R Notebook

```
library(readr)
library(tidyverse)
```

```
## -- Attaching packages --------- tidyverse 1.2.1 --
```

```
## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v ggplot2 3.2.1     v forcats 0.4.0
```

```
## -- Conflicts ------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
library(dplyr)
library(data.table)
```

```
##
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```

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

```
data<-read_csv("~/Desktop/mental_health.csv")
```

```
## Warning: Missing column names filled in: 'X1' [1]
```

```
## Parsed with column specification:
## cols(
##   .default = col_character(),
##   X1 = col_double(),
##   self_empl_flag = col_double(),
##   tech_comp_flag = col_double(),
##   tech_role_flag = col_double(),
##   prev_employers_flag = col_double(),
##   age = col_double(),
##   sex = col_double()
## )
```

```
## See spec(...) for full column specifications.
```

```
## Warning: 608 parsing failures.
##  row            col expected actual                          file
## 2607 tech_comp_flag a double   True '~/Desktop/mental_health.csv'
## 2607 tech_role_flag a double   True '~/Desktop/mental_health.csv'
## 2608 tech_comp_flag a double   True '~/Desktop/mental_health.csv'
## 2608 tech_role_flag a double   True '~/Desktop/mental_health.csv'
## 2609 tech_comp_flag a double   True '~/Desktop/mental_health.csv'
## .... .............. ........ ...... ..............................
## See problems(...) for more details.
```

```
data<-data[,2:29]
```

```
names(data)
```

```
##  [1] "self_empl_flag"                    "comp_no_empl"
##  [3] "tech_comp_flag"                    "tech_role_flag"
##  [5] "mh_employer_discussion"            "mh_anonimity_flag"
##  [7] "mh_diagnosed&reveal_clients_flag" "mh_diagnosed&reveal_cowork_flag"
##  [9] "mh_prod_impact"                    "mh_prod_impact_perc"
## [11] "prev_employers_flag"               "prev_mh_benefits"
## [13] "prev_mh_benefits_awareness"        "prev_mh_discussion"
## [15] "prev_mh_anonimity"                 "future_ph_specification"
## [17] "why/why_not"                       "mh_sharing_friends/fam_flag"
## [19] "mh_bad_response_workplace"         "mh_family_hist"
## [21] "mh_disorder_past"                  "mh_disorder_current"
## [23] "age"                               "sex"
## [25] "country_live"                      "live_us_teritory"
## [27] "country_work"                      "work_us_teritory"
```

```
#focus columns 10-14
data2<-data[, 10:14]
```

```
head(data2)
```

```
## # A tibble: 6 x 5
##   mh_prod_impact_~ prev_employers_~ prev_mh_benefits prev_mh_benefit~
##   <chr>                       <dbl> <chr>            <chr>
## 1 <NA>                            1 No, none did     N/A (not curren~
## 2 <NA>                            1 Yes, they all d~ I was aware of ~
## 3 <NA>                            1 No, none did     N/A (not curren~
## 4 1-25%                           1 Some did         N/A (not curren~
## 5 <NA>                            1 I don't know     N/A (not curren~
## 6 <NA>                            1 No, none did     Yes, I was awar~
## # ... with 1 more variable: prev_mh_discussion <chr>
```

```
skimr::skim(data2)
```

```
## Skim summary statistics
##  n obs: 2958
##  n variables: 5
##
```

```
## -- Variable type:character -----------------------
##                    variable missing complete    n min max empty n_unique
##         mh_prod_impact_perc    2591      367 2958   5   7     0        4
##            prev_mh_benefits     368     2590 2958   8  17     0        4
##  prev_mh_benefits_awareness     368     2590 2958  18  31     0        6
##          prev_mh_discussion     368     2590 2958   8  17     0        4
##
## -- Variable type:numeric -------------------------
##             variable missing complete    n mean   sd p0 p25 p50 p75 p100
##  prev_employers_flag       0     2958 2958 0.88 0.33  0   1   1   1    1
##      hist
##
```

#11

