

# STA303/1002: Final project feedback for The Hive

## Winter 2021

### Contents

<b>General feedback (whole class)</b>	<b>2</b>
<b>Rubric (personalized)</b>	<b>4</b>
Reminder of section weights and rubric . . . . .	4
Basic checklist . . . . .	4
Executive Summary . . . . .	5
Technical Report . . . . .	7
Ethics Statement . . . . .	10
Reproducibility . . . . .	10
<b>Overall comparison</b>	<b>11</b>
Brief comments (if any) . . . . .	11
<b>Regrade requests instructions</b>	<b>12</b>
This report provides a mix of general commentary and specific numeric feedback on your final project.	

## General feedback (whole class)

- There tended to be some quite obvious differences between projects where folks appeared to have read the rubric and those where they hadn't. In many cases this resulted in the loss of small points from not following the 'basic checklist' and including raw code outputs. In more important cases, this meant that a limitations section was not included in the discussion.
- For rubric items ranked on a 10 point scales, 9s and 10s were given for outstanding work, while 7s and 8s were the most common and given for good, competent work.
  - One way to think about this: if you were doing an internship and were submitting this to your manager/supervisor, a 7 or 8 means they'd be pretty happy with your work, but have some revisions, while a 9 or 10 would mean you'd really impressed them and they might only have a few minor changes. A 5 or 6 might be akin to them thinking you made a reasonable start, but they'd want to sit down to review and improve a fair amount with you.
- Higher scoring projects:
  - addressed hiring, promotion and salary in their research questions.
  - discovered that in the hiring phases, most ratings were similar between genders, except for leadership presence and speaking skills, where women tended to be rated lower. High scoring projects discussed how this may be due to bias introduced from training data for the AI.
  - identified disparities in salary, after controlling for team/seniority, as well as in promotion, with women earning less and being promoted less often.
  - noted as a limitation that data was only for currently employed employees and that this could result in a form of selection bias. Strong reports specifically noted that people feeling they are not fairly being remunerated/promoted may be more likely to leave a company and made recommendations about future data collection/availability.
  - noted that gender was the only metric on which to explore issues of bias and that future work could consider race/ethnicity and/or other identity features.
  - were carefully proofread.
  - made use of headings to guide the reader, especially in the executive summary.
  - had well-constructed and easy to read tables and visualizations that were relevant.
  - showed careful selection of which figures and tables to include, or modify and include, in the executive summary.
  - the purpose of analyses was made clear and what was included was carefully curated to tell a clear story of the data.
- Lower scoring projects:
  - had too many tables or figures. More careful thought of what to include would have helped.
  - had vague or unclear research questions.
  - did not justify the choice of the models/variables used.
  - used random effects for gender (inappropriate) and/or failed to use employee as a random effect when it would be appropriate.
  - made mistakes with the interpretation of model coefficients. Most often an issue with logistic regression.
  - were challenging to reproduce due to not including the full code, editing the data outside of R, using local paths, not including libraries for packages used. (It is okay if we had to install the package, but you still needed to load it explicitly in the code.)

