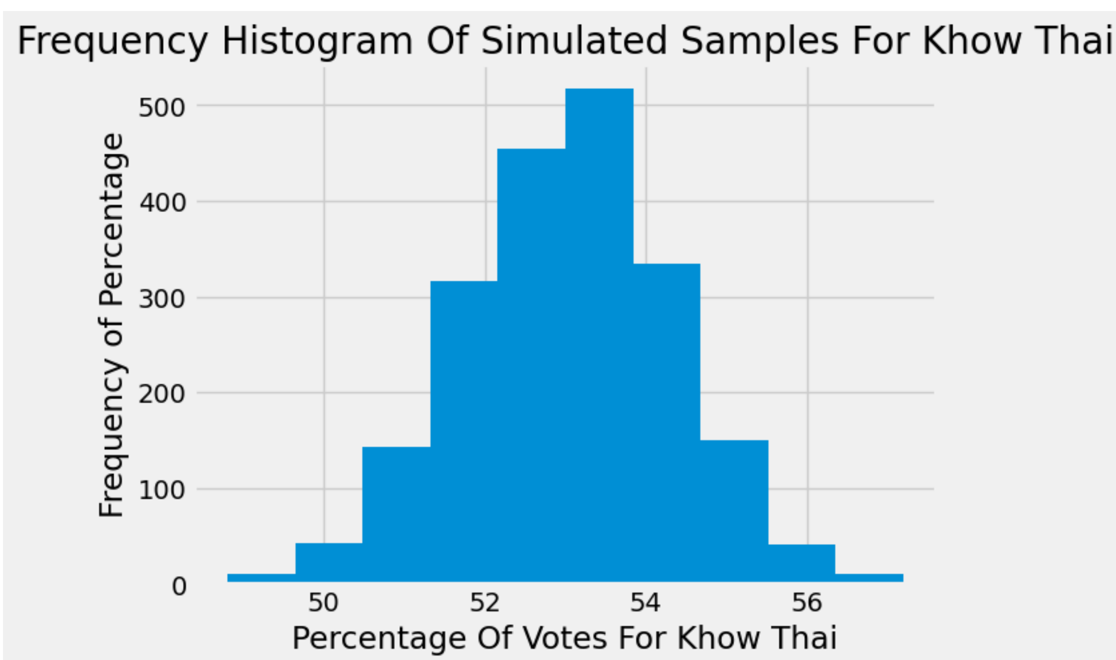
Question 1.2.

a). Complete the `percentages_in_resamples` function such that it simulates and returns a numpy array of 2022 elements, where each element represents a bootstrapped estimate of the percentage of voters who will vote for Khow Thai. You should use the `one_resampled_percentage` function you wrote above.

b). Then run your function `percentages_in_resamples` and store the results in a numpy array called `resampled_percentages`. Then create a density histogram of the entries in `resampled_percentages` array. Label your axes and include a title on your plot.

```
In [156]: def percentages_in_resamples():
              ret = np.array([one_resampled_percentage(votes) for _ in range(2022)])
              return ret
```

```
In [157]: resampled_percentages = percentages_in_resamples()
          plt.hist(resampled_percentages)
          plt.xlabel(f"Percentage Of Votes For Khow Thai")
          plt.ylabel(f"Frequency of Percentage")
          plt.title(f"Frequency Histogram Of Simulated Samples For Khow Thai")
          plt.show()
          # your code for histogram above this line
```



In [158]:
```
grader.check("q1_2")
```

Out[158]: q1_2 results: All test cases passed!

Question 2.5. The staff also created 70%, 90%, and 99% confidence intervals from the same sample, but we forgot to label which confidence interval represented which percentages! **First**, match each confidence level (70%, 90%, 99%) with its corresponding interval in the cell below

(e.g. ___ % CI: [52.1, 54] $\rightarrow$ replace the blank with one of the three confidence levels).

Then, explain your thought process and how you came up with your answers.

