1. **Project Description** – Write one or more paragraphs to provide relevant details such as:

* Who your client is
* The population of interest
* The type of study (e.g. observational vs. experimental)
* How the data were collected
* The sample size (mention here or below)
* The general goal(s) of the study
  1. **Research Questions** – Specifically state the research question(s) you will answer.

*Question 1: …*

*Question 2: …*

Etc.

* 1. **Variables** – Discuss variables in context of the study and its goals. Also, include a table to summarize the variables used in the report. E.g., “*The five variables in the dataset are summarized in Table 1 below.”*

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Type | Description | Levels and Ranges |
| Brand | Categorical | Oil brand used to cook fries | Four brands |
| OilAge | Categorical | How many times oil was already used | 1 to 5, where 5=5 or more |
| Sex | Categorical | Biological sex of participant | 1=Male and 2=Female |
| Age | Numeric | Age of participant in years | 18 to 65 years old |
| Taste | Numeric | Participant rating for fries | 1 to 9 where 9 is the best |

*Table 1: Summary of variables for the French Fries study*

1. **Exploratory Data Analysis (EDA)**

* Include numerical summaries (e.g., summary statistics and/or frequency tables)
* Include the most relevant graphs in terms of your statistical results and recommendations; place all other graphs in the Appendix
* You may note outliers, but they will need to be reassessed in your statistical analysis section below when fitting models
* Every table and graph need a descriptive caption (e.g., *Table 2: blah blah blah*) and needs to be referenced within the report (e.g., *“See Table 2 for…”*)
* Given your client has little statistical knowledge, interpret your EDA using language that can easily be understood
* Note when interpretations must be verified using statistical analysis (e.g., *“We can see from the boxplots shown in Figure 1 that there may be a significant difference in average taste scores for the oil brands. However, we need to run an analysis of variance to investigate this further.”*)

1. **Statistical Analysis**

* Note statistical software used
* If fitting a model, note the original model including the main effects and interaction terms that were included (e.g., *“We used an ANCOVA to model the taste response using the participant age covariate and the factors brand, oil age, and sex, including all of the two-way interaction terms.”*)
* Specify the model reduction method used to arrive at the final, hierarchical model (e.g. =0.05 with stepwise, forward selection, backward elimination, AIC, etc.) – your results must be reproducible by someone else
* Use a table to list the variables in your final, reduced model (include variable names and p-values):

|  |  |
| --- | --- |
| *Variable* | *P-value* |
| Brand | 0.001 |
| OilAge | 0.274 |
| Brand\*OilAge | 0.000 |

*Table 2: Summary of the final model for analyzing the taste variable*

* You do not need to formally write the hypotheses – your client likely will not understand them – but you need to clearly communicate how to interpret the results (e.g., *“Per Table 2, there is a significant interaction between oil brand and age. In other words… OR Therefore…”)*
* Mention assumption checks for your final model (residual plots can be placed in the Appendix)
* Per the assumption checks, note any outliers (e.g., using standardized residuals beyond ±2 or ±3); consider provide two analyses – with and without outliers
* Consider including any other relevant results (e.g., adjusted R-squared, VIFs, Tukey comparisons, etc.)
* Similar to EDA section, every table and graph needs a caption AND needs to be referenced in the report

1. **Recommendations**

List each research question and directly answer each one:

*Question 1: …*

*Question 2: …*

Etc.

1. **Resources** – List resources (website, textbooks, statistical software resources, etc.) your client may find useful. You must provide a few resources in your report.
2. **Additional Considerations** – Provide any important notes your client may want to consider for this analysis (e.g., limitations to your recommendations) and/or future analyses. Also, if the study was observational (rather than experimental), mention that we cannot conclude there is causation.
3. **Appendix**

* This section can be as technical as you want; it does not have to be written for a non-statistician
* Include software code or menus to ensure your results are reproducible
* Any results that you created and want to retain but decided to omit from the main body of the report:
  + Graphs, tables, etc. that you omitted from the EDA section
  + ANOVA tables for both your initial and final, reduced models
  + Residual plots for your final model(s)