Data Management Culminating Project: Statistical Analysis

The following outlines our expectation of what is to be handed in. Your due date is **April**. No extensions are available. Making efficient use of class time is essential.

Formatting

- Use Times New Roman or Arial fonts size 11/12 with 1½ or 2 spacing
- Make sure that everything is in the correct order and that there are no obvious errors – you will be marked for organization, layout, spelling and grammar
- Make sure all graphs/charts are properly labelled – do not squish graphs
- Graphs that are being compared should have the same scale
- Should be a minimum of 7 pages (including title page and graphs, etc.)

Content

1. **Title Page**

2. **Abstract/synopsis**
   This is a brief description of your whole project with a small summary of your results
   .
3. **Introduction**
   This is a one or two paragraph brief introduction including your idea/hypothesis and where it came from.

4. **Plan and methodology**
   This explains in detail exactly how you carried out your data collection (method of sampling), what data you collected and what you hoped it would show. Specifically you should describe how your data will answer your question. You should also highlight any difficulties and problems you had finding the data.

5. **Critical Analysis (minimum 300 words)**
   Describe in detail everything you have done to the data to prove your hypothesis. Include only the most relevant parts of your analysis. DO NOT include anything that is not worth discussing. Avoid repetitious graphs; choose a selection that will display the data to best support your comments. Include others in appendices and refer to them in their location. DO NOT be overly concerned if your hypothesis is not proved; rather, try to explain why this may have happened.
   Your critical analysis should at least include:
   (a) Printouts from Fathom
   - *Histograms of both your independent and dependent variables, including mean and standard deviation calculations on the graph.*
   - *Box and Whisker Plots of both your independent and dependent variables, including Q_1, Q_3, and the median.*
   - *A Scatterplot including the Line of Best Fit and r value.*

   (b) The following information
   - *presentation of your results for measures of central tendency and spread*
   - *your r value and a conclusion as to the type of correlation that exists*
   - *a discussion as to whether the r value and correlation are consistent with your hypothesis*
   - *the causal relationship that exists (using correct terminology) and justification of your choice*
   - *discussion of the effect of any outliers in you data*
   - *discussion/analysis of extraneous variables*

6. **Conclusions**
   Use your data analysis to answer your original question. If you have inconclusive evidence, explain why and suggest what you would do to develop your research further. Discuss whether your results appear biased in some way try to explain why. Include an evaluation of your techniques (would you do anything differently next time).

7. **Appendices**
   This section should include all the large tables/charts of analysis that were not appropriate to include in the main body of the report. Label each with a letter and refer to it in the main text if necessary. Your data should be referenced here with websites etc.

Note: Make use of your textbook. We have already completed all of the tasks outlined above at one time or another in this course.