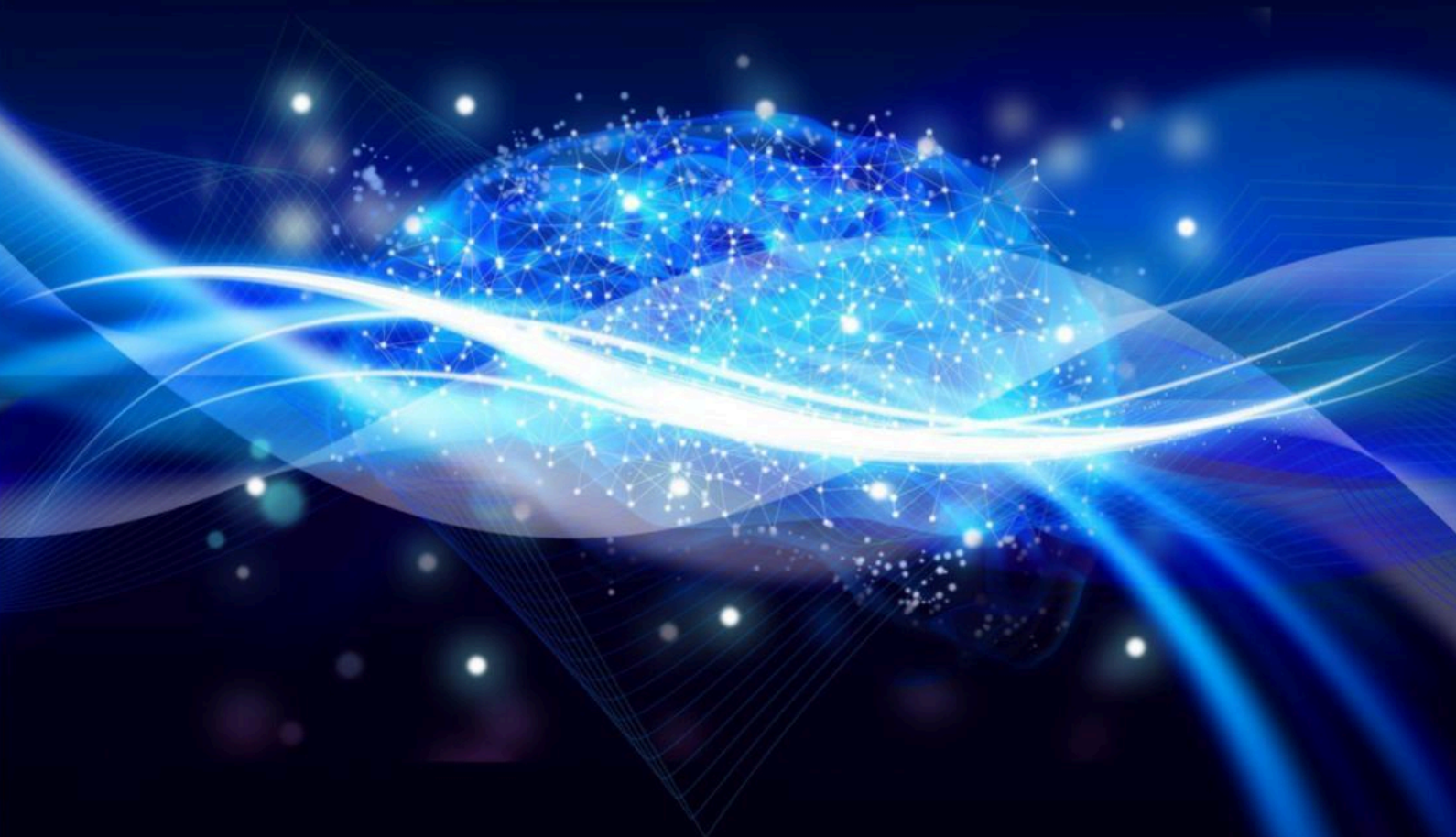
INTERSECT
RESEARCH FASTER

Intersect Australia

Research Digital Skills Training Program

2023 Report



Authors

Dr Anastasios Papaioannou[1], Dr Katherine Howard[1,2], Charlotte François[1], Peta Scott[1,3], Aidan Wilson[1], Dr Jonathan Arthur[1]

[1] *Intersect Australia, Sydney, Australia*
[2] *University of South Australia, Adelaide, Australia*
[3] *University of New England, Armidale, Australia*

Keywords

data literacy, digital skills, education, training, researcher training, digital tools, digital technologies, data science, data management, surveys, machine learning, statistics, databases, high-performance computing

Licence

CC BY-NC-ND

BY – Credit must be given to the creator
NC – Only noncommercial uses of the work are permitted
ND – No derivatives or adaptations of the work are permitted

Copyright

This report was published on 12 April 2024. Copyright © 2024 Intersect Australia Limited

Table of Contents

1. Executive summary **5**
2. Attendance overview **6**
2.1 Attendance by Tool/Technology 7
2.2 Attendance by Role/Position 8
2.3 Attendance by Faculty 9
2.4 Attendance by Field of Research (FoR) code 11
2.5 Return to attend more Intersect courses 13
2.6 Reason for attending 14
2.7 No shows 16
3. Evaluation **18**
3.1 Evaluation by Tool/Technology 18
3.2 Evaluation by month and class size 19
3.3 Metrics for measuring quality of delivery 20
3.4 Qualitative feedback and evaluation 20
What were the most useful aspects of the training? 21
Least useful aspects of training 22
Other suggestions or feedback 22
4. Communication **24**
5. Historical trends **26**
Appendix A - Intersect training catalogue **33**

ATTENDANCE KEY STATS

314 COURSES DELIVERED

5834 TOTAL ATTENDEES

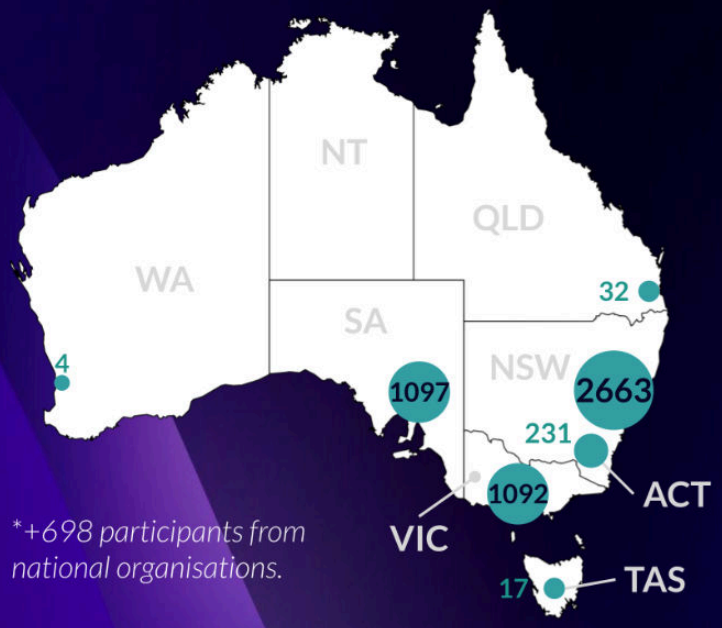
4659 ATTENDEES At their local university

1175 ATTENDEES At other Intersect members

1.2+ COURSES DELIVERED EVERY WORKING DAY

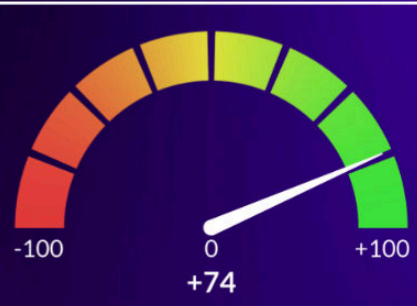
18.6 ATTENDEES PER COURSE

2023 ATTENDANCE BY STATE/TERRITORY



*+698 participants from national organisations.

EVALUATION



*Based on 1879 survey responses

How likely is it that you would recommend Intersect training courses to colleagues?

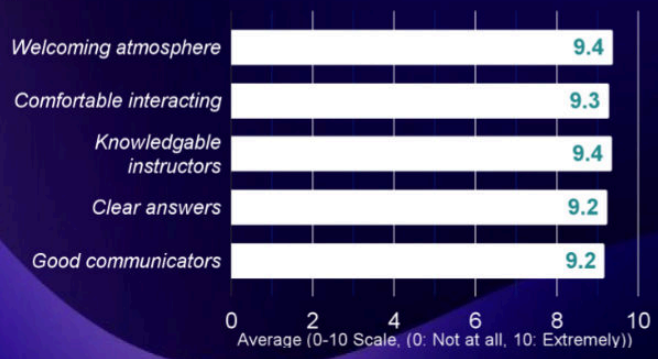
Promoters: 9 - 10

Passives: 7 - 8

Detractors: 0 - 6

Net Promoter Score = %Promoters - %Detractors

QUALITY



QUALITATIVE FEEDBACK

Intersect's pedagogical approach to training (hands-on, practical, scaffolded learning) **is recognised by participants** as providing a positive learning environment

Courses are **delivered by knowledgeable, expert and professional trainers** who provide a welcoming space in which to learn

"The practical elements where we acted out on sample data. Why? Because learning isn't a spectator sport ;) [...]"

1. Executive summary

Intersect's mission is to ***help researchers pursue a better future for all***. This includes helping researchers to be more efficient and effective in their research. As a leading provider of digital skills training for researchers in the Australasian region, Intersect provides an extensive range of technology-focused training courses to researchers and higher degree research (HDR) students across Australia. This training ranges from awareness to advanced levels; is delivered interactively either face-to-face or online; and covers categories such as Research Computing, Programming, Data Science, Data Analytics, Machine Learning (ML) & Artificial Intelligence (AI), Statistics, Data Visualisation, Data Collection, and Data Management. Intersect continually revises, updates, and expands its course catalogue, ensuring researchers always have access to the most relevant and useful digital research training.

Our hands-on, instructor-led, live, interactive training is delivered by over 25 highly experienced instructors and is targeted at enhancing the capabilities of researchers in digital tools and technologies. Our research and training expertise extends across various disciplines including, but not limited to: ICT, Data Science, Linguistics, Engineering, Statistics, Bioinformatics, Health & Medical Sciences, Materials Science, Sports Science, Spatial Analytics, Computational Chemistry, Numerical Modelling, Behavioural Science, and Social and Political Sciences.

Throughout 2023, Intersect facilitated training for 5,834 individuals across 314 courses, totaling 242.5 training days. Intersect members are located in NSW, VIC, SA, and ACT, where the majority of the training program takes place. However, our reach extends nationwide, due to the NCI-Intersect Training Partnership. The NCI-Intersect training partnership includes a series of training courses aimed at users who may be new to NCI's high-performance computing environment, as well as experienced users looking for a refresher.

Notably, our courses attract attendees from diverse backgrounds, including HDR students, researchers, and professionals from various sectors.

In 2023, Intersect received overwhelmingly positive feedback, with a Net Promoter Score (NPS) of +74, based on 1,879 responses (~32% of attendees). The average scores for training quality metrics exceeded 9.2 out of 10, reflecting excellent participant satisfaction.

This report summarises the Research Digital Skills Training Program provided by Intersect Australia across Australia for the calendar year 2023 as well as providing useful insights on historical trends.

2. Attendance overview

The vast majority of Intersect training courses are scheduled between March and November as shown in Figure 2.1, with a slightly smaller number of courses being scheduled in February, and only 1-2 in December due to the holiday period. The three busiest months were May, March and July during which a total number of 114 courses were delivered and around 2,100 attendees were trained.

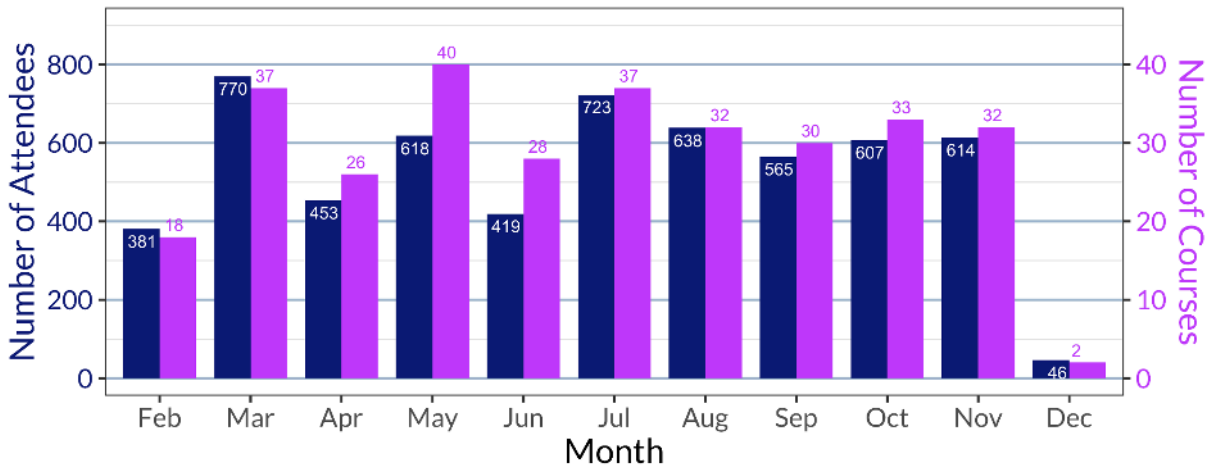


Figure 2.1: Number of attendees trained (left y axis) and number of courses delivered (right y axis) by month as part of the Intersect Research Digital Skills Training Program in 2023.

Figure 2.2 shows the distribution of courses delivered and number of attendees trained by the day of the week. Most Intersect courses were scheduled on Tuesday, followed by Wednesday and Thursday. This is consistent with the training delivery in 2022. A smaller number of courses was scheduled on Friday, and only a very small number of courses were delivered on Monday, primarily due to internal and other commitments for the Intersect Digital Research Analysts, who are responsible for the scheduling of courses.

