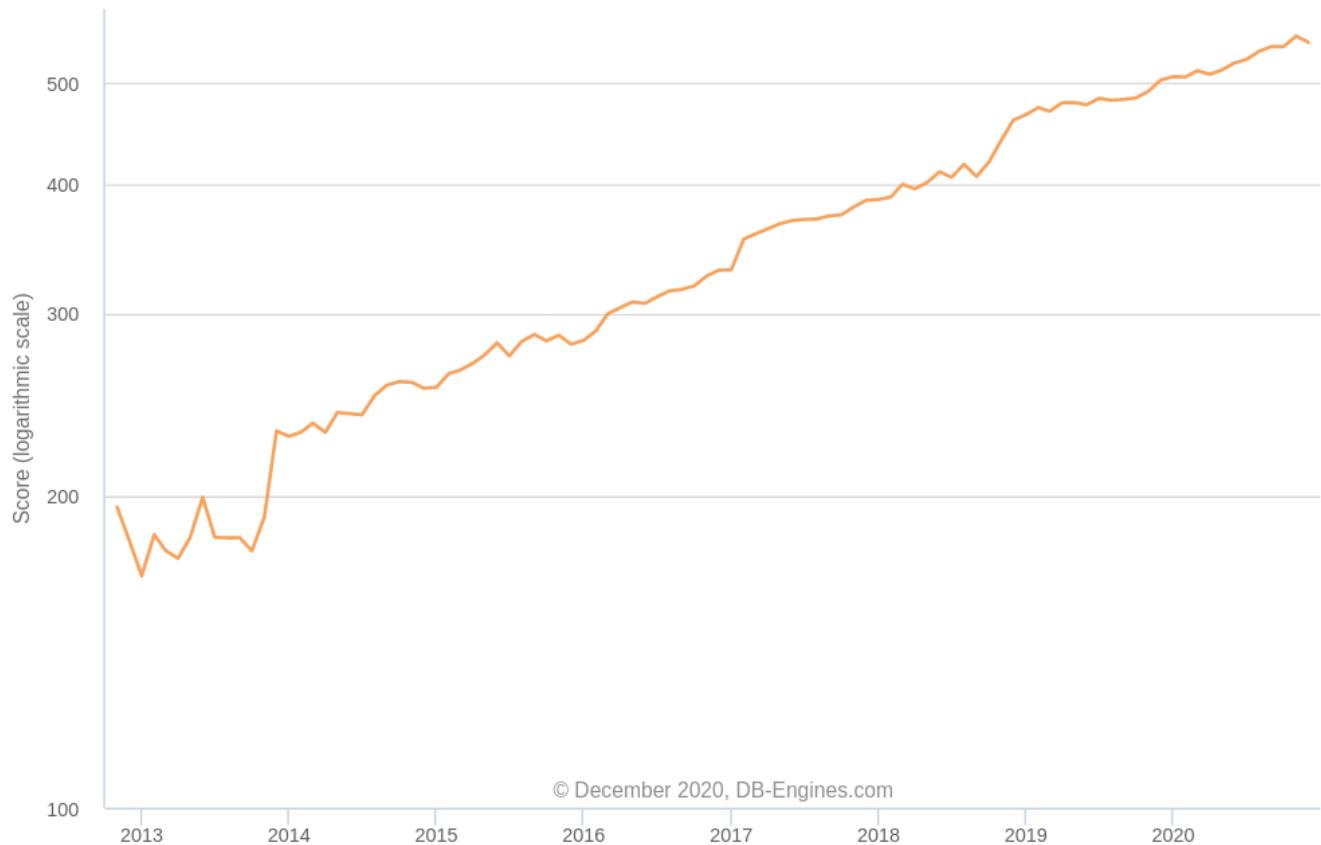
49,537 responses; select all that apply



Source: [Stack Overflow Developer Survey, 2020](#)

Trending

Because of its advanced features and innovations, PostgreSQL is getting lots of hype even after 30 years of its introduction, as shown below:



Source: [DB-Engines](#)

Google Trends also shows an increase in traction for PostgreSQL in recent years, which is an affirmation of increasing interest in PostgreSQL:



Source: [Google Trends](#)

4. Microsoft SQL Server

In 1988, Microsoft joined with Ashton-Tate and Sybase to create an RDBMS as a Sybase SQL Server variant. In 1989, they released the first version of Microsoft SQL Server. Although not as innovative or advanced as others, Microsoft SQL Server has gone through major updates and overhauls over the years. Initially released for IBM OS/2, it was soon released for Windows Systems. In the last three decades, it was the dominating commercial Mid-range database in Windows Systems.

Currently, it is one of the leading commercial database Systems with excellent Tooling support from Microsoft.

5 Key Features

- Proprietary RDBMS with diverse licenses: Developer and Express (free), Standard, and Enterprise (commercial).
- Offers ACID transactional guarantee. In terms of CAP, it offers immediate Consistency as a single Server.
- Supports Server-side scripting via T-SQL, .NET languages, R, Python, and Java.
- It is a multi-Model database supporting Structured Data (SQL), Semi-Structured Data(JSON), Spatial Data.
- Excellent Tooling Support for both On-premise and Cloud. Also, Azure offers tooling to move On-premise SQL Server to Cloud with ease.

When to Use SQL Server

- Data is structured (SQL) with an ACID transaction guarantee.
- The development platform is strongly coupled with other Microsoft Products.
- Azure Cloud is the preferred public Cloud Technology with a preference for Managed SQL database.
- A company already has MS-SQL databases and wants to lift-and-shift them in Cloud.

When not to Use SQL Server

- Not ideal for companies with low Budgets for databases.

- Multi-Master ACID transaction is a must-have feature.
- Data is Semi-structured, i.e., JSON data, and needs advanced query operations.
- Distributed SQL database is desired.
- A Converged database is needed.

SQL Server As a Service

- [Azure SQL](#) (horizontal scaling)
- [Amazon RDS for SQL Server](#)
- [SQL Server on Google Cloud](#)

Alternatives

- [MySQL](#)
- [Oracle](#)
- [PostgreSQL](#)
- [MariaDB](#)
- [IBM Db2](#)
- [Amazon Aurora](#)
- [Google Cloud Spanner](#)

Popularity:

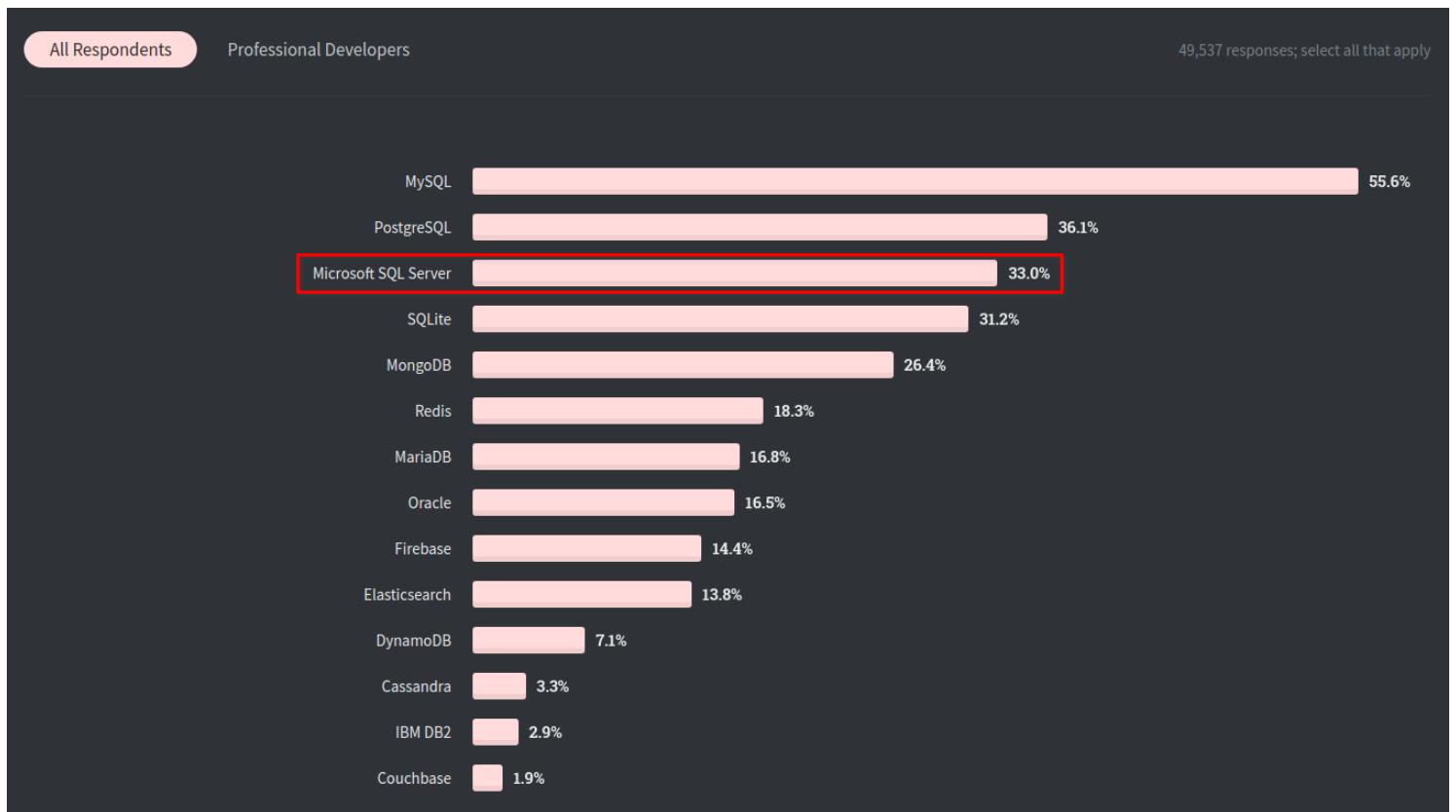
MS SQL is the most popular commercial database in the Windows platform and the preferred SQL database in Azure Cloud. Due to these facts, it remains as the third most popular database according to the DB-Engines ranking:





