Research Digital Skills Training - Impact on Researchers

Dr Anastasios Papaioannou & Dr Jianzhou Zhao

16/03/2022

```r
# Select only the people who responded and only the Research related Roles
Only_Responses <- Responses %>%
  filter(`Survey Timestamp` != "NULL") %>%
  filter(Role %in% c("Higher Degree Research Student (Masters)",
                     "Higher Degree Research Student (PhD)",
                     "Post-doc / Fellow",
                     "Academic",
                     "Professional (research-related)"))

# Add the Course level in the Only_Responses df
# test <- merge(Only_Responses,Courses_coeff[,c("Title","Level")],by.x="Course",by.y="Title",all.x = TR

# Recode the levels for 4 kirkpatrick's metrics and open material
Only_Responses <- Only_Responses %>%
  mutate(reaction_corrected = as.factor(fct_recode(`Do you feel that attending the course was worthwhil
        "5 - Extremely worthwhile"="Extremely worthwhile",
        "4 - Very worthwhile"="Very worthwhile",
        "3 - Slightly worthwhile"="Slightly worthwhile",
        "2 - Not very worthwhile"="Not very worthwhile",
        "1 - Not at all worthwhile"="Not at all worthwhile"))) %>%
  mutate(learning_corrected = as.factor(fct_recode(`Do you feel more or less confident using [technology
        "5 - Much more confident"="Much more confident",
        "4 - More confident"="More confident",
        "3 - No more confident"="No more confident",
        "2 - Less confident"="Less confident",
        "1 - Much less confident"="Much less confident"))) %>%
  mutate(behaviour_corrected = as.factor(fct_recode(`How frequently were you able to apply [technology]
        "5 - Frequently"="Frequently",
        "4 - Often"="Often",
        "3 - Sometimes"="Sometimes",
        "2 - Rarely"="Rarely",
        "1 - Never"="Never"))) %>%
  mutate(results_corrected = as.factor(fct_recode(`To what extent has [technology] been helpful to your
        "5 - Extremely helpful"="Extremely helpful",
        "4 - Very helpful"="Very helpful",
        "3 - Somewhat helpful"="Somewhat helpful",
        "2 - Not very helpful"="Not very helpful",
        "1 - Not at all helpful"="Not at all helpful"))) %>%
  mutate(materials_only_corrected = as.factor(fct_recode(`If you had access to Intersect's training mat
        "5 - Definitely"="Definitely",
        "4 - Probably"="Probably",
        "3 - Possibly"="Possibly",
```

```r
        "2 - Probably not"="Probably not",
        "1 - Definitely not"="Definitely not")))

# Sort the levels of the 4 corrected Kirkpatrick's metrics and open material
Only_Responses$reaction_corrected = factor(Only_Responses$reaction_corrected, levels = c(sort(levels(Onl
Only_Responses$learning_corrected <- factor(Only_Responses$learning_corrected, levels = c(sort(levels(On
Only_Responses$behaviour_corrected <- factor(Only_Responses$behaviour_corrected, levels = c(sort(levels
Only_Responses$results_corrected <- factor(Only_Responses$results_corrected, levels = c(sort(levels(Only
Only_Responses$materials_only_corrected <- factor(Only_Responses$materials_only_corrected, levels = c(s


# Recode the levels of eRA support helpful
Only_Responses <- Only_Responses %>%
  mutate(eRA_helpful = as.factor(fct_recode(`To what extent was/were the local eResearch Analyst/s help
        "5 - Extremely helpful"="Extremely helpful",
        "4 - Very helpful"="Very helpful",
        "3 - Somewhat helpful"="Somewhat helpful",
        "2 - Not very helpful"="Not very helpful",
        "1 - Not at all helpful"="Not at all helpful")))

# Sort the levels of eRA support helpful
Only_Responses$eRA_helpful <- factor(Only_Responses$eRA_helpful, levels = c(sort(levels(Only_Responses$
```

Introduction

- Intro about Research Digital Skills Training
- Intro about Intersect Training Program
- Intro about Intersect's Training Impact survey

Demographics

Intersect offered most training courses with attendance of between 5 to 25, with the highest number of courses
ranging between 10 to 15 attendees.

```r
Courses %>%
  group_by(Attended) %>%
  summarise(Course_number = n()) %>%
  filter(Attended >= 5) %>%
  filter(Attended <= 43) %>%
  ggplot(aes(x = Attended, y = Course_number)) +
  geom_area(fill = intersect_color1, alpha = 0.3) +
  geom_line(color = intersect_color1, size = 1) +
  geom_point(size = 2, color = "#0d1a75") +
  theme_bw() +
  theme(
    legend.position = "none",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
```

```
labs(x = "Attendance",
     y = "Number of courses")
```
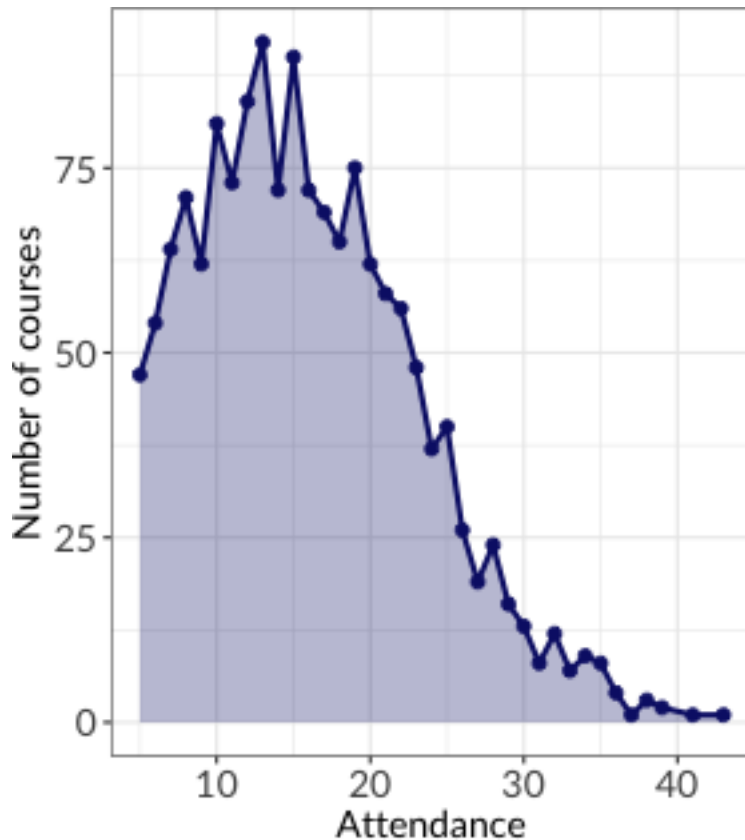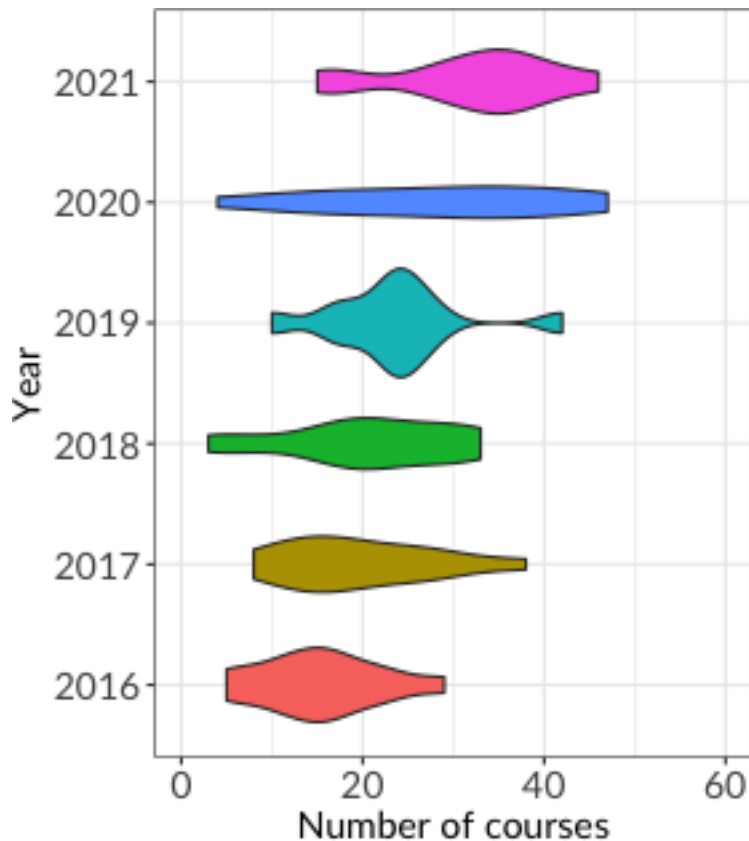

Fig. X shows the distribution of the number of courses offered per month for each year. 2016 to 2018 had a similar trend with about 15 to 20 courses been delivered per month. In 2019, there was an obvious increase of the number of courses per month. 20 to 30 courses were delivered in most months. Due to the pandemic, Intersect had undertaken the transition of all courses from in-person to online, and delivered a wide range number of courses throughout 2020. Since then, the number of courses per month has increased significantly, averaging around 30+ courses in each month.

```
Courses %>%
  mutate(Month_Courses = as.factor(month(`Course date`))) %>%
  group_by(Year, Month_Courses) %>%
  filter(Year != 2022) %>%
  summarise(Course_number = n()) %>%
  drop_na(Month_Courses) %>%
  mutate(Year = as.factor(Year)) %>%
  ggplot(aes(x = Course_number, y = Year, fill = Year)) +
    geom_violin() +
    # geom_jitter(color="black", size=0.4, alpha=0.9) +
    # scale_fill_viridis(name = "Temp. [F]", option = "C") +
    # theme_ipsum() +
    theme_bw() +
    theme(
      legend.position = "none",
      axis.title = element_text(size = 14),
      axis.text = element_text(size = 14),
```

```
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato"),
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  labs(x = "Number of courses",
       y = "Year") +
xlim(0,60)
```



```
Courses %>%
  mutate(Month_Courses = as.factor(month(`Course date`))) %>%
  group_by(Year, Month_Courses) %>%
  filter(Year != 2022) %>%
  summarise(Course_number = n()) %>%
  drop_na(Month_Courses) %>%
  mutate(Year = as.factor(Year)) %>%
  ggplot(aes(x = Course_number, y = Year, fill = Year)) +
    geom_boxplot() +
    # geom_jitter(color="black", size=0.4, alpha=0.9) +
    # scale_fill_viridis(name = "Temp. [F]", option = "C") +
    # theme_ipsum() +
    theme_bw() +
    theme(
      legend.position = "none",
      axis.title = element_text(size = 14),
      axis.text = element_text(size = 14),
```

```
      legend.text = element_text(size = 11),
      legend.title = element_text(size = 13),
      text = element_text(family = "Lato"),
      panel.spacing = unit(0.1, "lines"),
      strip.text.x = element_text(size = 8)
   ) +
   labs(x = "Number of courses",
        y = "Year") +
 xlim(0,60)
```
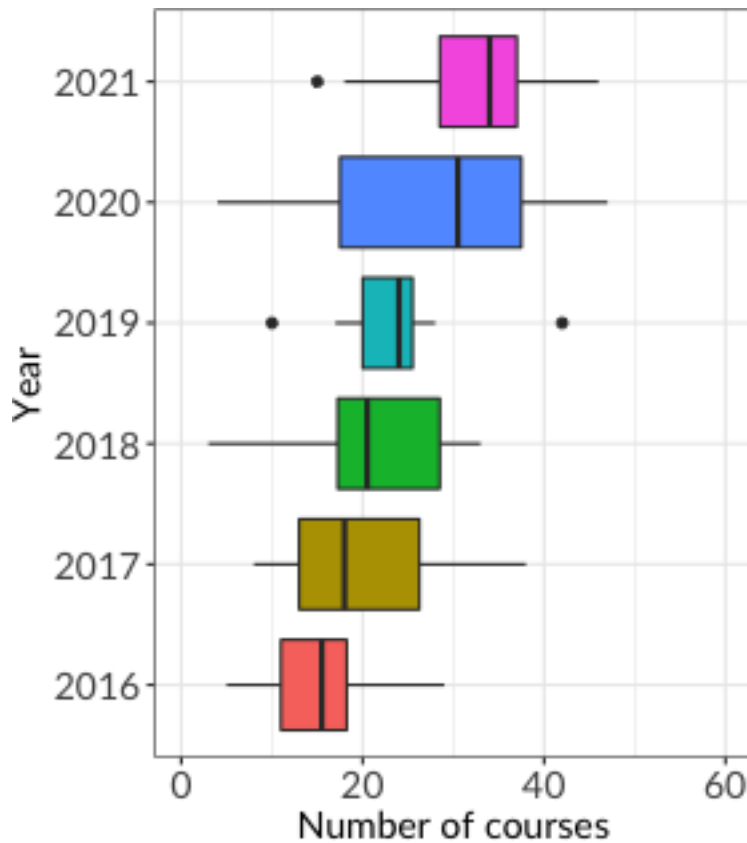

Fig. X presents an overview of the number of courses per month throughout all years. It is obvious that January, April and December had the lower number of courses because of the holidays whilst more courses were delivered between March and November, particularly in the second half of a year.
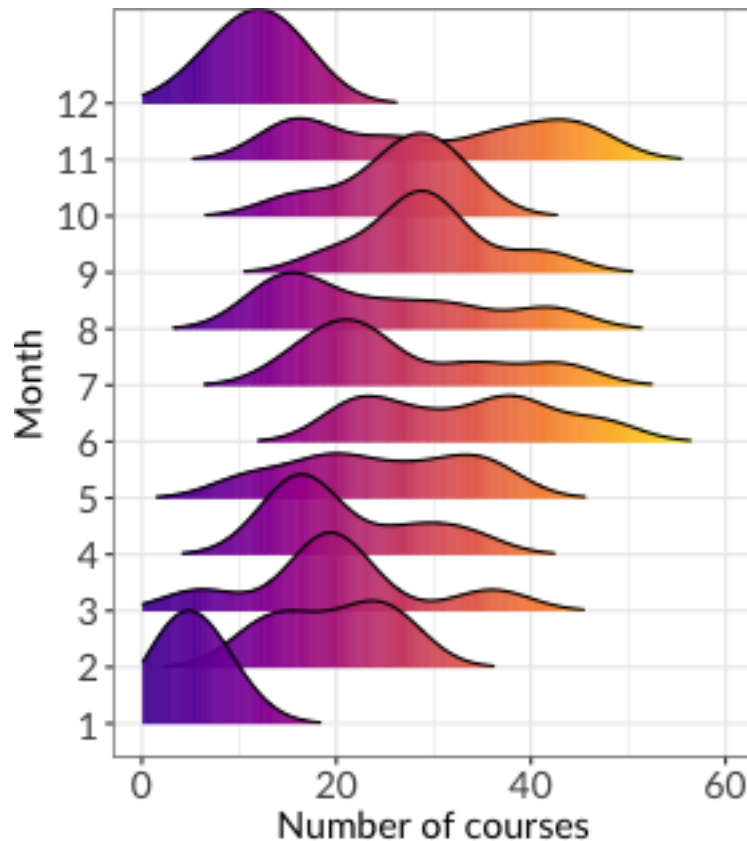
```
Courses %>%
  mutate(Month_Courses = as.factor(month(`Course date`))) %>%
  group_by(Year, Month_Courses) %>%
  filter(Year != 2022) %>%
  summarise(Course_number = n()) %>%
  drop_na(Month_Courses) %>%
  ggplot(aes(x = Course_number, y = Month_Courses, fill=..x..)) +
    geom_density_ridges_gradient(scale = 2, rel_min_height = 0.01) +
    scale_fill_viridis(name = "Temp. [F]", option = "C") +
    # theme_ipsum() +
    theme_bw() +
    theme(
      legend.position = "none",
      axis.title = element_text(size = 14),
```

```
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato"),
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  labs(x = "Number of courses",
       y = "Month") +
xlim(0,60)
```


From the heat map which shows the numbers of courses by month and year, it is clear that as Intersect expanded its course catalogue, significantly more courses were delivered beyond 2019 compared to previous years. Also, note that more than 30 courses have been delivered each month in the second half of 2020 as Intersect shifted its training delivery to online.

```
Courses %>%
  mutate(Month_Courses = as.factor(month(`Course date`))) %>%
  group_by(Year, Month_Courses) %>%
  filter(Year != 2022) %>%
  summarise(Course_number = n()) %>%
  drop_na(Month_Courses) %>%
  ggplot(aes(x = Year, y = Month_Courses, fill=Course_number)) +
    geom_tile() +
    scale_fill_viridis(name = "# courses", option = "B") +
    # theme_ipsum() +
    theme_bw() +
    theme(
      legend.position = "right",
```

```
        axis.title = element_text(size = 14),
        axis.text = element_text(size = 14),
        legend.text = element_text(size = 11),
        legend.title = element_text(size = 13),
        text = element_text(family = "Lato")
    ) +
    labs(x = "Year",
         y = "Month")
```
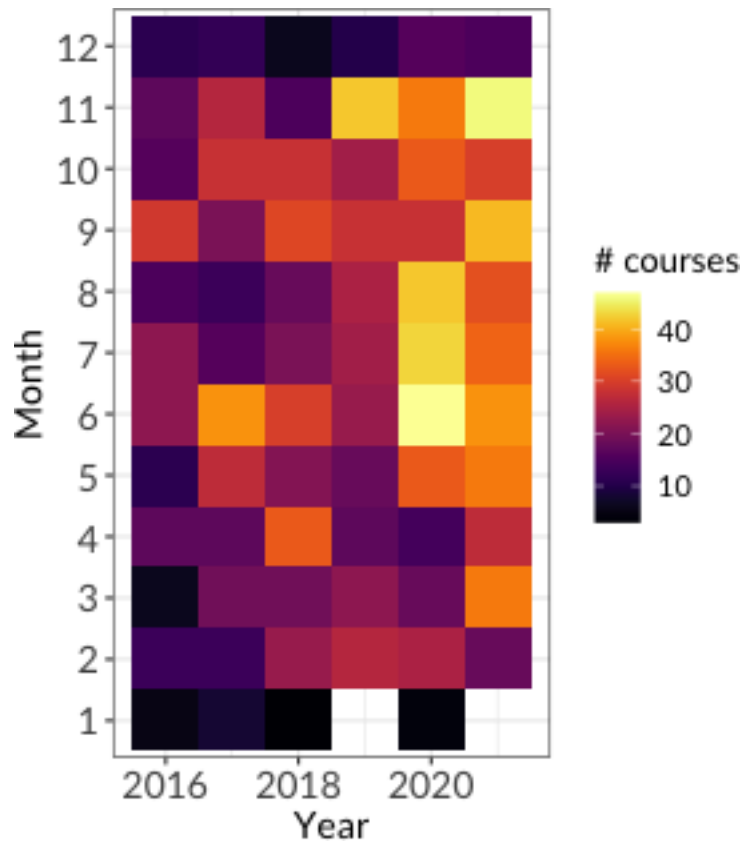


```
## Try the same graph but with normalised data per year
```

As per the distribution of attendees per tool/technology, R was the most popular (23.7%) followed by Python (19%). Excel came in third place at 13%. These three tools, in total, attracted more than 50% of the course attendees. Approximately 10% of the attendees learnt the data collection tools, 6% for REDCap and 4% for Qualtrics. It is also worth noting slightly over 4% of the attendees studied NVivo for qualitative data analysis, which is the fifth most popular tool in Intersect's course category.

```
Attendees_researchers %>%
  filter(Tool.Technology != "#N/A") %>%
  group_by(Tool.Technology) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(Frequency)) %>%
  top_n(8) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(Tool.Technology, ": ", Frequency)) %>%
```

```r
ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=Tool.Technology)) +
  geom_rect() +
  geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  scale_fill_brewer(palette=4) +
  coord_polar(theta="y") +
  xlim(c(1.5, 4)) +
  theme_void() +
  theme(legend.position = "none")
```
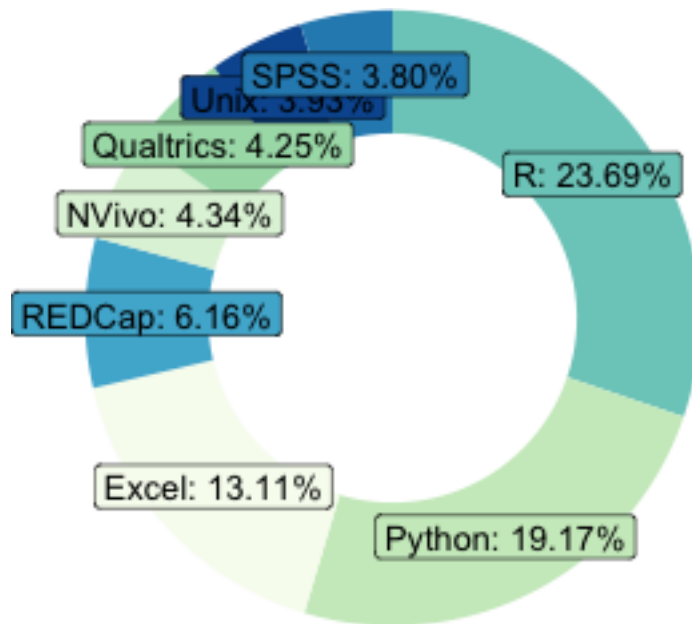

Fig. X shows the number of attendees per tool/technology (6 most popular) from 2016 to 2021. Excel was the most popular tool till 2018 when programming tools jumped to the top of the list. Since then, there has been a steady uptake of both R and Python by training attendees till 2020 followed by a significant increase in 2021. Around 3,000 attendees have learnt R and Python that year. In terms of the two data collection tools, they became popular in 2020, with approximately 500 attendees learnt each tool respectively. Intersect started training on NVivo in 2020 and it overtook REDCap in 2021, to be the fourth most popular course of the year, behind R, Python and Excel.

```r
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  group_by(year, Tool.Technology) %>%
  count() %>%
  filter(Tool.Technology != "#N/A") %>%
  drop_na() %>%
  arrange(year,desc(n)) %>%
  group_by(year) %>%
  slice_max(order_by = n, n = 6) %>%
  mutate(Tool.Technology = factor(Tool.Technology, levels=Tool.Technology)) %>%
  ggplot() +
  geom_bar(aes(x = year, y = n, fill = Tool.Technology, group = n),
           stat = 'identity',
           width = 0.7) +
  scale_fill_viridis(discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
```

```
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # text = element_text(family = "Lato")
  ) +
labs(x = "Year",
     y = "Number of attendees") +
scale_y_continuous(
  limits = c(0, 6000),
  expand = c(0.05, 0.05),
  breaks = seq(0, 6000, 1000),
  minor_breaks = seq(0, 10000, 500)
) +
scale_x_continuous(
  breaks = seq(2016, 2021, 1)
)
```
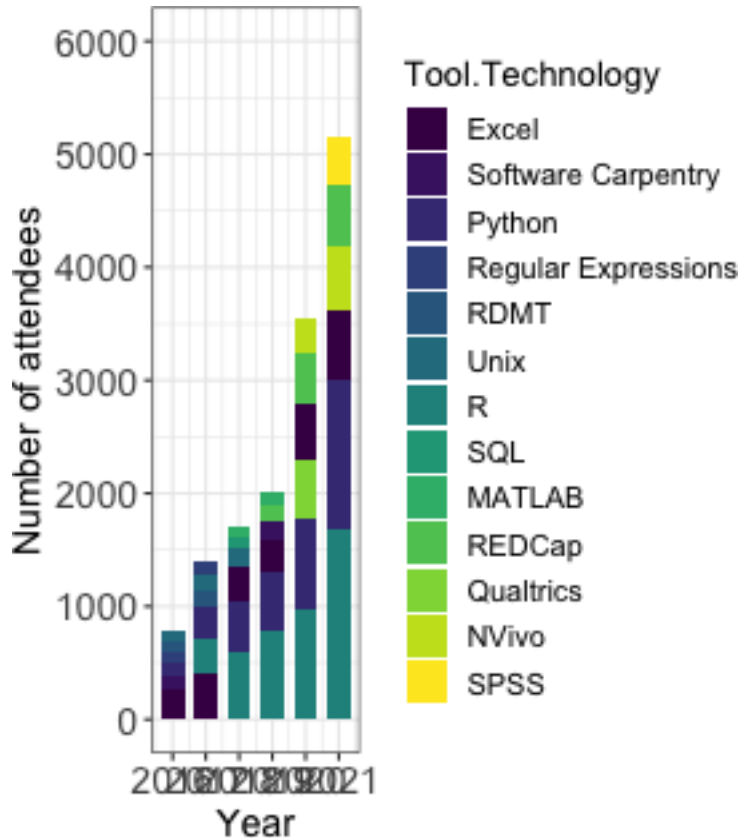