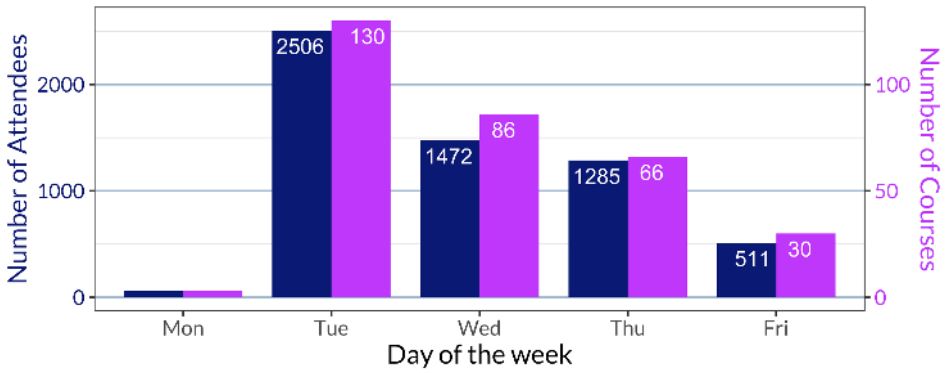


Figure 2.2: Number of attendees trained (left y axis) and number of courses delivered (right y axis) by day of the week as part of the Intersect Research Digital Skills Training Program in 2023.

There are two different types of courses that Intersect delivers: a) courses that are delivered by the Digital Research Analyst solo (usually with less than 15 attendees); b) courses that are delivered by 2 or 3 trainers (one lead instructor and one or two assistant trainers, usually with 25+ attendees). Intersect's trainer:trainee ratio is between 1:10 (one trainer per 10 attendees) and 1:15 (one trainer per 15 attendees). Therefore, the class size

varies from as few as 8 attendees to 30+ attendees per course based on the type of the courses and the attendance rate. The distribution of courses delivered in 2023 by class size is shown in Figure 2.3. Based on the distribution below, the majority of courses include 12 to 22 attendees.

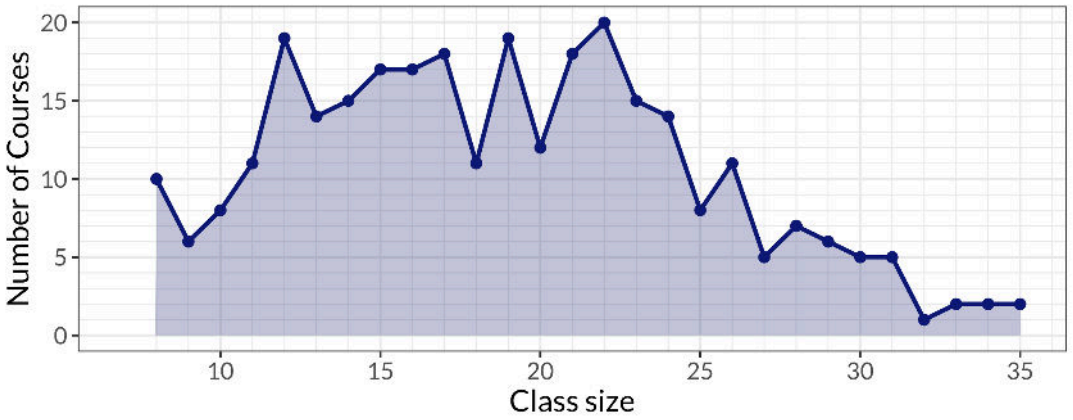


Figure 2.3: Distribution of number of courses by class size (number of people attended the course) for the Intersect Research Digital Skills Training Program run in 2023.

2.1 Attendance by Tool/Technology

Figure 2.1.1 shows the percentage of attendees that were trained in 2023 by Tool/Technology. The tools and technologies that are beyond the top ten with the highest attendance were aggregated and presented as "Other tool". Python and R programming courses were the most popular with more than 50% of attendees participating in these courses. REDCap, Excel, NVivo, and Qualtrics were the next most popular courses with a percentage of attendees between 5% and 10% each. The courses were spread across the year, however the highest number of the courses were delivered in Quarter 3. The most popular tool in Quarter 1 was R followed by Python, while this reversed in Quarter 2-4 with more attendees participating in Python courses.

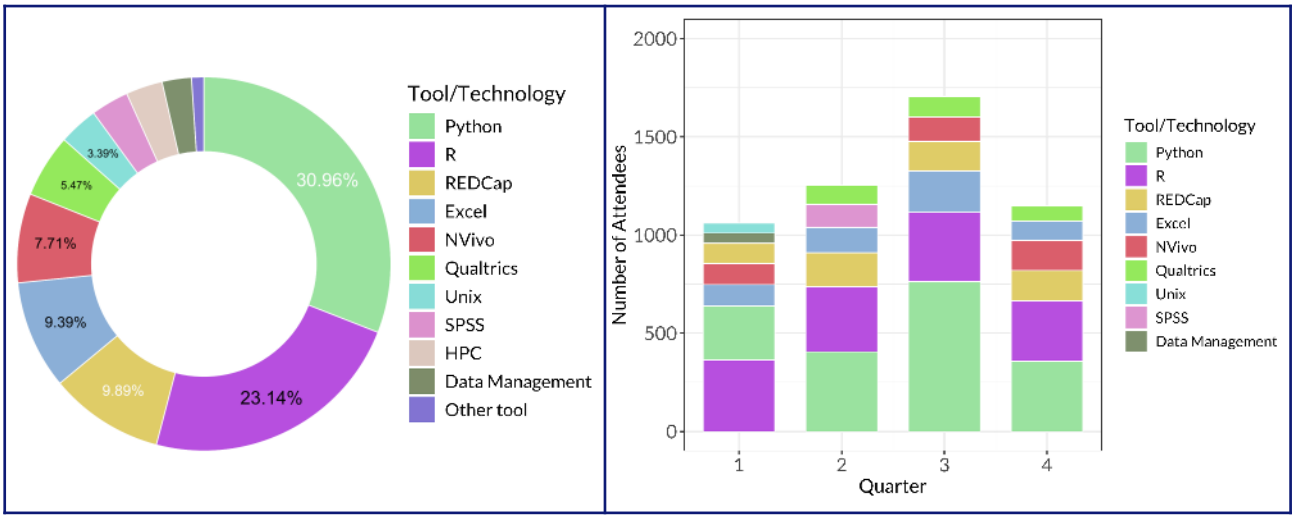


Figure 2.1.1: Left: Percentage of attendees by Tool/Technology; Right: Quarterly number of attendees by Tool/Technology (only top six tools/technologies per quarter are presented).

2.2 Attendance by Role/Position

Figure 2.2.1 (left) shows that Higher Degree Research (HDR) students (PhD) are the top consumer of Intersect's training program in 2023 making up 47% of the total attendees. The second highest consumer is Academics comprising approximately 14% of attendees followed by Post-doc/Fellow and Professional (research-related) with approximately 10.7% and 10.5%, respectively. PhD students were consistently the top consumer across all four quarters.

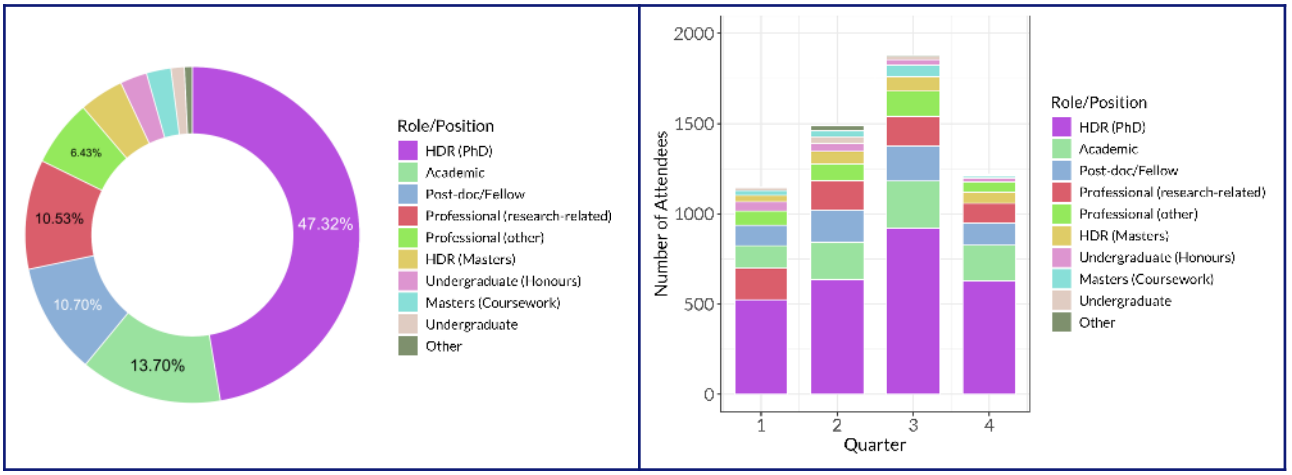


Figure 2.2.1: Left: Percentage of attendees by Role/Position; Right: Quarterly number of attendees by Role/Position.

Figure 2.2.2 shows the percentage of attendance by tool and technology and then Role. Interestingly, the distributions vary when comparing different roles with different tools/technologies. PhD students are the top consumers for all tools and technologies. Moreover, Academics show a greater uptake of the surveying tools (REDCap and Qualtrics), MATLAB, and Data Management courses. The highest percentage of attendance for Post-doc/Fellow is observed in SQL and programming (Python, R and MATLAB) courses. Regarding Professionals (research-related), the highest uptake is shown to be in Unix, HPC, and REDCap, mainly due to the training program delivered to the professional and scientific staff at NSW Department of Climate Change, Energy, the Environment and Water (DCCEEW).

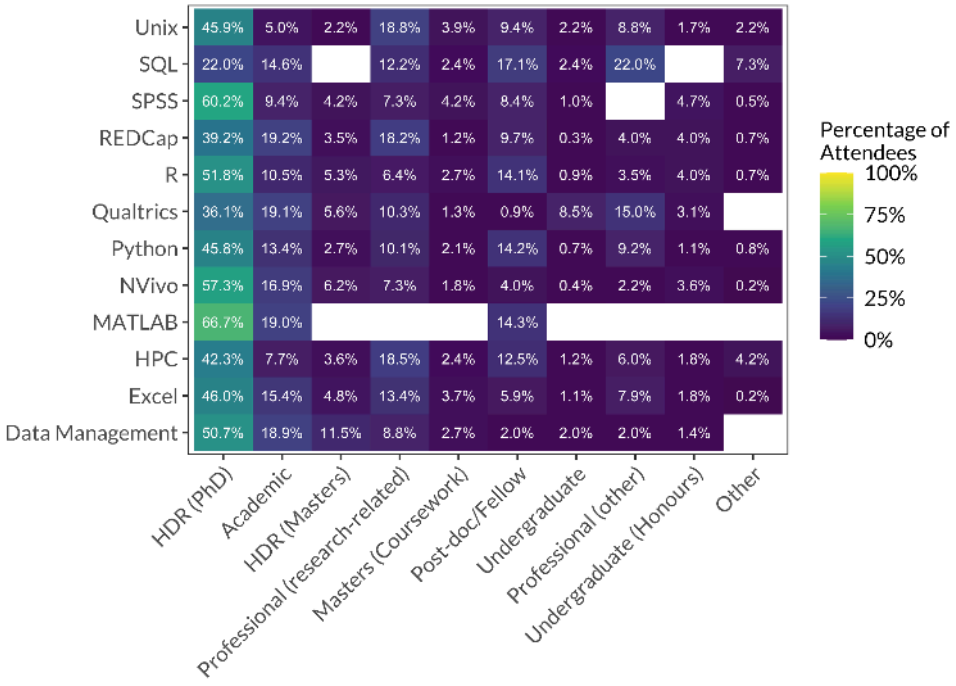


Figure 2.2.2: Percentage of attendees by tool/technology that is further split by Role.

2.3 Attendance by Faculty

Intersect consistently captures the Field of Research (FoR) code during course registration. As there is no controlled vocabulary or naming convention for university faculties, the FoR codes are mapped to generic faculty names to estimate attendance by faculty.

More than a third (~36.6%) of the training attendees are from the Faculty of Medicine and Health followed by the Faculty of Science, with 27% of the attendees coming from this faculty. In 2023, the Faculty of Engineering was the third biggest consumer closely followed by the Faculty of Arts and Social Sciences which had been the third highest consumer in 2020 and 2021 mainly due to the introduction of NVivo courses.

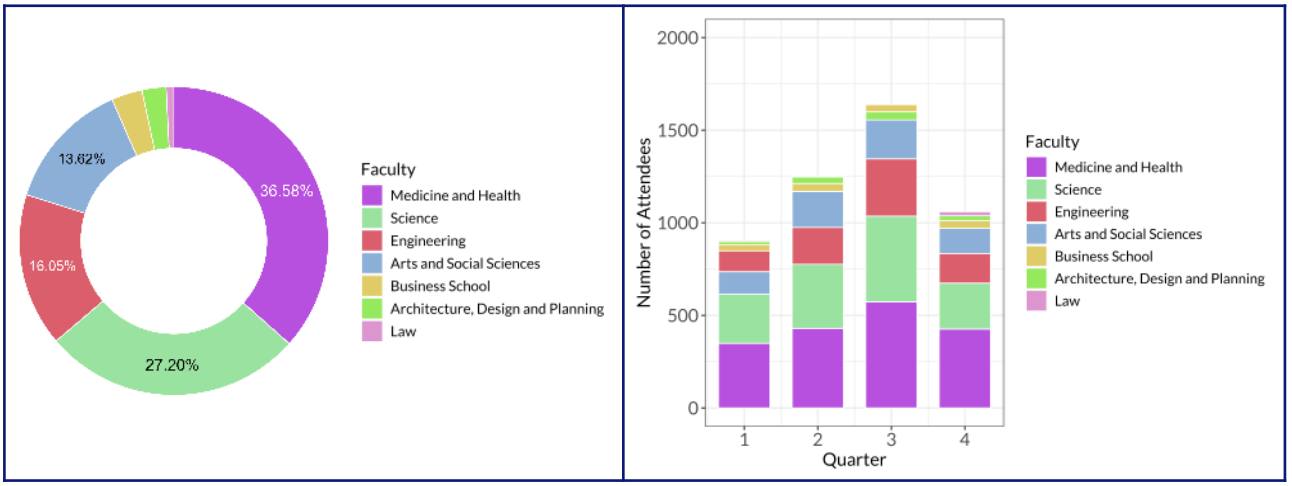


Figure 2.3.1: Left: Percentage of attendees by Faculty; Right: Quarterly number of attendees by Faculty.