Image by Author (Data: [DB-Engines](#))

It is also the third most popular database in the Stack Overflow Developer survey:



Source: [Stack Overflow Developer Survey, 2020](#)

Trending

MS SQL is one of the oldest databases on this list. Although it is a solid and stable database, it is not the most innovative or advanced one and has downward trending as shown by DB-Engines trending:



Source: [DB-Engines](#)

Google Trends showed a downward trending for MS SQL over the years, which will not change in 2021 as well:



Source: [Google Trends](#)

5. MongoDB

There are many use cases when we need to save Semi-Structured data (e.g., JSON, XML) in the database. Furthermore, loading and accessing data using Object Oriented programming languages into RDBMS need additional application-level mapping (Object-Relational Impedance Mismatch). To address those issues especially handling the Document Data, *the company 10gen* (currently **MongoDB Inc.**) released **MongoDB** as the first Document Database in 2009. In the next decade, MongoDB went through major improvements. It has addressed many of its shortcomings (e.g., Security) and innovated and pioneered many features.

Currently, it is the principal Document Database and the leading NoSQL Database.

5 Key Features

- Open-Core NoSQL Document Database (BSON) with various licenses: Community Server (free), Enterprise Server (commercial), Atlas (managed in Cloud).
- Offers horizontal scaling via Auto-Sharding and has built-in replication via primary-secondary nodes. In terms of CAP, it is CP (Consistent and Partition tolerant).
- Distributed multi-document ACID transactions with snapshot isolation.
- Rich and powerful query language with complex querying using aggregation pipeline. It also supports Map-Reduce query, text search, graph search, and geo-search.
- MongoDB Inc. also offers a full-text search engine (Atlas Search) and data lake (Atlas Data Lake) built on MongoDB.

When to Use MongoDB

- Data is Document or semi-structured(e.g. JSON, XML) with advanced query features.
- Fast application development where the schema is continuously changing.
- Large structured data is normalized to multiple SQL tables and suffers in terms of performance and complexity.
- Data is Schema-less.

- Consistency is preferred over availability for semi-structured data.

When not to Use MongoDB

- Data is structured data.
- Data is extremely relational (e.g., Social Media), i.e., Graph like data.
- For Data Warehousing (OLAP database).
- Multi-Master write is needed for write-heavy loads.
- High availability is preferred over Consistency.

MongoDB As a Service

- [MongoDB Atlas](#) (Global Scaling)
- [ScaleGrid MongoDB](#) (Managed MongoDB)
- [ApsaraDB for MongoDB](#)

Alternatives

- [Amazon DynamoDB](#)
- [Azure Cosmos DB](#)
- [Amazon DocumentDB](#)
- [Couchbase](#)
- [CouchDB](#)
- [Google Cloud Firestore](#)

Popularity

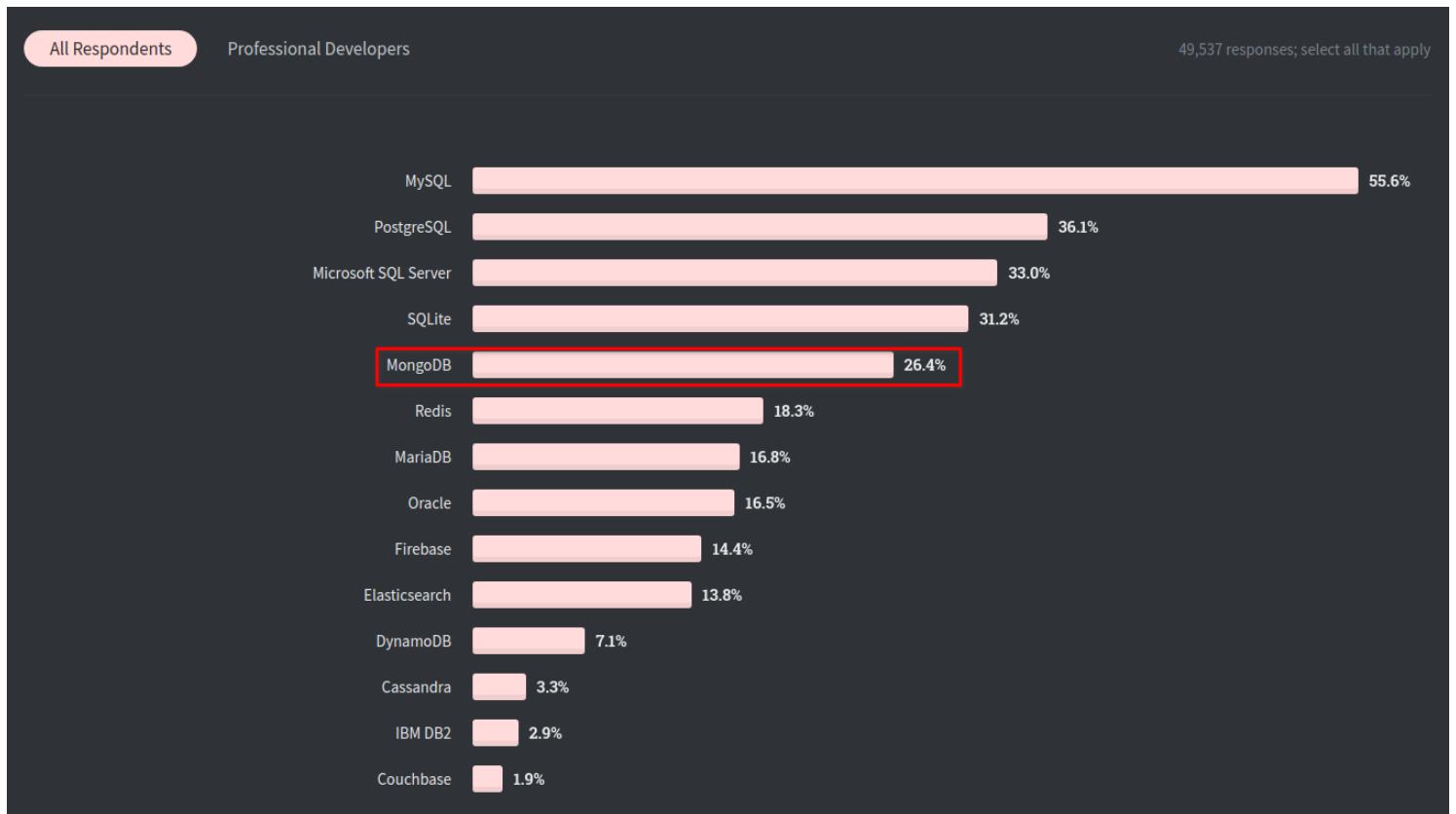
As the prominent Document Database, MongoDB is enjoying soaring popularity in recent years and the fifth most popular database according to the DB Engines ranking:





Image by Author (Data: [DB-Engines](#))

