Tomas Sandnes
tomas.sandnes@kristiania.no

Week 8

DB1102 / PGR 111 – DATABASES

Today's topics

(Today's chapters: 7 + 8.1 in Norwegian book, 6 in English)

Walkthrough of task 2-f from the mandatory coursework.

- (The "last name of tour guide" task. O_o)

- ER modelling
  - Part 2 (of 2)

Coursework

task walkthrough

Coursework task 2-f walkthrough

Mandatory coursework, Task 2-f:

f) Create a query that retrieves information about all those registered for hiking trips with a demanding level of difficulty. The following columns should be included:
   - The name of the cabin where the route begins.
   - Number of days the trip takes
   - When the hiking trip starts (date)
   - Last name of tour guide (if the hiking trip is set up with a tour guide)
   - First name of registered participant
   - Last name of registered participant
   - Email address of registered participant.

Coursework task 2-f walkthrough – cont.

- The solution: (walkthrough done live in class, see video)

```
SELECT Cabin.CabinName, Route.NrOfDays, HikingTrip.StartDate, TourGuide.LastName, Participant.FirstName,
       Participant.LastName, Participant.Email
FROM Registration
JOIN Route ON Registration.RouteNr = Route.RouteNr
JOIN Cabin ON Route.CabinNr = Cabin.CabinNr
JOIN HikingTrip ON Registration.RouteNr = HikingTrip.RouteNr AND Registration.StartDate = HikingTrip.StartDate
LEFT JOIN Person AS TourGuide ON HikingTrip.Email = TourGuide.Email
JOIN Person AS Participant ON Registration.Email = Participant.Email
WHERE Route.Difficulty = 'demanding';
```

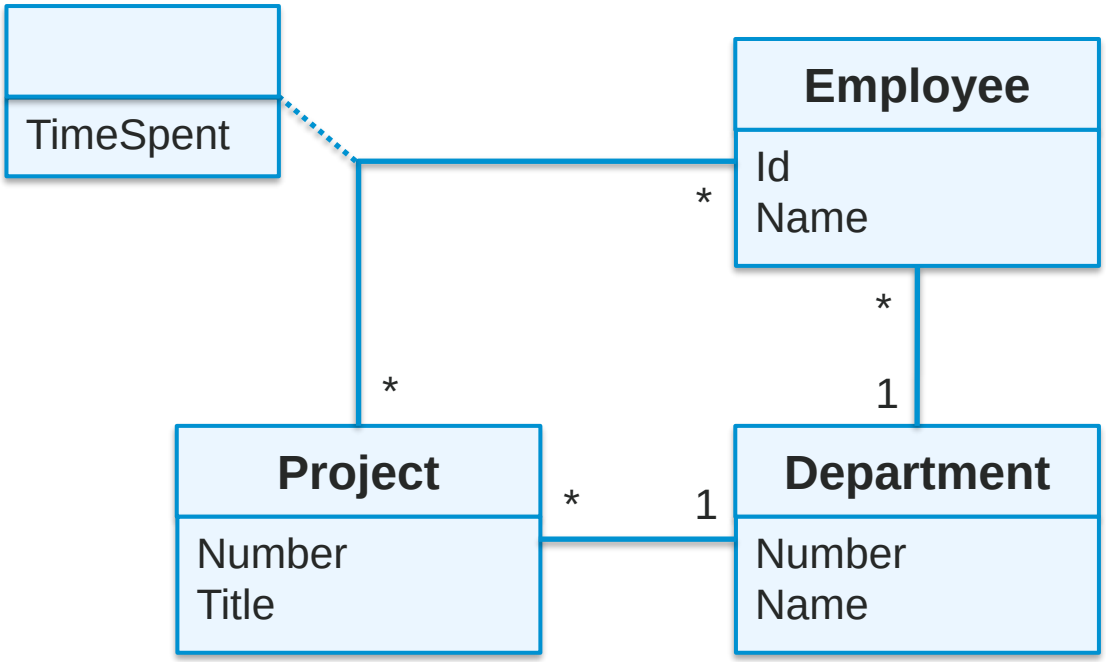
  - *NOTE:* Using "`JOIN HikingTrip`" with the following `ON`-column gives to many rows in the answer: ("cartesian product"-ish)

```
JOIN HikingTrip ON Route.RouteNr = HikingTrip.RouteNr
```

