

R Markdown

Load the data

Total 290 rows and 124 columnn

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.5.2
```

```
##  
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
## filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
## intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

```
## — Attaching packages —————  
————— tidyverse 1.2.1 —
```

```
## ✔ ggplot2 3.2.1    ✔ readr    1.1.1  
## ✔ tibble  2.1.3    ✔ purrr    0.3.1  
## ✔ tidyr   1.0.0    ✔ stringr  1.4.0  
## ✔ ggplot2 3.2.1    ✔ forcats  0.3.0
```

```
## Warning: package 'ggplot2' was built under R version 3.5.2
```

```
## Warning: package 'tibble' was built under R version 3.5.2
```

```
## Warning: package 'tidyr' was built under R version 3.5.2
```

```
## Warning: package 'purrr' was built under R version 3.5.2
```

```
## Warning: package 'stringr' was built under R version 3.5.2
```

```
## — Conflicts —————  
————— tidyverse_conflicts() —  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag()     masks stats::lag()
```

```
library(ggplot2)
library(scales)
```

```
##
## Attaching package: 'scales'
```

```
## The following object is masked from 'package:purrr':
##
##   discard
```

```
## The following object is masked from 'package:readr':
##
##   col_factor
```

```
library(knitr)
library(kableExtra)
```

```
## Warning: package 'kableExtra' was built under R version 3.5.2
```

```
##
## Attaching package: 'kableExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##   group_rows
```

```
library(magick)
```

```
## Warning: package 'magick' was built under R version 3.5.2
```

```
## Linking to ImageMagick 6.9.9.39
## Enabled features: cairo, fontconfig, freetype, lcms, pango, rsvg, webp
## Disabled features: fftw, ghostscript, x11
```

```
TA_data<-read.csv("/Users/clarawang/Dropbox/Survey-TA/Raw data_May 13_text response.csv", na.strings=c("", "NA"))
TA_data <-TA_data[-c(1,2),]
#head(TA_data)
dim(TA_data)
```

```
## [1] 290 124
```

Remove not TA data

total 280 rows after removing the not TA (Q1)

```
## [1] 280 124
```

```
## [1] "size of raw data: 280"
```

1. International Students Related Questions

1.1 Question: Does international TAs get same percentage of funding compared to local students as TA.

```
#Q6 Are you an international student?
#Q7 When you arrived at UMass, for how many years did you expect to have funding through your department or graduate program?
# create a summary dataset
plotdata <- df_ta %>%
  dplyr::filter(!is.na(Q6) & !is.na(Q7))%>%
  group_by(Q6, Q7) %>%
  dplyr::summarize(n = n()) %>%
  dplyr::mutate(pct = n/sum(n), lbl = scales::percent(pct))
plotdata
```

```
## # A tibble: 10 x 5
## # Groups:   Q6 [2]
##   Q6    Q7                                n    pct lbl
##   <fct> <fct>                                <int> <dbl> <chr>
## 1 No    1 year or less                        5 0.0292 2.9%
## 2 No    For the full expected length of my program 123 0.719 71.9%
## 3 No    I didn't have a clear expectation for my fundi... 16 0.0936 9.4%
## 4 No    More than a year, but less than the full expec... 22 0.129 12.9%
## 5 No    None                                    5 0.0292 2.9%
## 6 Yes   1 year or less                        5 0.0505 5.1%
## 7 Yes   For the full expected length of my program 79 0.798 79.8%
## 8 Yes   I didn't have a clear expectation for my fundi... 3 0.0303 3.0%
## 9 Yes   More than a year, but less than the full expec... 6 0.0606 6.1%
## 10 Yes  None                                    6 0.0606 6.1%
```

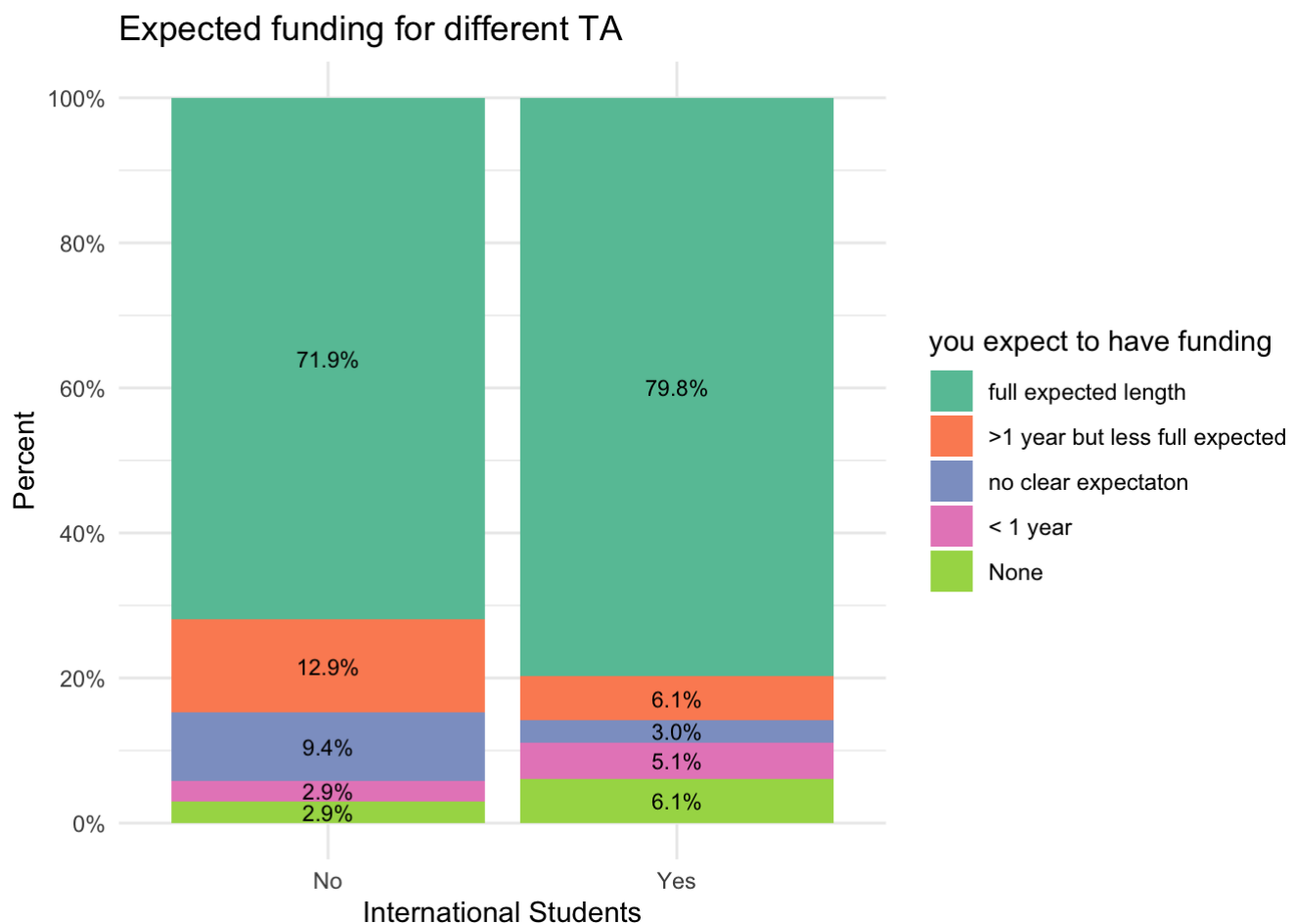
International students gets a little higher percentage funding compared to local students with full expected length of my program.

local TA has a higher percentage with more than a year, but less than the full expected length of my program.

```

# create segmented bar chart
# adding labels to each segment
ggplot(plotdata,
       aes(x = factor(Q6,
                      levels = c("No", "Yes")),
          y = pct,
          fill = factor(Q7,
                      levels = c("For the full expected length of my program", "More than a year, but less than the full expected length of my program", "I didn't have a clear expectation for my funding", "1 year or less", "None"),
                      labels = c("full expected length", ">1 year but less full expected", "no clear expectaton", "< 1 year", "None")))) +
  geom_bar(stat = "identity",
          position = "fill") +
  scale_y_continuous(breaks = seq(0, 1, .2),
                    label = percent) +
  geom_text(aes(label = lbl,
                size = 3,
                position = position_stack(vjust = 0.5)) +
  scale_fill_brewer(palette = "Set2") +
  labs(y = "Percent",
       fill = "you expect to have funding",
       x = "International Students",
       title = "Expected funding for different TA ") +
  theme_minimal()

```



1.2 Question: Is International students professionally communication in English enough?

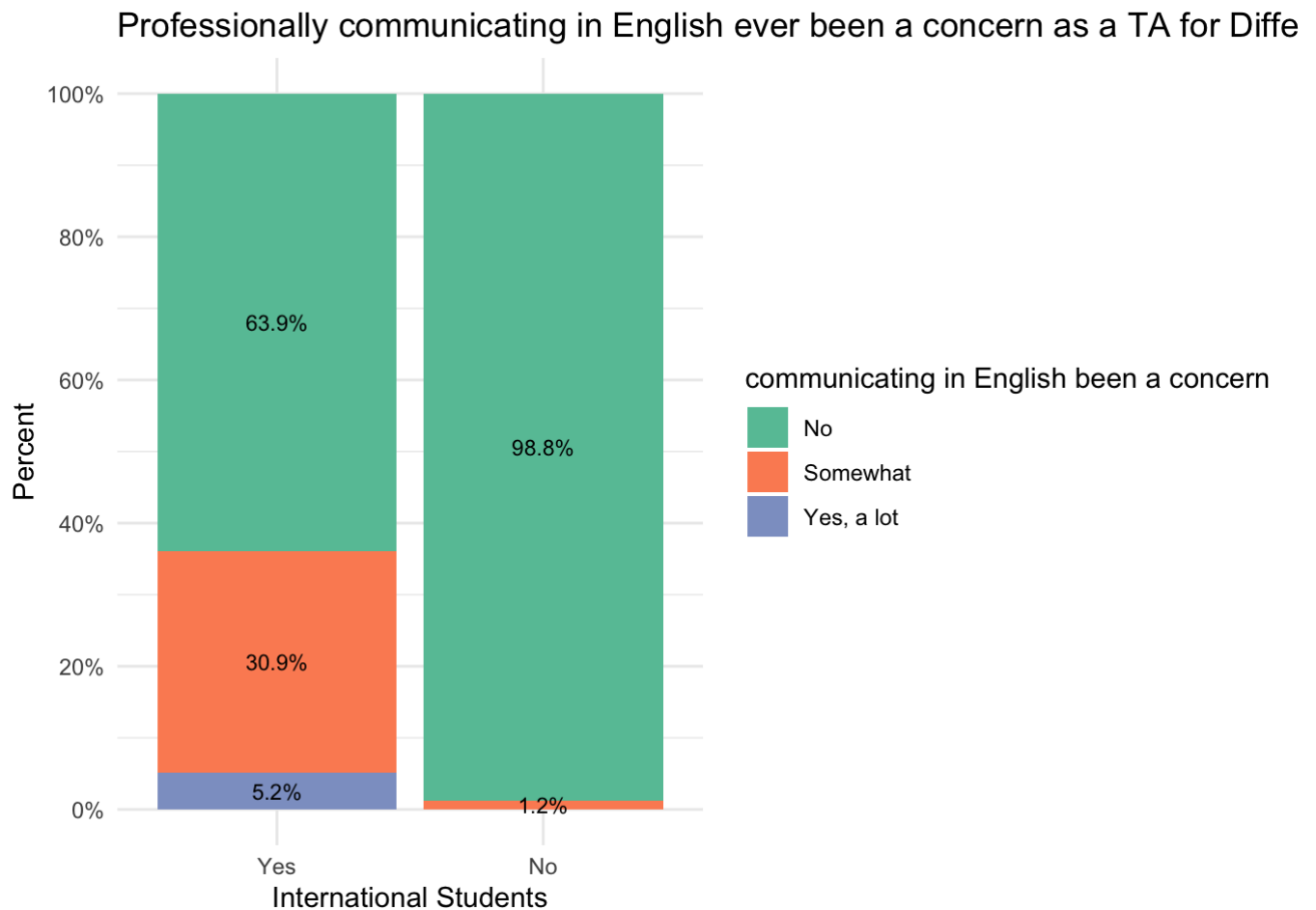
```
#Q6 Are you an international student?
#Q11 Has your level of comfort professionally communicating in English ever been a concern as a TA?

# create a summary dataset
plotdata <- df_ta %>%
  dplyr::filter(!is.na(Q6) & !is.na(Q11))%>%
  group_by(Q6, Q11) %>%
  dplyr::summarize(n = n()) %>%
  dplyr::mutate(pct = n/sum(n),lbl = scales::percent(pct))
plotdata
```

```
## # A tibble: 5 x 5
## # Groups:   Q6 [2]
##   Q6      Q11      n    pct lbl
##   <fct> <fct>   <int> <dbl> <chr>
## 1 No     No       169 0.988 98.8%
## 2 No     Somewhat    2 0.0117 1.2%
## 3 Yes    No        62 0.639 63.9%
## 4 Yes    Somewhat   30 0.309 30.9%
## 5 Yes    Yes, a lot    5 0.0515 5.2%
```