Figure 2.3.2 shows the distribution of each tool and technology attendance by faculty. Regarding the programming courses, Python shows a different distribution compared to R with the vast majority of attendees being from the Faculty of Science and Faculty of Engineering. Regarding R courses, the vast majority of participants (approx. 3 out of 4 participants) are from the Faculty of Medicine and Health and the Faculty of Science. The highest uptake for the MATLAB courses was observed in the Faculty of Engineering with about 43% of attendees coming from this faculty. The Faculty of Arts and Social Sciences is one of the largest consumers of the survey tool Qualtrics and the qualitative analysis tool NVivo. Interestingly, the Faculty of Medicine and Health is the largest consumer of NVivo courses surpassing the Faculty of Arts and Social Sciences. Attendees from the Faculty of Arts and Social Sciences have also shown an increased interest in attending SPSS courses. As for the other survey tool, REDCap, 66% of the attendees are from the Faculty of Medicine and Health.

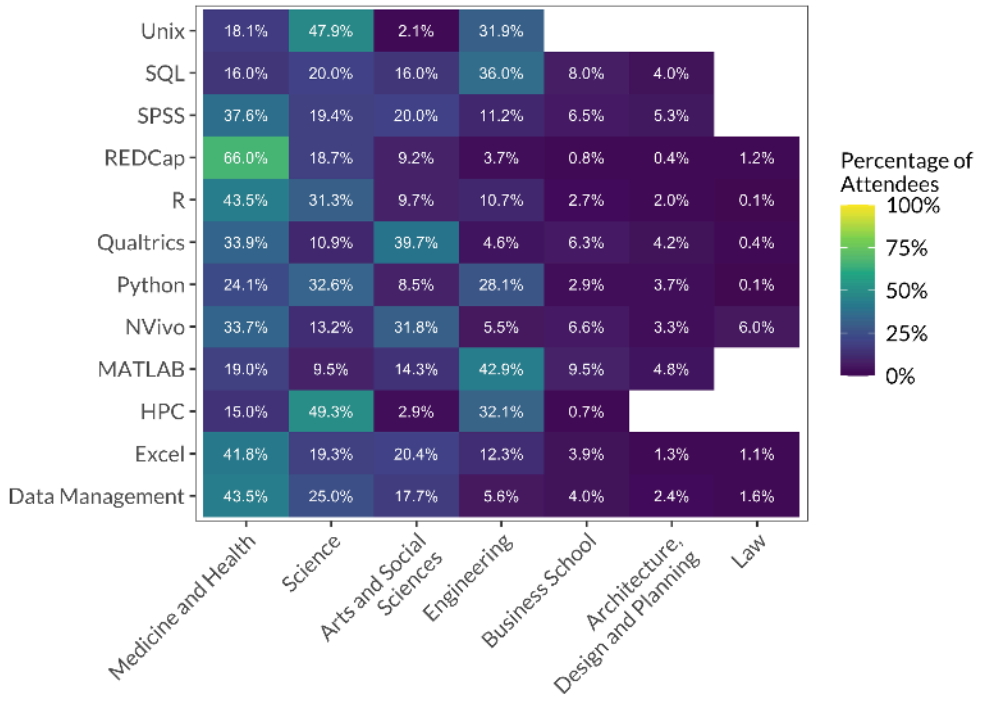


Figure 2.3.2: Percentage of attendees by tool/technology that is further split by Faculty.

Figure 2.3.3 shows the sankey diagram of the flow from Role/Position (top five) to Faculty and then to Tool/Technology (top ten). HDR (PhD) students is consistently the largest cohort among all faculties. Compared to the average percentages (see Figure 2.2.1 left), a higher number of Academics from the Faculty of Law, the Business School, the Faculty of Arts and Social Sciences, the Faculty of Architecture, Design, and Planning, and the Faculty of Medicine and Health have attended Intersect courses with a value between 16.5% and 39% of the total attendees of each faculty. The highest percentages of Post-doc/Fellows is seen in the Faculty of Science and the Faculty of Medicine and Health (a percentage between 12.5% and 15%), while for Professional (research-related) is in the Faculty of Law (approx. 18.5%). As mentioned, the largest consumer of Intersect training courses is the Faculty of Medicine and Health with about 36.5% of total attendees (see Figure 2.3.1 left). Participants from the Faculty of Medicine and Health attended mostly R courses (~31%), followed by Python (19.3%), REDCap (17.5%), and Excel (10.7%). HDR students and researchers from the Faculty of Science are predominantly interested in learning Python and R with a combined percentage of about 65% of total attendees from this faculty attending R and Python courses. Notably, over 50% of attendees from the Faculty of Engineering participate in Python courses, and to a lesser extent R courses (~17%). Interestingly, the

Faculty of Arts and Social Sciences shows a more diverse interest in learning different tools and technologies, with the top five tools being from highest to lowest (19.9% - 11.9%), Python, R, NVivo, Excel, and Qualtrics. This demonstrates the increased interest of HDR students and researchers from the Faculty of Arts and Social Sciences to learn programming languages, which might be attributed to Intersect's multiyear effort to enhance the adoption of various tools within this faculty.

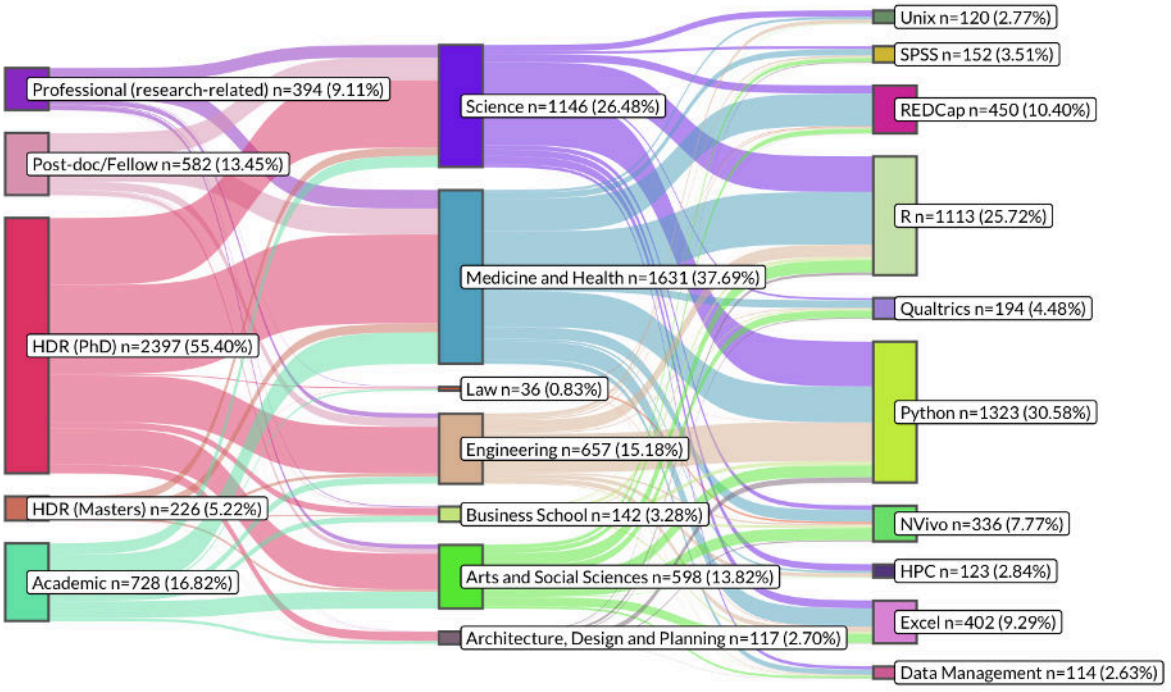


Figure 2.3.3: Sankey diagram depicting the flow from the top five Roles/Positions of attendees to the different Faculties and then to the ten most popular Tools/Technologies.

2.4 Attendance by Field of Research (FoR) code

Figure 2.4.1 shows the distribution of attendees based on the Field of Research (FoR) code. Note that the FoR codes that are not in the top ten highest number of attendees were aggregated and presented as “Other FoR Code” in Figure 2.4.1. Health Sciences is on top of the list with about 21% of the attendees, while Other/Not Applicable and Engineering are the second and third FoR codes with about 13.2% and 10.1% of the attendees, respectively. All the other FoR codes show a percentage between 4% and 7.5%.

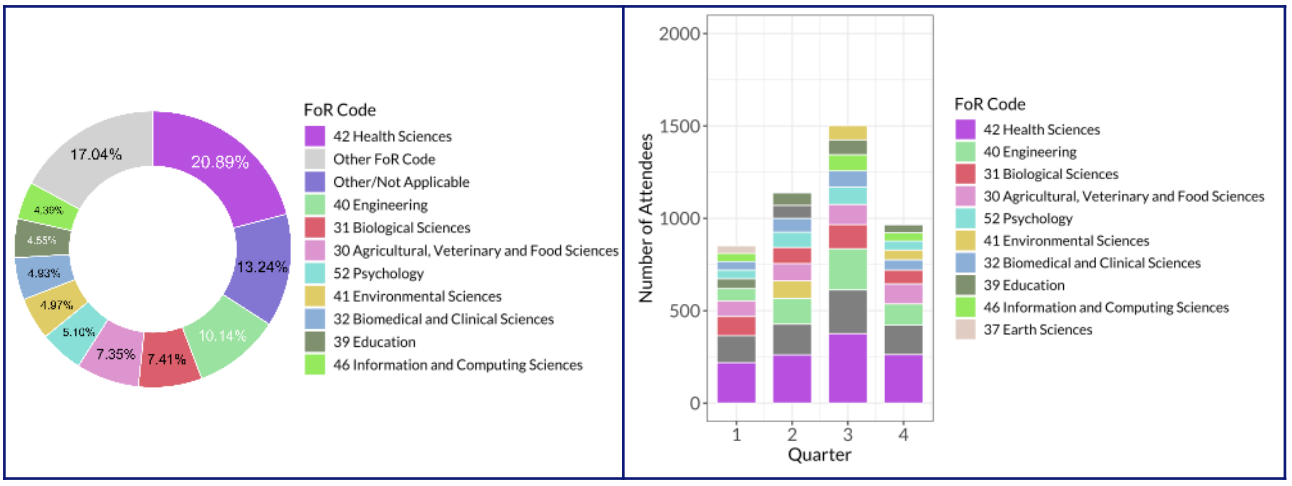


Figure 2.4.1: Left: Percentage of attendees by FoR Code; Right: Quarterly number of attendees by FoR Code.

Figure 2.4.2 shows the distribution of each tool and technology attendance by FoR code. The comparison of distributions for the two most popular tools, R and Python, shows some interesting insights. The highest uptake of R courses is observed in Health Sciences, Biological Sciences, Agricultural, Veterinary and Food Sciences, and Biomedical and Clinical Sciences. In the case of Python courses, the largest consumer is Engineering followed by Health Sciences, Biological Sciences, Agricultural, Veterinary and Food Sciences, and Information and Computing Sciences. It becomes apparent that researchers and HDR students from Medical and Health sciences as well Biological, Biomedical and Clinical sciences prefer to learn R whereas Python is the preferred programming language for Engineering and IT researchers and HDR students. Regarding NVivo, Health Sciences is the top consumer followed by Education, Psychology and Human Society. Regarding REDCap courses, the top consumer by a great extent is Health Sciences while Qualtrics courses are mostly attended by Health Sciences, Education and Economics.

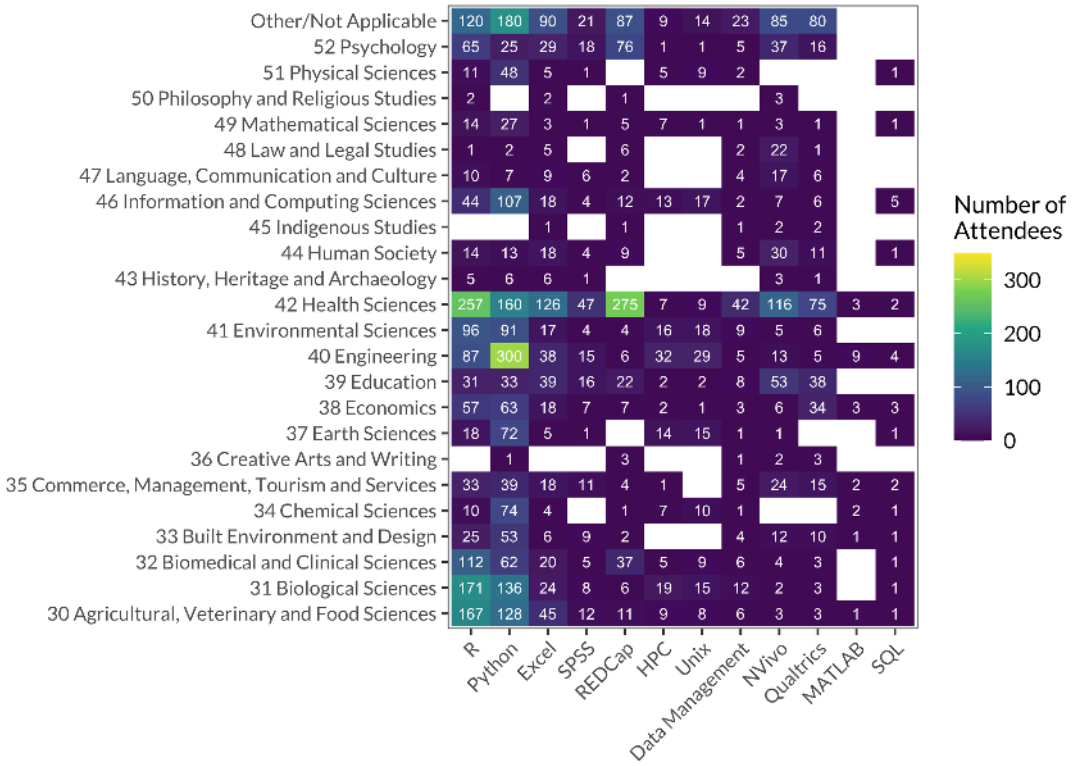


Figure 2.4.2: Number of attendees by tool/technology that is further split by FoR code.