The intervals are below:

- [50.03, 55.94]
- [52.1, 54]
- [50.97, 54.99]

The larger the confidence interval, the wider the range. For instance, if we have a 100% confidence interval, we are essentially including all ranges of values in the data set. With this logic, we can then say

- `[50.03, 55.94]`: 99% confidence interval, largest range of values.
- `[50.97, 54.99]`: 90% confidence interval, median (size) range of values.
- `[52.1, 54]`: 70% confidence interval, smallest range of values.

Question 4.1. Michelle wants to use 10,000 bootstrap resamples to compute a confidence interval for the proportion of all Colorado voters who will vote Yes.

a). Use bootstrap resampling to simulate 10,000 election outcomes, and assign the np.array `resample_yes_proportions` to contain the Yes proportion of each bootstrap resample.

b). Calculate the 95% bootstrapped confidence interval for the Yes proportion.

c). Then, plot a density histogram of `resample_yes_proportions`. Include a title and label both axes. **You should see a bell shaped curve centered near the proportion of Yes in the original sample.** We have provided code that overlays your confidence interval at the bottom of your histogram.

In [184]:
```
prop68votes = np.array(["Yes"] * 210 + ["No"] * 190)
resample_yes_proportions = []
for _ in range(10000):
    resample = np.random.choice(prop68votes, size = 400, replace=True)
    yes_prop = np.sum(resample == "Yes") / 400
    resample_yes_proportions.append(yes_prop)
# Your code to simulate empirical distribution above this line
resample_yes_proportions
```

Out[184]:
```
[0.4825,
 0.5625,
 0.5425,
 0.535,
 0.5475,
 0.5175,
 0.57,
 0.535,
 0.4875,
 0.5325,
 0.495,
 0.5475,
 0.4975,
 0.51,
 0.51,
 0.4975,
 0.57,
 0.525,
 0.4875,
 0.515,
 0.5875,
 0.55,
 0.525,
 0.5225,
 0.5475,
 0.545,
 0.5375,
 0.5625,
```

```
0.5575,
0.4975,
0.5475,
0.525,
0.5675,
0.5175,
0.555,
0.53,
0.5025,
0.56,
0.55,
0.4825,
0.5225,
0.5575,
0.52,
0.525,
0.49,
0.5525,
0.53,
0.485,
0.4725,
0.545,
0.57,
0.5075,
0.57,
0.5175,
0.5225,
0.53,
0.5175,
0.495,
0.5075,
0.4975,
0.5525,
0.5575,
0.5175,
0.53,
0.5625,
0.5225,
0.5275,
0.505,
0.4925,
0.54,
0.5025,
0.5325,
0.4775,
0.55,
0.5675,
0.535,
0.5675,
0.5425,
0.515,
0.5175,
0.5525,
0.4825,
```

```
0.525,
0.5375,
0.5475,
0.545,
0.4825,
0.485,
0.54,
0.535,
0.5125,
0.475,
0.54,
0.53,
0.5175,
0.5275,
0.495,
0.5375,
0.5175,
0.51,
0.5375,
0.56,
0.5625,
0.53,
0.4825,
0.53,
0.5175,
0.5475,
0.47,
0.5075,
0.5125,
0.565,
0.5225,
0.5475,
0.5275,
0.4925,
0.53,
0.5275,
0.5475,
0.5525,
0.5,
0.5325,
0.5525,
0.5425,
0.5725,
0.54,
0.5025,
0.4925,
0.525,
0.555,
0.535,
0.5725,
0.5175,
0.515,
0.515,
0.495,
```

```
0.5275,
0.505,
0.54,
0.52,
0.5025,
0.5425,
0.54,
0.4925,
0.5325,
0.52,
0.565,
0.51,
0.5225,
0.51,
0.5575,
0.5075,
0.53,
0.505,
0.5075,
0.5175,
0.5175,
0.5225,
0.475,
0.535,
0.5475,
0.5325,
0.5375,
0.5425,
0.54,
0.555,
0.5675,
0.5275,
0.5525,
0.5125,
0.5025,
0.5325,
0.5,
0.55,
0.535,
0.485,
0.53,
0.54,
0.54,
0.5725,
0.5075,
0.495,
0.5775,
0.5425,
0.47,
0.535,
0.535,
0.4875,
0.5125,
0.4925,
```