Fig. X shows the popularity of tool/technology for PhD students only, who are the biggest cohort in Intersect training. The distributions are similar to the whole attendee group with programming tools being the most popular, followed by data analysis (Excel, NVivo, SPSS) and data collection tools. It is interesting to see that SPSS is slightly more popular among the PhD students than REDCap.

```
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  filter(Role_Modified == "Higher Degree Research Student (PhD)") %>%
  group_by(year, Tool.Technology) %>%
```

```r
count() %>%
filter(Tool.Technology != "#N/A") %>%
drop_na() %>%
arrange(year,desc(n)) %>%
group_by(year) %>%
slice_max(order_by = n, n = 6) %>%
mutate(Tool.Technology = factor(Tool.Technology, levels=Tool.Technology)) %>%
ggplot() +
geom_bar(aes(x = year, y = n, fill = Tool.Technology, group = n),
         stat = 'identity',
         width = 0.7) +
scale_fill_viridis(discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  # text = element_text(family = "Lato")
) +
labs(x = "Year",
     y = "Number of attendees") +
scale_y_continuous(
  limits = c(0, 3000),
  expand = c(0.05, 0.05),
  breaks = seq(0, 6000, 1000),
  minor_breaks = seq(0, 10000, 500)
) +
scale_x_continuous(
  breaks = seq(2016, 2021, 1)
)
```
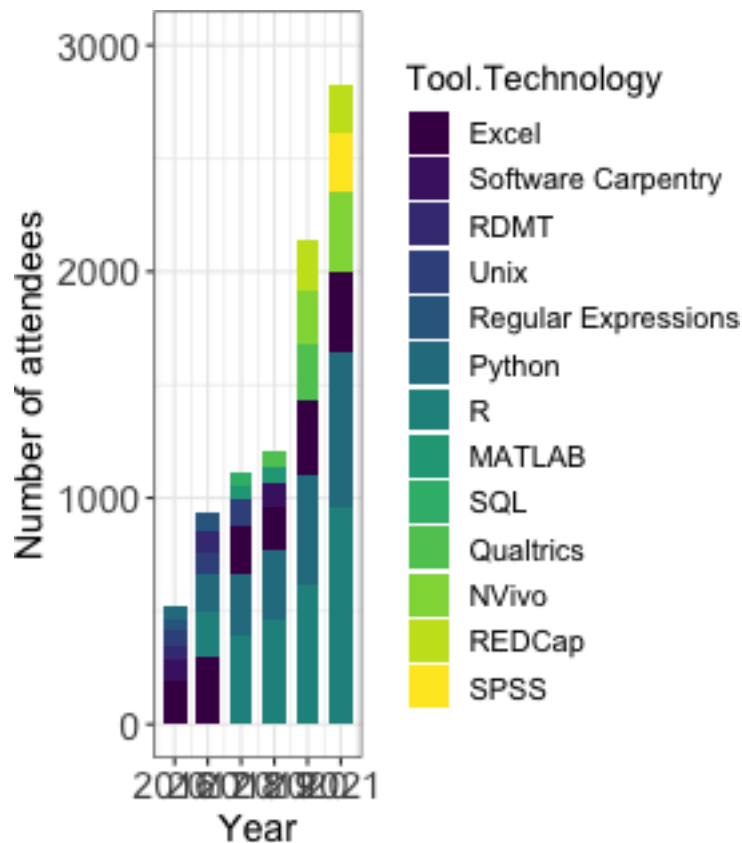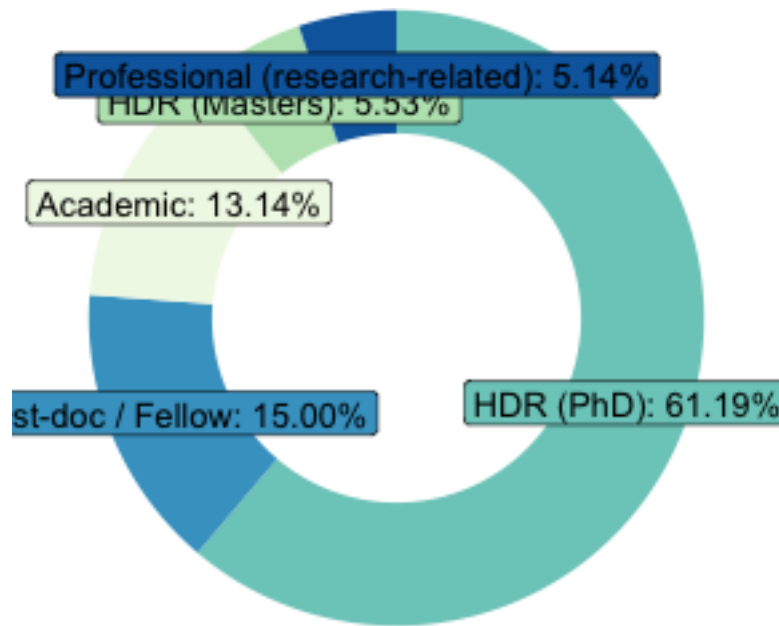
A majority of the training attendees were High Degree Researchers (HDR) PhD students (61%). Exactly 15% were Post-doc / Fellow and 13% were Academics. HDR Master students and Professional staff, who contribute to research shared the remaining 11%.

```
Attendees_researchers %>%
  filter(Role_Modified != "#N/A") %>%
  mutate(Role_Modified = as.factor(fct_recode(Role_Modified,
                             "HDR (PhD)"="Higher Degree Research Student (PhD)",
                             "HDR (Masters)"="Higher Degree Research Student (Masters)"))) %>%
  group_by(Role_Modified) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(Frequency)) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(Role_Modified, ": ", Frequency)) %>%
ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=Role_Modified)) +
  geom_rect() +
  geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  scale_fill_brewer(palette=4) +
  coord_polar(theta="y") +
  xlim(c(1.5, 4)) +
  theme_void() +
  theme(legend.position = "none")
```

Professional (research-related): 5.14%
HDR (Masters): 5.53%
Academic: 13.14%
st-doc / Fellow: 15.00%
HDR (PhD): 61.19%

It can be seen from Fig. X that during 2016 and 2018, a vast majority of attendees were HDR PhD students, Professional (research-related). Since 2019, Academics became the second largest cohort and its proportion steadily increased over recent years. Similarly, an substantial increase in the uptake of Intersect training courses was observed in the professional staff (research-related). This might be a result of ncreasingly number of higher profile researchers and professional staff are undertaking these training courses to up-skill themselves due to the fact that digital tools and technologies are becoming an integral part of the research lifecycle.

```r
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  group_by(year, Role_Modified) %>%
  count() %>%
  filter(Role_Modified != "#N/A") %>%
  drop_na() %>%
  arrange(year,desc(n)) %>%
  group_by(year) %>%
  # slice_max(order_by = n, n = 6) %>%
  mutate(Role_Modified = factor(Role_Modified, levels=Role_Modified)) %>%
  ggplot() +
  geom_bar(aes(x = year, y = n, fill = Role_Modified, group = n),
           stat = 'identity',
           width = 0.7) +
  scale_fill_viridis(name="Role/Position",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # text = element_text(family = "Lato")
  ) +
  labs(x = "Year",
       y = "Number of attendees") +
```

```
scale_y_continuous(
  limits = c(0, 6000),
  expand = c(0.05, 0.05),
  breaks = seq(0, 6000, 1000),
  minor_breaks = seq(0, 10000, 500)
) +
scale_x_continuous(
  breaks = seq(2016, 2021, 1)
)
```

6000 -

5000 -

4000 - Role/Position

           ■ Higher Degree Research Student (PhD)

           ■ Post-doc / Fellow
3000 -
           ■ Professional (research-related)

           ■ Academic
2000 -
           ■ Higher Degree Research Student (Masters

1000 -

   0 -
  2020
  Year

From the faculty distribution of the attendees shown in Fig. X, nearly a third of attendees came from the
Faculty of Medicine and Heath, whereas 28% and 18% were from the Faculty of Science and the Faculty of
Engineering respectively. The Faculty of Arts & Social Sciences accounted for nearly 14% of the attendees,
followed by the Business School, the Faculty of Architecture, Design and Planning, and the Faculty of Law.

```
Attendees_researchers %>%
  filter(Faculty_Modified != "#N/A") %>%
  mutate(Faculty_Modified = as.factor(fct_recode(Faculty_Modified,
                            "Medicine and Health"="Faculty of Medicine and Health",
                            "Engineering"="Faculty of Engineering",
                            "Arts & Social Sciences"="Faculty of Arts and Social Sciences",
                            "Law"="Faculty of Law",
                            "Science"="Faculty of Science",
                            "Architecture, Design & Planning"="Faculty of Architecture, Design a
  group_by(Faculty_Modified) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(Frequency)) %>%
```

```r
mutate(ymax = cumsum(Frequency),
       ymin = c(0, head(ymax, n=-1)),
       labelposition = (ymax+ymin)/2,
       label = paste0(Faculty_Modified, ": ", Frequency)) %>%
ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=Faculty_Modified)) +
  geom_rect() +
  geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  scale_fill_brewer(palette=4) +
  coord_polar(theta="y") +
  xlim(c(1.5, 4)) +
  theme_void() +
  theme(legend.position = "none")
```
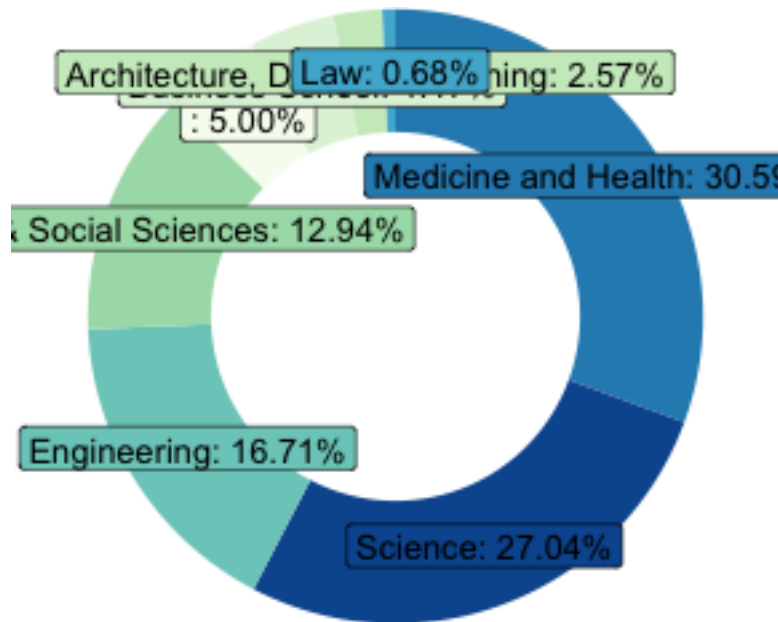

Fig. X presents the faculty breakdown of the attendees per year. It is interesting to see that the number of attendee from Faculty of Medicine and Health was similar to that of the Faculty of Science between 2016 and 2018. Since 2019, the Faculty of Medicine and Health has become the top consumer of Intersect training program. More importantly, the Faculty of Arts & Social Sciences has overceeded the Faculty of Engineering, taking the third place of the number of attendees per year in 2020 and 2021. This, together with the increase of the attendees from the Business School, the Faculty of Architecture, Design and Planning, and the Faculty of Law, indicates an good uptake of the Intersect's digital research raining program by faculties which are conventionally less technology-oriented.
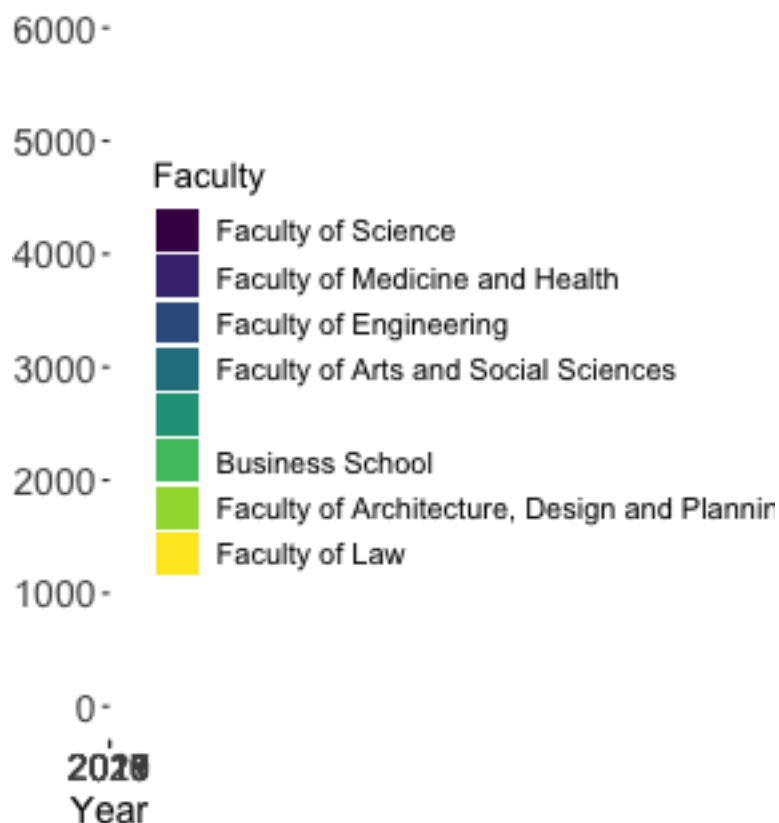
```r
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  group_by(year, Faculty_Modified) %>%
  count() %>%
  filter(Faculty_Modified != "#N/A") %>%
  drop_na() %>%
  arrange(year,desc(n)) %>%
  group_by(year) %>%
  # slice_max(order_by = n, n = 6) %>%
  mutate(Faculty_Modified = factor(Faculty_Modified, levels=Faculty_Modified)) %>%
  ggplot() +
```

```r
geom_bar(aes(x = year, y = n, fill = Faculty_Modified, group = n),
         stat = 'identity',
         width = 0.7) +
scale_fill_viridis(name="Faculty",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  # text = element_text(family = "Lato")
) +
labs(x = "Year",
     y = "Number of attendees") +
scale_y_continuous(
  limits = c(0, 6000),
  expand = c(0.05, 0.05),
  breaks = seq(0, 6000, 1000),
  minor_breaks = seq(0, 10000, 500)
) +
scale_x_continuous(
  breaks = seq(2016, 2021, 1)
)
```


Analysis of the data based on the Field of Research (FoR) code shows Health Sciences is on top of the list with approximately 26% of the attendees followed by Engineering, Biological Sciences, Psychology and

Environmental Sciences as shown in Fig.X.

```
Attendees_researchers %>%
  filter(FoR_Modified != "#N/A") %>%
  group_by(FoR_Modified) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(Frequency)) %>%
  top_n(10) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(FoR_Modified, ": ", Frequency)) %>%
  ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=FoR_Modified)) +
    geom_rect() +
    geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    scale_fill_brewer(palette=4) +
    coord_polar(theta="y") +
    xlim(c(1, 4)) +
    theme_void() +
    theme(legend.position = "none")
```
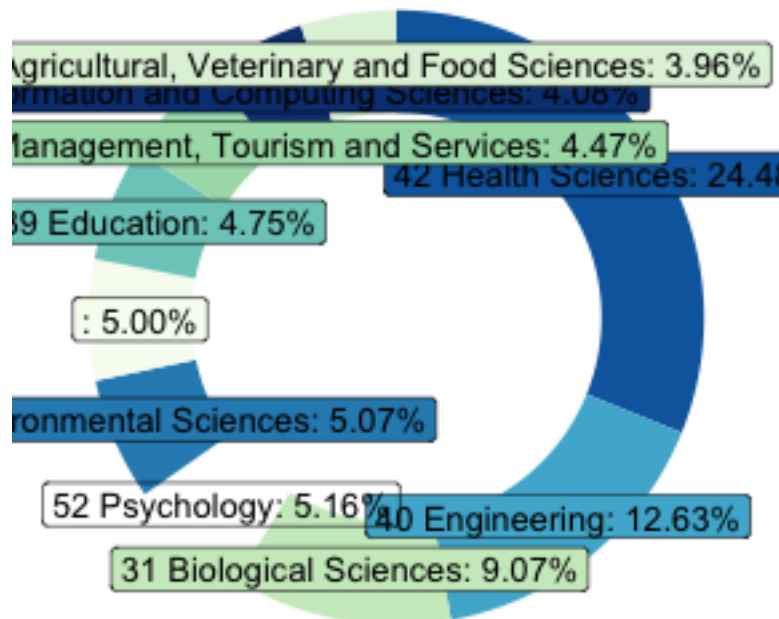

It is encouraging that the numbers of attendees by FOR code have been increasing over the years. Particularly, Biomedical and Clinical Sciences was the fourth on the list with almost all attendees participating in courses on REDCap and R programming.

```
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  group_by(year, FoR_Modified) %>%
  count() %>%
  filter(FoR_Modified != "#N/A") %>%
  drop_na() %>%
  arrange(year,desc(n)) %>%
  group_by(year) %>%
```

```r
slice_max(order_by = n, n = 8) %>%
mutate(FoR_Modified = factor(FoR_Modified, levels=FoR_Modified)) %>%
ggplot() +
geom_bar(aes(x = year, y = n, fill = FoR_Modified, group = n),
         stat = 'identity',
         width = 0.7) +
scale_fill_viridis(name="FOR code",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  # text = element_text(family = "Lato")
) +
labs(x = "Year",
     y = "Number of attendees") +
scale_y_continuous(
  limits = c(0, 6000),
  expand = c(0.05, 0.05),
  breaks = seq(0, 6000, 1000),
  minor_breaks = seq(0, 10000, 500)
) +
scale_x_continuous(
  breaks = seq(2016, 2021, 1)
)
```

FOR code

- 42 Health Sciences
- 40 Engineering
- 41 Environmental Sciences
- 31 Biological Sciences
- 51 Physical Sciences
- 39 Education
- 30 Agricultural, Veterinary and Food Sciences
- 52 Psychology
- 34 Chemical Sciences
- 35 Commerce, Management, Tourism and Ser
- 46 Information and Computing Sciences
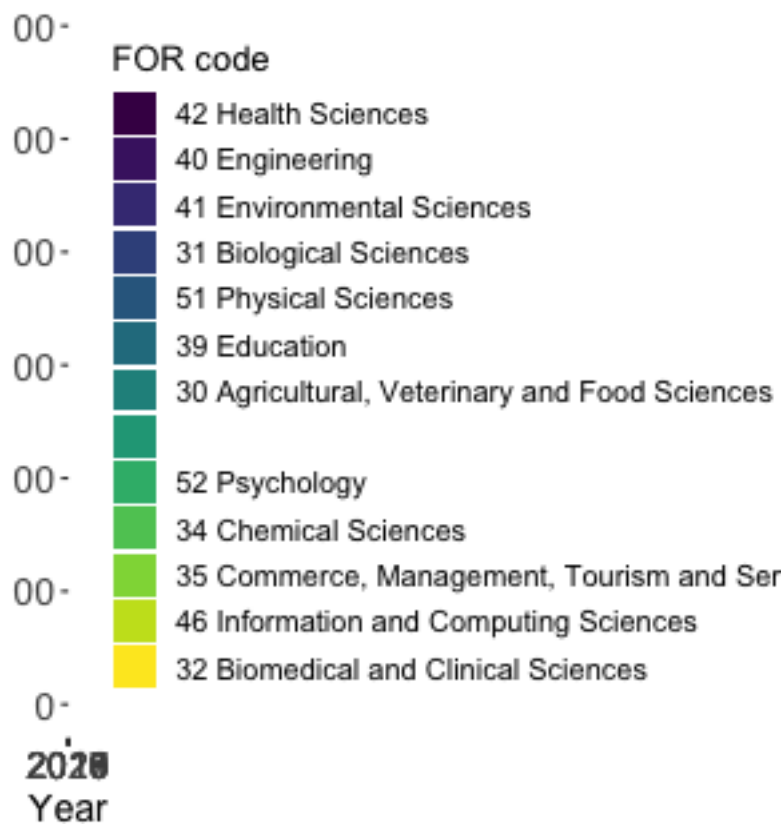- 32 Biomedical and Clinical Sciences

2020
Year

Fig. X presents the attendance breakdown by Role/Position for each Tool/Technology. As expected, HDR PhD student was the top consumer of all tools and technologies while Post-doc / Fellow was the second-highest except for tools and technologies that are related to research data management, qualitative analysis, and statistics. These include Research Data Management Techniques (RDMT), Web Scraping, Tableau, SPSS, REDCap, Qualtrics and NVivo.

```
### <------ NEED TO DO: Sort by most popular tool
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  group_by(Tool.Technology,Role_Modified) %>%
  count() %>%
  filter(Tool.Technology != "#N/A") %>%
  drop_na() %>%
  arrange(desc(n)) %>%
  group_by(Tool.Technology) %>%
  # slice_max(order_by = n,  n = 6) %>%
  mutate(Role_Modified = factor(Role_Modified, levels=Role_Modified)) %>%
  ggplot() +
  geom_bar(aes(x = Tool.Technology, y = n, fill = Role_Modified, group = n),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_viridis(name="Role/Position",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
```

```r
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # text = element_text(family = "Lato")
  ) +
  labs(x = "Tool/Technology",
       y = "Percentage of attendees") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```
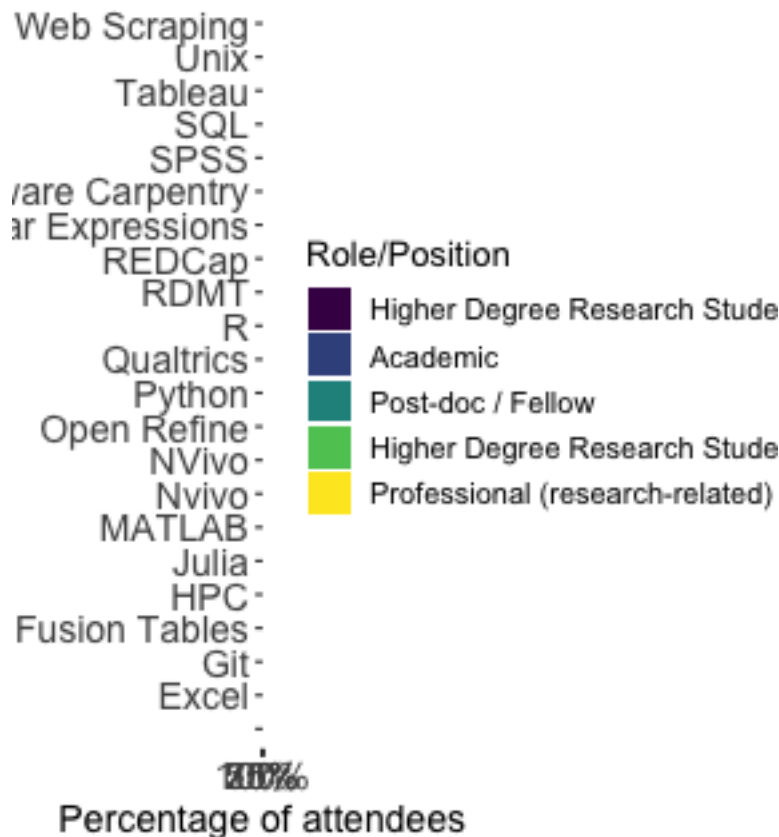

Percentage of attendees

Fig. X presents the attendance breakdown by Faculty for each Tool/Technology. It is worth noting that there is an increased uptake of NVivo and Qualtrics, as well as Tableau and Web Scraping courses, by the Faculty of Arts and Social Sciences, in comparison to the other faculties.

In terms of the surveying tools, REDCap attracted more attendees from the Faculty of Medicine and Health, followed by the Faculty of Science. On the other hand, the attendance of Qualtrics was almost equally distributed across the Faculty of Medicine and Health and the Faculty of Arts and Social Sciences. Also note the third-highest attendance for Qualtrics was Business School (~25%).

Regarding programming courses, Python showed a different distribution compared to R. A vast majority of Python course attendees were from the Faculty of Science and Faculty of Engineering whilst the Faculty of Medicine and Health and the Faculty of Science were more interested in learning R. Lastly, attendees from the Faculty of Law were attracted by the courses on data analysis and managed, e.g., survey tools, Excel, NVivo.
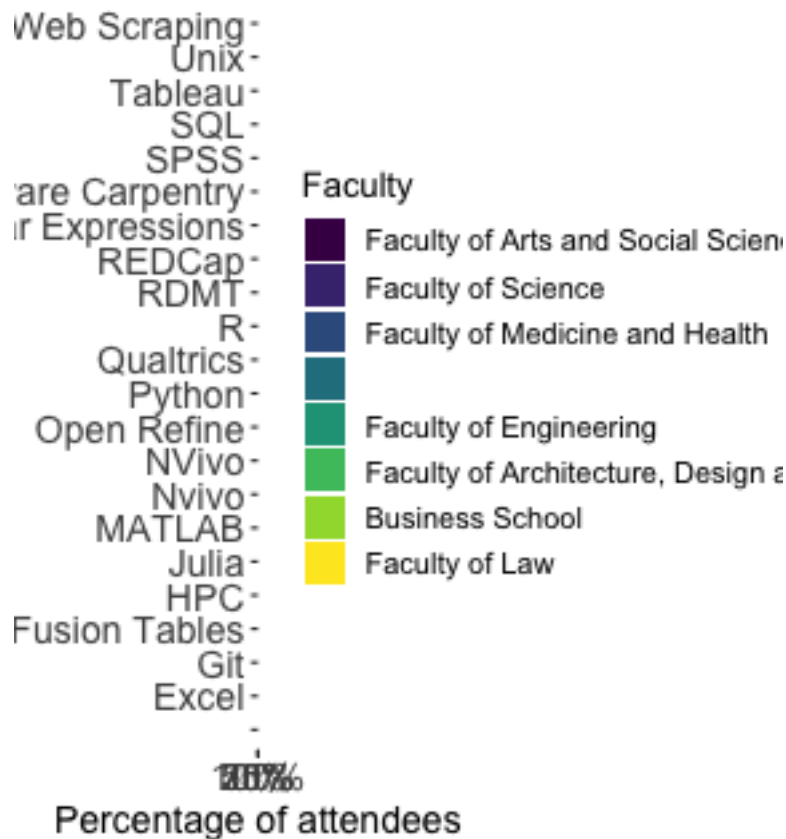
```r
### <------ NEED TO DO: Sort by most popular tool
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  group_by(Tool.Technology,Faculty_Modified) %>%
```

19

