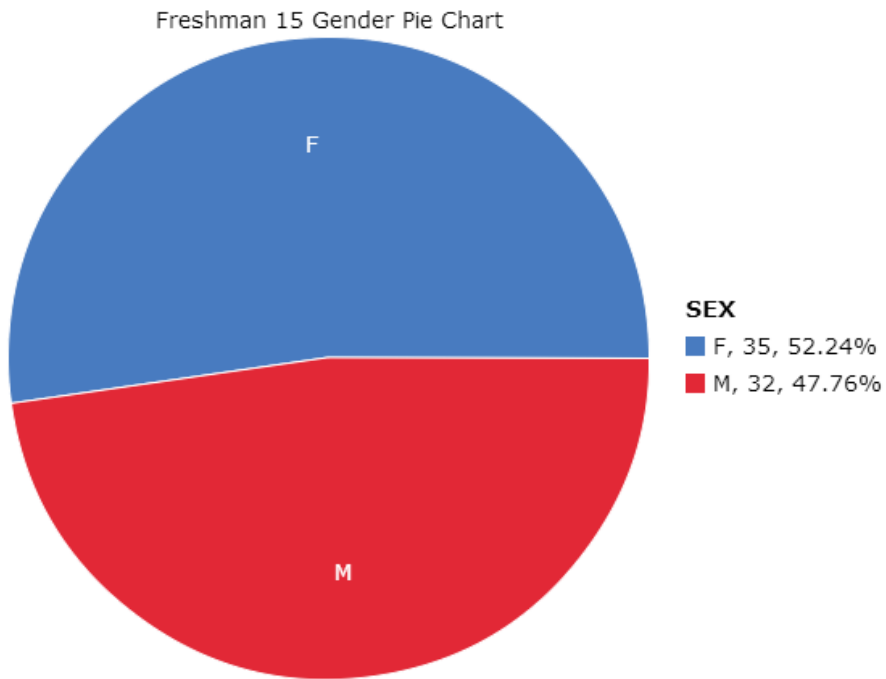
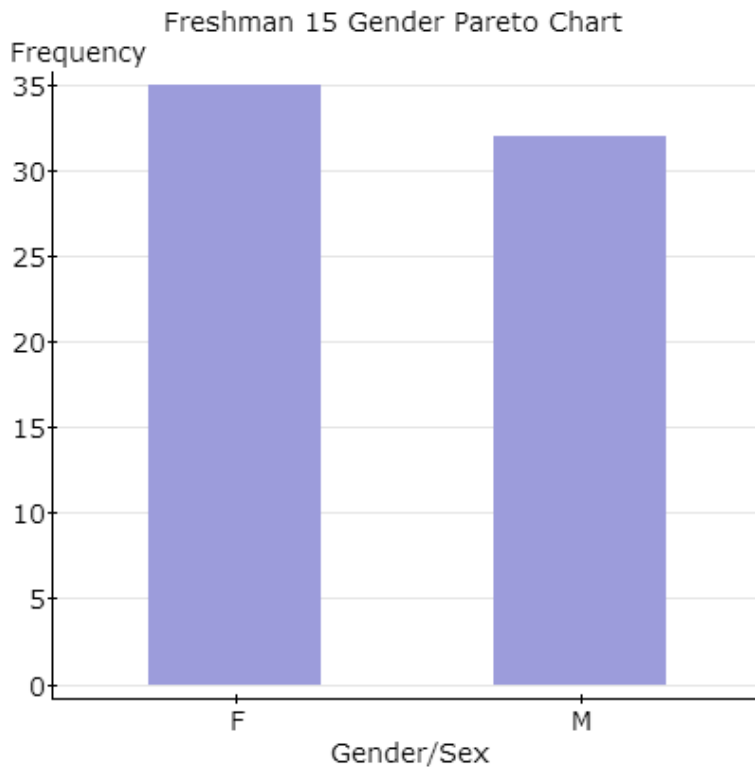