```
##personal advice, delete the column mh_prod_impact_perc
#'If yes, what percentage of your work time (time performing primary or secondary job functions) is aff
#mh_prod_impact_perc
x<-table(data2$mh_prod_impact_perc)
x
```

```
##
##   1-25%  26-50%  51-75% 76-100%
##     164     125      53      25
```

```
str(x)
```

```
##  'table' int [1:4(1d)] 164 125 53 25
##  - attr(*, "dimnames")=List of 1
##   ..$ : chr [1:4] "1-25%" "26-50%" "51-75%" "76-100%"
```

```
mh_impact_count<-as.vector(x)
mh_impact_percentage<-names(x)
mh_impact_count
```

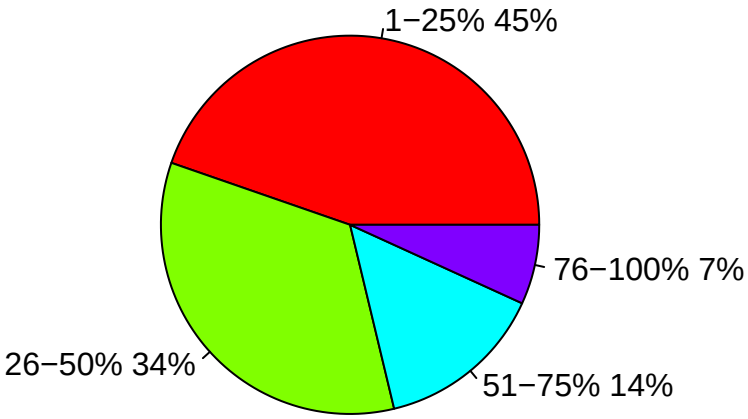
```
## [1] 164 125  53  25
```

```
mh_impact_percentage
```

```
## [1] "1-25%"   "26-50%"  "51-75%"  "76-100%"
```

```
pct <- round(mh_impact_count/sum(mh_impact_count)*100)
mh_impact_percentage <- paste(mh_impact_percentage, pct) # add percents to labels
mh_impact_percentage <- paste(mh_impact_percentage,"%",sep="") # ad % to labels
pie(mh_impact_count,labels = mh_impact_percentage, col=rainbow(length(mh_impact_percentage)),
   main="Pie Chart of Mental Health Impact")
```

Pie Chart of Mental Health Impact



#12

```
#'Do you have previous employers?'
#prev_employers_flag
#no missing value
y=table(data2$prev_employers_flag)
y
```

```
##
##    0    1
##  368 2590
```

```
str(y)
```

```
##  'table' int [1:2(1d)] 368 2590
##  - attr(*, "dimnames")=List of 1
##   ..$ : chr [1:2] "0" "1"
```

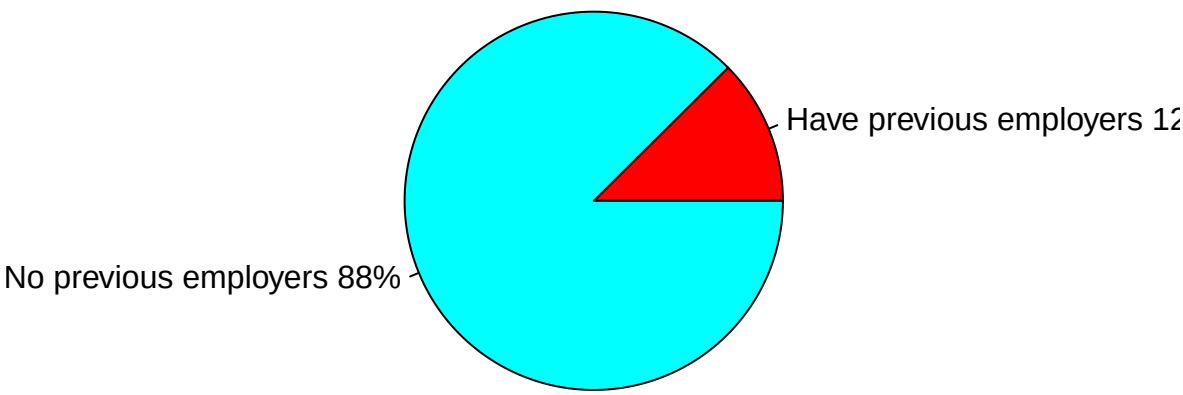
```
pre_employer_count<-as.vector(y)
pre_employer_percentage<-c("Have previous employers", "No previous employers")
pre_employer_count
```

