ST344: Professional Practice of Data Analysis

Word, figure and table limits for assignments, 2023-2024:
Definitions and penalties for failing to adhere to limits

Background

Please ensure that all work submitted as a pdf includes your student or group number in the heading.

Professional reports and publications

Professional reports and scientific publications are required to respect certain limits. If a company director or a minister of state asks for a concise 500 word summary, they do not expect to receive 600 words. In most real scenarios (such as writing grant applications, abstracts for conferences, reports and scientific papers for journals) such limits have to be taken seriously.

Some Statistical scientific articles have quite relaxed requirements. For example, for the Journals of the Royal Statistical Society "*Manuscripts that are longer than 24 journal pages are unlikely to be accepted for publication.*"
(https://rss.onlinelibrary.wiley.com/hub/journal/14679868/author-guidelines). The Journal of the American Statistical Association is also restricting page number but submissions have to follow their journal page template and any submission exceeding the page number cannot be uploaded for submission.

Most scientific journals have very detailed requirements. For example the Lancet states in their instructions for authors (https://www.thelancet.com/preparing-your-manuscript):
"*Be up to 3500 words (4500 for randomised controlled trials) with 30 references (the word count is for the manuscript text only)*"
"*When reporting Kaplan-Meier survival data, at each timepoint, authors must include numbers at risk, and are encouraged to include the number of censored patients.*" UK Research councils (UKRI) formulate rigorous requirements on grant applications and restrictions on the page number, with an exact specification of the margin size and the typesetting allowed. Papers exceeding these limits cannot be uploaded. We have encountered other funders who ask their applicants to fill in online forms with restrictions on the word count for each question. Error messages are issued in real time if the answer exceeds the given word limit and the submission cannot progress until the text provided is within the word limit.

Assignments for this module will be marked with strict attention to limits, to allow you to practice achieving a high standard of information and brevity. Assessment will be based on the presentation of your reports, your understanding, the competence of your analysis and discussion of the data or articles and your critical evaluation of your work.

Limits and penalties

Words

For each assignment you will given a word limit and **you must provide a word count** which does not include any references, or the required statements of the number of words. The word limit also does not include any words used in figures, tables, equations or R-script lines.

If you do not provide a word count, the count made by the examiner will be final.

The word limit is an upper limit, and there will be no penalty if you use fewer words. However, it is likely that you will need at least 80% of the words in order to provide a good report. The following penalties will apply to your mark if you exceed the word limit, X by using Y words. Your mark M, will be multiplied by the factor, f.

Proportional excess	Penalty factor (f)
Y/X < 1.01	1
$1.01 \leq Y/X < 1.02$	0.95
$1.02 \leq Y/X < 1.03$	0.90
$1.03 \leq Y/X < 1.04$	0.80
$1.04 \leq Y/X < 1.05$	0.70
$1.05 \leq Y/X < 1.10$	0.50
$1.10 \leq Y/X < 1.15$	0.25
$1.15 \leq Y/X$	0

For example, if the work limit is 500 words, and you use 529 words, Y/X=1.058. If your mark was 80% (of the total number of marks allowed), the recorded mark would be 40%.

Tables and Figures

Tables provide numerical information. Tables can have two related sections, such as summaries of two fitted models which are discussed in the text. Tables should usually fit on a single page.

Figures provide mainly visual information, though can include some numerical information. A figure with multiple subplots will count as a single figure if there are no more than five rows and four columns. As legends must be legible, do not try to evade the limit on figures by using excessive subplots.

The penalties for exceeding the limit on tables and figures will be that they will be ignored in marking: no credit will be given. For example, if there are marks for diagnostic plots, three figures are allowed, and your figure with diagnostic plots is Figure 4, you will get zero marks for diagnostic plots.

Peer marking

Reading and commenting on a fellow student's work can help you to refine your ideas and improve your own work. You will hand in reports on the first five weeks' practical sessions. For the practical reports in weeks 2 and 4, you should exchange a report with another student. You should mark the work and provide constructive feedback

R-script lines

For lab reports, you will generally be expected to submit your R-scripts or RMarkdown files. It is important that your code is well-documented by using explanatory comments and systematically laid out (for example, by using a pretty-printing utility). Please ensure that you delete any redundant code and do not include any confidential information. There is no limit on the number of lines of code but you should aim for your code to be concise and no longer than necessary. No credit will be given in marking if the marker has to make undue efforts to try to decipher your code.

Another reason for limits

The limits are intended not only to help you to spend an appropriate amount of time on the work (by not spending too much time on writing a lot or preparing too many figures and tables) but most importantly you will see that this improves your overall scientific writing style as it will train you to be focused and precise and to spot and get rid of redundancies.

The module leaders (Jane L Hutton, B F Finkenstadt, C Viaro).