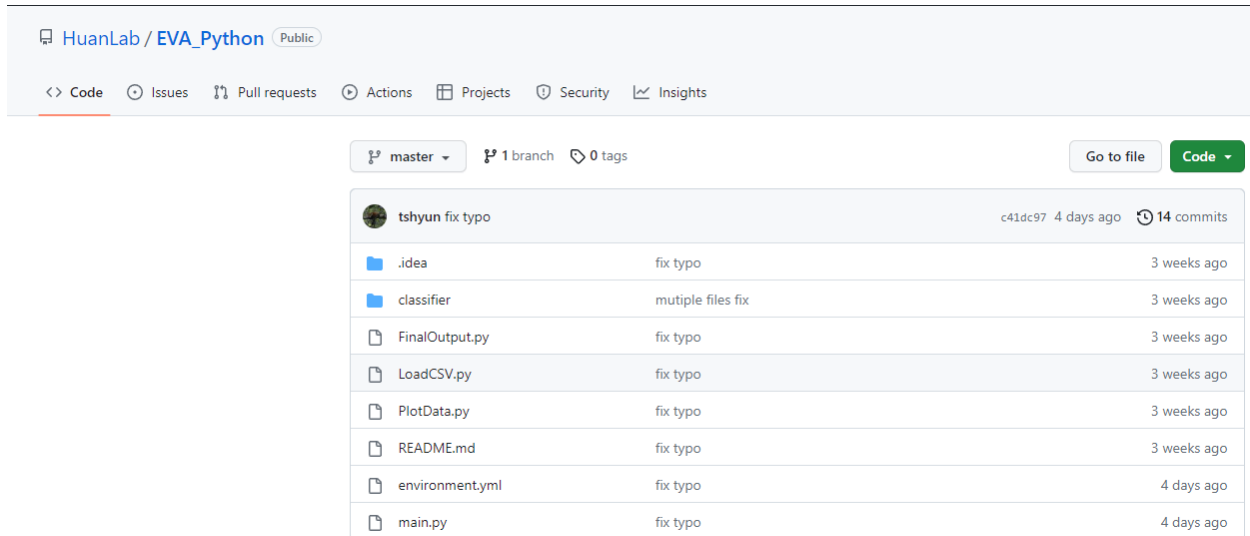# pyEVA Startup Guide

16/5/2023

Ethan Wong-Ma

University of British Columbia

1. Installation
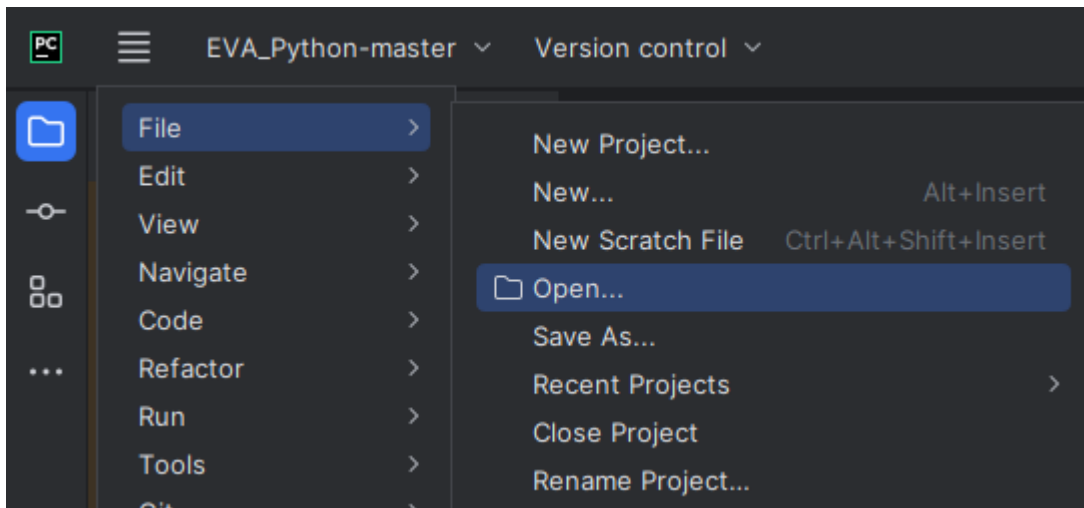
Download 'pyEVA' from the Huan Lab's Github repository.

2. Setup

2.1: Initializing the Environment