```
## [1]  368 2590
```

```
pre_employer_percentage
```

```
## [1] "Have previous employers" "No previous employers"
```

```
pct2 <- round(pre_employer_count/sum(pre_employer_count)*100)
pre_employer_percentage <- paste(pre_employer_percentage, pct2) # add percents to labels
pre_employer_percentage <- paste(pre_employer_percentage,"%",sep="") # ad % to labels
pie(pre_employer_count,labels = pre_employer_percentage, col=rainbow(length(pre_employer_percentage)),
   main="Pie Chart of Previous Employers Flag")
```



#13

```
#'Have your previous employers provided mental health benefits?',
#prev_mh_benefits
z<-table(data2$prev_mh_benefits)
z
```

```
##
##      I don't know      No, none did          Some did Yes, they all did
##               619               700               852               419
```

```
str(z)
```

```
##  'table' int [1:4(1d)] 619 700 852 419
##  - attr(*, "dimnames")=List of 1
##   ..$ : chr [1:4] "I don't know" "No, none did" "Some did" "Yes, they all did"
```

```
prev_mh_benefits_count<-as.vector(z)
prev_mh_benefits_percentage<-names(z)
prev_mh_benefits_count
```

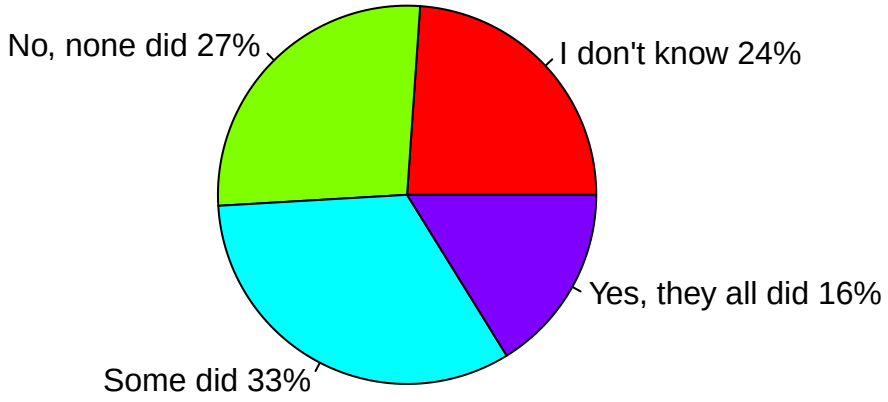
```
## [1] 619 700 852 419
```

```
prev_mh_benefits_percentage
```

```
## [1] "I don't know"      "No, none did"      "Some did"
## [4] "Yes, they all did"
```

```
pct3 <- round(prev_mh_benefits_count/sum(prev_mh_benefits_count)*100)
prev_mh_benefits_percentage <- paste(prev_mh_benefits_percentage, pct3) # add percents to labels
prev_mh_benefits_percentage <- paste(prev_mh_benefits_percentage,"%",sep="") # ad % to labels
pie(prev_mh_benefits_count,labels = prev_mh_benefits_percentage, col=rainbow(length(prev_mh_benefits_pe
   main="Pie Chart of Whether Previous Employers Provided Mental Health Benefits")
```

Pie Chart of Whether Previous Employers Provided Mental Health Ben




```
data2$prev_mh_benefits=
  ifelse(is.na(data2$prev_mh_benefits)==T,999999,
         ifelse(data2$prev_mh_benefits=='No, none did',1,
                ifelse(data2$prev_mh_benefits=="I don't know",2,
                       ifelse(data2$prev_mh_benefits=='Some did',3,4))))
```

