



Technische Universität Berlin  
Faculty VII (Economics and Management)  
Workgroup for Infrastructure Policy (WIP)

# Operations Research – Coding Lab

## Homework 4

Author: OR Team  
or@wip.tu-berlin.de  
Summer Semester 2022

## Homework 4

You are given two datasets of a European market. The first one contains production firms with a maximum production capacity of each plant and shipping costs to the markets, where demand is located. The second one contains the markets and their respective demand.

### Exercise 1 - Import and clean input data

- Import both datasets using the package *CSV*.
- In the producers dataset, deal with missing values by replacing the missing values with average values from the available data. To do so, subset the dataframe to only contain complete rows and calculate the average value for each column of the subsetting dataframe.
- In the full dataframe, replace the missing values by the average values that you just calculated. For instance, if the maximum production capacity is missing for one plant, it should be replaced by the average production capacity of all other plants. The same should be done for shipping costs (to each country individually). *Hint: The `coalesce`-function provides useful functionality.*

For further processing, the data should now be aggregated by region.

### Exercise 2 - Aggregate data

- Group the dataframe by region.
- For each region, sum up the maximum production capacity and calculate the mean shipping costs to each region. Store the aggregated values in a new dataframe.
- Add a new column to the dataframe that contains the mean shipping cost from that region to all other regions divided by the maximum production of that region. Name the new column "Export\_Production\_Ratio". *Hint: Use the `transform` function.*

Now we want to use the new data in the model from last week.

### Exercise 3 - Modeling

- Use the aggregated dataframe and the data given in the second dataset (markets) to model the European market using the model that you developed in your homework (Task 2 of Homework 3). It is not necessary to change the model - reshape the data provided to match the input format that you chose to use last week. We will also upload a solution for the model tomorrow (Tuesday).