ER modelling

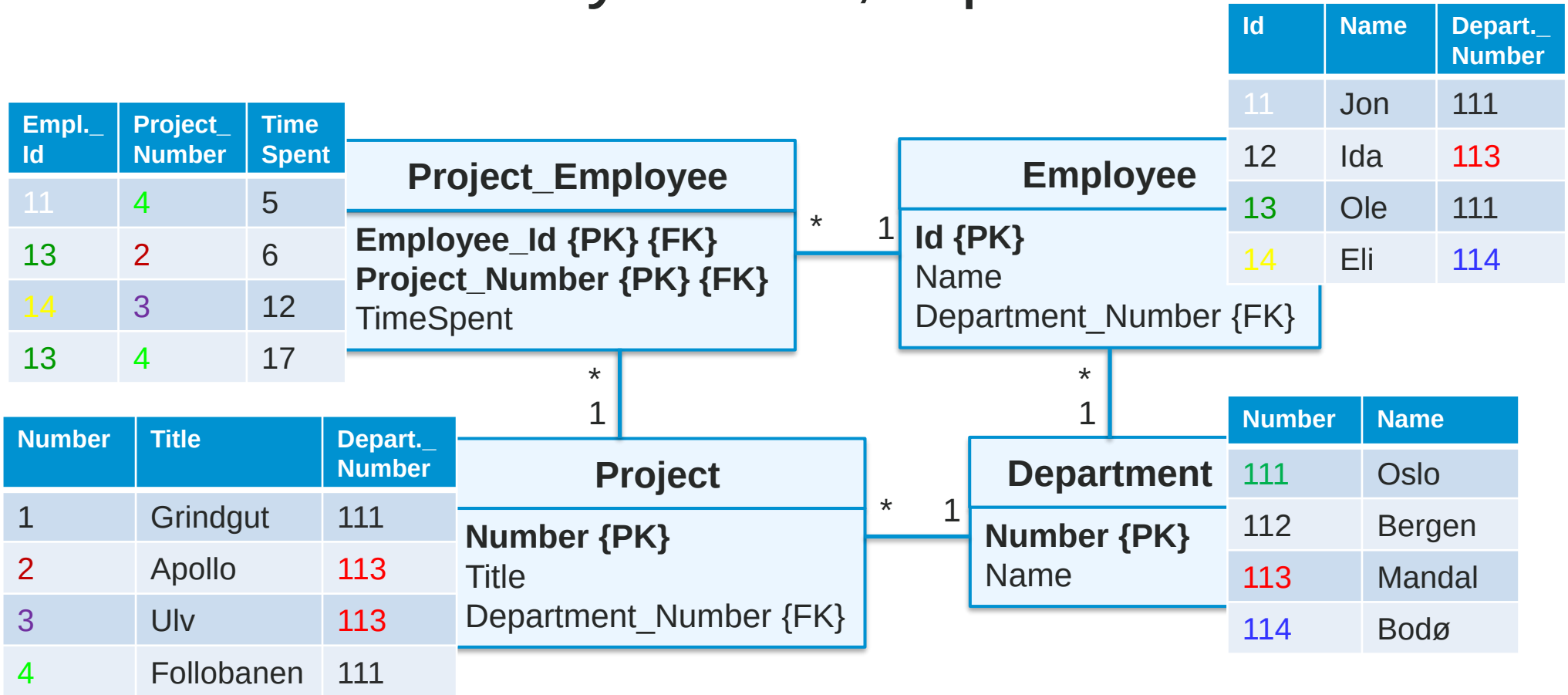
part 2 (of 2)