- included figures for which the code was not provided.
- didn't provide captions for figures and tables, or didn't use tables and figures at all.
- included unnecessary R-specific details (code, discussion of wrangling, etc.). We discussed that the technical report was meant for other statistically literate people, but that they were *not* necessarily R users.
- didn't report and/or **interpret** model results in context. Some did not report estimates at all and did not report confidence intervals where they would have been possible to produce. In some cases p-values were not reported at all, just whether or not they were less than 0.05.
- included too much technical detail in the executive summary that was not appropriate for the audience. In some cases the Executive Summary was treated more as an introduction and so did not include results and limitations. Executive summaries should be able to be read as stand-alone summaries of the report, for an 'executive' audience.
- were hard to follow, often due to organization issues. Headings were either not used or present but not informative for navigating the report.
- included raw R output.
- didn't round numbers sensibly.
- included z-values in tables where just estimates and confidence intervals would be much more appropriate.
- were inconsistent in how figures were presented, e.g. changing gender colours for no reason because default settings were used.
- had figures that were hard to read due to 'too much going on' and/or small text.
- used variable names as they appear in code (i.e. with underscores) in the report instead of making them 'human friendly'. In many cases, simply reading the rubric and changing `leadership_for_level` to "rating of leadership, for level" and then just calling it "leadership for level" thereafter would have made things much nicer for the reader.
- didn't explain data wrangling choices, e.g., when to include/not include those who had answered 'prefer not to say' for gender.
- had word-choice issues. This can sometimes be due to trying to sound more professional by using a thesaurus to find a 'smart word' instead of a simpler one. If the simpler word is correct, always go with that instead. Making your reader's life easier makes you look 'smarter' than a fancy word will, especially when it has odd or inappropriate connotations for the context.
- missed some of the areas the client asked them to investigate and/or entire sections of the report.

## Rubric (personalized)

### Reminder of section weights and rubric

Recall from [the rubric](#) the following section weights. Subsections within these sections also have relative weightings as you'll see below.

Component	Percentage of project
Basic checklist	5%
Executive summary	35%
Technical report	50%
Ethics statement	7%
Reproducibility	3%

### Basic checklist

Rubric item	Ticked?	Weight
A PDF report, created from Knitting to PDF from R Markdown is submitted.	Yes	1
An appropriate submission date is included.	Yes	1
An Rmd file is submitted.	Yes	1
Clear from cover page by which company the report has been prepared.	Yes	1
Clear from cover page for which company the report has been prepared.	Yes	1
Code is commented and clearly organized.	Yes	2
Cover page exists as a separate first page AND is professional in appearance (can use provided template for full marks).	Yes	1
Every team member has a short professional biography. It can be totally fictitious, it just needs to sound reasonable.	Yes	2
Executive summary no more than two pages.	Yes	2
No code or raw code output (e.g., anything fixed width font) is included.	Yes	2
Title and subtitle together provide an informative introduction to the report.	Yes	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.	Yes	2
Visualizations/tables have captions appropriately numbered and located (viz – bottom, table – top)	Yes	2

