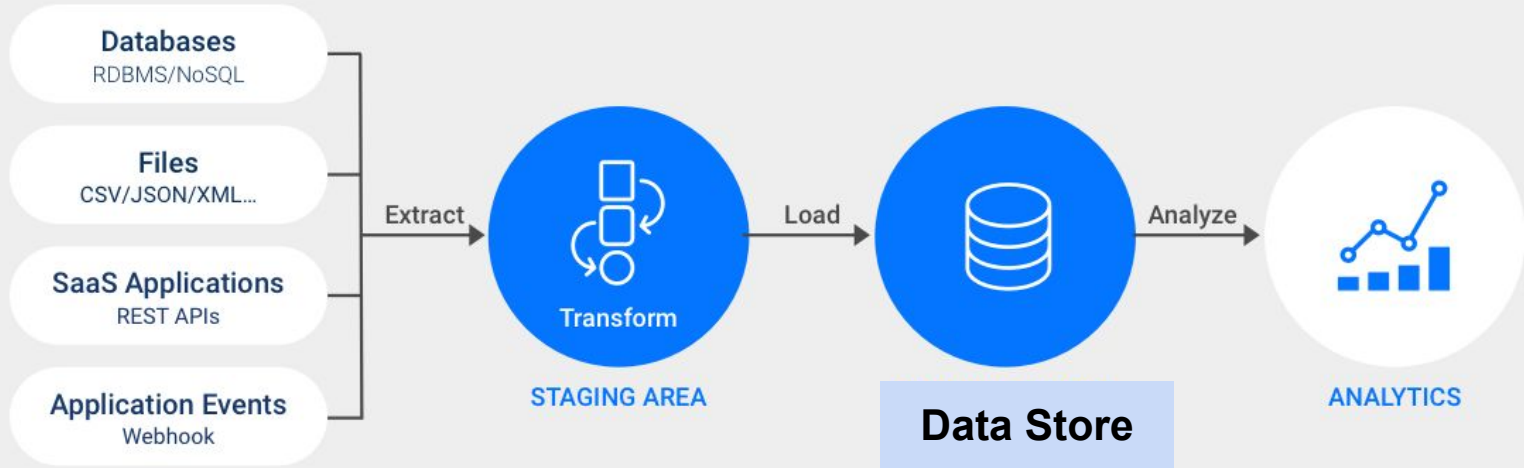# Data Organization: Databases and SQL

DATABASE

ETL PROCESS

Database: *Organized collection of (structured) data to make it easily accessible and manageable for users*

ETL PROCESS

Databases
RDBMS/NoSQL

Files
CSV/JSON/XML...