ER Model, repetition



ER Model ready for DB, repetition

Empl._Id	Project_Number	Time Spent
11	4	5
13	2	6
14	3	12
13	4	17



Id	Name	Depart._Number
11	Jon	111
12	Ida	113
13	Ole	111
14	Eli	114

Number	Title	Depart._Number
1	Grindgut	111
2	Apollo	113
3	Ulv	113
4	Follobanen	111

Number	Name
111	Oslo
112	Bergen
113	Mandal
114	Bodø

Surrogate key

- On the previous slide, Project_Number and Attribute_Id together form a primary key.
  - Instead of using these two attributes together, we could add a new attribute: a surrogate key.

- A surrogate key can be explained as:
  - *A serial number without "meaning" which is only used for unambiguous identification.*
  - Surrogate keys are usually generated automatically by the system.

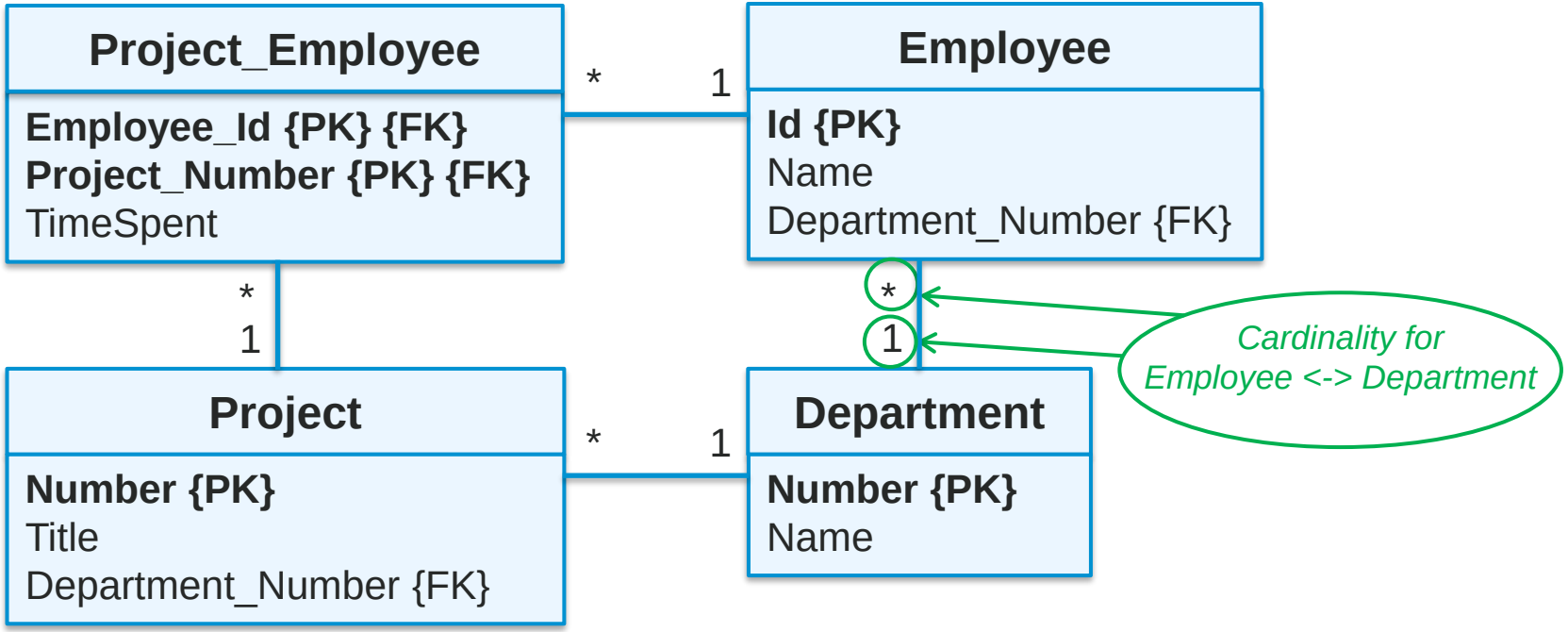
- Interesting!
  - What is best to use? A surrogate primary key or the "natural" composite primary key?
  - Is there a final answer?