From this plot, About 30% international TA shows somewhat comfort professionally communicating in English ever been a concern as a TA, so professionally English training may help for international TA.

```
# create segmented bar chart
# adding labels to each segment
ggplot(plotdata,
      aes(x = factor(Q6,
                    levels = c("Yes", "No")),
          y = pct,
          fill = factor(Q11,
                    levels = c("No", "Somewhat", "Yes, a lot"),
                    labels = c("No", "Somewhat", "Yes, a lot")))) +
  geom_bar(stat = "identity",
          position = "fill") +
  scale_y_continuous(breaks = seq(0, 1, .2), label = percent) +
  geom_text(aes(label = lbl),
          size = 3,
          position = position_stack(vjust = 0.5)) +
  scale_fill_brewer(palette = "Set2") +
  labs(y = "Percent",
       fill = "communicating in English been a concern",
       x = "International Students",
       title = "Professionally communicating in English ever been a concern as a TA f
or Different TA ") +
  theme_minimal()
```



1.3 Question: Does International students have a higher degree compared to local students ?

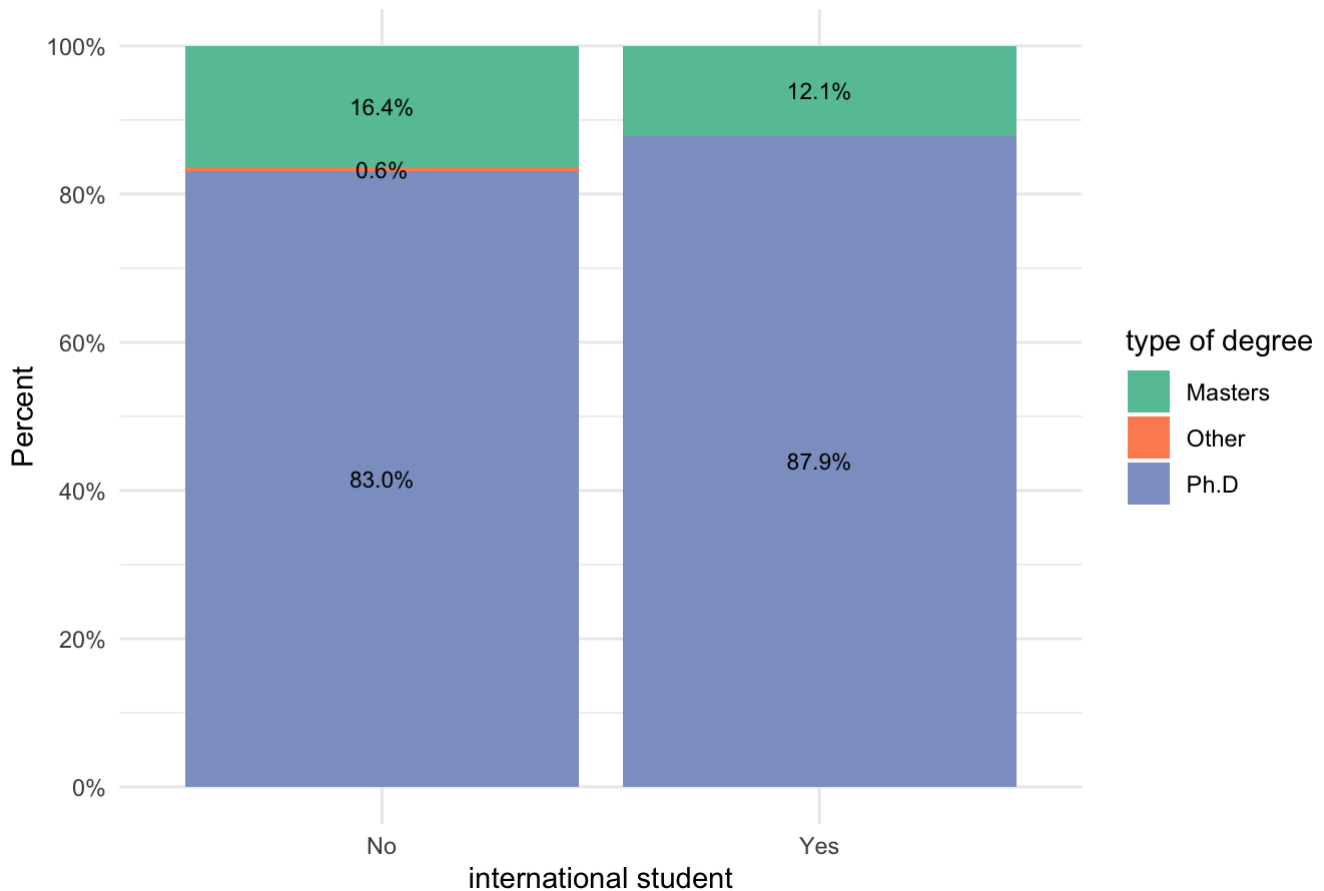
```
#Q6 Are you an international student?
#Q3 What type of degree are you working towards?
# create a summary dataset
plotdata <- df_ta %>%
  dplyr::filter(!is.na(Q6) & !is.na(Q3))%>%
  group_by(Q6, Q3) %>%
  dplyr::summarize(n = n()) %>%
  dplyr::mutate(pct = n/sum(n),lbl = scales::percent(pct))
plotdata
```

```
## # A tibble: 5 x 5
## # Groups:   Q6 [2]
##   Q6     Q3      n    pct lbl
##   <fct> <fct> <int> <dbl> <chr>
## 1 No     Masters    28 0.164 16.4%
## 2 No     Other      1 0.00585 0.6%
## 3 No     Ph.D     142 0.830 83.0%
## 4 Yes    Masters    12 0.121 12.1%
## 5 Yes    Ph.D      87 0.879 87.9%
```

Yes ,a little bit,Internaltional TA has hihger percentage of Ph.d degree.

```
# create segmented bar chart
# adding labels to each segment
ggplot(plotdata,
  aes(x = factor(Q6,
    levels = c("No","Yes")),
    y = pct,
    fill = factor(Q3,
      levels = c("Masters","Other","Ph.D"),
      labels = c("Masters","Other","Ph.D")))) +
  geom_bar(stat = "identity",
    position = "fill") +
  scale_y_continuous(breaks = seq(0, 1, .2), label = percent) +
  geom_text(aes(label = lbl),
    size = 3,
    position = position_stack(vjust = 0.5)) +
  scale_fill_brewer(palette = "Set2") +
  labs(y = "Percent",
    fill = "type of degree",
    x = "international student",
    title = "Are you an international student vs Type of degree ") +
  theme_minimal()
```

Are you an international student vs Type of degree



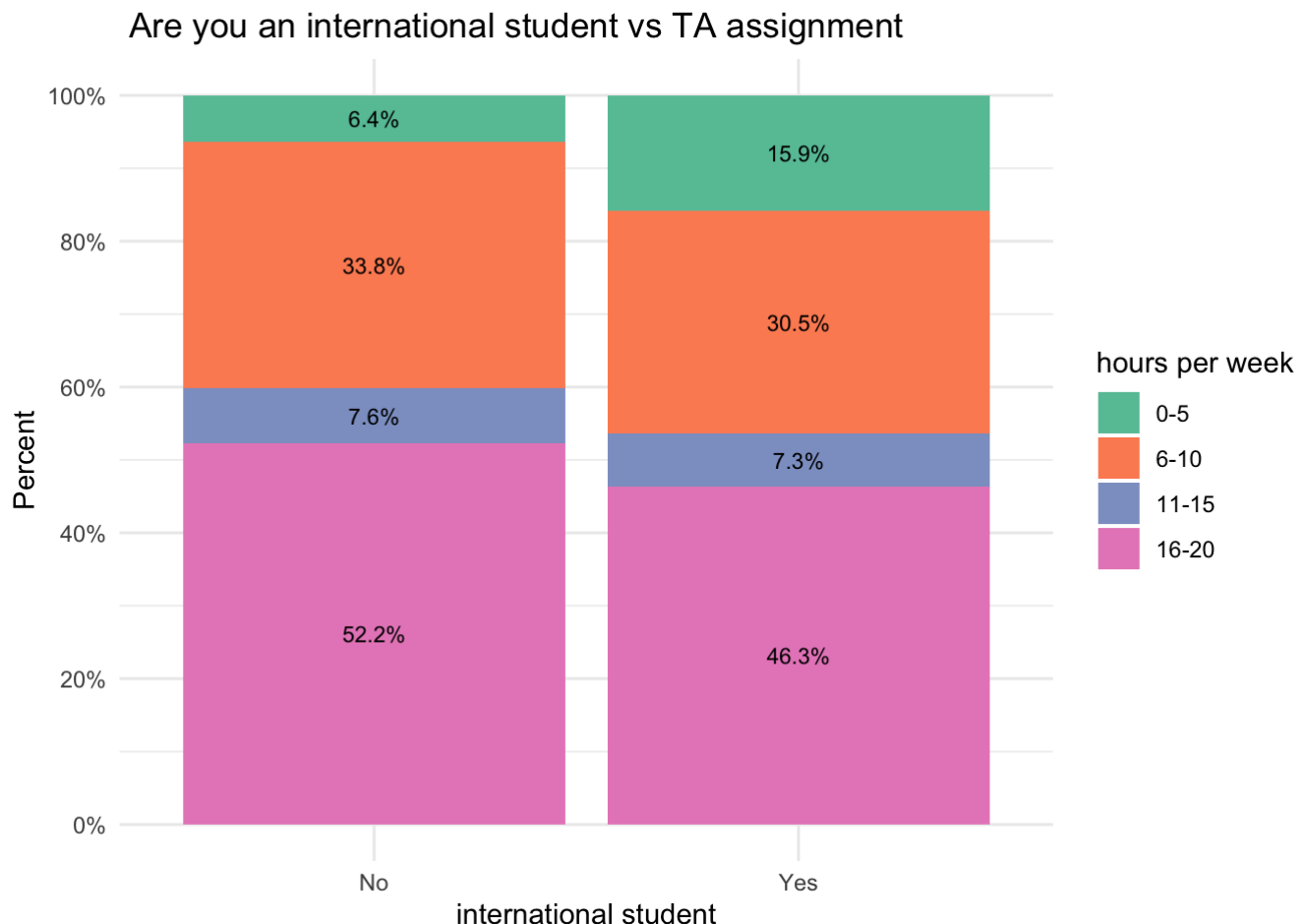
1.4 Question : Dose International students have higher hours assignment?

```
#Q6 Are you an international student?
#Q16 How many hours per week were you assigned for this TA assignment?
# create a summary dataset
plotdata <- df_ta %>%
  dplyr:: filter(!is.na(Q6) & !is.na(Q16))%>%
  group_by(Q6, Q16) %>%
  dplyr::summarize(n = n()) %>%
  dplyr::mutate(pct = n/sum(n),lbl = scales::percent(pct))
plotdata
```

```
## # A tibble: 8 x 5
## # Groups:   Q6 [2]
##   Q6    Q16      n    pct lbl
##   <fct> <fct> <int> <dbl> <chr>
## 1 No    0-5      10 0.0637 6.4%
## 2 No    10-Jun    53 0.338 33.8%
## 3 No    15-Nov    12 0.0764 7.6%
## 4 No    16-20    82 0.522 52.2%
## 5 Yes    0-5      13 0.159 15.9%
## 6 Yes    10-Jun    25 0.305 30.5%
## 7 Yes    15-Nov     6 0.0732 7.3%
## 8 Yes    16-20    38 0.463 46.3%
```


