AN INTRODUCTION TO SQL

Claremont Mckenna College

Pre-Requisite

- Workshop Environment
- Dbeaver Sqlite GUI environment
  - https://dbeaver.io/
  - Download and install Community version on local machine (installer available for Windows, Mac and Linux)

Outline

- What is SQL
- Creating a table and inserting data
- Querying the database
- Aggregating Data
- Trivia: How to pronounce - S-Q-L or SEQUEL?

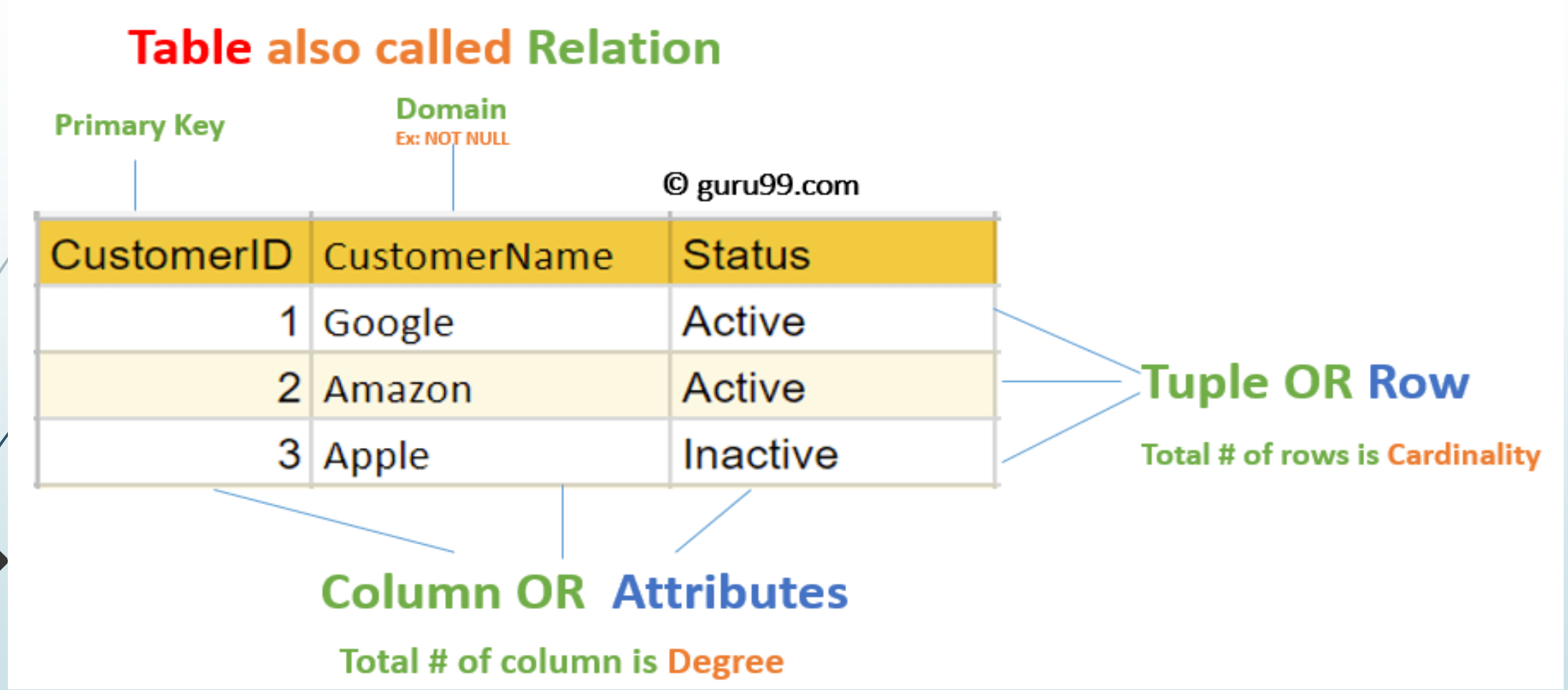
What is SQL

- SQL is acronym for Structured Query Language
- Language specifically designed to access and work with relational databases
- Relational databases are databases that store and provides access to data points that are related to one another. Relational databases are based on the relational model, an intuitive, straightforward way of representing data in tables.

