Instructions for Lab Report

Engineering Statistics I (MA223)

You are to answer the following questions for this lab; use the "MA223 Assignment Template" (in RStudio, `File > New File > R Markdown > From Template > MA223 Assignment`).

Question 1: Research Objective

State the research objective you hope to address in this lab.

> Criteria
>
> In order to successfully complete this question, your answer must adhere to the following requirements:
>
> - A research objective is stated.
> - The objective is stated in context (specific variables referenced, units included, no statistical jargon).
> - It is clearly stated whether you are seeking "evidence" for a statement, determining if a statement is "reasonable," or "estimating" an unknown quantity.

Question 2: Variables

List all variables that will be recorded on each unit of observation (subject) in the sample.

> Criteria
>
> In order to successfully complete this question, your answer must adhere to the following requirements:
>
> - Every variable recorded in your dataset is stated.

- The role (i.e., response, predictor, factor, block, or extraneous variable) of each variable recorded is identified; note that every variable recorded has one of these roles, but not every role needs to be present in a dataset.
- Units are specified, when applicable.

Question 3: Model

State an appropriate model for the data generating process.

Criteria

In order to successfully complete this question, your answer must adhere to the following requirements:

- The model is a function of both parameters and recorded variables.
- Names of variables are obvious, or unclear notation is clearly defined.
- Each variable is appropriately indexed.
- Error term is included and is appropriately indexed.
- The model is appropriate for addressing the research objective.

Question 4: Framing the Research Objective

Identify the parameter(s) of interest from the model for the data generating process will help you address your research objective. Depending on your objective, this may be presented in a couple of ways:

1. State the hypotheses associated with the research objective using the parameter(s) from the model for the data generating process you defined above.
2. Give an interpretation of the parameter(s) from the model for the data generating process in the context of the problem and state the confidence level you intend to use to estimate the parameter(s).

Criteria

In order to successfully complete this question, your answer must adhere to the following requirements:

- If stating hypotheses:

  - The hypotheses correspond to the research objective defined previously.
  - The hypotheses are stated using mathematical notation.
  - The hypotheses make use of the parameter(s) from the model for the data generating process.

- If defining parameters:
  - Parameters correspond to the research objective defined previously.
  - Parameters are defined in the context of the problem (with units, if applicable).
  - The confidence level you intend to use is stated.

Question 5: Data Collection

In one or two paragraphs, describe your data collection procedure.

Criteria

In order to successfully complete this question, your answer must adhere to the following requirements:

- Description includes how each measurement was obtained.
- State the number of observations collected.
- Provide enough detail that an outside party (not someone in the course) could recreate your collection process.
- Description includes rationale (how this helps construct measurements that will be representative of the population) for decisions. This includes appropriate use/description of randomization techniques, if used.
- Description is concise.

Question 6: Limitations and Sources of Error

State one limitation of the study design and state one source of error.

Note

A limitation of a study "limits" the scope of generalization; that is, it "limits" how the results apply to the intended population. Limitations often reflect decisions that result in all observations sharing (or omitting) some characteristic. Occasionally, limitations are not the result of decisions but simply bad luck. Certainly, in order to complete the study in the allotted time, limitations on the scope of the study are necessary. **"Sample size" is not a valid limitation.**

A source of error contributes to the variability in the response from one observation to the next. Sources of error reflect inherent variability in the units of observation (with respect to the response of interest). They can be the result of imprecise measurement practices. However, **"more precise equipment" is not a valid source of error.**

Criteria

In order to successfully complete this question, your answer must adhere to the following requirements:

- Reasonable limitation of the study is stated.
- Posit how the results might be impacted by the limitation.
- Reasonable source of error is stated.

Question 7: Graphic

Construct a graphical summary of the data collected that addresses the research objective. Then, write a short description of what you learned about the research objective from the graphic.

Note

Your description should tell the *story* behind the data. The description is **not** asking for a technical summary of how the graphic ("each bar represents the height of…"). Instead, in the context of the research objective being studied, the description summarizes what the graphic is communicating about the data. Descriptions often focus on trends in location, spread, or shape (any aspect of the distribution).

Criteria

In order to successfully complete this question, your answer must adhere to the following requirements:

- Graphic is appropriate for summarizing the raw data towards addressing the research objective.
- Graphic illustrates the variability in the data.
- All axes (x-, y-, color-, etc.) are clearly labelled.
- Description is written towards addressing the research objective.
- Key statistical terminology is used appropriately.

Question 8: Conditions

Describe how you determined which conditions/assumptions to impose on the model for the data generating process.

Criteria

In order to successfully complete this question, your answer must adhere to the following requirements:

- Each potential condition is stated and assessed.
- If applicable, an appropriate graphic is provided for each condition.
- For each condition, a statement is given indicating whether the data is consistent with the condition or not.
- Rationale is provided to support the conclusion about each condition.
- Rationale has consistent and appropriate logic.
- Key statistical terminology is used appropriately.

Question 9: Analysis

Conduct an appropriate statistical analysis to address the research objective, and provide relevant output.

> **Note**
>
> It may be that based on your assessment in Question 8, an appropriate analysis is not possible given the content of the course. In that case, state any additional conditions *required* to conduct an analysis and then proceed under those conditions.

> **Criteria**
>
> In order to successfully complete this question, your answer must adhere to the following requirements:
>
> - Code used to conduct the analysis is provided.
> - All output that will be used to address the research objective is provided (this could include, where applicable, p-values, confidence intervals, summary of model fit, etc.).
> - The analysis conducted is appropriate for addressing the research objective.
> - The analysis conducted is consistent with the conditions determined to be reasonable in Question 8; if any additional conditions (beyond those stated as reasonable in Question 8) were imposed to obtain results, state those additional conditions.

Question 10: Conclusion

Based on the results above, interpret your findings in context. That is, what can be concluded about the research objective/question?

> **Criteria**
>
> In order to successfully complete this question, your answer must adhere to the following requirements:
>
> - The conclusions are stated in the context of the problem; no statistical jargon (e.g., "reject the null hypothesis") is used, and no mathematical symbols are relied upon (even if previously defined).
> - Conclusions relate back to the research objective defined in Question 0.
> - Conclusions are consistent with the analysis provided in Question 9.
> - An appropriate p-value or CI is cited.
> - Key statistical terminology is used appropriately.