```r
count() %>%
filter(Tool.Technology != "#N/A") %>%
filter(Faculty_Modified != "#N/A") %>%
drop_na() %>%
arrange(desc(n)) %>%
group_by(Tool.Technology) %>%
# slice_max(order_by = n, n = 6) %>%
mutate(Faculty_Modified = factor(Faculty_Modified, levels=Faculty_Modified)) %>%
ggplot() +
geom_bar(aes(x = Tool.Technology, y = n, fill = Faculty_Modified, group = n),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_viridis(name="Faculty",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  # text = element_text(family = "Lato")
) +
labs(x = "Tool/Technology",
     y = "Percentage of attendees") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```

Web Scraping-
Unix-
Tableau-
SQL-
SPSS-
are Carpentry-
r Expressions-
REDCap-
RDMT-
R-
Qualtrics-
Python-
Open Refine-
NVivo-
Nvivo-
MATLAB-
Julia-
HPC-
Fusion Tables-
Git-
Excel-

Faculty

Faculty of Arts and Social Scien

Faculty of Science

Faculty of Medicine and Health

Faculty of Engineering

Faculty of Architecture, Design a

Business School

Faculty of Law

Percentage of attendees

According to Fig. X (Role/Position breakdown per Faculty), 50% or more of the attendees for all faculties were HDR PhD students. Regarding the three faculties with the highest number of attendees, (the Faculty of Medicine and Health, the Faculties of Science, and the Faculty of Engineering), the Post-doc / Fellow was the second consumer. However, significant more Academics than Post-doc / Fellows from the Faculty of Law, the Faculty of Arts and Social Sciences, the Faculty of Architecture, Design and Planning, and the Business School attended the courses.

```r
### <------ NEED TO DO: Sort by most popular tool
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  group_by(Faculty_Modified,Role_Modified) %>%
  count() %>%
  filter(Role_Modified != "#N/A") %>%
  filter(Faculty_Modified != "#N/A") %>%
  drop_na() %>%
  arrange(desc(n)) %>%
  group_by(Faculty_Modified) %>%
  # slice_max(order_by = n, n = 6) %>%
  mutate(Role_Modified = factor(Role_Modified, levels=Role_Modified)) %>%
  ggplot() +
  geom_bar(aes(x = Faculty_Modified, y = n, fill = Role_Modified, group = n),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_viridis(name="Role/Position",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
```

```
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # text = element_text(family = "Lato")
  ) +
  labs(x = "Faculty",
       y = "Percentage of attendees") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

Faculty of Science -

ty of Medicine and Health -

Faculty of Law -

Faculty of Engineering -

Arts and Social Sciences -

:ure, Design and Planning -

Business School -

-

Role/Position

■ Higher Degree R|
■ Post-doc / Fellow
■ Academic
■ Professional (res|
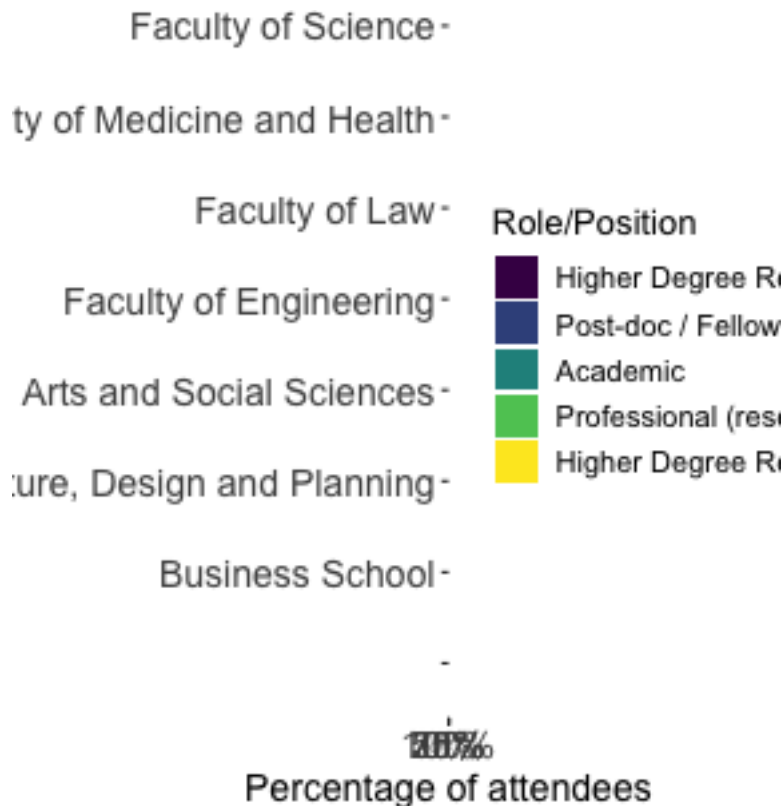■ Higher Degree R|

Percentage of attendees

Fig. X and Fig. X present the course competency (difficulty of the course) per Role/Position and per Faculty respectively. For all Role/Position groups, aproximately three quarters of the have participated in the introductory courses while the remaining have attended the intermediate level courses. Only a few percentages of Post-doc / Fellow and HDR PhD have selected the advanced courses.

It is even more interesting to see the various distribution of course levels across the faculties. More attendees from the Faculty of Engineering and the Faculty of Science (25%+) favoured intermediate and advanced courses, in comparison to the Faculty of Law and the Faculty of Arts and Social Sciences. As mentioned earlier, in Faculty of Law and the Faculty of Arts and Social Sciences, the introductory-level courses such as Excel, NVivo, and survey tools are more popular.

```
### <------ NEED TO DO: Sort by most popular tool
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  group_by(Role_Modified, Level) %>%
  count() %>%
  filter(Role_Modified != "#N/A") %>%
```

```r
filter(Level != "#N/A") %>%
drop_na() %>%
arrange(desc(n)) %>%
group_by(Role_Modified) %>%
# slice_max(order_by = n, n = 6) %>%
mutate(Level = factor(Level, levels=Level)) %>%
ggplot() +
geom_bar(aes(x = Role_Modified, y = n, fill = Level, group = n),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_viridis(name="Course level",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  # text = element_text(family = "Lato")
) +
labs(x = "Role/Position",
     y = "Percentage of attendees") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```

Professional (research-related)-

Post-doc / Fellow-

Cours

Int

gher Degree Research Student (PhD)-

Int

Ad

· Degree Research Student (Masters)-

Academic-

100%

Percentage of attend

```r
### <------ NEED TO DO: Sort by most popular tool
Attendees_researchers %>%
```

```
filter(status == "Checked In") %>%
group_by(Faculty_Modified, Level) %>%
count() %>%
filter(Faculty_Modified != "#N/A") %>%
filter(Level != "#N/A") %>%
drop_na() %>%
arrange(desc(n)) %>%
group_by(Faculty_Modified) %>%
# slice_max(order_by = n, n = 6) %>%
mutate(Level = factor(Level, levels=Level)) %>%
ggplot() +
geom_bar(aes(x = Faculty_Modified, y = n, fill = Level, group = n),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_viridis(name="Course level",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  # text = element_text(family = "Lato")
) +
labs(x = "Faculty",
     y = "Percentage of attendees") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
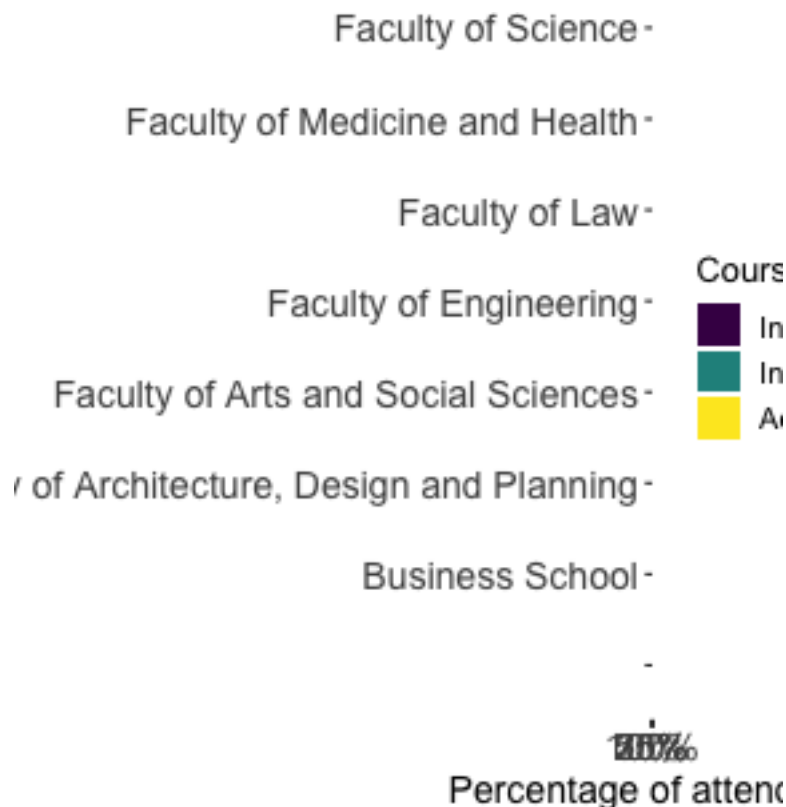
Fig. X shows the trend of attendee numbers by state since 2016. Prior to 2019, almost all attendees were affiliated with an institution in New South Wales, Australia. As Intersect transitioned its training delivery to online since the beginning of 2020, nearly 4,500 attendees joined the course, almost 2,000 higher than the previous year. The same trend continued to 2021, with an additional of 1,500 attendees being trained. The reason becuase might be that online training is more accessible to researchers and research professionals who are not based at main university campuses. Flexibility in terms of time commitment could be another factor as it is easier to commit two half-days online rather than a full day in person.

```
Attendees_researchers %>%
  filter(status == "Checked In") %>%
  group_by(year, State) %>%
  count() %>%
  filter(State != "#N/A") %>%
  drop_na() %>%
  arrange(year,desc(n)) %>%
  group_by(year) %>%
  slice_max(order_by = n, n = 8) %>%
  mutate(State = factor(State, levels=State)) %>%
  ggplot() +
  geom_bar(aes(x = year, y = n, fill = State, group = n),
           stat = 'identity',
           width = 0.7) +
  scale_fill_viridis(discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
```

```
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # text = element_text(family = "Lato")
  ) +
labs(x = "Year",
     y = "Number of attendees") +
scale_y_continuous(
  limits = c(0, 6000),
  expand = c(0.05, 0.05),
  breaks = seq(0, 6000, 1000),
  minor_breaks = seq(0, 10000, 500)
) +
scale_x_continuous(
  breaks = seq(2016, 2021, 1)
)
```


According to Fig. X, which summarizes the reasons for attending the course. A majority of participants expressed that they are interested in these courses to learn skills that they can either apply to their work now or in the future. Also, about 12% of the attendees chosed to learn these skills for better opportunities and employability in the future ("To learn skills that will help me get a job").

```
reason_clean %>%
  group_by(reason) %>%
  summarise(Count=n()) %>%
  # Combine last two categories
  mutate("Count_new" = ifelse(reason %in% c("undefined", "Other"), sum(Count[reason %in% c("undefined",
  filter(reason!="undefined") %>%
```

```
  select(-Count) %>%
  mutate(Frequency = formattable::percent(Count_new/ sum(Count_new))) %>%
  arrange(desc(Frequency)) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(reason, ": ", Frequency)) %>%
ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=reason)) +
  geom_rect() +
  geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  scale_fill_brewer(palette=4) +
  coord_polar(theta="y") +
  xlim(c(1, 4)) +
  theme_void() +
  theme(legend.position = "none")
```
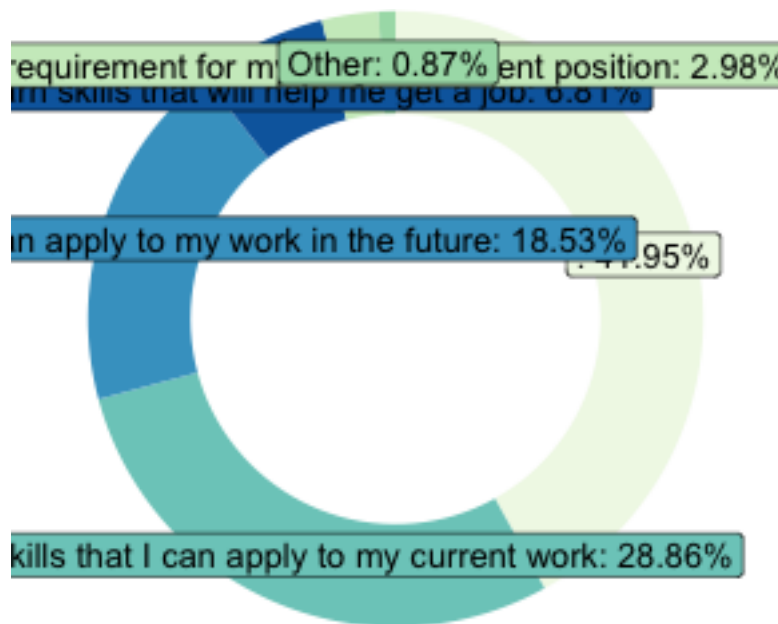

Fig. X below demonstrates the reasons for attending the course for by Role/Position and by Faculty. The percentage of participants who learn skills for current and future work is higher amongst Academics and Post-docs compared to Undergraduate, Masters and HDR students. For the students, more tended to learning these skills for better opportunities in the future. Analysis on the reason for being interested in attending this training course by the Faculty reveals a consistent trend across all faculties.

```
reason_by_role <- reason_clean %>%
  filter(Role_Modified != "#N/A") %>%
  group_by(reason, Role_Modified) %>%
  drop_na() %>%
  summarise(Count = n()) %>%
  group_by(Role_Modified) %>%
  slice_max(order_by = desc(Count), n=6) %>%
  mutate(Role_Modified = factor(Role_Modified, levels=unique(Role_Modified))) %>%
  mutate("Count_new" = ifelse(reason %in% c("undefined", "Other"), sum(Count[reason %in% c("undefined",
  filter(reason!="undefined") %>%
  select(-Count) %>%
```

```
    mutate(reason = factor(reason,levels = reason))

reason_by_role %>%
  ggplot() +
  # ordered by count per role/position
  geom_bar(aes(x = reorder(Role_Modified,Count_new,FUN=sum), y = Count_new, fill = reason),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_viridis(name="Role/Position",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # text = element_text(family = "Lato")
  ) +
  labs(x = "Role/Position",
       y = "Percentage of attendees") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```



```
reason_by_faculty <- reason_clean %>%
  filter(Faculty_Modified != "#N/A") %>%
  group_by(reason, Faculty_Modified) %>%
  drop_na() %>%
```

```r
    summarise(Count = n()) %>%
    group_by(Faculty_Modified) %>%
    slice_max(order_by = desc(Count), n = 6) %>%
    mutate(
      Faculty_Short = case_when(
        Faculty_Modified == "Faculty of Medicine and Health" ~ "Medicine and Health",
        Faculty_Modified == "Faculty of Architecture, Design and Planning" ~ "Architecture, Design & Plan
        Faculty_Modified == "Faculty of Engineering" ~ "Engineering",
        Faculty_Modified == "Faculty of Arts and Social Sciences" ~ "Arts & Social Sciences",
        Faculty_Modified == "Faculty of Law" ~ "Law",
        Faculty_Modified == "Faculty of Science" ~ "Science",
        Faculty_Modified == "Business School" ~ "Business"
      )
    ) %>%
    mutate("Count_new" = ifelse(reason %in% c("undefined", "Other"), sum(Count[reason %in% c("undefined",
    filter(reason!="undefined") %>%
    select(-Count) %>%
    mutate(reason = factor(reason,levels = reason))

reason_by_faculty %>%
  ggplot() +
  # ordered by count per faculty
  geom_bar(aes(x = reorder(Faculty_Short,Count_new,FUN=sum), y = Count_new, fill = reason),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_viridis(name="Role/Position",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # text = element_text(family = "Lato")
  ) +
  labs(x = "Faculty",
       y = "Percentage of attendees") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

NA -

cine and Health -

Science -

Role/Position

■ Other
■ As a requirement for my program
■ To learn skills that will help me ge
■ To learn skills that I can apply to r
■ To learn skills that I can apply to r
■

Social Sciences -

Engineering -

Business -

sign & Planning -

Law -

11.1%

Percentage of attendees

The following graphs illustrate the communication of Intersect training program to attendees.

26% of the participants found out about the Intersect training program through the Faculty/School newsletter. The second top option is via the Research office/division (website/email) with ~16% followed by the University newsletter and University website. Notably, a considerable number of participants learn about these courses through their supervisors (~7%).

```
hear_clean %>%
  group_by(hear_about_course) %>%
  summarise(Count=n())%>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(Frequency)) %>%
  top_n(8) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(hear_about_course, ": ", Frequency)) %>%
  ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=hear_about_course)) +
    geom_rect() +
    geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    scale_fill_brewer(palette=4) +
    coord_polar(theta="y") +
    xlim(c(1, 4)) +
    theme_void() +
    theme(legend.position = "none")
```
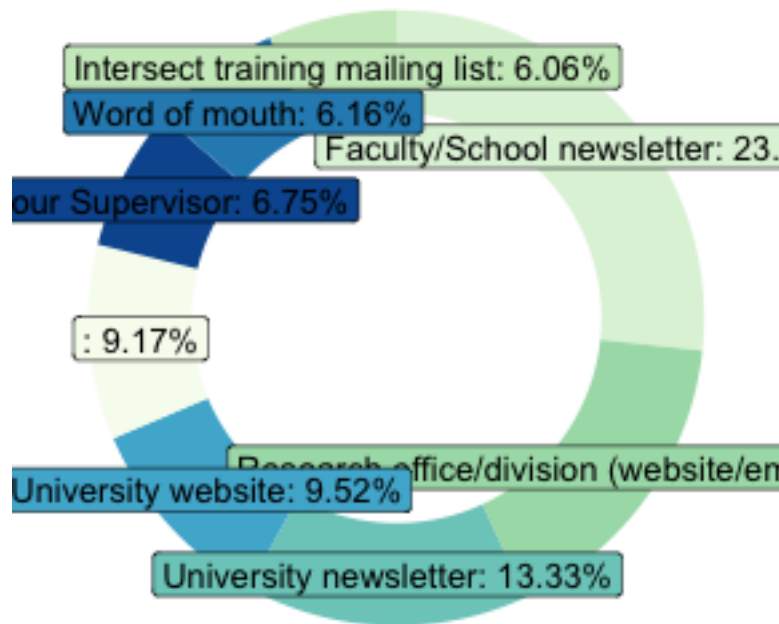
Fig. X shows further analysis of the communication methods by Role/Position and by Faculty. Faculty/School newsletter, Research office/division (website/email) and University newsletter were the major sources from which researchers acquired the training information. It is also interesting to note that more than 10% of HDR Masters students and Professional staff (research-related) heard about the courses from their supervisors.

The distribution of the communication methods shows a consistent trend across all faculties. It is interesting to know different to the other faculties, the Business School attendees.

```r
hear_by_role <- hear_clean %>%
  # select(hear_about_course,Role_Modified) %>%
  filter(Role_Modified != "#N/A") %>%
  group_by(hear_about_course, Role_Modified) %>%
  drop_na() %>%
  summarise(Count = n()) %>%
  group_by(Role_Modified) %>%
  slice_max(order_by = Count, n=8) %>%
  mutate(Role_Modified = factor(Role_Modified, levels=unique(Role_Modified))) %>%
  mutate(hear_about_course = factor(hear_about_course,levels = hear_about_course))

hear_by_role %>%
  ggplot() +
  # ordered by count per role/position
  geom_bar(aes(x = reorder(Role_Modified,Count,FUN=sum), y = Count, fill = hear_about_course),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_viridis(name="Role/Position",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
```

31

```
    # text = element_text(family = "Lato")
  ) +
  labs(x = "Role/Position",
       y = "Percentage of attendees") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```



```
hear_by_faculty <- hear_clean %>%
  filter(Faculty_Modified != "#N/A") %>%
  group_by(hear_about_course, Faculty_Modified) %>%
  drop_na() %>%
  summarise(Count = n()) %>%
  group_by(Faculty_Modified) %>%
  slice_max(order_by = Count, n = 7) %>%
  mutate(
    Faculty_Short = case_when(
      Faculty_Modified == "Faculty of Medicine and Health" ~ "Medicine and Health",
      Faculty_Modified == "Faculty of Architecture, Design and Planning" ~ "Architecture, Design & Plann
      Faculty_Modified == "Faculty of Engineering" ~ "Engineering",
      Faculty_Modified == "Faculty of Arts and Social Sciences" ~ "Arts & Social Sciences",
      Faculty_Modified == "Faculty of Law" ~ "Law",
      Faculty_Modified == "Faculty of Science" ~ "Science",
      Faculty_Modified == "Business School" ~ "Business"
    )
  ) %>%
  mutate(hear_about_course = factor(hear_about_course,levels = hear_about_course))

hear_by_faculty %>%
  ggplot() +
```

```r
# ordered by count per faculty
geom_bar(aes(x = reorder(Faculty_Short,Count,FUN=sum), y = Count, fill = hear_about_course),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_viridis(name="Role/Position",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 12),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  # text = element_text(family = "Lato")
) +
labs(x = "Faculty",
     y = "Percentage of attendees") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```

NA -

Medicine and Health -

Role/Position

Science -

■ Faculty/School newsletter

■

Engineering -

■ Research office/division (w

■ University newsletter

■ University website

rts & Social Sciences -

■ Word of mouth

Business -

■ Your Supervisor

■ Intersect training mailing li

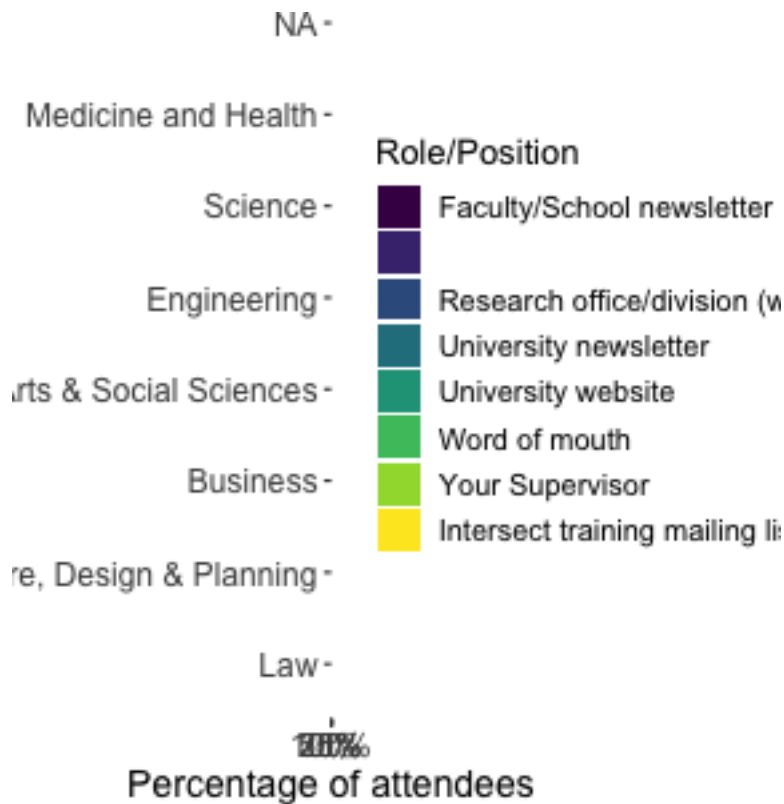re, Design & Planning -

Law -

100%

Percentage of attendees

Fig. X shows the distribution of the number of times that the same participant has attended the Intersect training program. Around 65% of the participants attended the course once. It is more encouraging to see nearly 20% have attended two courses, followed by ~8% have attended three, and 3.5% have attended four courses. Approximately 4% (1,000+) of the total number of attendees participated the program 5 times and more, which implies that the Intersect training courses were highly regarded by those attendees.