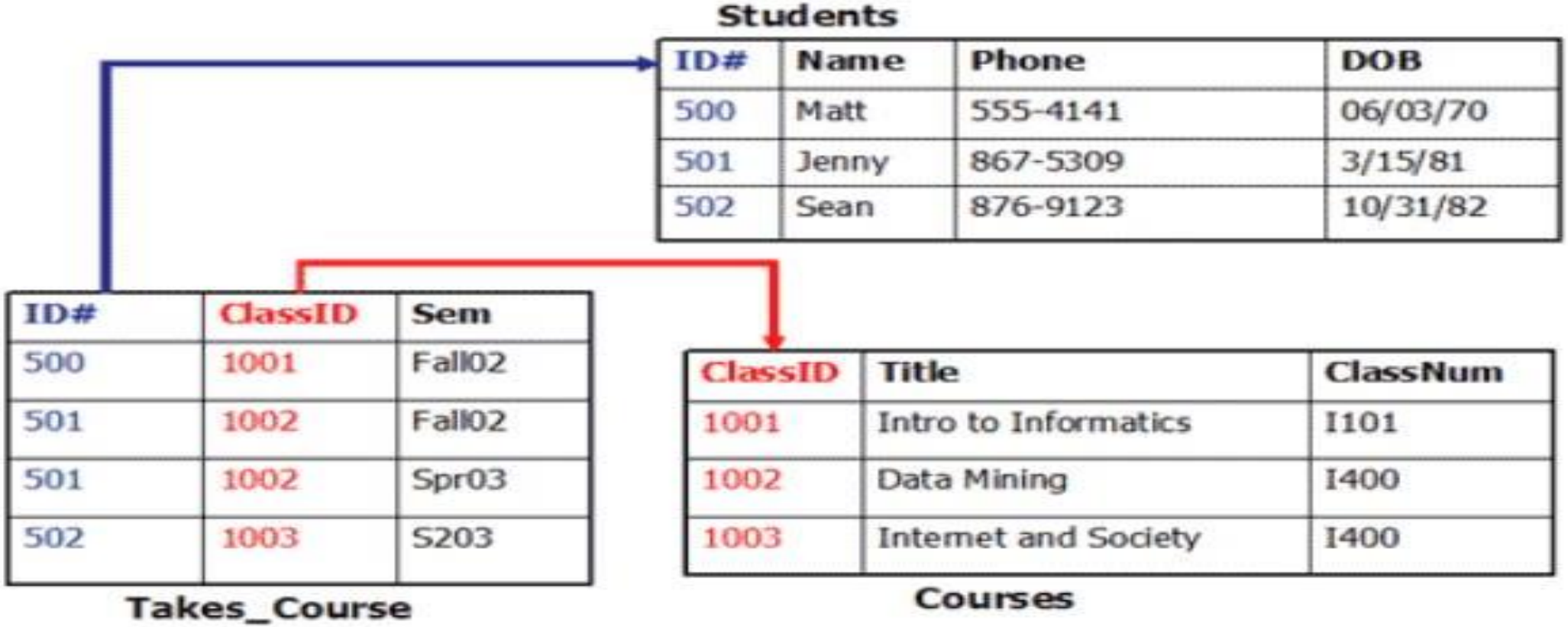
What are databases

- We are surrounded by data
- The current workshop has data about participants, their email address, what institution they are from
- How do we store and organize the data?
- In DATABASES
- Stores the data
- Provides functionalities for adding, removing and modifying data

The Relational Model



Relational DBMS



Students

ID#	Name	Phone	DOB
500	Matt	555-4141	06/03/70
501	Jenny	867-5309	3/15/81
502	Sean	876-9123	10/31/82

ID#	ClassID	Sem
500	1001	Fall02
501	1002	Fall02
501	1002	Spr03
502	1003	S203

Takes_Course

ClassID	Title	ClassNum
1001	Intro to Informatics	I101
1002	Data Mining	I400
1003	Internet and Society	I400

Courses

Relating Tables

DATABASES

- Like Excel, only organized into many more active ‘worksheets’ (Tables)
- The rows are the elements or items
- The columns are the attributes or properties of the items
  - The columns, or attributes determine the SCHEMA of the database – how many variables of each item can be stored, and what type of variables they are

What is SQL?

