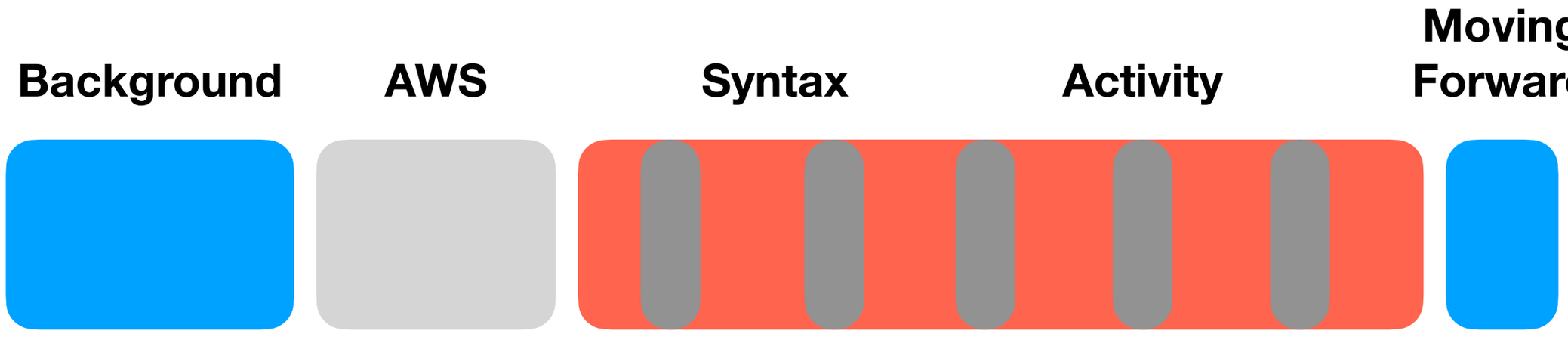
Introduction to SQL

https://github.com/la-process-and-theory/sql-db-setup

Today



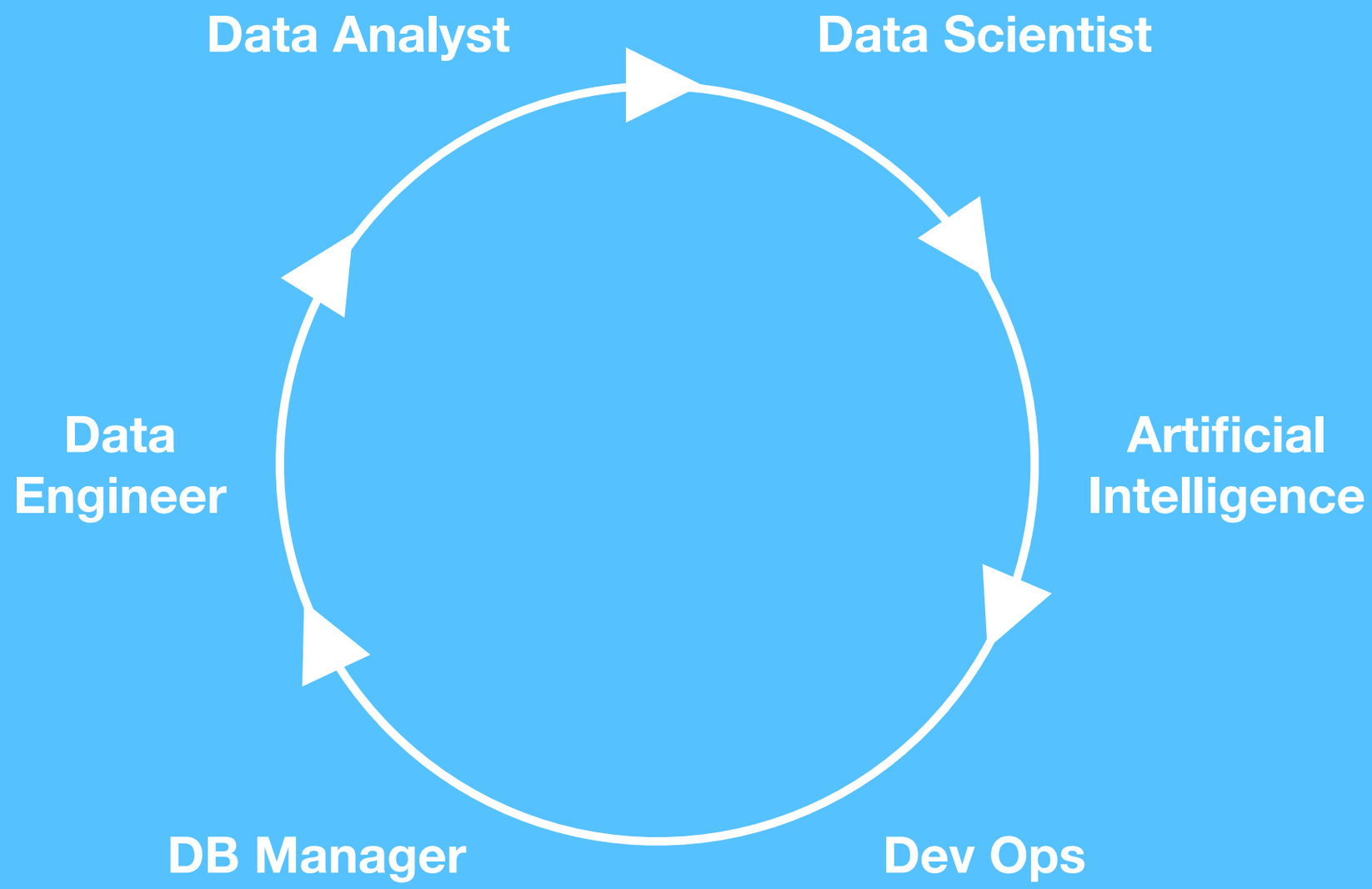
History
Purpose

Create DB

DBI
Creating
Reading
Updating
Deleting
NULL
Constraints
ID Columns
Filtering
Duplicates
Conditionals
Relationships

Resources

Data Science Workflow



DB Taxonomy

Relational (RDDBMS)

Old School

MySQL
Oracle
PostgreSQL

New School

Amazon Aurora
MySQL Cluster
Maria

Object-Oriented

Smalltalk

Non-Relational

NoSQL

Key-Value

Membrane

Document

MongoDB

Graph

InfiniteGraph

Distributed

Blockchain

Relational DBs

student — **Table**

student_id	username	email
1	AAA	emailA
2	BBB	emailB
3	CCC	emailC
4	DDD	emailD

Row/ Record

Unique Key/ Primary key*

Variable/ Column/ Field

lesson

id	student_id
Z	1
Z	2
Z	3
Z	4

Foreign key

game

student_id	lesson_id	level
1	Z	1
1	Z	2
1	Z	1
1	Z	1

* Can be hidden

Student - activities

studentAssessment
- id_student INT
- id_assessment INT
- date_submitted INT
- is_banked TINYINT
- score FLOAT

Module Presentation

assessments
- code_module VARCHAR(45)
- code_presentation VARCHAR(45)
- id_assessment INT
- assessment_type VARCHAR(45)
- date INT

studentRegistration
- code_module VARCHAR(45)
- code_presentation VARCHAR(45)
- id_student INT
- date_registration INT
- date_unregistration INT

studentVle
- id_site INT
- id_student INT