0.5025,
0.49,
0.535,
0.5275,
0.5875,
0.51,
0.5275,
0.5525,
0.56,
0.5475,
0.535,
0.5225,
0.5275,
0.5475,
0.56,
0.5275,
0.55,
0.5075,
0.5625,
0.5275,
0.4975,
0.55,
0.5325,
0.5275,
0.52,
0.495,
0.5125,
0.5125,
0.53,
0.52,
0.535,
0.515,
0.53,
0.515,
0.555,
0.5675,
0.6025,
0.57,
0.4825,
0.4975,
0.5375,
0.56,
0.535,
0.5275,
0.565,
0.49,
0.5475,
0.5325,
0.5475,
0.4975,
0.55,
0.56,
0.545,
0.57,

```
0.52,
0.5275,
0.555,
0.5225,
0.5275,
0.5125,
0.565,
0.5475,
0.525,
0.505,
0.525,
0.4775,
0.5125,
0.515,
0.525,
0.5425,
0.535,
0.515,
0.5725,
0.4775,
0.5125,
0.5775,
0.5075,
0.52,
0.5225,
0.5475,
0.5575,
0.535,
0.5425,
0.5125,
0.5125,
0.5025,
0.505,
0.54,
0.54,
0.4775,
0.52,
0.5075,
0.5425,
0.49,
0.5375,
0.53,
0.545,
0.54,
0.535,
0.5475,
0.5825,
0.545,
0.5175,
0.5575,
0.51,
0.5275,
0.5275,
0.5375,
```

```
0.5325,
0.5075,
0.475,
0.5275,
0.5075,
0.53,
0.5325,
0.475,
0.5825,
0.5,
0.4925,
0.4925,
0.525,
0.505,
0.5125,
0.505,
0.515,
0.5175,
0.5075,
0.5325,
0.5375,
0.565,
0.53,
0.5275,
0.55,
0.54,
0.5275,
0.4925,
0.51,
0.51,
0.51,
0.535,
0.505,
0.5,
0.525,
0.535,
0.5125,
0.5225,
0.5325,
0.5475,
0.485,
0.5475,
0.555,
0.55,
0.5125,
0.52,
0.5425,
0.52,
0.515,
0.485,
0.4825,
0.5375,
0.5175,
0.5,
```

