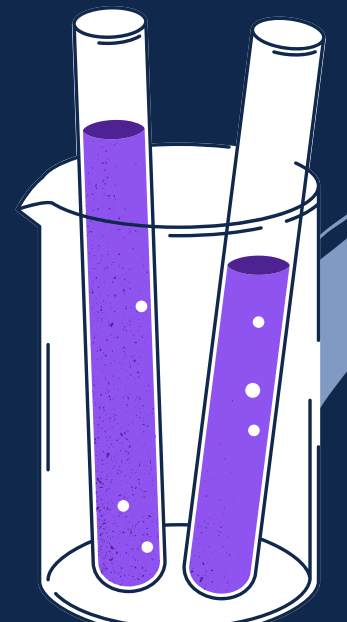


Project 2

# PROJECT DATABASE: SCIENCE



Karolina, Elnara,  
Jonathon





# CONTENT



1. Planning in Jira
  2. Methodology
  3. Data Sources
  4. ER, DataBase Schema
  5. MySQL Queries
  6. Challenges
  7. Highlights
- 
- 

STEP 1

JIRA PART

TO DO 6 ISSUES

Methodology

✓ P2-6

Data Retrieving

✓ P2-7

Index Calculation

✓ P2-9

Create SQL Database and empty tables

✓ P2-11

Populate Empty tables SQL with data

IN PROGRESS 5 ISSUES

Find 4indicators

✓ P2-2

get the data in csv file

✓ P2-10

Find ideas

✓ P2-5

Find in which region we would like to gain insights into.

✓ P2-3

CREATE PRESENTATION

✓ P2-15

DONE 3 ISSUES ✓

Find data sources

✓ P2-1 ✓

ER Model

✓ P2-8 ✓

Country

PK 16 INT NOT NULL

country\_name CHAR(50) NOT NULL

country\_code CHAR(50)

R&D Expenditure

PK 16 INT NOT NULL

FK1 country\_id 16 INT NOT NULL

country\_code CHAR(50) NOT NULL

YEAR YEAR NOT NULL

Tax Subsidy

PK 16 INT NOT NULL

FK1 country\_id 16 INT NOT NULL

country\_code CHAR(50) NOT NULL

YEAR YEAR NOT NULL

Researchers in R&D

PK 16 INT NOT NULL

FK1 country\_id 16 INT NOT NULL

country\_code CHAR(50) NOT NULL

YEAR YEAR NOT NULL

Journal Articles

PK 16 INT NOT NULL

FK1 country\_id 16 INT NOT NULL

country\_code CHAR(50) NOT NULL

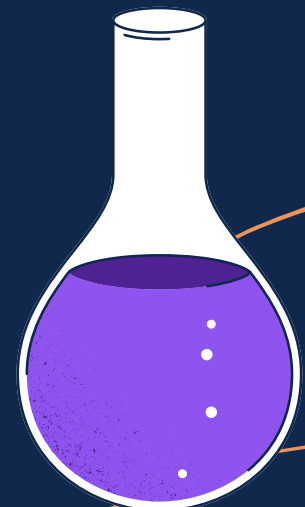
YEAR YEAR NOT NULL



## STEP 2

# METHODOLOGY

- Create a broad measure of R&D maturity
- Focus on Europe for Data availability
- Analyse changes over time
- Discern if joining the EU has an impact on R&D maturity



## STEP 3

# DATA SOURCES

### World Bank - API extraction

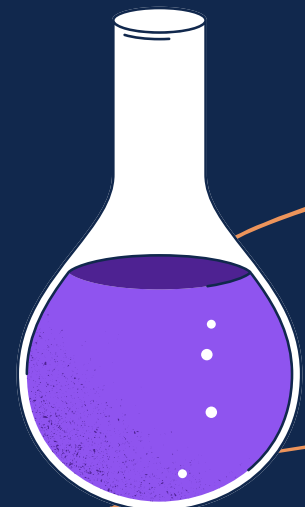
- R&D expenditure (% of GDP)
- Researchers in R&D (per million people)
- Scientific and technical journal articles (per 1000 people)