1) Variable Name: Sex/Gender

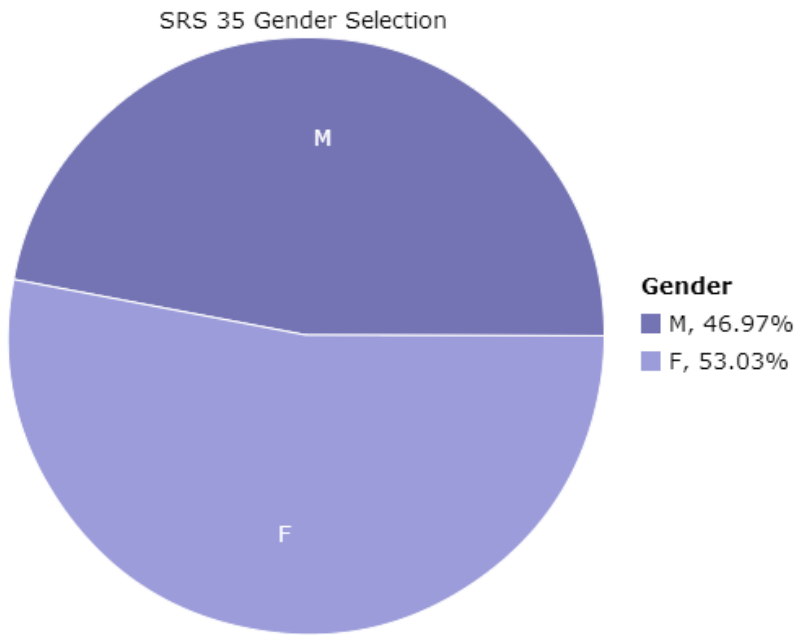
i) Freshman 15 Gender Pie Chart



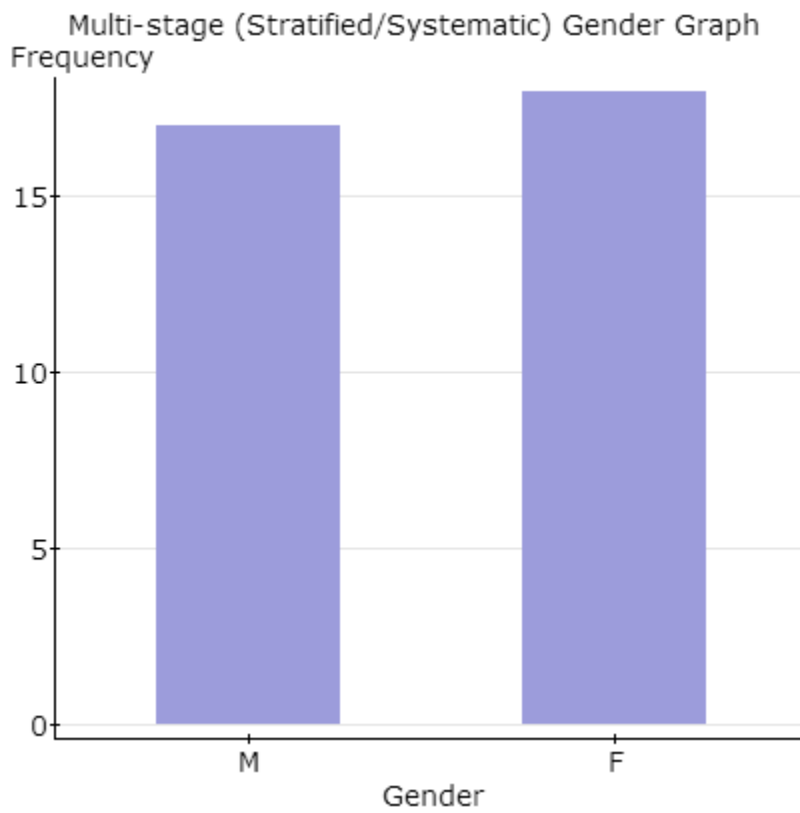
Freshman 15 Gender Pareto Chart



Simple Random Sample of 35 categorical data (Gender)