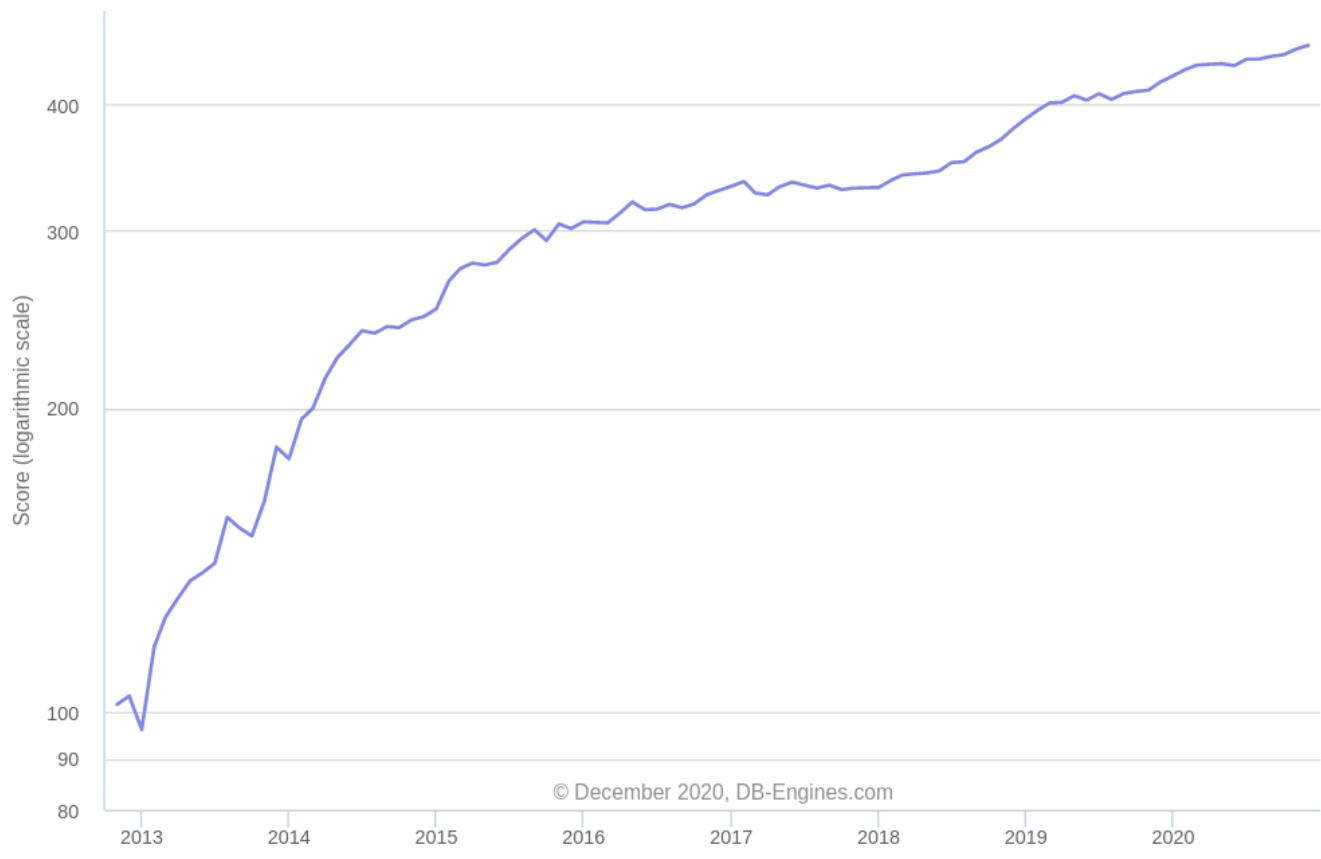
The Stack Overflow Developer Survey 2020 also puts MongoDB in 5th position:



Source: [Stack Overflow Developer Survey, 2020](#)

Trending

MongoDB is getting upsurging traction in recent years, as shown by the DB Engines:



Source: [DB-Engines](#)

Google Trends also shows rising trends of MongoDB:



Source: [Google Trends](#)

6. Redis

In 2009, an Italian Developer, Salvatore Sanfilippo, was trying to improve his startup's scalability and wanted to develop a real-time log analyzer. He encountered significant scaling issues using a traditional database system and created Redis as a distributed in-memory key-value store. Redis quickly got traction from the community and evolved as a highly-scalable in-memory data structure store used extensively in the industry.

5 Key Features

- Open Core distributed in-memory NoSQL data structure store with BSD 3-clause license.
- Used as a distributed, in-memory key-value database. Also used as a distributed cache (with or without database) or as message broker with optional durability.
- Supports a wide range of data structures: strings, hashes, lists, sets, sorted sets with range queries, bitmaps, hyper-logs, geospatial indexes, and streams.
- Highly scalable with built-in replication, automatic failover, and sharding via Redis Cluster. In terms of CAP, it is CP (Consistent and Partition tolerant).
- Advanced Cache functionality (LRU eviction policy, Keys with limited TTL).

When to Use Redis

- The distributed in-memory data structure is required.
- Distributed and enterprise-grade caching is required, e.g., to improve database access performance.
- For Session Management.
- Real-time use cases, e.g., Inventory systems.
- For a large amount of Dataset.

When not to Use Redis

- Single server in-memory data structure store is required for a small set of data.
- Primary data storage (OLTP Database) with an ACID transactional guarantee.
- Single Server caching is needed for a small set of data.

- Data is extremely relational (e.g., Social Media), i.e., Graph like data.
- When an OLAP database is needed.

Redis As a Service

- [Redis Enterprise Cloud](#) (Multi-Cloud)
- [ScaleGrid Redis](#) (Managed Redis)
- [Aiven for Redis](#) (Multi-Cloud)
- [Amazon ElastiCache](#)
- [Azure Cache for Redis](#)
- [Google Memorystore](#)

Alternatives

- [Memcached](#)
- [Hazelcast](#)
- [etcd](#)
- [Ehcache](#)
- [Riak KV](#)

Popularity:

Distributed In-memory data store and distributed caching is a common use-case for almost all enterprises. As a result, Redis is highly popular in the industry, as shown below:

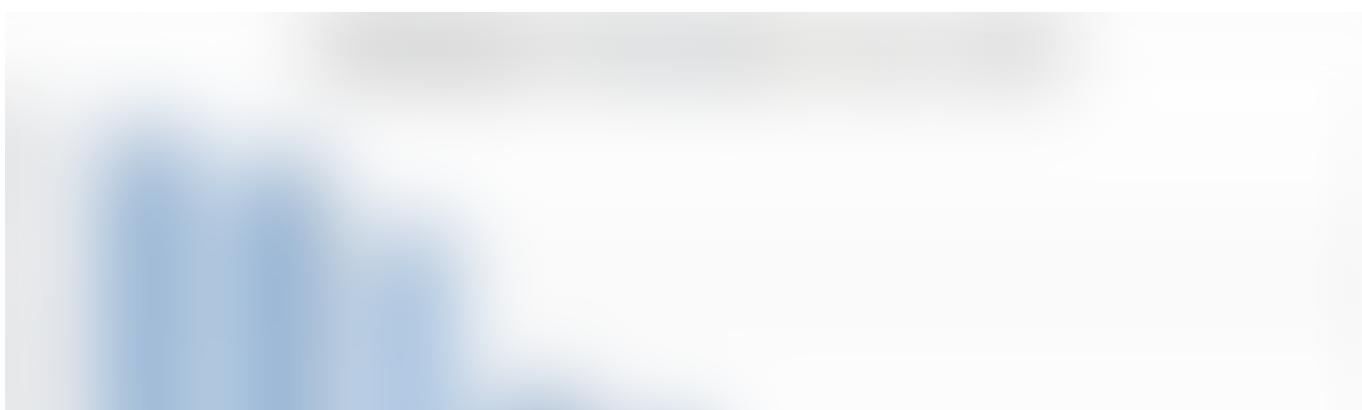
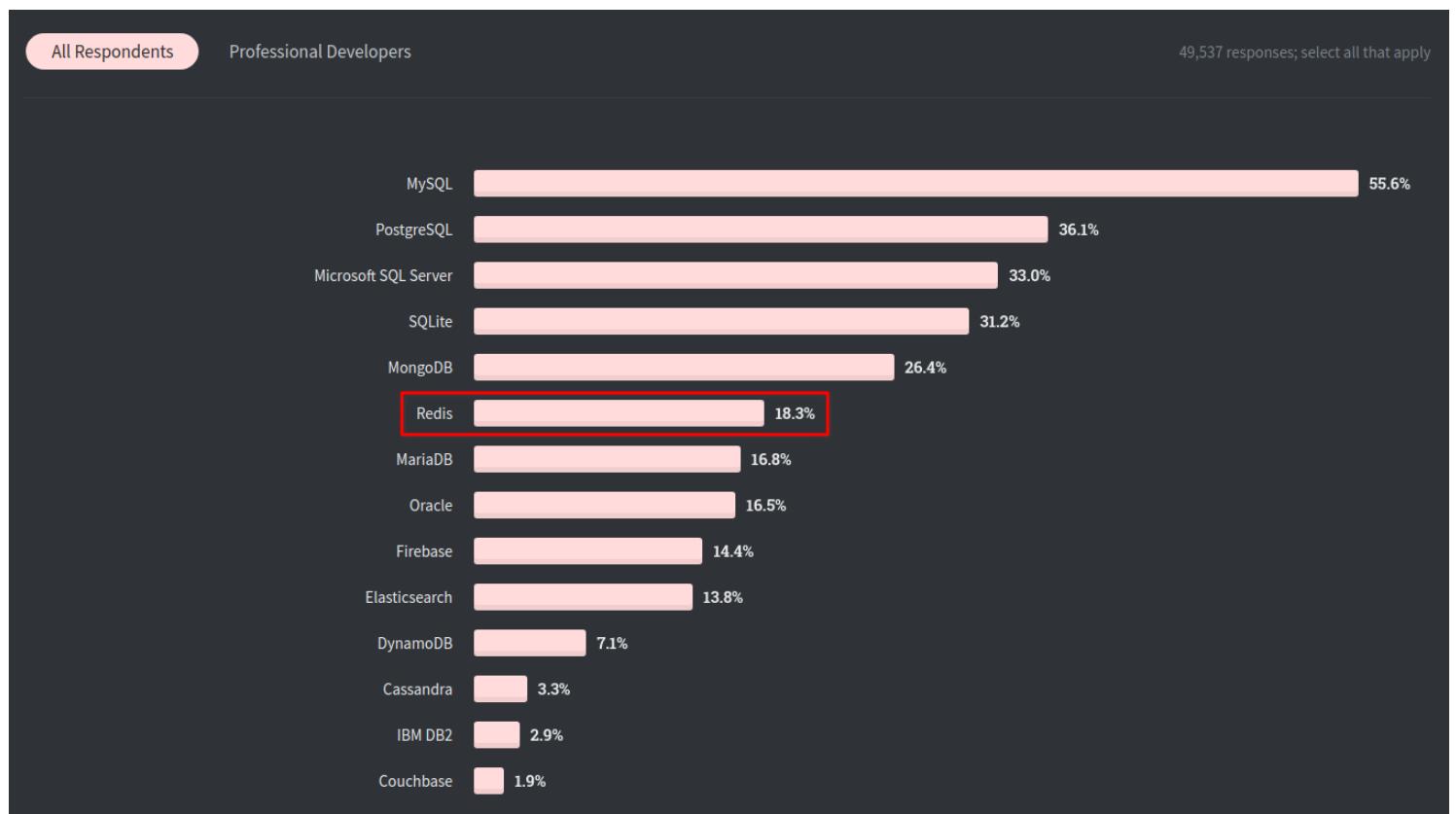




