Roles in DBS

User	Database Administrator	Programmer (software engineer)	(database, system) Designer
Data Engineer	Data Analyst (*Data Scientist*)	Product Manager	Machine learning engineer

Types of Database

The Relational Database Model

The dominant database model is the **relational database model**—all current major DBMS products are based on it.

It was created by IBM engineer **E. F. Codd** in 1970.

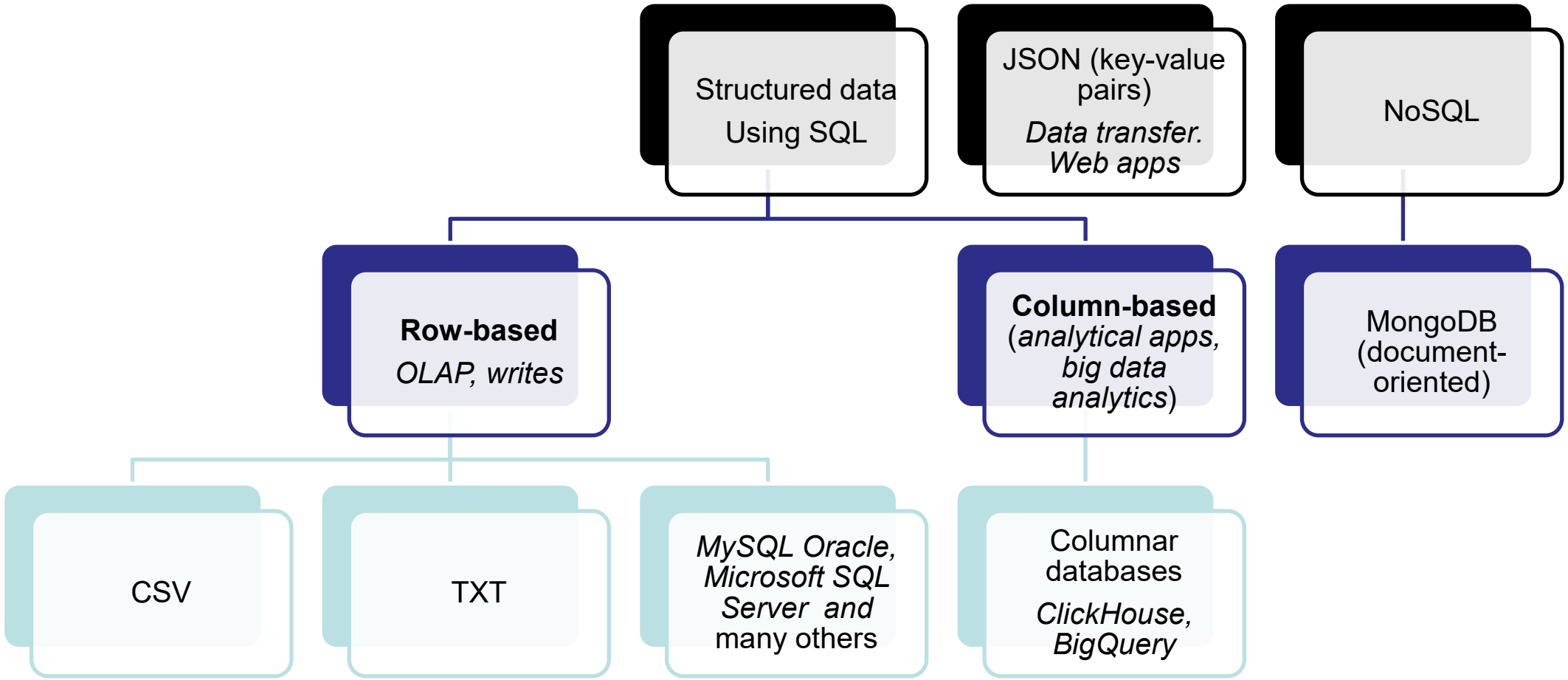
It was based on mathematics called **relational algebra**.

The text mainly examines and explains the relational database model.

The NoSQL Movement and Big Data

- Recent developments in Internet and mobile computing have resulted in the development of non-relational DBMSs.
  - NoSQL movement
  - Big Data
- These do not replace the relational model, but rather complement it.
- These topics are discussed in Chapter 12 and Appendix K.