0.5375,
0.5525,
0.5325,
0.51,
0.5125,
0.5275,
0.4975,
0.51,
0.5,
0.5475,
0.5,
0.515,
0.5775,
0.5225,
0.54,
0.5725,
0.5425,
0.49,
0.515,
0.555,
0.53,
0.5425,
0.5,
0.5025,
0.5175,
0.4975,
0.495,
0.55,
0.54,
0.5,
0.5725,
0.5325,
0.58,
0.5225,
0.5,
0.545,
0.5625,
0.5025,
0.5075,
0.5,
0.5675,
0.5375,
0.5125,
0.495,
0.5275,
0.5175,
0.5575,
0.5675,
0.5475,
0.5675,
0.505,
0.545,
0.5275,
0.5125,

```
0.4975,
0.5325,
0.48,
0.54,
0.59,
0.5475,
0.525,
0.5175,
0.5275,
0.5175,
0.53,
0.5075,
0.5225,
0.54,
0.5225,
0.5125,
0.52,
0.4875,
0.5525,
0.53,
0.52,
0.5475,
0.475,
0.4725,
0.5175,
0.5325,
0.5425,
0.5,
0.555,
0.505,
0.53,
0.5,
0.5,
0.5225,
0.5075,
0.5275,
0.54,
0.54,
0.5225,
0.515,
0.535,
0.5925,
0.56,
0.52,
0.5225,
0.5175,
0.55,
0.5225,
0.54,
0.4975,
0.5325,
0.54,
0.525,
0.4925,
```

```
0.5175,
0.5375,
0.525,
0.53,
0.5275,
0.5275,
0.53,
0.58,
0.5275,
0.465,
0.5325,
0.5325,
0.5775,
0.495,
0.54,
0.51,
0.5175,
0.5425,
0.5375,
0.5,
0.5075,
0.5425,
0.51,
0.55,
0.5225,
0.5025,
0.505,
0.5375,
0.5475,
0.5575,
0.515,
0.5375,
0.5225,
0.55,
0.525,
0.4975,
0.525,
0.53,
0.515,
0.525,
0.5025,
0.48,
0.5025,
0.505,
0.51,
0.5125,
0.565,
0.535,
0.495,
0.5325,
0.5225,
0.5275,
0.515,
0.52,
```

```
0.4875,
0.5475,
0.485,
0.54,
0.51,
0.5425,
0.5175,
0.5425,
0.485,
0.555,
0.5225,
0.5425,
0.535,
0.505,
0.4975,
0.51,
0.5475,
0.55,
0.54,
0.5275,
0.54,
0.545,
0.4925,
0.515,
0.5025,
0.525,
0.5,
0.56,
0.5675,
0.5275,
0.52,
0.505,
0.5575,
0.545,
0.5525,
0.565,
0.5525,
0.49,
0.52,
0.515,
0.5325,
0.5425,
0.5,
0.54,
0.535,
0.5725,
0.5375,
0.555,
0.5075,
0.5825,
0.52,
0.5275,
0.51,
0.525,
```

```
0.5725,
0.47,
0.485,
0.5275,
0.47,
0.4775,
0.5375,
0.58,
0.5025,
0.565,
0.525,
0.5175,
0.5725,
0.56,
0.5625,
0.55,
0.4925,
0.5225,
0.485,
0.535,
0.5125,
0.4825,
0.4925,
0.555,
0.51,
0.5425,
0.505,
0.5225,
0.495,
0.505,
0.55,
0.4925,
0.56,
0.525,
0.56,
0.5475,
0.5425,
0.5275,
0.5425,
0.5375,
0.525,
0.48,
0.5225,
0.5275,
0.5075,
0.485,
0.5325,
0.5075,
0.535,
0.525,
0.5425,
0.515,
0.5175,
0.525,
```

```
0.5,
0.5475,
0.5875,
0.505,
0.55,
0.5075,
0.5425,
0.5,
0.5375,
0.5525,
0.55,
0.52,
0.545,
0.6,
0.525,
0.55,
0.54,
0.5075,
0.4975,
0.5275,
0.5375,
0.545,
0.5425,
0.5275,
0.535,
0.5475,
0.53,
0.51,
0.5425,
0.5325,
0.545,
0.53,
0.52,
0.5375,
0.4875,
0.5125,
0.5025,
0.565,
0.5475,
0.525,
0.565,
0.505,
0.4875,
0.51,
0.5175,
0.5475,
0.5725,
0.535,
0.5075,
0.51,
0.5275,
0.535,
0.515,
0.545,
```

```
0.5725,
0.5,
0.5475,
0.545,
0.5425,
0.52,
0.5275,
0.4975,
0.505,
0.535,
0.5125,
0.4875,
0.53,
0.5025,
0.535,
0.5075,
0.525,
0.5125,
0.53,
0.5225,
0.5275,
0.515,
0.5025,
0.555,
0.49,
0.535,
0.52,
0.5225,
0.515,
0.5325,
0.47,
0.5225,
0.555,
0.525,
0.495,
0.5425,
0.505,
0.52,
0.52,
0.4925,
0.4975,
0.5225,
0.515,
0.4875,
0.515,
0.535,
0.495,
0.525,
0.5425,
0.515,
0.515,
0.5275,
0.4925,
0.525,
```

```
0.515,
0.52,
0.5225,
0.5375,
0.5275,
0.495,
0.5025,
0.4825,
0.5425,
0.48,
0.53,
0.545,
0.475,
0.525,
0.515,
0.5325,
0.5575,
0.5,
0.5225,
0.54,
0.5325,
0.6,
0.525,
0.4825,
0.5275,
0.5025,
0.5325,
0.5325,
0.51,
0.5,
0.535,
0.515,
0.5275,
0.5475,
0.5525,
0.565,
0.52,
0.53,
0.525,
0.5075,
0.5175,
0.5475,
0.5075,
0.535,
0.5,
0.51,
0.5225,
0.4975,
0.5125,
0.58,
0.5425,
0.505,
0.5525,
0.505,
```

```
0.475,
0.515,
0.4725,
0.5275,
0.5325,
0.535,
0.5175,
0.5125,
0.555,
0.525,
0.525,
0.475,
0.515,
0.52,
0.5375,
0.515,
0.52,
0.53,
0.525,
0.545,
0.48,
0.525,
0.4825,
0.555,
0.48,
0.54,
0.5125,
0.515,
0.5375,
0.51,
0.5425,
0.56,
0.5,
0.58,
0.51,
0.505,
0.49,
0.5275,
0.565,
0.4875,
0.5225,
0.57,
0.5075,
0.5475,
0.505,
0.51,
0.48,
0.5275,
0.485,
0.5275,
0.5775,
0.5325,
0.5475,
0.5,
```

```
0.53,
0.505,
0.5425,
0.5625,
0.505,
0.535,
0.545,
0.565,
0.505,
0.555,
0.5225,
0.5425,
0.5325,
0.4925,
0.555,
0.5575,
0.545,
0.5325,
0.54,
0.515,
0.4625,
0.57,
0.5,
0.5275,
0.4925,
0.54,
0.53,
0.5375,
0.5475,
0.5125,
0.5125,
0.5375,
0.5575,
0.4975,
0.515,
0.5175,
0.5675,
0.54,
0.5,
0.5175,
0.5475,
0.53,
0.535,
0.515,
0.565,
0.5375,
0.5175,
0.5075,
0.485,
0.52,
0.51,
0.5375,
0.505,
0.5275,
```

```
0.5125,
0.5025,
0.535,
0.5325,
0.5275,
0.51,
0.5025,
0.52,
0.57,
0.48,
0.4875,
0.505,
0.5325,
0.4875,
0.515,
0.5675,
0.53,
0.555,
0.5125,
0.5,
0.5475,
0.52,
0.4875,
0.5,
0.5275,
0.5475,
0.555,
0.5125,
0.51,
0.5625,
0.4825,
0.5225,
0.5025,
0.5175,
0.5575,
0.5025,
0.5625,
0.535,
0.4925,
0.5725,
0.49,
0.5225,
0.495,
0.5675,
0.5675,
0.5225,
0.515,
0.5175,
0.5575,
0.545,
0.4725,
0.535,
0.505,
0.5175,
```