```r
# Group by email to get # of returns
# # Group by number of returns to get # of attendees with same # of returns
```

```r
return_to_course <- Attendees_researchers %>%
  filter(status == "Checked In") %>%
  select(year, email, Faculty_Modified, Role_Modified, FoR_Modified) %>%
  group_by(email) %>%
  count() %>%
  arrange(desc(n)) %>%
  group_by(n) %>%
  summarise(Count = n())

# Aggregate 10+ returns
# Show as %
return_to_course_percent <- return_to_course %>%
  top_n(9) %>%
  rbind(c(10,sum(return_to_course[which(return_to_course$n>=10),2]))) %>%
  mutate(freq = formattable::percent(Count / sum(Count))) %>%
  arrange(n)

# Plot
# top_n(10) %>%
  ggplot(return_to_course_percent)+
  geom_bar(aes(x=n, y=freq),
           stat = 'identity',
           fill = intersect_color1) +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # text = element_text(family = "Lato")
  ) +
  scale_y_continuous(
    limits = c(0, 1),
    expand = c(0.05, 0.05),
    breaks = seq(0, 1, 0.2),
    minor_breaks = seq(0, 1, 0.1),
    labels = scales::percent) +
  scale_x_continuous(
    breaks = seq(1, 10, 1),
    labels = append(as.character(seq(1,9,1)),"10+")
  )+
  labs(x = "Returns to a course",
       y = "") +
  geom_text(aes(x=n, y=freq,label=freq),vjust=-0.25)
```
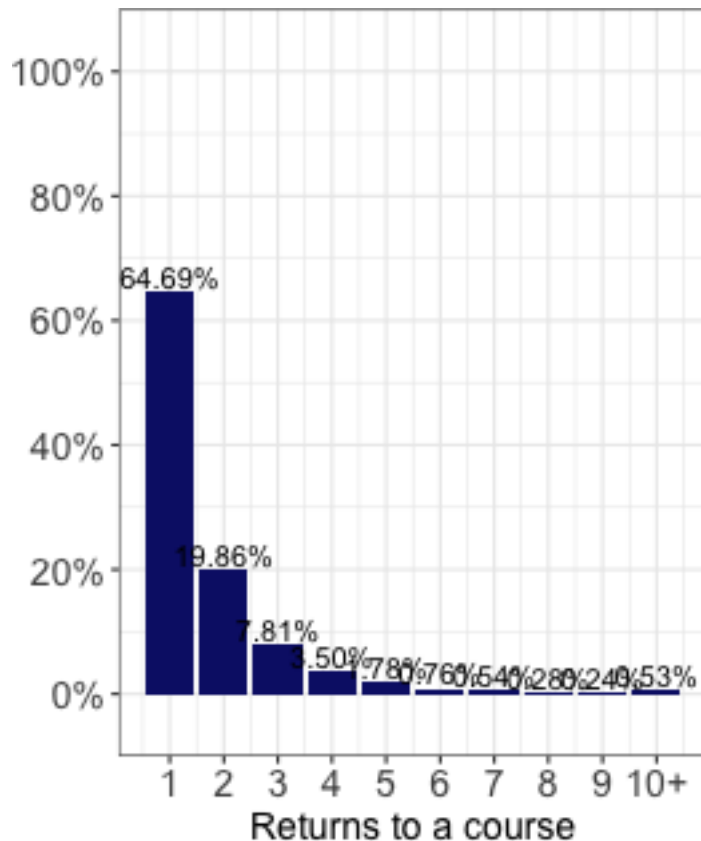
The number of times the attendees returned to a course was further splitted by Role and by Faculty.

As seen from the boxplots shown in Fig. X, the frequencies between 2016 and 2018 were almost consistent with slight shift observed for both HDR PhD students and Professional (research-related) staff. From 2019 and onwards, the frequencies of returning to a course increased significantly for all roles/positions. In particular, the HDR PhD students and Academics were more interested in the Intersect training program with most of the returning attendees joined the course at least 4 times. It is also interesting to note an academic had returned to a course 16 times and another HDR PhD students had attended the training 13 times.

```
### Return to a course
Group_Attendees_researchers_role <- Attendees_researchers %>%
  filter(status == "Checked In") %>%
  select(year, email, Faculty_Modified, Role_Modified) %>%
  group_by(email,Faculty_Modified, Role_Modified,year) %>%
  count() %>%
  arrange(desc(n)) %>%
  group_by(n, Role_Modified, year) %>%
  summarise(Count=n()) %>%
  arrange(desc(n))
  # mutate(Role_Modified = factor(Role_Modified, levels=unique(Role_Modified)))

# # rename factors
# levels(Group_Attendees_researchers_role$Role_Modified) <-
#   list(
#     "HDR (PhD)" = "Higher Degree Research Student (PhD)",
#     "Academic" = "Academic",
#     "HDR (Masters)" =
#       "Higher Degree Research Student (Masters)",
```

35

```r
#     "Post-doc/Fellow" =
#        "Post-doc / Fellow",
#     "Professional (research-related)" =
#        "Professional (research-related)"
#   )

# Boxplot
Group_Attendees_researchers_role %>%
  ggplot() +
  geom_boxplot(aes(
    x = reorder(Role_Modified, n, FUN = median),
    y = n,
    fill = Role_Modified
  )) +
  theme_bw() +
  theme(
    legend.position = "none",
    axis.title = element_text(size = 13),
    axis.text = element_text(size = 11),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # text = element_text(family = "Lato")
  ) +
  labs(x = "",
       y = "Returns to a course")+
  coord_flip() +
  facet_wrap( ~ year)
```
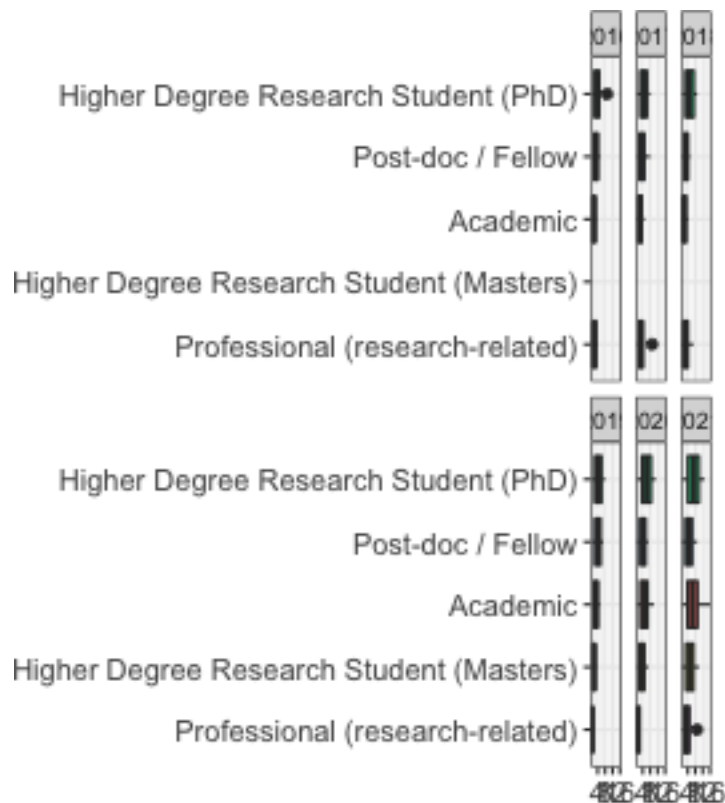
Fig. X below shows the distribution of returns to a course for each faculty. In terms of the top three faculties with the highest number of attendees (Science, Medicine and Health, Engineering), all have had a steady increase of the return frequencies, with the median number rised from round 3 in 2016 to 6 in 2021. In terms of the rest of the faulties, the number of returns to a course barely changed until 2019. However, with the introduction of online training delivery and targeted communication, the Faculty of Arts and Social Sciences, the Business School as well as the Faculty fo Architecture, Design & Planning had a notable increase in return frequencies since 2020. It is worth noting that the median number of returns to the course by attendees from Faculty of Arts and Social Sciences was nearly 6, the highest among all faculties in 2020.

Even though the Faculty of Law did not have similar returns to a course distribution as the other faculties, it is clear that more attendees from Law had returned the course multiple times in recent years.

```
### Return to a course
Group_Attendees_researchers_faculty <- Attendees_researchers %>%
  filter(status == "Checked In") %>%
  select(year, email, Faculty_Modified, Role_Modified) %>%
  group_by(email, Faculty_Modified, Role_Modified, year) %>%
  drop_na(Faculty_Modified) %>%
  filter(Faculty_Modified != "#N/A") %>%
  count() %>%
  arrange(desc(n)) %>%
  group_by(n, Faculty_Modified, year) %>%
  summarise(Count = n()) %>%
  arrange(desc(n)) %>%
  # mutate(Faculty_Modified = factor(Faculty_Modified, levels = Faculty_Modified))
  # rename faculty name
  mutate(Faculty_Short = case_when(
    Faculty_Modified == "Faculty of Medicine and Health" ~ "Medicine and Health",
    Faculty_Modified == "Faculty of Architecture, Design and Planning" ~ "Architecture, Design & Planni
```

```r
    Faculty_Modified == "Faculty of Engineering" ~ "Engineering",
    Faculty_Modified == "Faculty of Arts and Social Sciences" ~ "Arts & Social Sciences",
    Faculty_Modified == "Faculty of Law" ~ "Law",
    Faculty_Modified == "Faculty of Science" ~ "Science",
    Faculty_Modified == "Business School" ~ "Business"
  ))

# Rename factors
# levels(Group_Attendees_researchers_faculty$Faculty_Modified) <-
#   list(
#     "Medicine and Health" = "Faculty of Medicine and Health",
#     "Engineering" =
#        "Faculty of Engineering",
#     "Arts & Social Sciences" =
#        "Faculty of Arts and Social Sciences",
#     "Law" =
#        "Faculty of Law",
#     "Science" =
#        "Faculty of Science",
#     "Architecture, Design & Planning" =
#        "Faculty of Architecture, Design and Planning",
#     "Business" = "Business School"
#   )

# Boxplot
Group_Attendees_researchers_faculty %>%
  ggplot(aes(
    x = reorder(Faculty_Short,n,FUN=median), # ordered by median, desc
    y = n,
    fill = Faculty_Short)) +
  geom_boxplot(
  ) +
  theme_bw() +
  theme(
    legend.position = "none",
    axis.title = element_text(size = 13),
    axis.text = element_text(size = 12),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    # axis.text.x = element_text(size = 10),
    # text = element_text(family = "Lato")
  ) +
  labs(x = "Faculty",
       y = "Returns to a course")+
  coord_flip() +
  facet_wrap( ~ year)
```
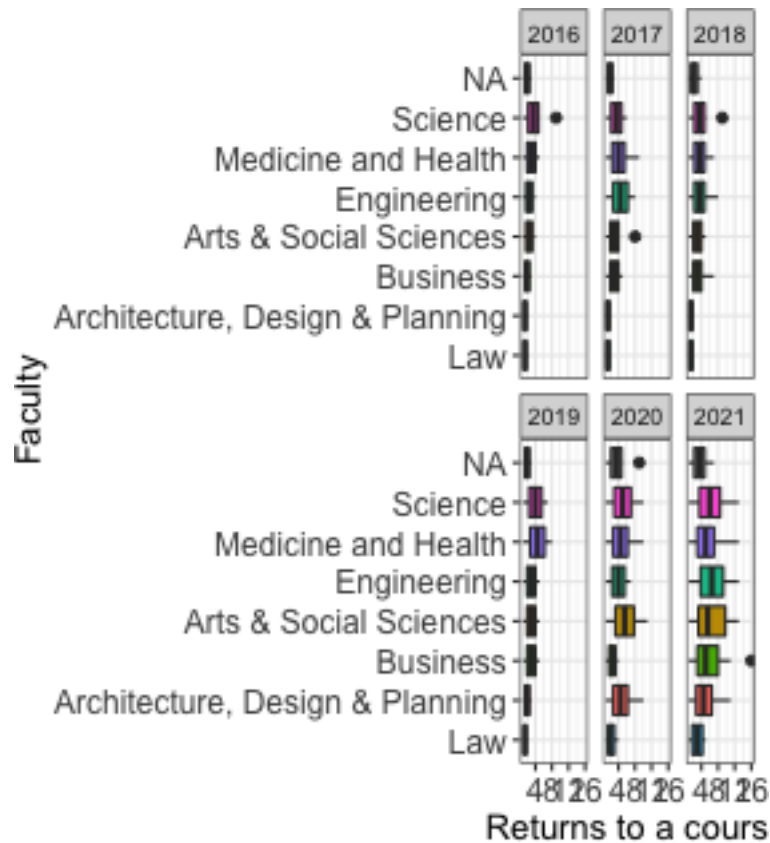
```
# JZ comments:
# - order has been changed after facet_wrap?
```

Training Impact

Demographics of training impact

Higher Degree Research Students (PhD) were the largest role group to respond to the training impact survey. This is unsurprising given the relatively large proportion of training attendees in this role.

```r
Only_Responses %>%
  group_by(Role) %>%
  count() %>%
  arrange(desc(n)) %>%
  mutate(Role = factor(Role, levels=Role)) %>%
  ggplot() +
  geom_bar(aes(x = reorder(Role,n), y = n),
           stat = 'identity',
           width = 0.7,
           fill=intersect_color1) +
  # scale_fill_viridis(discrete = TRUE, option="viridis") +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
```

```r
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
labs(x = "Role/Position",
     y = "Number of Responses") +
coord_flip()
```

Higher Degree Research Student (PhD) -

Academic -

Post-doc / Fellow -

Role/Position

Higher Degree Research Student (Masters) -

Professional (research-related) -

20

Number of

```r
# scale_y_continuous(
#   limits = c(0, 6000),
#   expand = c(0.05, 0.05),
#   breaks = seq(0, 6000, 1000),
#   minor_breaks = seq(0, 10000, 500)
# ) +
# scale_x_continuous(
#   breaks = seq(2016, 2021, 1)
# )
```

The Faculty of Medicine and Health were the largest faculty group to respond to the training impact survey. This group comprises all Medicine and Health faculties, such as Public Health and Psychology.

```r
Only_Responses %>%
  group_by(faculty) %>%
  count() %>%
  drop_na(faculty) %>%
  arrange(desc(n)) %>%
  mutate(faculty = factor(faculty, levels=faculty)) %>%
  ggplot() +
  geom_bar(aes(x = reorder(faculty, n), y = n),
           stat = 'identity',
           width = 0.7,
```
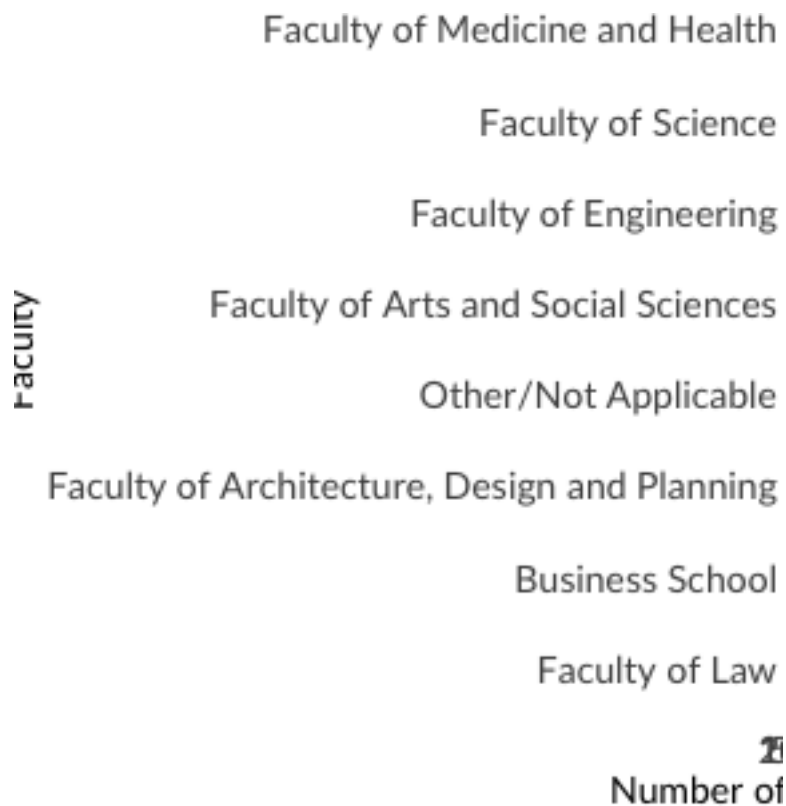
```
        fill = intersect_color1) +
# scale_fill_viridis(discrete = TRUE, option="viridis") +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Faculty",
     y = "Number of Responses") +
coord_flip()
```

Faculty of Medicine and Health

Faculty of Science

Faculty of Engineering

Faculty of Arts and Social Sciences

Other/Not Applicable

Faculty of Architecture, Design and Planning

Business School

Faculty of Law

Faculty

Number of

```
# scale_y_continuous(
#   limits = c(0, 6000),
#   expand = c(0.05, 0.05),
#   breaks = seq(0, 6000, 1000),
#   minor_breaks = seq(0, 10000, 500)
# ) +
# scale_x_continuous(
#   breaks = seq(2016, 2021, 1)
# )
```

Health Sciences were the largest Field of Research (FOR) group to respond to the training impact survey. This aligns with the distribution of responses by faculty.

```
Only_Responses %>%
  group_by(for_code) %>%
  count() %>%
  drop_na(for_code) %>%
  arrange(desc(n)) %>%
  mutate(for_code = factor(for_code, levels=for_code)) %>%
  ggplot(aes(x = reorder(for_code, n), y = n)) +
  geom_bar(stat = 'identity',
           width = 0.7,
           fill = intersect_color1) +
  # scale_fill_viridis(discrete = TRUE, option="viridis") +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "FOR Code",
       y = "Number of Responses") +
  coord_flip()
```

42 HEALTH SC
40 ENGIN
52 PSYCH
31 BIOLOGICAL SC
Other/Not Ap
39 EDU(
41 ENVIRONMENTAL SC
AGRICULTURAL, VETERINARY AND FOOD SC
33 BUILT ENVIRONMENT AND
MMERCE, MANAGEMENT, TOURISM AND SE
51 PHYSICAL SC
46 INFORMATION AND COMPUTING SC
38 ECON
44 HUMAN S
34 CHEMICAL SC
49 MATHEMATICAL SC
37 EARTH SC
48 LAW AND LEGAL S
47 LANGUAGE, COMMUNICATION AND CI
43 HISTORY, HERITAGE AND ARCHAE
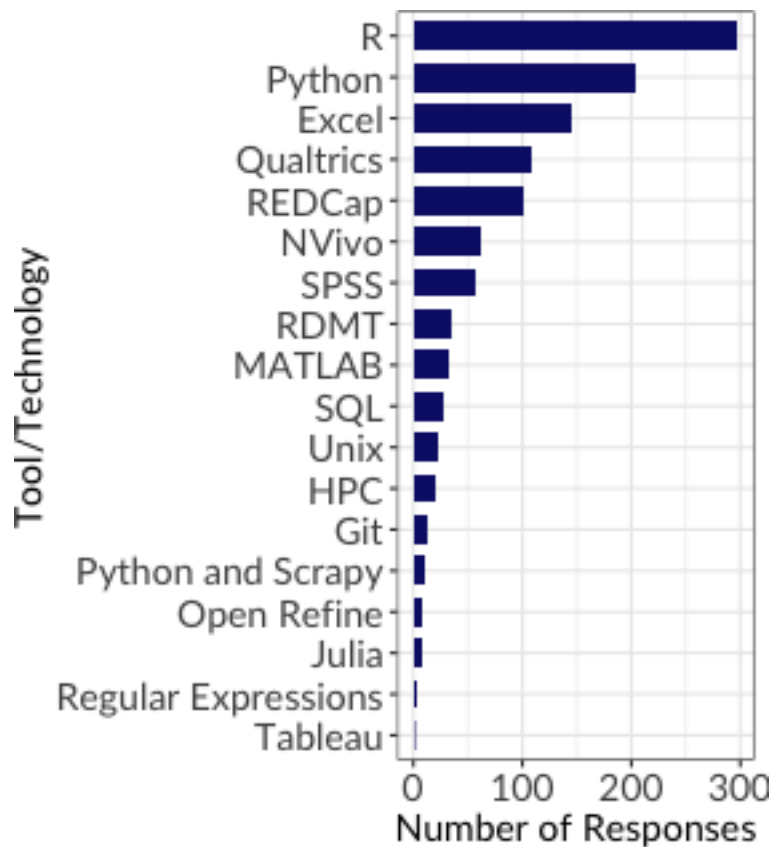36 CREATIVE ARTS AND W

N(

```
# scale_y_continuous(
#    limits = c(0, 6000),
#    expand = c(0.05, 0.05),
```

```
#    breaks = seq(0, 6000, 1000),
#    minor_breaks = seq(0, 10000, 500)
# ) +
# scale_x_continuous(
#    breaks = seq(2016, 2021, 1)
# )
```

The most frequent research technology tool associated with responses to the training impact survey was R (substantially more than the next most frequently reported) with over 250 responses.

```
Only_Responses %>%
  group_by(Technology) %>%
  count() %>%
  drop_na(Technology) %>%
  arrange(desc(n)) %>%
  mutate(Technology = factor(Technology, levels=Technology)) %>%
  ggplot() +
  geom_bar(aes(x = reorder(Technology, n), y = n),
           stat = 'identity',
           width = 0.7,
           fill = intersect_color1) +
  # scale_fill_viridis(discrete = TRUE, option="viridis") +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Tool/Technology",
       y = "Number of Responses") +
  coord_flip()
```

```
# scale_y_continuous(
#   limits = c(0, 6000),
#   expand = c(0.05, 0.05),
#   breaks = seq(0, 6000, 1000),
#   minor_breaks = seq(0, 10000, 500)
# ) +
# scale_x_continuous(
#   breaks = seq(2016, 2021, 1)
# )
```

Kirkpatrick's metrics

```
worthwhile_plot <- Only_Responses %>%
  filter(reaction_corrected != "#N/A") %>%
  group_by(reaction_corrected) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(reaction_corrected)) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(reaction_corrected, ": ", Frequency)) %>%
  ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=reaction_corrected)) +
    geom_rect() +
    geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
```

```r
    scale_fill_brewer(palette=4) +
    coord_polar(theta="y") +
    xlim(c(1.5, 4)) +
    theme_void() +
    theme(legend.position = "none")

frequency_plot <- Only_Responses %>%
  filter(behaviour_corrected != "#N/A") %>%
  group_by(behaviour_corrected) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(behaviour_corrected)) %>%
  top_n(10) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(behaviour_corrected, ": ", Frequency)) %>%
  ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=behaviour_corrected)) +
    geom_rect() +
    geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    scale_fill_brewer(palette=4) +
    coord_polar(theta="y") +
    xlim(c(1.5, 4)) +
    theme_void() +
    theme(legend.position = "none")

apply_plot <- Only_Responses %>%
  filter(learning_corrected != "#N/A") %>%
  group_by(learning_corrected) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(learning_corrected)) %>%
  top_n(10) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(learning_corrected, ": ", Frequency)) %>%
  ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=learning_corrected)) +
    geom_rect() +
    geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    scale_fill_brewer(palette=4) +
    coord_polar(theta="y") +
    xlim(c(1.5, 4)) +
    theme_void() +
    theme(legend.position = "none")

helpful_plot <- Only_Responses %>%
  filter(results_corrected != "#N/A") %>%
  group_by(results_corrected) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(results_corrected)) %>%
```

```
  top_n(10) %>%
mutate(ymax = cumsum(Frequency),
       ymin = c(0, head(ymax, n=-1)),
       labelposition = (ymax+ymin)/2,
       label = paste0(results_corrected, ": ", Frequency)) %>%
ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=results_corrected)) +
  geom_rect() +
  geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  scale_fill_brewer(palette=4) +
  coord_polar(theta="y") +
  xlim(c(2, 4)) +
  theme_void() +
  theme(legend.position = "none")
```

Approximately 79% of respondents indicated that the training course they attended was either "very worthwhile" or "extremely worthwhile. This reflects that the majority of attendees considered attending an Intersect course to be a valuable use of their time.

Approximately 80% of respondents have continued to use the research tool they learnt about during their training course at least "sometimes", with ~52% of attendees using the tool "often" or "frequently" following the course. This suggests that attendees feel comfortable using the tool and applying what they have learned during the training course subsequently in their research.

Over 86% of survey respondents indicated that they feel more or "much more" confident using the research tool than they did prior to attending the training course. This indicates good understanding of the content in a way that is applicable to those researcher's use cases, making them feel comfortable to apply the concepts and skills gained in the training course to their own research.
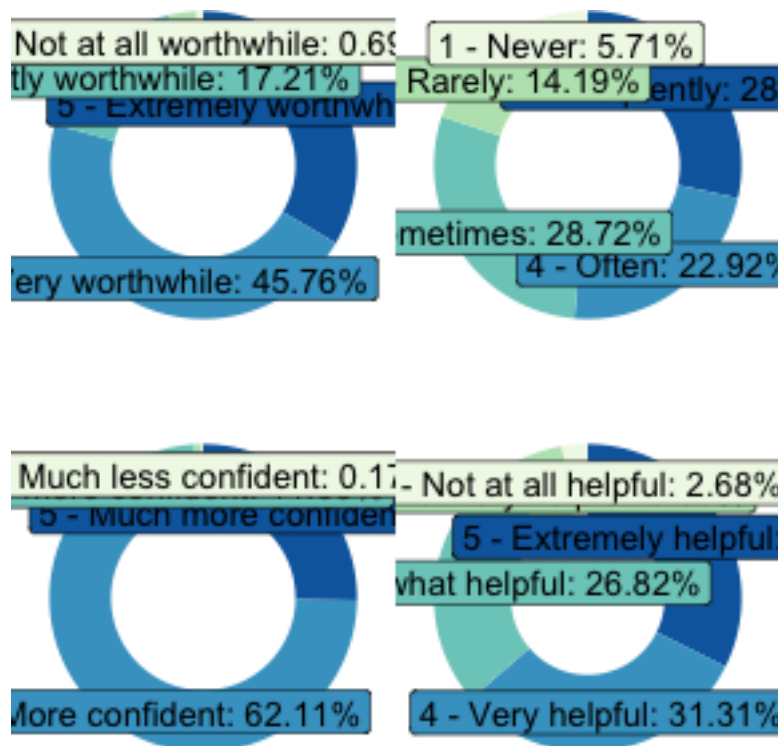
Nearly 63% of survey respondents indicated that the training course they attended was "very" or "extremely" helpful in producing results of their analysis in subsequent research.