- No final answer, but you can read more here:
  - stackoverflow.com/questions/963809/should-i-use-composite-primary-keys-or-not
  - techrepublic.com/article/the-great-primary-key-debate/

ER modelling, new terminology

- Cardinality for ER modelling:
- Cardinality indicates the largest number of connections an entity participating in a given relationship can have.
- Example: An employee can belong to a maximum of one department.
  - The cardinality of an employee in the employee <-> department relationship is therefore "one".
  - (For department in the same relationship it is "many".)

ER modelling, new terminology – cont.

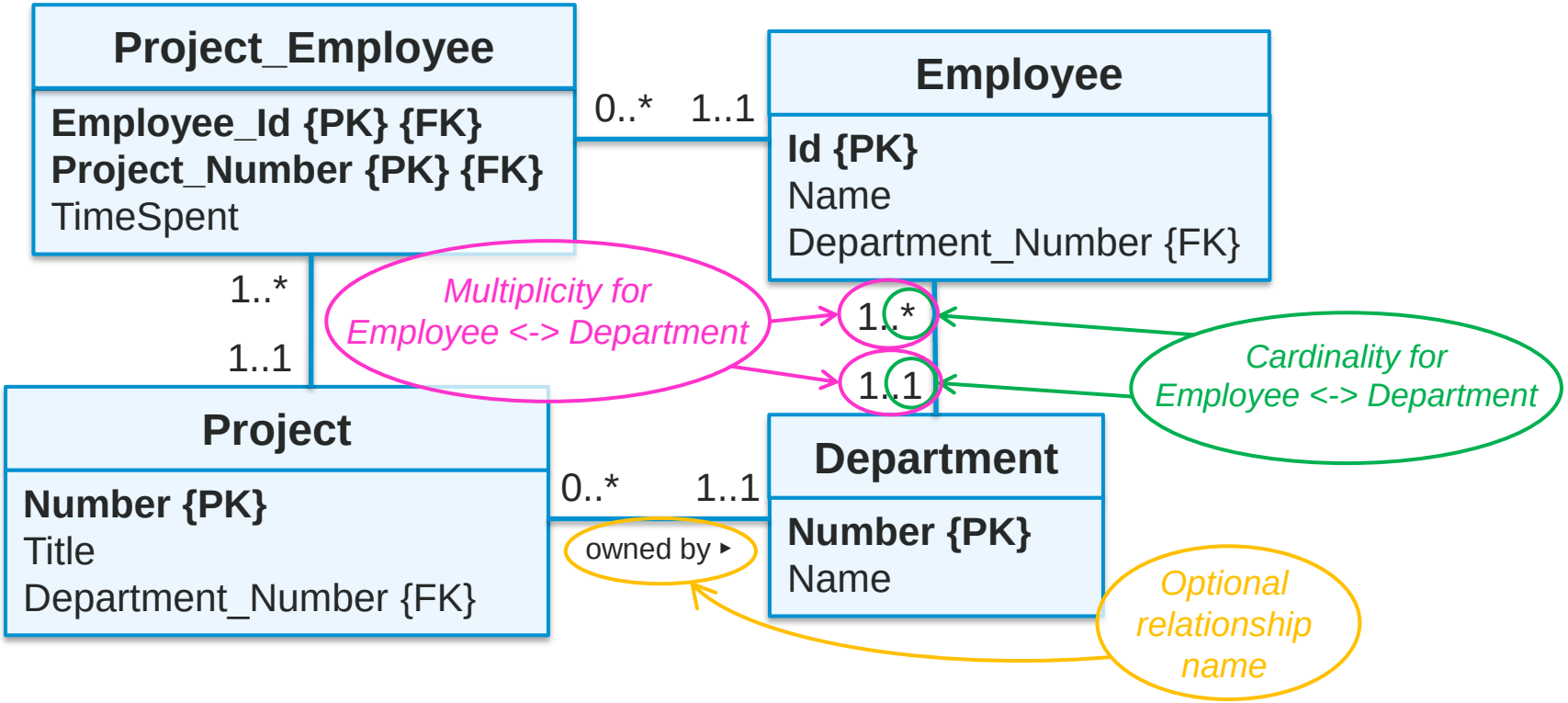


ER modelling, new terminology – cont.

- Participation: Where cardinality indicates the maximum number of connections an entity can have, participation indicates the minimum number of connections an entity can have.
  - *Example:* An employee can belong to a minimum of 1 department. The participation of an employee in the employee <-> department relationship is therefore 1. (For a department in the same relationship, it is also 1.)
  - Participation is indicated before cardinality.

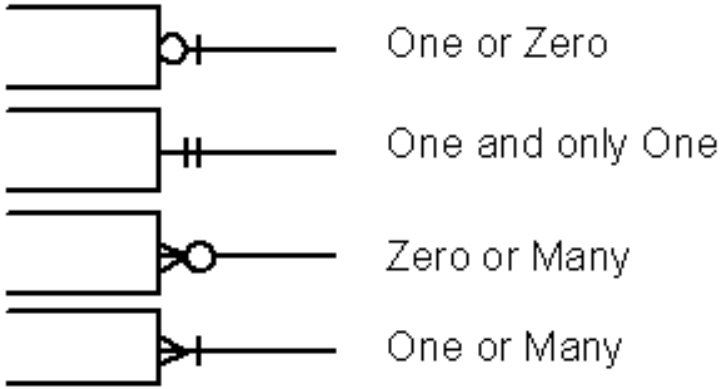
