# Stalker Challenge

Use the dataset here: https://snap.stanford.edu/data/loc-gowalla.html

- Assume a “stalker” is someone who, in this dataset, visits some of the same locations as another person, after the other person goes to that location.
- A “stalker score” for a pair of people, A & B, is the number of locations for which A has visited a location followed by B visiting that same location in the future.
- Any given location should be counted once in the score, so a stalker score can never be higher than the number of unique locations that A and B have in common.

Use the datasets from the web page above to answer the following questions:

1. Which friend pair has the highest “stalker score”?
2. Which non-friend pair has the highest “stalker score”?

A solution must be submitted using Python. We give points for “pure” and elegant solutions, so please do not use external libraries such as Dask and stick to the core library package with Python. For extra points you can submit a solution in the SQ. Please feel free to email at any time for any clarifications.

Hint: a good solution in a pure Python script (no external libraries e.g. Pandas) should finish in under 2 minutes on a single core 2GB machine. Using Pandas with a group by or iteration will be too expensive in terms of time and space.

Please give the winning user id pairs and “stalker score” for each question, and please *explain your solution methods, including any source code if you wrote any.*