```
table(data2$prev_mh_benefits)
```

```
##
##      1      2      3      4 999999
##    700    619    852    419    368
```

#14

```
#'Were you aware of the options for mental health care provided by your previous employers?',
#prev_mh_benefits_awareness
a<-table(data2$prev_mh_benefits_awareness)
a
```

```
##
##             I was aware of some              N/A (none offered)
##                             816                             231
##       N/A (not currently aware)            N/A (was not aware)
##                             582                             393
##   No, I only became aware later Yes, I was aware of all of them
##                             217                             351
```

```
str(a)
```

```
##  'table' int [1:6(1d)] 816 231 582 393 217 351
##  - attr(*, "dimnames")=List of 1
##   ..$ : chr [1:6] "I was aware of some" "N/A (none offered)" "N/A (not currently aware)" "N/A (was n
```

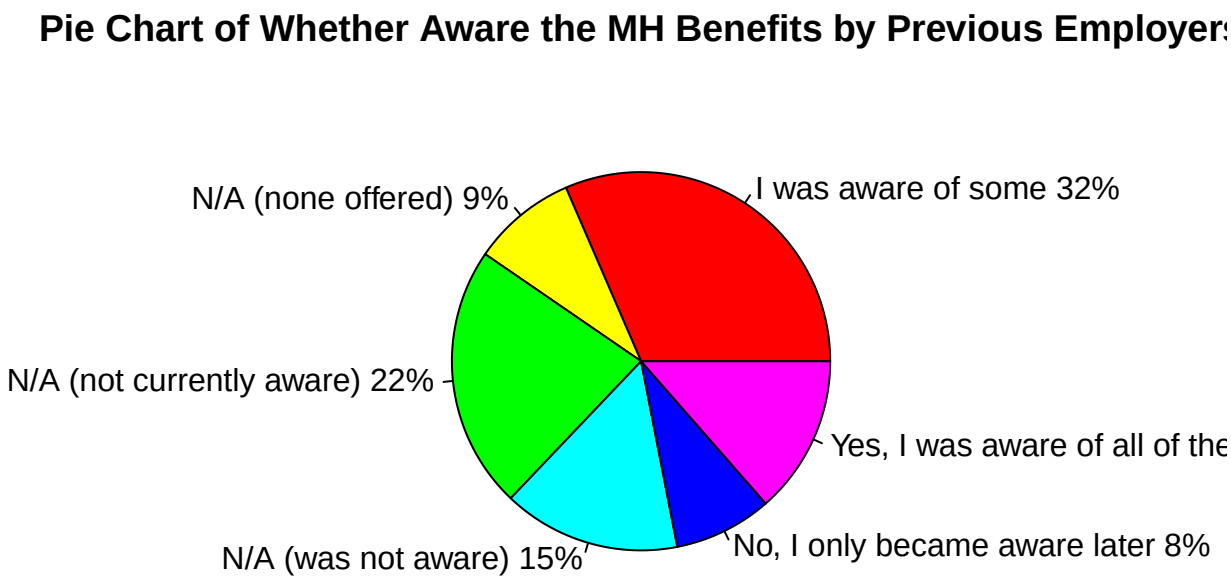
```
prev_mh_benefits_awareness_count<-as.vector(a)
prev_mh_benefits_awareness_percentage<-names(a)
prev_mh_benefits_awareness_count
```

```
## [1] 816 231 582 393 217 351
```

```
prev_mh_benefits_awareness_percentage
```

```
## [1] "I was aware of some"              "N/A (none offered)"
## [3] "N/A (not currently aware)"        "N/A (was not aware)"
## [5] "No, I only became aware later"    "Yes, I was aware of all of them"
```

```
pct4 <- round(prev_mh_benefits_awareness_count/sum(prev_mh_benefits_awareness_count)*100)
prev_mh_benefits_awareness_percentage <- paste(prev_mh_benefits_awareness_percentage, pct4) # add perce
prev_mh_benefits_awareness_percentage <- paste(prev_mh_benefits_awareness_percentage,"%",sep="") # ad %
pie(prev_mh_benefits_awareness_count,labels = prev_mh_benefits_awareness_percentage, col=rainbow(length
   main="Pie Chart of Whether Aware the MH Benefits by Previous Employers")
```




