Core Questions

```
##                 x freq
## 1 Every 1-2 hours  177
## 2 Every 3-4 hours  161
## 3      Every hour   36
## 4           Other   26
```

```
## [1] "About the same"     "Less"               "More"
## [4] "Significantly less" "Significantly more"
```

```
##         x freq
## 1    East   38
## 2 Midwest   75
## 3   South  209
## 4    West   78
```

```
  levels(factor(Core_Questions$State))
```

```
## [1] "East"    "Midwest" "South"   "West"
```

```
  levels(factor(Core_Questions$S_D_E))
```

```
## [1] "Other"        "Three months" "Two months"
```

```
  levels(factor(Core_Questions$RO_F_EG))
```

```
## [1] "1-25%"        "25-50%"       "50-75%"       "75-100%"
## [5] "Fully funded" "Not funded"
```

```
  levels(factor(Core_Questions$Deadline_today))
```

```
## [1] ""    "No"  "Yes"
```

```
  levels(factor(Core_Questions$Workload_today))
```

```
## [1] ""         "Heavy"    "Light"    "Standard"
```

```
  levels(factor(Core_Questions$Workplace))
```

```
## [1] "Home"   "Office" "Other"
```

```
  levels(factor(Core_Questions$R_Style))
```

```
## [1] "Hands-off" "Hands-on"
```

```
  levels(factor(Core_Questions$TW_W_H))
```

```
## [1] "< 30"  "> 50"  "30-40" "40-50"
```

```
  levels(factor(Core_Questions$Break))
```

```
## [1] "Every 1-2 hours" "Every 3-4 hours" "Every hour"      "Other"
```

```
  levels(factor(Core_Questions$Email))
```

```
## [1] "Reply instantly"  "Reply Once/Twice"
```

```
  levels(factor(Core_Questions$funding_proposal))
```

```
## [1] "No"  "Yes"
```

```
  levels(factor(Core_Questions$A_N_Pro))
```

```
## [1] ""     ">=10" "1-2"  "3-4"  "5-6"  "7-9"
```

```
  levels(factor(Core_Questions$funding_agency))
```

```
## [1] "DOD"   "DOE"   "NASA"  "NIH"   "NSF"   "Other"
```

```
  levels(factor(Core_Questions$Success))
```

```
## [1] ""       "< 10%"  "> 90%"  "10-20%" "20-30%" "30-50%" "50-75%" "75-90%"
```

```
  levels(factor(Core_Questions$Com_Proposal))
```

```
## [1] ""           "< 1 week"   "> 2 months" "1-2 months" "1-2 weeks"
## [6] "2-4 weeks"
```

```
  levels(factor(Core_Questions$L_Of_SR))
```

```
## [1] ""            "< 1 month"   "> 12 months" "1-3 months"  "3-6 months"
## [6] "6-12 months"
```

```
  levels(factor(Core_Questions$W_WB_PD))
```

```
## [1] ""                   "About the same"     "Less"
## [4] "More"               "Significantly less" "Significantly more"
```

```
  levels(factor(Core_Questions$Submit_P))
```

```
## [1] "1-3 hours before"      "1 day before"          "2 or more days before"
## [4] "3-6 hours before"      "Minutes before"
```

```
  levels(factor(Core_Questions$Stress_PD))
```

```
## [1] "About the same"     "Less"               "More"
## [4] "Significantly less" "Significantly more"
```

```
  levels(factor(Core_Questions$refereed_conference))
```

```
## [1] "No"  "Yes"
```

```
  levels(factor(Core_Questions$A_N_Conf_Pap))
```

```
## [1] ""      ">= 10" "1-2"   "3-4"   "5-6"   "7-9"
```

```
  levels(factor(Core_Questions$core_rank))
```

```
## [1] ""   "A"  "A*" "B"  "C"
```

```
  levels(factor(Core_Questions$if_you_submit_manuscripts))
```

```
## [1] ""       "< 10%"  "> 90%"  "10-20%" "20-30%" "30-50%" "50-75%" "75-90%"
```

```
  levels(factor(Core_Questions$far_in_advance_do_you))
```

```
## [1] ""           "< 1 week"   "> 2 months" "1-2 months" "1-2 weeks"
## [6] "2-4 weeks"
```

```
  levels(factor(Core_Questions$length_of_supp))
```

```
## [1] ""            "< 1 month"   "> 12 months" "1-3 months"  "3-6 months"
## [6] "6-12 months"
```

```
  levels(factor(Core_Questions$in_the_week_leading_to_a_c))
```

```
## [1] ""                   "About the same"     "Less"
## [4] "More"               "Significantly less" "Significantly more"
```

```
  levels(factor(Core_Questions$you_typically_subm))
```

```
## [1] "1-3 hours before"      "1 day before"          "2 or more days before"
## [4] "3-6 hours before"      "Minutes before"
```

```
  levels(factor(Core_Questions$ss_level_in_a_fundi))
```

```
## [1] "About the same"     "More"               "Significantly less"
## [4] "Significantly more"
```

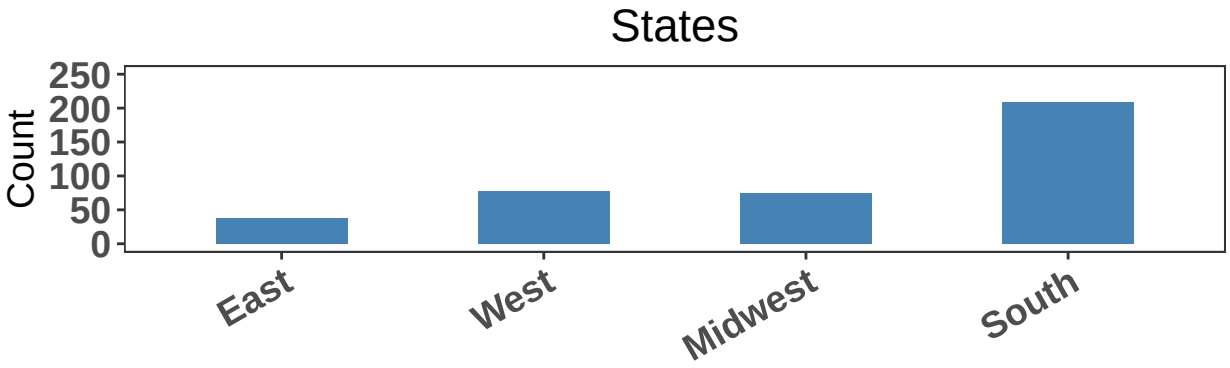
```
  count(Core_Questions$funding_agency)
```

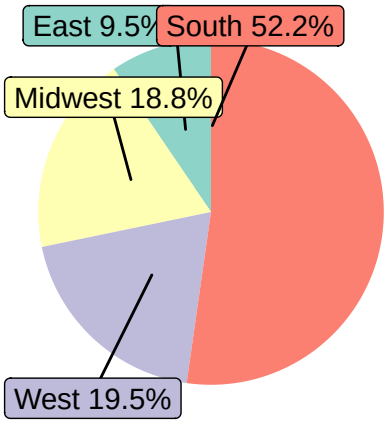
```
##       x freq
## 1   DOD   18
## 2   DOE   21
## 3  NASA   11
## 4   NIH   67
## 5   NSF  234
## 6 Other   49
```

```
# ######Test a single core questions
#   temp <- count(Core_Questions$Workload_today)
#   colnames(temp) <- c("item", "count")
#   temp <- temp[!(temp$item == ""),]
#
#  bar_plot <- ggplot(data = temp, aes(x = item, y = count)) +
#     geom_bar(stat = "identity",
#              width = 0.5,
#              fill = "steelblue") +
#     theme_minimal() +
#     scale_y_continuous(breaks = seq(0, ylimit, by = 5),
#                        limits = c(0, ylimit)) +
#     labs(x = "", y = "Participant count", title = title_list[i - 1]) +
#     theme(
#       panel.grid.major = element_blank(),
#       panel.grid.minor = element_blank(),
#       plot.title = element_text(hjust = 0.5),
#       axis.text.x = element_text(
#         face = "bold",
#         size = 10 ,
#         angle = 30,
#         hjust = 1
#       ),
#       axis.text.y = element_text(face = "bold", size = 10)
#     )+
#    scale_x_discrete( limits=list[[1]])
#
#  bar_plot
```

