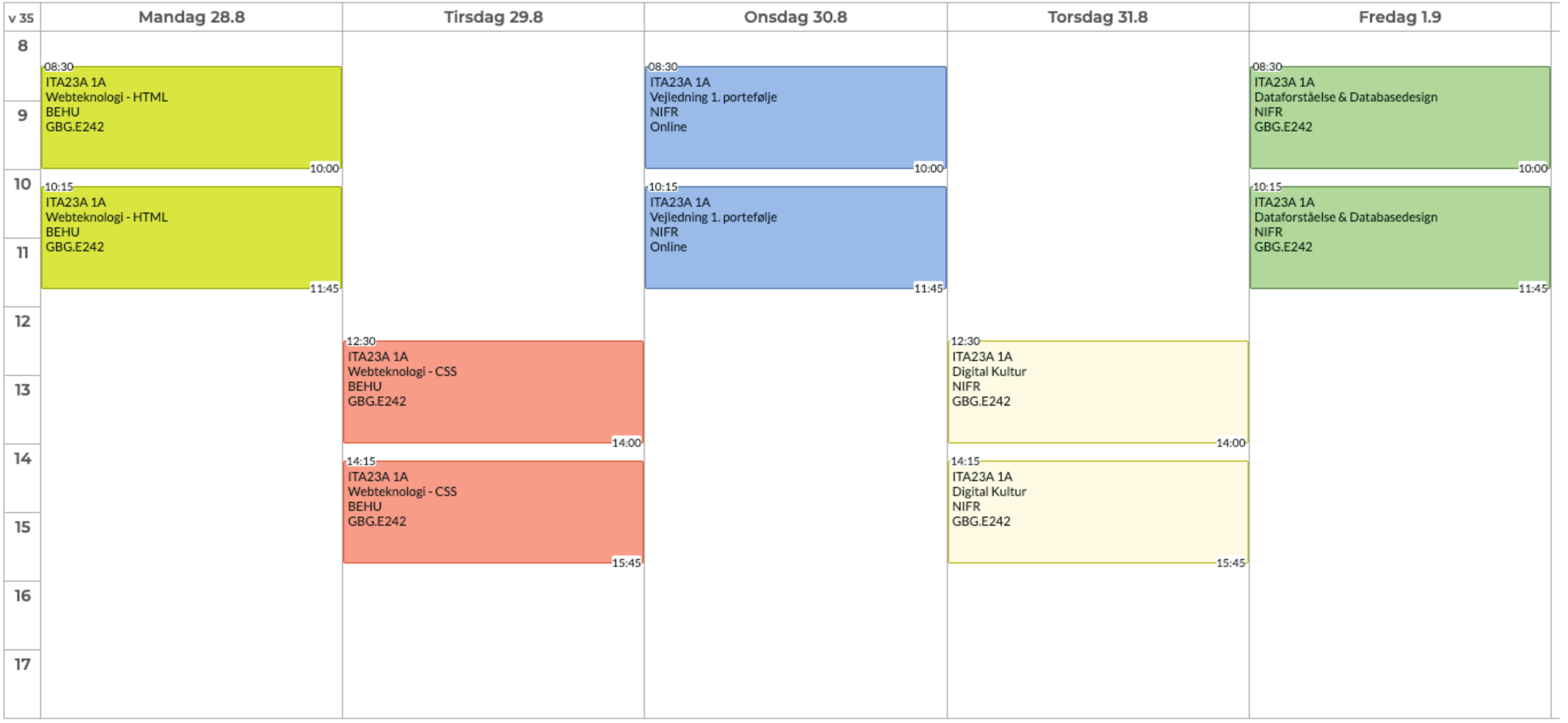
What is data + Datatyper & skalaer
Retrieving data with MySQL

Dataforståelse

NIFR 2023


v 35
Mandag 28.8
Tirsdag 29.8
Onsdag 30.8
Torsdag 31.8
Fredag 1.9
8
9
10
11
12
13
14
15
16
17
08:30
ITA23A 1A
Webteknologi - HTML
BEHU
GBG.E242
10:00
10:15
ITA23A 1A
Webteknologi - HTML
BEHU
GBG.E242
11:45
12:30
ITA23A 1A
Webteknologi - CSS
BEHU
GBG.E242
14:00
14:15
ITA23A 1A
Webteknologi - CSS
BEHU
GBG.E242
15:45
08:30
ITA23A 1A
Vejledning 1. portefølje
NIFR
Online
10:00
10:15
ITA23A 1A
Vejledning 1. portefølje
NIFR
Online
11:45
12:30
ITA23A 1A
Digital Kultur
NIFR
GBG.E242
14:00
14:15
ITA23A 1A
Digital Kultur
NIFR
GBG.E242
15:45
08:30
ITA23A 1A
Dataforståelse & Databasedesign
NIFR
GBG.E242
10:00
10:15
ITA23A 1A
Dataforståelse & Databasedesign
NIFR
GBG.E242
11:45

Agenda

Dataforståelse & Databasedesign

- Recap: 01 - What is data
- Datarepræsentation i en computer
- Datatyper
- Dataskalaer
- Datatyper repræsentation i MySQL
- Retrieving data with MySQL

- SQL Workshop
- Introduktion til 1. porteføljeopgave

FORMÅL OG LÆRINGSMÅL

Dataforståelse indeholder analyse af data og databasemodellering. Der bliver arbejdet med at lave de rigtige udtræk af data og forståelsen for hvordan data bedst struktureres ved hjælp af datamodeller. Der bliver arbejdet med relationelle databaser og data i forskellige dataformater. Fagelementet indeholder også hvordan man implementerer en relationel database baseret på en datamodel.

Formål og overordnet læringsmål

Dataforståelse

- Data & Databasemodeller
  - Udtræk af data
  - Relationelle Databaser
  - Dataformater
  - Implementere en relationel database

Viden

Den studerende har:

• viden om hvordan forskellige dataformater bruges, og deres styrker og svagheder

• viden om teori og metoder inden for dataforståelse, metadata, dataanalyse og datakvalitet

• viden om teori og metoder inden for datamodellering og -typer

• forståelse for praksis, anvendt teori og metode samt kan reflektere over data, datakvalitet og databasemodellering

Færdigheder

Den studerende kan:

• anvende metoder og redskaber inden for datamodellering og dataanalyse

• vurdere praksisnære og teoretiske problemstillinger inden for data og dataanalyse herunder hvilke dataudtræk der giver mening for applikationen, samt programmere disse dataudtræk

• analysere og vurdere databehov i forhold til applikationen der skal udvikles

Kompetencer

Den studerende kan:

• håndtere udarbejdelse af en datamodel baseret på de behov og krav den givne opgave kræver samt implementere en database baseret på en datamodel

• identificere egne læringsbehov og udvikle egen viden, færdigheder og kompetencer i relation til fagområdet dataforståelse

https://katalog.kea.dk/course/4111101/2023-2024

Formål og overordnet læringsmål

Dataforståelse

- Data & Databasemodeller
  - Udtræk af data
  - Relationelle Databaser
  - Datatyper & skalaer
  - SQL & den relationelle databasemodel

Dictionary

Definitions from Oxford Languages · Learn more

data

/ˈdeɪtə/

noun

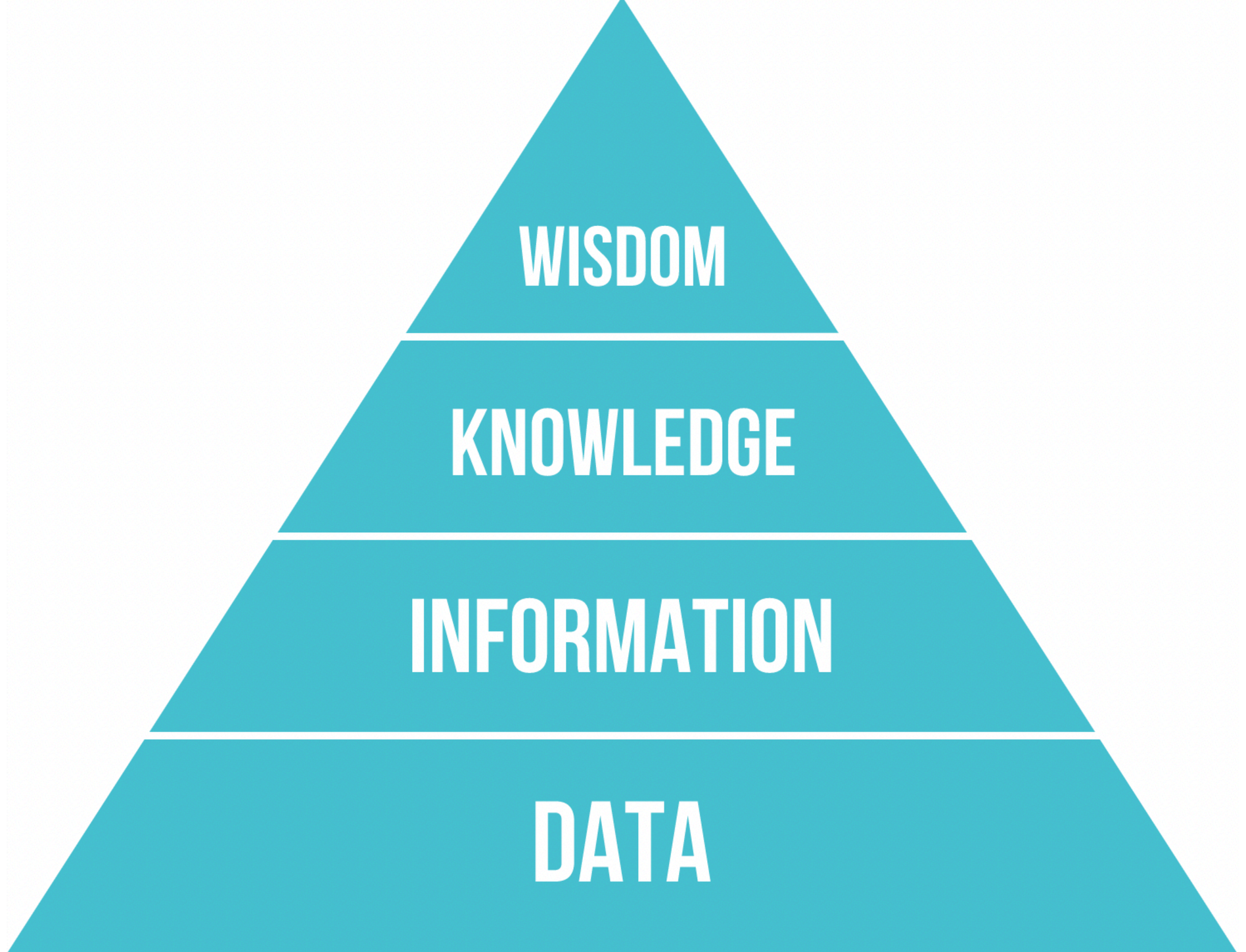
facts and statistics collected together for reference or analysis.
"there is very little data available"

Similar: facts | figures | statistics | details | particulars | specifics | features

- the quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media.
- **PHILOSOPHY**
  things known or assumed as facts, making the basis of reasoning or calculation.

Data is factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation.

