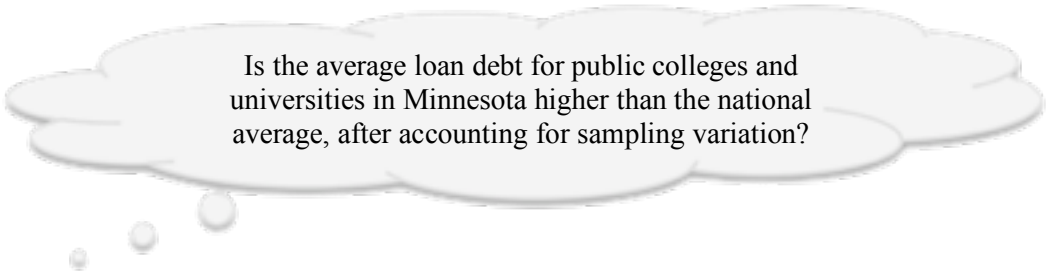
College Debt

The Department of Education recently released its College Scorecard data. One of the variables in this data is the typical amount of loan debt accumulated at the institution by student borrowers. The average loan debt for the 52 public colleges and universities in Minnesota is $11,022. In this assignment, you will use Monte Carlo simulation to answer the following research question:



To answer this question, you will use the population data from the *college-debt.tp* file. This file contains data on 1,732 public colleges and universities in the United States.

Explore the Population

1. Plot the debt variable for the 1,732 public colleges and universities. Add the mean debt and the numerical value of the mean debt to the plot. Add a reference line in the plot at the value of the typical student debt for a student at the University of Minnesota–Twin Cities. Copy and paste the plot into your word-processed document.

2. Describe the shape of the population distribution. Also, compute and report the mean and standard deviation of this distribution.

3. The typical University of Minnesota–Twin Cities student is higher than average. How many standard deviations higher is the typical University of Minnesota–Twin Cities student debt than the population average? Compute this to the nearest tenth.

Model

Although examining the typical debt for students attending a Minnesota school is interesting (and maybe scary?) it does not help us answer the research question. We also cannot just put the Minnesota average of $11,022 into the population plot and see if it is higher or lower. The reason we cannot do that is because, as the saying goes, you need to compare apples to apples. The cases in the population plot are individual institutions, so we can use the population plot to evaluate individual institutions' student debt.

The $11,022 is based on the average from 52 institutions. As such, it needs to be evaluated in a plot where the cases are averages from 52 institutions. To obtain such a plot, we will randomly sample 52 institutions from the population of institutions and compute their average debt. After we carry out many, many trials of this simulation, each time collecting the average debt for the 52 sampled institutions, then we can evaluate Minnesota's mean debt of $11,022 and answer our research question.

4. Set up a TinkerPlots™ sampler that includes the 1,732 debt values from the population (i.e., your mixer will include 1,732 balls). To do this, you may want to re-read through the *Cat Factory* activity (especially the section on *Modeling Cat Names*). To be consistent with the models we have been using, sample cases *with* replacement. Copy and paste the sampler window into your word-processed document.

The statistical *null hypothesis* that mathematically describes the model in your TinkerPlots™ sampler is

H_0: The average student debt is equal to $10,025 (the population average).

Simulate from the Model

Use TinkerPlots™ to generate a random sample of 52 debt values. Plot these randomly generated values and add the mean to the plot. Collect the results (mean) from 500 trials of the simulation. To collect the mean, right-click the **mean triangle** (not the mean value) and choose `Collect Statistic`.

Evaluate the Observed Results

5. Describe the shape of the distribution of the 500 simulated results. Also compute and report the mean and standard deviation of this distribution.

6. Explain, by referring to the null hypothesis, why we could have expected the mean you computed in Question #5.

7. Describe (using the context of the problem) what each case in this distribution represents.

8. Using TinkerPlots™, add a reference line to the plot of the distribution of the 500 simulated results at Minnesota's average debt of $11,022. Also add a divider to the plot. Place the endpoints of the divider at two standard deviations above and below the mean. Copy and paste the plot (with the reference line and divider) into your word-processed document.

9. How compatible is Minnesota students' average debt of $11,022 with the simulated results from hypothesized model? Explain based on the range of likely values you computed in Question #8.

Answer the Research Question

10. Based on your answer to Question #9, provide an answer to the research question posed at the beginning of the assignment.