

# COVID Survey 2

Christopher Salahub

12/01/2021

## Introduction

### Preface

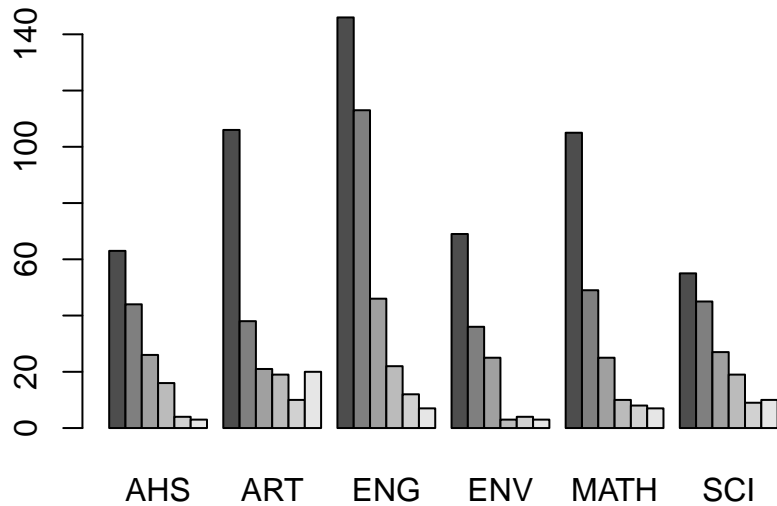
In order to ease the interpretation of results, a number of simple graphics are used repetitively throughout this report. The goal of this uniformity is not to bore the reader, but rather to make the effort of understanding somewhat simpler. The two main graphics repeated are the bar chart, which is hopefully familiar to the reader, and the eikosogram or mosaic plot, which may not be.

Both of these plots use a simple concept to convey information: that of differing lengths of aligned objects. Consider the construction of a bar chart, in which a series of bars of the same width are displayed side by side. The simplest manifestation of such a plot will have these bars equally spaced apart and labeled. The height of each bar and a scale provide the means for a viewer to interpret the value that the heights correspond to. Differences in spacing, placing some bars adjacent and others farther apart, provide a means of grouping the bars and suggesting which should be more directly compared.

An eikosogram uses lengths in a similar way to split an area into rectangles which have areas proportional to counts. Consider a contingency table, that is a table of counts of the form

##		Faculty					
##	Year	AHS	ART	ENG	ENV	MATH	SCI
##	1	63	106	146	69	105	55
##	2	44	38	113	36	49	45
##	3	26	21	46	25	25	27
##	4	16	19	22	3	10	19
##	5	4	10	12	4	8	9
##	6 or more	3	20	7	3	7	10

Suppose we would like to represent this using a similar concept to a bar chart. We could simply display all of these counts as bars beside each other, perhaps grouping by faculty, giving a plot like the following:



This is a perfectly good way to show the counts of year and faculty combinations, but is not very useful if we are interested in conditional questions. Rather than an interest in the raw values of each entry in the table, we might be interested in their conditional relationship. That is, we might be interested in the proportion of people within a faculty in a given year in their program. The bar chart above does not make doing so easy. An eikosogram aims to fix this.

Consider adding some marginal counts to the table from earlier. That is, consider adding the sum of values in the rows or columns to the table, like so

##	Year	Faculty					
##		AHS	ART	ENG	ENV	MATH	SCI
##	1	63	106	146	69	105	55
##	2	44	38	113	36	49	45
##	3	26	21	46	25	25	27
##	4	16	19	22	3	10	19
##	5	4	10	12	4	8	9
##	6 or more	3	20	7	3	7	10
##	Sum	156	214	346	140	204	165

This final “Sum” row gives us the count of people in each faculty. In order to get the proportion of each year in each faculty, we need to divide each count in each column by the sum of that column. In the first column, for example, we would go from 63, 44, 26, 16, 4, 3 to  $63/156$ ,  $44/156$ ,  $26/156$ ,  $16/156$ ,  $4/156$ ,  $3/156$ . By taking these fractions, we would obtain

##	Year	Faculty					
##		AHS	ART	ENG	ENV	MATH	SCI
##	1	0.40	0.50	0.42	0.49	0.51	0.33
##	2	0.28	0.18	0.33	0.26	0.24	0.27
##	3	0.17	0.10	0.13	0.18	0.12	0.16
##	4	0.10	0.09	0.06	0.02	0.05	0.12
##	5	0.03	0.05	0.03	0.03	0.04	0.05

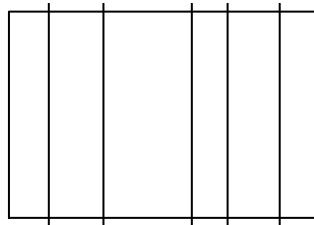
## 6 or more 0.02 0.09 0.02 0.02 0.03 0.06

Here, we can see exactly what we want, the proportions of each faculty in each year. Notice, however, that we've lost the sum column! There's no good way to keep that column there and show the relative sizes of the faculties while keeping these as simple numbers. We need to be a bit more creative.

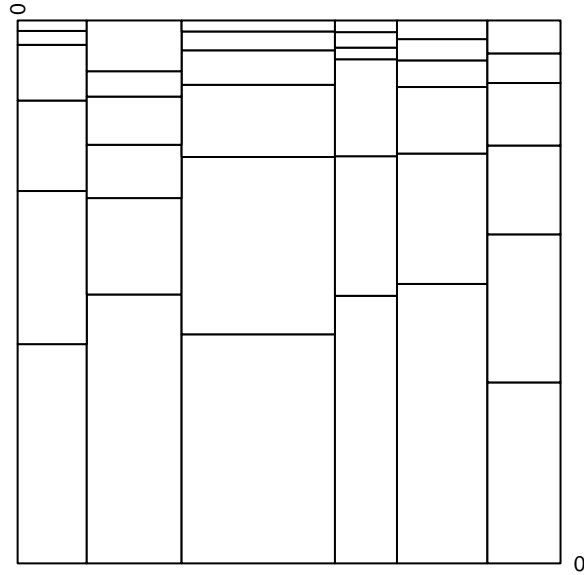
So let's start with a blank box.



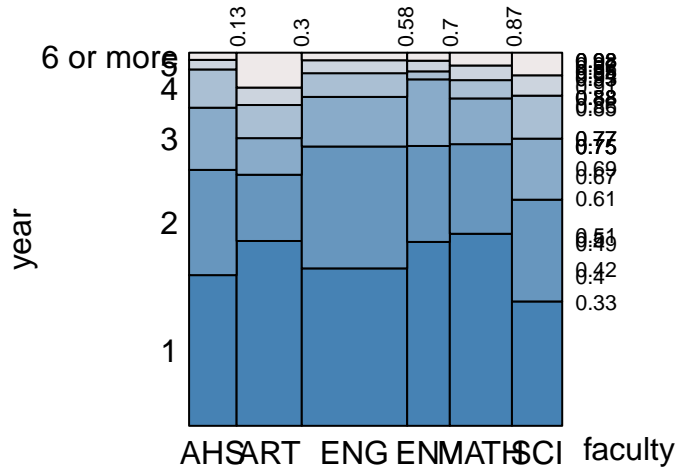
In order to show that "Sum" column, let's use lengths just like in the bar chart, but let's place them end on end along the bottom of this box and separate them with lines along the entire box. That is, let's divide the box into strips with widths that correspond to the relative sizes of each faculty. This gives us



We can now imagine that each of these strips represents a faculty, and consider repeating this process again. We can divide each strip vertically using the proportions from the earlier table to place our lines. In this way, the heights of each segment will be proportional to the corresponding value in the table, and the widths will show that "Sum" row so we don't lose track of the relative size of each faculty. This gives the eikosogram:



In order to make this readable, we can then add labels and colours.

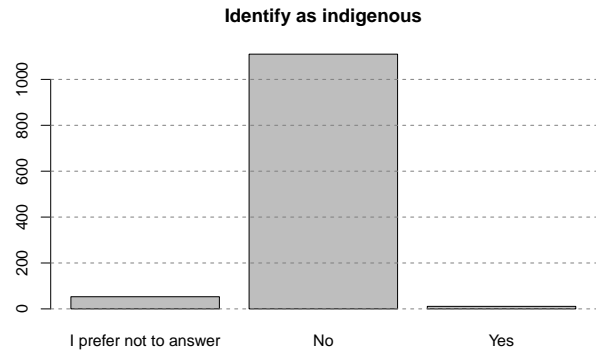
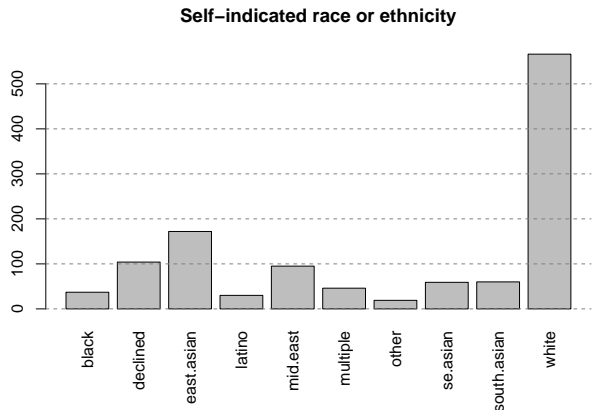
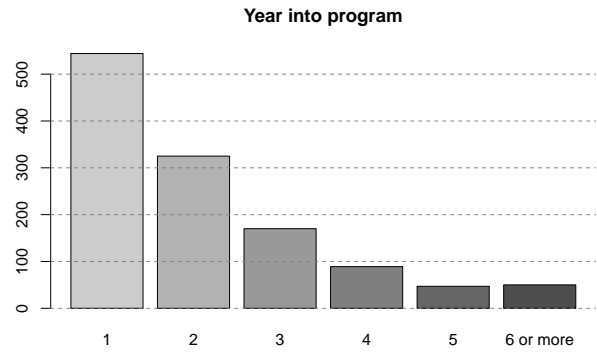
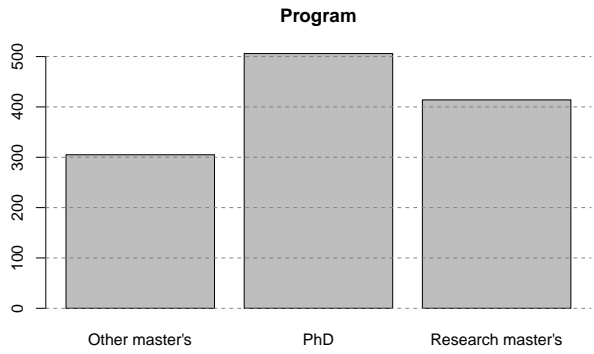
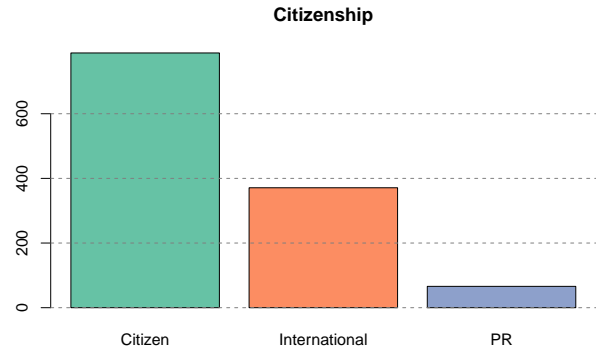
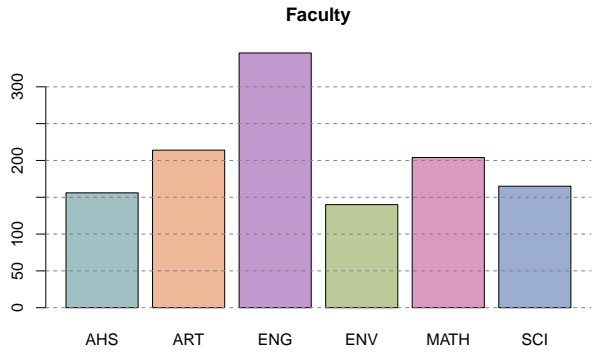


For a more detailed introduction to the eikosogram, refer to the tutorial at <https://rwoldford.github.io/eikosograms/articles/Introduction.html>. It is better conceived and more detailed than what has been presented here.

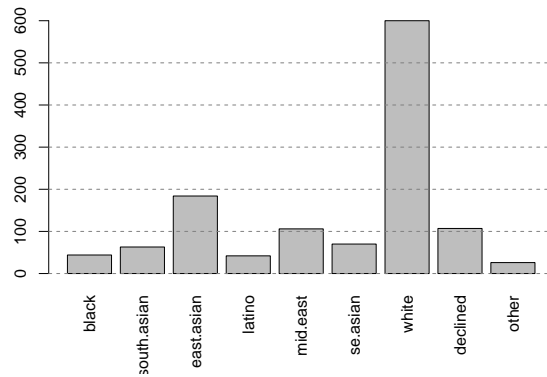
### The survey

The second GSA COVID survey was launched in November of 2020 as a follow up to April survey released in the first Ontario provincial lock down due to COVID. There were 1256 respondents, with demographics

broken down below:

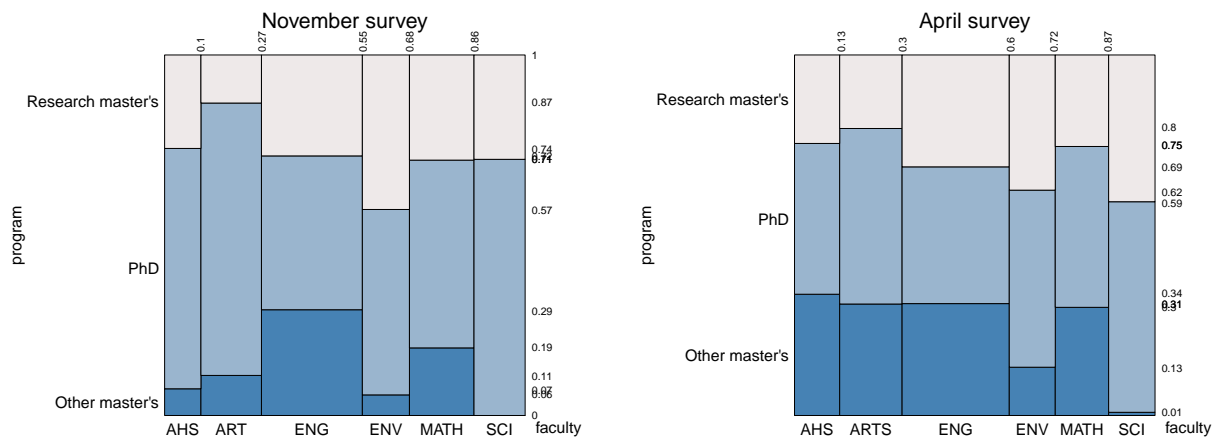


It must be noted that the last plot of race and ethnicity is a processed version of the original response. The original response allowed respondents to indicate multiple racial identities. The counts of each are displayed below.