Data har form & formål


WISDOM
KNOWLEDGE
INFORMATION
DATA

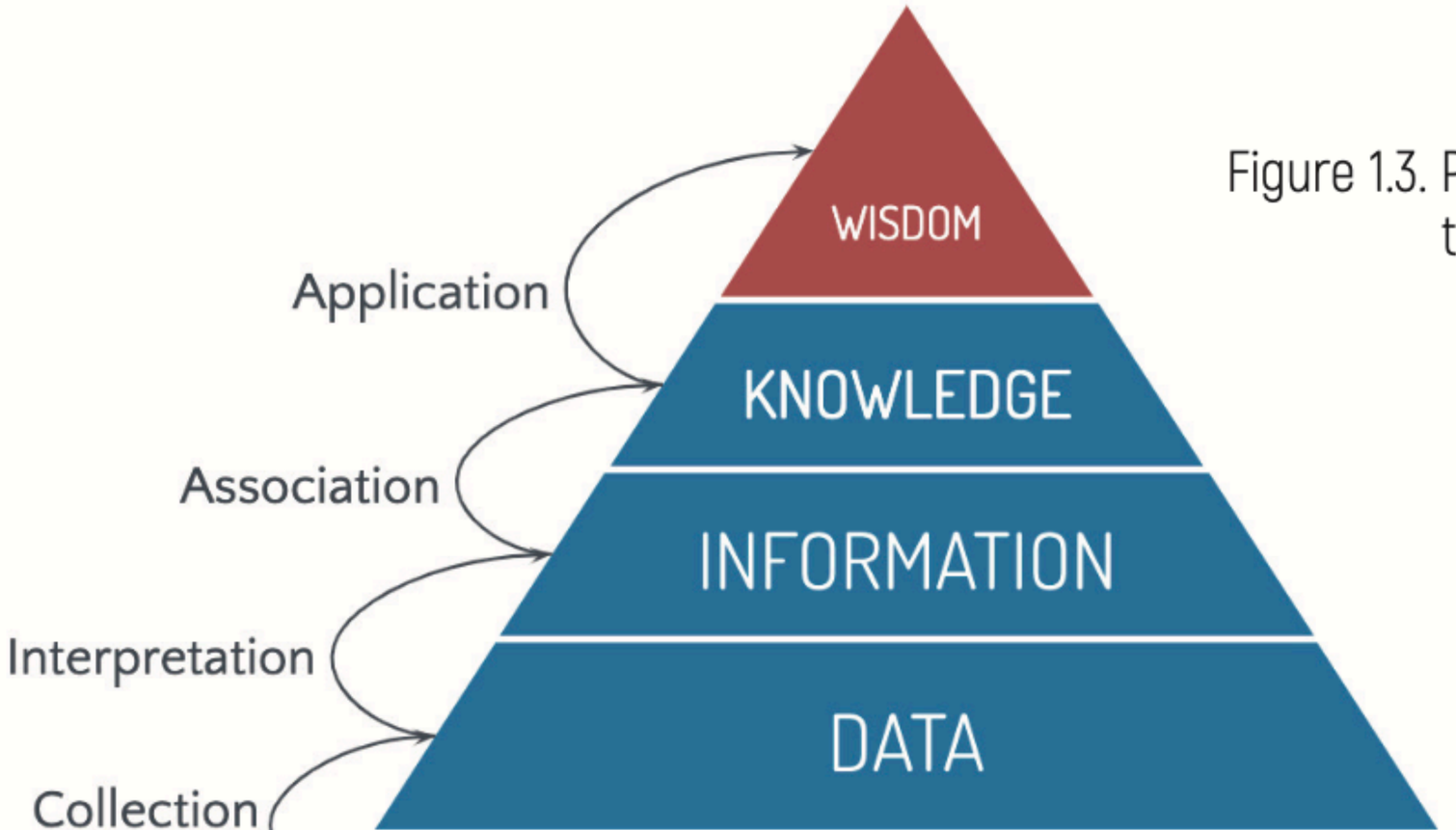
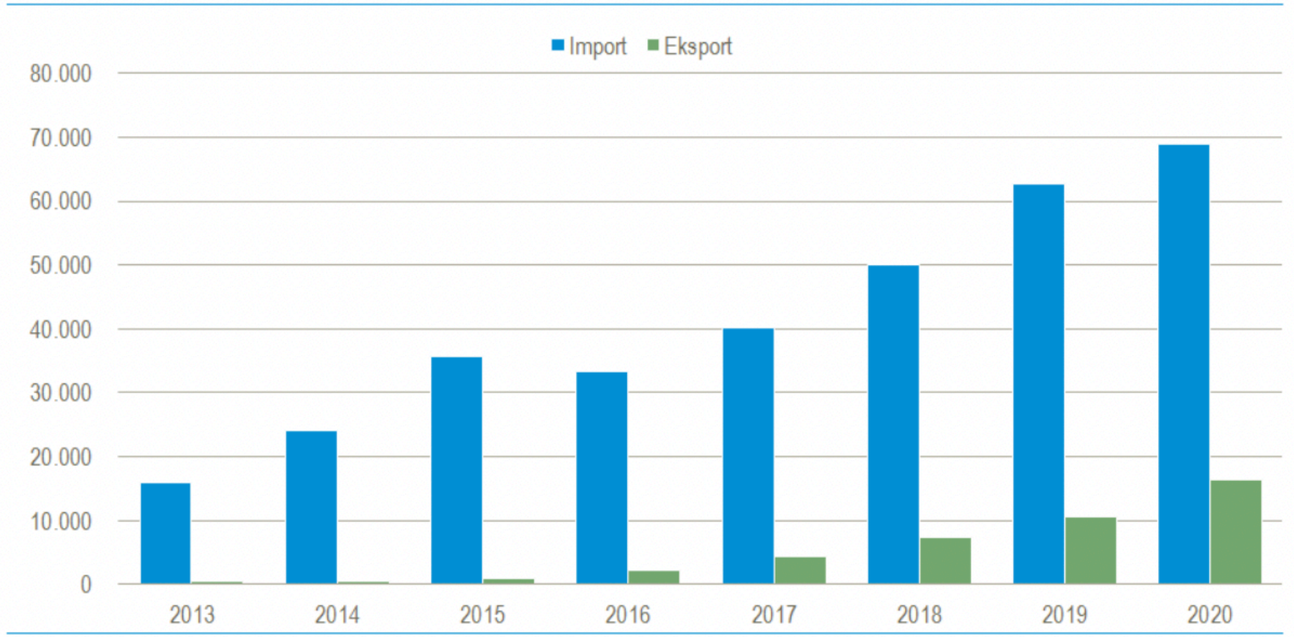


Figure 1.3. Progressing up the DIKW Pyramid.

Kildemoes	01/02/2022	Mountainbike	Rød	Ikke
Kildemoes	01/02/2002	City-bike	Blå	El-drevet
Kildemoes	01/02/2005	Mountainbike	Rød	Ikke
Kildemoes	01/02/2020	City-bike	Blå	El-drevet
Kildemoes	01/02/2022	Mountainbike	Rød	Ikke

Råmateriale

DATA



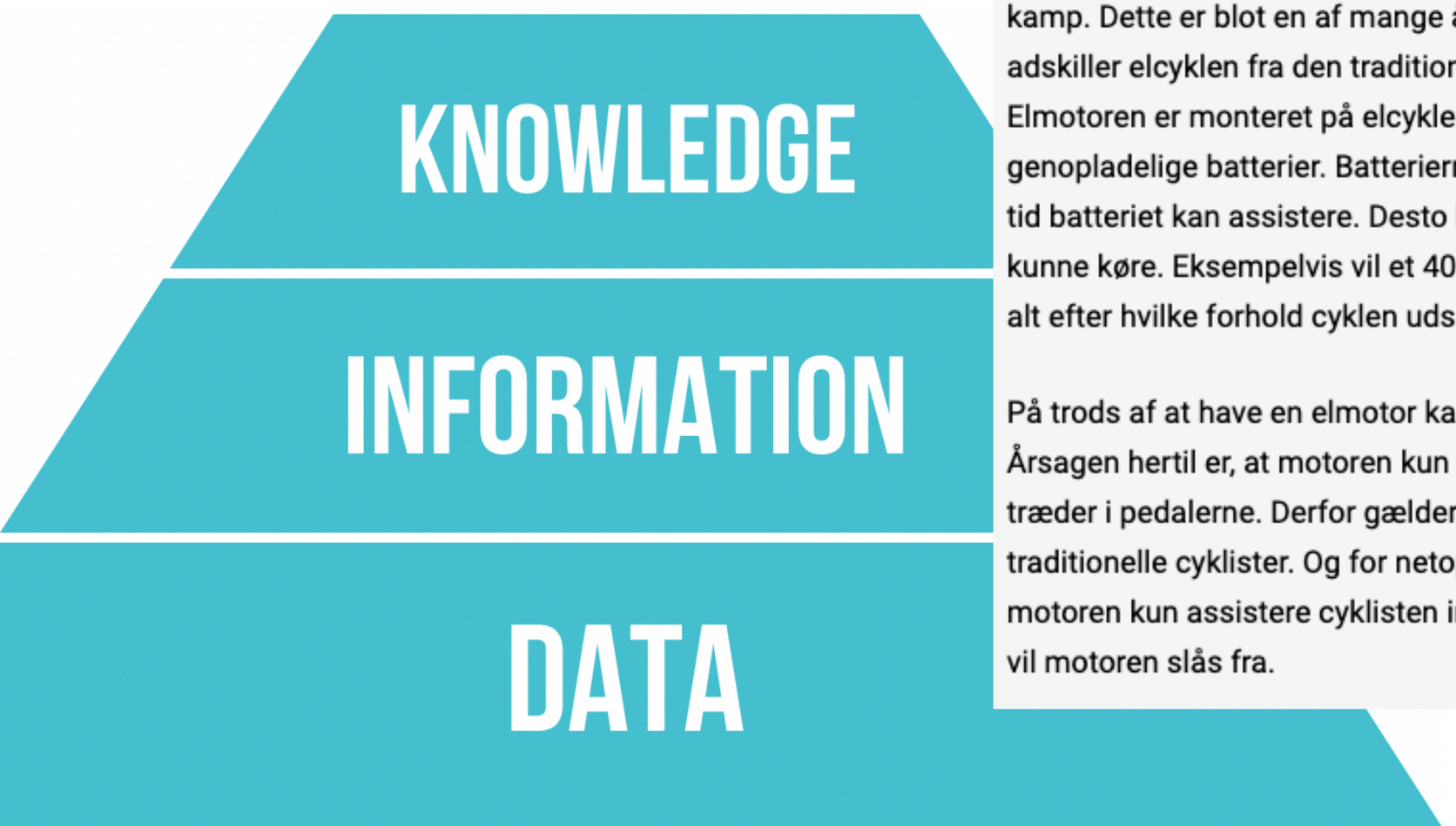
Anvendeligt

INFORMATION

DATA

- What are the **units of measurement** of the data value or values? Are the values representative of units ("eaches"), boxes of 10, or grams, dollars, or something else? If dollars, are we talking about US dollars or Canadian dollars? And so on.
- Do we need to **aggregate multiple values** into a more meaningful quantity, such as total sales over a period of time, or average height of the students?

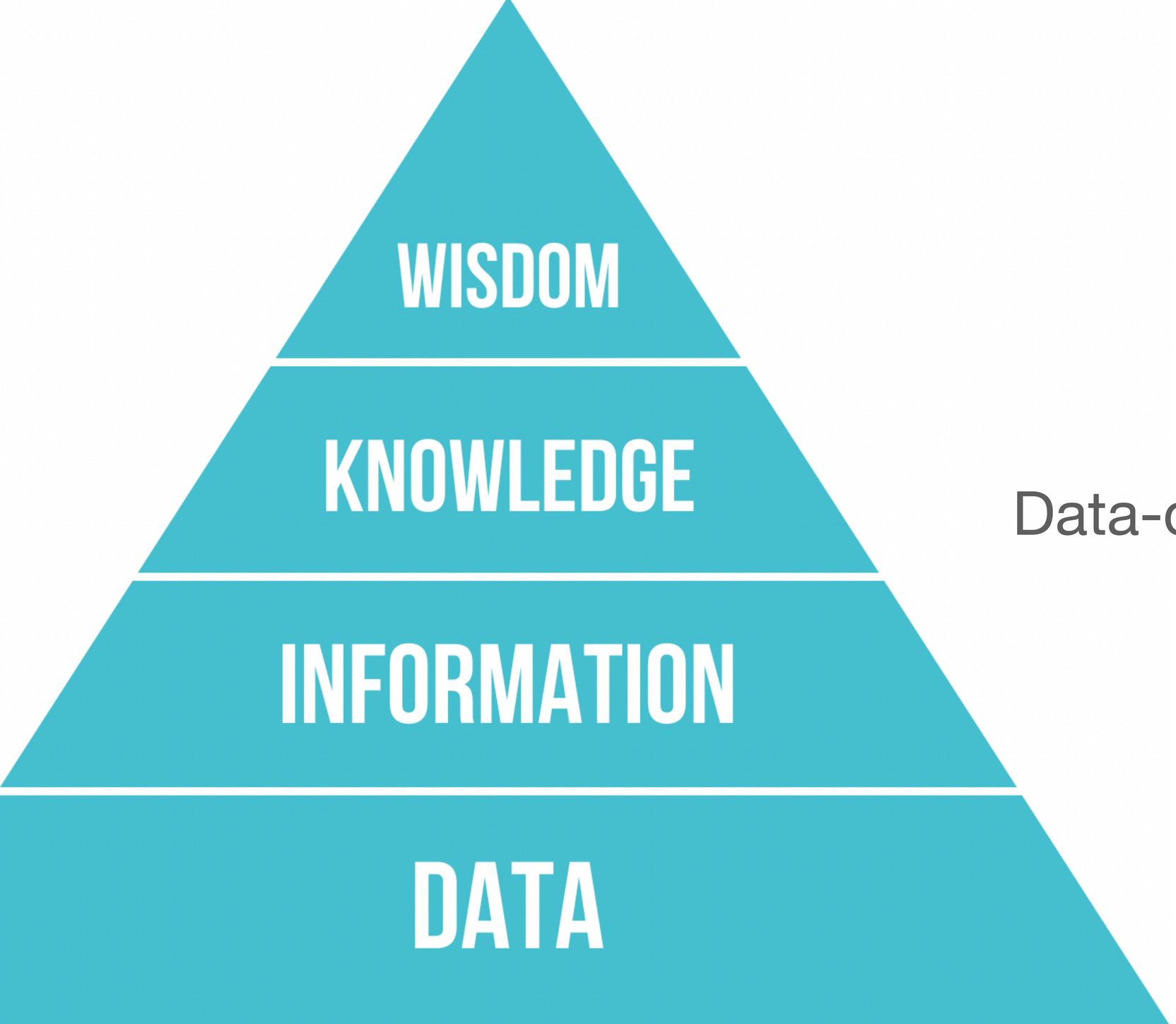
Kontekst



Elcyklen kan blive din bedste ven op ad bakkerne

Danmark har et ujævnt terræn med mange bakker, hvilket kan forårsage en del sved på panden for cyklisten. Men på grund af de mange elcykler på cykelstierne, er det ikke længere unormalt at blive overhalet af cyklister på en elcykel - elcyklister som uden kvaler suser op ad bakken mens de traditionelle cyklister kæmper en hed kamp. Dette er blot en af mange årsager til elcyklens voksende popularitet. Det der adskiller elcyklen fra den traditionelle cykel er naturligvis, at den kører på el. Elmotoren er monteret på elcyklens for- eller baghjul. Derudover har de genopladelige batterier. Batterierne har forskellige wattmeter, som afgør hvor lang tid batteriet kan assistere. Desto højere watt du vælger, desto flere kilometer vil du kunne køre. Eksempelvis vil et 400 watts batteri række til mellem 80 – 120 kilometer, alt efter hvilke forhold cyklen udsættes for.

På trods af at have en elmotor kan elcykler stadig godt kategoriseres som en cykel. Årsagen hertil er, at motoren kun hjælper ved at trække cyklen, når cyklisten selv træder i pedalerne. Derfor gælder der også de samme regler for elcyklister som for traditionelle cyklister. Og for netop at tage hensyn til de traditionelle cyklister må motoren kun assistere cyklisten indtil, at der nås en hastighed på 25 km/t, herefter vil motoren slås fra.