Multi-stage sample of 35 categorical data (Gender)



ii) Simple Random Sample n 35 Gender selection

Sample(SEX) row id	Sample(SEX)
66	M
37	F
56	F
47	F
53	M
20	M
3	M
2	M
54	M
57	F
17	F
49	F
18	M
30	F
15	F
48	F
64	M
31	F
51	M
59	M
8	F
60	F
65	F
61	M
63	M
19	F
43	M
6	M
39	F
55	F
42	M
5	F
22	F
14	F
9	F



erroneous because the likelihood of getting a higher female sampling is reasonable to expect, especially in any other sampling method other than just simple random sampling. (Going with probability concepts on that as well.)

- 4) My results compared to the population results really only varied based on the mean itself. The graphs made relatively true statements in comparison to the actual population, in that the females were slightly higher than the males, but not significantly so (there was less than 5% difference between the female and male population with only about 52% of the population being female). I feel that all the graphs that were created using sample data were still very spot on to the actual population data.

1) Sample Statistics

a. Random Sample

**Summary statistics:**

Column	Mean	Std. dev.	Min	Q1	Median	Q3	Max
Sample WTSEP	65.142857	11.463864	50	56	63	72	97

Random Sample Five Number Summary:

Min: 50                  Q1: 56                  Median: 63                  Q3: 72                  Max: 97

b. Stratified Sample

**Summary statistics:**

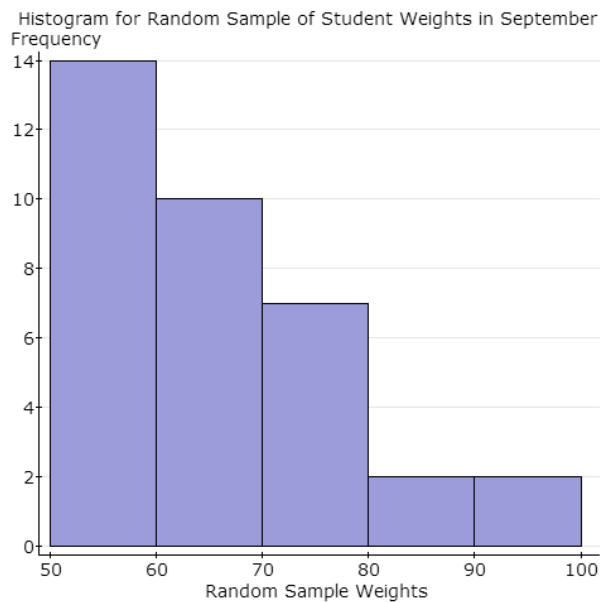
Column	Mean	Std. dev.	Min	Q1	Median	Q3	Max
WTSEPT Stratified	64.942857	12.160551	42	56	64	71	94

Stratified Sample Five Number Summary:

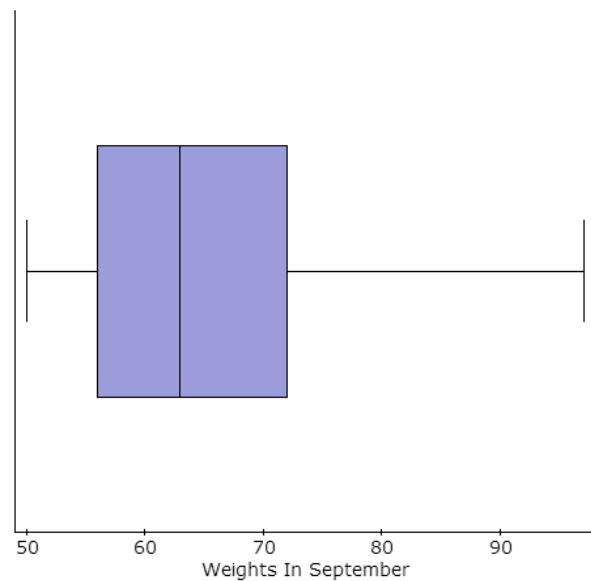
Min: 42                  Q1: 56                  Median: 64                  Q3: 71                  Max: 94