### OECD - CSV extraction

- Implied tax subsidy rates on R&D expenditures



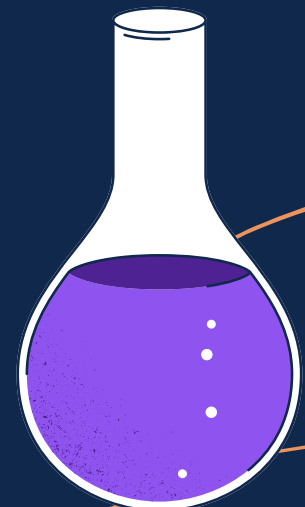
R&D Maturity INDEX ( RDMI)



STEP 4

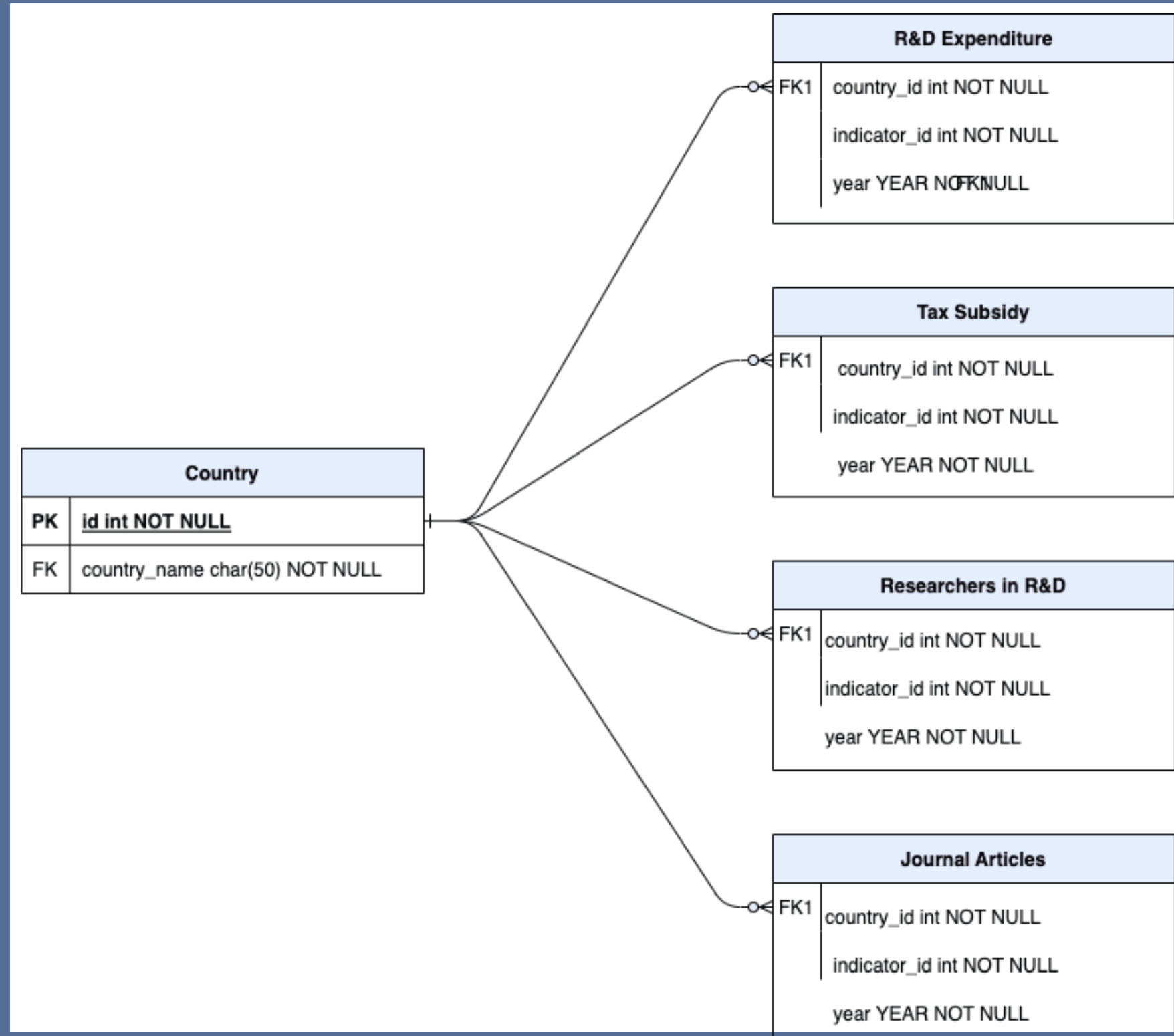
DATA -> CSV

- DATA SELECTION
- DATAFRAME
- TRANSPOSE



## STEP 4

# ER MODEL

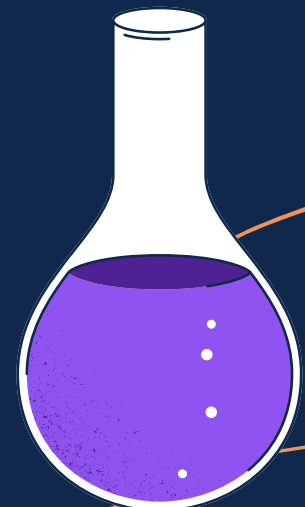


## STEP 5

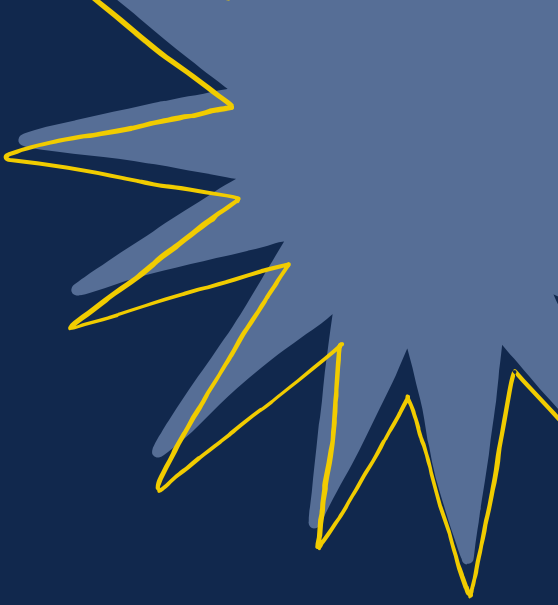

# MySQL QUERIES

- CREATE
- INSERT
- NORMALISE
- JOIN >>>>
- INSERT

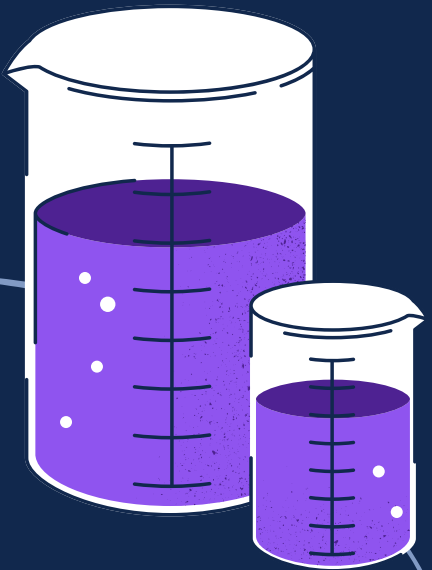

```
CREATE TEMPORARY TABLE joined_indicators_4
SELECT journal_articles_2.country_code,
journal_articles_2.year,
journal_articles_2.norm_ind,
rd_expenditure_2.norm_ind AS norm_ind2,
researchers_rd_2.norm_ind AS norm_ind3,