Install Anaconda and PyCharm. Then, open the Anaconda Navigator, and through it, launch PyCharm and open pyEVA as a project.

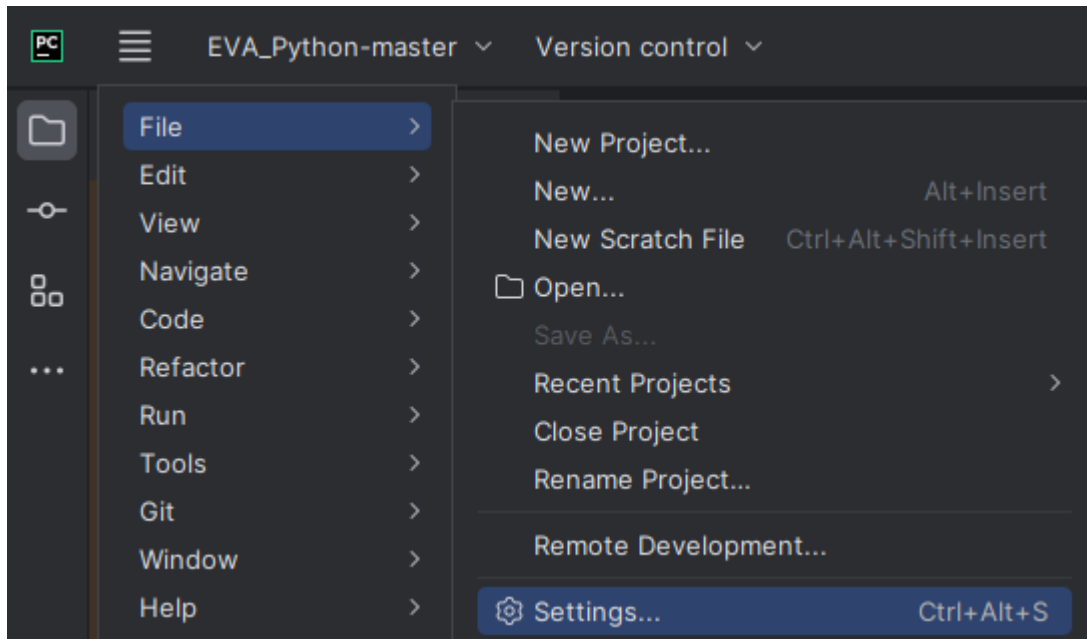Open the PyCharm terminal, and type the following:

```
conda env create --file environment.yml
```

