Marking scheme for the STA303/1002 project

Due April 07, 2022, at 3:03 p.m. ET for 2% point bonus, OR April 11, 2022 at 3:03 p.m. ET.
Worth 15% of final STA303 grade.
Submit on **MarkUs**

Overview

Component	Percentage of project
Basic checklist	5%
Executive summary	35%
Technical report	50%
Code appendix	5%
Ethics statement	2%
Reproducibility	3%

Basic checklist – 5%

Item	Tick?	Relative weight
Cover page		
Cover page exists as a separate first page AND is professional in appearance (can use provided template for full marks).		1
Title and subtitle together provide an informative introduction to the report.		2
Clear from cover page **for** which company the report has been prepared.		1
Clear from cover page **by** which company the report has been prepared.		1
An appropriate submission date is included.		1
Overall		
Visualizations/tables have captions appropriately numbered and located (viz – bottom, table – top)		2
No code or raw code output (e.g., anything fixed width font) is included.		2
Variables/data are described in words, not using variable names (e.g., please no "the mean of `promotion_var_1` is …") in text, tables and visualizations.		2
Executive summary		
No more than two pages		2
Consultant profiles		
Every team member has a short professional biography. It can be totally fictious, it just needs to sound reasonable.		2
References		
Organized clearly and includes relevant data sources, packages and any external articles/resources. Appropriate in-text citations exist for each resource also.		2
Code		
Code is commented and clearly organized.		2
Submission		
A PDF report, created from Knitting to PDF from R Markdown is submitted.		1
The following files and folders are submitted: - sta303-final-project.Rmd - sta303-final-project.pdf - data-prep.Rmd - data (a folder with Rds files in it)		1

Executive summary (35%)

	Poor/Missing	Basic (5–6)	Proficient (7–8)	Outstanding (9–10)
Structure	No or limited use of headings. Structure is not easy to follow.	Some organization under sections and headings but lacks some clarity. Executive summary is somewhat difficult to scan, lack of emphasis on key information.	Some use of headings but may lack clarity/organization. Executive summary is mostly easy to scan and key information is emphasised.	Clear and appropriate headings. Structure is logical and well organized. Executive summary is easy to scan and key information emphasised.
Level of detail appropriate for audience	Overall, the level of detail is too high or too low for the intended audience.	Level of detail is too low or too high for the intended audience in some sections	Level of detail is mostly appropriate for the intended audience	Appropriate level of detail for the intended audience
Clarity and appropriateness: Visualizations and tables	Tables / figures are inappropriate for the context.	Some tables/figures included are not directly relevant to answering the research question(s), and do not effectively communicate the story of the results as they relate to the research question(s). Too many visualisations / tables included that are not directly relevant and/or should have been further curated.	Tables and figures are relevant, but design could be improved to better communicate results as they relate to the research question. Choices of visualisations / tables may need to have been curated further.	Tables and figures are relevant, well-constructed, and communicate the results clearly (in the context of the research questions). They are appropriately chosen for the audience of the executive summary. With only the most important and interesting visualisations and tables included. Use of colour is appropriate and consistent.
Clarity and appropriateness: Written comments	Results are reported with insufficient information and context AND/OR too much jargon for audience. Too many unimportant or irrelevant results included.	Results are reported but with insufficient information for some important results AND/OR not interpreted in context. Reported results are not connected to the research questions. Language used is inappropriate for the audience. Some unimportant or irrelevant results included.	Size, direction and confidence (as appropriate) of most significant results are given. Context is stated for most results. Relevance to the research questions mostly appropriate.	Size and direction of significant results are given in the context of the research questions. Reported results are clearly related to the research questions. No unimportant or irrelevant results included.
Writing mechanics	Considerable writing and grammatical issues that obscure the meaning AND/OR lots of slang and inappropriate word choice. No bullet points used.	Multiple difficulties in understanding due to word choice, grammar and or sentence construction, but is overall otherwise understandable. Paragraphing / bullet points not well done.	Slight difficulty in understanding one or two sections but overall easy to understand. Mostly appropriate use of bullet points and paragraphing. Writing is in full sentences.	Executive summary can be read and followed with minimal effort. Some grammatical or word choice errors are allowable but must not obstruct meaning. Writing is in full sentences. Makes appropriate use of bullet points and paragraphing.