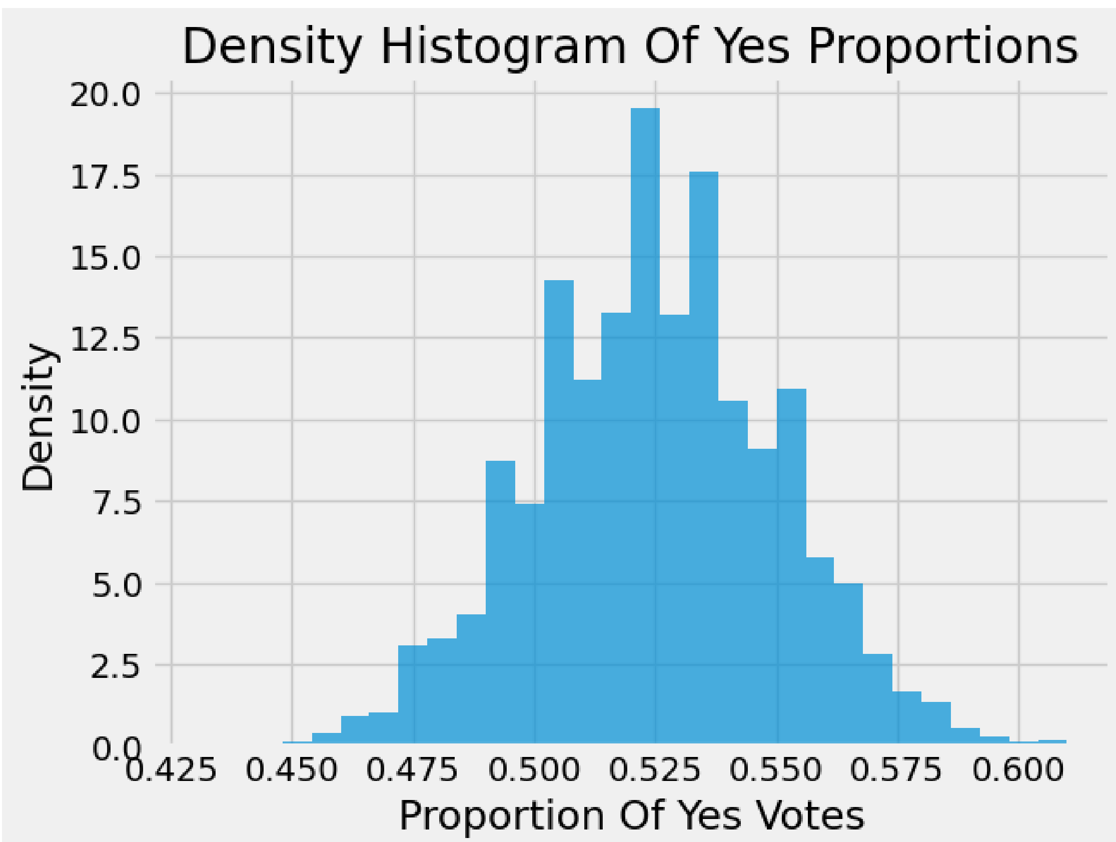
```
0.4925,
0.5075,
0.54,
0.5275,
0.56,
0.5175,
0.5325,
0.55,
0.5125,
0.5125,
0.5075,
0.515,
0.5325,
0.545,
0.49,
0.5275,
0.5675,
0.515,
0.5075,
0.5125,
0.5325,
0.5375,
0.5225,
0.515,
0.485,
0.5325,
0.535,
0.5175,
0.5125,
0.505,
0.5225,
0.5175,
0.4975,
0.5125,
0.5025,
0.535,
0.51,
0.52,
0.54,
0.48,
0.4975,
0.525,
0.52,
0.5425,
0.565,
0.5075,
0.5575,
0.5275,
0.525,
0.54,
0.5325,
0.5475,
0.515,
0.5725,
```

