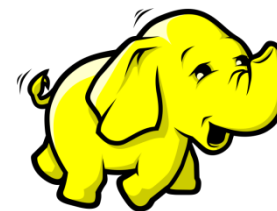


# Relational Databases (SQL)

Michael Enudi

Journey through the world of databases and data engineering



Note

# Scope

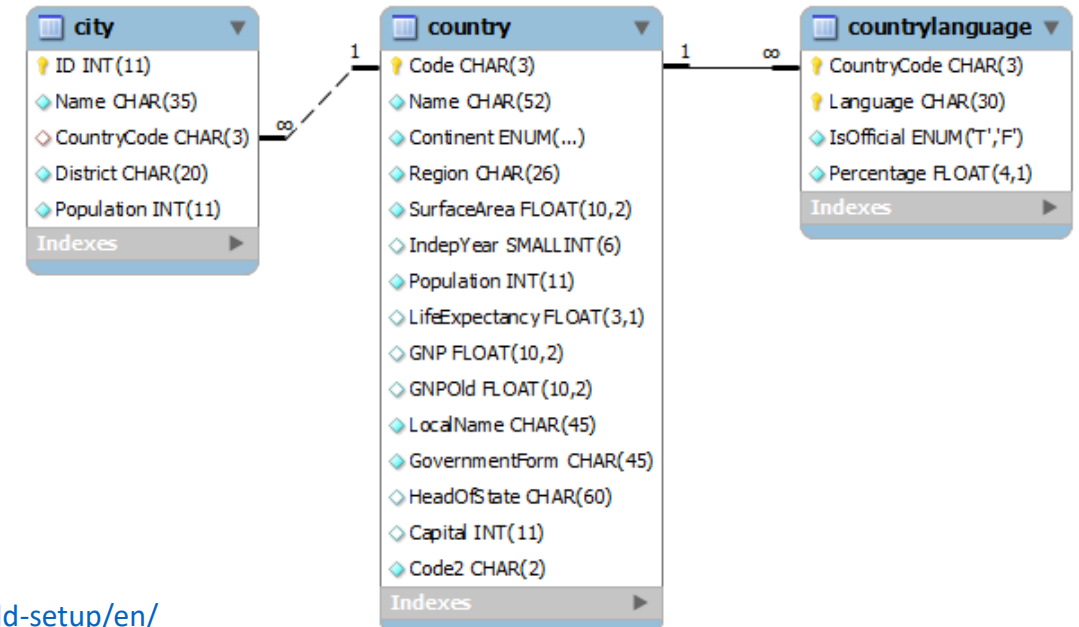
- Introduction to Relational Databases
- SQL
- MySQL
- Movielens in MySQL
- Data Processing
- Indexes
- OLTP Systems
- Data Warehousing
- Analytical Processing (OLAP)
- Database Logs
- Relational Databases - To be or Not to Be



# Relational Databases

- Relations / Tables
- Record / Row / Tuple
- Primary Key(s)
- Fields / Columns / Properties
- Data types
- NULL
- Constraints
- Uniqueness
- Relationships
- Transactions (ACID)
- Index
- ERD
- RDBMS

Code	Name	Continent	Region	SurfaceArea	IndepYear	Population	LifeExpectan...	GNP	GNPOld
ABW	Aruba	North America	Caribbean	193.00	NULL	103000	78.4	828.00	793.00
AFG	Afghanistan	Asia	Southern and Central Asia	652090.00	1919	22720000	45.9	5976.00	NULL
AGO	Angola	Africa	Central Africa	1246700.00	1975	12878000	38.3	6648.00	7984.00
AIA	Anguilla	North America	Caribbean	96.00	NULL	8000	76.1	63.20	NULL
ALB	Albania	Europe	Southern Europe	28748.00	1912	3401200	71.6	3205.00	2500.00
AND	Andorra	Europe	Southern Europe	468.00	1278	78000	83.5	1630.00	NULL
ANT	Netherlands Antilles	North America	Caribbean	800.00	NULL	217000	74.7	1941.00	NULL
ARE	United Arab Emirates	Asia	Middle East	83600.00	1971	2441000	74.1	37966.00	36846.00
ARG	Argentina	South America	South America	2780400.00	1816	37032000	75.1	340238.00	323310.00
ARM	Armenia	Asia	Middle East	29800.00	1991	3520000	66.4	1813.00	1627.00
ASM	American Samoa	Oceania	Polynesia	199.00	NULL	68000	75.1	334.00	NULL
ATA	Antarctica	Antarctica	Antarctica	13120000.00	NULL	0	NULL	0.00	NULL
ATF	French Southern ter...	Antarctica	Antarctica	7780.00	NULL	0	NULL	0.00	NULL
ATG	Antigua and Barbuda	North America	Caribbean	442.00	1981	68000	70.5	612.00	584.00
AUS	Australia	Oceania	Australia and New Zealand	7741220.00	1901	18886000	79.8	351182.00	392911.00
AUT	Austria	Europe	Western Europe	83859.00	1918	8091800	77.7	211860.00	206025.00
AZE	Azerbaijan	Asia	Middle East	86600.00	1991	7734000	62.9	4127.00	4100.00
BDI	Burundi	Africa	Eastern Africa	27834.00	1962	6695000	46.2	903.00	982.00
BEL	Belgium	Europe	Western Europe	30518.00	1830	10239000	77.8	249704.00	243948.00
BEN	Benin	Africa	Western Africa	112622.00	1960	6097000	50.2	2357.00	2141.00