When prompted to, type the following (or whatever it prompts you to type to activate it):
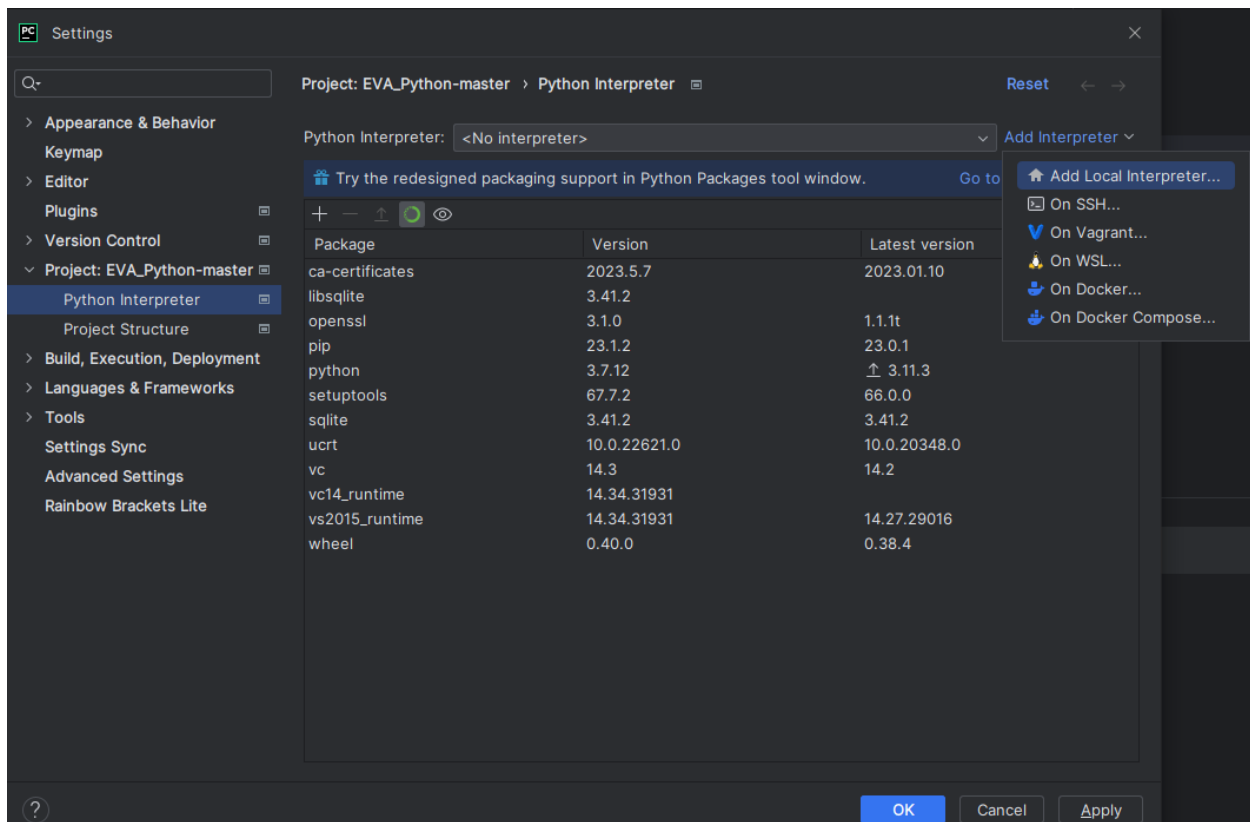
```
conda activate EVA_python
```

2.2 Setting up the Interpreter

Under the menu, select File, then Settings (or press CTRL + ALT + S on Windows).

In the Settings pop-out, navigate to the pyEVA project, then Python Interpreter. Click on 'Add Interpreter', then 'Add Local Interpreter'. Your Python Interpreter at this point will be blank, as it has not been set yet.