Score on basic checklist: 20/20

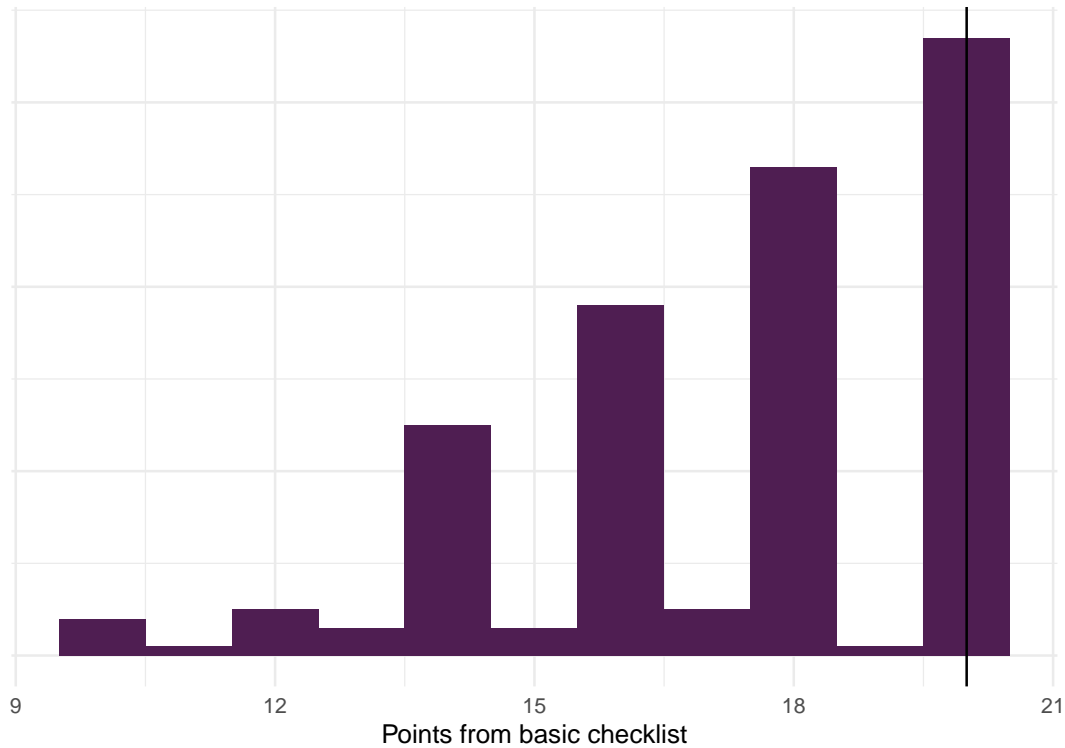


Figure 1: Distribution of overall scores on the basic checklist. Your score is shown as a black vertical line.

## Executive Summary

Rubric item	Score	Weight
Structure	10	0.2
Level of detail appropriate for audience	9	0.2
Clarity and appropriateness visualizations and tables	7	0.3
Clarity and appropriateness written comments	10	0.2
Writing mechanics	9	0.1

Score on executive summary: 88%.

