



Bases de Datos 1

Alejandra Beatriz Lliteras

Contenidos de la materia

- ▶ Modelo de datos
 - ▶ Conceptos generales
 - ▶ Algunos modelos en particular
 - ▶ Modelo de Entidades y Relaciones
 - ▶ Modelo relacional
- ▶ Transformación entre modelos de datos
- ▶ Álgebra Relacional
 - ▶ Operaciones y Consultas
- ▶ **Teoría de diseño de bases de datos**
 - ▶ **Conceptos generales**
 - ▶ Proceso de Normalización
- ▶ SGBD Relacional
- ▶ Conceptos generales de bases de datos

Motivación

Albarak, M., Bahsoon, R., Ozkaya, I., & Nord, R. L. (2020). Managing Technical Debt in Database Normalization. *IEEE Transactions on Software Engineering*.

- ▶ Uno de los principales principios del diseño de bases de datos es la **NORMALIZACION**
 - ▶ **Organiza** los datos en relaciones o tablas siguiendo **reglas** específicas para minimizar la redundancia de los datos y en consecuencia, favorecer a la consistencia reduciendo anomalías
 - ▶ Un mal diseño puede llevar a la necesidad de **refactorizar** la base de datos y eso suele ser muy caro
 - ▶ Los beneficios de la normalización van mas allá de la **calidad** de los datos, pudiendo mejorar la **mantenibilidad** y **performance** de una base de datos
 - ▶ En general se ignora por ser un **proceso que requiere tiempo y conocimiento específico**

Motivación

Jadhav, R., Dhabe, P., Gandewar, S., Mirani, P., & Chugwani, R. (2020). A New Data Structure for Representation of Relational Databases for Application in the Normalization Process. In *Machine Learning and Information Processing* (pp. 305-316). Springer, Singapore.

- ▶ El **diseño de un esquema** de bases de datos es uno de los factores mas importantes de los que **depende el éxito** de la base de datos relacional
- ▶ La normalización es un **paso clave** en el diseño de la base de datos relacional
 - ▶ Toma una **relación grande** como entrada y la **descompone** en relaciones mas pequeñas las cuales están libres de redundancia de datos y otras **anomalías** como la de inserción/eliminación
 - ▶ La descomposición se realiza siguiendo **reglas/pasos**
- ▶ Puede ser **manual o automática**
 - ▶ La automatización de la normalización es de **gran aplicación en la industria** y reduce el costo de la normalización manual