Image by Author (Data: [DB-Engines](#))

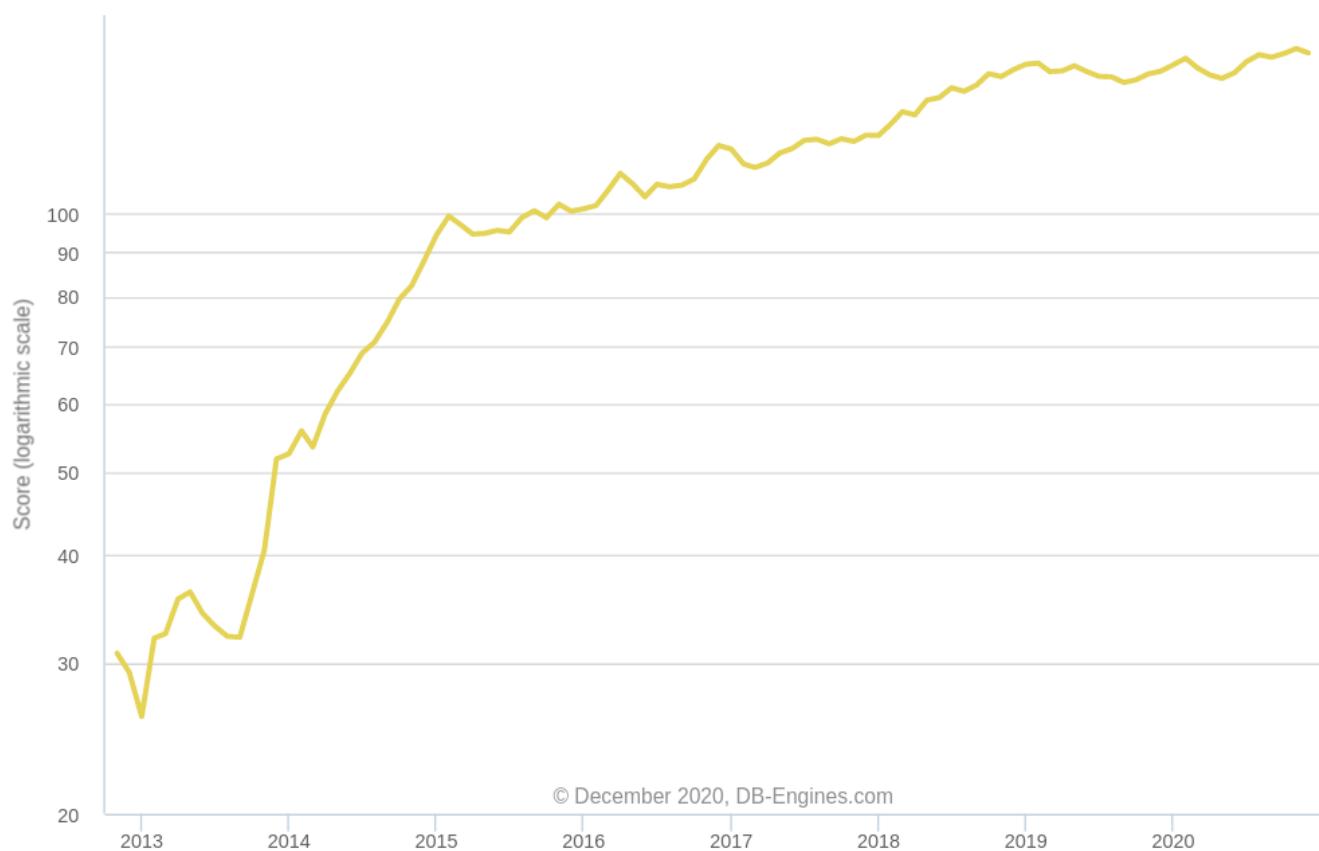
The Stack Overflow Developer Survey also puts Redis as the 6th most popular database overall:



Source: [Stack Overflow Developer Survey, 2020](#)

Trending

With the rise of Cloud and Microservices, Redis is getting remarkable traction in recent years, as shown below:



Source: [DB-Engines](#)

Google Trend also shows liner rise of trending for Redis since its inception:



Source: [Google Trends](#)

7. Elasticsearch

Whether it is a small startup company or a Web-Scale e-commerce company, it needs full-text search. Doug Cutting developed the Java library Lucene for the full-text search purpose in 1999. Since then, Lucene is the industry standard for full-text search.

Elasticsearch is a full-text search engine based on Lucene and originally developed by Shay Banon and released in 2010. It is a distributed, multi-tenant capable full-text search engine offering REST API.

5 Key Features

- Open Core, distributed search engine.
- Offers horizontal scaling via automatic sharding. It is the most scalable search engine. In terms of CAP, it is AP (Available and Partition Tolerant).
- Offer REST API and supports structured and schemaless data (JSON). As a result, it is especially suited to analyze Logging or Monitoring data.
- Support automatic replication and Cross-cluster Replication (CCR).
- Part of the formidable “Elastic Stack” (Elasticsearch, Kibana, Logstash, Beats). Organizations use it massively in the Observability landscape.

When to Use Elasticsearch

- Moderate to advanced full-text search is needed.
- Near real-time full-text search is required with advanced machine learning features (e.g., fraud detection).
- Search engine with advanced tooling support (Kibana, Logstash, Beats).
- For a large volume of data sets.
- Semi-structured data (e.g., JSON, Log Data) is used.

When not to Use Elasticsearch

- Only basic full-text search features are required.
- ACID-compliant primary database (OLTP Database) is required.
- Dataset is small and only structured data.

- Only full-text search is required without sophisticated tooling (Kibana, Logstash).
- When cutting-edge but expensive features like advanced machine learning is not a key requirement.

