**ReadMe**

The dataset presented here draws on several French institutional database sources. The aim is to develop a nationwide database of food outlet locations that can be automatically enriched and updated with a high level of reliability. To this end, we created a flexible and dynamic Python script that can be widely utilized by a maximum of users.

Data can be updated either on a daily basis to create a series of files within a limited timeframe, or on a monthly basis to reflect updates to the SIRENE 3.11 database of the national statistical institute INSEE. Starting from SIRENE 3.11, we enriched the data using the *Alim’ Confiance* (ministry of agriculture) and the *Les Professionnels Engagés en Bio* (organic farmers) databases.

We were also able to enhance the accuracy of food environment description by identifying food outlets currently closed, exploiting the BODACC (official trade bulletin) database with its almost daily updating. These data provide valuable support for studies on the food environment at both macro and micro levels.

The database’s wide range of scales makes it possible to map food sales/distribution outlets, a prerequisite to any investigation of food environments in a given territory, whatever the scale.

We used the Polars library, which accelerates the automation of import, filtering, and cleaning tasks. A pre-filtering mechanism reduces the amount of data processed, further optimizing performance by processing around 500 000 entities instead of 40 million.