Data-drevne beslutninger

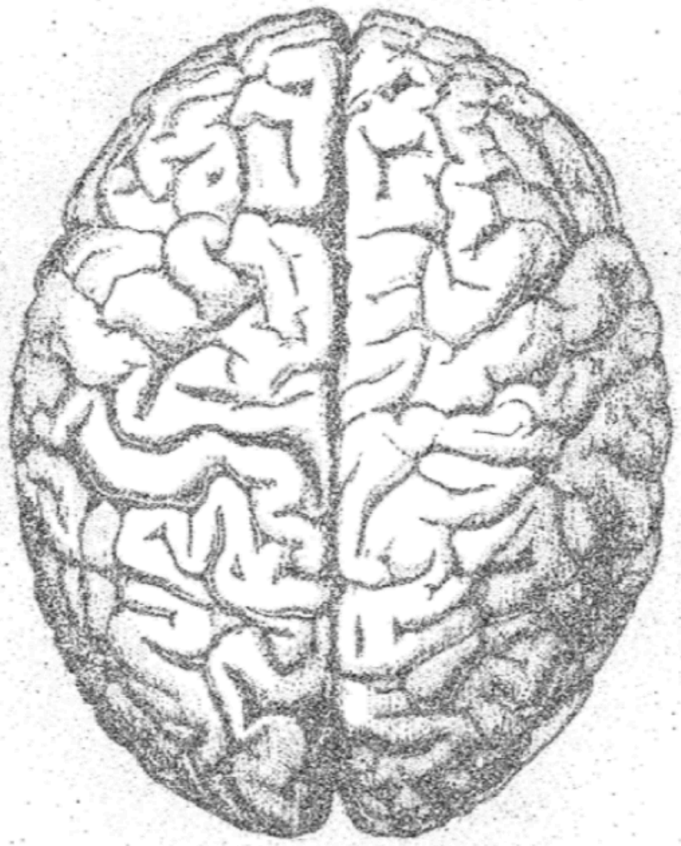
Why is data useful?

We used to use intuition; now we use analytics.

System 1

Hurtig
Impulsiv
Emotionel
Ubevidst
Hverdag

90%



System 2

Langsom
Analytisk
Rationel
Bevidst
Komplekst

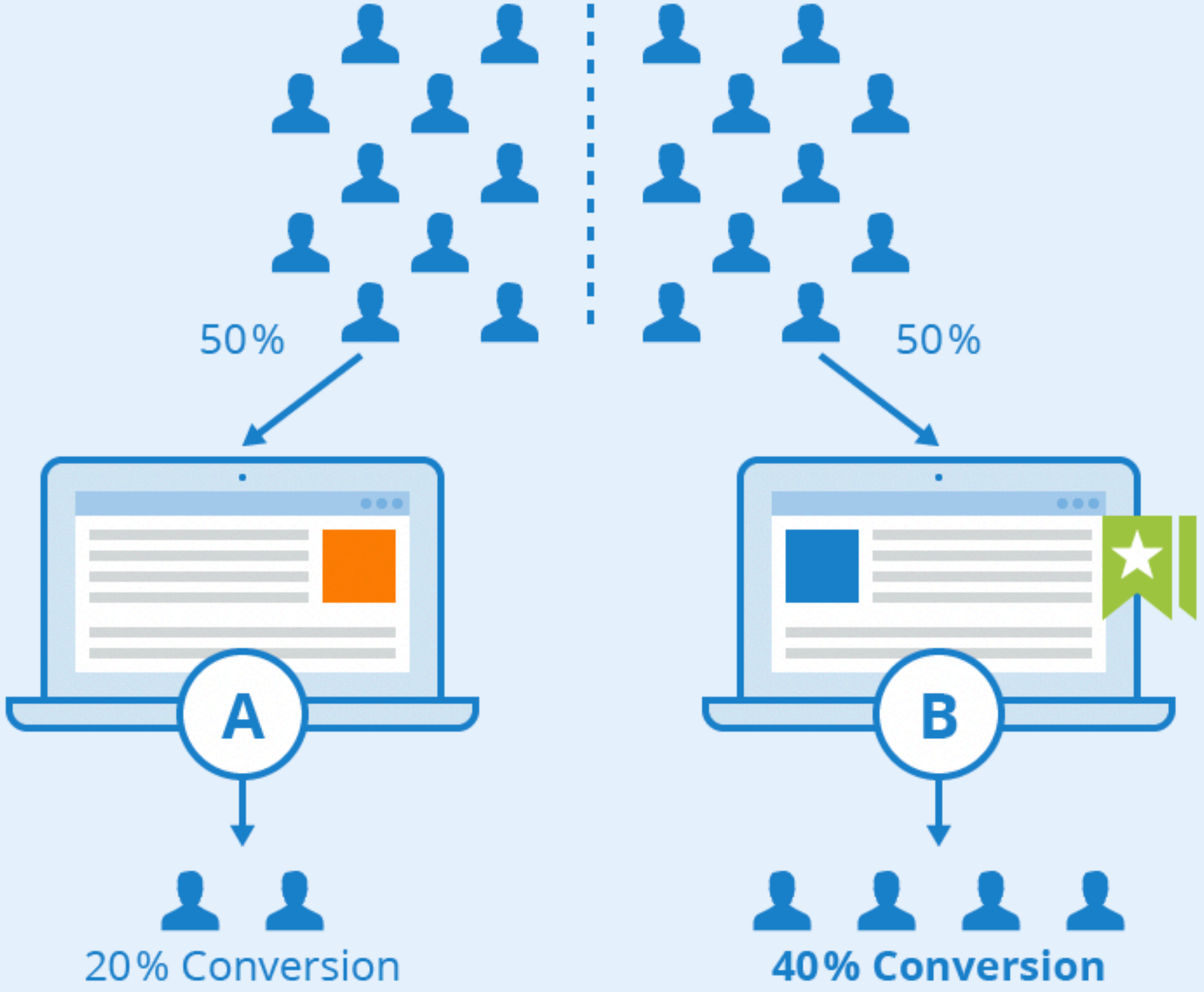
10%

The ball costs 10 cents.
The bat costs 1 dollar more than the ball.

How much does the bat cost?

Confirmation bias

The tendency to seek out and prefer information that supports our preexisting beliefs

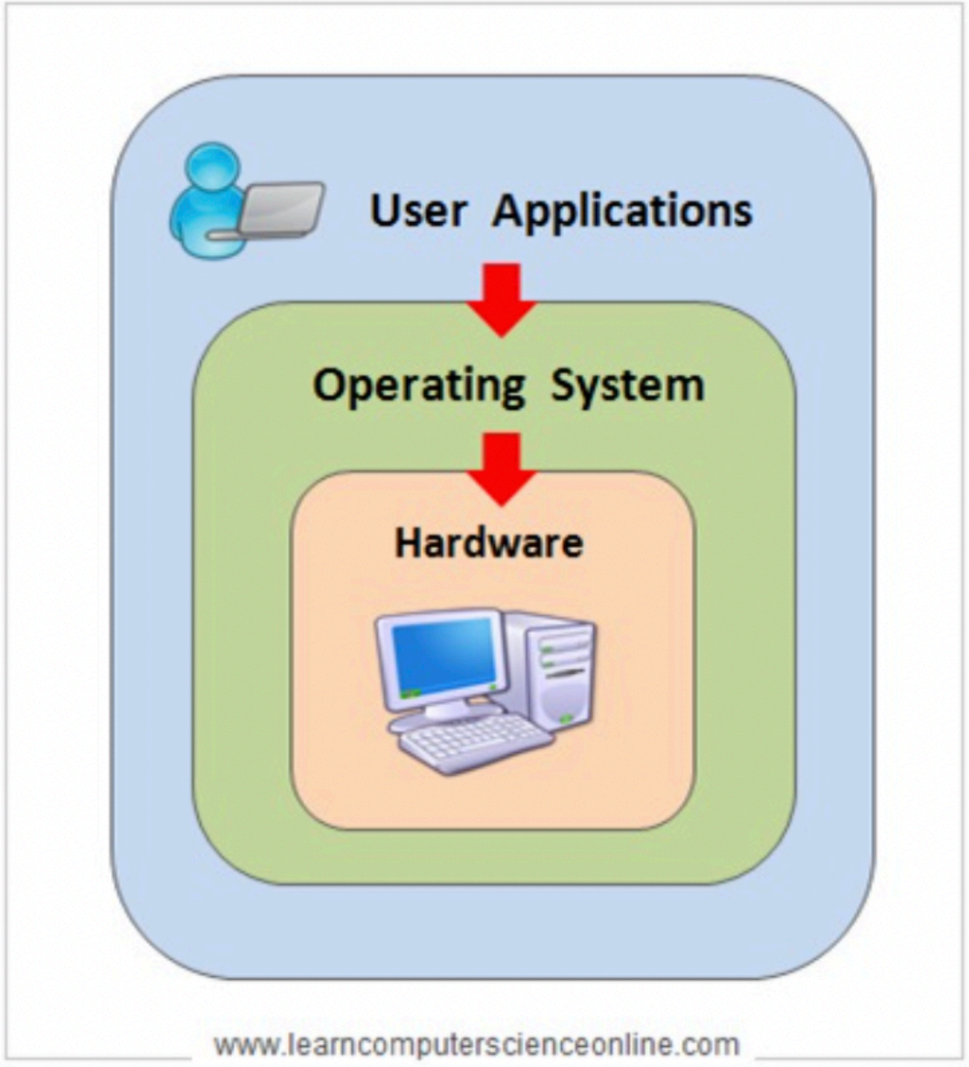

50%
50%
A
B
20% Conversion
40% Conversion

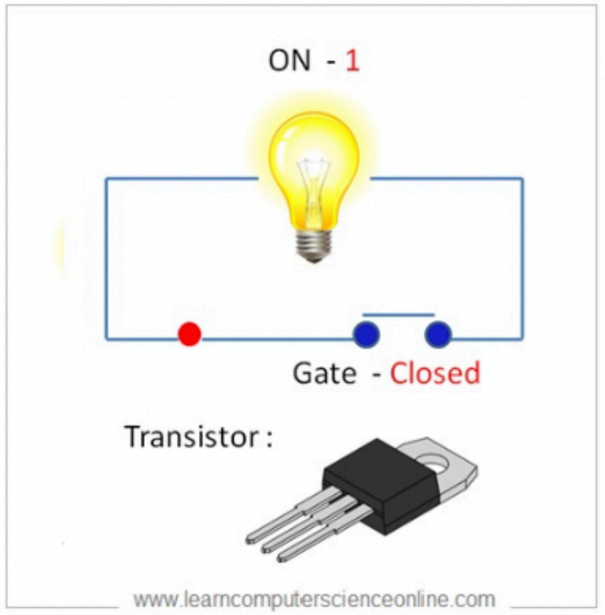
Data is factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation.