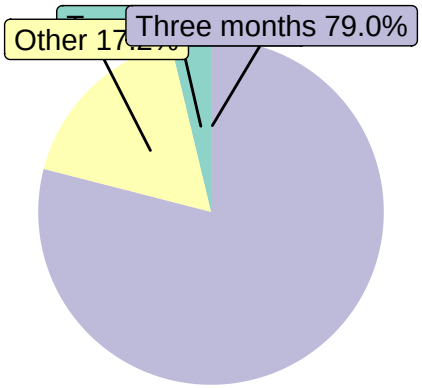
```
## [1] "state_do_you_reside"
```

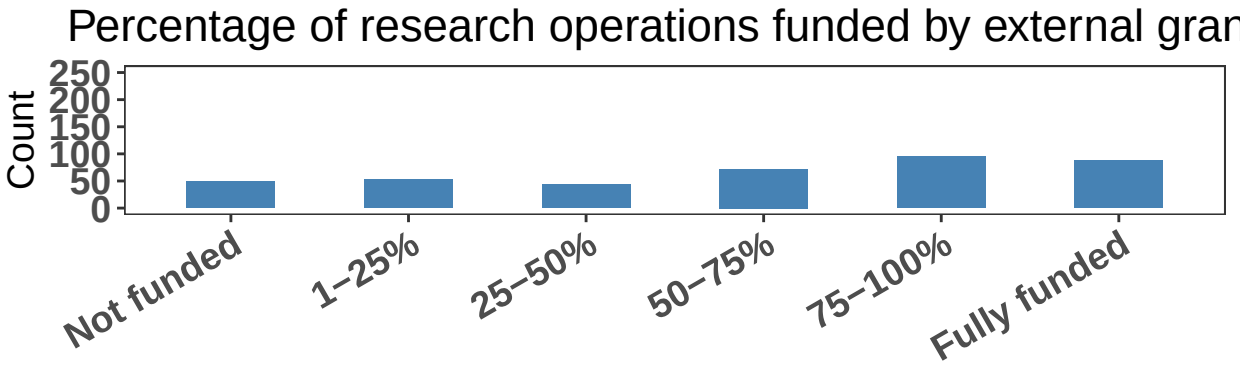
```
## Warning: Removed 1 rows containing missing values (position_stack).
```

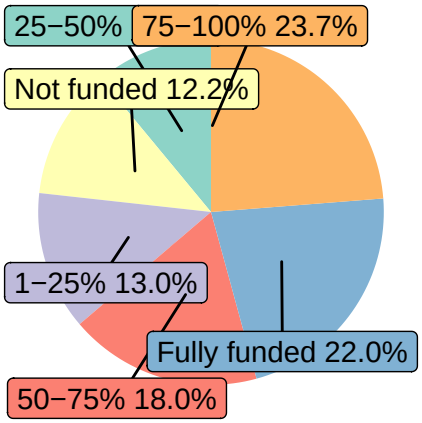

States
Count
250
200
150
100
50
0
East
West
Midwest
South

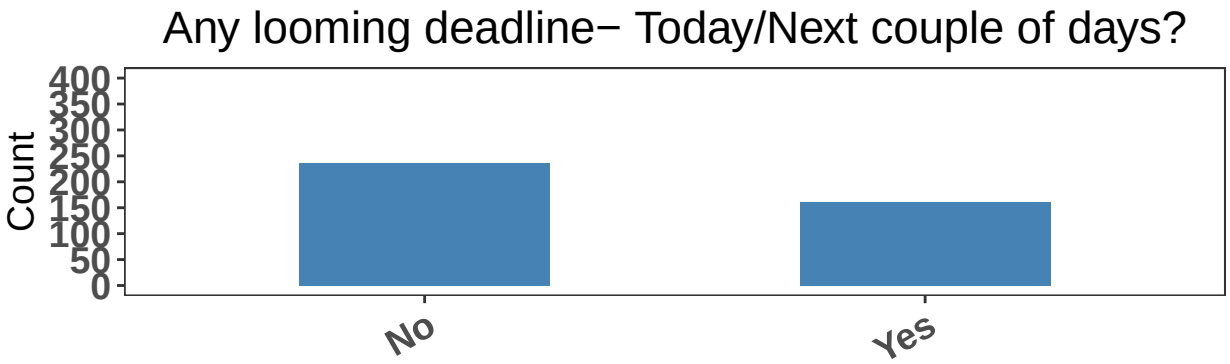

East 9.5%
South 52.2%
Midwest 18.8%
West 19.5%

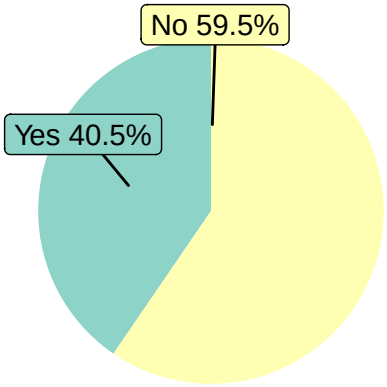

Salary dependance on external grants
Count
400
0
Two months
Three months
Other

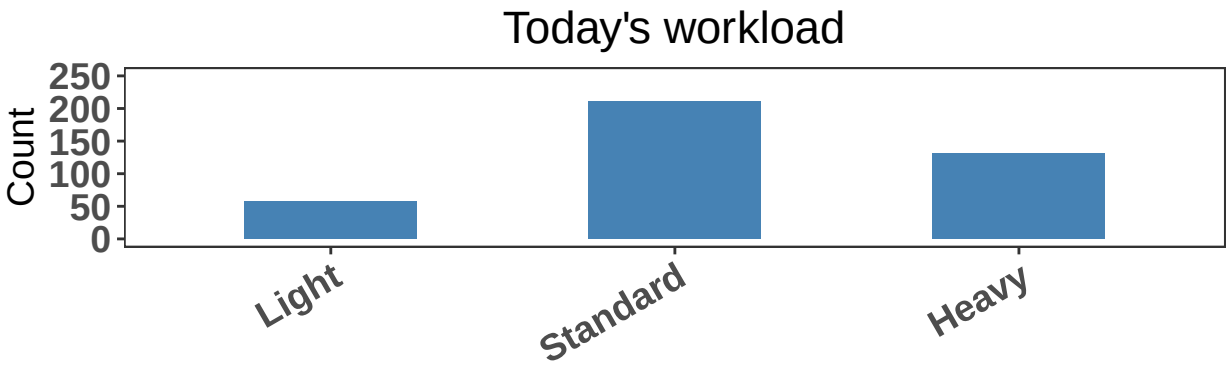

Three months 79.0%

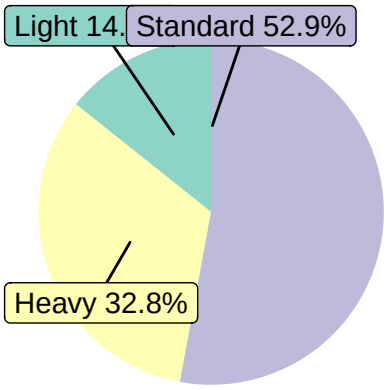

Percentage of research operations funded by external grar
Count
250
200
150
100
50
0
Not funded
1–25%
25–50%
50–75%
75–100%
Fully funded


25–50%
75–100% 23.7%
Not funded 12.2%
1–25% 13.0%
Fully funded 22.0%
50–75% 18.0%

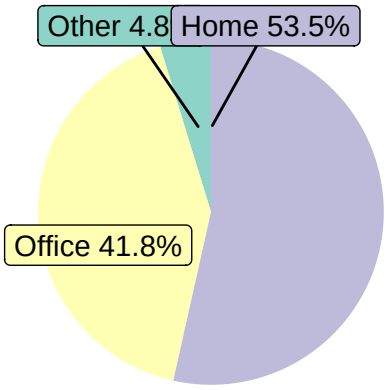

Any looming deadline− Today/Next couple of days?
Count
400
350
300
250
200
150
100
50
0
No
Yes

