Question 1:



Question 2:

a)



b) The Census glossary states the following regarding the population size of census tracts: “Census tracts generally have a population size between 1,200 and 8,000 people, with an optimum size of 4,000 people”. The grand mean and standard deviations, and the measures of variation and central tendency within each CBSA, seem to fit within this range. Obviously, the min and max point to the existence of outliers beyond the target range, sometimes by multiples. It would be useful to produce boxplots of census tract populations to better understand the extent of the outlier problem.

c) It’s tough to identify patterns of tract size by CBSA size with only five CBSAs to observe. It seems like more densely populated areas (Chicago, San Francisco) may be more likely to have overpopulated tracts, since both have maximum tract sizes that vastly exceed the Census Bureau’s target size. Their means are also higher, although still within the target range. The smallest CBSA, Santa Fe, also has substantially fewer people per tract.

d) I took a mean across CBSAs and weighting each CBSA mean by its tract count. Looks like weighting CBSA means by tract count leads to a higher overall mean. You could also take the mean over all of the tracts. This should yield the same result as weighting CBSA means by tract (I think).

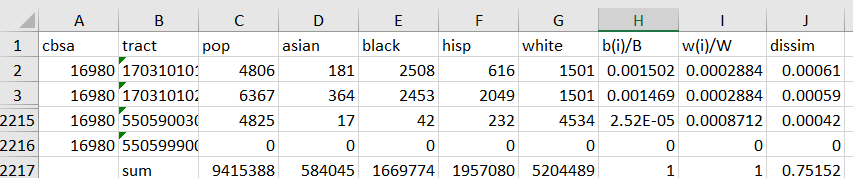
3.

a) I calculated dissimilarity for 16980 (Chicago) and got 0.752.

b) 75.2% of the people in either of these two groups would need to relocate to construct an even distribution of white and black people across all the tracts in this CBSA. Specifically, I think black people in predominantly black tracts would need to move to predominantly white tracts, and vice versa, until nobody else could move.

c) From figure 2 it seems the average dissimilarity in the 2010 Census is 0.59. So, this Chicago CBSA is more segregated than the average CBSA.

d)



4.

a) For the CBSA 16980 I calculated that the exposure of whites to blacks is 0.063. This means the probability that a randomly selected white person shares a tract with a black person equals 6.3%. This seems unbelievably low.

b) If we didn’t include other racial groups in this analysis, the results would overstate the probability that white person lives in the same tract as a black person. The denominator of the term b(i)/t(i) would shrink, making the term larger overall and growing the probability. I can’t think of any obvious situation in which this would have utility. It just doesn’t reflect the empirical reality as well.