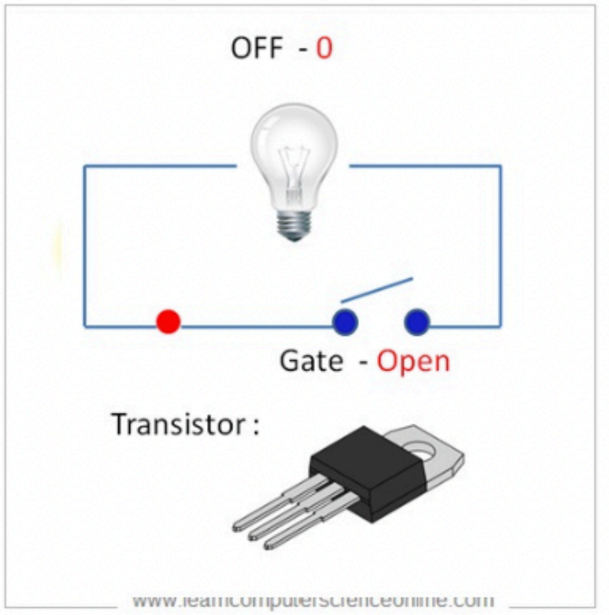
Data har form & formål

Typer & Formater

Computer Architecture

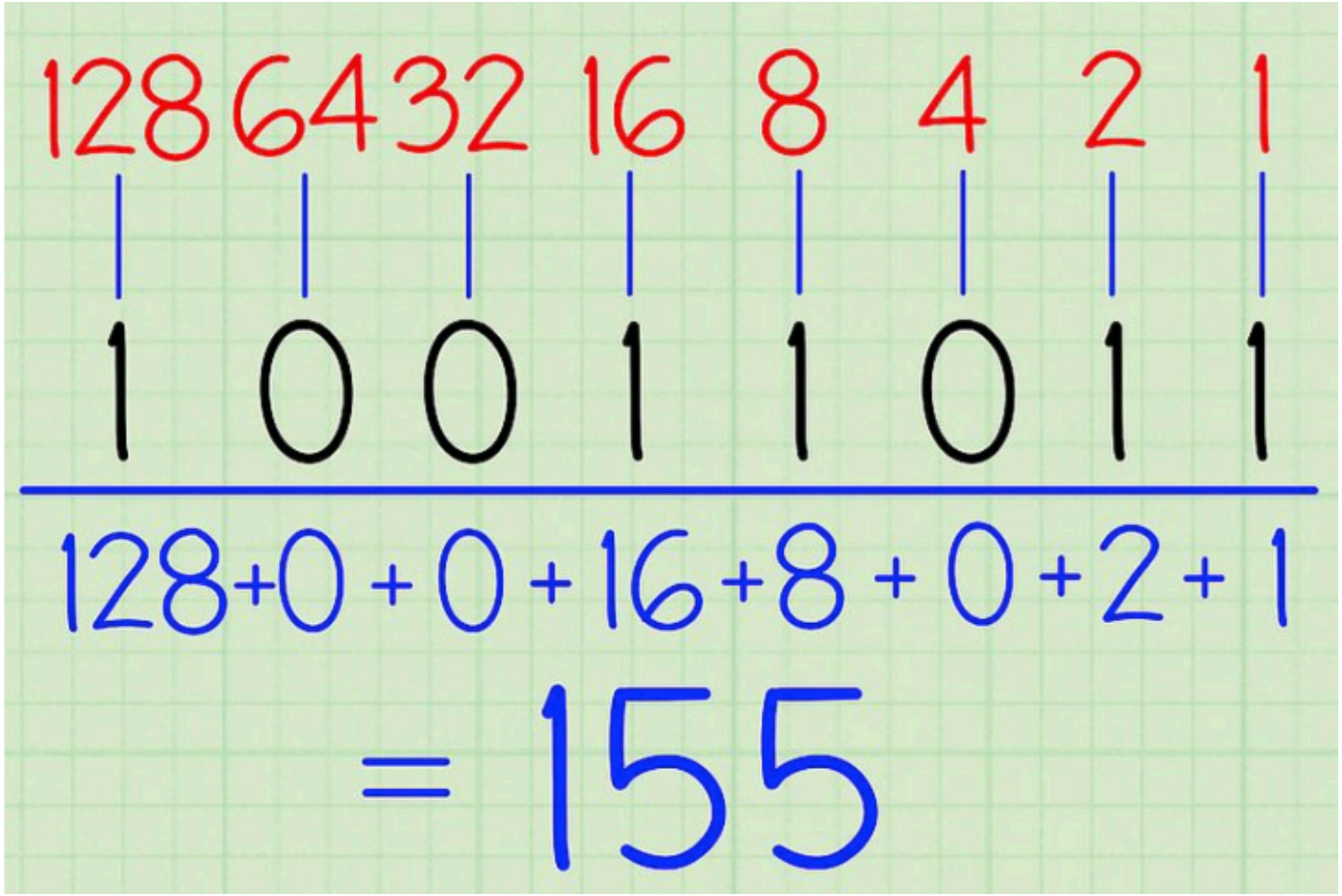






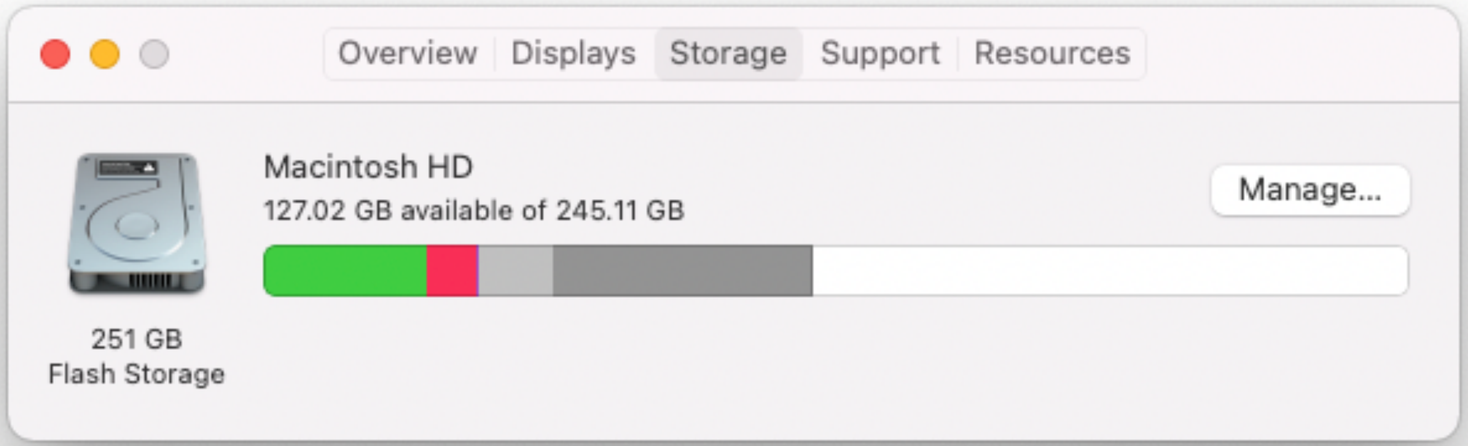
Number Systems

System	Base	Digits
Binary	2	0 1
Octal	8	0 1 2 3 4 5 6 7
Decimal	10	0 1 2 3 4 5 6 7 8 9
Hexadecimal	16	0 1 2 3 4 5 6 7 8 9 A B C D E F


128 64 32 16 8 4 2 1
1 0 0 1 1 0 1 1
128+0+0+16+8+0+2+1
= 155

Code	Character	Code	Character	Code	Character	Code	Character	Code	Character
00	0	09	9	18	I	27	R	36	Space
01	1	10	A	19	J	28	S	37	$
02	2	11	B	20	K	29	T	38	%
03	3	12	C	21	L	30	U	39	*
04	4	13	D	22	M	31	V	40	+
05	5	14	E	23	N	32	W	41	–
06	6	15	F	24	O	33	X	42	.
07	7	16	G	25	P	34	Y	43	/
08	8	17	H	26	Q	35	Z	44	:

Data takes up space



- Elektroniske systemer
- Gemme data med høj præcision
- Infrastrukturen (internettet, elnettet)
- Videnssamfund
- Databaseteknologier

Data types

Numbers, words, boolean, null

Objects/complex data types

Numbers, words, boolean, null

123,5

Numbers, words, boolean, null

“Nicklas”

Numbers, words, boolean, null

“123,5”

Numbers, words, boolean, null

2+2 = ?

Numbers, words, boolean, null

“2”+”2” =“22”

2+2 = 4

“2”+”2” = “22”

Numbers, words, boolean, null

$22 > 10 = ?$

Numbers, words, boolean, null

(22 > 10) = true

null

Object/complex type

“Nicklas”, 33, true

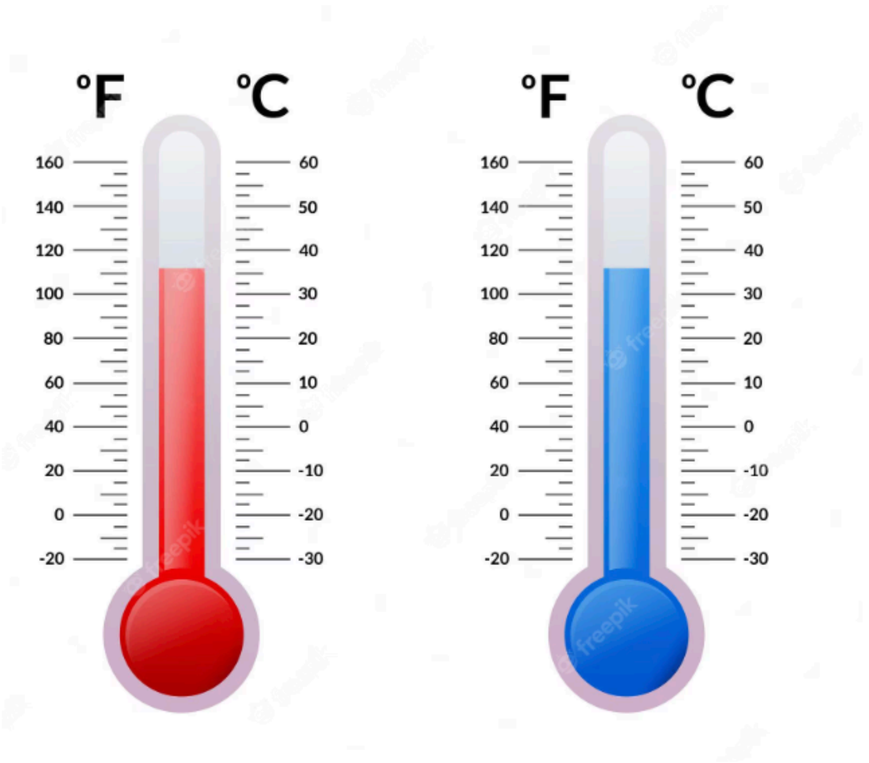
Object/complex type entity/tupple

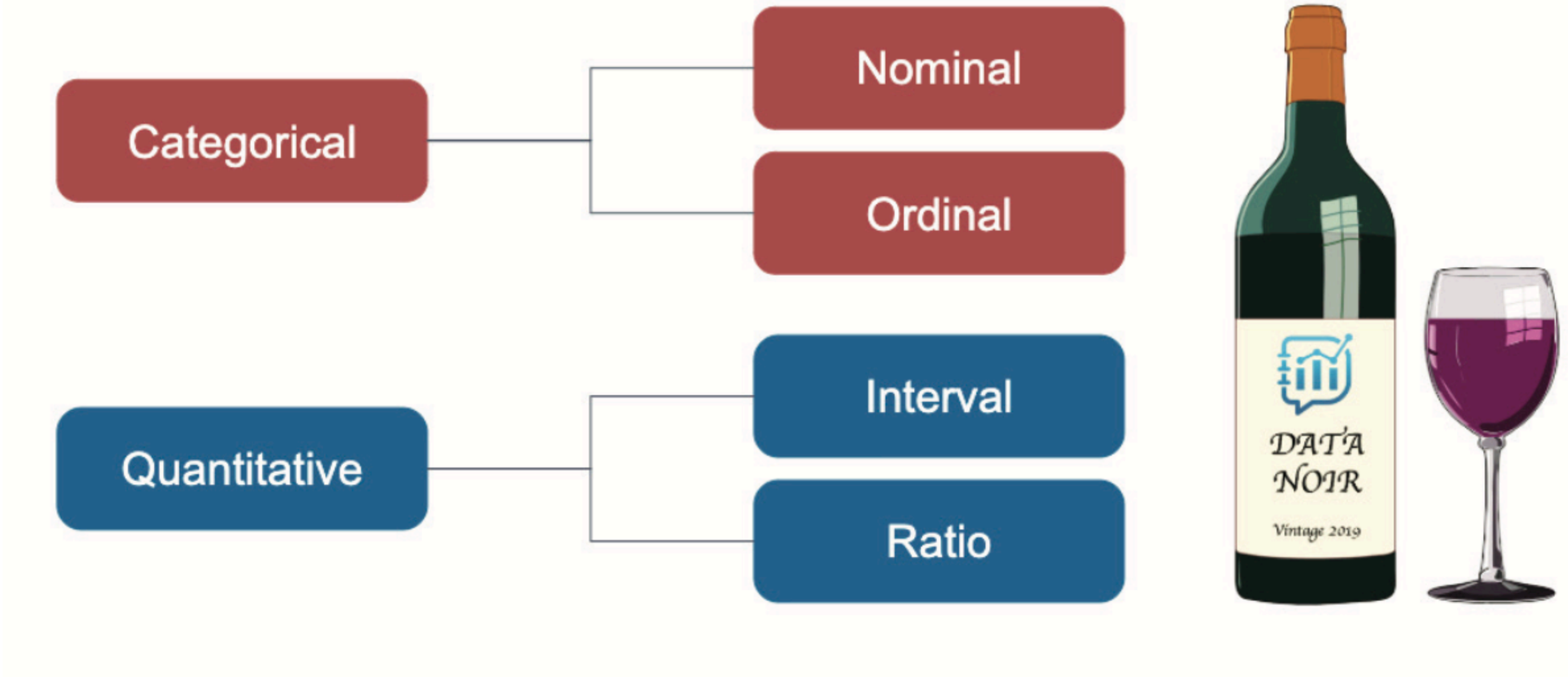
“Nicklas”, 33, true

“Benjamin”, 37, true

“Trine”, 27, true

Data scales


°F
°C
160
140
120
100
80
60
40
20
0
-20
60
50
40
30
20
10
0
-10
-20
-30
°F
°C
160
140
120
100
80
60
40
20
0
-20
60
50
40
30
20
10
0
-10
-20
-30


Categorical
Nominal
Ordinal
Quantitative
Interval
Ratio
DATA
NOIR
Vintage 2019

Nominal

Data-scale

Figure 4.2. Examples of nominal data scales: jersey numbers, fruit labels, license plates.