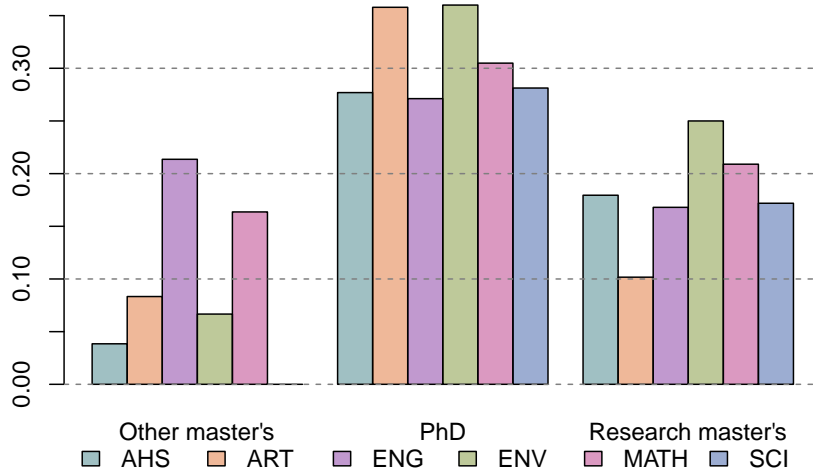
You may note that these counts look very similar to the earlier table, but the “multiple” count is missing. Across all respondents, only 46 identified with more than one racial group (see Appendix). In order to simplify analysis and interpretation, a variable “multiple” was created to indicate those respondents that selected more than one racial or ethnic identity.

For context, we can compare some aspects of these two surveys. For example, looking at repeat respondents by their faculty and program:



It’s obvious there are some substantial differences, with a greater number of course-based and professional master’s students present in the April survey. In total, roughly one third of the PhD students in the November survey are return respondents that also completed the April survey, while a smaller proportion of research and other master’s degrees represent measurements on the same individuals. More on these repeated respondents will come later when repeated questions are addressed.

**Proportion of return respondents (November)**

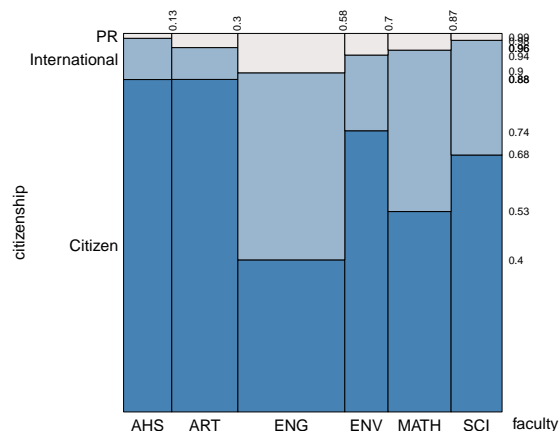


## Controlling for confounders

### Demographic confounding

A key problem with any survey is attribution. We would like to take a question of interest and determine which demographic factors or other questions are most strongly related to this question. The goal is to find a robust association which can be extended to the population as a whole.

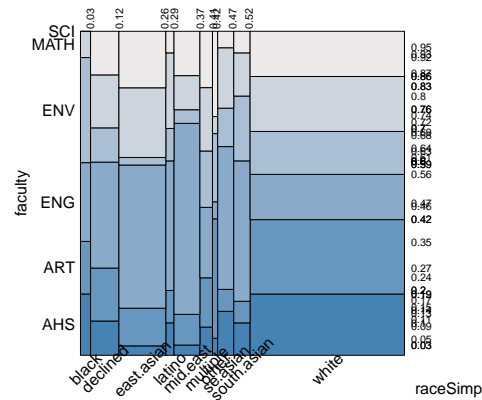
Two aspects of survey data make this difficult. The first, and most challenging, is confounding. The demographic factors in this survey are already related to each other. Consider, for example, the relationship of student citizenship and faculty:



Engineering and mathematics have a far greater prevalence of international students than the other faculties. As a consequence of this, any association we find between faculty and a question could really be an association between citizenship and that question. Indeed, separating these two factors to determine which is the most important for a given response is quite difficult.

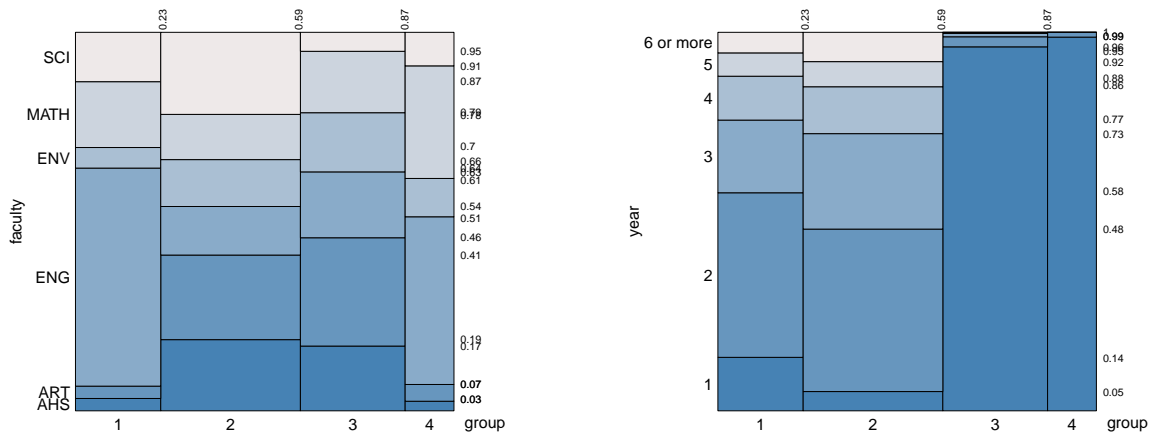
Generally, this is accomplished by the use of conditioning. That is, by viewing the impact of faculty while controlling for citizenship. Unfortunately, in the presence of demographic variables which have complicated

associations among each other, this can lead to problems. Consider, for example, the relationship between a respondent's race and ethnicity and their faculty of study.

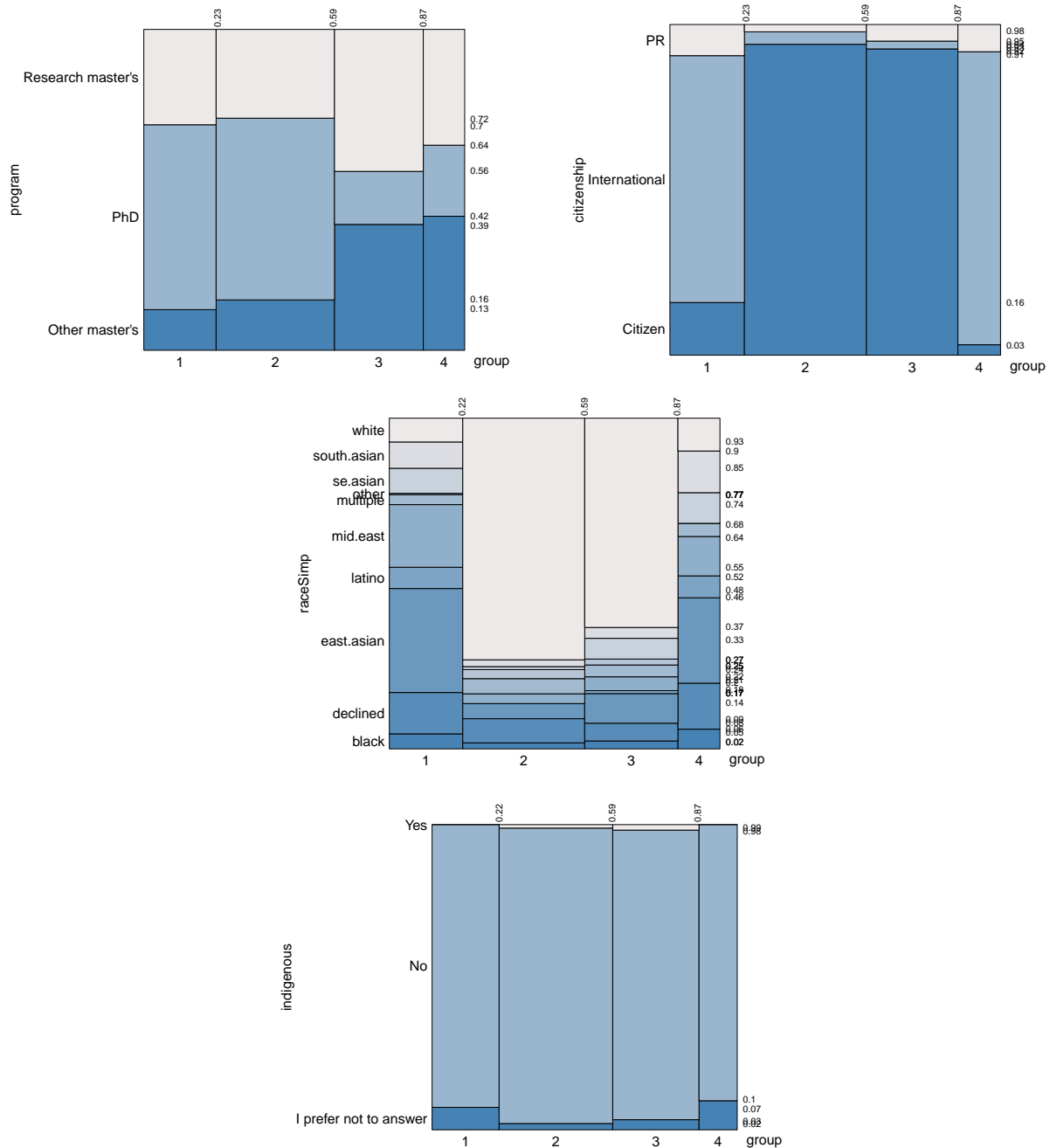


We would like to measure the effect of one while controlling the other, but we have a slight problem. Look at the first column of this eikosogram, and note that not a single respondent who identified as Black and their faculty is science. As a result of this, even though we may, mathematically, be able to estimate what we think a Black science student thinks of the questions on this survey, we have no actual data with which to base that result. Even if we control well, we will have this problem repeat itself again and again. That is, there will be subgroups of respondents for whom we have no observations.

In order to solve this problem, we need to think additively. Rather than removing and separating small groups of individuals to look at the way they work, we want to group similar students to create a few groups of students with meaningful things in common. The way this is done is some rather complex math, and a primer is in the appendix, but the end result of this process (clustering on the main dimensions of a multiple correspondence analysis) are four groups of students. Using eikosograms, we can examine the distributions of students in each.







This series of plots leads to some quick demographic descriptions of each group:

1. Group one is primarily upper year, non-white, international PhD students in STEM faculties, particularly engineering and math
2. Group two is primarily upper year, white, domestic PhD students in non-STEM faculties
3. Group three is primarily first year, white, domestic master's students in non-STEM faculties
4. Group four is primarily first year, non-white, international master's students in STEM faculties, particularly engineering and math.

These stark demographic differences provide a means by which many broad questions of interest might be addressed by the comparison of these groups. If, for example, we wanted to look at the impacts of COVID on STEM students compared to non-STEM students, we can compare the responses within groups one and four

to those of two and three. If we wanted to perform this comparison only for first years, we could compare group three to group four. For upper years? We can compare one and two. By such comparisons, a great deal can be learned with a relatively small number of groups kept in mind.

Notice that certain comparisons cannot be made with these groups. For example, we cannot separate STEM and international students. These comparisons are exactly those which are difficult to perform given the structure of our data due to groups which were not sampled or which are uncommon, and so are restrictions which would be present under any analysis.

To demonstrate the predictive and descriptive power of these groups, consider the example of viewing the distribution of where students are living by these groupings.

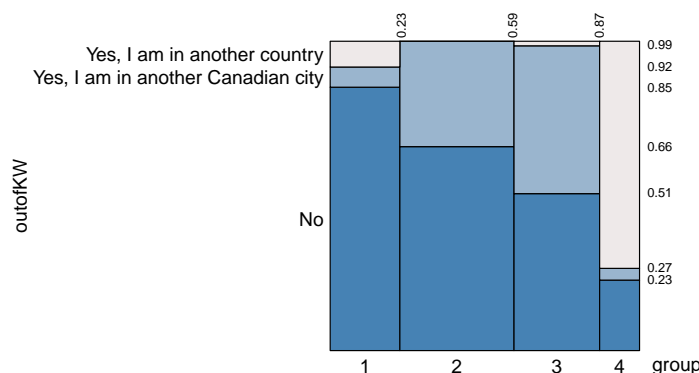


Figure 1: Residence by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

By comparing groups three and four to groups one and two, we can immediately surmise that first year students are much less likely to be in the KW region than upper year students. Furthermore, by noting that group three is dominated by domestic first year students and group four by international first years, we can explain almost entirely the important difference in whether first year students are living abroad or in Canada. By noting that the final group is dominated by engineering and mathematics students, we can gain a final piece of useful insight: a majority of international STEM first years are not living in Canada, and may never have. This has important implications for the administration of funding and employment for these students.

Additionally, we might note that the group that is most likely to still be present in KW is that of upper year international STEM students. This has immediate implications for planned programming which may be organized for the immediate end of lock down measures.

It must be stressed that this example of the groupings being right should not be taken to imply that these groupings will not account for all differences present in the responses of students. They are, ultimately, still a limited tool. However, by condensing the largest differences in nine demographic variables into a single variable indicating membership to one of four groups, they provide a means by which our analysis of the data might be cohesive and narrative in nature.

Rather than a series of simple comparisons between a pair of questions or demographic variables, we have the opportunity to learn more about these four artificial groupings which seem to capture important demographic differences. This report aims to learn about these groups, rather than focus narrowly on any one demographic variable. More detailed analysis of any of these, of course, might still provide additional insight on top of this

summary.

