CASE STUDY

Designing an introductory statistics subject for students with diverse educational backgrounds and chosen qualifications

Rupert E. H. Kuveke, Department of Mathematical and Physical Sciences, La Trobe University, Melbourne, Australia. Email: r.kuveke@latrobe.edu.au
Amanda J. Shaker, Department of Mathematical and Physical Sciences, La Trobe University, Melbourne, Australia. Email: a.shaker@latrobe.edu.au
Luke Prendergast, Department of Mathematical and Physical Sciences, La Trobe University, Melbourne, Australia. Email: luke.prendergast@latrobe.edu.au

Abstract

This is a case study on the design of a first-year undergraduate statistics subject at La Trobe University, entitled Making Sense of Data, which is taken by students from various disciplines. To account for students' diverse educational backgrounds and chosen qualifications, this subject is designed such that all students complete core statistics concepts, while a third of the subject contains stream-specific content. This subject design provides students with a solid foundation in statistics, while addressing the demand for a flexible first-year statistics subject which is accessible and relevant for students enrolled in a variety of tertiary degrees. This structure allows for stream-specific lectures, computer lab material, assessments, and even statistical software programs to be used across different streams. The design also incorporates strategies for addressing statistics anxiety within the curriculum. In this paper, we present the outcomes of this subject design in terms of student performance, engagement and satisfaction. We also present iterative and reflective changes that have been made to the subject over time, in response to student and staff feedback, and discuss the impact these changes have had on student outcomes.

Keywords: higher education, statistics anxiety, flexible curriculum, student engagement, undergraduate statistics.

1. Introduction

Statistical literacy is an increasingly vital and necessary skill in today's data-driven world. In recognition of this, a wide variety of higher education courses often contain a compulsory introductory statistics component in their curriculum (see e.g. Tishkovskaya and Lancaster, 2012). While the demand for quality tertiary-level statistics teaching continues to increase, numerous challenges in the teaching and learning of statistical content remain (Bromage et al., 2022). Students undertaking a first-year undergraduate statistics subject (known as a course or module in UK terminology) may have diverse educational backgrounds and chosen tertiary qualifications, and indeed for many students, such a subject may be their only formal tertiary-level exposure to statistics.

As a result, these students can often fail to appreciate the importance or relevance of statistics to their chosen degree(s), and can approach learning statistics with low motivation and negative attitudes (see e.g. Gal and Ginsburg, 1994; Conners, McCown and Roskos-Ewoldsen, 1998). Indeed, learning statistics has been argued to be similar to learning a second language (Lalonde and Gardner, 1993), with all the attendant difficulties. This can be compounded by the fact that large percentages of students can also experience statistics anxiety (Onwuegbuzie and Wilson, 2003), a long-acknowledged form of anxiety (see Cruise, Cash and Bolton, 1985) which is related to, but distinct from, mathematics anxiety (see e.g. Zeidner, 1991, Paechter et al., 2017). Statistics anxiety can lead to a lack of enjoyment of and engagement with the content, reduced knowledge retention, and lower pass rates (see e.g. Macher et al., 2012, González et al., 2016 and Marshall et al., 2022). These poor
attitudes towards learning statistics may in part be promoted by prior negative experiences, with Koparan (2015, p.103) noting that there are “serious problems experienced in teaching and learning statistics” in schools.

Given these challenges, in order for a first-year undergraduate statistics subject to be successful, it must be flexible enough to account for students’ various educational backgrounds, and be able to address and mitigate barriers to learning such as statistics anxiety, while also ensuring students receive a solid foundation in statistical knowledge, regardless of their chosen tertiary qualification(s).

In recognition of these requirements, a new first-year La Trobe University subject entitled Making Sense of Data (with subject code STM1001) was designed during 2021 and first offered in 2022. STM1001 consists of a core statistics module (about two-thirds of the subject), and stream-specific modules (about one-third of the subject) which focus on the use of statistics in specific disciplines. Students complete both the core module, and a stream-specific module which best fits their chosen qualification. This structure ensures that students receive a solid background in foundational statistics concepts, while also learning statistical concepts which more closely align with their chosen qualifications. This novel design structure facilitates content delivery to large cohorts of students from numerous disciplines, while being flexible enough to allow for the learning materials, assessments, and statistical software programs used to differ between streams. Similar approaches that cater for diverse cohorts do exist: for example, by using examples relevant to all disciplines and/or discipline-specific tutorial support (e.g. Hilliam and Calvert, 2019, Calvert et al., 2022). As well as including these elements in its design, STM1001 additionally includes stream-specific lectures, assessment and use of statistical software packages in its design.