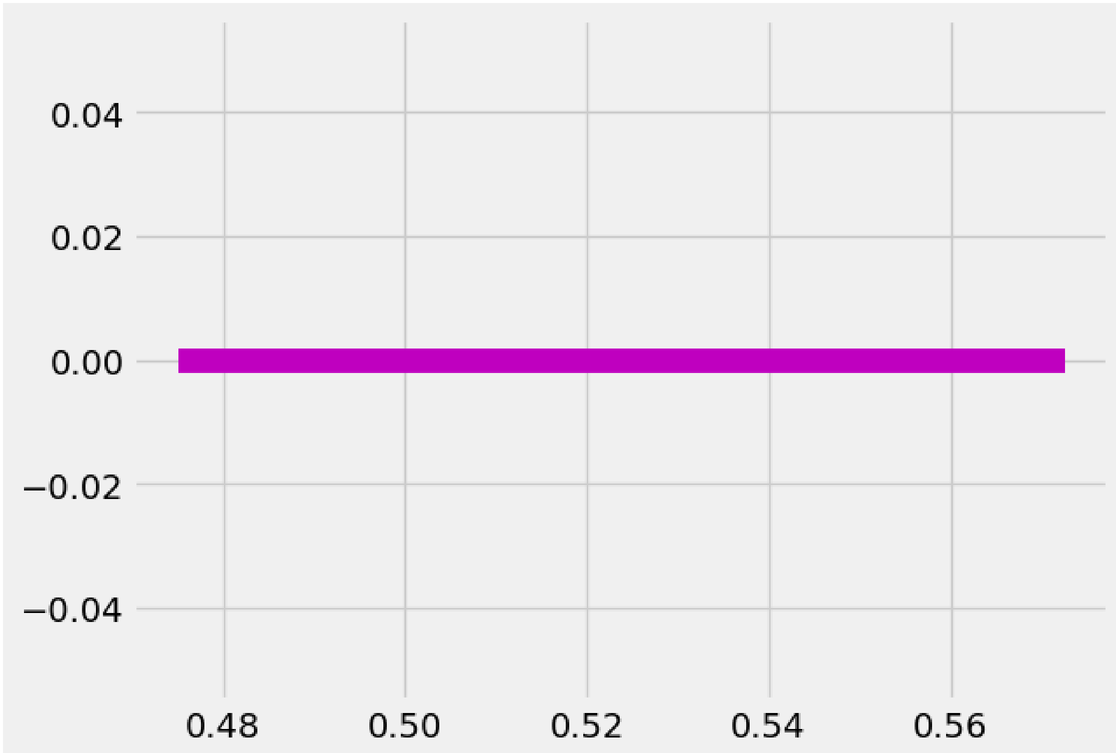
…]

```
In [185]: CI_lower = np.percentile(resample_yes_proportions, 2.5)
          CI_upper = np.percentile(resample_yes_proportions, 97.5)
          [CI_lower,CI_upper]
```

```
Out[185]: [0.475, 0.5725624999999991]
```

```
In [186]: plt.hist(resample_yes_proportions, bins=30, density=True, alpha=0.7)
          plt.title('Density Histogram Of Yes Proportions')
          plt.xlabel('Proportion Of Yes Votes')
          plt.ylabel('Density')
          plt.show()
          # Your code to plot histogram above this line
          plt.plot(np.array([CI_lower, CI_upper]), np.array([0, 0]), c='m', lw=10);
```