## Key findings

### Finance

- a greater proportion of students across faculties reported decreased income than in the April survey
- group four (international first years in engineering and math) were mostly likely to report a decrease in income
- first year students and students in their fifth year or above were more likely to report decreased income
- international students were more likely to report decreased income
- the most common reason for income reduction was a depletion of savings
- groups one and four (international non-white students in engineering and math) were most likely to cite funding concerns
- first year students were most likely to cite quitting as a reason for decreased income

### Academics

- upper year students are more likely to have experienced setbacks, and typically have experienced setbacks of a greater number of terms
- a majority of students have been setback by two terms or fewer
- among respondents unsatisfied with their work from home situation, a lack of daily structure and inappropriate workspace were the most cited reasons
- respondents requiring lab equipment found the reduction in hours the primary hindrance across all groups
- international students experienced poorer communication in online administered courses
- first year students felt that feedback, support, and discussion in online courses was poor

### TAs

- international first years in engineering and mathematics are much less likely to receive a TAship than other groups
- across all groups, students find being a TA more challenging online
- the most cited challenges to online TA work were issues providing guidance to students, time spent answering emails, and the difficulty communicating with students
- international students were much more likely to cite work converting course material to an online format as a challenge

### Supervisors

- most students report no change in their supervisor relationship
- groups one and four (international STEM students) were much more likely to cite poorly communicated expectations as a problem between them and their supervisor
- group four (first year, international, STEM master's students) was much more likely to cite personal priorities they placed above their supervisor relationship, while group two (upper year, domestic, non-STEM PhD students) was much more likely to cite supervisor priorities being placed above them

## Wellbeing

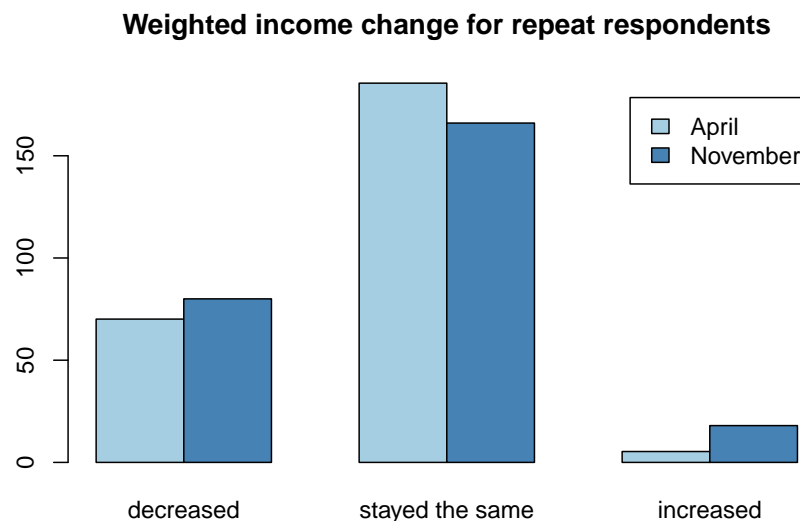
- student mental health has deteriorated somewhat since April: more students state a significant decline in their wellbeing has occurred
- a lack of physical activity and social contact were top contributors to this across all groups
- upper year students were more concerned with their progress than first years, and first years were more concerned with course administration
- international and non-white students were slightly more likely to have cited income changes and discrimination as a cause of mental decline
- domestic students tended to be more concerned about friends and networking, while international students were more concerned about funding

## Supports

- in general, few people apply to supports relative to the number of people aware of them
- financial supports are less well known and less used than other supports

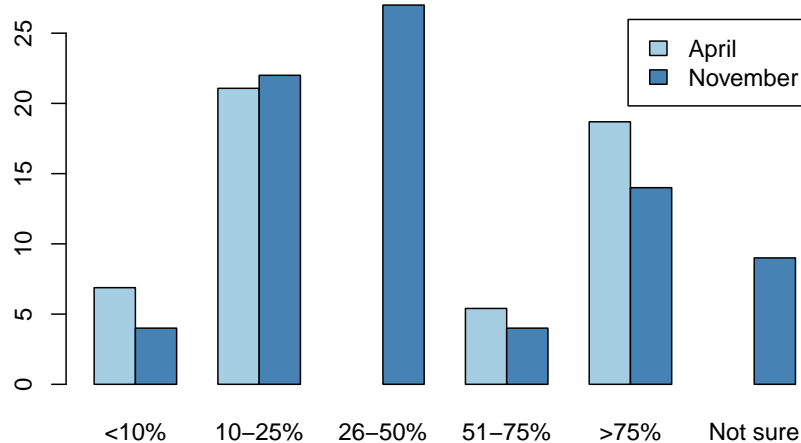
## Finances

The groups are not the only interesting avenue for analysis, however. Consider the repeated measures presented by the repeat respondents on finances. In both surveys, questions were asked about the change in income experienced by students. The barplot below presents a comparison of the responses for repeat respondents, weighted to reflect the difference in the distribution of demographics from April to November.



This plot indicates two things. The first is that students which did not experience an income change in the April survey are more likely to have had a change in income by November. The second is that more of these students increased their income than experienced a decrease. Viewing those decreases:

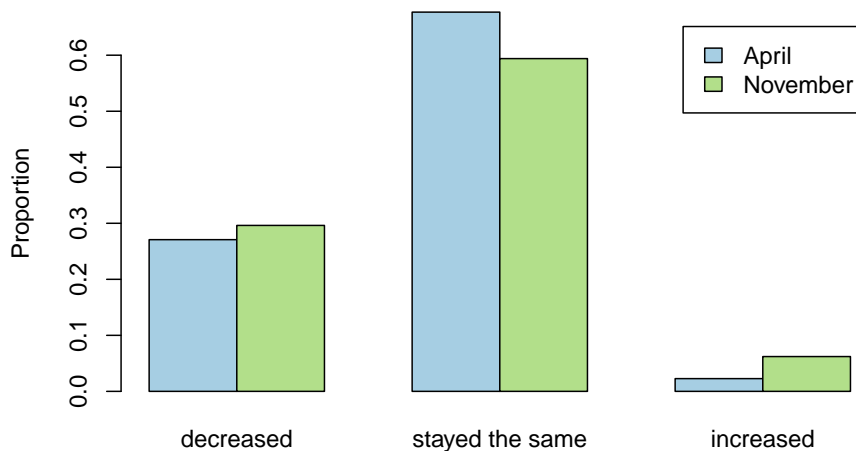
**Weighted income decrease for repeat respondents**

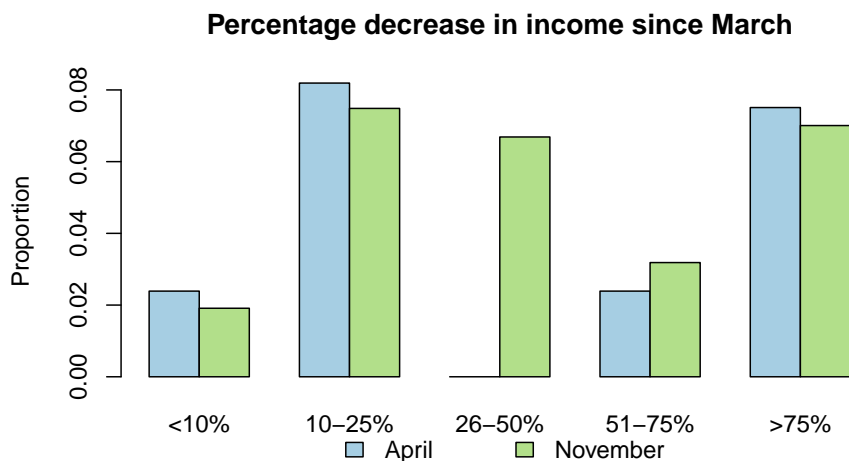


A stark difference is apparent. Those that had a decreased income in November were most likely to have lost a quarter to half of their income, while in April the plurality lost less than a quarter. Note the small sample sizes present in this plot should give pause. Extending these results, from a non-random sample, to the general student body should be done only cautiously.

We can now look at these answers across both surveys for the entirety of respondents.

**Income change since March**





These do not tell much of a different story, though we can see that repeat respondents appear to have fared better than the average respondent in the November survey, being less likely to report to a loss of more than 75% of their income.

We can now compare the changes across our four demographic groups.

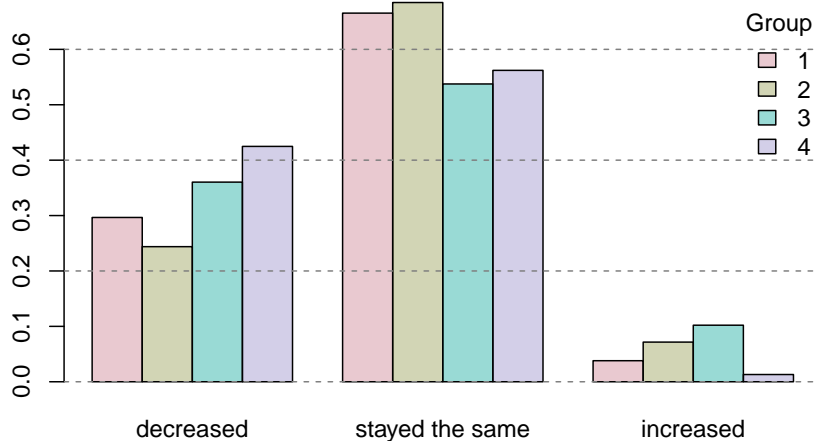
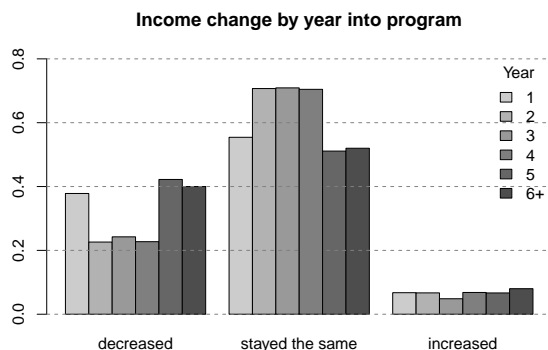


Figure 2: Income change by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master's students not in engineering or math; Group 4: international, first year master's students in engineering and math)

Recall the composition of these groups: the first two are primarily students who have been at Waterloo

for more than a year, the second and third are primarily domestic students, and the first and fourth are dominated by engineering and mathematics students. The above bars show quite clearly that citizens have fared somewhat better than international students, but it seems the most salient effect is that of the year into program. Viewing the first two groups of bars, it is clear that the students who are in their first year are far more likely to have experienced a loss of income, while those who are several years in are much more likely to have experienced no change. Across all groups, more than half of respondents reported no income change since March due to lockdown measures.

Indeed, the year into one's program appears to be the main determinant of their income change, as can be seen below.



First year and students beyond their fourth year of study are much more likely to have had a decrease in income than those in their second to fourth terms.

We need not wonder about the reasons for this pattern, however, as respondents provided information on the reasons for their lost income, which we can view across our groups.

The shortened names here stand, respectively, for the eight responses:

1. savings used to supplement income depleted
2. did not obtain a planned position
3. being laid off
4. reduction in family income
5. reduction in scholastic funding
6. reduction in earnings at work
7. quitting a job
8. changing jobs at the same employer to a lower-paid position

We can see that the proportion of respondents reporting these income changes varies widely between the groups. Across all groups, the depletion of savings was a critical reason for a decreased income. Indeed, it is the top reason for all groups except group three, which experienced changes in planned jobs and layoffs at a much higher rate than the others. We might also note that groups three and four, the first year students, tend to have experienced business-related and economic reasons at a higher rate than their upper-year peers. This can be seen by viewing the major factors they reported for the loss of their income.

Where the labels here are, respectively, shorthand for the reported factors:

1. change in life circumstances during the pandemic
2. business impacts of public health measures
3. mental health reasons
4. impacts within family and personal networks
5. physical health reasons
6. discomfort with exposure risk
7. increased caregiving duties

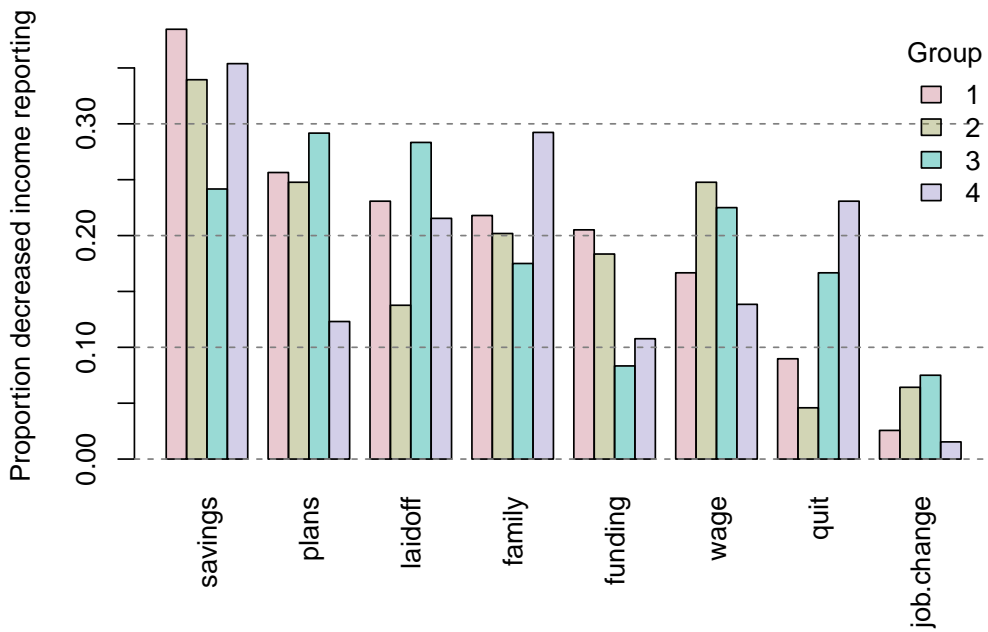


Figure 3: Income decrease reasons given by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master's students not in engineering or math; Group 4: international, first year master's students in engineering and math)



