

STAT 218 Final Project

Team Member 1: NAME HERE Team Member 2: NAME HERE
Team Member 3: NAME HERE Team Member 4: NAME HERE
Team Member 5: NAME HERE

March 8th, 2024

Directions: In this project, you will apply various statistical procedures covered in our course. Ensure to load the necessary R packages, including ‘openintro’ and ‘tidyverse’, render the `library()` functions.

This code chunk below indicates how to load your data. If you use an OpenIntro data set, simply use `data()` function to load your data. If your data set is coming from any website as a .csv file, use `read_csv()` function to load it. Keep in mind that you can use multiple data sets as for this project (please do not try to merge them in a single file).

```
# Load necessary library functions here  
# data(datasetname) OR  
# midtermdata <- read_csv("data-set-name.csv")
```

Part 1: Getting to Know Your Data Set

Question 1: Use the `glimpse()` function to introduce your data set(s). Write a short paragraph explaining key aspects.

```
# Type your codes here
```

Type your answer here.

Question 2: Determine the number of variables and cases in the data set(s).

```
# Type your codes here
```

Type your answer here.

Question 3: Identify the types of variables in the data set (numeric or categorical).

Type your answer here.

Question 4: Visualize your numeric variables that you will use in this project and **DESCRIBE** what you see.

```
# Type your codes here
```

Type your answer here.

Question 5: Visualize your categorical variables that you will use in this project and **DESCRIBE** what you see.

```
# Type your codes here
```

Type your answer here.

Question 6: Identify and explain the handling of any outliers in your dataset, detailing whether they were removed and the rationale behind this decision.

```
# Type your codes here
```

Type your answer here.

Part 2: Statistical Inference

In this section, you'll perform inferential statistics tests using your data set.

The Chi-Square Test for Independence

Directions: In this section, you are expected to write a case study based on your data set. Choose TWO categorical variables (with at least 2 CATEGORIES) conduct The Chi-Square Test for Independence. Show your work.

Question 7.1. Introduce your case here.

Type your answer here.

Question 7.2. What is the research question of this study?

Type your answer here.

Question 7.3. What type of variables do we have in this study?

Type your answer here.

Question 7.4. Please list the assumptions required for this test.

List the assumptions here.

Directions for Question 5: Use a Chi-Square Test for Independence test to investigate your research hypothesis. Answer the questions given below. Show your work.

Question 7.5.1 Write out the null and the alternative hypotheses:

Type your answer here.

Question 7.5.2. Calculate Chi-square statistic and find the p-value and explain what you see in the output.

```
# Type your codes here
```

Type your answer here.

Question 7.5.3. Explain what each function/argument does in this question.

Type your answer here.

Question 7.5.4. Compute expected counts for each cell. Interpret the assumptions overall that you listed in Question 7.3.

```
# Type your codes here
```

Type your answer here.

Question 7.5.5. Draw a conclusion and write a conclusion statement.

Type your answer here.

Question 7.5.6. Describe the population to which you would feel comfortable generalizing the results of this study.

Type your answer here.

Question 7.5.7. Were the experimental/observational units in the study randomly assigned or randomly sampled? How does this affect the scope of conclusion that you can draw? Can you make a causal conclusion, why or why not?

Type your answer here.

Question 7.5.8. Looking back: Did anything about the design and conclusions of this study concern you? In particular, are there things that could have been done to give a better chance finding strong evidence against the null hypothesis? Issues you may want to critique include:

- Any mismatch between the research question and the study design
- How the experimental/observational units were selected
- How the treatments were assigned to the experimental units (if this is an experiment)
- How the measurements were recorded
- The number of experimental/observational units in the study
- Whether what we observed is of practical value

Type your answer here.

Question 7.5.9. Looking ahead: What should the researchers' next steps be to fix the limitations or build on this knowledge?

Type your answer here.

One Way ANOVA

Direction: Write a case study based on your data set. Choose a numeric variable and one categorical variable (The categorical variable should be preferably more than 2 categories) conduct One Way ANOVA. Show your work.

Question 8.1. Introduce your case here.

Type your answer here.

Question 8.2. What is the research question of this study?

Type your answer here.

Question 8.3. What type of variables do we have in this study?

Type your answer here.

Question 8.4. List the assumptions required for this test and check the assumptions for this study. Interpret the assumptions overall.

List the assumptions here and explain them.

```
# Type your codes here to check assumptions
```

Interpret the assumptions and type your reasoning.

Directions for Question 5: Conduct a One Way ANOVA to investigate your research hypothesis. Answer the questions given below. Show your work.

Question 8.5.1 Write out the null and alternative hypotheses, in the context of this study.

Type your answer here.

Question 8.5.2. Generate an ANOVA Table and explain what you see in the output.

```
# Type your codes here
```

Type your answer here.

Question 8.5.3. If your results imply a pairwise comparison, conduct your pairwise comparison here and summarize the output.

```
# Type your codes here
```

Type your answer here.

Question 8.5.4. Based on your analysis, draw a conclusion regarding the observed patterns or effects. Summarize your findings by providing a conclusion statement that succinctly communicates the key results of your study. Additionally, use the `summarize()` function to create a table that presents the mean values and standard deviation values for each group involved in your analysis.

Type your answer here.

Question 8.5.5. Describe the population to which you would feel comfortable generalizing the results of this study.

Type your answer here.

