University of California
Department of Civil Engineering

Pozdnukhov, Instructor
Fall 2016

CE 263 Problem #5

New York, New York

You are given a set of time-stamped locations of 30 citizens of New York in a period of 3 months. Location history for each person is contained in a separate file named **Name.csv**. File name is a person's pseudonym. The data format for each record is comma-separated fields containing:

Time (as a text string),
Latitude,
Longitude,
Text String.

You must use a Gaussian Mixture Model (you can use the scikit-learn or other implementations) to solve the problems below. For each person, determine the following:

Q1 (30 points). If the person is a commuter (for example, a student or a working professional with regular return trips from home to work/school).

Q2 (30 points). Coodinates of Home and Work/School (if work/schools exists).

Q3 (30 points). The most likely time to find the person at Home and at Work/School (if work/schools exists).

Your submission has to contain:

1. A report in (HTML/PDF/Word) on the approach you have taken, with any plots summarizing your results, and a table with answers to the questions **Q1, Q2, Q3**.

2. Your code.

Up to (10 points) will be awarded for the clarity of your report and the correct use of a GMM.

Good Luck!