1. The weights of adult golden retrievers are approximately normally distributed about a mean of 71 lbs with a standard deviation of 7 lbs. If Max is at the 96th percentile in weight for adult golden retrievers, then his weight, in lbs, is closest to:

   (a) 81.5
   (b) 83.5
   (c) 85
   (d) 86
   (e) 87

   Use the formula $z-\text{score} = \frac{\text{Observation-mean}}{\text{SE}}$. Notice that $z_{0.96} = 1.75$, so $Observation = z-\text{score}\cdot\text{SE}+\text{mean} = 1.75 \cdot 7 + 71 = 83.25$.

2. Bonita's tech service business is booming and she wants to reward her employees. She selects a random sample of 150 employees from the 2000 people she employs to investigate their preference for how to be rewarded. The results are shown in the table below.

Preference	Number of Employees
Increase number of paid vacation days for all employees	64
Increase salary for all employees	86

   Bonita incorrectly performed a large sample test of the difference in two proportions 64/150 and 86/150 and calculated a p-value of 0.011. Consequently, she concluded that there was a significant difference preference for the two options. Which of the following best describes the error Bonita made?

   (a) No statistical test was necessary because 64 is clearly lower than 86
   (b) The results of the test were invalid because less than 10% of the population was sampled
   (c) Bonita performed a two-tailed test and should have performed a one-tailed test
   (d) A one-sample test for a proportion should have been performed because only one sample was used. Also, there is no significance level identified so that Bonita cannot make a conclusion based on p-value.

   Identify what are the hypotheses. H_0: P(prefer paid vacation)$= \frac{1}{2}$. H_A: P(prefer paid vacation)$\neq \frac{1}{2}$. So, only one sample was used.

3. Miriam's boss told her that the standardized score (z-score) for her salary compared to the salary of employees at the company is 1.60. Which of the following is the best interpretation of this standardized score?

   (a) Miriam's salary is $160,000.
   (b) Miriam's salary is 1.60 above the average salary of employees at the company.
   (c) Miriam's salary is 1.60 times above the average salary of employees at the company.
   (d) Miriam's salary is 1.60 standard deviations above the average salary of employees at the company.
   (e) 16% of people at Miriam's company have a higher salary than she does.

You should understand what the $z - value$ means. Recall the formula $Observation = z - \text{score} \cdot \text{SE} + \text{mean}$.

4. A quality control specialist at a plate glass factory must estimate the mean clarity rating of a new batch of glass sheets being produced using a sample of 18 sheets of glass. The actual distribution of this batch is unknown, but preliminary investigations show that a normal approximation is reasonable. The specialist decides to use a t-distribution rather than a z-distribution because

   (a) The z-distribution is not appropriate because the sample size is too big.

   (b) The sample size is large compared to the population size.

   (c) The data comes from only one batch.

   (d) the standard deviation of the batch is unknown and the sample size is smaller than 30.

   (e) The t-distribution results in a narrower confidence interval.

   In fact, the principle you decide whether to take t-test or z-test is by looking at the sample size. When the sample size is small or unknown, then you should definitely choose t-test. When sample size is large, then you can either choose t-test or z-test because the behavior of these two tests will be similar. However, t-test will require additional *normal* assumption.

5. For each of the following research questions, identify at least one analysis that would be appropriate for the situation.

   (a) Do seniors earn higher semester grade point averages than freshmen? We have two independent groups, seniors and freshmen. The response is numeric (continuous). We are interested in determining whether or not the average of seniors' grades is higher than the freshmen's. Therefore, conduct a one-sided, two-sample t-test or z-test (if we have large enough sample sizes) for testing the difference in the mean GPA for seniors and the mean GPA for freshmen. Notice that if you want to use t.test for the small sample, then you need to assume that the data following a normal distribution from the population.

   If the students are from the different universities and there is a paired relationship, you can even consider the one-sided paired t.test.

   (b) Is there a difference in the percentage of NCAA basketball players who graduate and NCAA football players who graduate? The response is binary (graduate or not), and there are two independent groups being compared (NCAA basketball players and NCAA football players). Therefore, we could conduct a two-sided Z-test for comparing two proportions (we did not cover how to do this, so won't have to perform it). We could also conduct a chi squared test for independence where graduate/not graduate is one variable and basketball/football is the other variable.

   (c) How many hours per week do Stat 301 students study outside of class? We have a continuous response variable (number of hours per week studied) and one group (301 students). Therefore, calculate a one-sample t-interval for the mean or 1 sample z-interval for mean if sample size is large enough (it always will be in this class)

(d) Do Stat 301 students drink, on average, more than one cup of coffee per day during finals week? (During finals week, a sample of 301 students will record how many cups of coffee they drink each day.) We have a quantitative, continuous response variable (number of cups of coffee consumed per day during finals week) and one group (stat 301 students). Therefore, conduct a one-sided, one-sample t-test for testing $H_0 : mean = 1$ against $H_A : mean > 1$. If the sample size is large, a one-sample z-test would give similar results.

(e) Is there a relationship between political affilition (Democrat, Republican, Independent) and income level (Low Income, Middle Class, Wealthy)? We have two categorical variables for which we are interested in determining whether or not a relationship exists. Therefore, conduct a chi-square test of independence. Still, you need to be careful that chi-square test can be used when the sample size is sufficiently large.