World database : <https://dev.mysql.com/doc/world-setup/en/>

# RDBMS

## MySQL

⌵ (C:) > Program Files > MySQL > MySQL Server 5.7 > bin

	Name	Date modified	Type	Size
★	innochecksum.exe	12/28/2017 5:03 AM	Application	4,717 KB
★	libmecab.dll	11/16/2017 4:51 PM	Application extens...	1,797 KB
★	lz4_decompress.exe	12/28/2017 5:03 AM	Application	105 KB
★	my_print_defaults.exe	12/28/2017 5:03 AM	Application	4,190 KB
★	mysam_ftdump.exe	12/28/2017 5:04 AM	Application	4,487 KB
	mysamchk.exe	12/28/2017 5:04 AM	Application	4,599 KB
	mysamlog.exe	12/28/2017 5:04 AM	Application	4,437 KB
	mysampack.exe	12/28/2017 5:04 AM	Application	4,511 KB
	mysql.exe	12/28/2017 5:04 AM	Application	4,958 KB
	mysql_config.pl	12/28/2017 4:59 AM	Perl Source File	8 KB
	mysql_config_editor.exe	12/28/2017 5:04 AM	Application	4,594 KB
	mysql_embedded.exe	12/28/2017 5:06 AM	Application	24,144 KB
	mysql_plugin.exe	12/28/2017 5:04 AM	Application	4,204 KB
	mysql_secure_installation.exe	12/28/2017 5:04 AM	Application	4,829 KB
	mysql_ssl_rsa_setup.exe	12/28/2017 5:03 AM	Application	4,272 KB
	mysql_tzinfo_to_sql.exe	12/28/2017 5:03 AM	Application	4,129 KB
	mysql_upgrade.exe	12/28/2017 5:04 AM	Application	5,631 KB
	mysqldadmin.exe	12/28/2017 5:04 AM	Application	4,844 KB
	mysqldbinlog.exe	12/28/2017 5:04 AM	Application	5,131 KB
	mysqlcheck.exe	12/28/2017 5:04 AM	Application	4,872 KB
	mysqld.exe	12/28/2017 5:06 AM	Application	38,625 KB
	mysqld_multi.pl	12/28/2017 4:59 AM	Perl Source File	28 KB
	mysqldump.exe	12/28/2017 5:04 AM	Application	4,911 KB
	mysqldumpslow.pl	12/28/2017 4:59 AM	Perl Source File	8 KB
	mysqlimport.exe	12/28/2017 5:04 AM	Application	4,834 KB
	mysqlpump.exe	12/28/2017 5:04 AM	Application	5,566 KB
	mysqlshow.exe	12/28/2017 5:04 AM	Application	4,831 KB
	mysqlslap.exe	12/28/2017 5:04 AM	Application	4,852 KB
	mysqlxtest.exe	12/28/2017 5:04 AM	Application	7,190 KB
	perror.exe	12/28/2017 5:03 AM	Application	4,355 KB
	resolveip.exe	12/28/2017 5:03 AM	Application	4,188 KB