No, actually international students have lower TA assignment except below 5 hours assignment per week

```
# create segmented bar chart
# adding labels to each segment
ggplot(plotdata,
       aes(x = factor(Q6,
                      levels = c("No", "Yes")),
          y = pct,
          fill = factor(Q16,
                       levels = c("0-5", "10-Jun", "15-Nov", "16-20"),
                       labels = c("0-5", "6-10", "11-15", "16-20")))) +
  geom_bar(stat = "identity",
           position = "fill") +
  scale_y_continuous(breaks = seq(0, 1, .2), label = percent) +
  geom_text(aes(label = lbl),
            size = 3,
            position = position_stack(vjust = 0.5)) +
  scale_fill_brewer(palette = "Set2") +
  labs(y = "Percent",
       fill = "hours per week",
       x = "international student",
       title = " Are you an international student vs TA assignment ") +
  theme_minimal()
```



2. TA Degree vs Hours per week assignment

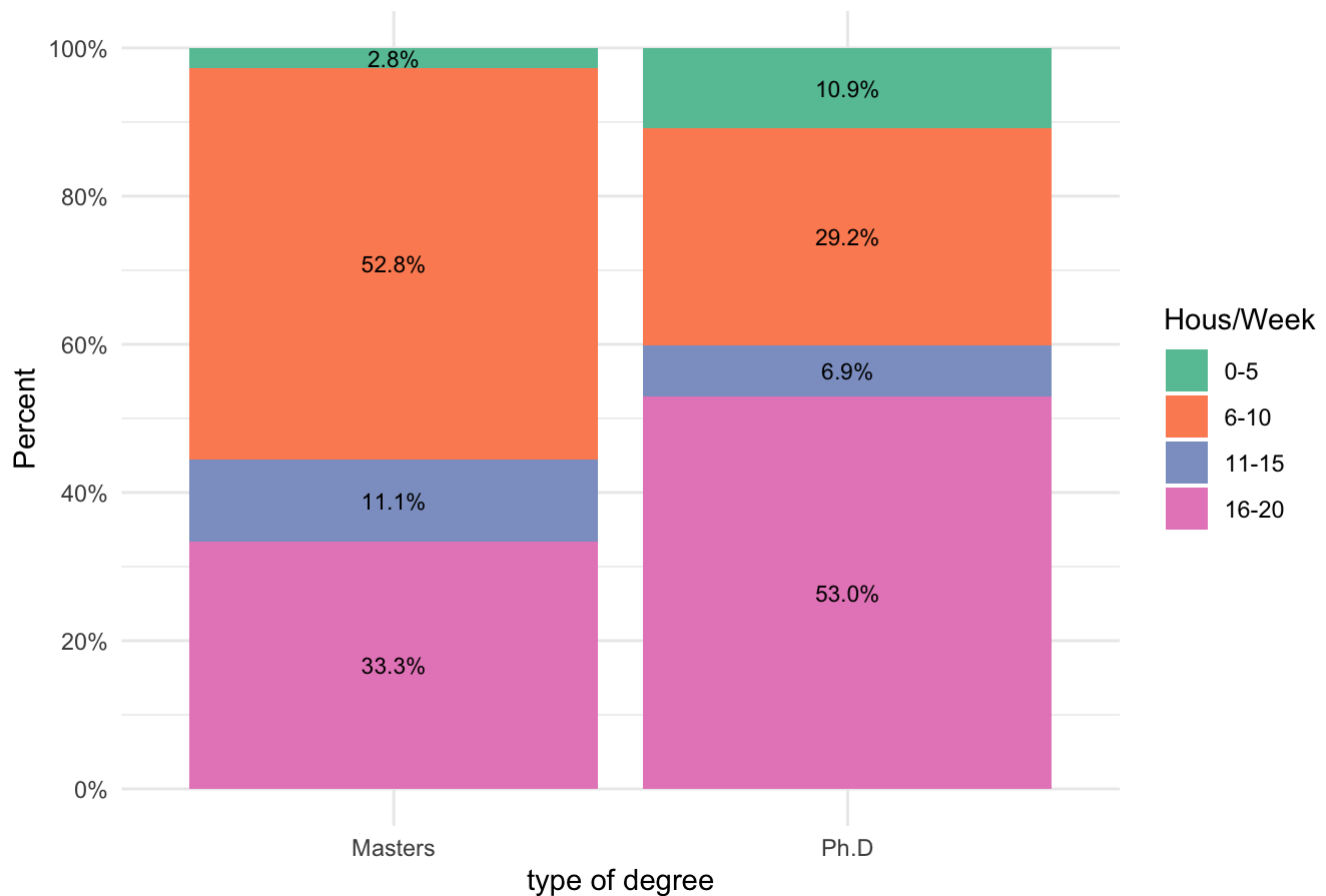
```
#Q3 What type of degree are you working towards?
#Q16 How many hours per week were you assigned for this TA assignment?
# create a summary dataset
plotdata <- df_ta %>%
  dplyr:: filter(!is.na(Q3) & !is.na(Q16))%>%
  dplyr:: filter(!Q3=="Other") %>% # only one data for other option
  group_by(Q3, Q16) %>%
  dplyr::summarize(n = n()) %>%
  dplyr::mutate(pct = n/sum(n),lbl = scales::percent(pct))
plotdata
```

```
## # A tibble: 8 x 5
## # Groups:   Q3 [2]
##   Q3      Q16      n    pct lbl
##   <fct>   <fct> <int> <dbl> <chr>
## 1 Masters 0-5      1 0.0278 2.8%
## 2 Masters 10-Jun     19 0.528 52.8%
## 3 Masters 15-Nov      4 0.111 11.1%
## 4 Masters 16-20     12 0.333 33.3%
## 5 Ph.D    0-5     22 0.109 10.9%
## 6 Ph.D    10-Jun     59 0.292 29.2%
## 7 Ph.D    15-Nov     14 0.0693 6.9%
## 8 Ph.D    16-20    107 0.530 53.0%
```

Master usually assign 6-10 hours per week, while ph.d often are assigned 16-20 hours per week.

```
ggplot(plotdata,
  aes(x = factor(Q3,
    levels = c("Masters", "Ph.D")),
    y = pct,
    fill = factor(Q16,
      levels = c("0-5", "10-Jun", "15-Nov", "16-20"),
      labels = c("0-5", "6-10", "11-15", "16-20")))) +
  geom_bar(stat = "identity",
    position = "fill") +
  scale_y_continuous(breaks = seq(0, 1, .2),
    label = percent) +
  geom_text(aes(label = lbl),
    size = 3,
    position = position_stack(vjust = 0.5)) +
  scale_fill_brewer(palette = "Set2") + labs(y = "Percent",
    fill = "Hous/Week",
    x = "type of degree",
    title = "Degree vs Hours per week for this TA assignment ") +
  theme_minimal()
```

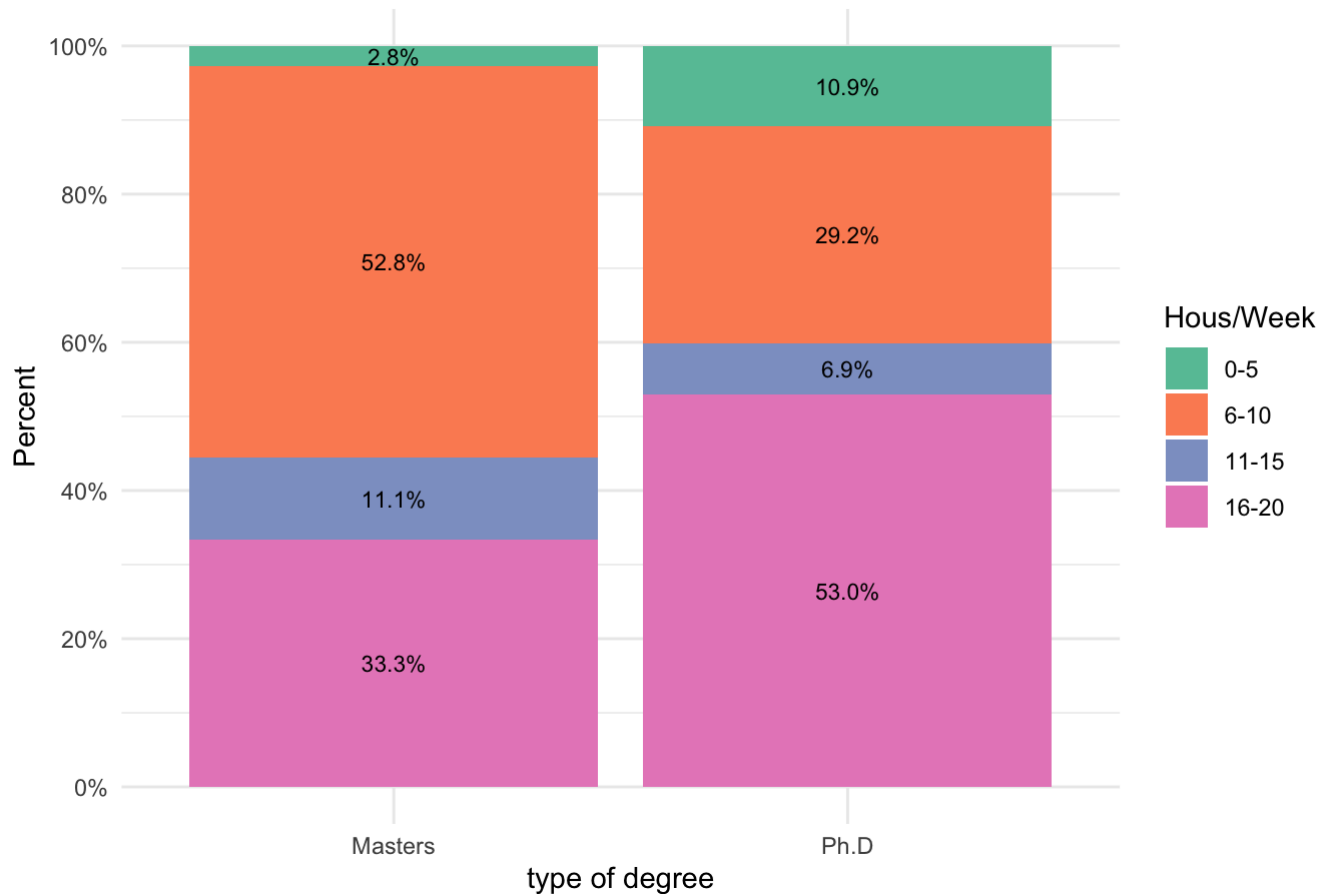
Degree vs Hours per week for this TA assignment



3 Feedback and performance related questions

```
ggplot(plotdata,
  aes(x = factor(Q3,
    levels = c("Masters", "Ph.D")),
    y = pct,
    fill = factor(Q16,
      levels = c("0-5", "10-Jun", "15-Nov", "16-20"),
      labels = c("0-5", "6-10", "11-15", "16-20")))) +
  geom_bar(stat = "identity",
    position = "fill") +
  scale_y_continuous(breaks = seq(0, 1, .2),
    label = percent) +
  geom_text(aes(label = lbl),
    size = 3,
    position = position_stack(vjust = 0.5)) +
  scale_fill_brewer(palette = "Set2") + labs(y = "Percent",
    fill = "Hous/Week",
    x = "type of degree",
    title = "Degree vs Hours per week for this TA assignment ") +
  theme_minimal()
```

Degree vs Hours per week for this TA assignment



```
#Q37 How often did you receive feedback about your performance from your TA supervisor or other faculty?
```

```
#Q38 How helpful was the performance feedback that you received?
```

```
# create a summary dataset
```

```
plotdata <- df_ta %>%
```

```
  dplyr:: filter(!is.na(Q37) & !is.na(Q38)) %>%
```

```
  dplyr:: filter(!Q3=="Other") %>% # only one data for other option
```

```
  group_by(Q37, Q38) %>%
```

```
  dplyr::summarize(n = n()) %>%
```

```
  dplyr::mutate(pct = n/sum(n), lbl = scales::percent(pct))
```

```
plotdata
```

```
## # A tibble: 13 x 5
```

```
## # Groups:   Q37 [4]
```

