

Sandbox setup and basic workflow

Adrian Jörg

2022-03-11

Contact

Organisational: Joel

- Requirements
- Date
- Etc.

Technical: Adrian

Telegram > Slack (easiest for code) > Mail

- Sandbox
- Tools
- Code

Requirements

- **GitHub Account to Fork**
- **Uni-VPN connection**
- **Terminal (e.g. Ubuntu subsystem for Windows)**
- **IDE (Visual Studio Code)**
- **Postgres (if working locally)**
- **Conda + Python (if working locally)**

Sandbox login

- **Connect to University VPN**
- *ssh <username>@fdn-sandbox3.inf.unibe.ch*

If you have not received your credentials or you can't access the sandbox message me!

Database setup local

- `sudo -u postgres -i`
- `createdb -T template0 scrc`
- `gunzip <PATH_TO_scrc.gzip> | psql scrc`

[All tables](#)

Database queries: Terminal

- *psql -d src -U readonly*
 - *Password: readonly*
- ***Then you can use normal psql commands and statements***
 - *E.g. \d to view all tables*

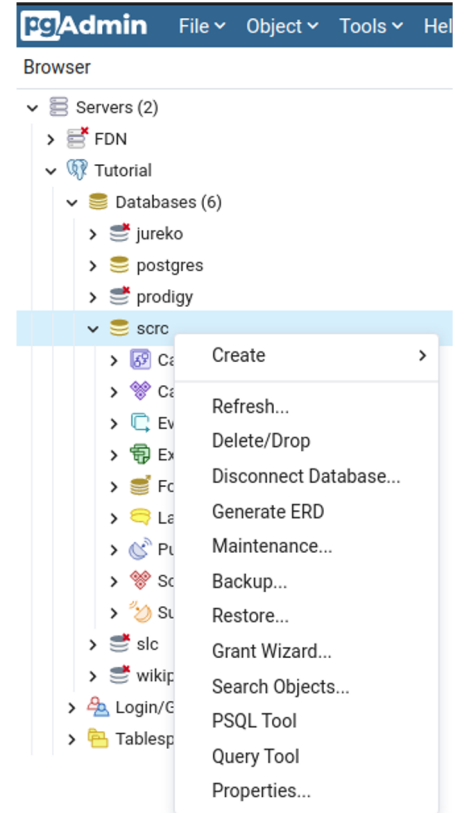
Database queries: pgAdmin

Create new Server

- **Connection**
 - Host: localhost
 - Port: 5432
 - Username/Password: readonly/readonly
- **SSH Tunnel (if not working locally)**
 - Use SSH tunneling: true
 - Tunnel host: fdn-sandbox3.inf.unibe.ch
 - Authentication: According to your preference; password would be the sandbox password

Database queries: pgAdmin

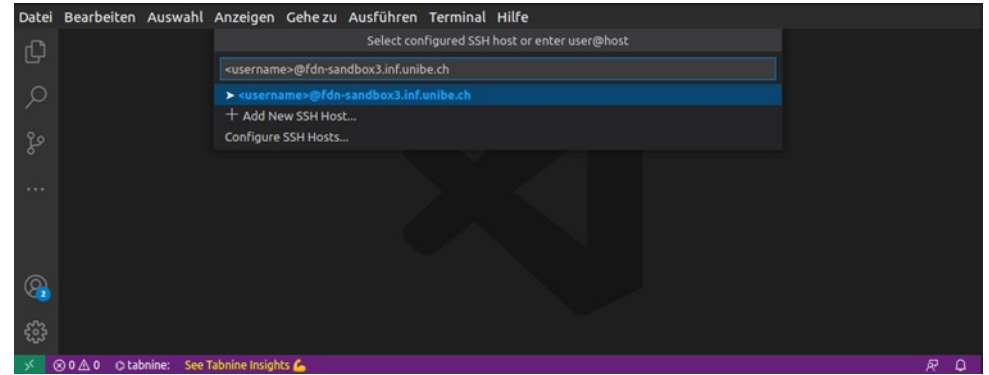
- PSQL tool for terminal use
- Query tool for queries (for most use cases)



Visual Studio Code

If working on the sandbox →

Many helpful things in the
[Contributing.md](#) file



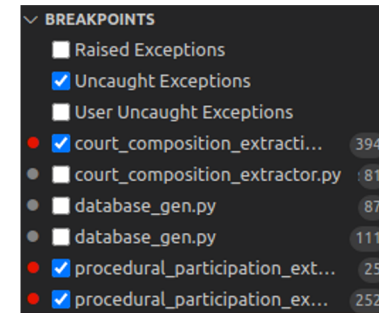
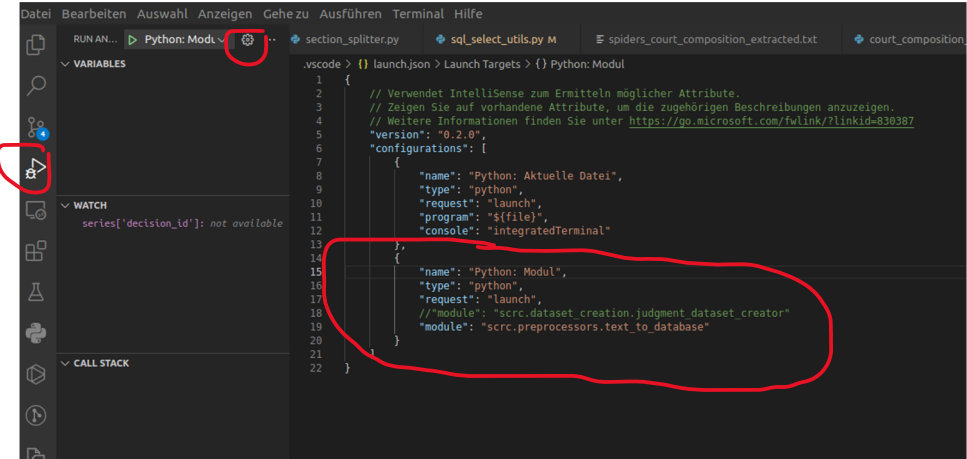
Visual Studio Code Debugging

