DNEXT API – Python Code Example - Download a dataset

Below is a python code snippet that can be used to download a dataset using DNEXT API:

```python
import time
from io import BytesIO
import pandas as pd
import json
import requests

# Login to API and get token
def connect(email, pwd, env):
    '''
    email: email used to connect to your dnext environment
    pwd: password
    env: name of your environment, first part in your dnext URL before the dot
    ENV.dnext.io -> env = 'ENV'
    '''
    url = f'https://api.dnext.io/v1.0/auth/custom-login?org={env}'

    payload = json.dumps({
        'email': f'{email}',
        'password': f'{pwd}',
        'organization': f'{env}'
    })

    headers = {
        'Content-Type': 'application/json'
    }

    response = requests.post(url, headers=headers, data=payload)
    token = response.json()['token']
    return token

email = "email@example.com"      # your email
password = "yourpassword"        # your password
environment = "datahub"          # this remains the same

token = connect(email, password, environment)

def _get_task_status(task_id, token):
    my_headers = {
        'Authorization': 'Bearer ' + token,
        'Content-Type': 'application/json'
    }

    success = False
    wait_time = 2

    while not success and wait_time < 10:
        time.sleep(wait_time)

        status_res = requests.get(f'https://api.dnext.io/v1.0/tasks/{task_id}', headers=my_headers)

        try:
            success = status_res.json()['status'] == 'SUCCEEDED'
        except:
            wait_time += 1
            continue

        if not success:
            wait_time += 1

    return success

def download_dataset(code, token):
    my_headers = {
        'Authorization': 'Bearer ' + token,
        'Content-Type': 'application/json'
    }
```

```
    dl_url = f'https://api.dnext.io/v1.0/data/datasets/{code}/download'
    res = requests.post(dl_url, headers=my_headers)

    task_id = res.json()['task']['id']
    data_ready = _get_task_status(task_id, token)

    if data_ready:
        data_url = res.json()['result']['url']
        data_resp = requests.get(data_url)
        df = pd.read_csv(BytesIO(data_resp.content))
        return df
    else:
        return None


dataset_code = "datahub-7623bd11-6c99-11f0-8b51-02abcd39fb87" #dataset code for the brazilian lineups
dataset

df = download_dataset(dataset_code, token)
```
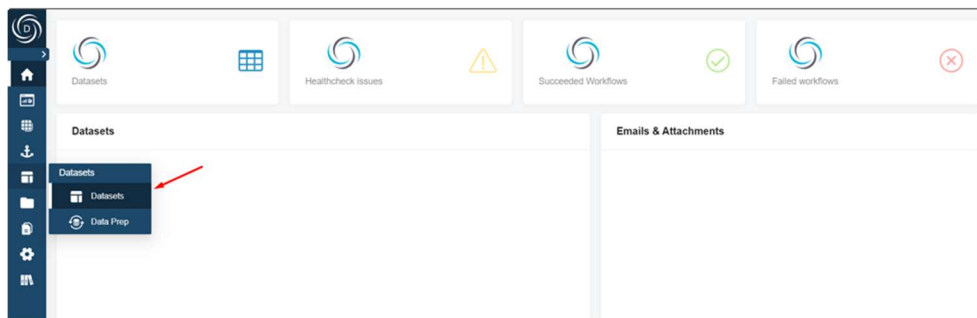
⚠ **Note about the dataset code**:

For now, the dataset code should remain unchanged as you currently have access to only one dataset.

If access to additional datasets is granted in the future, you can locate the dataset code by:

1. Logging into the platform

2. Hovering over the vertical navigation bar on the left

3. Navigating to **Datasets → Datasets**

4. Locating the dataset you want and checking the **"Code"** column