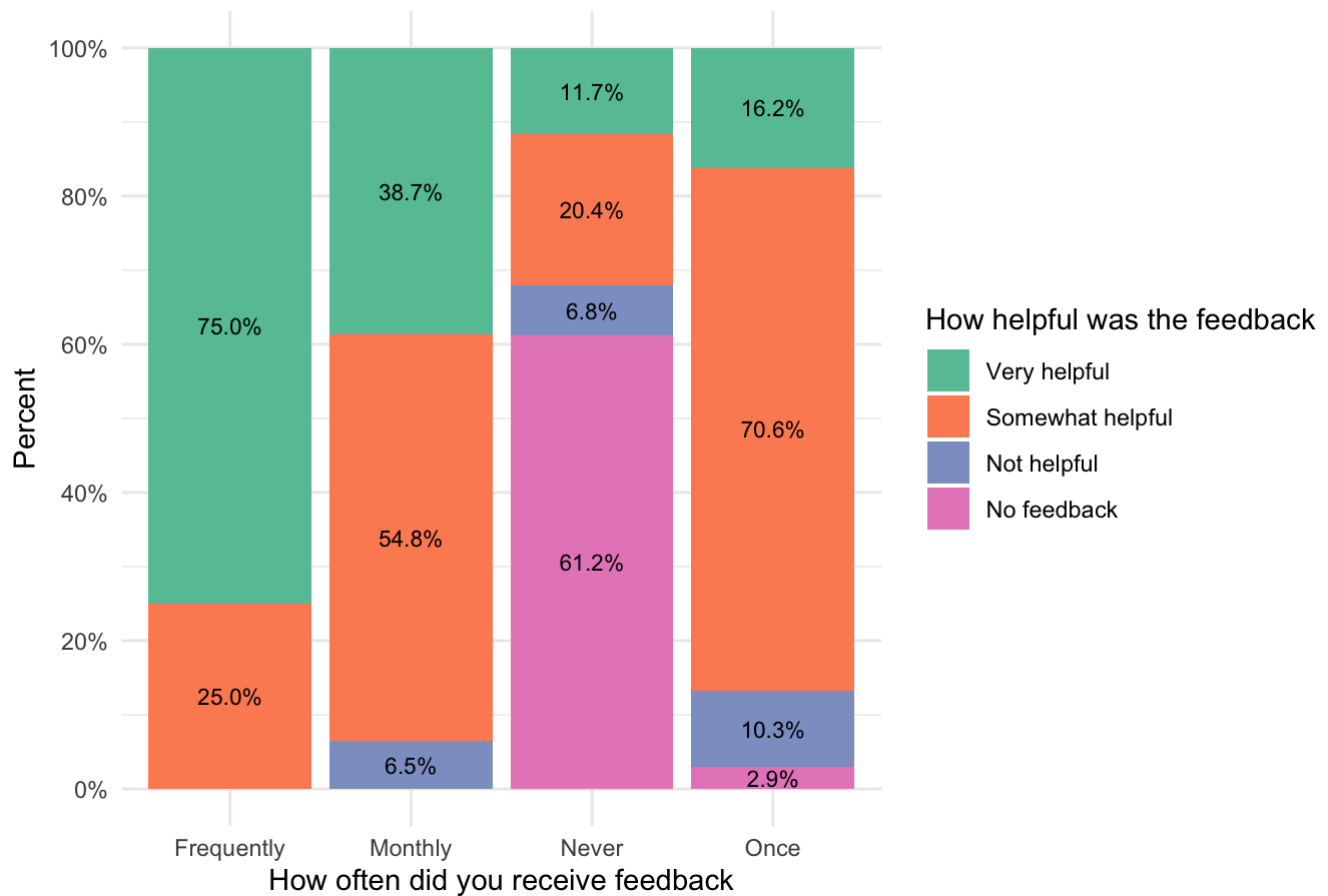
Q37	Q38	n	pct	lbl
<fct>	<fct>	<int>	<dbl>	<chr>
1 Monthly	Not helpful	2	0.0645	6.5%
2 Monthly	Somewhat helpful	17	0.548	54.8%
3 Monthly	Very helpful	12	0.387	38.7%
4 Never	I didn't receive feedback	63	0.612	61.2%
5 Never	Not helpful	7	0.0680	6.8%
6 Never	Somewhat helpful	21	0.204	20.4%
7 Never	Very helpful	12	0.117	11.7%
8 Once	I didn't receive feedback	2	0.0294	2.9%
9 Once	Not helpful	7	0.103	10.3%
10 Once	Somewhat helpful	48	0.706	70.6%
11 Once	Very helpful	11	0.162	16.2%
12 Weekly or more frequently	Somewhat helpful	7	0.25	25.0%
13 Weekly or more frequently	Very helpful	21	0.75	75.0%

3.1 Question: Does receiving more frequent feedback will help increase helpful feedback

from the plot ,we can conclude that receiving feedback about your performance frequently will increase positive(very help) feedback about the performance feedback.

```
ggplot(plotdata,
      aes(x = factor(Q37,
                    levels = c("Weekly or more frequently", "Monthly", "Never", "Once", "Never"),
                    labels=c("Frequently", "Monthly", "Never", "Once", "Never")),
      y = pct,
      fill = factor(Q38,
                    levels = c("Very helpful", "Somewhat helpful", "Not helpful", "I didn't receive feedback"),
                    labels = c("Very helpful", "Somewhat helpful", "Not helpful", "No feedback")))) +
  geom_bar(stat = "identity",
           position = "fill") +
  scale_y_continuous(breaks = seq(0, 1, .2),
                     label = percent) +
  geom_text(aes(label = lbl),
            size = 3,
            position = position_stack(vjust = 0.5)) +
  scale_fill_brewer(palette = "Set2") + labs(y = "Percent",
      fill = "How helpful was the feedback",
      x = "How often did you receive feedback",
      title = "How helpful was the feedback vs How often did you receive feedback ")
+
  theme_minimal()
```

How helpful was the feedback vs How often did you receive feedback

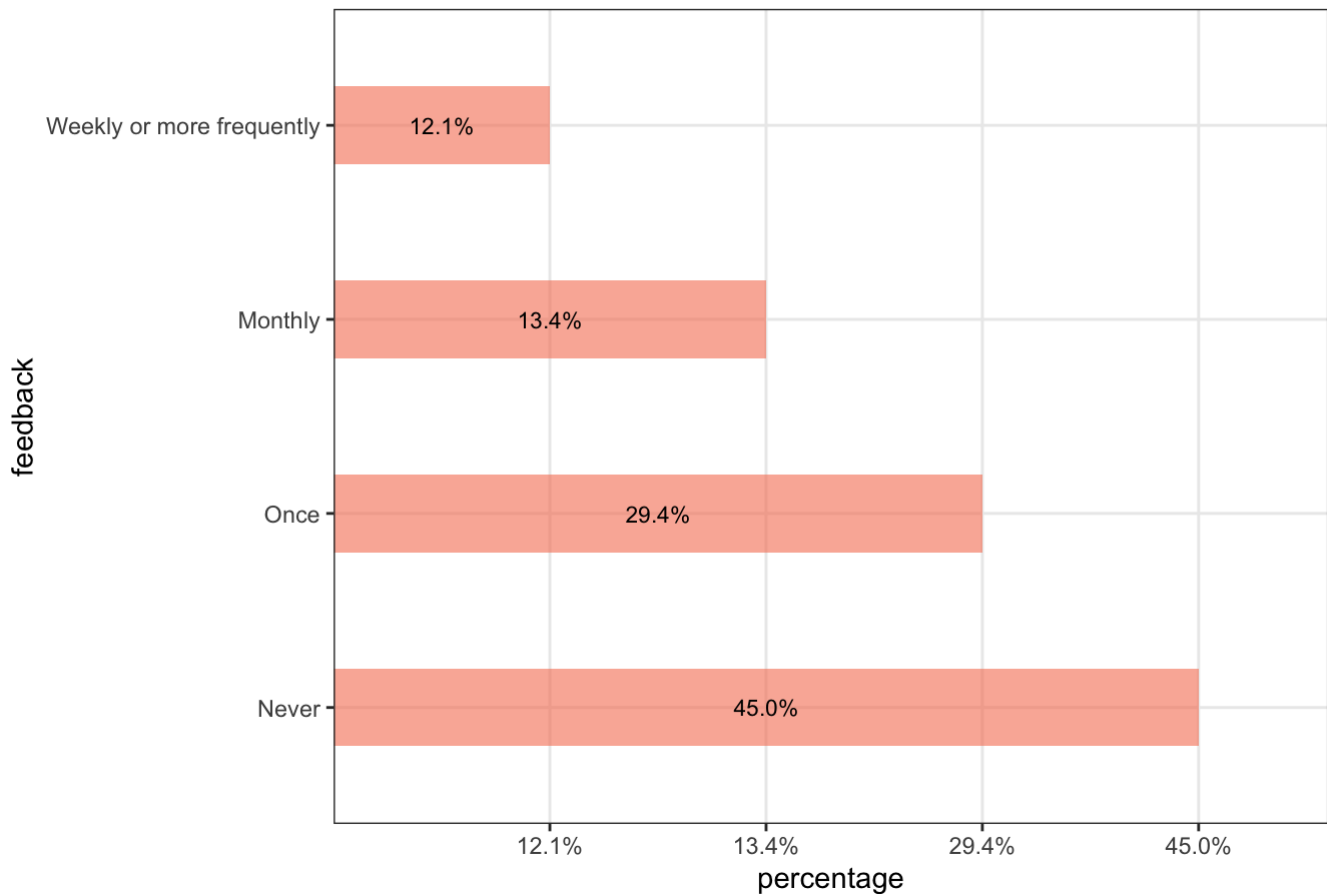


3.2 Most of TAs' performance shows they did not receive feedback about their performance

```
plotdata <- df_ta %>%
  dplyr::filter(!is.na(Q37)) %>% # remove NA
  group_by(Q37) %>%
  dplyr::summarize(n = n()) %>%
  dplyr::mutate(pct = n/sum(n), lbl = scales::percent(pct))
```

```
# Reverse side
plotdata %>%
  mutate(name = fct_reorder(Q37, desc(lbl))) %>%
  ggplot(aes(x=name, y=lbl)) +
    geom_bar(stat="identity", fill="#f68060", alpha=.6, width=.4) +
    coord_flip() +
    geom_text(aes(label = lbl,
                  size = 3,
                  position = position_stack(vjust = 0.5)) +
    xlab("feedback") +
    ylab("percentage")+
    labs(title="Q37 How often did you receive feedback about your performance")+
    theme_bw()
```

Q37 How often did you receive feedback about your performance

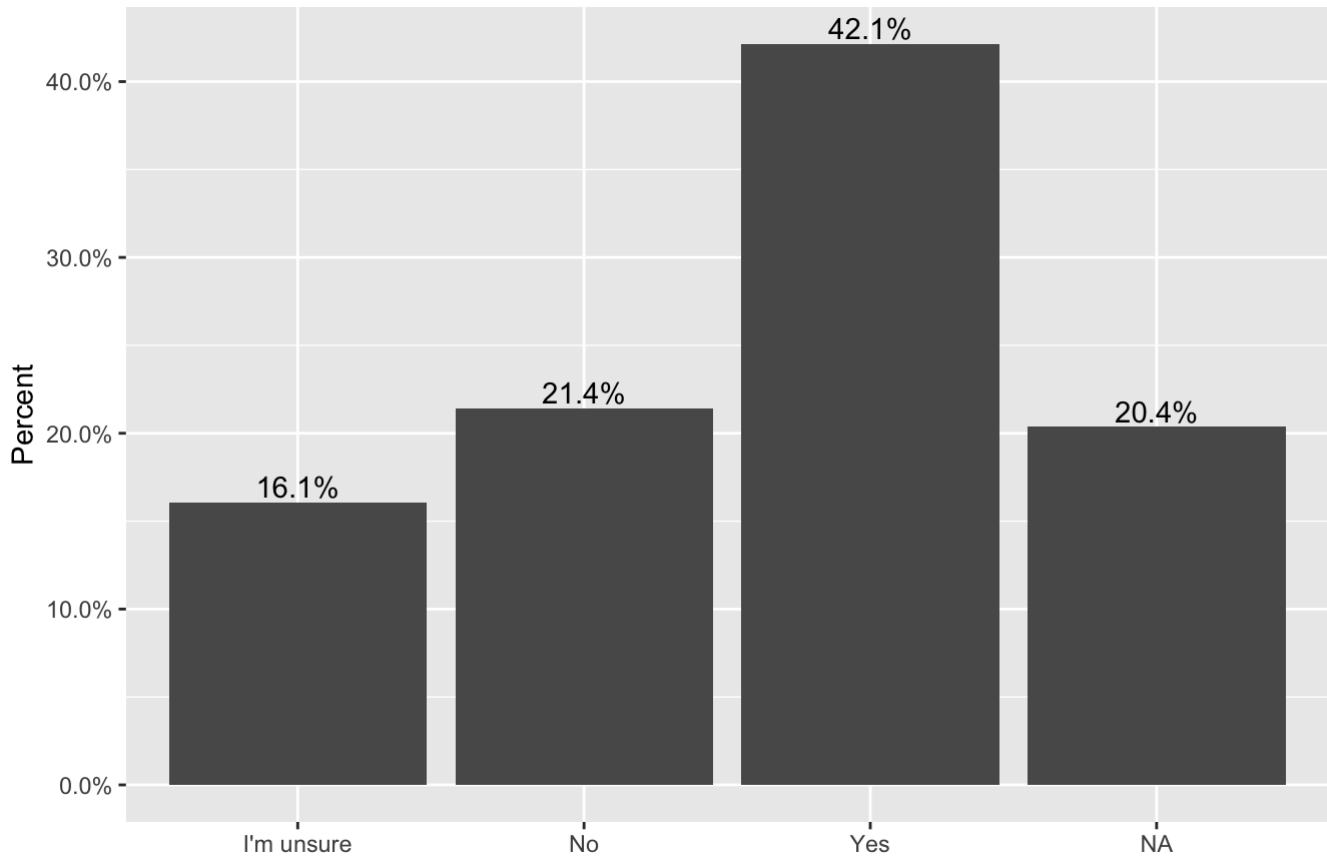


4. Training

4.1 21.4 % shows they did not receive adequate training.

```
#54
ggplot(df_ta, aes(x = as.factor(Q54))) +
  geom_bar(aes(y = (..count..)/sum(..count..))) +
  geom_text(aes(y = ((..count..)/sum(..count..)), label = scales::percent((..count..)/sum(..count..)), stat = "count", vjust = -0.25) +
  scale_y_continuous(labels = percent) +
  labs(title = "Q54 Did you receive adequate training for your most recent TA assignment?", y = "Percent", x = "")
```

