# Part I:

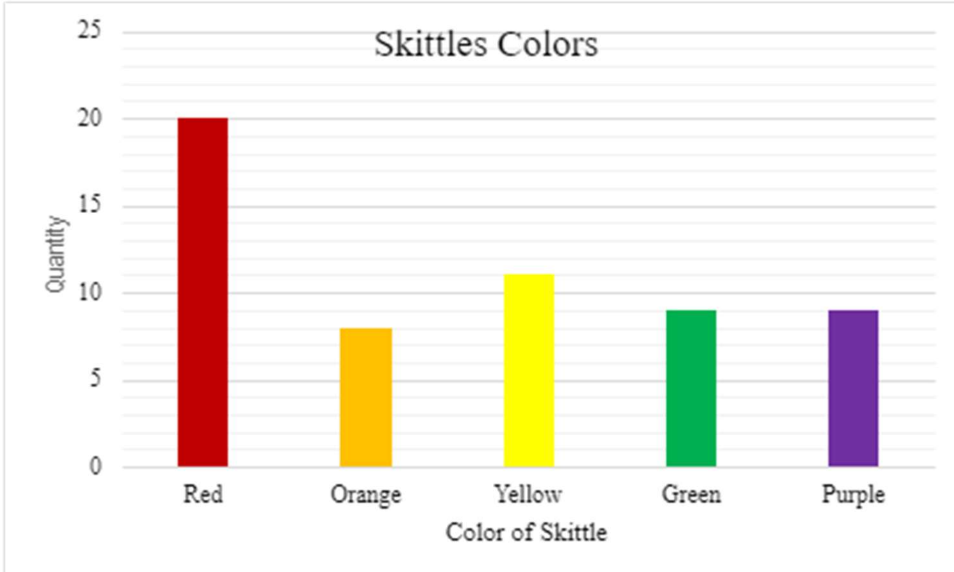
Red: 20

Orange: 8

Yellow: 11

Green: 9

Purple: 9



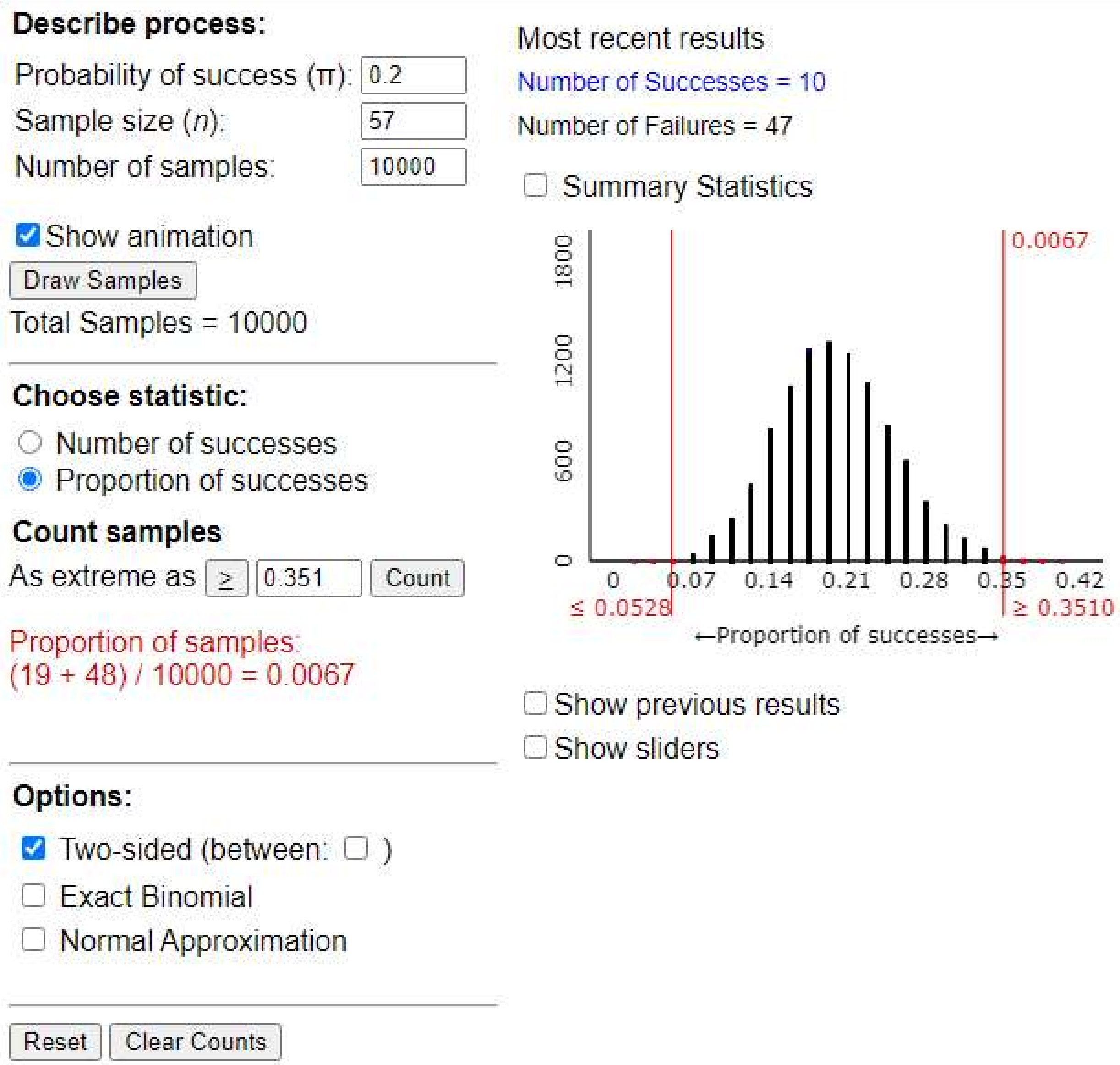

C:

1. The most common color of skittles I examined was red at 20 red skittles out of the 57 total skittles. The least common color was orange at 8 out of 57.
2. The relative frequency of red skittles in my sample was 0.351, $\hat{p} = 0.351$.
3. If I had a bag of 5,000 skittles, I would have 1,754 skittles of my most common color, red.

D:

The skittles I used should be from a random sample. During the production of the skittles, the amount of each color will be about the same, and since there is an unequal amount of each color in the bags of skittles, one can infer that the colors of the skittles are distributed randomly. Furthermore, if the amount of each color was standardized to be unequal, one would expect to see an equal distribution between bags. Since the amount of each color in different bags is different, one can conclude that the skittles are randomized.

E:



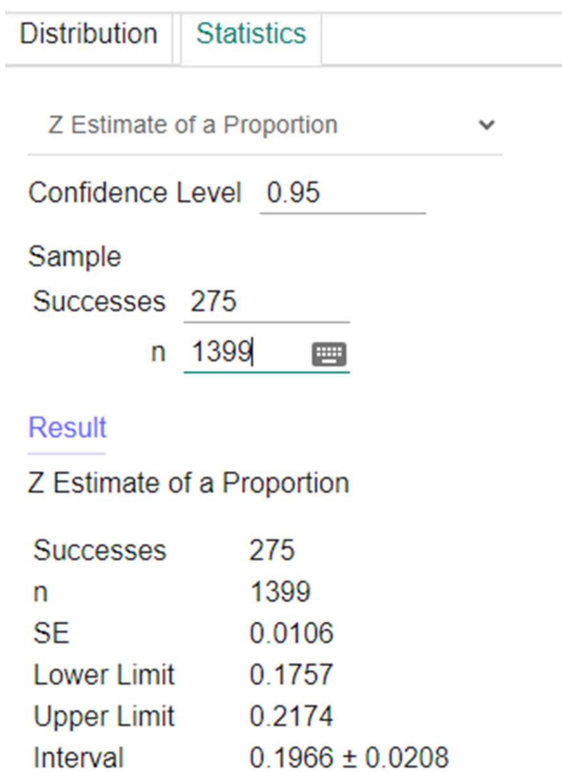

With an α of 0.05 and a p-value of 0.0070, I should reject the null hypothesis. This means that there is sufficient evidence that the probability of getting a red skittle is not 0.2.

# Part II:

A:

The sample can be considered random because each bag was compiled by random means and each student chose the bags from different stores. The sample is large enough because $np_0$ and $n(1- np_0)$ are both greater than or equal to 10. The value of $np_0$ is 275 and the value of $n(1- np_0)$ is 1,124. For the amount of skittles in the study to be 5% of the total population there would need to be at least 27,980 skittles in the total skittles population. One can assume that this number is less than the total skittle population. Because the sample is random, $np_0$ and $n(1- np_0)$ are $\geq 10$, and this sample is less than 5% of the total population, the confidence interval in part B will be valid.

B:




With a confidence interval of (0.1757, 0.2174). This means that with a 95% confidence level, the true proportion of red Skittles is between 0.1757 and 0.2174.