## PostgresSQL

> PostgreSQL > pg96 > bin

	Name	Date modified	Type	Size
★	libgcc_s_seh-1.dll	5/9/2018 4:03 PM	Application extens...	80 KB
★	libiconv-2.dll	5/9/2018 4:03 PM	Application extens...	1,038 KB
★	libintl-8.dll	5/9/2018 4:03 PM	Application extens...	129 KB
★	libpgtypes.dll	5/9/2018 4:01 PM	Application extens...	285 KB
★	libpq.dll	5/9/2018 4:01 PM	Application extens...	338 KB
	libwinpthread-1.dll	5/9/2018 4:03 PM	Application extens...	88 KB
	libxml2-2.dll	5/9/2018 4:03 PM	Application extens...	5,746 KB
	libxslt-1.dll	5/9/2018 4:03 PM	Application extens...	1,216 KB
	oid2name.exe	5/9/2018 4:03 PM	Application	199 KB
	pg_archivecleanup.exe	5/9/2018 4:01 PM	Application	210 KB
	pg_basebackup.exe	5/9/2018 4:01 PM	Application	397 KB
	pg_config.exe	5/9/2018 4:01 PM	Application	201 KB
	pg_controldata.exe	5/9/2018 4:01 PM	Application	232 KB
	pg_ctl.exe	5/9/2018 4:01 PM	Application	253 KB
	pg_dump.exe	5/9/2018 4:01 PM	Application	618 KB
	pg_dumpall.exe	5/9/2018 4:01 PM	Application	285 KB
	pg_isready.exe	5/9/2018 4:01 PM	Application	248 KB
	pg_receivevlog.exe	5/9/2018 4:01 PM	Application	265 KB
	pg_recvlogical.exe	5/9/2018 4:01 PM	Application	259 KB
	pg_resetxlog.exe	5/9/2018 4:01 PM	Application	258 KB
	pg_restore.exe	5/9/2018 4:01 PM	Application	385 KB
	pg_rewind.exe	5/9/2018 4:01 PM	Application	297 KB
	pg_standby.exe	5/9/2018 4:03 PM	Application	216 KB
	pg_test_fsync.exe	5/9/2018 4:01 PM	Application	214 KB
	pg_test_timing.exe	5/9/2018 4:01 PM	Application	186 KB
	pg_upgrade.exe	5/9/2018 4:01 PM	Application	334 KB
	pg_xlogdump.exe	5/9/2018 4:01 PM	Application	262 KB
	pgbench.exe	5/9/2018 4:01 PM	Application	325 KB
	pltcl_delmod	5/9/2018 4:01 PM	File	3 KB
	pltcl_listmod	5/9/2018 4:01 PM	File	3 KB

# SQL

Article [Talk](#)

[Read](#)

[Edit](#)

[View history](#)

## Relational algebra

From Wikipedia, the free encyclopedia

*Not to be confused with [Relation algebra](#).*



This article's **lead section does not adequately summarize key points of its contents**. Please consider expanding the lead to [provide an accessible overview](#) of all important aspects of the article. Please discuss this issue on the article's [talk page](#). *(April 2015)*

**Relational algebra**, first created by [Edgar F. Codd](#) while at IBM, is a family of algebras with a [well-founded semantics](#) used for modelling the data stored in relational databases, and defining queries on it.

The main application of relational algebra is providing a theoretical foundation for [relational databases](#), particularly [query languages](#) for such databases, chief among which is [SQL](#).

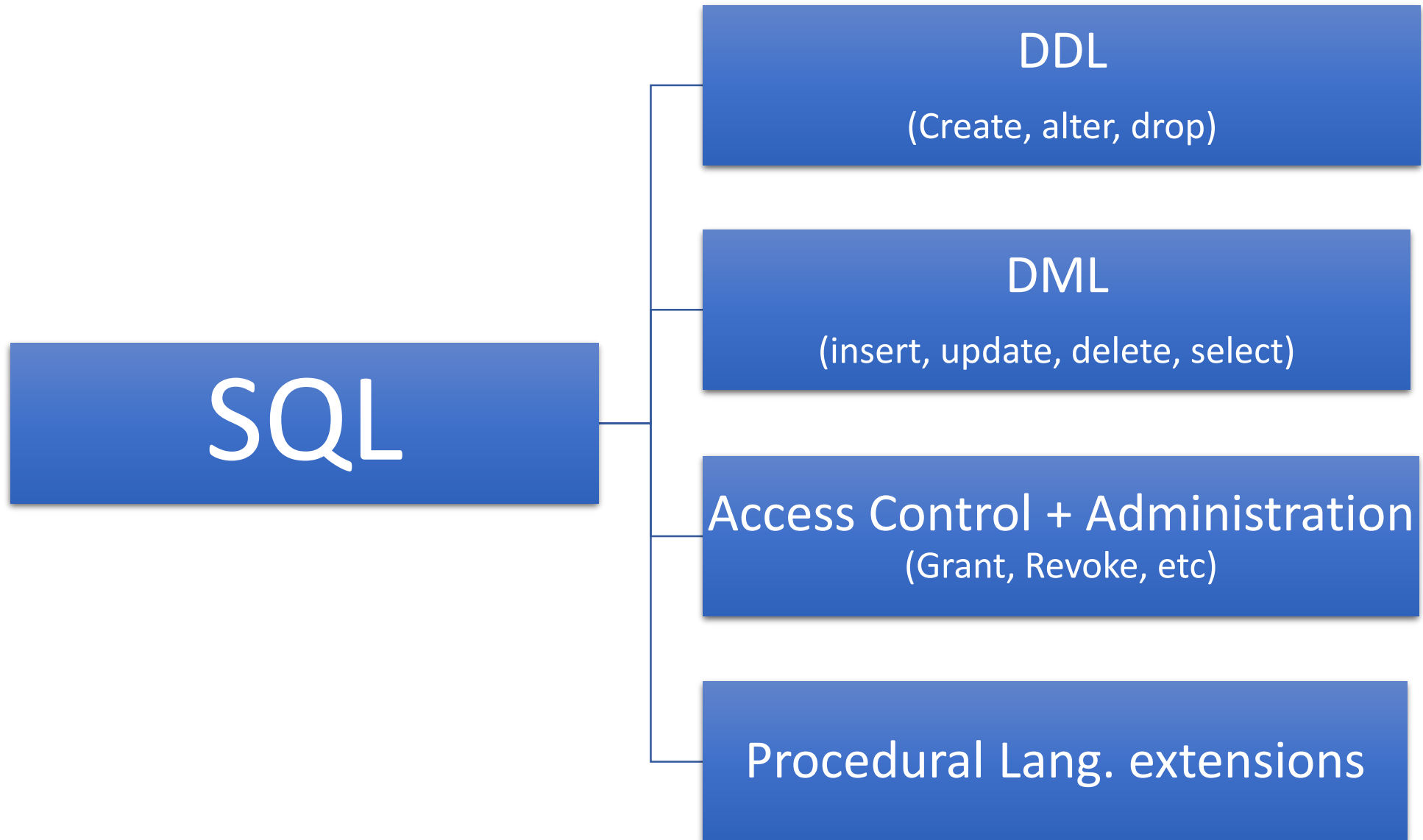
### Contents [\[hide\]](#)