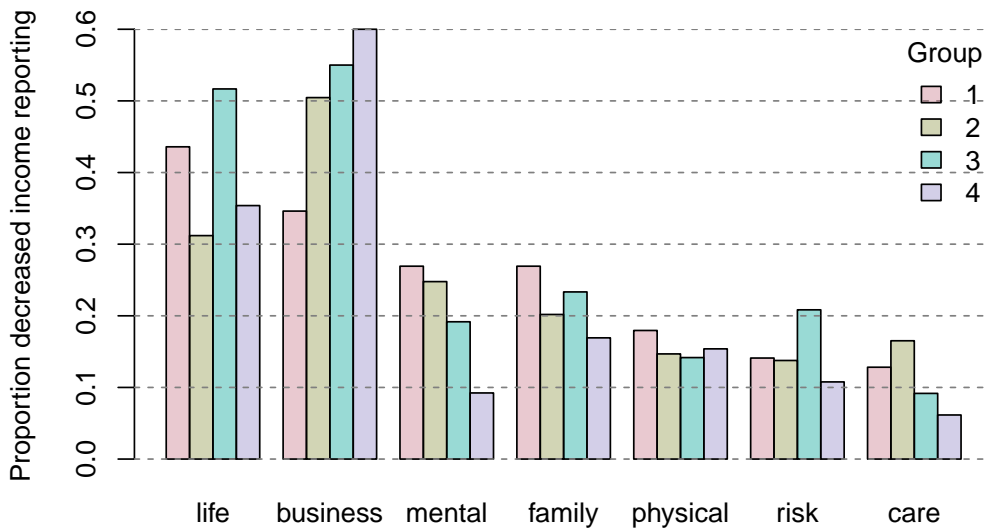
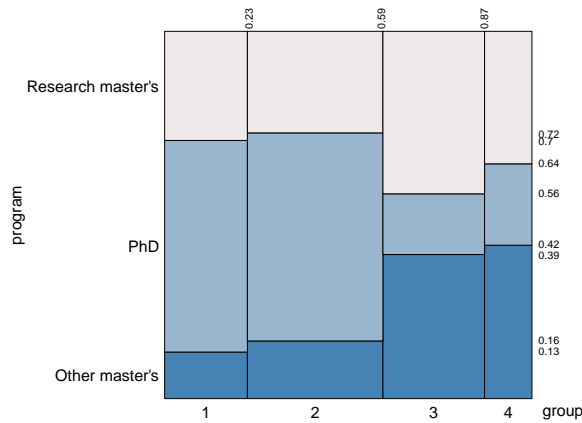


Figure 4: Income factors given by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master's students not in engineering or math; Group 4: international, first year master's students in engineering and math)

We can see that life and business factors dominate among the first year students in groups three and four. One might hypothesize that many of these students started their graduate programs as a result of these factors. Changing life circumstances, layoffs, and worsened job conditions alongside the flexibility to gain a degree from an internationally-recognized Canadian institution completely online may have encouraged these individuals to pursue graduate degrees. Such a trend, while it speaks to an aspirational struggle all of us can relate to, suggests incoming graduate students with different goals and needs than a typical year.

## Academic plans

Recall our breakdown of groups by program type:



Some of the following analyses are primarily of conditional questions which will address certain subgroups specifically based on their program. It is useful to keep the above picture in mind when these questions are analyzed.

First, however, the general question of setbacks experienced by students was posed, and the groups responded thus:

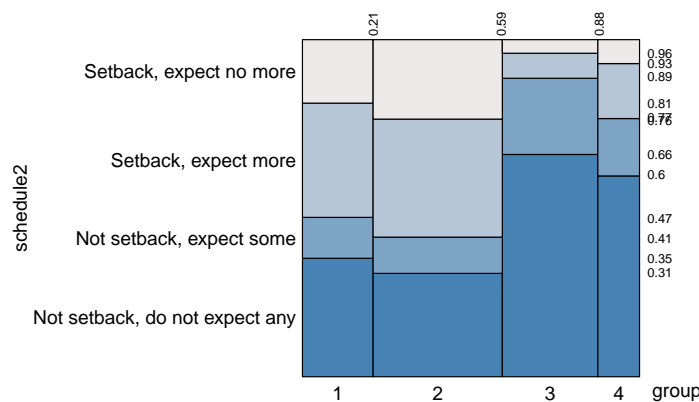


Figure 5: Setback by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master's students not in engineering or math; Group 4: international, first year master's students in engineering and math)

Unsurprisingly, the groups dominated by first year students have yet to experience setbacks, though this may be true of most students that are so fresh in their programs. Just over half of the upper year students are reporting setbacks, with almost a third expecting more setbacks in the future. Contrast this with the first year groups, who are largely optimistic and anticipate no setbacks to their program schedule.

Despite this, it seems that across all groups, those that have experienced, or expect to experience, a setback do not anticipate that they will be setback more than one term.

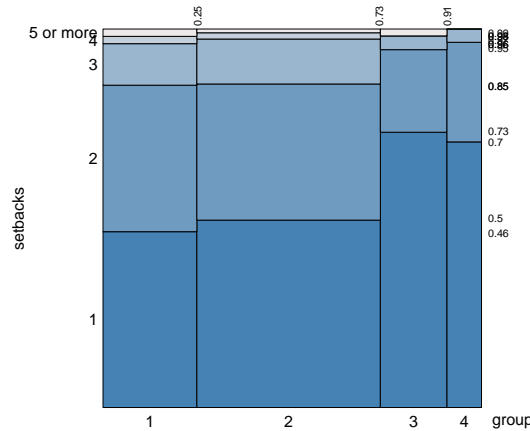


Figure 6: Amount setback by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

Again, first years are more optimistic (or perhaps naive) than upper years, but most students expect the setbacks in their program to be minimal and two terms at most.

When asked about their satisfaction with their workspace, there are less distinct differences between the groups.

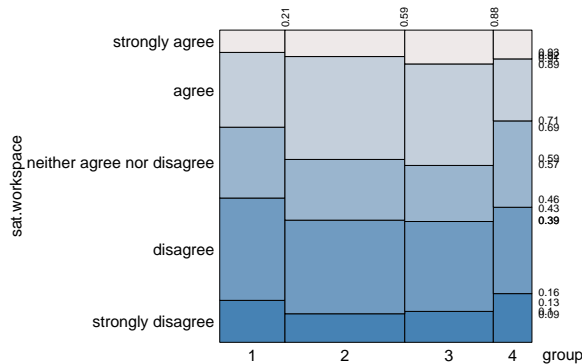


Figure 7: Work from home satisfactory by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

Students seem split across all groups, with a slight tendency to disagree for groups one and four (primarily international STEM students). When asked about the factors causing dissatisfaction:

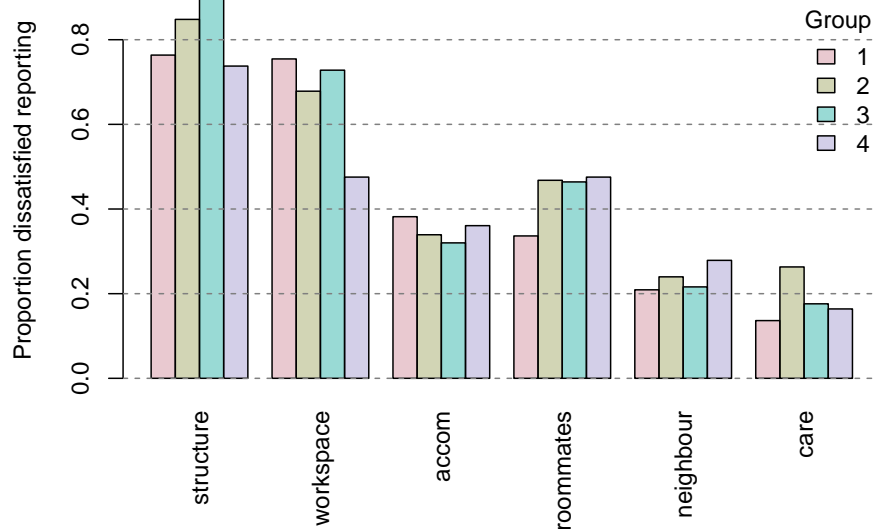


Figure 8: Work from home factors by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

The above categories are abbreviations of:

1. lack of a structured work day
2. lack of an appropriate work space
3. accommodation reasons
4. roommate or family issues
5. caregiving duties

We can see that the dominant factor, reported by a majority of those experiencing dissatisfaction across all groups, is the loss of structure in their work day, followed reasonably closely by the lack of a dedicated workspace. Unfortunately, these barriers to productive work seem to be a necessary consequence of the restriction of social engagement followed by governments in the response to the pandemic.

The following questions, about equipment and lab access, were posed only to PhD and research master’s students. Remembering our breakdown of the groups by program, we should pay extra attention to the response of the first two groups to these questions, as these groups were dominated by research students. First, students were asked if they require specific, non-lab, equipment for their research.

Most students across most groups, do not require any such equipment. Those who did, however, had the following challenges related to accessing this equipment.

While the need to access equipment is spread roughly evenly across groups, there are some clear patterns by group present in these responses. International students in engineering and mathematics in group one rate the reduced hours and closure of campus as their key gripes, while citizens of other faculties in group two seem to be experiencing fewer difficulties overall. Viewing the access to lab-specific equipment by surveying students if they require labs:

We see most students do not require labs for their research, with group one having the greatest dependence and group three the least. The factors affecting lab access were then given by the respondents that require this equipment.

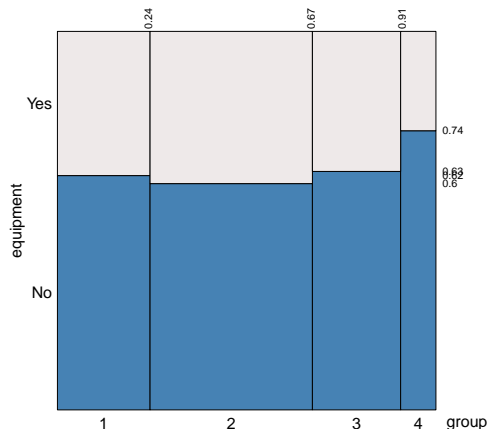


Figure 9: Equipment use by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master's students not in engineering or math; Group 4: international, first year master's students in engineering and math)

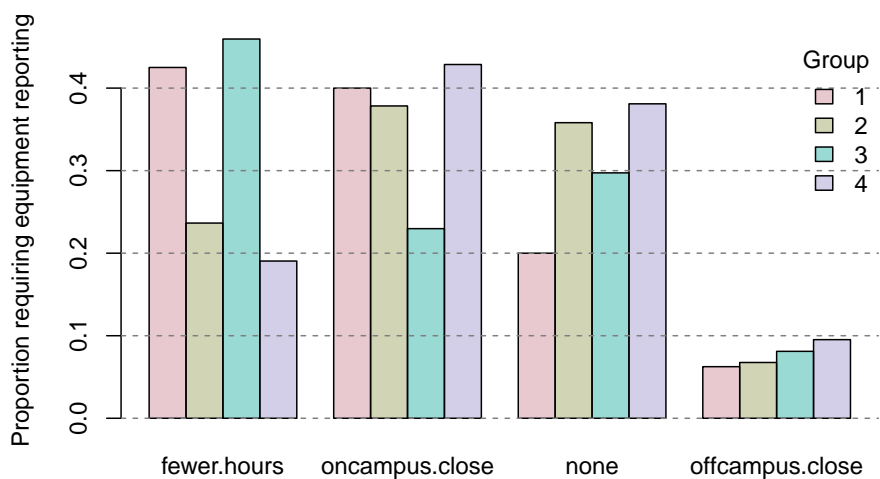


Figure 10: Equipment challenges by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master's students not in engineering or math; Group 4: international, first year master's students in engineering and math)

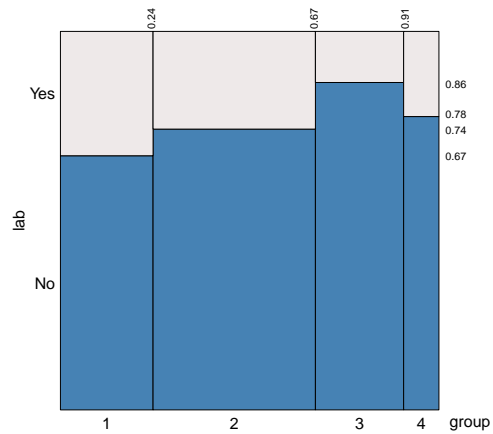


Figure 11: Lab use by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master's students not in engineering or math; Group 4: international, first year master's students in engineering and math)

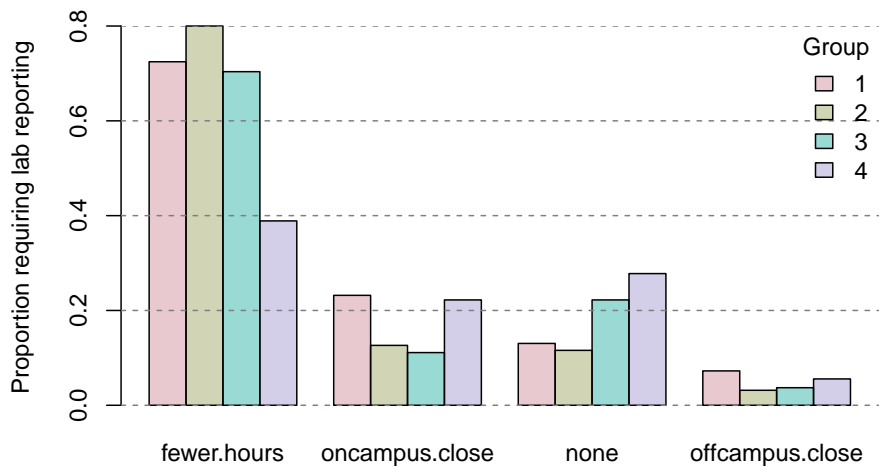


Figure 12: Lab challenges by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master's students not in engineering or math; Group 4: international, first year master's students in engineering and math)

Unlike the non-lab equipment, there is a clear problem primarily with the hours allowed in labs across all groups, and in particular one, two and three. Closures on and off campus seem to be far less relevant for lab access. Relating to research, students were asked about the impact of the pandemic on their thesis project plan.

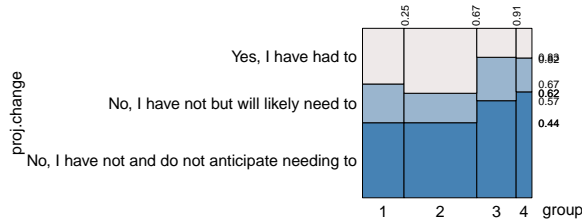


Figure 13: Thesis project changes by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

Here we see the first year optimism once again, as groups three and four are more confident they will not need to adapt their thesis plan for the pandemic at all. For those who have already had their dreams of sticking to their original plan shattered, an additional question was asked to gauge their satisfaction with the changes that have been made:

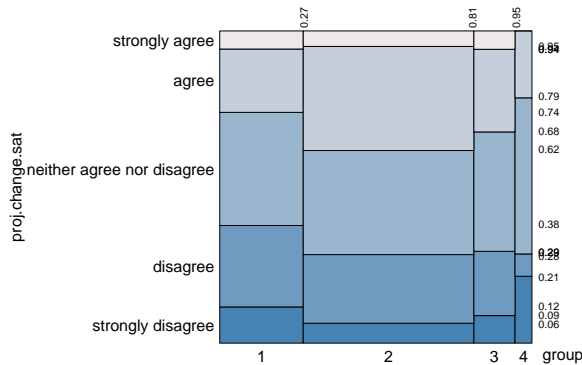


Figure 14: Project change satisfaction by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

Generally, it seems groups two and three are more likely to be satisfied with the change in their thesis project plan. This could be a result of the largely domestic character of these groups or the large engineering representation in the others. A final question inquired about the completion of in-person milestones such as a co-op placement, practicum, or internship. First, we had to see which students have such milestones.