Question 8.5.6. Were the experimental/observational units in the study randomly assigned or randomly sampled? How does this affect the scope of conclusion that you can draw? Can you make a causal conclusion, why or why not?

Type your answer here.

Question 8.5.7. Looking back: Did anything about the design and conclusions of this study concern you? In particular, are there things that could have been done to give a better chance finding strong evidence against the null hypothesis? Issues you may want to critique include:

- Any mismatch between the research question and the study design
- How the experimental/observational units were selected

- How the treatments were assigned to the experimental units (if this is an experiment)
- How the measurements were recorded
- The number of experimental/observational units in the study
- Whether what we observed is of practical value

Type your answer here.

Question 8.5.8. Looking ahead: What should the researchers' next steps be to fix the limitations or build on this knowledge?

Type your answer here.

Two Way ANOVA

Directions: Write a case study based on your data set. Choose a numeric variable and two categorical variables (one of the categorical variables should be more than 2 categories) conduct Two Way ANOVA. Show your work.

Question 9.1. Introduce your case here.

Type your answer here.

Question 9.2. What is the research question of this study?

Type your answer here.

Question 9.3. What type of variables do we have in this study?

Type your answer here.

Question 9.4. List the assumptions required for this test and check the assumptions for this study. Interpret the assumptions overall.

List the assumptions here and explain them.

```
# Type your codes here to check assumptions
```

Interpret the assumptions and type your reasoning.

Directions for Question 9.5: Conduct a Two Way ANOVA to investigate your research hypothesis. Answer the questions given below. Show your work.

Question 9.5.1 Write out the null and alternative hypotheses in words, in the context of this study.

Type your answer here.

Question 9.5.2. Determine the (α) level.

Type your answer here.

Question 9.5.3. Generate an ANOVA Table and explain what you see in the output.

```
# Type your codes here
```

Type your answer here.

Question 9.5.4. Analyzing Interaction and Main Effects: In your statistical analysis, evaluate the presence of interaction effects. If an interaction effect is observed, proceed to conduct a post-hoc analysis. On the other hand, if no interaction effect is identified, shift the focus to the main effects. If any of the main effects are statistically significant, perform a post-hoc analysis for those effects. Provide a comprehensive explanation of the outputs obtained during these analyses.

```
# Type your codes here
```

Type your answer here.

Question 9.5.5. Based on your analysis, draw a conclusion regarding the observed patterns or effects. Summarize your findings by providing a conclusion statement that succinctly communicates the key results of your study. Additionally, use the `summarize()` function to create a table that presents the mean values and standard deviation values for each group involved in your analysis.

Type your answer here.

Question 9.5.6. Describe the population to which you would feel comfortable generalizing the results of this study.

Type your answer here.

Question 9.5.7. Were the experimental/observational units in the study randomly assigned or randomly sampled? How does this affect the scope of conclusion that you can draw? Can you make a causal conclusion, why or why not?

Type your answer here.

Question 9.5.8. Looking back: Did anything about the design and conclusions of this study concern you? In particular, are there things that could have been done to give a better chance finding strong evidence against the null hypothesis? Issues you may want to critique include:

- Any mismatch between the research question and the study design

- How the experimental/observational units were selected
- How the treatments were assigned to the experimental units (if this is an experiment)
- How the measurements were recorded
- The number of experimental/observational units in the study
- Whether what we observed is of practical value

Type your answer here.

Question 9.5.9. Looking ahead: What should the researchers' next steps be to fix the limitations or build on this knowledge?

Type your answer here.

Least Squares Regression

Write a case study based on your data set. Choose 2 numeric variables conduct bivariate regression to model the relationship between these two variables. Show your work.

Question 10.1. Introduce your case here.

Type your answer here.

Question 10.2. What is the research question of this study?

Type your answer here.

Question 10.3. What type of variable do we have in this study?

Type your answer here.

Question 10.4. List the assumptions required for this test.

List the assumptions here and explain them.

Directions for Question 10.5. Conduct a least squares regression to investigate your research question. Answer the questions given below. Show your work.

Question 10.5.1 Write out the null & the alternative hypotheses:

Type your answer here.

Question 10.5.2. Fit a least squares regression model and explain what you see in the output.


```
# Type your codes here
```

Type your answer here.

Question 10.5.3. Go back to the assumptions. Check the assumptions for this study and interpret the assumptions overall.

```
# Type your codes here to check assumptions
```

Interpret the assumptions and type your reasoning.

Question 10.5.4. Based on your analysis, draw a conclusion. Summarize your findings by providing a conclusion statement that succinctly communicates the key results of your study.

Type your answer here.

Question 10.5.5. Describe the population to which you would feel comfortable generalizing the results of this study.

Type your answer here.

Question 10.5.6. Were the experimental/observational units in the study randomly assigned or randomly sampled? How does this affect the scope of conclusion that you can draw? Can you make a causal conclusion, why or why not?

Type your answer here.

Question 10.5.7. Looking back: Did anything about the design and conclusions of this study concern you? In particular, are there things that could have been done to give a better chance finding strong evidence against the null hypothesis? Issues you may want to critique include:

- Any mismatch between the research question and the study design
- How the experimental/observational units were selected
- How the treatments were assigned to the experimental units (if this is an experiment)
- How the measurements were recorded
- The number of experimental/observational units in the study
- Whether what we observed is of practical value

Types your answer here.

Question 10.5.8. Looking ahead: What should the researchers' next steps be to fix the limitations or build on this knowledge?

Types your answer here.