ENCLOSE: MIDAS imputation methods

Workflow: Harmonization, Imputation, and Analysis

1. Data Preparation & Latent Trait Estimation

Input files:

- `EES10.sav`
- `ZA7649_v2-1-0.sav`

Code: `ENCLOSE_data_harmonization_clean.Rmd`
Output: `ENCLOSE_harmonized_data.RData`

Description:
Harmonizes datasets, aligns variable coding, estimates latent traits.

---

2. Common Variable Distribution Comparison

Input: `ENCLOSE_harmonized_data.RData`
Code: `ENCLOSE_common_variable_comparison_distances_clean.Rmd`
Output: `covariate_distance.xlsx`

Description:
Computes distances between distributions of common variables across datasets.

---

3. Effect of Common Variables on Target Variables

Input: `ENCLOSE_harmonized_data.RData`
Code: `ENCLOSE_regression_on_common_variables_clean.Rmd`
Output: Regression plots and graphs.

---

4. Matching (MIDAS imputation)

Input: `ENCLOSE_harmonized_data.RData`
Code: `Midas_imputations.Rmd`
Output: `Rmidas_ENV.RData`

Description:
Performs imputations using `Rmidas` for deep learning approaches.

---

5. Imputation Result Analysis – Regression & Correlation

Input: `Rmidas_ENV.RData`
Code: `Midas_imputations.Rmd`
Output:
- Density plots (original vs imputed data)
- (regression and correlation coefficients)

---

6. Imputation Result Analysis – Distance Metrics

Input: `Rmidas_ENV.RData`
Code: `Midas_distances.Rmd`
Output:
- Plots of Hellinger distance and Overlap index
- (metrics by method and country)

---

7. Summary Workflow Diagram