In this paper, we present a case study on the creation, design and implementation of STM1001, and explain the pedagogical rationales behind our design processes, decision-making and delivery styles. We then present an evaluation of the subject’s success, in terms of student performance and satisfaction, based on its operation since teaching began in 2022. The design and implementation of STM1001 is an ongoing, dynamic process, and so we also detail some recent modifications made to the subject, in response to student and staff feedback, along with preliminary evaluation of these changes. We conclude by discussing the current state of STM1001, and our future intentions.

2. Background

2.1 Motivation from an institutional perspective

In early 2020, the University sought to reduce duplication of teaching efforts into subjects with overlapping content that were offered either independently within different areas (e.g., different departments and disciplines), or within the same area with content tailored to the needs of students from within different degrees. The key objectives were to (i) create greater efficiency whereby workloads associated with duplication could be reduced, (ii) implement targeted efforts for continuous subject improvements in a single subject to ensure that each student’s educational needs are met and (iii) ensure that subject content was created and delivered by discipline experts.

Led by the Pro Vice-Chancellor Learning & Teaching in the then called College of Science, Health and Engineering, the discipline of statistics was identified as part of this effort. At that time, the Department of Mathematics and Statistics (now the Department of Mathematical and Physical Sciences), offered four introductory statistics subjects: Statistics for Psychology, Statistics for the Life Sciences, Statistical Methods, and Statistical Science. Further, introductory statistics was also taught, either in whole, or in part, in many other departments, e.g., in the health sciences and in business. An initial stakeholder meeting was held in May, 2020, and was attended by the two Pro Vice-Chancellors of Learning & Teaching from the university’s two colleges, the Head of Department of Mathematics and Statistics,
and academic representatives from science and business disciplines. The meeting agreed on two things. Firstly, there was substantial overlap between various introductory statistics subjects across the university. Secondly, that there were discipline-specific needs within many of those subjects.

Concurrently during 2020, the university was undergoing extensive course re-structuring in alignment with a guiding set of principles called Course Architecture. These principles provided structural parameters for all courses that provided greater opportunities for subjects, minors, majors and specialisations to be shared across courses. Course Architecture also included the provision for a subject to be constructed out of two or more modules. For example, a typical single subject is 15 credit points (8 subjects make up a full-time load for one year). According to a set of rules and where it is a suitable to do so, a 15 credit point subject could be made up of modules which collectively contribute to the student’s learning needs.

Following the initial meeting, and in full consideration of Course Architecture, the Department of Mathematics and Statistics proposed a university-wide modularised introductory statistics subject consisting of a common core module for all, as well as stream-specific modules to cater for the different needs of students across different degrees and qualifications. This proposed modularised structure meant that new modules could be created and included in the future to suit new needs (e.g., new qualifications or identified gaps in existing qualifications). This proposal was agreed to by stakeholder disciplines and development started in 2021.

2.2 Core and Stream-specific design

STM1001 content has been designed to ensure that (a) all students develop a strong understanding of foundational statistical concepts which are broadly applicable, while (b) also learning statistical skills which relate more directly to their chosen area(s) of study. As explained in the previous section, this is achieved by the use of modules. All students complete the core module, which accounts for 70% of the subject content and assessment. Students also complete one stream-specific module, which accounts for the remaining 30% of the subject content and assessment. The robust architecture of the streams allows for significant differences between streams, such as the use of different software and assessments, while still being flexible enough to support crossover of content where relevant.