SaaS Applications
REST APIs

Application Events
Webhook

Extract

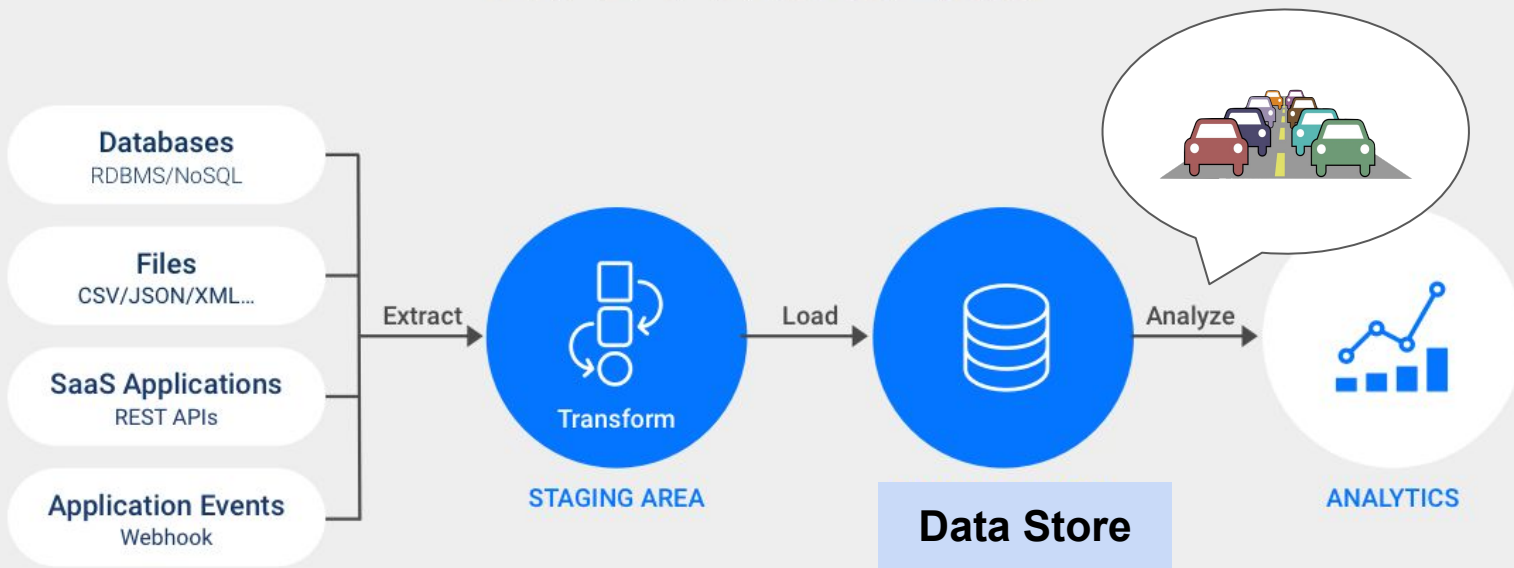**Transform**

STAGING AREA
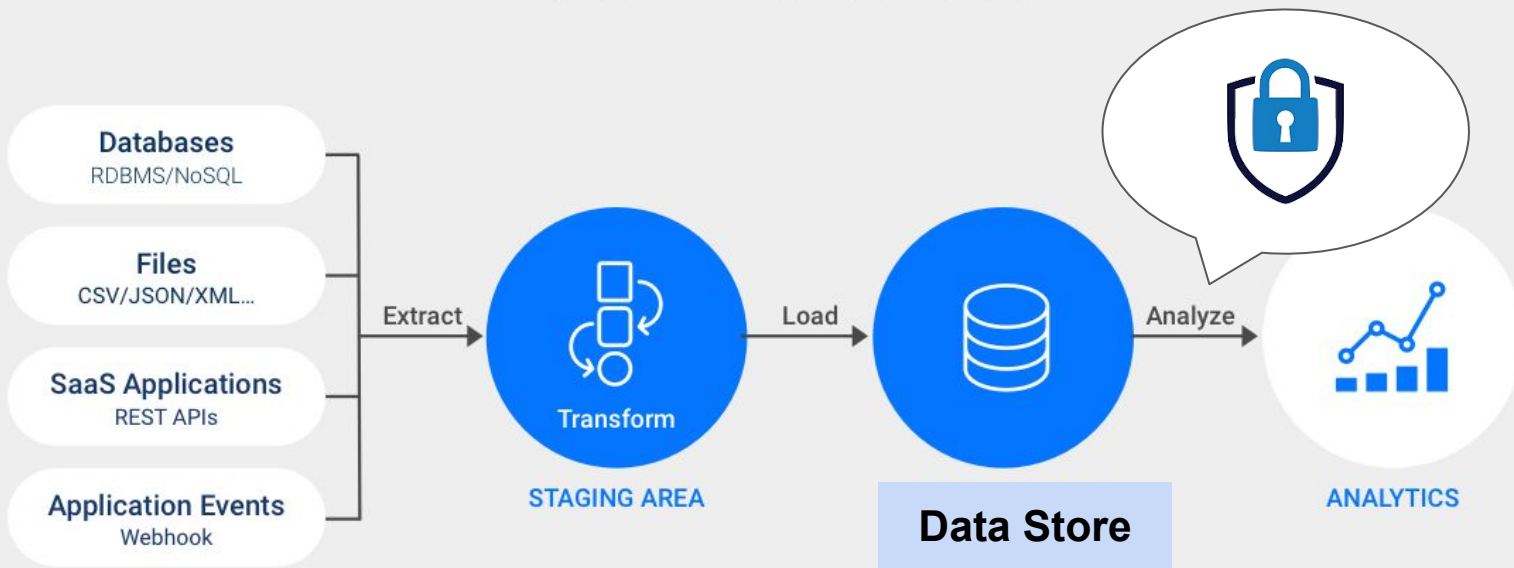
Load

Data Store

Analyze

ANALYTICS

Database: *Organized collection of (structured) data to make it easily accessible and manageable for users*

ETL PROCESS

Database: *Organized collection of (structured) data to make it easily accessible and manageable for users*

CREATE READ UPDATE DELETE

CRUD

Database

raw

structured

updated
frequently

current state

transactions

Database	Data Warehouse
raw	transformed
structured	structured
updated frequently	updated In batch
current state	historical
transactions	**analytics**

Database	Data Warehouse	Data Mart
raw	transformed	aggregated
structured	structured	structured
updated frequently	updated In batch	updated In batch
current state	historical	views
transactions	**analytics**	**scoped analytics**

Database	Data Warehouse	Data Mart	Data Lake
raw	transformed	aggregated	repository
structured	structured	structured	any
updated frequently	updated In batch	updated In batch	self-updating
current state	historical	views	all
transactions	**analytics**	**scoped analytics**	**archiving, machine learning, ??? etc.**

LANGUAGE

Language

- The principal method of human communication, consisting of words used in a structured and conventional way and conveyed by speech, writing, or gesture (*Oxford*)

- A system of communication which consists of a set of sounds and written symbols which are used by the people of a particular country or region for talking or writing. (*Collins*)

- Language is the use of a system of communication which consists of a set of sounds or written symbols. (*Collins*)

BRITISH VS AMERICAN SPELLING
FOOD EDITION!

British	American
aubergine	eggplant
biscuit	cookie