Run module in debug mode

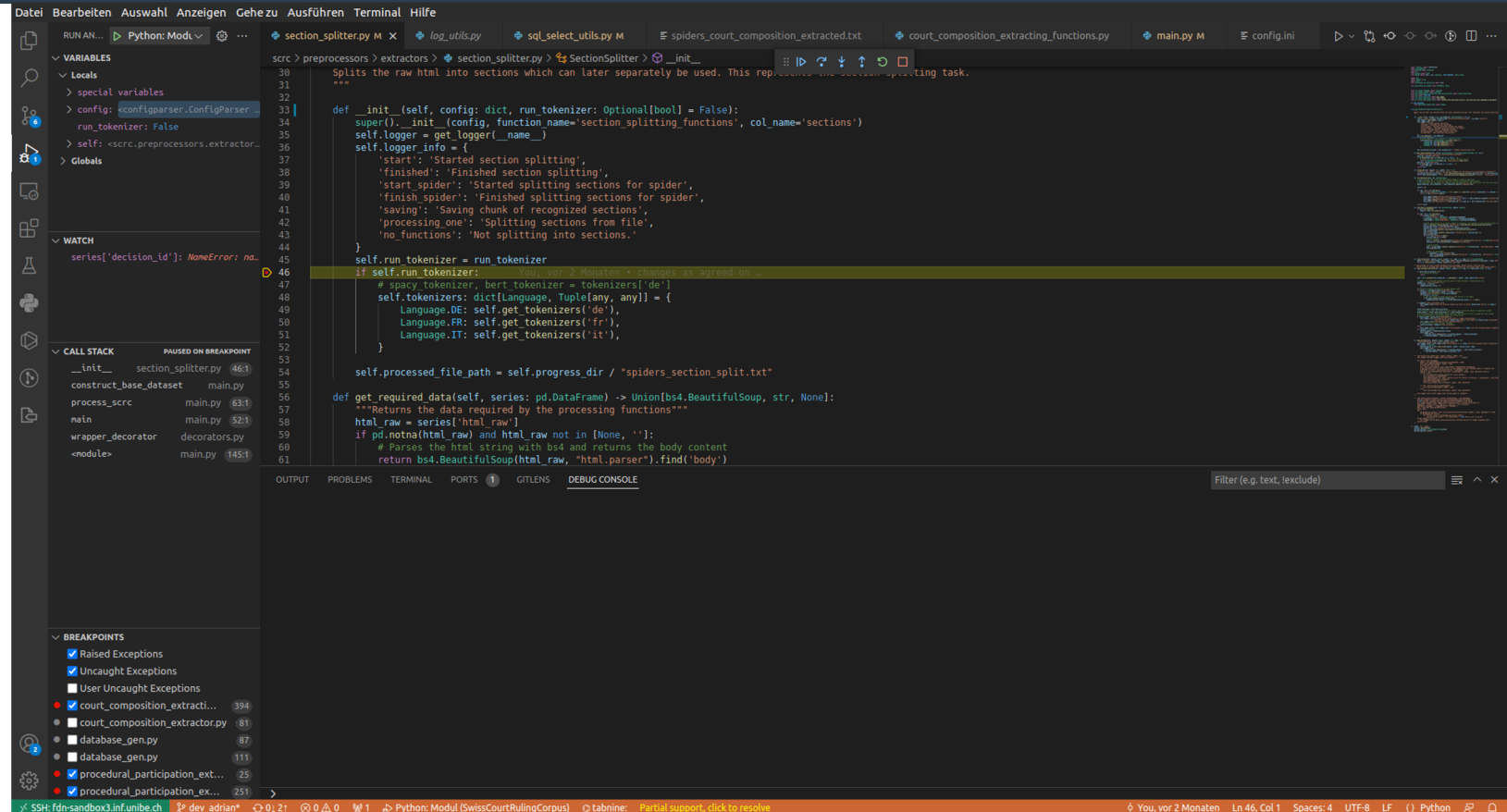
```
{  
    "name": "Python: Modul",  
    "type": "python",  
    "request": "launch",  
    "module": "src.preprocessors.text_to_database"  
}
```

Set breakpoints by click left of the line numbers

Enable/Disable breakpoints



Visual Studio Code Debugging



Tmux

Terminal multiplexing

Can be used to let commands run in the background

- `tmux new -s <username>-<task-running>`
- `tmux a -t <username>-<task-running>`

[Cheatsheet](#)

General Tips

- .Read through Contributing.md**
- .Only run the required step with the required spider**
 - . Do not run src.main if not needed*
- .Get familiar with the debugging tool!**

Errors

Errors can occur in the code we provide.

Let us know about it as soon as possible and we try to fix it, meanwhile check if you can still work.

Communicating errors with error messages, file name and line number really helps.