- code_module VARCHAR(45)
- code_presentation VARCHAR(45)
- date INT
- sum_click INT

courses
- code_module VARCHAR(45)
- code_presentation VARCHAR(45)

Student - demographics

studentInfo
- code_module VARCHAR(45)
- code_presentation VARCHAR(45)
- id_student INT
- gender VARCHAR(3)
- imd_band VARCHAR(16)
- highest_education VARCHAR(45)
- age_band VARCHAR(16)
- num_of_prev_attempts INT
- studied_credits INT
- region VARCHAR(45)
- disability VARCHAR(3)
- final_result VARCHAR(45)

vle
- id_site INT
- code_module VARCHAR(45)
- code_presentation VARCH...
- activity_type VARCHAR(45)
- week_from INT
- week_to INT

OU LA Data ER Diagram

DB Popularity

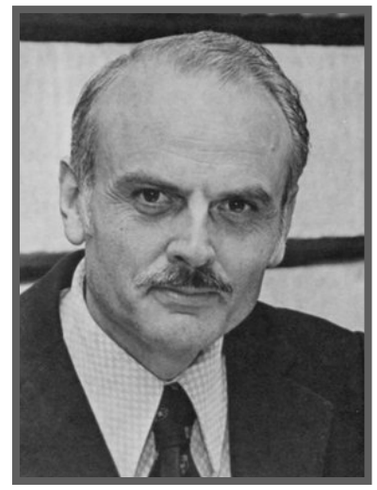
Rank			DBMS	Database Model	Score		
Oct 2019	Sep 2019	Oct 2018			Oct 2019	Sep 2019	Oct 2018
1.	1.	1.	Oracle	Relational, Multi-model	1355.88	+9.22	+36.61
2.	2.	2.	MySQL	Relational, Multi-model	1283.06	+3.99	+104.94
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	1094.72	+9.66	+36.39
4.	4.	4.	PostgreSQL	Relational, Multi-model	483.91	+1.66	+64.52
5.	5.	5.	MongoDB	Document, Multi-model	412.09	+2.03	+48.90
6.	6.	6.	IBM Db2	Relational, Multi-model	170.77	-0.79	-8.91
7.	7.	↑ 8.	Elasticsearch	Search engine, Multi-model	150.17	+0.90	+7.85
8.	8.	↓ 7.	Redis	Key-value, Multi-model	142.91	+1.01	-2.38
9.	9.	9.	Microsoft Access	Relational	131.18	-1.53	-5.62
10.	10.	10.	Cassandra	Wide column	123.22	-0.18	-0.17