tax_subsidy_2.norm_ind AS norm_ind4
FROM journal_articles_2
LEFT JOIN rd_expenditure_2
ON journal_articles_2.country_code=rd_expenditure_2.country_code
AND journal_articles_2.year=rd_expenditure_2.year
LEFT JOIN researchers_rd_2
ON rd_expenditure_2.country_code=researchers_rd_2.country_code
AND rd_expenditure_2.year=researchers_rd_2.year
LEFT JOIN tax_subsidy_2
ON researchers_rd_2.country_code=tax_subsidy_2.country_code
AND researchers_rd_2.year=tax_subsidy_2.year
;
```







FINAL\_INDEX =  
AVG(DATA\_POINT +  
DATA\_POINT2 +  
DATA\_POINT3 +  
DATA\_POINT4 )



STEP 6

# CHALLENGES

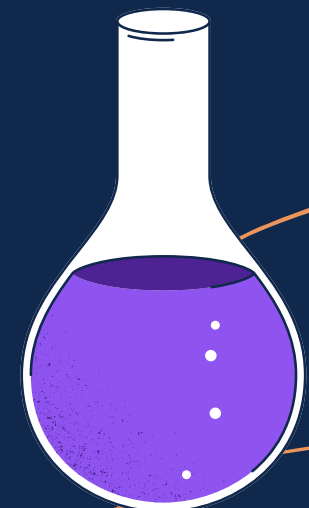
- Find the data
- Find more data
- Placing data in MySQL
- Reaching the final values