In [187]:
```
grader.check("q4_1")
```

Out[187]: q4_1 results: All test cases passed!

Question 4.2.

a). Why does the Central Limit Theorem (CLT) apply in this situation, and how does it explain the distribution we see above?

b). Prove the following: In a population whose members are 0 or 1, the **standard deviation** of that population is:

$$\text{standard deviation of population} = \sqrt{(\text{proportion of 0s}) \times (\text{proportion of 1s})}$$

Write up your answers to both parts in the same Markdown cell below:

0.0.1 Part A

The central limit theorem in this example applies to our data set because we have sufficient data points. Our sample is large enough that the population when plotted is then in turn normally distributed.

0.0.2 Part B

We know that the variation of Bernoulli trials is

$$\text{Var}(X) = E[X^2] - (E[X])^2.$$

Since a Bernoulli trial is a binary outcome, the former expression can then be simplified to

$$\text{Var}(X) = p - p^2 = p(1-p).$$

Furthermore, we can then say that the standard deviation (σ) is then

$$\sigma = \sqrt{p(1-p)}.$$

If we call the proportion of 0s p then the proportion of 1s is then $1 - p$. Putting this into the previous expression for the standard deviation of the population we show

$$\sigma = \sqrt{(\text{proportion of 0s})(\text{proportion of 1s})}.$$