Currently, two stream-specific modules, the Science/Health (SH) module and the Data Science (DS) module, are offered, with the potential for further modules to be added in future. The SH module is for students enrolled in Health Sciences-, Life Sciences-, Applied Life Sciences- and Social Sciences-focused courses, while the DS module is for students enrolled in Statistics-, Mathematics- or Information and Communications (ICT)-focused courses. Initially, the primary software package used for all students was the R software environment for statistical computing and graphics, via the RStudio integrated development environment, in accordance with stakeholder requests. Further discussion on statistical software packages used is provided in section 2.5.

Stream-specific content and software skills are assessed via two stream-specific assignments that students complete throughout the teaching period. These stream-specific assignments contain unique questions relevant to the specific stream, but utilise the same questions and data sets across streams where possible (i.e. where content overlaps between streams, see figure 1), for efficiency. Each stream-specific assignment contributes 15% to a student’s overall mark for STM1001. The remaining 70% of assessments are stream-agnostic and are completed by all students.

Figure 1 outlines the different STM1001 core and stream-specific topics covered throughout the teaching period. The topics included were chosen following extensive consultation with key stakeholders, including program coordinators of Science and Health degrees whose students would be taking STM1001. Following consultation, it became clear that the needs of students in both Health
and Science degrees were very similar, which led to the Science/Health stream being designed as one module rather than two separate modules. The core module covers typical introductory statistics content including descriptive statistics and plots, t-tests (one-sample, paired and independent samples), ANOVAs, linear regression, and chi-squared tests. In the stream-specific modules, SH students focus predominantly on study design, while DS students focus on R coding, data visualisation and machine learning. The specific examples and data sets discussed in each of these streams have been carefully selected, to ensure they are relevant to the students (see e.g. Cornock, 2016). This supports students in developing an appreciation of the benefits of statistics in the context of their chosen tertiary qualification. For stream-specific Topics 7 and 8, students from both streams cover the same material.

![Figure 1. An overview of the current STM1001 core and stream-specific content covered throughout a teaching period.](image)

### 2.3 Statistics Anxiety

Previous research has highlighted that 70% to 90% of students may experience anxiety when confronted with the prospect of learning statistics (Zeidner, 1991; Marshall et al., 2021). When designing STM1001, a concerted effort was made to integrate measures to address and mitigate statistics anxiety. Sessions to discuss and normalise the phenomenon of statistics anxiety are built into the early weeks of the subject curriculum, to help defuse students' concerns, and to highlight that they are not alone in feeling anxiety about learning statistics. Raising awareness of statistics anxiety early in the subject allows students to begin considering methods for tackling this issue, as advised by e.g Marshall et al. (2017). These interventions provide opportunities to encourage and reassure students about their learning (Wilson, 1999), and reinforce the message that anxiety about statistics and mathematics is a valid and common scientific phenomenon (see e.g. Lyons and Beilock, 2012). These interventions also serve as an opportunity to remind students of the support services that are available to them. We provide weekly one-on-one and small group support sessions with experienced tutors, following e.g. Lalayants (2012). In addition, students have access to the Maths Hub (see e.g. Jackson, 2021), a La Trobe University student support initiative which provides free mathematics and statistics support to all La Trobe University students throughout the university’s teaching periods.

A number of other elements recommended in the statistics anxiety literature (e.g. Chew and Dillon, 2014) have been incorporated into the design. These include: less emphasis on mathematical calculations and more emphasis on conceptual understanding and use of software; weekly low-stakes quizzes that have generous time limits and multiple attempts available; humour and enjoyable in-class activities including classroom polling (e.g. Kahoot [https://kahoot.com/] and Mentimeter
immediacy behaviours in class such as using students’ names and being approachable; collecting data in class; and student support provided by personalised emails sent via the Student Relationship Engagement System (SRES, Lui et al., 2017).

STM1001 is taught both in blended and fully online instances. While students in online subjects may experience lower levels of statistics anxiety on average than those learning in face-to-face settings (see e.g. DeVaney, 2010; Marshall, Mahmood et al., 2022), online learning presents its own challenges, and can be an isolating experience. To help reduce statistics anxiety and foster a sense of community and belonging (see Lear, Ansorge and Steckelberg, 2010; O’Shea, Stone and Delahunty, 2015), we host a weekly Zoom session for STM1001 online students. To promote interaction between students and build rapport between staff and students, these sessions are used to review current subject information, discuss content-related issues, and revise recent content in a fun, low-stakes atmosphere using Kahoot quizzes (see Shaker, Hurst and Marshall, 2021; Russell, 2022).