Q54 Did you receive adequate training for your most recent TA assignment?

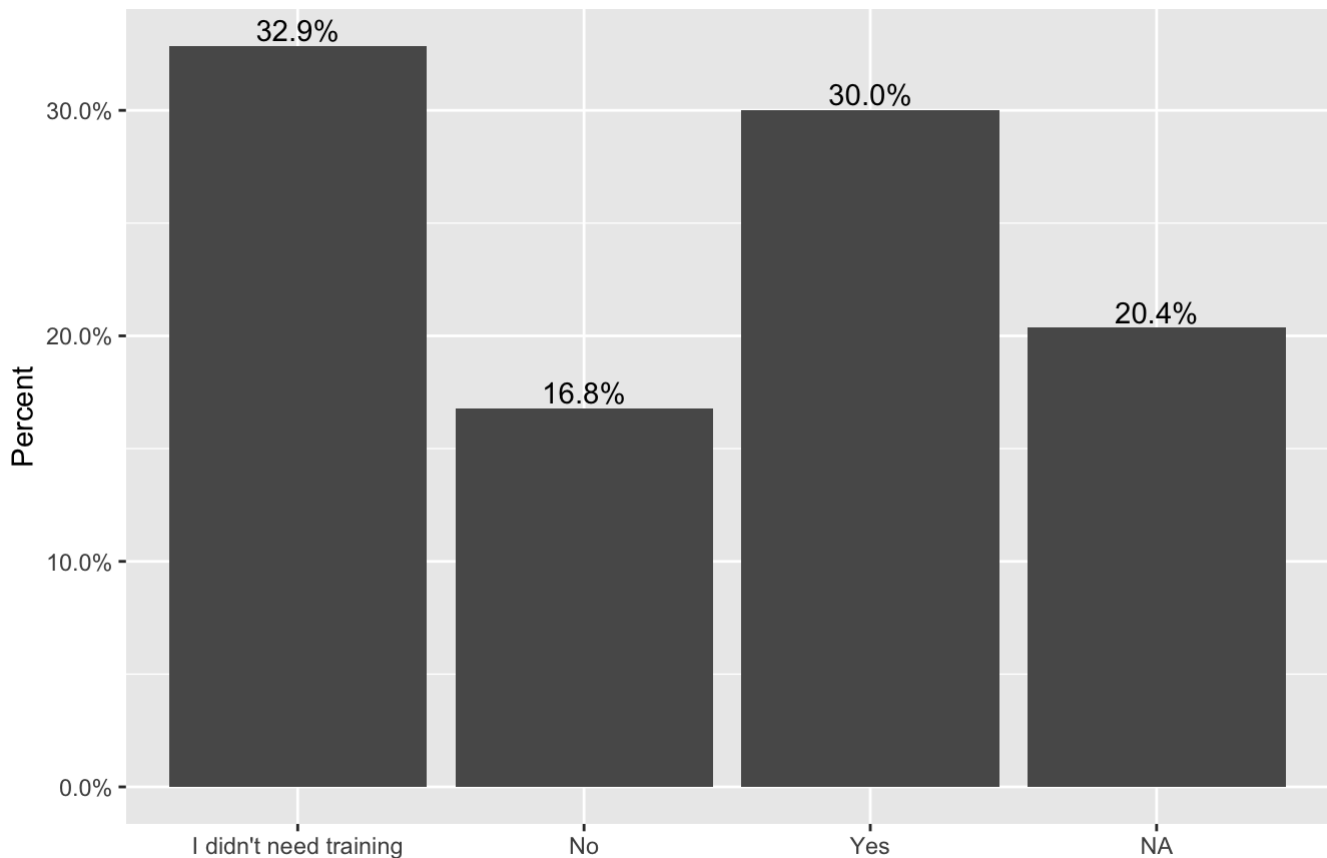


4.2 16.8 % TA need technology training to fulfill TA duties.

#64

```
ggplot(df_ta, aes(x = as.factor(Q64))) +
  geom_bar(aes(y = (..count..)/sum(..count..))) +
  geom_text(aes(y = ((..count..)/sum(..count..)), label = scales::percent((..count..)/sum(..count..)), stat = "count", vjust = -0.25) +
  scale_y_continuous(labels = percent) +
  labs(title = "Q64 Did you have training in any technology that you needed to fulfill TA duties ?", y = "Percent", x = "")
```


Q64 Did you have training in any technology that you needed to fulfill TA duties?

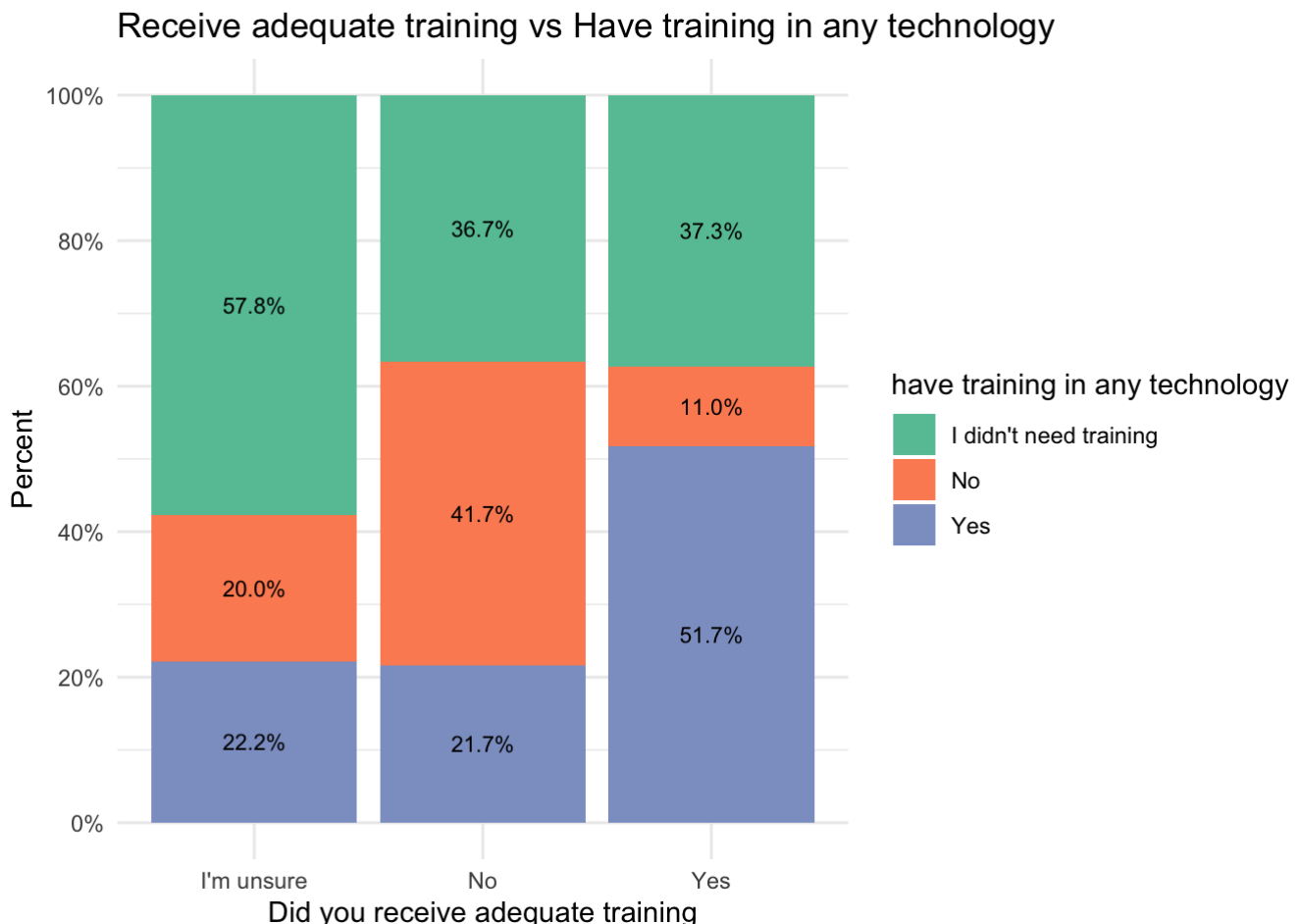


```
#Q54 Did you receive adequate training for your most recent TA assignment?
#Q64 Did you have training in any technology that you needed to fulfill your TA duties?
# create a summary dataset
plotdata <- df_ta %>%
  dplyr::filter(!is.na(Q54) & !is.na(Q64))%>%
  #dplyr::filter(!Q3=="Other") %>% # only one data for other option
  group_by(Q54, Q64) %>%
  dplyr::summarize(n = n()) %>%
  dplyr::mutate(pct = n/sum(n),lbl = scales::percent(pct))
plotdata
```

```
## # A tibble: 9 x 5
## # Groups:   Q54 [3]
##   Q54      Q64          n    pct lbl
##   <fct>    <fct>    <int> <dbl> <chr>
## 1 I'm unsure I didn't need training    26 0.578 57.8%
## 2 I'm unsure No                        9 0.2    20.0%
## 3 I'm unsure Yes                      10 0.222 22.2%
## 4 No      I didn't need training    22 0.367 36.7%
## 5 No      No                       25 0.417 41.7%
## 6 No      Yes                       13 0.217 21.7%
## 7 Yes     I didn't need training    44 0.373 37.3%
## 8 Yes     No                       13 0.110 11.0%
## 9 Yes     Yes                       61 0.517 51.7%
```

4.3 41.7% shows TA did not get enough training is technology, we may increase technology training.

```
ggplot(plotdata,
  aes(x = factor(Q54,
    levels = c("I'm unsure", "No", "Yes")),
    y = pct,
    fill = factor(Q64,
      levels = c("I didn't need training", "No", "Very helpful",
        "Yes"),
      labels = c("I didn't need training", "No", "Very helpful",
        "Yes")))) +
  geom_bar(stat = "identity",
    position = "fill") +
  scale_y_continuous(breaks = seq(0, 1, .2),
    label = percent) +
  geom_text(aes(label = lbl),
    size = 3,
    position = position_stack(vjust = 0.5)) +
  scale_fill_brewer(palette = "Set2") + labs(y = "Percent",
    fill = "have training in any technology",
    x = "Did you receive adequate training",
    title = "Receive adequate training vs Have training in any technology") +
  theme_minimal()
```