Motivación

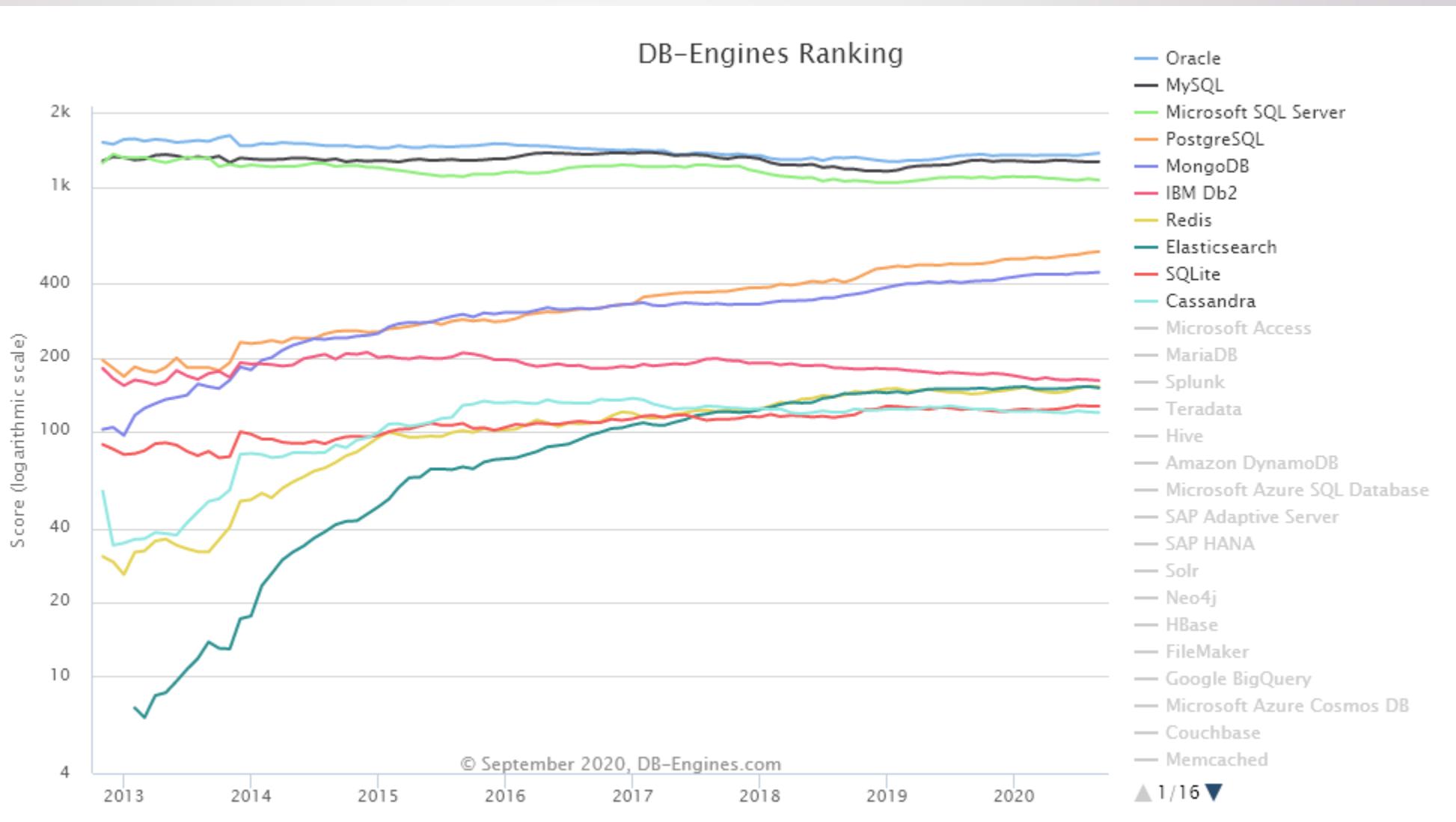
Albarak, M., Bahsoon, R., Ozkaya, I., & Nord, R. L. (2020). Managing Technical Debt in Database Normalization. *IEEE Transactions on Software Engineering*.

- ▶ Normalización se aplica a las bases de datos relacionales (no así a otros modelos como por ejemplo NoSQL)
- ▶ El modelo relacional, sigue siendo el modelo dominante en la industria
- ▶ Utilizado por los 4 principales DBMS
 - ▶ Oracle, MySQL, MS SQL Server y PostgreSQL
- ▶ También utilizado por DBMS como IBM Db2 y SQLight

Motivación

Rank			DBMS	Database Model
Sep 2020	Aug 2020	Sep 2019		
1.	1.	1.	Oracle 	Relational, Multi-model 
2.	2.	2.	MySQL 	Relational, Multi-model 
3.	3.	3.	Microsoft SQL Server 	Relational, Multi-model 
4.	4.	4.	PostgreSQL 	Relational, Multi-model 
5.	5.	5.	MongoDB 	Document, Multi-model 
6.	6.	6.	IBM Db2 	Relational, Multi-model 
7.	7.	↑ 8.	Redis 	Key-value, Multi-model 
8.	8.	↓ 7.	Elasticsearch 	Search engine, Multi-model 
9.	9.	↑ 11.	SQLite 	Relational
10.	↑ 11.	10.	Cassandra 	Wide column