2.4 Open access materials

STM1001 content was designed to maximise interactivity and open access for all students. Learning resources are provided as open access material on RPubs.com, bookdown.org, and YouTube, and can be accessed via links provided on the subject Learning Management System (LMS). Topic readings, lectures, supplementary guides, computer lab content, assignments and the exam were all written in R v4.0.4+, using packages such as bookdown, RMarkdown, xaringan, and exams (Grün and Zeileis, 2009).

There are numerous advantages to providing access to subject materials on these platforms, including the ability to update and re-publish content in real time, and seamless integration of videos into online readings and computer labs, as well as interactive elements (Dunn, Brunton and Farrar, 2022). In addition, the open access availability of materials leads to them having a wider reach, contributing to the public good, as well as the opportunity for readers to provide feedback. For example, since 1 January 2022, the STM1001 reading materials have been accessed by over 50,420 users from 188 countries around the world.

2.5 Iterative changes

We regularly implement iterative changes to STM1001 content and delivery methods, based on teaching experiences and student and staff feedback. In the first semester of teaching, STM1001 used interactive workshops rather than lectures, following a ‘flipped learning’ approach (Bergmann and Sams, 2014). However, feedback suggested students required more structured sessions when learning new statistical concepts, and so in subsequent teaching periods, lectures on both core and stream-specific content have been gradually added to the subject curriculum. Additional short video guides have also been created to support students in learning key statistical concepts. While these changes have meant an increase in weekly contact hours, feedback on these changes has been positive overall. We have balanced this for students by reducing the amount of content in readings and supplementary material over time, and by replacing some written sections of content with short videos.

We have also made adjustments to statistical software packages used over time and in response to student and staff feedback. Originally, SH students learnt R, in accordance with requests from stakeholders. However, despite additional support resources being created for SH students in the second teaching period, feedback and teaching experiences suggested that for this cohort, jamovi would be a more suitable choice of statistical software package. As a result, in 2023 R was replaced by jamovi in the SH stream, with an immediate improvement in student experience observed. The flexible design of the subject has meant that while SH students use jamovi throughout the subject (core
and stream-specific content), the DS students use R throughout the subject, thus optimising the learning experience for all students.

Finally, to support student success, at-risk students are now emailed earlier in the semester so that contact is made and to advise regarding support services they can access. In addition, all students are contacted earlier in the semester advising marks required in remaining assessments to pass the subject. As discussed in section 2.3, all emails are personalised, and sent via SRES (Lui et al., 2017).

2.6 STM1001 Students

The STM1001 cohort is diverse not only in terms of students’ educational backgrounds and chosen qualifications, but also location, being taught at three different campuses as well as online. During the four semesters offered so far, 1754 students have completed the subject, as shown in table 1.

Most students (1546) were in the SH stream, while the DS stream has been taken by 208 students so far. Figure 2 shows a further breakdown of students by discipline group over time. Most notably, the number of students from the Health Sciences increased dramatically starting from Semester 1, 2023, at which time STM1001 became a core subject in the Bachelor of Health Sciences.

Table 1. Overview of STM1001 students by stream.

<table>
<thead>
<tr>
<th>Stream</th>
<th>Semester 1, 2022</th>
<th>Semester 2, 2022</th>
<th>Semester 1, 2023</th>
<th>Semester 2, 2023</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Science</td>
<td>19</td>
<td>62</td>
<td>44</td>
<td>83</td>
<td>208</td>
</tr>
<tr>
<td>Science/Health</td>
<td>126</td>
<td>392</td>
<td>464</td>
<td>564</td>
<td>1546</td>
</tr>
<tr>
<td>Total</td>
<td>145</td>
<td>454</td>
<td>508</td>
<td>647</td>
<td>1754</td>
</tr>
</tbody>
</table>

3. Evaluation

In this section, we provide an evaluation of the subject in terms of the core and stream-specific design, followed by a more general evaluation, particularly with respect to iterative changes that have been made to the subject over time.

