Homework 6: Has average body temperature changed?

Inference for a Single Numerical Variable

Many researchers claim that the value of 98.6 degrees Fahrenheit for the average body temperature for adult humans is out-of-date, and doesn't apply to present day adult humans. A random sample of students in Stat 218 classes at Cal Poly were asked to report their body temperature in degrees Fahrenheit.

```
head(bodytemp_data)
```

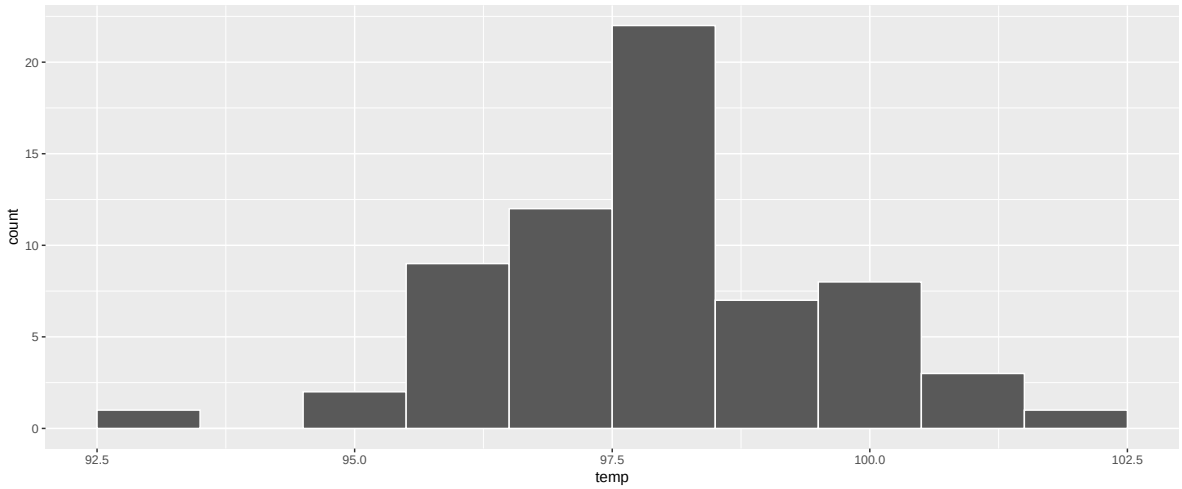
```
# A tibble: 6 x 4
  class     age major                 temp
  <chr>   <dbl> <chr>                <dbl>
1 Stat218    18 Marine Science        99.4
2 Stat218    19 Horticulture          98.7
3 Stat218    20 Wine and Viticulture  97.5
4 Stat218    21 Wine and Viticulture  96.5
5 Stat218    20 Wine and Viticulture  97.2
6 Stat218    22 Biology               96.1
```

The following summary statistics and visualizations for the average body temperature from our observed data are given:

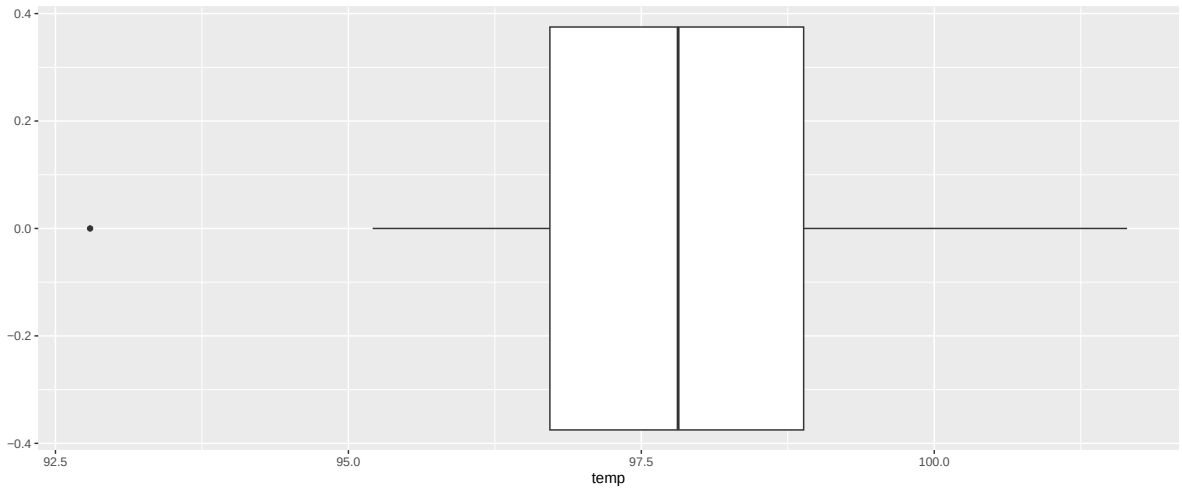
```
favstats( ~ temp, data = bodytemp_data)
```

```
     min       Q1   median       Q3      max     mean       sd  n missing
92.79559 96.71969 97.81392 98.88526 101.6455 97.90885 1.624018 65       0
```

```
ggplot(data = bodytemp_data,
       mapping = aes(x = temp)) +
  geom_histogram(binwidth = 1,
                 color = "white")
```




```
ggplot(data = bodytemp_data,
       mapping = aes(x = temp)) +
  geom_boxplot()
```



Research Question: Is there evidence the average body temperature for all adult humans is different from 98.6 degrees Fahrenheit?

