Test assignment for a Data Analyst vacancy

Initial data

They are a .csv file with 4 columns:

- time_start_local — start time interval
- time_end_local — end of time interval
- time_offset — time zone offset in seconds relative to UTC
- steps — the number of steps the user took during this time interval

The file contains about a month's data of one Welltory user (we received his direct permission to publish the data). We believe that they will be sufficient for an acceptable solution to the problem.

General task

Come up with an algorithm that would find this type of walk in the data. The concept of "walk" is not formally defined, but we imagine that this is approximately constant walking for at least 5 minutes. The algorithm must be implemented in python (version 3.10 or earlier).

Expected Result

Archive named <your telegram nickname or last name>.zip with the following 4 files:

- requirements.txt — enumeration of dependencies
- walkings.py — file with find_walkings(data: pandas.DataFrame) -> dict function that returns a dictionary of the form

```
{
    “<day dddd-mm-dd>”: [  # list of walkings for this day
        {
            “start”: “<the beginning of the walk in the format dddd-mm-dd HH:MM:SS>”,
            “end”: “<the end of the walk in the format dddd-mm-dd HH:MM:SS>”,
            “steps”: <number of steps while walking int>
        },
        …
    ]
}
```

- walkings.json — file with the results of running the find_walkings function on the proposed data
- walkings.ipynb — jupyter file c EDA and a description of how you arrived at the decision

What will we evaluate

- The formal side of the solution: the find_walkings function should give an adequate answer on the proposed data
- Completeness of data research
- The quality of your reasoning

Deadlines
It's an open problem in the sense that it can take any amount of time, depending on how deep you go into it. But we believe that for a person with extensive experience (commercial or educational) using python + pandas, a fairly good solution to this task can take 2-4 hours. Well, or we want to find a person who can find acceptable solutions to such problems in a similar time.

We understand that you may be working, away, dealing with personal matters and not having enough free time to take the test right now. It will be great if you send a solution within a few days after receiving the test task. Please let me know if you need more time.