candy floss	cotton candy
chips	french fries
crisps	potato chips
maize	corn
coriander	cilantro
spring onion	green onion
soft drink	soda
sweets	candy

CREEDO
@Alonzo_Creed

The five love languages:

Words of Affirmation: This is a good burrito.

Acts of Service: I made you a burrito.

Receiving Gifts: Here's a burrito.

Quality time: Let's go get some burritos together.

Physical touch: Arms around a person wrapped in a warm hug, like a burrito.

CREEDO
@Alonzo_Creed

The five love languages:

Words of Affirmation: This is a good burrito.

Acts of Service: I made you a burrito.

Receiving Gifts: Here's a burrito.

Quality time: Let's go get some burritos together.

Physical touch: Arms around a person wrapped in a warm hug, like a burrito.

HANDS

The **American** "goodbye" wave can be interpreted as the signal for "no" in many parts of **Europe** and **Latin America**.

The **Italian** goodbye wave can be interpreted by **Americans** as "come here." The **American** "come here" gesture can be seen as an insult in most **Asian** countries.

In **Malaysia** it is taboo to point with your index finger, but you can point with your thumb. "Thumbs up" is used in many cultures, none more so than **Brazil** however,

...where the signal is used everywhere. However, it's a rude sexual signal in some **Islamic** countries, **Sardinia**, and **Greece**, and can signify the number "1" in **France**.

Hello World in Different
Programming Languages

- **Python:**

print("Hello World")

- **PHP :**

<?php echo "Hello World"; ?>

- **Javascript :**

console.log("Hello World")

- **LISP :**

(print "Hello World")

- **Swift:**

print("Hello World")

- **C++:**

int main() { std::cout<<"HelloWorld"; return 0; }

- **Ruby :**

puts "Hello World"

- **R :**

print("Hello World")

- **MatLab :**

disp('Hello, world!')

