Summary of Joining Exercise

Ronnie Bailey-Steinitz

2025-10-16

Summary of the Join Workflow

Each join in this workflow adds new context to the forest area data, helping us move from raw yearly numbers to interpretable ecological relationships.

1. **Join 1: `forest_long` + `climate` (by country and year)**
   This join merges forest area and annual temperature data for the same country–year combinations.
   It allows us to explore whether forest cover changes correspond to changes in climate variables through time.
   → *Used in:* line plots showing forest cover over time, and scatter plots of forest cover vs. temperature.

2. **Join 2: Add `area` (by country only)**
   The country area dataset doesn't vary by year, but it provides the denominator needed to calculate **percent forest cover**.
   By dividing each year's forest area by total land area, we can standardize across countries of different sizes.
   → *Used in:* all plots showing "% forested area," making results comparable across countries.

3. **Join 3: Add `cont` (country–continent lookup)**
   The continent dataset provides a categorical grouping variable that enables regional comparisons.
   By joining continent information to the main table, we can facet or color plots by continent to reveal geographic patterns.
   → *Used in:* faceted plots of forest cover through time (by continent) and scatter plots colored by continent.

Together, these joins produce the **`joined_cont`** dataset — a single, integrated table containing forest area, country size, temperature, and continent information.
This dataset forms the basis for the final visualizations: - *Forest Cover Through Time* (faceted by continent) - *Forest Cover vs. Mean Temperature* (colored by continent)

These plots highlight both temporal and spatial dimensions of global forest dynamics.