#### 1 [Introduction](#)

##### 1.1 [Set operators](#)

##### 1.2 [Projection \( \$\Pi\$ \)](#)

# SQL



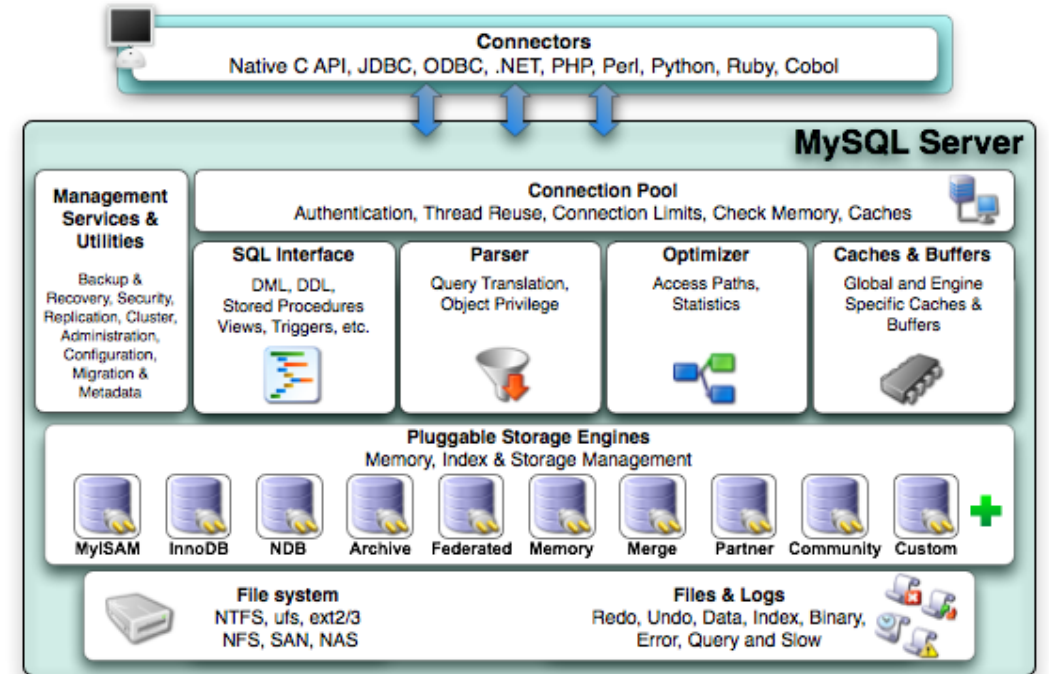
# SQL History

Year	Name	Alias	Comments
1986	SQL-86	SQL-87	First formalized by ANSI.
1989	SQL-89	FIPS 127-1	Minor revision that added integrity constraints, adopted as FIPS 127-1.
1992	SQL-92	SQL2, FIPS 127-2	Major revision (ISO 9075), <i>Entry Level</i> SQL-92 adopted as FIPS 127-2.
1999	SQL:1999	SQL3	Added regular expression matching, <a href="#">recursive queries</a> (e.g. <a href="#">transitive closure</a> ), <a href="#">triggers</a> , support for procedural and control-of-flow statements, non-scalar types (arrays), and some object-oriented features (e.g. <a href="#">structured types</a> ). Support for embedding SQL in Java ( <a href="#">SQL/OLB</a> ) and vice versa ( <a href="#">SQL/JRT</a> ).
2003	SQL:2003		Introduced <a href="#">XML</a> -related features ( <a href="#">SQL/XML</a> ), <i>window functions</i> , standardized sequences, and columns with auto-generated values (including identity-columns).
2006	SQL:2006		ISO/IEC 9075-14:2006 defines ways that SQL can be used with XML. It defines ways of importing and storing XML data in an SQL database, manipulating it within the database, and publishing both XML and conventional SQL-data in XML form. In addition, it lets applications