**2055 Research & Policy Seminar: Data Science and Machine Learning**

1. **Data Processing and Visualization Module**

*Due: 24 October 2025*

## Overview

In this assignment you will build production‑ready R utilities for cleaning, merging, and analyzing economic datasets. You will also implement a reusable UNCTAD API downloader for an indicator other than FDI or GDP. Notes! What to do

## Learning Goals

* Design clear, reusable R functions with arguments and return values.
* Apply robust methods to handle missing data and outliers.
* Aggregate and combine datasets consistently across countries and years.
* Use parallel computing in R with the foreach and doParallel packages.
* Compute and visualize results based on configurable settings.
* Query the UNCTAD API for a chosen indicator and document your workflow.

## Exercises

1. Write your own R functions for the data‑processing:

* fillMissing( ...): fill gaps using methods such as NOCB, linear interpolation, or group medians.
* fillOutlier(...): detect outliers (e.g., z‑score, Hampel, or IQR, Median absolute deviation) and handle them
* aggregateCountry( ...): aggregate by Country with clear handling of missing values.
* joinData(...): combine datasets with inner/left/right/full joins and explicit key management.

1. Use parallel programming to speed up your R-codes:

* R packages: foreach, doParallel.
* Parallelize alle for loops in Data Processing and Visualization Module
* Measure and briefly report the speed‑up versus a single‑core computing

1. Calculate and visualize analyses according to configurable settings

* Propose one supplementary metric or transformation relevant to your dataset
* Generate at least two charts that reflect those settings.
* Compute it and show at least one visualization.

1. Implement a UNCTAD API function for another economic indicator (not FDI or GDP):

* Model your solution after getUnctadFdiData.R and getUnctadGdpData.R, but target a different economic indicator.
* Allow filtering by years and economies; return a clean data.table.
* Document the indicator code used and include a short usage example.

## Submission

Upload your R-Codes and Data

<https://canvas.wu.ac.at/courses/15283/files>