Click on 'Conda Environment', then 'Use Existing Environment'. Follow up by selecting on pyEVA as the chosen environment for the interpreter. Click OK.
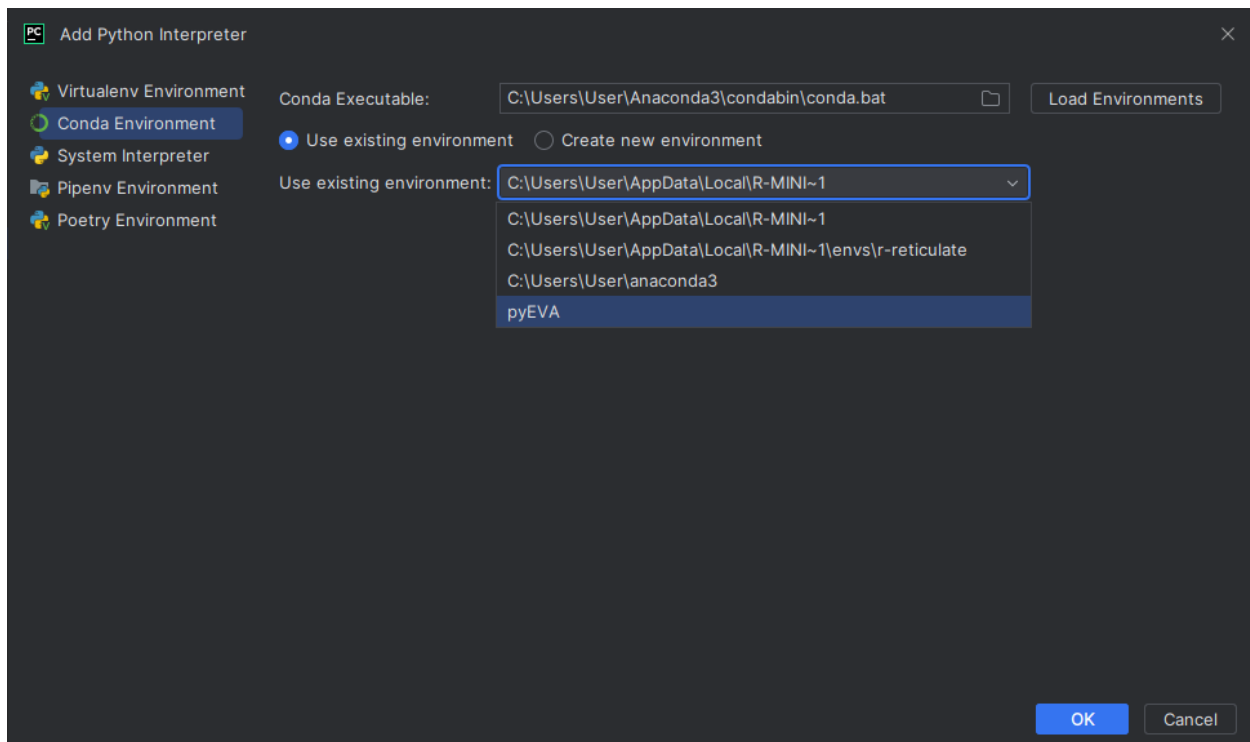

2.3 Setting up Packages

Install the newest version of R, and R Studio. In R Studio, install necessary packages by typing the following code, and then running the scripts:

```
if (!requireNamespace("BiocManager", quietly = TRUE))
install.packages("BiocManager")
BiocManager::install("xcms")
install.packages("ggplot2")
install.packages("scales")
install.packages("dplyr")
```
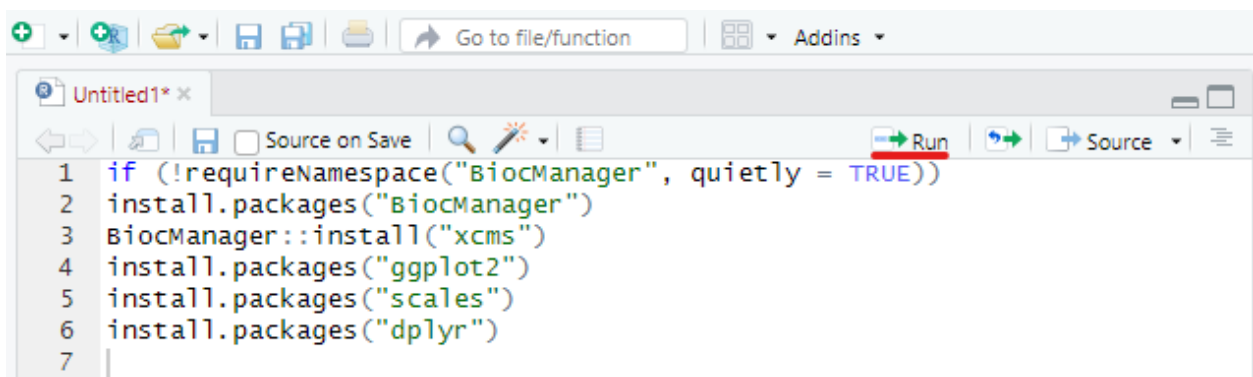