SQL

Structured Query Language

SQL or MySQL?

SQL (Structured Query Language)

- SQL is a syntax, a language used to communicate with the database
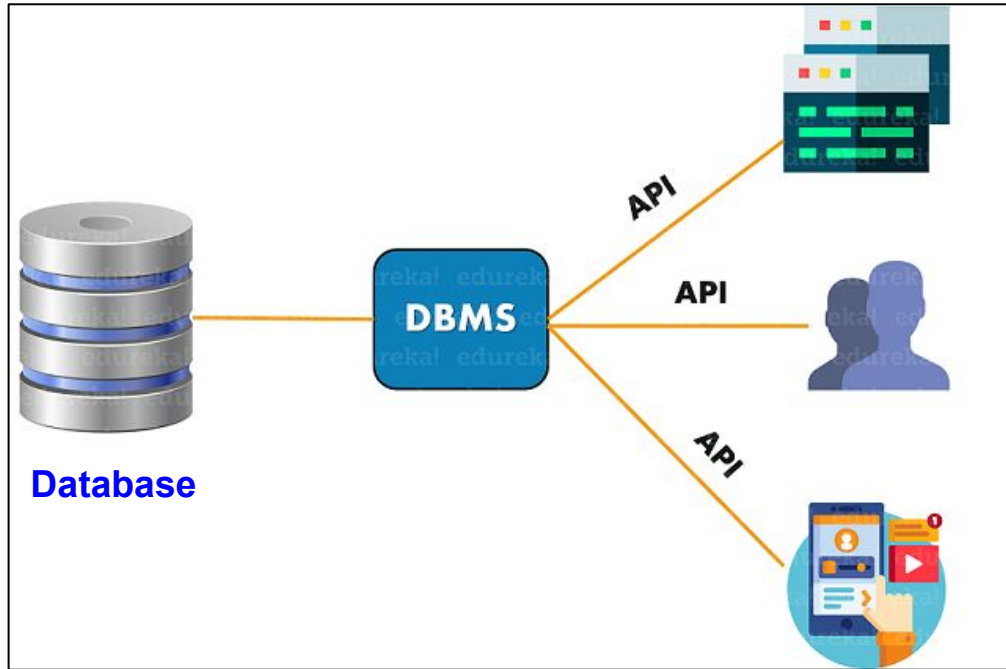
Relational Database

- Organization of data by tables. A **table** is a collection of related data entries and consists of numerous columns and rows
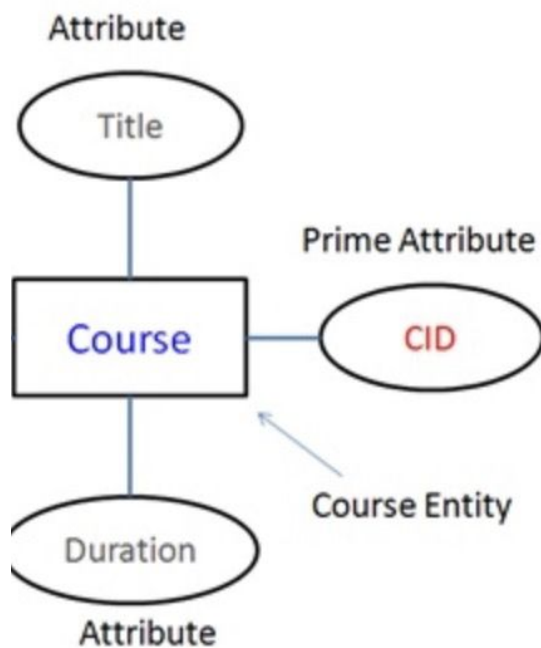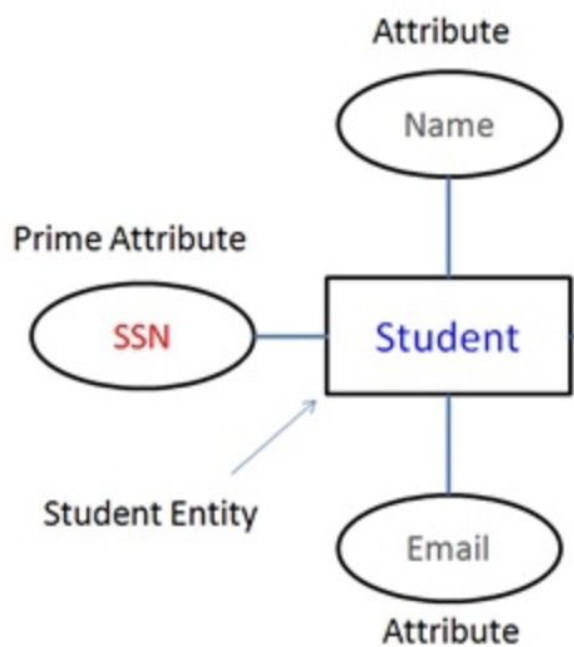
MySQL