Data stores in a variety of formats.



Columnar databases for big data

ClickHouse

- By Yandex in 2016
- A fast open-source OLAP database management system.
- It is column-oriented and allows to generate analytical reports using SQL queries in real-time.

Amazon Redshift

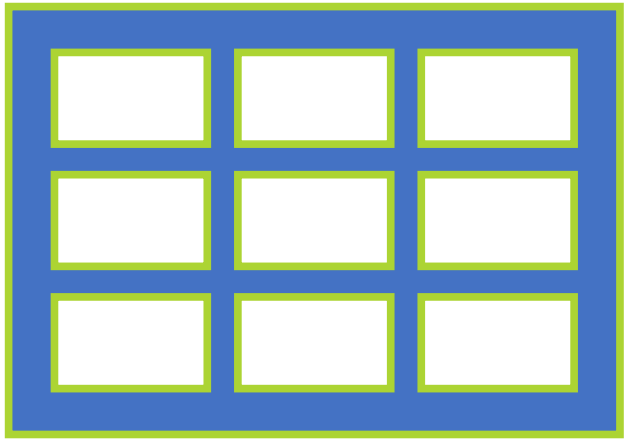
- Amazon Redshift is a data warehouse product which forms part of the larger cloud-computing platform Amazon Web Services.
- The name means to shift away from Oracle, red being an allusion to Oracle, whose corporate color is red and is informally referred to as "Big Red."

Google BigQuery

- BigQuery is a fully-managed, serverless data warehouse that enables scalable analysis over petabytes of data.

Acronym
OLAP

- **OLAP** is an acronym for **Online Analytical Processing**.
- **OLAP** performs multidimensional analysis of business data and provides the capability for complex calculations, trend analysis, and sophisticated data modeling.



Row-based vs. Column-based

- Store data by rows or by columns together
- Indexed by row or by column

- Transaction-focused
  - Row-based database
- Online analytics-focused
  - Columnar database

In-process, Embedded Database

- Database management systems (DBMS) built or integrated into an application, effectively hiding or minimizing interaction with the database by an application's end users.
  - Dramatically reduce latency and network I/O load; faster responses.
  - A better user experience since the database is invisible to the user.

In-memory Database

- A database management system that primarily relies on main memory for computer data storage
- Volatile storage

SQLite

- https://www.sqlite.org/
  - The most used database engine in the world (2021-6-18)
- A *C-language library* that implements a small, fast, self-contained, high-reliability, full-featured, SQL database engine.
- built into all cell phones and most computers
- bundled inside countless other applications

When to use SQLite?

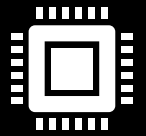
- NOT client/server SQL database engines like:
  - *Client/Server engines*: MySQL, Oracle, PostgreSQL, SQL Server
- https://www.sqlite.org/whentouse.html
  - The Internet of Things (IoT): Embedded devices, In-process databases, In-memory databases
  - Cellphones, game consoles, cameras…
  - Testing, demos

Acronym

- ACID
- ETL
- EDA
- RDBMS
- ERD
- CRUD
- KISS
- DRY

What is SQL and MySQL?