2.5 Return to attend more Intersect courses

Figure 2.5.1 shows the distribution of the number of courses that the same participant has attended in 2023. Around 66% of the participants attended an Intersect training course once in 2023, which means one out of three participants returned to attend another Intersect training course. Approximately 19% have attended two courses, followed by 7.6% who have attended three, and 3.4% have attended four courses. Approximately 4.2% of the total number of attendees have participated in the training program 5 times or more in courses delivered in 2023. Although we present only 2023 data in Figure 2.5.1, please note that it is also very common for the same person to attend various Intersect courses across several consecutive years (see Section 5).

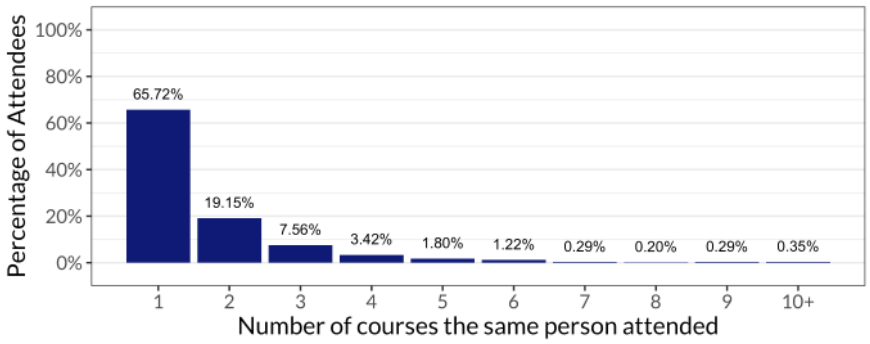


Figure 2.5.1: Number of courses that the same person attended in 2023.

2.6 Reason for attending

When participants register, they are asked to provide reasons for attending the course. Most participants answered that they are interested in these courses to learn skills that they can either apply to their work now (approx. 46%) or in the near future (approx. 32%). Approximately 13% of the attendees are learning these skills for better opportunities and employability in the future ("To learn skills that will help me get a job").

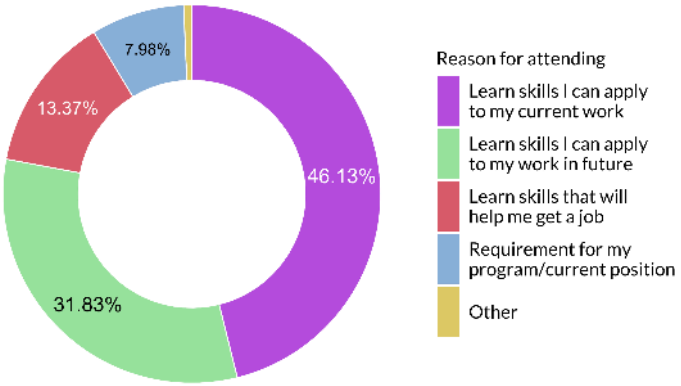


Figure 2.6.1: Reason for being interested in attending this training course.

Analysis of the reason for being interested in attending this training course by the Tool/Technology reveals a similar trend across all tools/technologies. The highest percentage for learning skills that can be applied to their work now is observed in NVivo, SQL, REDCap, and Qualtrics. Regarding learning skills that can help them get a job, all programming courses (R, Python, and MATLAB) show a high percentage compared to other tools/technologies.

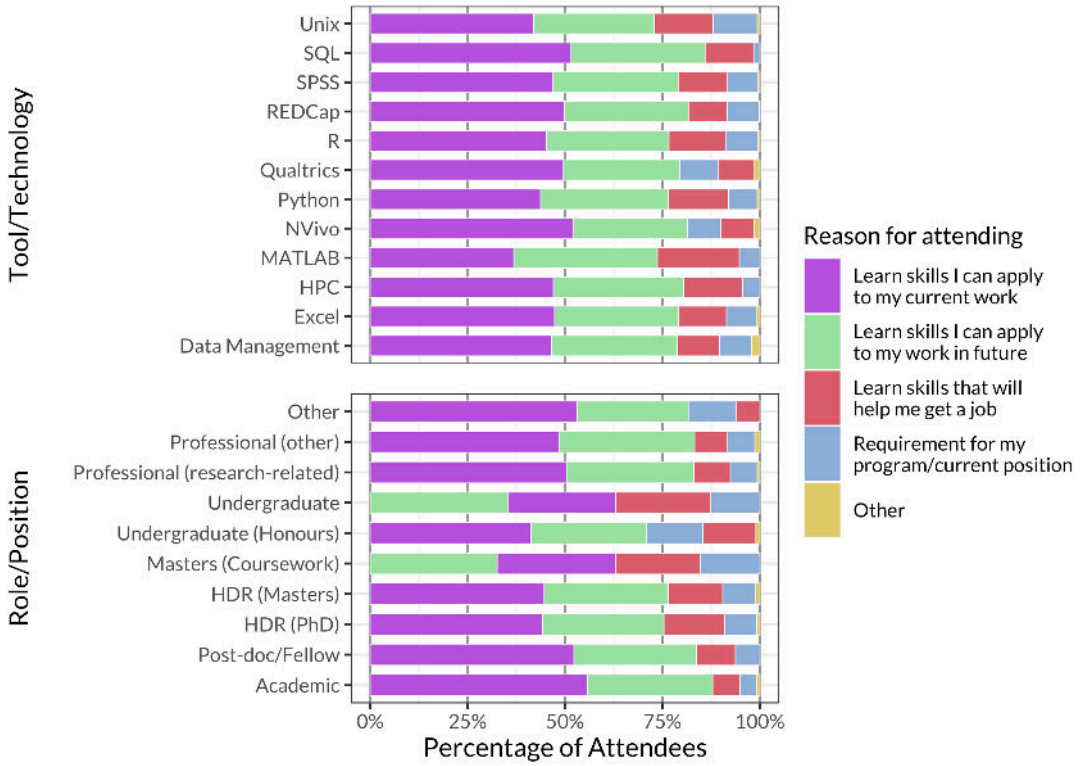


Figure 2.6.2: Top: Reason for being interested in attending Intersect training course by Tool/Technology; Bottom: Reason for being interested in attending Intersect training course by Role/Position.

Further analysis by Role/Position showed that the percentage of participants who learn skills for current work is higher among Academics and Post-docs compared to HDR Masters and PhD students where the percentage of those learning these skills for better opportunities and as a requirement for their program of study is higher.

Figure 2.6.3 shows the sankey diagram of the flow from Role/Position (top five) to the reason for attending the Intersect course and then to Tool/Technology (top ten). All top five roles consistently answered that the top reason for attending an Intersect training course is to learn skills that they can apply to their current work with a percentage that fluctuates between approximately 44% (HDR PhD and HDR Masters) and 56% (Academics). Attending the courses to learn skills that can help in their work in the future was consistently the second top reason among all roles with 31%-33% of responses. Interestingly, 15.7% and 14.0% of HDR PhD and Masters students responded that the reason for attending these training courses is also to gain skills that can help them get a job, and therefore they attend the courses to improve their employability. The percentage for this reason for Post-docs was 10% followed by Professional staff (research-related) and Academic with 9.4% and 7.0%, respectively. Analysing the reason for attending a course by Tool/Technology, a similar trend and percentages are observed for the top two reasons, *e.g.*, to learn skills they can apply to their current work now or in the future. However, several tools/technologies stood out when analysing the response "Learn skills that will help me get a job" by tool with more than 14.5% of registrants mentioning that they want to learn MATLAB, Python, HPC, Unix, and R to help get better job opportunities in the future. This also demonstrates how popular skills programming languages and large-scale infrastructure are when it comes to employability, and in particular among HDR students who are looking for opportunities to upskill themselves in emerging tools and technologies.

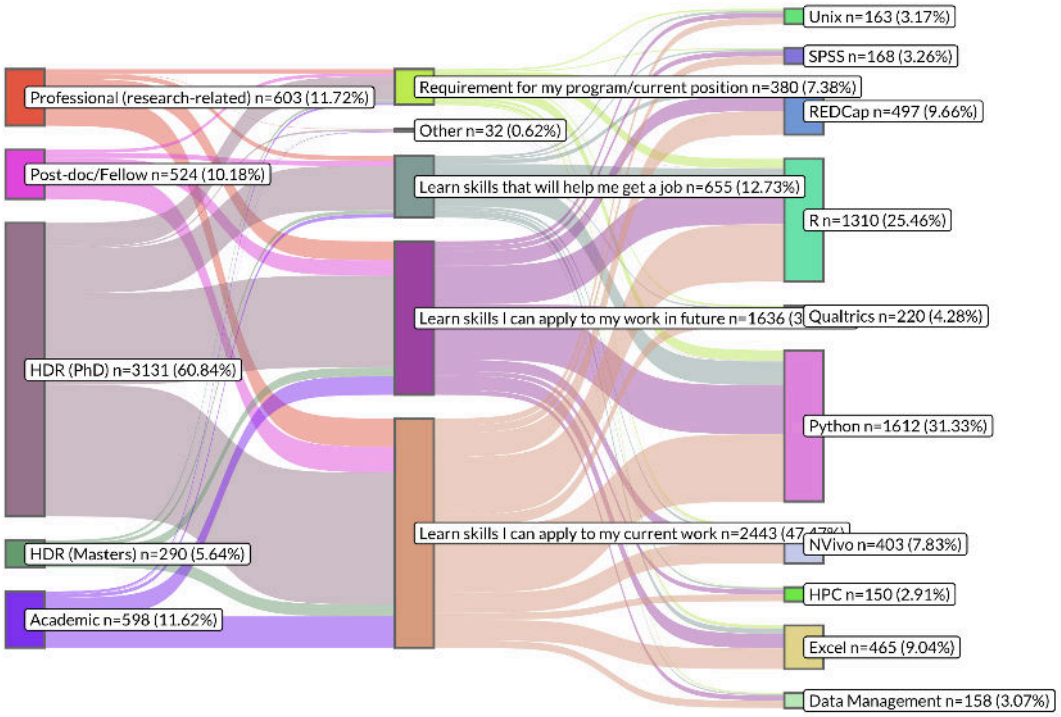


Figure 2.6.3: Sankey diagram depicting the flow from the top five Roles/Positions of registrants to the different reasons for attending a course and then to the ten most popular Tools/Technologies.

2.7 No shows

"No shows", *i.e.*, people who register for a course and then fail to attend, are a common issue faced by Intersect and other similar training providers. Intersect closely monitors the percentage of no shows and has applied various methods, in different use cases, to tackle the issue and minimise the impact.

Figure 2.7.1. shows the average percentage of no shows by month and day of the week. The highest average percentage of no shows was in December followed by June, October, and September during which more than 26% (1 out of 4) of registrants did not attend the training course without notifying the Intersect training team. The months with the lowest percentage of no shows are February and August. Analysing the percentage of no shows by the day of the week on which the courses run, we notice that the lowest percentage no shows rate is observed in courses delivered on Monday, while the percentages in other days of the week are similar, fluctuating between 21.6% and 26.6%.

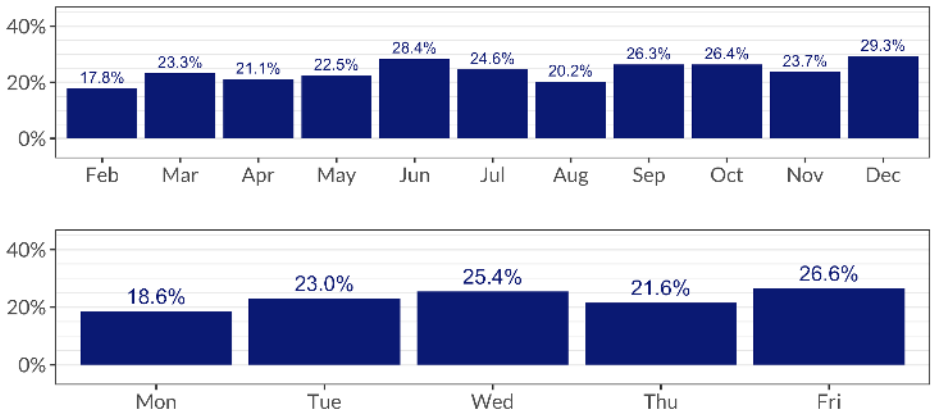


Figure 2.7.1: Top: Average percentage of no shows by the month the course was delivered; Bottom: Average percentage of no shows by the day of the week the course was delivered.

Some more useful insights can be derived by plotting the distribution of no shows by month and day of the week as shown in Figure 2.7.2. Monday is the day with the lowest median percentage of no shows, while Friday is the highest one. When comparing the distributions by month, the first 5 months of the year display lower median percentages of no shows compared to months after, with August being an exemption. February is shown to have the lowest median percentage of no shows while June, October and December have the highest one.

