

Дата-инженер

Настройки Airflow. Файл конфигурации

Ася Гайламазян



Дефолтный конфиг (Default config)

```
sudo docker cp  
airflow_airflow-scheduler_1:/opt/airflow/airflow.cfg .
```

Но изменения в контейнер вноситься не будут

[Документация по параметрам](#)



Таймзоны

```
default_timezone = utc
```

```
import pendulum
```

```
...
```

```
start_date = pendulum.datetime(2023, 1, 1, tz = "UTC")
```



Pools

```
aggregate_db_message_job=BashOperator(  
    task_id='task_1',  
    pool='non_default_pool',  
    bash_command="echo 1",  
    dag=dag,  
)
```



Переменные в Airflow. Задание переменных

Через WebUI

Через CLI командой `airflow variables set`

В коде DAG

```
from airflow.models import Variable  
  
my_var=Variable.set("my_key", "my_value")
```

Считывание кода DAG

`min_file_process_interval`

Не следует писать код вне задач, Airflow будет его запускать каждый раз при парсинге (по умолчанию каждые 30 секунд)



Источники переменных

