AoCMM 2017

Keep in mind that you are not expected to complete both problems. Try your best on one of them before moving on to the other, since prizes will be given to not only the teams that score the highest total of the two problems, but also the teams that place among the highest for one of the two problems.

1. It has been shown typing patterns can be used to identify a person. To confirm this idea, we collected typing patterns from eleven of our officers using two different typing methods: fourteen short English quotes and six paragraphs composed of random characters. The quotes and paragraphs are grouped into three categories: eight quotes and three paragraphs where the officer's identity is known, six quotes where the officer's identity is unknown, and another three paragraphs where the officer's identity is also unknown. Based on the first category, how would one match the second and third categories to the officers?
The data can be downloaded at https://drive.google.com/open?id=0B7h3lxLzcvgCdzVXbF9Tb0pab0k or https://www.dropbox.com/s/o5cqqitjib6niy5/AoCMM-Typing.zip?dl=0, which was generated using https://www.keyhero.com/. To type the random characters, click on typing practice -> typing lessons -> option 5) letters asdfqwerjkl;uiop.
Note that person K typed a different quote for test 1 from the rest of the ten people. The third category was only collected from ten of the eleven officers. You may try generating more typing data if necessary. Any information available on the website can be used towards your analysis.

2. There will always be times when taxis are vacant. Some drivers say that you should head to the city center to find more customers, but is that always true?
1) Suppose you are a taxi driver in NYC, what should you do when your car is vacant? 2) If you are the head of a taxi company, what would you advise your
drivers do?

You may find taxi trip record data for Green Taxi in NYC at https://data.cityofnewyork.us/Transportation/2016-Green-Taxi-Trip-Data/hvrh-b6nb. If you are unable to download the entire file, we have compiled a smaller version of the first 500,000 rows which can be downloaded at
https://drive.google.com/open?id=0B7h3lxLzcvgCOEx6OS1CTVdnRFE or https://www.dropbox.com/s/hdc1xwd0dn450i8/2016_Green_Taxi_Trip_Data.xlsx?dl=0. More information about this file can be found at https://drive.google.com/open?id=0B7h3lxLzcvgCWE5oUmZMclFtNWs or https://www.dropbox.com/s/gp0ogk6j24vgehj/data_dictionary_trip_records_green.pdf?dl=0.