To ensure that all the necessary packages are installed, follow up by typing the following into the PyCharm terminal:
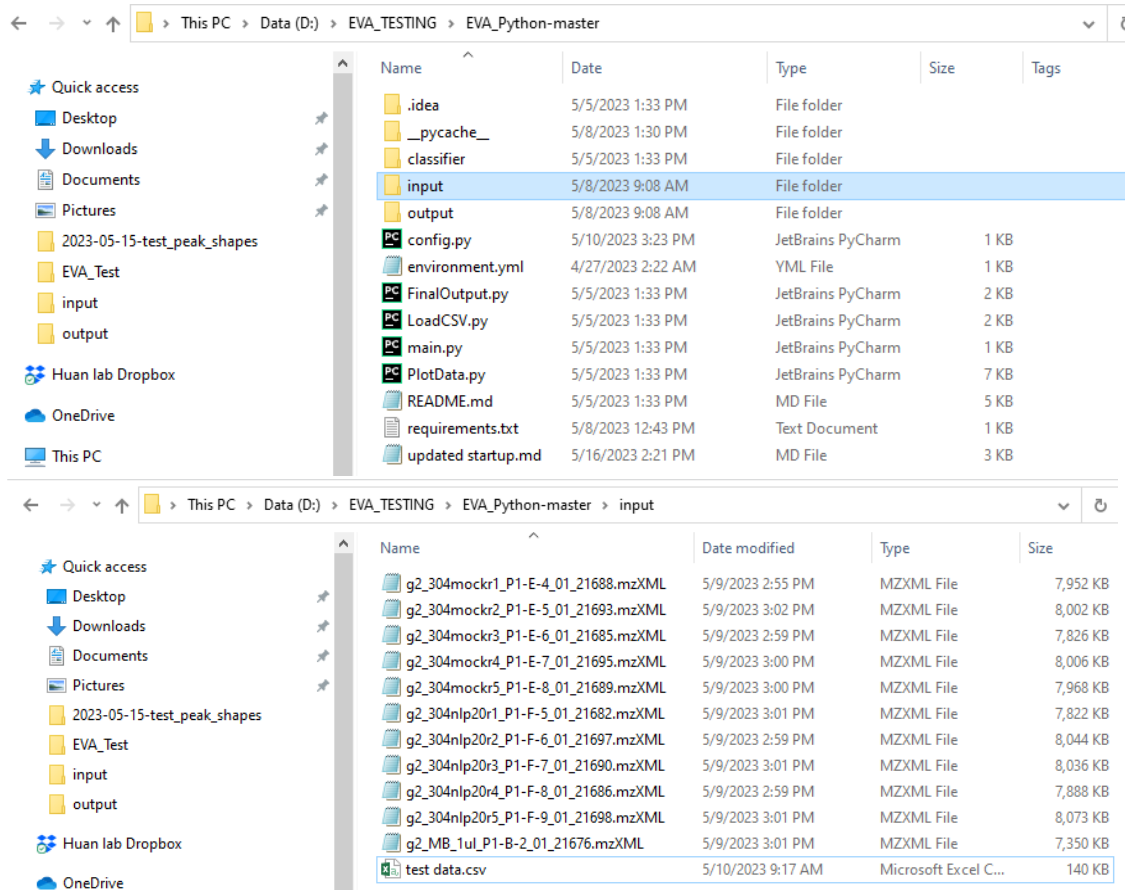
```
pip install -r requirements.txt
```
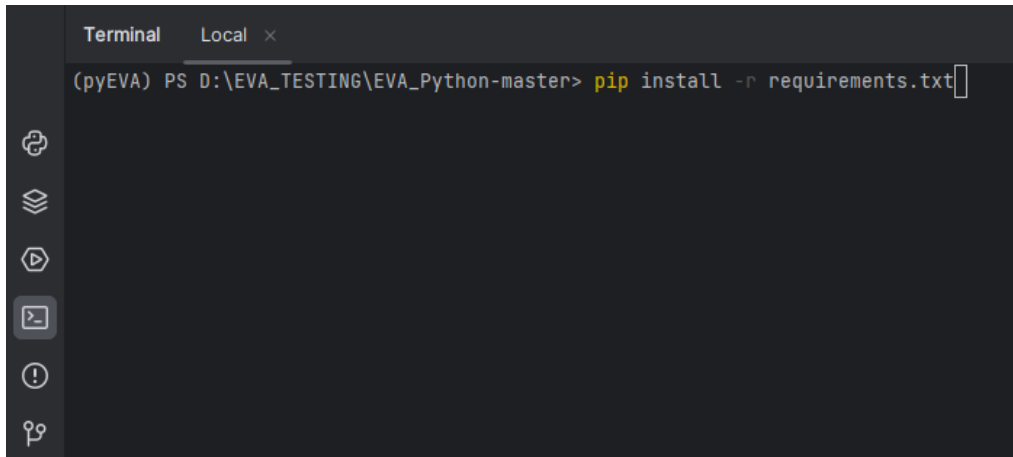