Technical Report (50%)

	Poor/Missing	Basic (5–6)	Proficient (7–8)	Outstanding (9–10)
Intro: Background	No introduction.	Limited introduction to the report.	A reasonable introduction is provided, though may lack some clarity.	A succinct introduction to the report is provided that sets the scene for the rest of the report.
Intro: Research questions	Research questions are not present or insufficiently address the full scope of the client's needs.	Research questions are present, but only address some of the areas the client asked for AND/OR lack clarity of scope AND/OR cannot be addressed with the available data.	Research questions mostly satisfy the criteria for excellence, but may not be distinct, or not fully address the client's needs or are only have a mostly clear scope.	Appropriate, distinct research questions have been developed that: - fully address the client's needs, - have a clear and specific scope, - are able to be addressed with the available data.
Methods: Data description and wrangling	Data not adequately described. No evidence of meaningful data wrangling.	Basic description of the data and basic data wrangling choices. Lacking completeness and/or justification.	Reasonable descriptions of data and appropriate data wrangling choices. Justifications are present, though may lack some clarity or completeness.	Data is described sufficiently to provide context for the chosen methods. Data wrangling choices are explained clearly, are appropriate for preparing this data, and are well-justified. *This can be as a separate section or discussed as appropriate within a methods section.*
Methods: Purpose	No justification of choice of statistical methods.	Limited explanation of purpose for chosen statistical methods. Justification of choice not persuasive.	Purpose of statistical methods is stated, but some points of confusion remain, lack of clarity of description.	Purpose of each statistical method is clearly explained.
Methods: Appropriateness	Serious issues with methods chosen.	Some methods are appropriate, but others are not appropriate for given research questions and data.	Methods are mostly appropriately chosen to address the research questions with this data. Appropriateness is discussed in terms of model assumptions and context/structure of the data.	Methods are appropriately chosen to address the research questions with this data. *Only methods learned in STA303 (with relevant knowledge from STA302 and hypothesis testing from earlier courses as relevant) are expected.* Appropriateness is discussed in terms of model assumptions and the context/structure of the data.
Methods: Accuracy of description	Statistical methods are not described accurately.	Description of statistical methods is mostly accurate, but the level of detail is too basic or too technical.	Description of statistical methods is accurate, although some points would be confusing to the intended audience.	Statistical methods are described clearly and accurately with the appropriate level of detail.

Results: Clarity and appropriateness of tables and figures	Tables / figures are inappropriate for the context.	Some tables/figures included are not directly relevant to answering the research question(s), and do not effectively communicate the story of the results as they relate to the research question(s)	Tables and figures are relevant, but design could be improved to better communicate results as they relate to the research question(s). Use of colour is mostly appropriate and consistent.	Tables and figures are relevant, well-constructed, and communicate the results clearly (in the context of the research questions). Use of colour is appropriate and consistent.
Results: Clarity and appropriateness of written comments	Results are reported with insufficient information and context.	Results are reported but with insufficient information for some important results AND/OR not interpreted in context. Reported results are not connected to the research questions.	Size, direction and confidence (as appropriate) of most significant results are given. Context is stated for most results. Mostly relevant to the research questions.	Size, direction and confidence (as appropriate) of significant results are given in the context of the research questions. Reported results are clearly related to the research questions.
Results: Accuracy	There are serious errors in the accuracy and interpretation of results.	There are several errors in the interpretation of statistical tools/measures.	P-values, confidence intervals and any other statistical tools/measures are mostly accurately communicated.	P-values, confidence intervals and any other statistical tools/measures are accurately communicated.
Results: Research questions are answered	Questions unanswered or inadequately supported by results.	Basic answers to research question(s), may be incomplete and/or not adequately supported by results.	Most of the research questions are answered and supported by results of data exploration and statistical analyses.	Research question(s) clearly answered and supported by results of data exploration statistical analyses.
Conclusion/ discussion	Results of data exploration and statistical analyses are not summarized.	Basic summary results, but not complete.	Results of data exploration and statistical analyses are summarized, although some points are not clear.	Results of data exploration and statistical analyses are well summarized with an appropriate level of detail.
Limitations/ concerns	No discussion of limitations. No suggestions for future consideration.	Limited discussion of limitations. Limited suggestions for future considerations.	Good discussion of limitations and/or concerns related to the data and/or methods. Description of suggestions for future consideration.	Limitations and/or concerns related to the data and/or methods are thoughtfully and clearly described. Suggestions for future consideration sensible and clearly stated.
Professionalism	Inappropriate language/tone used. Client's needs/context not addressed.	Issues with appropriateness and/or consideration of client and/or professionalism of the tone.	Demographic/identity variables are *mostly* handled with / communicated about with appropriate professionalism. Findings are *mostly* delivered truthfully and with awareness of client's needs/context. Report has *mostly* appropriate tone.	Demographic/identity variables are handled with /communicated about with appropriate professionalism. Findings are delivered truthfully and with awareness of client's needs/context. Report has appropriate tone.
Organization	Order of ideas is very difficult to follow.	Order of ideas is somewhat difficult to follow.	Order of ideas is logical, although some parts could be organized more effectively to improve flow.	Ideas presented in a very logical order, which is easy to follow.