```
# grid plot to include all four plots, e.g. worthwhile, frequency, apply, helpful

grid.arrange(worthwhile_plot, frequency_plot, apply_plot, helpful_plot,
             ncol=2, nrow=2)
```

Not at all worthwhile: 0.6% 1 - Never: 5.71%
tly worthwhile: 17.21% Rarely: 14.19% ently: 28
5 - Extremely worthwh

metimes: 28.72%
4 - Often: 22.92%

ery worthwhile: 45.76%

Much less confident: 0.1 - Not at all helpful: 2.68%
5 - Much more confiden 5 - Extremely helpful
vhat helpful: 26.82%

More confident: 62.11% 4 - Very helpful: 31.31%

Reaction: Worthwhile

Interestingly, more than 35% of Post-doc attendees found their training course to be "extremely worthwhile" compared to other roles (~30%). Apart from Masters attendees, at least 75% of attendees from all roles found their training course to be either "very" or "extremely" worthwhile.

```r
# Only_Responses$reaction_corrected <- as.factor(Only_Responses$reaction_corrected)

Only_Responses %>%
  group_by(reaction_corrected, Role) %>%
  count() %>%
  drop_na(reaction_corrected) %>%
  arrange(desc(reaction_corrected)) %>%
  mutate(Role = factor(Role, levels=Role)) %>%
  mutate(reaction_corrected = factor(reaction_corrected, levels=reaction_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = Role, y = n, fill = reaction_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  # scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
  scale_fill_brewer(palette=4, name = "Worthwhile") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
```

```r
  labs(x = "Role/Position",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

Professional (research-related)-

Post-doc / Fellow-   **Worthwhile**

                                    ☐ 1 - Not at
                                    ☐ 2 - Not ve
gree Research Student (Masters)-   ☐ 3 - Slightl
                                    ☐ 4 - Very v
                                    ■ 5 - Extren

Degree Research Student (PhD)-

Academic-

TﬥiⅠⅡ%
Percentage of Responses

Over 75% of attendees from the Faculty of Medicine and Health found their training course to be "very" or "extremely" worthwhile. This is notable as nearly half of all survey respondents were from this faculty. Interestingly, just under 50% of attendees from the Faculty of Architecture, Design and Planning found their course to be "extremely worthwhile", this proportion being more than any other faculty.

```r
# Only_Responses$reaction_corrected <- as.factor(Only_Responses$reaction_corrected)

Only_Responses %>%
  group_by(reaction_corrected, faculty) %>%
  count() %>%
  drop_na(faculty) %>%
  drop_na(reaction_corrected) %>%
  arrange(desc(reaction_corrected)) %>%
  mutate(faculty = factor(faculty, levels=faculty)) %>%
  mutate(reaction_corrected = factor(reaction_corrected, levels=reaction_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = faculty, y = n, fill = reaction_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name="Worthwhile") +
  # scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
```

```
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Faculty",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
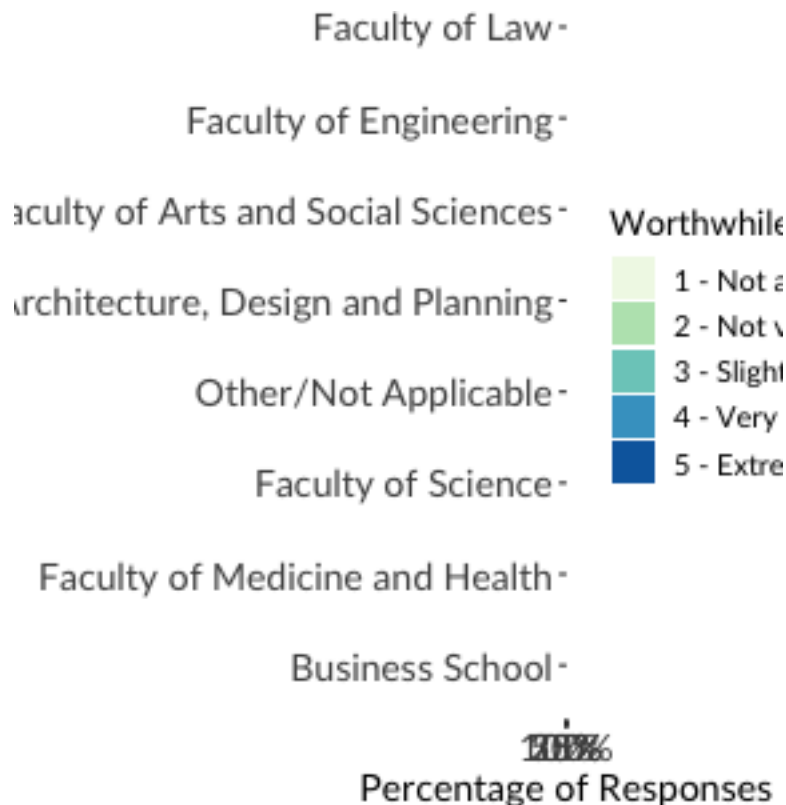

According to survey respondents, the courses that were most worthwhile (i.e. "very" or "extremely worthwhile") were REDCap (~85%), Excel (~85%), R (~80%), Qualtrics (~80%) and SPSS (~80%). Note that the sample of respondents who attended the Tableau training course was very small, hence it is difficult to infer conclusions from their responses.

```
# Only_Responses$reaction_corrected <- as.factor(Only_Responses$reaction_corrected)

Only_Responses %>%
  group_by(reaction_corrected, Technology) %>%
  count() %>%
  drop_na(Technology) %>%
  drop_na(reaction_corrected) %>%
  arrange(desc(reaction_corrected)) %>%
  mutate(Technology = factor(Technology, levels=Technology)) %>%
  mutate(reaction_corrected = factor(reaction_corrected, levels=reaction_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = Technology, y = n, fill = reaction_corrected),
```

```
        stat = 'identity',
        position = "fill",
        width = 0.7) +
scale_fill_brewer(palette=4, name = "Worthwhile") +
# scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Tool/Technology",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
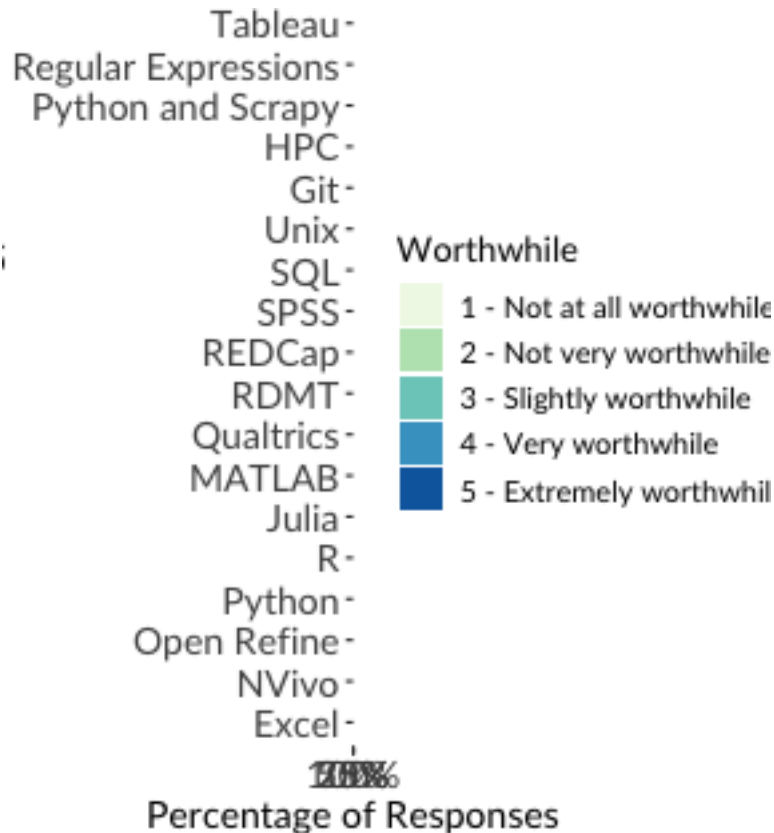

Behaviour: Frequency

Post doc attendees reported frequent use of the research technology of the course they attended in over 30% of cases, while over 60% of Professionals reported "frequent" or "often" use of their research technology. This is aligned with the fact that Post doc attendees found attendance at their training course to be "extremely worthwhile" in ~35% of cases, more than any other role.

```r
Only_Responses %>%
  group_by(behaviour_corrected, Role) %>%
  count() %>%
  drop_na(behaviour_corrected) %>%
  arrange(desc(behaviour_corrected)) %>%
  mutate(Role = factor(Role, levels=Role)) %>%
  mutate(behaviour_corrected = factor(behaviour_corrected, levels=behaviour_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = Role, y = n, fill = behaviour_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name = "Frequency") +
  # scale_fill_viridis(name="Frequency",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Role/Position",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

Attendees from the Faculty of Law and Faculty of Architecture, Design and Planning reported "frequent" use of the research technology from their training course in nearly 40% of cases. This suggests that they found the course content to be applicable to their work or studies and they felt comfortable applying this content to their specific projects.

```r
Only_Responses %>%
  group_by(behaviour_corrected, faculty) %>%
  count() %>%
  drop_na(faculty) %>%
  drop_na(behaviour_corrected) %>%
  arrange(desc(behaviour_corrected)) %>%
  mutate(faculty = factor(faculty, levels=faculty)) %>%
  mutate(behaviour_corrected = factor(behaviour_corrected, levels=behaviour_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = faculty, y = n, fill = behaviour_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name="Frequency") +
  # scale_fill_viridis(name="Frequency",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Faculty",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

of Architecture, Design and Planning -

Other/Not Applicable -

Faculty of Science -

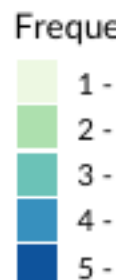Faculty of Medicine and Health -

Faculty of Law -

Faculty of Engineering -

Faculty of Arts and Social Sciences -

Business School -

Freque

1 -
2 -
3 -
4 -
5 -

Percentage of Respor

Over 35% of Excel and REDCap training course attendees reported "frequently" using this technology following the course, and nearly 75% of Excel course attendees reported they used the technology "frequently" or "often". This highlights the applicability of the Excel for Researchers course and the direct relevance of the course content to a variety of research projects and use cases.

```r
Only_Responses %>%
  group_by(behaviour_corrected, Technology) %>%
  count() %>%
  drop_na(Technology) %>%
  drop_na(behaviour_corrected) %>%
  arrange(desc(behaviour_corrected)) %>%
  mutate(Technology = factor(Technology, levels=Technology)) %>%
  mutate(behaviour_corrected = factor(behaviour_corrected, levels=behaviour_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = Technology, y = n, fill = behaviour_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name="Frequency") +
  # scale_fill_viridis(name="Frequency",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
```

53

```
) +
labs(x = "Tool/Technology",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
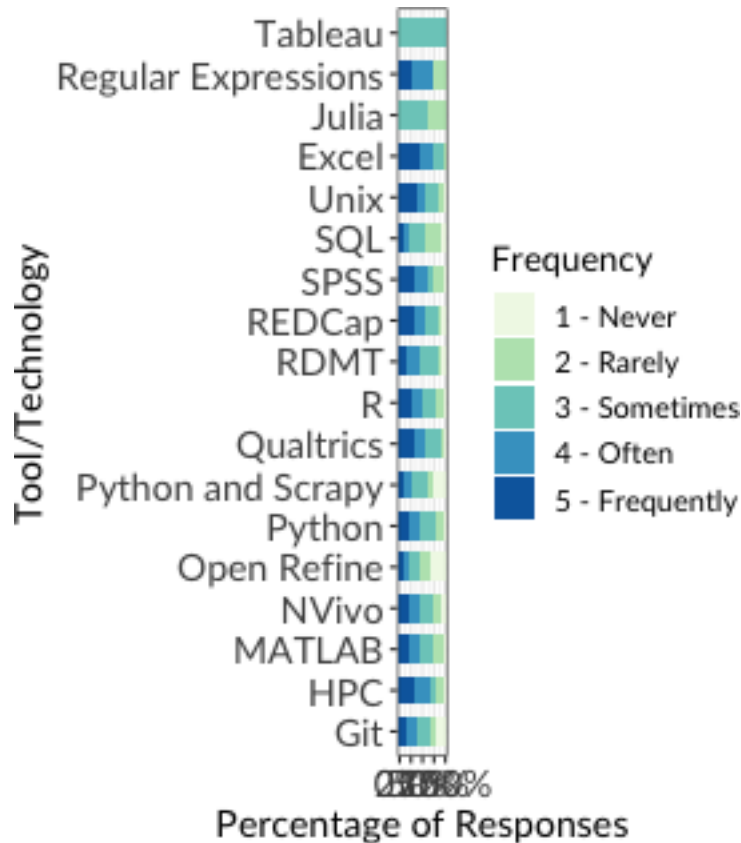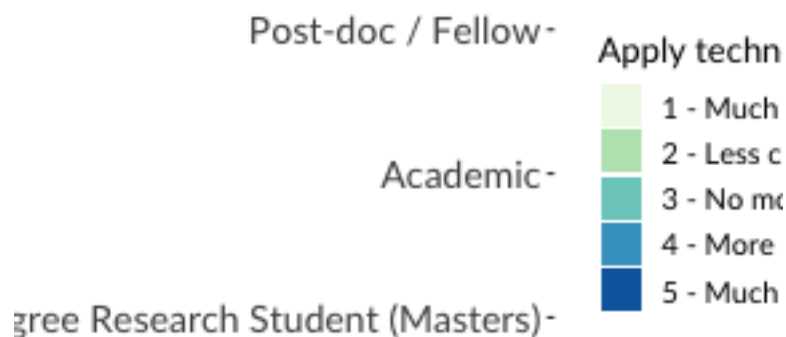

Learning: Apply technology

More than 90% of Professional attendees reported that they felt more confident using the technology after attending their training course. This was the case for more than 85% of all other attendees (Post doc, HDRs and Academics). This emphasises the utility of formative assessments used by instructors during the course for attendees to test their knowledge and application of the content so that they feel comfortable using the technology outside of the training environment.

```
Only_Responses %>%
  group_by(learning_corrected, Role) %>%
  count() %>%
  drop_na(learning_corrected) %>%
  arrange(desc(learning_corrected)) %>%
  mutate(Role = factor(Role, levels=Role)) %>%
  mutate(learning_corrected = factor(learning_corrected, levels=learning_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = Role, y = n, fill = learning_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name="Apply technology") +
```

```
# scale_fill_viridis(name="Apply technology",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Role/Position",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```

Professional (research-related) -

Post-doc / Fellow -

Apply techn

1 - Much

2 - Less c

Academic -

3 - No mo

4 - More

5 - Much

gree Research Student (Masters) -

Degree Research Student (PhD) -

100%

Percentage of Responses

Faculty of Medicine and Health and Faculty of Engineering attendees reported the most confidence in applying the content of their training course to their use cases following training (~35%). Close to 100% of Faculty of Engineering attendees reported being more confident in applying their knowledge subsequent to the training course. Attendees from all faculties reported being more confident in at least 75% of cases.

```
Only_Responses %>%
  group_by(learning_corrected, faculty) %>%
  count() %>%
  drop_na(faculty) %>%
  drop_na(learning_corrected) %>%
  arrange(desc(learning_corrected)) %>%
  mutate(faculty = factor(faculty, levels=faculty)) %>%
  mutate(learning_corrected = factor(learning_corrected, levels=learning_corrected)) %>%
```

```r
ggplot() +
geom_bar(aes(x = faculty, y = n, fill = learning_corrected),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_brewer(palette=4, name="Apply technology") +
# scale_fill_viridis(name="Apply technology",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Faculty",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```


Nearly 50% of attendees of REDCap and Qualtrics training courses reported to be "much more confident" in using their chosen technology subsequent to training. At least 65% of attendees reported to be more confident across all technologies taught in the training courses.

```r
Only_Responses %>%
  group_by(learning_corrected, Technology) %>%
```

```r
count() %>%
drop_na(Technology) %>%
drop_na(learning_corrected) %>%
arrange(desc(learning_corrected)) %>%
mutate(Technology = factor(Technology, levels=Technology)) %>%
mutate(learning_corrected = factor(learning_corrected, levels=learning_corrected)) %>%
ggplot() +
geom_bar(aes(x = Technology, y = n, fill = learning_corrected),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_brewer(palette=4, name="Apply technology") +
# scale_fill_viridis(name="Apply technology",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Tool/Technology",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
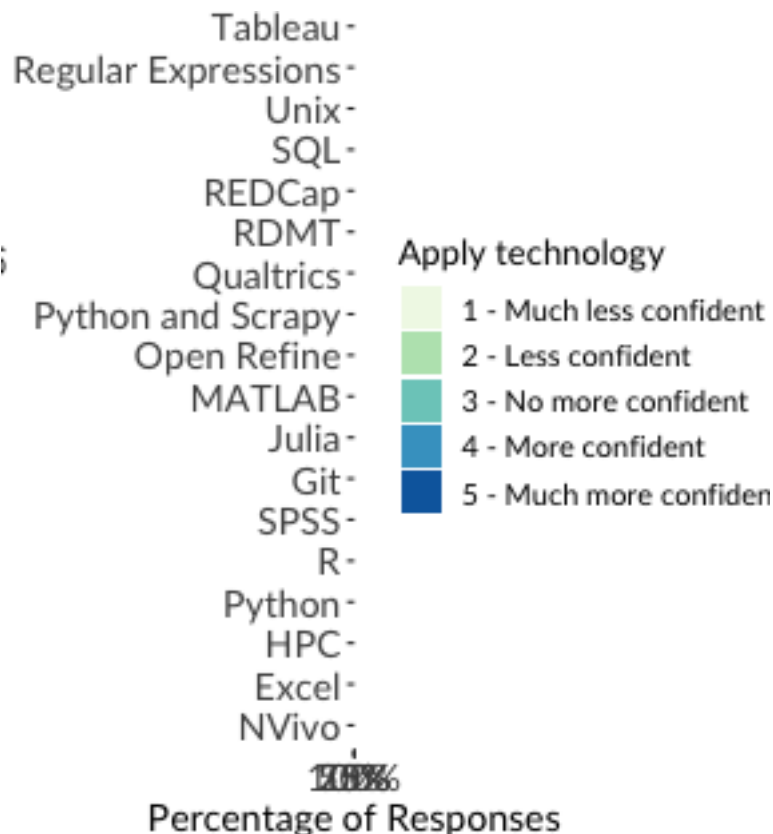
Results: Helpful

Professional attendees reported the training course to be "very" or "extremely" helpful in ~90% of cases. Over 65% of attendees across all roles reported the training course to be at least "very" helpful. This suggests the courses are a valuable investment of attendee time and assist them in their research projects following training.

```r
Only_Responses %>%
  group_by(results_corrected, Role) %>%
  count() %>%
  drop_na(results_corrected) %>%
  arrange(desc(results_corrected)) %>%
  mutate(Role = factor(Role, levels=Role)) %>%
  mutate(results_corrected = factor(results_corrected, levels=results_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = Role, y = n, fill = results_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name="Helpful") +
  # scale_fill_viridis(name="Helpful",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Role/Position",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

Professional (research-related) -

Post-doc / Fellow -

Helpful

1 - Not
2 - Not
3 - Som
4 - Very
5 - Extr

r Degree Research Student (PhD) -

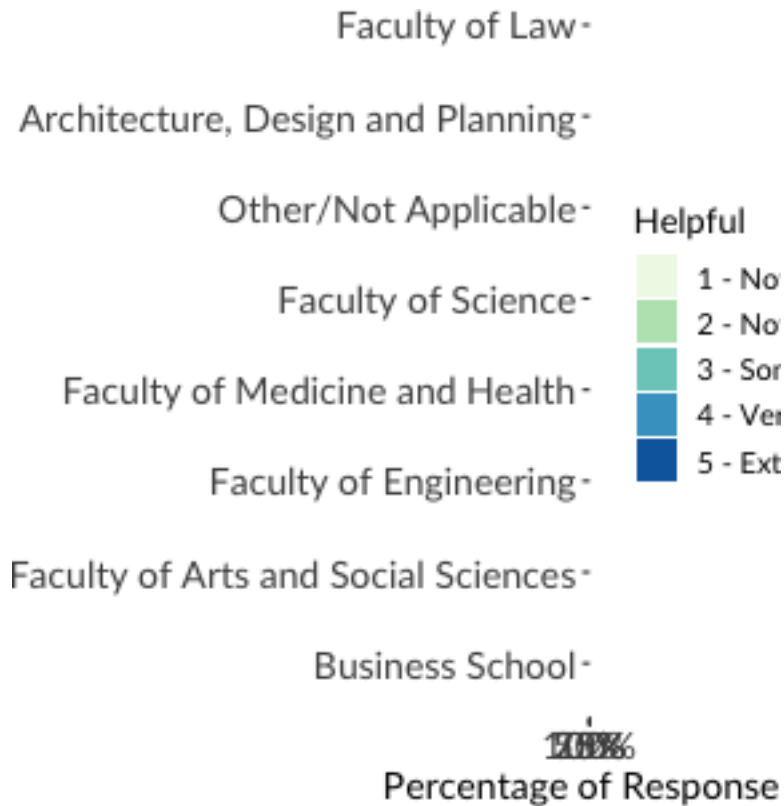egree Research Student (Masters) -

Academic -

10.0% 50%

Percentage of Responses

With the exception of the Faculty of Arts and Social Sciences, attendees from all faculties reported their training course to be "very" or "extremely" helpful in the subsequent research projects following training. Close to 90% of attendees stated that the training course they attended was at least "somewhat helpful".