C:

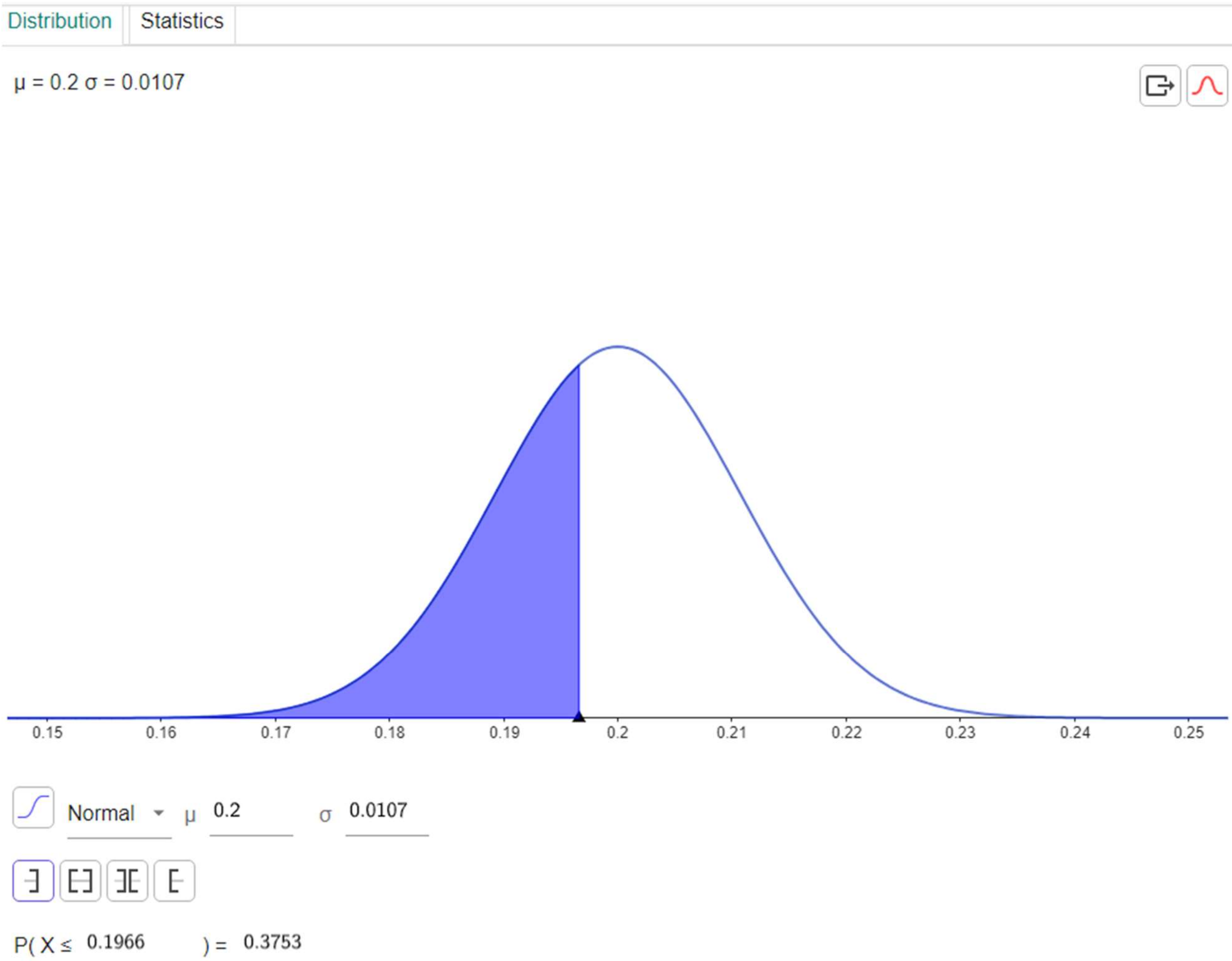
The proportion of red Skittles in my bag from part I was 0.351. Based on my confidence interval this is not a likely proportion. This is not only because my value is outside the range of the confidence interval but is very far out of the range. Due to how far out this is, the proportion of red skittles I got was very unlikely.

D:

The skittles distribution would still be considered random, and the sample would be considered independent, but we need to measure against a p of 0.2 for this test. The value of np is 279.8 and the value of n(p-1) is 1,119.2. Since the data is random, independent, and np and n(p-1) are both ≥ 10, the conditions for a two-tailed test are met.

E:

The null hypothesis is $H_0 = 0.2$ and the alternate hypothesis is $H_a \neq 0.2$. Since the p-value of 0.3753 is greater than 0.05 we can fail to reject the null hypothesis. The conclusion that we reach based on this is that there is not sufficient evidence to support the claim that the amount of red Skittles is different than 20%. Picture of GeoGebra is on the next page.