These programs seem to be more common outside of mathematics and engineering, as groups two and three

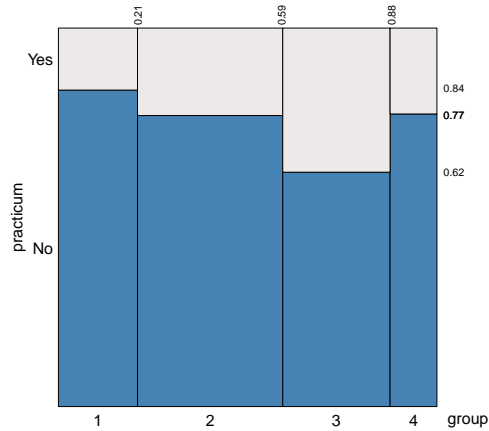


Figure 15: In person milestone requirement by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

are more likely to require them. How many students that require these components have had to change them?

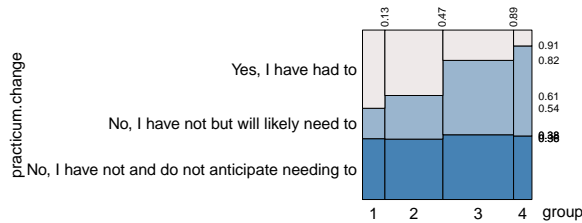


Figure 16: In person milestone change by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

While there is no difference in the proportion of respondents that have not, and do not expect, an interruption between the groups, we can see that most students that require such in-person milestones expect, or have already experienced, delays. Furthermore, the follow up questions for those who have already experienced delays and those that expect to experience delays indicate either neutrality or satisfaction with the changes.

A more universal set of questions involves the administration of courses. We begin by considering the fall 2020 term and checking the enrollment by group.

As might be expected, first year students in more course-focused programs are far more likely than upper year students to have been enrolled in classes, as indicated by the greater proportion of “Yes” responses in groups three and four. For those individuals who are enrolled, the next question was of their satisfaction.

There is a tendency in all groups towards neutrality or satisfaction with the adjustment of courses. For those who were dissatisfied, the primary reasons were then provided by respondents, and are shown in Figure 20.



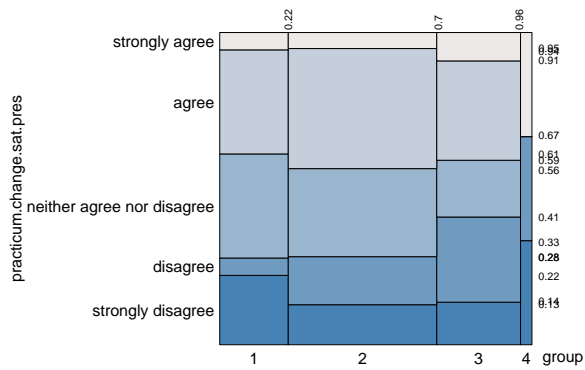


Figure 17: Satisfaction with in person milestone change by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

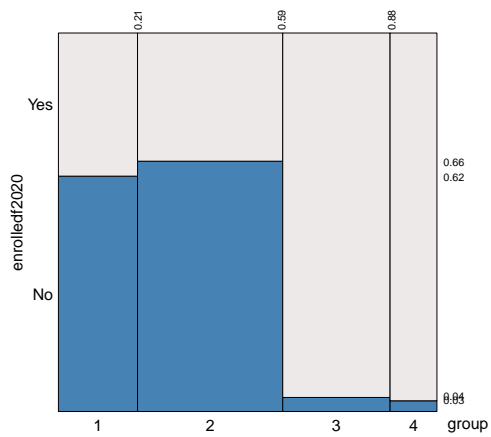


Figure 18: Enrolled fall 2020 by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

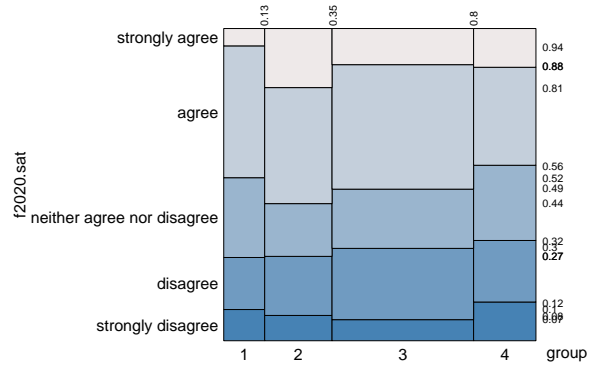


Figure 19: Fall course satisfaction by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

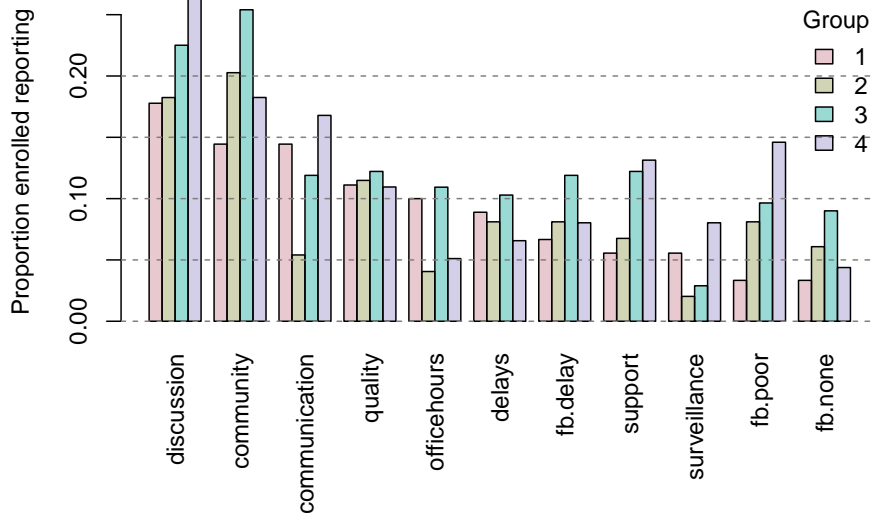


Figure 20: Course dissatisfaction factors by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

Once again, these factors are, in order:

1. poorer quality of online discussions
2. lack of online class community
3. difficulty communicating with the class instructor
4. reduced quality of course material
5. absence of office hours
6. delays in uploading content
7. delays in feedback
8. lack of support beyond course content
9. concerns over exam surveillance and checking methods
10. poor quality feedback
11. no feedback provided

We can see that the dominant factors in all groups are the lower quality discussion, lack of community, and difficulty communicating with the instructor in an online setting. We can also see that in group four and group three, the students that have just started, feel a lack of support in their studies more acutely than do groups one and two.

## Teaching assistantships and supervisors

A typical part of the graduate student experience is teaching assistantships. Let's see how the groups differ in the prevalence of TAs.

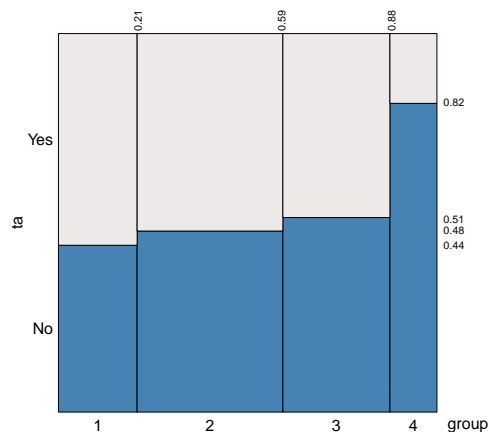


Figure 21: Teaching assistantships by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master's students not in engineering or math; Group 4: international, first year master's students in engineering and math)

TAs seem to be offered rather evenly across these groups, with the exception of group four, where respondents are unlikely to have been a TA. This is, perhaps, unsurprising. Recall that this group was dominated by international students that started this year and largely still live abroad. Many of these students may not have a Canadian social insurance number or bank account, and so are not eligible to be employees of the university.

For those without that limitation who have been granted a TAship, the next question was whether they agreed that TA duties have been made more challenging by COVID.

Which very few students disagreed with. The primary factors for those that find it more difficult reported by each group are displayed in Figure 23.

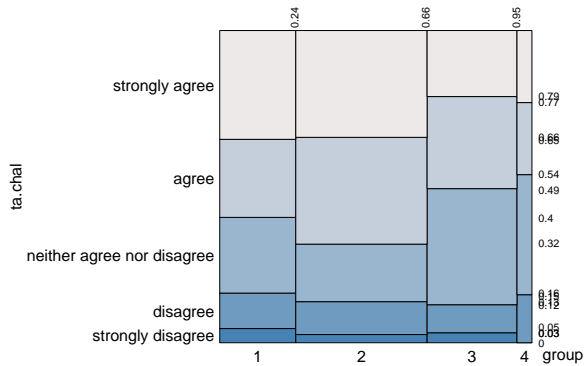


Figure 22: Increased TA difficulty by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

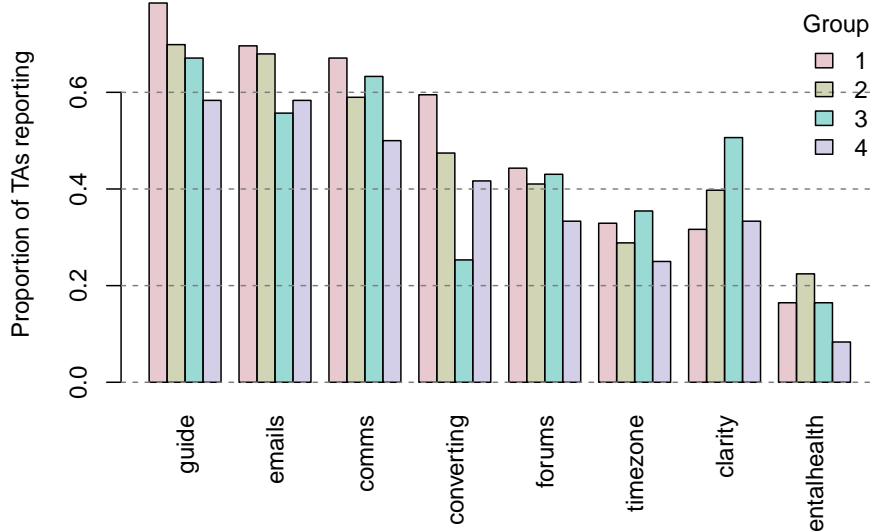


Figure 23: TA difficulty factors by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

The shorthand on this question is expanded in order as:

1. increased workload providing guidance to students
2. increased workload answering student emails
3. difficult communicating with students online
4. increased workload converting in-person material to online material
5. increased workload monitoring forums
6. time zone differences between teaching staff and students
7. decreased clarity on duties by instructor
8. increased reliance on TAs to handle student mental health

In sharp contrast to many of the other factor questions asked, most factors here have a rather large representation, and a relatively consistent pattern across groups. The largest factors are clearly related to communication with students to provide guidance, especially answering student emails.

Next, respondents were asked about the relationship they have with their supervisor

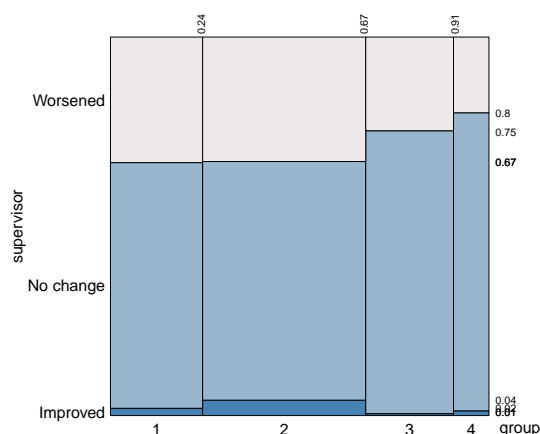


Figure 24: Supervisor relationship change by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

Here, a large majority of students across all groups have felt no change in the relationship between them and their supervisor, in particular this is the case for first year students. As always, the respondents who indicated a worsened relationship were asked to provide more information about the major factors contributing to the decline in their relationship, which are displayed in Figure 25.

The options provided to respondents were, in order:

1. difficulties in communication
2. poor communication of expectations
3. supervisor focused on competing priorities
4. respondent focused on competing priorities

While every group indicated that communication is a key reason for the decline in their relationship, we can see that groups one and four are much more likely to report the poor communication of expectations as an issue, and groups two and three were much more likely to cite competing supervisor priorities. This suggests that students in STEM programs have had greater difficulty adapting to new expectations, which students in non-STEM programs feel neglected in favour of other priorities.

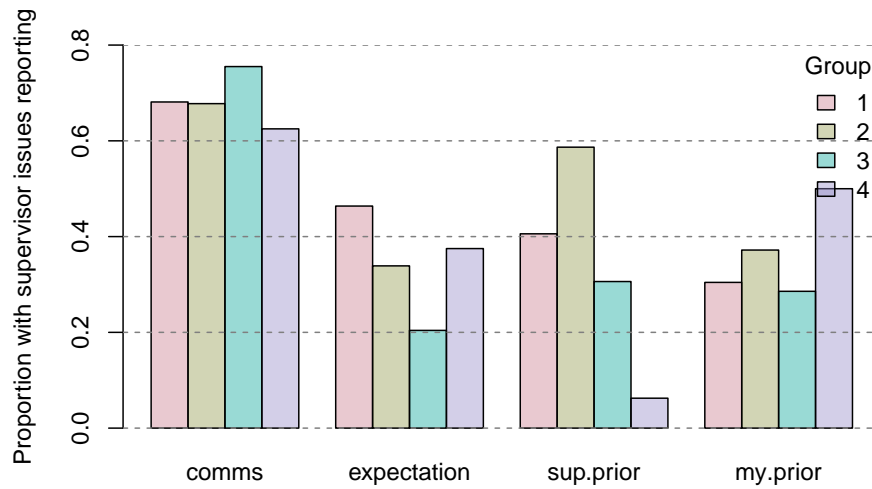


Figure 25: Supervisor relationship decline factors by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

## Student wellbeing and supports

All of the above can be imagined as a means to colour our real interest: that of the wellbeing of students. Questions aimed at this topic made up the final section of the November COVID survey. First, we can see the self reported change in mental health by our different groups.