Elasticsearch As a Service

- [Elastic Cloud](#) (Multi-Cloud)
- [Aiven for Elasticsearch](#) (Multi-Cloud)
- [Amazon Elasticsearch Service](#)

Alternatives

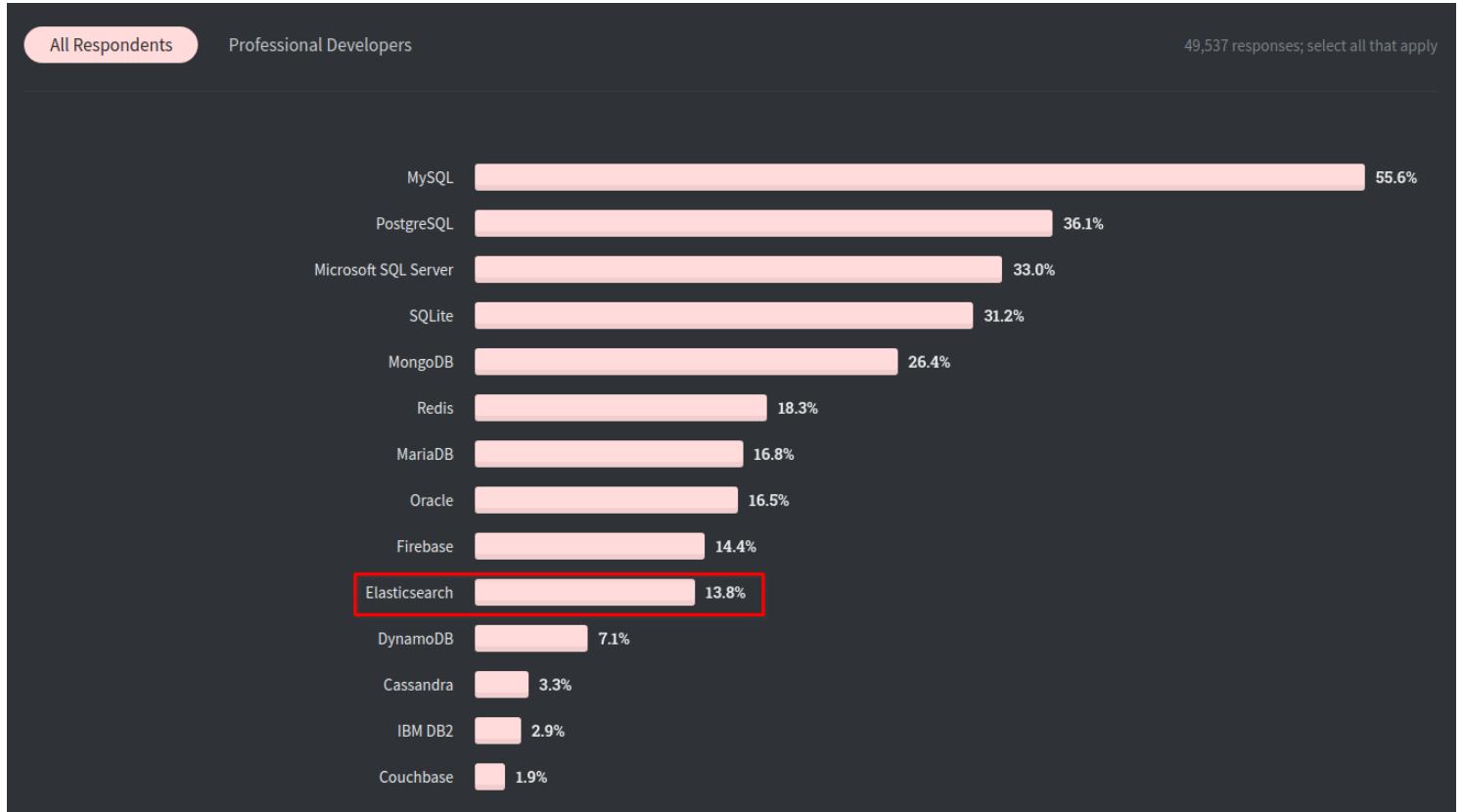
- [Apache Solr](#)
- [Splunk](#)
- [Algolia](#)

Popularity:

In recent times, Elasticsearch is used massively as a search engine and for the observability of applications. As a result, it is the 8th most popular database and the most popular search engine in the DB-Engines ranking:



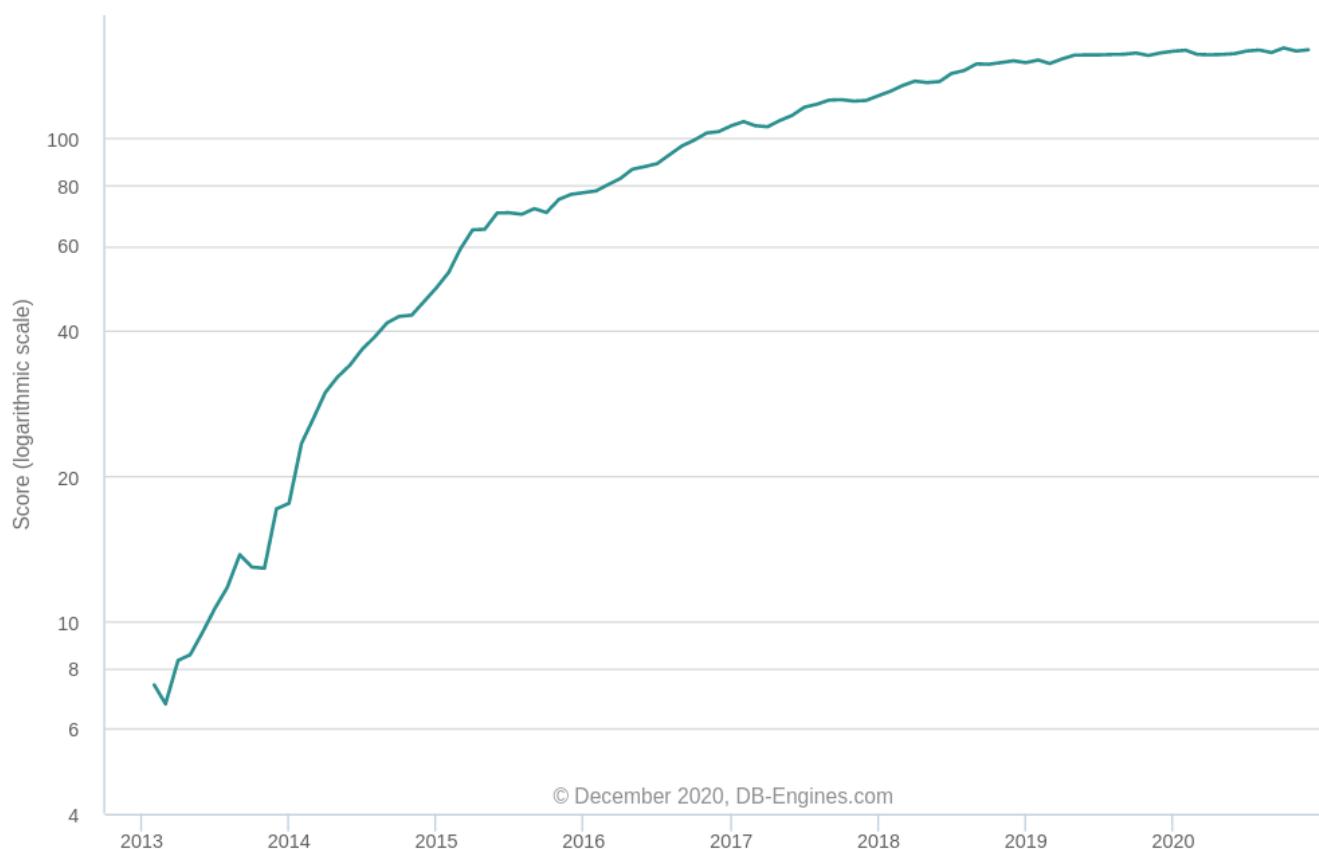
In Stack Overflow Developer Survey 2020, Elasticsearch ranked 10th positions which is a tremendous feat as a full-text search engine:



Source: [Stack Overflow Developer Survey, 2020](#)

Trending

In recent times, observability is gaining significant interest because of the rise of Cloud computing and Microservices. As a cornerstone of the modern observability stack, Elasticsearch is getting tremendous traction in recent years:



Source: [DB-Engines](#)

Data from Google Trends also shows a significant rise of interest for Elasticsearch in recent years:



Source: [Google Trends](#)

8. Cassandra

During the mid-2000s, Facebook was rapidly growing and needed a database with massive scaling. Unfortunately, there was no database in the market which could meet Facebook's scaling requirement. Taking influence from Google's [Big Table](#) paper (2006) and Amazon's [Dynamo](#) paper (2007), two Facebook Engineers [Avinash Lakshman](#) (co-author of Amazon Dynamo paper) and [Prashant Malik](#), developed [Cassandra](#) as a distributed [Wide Column Store](#) database in 2008. It is a highly scalable database and widely used in the industry to handle massive data.

5 Key Features

- Open Core, distributed, Wide Column Store database with Apache License 2.0.
- Used as OLAP Database (e.g., Data Warehouse) to handle the gigantic volume of data. Also used as a time-series database.
- Offers linear horizontal scaling and one of the most scalable databases with automatic sharding. In terms of CAP, it is AP (Available and Partition Tolerant).
- Decentralized database (Leaderless) with automatic replication and multi-datacenter replication. As a result, it is fault-tolerant with no single point of failure.
- Has user-friendly and SQL like query language: Cassandra Query Language (CQL).