- Open-source RDBMS (Relational Database Management System)
- Uses the SQL language to query the database
- Other popular RDBMS - SQL Server, Oracle DB, PostgreSQL
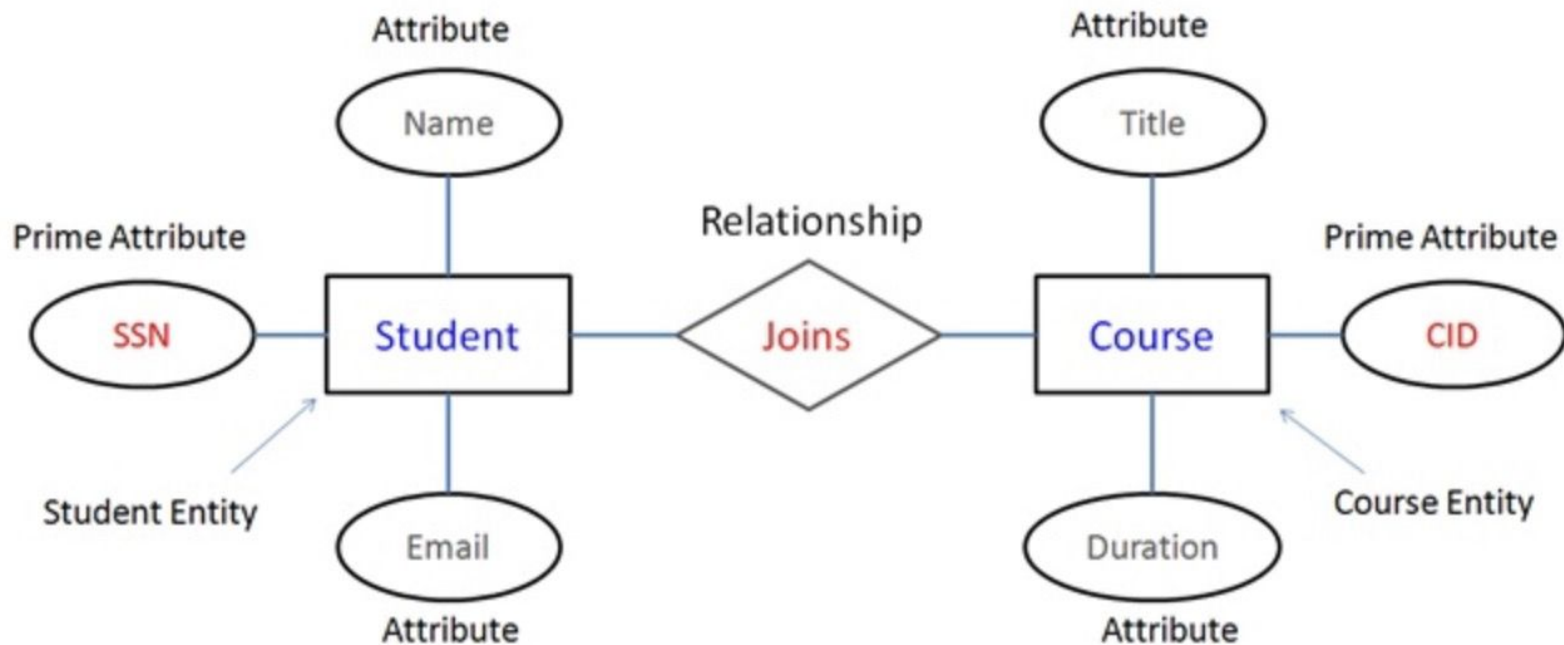
MySQL RDBMS Ecosystem