```r
Only_Responses %>%
  group_by(results_corrected, faculty) %>%
  count() %>%
  drop_na(faculty) %>%
  drop_na(results_corrected) %>%
  arrange(desc(results_corrected)) %>%
  mutate(faculty = factor(faculty, levels=faculty)) %>%
  mutate(results_corrected = factor(results_corrected, levels=results_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = faculty, y = n, fill = results_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4,name="Helpful") +
  # scale_fill_viridis(name="Helpful",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Faculty",
```

```
      y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
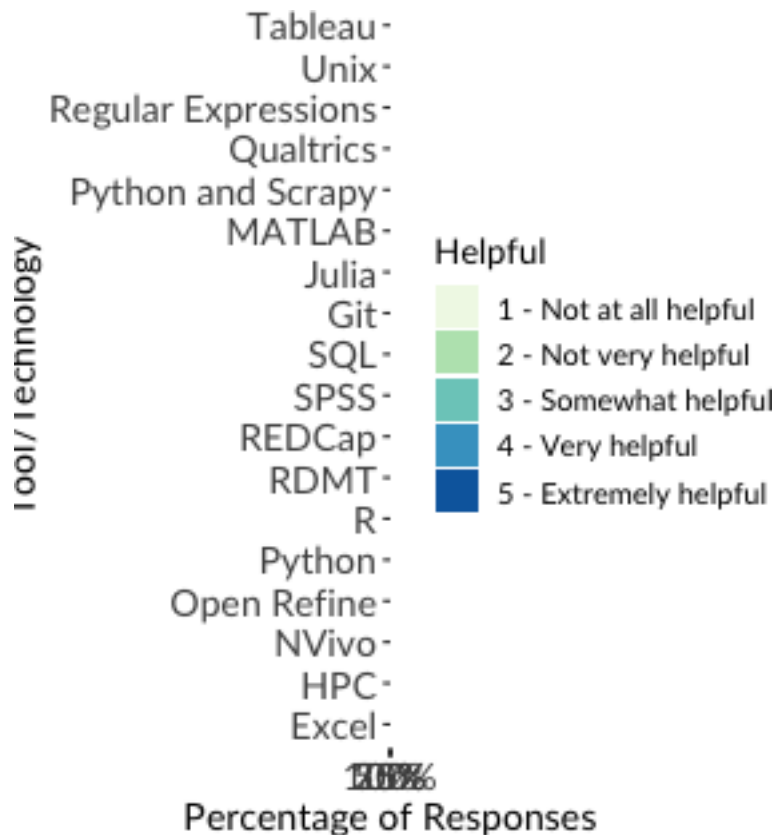

Nearly 80% of Excel, REDCap and Qualtrics training course attendees reported their course to be "very" or "extremely" helpful. This reflects the ubiquity of these research tools across a wide range of research fields, but also the applicability of content covered in these courses to equip attendees from diverse areas of expertise with skills for their specific research projects.

```
Only_Responses %>%
  group_by(results_corrected, Technology) %>%
  count() %>%
  drop_na(Technology) %>%
  drop_na(results_corrected) %>%
  arrange(desc(results_corrected)) %>%
  mutate(Technology = factor(Technology, levels=Technology)) %>%
  mutate(results_corrected = factor(results_corrected, levels=results_corrected)) %>%
  ggplot() +
  geom_bar(aes(x = Technology, y = n, fill = results_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4,name="Helpful") +
  # scale_fill_viridis(name="Helpful",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
```

```
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Tool/Technology",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```


Attend if you had access to the material

Nearly 83% of attendees stated that they would at least probably still attend the course if they were provided with the course material prior to training. This highlights the value of the instructor's explanations of concepts and the formative assessments throughout the course that are not included in the course notes.

```
Only_Responses %>%
  filter(materials_only_corrected != "#N/A") %>%
  group_by(materials_only_corrected) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(materials_only_corrected)) %>%
  top_n(10) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
```

```
        label = paste0(materials_only_corrected, ": ", Frequency)) %>%
ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=materials_only_corrected)) +
  geom_rect() +
  geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
  scale_fill_brewer(palette=4) +
  coord_polar(theta="y") +
  xlim(c(1.5, 4)) +
  theme_void() +
  theme(legend.position = "none")
```
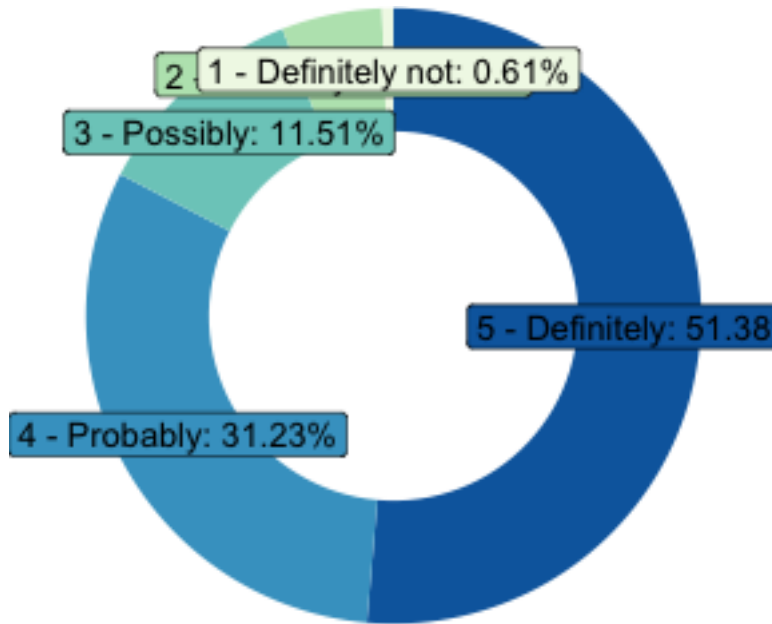

More than 50% of Post doc/fellow attendees stated that they would definitely still attend the course if they were provided with the course material prior to training. This suggests that attendees in that role highly valued the interactivity of the course, the examples and the formative assessments used by instructors that are not included in the course material. Apart from Masters student attendees, over 80% of all attendees indicated that they would at least "probably" attend the course if they had the course materials.

```
Only_Responses %>%
  group_by(materials_only_corrected, Role) %>%
  count() %>%
  drop_na(materials_only_corrected) %>%
  arrange(desc(materials_only_corrected)) %>%
  mutate(Role = factor(Role, levels=Role)) %>%
  mutate(materials_only_corrected = factor(materials_only_corrected, levels=materials_only_corrected)) %
  ggplot() +
  geom_bar(aes(x = Role, y = n, fill = materials_only_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  # scale_fill_viridis(name="Attend with open material",discrete = TRUE, option="viridis") +
  scale_fill_brewer(palette=4, name=paste("Attend if people had\n access to material")) +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
```

```r
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Role/Position",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

Professional (research-related) -

Post-doc / Fellow -

Attend if
access to

1 - De

Degree Research Student (Masters) -

2 - Pro

3 - Pos

4 - Pro

er Degree Research Student (PhD) -

5 - De

Academic -

Percentage of Response

Faculty of Arts and Social Sciences and Faculty of Business attendees reported in over 75% of cases that they would attend the training course even if they were provided with the course material. This suggests that attendees from those faculties obtained value from attending the course itself and the extra learning and interactivity provided by the instructor(s) that cannot be delivered via the course materials.
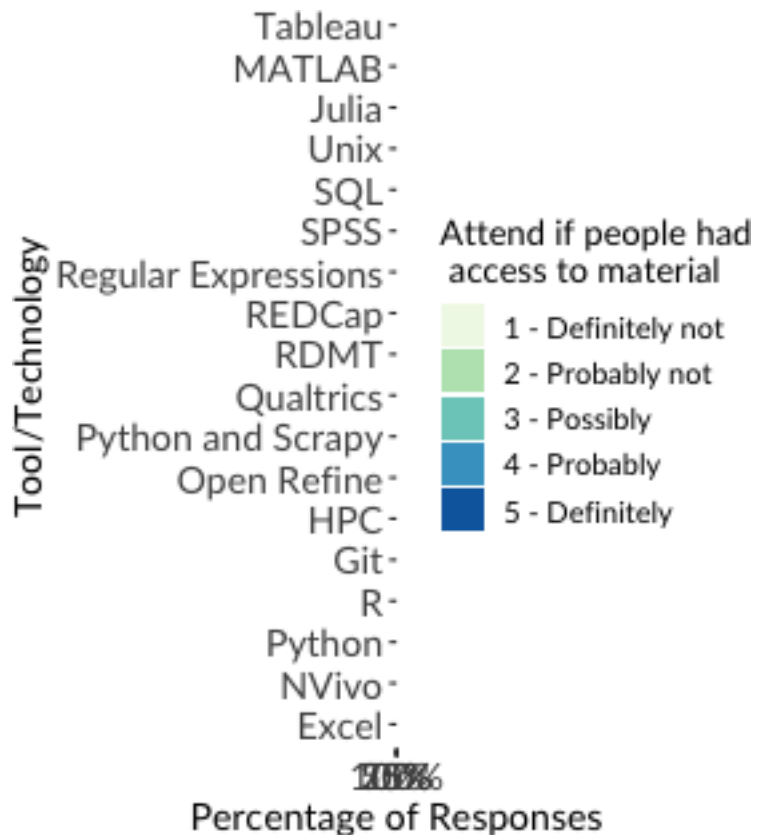
```r
Only_Responses %>%
  group_by(materials_only_corrected, faculty) %>%
  count() %>%
  drop_na(faculty) %>%
  drop_na(materials_only_corrected) %>%
  arrange(desc(materials_only_corrected)) %>%
  mutate(faculty = factor(faculty, levels=faculty)) %>%
  mutate(materials_only_corrected = factor(materials_only_corrected, levels=materials_only_corrected)) %
  ggplot() +
  geom_bar(aes(x = faculty, y = n, fill = materials_only_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
```

```
# scale_fill_viridis(name="Attend with open material",discrete = TRUE, option="viridis") +
scale_fill_brewer(palette=4,name=paste("Attend if people had\n access to material")) +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Faculty",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```


In nearly all technologies, attendees said that they would at least "probably" still attend the course even if they had access to the course materials in 60% of more of cases. In particular, attendees of SPSS, R, Python, NVIVO, HPC and MATLAB training reported that they would at least "probably " still attend training in more than 80% of cases.

```
Only_Responses %>%
  group_by(materials_only_corrected, Technology) %>%
  count() %>%
  drop_na(Technology) %>%
  drop_na(materials_only_corrected) %>%
  arrange(desc(materials_only_corrected)) %>%
```

```
mutate(Technology = factor(Technology, levels=Technology)) %>%
mutate(materials_only_corrected = factor(materials_only_corrected, levels=materials_only_corrected))
ggplot() +
geom_bar(aes(x = Technology, y = n, fill = materials_only_corrected),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
# scale_fill_viridis(name="Attend with open material",discrete = TRUE, option="viridis") +
scale_fill_brewer(palette=4,name=paste("Attend if people had\n access to material")) +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Tool/Technology",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```

Post-training support

Intersect provided ~35% of post-training support (via extra training or eRA consultation). Peer Support accounted for close to a third of the post-training support that attendees used.

```
Support_checkbox_long %>%
  group_by(Support_Type) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(Support_Type)) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(Support_Type, ": ", Frequency)) %>%
  ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=Support_Type)) +
    geom_rect() +
    geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    scale_fill_brewer(palette=4) +
    coord_polar(theta="y") +
    xlim(c(1.5, 4)) +
    theme_void() +
    theme(legend.position = "none")
```
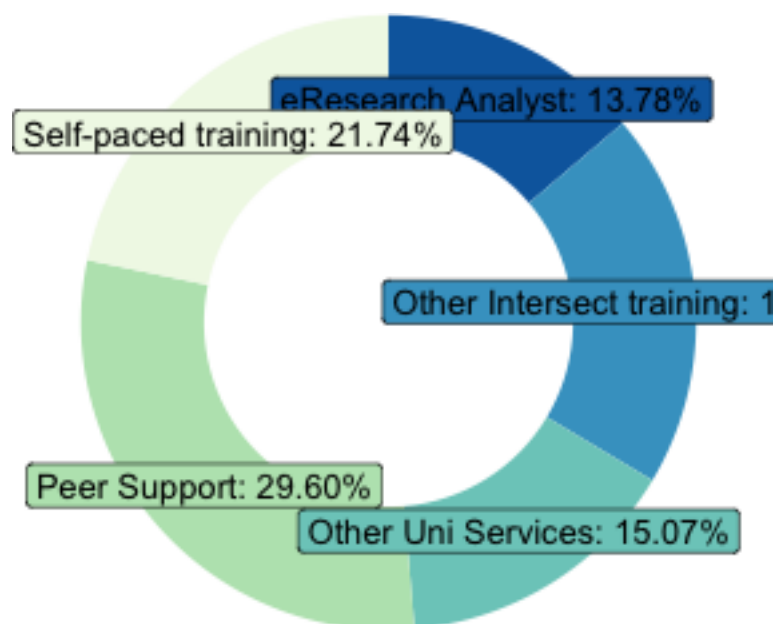

Professionals used eRA support extensively (in ~30% of cases) which was the highest of all role categories. Support provided by Other University Services was reported in ~10% or less of cases across all roles, suggesting an absence of such support at universities beyond the services provided by Intersect.

```
Support_checkbox_long %>%
  group_by(Support_Type, Role) %>%
  count() %>%
  drop_na(Support_Type) %>%
  arrange(desc(Support_Type)) %>%
  mutate(Role = factor(Role, levels=Role)) %>%
  mutate(Support_Type = factor(Support_Type, levels=Support_Type)) %>%
```

```
ggplot() +
geom_bar(aes(x = Role, y = n, fill = Support_Type),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_brewer(palette=4,name="Post-training support") +
# scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Role/Position",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```

Professional (research-related)-

Post-doc / Fellow-   Post-trainin

                                       Self-pace
                                       Peer Sup
Degree Research Student (PhD)-   Other Ur
                                       Other In
                                       eResearc
gree Research Student (Masters)-

                    Academic-

                    100%
              Percentage of Responses

Peer Support or Self-paced training represented ~50% of post-training support reported by attendees. eRA support was used by attendees from the Faculty of Arts and Social Sciences the most (~20%).
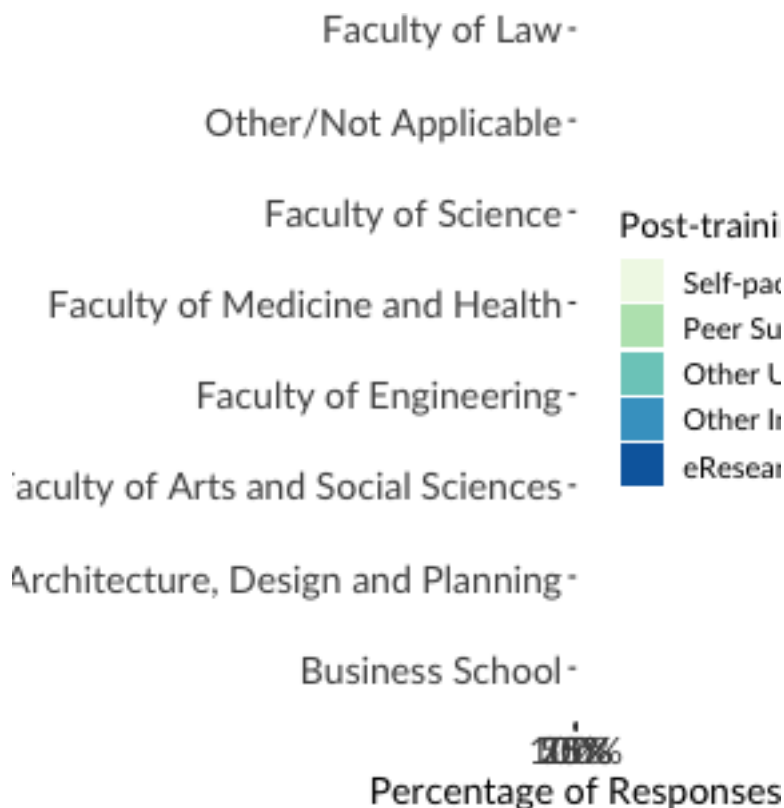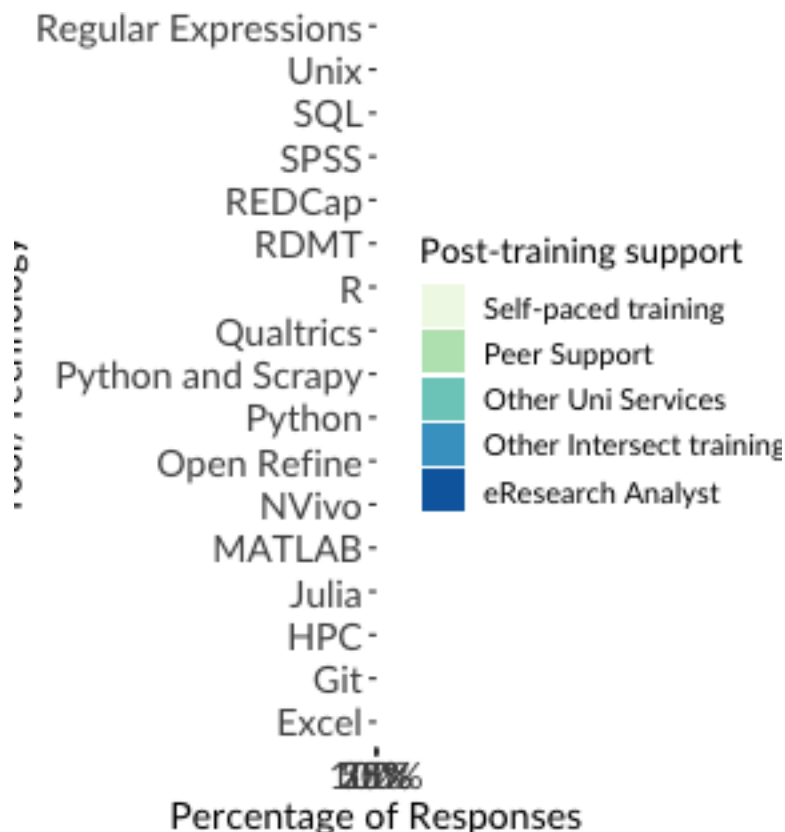
```
Support_checkbox_long %>%
  drop_na(faculty) %>%
  group_by(Support_Type, faculty) %>%
  count() %>%
```

```
drop_na(Support_Type) %>%
arrange(desc(Support_Type)) %>%
mutate(faculty = factor(faculty, levels=faculty)) %>%
mutate(Support_Type = factor(Support_Type, levels=Support_Type)) %>%
ggplot() +
geom_bar(aes(x = faculty, y = n, fill = Support_Type),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_brewer(palette=4,name="Post-training support") +
# scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Faculty",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```

Faculty of Law -

Other/Not Applicable -

Faculty of Science -   Post-traini

Faculty of Medicine and Health -   Self-pac
                                    Peer Su
Faculty of Engineering -   Other U
                                    Other I
aculty of Arts and Social Sciences -   eResea

Architecture, Design and Planning -

Business School -

100%

Percentage of Responses

Attendees of REDCap courses used eRA support the most of any technology (close to 50% of cases). This
was similar for NVivo (~40%). Interestingly, less than 10% of attendees of R and Python courses reported

68

receiving support from eRAs, instead predominantly supporting themselves through self-paced learning and peer support.

```r
Support_checkbox_long %>%
  drop_na(Technology) %>%
  group_by(Support_Type, Technology) %>%
  count() %>%
  drop_na(Support_Type) %>%
  arrange(desc(Support_Type)) %>%
  mutate(Technology = factor(Technology, levels=Technology)) %>%
  mutate(Support_Type = factor(Support_Type, levels=Support_Type)) %>%
  ggplot() +
  geom_bar(aes(x = Technology, y = n, fill = Support_Type),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4,name="Post-training support") +
  # scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Tool/Technology",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

Research outputs

Close to 40% of attendees reported that their training attendance directly resulted in a Journal Article publication or contributed to their Thesis. It should be noted that the Thesis category only applied to HDR students. Interestingly, very few attendees reported training attendance contributing to a book chapter (less than 5%). Conferences and Presentations made up close to a third of research outputs following training.

```r
# Create a palette from library(rcartocolor)
vivid_pal <- carto_pal(12, "Vivid")
safe_pal <- carto_pal(12, "Safe")
pastel_pal <- carto_pal(12, "Pastel")
prism_pal <- carto_pal(12, "Prism")

ResearchOutputs_checkbox_long %>%
  group_by(`Research Output`) %>%
  summarise(Count = n()) %>%
  mutate(Frequency = formattable::percent(Count/ sum(Count))) %>%
  arrange(desc(Frequency)) %>%
  mutate(ymax = cumsum(Frequency),
         ymin = c(0, head(ymax, n=-1)),
         labelposition = (ymax+ymin)/2,
         label = paste0(`Research Output`, ": ", Frequency)) %>%
  ggplot(aes(ymax=ymax, ymin=ymin, xmax=4, xmin=3, fill=`Research Output`)) +
    geom_rect() +
    geom_label( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    # geom_text( x=3.5, aes(y=labelposition, label=label), size=4.5) +
    # scale_fill_brewer(palette="Paired") + #Set3 palette
```

```r
scale_fill_manual(values = prism_pal) +
# scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
coord_polar(theta="y") +
xlim(c(1.5, 4)) +
theme_void() +
theme(legend.position = "none")
```
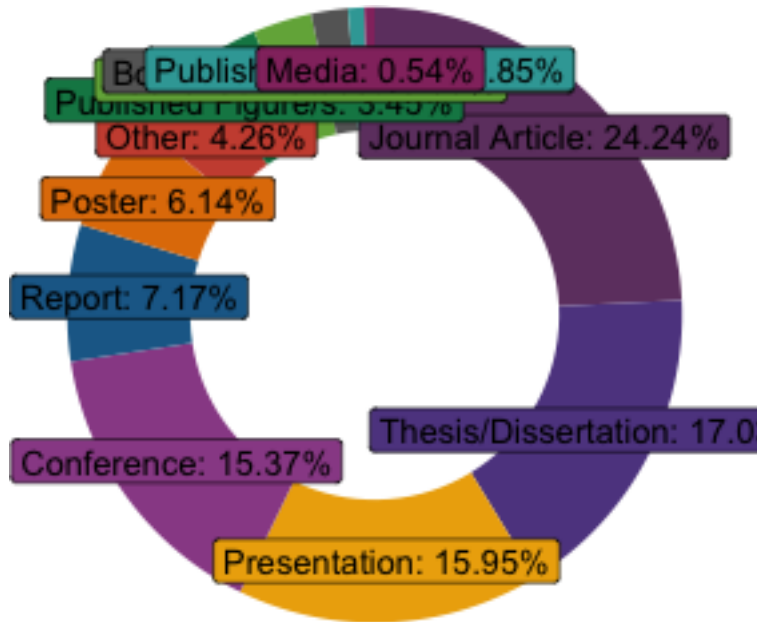

Professionals reported that training contributed to Journal Article publication in ~30% of cases, the most
of any role category. Contribution of training attendance to conference presentations was also close to 20%
across all roles, and when combined with Presentation was close to 40% in all categories.

