```
---
title: "Station 3: Inference with R"
author: "Your name here"
output: html_document
---

```{r, include = F}

# Clear Workspace
rm(list = ls())

# load packages we typically use for this class.
library(mosaic)
library(ggformula)
library(Stat2Data)
library(tidyverse)
```
```

Introduction

You should remember the basics of hypothesis testing from your introductory statistics class, whether it was MATH 102, AP statistics, or a course through a different department. **You can find the R code to perform all the necessary inference procedures on the "Inference Summary Sheet" handout on Moodle.**

At this station, you will work with your partner to pose and answer research questions from beginning to end, using the US States data (from Station 2). You will have the opportunity to ask and answer three different research questions, one using only one variable and two exploring the relationship between two variables. Although not required, **you may re-use your work from Station 2.**

Please reference the data codebook (available on the Data sets tab on Moodle) to decide which variables you'd like to use when forming your research question.

```
```{r, include = F}
#load the US_States data set here. Remember that even if it's already loaded in your
"Environment"/Console, you'll still need to load it into the RMarkdown document here!

```
```

A. Univariate: Quantitative

**TASK 1.1.1--Choose a quantitative variable from the US States dataset you'd like to explore. What is the variable? **

Pose a research question

TASK 1.1.2--What would be a reasonable research question to ask about this parameter? Your answer should make sense practically.

TASK 1.1.3--Write the null and alternative hypotheses in terms of the population parameter. If your research question is better suited for a confidence interval, you do not need to state hypotheses.

Perform EDA

A good EDA should attempt to answer the research question and also identify any issues that might impact eventual inference.

**TASK 1.1.4--Generate a relevant plot for your research question. **
```{r fig.width=5, fig.height=5}

```

TASK 1.1.5-- Calculate the summary statistic that is most relevant to your CI/test.

```{r}

```

TASK 1.1.6--Write a few sentences summarizing what you learned from your EDA

Answer the question using inference

To answer your research question, you'll need to test the hypotheses you defined earlier (or construct the relevant CI) .

TASK 1.1.7--Which statistical test will you perform to test the hypotheses? OR Which type of interval will you construct?

TASK 1.1.8--Did your EDA indicate any problems with performing this test/CI? That is, are the conditions for inference met?

TASK 1.1.9--Use R to generate the test statistic and p-value. OR Use R to build the confidence interval.
```{r}

```

TASK 1.1.10--Make a conclusion about the null and alternative hypotheses, being sure to answer your research question in context. OR Provide an interpretation of the confidence interval in context.

B. Relationship between two variables 1: Numerical/Binary

Choose two variables from the Codebook: one numerical, one binary. In this section you will investigate the relationship between these two variables.

TASK 2.1.1--What are the two variables? Your answer should include a brief description and be more than simply the variable name.

Pose a research question

TASK 2.1.2--What would be a reasonable research question to ask about this parameter?

Your answer should make sense practically.

TASK 2.1.3--Write the null and alternative hypotheses in terms of the population parameter.

Perform EDA

Remember, a good EDA should attempt to answer the research questions and also identify any issues that might impact eventual inference.

TASK 2.1.4--Generate a relevant plot for your research question.
```{r fig.width=5, fig.height=5}

```


TASK 2.1.5--Calculate the summary statistics that are most relevant to your hypothesis test.
```{r}

```


TASK 2.1.6--Write a few sentences summarizing what you learned from your EDA

Answer the question using inference

To answer your research question, you'll need to test the hypotheses you defined earlier.

TASK 2.1.7--Which statistical test will you perform to test the hypotheses?

TASK 2.1.8--Did your EDA indicate any problems with performing this test? That is, are the conditions for inference met?

TASK 2.1.9--Use R to generate the test statistic and p-value.
```{r}

```


TASK 2.1.10--Make a conclusion about the null and alternative hypotheses, being sure to answer your research question in context.

TASK 2.1.11--Use R to calculate a 95% confidence interval for the parameter of interest, and provide an interpretation in context.
```{r}

```
```

C. Relationship between two variables 2: Binary/Binary

Choose two binary variables from the Codebook. In this section you will investigate the
relationship between these two variables.

TASK 2.2.1--What are the two variables? Your answer should include a brief description
and be more than simply the variable name.

Pose a research question

TASK 2.2.2--What would be a reasonable research question to ask about this parameter?
Your answer should make sense practically.

**TASK 2.2.3--Write the null and alternative hypotheses in terms of the population
parameter.**

Perform EDA

Remember, a good EDA should attempt to answer the research questions and also identify any
issues that might impact eventual inference.

TASK 2.2.4--Generate a relevant plot for your research question.
```{r fig.width=5, fig.height=5}

```

**TASK 2.2.5--Calculate the summary statistics that are most relevant to your hypothesis
test.**
```{r}

```

TASK 2.2.6--Write a few sentences summarizing what you learned from your EDA

Answer the question using inference

To answer your research question, you'll need to test the hypotheses you defined earlier.

TASK 2.2.7--Which statistical test will you perform to test the hypotheses?

**TASK 2.2.8--Did your EDA indicate any problems with performing this test? That is, are the
conditions for inference met?**

TASK 2.2.9--Use R to generate the test statistic and p-value.
```{r}

```


TASK 2.2.10--Make a conclusion about the null and alternative hypotheses, being sure to answer your research question in context.

TASK 2.2.11--Use R to calculate a 95% confidence interval for the parameter of interest, and provide an interpretation in context.
```{r}

```