Prerequisites for this chapter

**Install:**[**MySQL Community Server**](https://dev.mysql.com/downloads/)**,**[**MySQL Workbench**](https://dev.mysql.com/downloads/)

SQL in 6 minutes

**Browse to**[**https://www.w3schools.com/sql/trysql.asp?filename=trysql\_op\_in**](https://www.w3schools.com/sql/trysql.asp?filename=trysql_op_in)

**Query 1**

**SELECT** \* **FROM** customers;

**Query 2**

**SELECT** supplierName, **COUNT**(\*) **AS** 'number of products' **FROM** suppliers

**INNER** **JOIN** products

**ON** products.SupplierID = suppliers.SupplierID

**GROUP** **BY** suppliers.SupplierID;

**Query 3**

**SELECT** o.OrderDate,

o.OrderID,

o.ShipperID,

ROUND(**SUM**(od.Quantity \* p.Price), 0) **AS** Basket,

**CASE** **WHEN** od.Quantity < 30 **THEN** 'SMALLQ' **ELSE** 'HIGHQ' **END** **as** QuantityLabel

**FROM** Orders **AS** o

**LEFT** **JOIN** OrderDetails **AS** od **ON** od.OrderID = o.OrderID

**INNER** **JOIN** ( **SELECT** \* **FROM** Products **WHERE** Price >= 100 ) **AS** p **on** p.ProductID = od.ProductID

**GROUP** **BY** o.OrderID,

o.ShipperID

**HAVING** Basket >= 1000

**ORDER** **BY** o.OrderDate **DESC**, o.OrderID;

First look on MySQL and MySQL Workbench

[Screenshot help](https://ceu-economics-and-business.github.io/ECBS-5146-Different-Shapes-of-Data/artifacts/intro/connect.png)

Your first local MySQL Database

**Create your first database / schema**

**CREATE** **SCHEMA** firstdb;

**SQL is not case sensitive:**

**create** **schema** FIRSTDB;

**For the next commands, make sure the created db is selected**

**USE** firstdb;

**Deleting a database**

**Execute twice**

**DROP** **SCHEMA** firstdb;

**Second time you will get and error because the db is already deleted with the first one.**

**Try this instead**

**DROP** **SCHEMA** IF **EXISTS** firstdb;

**Let’s create a db again**

**CREATE** **SCHEMA** birdstrikes;

**USE** birdstrikes;

Loading CSV into a table

**Note:** If you are not familiar with CSV file format, read the CSV section [here](https://ceu-economics-and-business.github.io/ECBS-5146-Different-Shapes-of-Data/08-dsd/index.html#csv)

**Let’s create a table:**

**CREATE** **TABLE** birdstrikes

(id INTEGER **NOT** **NULL**,

aircraft VARCHAR(32),

flight\_date DATE **NOT** **NULL**,

damage VARCHAR(16) **NOT** **NULL**,

airline VARCHAR(255) **NOT** **NULL**,

**state** VARCHAR(255),

phase\_of\_flight VARCHAR(32),

reported\_date DATE,

bird\_size VARCHAR(16),

cost INTEGER **NOT** **NULL**,

speed INTEGER,**PRIMARY** **KEY**(id));

This table is empty, we need to fill in with data.

**This time we will load a csv file into the table. For security reason, CSV loading is limited, so you need to copy the CSV file in a place indicated by this command:**

**SHOW** VARIABLES **LIKE** "secure\_file\_priv";

**also the next command should give you “ON”**

**SHOW** VARIABLES **LIKE** "local\_infile";

Plan A

**If “local\_infile” is “ON” and “secure\_file\_priv” is not “NULL”**

**Copy**[**birdstrikes\_small.csv**](https://ceu-economics-and-business.github.io/ECBS-5146-Different-Shapes-of-Data/artifacts/intro/birdstrikes_small.csv)**in the folder resulted in the previous command.**

**Then load CSV data into the table with this command:**

**LOAD** **DATA** INFILE 'c:/ProgramData/MySQL/MySQL Server 8.0/Uploads/birdstrikes\_small.csv'

**INTO** **TABLE** birdstrikes

FIELDS TERMINATED **BY** ';'

LINES TERMINATED **BY** '**\r\n**'

**IGNORE** 1 LINES

(id, aircraft, flight\_date, damage, airline, **state**, phase\_of\_flight, @v\_reported\_date, bird\_size, cost, @v\_speed)

**SET**

reported\_date = **nullif**(@v\_reported\_date, ''),

speed = **nullif**(@v\_speed, '');

Plan B

**If “local\_infile” is not “ON” or “secure\_file\_priv” is NULL, you need to change my.cnf (Mac,Linux) or my.ini (Windows). This is an advanced operation, so as plan B: download**[**birdstrikes.sql**](https://ceu-economics-and-business.github.io/ECBS-5146-Different-Shapes-of-Data/artifacts/intro/birdstrikes.sql)**and Open SQL Script in MySQL Workbench, then execute.**

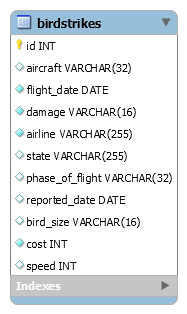
Exploring your first database

**List the table(s) of your database**

**SHOW** TABLES;

**List the structure of a table**

**DESCRIBE** birdstrikes;



**Retrieving data stored in birdstrikes**

**SELECT** \* **FROM** birdstrikes;

**Select certain field(s)**

**SELECT** cost **FROM** birdstrikes;

**SELECT** airline,cost **FROM** birdstrikes;

Dumping a database with MySQL Workbench

[Screenshot help](https://ceu-economics-and-business.github.io/ECBS-5146-Different-Shapes-of-Data/artifacts/intro/dump.png)

Ninja challenge

* Load lines starting with “Data:” in [ninja.txt](https://ceu-economics-and-business.github.io/ECBS-5146-Different-Shapes-of-Data/artifacts/intro/ninja.txt) into a table.
* Further requirements:
  + Divide the last column with 1000 during the load
  + ” signs should be removed during the load
  + Dump the table and send to me in mail with the sql script.

Homework 1

* Import a relational data set of your choosing into your local instance.
* Requirements:
  + find a data set worth to analyze later (prepares you for the term project)
  + no restriction on the type of data source, can be excel, csv, another db, sql file etc
  + pay attention on the relational nature of the set, advised to find a structure of 3+ interlinked table
  + do not use this: https://www.mysqltutorial.org/mysql-sample-database.aspx (because we will use it later in the course)
  + hint: you can find various open datasets on the internet, like here: https://data.worldbank.org/
* Create a public GitHub repo. This repo will be used for all homeworks and term project in this course.
* Save your artifacts (possible sources like csv, sql file ) in a folder called HW1.
* Submit GitHub repo link to moodle when you are ready