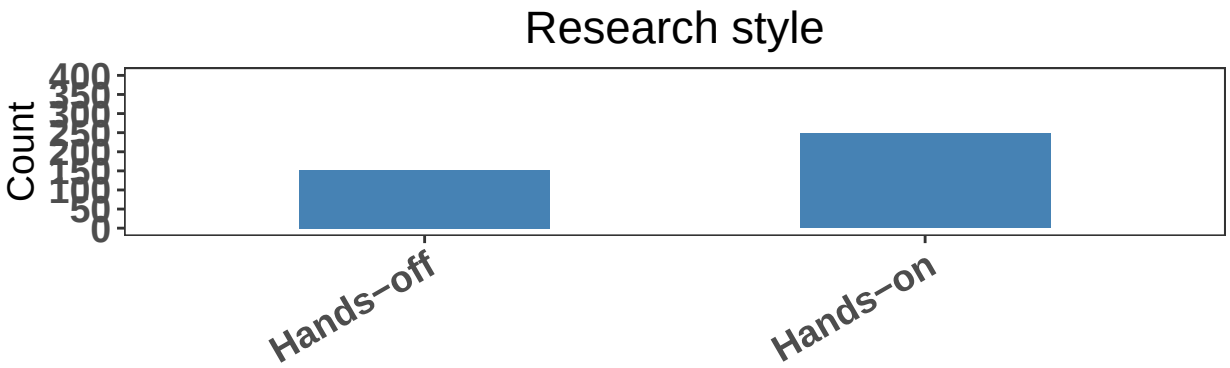

No 59.5%
Yes 40.5%

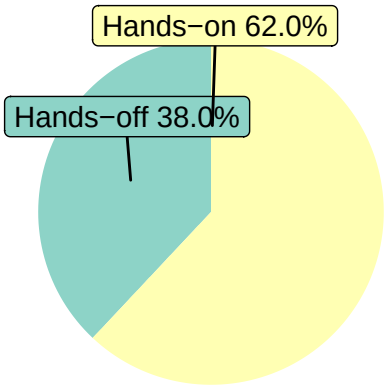

Today's workload
Count
250
200
150
100
50
0
Light
Standard
Heavy

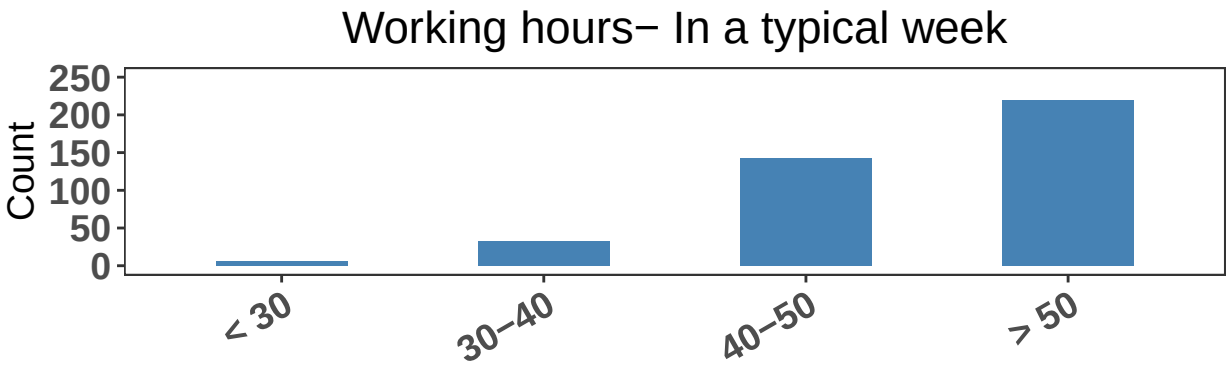

Light 14.
Standard 52.9%
Heavy 32.8%

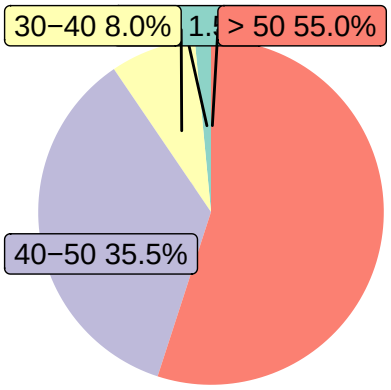

Workplace
Count
250
200
150
100
50
0
Home
Office
Other

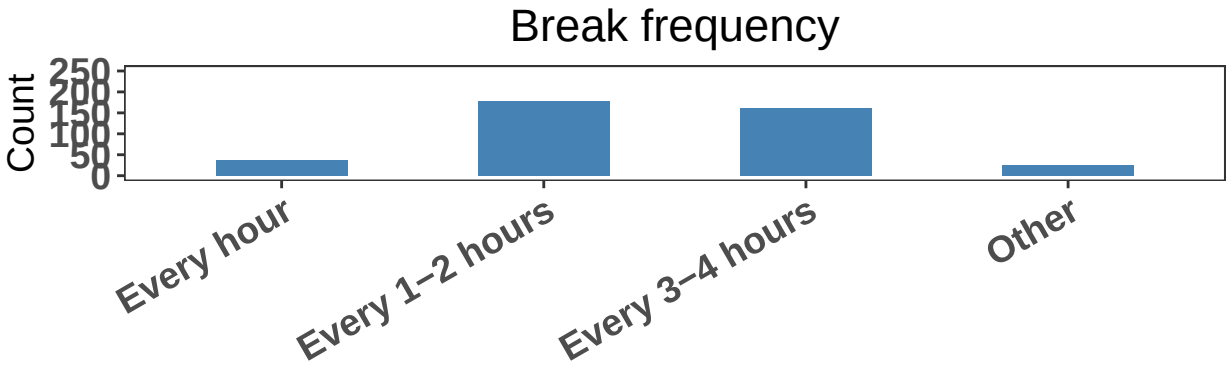

Other 4.8
Home 53.5%
Office 41.8%

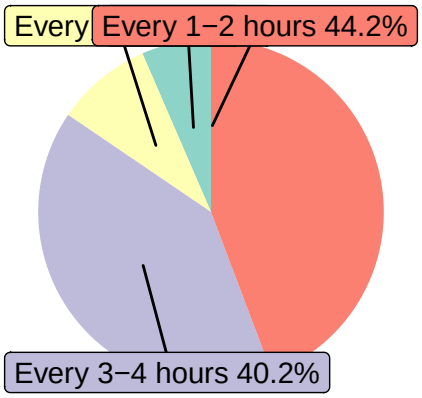

Research style
Count
400
350
300
250
200
150
100
50
0
Hands−off
Hands−on

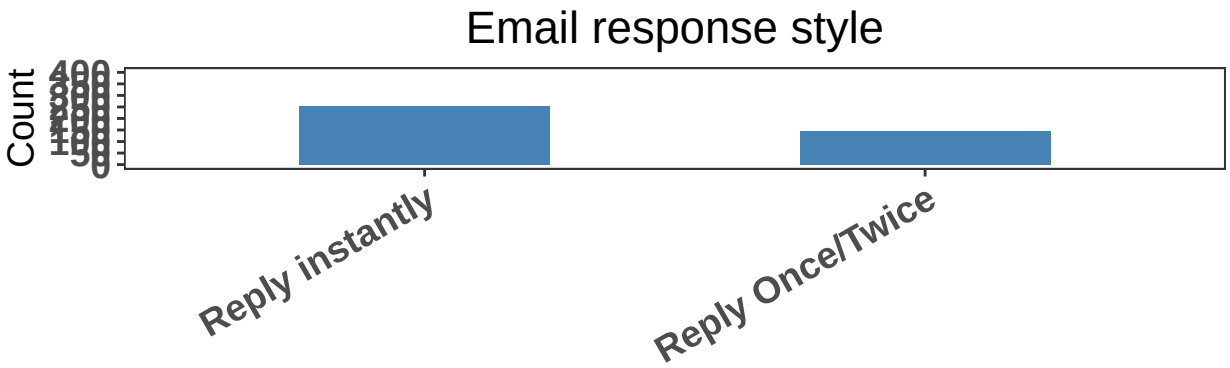

Hands−on 62.0%
Hands−off 38.0%

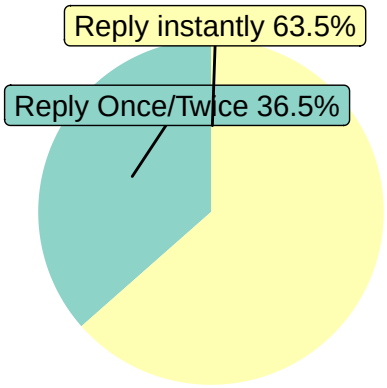

Working hours− In a typical week
Count
250
200
150
100
50
0
< 30
30−40
40−50
> 50

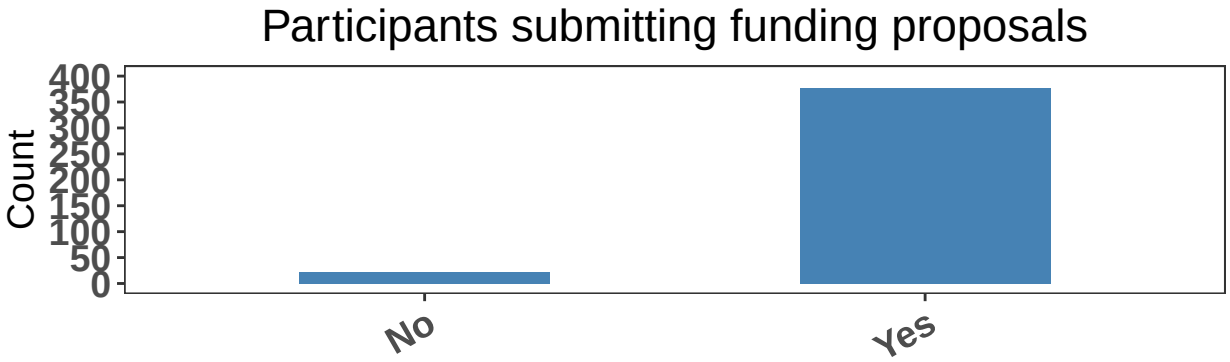

30−40 8.0%
1.5
> 50 55.0%
40−50 35.5%

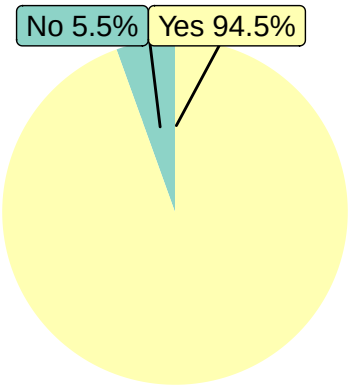

Break frequency
Count
250
200
150
100
50
0
Every hour
Every 1–2 hours
Every 3–4 hours
Other

