

Installation

You should install giza-datasets in a [virtual environment](#) . If you're unfamiliar with Python virtual environments, take a look at this [guide](#) . A virtual environment makes it easier to manage different projects, and avoid compatibility issues between dependencies.

Start by creating a virtual environment in your project directory:

...

Copy python-mvenv.env

...

Activate the virtual environment. On Linux and MacOS:

...

Copy source.env/bin/activate

...

Activate Virtual environment on Windows:

...

Copy .env/Scripts/activate

...

Now you're ready to install giza-datasets with the following command:

...

Copy pipinstallgiza-datasets

...

Install from source

Install giza-datasets from source with the following command:

...

Copy pipinstallgit+https://github.com/gizatechxyz/datasets

...

This command installs the bleeding edge main version rather than the latest stable version. The main version is useful for staying up-to-date with the latest developments. For instance, if a bug has been fixed since the last official release but a new release hasn't been rolled out yet. However, this means the main version may not always be stable. We strive to keep the main version operational, and most issues are usually resolved within a few hours or a day. If you run into a problem, please open an [issue](#) so we can fix it even sooner!

Editable install

You will need an editable install if you'd like to:

- Use the main
- version of the source code.
- Contribute to giza-datasets
- and need to test changes in the code.
-

Clone the repository and install giza-datasets with the following commands:

...

Copy git clone https://github.com/gizatechxyz/datasets.git cd datasets pip install -e .

...

These commands will link the folder you cloned the repository to and your Python library paths. Python will now look inside the folder you cloned to in addition to the normal library paths. For example, if your Python packages are typically installed in `~/anaconda3/envs/main/lib/python3.7/site-packages/`, Python will also search the folder you cloned to: `~/datasets/`.

You must keep the `datasets` folder if you want to keep using the library.

Now you can easily update your clone to the latest version of `giza-datasets` with the following command:

```
...
```

```
Copy cd ~/datasets/ git pull
```

```
...
```

Your Python environment will find the main version of `giza-datasets` on the next run.

[Previous Giza Datasets](#) [Next Quickstart](#)

Last updated 1 month ago