When to Use Cassandra

- As typical OLAP database, e.g., as Data Warehouse.
- As a time series database with massive write operations.
- Web-Scale applications need to handle a massive amount of write and read operations with linear scalability.
- For high durable use cases when data-loss is acceptable.

When not to Use Cassandra

- As a primary database (OLTP) needing an ACID transaction guarantee.
- If linear scalability is not required or high-availability is not required.
- If Data is frequently deleted, then its performance degrades ("Tombstones").

- It is JVM based and needs more Node, and causes a hefty Cloud bill. The C++ based Cassandra compliant [ScyllaDB](#) is a better option to save Cloud bill.

Cassandra As a Service

- [DataStax Astra \(Multi-Cloud\)](#)
- [Aiven for Apache Cassandra \(Multi-Cloud\)](#)
- [Amazon Managed Apache Cassandra Service](#)

Alternatives

- [HBase](#)
- [ScyllaDB](#)
- [Azure Cosmos DB](#)
- [Google Cloud Bigtable](#)
- [Amazon Keyspaces](#)

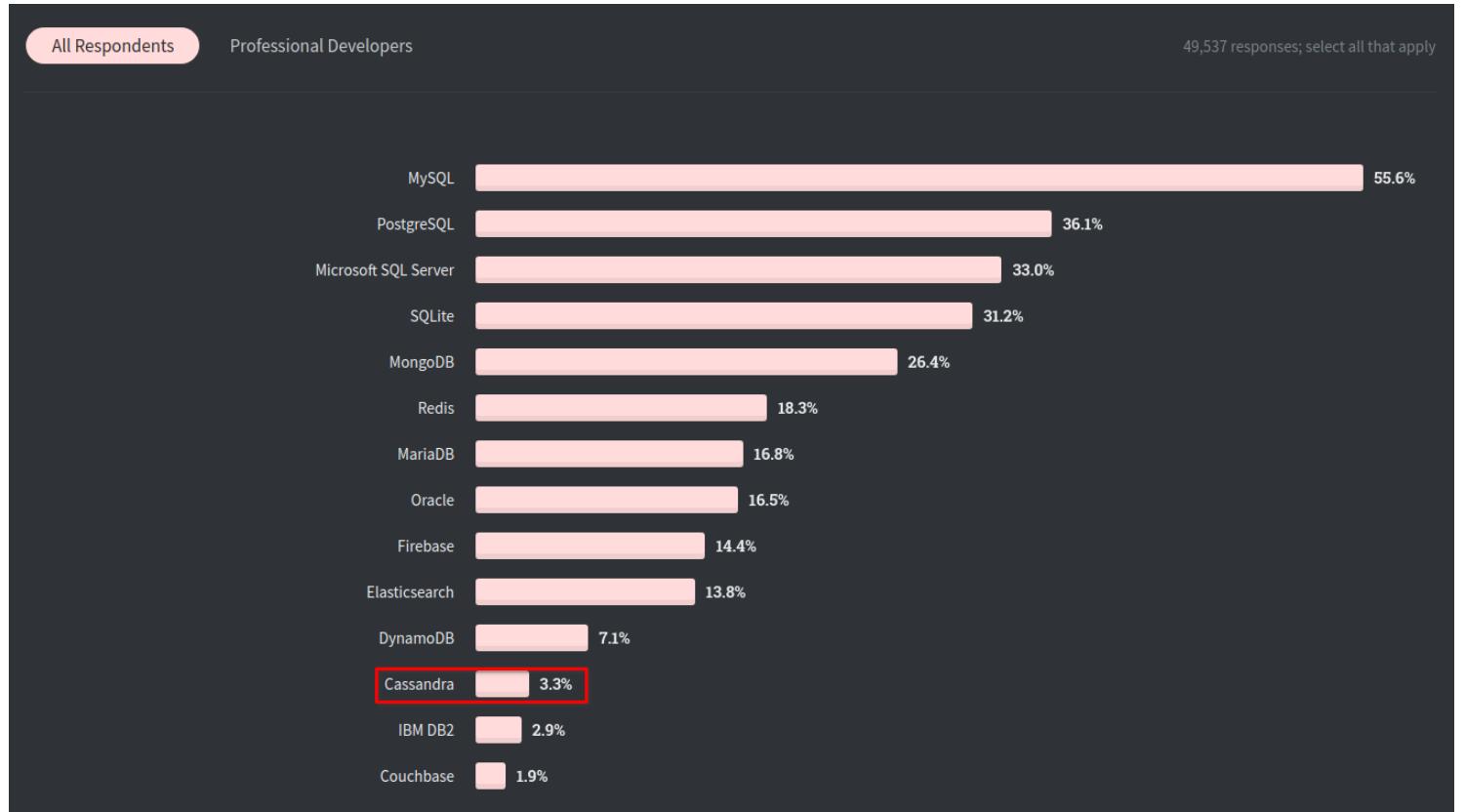
Popularity:

Cassandra is one of the most popular NoSQL database and the most popular Wide Column Store database. It is the 9th most popular database shown below:



Image by Author (Data: [DB-Engines](#))

The Stack Overflow Developer Survey also ranked Cassandra in the 12th position for 2020:



Source: [Stack Overflow Developer Survey, 2020](#)

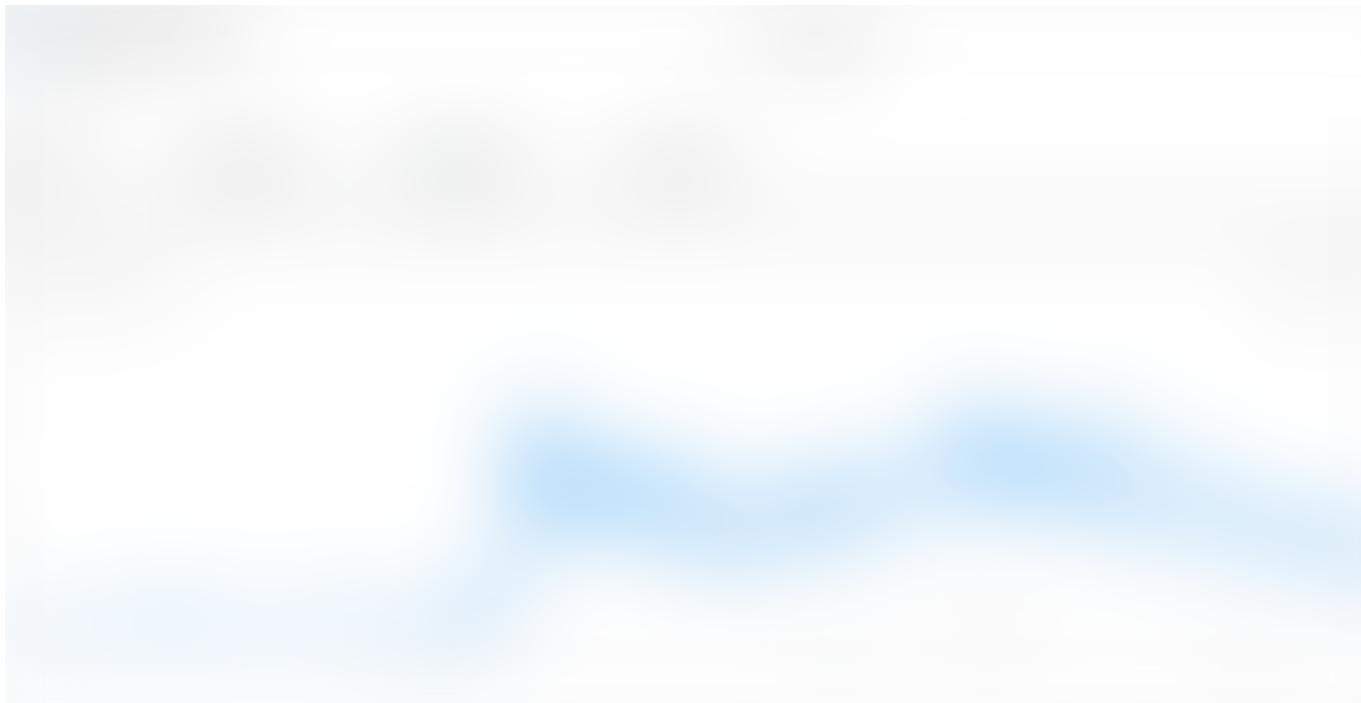
Trending

As the leading database for massive OLAP workload, Cassandra got huge traction in 2016 and still maintaining that level:



Source: [DB-Engines](#)

Google Trending also shows steady interest for Cassandra for a long period of time:



Source: [Google Trends](#)

9. MariaDB

When Oracle acquired Sun Microsystems, the future of MySQL was unclear as Sun Microsystems owned it. The co-creator of MySQL, Michael Widenius, forked the MySQL and created **MariaDB**. MariaDB is a Relational Database Management System and compatible with MySQL protocol and MySQL clients. MySQL server can be replaced with MariaDB server effortlessly with no Code change. MariaDB is more community-driven compared to MySQL.