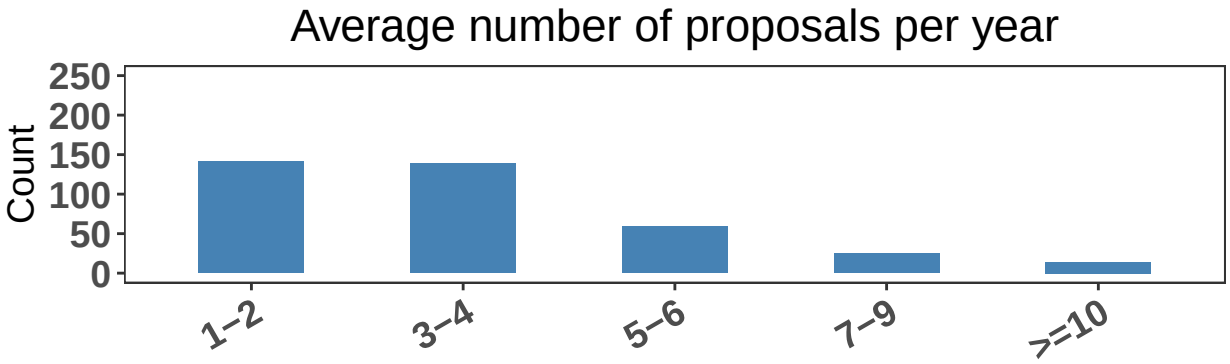

Every
Every 1−2 hours 44.2%
Every 3−4 hours 40.2%

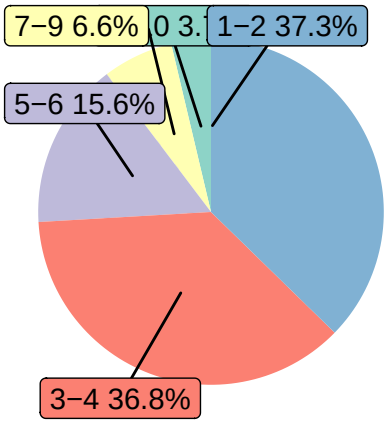

Email response style
Count
Reply instantly
Reply Once/Twice

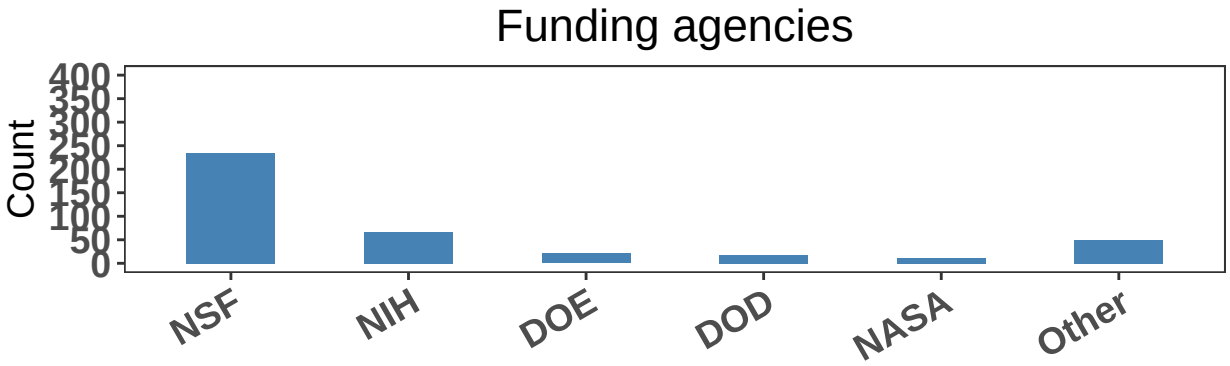

Reply instantly 63.5%
Reply Once/Twice 36.5%

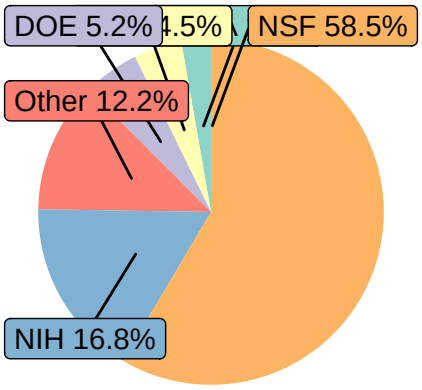

Participants submitting funding proposals
Count
400
350
300
250
200
150
100
50
0
No
Yes

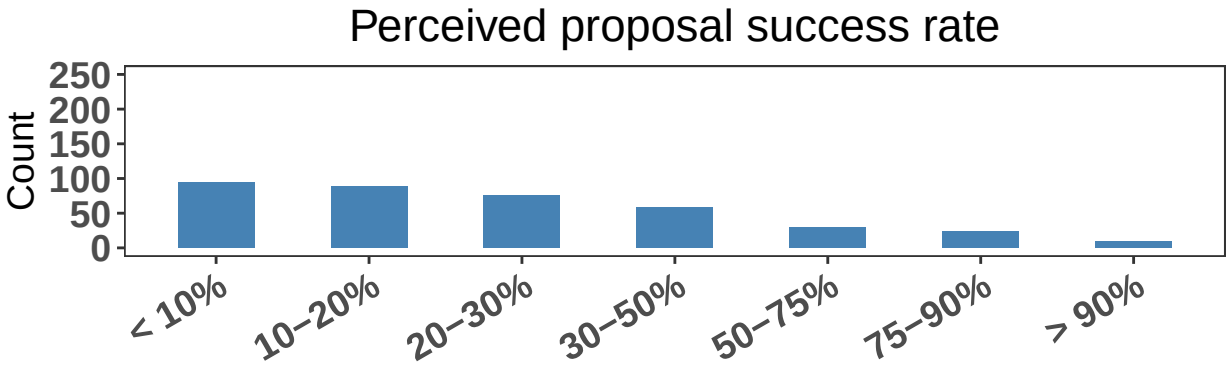

No 5.5%
Yes 94.5%

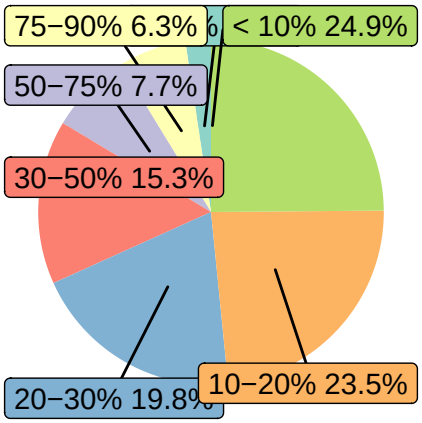

Average number of proposals per year
Count
250
200
150
100
50
0
1–2
3–4
5–6
7–9
>=10