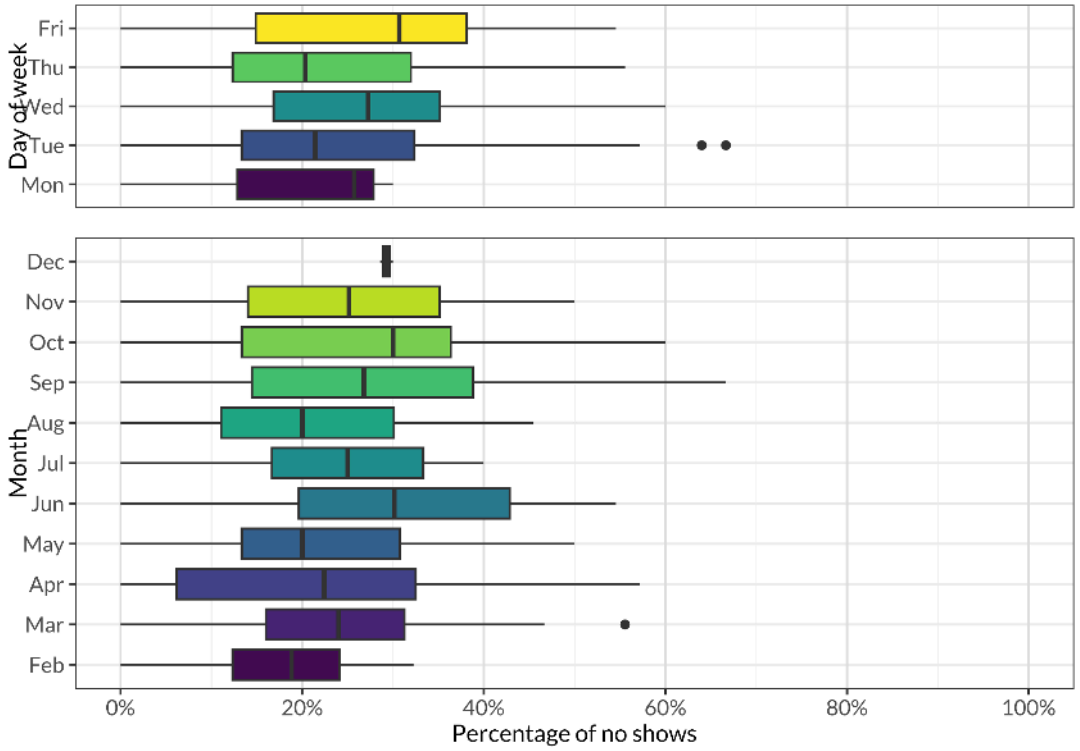


Figure 2.7.2: Top: Distribution of no shows by the day of the week the course was delivered; Bottom: Distribution of no shows by the month the course was delivered.

Further analysing the data about no shows by Role/Position and Faculty, we observed that the average highest no show rate is among Masters (Coursework) students with approx. 40% of registrants failing to attend, followed by Undergraduate and Other with a percentage of no shows higher than 32%, respectively. The lowest average percentage is shown in Post-doc/Fellow and Professional staff. Participants from the Business School, Faculty of Law, and Faculty of Engineering have the highest average percentage of no shows with 32-33%, respectively, while participants from the Faculty of Science, the Faculty of Medicine and Health, and the Faculty of Architecture, Design and Planning display the lowest average percentage of no shows.

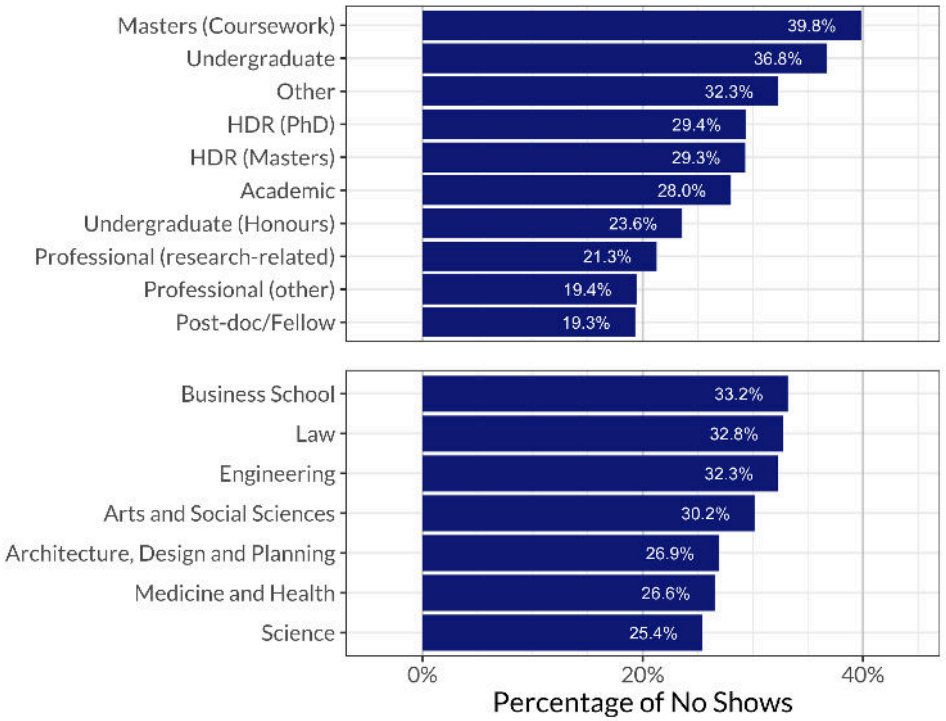


Figure 2.7.4: Top: Average percentage of no shows by Role/Position; Bottom: Average percentage of no shows by Faculty.

3. Evaluation

To evaluate the quality of the training delivery, Intersect instructors ask the attendees to fill in a course evaluation survey at the end of each course. In our course evaluation survey, a scale of 0 (worst) to 10 (best) is used. In addition, a Net Promoter Score, in which attendees are asked how likely they are to recommend Intersect Training courses to others, is also measured. A positive NPS is considered great, while achieving an NPS of +50 or higher is considered outstanding and seldom achieved commercially. The Intersect training program demonstrates high quality and is well received by the attendees. We have analysed all the survey responses in 2023 and the analysis is presented in the following subsections.

3.1 Evaluation by Tool/Technology

The Intersect training program demonstrates high quality and is well received by the attendees. In 2023, Intersect scored an NPS of +74 and all courses by tool/technology have independently achieved an average NPS of over +50, except for MATLAB, which is considered outstanding. This demonstrates that while we deliver digital skills training for researchers at scale, we still maintain, and where possible improve, the quality of delivery to provide the best possible learning experience to attendees.

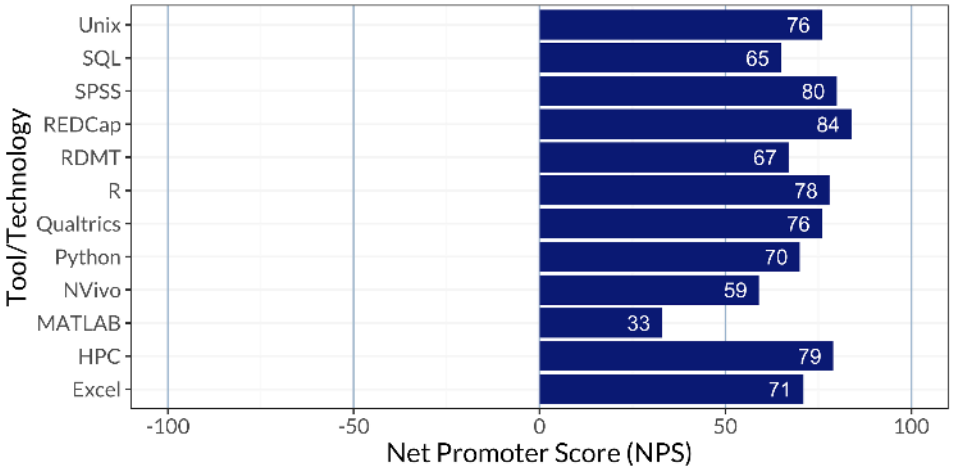


Figure 3.1.1: Average Net Promoter Score (NPS) by Tool/Technology. The course data was aggregated by Tool/Technology.

3.2 Evaluation by month and class size

The NPS score was further calculated by month (Figure 3.2.1) and class size (Figure 3.2.2). A very high NPS score is observed in all months, with all months consistently achieving an average NPS of more than 50.

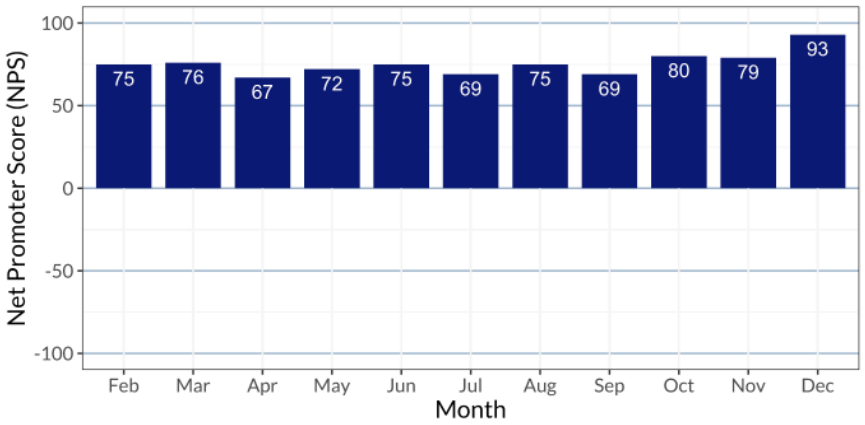


Figure 3.2.1: Average Net Promoter Score (NPS) by month.

When analysing the survey by class size, *e.g.*, the number of participants in a course, although there is a slight fluctuation on the NPS score, we notice a consistently high NPS of 50+ among all class sizes, except for a class size of 34 (NPS is +42). It is worth noting that even classes in 2023 with more than 30 attendees achieved a consistently high NPS score (+40 and above), indicating that the quality of the course delivery does not drop in courses with higher attendance. It is important to note that analysis on historical data (based on over 10,000 responses) showed that there is a very slight decrease in the NPS as we increase the number of participants in the training courses, and therefore the number of attendees per course should be carefully considered.

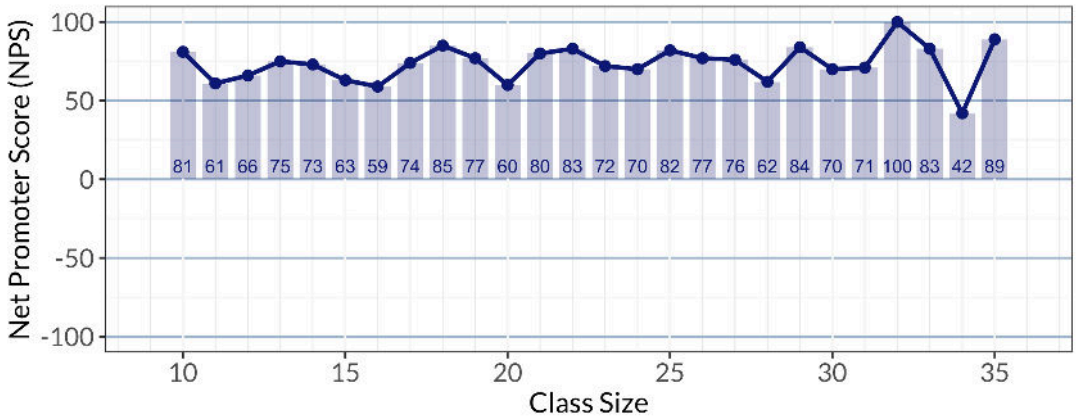


Figure 3.2.2: Average Net Promoter Score (NPS) by class size.

3.3 Metrics for measuring quality of delivery

Figure 3.3.1 shows the five primary metrics for measuring the quality of the training delivery per quarter, namely:

- **Clear answers**: "*Do you think that the instructors gave clear answers to questions?*"
- **Comfortable interacting**: "*Did you feel comfortable interacting with the instructors?*"
- **Good communicators**: "*Do you think that the instructors were good communicators?*"
- **Knowledgeable instructors**: "*Do you think that the instructors were knowledgeable about the material being taught?*"
- **Welcoming atmosphere**: "*Did you feel that the training course atmosphere was welcoming?*"

All metrics are above 9.1 out of 10 in each quarter, which indicates that the training attendees appreciate the comfortable training environment, interactive teaching style, and excellent communication from knowledgeable instructors.

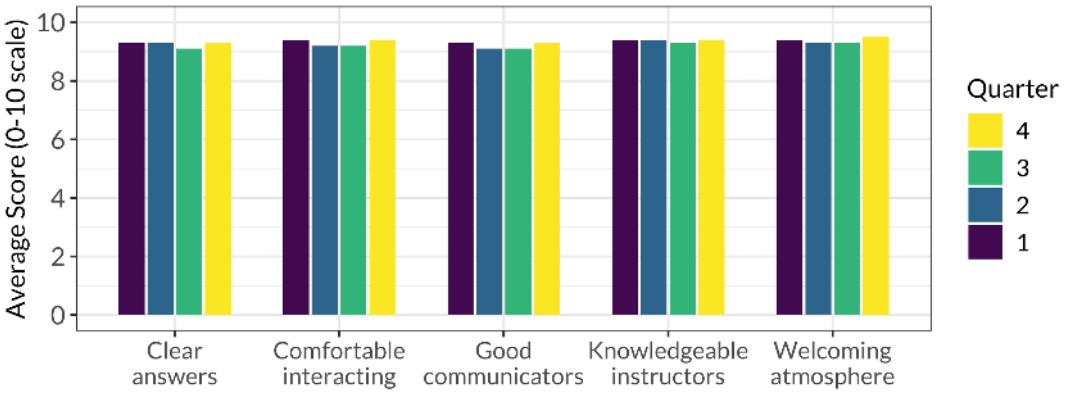


Figure 3.3.1: Primary metrics for measuring the quality of the training delivery per quarter.

3.4 Qualitative feedback and evaluation

For the second consecutive year, a high-level thematic analysis was undertaken on the qualitative data collected from participants who responded to the three open-ended questions on the Intersect course

evaluation survey.

1. "*Which parts of the course did you find most useful? Why?*"
2. "*Which parts of the course did you find least useful? Why?*"
3. "*Do you have any other suggestions or feedback on this course or Intersect Training in general?*"

The same approach was taken with the 2023 data as the 2022 data - that is, the data was analysed inductively rather than deductively using the 2022 themes, in order to minimise bias as much as possible. Three independent coders analysed different subsets of the data, applying open coding techniques to uncover relevant themes. A maximum of two themes per comment was applied. Not only was a high level of consensus gained between the coders, but the emerging themes were not dissimilar to those developed from the 2022 data. For consistency and ease of comparison, therefore, the 2022 themes were used when appropriate. In some cases however, it was prudent to use a more refined code to better reflect the feedback that was provided (see for example the 'Specific content' theme in the 'Least useful' section). The top three themes in each of the three open-ended questions are discussed in the following sections, along with any relevant adjacent themes.