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

Structured Query Language (SQL)

- Developed SEQUEL at IBM ~1970
- Domain-specific language for querying relational DBs
- Based on relational algebra
- Works with structured data
- Main benefit: don't need to specify how to reach a given record (no file path, index number)
- Is *loosely standardized* across products

Donald Chamberlin



Todd Codd

Ray Boyce

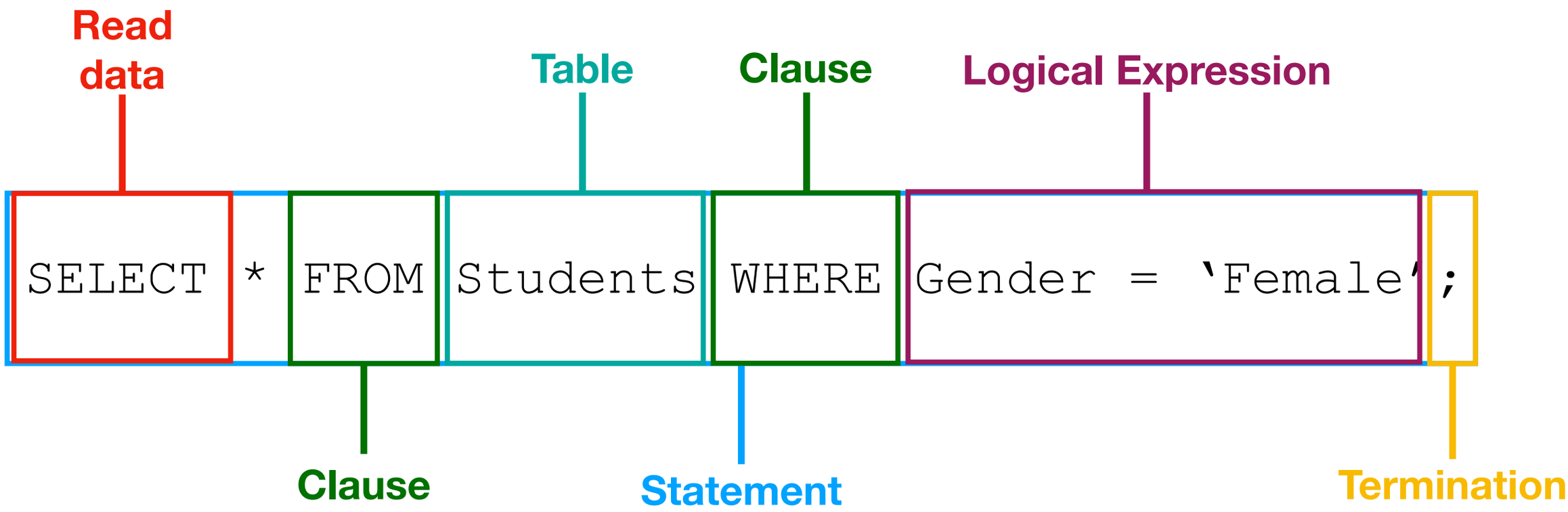
DB Fundamental Functions

- Create
- Read
- Update
- Delete

DB Fundamental Functions

- **C**reate: `INSERT` - Add rows
- **R**ead: `SELECT` - Get data
- **U**pdate: `UPDATE` - Change data
- **D**elete: `DELETE` - Remove rows

