Soc 756
Problem Set 1

Use R to program the answers to question 1 and graph the lifetable functions. Submit your program stapled to your answers. Question 2 does not require programming.

 1. Table 1 (next page) contains deaths by age for French males in 1985. These data also include mid-year population estimates and a set of $_na_x$ values for French males for 1985. Table 1 is on our webpage in .csv format.

    a. Use these data to construct a life table for the male population. Do this by performing operations on the vectors. You will need to calculate and fill in the following life table columns: $_nq_x$; l_x; $_nd_x$; $_nL_x$; $_nm_x$; T_x; and e_x.
    b. Graph the following life table functions using either plot() or ggplot(): l_x; $_nd_x$; and $_nm_x$. What do you observe?
    c. What was life expectancy at age 40? How would you interpret this number?
    d. What was the probability of surviving from birth to age 30?
    e. What was the probability of surviving to age 65 for those who survived to age 30?
    f. What was the probability that a newborn would die between 50 and 55?
    g. How many years could a newborn expect to live in the interval 15-65?
    h. If you only had the fourth column of Table 1, would you be able to distinguish this population as one with high mortality or low mortality? (What $_na_x$ value in particular might help distinguish between the two?)
    i. If the French population were stationary, what would be the crude death rate?
    j. **Extra credit part 1:** push your code to your Github page and list the URL in your submitted answers.
    k. **Extra credit part 2:** install the Lifetables package in R. With $_nm_x$ in hand, use lt.mx() to populate the other functions. Check your work in 1(a), noting discrepancies if you set nax=NULL.

2. Think about the social phenomena / processes that most interest you. Might any of these processes be measured in the form of a lifetable?

If yes:
    a. What events would constitute "births" and "deaths"?
    b. What could you learn from using a lifetable?
    c. Where might you start looking for data to identify the size of the population at risk and the age-specific "death" rates or probabilities?
    d. What issues might limit how the information produced in your lifetable can be interpreted?

If no, describe the issues that would make the lifetable an inappropriate analytical tool for the social processes that you study.

Table 1.

x	Mid-year population ($_nN_x$)	Number of deaths ($_nD_x$)	$_na_x$
0	379,985	3,741	0.087
1	1,559,722	770	1.5
5	1,896,295	532	2.5
10	2,160,190	673	2.966
15	2,179,837	2,138	2.769
20	2,159,556	3,432	2.574
25	2,106,750	3,291	2.512
30	2,147,845	3,657	2.586
35	2,165,387	4,956	2.657
40	1,516,952	5,269	2.697
45	1,498,630	8,654	2.695
50	1,552,746	14,490	2.663
55	1,476,770	20,831	2.625
60	1,350,479	26,805	2.601
65	722,430	20,233	2.615
70	842,589	38,315	2.598
75	636,848	46,903	2.538
80	372,059	44,443	2.466
85+	175,169	37,759	4.639