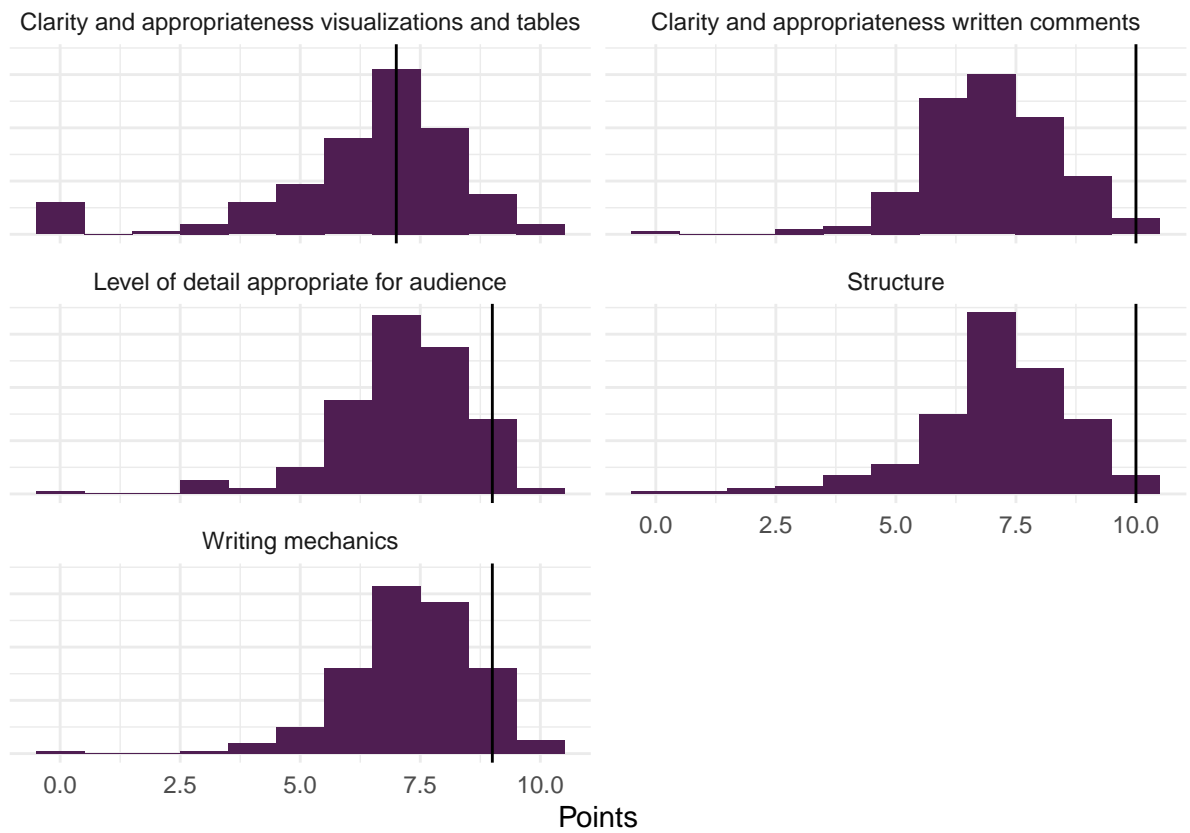


Figure 2: Distribution of points by rubric item for the executive summary. Your score is indicated by the black line.

## Technical Report

Rubric item	Score	Weight
Introduction: background	8	0.08
Introduction: research questions	10	0.03
Methods: data description and wrangling	9	0.05
Methods: purpose	7	0.15
Methods: appropriateness	9	0.05
Methods: accuracy of description	9	0.08
Results: clarity and appropriateness tables and figures	7	0.15
Results: clarity and appropriateness written comments	9	0.05
Results: accuracy	8	0.08
Results: research questions are answered	9	0.03
Conclusion and discussion: conclusion/discussion	9	0.05
Conclusion and discussion: limitations/concerns	9	0.05
Structure and organization organization	9	0.05
Structure and organization sectioning and completeness	8	0.05
Structure and organization level of detail for audience	8	0.05

**Score on technical report: 81.7%.**

See Figures 3 and 4 on the following pages for comparative summaries.