Ordinal

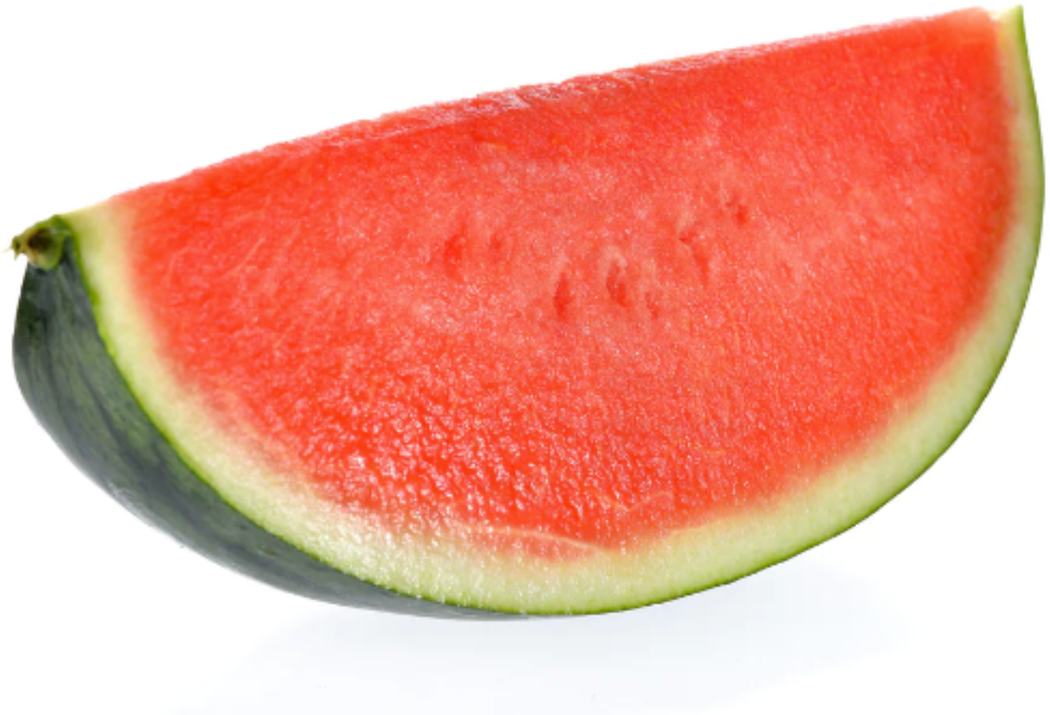
Data-scale

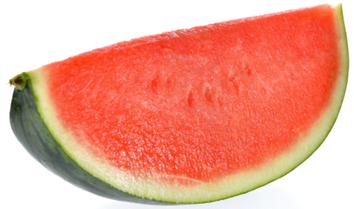
Interval

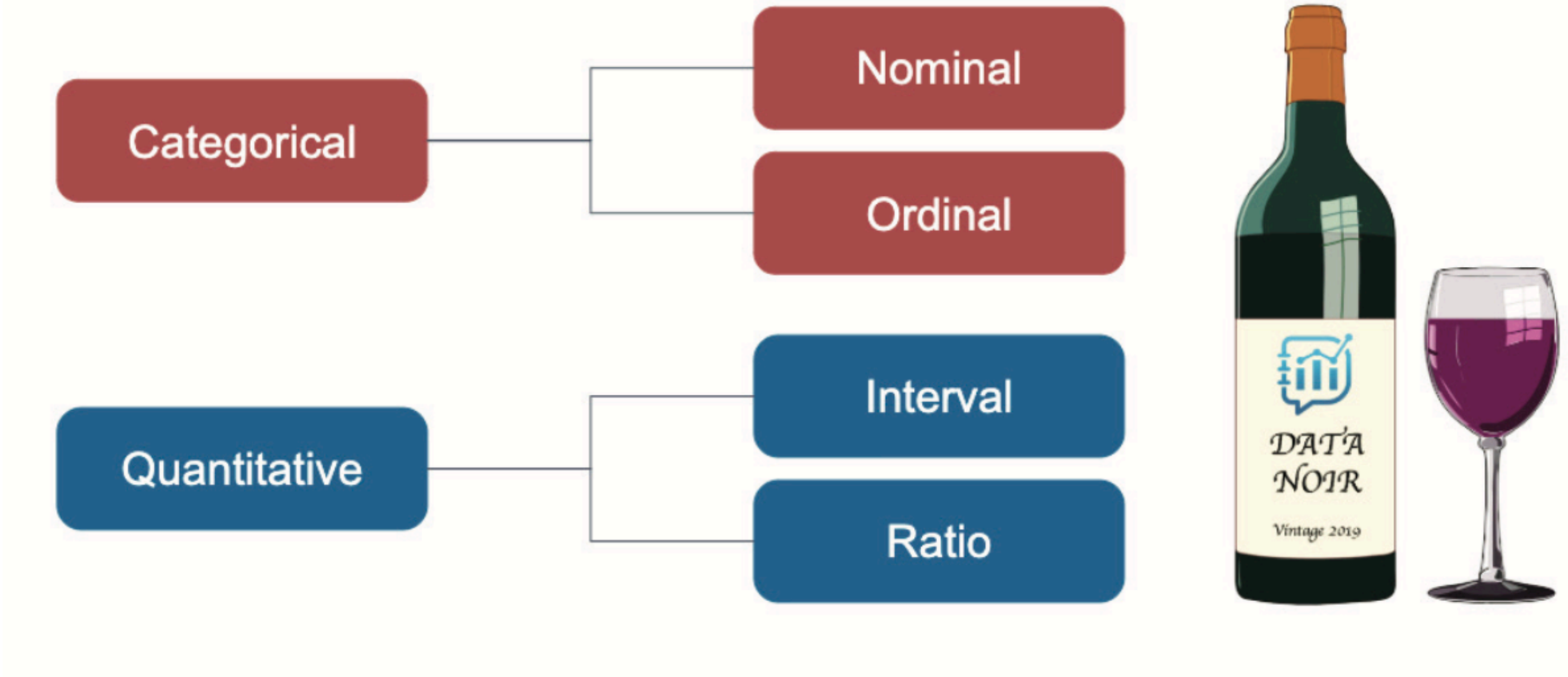
Data-scale

Ratio

Data-scale



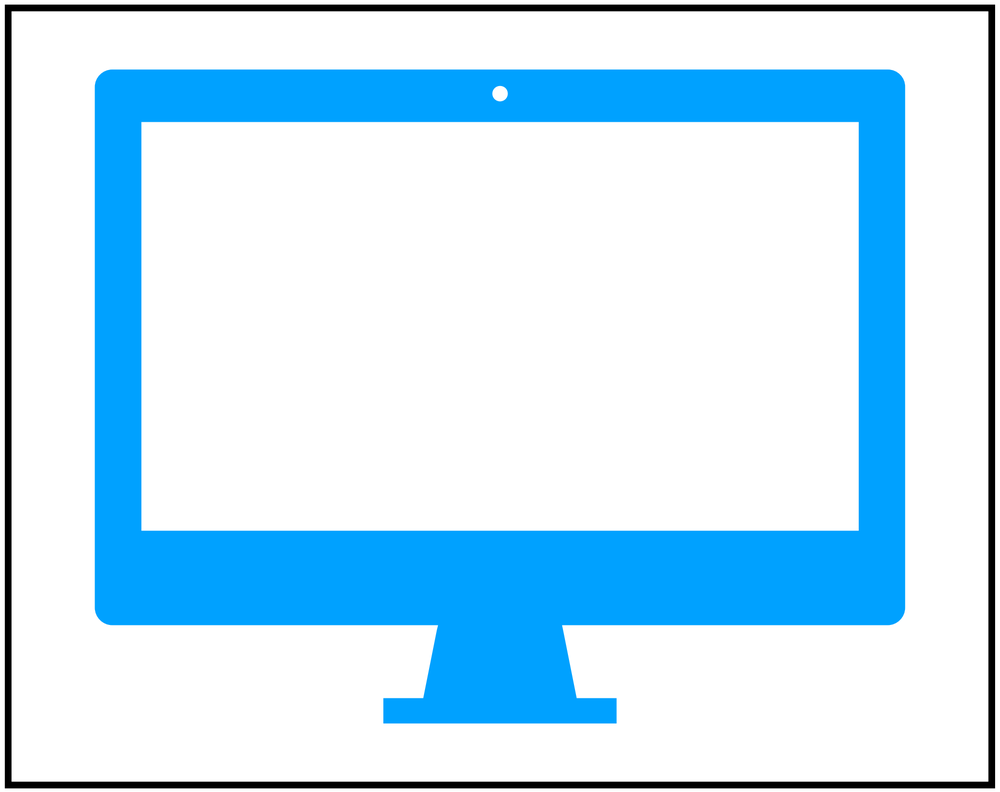



Categorical
Nominal
Ordinal
Quantitative
Interval
Ratio
DATA
NOIR
Vintage 2019

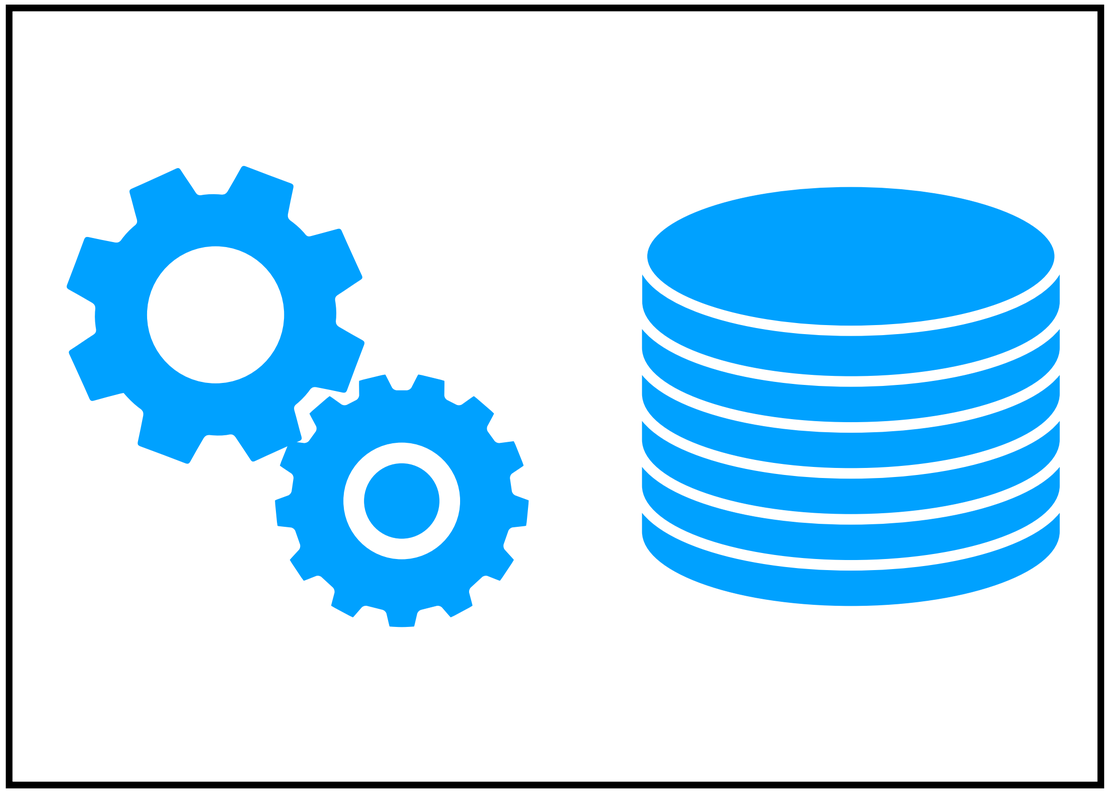
Exercise 1 & 2 Gitbook (30 min)

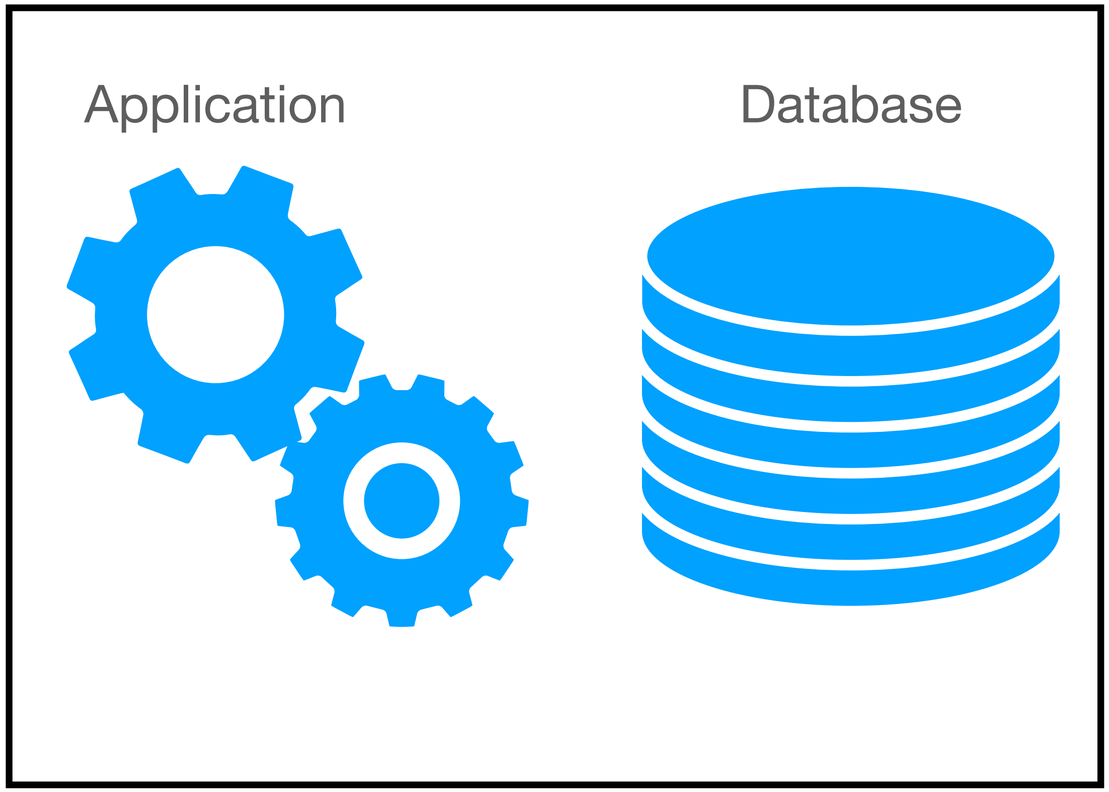
MySQL introduction

Front-end



Backend




Application
Database

Kildemoes	01/02/2022	Mountainbike	Rød	Ikke
Kildemoes	01/02/2002	City-bike	Blå	El-drevet
Kildemoes	01/02/2005	Mountainbike	Rød	Ikke
Kildemoes	01/02/2020	City-bike	Blå	El-drevet
Kildemoes	01/02/2022	Mountainbike	Rød	Ikke

Råmateriale

DATA

Praktiske element

MySQL Database opsætning

- Installation af MySQL Workbench
- Installation af MySQL server
- Oprettelse af database
- Første query

MySQL Database

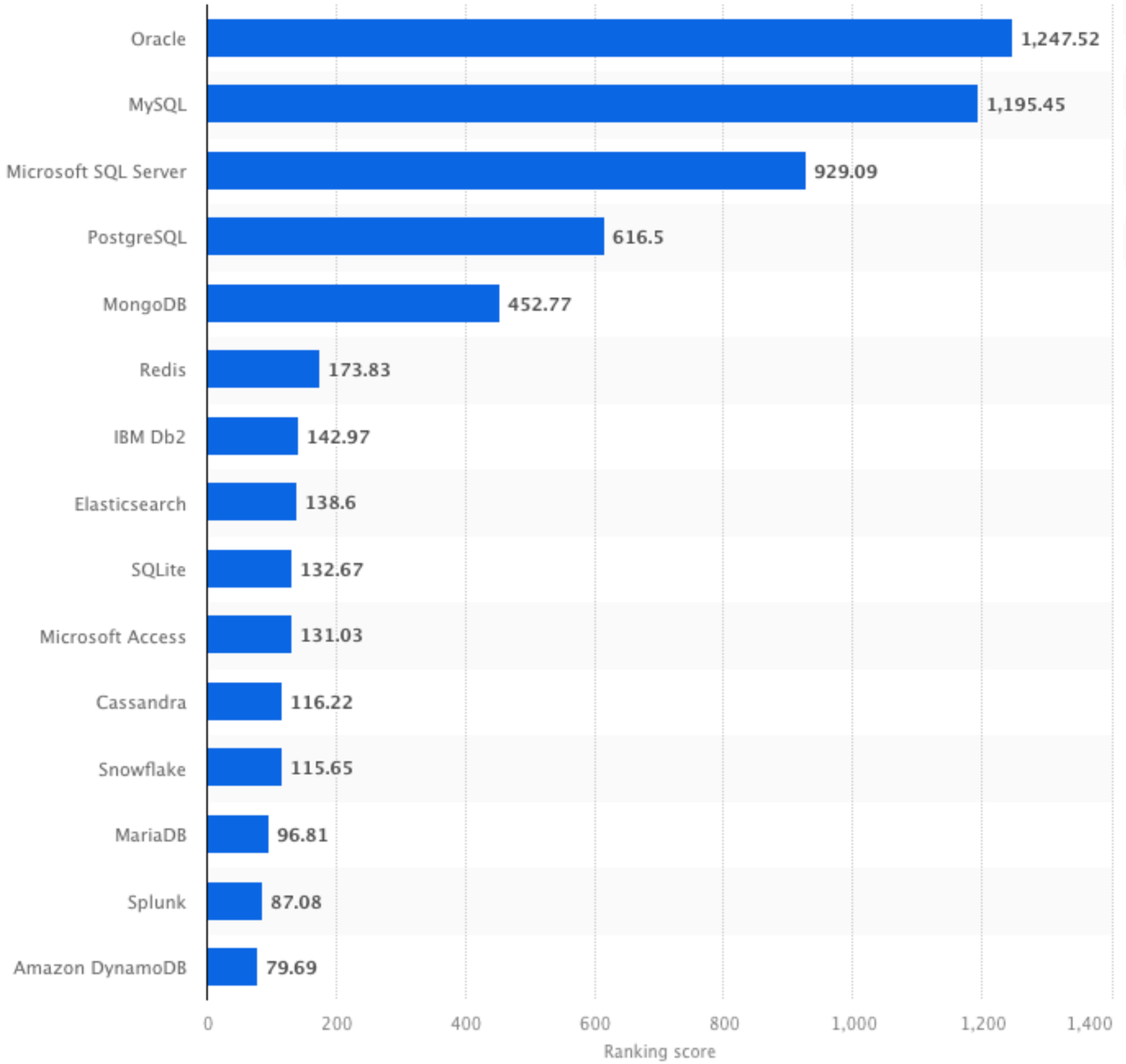
Structured Query Langauge (SQL)

```sql
SELECT MIN(popularity)
FROM spotify
WHERE popularity >
    (SELECT AVG(popularity)
    FROM spotify)
GROUP BY year