integrate queries into their SQL code with <a href="#">XQuery</a> , the XML Query Language published by the World Wide Web Consortium ( <a href="#">W3C</a> ), to concurrently access ordinary SQL-data and XML documents. <sup>[34]</sup>
2008	SQL:2008		Legalizes ORDER BY outside cursor definitions. Adds INSTEAD OF triggers, TRUNCATE statement, <sup>[35]</sup> FETCH clause.
2011	SQL:2011		Adds temporal data (PERIOD FOR) <sup>[36]</sup> (more information at: <a href="#">Temporal database#History</a> ). Enhancements for <i>window functions</i> and FETCH clause. <sup>[37]</sup>
2016	SQL:2016		Adds row pattern matching, polymorphic table functions, JSON.



# MySQL

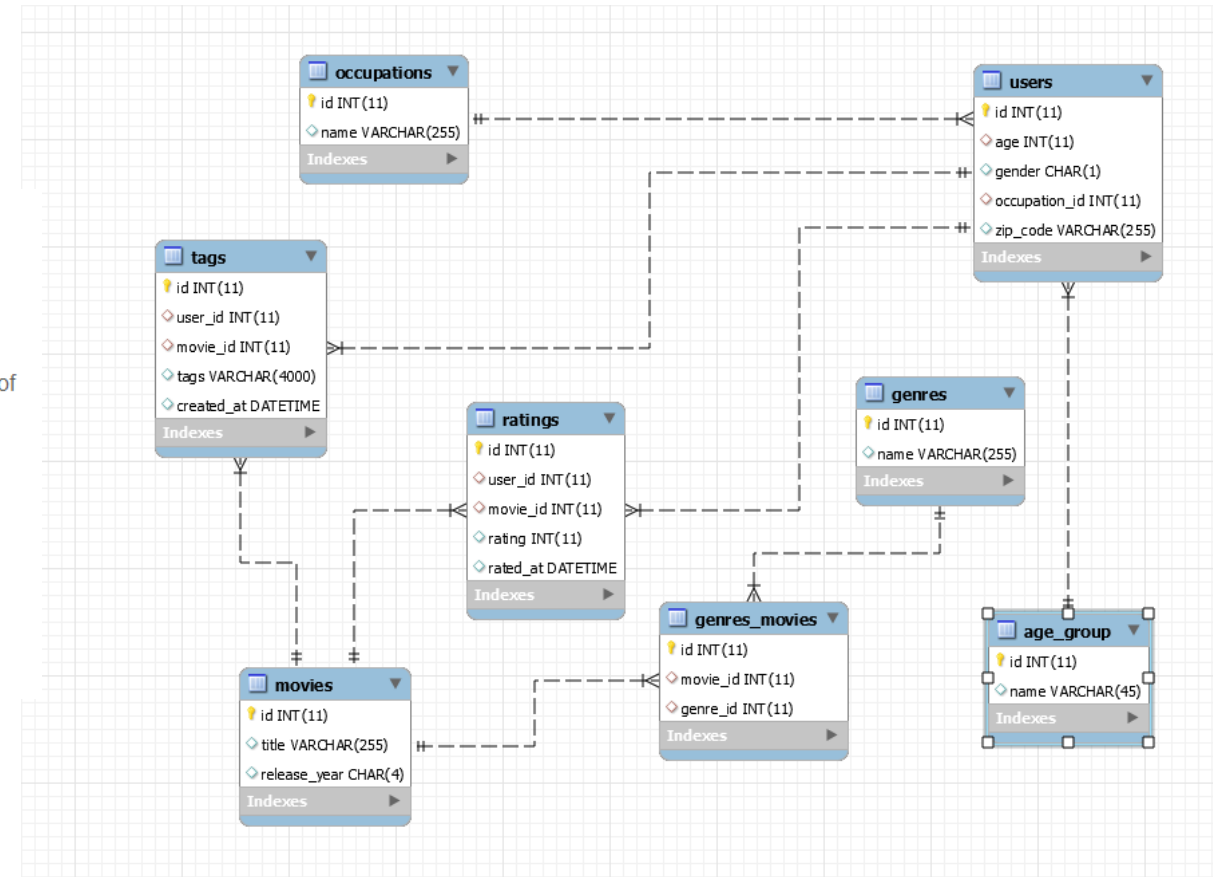
- ❖ Authored in 1995 by a Swedish company MySQL AB
- ❖ Currently developed by Oracle Incorporated
- ❖ Current version – 8.0.15 as at Feb. 2019
- ❖ According to db-engines
  - ❖ The second most used database in the world
  - ❖ The most deployed database in the world for web servers
- ❖ Written C and C++
- ❖ Vast language support
- ❖ Facebook has one of the known largest deployment of MySQL with over 1.3 billion users.  
<https://www.itworld.com/article/2831999/facebook--other-web-giants-unite-to-scale-mysql.html>
- ❖ Pluggable and flexible storage engine architecture. Supports a number out of the box
- ❖ Default storage engine today is InnoDB