3. Analysis of Data

3.1 Navigating to your data

In Windows Explorer or Finder, open the pyEVA folder. In the folder named 'input', place your .mzXML or .mzML files, along with the single corresponding .csv file.
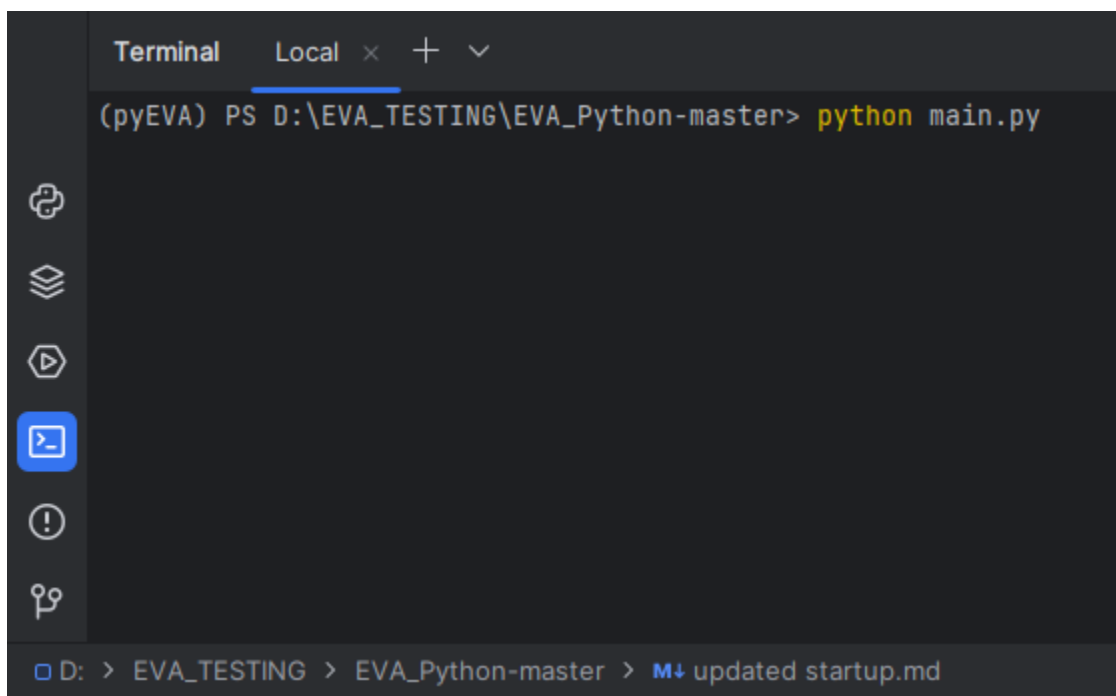
Note that the .csv file that is uploaded must only contain columns in the following order: m/z, retention time (in seconds), and then the intensity of the features of each metabolite in the sample(s). Following is an example of a snippet of a .csv formatted properly. A single sample can be run, or multiple if desired.

