Math 152, Fall 2022
Jo Hardin
WU # 12
Thursday 10/20/22

Your name: _____

Names of people you worked with: _____

Task:

1. What is the primary reason we bootstrap? Or said differently, what do we get from bootstrapping (that we can't get elsewhere)?
2. When is it a bad idea to bootstrap?

Solution:

1. The primary (overarching) reason we bootstrap is to understand the sampling distribution of our statistic, $\hat{\theta}$. Different aspects of the distribution contribute to different parts of statistical inference.
   - If the sampling distribution is normal, we can use normal theory (!) to create CIs (and find p-values), and we wouldn't need to bootstrap.
   - If the statistic is biased, we should adjust our methods in particular ways (e.g., see BCa intervals, which we won't cover this semester).
   - If the SE of the statistic is intractable (i.e., there is no formula for the SE), we can estimate the SE using bootstrapping. Estimating the SE is particularly important for arbitrary statistics that don't have corresponding theory which can be used to estimate the SE.
   - The shape of the distribution of $\hat{\theta}^*$ (or $T^*(b)$) provides an estimate of the tails of the sampling distribution to use for creating CIs.
2. We cannot bootstrap if the data are not a good proxy for the population. One reason the data would be a bad proxy is if the dataset is extremely small. Other reasons would be if the data were not independent of one another.