- Databases need a method for accessing their content, manipulating and retrieving the information
- These methods called database language
- SQL is the most common and widely uses language for accessing and interacting with databases

General SQL Vocabulary

- Data Definition Language (DDL): CREATE, DROP, ALTER, TRUNCATE
- Data Manipulation Language (DML): INSERT, UPDATE, DELETE
- Data Control Language (DCL): GRANT, REVOKE
- Data Query Language (DQL): SELECT, JOIN

SQL Data Types

SQL Numeric Data Types		
Datatype	From	To
bit	0	1
tinyint	0	255
smallint	-32,768	32,767
int	-2,147,483,648	2,147,483,647
bigint	-9,223,372,036, 854,775,808	9,223,372,036, 854,775,807
decimal	1E+38	10^38 -1
numeric	1E+38	10^38 -1
float	-1.79E + 308	1.79E + 308
real	-3.40E + 38	3.40E + 38

SQL Data Types

SQL Date and Time Data Types	
Datatype	Description
DATE	Stores date in the format YYYY-MM-DD
TIME	Stores time in the format HH:MI:SS
DATETIME	Stores date and time information in the format YYYY-MM-DD HH:MI:SS
TIMESTAMP	Stores number of seconds passed since the Unix epoch ('1970-01-01 00:00:00' UTC)
YEAR	Stores year in 2 digit or 4 digit format. Range 1901 to 2155 in 4-digit format. Range 70 to 69, representing 1970 to 2069.

SQL Character and String Data Types	
Datatype	Description
CHAR	Fixed length with maximum length of 8,000 characters
VARCHAR	Variable length storage with maximum length of 8,000 characters
VARCHAR(max)	Variable length storage with provided max characters, not supported in MySQL
TEXT	Variable length storage with maximum size of 2GB data

SQL Data Types

SQL Unicode Character and String Data Types	
Datatype	Description
NCHAR	Fixed length with maximum length of 4,000 characters
NVARCHAR	Variable length storage with maximum length of 4,000 characters
NVARCHAR(max)	Variable length storage with provided max characters
NTEXT	Variable length storage with maximum size of 1GB data

SQL Binary Data Types	
Datatype	Description
BINARY	Fixed length with maximum length of 8,000 bytes
VARBINARY	Variable length storage with maximum length of 8,000 bytes
VARBINARY(max)	Variable length storage with provided max bytes
IMAGE	Variable length storage with maximum size of 2GB binary data

SQL or SEQUEL?

- Originally developed in the 1970s at IBM
- Original acronym – SEQUEL – **S**tructured **E**nglish **QUE**ry **L**anguage
- Acronym changed to SQL because SEQUEL trademark belonged to British aircraft production company Hawker Siddeley Dynamics Engineering Limited (later part of British Aerospace, now BAE Systems)

Workshop Activities

- Create a data table called Covid_Data
- Insert data from **Our World in Data (owid)** on COVID by country and by continent into the Covid_Data table
- Query the Covid_Data table using the SELECT statement
- Filter your query using the WHERE clause
- Create another data table called Stability_Data
- Insert World Stability index data from **Fund for Peace**
- Relate the two tables using the JOIN connector

Additional Resources

- Dbeaver Wiki - https://github.com/dbeaver/dbeaver/wiki
- W3schools - https://www.w3schools.com/sql/default.asp
- CodeAcademy - https://www.w3schools.com/sql/default.asp
- Learn SQL - https://learnsql.com/
- Khan Academy - https://www.khanacademy.org/ (fee-based learning platform)
- Tutorialspoint - https://www.tutorialspoint.com/sql/?utm_source=7_&utm_medium=affiliate&utm_content=5f34cd37cdf1050001b09537&utm_campaign=Admitad&utm_term=0bf9452ff944206eb94731a49408c3b8
- SQLZoo - https://sqlzoo.net/
- Udacity - https://www.udacity.com/course/intro-to-relational-databases--ud197
- Essential SQL - https://www.essentialsql.com/
- Learn SQL the hard way - https://learncodethehardway.org/sql/
- Udemy SQL Tutorials - https://www.udemy.com/courses/search/?q=sql (fee-based)
- SoloLearn SQL Fundamentals - https://www.sololearn.com/Course/SQL/