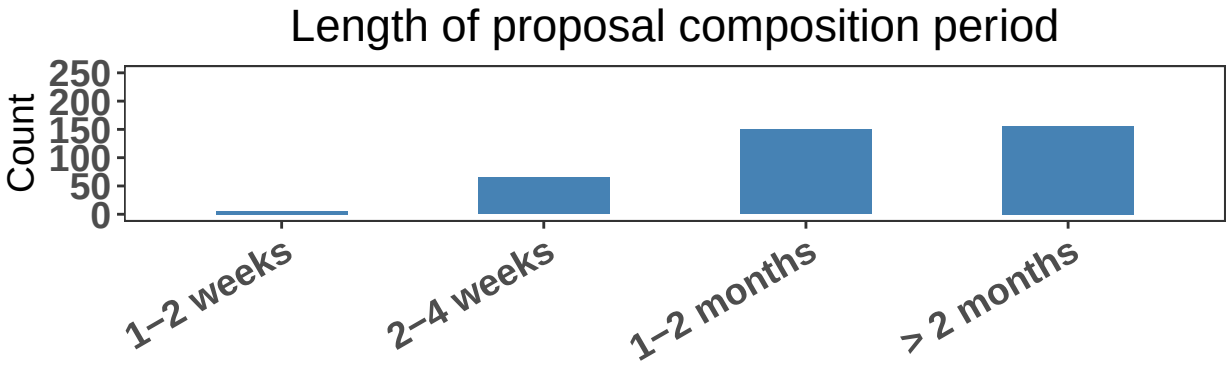

7–9 6.6%
1–2 37.3%
5–6 15.6%
3–4 36.8%

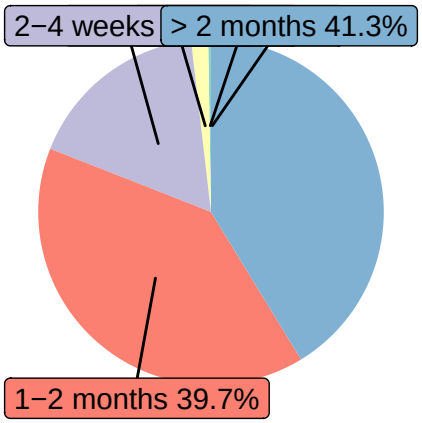

Funding agencies
Count
400
350
300
250
200
150
100
50
0
NSF
NIH
DOE
DOD
NASA
Other

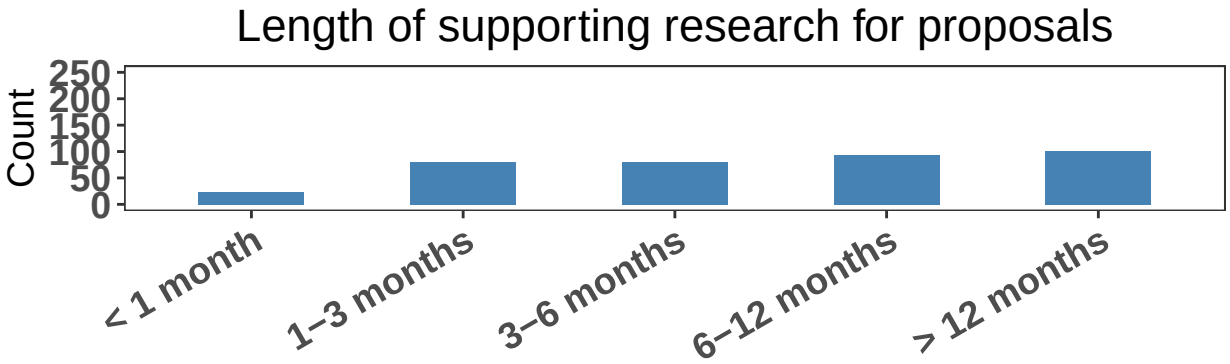

DOE 5.2%
4.5%
NSF 58.5%
Other 12.2%
NIH 16.8%

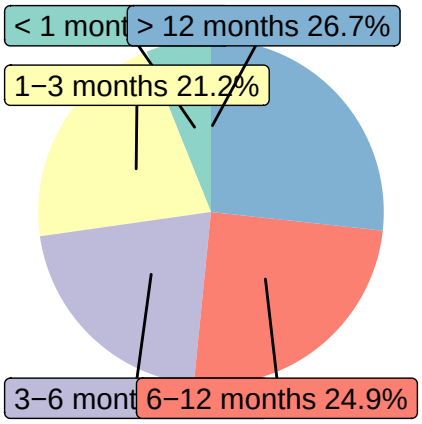

Perceived proposal success rate
Count
250
200
150
100
50
0
< 10%
10–20%
20–30%
30–50%
50–75%
75–90%
> 90%

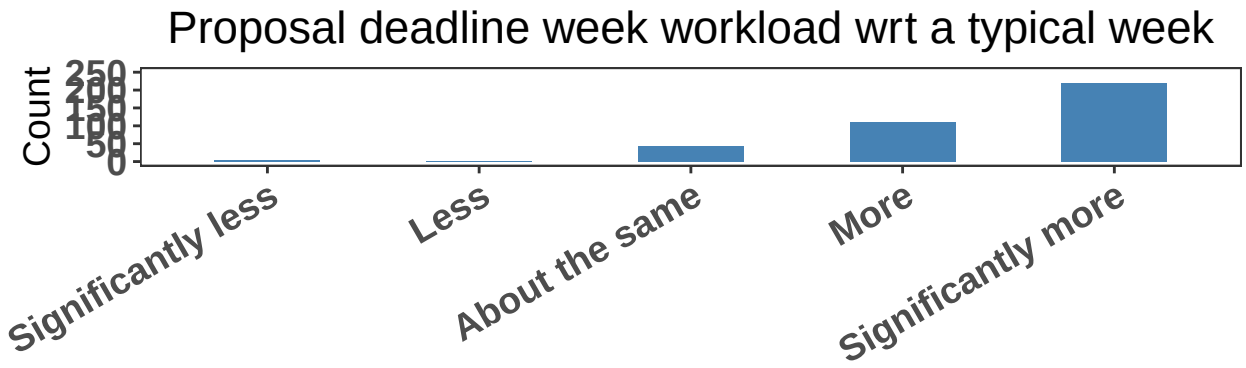

75–90% 6.3%
< 10% 24.9%
50–75% 7.7%
30–50% 15.3%
10–20% 23.5%
20–30% 19.8%

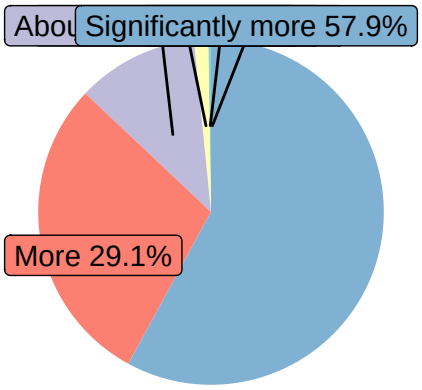

Length of proposal composition period
Count
250
200
150
100
50
0
1–2 weeks
2–4 weeks
1–2 months
> 2 months

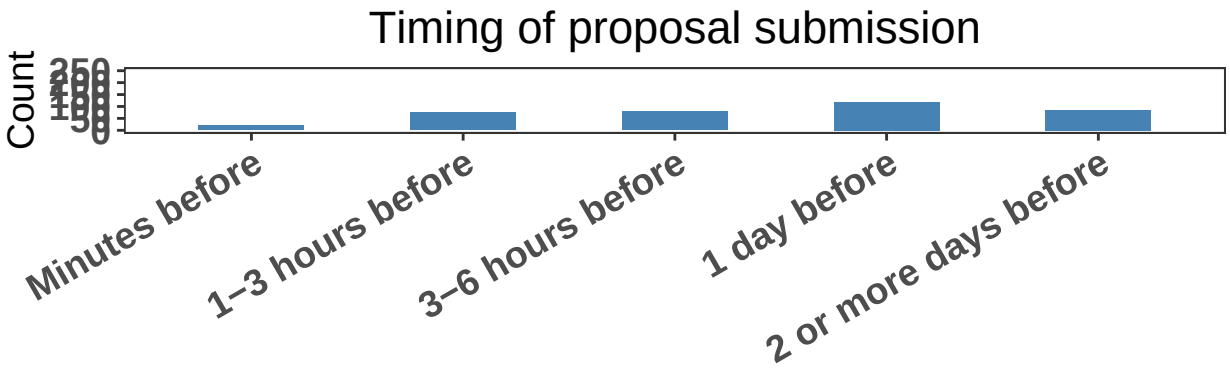

2–4 weeks
> 2 months 41.3%
1–2 months 39.7%

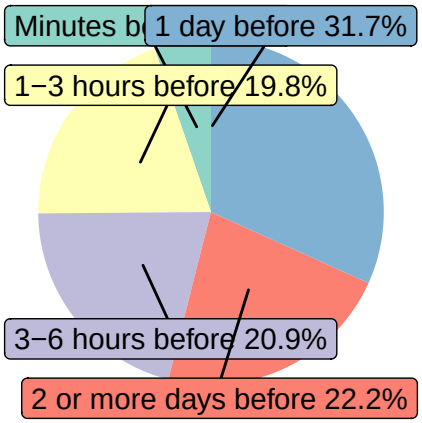

Length of supporting research for proposals
Count
250
200
150
100
50
0
< 1 month
1–3 months
3–6 months
6–12 months
> 12 months