```

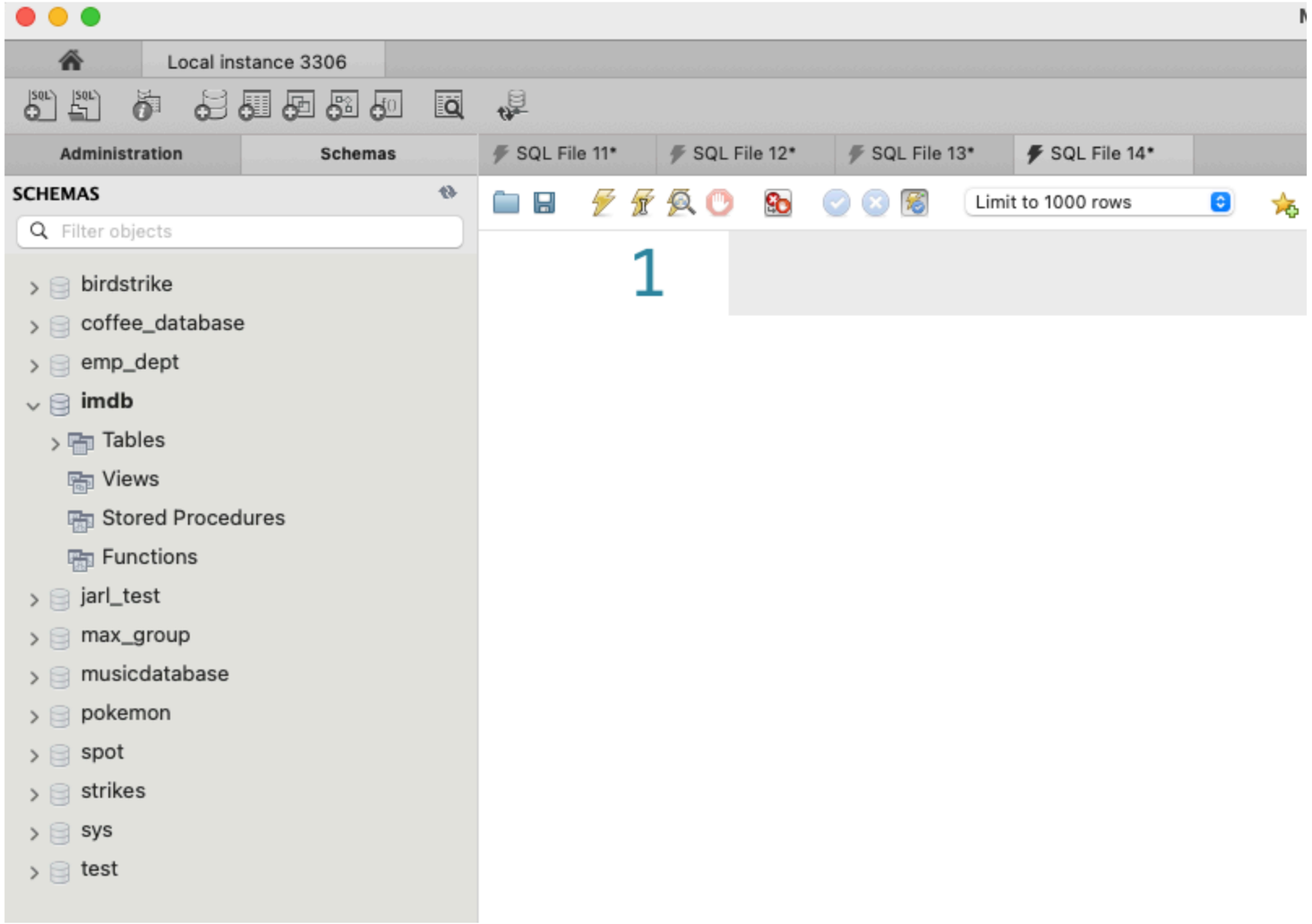
SQL is
universal

MySQL is popular

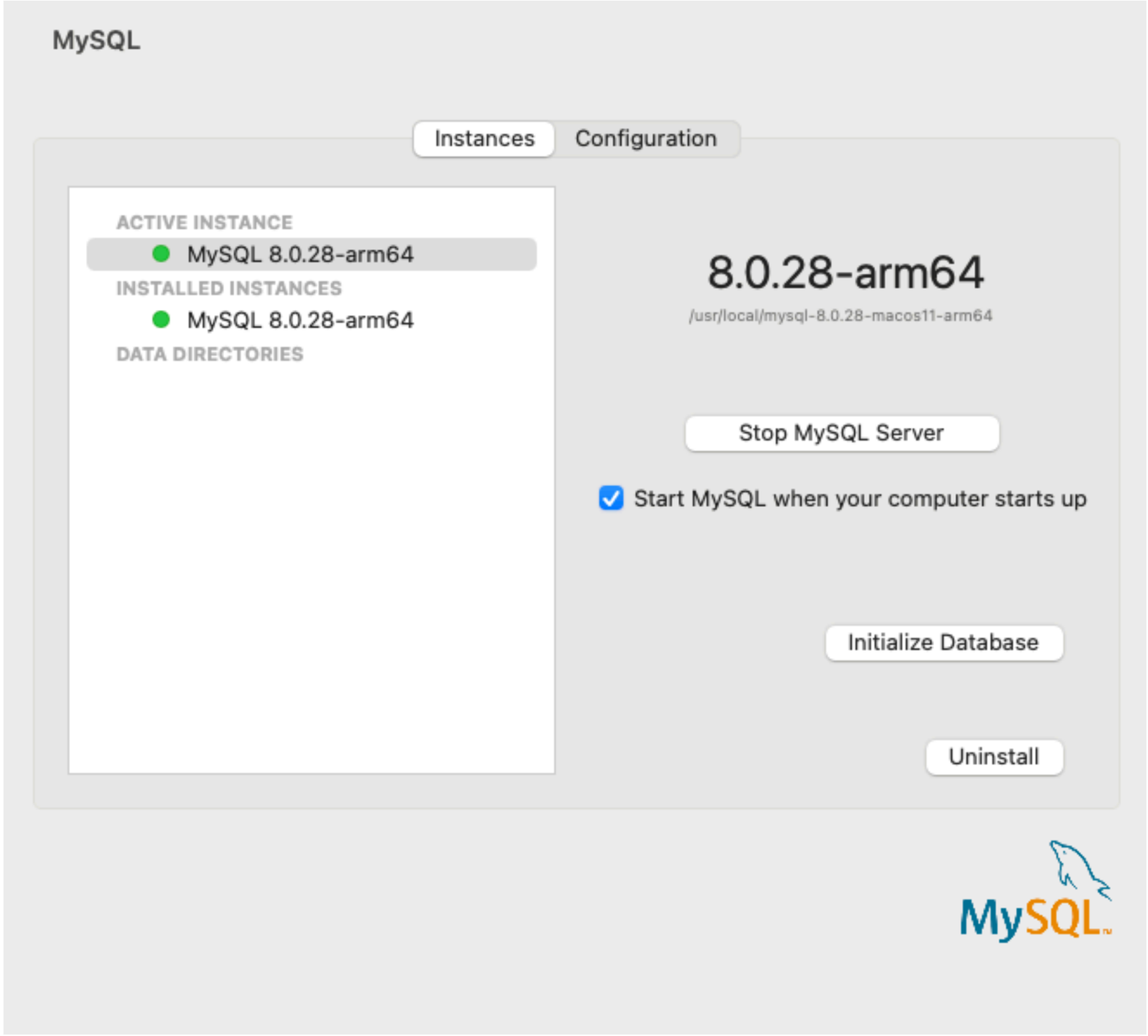


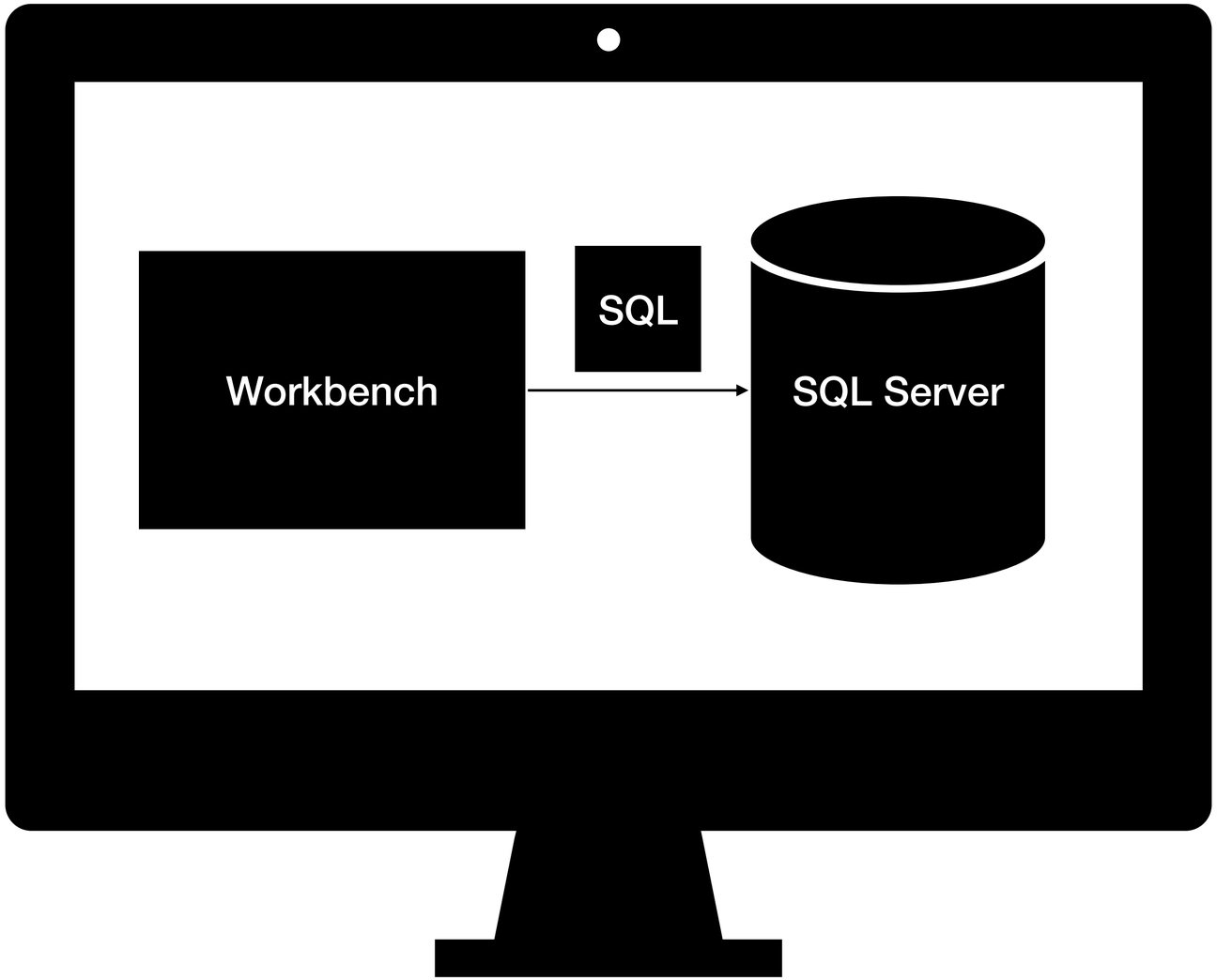
MySQL is
open source

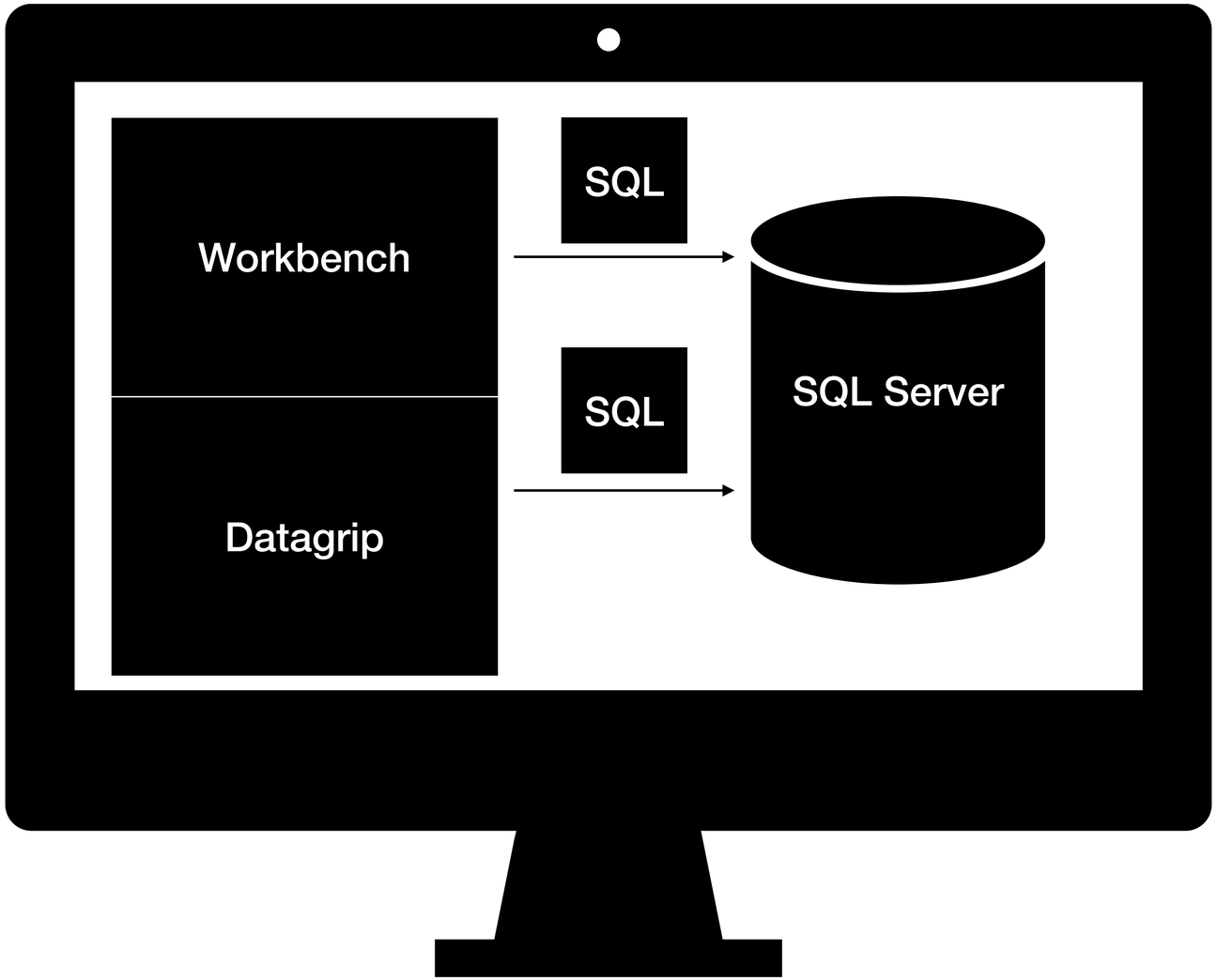
MySQL workbench



MySQL Server




Workbench
SQL
SQL Server


Workbench
Datagrip
SQL
SQL
SQL Server

Installing

MySQL Server + Workbench

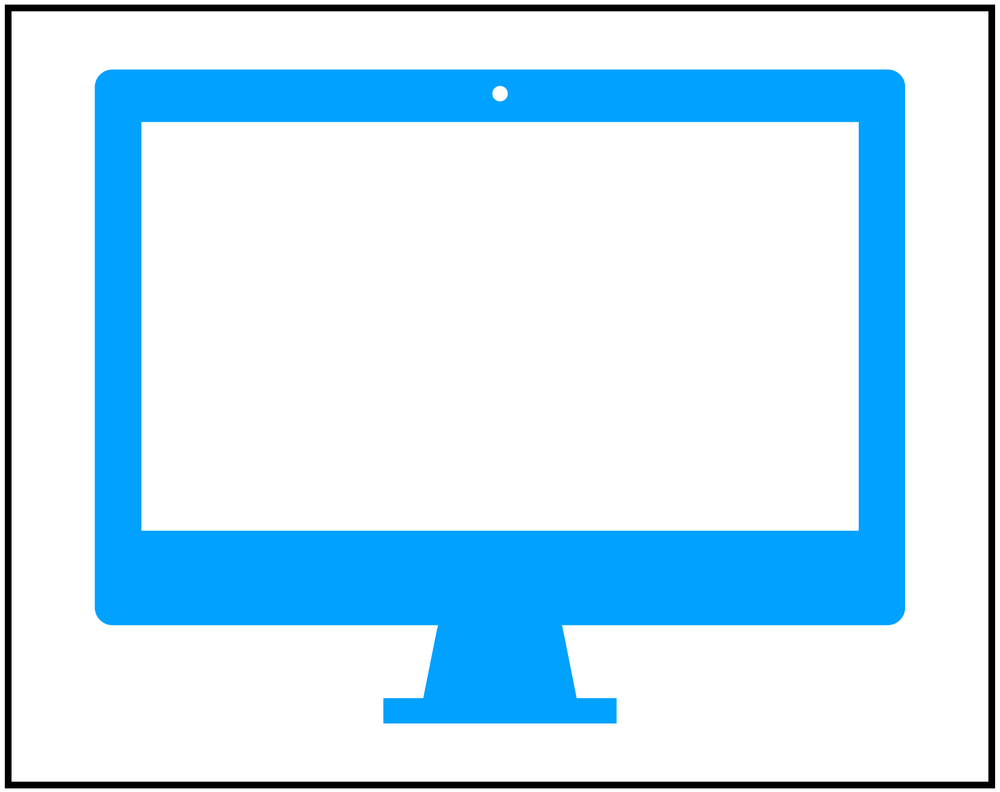
- 1. Installer MySQL Server
  - **HUSK DIT ROOT PASSWORD**
- 2. Installer MySQL workbench
- 3. Indsæt data
- 4. Test det virker

https://nicklasdean.gitbook.io/ita2023-1.-semester/data-literacy/01-what-is-data

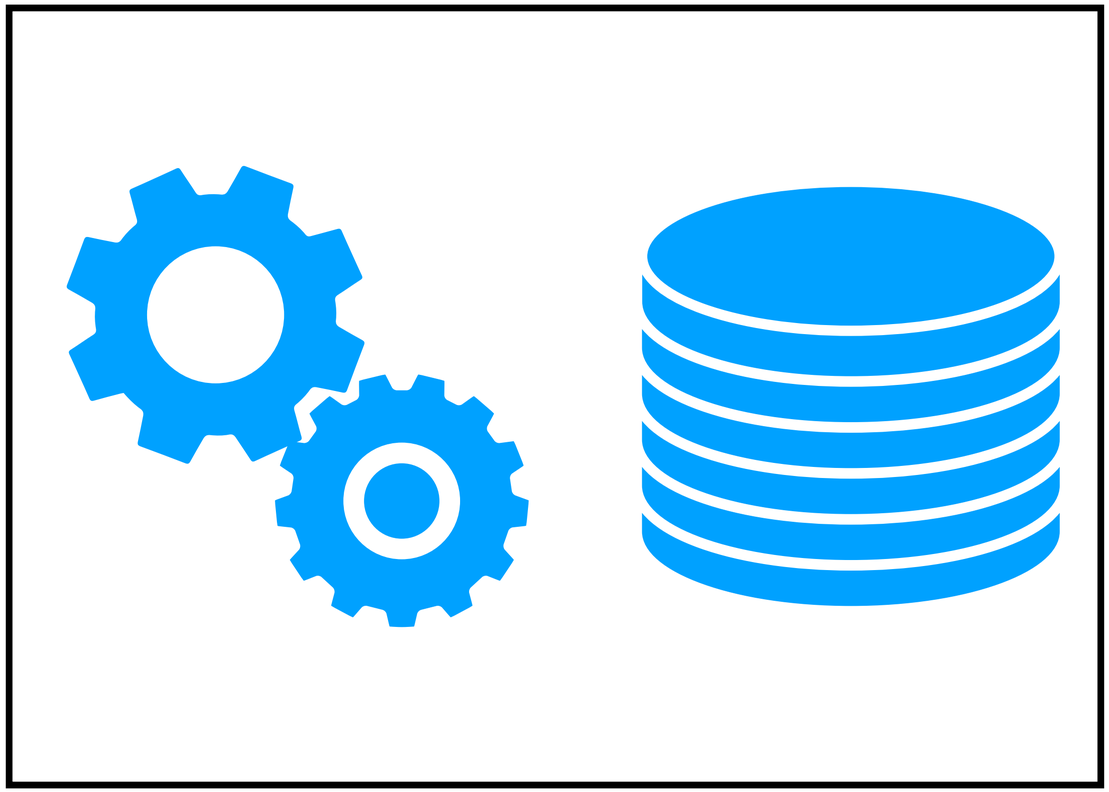
Install MySQL now

Maybe rest
for next time

Front-end