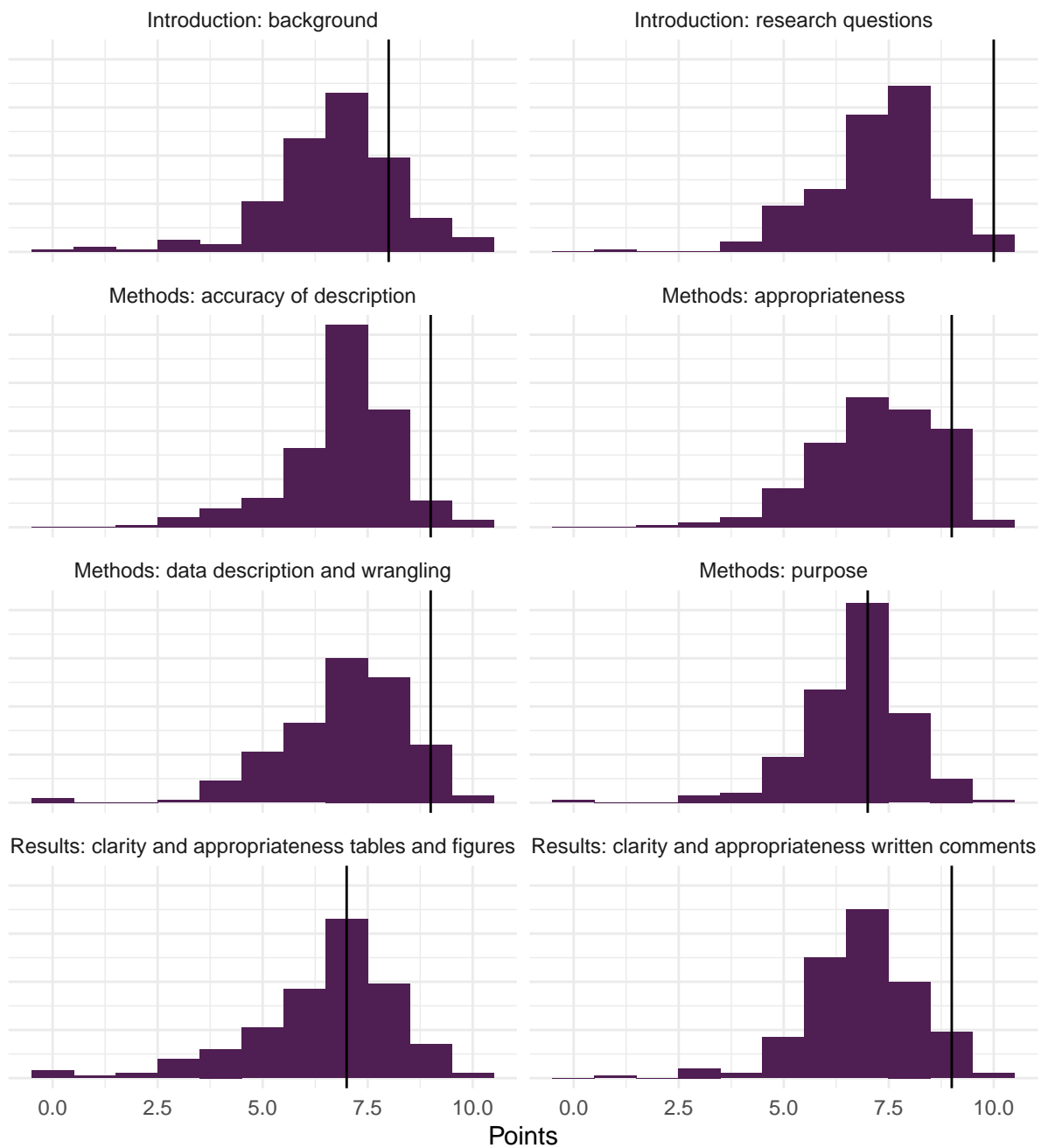


Figure 3: Distribution of points by rubric item for the technical report (part 1). Your score is indicated by the black line.