DBMS provides a variety of administrative operations such as session management, tuning, performance monitoring, and backup recovery

Entity Relationship Diagram (ERD)

Entity Relationship Diagram (ERD)

Attribute — Name

Attribute — Title

Prime Attribute — SSN

Relationship — Joins

Prime Attribute — CID

Student Entity

Student

Course

Course Entity

Email — Attribute

Duration — Attribute

Customer / Product / Ordering Table

customer

cid	name	city	discount	contact
c001	TipTop	London	10.00	Mr Smith
c002	Basic	Paris	12.00	M. Parot
c003	Allied	London	8.00	Ms Evans

product

pid	name	quantity	price
p01	comb	11140	0.50
p02	brush	30300	0.50
p03	razor	15060	1.00

Customer / Product / Ordering Table

customer

cid	name	city	discount	contact
c001	TipTop	London	10.00	Mr Smith
c002	Basic	Paris	12.00	M. Parot
c003	Allied	London	8.00	Ms Evans

product

pid	name	quantity	price
p01	comb	11140	0.50
p02	brush	30300	0.50
p03	razor	15060	1.00

ordering

ordno	month	quantity	cid	pid
1011	Jan	1000	c001	p01
1012	Feb	1000	c002	p03
1013	Jun	1000	c003	p03

references to customer and product table
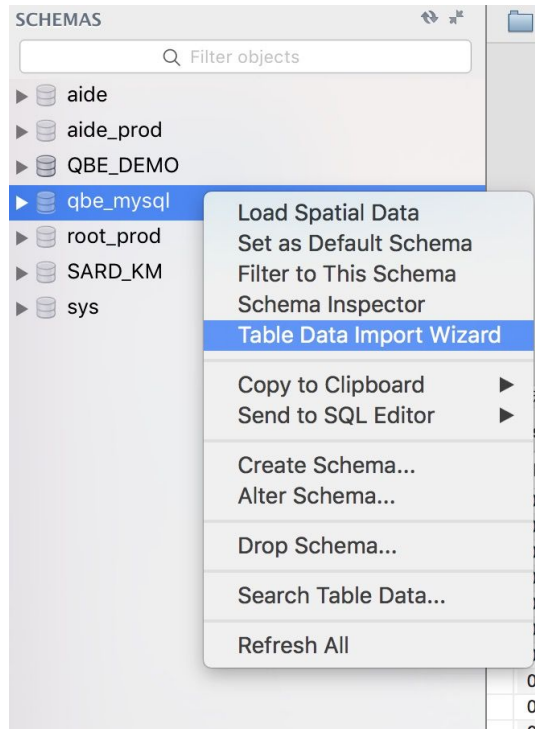
3

MySQL Hands-on

MySQL Basics

TMDB Dataset (5000 movie dataset)

MySQL Basics Exercises

Table Data Import

1. Right click on your created database
2. Select Table Data Import Wizard