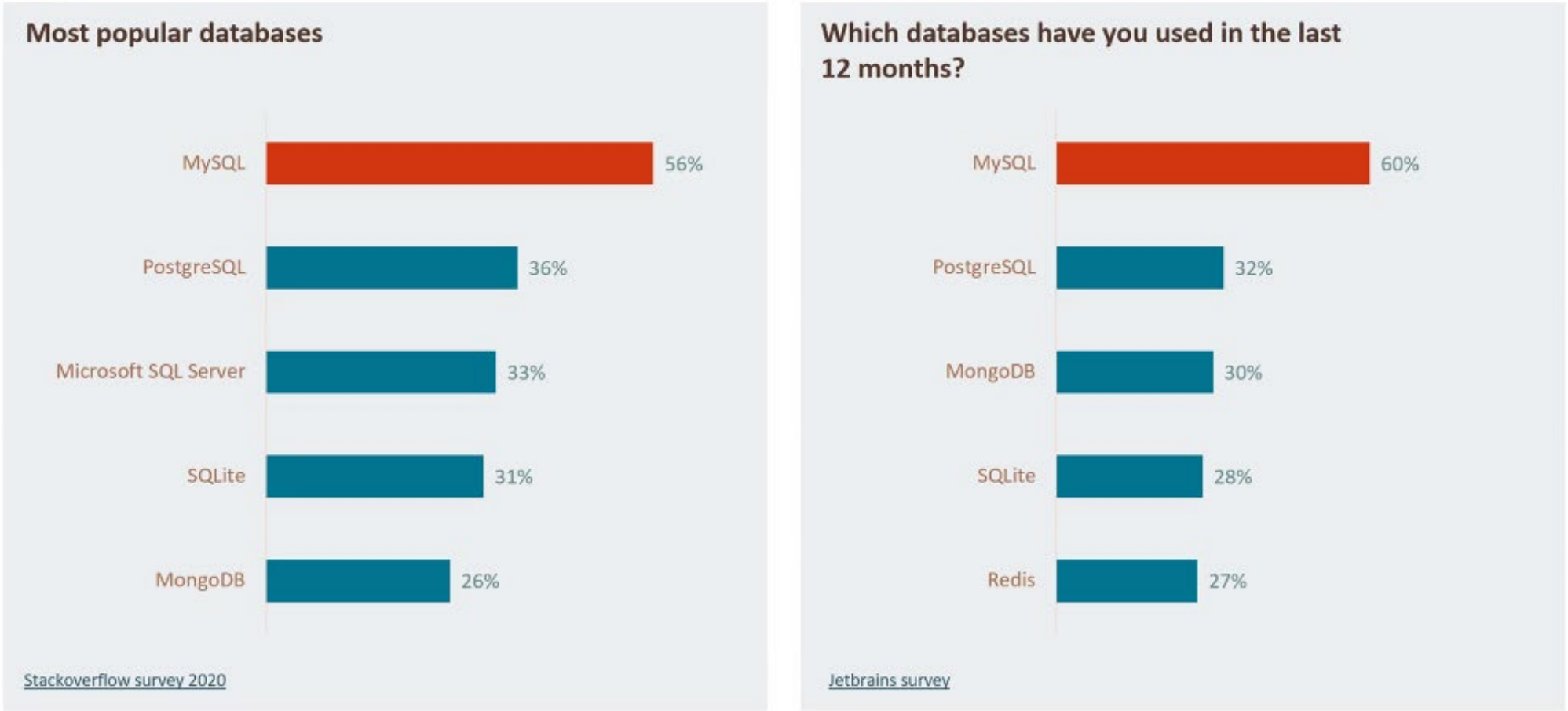
SQL = Structured Query Language



MySQL – world's largest open-source SQL platform

Allows us to store structured data and retrieve the data from it using SQL

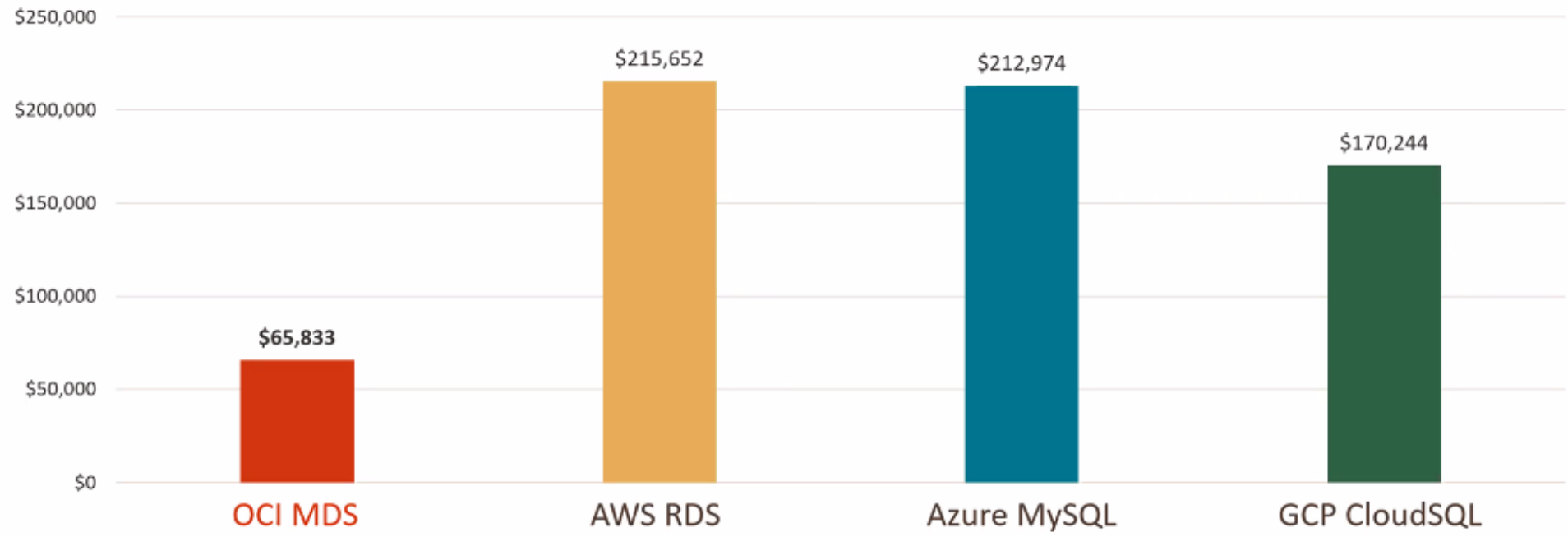
MySQL is the most popular database for developers



Copyright © 2021 Oracle and/or its affiliates.

MySQL Database Service Costs Less

Annual cost for 100 OCPUs, 1TB Storage configuration



Configuration: 100 OCPUs, 1 TB Storage
MySQL Database Service: Standard E3 AMD 16GB/Core, all regions have the same price
Amazon RDS: Intel R5 16GB/Core, AWS US East.
Azure: Memory Optimized Intel 20GB/Core, MS Azure US East.
Google: High Memory N1 Standard Intel 13GB/Core, GCP Northern Virginia.

Copyright © 2021 Oracle and/or its affiliates.

OCPU

An **OCPU** provides **CPU** capacity equivalent of one physical **core** of an Intel Xeon processor with hyper threading enabled.

SQL
Answer questions

Find		Select	Select	Select
Find all the customers	Find the customers matching certain criteria	Select * from customers where name like ‘peter’	Select * from customers where name like ‘peter%’	Select * from customers where name like ‘peter*’

Where is MySQL?

MySQL **server**

- Installed on most UNIX/Linux systems by default.
- can be installed on Mac or PC
  - MySQL (***community edition***)
  - Standard
  - enterprise

MySQL **client**

A software/tool that can be run using command-line or available through IDE like MySQL Workbench or Sequel Pro

- **MySQL Workbench** and many others
- You can develop your own client program.
  - Browse tables
  - Enter SQL select and run it
  - Display the query result
  - Even modify the database depending on the access rights.

What you need to know

- **MySQL database**
  - Query database
  - Design database
  - Create database
  - Populate database
  - Update database
  - Manage users
- Access MySQL in client programs: Java, Python, C++ and others
- Access MySQL in Web apps

They all belong to Oracle!

- Java
- MySQL database
- Oracle database

Software environment

- Submission required:
  - Install software
  - Besides, you need to duplicate the Cape Codd database in your local MySQL server

Connecting to MySQL

txt command

Using command-line:

- mysql -h <servername> -u <username> -p

GUI

Using IDE:

- MySQL Workbench
- Sequel Pro

Basic Operations (CRUD) by SQL

Create a new database (tables and relationships)

Import data to tables (populate tables)