Here, every group has reported roughly the same decline in mental health. As this question was repeated in both the April and November surveys, we can forget the groups for a moment and simply compare the results between these two surveys.

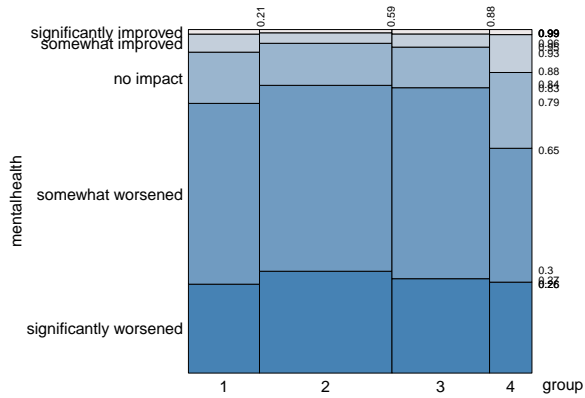
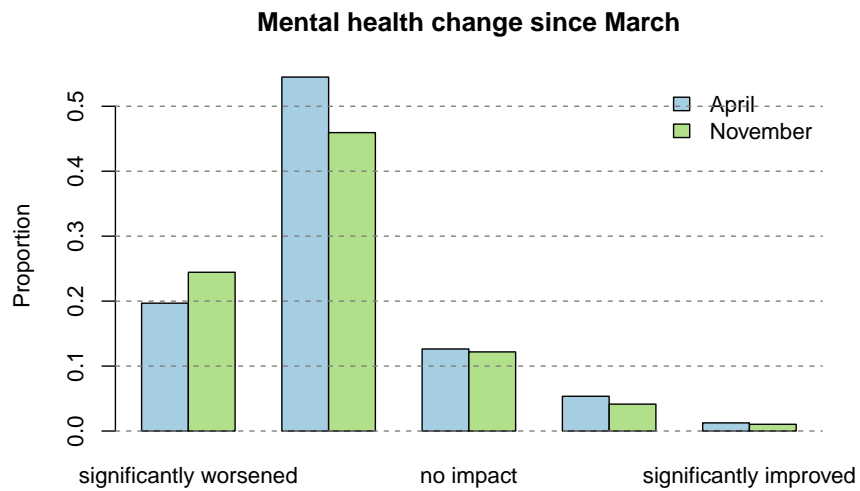
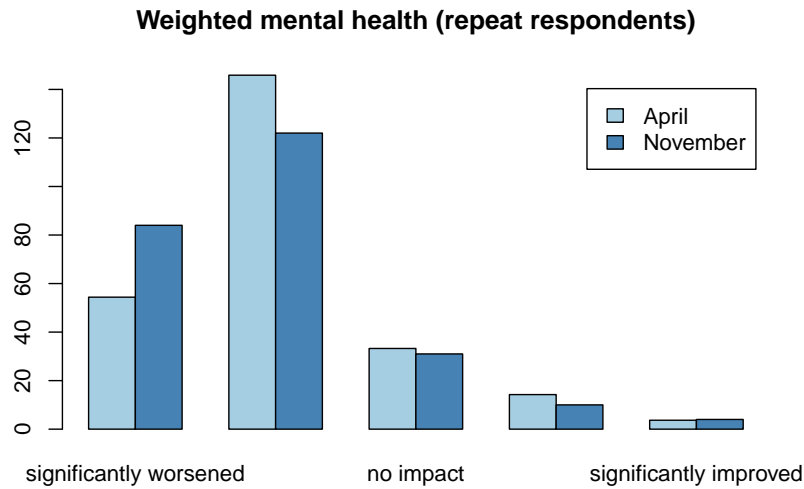


Figure 26: Mental health changes by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)



Compared to April, there are more students reporting a significant decline in mental wellbeing in the November survey. While the proportion reporting any decline is not much different than in April, it seems the magnitude of that decline tended to be larger in November than in April. Indeed, this trend is even more exaggerated among the respondents that completed both surveys, where a larger change has occurred.



One hypothesis for this decline is the slow degradation of mental health that follows from the protracted period of uncertainty of the pandemic. Individuals that were stable enough to handle the shock of one month of social isolation and disruption may not have been mentally prepared for an entire year of such measures. This hypothesis can be tested somewhat by viewing the self-reported factors for the decline in mental health given by those who have experienced a decline, again introducing our groups.

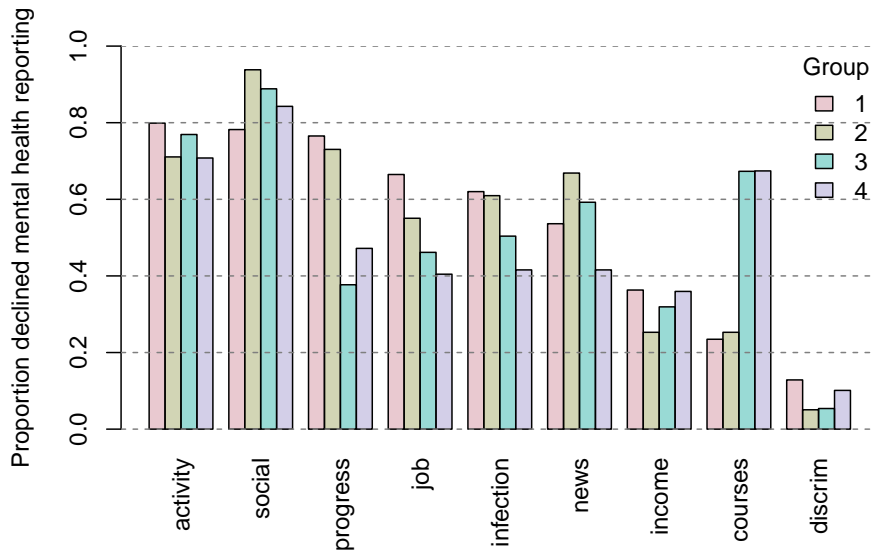


Figure 27: Mental decline factors by group (Group 1: international, non-white, upper-year PhD students in engineering and math; Group 2: domestic, white, upper-year PhD students not in engineering or math; Group 3: domestic, white, first year master’s students not in engineering or math; Group 4: international, first year master’s students in engineering and math)

Shorthand is again used in the above plot, where the labels stand for, in order:

1. lack of regular physical activity



2. reduced in-person interaction with others
3. concerns with research progress
4. concerns with obtaining a job after graduating
5. concerns or experience with COVID infection personally or in others
6. overload of pandemic-related news
7. reduced income due to the pandemic
8. concerns with course administration
9. experiencing discrimination

Across all groups, the dominant factors are social and physical. Most students report a lack of adequate physical activity and a lack of social interaction as factors impacting their mental health negatively. Groups one and two (the upper year students) are very concerned with their research progress, while groups three and four (the first year students) are more concerned with courses. These distinctions make sense considering the differences in program and year into program for these groups. Reported discrimination was experienced only by a small minority of students across all groups, but is more common in the international student groups than the domestic ones.

Students were also asked what aspects of their experience were most concerning to them in light of the pandemic, with domestic and international students receiving slightly different questions.

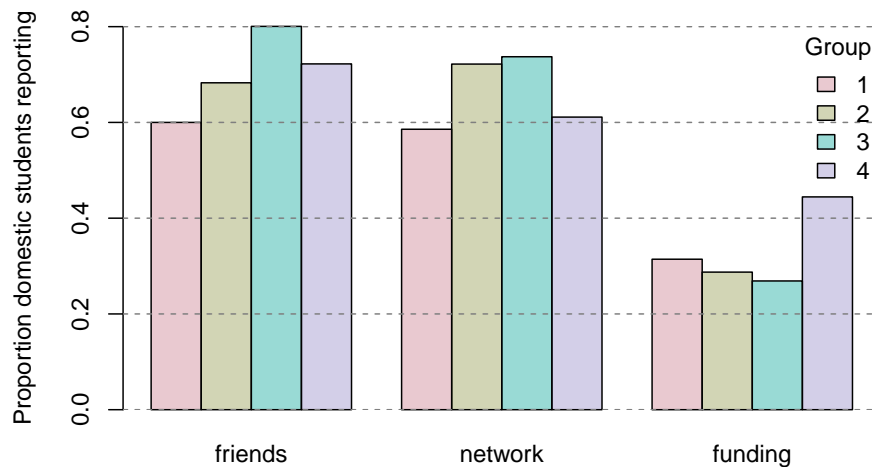


Figure 28: Domestic student experience concerns by group (Group 1: non-white, upper-year PhD students in engineering and math; Group 2: white, upper-year PhD students not in engineering or math; Group 3: white, first year master’s students not in engineering or math; Group 4: first year master’s students in engineering and math)

Among domestic students, the dominant aspects which students are concerned about are making friends and networking professionally. In short, social aspects of their degree. Both of these social aspects seem to be felt harder among respondents in groups two and three, while group four feels the greatest concern for funding. For international students:

A similar story is told by the international students. Most respondents are concerned about the loss of social aspects of their degree. Unsurprisingly, group four is also concerned about their ability to experience Canadian culture. Recall that this group is predominantly international and consists of many students that are still living abroad. We can also see that funding feels more precarious to international students in groups

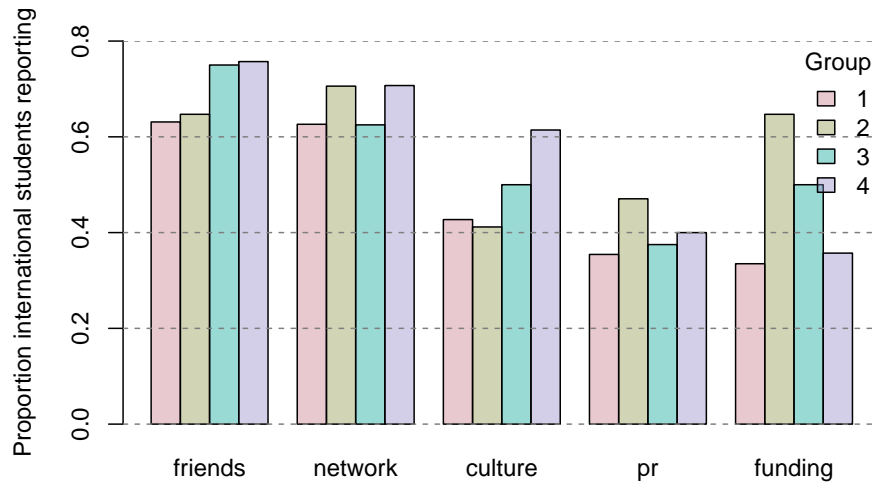


Figure 29: International experience concerns by group (Group 1: non-white, upper-year PhD students in engineering and math; Group 2: white, upper-year PhD students not in engineering or math; Group 3: white, first year master’s students not in engineering or math; Group 4: first year master’s students in engineering and math)

two and three, representing students not in engineering or math.

Finally, we might consider the use of different supports by each group.

Across all groups, we see similar patterns present. While group two demonstrates the greatest awareness and application numbers, the shape of these stacked bar plots is similar across all groups. The names here stand for:

1. CCA advisors
2. Regular CCA workshops
3. CCA job shadow program
4. CAA conferences
5. CTE workshops
6. WCC workshops
7. WCC programs
8. EDGE by Mitacs
9. GSA-UW virtual speed networking and meet and greets
10. GSPA events
11. HREI training
12. SSO student leadership program

In general, students are most aware of GSA activities, WCC, CTE, and CCA regular programming, but tend to participate most in the CTE programming.

Moving on to financial supports:

We can rather clearly see that the dominant financial support used by students is OSAP, and that OSAP is more utilized by groups two and three, which have a greater proportion of domestic students. Upper years tend to be more likely to have used or be aware of supports, and domestic and white students are more likely to do so than international and non-white students. As well, we should observe the generally lower awareness

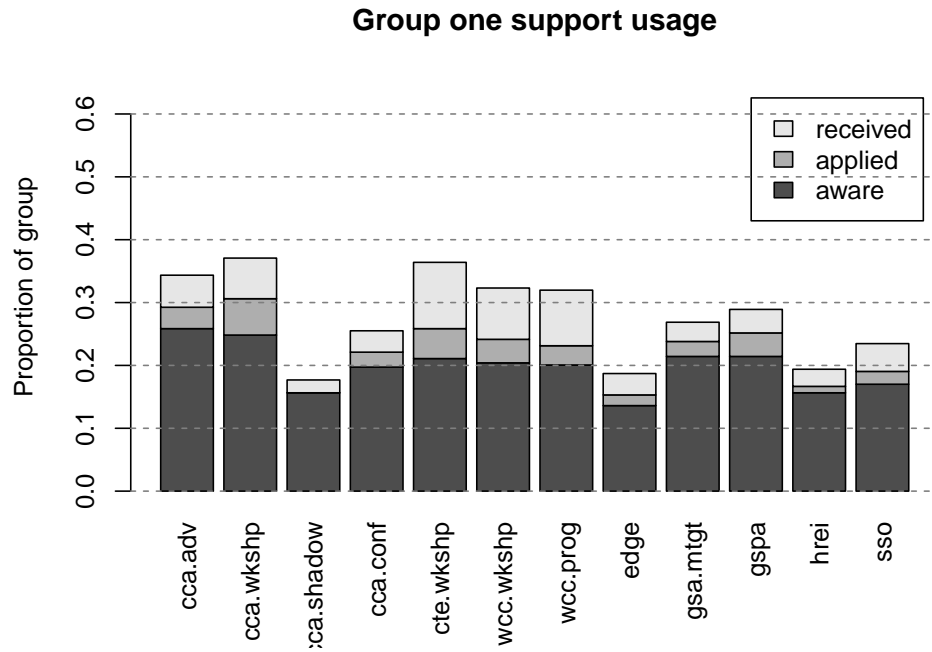


Figure 30: Support usage by group 1: international, non-white, upper-year PhD students in engineering and math