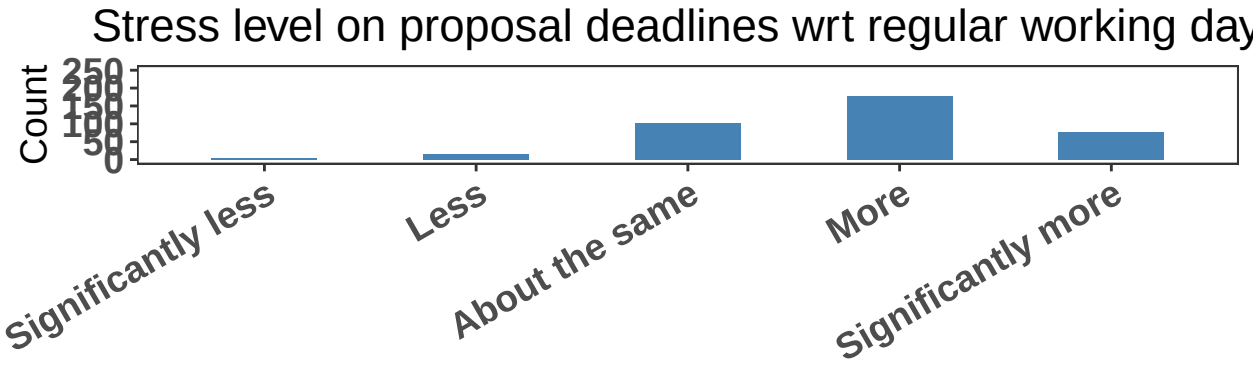

< 1 mont
> 12 months 26.7%
1–3 months 21.2%
3–6 mont
6–12 months 24.9%

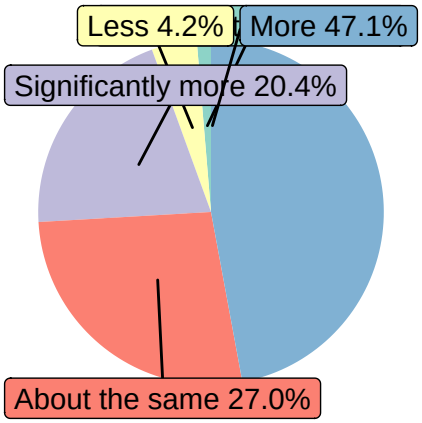

Proposal deadline week workload wrt a typical week
Count
250
200
150
100
50
0
Significantly less
Less
About the same
More
Significantly more

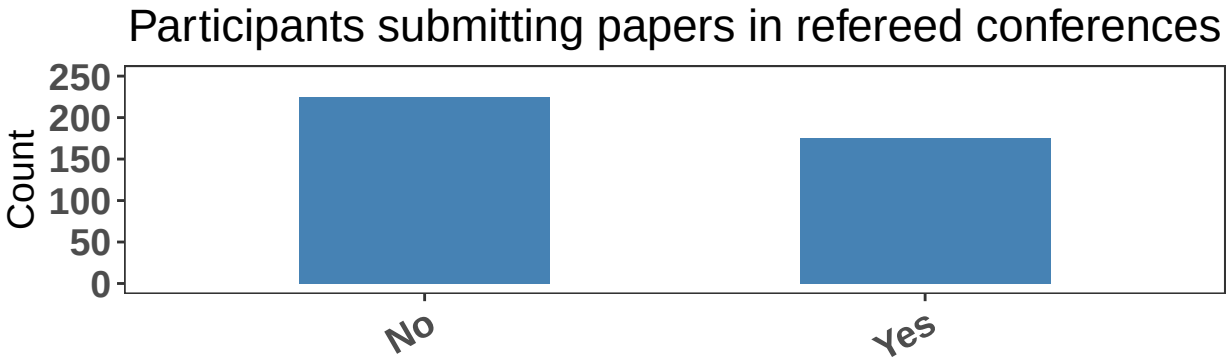

Abou
Significantly more 57.9%
More 29.1%

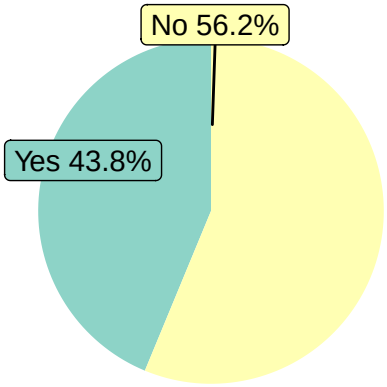

Timing of proposal submission
Count
Minutes before
1–3 hours before
3–6 hours before
1 day before
2 or more days before

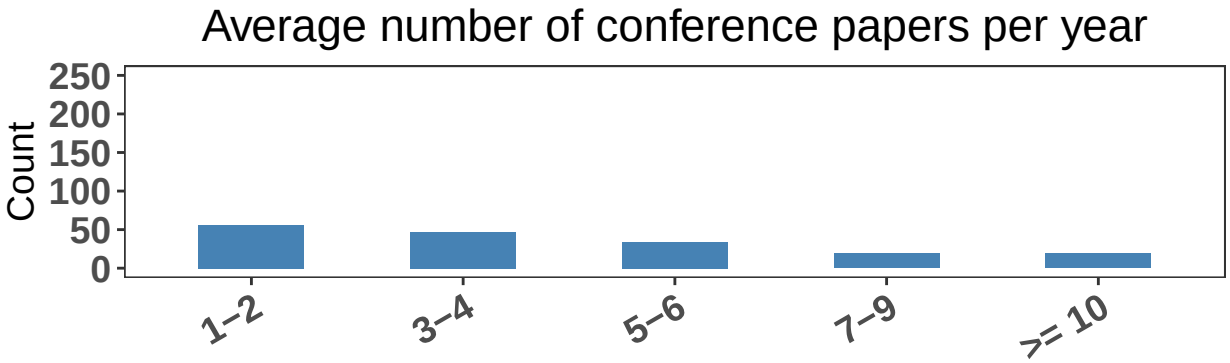

1 day before 31.7%
1–3 hours before 19.8%
3–6 hours before 20.9%
2 or more days before 22.2%

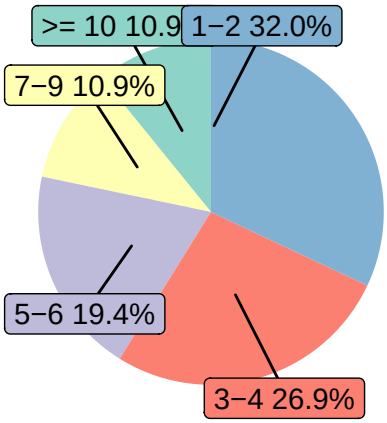

Stress level on proposal deadlines wrt regular working day
Count
250
200
150
100
50
0
Significantly less
Less
About the same
More
Significantly more

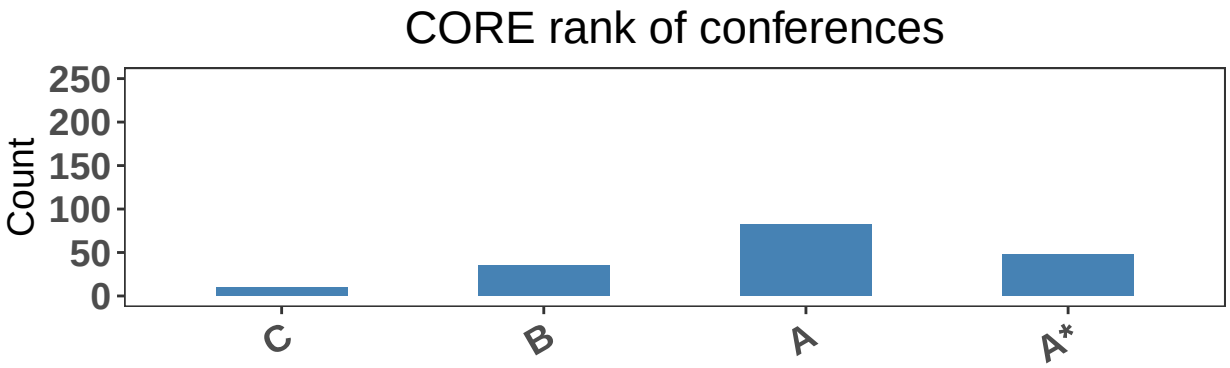

Less 4.2%
More 47.1%
Significantly more 20.4%
About the same 27.0%

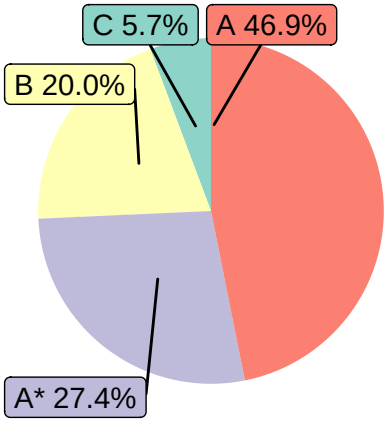

Participants submitting papers in refereed conferences
Count
250
200
150
100
50
0
No
Yes