5 Key Features

- Open source RDBMS with GPL v2 license. It has two licensing models: free Community Server and commercial Enterprise Server.
- Offers ACID transactional guarantee (with InnoDB engine). In terms of CAP, it offers immediate Consistency as a single Server.
- With its Storage Engine “ColumnStore,” MariaDB offers columnar storage with massively parallel distributed data architecture. It also offers horizontal partitioning (sharding) via its **MaxScale** and **Spider** Engine. Thus MariaDB is also suitable as an OLAP database.
- With its Graph Computation Engine “OQGRAPH,” it supports hierarchical and complex graph data types.
- With its **Xpand** smart Storage, it offers the “**Distributed SQL**” features out-of-the-box. As a result, it offers Millions of transactions per second.

When to Use MariaDB

- Data is structured (SQL), and ACID transaction guarantee is a key criterion.
- “**Distributed SQL**” is required where millions of transactions should be handled in a globally distributed database.
- Multi-Master clustering and Multi-Node data warehousing are required (OLAP).
- A multi-model database is required, i.e., one database to handle structured, semi-structured, graph, and columnar data.

- A converged database is required, i.e., one database for OLTP, OLAP, and Graph workload.

When not to Use MariaDB

- Multi-Master ACID transaction is a must-have feature.
- Data is Semi-structured, i.e., JSON data and highly advanced query functions are needed.
- Advanced Indexing like descending, functional and invisible are expected.
- Advanced Authentication and Authorization (e.g., Password reuse policies, partial revokes) are desired.

MariaDB As a Service

- [SkySQL](#) (Distributed SQL)
- [Amazon RDS for MariaDB](#)
- [Azure Database for MariaDB](#)

Alternatives

- [MySQL](#)
- [Oracle](#)
- [PostgreSQL](#)
- [Microsoft SQL Server](#)
- [IBM Db2](#)
- [Amazon Aurora](#)
- [Google Cloud Spanner](#)

Popularity:

The big four SQL databases dominate the SQL database domain. As the youngest SQL database in this list, it was not easy for MariaDB to gain popularity.

But MariaDB has done a pretty good job and already one of the most popular databases, as shown below:



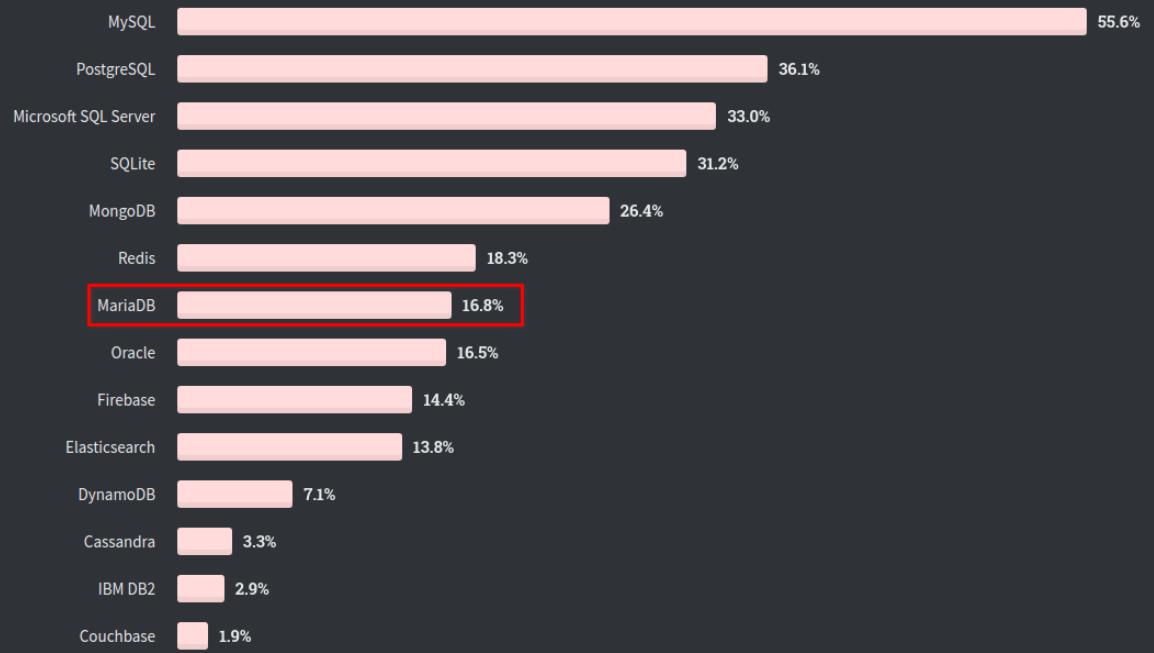
Image by Author (Data: [DB-Engines](#))

According to the Stack Overflow Developer Survey 2020, it has already surpassed many established databases like Oracle and IBM DB2:

All Respondents

Professional Developers

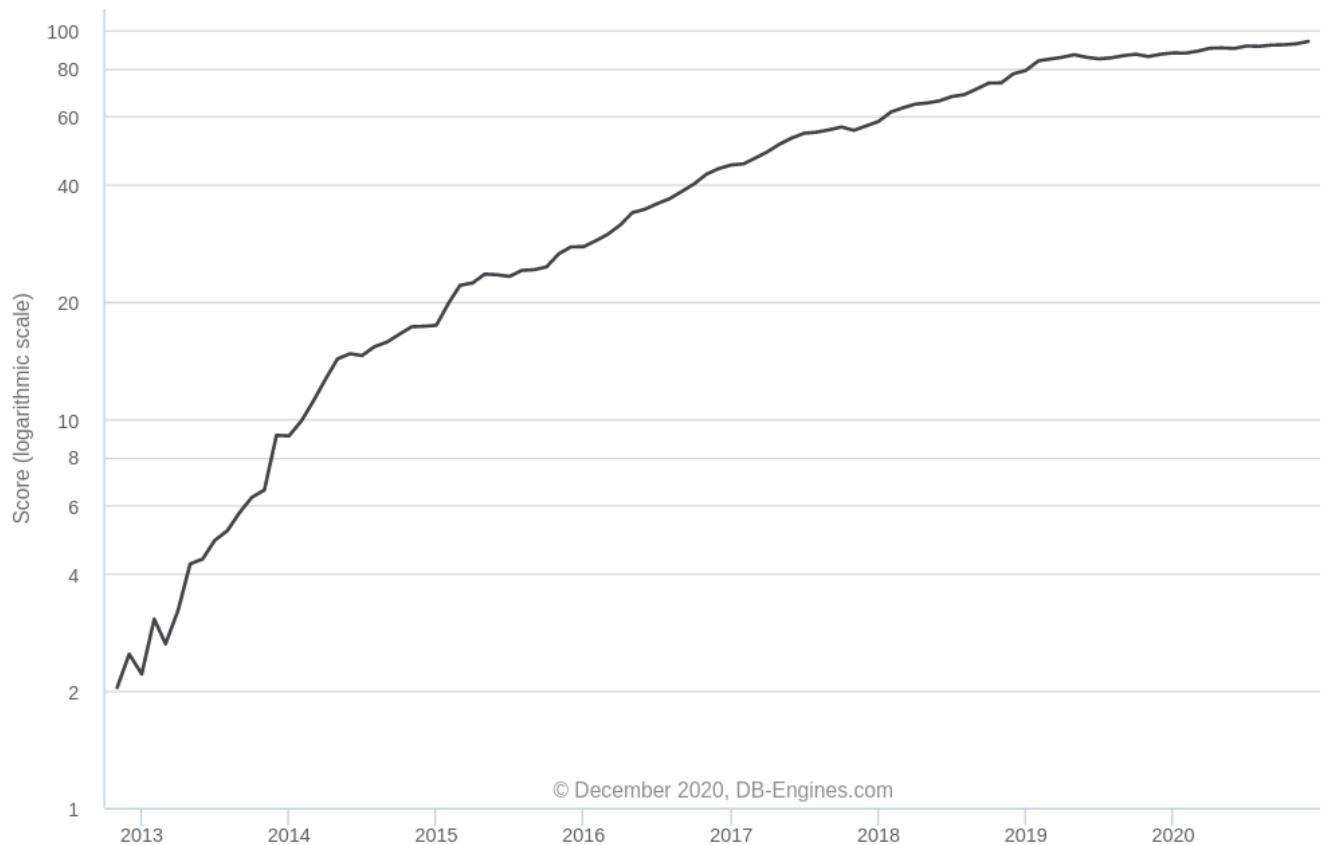
49,537 responses; select all that apply



Source: [Stack Overflow Developer Survey, 2020](#)

Trending

With Many innovations and Distributed SQL features, MariaDB is one of the trendiest SQL databases, as shown below:



Source: [DB-Engines](#)

Also, it is getting more traction in recent years, as shown by Google Trends:



Source: [Google Trends](#)

10. IBM Db2

In 1966, IBM developed one of the first database Management System IMS (a hierarchical database) for the Apollo program. When Edgar F. Codd, an IBM Engineer, published his seminal paper on Relational Database, IBM was reluctant to abandon their IMS database. Once Oracle released the first Relational Database, IBM quickly corrected its path and released Db2 as the commercially supported RDBMS in 1983. Initially, IBM released Db2 for IBM's mainframe machines. In 1987, IBM released Db2 LUW for Windows, Linux, Unix systems as well. Here I will discuss Db2 LUW as the Db2 family has other databases.