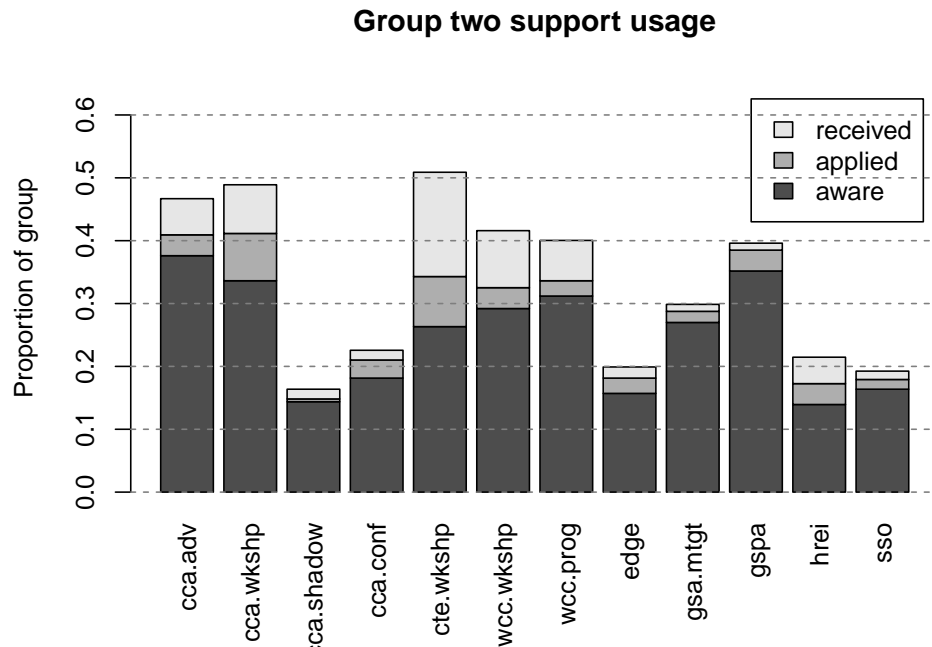


Figure 31: Support usage by group 2: domestic, white, upper-year PhD students not in engineering or math

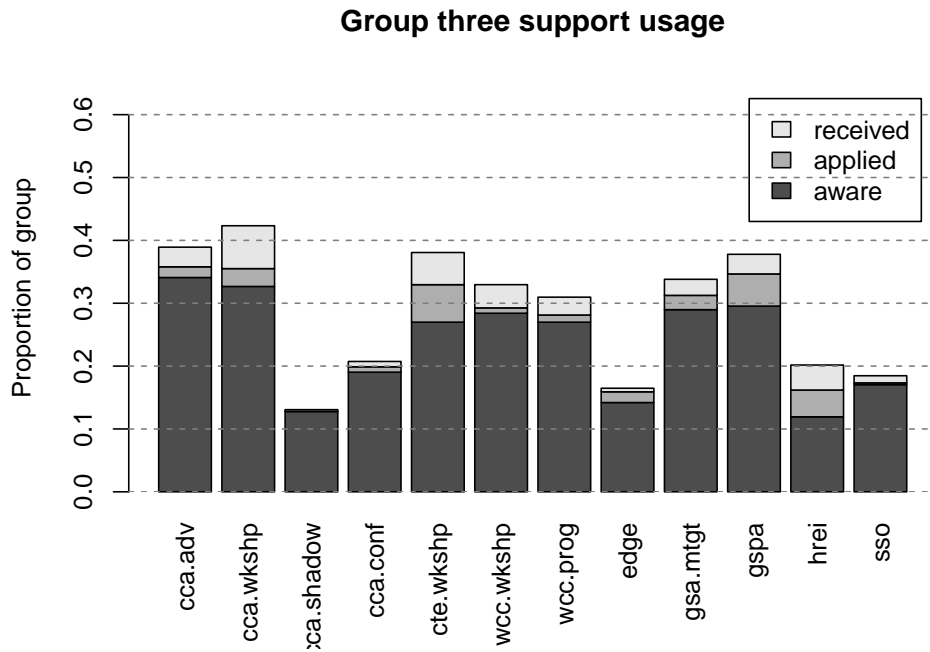


Figure 32: Support usage by group 3: domestic, white, first year master's students not in engineering or math

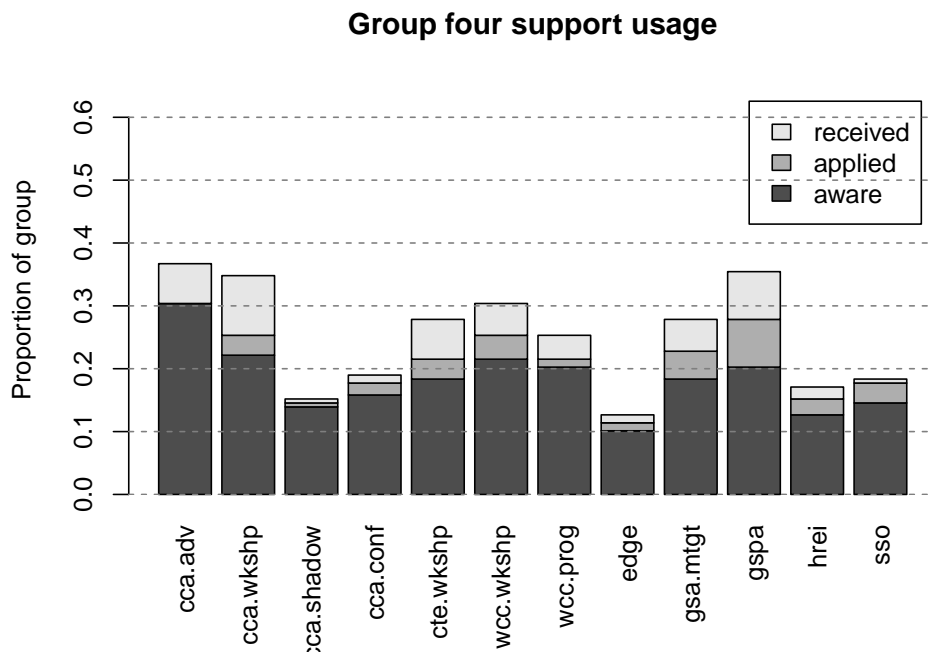


Figure 33: Support usage by group 4: international, first year master's students in engineering and math

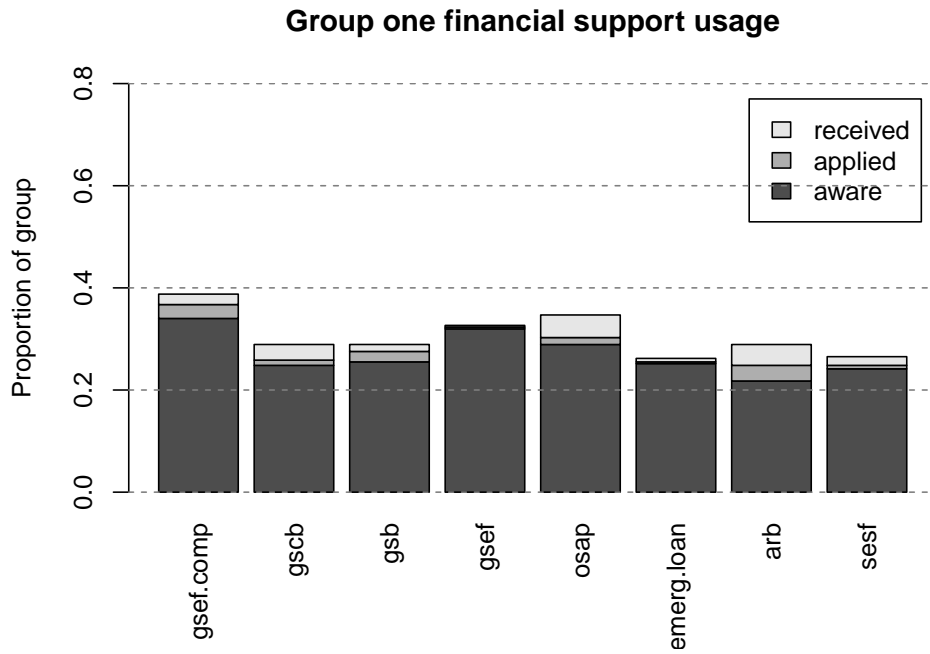


Figure 34: Financial support usage by group 1: international, non-white, upper-year PhD students in engineering and math

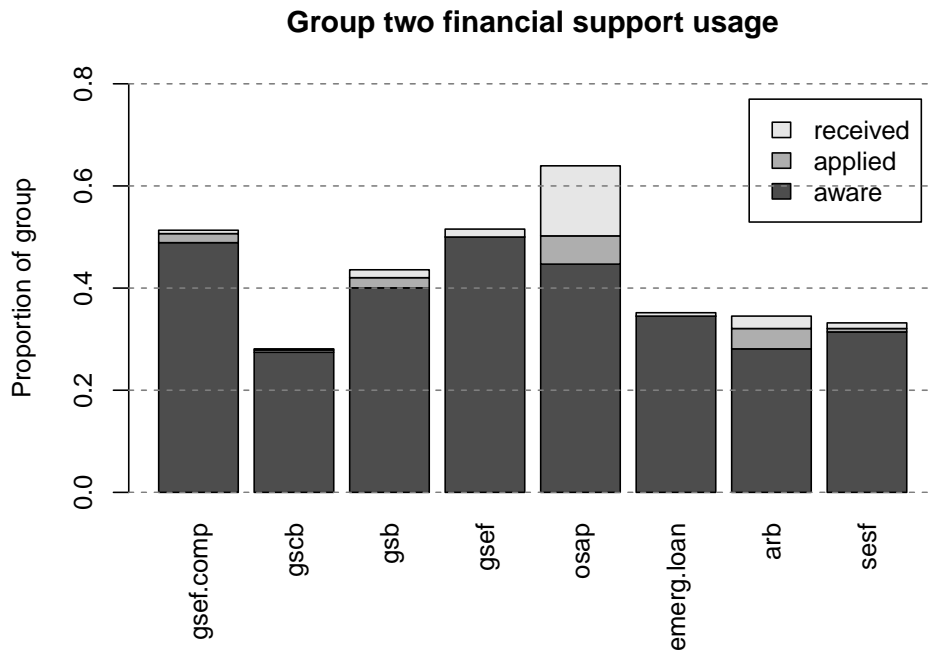


Figure 35: Financial support usage by group 2: domestic, white, upper-year PhD students not in engineering or math

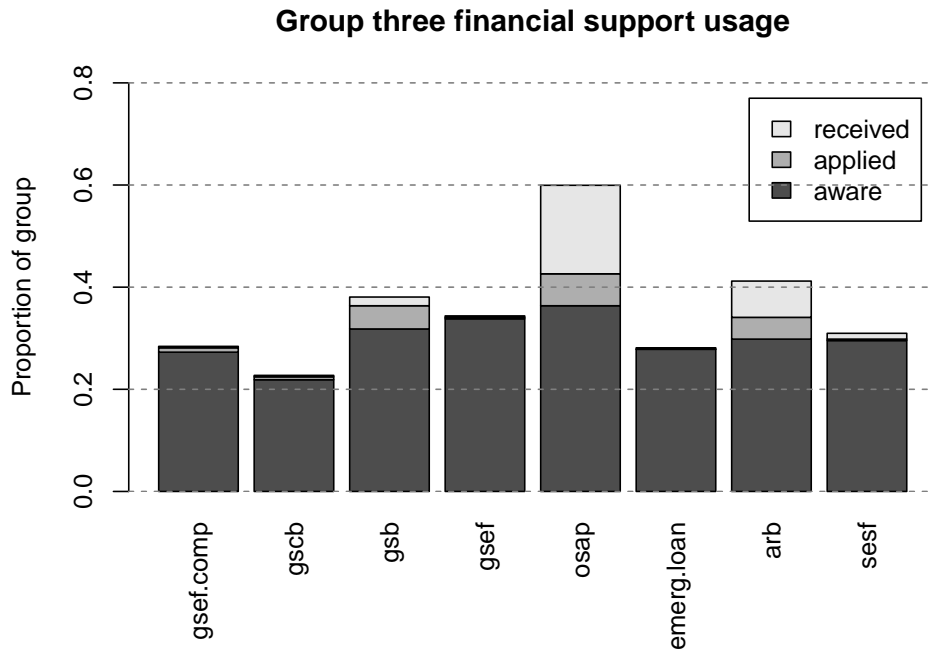


Figure 36: Financial support usage by group 3: domestic, white, first year master's students not in engineering or math

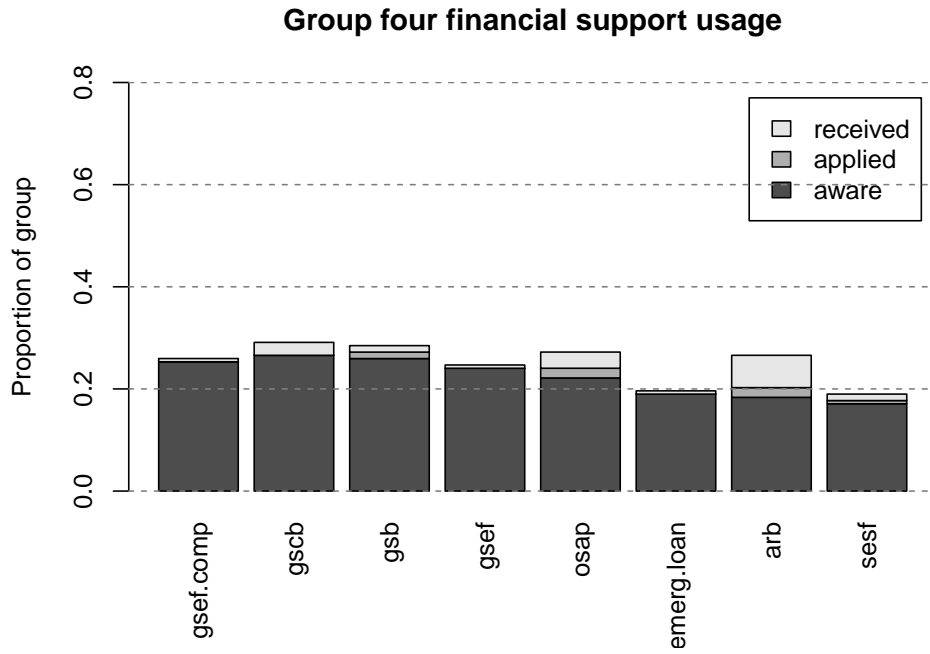


Figure 37: Financial support usage by group 4: international, first year master's students in engineering and math

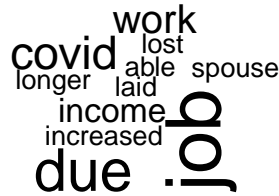
and tendency to apply for these financial supports compared to the general supports. The abbreviations here stand for

1. GSEF Graduate Student Thesis Completion Award
2. Graduate Student Contingency Bursary
3. Graduate Student Bursary
4. Graduate Studies Endowment Fund
5. Ontario Student Assistance Program
6. Emergency loans
7. Academic Readiness Bursary
8. Student Emergency support Fund

## Text analysis

All of the above analysis was performed on the “simple” questions. These are survey options which did not allow text input from respondents. However, there were many points on the survey where questions which required text input were asked, and the final piece of the puzzle for analyzing this survey requires some examination of these questions.