Basic Statement



***Capitalization of SQL keywords is not required but is useful**

Moving Forward

- SoloLearn
- LeetCode.com
- Test
- Projects

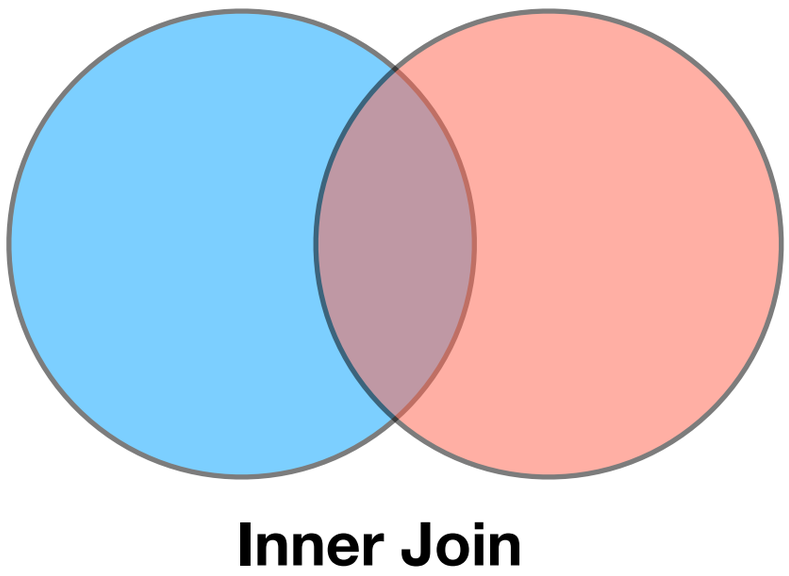
Create a MySQL DB

- Log into your AWS Management Console
- Locate `RDS` under the `Databases` heading
- Within Amazon RDS click `Create database`
- Under `Choose a database creation method` click `Standard Create`
- Under `Engine options` choose `MySQL`
- Under `Templates` choose `Free tier`
- Under `Settings` name your `DB instance identifier` as `database-1`
- Under `Credential settings` create a username and password combination and write it down (you will need it later)
- Under `Connectivity` expand `Additional connectivity configuration` to show additional menu items and make sure that `Publicly accessible` is checked `Yes`
- Expand the `Additional configuration` menu
- Under `Initial database name` write `oudb`
- Uncheck `Automatic backups`
- Click `Create database`

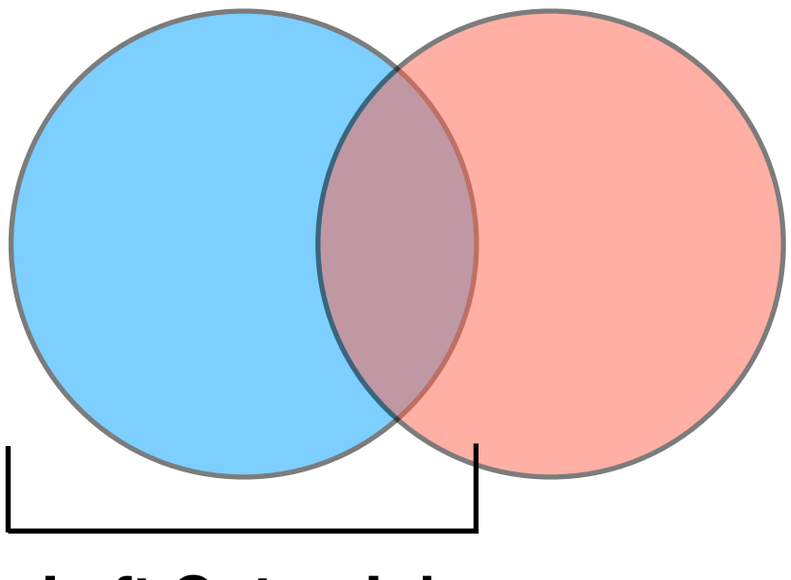
Create a MySQL DB

- Once your DB has been created
- Under `Security groups rules` click the hyperlink
- Click `Inbound` and then `Edit`
- Do not delete any rules!
- Add the rule `SQL/Aurora` on `Port 3306` with the `Connection` of `MyIP`

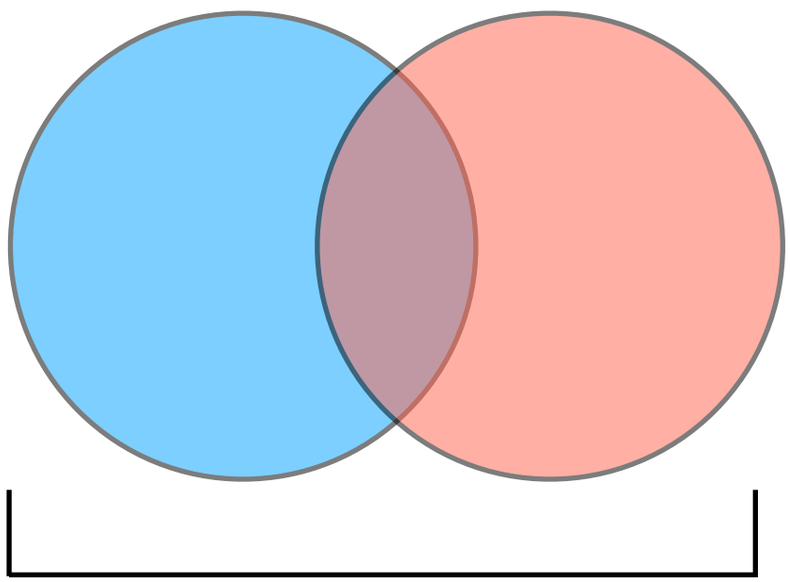
Joins



Inner Join



Left Outer Join



Full Join