3.1 Core and Stream-specific design

A total of 547 comments have been received from institutional Student Feedback on Subject (SFS) surveys carried out over the four teaching periods. 27 of these comments referred directly or indirectly to the core and stream-specific design of the subject. Of these comments, 13 may be considered positive, 13 negative, and one neutral. The main theme arising from the positive comments was that students appreciated learning statistics within a context related to their chosen degree, for example: “I found the content to be interesting, in how it involves stats and science as a whole. I can see how the skills learnt can be applied later on in my course and or career”. One student also commented on the benefit of studying with peers who are studying towards similar degrees: “… it was nice to be grouped with students studying similar degrees to me”. The main themes arising from the negative comments were related to not finding the content to be interesting, as well as disadvantages to having the subject split into two parts. For example, “The two parts of the course for Health students made it feel like doing 2 separate subjects in one”.

22 MSOR Connections 22(3) – journals.gre.ac.uk
Overall, students appear to benefit from the stream-specific design of the subject, as they are more able to see the relevance of the statistical content they are learning to their chosen degrees. Further benefit may also be gained by further integration of core and stream material, or by making the links between the two more explicit.

As discussed in section 2.1, from an institutional perspective, there are clear efficiency gains, since all introductory statistics can be taught within one subject rather than 2-3 subjects or even up to 6 or more. Some of the efficiency gains made by the institution are offset by the increased academic workload required to run a subject that contains multiple streams. Part of this academic workload increase includes the administration load of placing students into streams and stream-specific classes. In addition multiple sets of curriculum and assessment must be maintained for approximately half of the subject, due to the stream-specific modules and the stream-specific software used throughout the subject in both core and stream classes. Overall though, despite the challenges and additional complexities, the flexible nature of the design affords worthwhile benefits.

3.2 Iterative changes

As discussed in section 2.5, the main iterative changes made to STM1001 over time have included replacing R with jamovi for the SH cohort, gradual incorporation of lectures for less reliance on readings and self-guided learning, reduction of content, and earlier intervention for at-risk students.

Figure 3 shows STM1001 pass rates over time. More specifically, figure 3A shows that the overall pass rate for all students increased from Semester 1, 2022 (64%) to Semester 2, 2022 (71%), and then
decreased slightly in Semester 1, 2023 (70%), before increasing again in Semester 2, 2023 (78%). However, as noted previously, STM1001 did not become a core subject in the Bachelor of Health Sciences until Semester 1, 2023, at which time there was a sharp increase in the number of students from this course. Therefore, for a more accurate comparison, the purple line shows the pass rate of the STM1001 cohort excluding students from the Bachelor of Health Sciences. This comparison shows an increase in pass rate over time, starting with 64% in Semester 1, 2022, followed by 71% in Semester 2, 2022, 77% in Semester 1, 2023 and 79% in Semester 2, 2023.

Figure 3B shows the pass rates over time separated by stream, and shows that in general, the pass rate for the DS stream is slightly higher than that of the HS stream.

Figure 3. STM100 pass rates over time.

Figure 4. STM100 student feedback over time. Mean score (out of 5) in response to the question, “Overall I was satisfied with the quality of this subject”.

Figure 4. STM100 student feedback over time. Mean score (out of 5) in response to the question, “Overall I was satisfied with the quality of this subject”.
In terms of student satisfaction, figure 4 shows a general increase over time. The chart displays data from institutional Student Feedback on Subject (SFS) surveys. In particular, it shows the mean score (out of 5) in response to the question, “Overall I was satisfied with the quality of this subject”. In Semester 1, 2022, the mean response was 3.71, followed by 3.78 in Semester 2, 2022, 4.02 in Semester 1, 2023 and 3.97 in Semester 2, 2023.

The evaluation of both student performance and student satisfaction over time seems to indicate that the iterative changes made over time have had a positive impact on the student experience. While it is difficult to assign causation to particular changes, anecdotally, the change from R to jamovi for the SH students from Semester 1, 2023 was significant and resulted in an immediate improvement in the student experience. This has been evidenced by a reduction in email traffic regarding technical issues, less reliance on support services for statistical software-related questions, and 16 student comments from the SFS survey specifically mentioning positive experiences with jamovi. A representative student comment from the SFS survey expresses this sentiment: “I did this subject last year and failed. This year I’ve really enjoyed aspects such as the transition to jamovi for Sci/health students as I had mainly struggled with the production of graphs previously but this is made for better with the new program”.