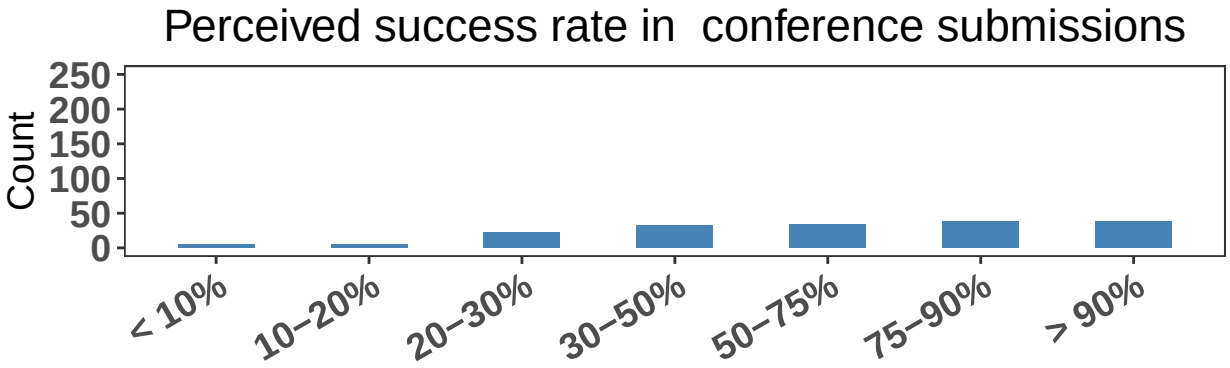

No 56.2%
Yes 43.8%

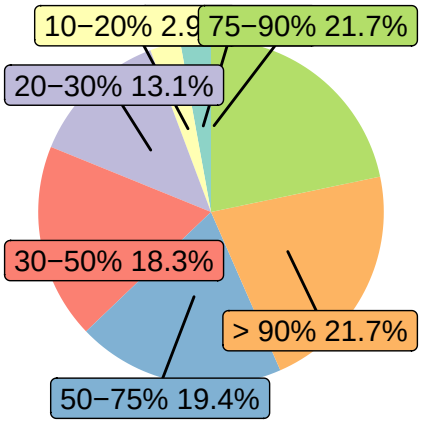

Average number of conference papers per year
Count
250
200
150
100
50
0
1–2
3–4
5–6
7–9
>= 10

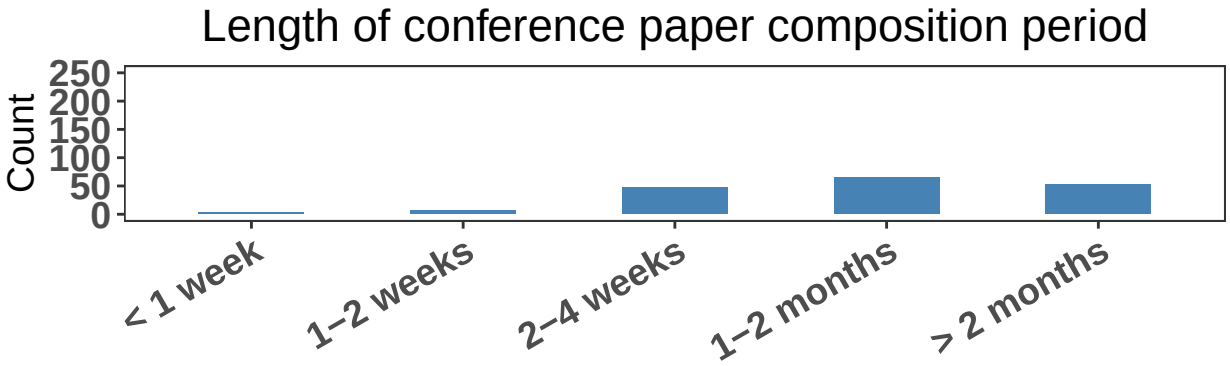

>= 10 10.9
1–2 32.0%
7–9 10.9%
5–6 19.4%
3–4 26.9%

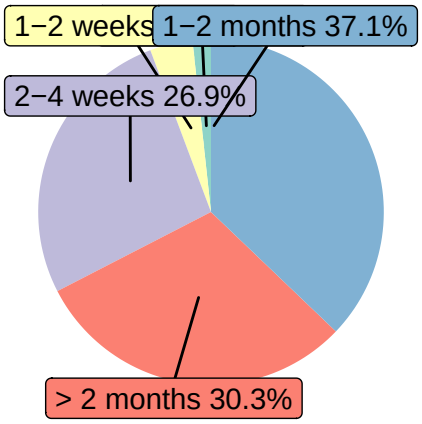

CORE rank of conferences
Count
250
200
150
100
50
0
C
B
A
A*

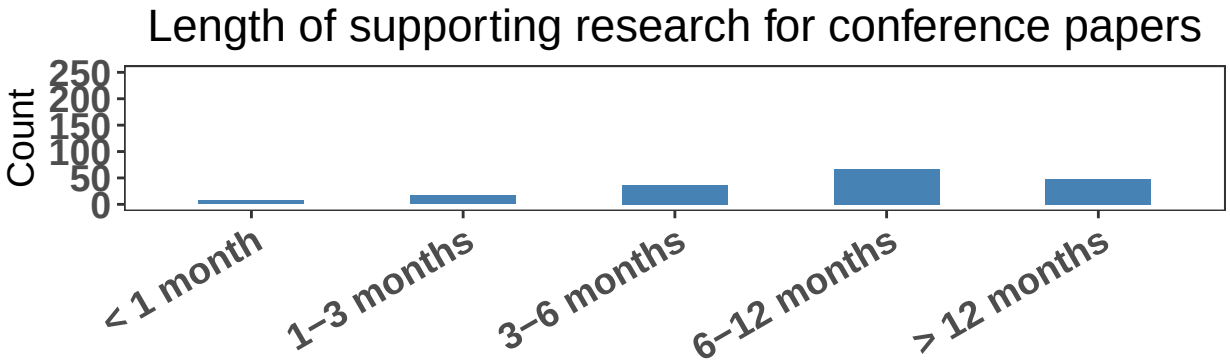

C 5.7%
A 46.9%
B 20.0%
A* 27.4%

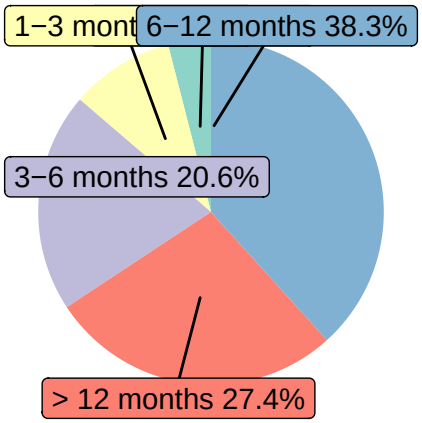

Perceived success rate in conference submissions
Count
250
200
150
100
50
0
< 10%
10–20%
20–30%
30–50%
50–75%
75–90%
> 90%

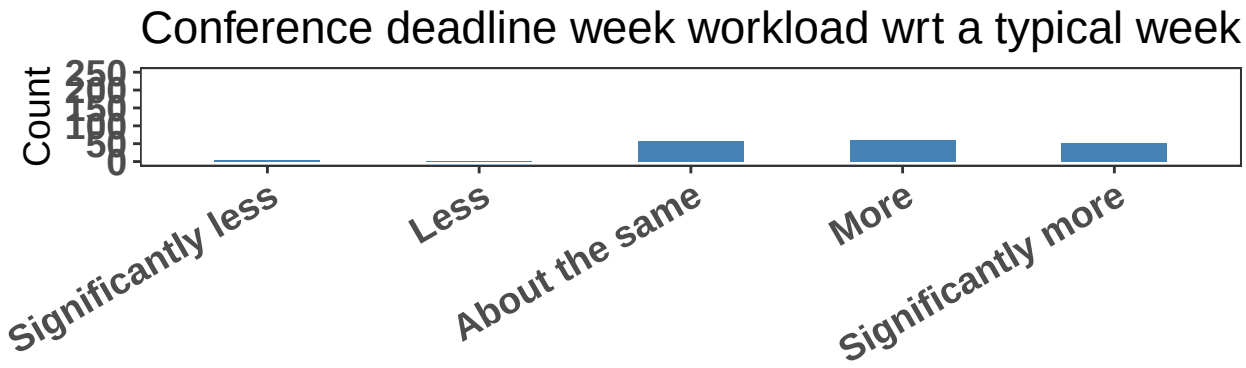

10–20% 2.9
75–90% 21.7%
20–30% 13.1%
30–50% 18.3%
> 90% 21.7%
50–75% 19.4%

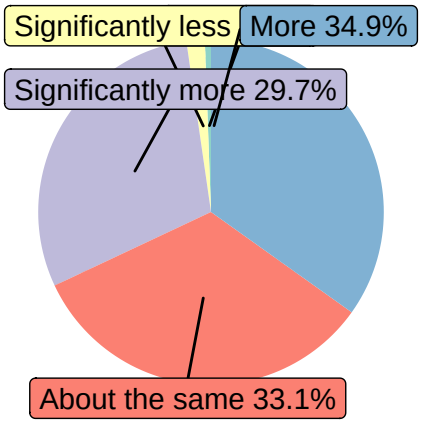

Length of conference paper composition period
Count
250
200
150
100
50
0
< 1 week
1–2 weeks
2–4 weeks
1–2 months
> 2 months