- Multiplicity: Participation and cardinality together constitute multiplicity.
  - *Example:* An employee can belong from 1 to 1 departments.
    (And a department can have from 1 to many employees.)
  - Multiplicity is written in UML as: "0..1", "1..1" (possibly only "1"), "0..*","1..*".

ER modelling, new terminology – cont.



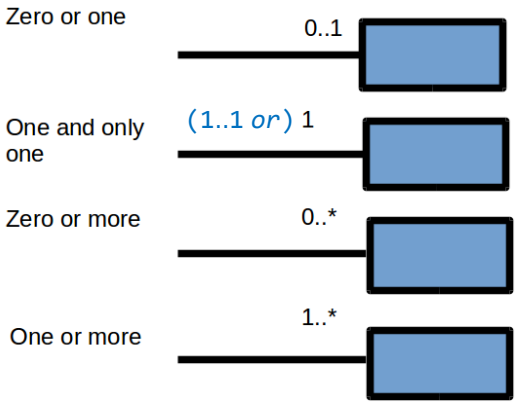
Crow's Foot and UML, repetition

Crow's Foot



Source: tdan.com/crows-feet-are-best/7474

UML



Source: vertabelo.com/blog/uml-notation/

Keys

some repetition and some new

Keys – some repetition and some new

- These we should know from before:
  - Primary key (PK). The column(s) we choose to use as our unique row-identifier in a table. (Can be a surrogate key, as talked about earlier today.)
  - Foreign key (FK). The column(s) referring to a primary key in another (well, actually, possibly the same) table.

- These are NEW NOW:
  - Super key. One or more columns that form a unique identifier for a row in a table. *Note:* These may consist of more columns than the smallest, unique selection.
  - Candidate key. A super key that cannot be reduced to fewer columns if it is to remain a unique identifier.