F:

With this hypothesis test, I failed to reject the null hypothesis, and with my part I I rejected the null hypothesis. My results were different between the different samples with the different sample sizes. The second sample is more reliable than the first sample because there were more Skittles to be observed. With a confidence level of 0.95 and a margin of error of 5% we would want a sample size of at least 246 skittles. Since the first sample did not have that many skittles and the second did, the second sample should be considered more reliable. Part I sample results on next page.

**Describe process:**

Probability of success (π): 0.2

Sample size ($n$): 57

Number of samples: 10000

☑ Show animation

Draw Samples

Total Samples = 10000

**Choose statistic:**

○ Number of successes

◉ Proportion of successes

**Count samples**

As extreme as ≥ 0.351 Count

Proportion of samples:
(19 + 48) / 10000 = 0.0067

**Options:**

☑ Two-sided (between: ☐ )

☐ Exact Binomial

☐ Normal Approximation

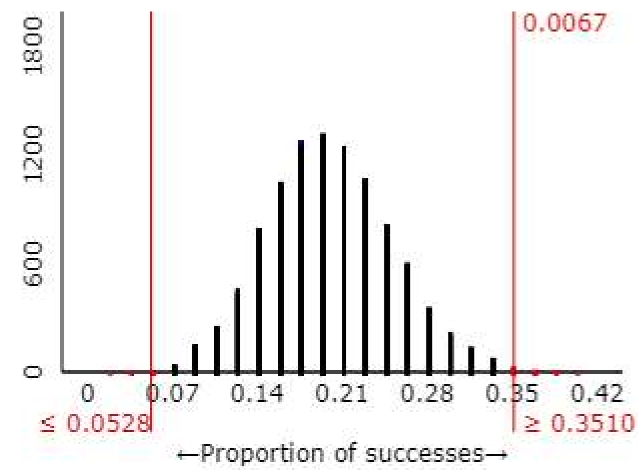
Reset Clear Counts

Most recent results

Number of Successes = 10

Number of Failures = 47

☐ Summary Statistics

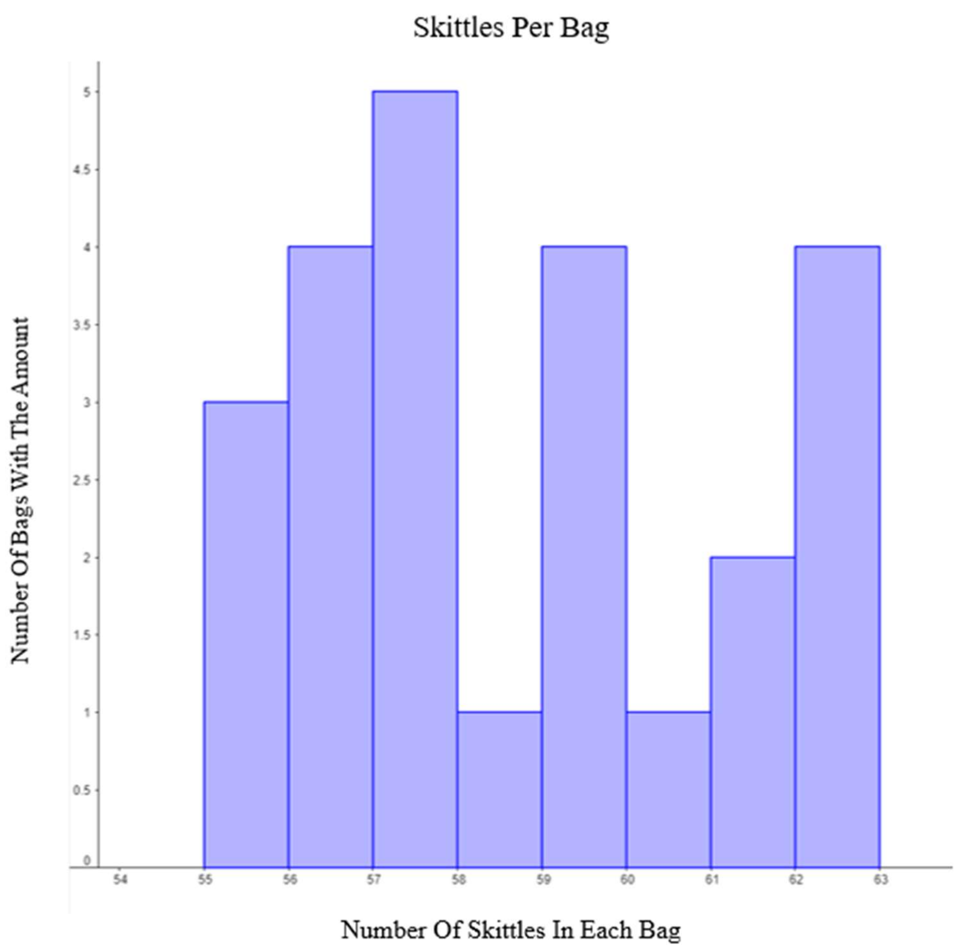



☐ Show previous results

☐ Show sliders

# Part III

A:




B:

The data are not normal or skewed. Because of this, the median and interquartile ranges are the best measures of center and variability. The median is 57.5 and the interquartile range is 4.5.