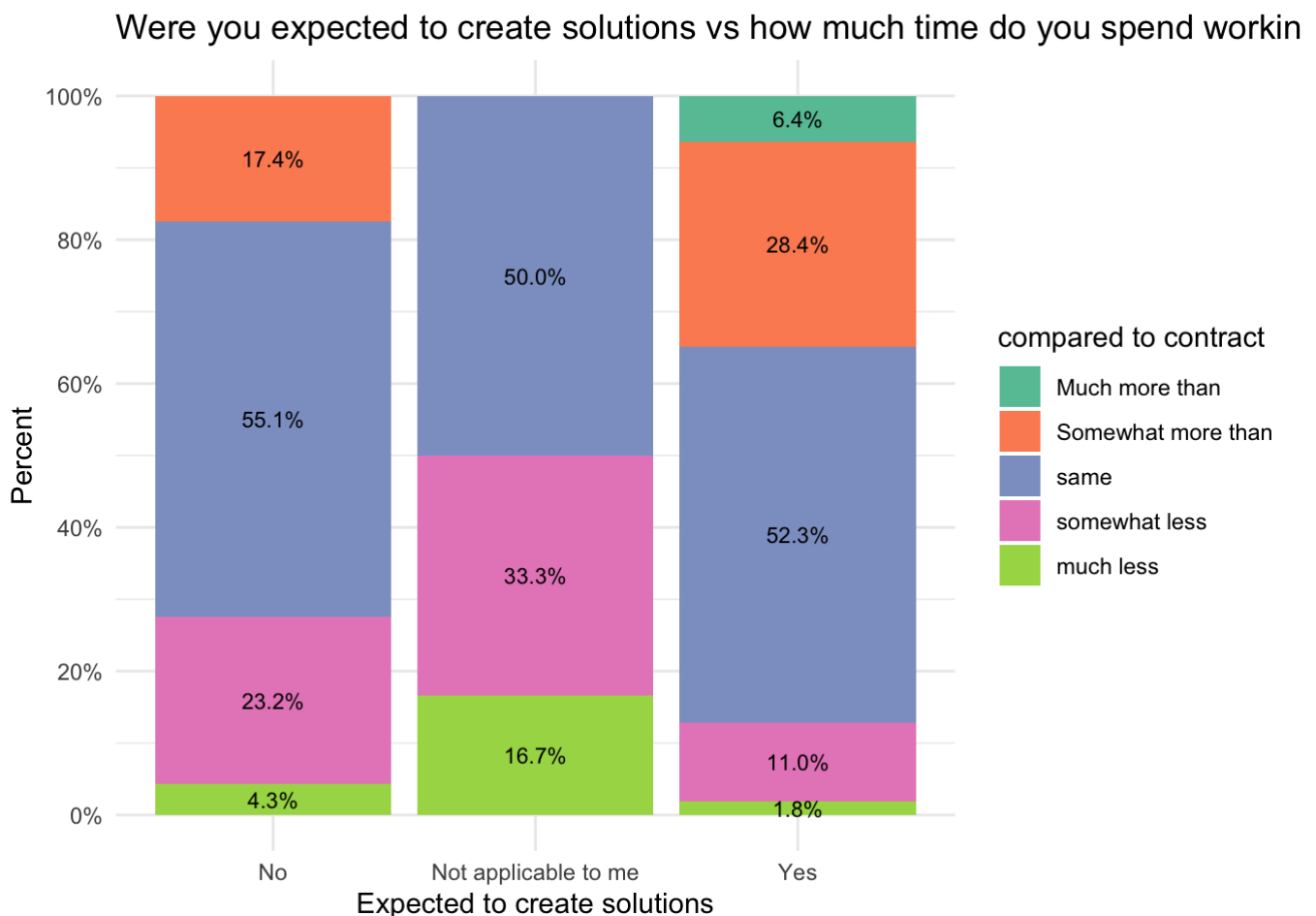
5. Provide solution effects the time spending for TA

```
##Q27 Were you expected to create solutions/define correct answers for the assignment
s you graded?
##Q33 On average, how much time do you spend working compared to the amount of time y
ou are contracted to work?
# create a summary dataset
plotdata <- df_ta %>%
  dplyr:: filter(!is.na(Q27) & !is.na(Q33))%>%
  #dplyr:: filter(!Q3=="Other") %>%    # only one data for other option
  group_by(Q27, Q33) %>%
  dplyr::summarize(n = n()) %>%
  dplyr::mutate(pct = n/sum(n),lbl = scales::percent(pct))
plotdata
```

```
## # A tibble: 12 x 5
## # Groups:   Q27 [3]
##   Q27          Q33          n    pct lbl
##   <fct>      <fct>    <int> <dbl> <chr>
## 1 No        About the same as I was contracted... 38 0.551 55.1%
## 2 No        Much less than I was contracted to...  3 0.0435 4.3%
## 3 No        Somewhat less than I was contracte... 16 0.232 23.2%
## 4 No        Somewhat more than I was contracte... 12 0.174 17.4%
## 5 Not applicable t... About the same as I was contracted...  3 0.5    50.0%
## 6 Not applicable t... Much less than I was contracted to...  1 0.167 16.7%
## 7 Not applicable t... Somewhat less than I was contracte...  2 0.333 33.3%
## 8 Yes       About the same as I was contracted... 57 0.523 52.3%
## 9 Yes       Much less than I was contracted to...  2 0.0183 1.8%
## 10 Yes      Much more than I was contracted to...  7 0.0642 6.4%
## 11 Yes      Somewhat less than I was contracte... 12 0.110 11.0%
## 12 Yes      Somewhat more than I was contracte... 31 0.284 28.4%
```

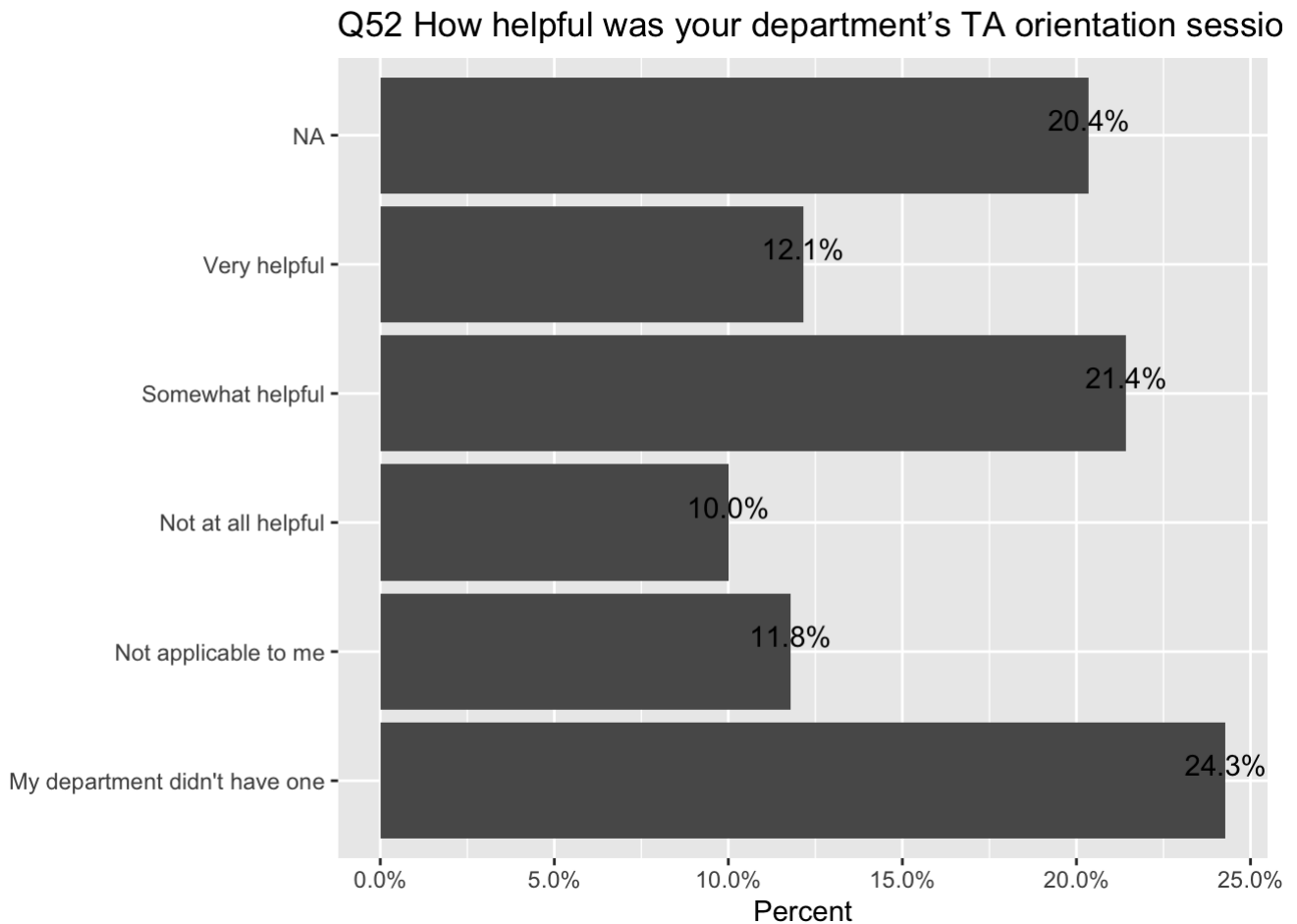
5.1 If require TA provide solution ,the percentage of higher of working time compared to the amount of time you are contracted than no solution require.

```
ggplot(plotdata,
  aes(x = factor(Q27,
    levels = c("No", "Not applicable to me", "Yes")),
    y = pct,
    fill = factor(Q33,
      levels = c("Much more than I was contracted to work", "Somewhat
at more than I was contracted to work", "About the same as I was contracted to work",
"Somewhat less than I was contracted to work", "Much less than I was contracted to wor
k")),
    labels = c("Much more than", "Somewhat more than", "same", "so
mewhat less", "much less")) +
  geom_bar(stat = "identity",
    position = "fill") +
  scale_y_continuous(breaks = seq(0, 1, .2),
    label = percent) +
  geom_text(aes(label = lbl,
    size = 3,
    position = position_stack(vjust = 0.5)) +
  scale_fill_brewer(palette = "Set2") + labs(y = "Percent",
    fill = "compared to contract",
    x = "Expected to create solutions",
    title = "Were you expected to create solutions vs how much time do you spend w
orking ") +
  theme_minimal()
```



5.2 question : TA orientation session is useful or not

```
#52 # result is not good enough
ggplot(df_ta, aes(x = as.factor(Q52))) +
  geom_bar(aes(y = (..count..)/sum(..count..))) +
  geom_text(aes(y = ((..count..)/sum(..count..)), label = scales::percent((..count..)/sum(..count..)), stat = "count", vjust = -0.25) +
  scale_y_continuous(labels = percent) +
  labs(title = "Q52 How helpful was your department's TA orientation session?", y = "Percent", x = "") + coord_flip()
```

