

Homework Assignment #6 Setup

(Flink)

Welcome to Assignment #6! In this assignment, you will use PyFlink to perform real-time data processing and analysis on ZotMusic's datasets on your local desktop.

Template Jupyter Notebook and Datasets

Download the [hw6_template.ipynb](#) and [zot-music-dataset-hw6.zip](#).

Runtime Dependencies

Java 8 or 11 and [Pyflink](#) are the required packages for this assignment. To install the dependency, you can execute the following command:

```
pip install apache-flink
```

If you prefer to use services like Colab, you can install the above dependencies and get started.

If you want to setup the local python development environment, you can refer to the following steps:

Setup Local Development Environment

1. Install Java 8 or 11, and Python 3.10 or newer on your local machine. Follow the appropriate guide for your operating system.
2. pip should be included in your Python installation. Verify by running:

```
pip --version
```

3. (Optional but recommended) A [virtual environment](#) is recommended to manage dependencies for this assignment. To create and activate such virtual environment, please follow the instructions in this document:
<https://docs.python.org/3/library/venv.html>
4. Install required dependencies using pip by executing the following command:

```
pip install apache-flink
```

[Pyflink](#) and `json` are the required packages for this assignment. If you decide to use the python virtual environment, please run the command with the virtual environment activated. You can install other packages if needed while working on the project.

5. Launch the Jupyter and open the given HW6_template.ipynb. Please [install Jupyter](#) first if you don't have it locally. By running the command:

```
jupyter lab
```

The jupyter lab should be running at `http://localhost:8888` . On the jupyter lab, open the `hw6_template.ipynb`.

If you are using the virtual environment, run the following command to let the jupyter lab recognize your environment:

```
pip install ipykernel
```

```
python -m ipykernel install --user --name={your_venv_name} --display-name  
"cs224p_hw6"
```