```r
ResearchOutputs_checkbox_long %>%
  group_by(`Research Output`, Role) %>%
  count() %>%
  drop_na(`Research Output`) %>%
  arrange(desc(`Research Output`),desc(n)) %>%
  mutate(Role = factor(Role, levels=Role)) %>%
  mutate(`Research Output` = factor(`Research Output`, levels=`Research Output`)) %>%
  ggplot() +
  geom_bar(aes(x = Role, y = n, fill = `Research Output`),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
# scale_fill_brewer(palette="Paired",name="Research outputs") +
# scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
  scale_fill_manual(values = prism_pal,name="Research outputs") +
#theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
```
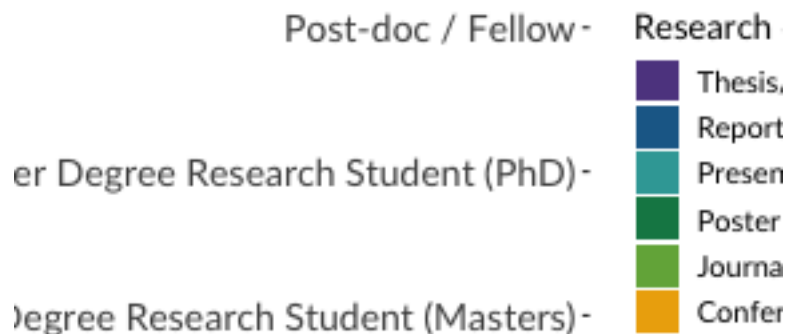
```
  text = element_text(family = "Lato")
) +
labs(x = "Role/Position",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```


Text here

```
ResearchOutputs_checkbox_long %>%
  group_by(`Research Output`, Role) %>%
  count() %>%
  drop_na(`Research Output`) %>%
  arrange(desc(`Research Output`),desc(n)) %>%
  mutate(Role = factor(Role, levels=Role)) %>%
  mutate(`Research Output` = factor(`Research Output`, levels=`Research Output`)) %>%
  filter(`Research Output` %in% c("Thesis/Dissertation","Report","Presentation",
                          "Poster","Journal Article","Conference")) %>%
  ggplot() +
  geom_bar(aes(x = Role, y = n, fill = `Research Output`),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  # scale_fill_brewer(palette="Paired",name="Research outputs") +
  # scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
  scale_fill_manual(values = prism_pal,name="Research outputs") +
  #theme_ipsum() +
  theme_bw() +
  theme(
```

```r
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
labs(x = "Role/Position",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
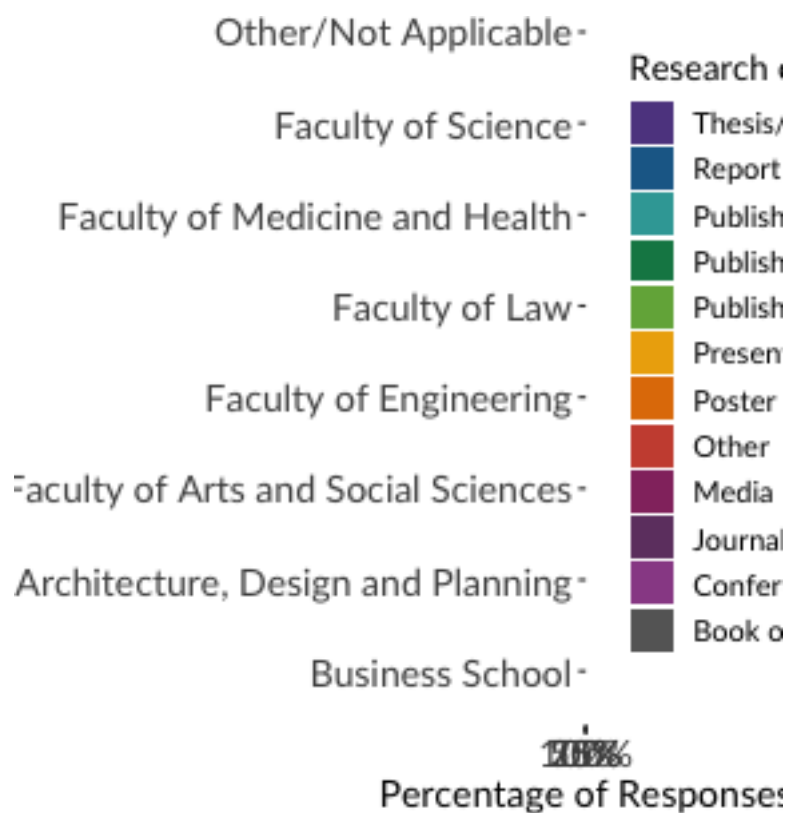
Professional (research-related) -

Post-doc / Fellow -   Research

                                       ▮ Thesis,
                                       ▮ Report
er Degree Research Student (PhD) -   ▮ Presen
                                       ▮ Poster
                                       ▮ Journa
)egree Research Student (Masters) -   ▮ Confer

Academic -

Percentage of Response:

The Faculty of Arts and Social Sciences and Faculty of Business reported contribution of training attendance to Journal Article publications in ~25% of cases, the most across all faculties.
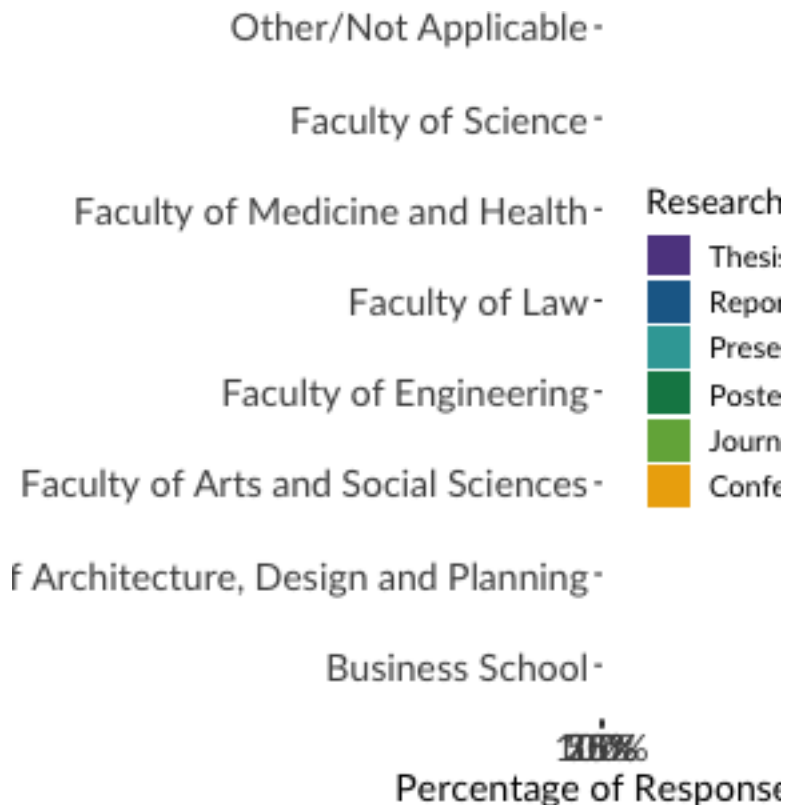
```r
ResearchOutputs_checkbox_long %>%
  drop_na(faculty) %>%
  group_by(`Research Output`, faculty) %>%
  count() %>%
  drop_na(`Research Output`) %>%
  arrange(desc(`Research Output`)) %>%
  mutate(faculty = factor(faculty, levels=faculty)) %>%
  mutate(`Research Output` = factor(`Research Output`, levels=`Research Output`)) %>%
  ggplot() +
  geom_bar(aes(x = faculty, y = n, fill = `Research Output`),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  # scale_fill_brewer(palette="Paired",name="Research outputs") +
  scale_fill_manual(values = prism_pal,name="Research outputs") +
```

```
# scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Faculty",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
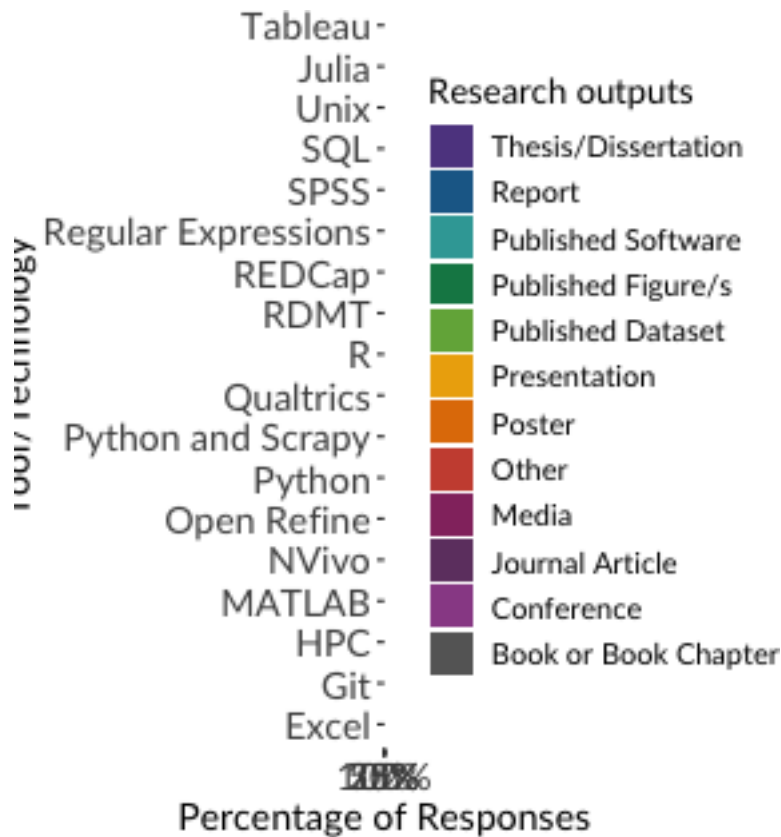

Text here

```
ResearchOutputs_checkbox_long %>%
  drop_na(faculty) %>%
  group_by(`Research Output`, faculty) %>%
  count() %>%
  drop_na(`Research Output`) %>%
  arrange(desc(`Research Output`)) %>%
  mutate(faculty = factor(faculty, levels=faculty)) %>%
  mutate(`Research Output` = factor(`Research Output`, levels=`Research Output`)) %>%
  filter(`Research Output` %in% c("Thesis/Dissertation","Report","Presentation",
                          "Poster","Journal Article","Conference")) %>%
```

```
ggplot() +
geom_bar(aes(x = faculty, y = n, fill = `Research Output`),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
# scale_fill_brewer(palette="Paired",name="Research outputs") +
scale_fill_manual(values = prism_pal,name="Research outputs") +
# scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Faculty",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```

Other/Not Applicable -

Faculty of Science -

Faculty of Medicine and Health -

Faculty of Law -

Faculty of Engineering -

Faculty of Arts and Social Sciences -

f Architecture, Design and Planning -

Business School -

Research

Thesi:
Repo
Prese
Poste
Journ
Confe

Percentage of Response

There was minimal difference between research technologies in terms of research outputs. Attendees of training in most technologies reported contribution to Journal Articles in ~25% of cases, Conferences in ~15% of cases and Presentations in ~15% of cases.
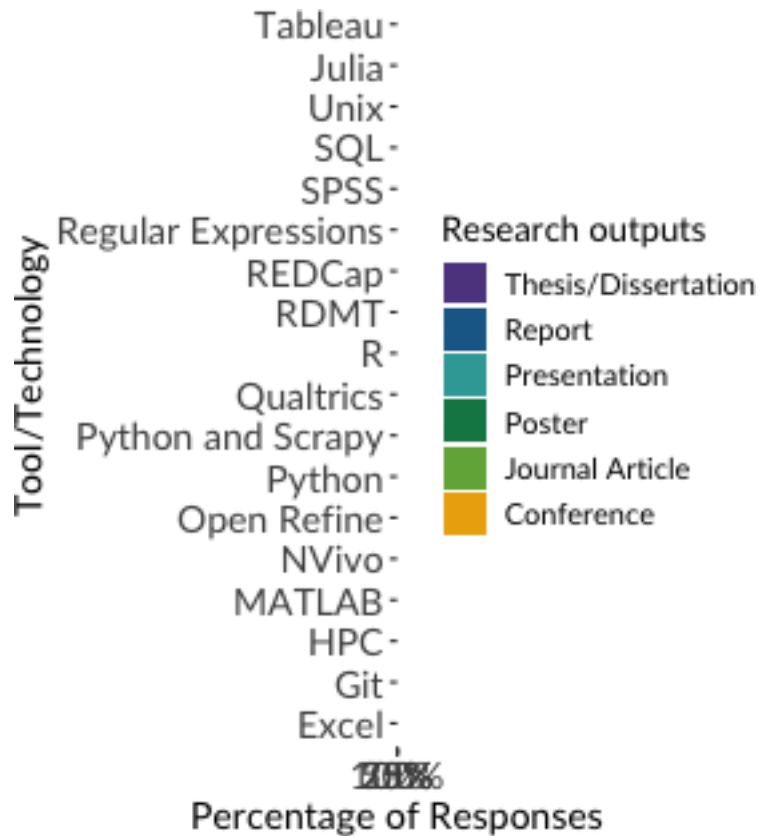
```r
ResearchOutputs_checkbox_long %>%
  drop_na(Technology) %>%
  group_by(`Research Output`, Technology) %>%
  count() %>%
  drop_na(`Research Output`) %>%
  arrange(desc(`Research Output`)) %>%
  mutate(Technology = factor(Technology, levels=Technology)) %>%
  mutate(`Research Output` = factor(`Research Output`, levels=`Research Output`)) %>%
  ggplot() +
  geom_bar(aes(x = Technology, y = n, fill = `Research Output`),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  # scale_fill_brewer(palette="Paired",name="Research outputs") +
  scale_fill_manual(values = prism_pal,name="Research outputs") +
  # scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Tool/Technology",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

Text here

```r
ResearchOutputs_checkbox_long %>%
  drop_na(Technology) %>%
  group_by(`Research Output`, Technology) %>%
  count() %>%
  drop_na(`Research Output`) %>%
  arrange(desc(`Research Output`)) %>%
  mutate(Technology = factor(Technology, levels=Technology)) %>%
  mutate(`Research Output` = factor(`Research Output`, levels=`Research Output`)) %>%
  filter(`Research Output` %in% c("Thesis/Dissertation","Report","Presentation",
                                  "Poster","Journal Article","Conference")) %>%

  ggplot() +
  geom_bar(aes(x = Technology, y = n, fill = `Research Output`),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  # scale_fill_brewer(palette="Paired",name="Research outputs") +
  scale_fill_manual(values = prism_pal,name="Research outputs") +
  # scale_fill_viridis(name="Post-training support",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
```

```
  text = element_text(family = "Lato")
) +
labs(x = "Tool/Technology",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```


Competency

More than 75% of training attendees who attended the introductory courses considered the courses very
worthwhile ("Extremely and Very worthwhile"). Nearly 90% of ther intermediate course attendees and all of
the advanced course attendees found it was very worthwhile taking part in the courses.

```
Only_Responses %>%
  group_by(Level,reaction_corrected) %>%
  count() %>%
  drop_na(Level) %>%
  drop_na(reaction_corrected) %>%
  arrange(desc(reaction_corrected)) %>%
  mutate(Level = factor(Level, levels=Level)) %>%
  ggplot() +
  geom_bar(aes(x = Level, y = n, fill = reaction_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name = "Worthwhile") +
  # scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
```

```
  #theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Course Level",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
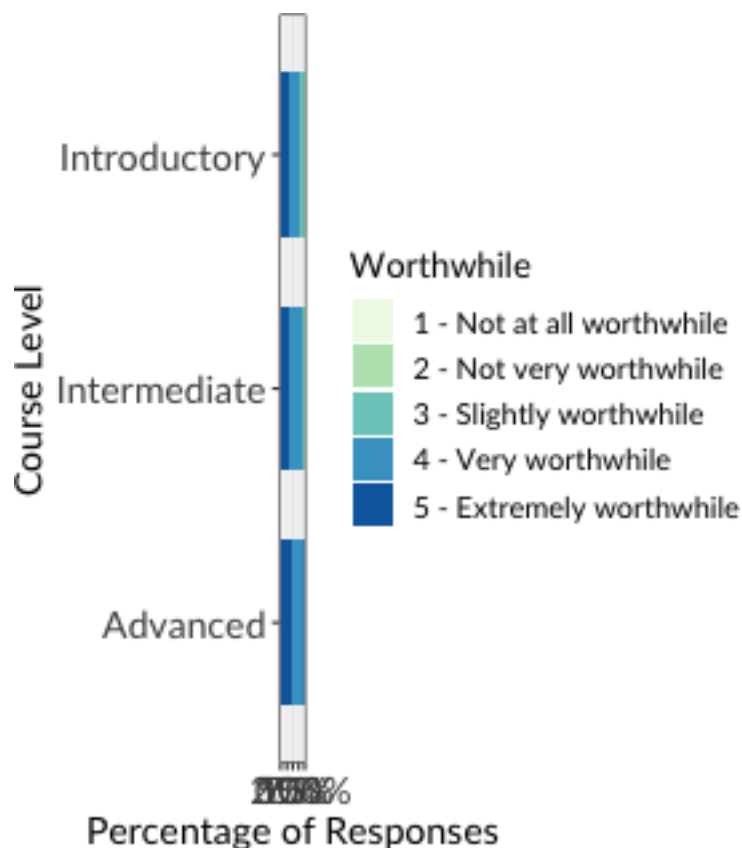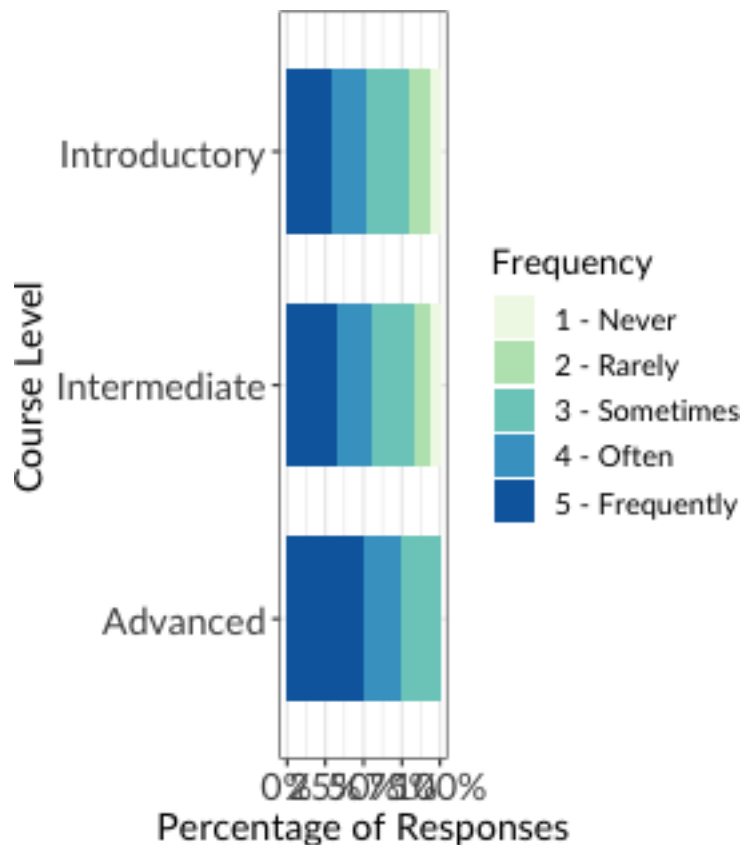

Fig.X shows how frequently the attendees were abile to apply the technology/knowledge since the course. Over a half of the attendees of introductory and intermediate courses and three quarters of the advanced course participants were indicated that they apply the technology learn through the training either often or frequently. Approximately 25% of all course participants claimed that they sometimes applied the technology.

```
Only_Responses %>%
  group_by(Level,behaviour_corrected) %>%
  count() %>%
  drop_na(Level) %>%
  drop_na(behaviour_corrected) %>%
  arrange(desc(behaviour_corrected)) %>%
  mutate(Level = factor(Level, levels=Level)) %>%
```

```
ggplot() +
geom_bar(aes(x = Level, y = n, fill = behaviour_corrected),
          stat = 'identity',
          position = "fill",
          width = 0.7) +
scale_fill_brewer(palette=4, name = "Frequency") +
# scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Course Level",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
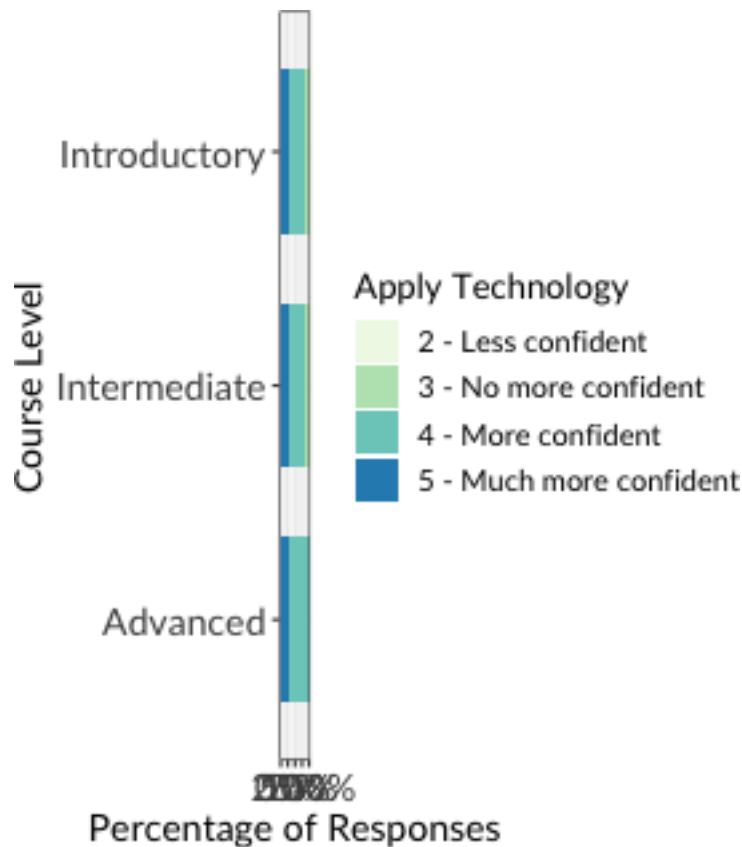

A great majority of the the course attendees expressed that they were more confident applying the technology learnt during the course, 90% of the participants who attended the introductory and intermediate courses, and 100% of who attended the advanced courses.

```r
Only_Responses %>%
  group_by(Level,learning_corrected) %>%
  count() %>%
  drop_na(Level) %>%
  drop_na(learning_corrected) %>%
  arrange(desc(learning_corrected)) %>%
  mutate(Level = factor(Level, levels=Level)) %>%
  ggplot() +
  geom_bar(aes(x = Level, y = n, fill = learning_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name = "Apply Technology") +
  # scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Course Level",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```

Fig. X shows the responses on whether the training courses were helpful to the research by course level. Over 90% of the participants of introductory and intermediate courses found the training helpful with 30% - 40% indicating the courses were extremely helpful. In terms of the advanced courses, all the participants said the courses were either "Very helpful" or "Extremely helpful".

```
Only_Responses %>%
  group_by(Level,results_corrected) %>%
  count() %>%
  drop_na(Level) %>%
  drop_na(results_corrected) %>%
  arrange(desc(results_corrected)) %>%
  mutate(Level = factor(Level, levels=Level)) %>%
  ggplot() +
  geom_bar(aes(x = Level, y = n, fill = results_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name = "Helpful") +
  # scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
```

```
    text = element_text(family = "Lato")
) +
labs(x = "Course Level",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
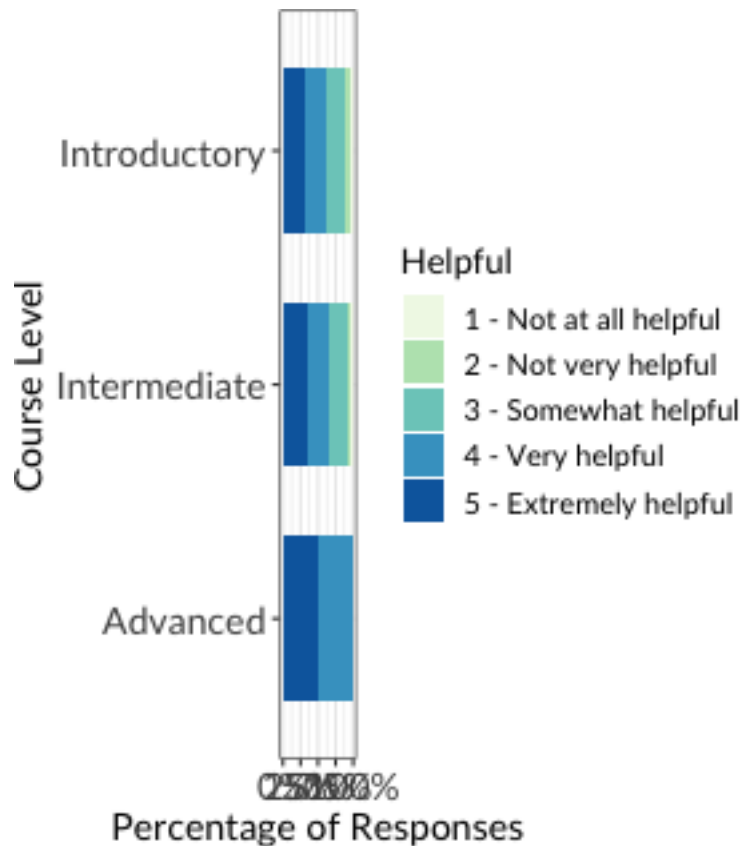

Fig.X shows the type of the post-training support that the attendees have obtained by course level. The distribution is similar for the introductory and intermediate courses. About a quarter of the course attendees conducted self-paced training, with another quarter sought support from the peers. Among the rest of participants, 10% - 15% received support from the Intersect eResearch Analyst, 20% attended other Intersect training courses. As for the advanced course attendees, a half have used other services available at the university and another half have taken the advantage of the Intersect services (other training and support from the eResearch Analyst).