	A	B	C	D	E
1	Average N	rt sec	g2_304mo	g2_304mo	g2_304mo
2	67.01974	445.56	848	450	692
3	68.99612	163.98	674	462	1392
4	71.01375	504.54	1052	499	666
5	71.01388	676.92	3663	4590	7078
6	71.01398	290.16	1068	731	836
7	71.01401	520.2	9015	9423	8309
8	71.01414	653.46	237178	96094	159528
9	71.01415	448.62	1856	1422	1942

3.2 Running the data
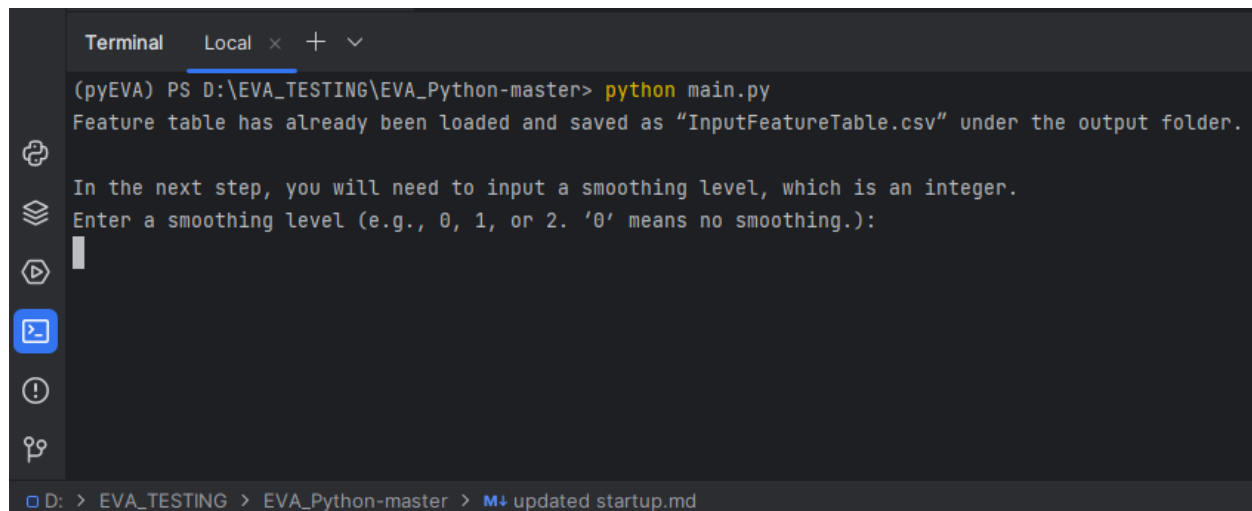
In the PyCharm terminal, type the following:

```
python main.py
```

```
Terminal    Local  ×  +  ∨

(pyEVA) PS D:\EVA_TESTING\EVA_Python-master> python main.py
```

□ D: > EVA_TESTING > EVA_Python-master > M↓ updated startup.md

When prompted, enter the desired level of smoothing for the program (0, 1, 2). It is suggested to choose the same level of smoothing as your sample's spectral rate.

When it is finished, the program will display the following message. This indicates the data has been processed and that the results have been produced in the output folder, with the names FinalTable.csv and PredictionOutcomes.csv.


```
Data processing is complete. The EIC evaluation result is in "PredictionOutcome.csv" and the final result is in "FinalTable.csv". Both files are found under the output folder.
Thank you for using pyEVA!
```

Ignore any tensorflow warnings produced.