4. Discussion and Conclusions

STM1001 was designed to meet the demand for a robust yet flexible first-year undergraduate statistics subject which could be taught to students with diverse educational backgrounds and chosen tertiary qualifications. The subject covers both typical introductory statistics material, and stream-specific content, in order to provide students with a solid foundational understanding of statistics, while also contextualising the use of statistics in the students’ chosen qualifications. The novel design of the subject allows for the use of different statistical software, learning materials and assessments across the streams, while also supporting synergistic learning between the core and stream-specific content, whereby students’ learning and adaption of skills in each module is enhanced and contextualised by the learning they have attained in the other.

STM1001 has now been offered for four semesters. Over that time, student performance and satisfaction has generally increased as iterative changes have been made in response to student feedback and teaching experience. Demand for the subject continues to grow, with the STM1001 student cohort in Semester 2 2023 being the largest yet, with a total of 647 active students across the three campuses and online.

We will continue to implement evidence-based, student-centric modifications in order to streamline the subject and to ensure it continues to address the diverse needs of the student cohort and institutional stakeholders.

For future teaching periods, we are considering adjustments including (i) combining the weekly interactive workshops and lectures into an integrated, interactive-style lecture, (ii) increasing the amount of short instructional videos for asynchronous learning, and (iii) merging the weekly 1-hour core and 1-hour stream-specific computer labs into a single combined 2-hour class. We are also planning on modifying assessments to promote greater weighting of early assessments, and to ensure they continue to provide a range of authentic learning experiences. These adjustments are intended to support students’ learning, performance and satisfaction and to help mitigate statistics anxiety, and are based on student and staff feedback, recent literature (see e.g. Cox & Taylor, 2018, Thomas et al., 2019, McArthur, 2023) and in consideration of efficiency gains from an academic workload perspective.

Designing, implementing and delivering a tertiary education subject like STM1001 involves a range of challenges. These include developing one subject with a design that is sufficiently complex and flexible such that it can meet multiple disciplinary and student needs, as well as staffing and workload...
challenges. As such, we offer the following reflections which may prove useful to others embarking on developing a similarly structured subject.

From an institutional perspective, clear and effective communication and discussion with stakeholders is vital to ensure the process and end product align with stakeholder expectations and the overarching strategic vision of the university. Staffing requirements must be considered, to ensure academic workloads are feasible and reasonable for subject and instance coordinators who face the added load required to run a subject that contains multiple streams. Training, support and guidance for large teaching teams within the subject is also key to the subject running successfully. Members of the teaching team may be predisposed to teaching in one specific stream, and this must be taken into account before each period of teaching to ensure staffing numbers are adequate. Assessments must be designed to effectively meet the subject intended learning outcomes and be relevant and engaging to students from a variety of disciplines, while also being scalable and robust to academic integrity violations. Fundamentally, careful planning of the overall subject design, creation and presentation of content, and day-to-day teaching that acknowledges and accounts for complex requirements such as those discussed here, can lead to a subject that successfully caters to students with diverse educational backgrounds and chosen tertiary qualifications.

5. References


Mathematical Education in Science and Technology, 53(3), pp.582-593. 
[https://doi.org/10.1080/0020739X.2021.1983051](https://doi.org/10.1080/0020739X.2021.1983051)

[https://doi.org/10.1080/10691898.1994.11910471](https://doi.org/10.1080/10691898.1994.11910471)

[https://doi.org/10.1016/j.lindif.2015.12.019](https://doi.org/10.1016/j.lindif.2015.12.019)


O'Shea, S., Stone, C. and Delahunty, J. (2015). “I ‘feel’ like I am at university even though I am online." Exploring how students narrate their engagement with higher education institutions in an online learning environment. Distance Education, 36(1), pp.41-58. https://doi.org/10.1080/01587919.2015.1019970