```
data2$prev_mh_benefits_awareness=
  ifelse(is.na(data2$prev_mh_benefits_awareness)==T,999999,
         ifelse(data2$prev_mh_benefits_awareness=="N/A (none offered)",1,
                ifelse(data2$prev_mh_benefits_awareness=="N/A (was not aware)",2,
                       ifelse(data2$prev_mh_benefits_awareness=="N/A (not currently aware)",3,
                              ifelse(data2$prev_mh_benefits_awareness=="No, I only became aware later",4
                                     ifelse(data2$prev_mh_benefits_awareness=="I was aware of some",5,6
```

```
table(data2$prev_mh_benefits_awareness)
```

```
##
##      1      2      3      4      5      6 999999
##    231    393    582    217    816    351    368
```

#15

```
#'Did your previous employers ever formally discuss mental health (as part of a wellness campaign or ot
#prev_mh_discussion
b<-table(data2$prev_mh_discussion)
b
```

```
##
##      I don't know          None did          Some did Yes, they all did
##               172              1794               559                65
```

```
str(b)
```

```
##  'table' int [1:4(1d)] 172 1794 559 65
##  - attr(*, "dimnames")=List of 1
##   ..$ : chr [1:4] "I don't know" "None did" "Some did" "Yes, they all did"
```

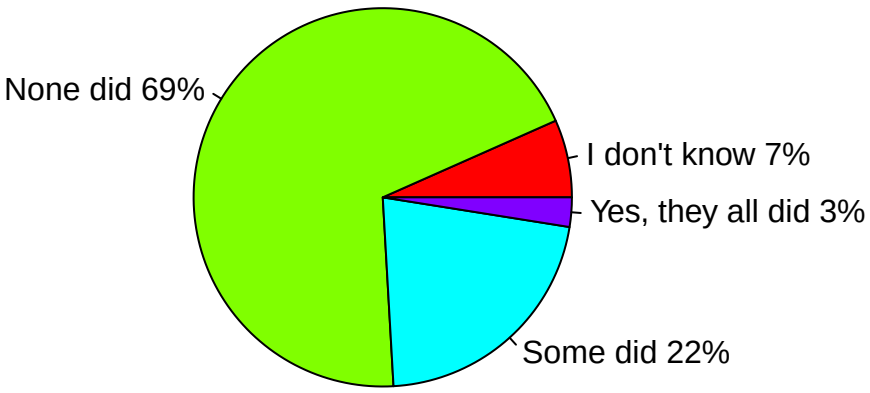
```
prev_mh_discussion_awareness_count<-as.vector(b)
prev_mh_discussion_awareness_percentage<-names(b)
prev_mh_discussion_awareness_count
```

```
## [1]  172 1794  559   65
```

```
prev_mh_discussion_awareness_percentage
```

```
## [1] "I don't know"      "None did"          "Some did"
## [4] "Yes, they all did"
```

```
pct5 <- round(prev_mh_discussion_awareness_count/sum(prev_mh_discussion_awareness_count)*100)
prev_mh_discussion_awareness_percentage <- paste(prev_mh_discussion_awareness_percentage, pct5)
prev_mh_discussion_awareness_percentage <- paste(prev_mh_discussion_awareness_percentage,"%",sep="")
pie(prev_mh_discussion_awareness_count,labels = prev_mh_discussion_awareness_percentage, col=rainbow(le
   main="Pie Chart of Whether Discuss Mental Health with Previous Employers")
```



```r
data2$prev_mh_discussion=
  ifelse(is.na(data2$prev_mh_discussion)==T,999999,
         ifelse(data2$prev_mh_discussion=="None did",1,
                ifelse(data2$prev_mh_discussion=="I don't know",2,
                       ifelse(data2$prev_mh_discussion=="Some did",3,4))))
```

```r
table(data2$prev_mh_discussion)
```

```
##
##      1      2      3      4 999999
##   1794    172    559     65    368
```