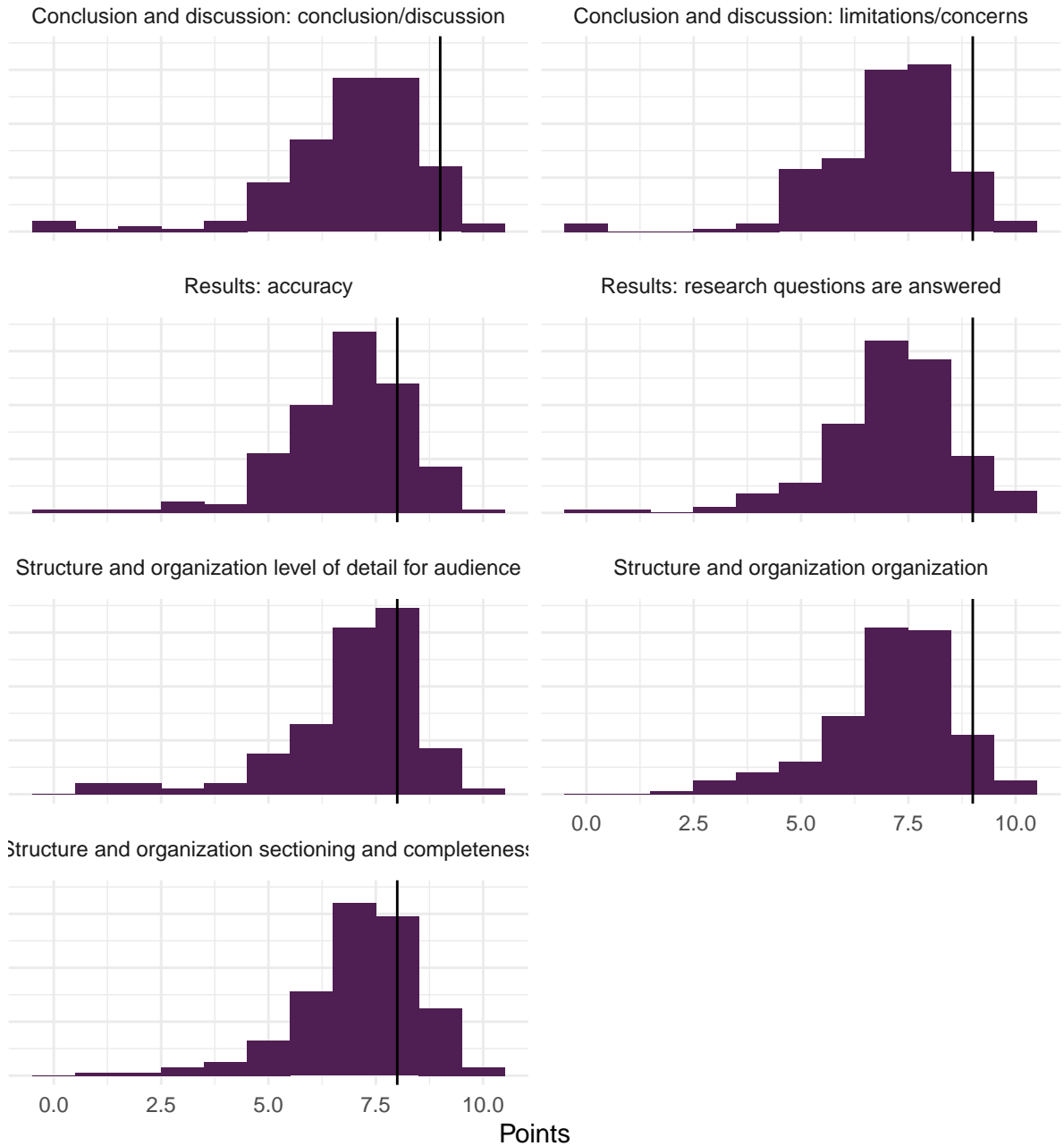


Figure 4: Distribution of points by rubric item for the technical report (part 2). Your score is indicated by the black line.

## Ethics Statement

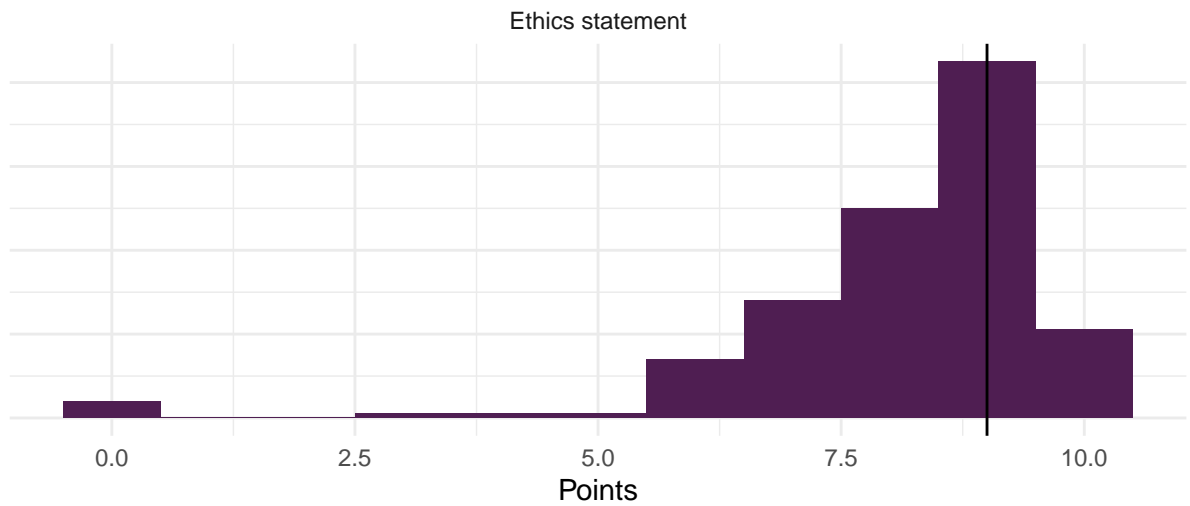


Figure 5: Distribution of points for the ethics statement. Your score is indicated by the black line.