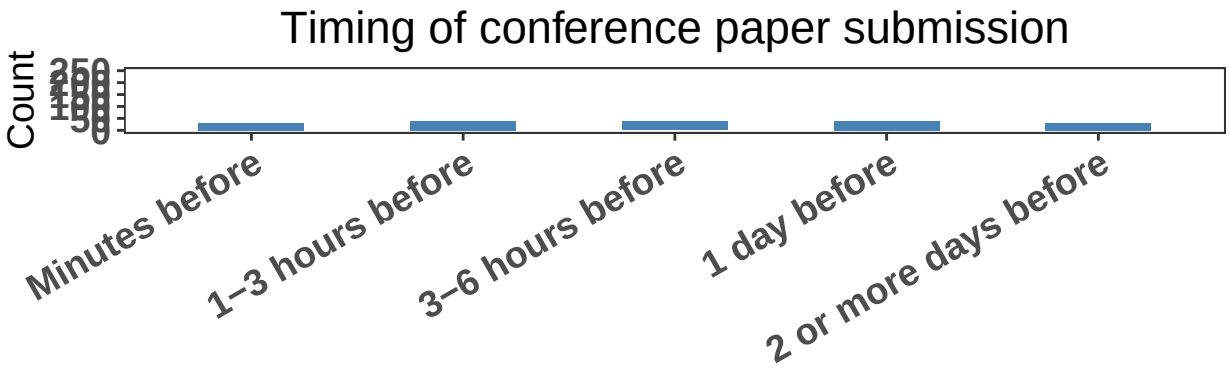

1–2 weeks
1–2 months 37.1%
2–4 weeks 26.9%
> 2 months 30.3%

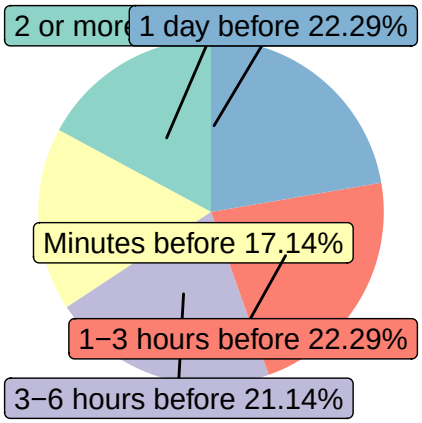

Length of supporting research for conference papers
Count
250
200
150
100
50
0
< 1 month
1–3 months
3–6 months
6–12 months
> 12 months

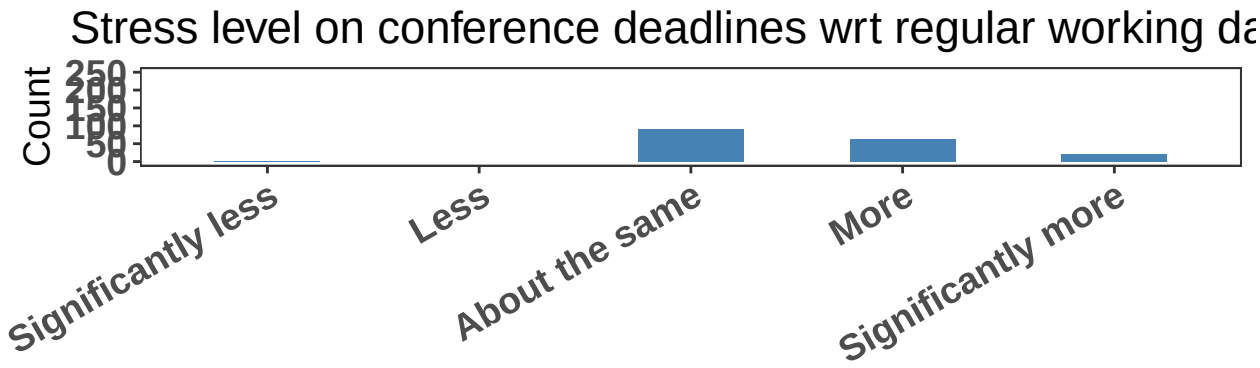

1–3 mont
6–12 months 38.3%
3–6 months 20.6%
> 12 months 27.4%

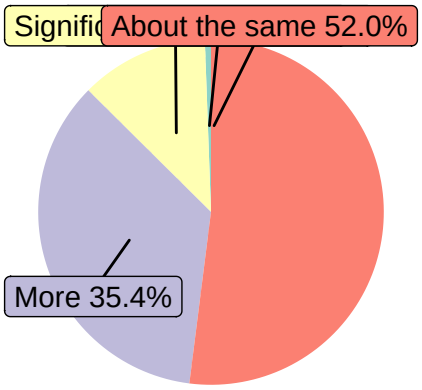

Conference deadline week workload wrt a typical week
Count
250
200
150
100
50
0
Significantly less
Less
About the same
More
Significantly more


Significantly less
More 34.9%
Significantly more 29.7%
About the same 33.1%


Timing of conference paper submission
Count
Minutes before
1–3 hours before
3–6 hours before
1 day before
2 or more days before


2 or mor
1 day before 22.29%
Minutes before 17.14%
1–3 hours before 22.29%
3–6 hours before 21.14%






```
Data_CQ <- Core_Questions[,c(13, 16, 9, 10, 18, 20)]

file_name='Selected_Core_Questions.csv'
write.csv(Data_CQ,file.path(curated_data_dir, file_name), row.names = FALSE)
```

```
# ylimit= 150
#   temp <- count(Core_Questions[14])
#   colnames(temp) <- c("item", "count")
#   temp <- temp[!(temp$item == ""),]
#   temp <- temp[order(temp$count), ]
#   temp <- temp[complete.cases(temp),]
#
#
#   bar_plot <- ggplot(data = temp, aes(x = item, y = count)) +
#     geom_bar(stat = "identity",
#              width = 0.5,
#              fill = "steelblue") +
#     theme_bw() +
#     scale_y_continuous(breaks = seq(0, ylimit, by = 50),
#                        limits = c(0, ylimit)) +
#     labs(x = "", y = "Count", title = "Funding Agencies") +
#     theme(
#       panel.grid.major = element_blank(),
#       panel.grid.minor = element_blank(),
#       plot.title = element_text(hjust = 0.5),
#       axis.text.x = element_text(
```

```
#         face = "bold",
#         size = 10 ,
#         angle = 30,
#         hjust = 1
#       ),
#       axis.text.y = element_text(face = "bold", size = 10)
#     )
#   # +scale_x_discrete( limits=Ticks_list[[i-1]])
#
#   bar_plot
#
#   # plot_list[[length(plot_list) + 1]] <- bar_plot
#
#   temp <- temp %>%
#   mutate(
#     cs = rev(cumsum(rev(count))),
#     prop = percent(count / sum(count)),
#     pos = count/2 + lead(cs, 1),
#     pos = if_else(is.na(pos), count/2, pos))
# temp$pos[is.na(temp$pos)] <- 1
#
# # temp<-temp[match(order_list[[1]], temp$item),]
#
# pichart <-ggplot(temp, aes(x = "" , y = count, fill = item)) +
#   geom_col(width = 1) +
#   coord_polar(theta = "y", start = 0 ) +
#   scale_fill_brewer(palette = "Set3", direction = -4) +
#   geom_label_repel(aes(y = pos, label =paste0(item," " ,prop)), data = temp, size=4, show.legend = F,
#   theme_void() +
#   theme(legend.position =  "none", legend.title = element_blank())+
#   labs(title = "")+
#   theme(panel.grid.major = element_blank(),panel.grid.minor = element_blank(),plot.title = element_te
#
# pichart
#
#
#
# plots_act=ggarrange(bar_plot, pichart, nrow = 2, ncol = 1)
#
#       # final_plot<-plot_grid(plots_act[[i]])
#       filename<-"FA.pdf"
#       full_path<-file.path(plot_dir, filename)
#       ggsave(full_path, plots_act, width = 8.5, height = 11, units = "in")
#       # print(final_plot)
```