Retrieve (Update, Delete) records from tables

Search data

Filter data

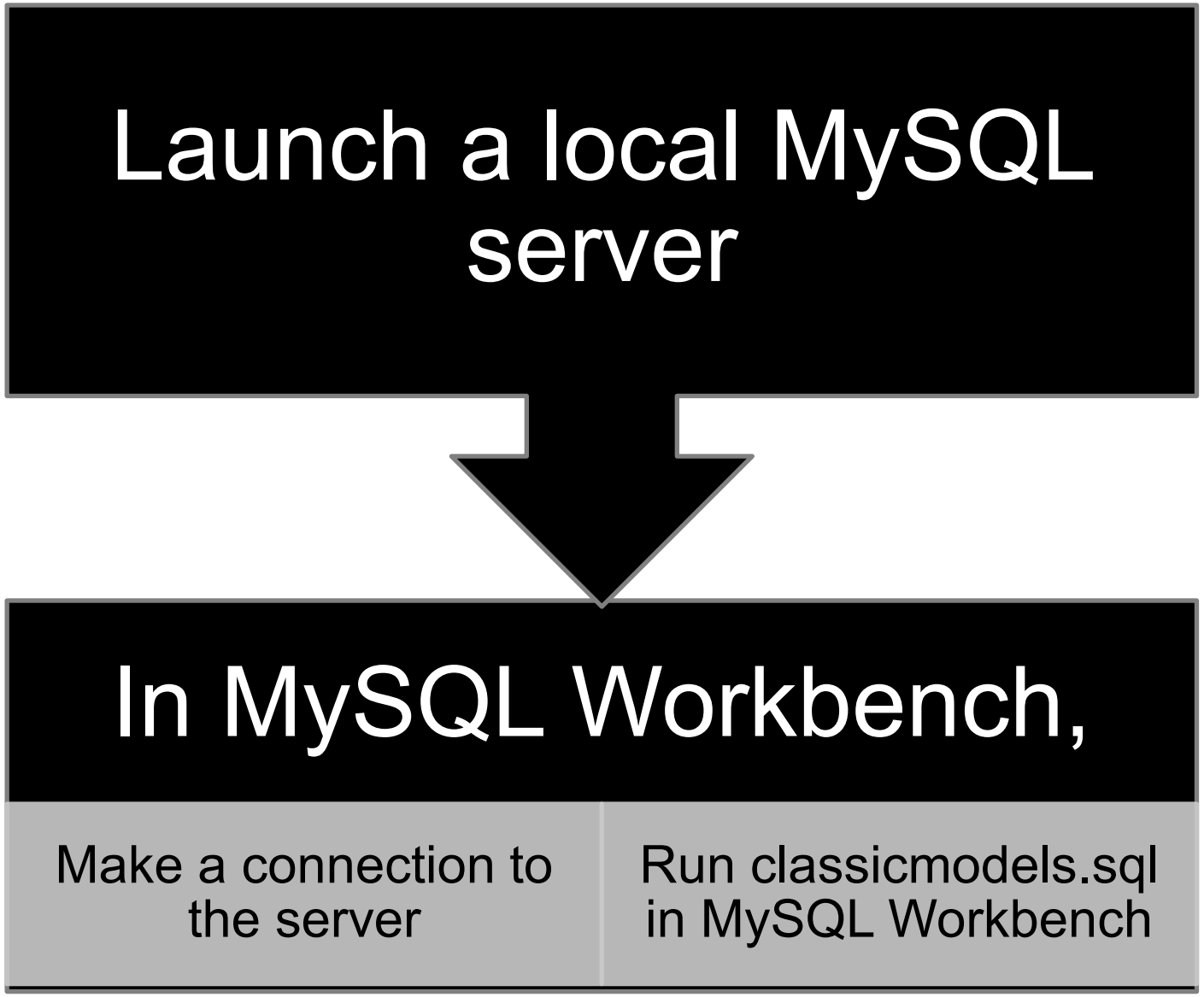
Database Operations vs. Rest API Methods

Rest API

- Application Programming Interface
- Representational State Transfer

CRUD:

- POST
- GET
- PUT
- DELETE



Create a database *from sql scripts*

classicmodels

Customers: stores customer data.

Products: stores a list of scale model cars.

ProductLines: stores a list of product line categories.

Orders: stores sales orders placed by customers.

OrderDetails: stores sales order line items for each sales order.

Payments: stores payments made by customers based on their accounts.

Employees: stores all employee information as well as the organization structure such as who reports to whom.

Offices: stores sales office data.