Over the years, IBM Db2 has gone through many major changes and evolved a lot like Oracle.

5 Key Features

- Db2 is a proprietary RDBMS with an ACID transactional guarantee. It has different licenses: Community (free), Standard, and Advanced (commercial).
- One of the most advanced Multi-Model databases supporting structured data (SQL), semi-structured data (JSON), and Graph Data.
- It is also a Converged database (or Master database) and has excellent OLAP support via IBM BLU Acceleration (In-memory computing, Massively Parallel Computing, Column-based Db2 Shadow tables).
- Native support for AI.
- Offer horizontal scaling via Db2 pureScale.

When to Use Db2

- If Converged database or Master Database (One database for all-purpose) is a key requirement.
- As a primary Database (OLTP Database) with an ACID transaction guarantee.
- AI functionality in the database is required for competitive advantage.
- As OLAP database (e.g., Data Warehouse).
- The company is already using IBM Cloud or IBM Cloud Pack for Data.

When not to Use Db2

- If the company wants to save money or on budget.
- Multi-Master ACID transaction is a must-have feature.
- Data is Semi-structured, i.e., JSON data with advanced query functions.
- A pure SQL database is required without AI functionality.
- When only Azure or Google Cloud is used as Cloud provider.

Db2 As a Service

- [IBM Db2 on Cloud](#)
- [Db2 Workgroup on AWS Marketplace](#)

Alternatives

- [MySQL](#)
- [Oracle](#)
- [PostgreSQL](#)
- [Microsoft SQL Server](#)
- [MariaDB](#)
- [Amazon Aurora](#)
- [Google Cloud Spanner](#)

Popularity:

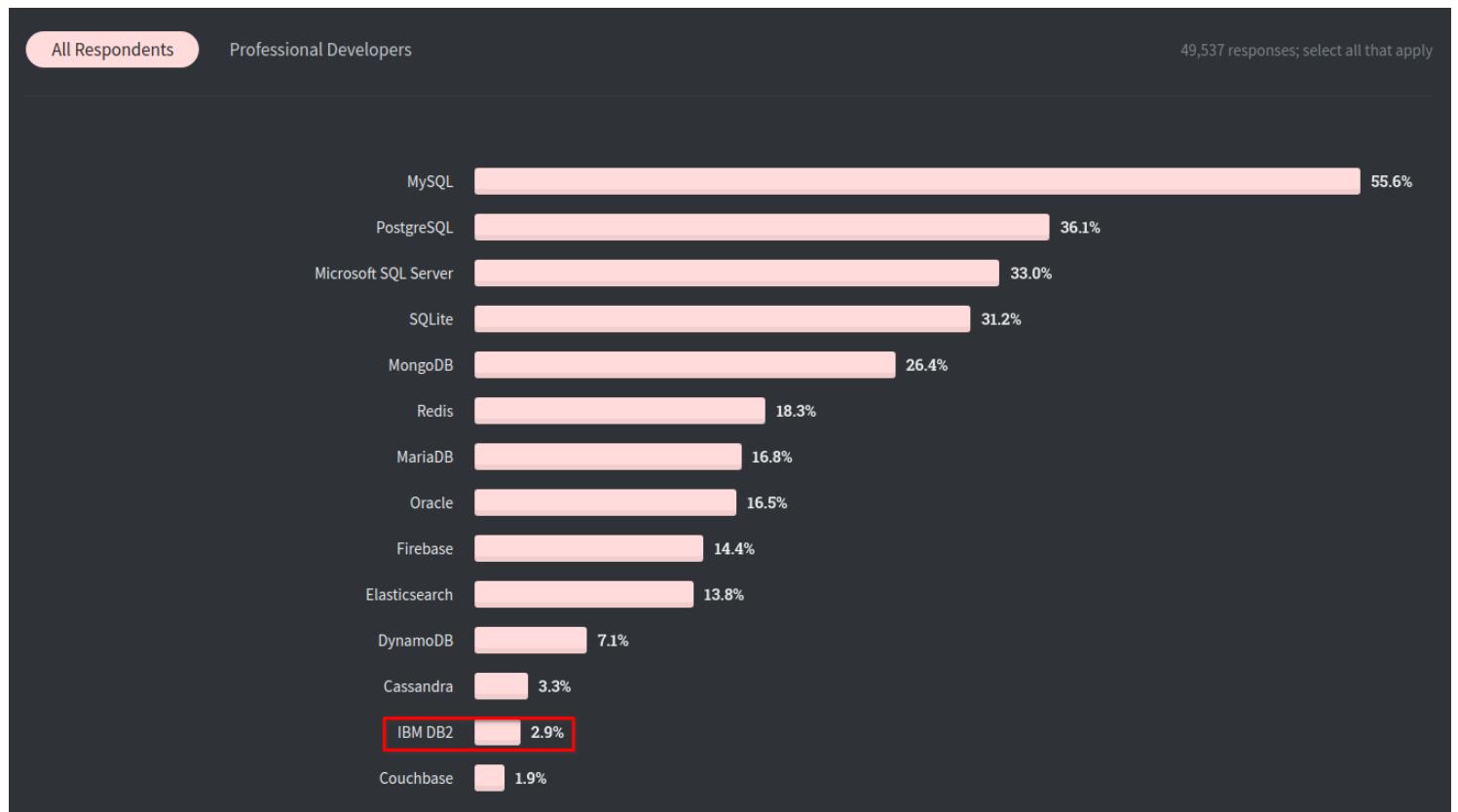
Once a market leader, Db2 has lost its popularity in recent years as the database landscape competition is multi-faceted now.

Nevertheless, it is still one of the most used databases in medium to large enterprises and enjoys a prominent position in popularity:



Image by Author (Data: [DB-Engines](#))

In the Stack Overflow Developer survey, which has recency bias, it is ranked quite low:



Source: [Stack Overflow Developer Survey, 2020](#)

Trending

Db2 is also not getting lots of traction in recent years, as shown below:



Source: [DB-Engines](#)

Google Trends also confirms its continuous downward trend:



Source: [Google Trends](#)

Conclusion

In this list, I have shown ten databases to use in 2021. The RDBMS systems dominate the list. MySQL and PostgreSQL are the leaders from the open-source and free database, closely followed by MS SQL. The MySQL compatible MariaDB is increasingly getting popular.

Among the commercial databases, Oracle has added many value-added features in its latest release and has taken a giant step towards the future. Although IBM Db2 is still used in large organizations, it will slowly decline.

Among the NoSQL databases, MongoDB (Document Database), Redis (in-memory Key-Value Store), Cassandra (Wide-Column Database), and Elasticsearch (full-text Search Engine) are the leaders in their domains.

The list is not all-inclusive, and many important database types could not make in the top ten list like NewSQL databases, Cloud Native Databases. But it will give a high-level overview of the top databases for 2021.

• • •

Similar Articles

5 Best Public Cloud Database to Use in 2021

Amazon DynamoDB, Google BigQuery, Azure SQL Server, Azure Cosmos DB, Amazon Redshift

[towardsdatascience.com](https://towardsdatascience.com/5-best-public-cloud-database-to-use-in-2021-1a2a2f3a2)



21 Predictions about the Software Development Trends in 2021

Cloud, Edge, Container, Quantum, Blockchain, AI, Deep Learning, Batch, Streaming, Database, Programming, Software...

[towardsdatascience.com](https://towardsdatascience.com/21-predictions-about-the-software-development-trends-in-2021-1a2a2f3a2)



Microservice Architecture and its 10 Most Important Design Patterns

Microservice Architecture, Database per Microservice, Event Sourcing, CQRS, Saga, BFF, API Gateway, Strangler, Circuit...

towardsdatascience.com



Sign up for The Variable

By Towards Data Science

Every Thursday, the Variable delivers the very best of Towards Data Science: from hands-on tutorials and cutting-edge research to original features you don't want to miss. [Take a look.](#)

+ Get this newsletter

Data

Database

NoSQL

Big Data

Programming



[About](#) [Write](#) [Help](#) [Legal](#)

Get the Medium app