Score on ethics statement: 90%.

## Reproducibility

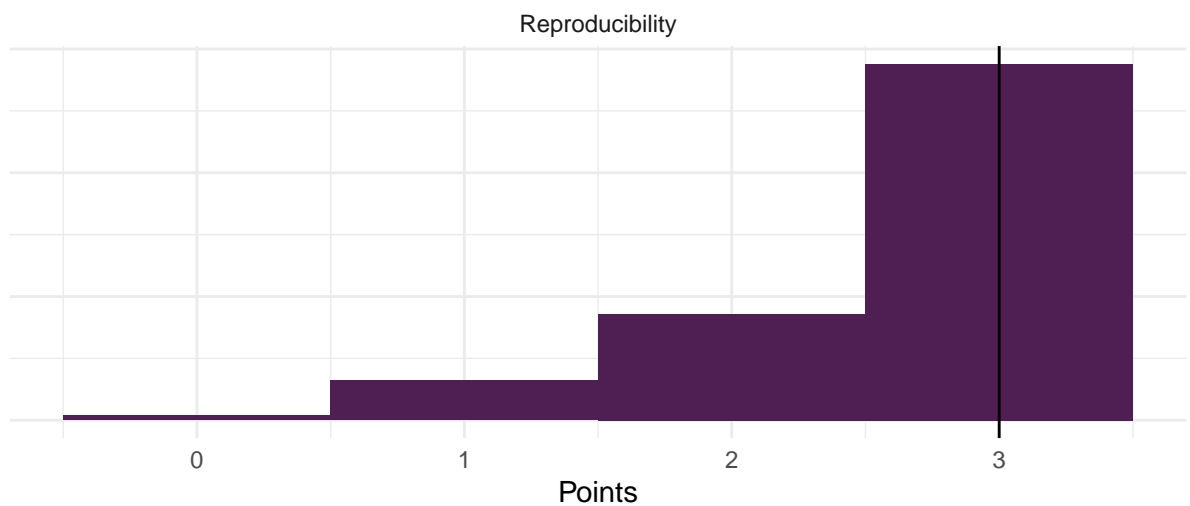


Figure 6: Distribution of points for reproducibility. Your score is indicated by the black line.

Score on basic reproducibility: 3 out of 3 points.