# MovieLens Dataset

movielens

MovieLens is a web site that helps people find movies to watch. It has hundreds of thousands of registered users. We conduct online field experiments in MovieLens in the areas of automated content recommendation, recommendation interfaces, tagging-based recommenders and interfaces, member-maintained databases, and intelligent user interface design.



MovieLens dataset : <https://grouplens.org/datasets/movielens/> from GroupLens Research

# Data Processing in RDBMS

CRUD

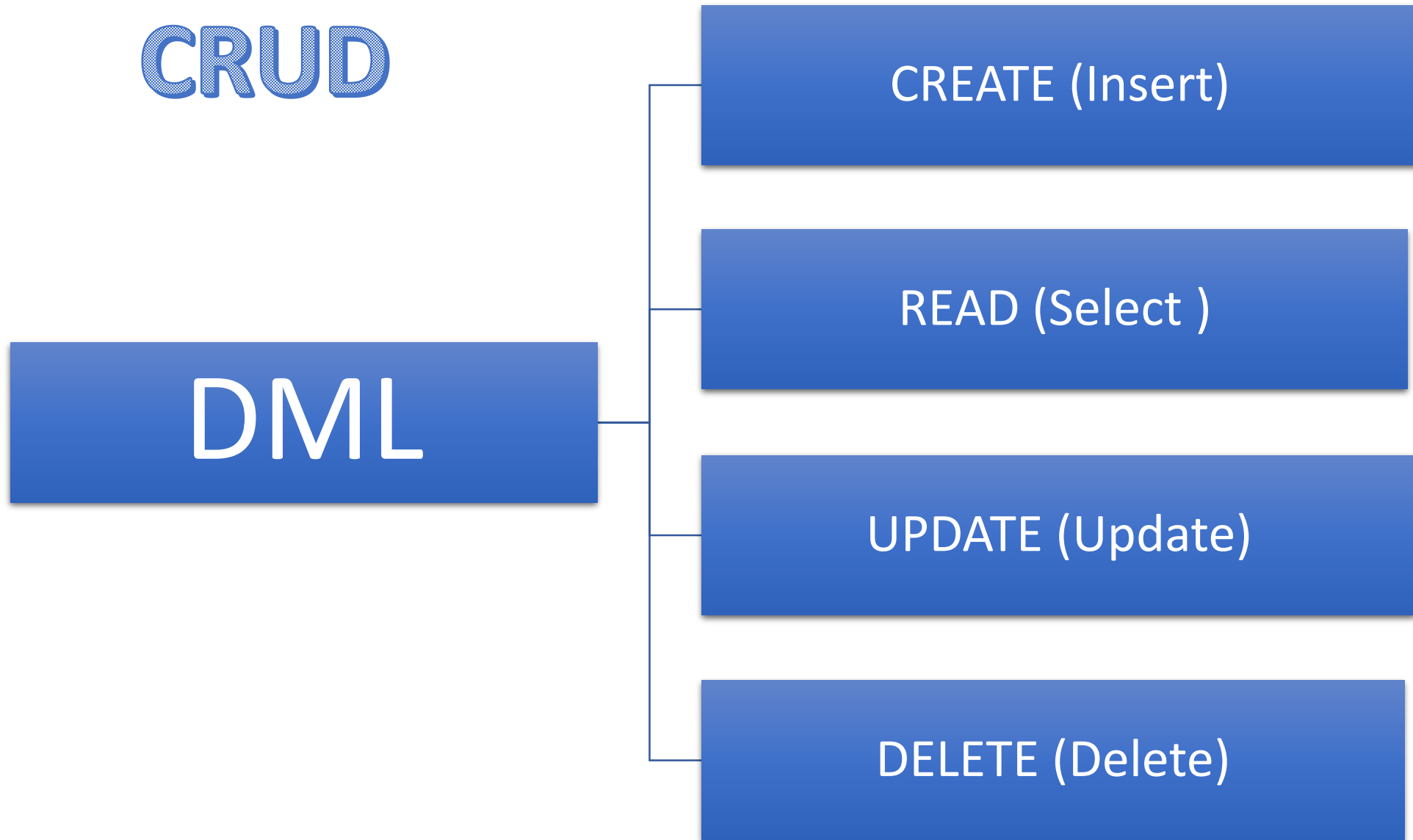
DML

CREATE (Insert)

READ (Select )

UPDATE (Update)

DELETE (Delete)



