**Deadline**: 11:59 pm, Oct 3 (Monday)

**Q1:** We have household survey data and population density data of Côte d'Ivoire. Merge departmente-level density data from the excel sheet (CIV\_populationdensity.xlsx) into the household data (CIV\_Section\_O.dta) (Note: We did this during the class)

**Q2**: We have the GPS coordinates for 111 households from a particular village. You are a field manager and your job is to assign these households to 19 enumerators (~6 surveys per enumerator per day) in such a way that each enumerator is assigned 6 households that are close to each other. Manually assigning them for each village will take you a lot of time. Your job is to write an algorithm that would auto assign each household (add a column and assign it a value 1-19 which can be used as enumerator ID). (Note: Your code should still work if I run it on data from another village.)

**~~Q3~~**~~: 2010 election data (Tz\_election\_2010\_raw.xlsx) from Tanzania is not usable in its current form. You have to create a dataset in the wide form, where each row is a unique ward and votes received by each party are given in separate columns. You can check the following dta file as a template for your output: Tz\_elec\_template.~~ (Note: This is optional)

**~~Q4~~**~~: Between 2010 and 2015, the number of wards in Tanzania went from 3,333 to 3,944. This happened by dividing existing ward into 2 (or in some cases more) new wards. You have to create a dataset where each row is a 2015 ward corresponding parent ward from 2010. You can use region, district and ward names to match 2010-2015 wards that weren’t divided. In case of wards that were divided (and given new names), you can use the following dataset to see 2010-2015 mapping: Tz\_GIS\_2015\_2010\_intersection.dta~~

Note: We were unable to go over reclink2 in detail so don’t attempt the question above (unless you really want to) You can instead do the following version of Q4:

**Q4.1**: Between 2010 and 2015, the number of wards in Tanzania went from 3,333 to 3,944. This happened by dividing existing ward into 2 (or in some cases more) new wards. Create a final dataset that includes the following types of wards 1) wards that are both in 2010 and 2015 2) wards that are only in 2010 (i.e. “childless ward”) 3) 2015 wards that do not have a corresponding 2010 ward (i.e. “parentless ward”) Generate a categorical variable to describe these 3 types of wards.