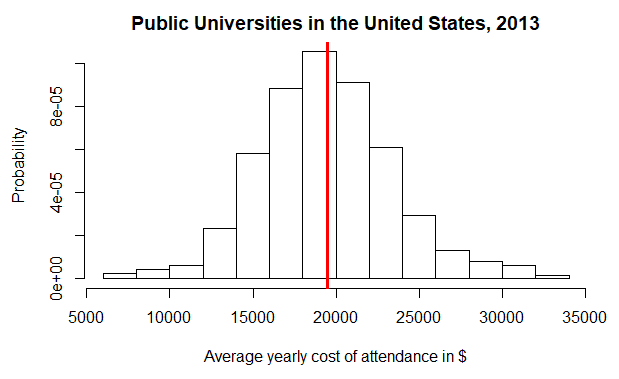
# Part A: Start small

**The college cost data**

  
**Note that μ = 19719.76, shown on the histogram as a thick, vertical line.**  
**Task 1:** In a few words, describe this **data** distribution. What is its shape? Where is it centered? How wide is it?

**Task 2: Take 5 samples of size n=5**

|  |  |  |
| --- | --- | --- |
| Sample number | Sample mean | Absolute value of difference |
| Sample 1 |  |  |
| Sample 2 |  |  |
| Sample 3 |  |  |
| Sample 4 |  |  |
| Sample 5 |  |  |
| Average of 5 sample means |  |  |

Mark each sample mean on the histogram above. Calculate the average of your sample means and the average of your absolute differences. How close are the sample means to the population mean?

**Task 3: Take 5 samples of size n=20**

|  |  |  |
| --- | --- | --- |
| Sample number | Sample mean | Absolute value of difference |
| Sample 1 |  |  |
| Sample 2 |  |  |
| Sample 3 |  |  |
| Sample 4 |  |  |
| Sample 5 |  |  |
| Average of 5 sample means |  |  |

Use a different symbol to mark each sample mean on the histogram. Calculate the average of your sample means and the average of your absolute differences. How close are the sample means to the population mean? Compare your results to when n = 5.

**Task 4: Take 5 samples of size n=50**

|  |  |  |
| --- | --- | --- |
| Sample number | Sample mean | Absolute value of difference |
| Sample 1 |  |  |
| Sample 2 |  |  |
| Sample 3 |  |  |
| Sample 4 |  |  |
| Sample 5 |  |  |
| Average of 5 sample means |  |  |

Use a different symbol to mark each sample mean on the histogram. Calculate the average of your sample means and the average of your absolute differences. How close are the sample means to the population mean? Compare your results to when n = 5 and n = 20.

# Part B: Go big or go home

**Task 5: Take 500 samples of size n = 5**

Run the R code to create a histograms of 500 sample means of size n = 5. Visually compare this sample means histogram to the data histogram. What do they have in common? How are they different?

**Task 6: Take 500 samples of size n = 20**

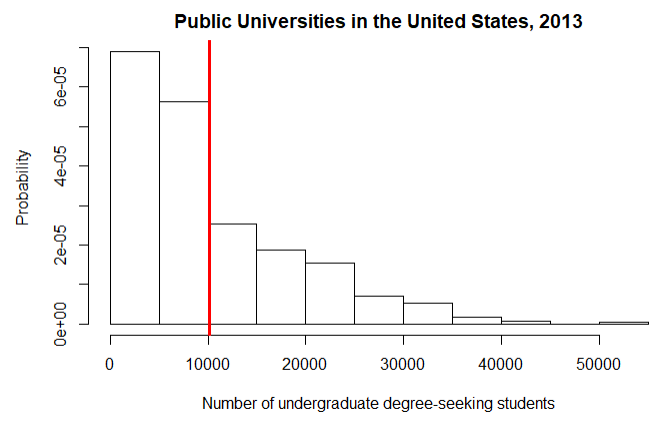
Run the R code to create a histograms of 500 sample means of size n = 20. Visually compare this sample means histogram to the data histogram and to the histogram(s) you created in task 5. What do they have in common? How are they different?

**Task 7: Take 500 samples of size n = 50**

Run the R code to create a histograms of 500 sample means of size n = 20. Visually compare this sample means histogram to the data histogram and to the histogram(s) you created in tasks 5 and 6. What do they have in common? How are they different?

# Part C: Repeat with a different variable

**For the college students data**



Note that μ = 10181.91, shown on the histogram as a thick vertical line.

**Task 8:** In a few words, describe this **data** distribution. What is its shape? Where is it centered? How wide is it?

**Task 9: Take 500 samples of size n = 5**

Run the R code to create a histograms of 500 sample means of size n = 5. Visually compare this sample means histogram to the data histogram. What do they have in common? How are they different?

**Task 10: Take 500 samples of size n = 20**

Run the R code to create a histograms of 500 sample means of size n = 5. Visually compare this sample means histogram to the data histogram. What do they have in common? How are they different?

**Task 11: Take 500 samples of size n = 50**

Run the R code to create a histograms of 500 sample means of size n = 5. Visually compare this sample means histogram to the data histogram. What do they have in common? How are they different?

True or False: For the students variable, as the sample size n got bigger, the sample means appeared to get closer to the population mean.

True or False: For the students variable, the histogram of sample means looked roughly like a normal distribution for each of the sample sizes n=5, n=20, and n=50.

True or False: For the students variable, the width of the histogram of sample means got narrower as the sample size got larger.