# Data Processing in RDBMS



## In the application space

Example is LAMP stack or Django

Application data structures

Data logic is implemented in a programming language

Prevalent in multi-tiered architecture

Suited for OLTP kind of application where the amount of data transferred is relatively small.



## In the database space

Example Microsoft Access Forms or Oracle Apex or Oracle Forms

Database objects

Data logic is implemented in the database procedure language

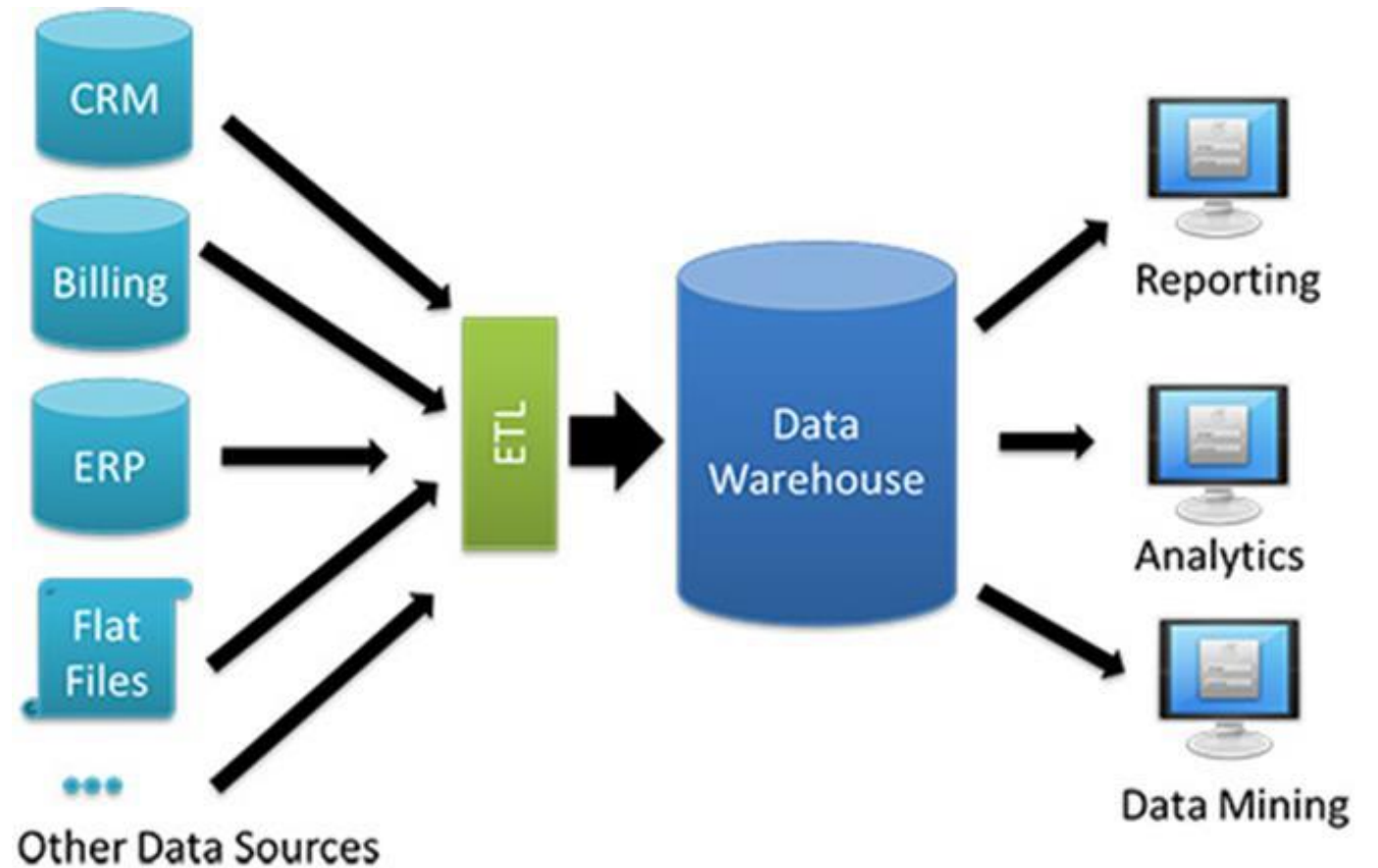
Prevalent in two-tiered architecture

Suited for data integration kind of work where the amount of data movement is usually large.