What were the most useful aspects of the training?

Of the 1322 responses, 1270 of them were coded to at least one theme. The remaining 52 responses did not fit within any of the themes, and therefore were not assigned a code. The top 3 Most useful themes were:

1. Course specifics (590 responses, 45%)
2. Structure of course (351 responses, 27%)
3. All of it (283 responses, 21%)

Course specifics refers to specific elements of the course content, for example, pivot tables in Excel; functions and loops in R and Python; and designing surveys in REDCap and Qualtrics. This theme was also the most useful theme of 2022 (refer Table 3.1). Other specific elements that were mentioned included min, max and average functions, nested functions, and box plots (Excel for Researchers); variables, pandas, and seaborn (Data Manipulation and Visualisation in Python); creating branching blocks and end of survey messages (Surveying with Qualtrics); and data analysis, visualisation and syntax (Data Entry and Processing in SPSS).

The theme **Structure of course** refers to the way the course is structured, including but not limited to being able to follow the instructor, the interactive nature of our training, going through the basics, and the step-by-step set up. Appreciating the learn-by-doing and learning from mistakes approach, one participant in Learn to Code: Python said that *"Being able to code live was great. Also, being shown errors and why they happen was really useful."* The practical aspect of live coding or working through the examples together with the tutor were helpful in being able to *"really understand and remember how to do things when I was actually doing them in real time"* (Getting Started with NVivo for Windows participant). This theme, being the second most frequent theme in 2023, is a positive indication that Intersect's pedagogical, hands-on approach is recognised among participants.

The number of participants responding that they found the entire course useful (**All of it**) increased over fourfold from 2022 (78 responses, 5.2%) to 283 responses (21%) (refer Table 3.1). **Style of delivery**, relating to the training being well organised and run in a supportive non-judgemental environment, and **Useful for my work** were also frequent themes, matching the 2022 results (refer Table 3.1).

Least useful aspects of training

A total of 1067 responses were received in response to the Least useful aspect of training question, with 805

responses being coded to at least one theme. The top 3 Least useful themes were:

1. None of it (614 responses, 58%)
2. Specific content (98 responses, 9%)
3. Level of course (30 responses, 3%)

The most prominent response (614 responses, 58%) was **None of it**, meaning that all of the course was useful, an increase on the 2022 results for this theme (508, 42%). Combined with the Most useful results, 897 (38%) of participants feel that every aspect of our training is useful to them in some way.

Demonstrating a maturing of the qualitative analysis process, the 2022 theme of Content developed into the more precise **Specific content** theme. Reflecting the Course specifics theme in the Most useful responses, this theme was applied when participants noted particular topics that shouldn't be included in the course. However, the comment that something wasn't useful was often accompanied by an acknowledgement that they understood why it was important to be included in the course. For example, an Excel for Researchers (our introductory Excel course) participant commented: *"Data plotting - I've used Excel and SPSS alot so perhaps I had more practice than other participants."*. A total of 98 responses (9%) were assigned to this theme, which gives Intersect more targeted feedback to be considered when reviewing and updating courses. Respondents who noted that there were aspects of the course that were not useful subsequently commented that the pace was slow. While the **Pace** theme was not in the Top 3 in terms of number of responses, there were 19 comments that were coded to this theme (2%). Of those 19, five (26%) were coded to both Specific content and Pace. Again, these comments were tempered with an acknowledgement that *"... I knew it was helpful for others so I wasn't annoyed by this at all"* (Data Manipulation and Visualisation in R participant).

The final theme in the Least useful category concerned the **Level of course** (30 responses, 3%). Like the Specific content theme, most comments were accompanied by an acknowledgement that they understood why the course was pitched at the level it was. One participant at Learn to Program: Python noted that the course was *"Too basic - but I knew that, as it was target [sic] to beginners."*. Similarly, another participant said that the least useful part of the course was the beginning section, "*[...] but understand the instructor needs to have a common start point."*.

Other suggestions or feedback

Perhaps not unexpectedly, the 858 responses given to the Other feedback question were wide and varied, resulting in a relatively low number of responses being coded to at least one theme (268 responses, 31%). The Top 3 themes to emerge were:

1. Trainer (70 responses, 8%)
2. Pace: Negative (30 responses, 3.5%)
3. Style of delivery (30 responses, 3.5%)

The **Trainer** theme was new for 2023 (70 responses, 8%), and was applied to responses that specifically mentioned the course trainer. The comment didn't have to mention the trainer by name necessarily - a reference to the training team was also coded to this theme. The expertise of the trainers was particularly appreciated by participants with little previous knowledge of the topic. Two examples of such comments are: *"Emran did a beautiful job and he [is] very warm and patient and an absolute beginner like me learned a lot"* (Introduction to Programming: Python); "*A big shout out to Esmaeel and Jianzhou who were both awesome - content experts who are also good at explaining things is pretty rare!"* (Learn to Program: R). These comments were not limited to introductory level courses. A participant in our second-level (intermediate) course, Data Manipulation in Python, gave the following feedback for the trainer:

> *The instructor was amazing. I felt I could follow well, and she explained topics in a complete/detailed and*

> *simple way. Thank you for explaining Python to me so that I didn't feel like I was small or stupid for not knowing something. It's a talent to be able to do that, so thank you.*

Comments such as these suggest that Intersect's recruitment and trainer induction processes are attracting high-level candidates inherently suitable for the training role.

The second most frequent theme to emerge from the Other feedback question was **Pace: negative** (30 responses, 3.5%). This theme was another new theme for 2023, although derived from, and a more refined version of the 2022 theme, Pace (refer Table 3.1). Some comments suggested that the entire course was too fast or too slow, while others noted specific points of the course - usually either the first or last sections - were too fast or slow. Generally speaking, comments relating to the training being too slow were from participants who had some experience or prior knowledge of the topic being trained; or who had experienced a training session where some extra attention was required by other participants, thus slowing the overall pace of training. An Excel for Researchers participant noted the following:

> *Maybe if somebody is struggling with the topic don't stop the course [but rather] ask the person to stay at the end or get an individual session because we all had to wait and we missed some topics that were not covered because we did not have enough time.*

While the frustration with this situation is understandable, this highlights the difficulty of having participants with differing levels of understanding of the topic at the same training session. As the Intersect approach to training is to only go as fast as the slowest person, those with more understanding will naturally feel the pace is too slow. Further, the number of topics covered in any given training tends to differ from session to session, depending on the pre-existing knowledge of the participants. The Course Material provided to participants contains more than what we are able to cover in our three- or six-hour training sessions for precisely this reason, as it gives trainers the flexibility to cover more advanced topics when and if required. If the training is at a slower pace, the participants are able to access the more advanced modules in the Course Materials at any time.

Conversely, a participant at an introductory HPC course noted that the pace was too fast, especially when correcting errors, because *"as a beginner it would take me a couple of seconds to think how to open up batch scripts in nano and then it was easy to miss changes."* Intersect tries to alleviate issues with pace with anonymous Zoom polls to gauge how participants are finding the training, whether too fast, too slow, or just right, and these are typically launched at the major breaks in the training (approximately 1 hour in). Additional polls to check participants' progress during set tasks and exercises also helps trainers to determine whether moving on to the next section of the training is warranted. The above quote is a good example of when this poll could have been used to great effect. As a relatively recent addition to the suite of polls, it is likely that this poll was not run in the HPC training course.

The **Style of delivery** theme (30 responses, 3.5%) in third place within the top three Other feedback themes is a repeat of the 2022 rank. Comments relevant to this theme indicated that the course content was explained well by trainers who were *"clear, patient and helpful"* (Getting started with NVivo for Mac) and *"welcoming"* (Data Capture and Surveys with REDCap, Data Manipulation and Visualisation in R). The ease in which questions could be asked (Data Capture and Surveys with REDCap, Unix Shell and Command Line Basics), the high level of knowledge of the trainers (Surveying with Qualtrics, Excel for Researchers, Data Manipulation and Visualisation in R), and the overall expertise and professionalism of the trainers (Excel for Researchers, Beyond Basics: Conditionals and Visualisation in Excel, Introduction to Machine Learning using Python, Research Data Management Techniques, Data Manipulation and Visualisation in R) were also common responses.

Table 3.1 below shows a comparison of the top 5 themes to emerge from the three qualitative questions in 2022 and 2023. With few exceptions, the themes that emerged in 2023 were overall the same as those in 2022. As with the responses in 2022, those in 2023 were also overwhelmingly positive. Those comments that are

more critical often have an acknowledgement of the circumstances around the criticism (as discussed in relation to the Specific content and Level of course themes). As always, these valuable responses provide opportunities for Intersect to continue refine and improve their training to continue to meet universities' and researchers' expectations.

Table 3.1: Thematic comparison between 2022 and 2023

2023	2022
Most useful themes	
Course specifics	Course specifics
Structure of course	Style of delivery
All of it	Structure of delivery
Style of delivery	Hints, tips, and shortcuts
Useful for my work	All of it
Least useful themes	
None of it	None of it
Specific content	Content
Level of course	Context
Pace	Pace
Structure of course	Structure of the course
Other feedback	
Trainer	Level of course
Pace (negative)	Context
Style of delivery	Style of delivery
Context	Pace
Level	Structure of the course

4. Communication

Most participants found out about the Intersect training program and the courses scheduled through the "Faculty/School newsletter" (~22%). The second and third top options are through the "Research office/division email" and "University newsletter" with approximately 14.8%, respectively. The "University website" and "Your Supervisor" were another two popular ways of raising awareness about Intersect training courses with about 11.8% and 8.9%, respectively. This demonstrates the importance of universities adopting a proactive approach to facilitating the dissemination of information about training through their internal communication channels, as these vastly outweigh Intersect or third party communication channels.

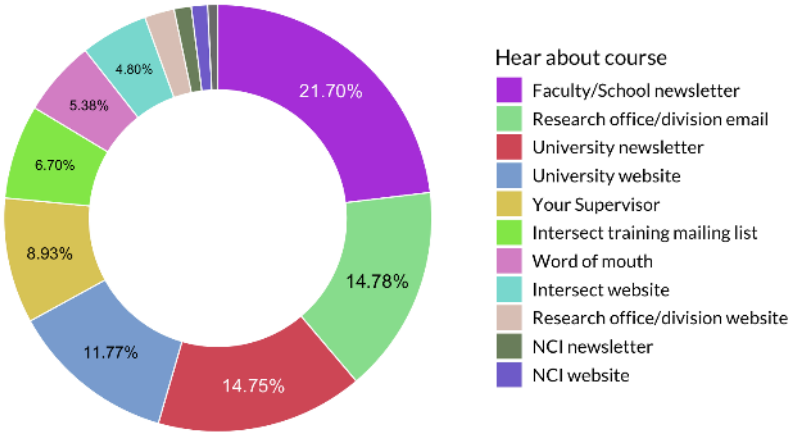


Figure 4.1: How did participants hear about the Intersect courses.

Figure 4.2 shows further analysis of the communication methods by Role/Position and by Faculty. HDR (PhD) students who are the biggest cohort attending Intersect training courses learn about Intersect training and the course scheduling through "Faculty/School newsletter" followed by "Research office/division email" and "University newsletter". These three communication channels also work best for Academic and Post-doc/Fellows who learn about Intersect training courses mostly through these channels. "Your supervisor" is a very popular communication channel among Undergraduate (Honours) but also among HDR students (PhD and Masters) as well as Professional staff (research-related and other).

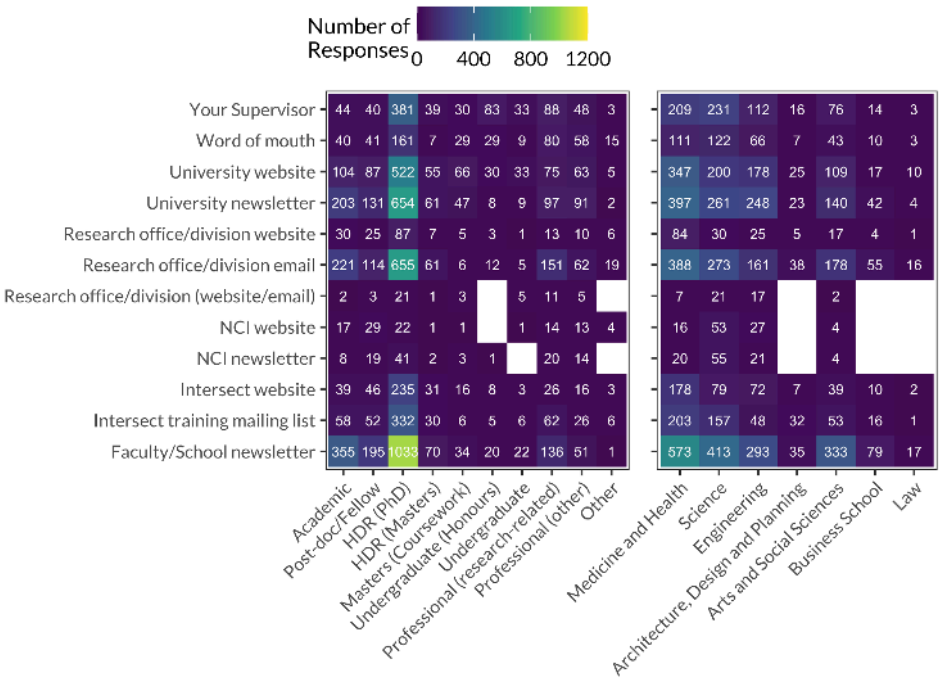


Figure 4.2: How did participants hear about the Intersect courses by Role/Position and by Faculty.

Participants from the Faculty of Medicine and Health hear about Intersect courses mostly via "Faculty/School newsletter" and to a lesser extent via "Research office/division email", "University newsletter", and "University website". Participants from the Faculty of Engineering, the Faculty of Science, and the Faculty of Arts and Social

Sciences hear about Intersect courses mostly through the same channels.