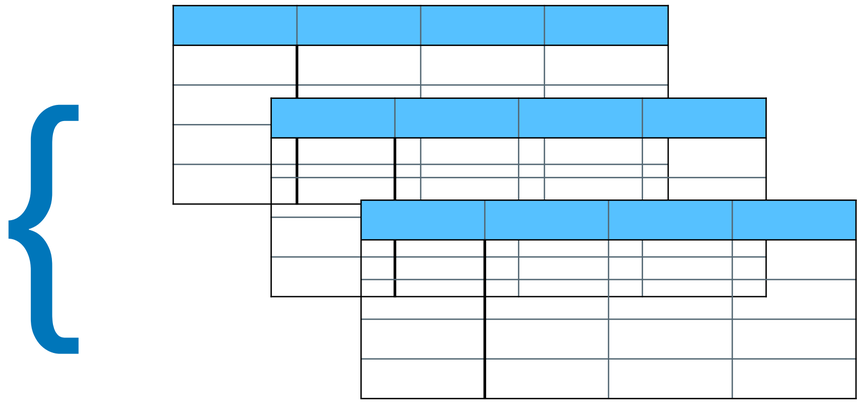


Backend



Database / Schema

Tables



Rows and columns (data)

Tabular Data - data in tables

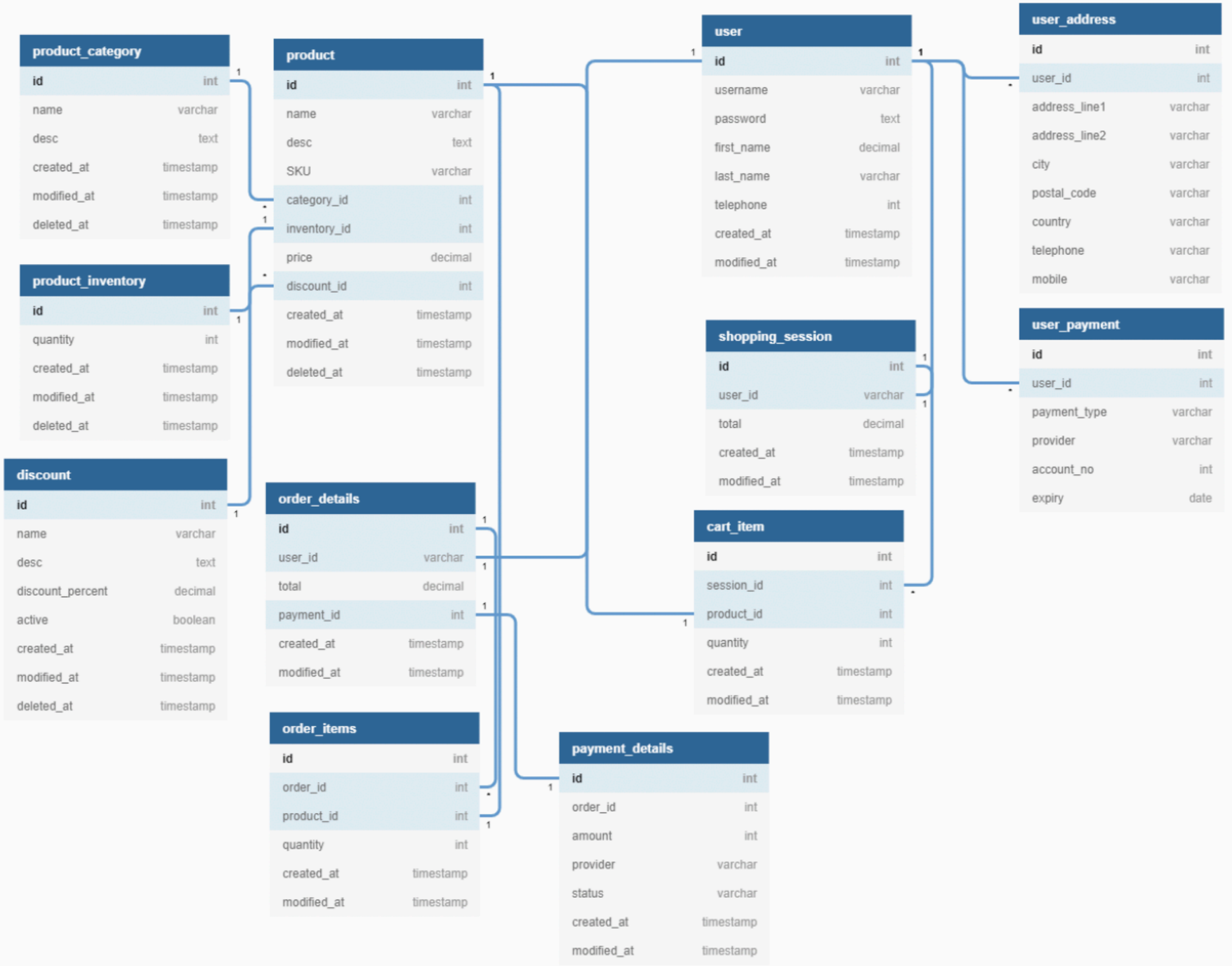
pokedex_number	name	speed	special_defence	special_attack	defence	attack	hp	primary_type	secondary_type
1	Bulbasaur	45	65	65	49	49	45	Grass	Poison
2	Ivysaur	60	80	80	63	62	60	Grass	Poison
3	Venusaur	80	100	100	83	82	80	Grass	Poison
4	Charmander	65	50	60	43	52	39	Fire	null
5	Charmeleon	80	65	80	58	64	58	Fire	null
6	Charizard	100	85	109	78	84	78	Fire	Flying
7	Squirtle	43	64	50	65	48	44	Water	null
8	Wartortle	58	80	65	80	63	59	Water	null
9	Blastoise	78	105	85	100	83	79	Water	null
10	Caterpie	45	20	20	35	30	45	Bug	null
11	Metapod	30	25	25	55	20	50	Bug	null