STEP 7

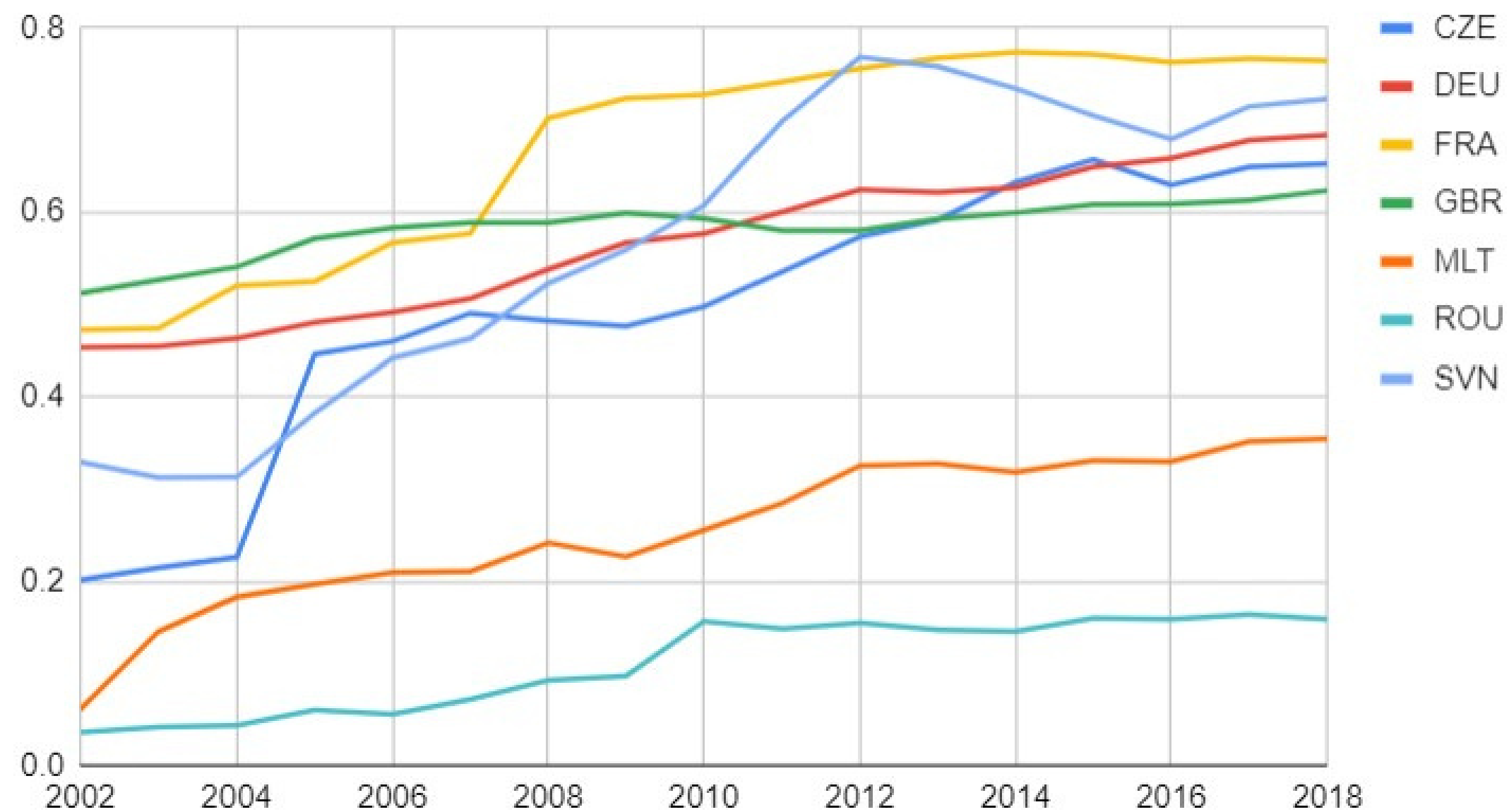
# HIGHLIGHTS

- Use of knowledge gained for the past fortnight
- Implanting database in MySQL
- Overcoming challenges of using panel data in MySQL



# FINAL RESULT

CZE, DEU, FRA, GBR, MLT...



LET'S CHECK THE CODE!