5. Historical trends

Since the inception of Intersect's Research Digital Skills Training Program in 2012, Intersect Australia has delivered over 2,450 courses and trained more than 37,500 researchers, HDR students, and professional staff in various universities, government departments, and organisations across Australia. In the past, there was a consistent annual increase in both the number of courses and training days delivered from 2012 to 2021. However, since 2022, there have been fluctuations observed in both metrics (see Figure 5.1). In 2022, Intersect delivered 314 courses and 241.5 training days which indicates that Intersect delivered more than 6 courses on average every working week in 2023 (more than 1 course per working day).

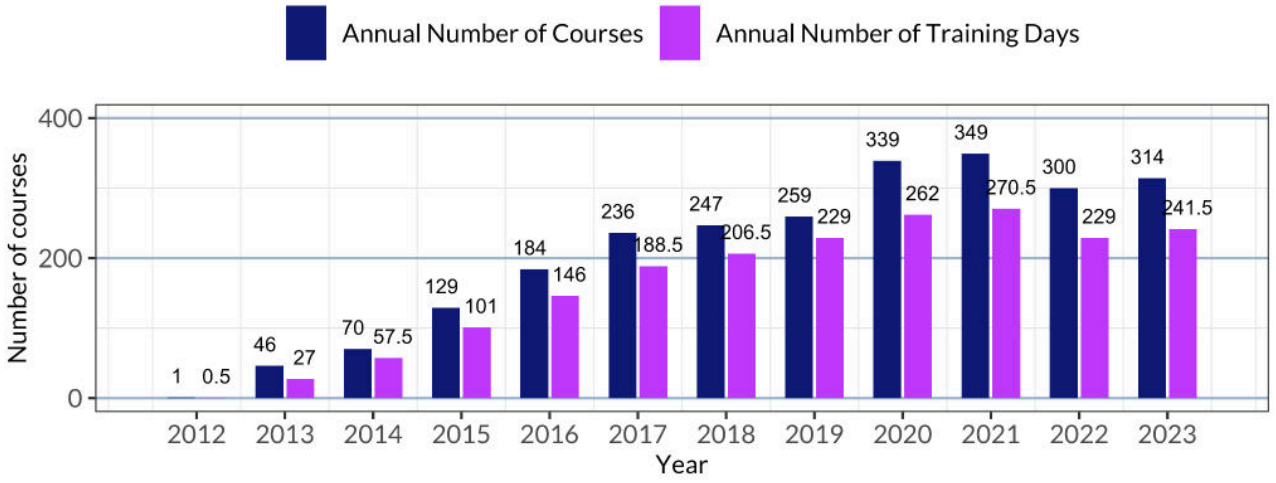


Figure 5.1: Total number of Intersect courses and training days delivered by Intersect Australia per year since 2012.

The number of researchers trained per annum has been consistently increasing from 2012 until 2021, while it has been fluctuating since 2022 as shown in Figure 5.2 (bars and right y axis). The cumulative number of attendees per year shows an exponential increase until 2020 after which it changes to a linear increase (see line/points and left y axis in Figure 5.2). Intersect has trained over 37,500 participants since the inception of our training program. In the coming 1-2 years, we anticipate reaching another huge milestone by successfully training over 20,000 *unique* researchers and HDR students across Australia, helping the Australian research workforce to upskill in emerging digital tools and technologies that are widely used in academia.

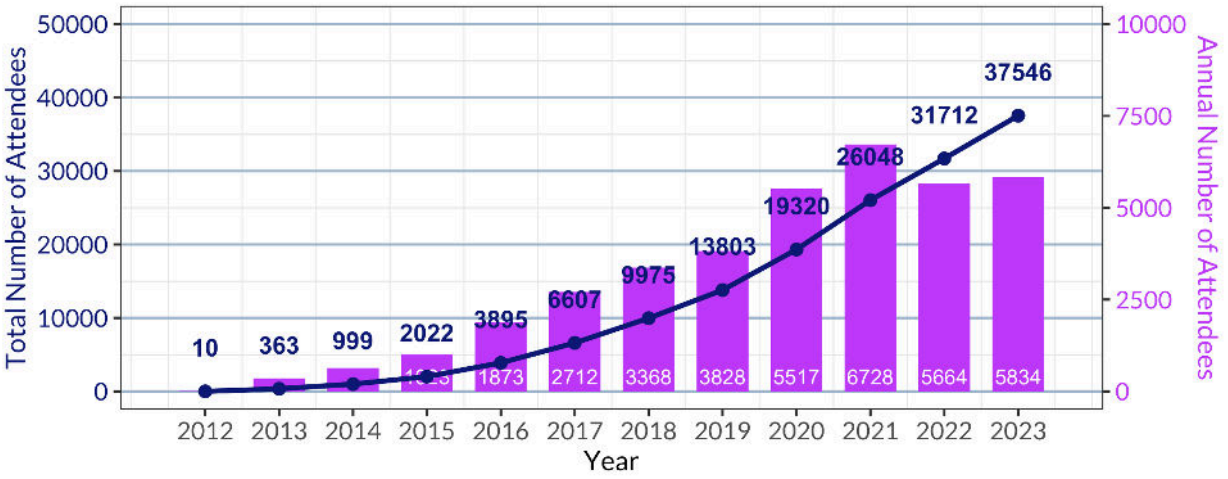


Figure 5.2: Annual and Total number of attendees since 2012.

Figure 5.3 shows the annual number of unique attendees and the cumulative number of unique attendees per year. At the end of 2023, Intersect has successfully trained over 16,900 unique HDR students and researchers across Australia, which makes Intersect one of the leading providers of research digital skills training in the Australasian region.

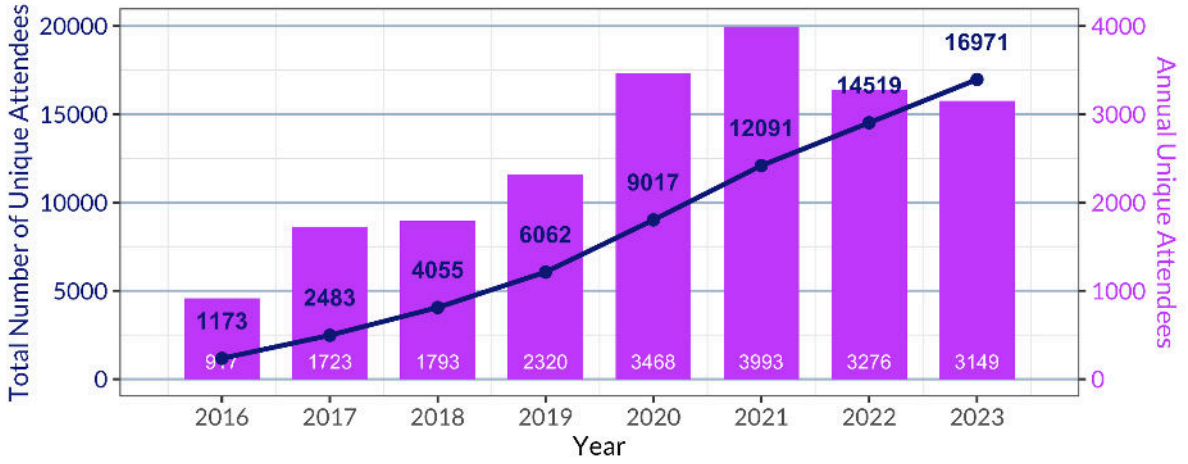


Figure 5.3: Annual and Total number of unique attendees since 2016.

Figure 5.4 compares the distribution of participants returning to attend more Intersect courses. The data presented in this figure include five time periods:

- Only 2018 attendees data
- Attendees data from 2018 to 2019
- Attendees data from 2018 to 2020
- Attendees data from 2018 to 2021
- Attendees data from 2018 to 2022
- Attendees data from 2018 to 2023

Analysing the 2018 attendees data only, 75.6% of participants attended only one Intersect course while 24.4% of participants returned to attend two or more Intersect courses. When taking into account a longer time period, *e.g.*, data from 2018 to 2023, the number of people who attended only a single Intersect course throughout these years drops to approximately 69.5% meaning that approximately 30% participants return to

attend more Intersect courses in the same or following years. Comparing 2018 only data with attendees data from 2018 to 2023, we observe a higher increase in the number of participants attending more Intersect courses while the percentage of participants attending a single course is decreasing significantly (see Table 5.1).

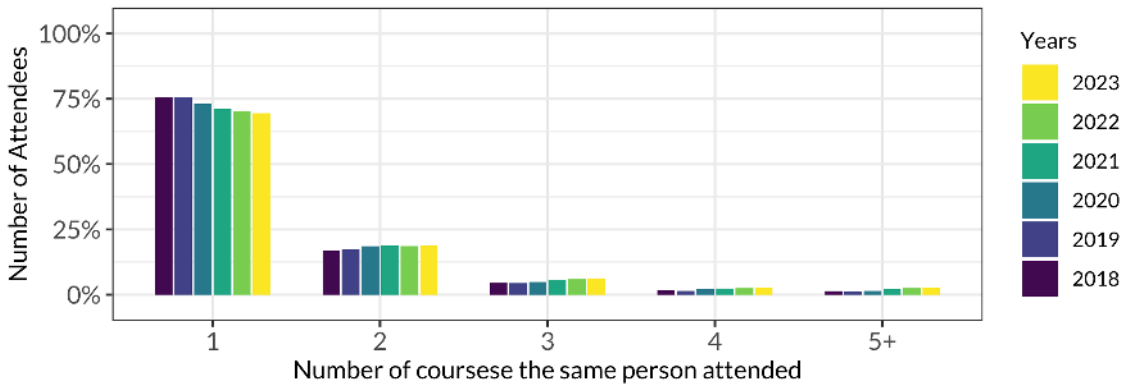


Figure 5.4: Number of courses that the same person attended. The data are presented in 2018 and then in consecutive years until 2023.

Table 5.1. Comparison of distributions of participants attending 1 to 5+ Intersect courses for 2018 only attendees data and attendees data for the 2018 to 2023 time period.

# of Courses	Only 2018 data	Data from 2018 to 2023	Difference
1	75.6%	69.5%	-8.8%
2	16.9%	18.8%	10.1%
3	4.5%	6.3%	28.6%
4	1.6%	2.6%	38.5%
5+	1.3%	2.8%	53.6%

Figure 5.5 shows the percentage of attendees by State/Territory in Australia. Although Intersect training was predominantly delivered in New South Wales (NSW) universities and organisations in 2016, in 2023, 47.2% of attendees are researchers and HDR students from NSW universities and organisations. South Australia and Victoria (VIC) are the second and third largest states in terms of number of attendees participating in Intersect training courses with 1097 and 1092 attendees in 2023 or 19.4% and 19.3% of the total Intersect training attendees, respectively. Notably, in 2023, South Australia surpassed Victoria and became the second largest state regarding the number of attendees for the first time. The training partnership between Intersect and the National Computational Infrastructure (NCI) has also increased Intersect training activities in Australian wide organisations and universities as well as in other States and Territories across Australia.

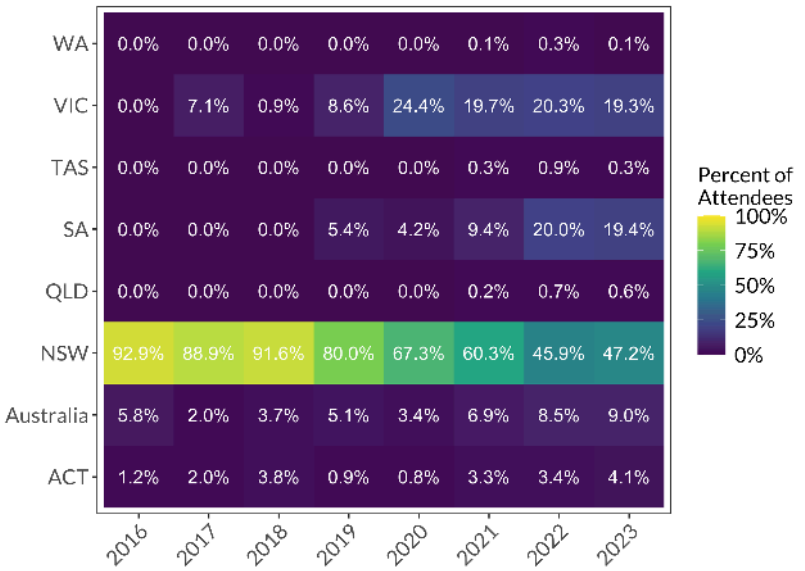


Figure 5.5: Percentage of attendees by state/territory since 2016.

Further insights can be derived by analysing the historical attendance data across several years and by Role/Position. Figure 5.6 presents the percentage of total attendees by Role/Position since 2016. The top consumer of Intersect training courses is consistently the HDR (PhD) students with the percentage fluctuating between 41% and 57%. Notably, although Academics constituted less than 3% in 2016-2018, a substantial increase was observed in 2019 with Academics constituting over 13% of attendees consistently since then. It is important to note that we restructured our Roles/Positions in our registration process in 2019 and therefore some inconsistencies occur in some of the Roles/Positions such as HDR (Masters) and Professional staff.