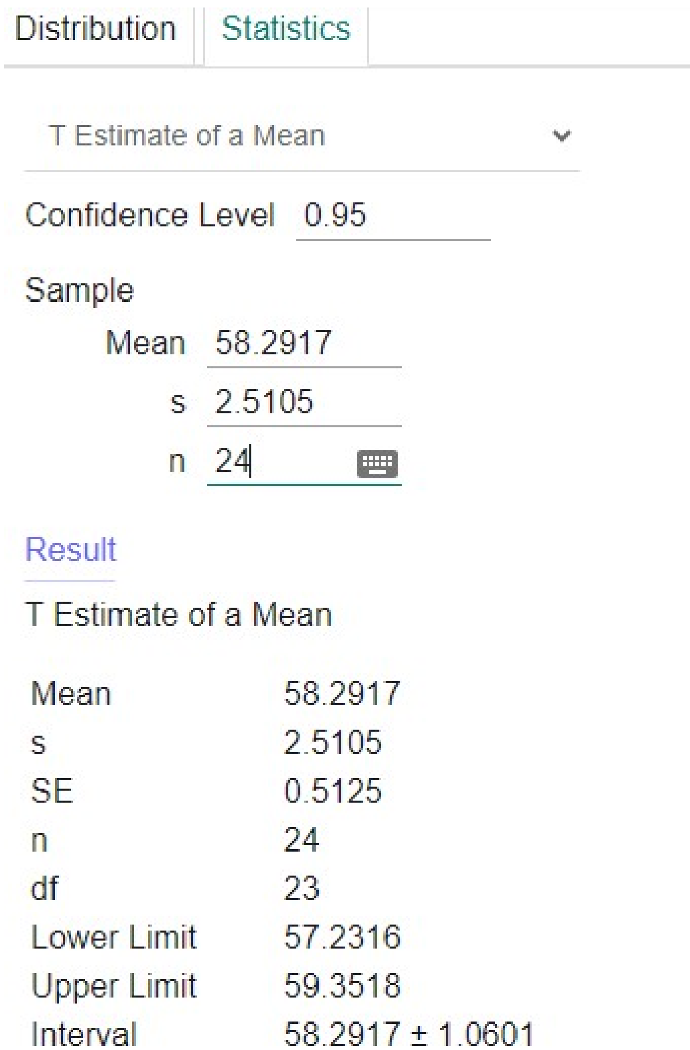
C:

There are no outliers. The interquartile range of the data is 4.5. To find the value to add or subtract from the Q1 and Q3 I can multiply 4.5 by 1.5 is 6.75. Q1-6.75 is 49.25. Since the lowest value is 55, there are no lower outliers. Q3+4.5 is 65. Since the maximum value is 63, there are no upper outliers.

D:

For the normal model to create a hypothesis test for $\mu$, the sample must be random or from a randomized experiment, the sample is normal, or the sample size is more than 30, and the sample is independent. This sample is from a randomized scenario. The sample of 24 is smaller than 30 and the sample is not normal. The sample is independent as the 24 pack are less than 5% of the total Skittles population. For this sample size to be more than 5% there would need to be 480 skittle packs or fewer in the world, and there are more than that. Because the sample size is less than 30 and the data are not normal, the outcomes from E and F will not be accurate.

E:




With the interval of 57.2316 to 59.3518 it means that, “We can be 95% confident that the true mean number of skittles per bag is between 57.2316 and 59.3518.”

F:

Hypothesis Test:

$\mu = 57$

$\mu \neq 57$

T Test of a Mean

Null Hypothesis μ = 57

Alternative Hypothesis ○ < ○ > ◉ ≠

Sample

Mean 58.2917

s 2.5105

n 24

Result

T Test of a Mean

| | |
|---|---|
| Mean | 58.2917 |
| s | 2.5105 |
| SE | 0.5125 |
| n | 24 |
| df | 23 |
| t | 2.5206 |
| p | 0.0191 |

With a p-value of 0.0191 and a significance level of 0.05, I should reject the null hypothesis. This means that there is sufficient evidence to conclude that the true mean of the amount of skittles per bag is different than 57.