**Assignment for Python Developer**

You are given a dataset which consists of entries with latitude, longitude, and name (link to the dataset file is given below along with some helpful materials). Your task is to write a Python program that will identify entries which are within 200 meters of each other and have similar names i.e. strings that are similar, but not necessarily same

For example:

Bangalore and Bangaloore

new delhi and NewDelhi

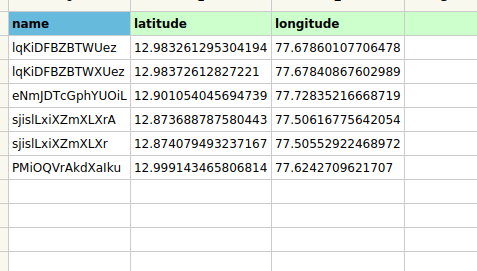
Similarity Criteria:

1. Similar points should be within 200 meters distance from each other
2. Maximum number of single-character edits (insertions, deletions or substitutions) required to change one word into the other should be less than 5

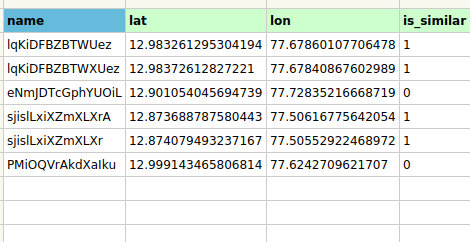
Here are the requirements for the solution:

1. The program should be written in Python and use appropriate data structures to store the dataset.
2. Output should be a csv file with all the entries which satisfy given criteria of similarity marked as True / 1 in a separate column named is\_similar
3. Submission files both python program and output csv file should be uploaded to a public github repository

Here's a sample of the input dataset:



And here's a sample of the expected output:



* Please submit your solution as a .py file, along with any relevant comments or explanations. The code should be well-documented and easy to understand. Good luck!
* Please submit your solutions within 5 days of receiving the assignment.

Link to Dataset

<https://drive.google.com/file/d/1fcfmdshOYn0A_X6D7hW49ioSL8VhYQby/view?usp=sharing>

Additional Helpful Material

[Levenshtein\_distance](https://en.wikipedia.org/wiki/Levenshtein_distance#:~:text=Informally%2C%20the%20Levenshtein%20distance%20between,considered%20this%20distance%20in%201965.)

[geopy library](https://pypi.org/project/geopy/)