# Indexes

# Data Warehousing



# OLTP Challenges

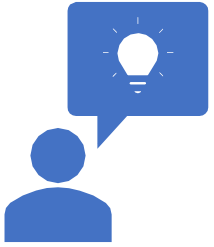
OLTP systems are not always good for handling aggregates over large amounts of data. OLTP should be doing quick short-live queries over transaction. Should not contend with OLAP requirements at the same time.

OLAP queries will tend to be quite complex when conducting analytics and reporting on data that is highly normalized.

Many times, we might want to integrate our business data with data from other sources. If we did this in OLTP, we will have a management issue of missing disparate data with our business data model.

Most OLTP are not designed to keep history. So performing analytics over historical data will face the shortcoming of having updated the historical data in an OLTP system.

As data begins to grow, we need different environments that we can tune separately for the various kinds of workload.



# Analytical Processing

---

- ☐ Aggregation Queries
- ☐ Window Queries

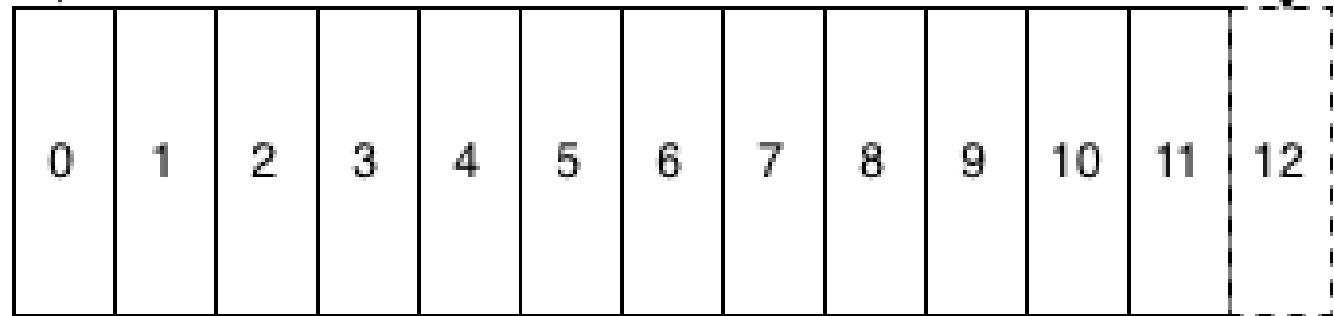





Logs

1st Record

Next  
Record  
Written



A thin vertical black line is positioned to the left of the text.

# Relational Database Wrap-up



**Gwen (Chen) Shapira**

@gwenshap

Follow



I thought that every engineer knows SQL. Today I discovered that since we hire people who spent most of their career optimizing storage drivers and kernels, they actually don't know SQL. I hope it isn't too challenging to learn. Should be a friendly language, right?

8:20 PM - 18 Jun 2019

10 Retweets 149 Likes



60 10 149



Tweet your reply



**Tanel Pöder** @TanelPoder · Jun 18

Replying to @gwenshap

Too friendly (makes me suspicious)

1 7



**Tanel Pöder** @TanelPoder · Jun 18

writing two books on SQL can't say I miss it. Who wants to write 5 lines of SQL when you can write 30 lines of JavaScript to do the same thing?

5 7



**Gwen (Chen) Shapira** @gwenshap · 20h

You mean 3000 lines in C++?

2



**db** @el\_db · Jun 18

Replying to @gwenshap

I would argue that In a world where analysts don't bother learning SQL (yes, it's a thing unfortunately) , engineers especially SREs will only pickup what's needed to solve a problem as itsbescakates to them.

I recall some tough mornings looking at redshift queries going WTF

6



**Timothy Perrett** @timperrett · Jun 18

Replying to @gwenshap

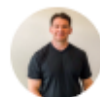
Friendly language with absolutely no edge cases or strange grammar, implemented in exactly the same manner by all vendors. #Sarcasm

1 16



**Ex falso quodlibet** @philderome · 21h

write once run everywhere #doublesarcasm (files, tasks)



**Nick Heudecker** @nheudecker · Jun 18

Replying to @gwenshap

Many new devs don't want to learn SQL because they (incorrectly) see it as a