```
Support_checkbox_long %>%
  group_by(Level,Support_Type) %>%
  count() %>%
  drop_na(Level) %>%
  drop_na(Support_Type) %>%
  arrange(desc(Support_Type)) %>%
  mutate(Level = factor(Level, levels=Level)) %>%
  ggplot() +
  geom_bar(aes(x = Level, y = n, fill = Support_Type),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
```

```
scale_fill_brewer(palette=4, name = "Post-training support") +
# scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Course Level",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
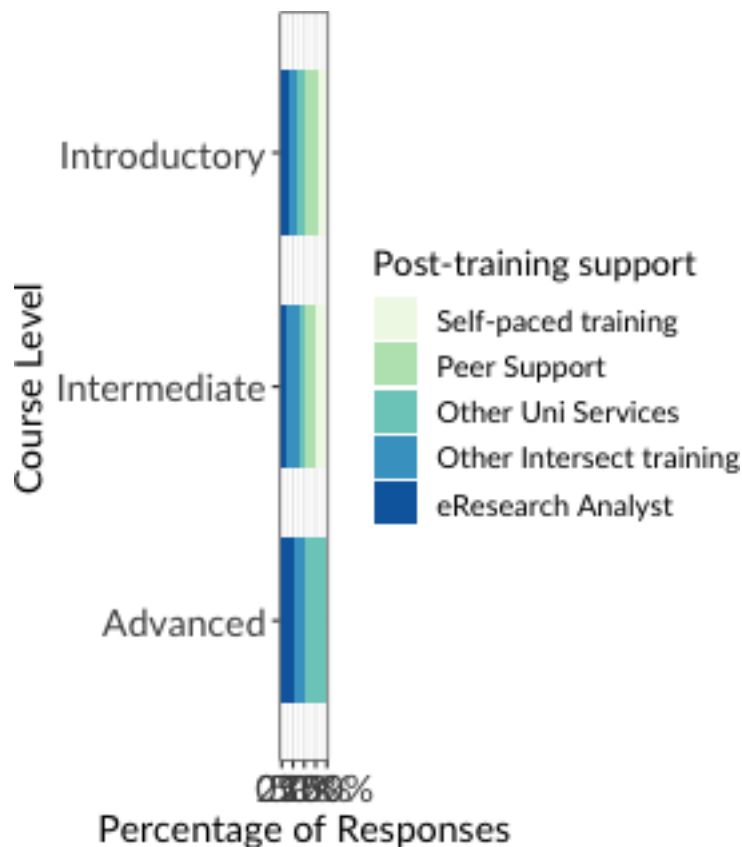

Course attendees were also asked what research output the knowledge they acquired during the course contributed to and Fig.X presents this information by the course level. For all three course levels, Thesis/Dissertation, Presentation, Journal Article, and Conference are the main research outputs the courses have contributed to. It is also worth highlighting that introductory and intermediate courses participants have applied the learning outcome on more research outputs, such as Published Software, Figures/s, Dataset, etc.

```
ResearchOutputs_checkbox_long %>%
  group_by(Level,`Research Output`) %>%
  count() %>%
```

```
drop_na(Level) %>%
drop_na(`Research Output`) %>%
arrange(desc(`Research Output`)) %>%
mutate(Level = factor(Level, levels=Level)) %>%
ggplot() +
geom_bar(aes(x = Level, y = n, fill = `Research Output`),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_manual(values = prism_pal,name="Research outputs") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Course Level",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
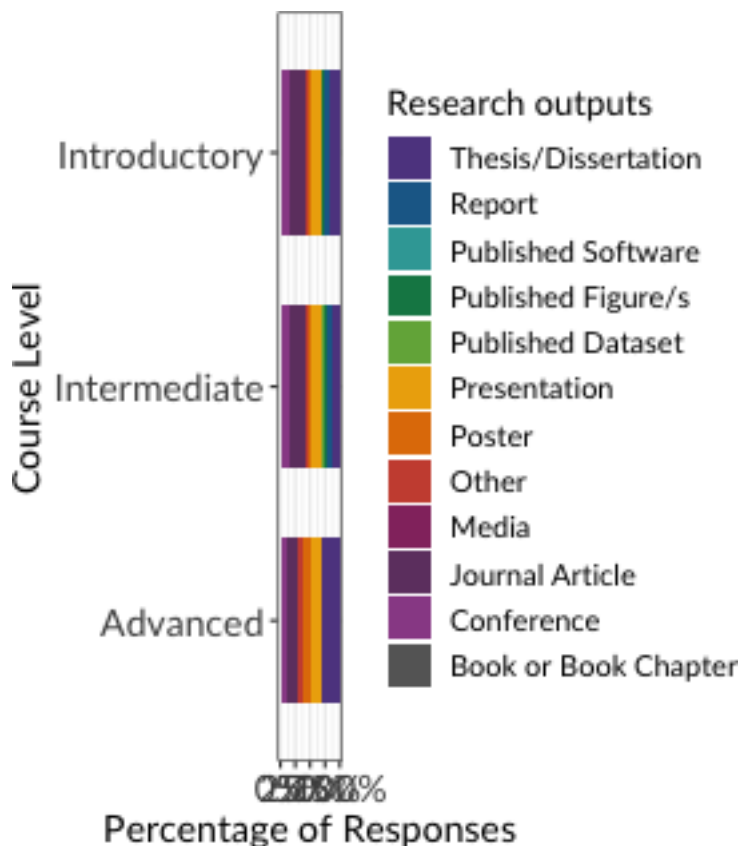

Course attendees were asked to provide the type of support they have received from the Intersect eRAs. For both introductory and intermediate courses, nearly a half of the attendees have received one-time consultation

with the eRAs and around a third received ongoing support. The remaining have had research collaboration with the eRAs. As for the advanced courses, all attendees obtained one-time support from the eRAs.
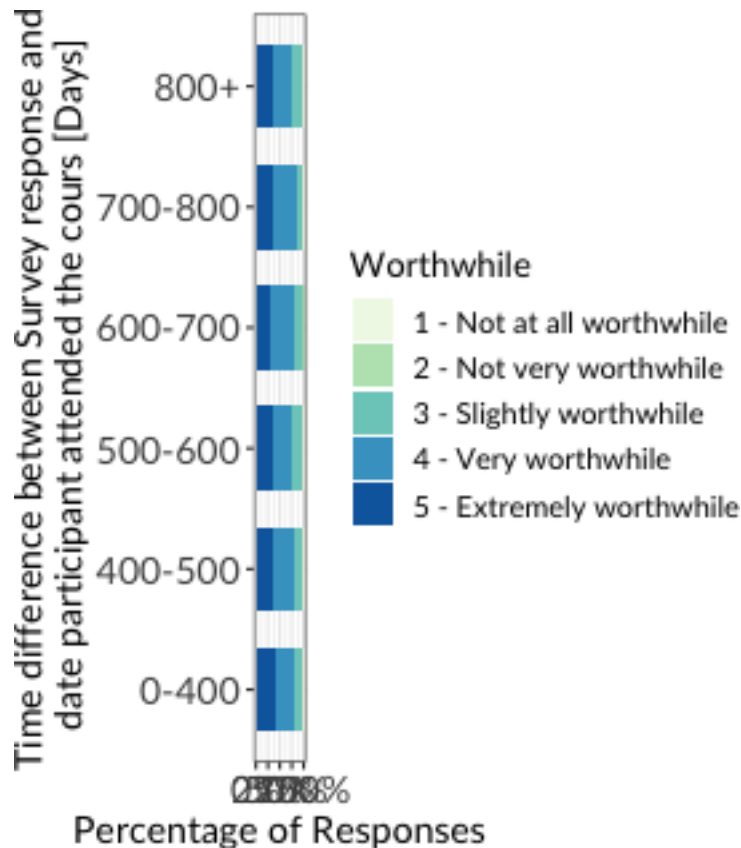
```r
Only_Responses %>%
  mutate(Type_support = `In what capacity have you been working with your local eResearch Analyst(s)?`)
  drop_na(Level, Type_support) %>%
  group_by(Level, Type_support) %>%
  count() %>%
  drop_na(Level) %>%
  arrange(desc(Level)) %>%
  # mutate(Type_support = factor(Type_support, levels=Type_support)) %>%
  mutate(Level = factor(Level, levels=Level)) %>%
  ggplot() +
  geom_bar(aes(x = Level, y = n, fill = Type_support),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4,name="Type of eRA support") +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Faculty",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```
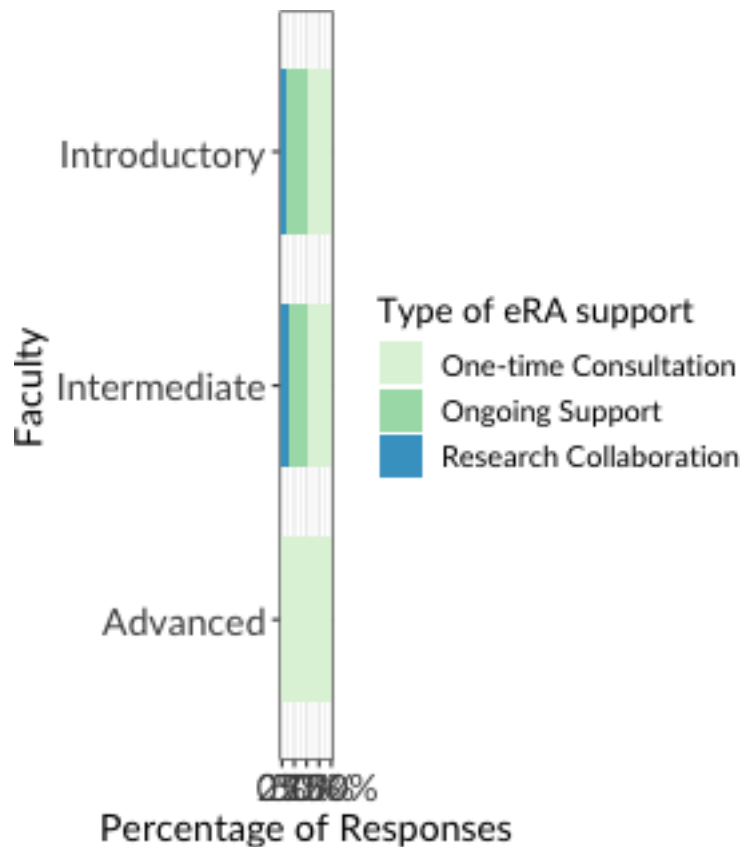
Time difference between Survey and course - Worthwhile

```
Time_diff_df %>%
  group_by(Time_diff_case,reaction_corrected) %>%
  count() %>%
  drop_na(Time_diff_case) %>%
  drop_na(reaction_corrected) %>%
  arrange(desc(reaction_corrected)) %>%
  # mutate(Time_diff_case = factor(Time_diff_case, levels=Time_diff_case)) %>%
  ggplot() +
  geom_bar(aes(x = Time_diff_case, y = n, fill = reaction_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name = "Worthwhile") +
  # scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
```

```
labs(x = "Time difference between Survey response and\n date participant attended the cours [Days]",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```


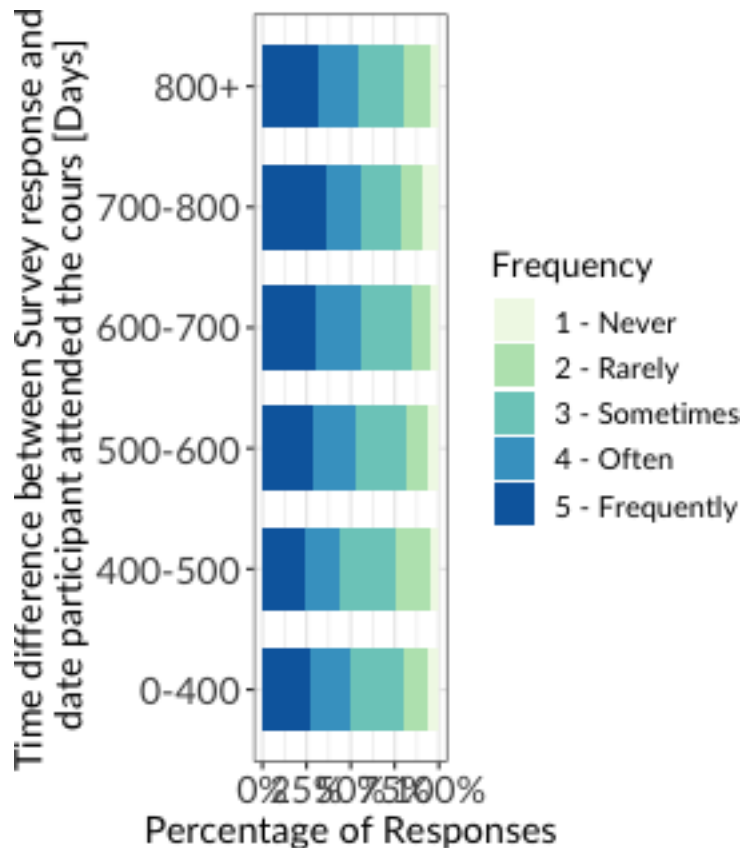## Time diff between Survey and course - Frequency

```
Time_diff_df %>%
  group_by(Time_diff_case,behaviour_corrected) %>%
  count() %>%
  drop_na(Time_diff_case) %>%
  drop_na(behaviour_corrected) %>%
  arrange(desc(behaviour_corrected)) %>%
  # mutate(Time_diff_case = factor(Time_diff_case, levels=Time_diff_case)) %>%
  ggplot() +
  geom_bar(aes(x = Time_diff_case, y = n, fill = behaviour_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name = "Frequency") +
  # scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
```

```
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Time difference between Survey response and\n date participant attended the cours [Days]",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```
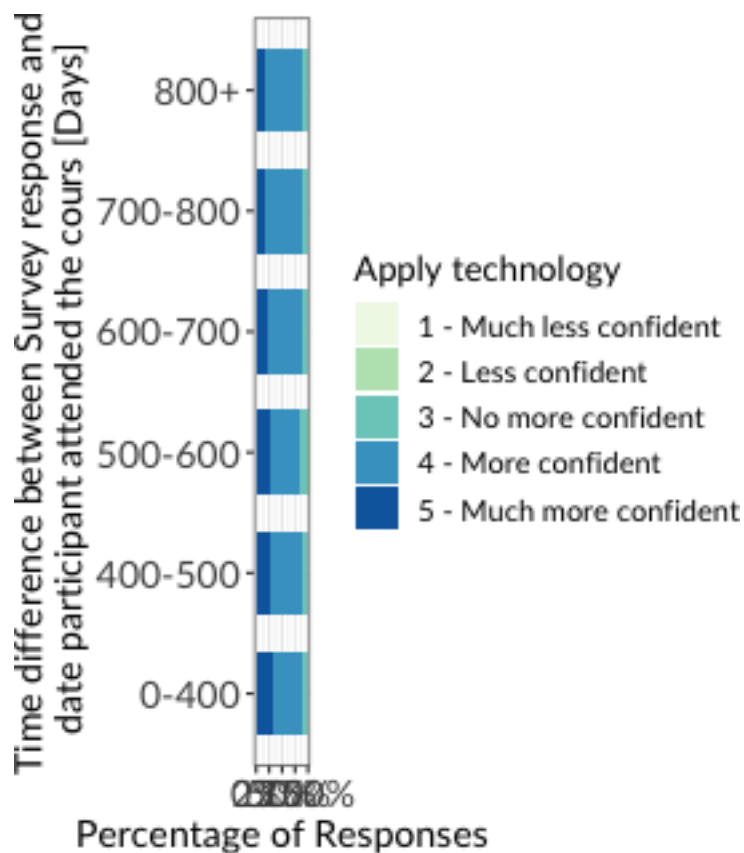

Time diff between Survey and course - Apply technology

```
Time_diff_df %>%
  group_by(Time_diff_case,learning_corrected) %>%
  count() %>%
  drop_na(Time_diff_case) %>%
  drop_na(learning_corrected) %>%
  arrange(desc(learning_corrected)) %>%
  # mutate(Time_diff_case = factor(Time_diff_case, levels=Time_diff_case)) %>%
  ggplot() +
  geom_bar(aes(x = Time_diff_case, y = n, fill = learning_corrected),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name = "Apply technology") +
```

```
# scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Time difference between Survey response and\n date participant attended the cours [Days]",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```
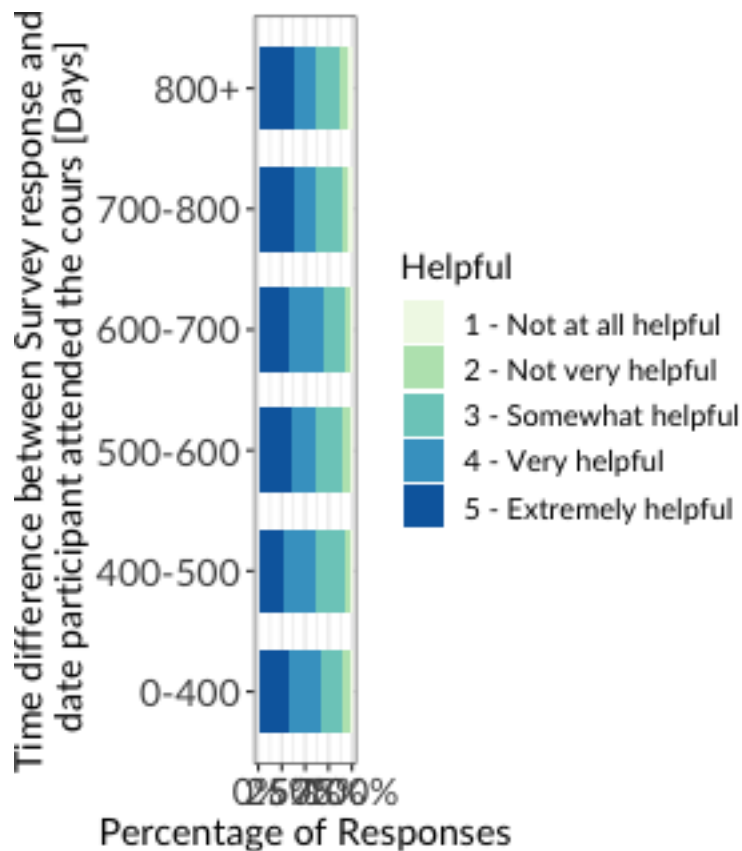

Time diff between Survey and course - Helpful

```
Time_diff_df %>%
  group_by(Time_diff_case,results_corrected) %>%
  count() %>%
  drop_na(Time_diff_case) %>%
  drop_na(results_corrected) %>%
  arrange(desc(results_corrected)) %>%
  # mutate(Time_diff_case = factor(Time_diff_case, levels=Time_diff_case)) %>%
```
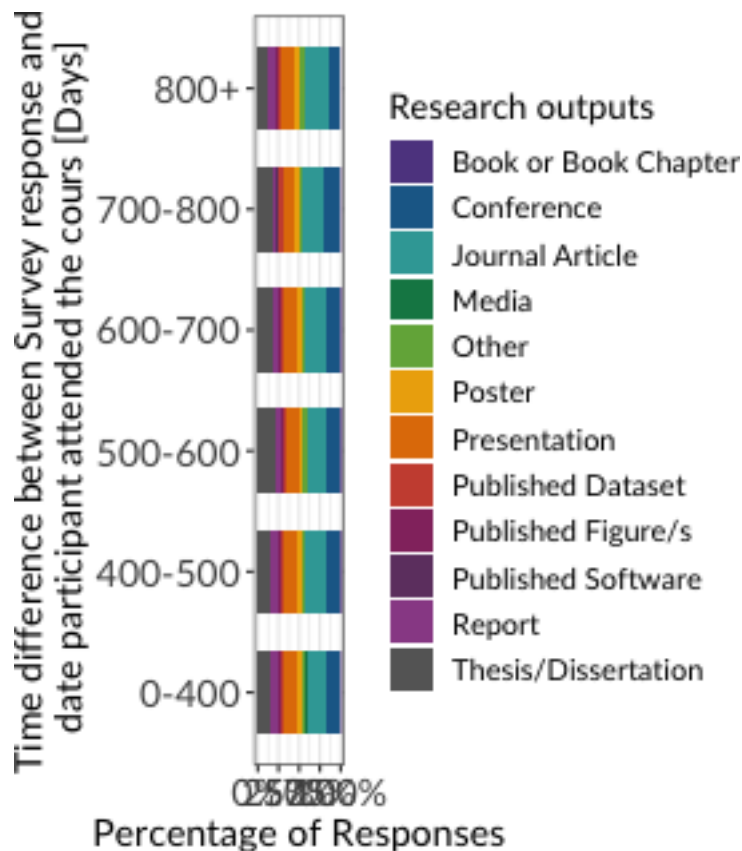
```r
ggplot() +
geom_bar(aes(x = Time_diff_case, y = n, fill = results_corrected),
         stat = 'identity',
         position = "fill",
         width = 0.7) +
scale_fill_brewer(palette=4, name = "Helpful") +
# scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
#theme_ipsum() +
theme_bw() +
theme(
  legend.position = "right",
  axis.title = element_text(size = 14),
  axis.text = element_text(size = 14),
  legend.text = element_text(size = 11),
  legend.title = element_text(size = 13),
  text = element_text(family = "Lato")
) +
labs(x = "Time difference between Survey response and\n date participant attended the cours [Days]",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```

Time diff between Survey and course - Research Outputs

```r
ResearchOutputs_checkbox_long_time_diff %>%
  group_by(Time_diff_case,`Research Output`) %>%
  count() %>%
  drop_na(Time_diff_case) %>%
  drop_na(`Research Output`) %>%
  arrange(desc(`Research Output`)) %>%
  mutate(Time_diff_case = factor(Time_diff_case, levels=Time_diff_case)) %>%
  ggplot() +
  geom_bar(aes(x = Time_diff_case, y = n, fill = `Research Output`),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_manual(values = prism_pal,name="Research outputs") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Time difference between Survey response and\n date participant attended the cours [Days]",
       y = "Percentage of Responses") +
  scale_y_continuous(labels = scales::percent_format()) +
  coord_flip()
```
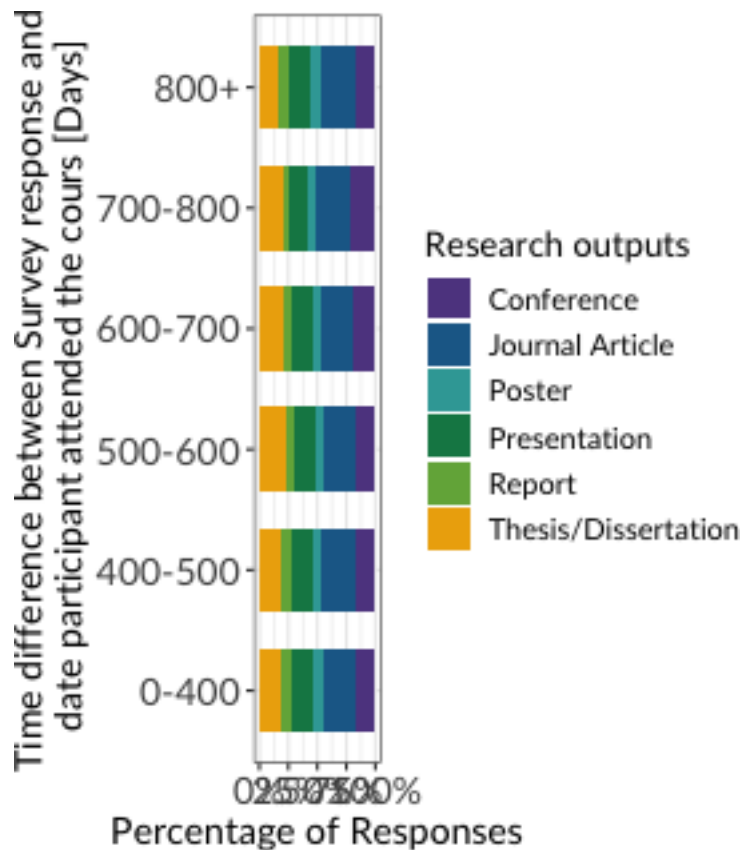
```
ResearchOutputs_checkbox_long_time_diff %>%
  group_by(Time_diff_case,`Research Output`) %>%
  count() %>%
  drop_na(Time_diff_case) %>%
  drop_na(`Research Output`) %>%
  arrange(desc(`Research Output`)) %>%
  mutate(Time_diff_case = factor(Time_diff_case, levels=Time_diff_case)) %>%
  filter(`Research Output` %in% c("Thesis/Dissertation","Report","Presentation",
                                  "Poster","Journal Article","Conference")) %>%
  ggplot() +
  geom_bar(aes(x = Time_diff_case, y = n, fill = `Research Output`),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_manual(values = prism_pal,name="Research outputs") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
  ) +
  labs(x = "Time difference between Survey response and\n date participant attended the cours [Days]",
       y = "Percentage of Responses") +
```

```
    scale_y_continuous(labels = scales::percent_format()) +
    coord_flip()
```


Time diff between Survey and course - Post-training support

```
Support_checkbox_long_time_diff %>%
  group_by(Time_diff_case,Support_Type) %>%
  count() %>%
  drop_na(Time_diff_case) %>%
  drop_na(Support_Type) %>%
  arrange(desc(Support_Type)) %>%
  mutate(Time_diff_case = factor(Time_diff_case, levels=Time_diff_case)) %>%
  ggplot() +
  geom_bar(aes(x = Time_diff_case, y = n, fill = Support_Type),
           stat = 'identity',
           position = "fill",
           width = 0.7) +
  scale_fill_brewer(palette=4, name = "Post-training support") +
  # scale_fill_viridis(name="Worthwhile",discrete = TRUE, option="viridis") +
  #theme_ipsum() +
  theme_bw() +
  theme(
    legend.position = "right",
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 14),
```

```
    legend.text = element_text(size = 11),
    legend.title = element_text(size = 13),
    text = element_text(family = "Lato")
) +
labs(x = "Time difference between Survey response and\n date participant attended the cours [Days]",
     y = "Percentage of Responses") +
scale_y_continuous(labels = scales::percent_format()) +
coord_flip()
```