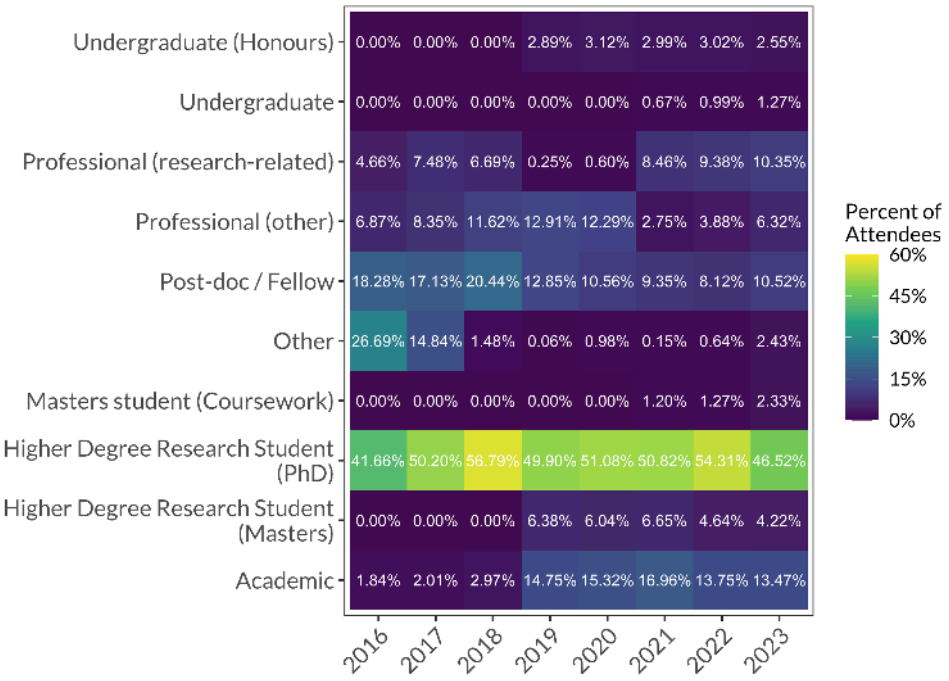


Figure 5.6: Percentage of attendees by Role/Position since 2017.

Figure 5.7 displays the percentage of attendees by Faculty since 2016. The Faculty of Medicine and Health is historically the top consumer of Intersect training with a percentage of attendees between approximately 25%

and 36%. The Faculty of Science followsg very closely as the second largest cohort of attendees (approx. 23% - 35%), with this faculty becoming the top consumer only in 2018. Although the third top consumer is historically the Faculty of Engineering, after 2019, a high rise of attendees from the Faculty of Arts and Social Sciences is observed surpassing the Faculty of Engineering during the pandemic (2020 - 2021) before dropping again to be the fourth top consumer after 2022. Since 2020, Intersect has invested time and effort to raise awareness of how digital tools and technologies can be beneficial and relevant to HASS researchers and HDR students and we are pleased to observe a higher uptake of Intersect training courses since then.

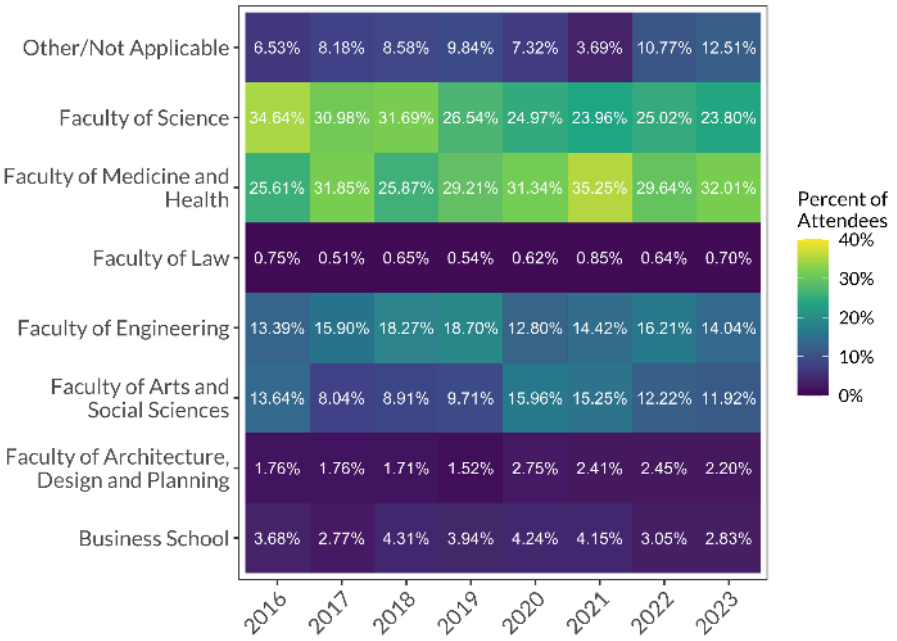


Figure 5.7: Percentage of attendees by Faculty since 2016.

Figure 5.8 shows the percentage of attendees per year split by the tool/technology. Until 2017, Excel courses were the most popular. However, in 2018, R and Python programming courses overtook Excel and became the most popular training courses Intersect delivers every year. Last year, despite Intersect introducing an intermediate Excel course, the percentage of attendance for Excel courses continued its year-on-year decrease, reaching less than 10% of total attendees for the first time in 2022. Since 2018, R programming courses attracted the most attendees; very closely followed by Python. However, in 2023, Python became the most popular tool/technology that Intersect taught. Courses on surveying tools such as REDCap and Qualtrics started gaining popularity after 2018 and, in particular, in 2020 (the beginning of the pandemic), interest in these courses surged, reaching their highest ever percentage of attendance, with 12% and 11.7%, respectively. Since then, the percentage of attendance for REDCap courses has dropped but remains higher than in pre-pandemic years. In contrast, Qualtrics courses have dropped to similar percentages compared to the pre-pandemic years. Another interesting case is NVivo courses for qualitative analysis. It was introduced in 2020 and quickly became a very popular tool.

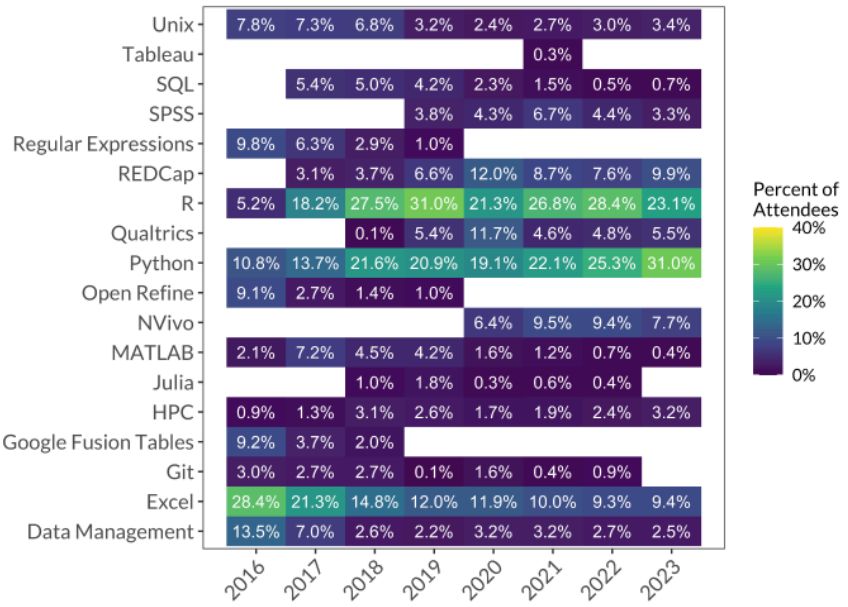


Figure 5.8: Percentage of attendees by Tool/Technology since 2016.

Intersect's annual NPS has increased rapidly over the years, from +46 in 2017, when NPS was first measured, to an NPS of above +70 since 2020. This demonstrates our commitment to the improvement of training quality and providing the best possible training experience to participants. Furthermore, Intersect managed to successfully move the entire course catalogue and delivery online after a tremendous effort from the Intersect team in the beginning of the pandemic in 2020. The consistently high NPS scores and metrics for measuring the quality of training suggest that moving our entire course catalogue and delivery online did not compromise the quality of delivery and training experience.

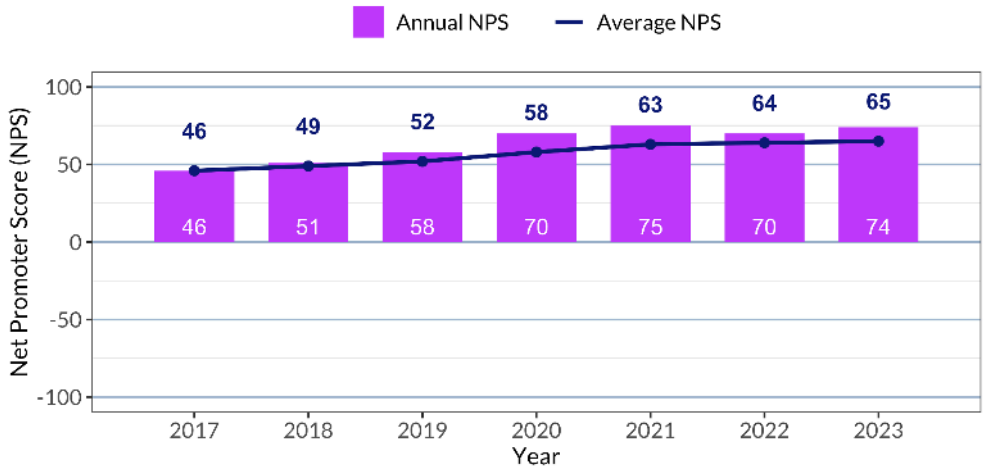


Figure 5.9: Average and Annual NPS per year since 2017.

As shown in Figure 5.10, all primary quality metrics consistently exceed on average 9.1 out of 10 since 2020 (when we introduced these metrics), demonstrating a very high quality training delivery from very knowledgeable and experienced instructors in a training atmosphere that is welcoming participants to interact with the instructors and ask questions during the delivery.

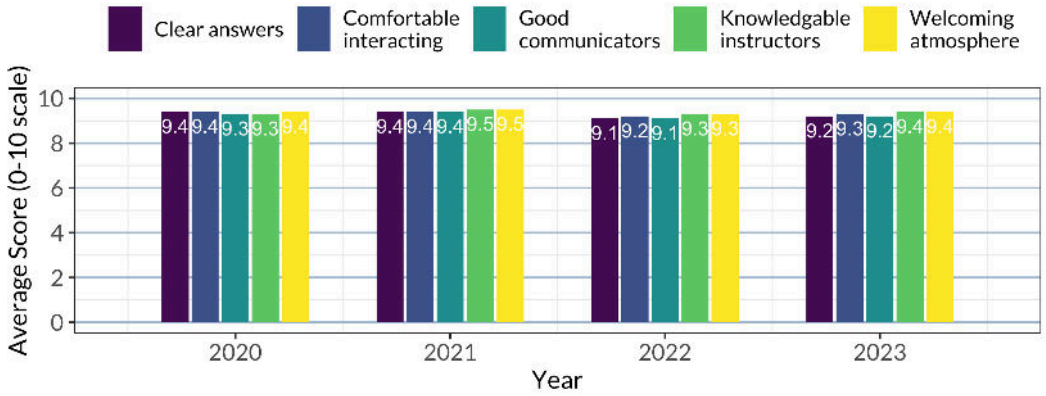


Figure 5.10: Primary metrics for measuring the quality of the training delivery per year since 2020.

Appendix A - Intersect training catalogue

Table A.1: List of Intersect training courses as in April 2023.

#	Course Name	Course Code	Duration (Days)	Tool/Technology being taught
1	Excel for Researchers	EXCEL101	1	Excel
2	Beyond Basics: Conditionals and Visualisation in Excel	EXCEL201	0.5	Excel
3	Version Control with Git	GIT101	0.5	Git
4	Getting started with HPC using PBS Pro	HPC201	1	HPC
5	Getting started with HPC using Slurm	HPC202	1	HPC
6	Parallel Programming for HPC	HPC301	1	HPC
7	Learn to Program: Julia	JULIA101	0.5	Julia
8	Beyond the Basics: Julia	JULIA201	0.5	Julia
9	Learn to Program: MATLAB	MATLAB101	1	MATLAB
10	Getting started with NVivo for Windows	NVIVO101	0.5	NVivo
11	Getting Started with NVivo for Mac	NVIVO102	0.5	NVivo
12	Learn to Program: Python	PYTHON101	1	Python
13	Python for Research	PYTHON110	0.5	Python
14	Data Manipulation in Python	PYTHON201	0.5	Python
15	Data Visualisation in Python	PYTHON202	0.5	Python
16	Data Manipulation and Visualisation in Python	PYTHON203	1	Python
17	Introduction to Machine Learning using Python: Introduction & Linear Regression	PYTHON205	1	Python
18	Introduction to Machine Learning using Python: Classification	PYTHON206	1	Python
19	Introduction to Machine Learning using Python: SVM & Unsupervised Learning	PYTHON207	0.5	Python
20	Surveying with Qualtrics	QLTRICS101	0.5	Qualtrics
21	Learn to Program: R	R101	1	R
22	R for Social Scientists	R103	1	R
23	R for Research	R110	0.5	R
24	Data Manipulation in R	R201	0.5	R
25	Data Visualisation in R	R202	0.5	R
26	Data Manipulation and Visualisation in R	R203	1	R
27	Introduction to Machine Learning using R: Introduction & Linear Regression	R205	1	R
28	Introduction to Machine Learning using R: Classification	R206	1	R
29	Introduction to Machine Learning using R: SVM & Unsupervised Learning	R207	0.5	R
30	Exploring Chi-square and correlation in R	R210	0.5	R

31	Traversing t tests in R	R211	0.5	R
32	Exploring ANOVAs in R	R212	0.5	R
33	Research Data Management Techniques	RDMT001	0.5	RDMT
34	Data Capture and Surveys with REDCap	REDCAP101	0.5	REDCap
35	Longitudinal Trials with REDCap	REDCAP201	0.5	REDCap
36	Cleaning Data with Open Refine	REFINE101	0.5	Open Refine
37	Mastering text with Regular Expressions	REGEX101	0.5	Regexes
38	Regular Expressions on the Command Line	REGEX201	0.5	Regexes
39	Software Carpentry (Python)	SC101	2	Python, Git, Unix
40	Software Carpentry (R)	SC102	2	R, Git, Unix
41	Data Entry and Processing in SPSS	SPSS101	1	SPSS
42	Exploring Chi-Square and correlation in SPSS	SPSS102	0.5	SPSS
43	Databases and SQL	SQL101	0.5	SQL
44	Getting Started with Tableau for Data Analysis and Visualisation	TABLEAU101	0.5	Tableau
45	Unix Shell and Command Line Basics	UNIX101	0.5	Unix
46	Collecting Web Data	WEBDATA201	1	Python