**Overview**

This Jupyter Notebook processes and enriches data from the SIRENE 3.11 database to create a comprehensive dataset of active food-related establishments in France. The script integrates additional data sources, such as Alim Confiance, Agence Bio, and BODACC, to provide a detailed view of food outlets, including their operational status.

# **Features**

1**. Data Import and Filtering**:

- Download the latest SIRENE 3.11 dataset.

- Filter establishments based on specific APE codes related to food commerce.

- Retain only active establishments.

2. **Integration with Alim Confiance**:

- Download and process the Alim Confiance dataset.

- Merge it with the filtered SIRENE data to enrich the dataset.

3. **Integration with Agence Bio**:

- Download and process the Agence Bio dataset.

- Identify and integrate food distributors into the dataset.

4. **BODACC Data Integration**:

- Download BODACC data to identify establishments that have been officially closed.

- Update the dataset with closure information.

5. **Final Dataset Creation**:

- Combine all sources into a single enriched dataset.

- Output the final dataset as a CSV file.

# **Requirements**

The following Python libraries are required to run the notebook:

- pandas

- numpy

- plotly

- polars

- requests

- datetime

- pathlib

- zipfile

# **Usage**

1. Clone the repository or download the notebook file.

2. Install the required Python libraries using ***pip:***

pip install pandas numpy plotly polars requests

# **Output**

The final enriched dataset is saved as a CSV file in the following format:

***Fichier\_final\_<current\_date>.csv***

# **Notes**

Ensure you have an active internet connection to download the datasets.

Modify the file paths in the notebook if necessary to match your local directory structure.

Person Downloading the data: Creurer Quentin/INRAE

First Downloading: 10/09/2024

Last Downloading: 04/04/2025

# **License**

Original license of the downloaded data is an Open License.

See the ORIGINAL LICENSE URL for details: <https://www.etalab.gouv.fr/licence-ouverte-open-licence/>

This project is licensed under the Creative Commons License