- Why learn these extra keys?
  - We need to understand what they are when we start normalizing our databases. (Content for the upcoming lessons.)

2 quick questions regarding keys

- What super keys do we have here?
- What candidate keys do we have here?

Code	Name	Continent	Region	SurfaceArea	IndepYear	Population	LifeExpectancy	GNP	GNPOld	LocalName	GovernmentForm	HeadOfState	Capital	Code2
ABW	Aruba	North America	Caribbean	193.00	NULL	103000	78.4	828.00	793.00	Aruba	Nonmetropolitan T...	Beatrix	129	AW
AFG	Afghani...	Asia	Souther...	652090.00	1919	22720000	45.9	5976.00	NULL	Afganistan/A...	Islamic Emirate	Mohammad Omar	1	AF
AGO	Angola	Africa	Central ...	1246700.00	1975	12878000	38.3	6648.00	7984.00	Angola	Republic	JosÃ© Eduard...	56	AO
AIA	Anguilla	North America	Caribbean	96.00	NULL	8000	76.1	63.20	NULL	Anguilla	Dependent Territor...	Elisabeth II	62	AI
ALB	Albania	Europe	Souther...	28748.00	1912	3401200	71.6	3205.00	2500.00	ShqipÃ«ria	Republic	Rexhep Mejdani	34	AL
AND	Andorra	Europe	Souther...	468.00	1278	78000	83.5	1630.00	NULL	Andorra	Parliamentary Copri...		55	AD
ANT	Netherl...	North America	Caribbean	800.00	NULL	217000	74.7	1941.00	NULL	Nederlandse...	Nonmetropolitan T...	Beatrix	33	AN
ARE	United ...	Asia	Middle East	83600.00	1971	2441000	74.1	37966...	36846.00	Al-Imarat al-...	Emirate Federation	Zayid bin Sulta...	65	AE
ARG	Argentina	South Amer...	South A...	2780400.00	1816	37032000	75.1	34023...	323310.00	Argentina	Federal Republic	Fernando de la...	69	AR
ARM	Armenia	Asia	Middle East	29800.00	1991	3520000	66.4	1813.00	1627.00	Hajastan	Republic	Robert KotÂšar...	126	AM
ASM	America...	Oceania	Polynesia	199.00	NULL	68000	75.1	334.00	NULL	Amerika Sam...	US Territory	George W. Bush	54	AS
ATA	Antarctica	Antarctica	Antarctica	13120000.00	NULL	0	NULL	0.00	NULL	Â–	Co-administrated		NULL	AQ
ATF	French ...	Antarctica	Antarctica	7780.00	NULL	0	NULL	0.00	NULL	Terres austr...	Nonmetropolitan T...	Jacques Chirac	NULL	TF
ATG	Antigua...	North America	Caribbean	442.00	1981	68000	70.5	612.00	584.00	Antigua and ...	Constitutional Mona...	Elisabeth II	63	AG
AUS	Australia	Oceania	Australia ...	7741220.00	1901	18886000	79.8	35118...	392911.00	Australia	Constitutional Mona...	Elisabeth II	135	AU
AUT	Austria	Europe	Western...	83859.00	1918	8091800	77.7	21186...	206025.00	Ã–sterreich	Federal Republic	Thomas Klestil	1523	AT
AZE	Azerbaijan	Asia	Middle East	86600.00	1991	7734000	62.9	4127.00	4100.00	AzÃ¤rbaycan	Federal Republic	HeydÃ¤r Ã„liyev	144	AZ
BDI	Burundi	Africa	Eastern ...	27834.00	1962	6695000	46.2	903.00	982.00	Burundi/Ubu...	Republic	Pierre Buyoya	552	BI

Today's exercises & looking ahead

- Now: 2 hours of exercises.

- Exercises are on Canvas, as usual. Short summary:
  - ER modelling, based on the descriptions given in the exercises.

- Main contents for the next lesson:
  - Normalization ("ER model optimization"), part 1.

The