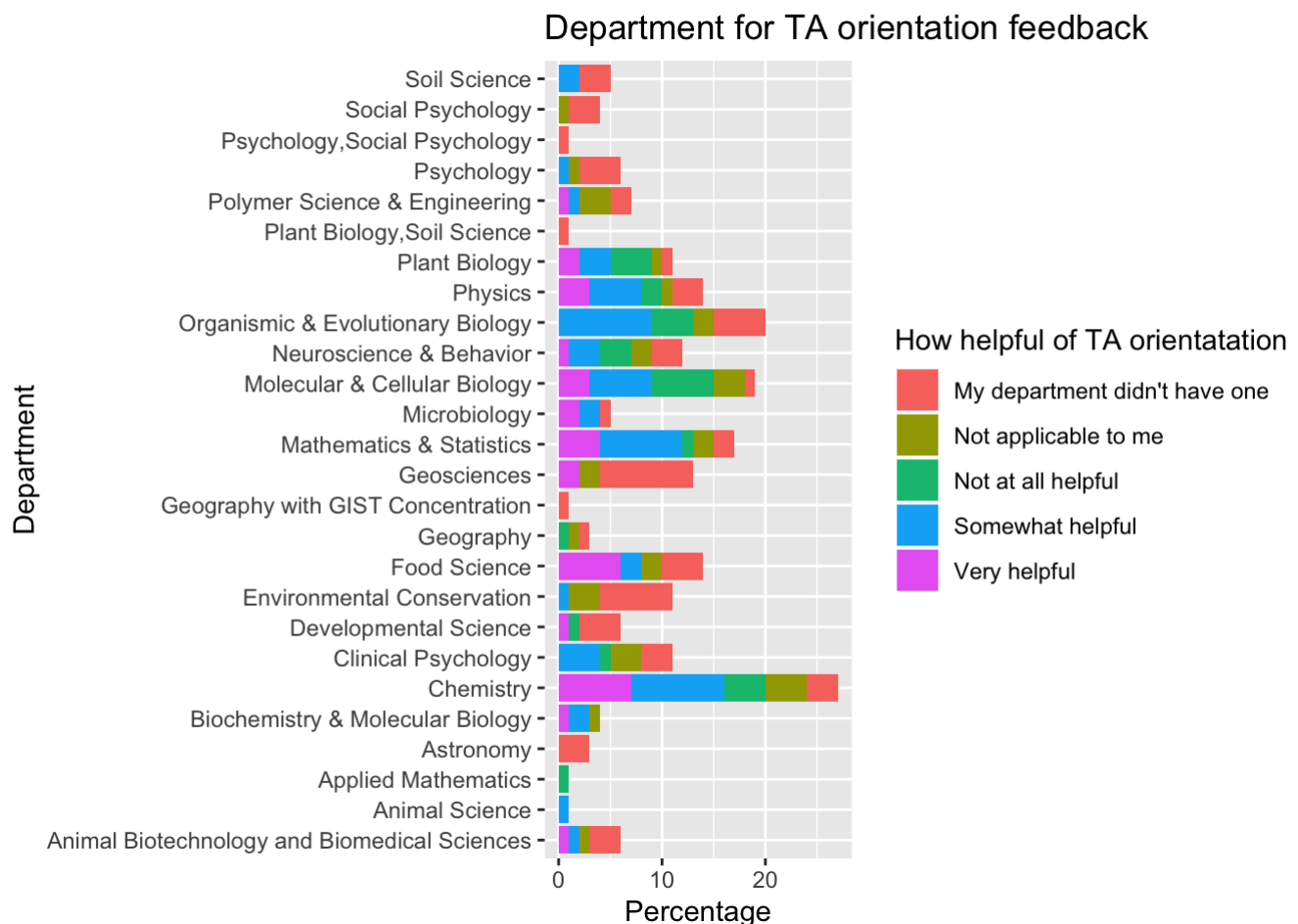


```
plotdata <- df_ta %>%
  dplyr:: filter(!is.na(Q5) & !is.na(Q52))
```

5.3 from the plot, we found not all of department has orientatation and positive feedback is not very higt.

```
#Q52 How helpful was your department's TA orientation session?
# rename the levels in Q5 Program
levels(plotdata$Q5)[levels(plotdata$Q5)=="Environmental Conservation,Organismic & Evo
lutionary Biology"] <-
  "Envir Cons,Organ & Evolu Biology"

# plot
ggplot(data = plotdata) +
  geom_bar(mapping = aes(x = Q5, fill = Q52)) +
  ggtitle("Department for TA orientation feedback") +
  labs(fill = "How helpful of TA orientatation")+
  xlab("Department") +
  ylab("Percentage") +
  coord_flip()
```



check the nubmer of different of feedback base on different department

```
#Q52 How helpful was your department's TA orientation session?
plotdata <- df_ta %>% dplyr:: filter(!is.na(Q52) )
library(reshape)
```

```
##
## Attaching package: 'reshape'
```

```
## The following objects are masked from 'package:tidyr':  
##  
##   expand, smiths
```

```
## The following object is masked from 'package:dplyr':  
##  
##   rename
```

```
cast(plotdata, Q5 ~ Q52) # Q5 departments
```

```
## Using Q74 as value column. Use the value argument to cast to override this choice
```

```
## Aggregation requires fun.aggregate: length used as default
```

```

## Q5
## 1 Animal Biotechnology and Biomedical Sciences
## 2 Animal Science
## 3 Applied Mathematics
## 4 Astronomy
## 5 Biochemistry & Molecular Biology
## 6 Chemistry
## 7 Clinical Psychology
## 8 Developmental Science
## 9 Environmental Conservation
## 10 Food Science
## 11 Geography
## 12 Geography with GIST Concentration
## 13 Geosciences
## 14 Mathematics & Statistics
## 15 Microbiology
## 16 Molecular & Cellular Biology
## 17 Neuroscience & Behavior
## 18 Organismic & Evolutionary Biology
## 19 Physics
## 20 Plant Biology
## 21 Plant Biology, Soil Science
## 22 Polymer Science & Engineering
## 23 Psychology
## 24 Psychology, Social Psychology
## 25 Social Psychology
## 26 Soil Science
## My department didn't have one Not applicable to me Not at all helpful
## 1 3 1 0
## 2 0 0 0
## 3 0 0 1
## 4 3 0 0
## 5 0 1 0
## 6 3 4 4
## 7 3 3 1
## 8 4 0 1
## 9 7 3 0
## 10 4 2 0
## 11 1 1 1
## 12 1 0 0
## 13 9 2 0
## 14 2 2 1
## 15 1 0 0
## 16 1 3 6
## 17 3 2 3
## 18 5 2 4
## 19 3 1 2
## 20 1 1 4
## 21 1 0 0
## 22 2 3 0
## 23 4 1 0
## 24 1 0 0
## 25 3 1 0
## 26 3 0 0
## Somewhat helpful Very helpful
## 1 1 1
## 2 1 0

```


## 3	0	0
## 4	0	0
## 5	2	1
## 6	9	7
## 7	4	0
## 8	0	1
## 9	1	0
## 10	2	6
## 11	0	0
## 12	0	0
## 13	0	2
## 14	8	4
## 15	2	2
## 16	6	3
## 17	3	1
## 18	9	0
## 19	5	3
## 20	3	2
## 21	0	0
## 22	1	1
## 23	1	0
## 24	0	0
## 25	0	0
## 26	2	0

5.4

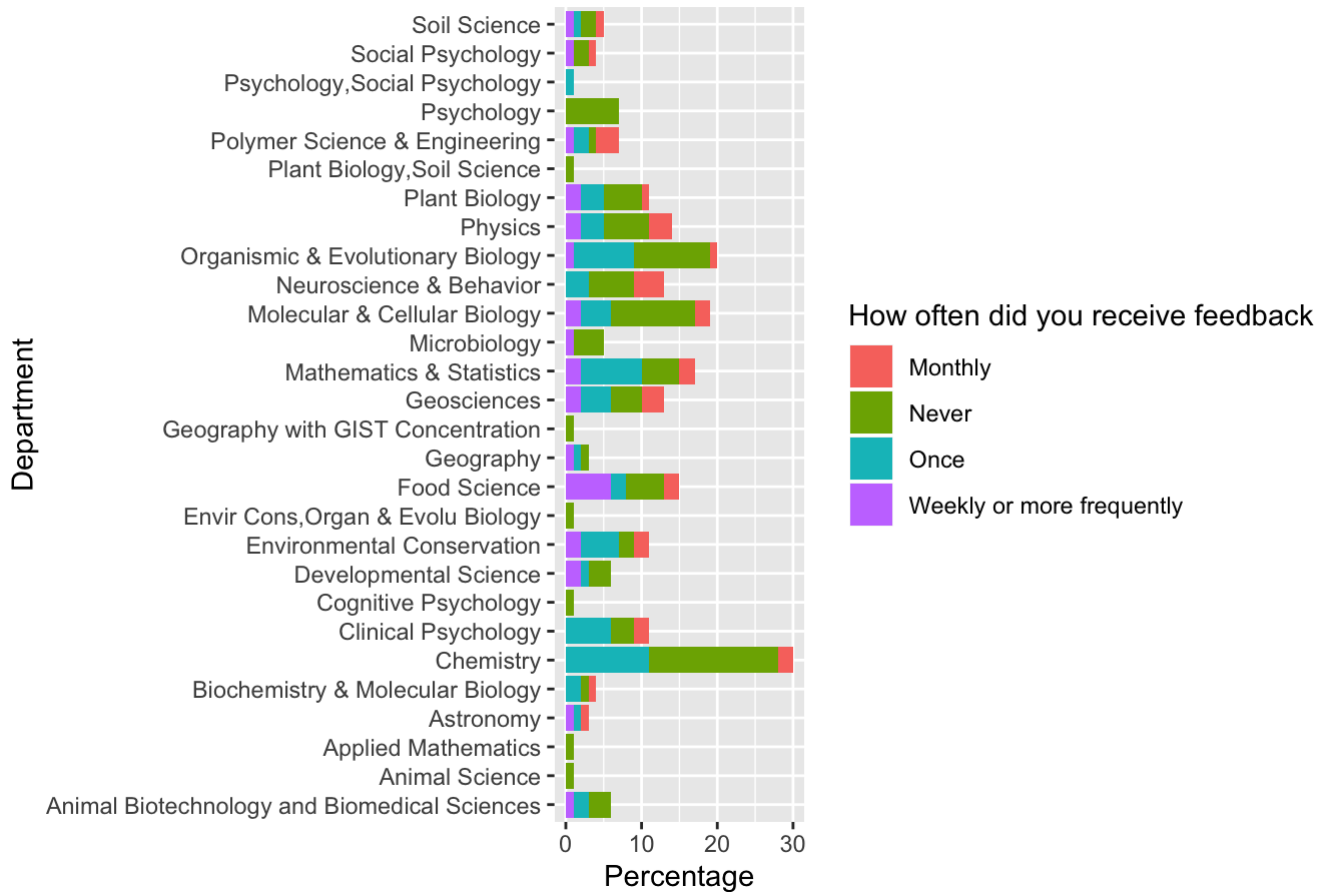
Most of departments has a higher percentage of “Never” feedback of TA’s performance.

```
plotdata <- df_ta %>%
  dplyr:: filter(!is.na(Q5) & !is.na(Q37))
```

```
##Q37 How often did you receive feedback about your performance from your TA supervis
or or other faculty?
# rename the levels in Q5 Program
levels(plotdata$Q5)[levels(plotdata$Q5)=="Environmental Conservation,Organismic & Evo
lutionary Biology"] <-
  "Envir Cons,Organ & Evolu Biology"

# plot
ggplot(data = plotdata) +
  geom_bar(mapping = aes(x = Q5, fill = Q37)) +
  ggtitle("Departments for how often TA receives feedback") +
  labs(fill = "How often did you receive feedback")+
  xlab("Department") +
  ylab("Percentage") +
  coord_flip()
```

Departments for how often TA receives feedback



```
#library(tigerstats)
#favstats(Q5 ~ Q37, data = df_ta)
```

```
#Q37 How often did you receive feedback about your performance from your TA supervisor or other faculty?
library(reshape)
cast(df_ta, Q5 ~ Q37) # Q5 departments
```

```
## Using Q74 as value column. Use the value argument to cast to override this choice
```

```
## Aggregation requires fun.aggregate: length used as default
```

##	Q5 Monthly			
## 1	Animal Biotechnology and Biomedical Sciences			0
## 2	Animal Science			0
## 3	Applied Mathematics			0
## 4	Applied Molecular Biotechnology			0
## 5	Astronomy			1
## 6	Biochemistry & Molecular Biology			1
## 7	Chemistry			2
## 8	Clinical Psychology			2
## 9	Cognitive Psychology			0
## 10	Developmental Science			0
## 11	Environmental Conservation			2
## 12	Environmental Conservation,Organismic & Evolutionary Biology			0
## 13	Food Science			2
## 14	Geography			0
## 15	Geography with GIST Concentration			0
## 16	Geosciences			3
## 17	Mathematics & Statistics			2
## 18	Microbiology			0
## 19	Molecular & Cellular Biology			2
## 20	Neuroscience & Behavior			4
## 21	Organismic & Evolutionary Biology			1
## 22	Physics			3
## 23	Plant Biology			1
## 24	Plant Biology,Soil Science			0
## 25	Polymer Science & Engineering			3
## 26	Psychology			0
## 27	Psychology,Social Psychology			0
## 28	Social Psychology			1
## 29	Soil Science			1
## 30	<NA>			0
##	Never Once Weekly or more frequently NA			
## 1	3	2	1	2
## 2	1	0	0	0
## 3	1	0	0	1
## 4	0	0	0	2
## 5	0	1	1	1
## 6	1	2	0	0
## 7	17	11	0	6
## 8	3	6	0	0
## 9	1	0	0	1
## 10	3	1	2	0
## 11	2	5	2	2
## 12	1	0	0	0
## 13	5	2	6	2
## 14	1	1	1	1
## 15	1	0	0	0
## 16	4	4	2	3
## 17	5	8	2	3
## 18	4	0	1	1
## 19	11	4	2	0
## 20	6	3	0	0
## 21	10	8	1	3
## 22	6	3	2	3
## 23	5	3	2	4
## 24	1	0	0	0
## 25	1	2	1	3

```
## 26      7      0      0      1
## 27      0      1      0      0
## 28      2      0      1      1
## 29      2      1      1      2
## 30      0      0      0      7
```