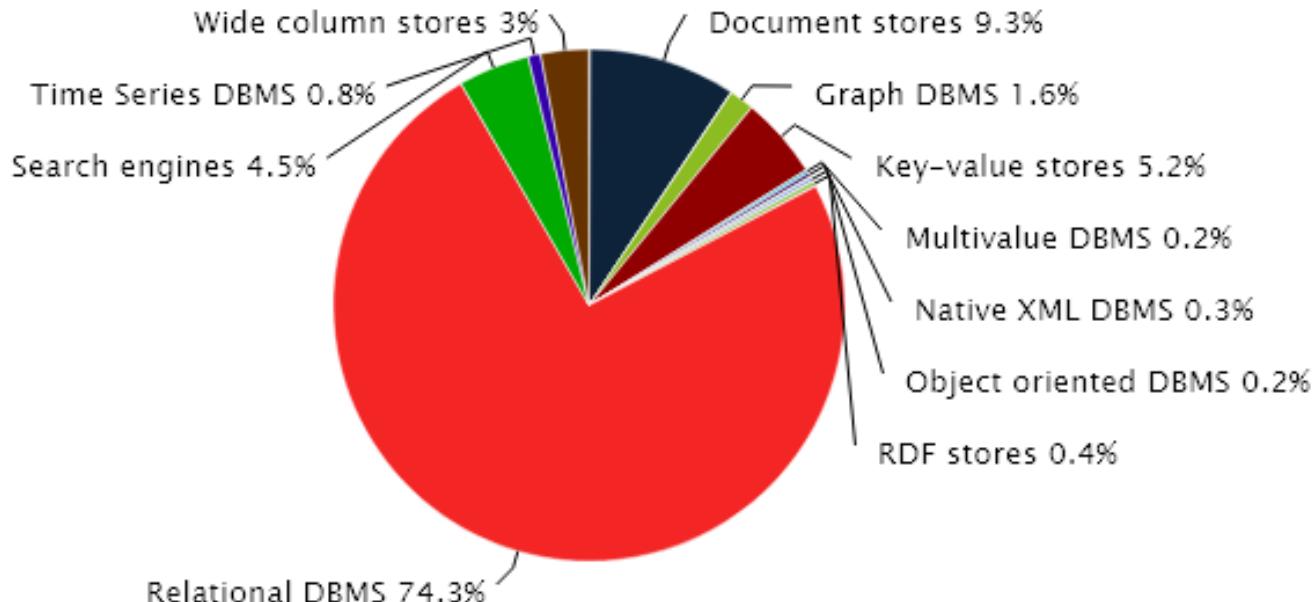
Motivación



https://db-engines.com/en/ranking_trend

Motivación

Ranking scores per category in percent, September 2020



© 2020, DB-Engines.com

https://db-engines.com/en/ranking_trend

Motivación

► En síntesis

- El éxito de una base de datos relacional depende del diseño de su esquema
- Existen procesos manuales y automáticos
- Con impacto en la industria, además de serlo en la academia
- Tema relevante y actual
- El modelo relacional domina el mercado

Bibliografía de los temas abordados en esta clase

- Date, C. J. (2019). *Database design and relational theory: normal forms and all that jazz*. Apress.
- Garcia-Molina, H. (2008). *Database systems: the complete book*. Pearson Education India.
- Ullman, J. D. (1988). Principles of database and knowledge-base systems.
- Albarak, M., Bahsoon, R., Ozkaya, I., & Nord, R. L. (2020). Managing Technical Debt in Database Normalization. *IEEE Transactions on Software Engineering*.
- Jadhav, R., Dhabe, P., Gandewar, S., Mirani, P., & Chugwani, R. (2020). A New Data Structure for Representation of Relational Databases for Application in the Normalization Process. In *Machine Learning and Information Processing* (pp. 305-316). Springer, Singapore.
- Ghawi, R. (2019, May). Interactive Decomposition of Relational Database Schemes Using Recommendations. In *International Conference: Beyond Databases, Architectures and Structures* (pp. 97-108). Springer, Cham.
- Stefanidis, C., & Koloniari, G. (2016, November). An interactive tool for teaching and learning database normalization. In *Proceedings of the 20th Pan-Hellenic Conference on Informatics* (pp. 1-4).
- Knowledge Base of Relational and NoSQL Database Management Systems https://db-engines.com/en/ranking_trend



IMPORTANTE: los slides usados en las clases teóricas de esta materia, no son material de estudio por sí solos.