2) Frequency Histogram and Box Plot for each Sample

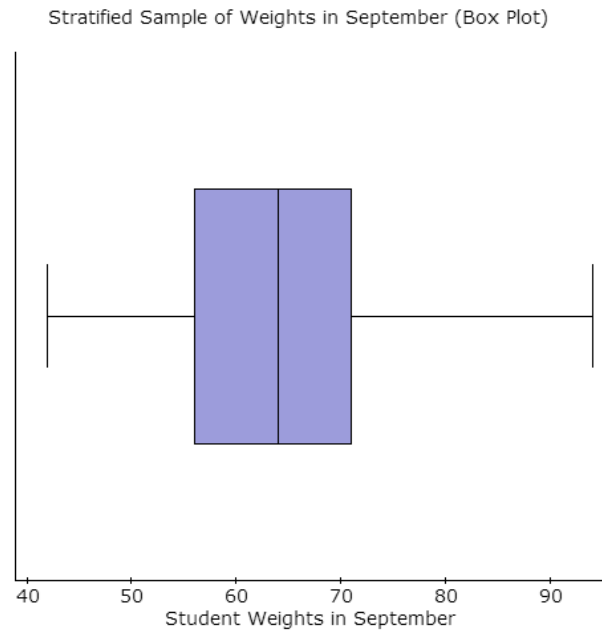
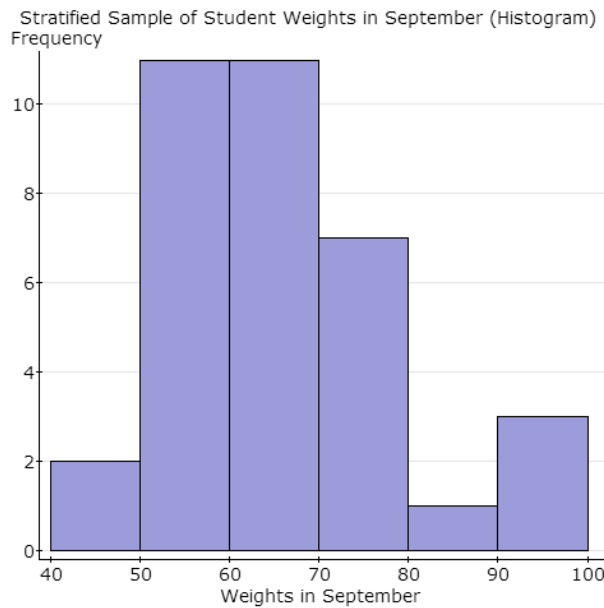
a. Random Sample Histogram and Box Plot



Random Sample of Student Weight in September (Box Plot)



## b. Stratified Sample Histogram and Box Plot



- 3) Both histograms appear to be slightly skewed to the right, although the Stratified sample does appear to be more bell shaped than the random sample. In either case, both have higher frequencies in the age ranges of 50-69. By comparison with the entire population of weights in September, the stratified sample appears to be almost a mirror image of the original population data, where as the random sample definitely appears to be slightly more skewed to the right. It would seem that a stratified sample would be the best way in order to get a true sample estimate of the population in this instance.

Part 1

2. For your group simple random sample from part 2 (categorical variable), create a confidence interval for the population proportion at your selected level of confidence, selecting one of the levels of your variable. Show the computations, including reporting the margin of error.

Our categorical variable is Sex. We are creating a confidence interval for the proportion of Men.

Confidence interval for "Sex"

$$E = z_{\alpha/2} \sqrt{(1-p)/n} = 1.96 \sqrt{0.457(1-0.457)/35} = 0.1651$$

(0.2919, 0.6221) is the 95% confidence interval for the population proportion of Men.