6. multiple choices one variable plot

```
#load the data and replace blank answers with NA
data<-read.csv("/Users/clarawang/Dropbox/Survey-TA/Raw data_May 13_numeric response.csv",na.strings=c("", "NA"))
library(ggplot2)
library(grid)
TA_data <- data[-c(1,2),]
TA_data<-subset(TA_data,Q1 == '1') # remove not TA data
dim(TA_data)
```

```
## [1] 280 124
```

```
#remove NA
sum(is.na(TA_data$Q56))
```

```
## [1] 57
```

```
df<-TA_data
choice <-c(1:7)
counts<-rep(0,7)
for (i in choice) {
  counts[i]=sum(grepl(i,df$Q56)==TRUE)
}
Q56_table<-data.frame(choice,counts)
Q56_table
```

```
##      choice counts
## 1         1     158
## 2         2     186
## 3         3      33
## 4         4      73
## 5         5      17
## 6         6      16
## 7         7      23
```

```
Q56_table$percent = Q56_table$counts / (280-57)
Q56_table
```

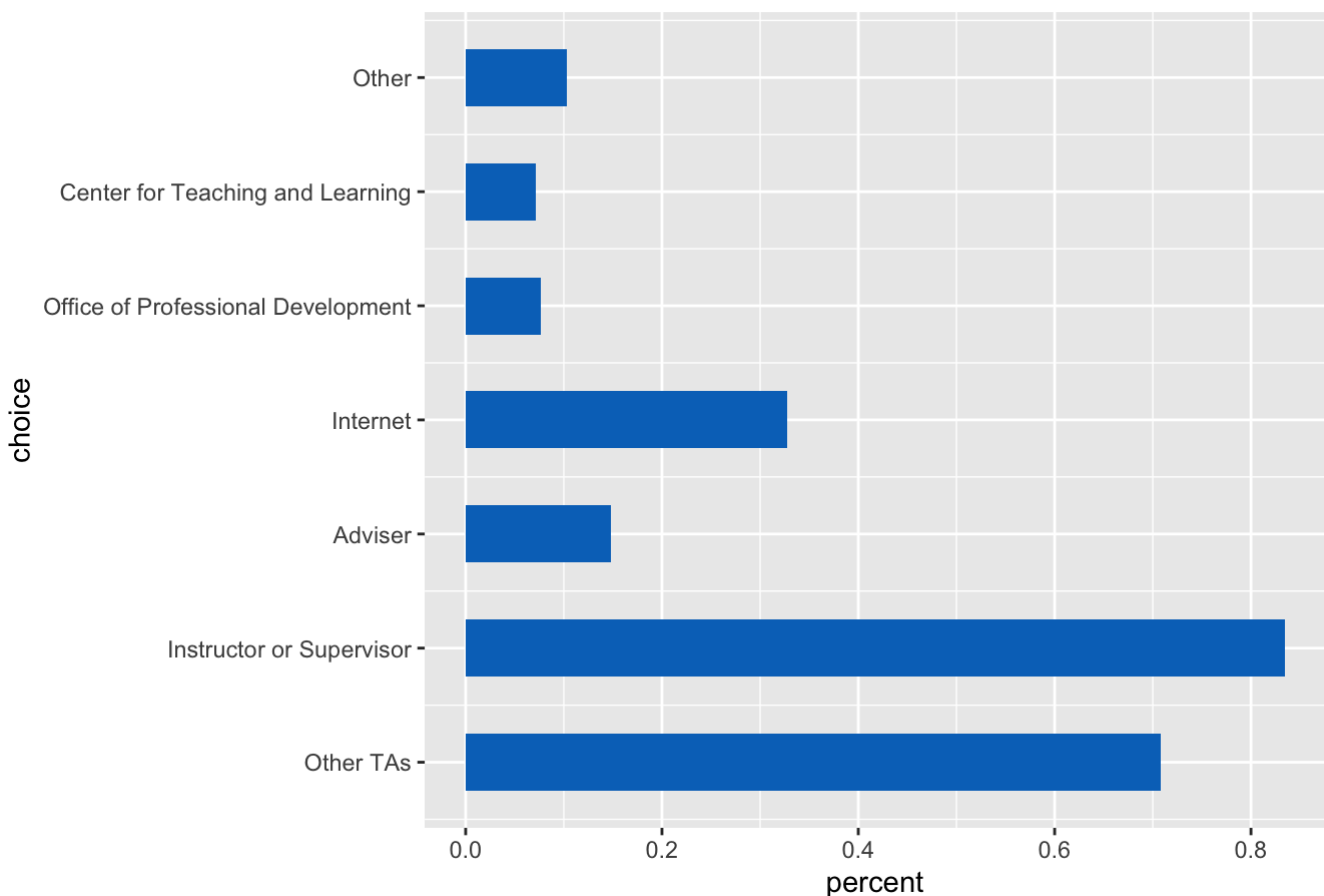
##	choice	counts	percent
## 1	1	158	0.70852018
## 2	2	186	0.83408072
## 3	3	33	0.14798206
## 4	4	73	0.32735426
## 5	5	17	0.07623318
## 6	6	16	0.07174888
## 7	7	23	0.10313901

6.1

```
plot56 <- ggplot(data=Q56_table, aes(x=choice, y=percent))
plot56 <- plot56 + geom_bar(fill = "#0073C2FF", stat = "identity", width=0.5)
plot56 <- plot56 + scale_x_continuous(breaks=1:7,
                                     labels=c("Other TAs", "Instructor or Sup
ervisor", "Adviser", "Internet", "Office of Professional Development", "Center for Teachi
ng and Learning", "Other"))+
  coord_flip()+

  ggtitle("Q56 Which of the following resources were helpful to you?")
plot56
```

Q56 Which of the following resources were helpful to you?



```
#remove NA
sum(is.na(TA_data$Q71))
```

```
## [1] 58
```

```
df<-TA_data
choice <-c(1:8)
counts<-rep(1,8)
for (i in choice) {
  counts[i]=sum(grepl(i,df$Q71)==TRUE)
}
Q71_table<-data.frame(choice,counts)
Q71_table
```

```
##   choice counts
## 1      1      4
## 2      2      9
## 3      3     64
## 4      4    140
## 5      5      2
## 6      6      4
## 7      7     10
## 8      8      5
```

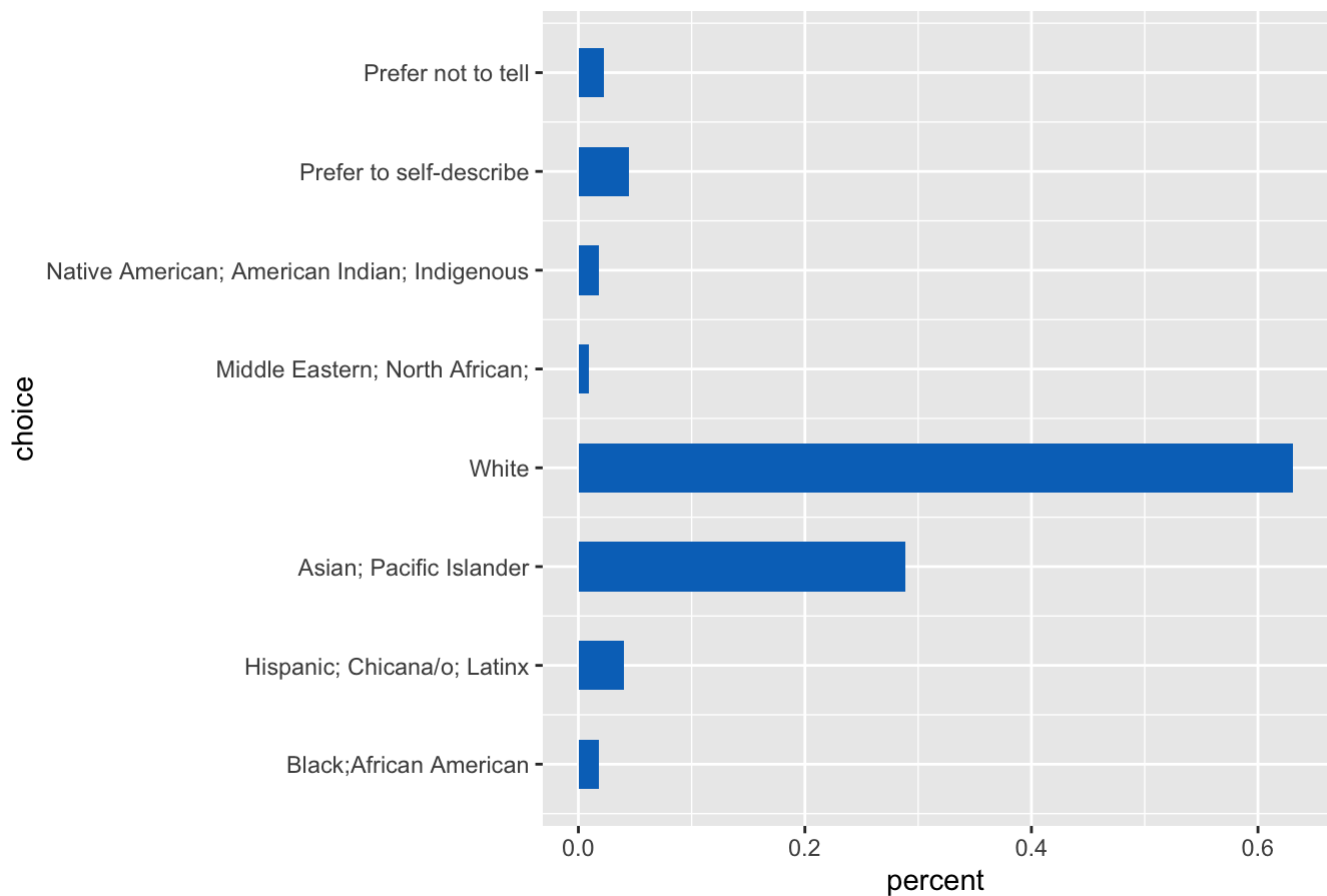
```
Q71_table$percent = Q71_table$counts / (280-58)
Q71_table
```

```
##   choice counts    percent
## 1      1      4 0.018018018
## 2      2      9 0.040540541
## 3      3     64 0.288288288
## 4      4    140 0.630630631
## 5      5      2 0.009009009
## 6      6      4 0.018018018
## 7      7     10 0.045045045
## 8      8      5 0.022522523
```

6.2

```
plot71 <- ggplot(data=Q71_table, aes(x=choice, y=percent))
plot71 <- plot71 + geom_bar(fill = "#0073C2FF",stat = "identity",width=0.5)
plot71 <- plot71 + scale_x_continuous(breaks=1:8,
                                     labels=c("Black;African American","Hispanic; Chicana/o; Latinx","Asian; Pacific Islander","White","Middle Eastern; North African","Native American; American Indian; Indigenous","Prefer to self-describe","Prefer not to tell"))+
  coord_flip()+
  ggtitle("Q71 What is your race/ethnicity?")
plot71
```

Q71 What is your race/ethnicity?



```
#remove NA
sum(is.na(TA_data$Q72))
```

```
## [1] 59
```

```
df<-TA_data
choice <-c(1:5)
counts<-rep(1,5)
for (i in choice) {
  counts[i]=sum(grepl(i,df$Q72)==TRUE)
}
Q72_table<-data.frame(choice,counts)
Q72_table
```

```
##   choice counts
## 1      1     127
## 2      2      86
## 3      3       3
## 4      4       1
## 5      5       5
```

```
Q72_table$percent = Q72_table$counts / (280-59)
Q72_table
```

##	choice	counts	percent
## 1	1	127	0.574660633
## 2	2	86	0.389140271
## 3	3	3	0.013574661
## 4	4	1	0.004524887
## 5	5	5	0.022624434

6.3

```
plot72 <- ggplot(data=Q72_table, aes(x=choice, y=percent))
plot72 <- plot72 + geom_bar(fill = "#0073C2FF",stat = "identity",width=0.5)
plot72 <- plot72 + scale_x_continuous(breaks=1:5,
                                     labels=c("Woman", "Man", "Non-binary; Gen
der non-conforming", "Prefer to self-describe ", "Prefer not to tell"))+
  coord_flip()+
  ggtitle("Q72 What is your gender?")
plot72
```