Set-up

1. Identify the following in context of the study:

- Population:

- Sample:

- Variable of Interest:

- Data Type:

2. How many students were in the observed data set? What is the mean of the observed data? What is the standard deviation? Assign a symbol to each.

3. Describe the distribution of the observed data.

- Shape:

- Center:

- Spread:

- Outliers:

t-test

4. State the parameter of interest (in words, and assign a symbol).

5. Set up the null and alternative hypotheses in both words and symbols.

Remember what we would like to know is: If the average body temperature for adult humans is 98.6 degrees Fahrenheit, then how unusual/unlikely would it be for us to observe the data that we obtained on the sample?

6. Assuming that the average body temperature for adult humans is 98.6 degrees Fahrenheit (i.e., the null is true), use the central limit theorem to predict the shape, center (mean), and variability (s) of the distribution of sample means (i.e., null distribution).

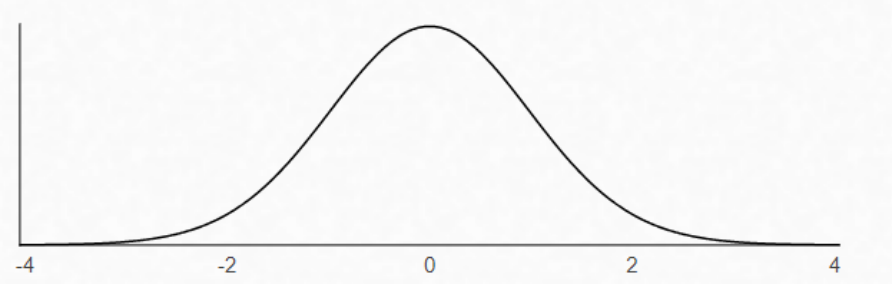
- Check the normality / large sample size assumptions for using a t-test.

- Shape of the distribution of sample means (i.e., null distribution)?

- Center (mean) of the distribution of sample means (i.e., null distribution)?

- Standard deviation distribution of sampling mean (i.e., standard error)?

I have provided the correct t-distribution for testing this research question below.



7. How many degrees of freedom does the t-distribution above have?

8. Calculate the T test statistic for the observed data.

9. On the t-distribution curve above, show how you would estimate/calculate the p-value to test the research question.

10. Select the appropriate R code below for testing the research question.

    a.

```
t_test(x = bodytemp_data,
       response = temp,
       mu = 98.6,
       alternative = "less",
       conf_int = FALSE
       )
```

```
# A tibble: 1 x 5
  statistic  t_df  p_value alternative estimate
      <dbl> <dbl>    <dbl> <chr>          <dbl>
1     -3.43    64 0.000529 less            97.9
```

b.

```
t_test(x = bodytemp_data,
       response = temp,
       mu = 98.6,
       alternative = "two-sided",
       conf_int = FALSE
       )
```

```
# A tibble: 1 x 5
  statistic  t_df p_value alternative estimate
      <dbl> <dbl>   <dbl> <chr>          <dbl>
1     -3.43    64 0.00106 two.sided       97.9
```

c.

```
t_test(x = bodytemp_data,
       response = temp,
       mu = 97.81,
       alternative = "two-sided",
       conf_int = FALSE
       )
```

```
# A tibble: 1 x 5
  statistic  t_df p_value alternative estimate
      <dbl> <dbl>   <dbl> <chr>          <dbl>
1     0.491    64   0.625 two.sided       97.9
```

d.

```
t_test(x = bodytemp_data,
       response = age,
       mu = 20,
       alternative = "two-sided",
       conf_int = FALSE
       )
```

```
# A tibble: 1 x 5
  statistic  t_df p_value alternative estimate
      <dbl> <dbl>   <dbl> <chr>          <dbl>
1    -0.656    64   0.514 two.sided       19.9
```

11. Using the correct output above, write a conclusion in context of the study.

12. Do you think it makes sense to conclude that the body temperature for *all* adults differs from 98.6 degrees Fahrenheit? Why or why not? Think about who was in the sample of our observed data.

i Key Idea: Scope of Inference

Obtaining a *random sample* from the population of interest allows us to *generalize* our results to the entire population.
If we do not have a *random (or representative) sample* from the population of interest, we can only draw conclusions about the sample.

Confidence Interval

Suppose instead of testing whether the body temperature differs from 98.6, we want to provide an interval estimate of the current average body temperature all intro statistics students.

Use the table below to answer the following four questions.

R code	Output
`qt(0.90, df = 45)`	1.3
`qt(0.95, df = 45)`	1.68
`qt(0.975, df = 45)`	1.68
`qt(0.99, df = 45)`	2.01
`qt(0.995, df = 45)`	2.69
`qt(0.90, df = 46)`	1.3
`qt(0.95, df = 46)`	1.68
`qt(0.975, df = 46)`	2.01
`qt(0.99, df = 46)`	2.41
`qt(0.995, df = 46)`	2.69

13. Calculate the 95% confidence interval.

14. Does it seem reasonable that 98.6 does not fall within the confidence interval? Explain.

15. What margin of error multiplier would you use for a 99% confidence interval? What happens to the width of the confidence interval?

16. How might your 95% confidence interval change if you, instead, had an observed standard deviation of 3.25?

Canvas Quiz

Make sure to complete the Homework Quiz on Canvas.