*The confidence interval .2919 < u < .6221 is interpreted to mean that we are 95% confident that the limits of .2919 and .6221 contain the true value of the population mean.*

3) From part 3, simple random sample, create a confidence interval for the population mean. (95% confidence) Show margin of error and computations

For this data we used the value of WTSEP or Weight in September.

$$\bar{x} = 65.14285714 \quad s = 11.46386358 \quad n = 35$$

$$\text{stat} \rightarrow \text{Test} \rightarrow 8 \rightarrow \hat{\mu} = (61.205, 69.081)$$

$$E = 3.938$$

$$t_{\alpha/2} = 0.05/2 = 0.025 = t_{0.025} = 2.032$$

$$2.032 * (11.46386358/\sqrt{35}) = 3.937501124$$

$$- E < \mu < + E$$

$$65.14285714 - 3.937501124 < \mu < 65.14285714 + 3.937501124$$

Confidence Interval: 61.20535602 <  $\mu$  < 69.08035826 Margin of Error (E): 3.938

*The confidence interval 61.20535602 <  $\mu$  < 69.08035826 is interpreted to mean that we are 95% that the limits of 61.205 and 69.080 contain the true value of the population mean. The margin of error tells us that our true population mean falls within 3.938 points of the sample population mean.*

4) For your group simple random sample from part 3 (quantitative variable), create a confidence interval for the population standard deviation at your selected level of confidence. Show the computations.

For this data we used the value of WTSEP or Weight in September.

$$n = 35; df = 34$$

$$CI = 0.05 \text{ *confidence level*}$$

$$= 9.86$$

$$s = 3.36$$

$$t_{\alpha/2} = 2.032$$

$$E = t_{\alpha/2} s/\sqrt{n} = 2.032 \times 3.36/\sqrt{35} = 1.154$$

$$- E < \sigma < + E \rightarrow 8.706 < \sigma < 11.014$$

*The confidence interval  $8.706 < \sigma < 11.014$  based on a 95% confidence is interpreted as we are 95% confident that the limits 8.706 and 11.014 contain the true value of the population mean.*

Part 2

Compare your confidence intervals to the values of the population parameters in your data sets. Did the intervals capture the population parameters?

The interval for the categorical data of gender was as follows: (0.2919, 0.6221) and based on proportions, should have captured the parameter.

$\bar{x} = 65.14285714$   $s = 11.46386358$  The population parameters for part 3 simple random sample were as follows: mean 65.143 and standard deviation = 11.464. The interval for this sample was as follows:  $61.20535602 < \mu < 69.08035826$  with a margin of error of 3.938

The intervals in this portion did capture the population parameter mean.

For the interval  $8.706 < \sigma < 11.014$  the mean was 9.86 so the interval did contain the mean.

There are many situations to which you rely on other people in order to accomplish regular tasks, possibly 90% of what you do even. However, when it comes to something that *you* have to pay for on an educational basis, I have never agreed with having to rely on other people in order to satisfy a requirement to achieve a long term accomplishment. That being said, this project has been a frustrating portion of this course for me in terms of group work. On the other hand, I have to admit that the actual application of work was quite interesting and definitely helped me get a better understanding of several of the concepts.

Being able to create graphs needed *and* see the comparison between what a group accomplished as well as yourself was of the most benefit. With the field I am seeking to join in the long run, I definitely believe that the ability to develop and understand graphs will benefit me because there are so many ways that these graphs can be used. Graphs by themselves are a great way to present statistical information in a way that is appealing to the viewer eye. I was about to say that I didn't think that qualitative data like gender would be as applicable to where I'm headed, I stopped and realized that I was dead wrong. Knowing the gender of a Client as opposed to the regular emotional levels on a graph could definitely be beneficial. Oh, there are so many things that just started going through my mind for the application of the way you could collect data and demonstrate it in a fabulous way (graphs) to which people would understand, actually be interested in, and be visually pleasing!

My only regret for this project is that we were not able to learn this material in programs that we will actually use in the future, for example, Microsoft Excel or the equivalent so that we would have been able to use these skills in actual jobs. I suppose that is all learnable through the internet, but application to real world usage seems so much more important. Overall, I am glad that I did have the opportunity to learn how to create graphs that visually explain data to the point to which you can actually interpret them in a socially applicable way.