## Overall comparison

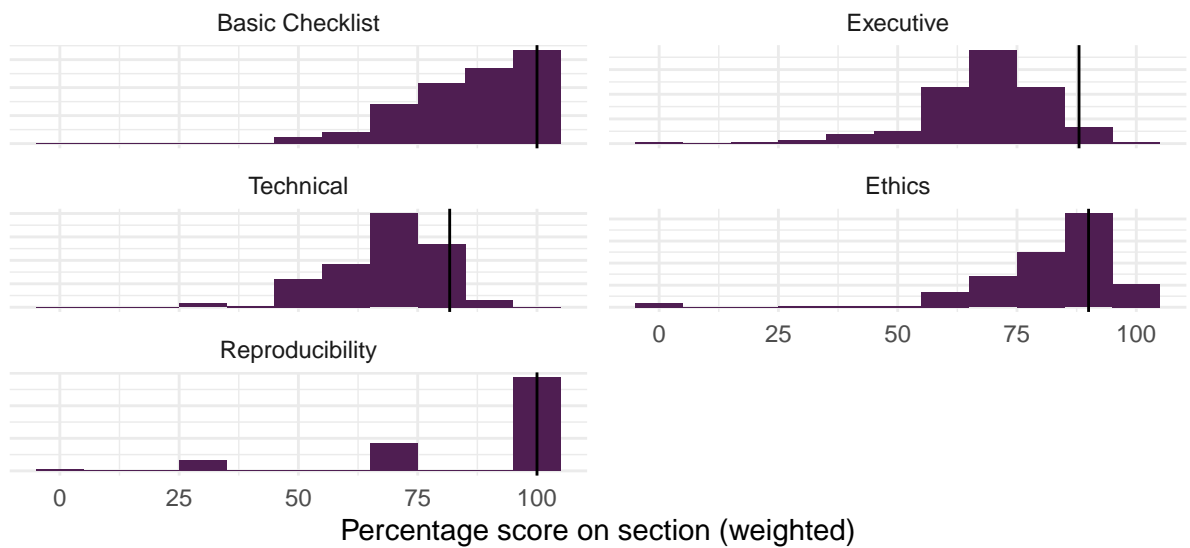


Figure 7: Weighted score on each section. Your section score indicated by the black vertical line.

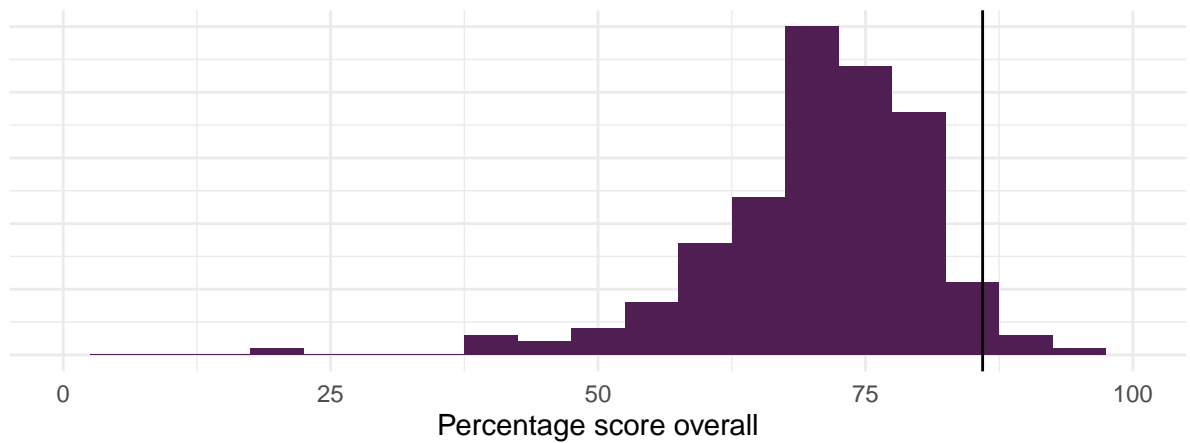


Figure 8: Distribution of final grades on project.

## Brief comments (if any)

A lot of tables, some could be removed and information added in-text; Figure 1 has small text and Figure 4 needs editing.

## Regrade requests instructions

While I know more detailed personalized comments would be desirable, this is what we're able to provide. Note that in a 'normal' semester you'd have to be booking an appointment to view an exam paper with little to no feedback on it, and none of this comparative feedback. As such, a regrade request that says you're sure the mark is wrong but don't have enough information to tell me why, will not be considered, nor can further feedback be provided before you decide.

- For your request to be considered, you must provide detailed and specific written justifications for what you think has been incorrectly graded. Make reference to course materials and/or assessment instructions, rubric, and feedback.
- Every member of the team must complete and sign a **declaration** and then one member will submit on behalf of everyone.
  - Either format is fine, but save your final version as a PDF
    - \* [DOCX template](#)
    - \* [PDF template](#)
- The regrade request must be submitted to [this form](#) by **6:00 p.m. ET on Friday, May 7**. Only one form should be submitted on behalf of the group and must include the justification (that everyone must have read and agreed to) and all of the group's declarations. (If you worked as an individual, you just need to submit one for yourself).