Sectioning and completeness	Sectioning not present AND/OR multiple sections missing.	Report is missing at least one section OR no section headings are used.	Report is sectioned mostly appropriately but could be section titles could have been better labelled/ordered more appropriately. Report is complete.	Report includes clearly and appropriately labelled sections that are well organized. A complete report should have sections that introduce the report, explain methods, present results, and discuss conclusions, limitations and next steps as appropriate.
Level of detail for audience	Overall, the level of detail is too high or too low for the intended audience.	Level of detail is too low or too high for the intended audience in some sections.	Level of detail is mostly appropriate for the intended audience.	Appropriate level of detail for an audience with statistical knowledge, though not necessarily R users.
Writing mechanics	Considerable writing and grammatical issues that obscure the meaning AND/OR lots of slang and inappropriate word choice. No bullet points used.	Multiple difficulties in understanding due to word choice, grammar and or sentence construction, but is overall otherwise understandable. Paragraphing / bullet points not well done.	Slight difficulty in understanding one or two sections but overall easy to understand. Mostly appropriate use of bullet points and paragraphing. Writing is in full sentences.	Technical report can be read and followed with minimal effort. Some grammatical or word choice errors are allowable but must not obstruct meaning. Writing is in full sentences. Makes appropriate use of bullet points and paragraphing.

Code appendix (2%)

	Poor/Missing	Basic (5–6)	Proficient (7–8)	Outstanding (9–10)
Code appendix	Incomplete or inappropriate.	Web scraping, API access and accessing the licensed postcode conversion are all discussed with limited clarity/completeness OR one of the topics is missing or has significant errors but the other two are at proficient level.	Web scraping, API access and accessing the licensed postcode conversion data are described mostly clearly and correctly in writing (minor mistakes/omissions) AND ethical professional practice considerations are highlighted in each case. Appropriate code is included and commented clearly.	Web scraping, API access and accessing the licensed postcode conversion data are described clearly and correctly in writing AND ethical professional practice considerations are discussed, demonstrating clear understanding, in each case. Appropriate code is included and commented clearly.

Ethics statement (2%)

	Poor/Missing	Basic (5–6)	Proficient (7–8)	Outstanding (9–10)
Ethics statement	Incomplete or inappropriate. *(Be very careful about your academic integrity. Don't plagiarize an ethics statement…)*	Two sensible statements were made but they lack clarity, appropriateness or appropriately professional language.	Only two statements about ethical statistical consulting are at the outstanding level OR there are three proficient statements but they lack clarity or appropriateness or professional language.	At least three relevant statements about ethical statistical consulting are made. These are appropriately in line with professional conduct advice and are clearly stated and use professional language.

Reproducibility (3%)

	Poor/Missing	(1)	(2)	(3)
Reproducibility	Unable to reproduce PDF from Rmd even with significant revisions OR .Rmd not submitted.	Several revisions required to reproduce PDF from Rmd. This may be caused by not including external images or all data manipulation code.	Only minor revisions are required to recreate PDF from Rmd.	No revisions required to run code to produce final PDF report. (*This doesn't include installing libraries, but all libraries used should be clearly introduced in the code.*)