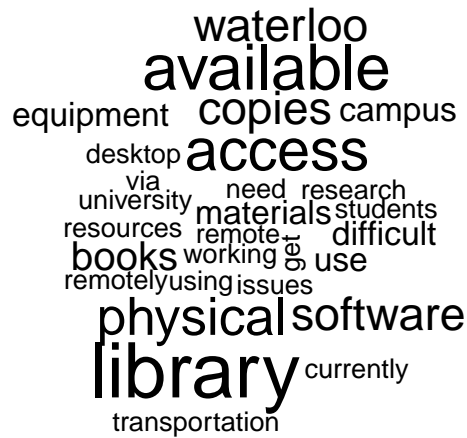
Consider the first such instance of such a question outside of the demographics, that of providing a reason for lost income other than the ones provided in the survey by default. Across all groups, only 36 individuals provided such text input, and so not breakdown into the groups will be performed. Instead, a word cloud of the most common terms is displayed below.



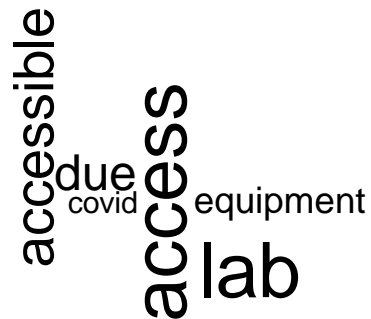
It would seem that these custom responses all fit under the broader categories supplied earlier, and people used this box as a place to tell their individual stories. These ranged from musicians that could no longer perform to people who lost scholarships. A similar story is present for the custom income factors we allowed respondents to report, where only 33 custom answers were provided.







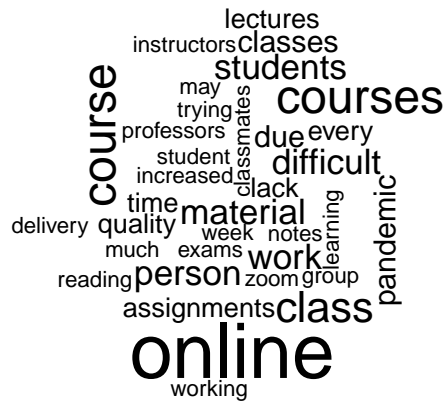
Books and software, which have either been unavailable or only accessible via remote desktop, dominate this section. This provides some degree of detail to the default options provided, though it is not quite distinct from the default options. The 22 lab equipment challenge responses, naturally, had different concerns.



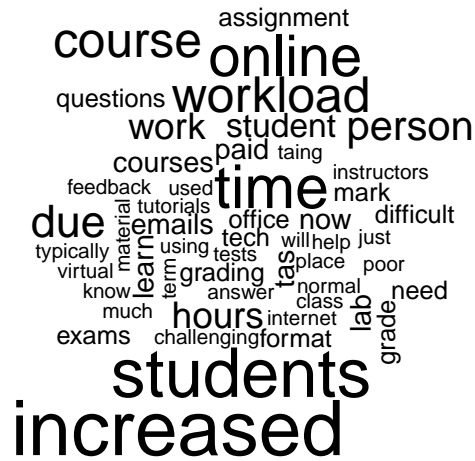
Primarily, the concerns provided here are around access to equipment. This is, categorically, one of the provided options. We can see the same desire to share specific stories in these free-response questions. A completely understandable one given the current state of things.

An opportunity was provided for students who have had to change their project plans to provide feedback about the challenges they have faced in this process. 79 students decided to provide such feedback, predominantly from groups one and two. The wordcloud below displays “stemmed” words, that is words with endings removed to associate like terms.





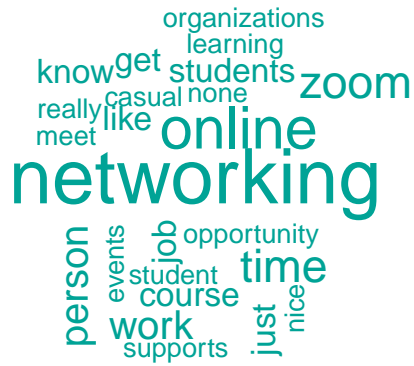
The same is again true of the 52 responses to the TA challenges text input question.



It would seem these 52 respondents have experienced an increased workload marking assignments and providing feedback online. As with all of these other questions, however, we must remember that these word clouds represent a sliver of all respondents, and do not change our interpretation of the default provided options. The default options provided the ability to select an increased workload, and the numbers from those figures should be taken over the impression left from this word cloud.

The 81 responses to the mental health text input prove more informative.





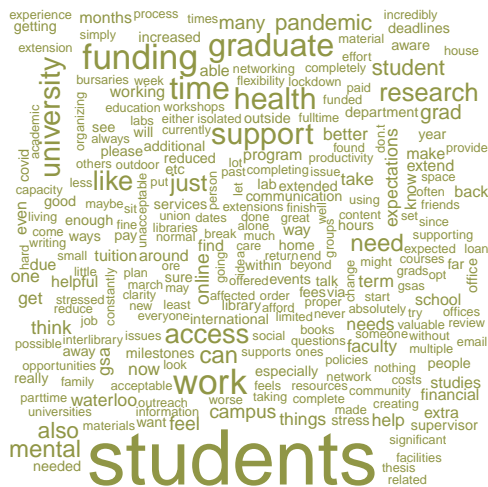
events  
meet

Rather unambiguously, the dominant request among students is for more networking supports. Groups two and three, in particular, seem to want more opportunities provided to network professionally with their peers and external organizations.





Finally, we can consider the ideas suggested by each group.







```

##                                     172
##                                east.asian, declined
##                                     1
##                                east.asian, se.asian
##                                     3
##                                east.asian, white
##                                     6
##                                     latino
##                                     30
##                                latino, mid.east, white
##                                     1
##                                latino, se.asian, declined
##                                     1
##                                latino, white
##                                     8
##                                mid.east
##                                     95
##                                mid.east, white
##                                     5
##                                other
##                                     19
##                                se.asian
##                                     59
##                                se.asian, white
##                                     3
##                                se.asian, white, other
##                                     2
##                                south.asian
##                                     60
##                                south.asian, se.asian
##                                     1
##                                south.asian, white
##                                     2
##                                white
##                                     566
##                                white, declined
##                                     1
##                                white, other
##                                     5

```

All of the above combinations with comma separated racial and ethnic identities were given “multiple” in analysis.

## Appendix: MCA

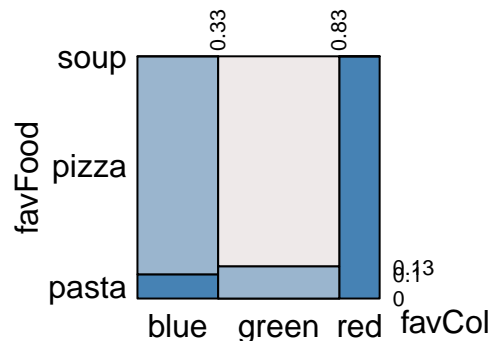
The core idea of multiple correspondence analysis is that a large number of interrelated factors can be represented by a smaller number of synthetic factors. Rather than view faculty, or program, or race directly, we introduce a smaller number of mathematical objects which reflect a majority of the original variation.

Consider an illustrative example. Suppose, rather than demographics, we were collecting a sample of the favourite colour and favourite food of a group of people. Let’s say that when all the collection is done, we end up with the following table:

```
##          favCol
```

```
## favFood blue green red
##  pasta    1    0  5
##  pizza    9    2  0
##  soup     0   13  0
```

Which has the eikosogram



The important thing to notice about this data is that the choice of favourite food and choice of favourite colour have a pretty strong relationship. Most people that like blue prefer pizza, most people that like red prefer pasta, and most people that like green prefer soup. We could, then, try introducing a new variable without any real meaning. For each person, simply paste their two favourites together, and call this “favourite”. When we look at this combined variable, we get the following one-dimensional table.

```
## favourite
##  pasta blue  pasta red  pizza blue  pizza green  soup green
##           1           5           9           2           13
```

We might then further simplify this by noting that most people answered “pasta red”, “pizza blue”, and “soup green”, and so we can capture most answers (but not all answers) by dropping the two options “pasta blue” and “pizza green”. For the people that answered “pasta blue” and “pizza green”, we can match one of their favourites to these new categories, say colour. We could then add a second variable, call it “realFood”, which indicates the real favourite of these three people we have deliberately placed in the wrong group. By doing this we would gain a new pair of variables that look like

```
##           realFood
## favourite  pasta pizza
##  pasta red  5    0    0
##  pizza blue 9    1    0
##  soup green 13   0    2
```

Note how most of the responses now fit in the first column of the table, and we can still determine the favourite food and colour of any person we might want! What we have done is taken a response which required two variables before, both of which varied quite a bit among respondents, and created two new variables. Critically, however, we have taken most of the difference in responses and put it in the first variable “favourite”.

“favourite” on its own can explain much more about the data than either of “favFood” or “favCol” could have on their own. We can see that 27 of the 30 responses are given perfectly by “favourite” on its own. What this means is that we have condensed two variables with three levels each into one variable with three levels with

minimal loss of information. In some sense, we have identified a “direction” or “dimension” in which the data predominantly varies.

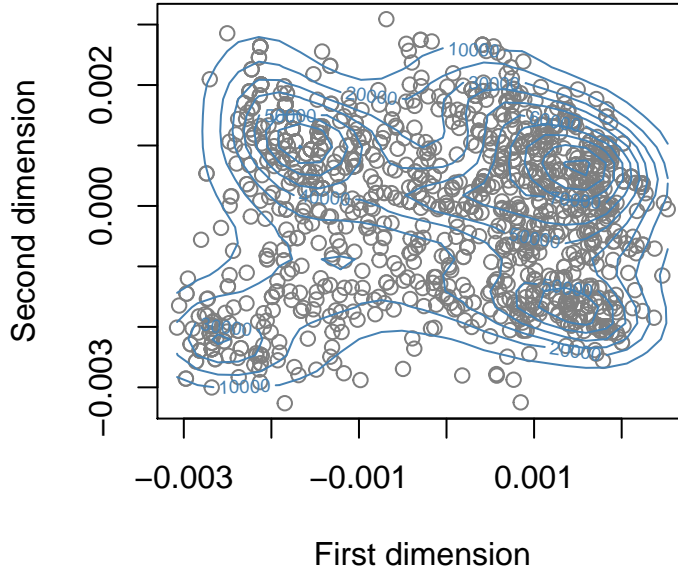
We also have made clear to ourselves and anyone else that there are certain aspects of the data we cannot comment on. We have no idea what people who like the colour red and prefer soup are like, because we did not see any of them. This simplification prevents such unfounded extrapolations to these unobserved combinations by limiting the variation we can model. Hopefully, this can help us to avoid fooling ourselves into claims we cannot support.

Of course, this simple sketch of the idea does not at all reflect the mathematics going on to create these dimensions in our much more complex real data. The new variables created by multiple correspondence analysis are nowhere near as interpretable or simplistic as “favourite” is above. The point is to give a sense of what is meant by “dimension reduction” as a result of multiple correspondence analysis. Dimension reduction is the replacement of many variables with complicated relationships with a few variables that have simpler ones.

### A natural grouping

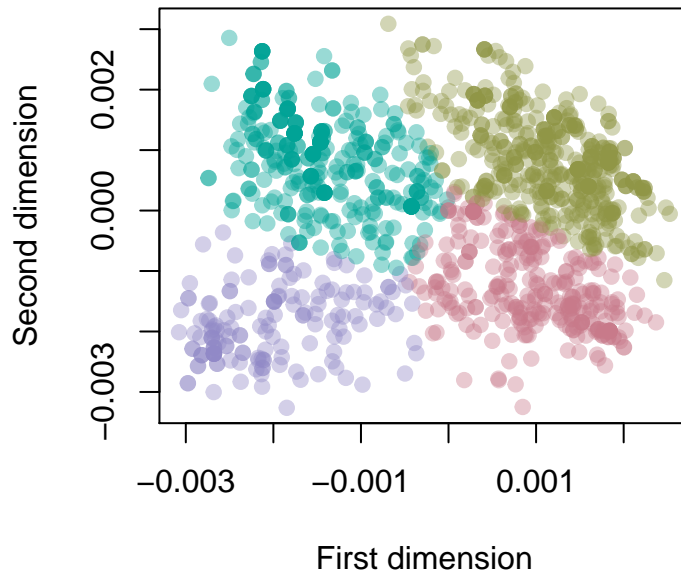
An early exploration of this method applied to our data, in the hope that it might simplify our confounding problem had a rather interesting result. It turns out, when you apply this analysis to the demographic variables of the November survey, a rather interesting structure can be found in the data. A plot of the first two reduced dimensions shows a rather interesting clustering of responses into four regions!

### Responses in reduced dimension



Of course, these regions do not have crisp borders. Just as our respondents are all unique and would resent being forced into any one box, so are their responses.

Nonetheless, the demographic variables, when placed into a reduced dimension representation using multiple correspondence analysis, naturally have this rather intriguing structure. Natural curiosity over this finding entreats us to look deeper. Consider using these four areas of concentration as the basis of four groups.



The borders of these groups, while mathematically optimal, are still somewhat arbitrary. We should expect that the gross simplification of the diversity of ten demographic variables to four groups will have some mis-attribution. We must simply hope that whatever complexity remains in these groups does not impact our ability to gain actionable insight. The loss of detail is in the hopes of gaining statistical robustness while controlling for confounding as much as possible.

We must therefore check whether these groups make any sense in the original demographic variables. While all may be good in the land of mathematics, we cannot sacrifice interpretability. That has been the point of this whole exercise. The following several eikosograms display how our demographic variables are distributed across the groups. By viewing these together, we can see whether these groups capture the features of interest in the data.