1. We are interested in determining the average height of all males in Europe. You randomly select 100 males from the database of males in Europe. The average height of the 100 males was calculated to be 70 inches, with a standard deviation of 8 inches.

   (a) Determine the population, sample, parameter, statistic, and point estimate from the scenario above. *The population is all males in Europe and the sample is 100 randomly selected males from that population. The parameter is true average height and the statistic is the average height of the sample with a point estimate of 70 inches.*

   (b) Calculate a 95% confidence interval for the average height of European males. *To construct a confidence interval, we first need to find the standard error, which is given by SE=*$\frac{8}{\sqrt{100}} = \frac{8}{10} = .8$*. We also note that the sample size is large enough (*$n > 30$ *since* $n = 100$*) to invoke the Central Limit Theorem. Therefore, we can construct the confidence interval as follows using the following formula:* $\bar{X} \pm z_{1-\alpha/2} \times SE$ *where* $\bar{X} = 70$*,* $\alpha = 100\% - 95\% = 5\%$ *so* $z_{0.975} \approx 1.96 \approx 2$*, and SE as calculated above. Plugging this in to the formula, we have* $70 \pm 2 * (0.8) = 70 \pm 1.6 = (68.4, 71.6)$ *inches as the* 95% *confidence interval.*

   (c) Interpret this confidence interval. *We are 95% confident that the interval from 68.4 inches and 71.6 inches covers the true average height of all males in Europe.*

   (d) Explain why someone may prefer a 99.7% confidence interval to a 95% one. Explain why someone else may find a 95% one more useful. *A 99.7% confidence interval will be wider:* $70 \pm 3(0.8) = 70 \pm 2.4$ *inches* $= (67.6, 72.4)$ *so we will be more confident that it covers the true average height of all males in Europe. The 99.7% CI however, is less precise with a spread of 5. The 95% CI gives a smaller interval of plausible values. This is a common theme in confidence intervals - while a higher percentage gives you more confidence that the interval covers the true mean, the interval can be significantly wider as a result of this higher level of confidence. Indeed, only way to be 100% confident that the interval contains the true value is to have the line go from* $(-\infty, \infty)$*, which isn't very useful!*

2. 80 UW-Madison students are randomly selected and asked if they regularly eat cereal for breakfast. 50 responded yes.

   (a) Calculate a 98% confidence interval for the proportion of students who eat cereal for breakfast. Interpret it in context. *Using the formula we had in part a, let's find the average (50/80=.625 or 62.5%), and standard error (*$\sqrt{.625 * .375}/\sqrt{80} = .054 = 5.4\%$*). Now we need to find the correct z-score for the 99th percentile, which will mean there is 1% remaining in each tail. Looking this up, we have* $P(Z < 2.33) = 0.99$*. Thus, our 98% confidence interval is* $62.5\% \pm (2.33 * 5.4\%) = (49.9\%, 75.1\%)$*. What this means is that we are 98% confident that the interval from 49.9% to 75.1% covers the true percentage of UW-Madison students who regularly eat cereal for breakfast.*

   (b) T/F there is a 98% chance the population proportion is in this confidence interval. *False, the probability of the true population proportion being between 49.9 and 75.1 is 0 or 1. We don't know for sure if it is in there or not.*

(c) T/F if we repeatedly found random samples and calculated 100 95% confidence intervals, we'd expect exactly 95 of them to capture the true population proportion. *False, we'd expect about 95 of the 100 95% confidence intervals to capture the true mean, based on how we built the 95% CI.*