

ECOWHEATALY Database Generation Report

Arianna Di Poola
Gianfranco Giulioni

April 4, 2025

Objective

The script builds the **ECOWHEATALY** database, a structured JSON file containing detailed farm-level information on wheat production in Italy, based on data from the RICA (Italian FADN) dataset. The goal is to consolidate information about crop production, fertilization, pesticide usage, machinery and labor use, and economic indicators for durum and common wheat producers.

Input Data

The script loads multiple CSV files from the `Stats/RICA_DATA` directory:

- `aziende_grano.csv`: General farm-level characteristics
- `colture_grano.csv`: Crop production and costs
- `fertilizzanti_grano.csv`: Fertilizer usage
- `fitofarmaci_grano.csv`: Pesticide usage
- `bilancio_grano.csv`: Economic balances
- `certificazioni_grano.csv`: Certifications

Processing Steps

1. **Province Name Standardization:** Harmonizes province names for consistency.
2. **Outlier Removal:** (Optional, not implemented here) Uses adjusted boxplot filtering to clean input variables.
3. **Farm Metadata:** For each farm, general info is stored (region, province, orientation, gender, youth).
4. **Yearly Farm Data:** Includes:
 - Standard output, farm acreage (SAU), machine power
5. **Wheat Crop Data:** For durum and common wheat:
 - Quantity, acreage, wheat price, hours of machine use per hectare
 - Labor, machinery, fertilizer and pesticide costs

6. **Fertilizer Classification:** Types grouped (e.g., mineral, organo-mineral) with NPK content.
7. **Pesticide Classification:** Categorized by type and toxicity; only selected classes retained.

Database Structure

The final JSON `ecowheataly_database.json` is a nested dictionary with the structure:

```
{
  "farm_id": {
    "region": "...",
    "province": "...",
    "technical-economic_orientation": "...",
    "years": {
      "2016": {
        "farm_acreage": ...,
        "standard_gross_output": ...,
        "durum_wheat": {
          "produced_quantity": ...,
          "fertilizers": {
            "Mineral": {
              "nitrogen_ha": ...,
              ...
            }
          },
          "phytosanitary": {
            "Herbicide": {
              3: { "distributed_quantity_ha": ..., ... }
            }
          }
        }
      }
    }
  }
}
```

Output

The JSON file is saved to:

`Stats/ecowheataly_database.json`

It can be used for statistical analysis, sustainability modelling, or visualization.

Authors

Arianna Di Poola
Gianfranco Giulioni