Column/Attribute/Property

Row/ Tupple/ Entity/

pokedex_number	name	speed	special_defence	special_attack	defence	attack	hp	primary_type	secondary_type
1	Bulbasaur	45	65	65	49	49	45	Grass	Poison
2	Ivysaur	60	80	80	63	62	60	Grass	Poison
3	Venusaur	80	100	100	83	82	80	Grass	Poison
4	Charmander	65	50	60	43	52	39	Fire	null
5	Charmeleon	80	65	80	58	64	58	Fire	null
6	Charizard	100	85	109	78	84	78	Fire	Flying
7	Squirtle	43	64	50	65	48	44	Water	null
8	Wartortle	58	80	65	80	63	59	Water	null
9	Blastoise	78	105	85	100	83	79	Water	null
10	Caterpie	45	20	20	35	30	45	Bug	null
11	Metapod	30	25	25	55	20	50	Bug	null

Relational Database Management System



```
1 SELECT *
2 FROM employees;
```

id	employee_name	job	manager	hiredate	salary	commission	department_number
7369	SMITH	CLERK	7902	1980-12-17	800	NULL	20
7499	ALLEN	SALESMAN	7698	1981-02-20	1600	300	30
7521	WARD	SALESMAN	7698	1981-02-22	1250	500	30
7566	JONES	MANAGER	7839	1981-04-02	2975	NULL	20
7654	MARTIN	SALESMAN	7698	1981-09-28	1250	1400	30
7698	BLAKE	MANAGER	7839	1981-05-01	2850	NULL	30
7782	CLARK	MANAGER	7839	1981-06-09	2450	NULL	10
7788	SCOTT	ANALYST	7566	1987-04-19	3000	NULL	20
7839	KING	PRESIDENT	NULL	1981-11-17	5000	NULL	10
7844	TURNER	SALESMAN	7698	1981-09-08	1500	0	30
7876	ADAMS	CLERK	7788	1987-05-23	1100	NULL	20
7900	JAMES	CLERK	7698	1981-12-03	950	NULL	30
7902	FORD	ANALYST	7566	1981-12-03	3000	NULL	20
7934	MILLER	CLERK	7782	1982-01-23	1300	NULL	10

```
1 SELECT *
2 FROM departments;
```

department_number	department_name	location
10	ACCOUNTING	NEW YORK
20	RESEARCH	DALLAS
30	SALES	CHICAGO
40	OPERATIONS	BOSTON

Relationship

id	employee_name	job	manager	hiredate	salary	commission	department_number
7369	SMITH	CLERK	7902	1980-12-17	800	NULL	20
7499	ALLEN	SALESMAN	7698	1981-02-20	1600	300	30
7521	WARD	SALESMAN	7698	1981-02-22	1250	500	30
7566	JONES	MANAGER	7839	1981-04-02	2975	NULL	20
7654	MARTIN	SALESMAN	7698	1981-09-28	1250	1400	30
7698	BLAKE	MANAGER	7839	1981-05-01	2850	NULL	30
7782	CLARK	MANAGER	7839	1981-06-09	2450	NULL	10
7788	SCOTT	ANALYST	7566	1987-04-19	3000	NULL	20
7839	KING	PRESIDENT	NULL	1981-11-17	5000	NULL	10
7844	TURNER	SALESMAN	7698	1981-09-08	1500	0	30
7876	ADAMS	CLERK	7788	1987-05-23	1100	NULL	20
7900	JAMES	CLERK	7698	1981-12-03	950	NULL	30
7902	FORD	ANALYST	7566	1981-12-03	3000	NULL	20
7934	MILLER	CLERK	7782	1982-01-23	1300	NULL	10

department_number	department_name	location
10	ACCOUNTING	NEW YORK
20	RESEARCH	DALLAS
30	SALES	CHICAGO
40	OPERATIONS	BOSTON

Structured Query Language

Database: MySQL

Language: SQL

Declarative nature of SQL

What I want - not how

Loading data . . . DDL

- Creating a database
- Creating a table
- Inserting data

```sql
CREATE DATABASE pokemon;

CREATE TABLE pokemon (
    pokedex_number int,
    name VARCHAR(45),
    speed int,
    special_defence int,
    special_attack int,
    defence int,
    attack int,
    hp int,
    primary_type VARCHAR(45),
    secondary_type VARCHAR(45)
);


INSERT INTO pokemon (pokedex_number, `name`, speed, special_defence,
special_attack, defence, attack, hp, primary_type, secondary_type)
VALUES (1, 'Bulbasaur', 45, 65, 65, 49, 49, 45, 'Grass', 'Poison');
```

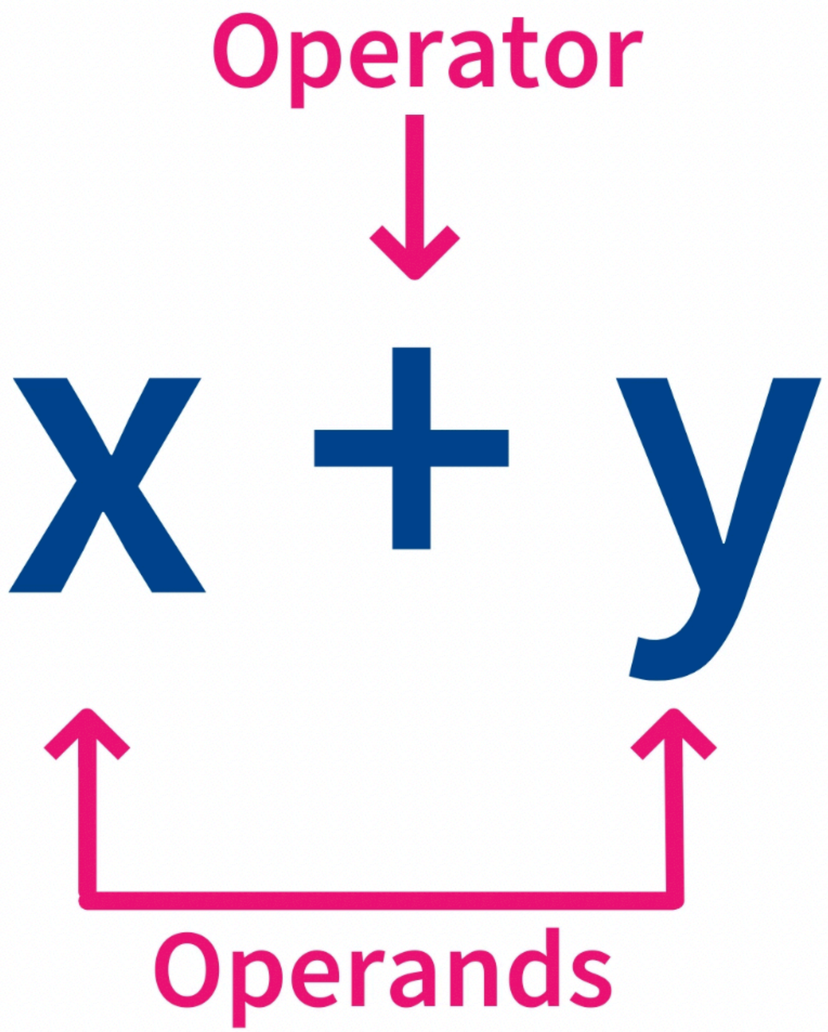
Fetch all the data

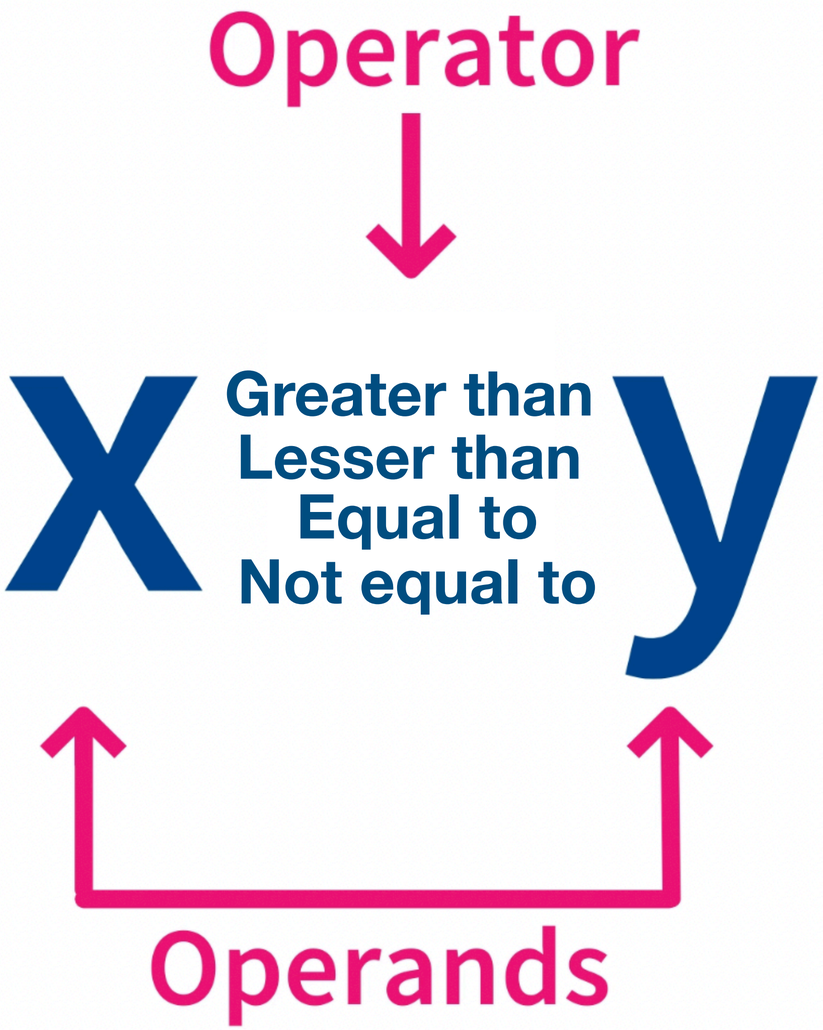
```
SELECT [Columns]
FROM [Table Name];
```

Fetch some of the data

```
SELECT [Columns]
FROM [Table Name]
WHERE [Filter Expression]
```

Filter / Boolean Expression

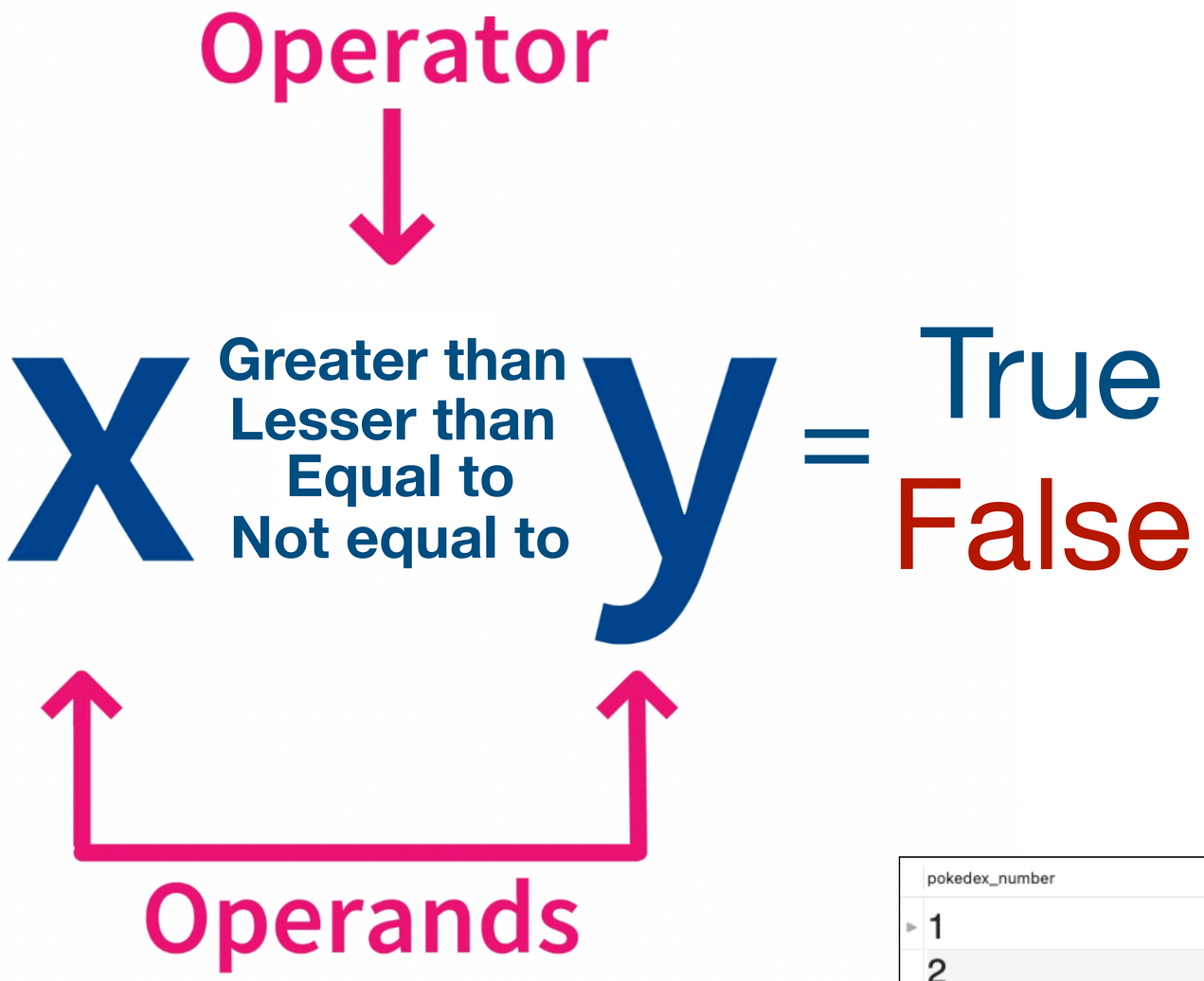

Operator
x + y
Operands


Operator
x
Greater than
Lesser than
Equal to
Not equal to
y
Operands

=

True

False




```
SELECT *
FROM pokemon
WHERE attack > 50
```

pokedex_number	name	speed	special_defence	special_attack	defence	attack
1	Bulbasaur	45	65	65	49	49
2	Ivysaur	60	80	80	63	62
3	Venusaur	80	100	100	83	82

Install + Exercises 3+4
Gitbook types and scales

https://nicklasdean.gitbook.io/ita2023-1.-semester/data-literacy/02-types-and-scales