Mental Health Survey

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Cmd+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.

library(readxl)  
data <- read\_excel("/Users/rishavdas/Downloads/External\_data/survey.xlsx")

data

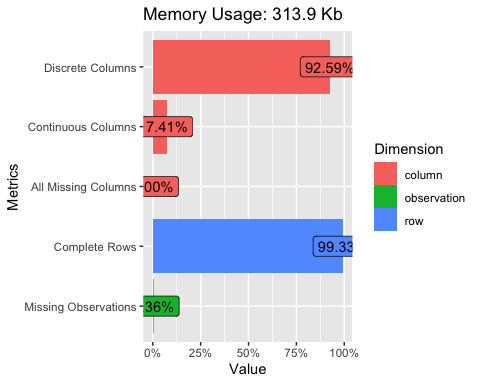
## # A tibble: 1,190 × 27  
## Timestamp Age Gender Country state self\_employed family\_history treatment  
## <dbl> <dbl> <chr> <chr> <chr> <chr> <chr> <chr>   
## 1 41879. -1726 male United… NA No No Yes   
## 2 41879. -29 Male United… MN No No No   
## 3 41882. -1 p United… AL Yes Yes Yes   
## 4 41878. 0 m Canada NA Yes No No   
## 5 41879. 0 Male United… NY No No Yes   
## 6 41879. 5 Male United… OH No No No   
## 7 41880. 8 A littl… Bahama… IL Yes Yes Yes   
## 8 41881. 11 male United… OH Yes No No   
## 9 41878. 18 Male Nether… NA No No No   
## 10 41879. 18 somethi… Russia NA No No No   
## # … with 1,180 more rows, and 19 more variables: work\_interfere <chr>,  
## # no\_employees <chr>, remote\_work <chr>, tech\_company <chr>, benefits <chr>,  
## # care\_options <chr>, wellness\_program <chr>, seek\_help <chr>,  
## # anonymity <chr>, leave <chr>, mental\_health\_consequence <chr>,  
## # phys\_health\_consequence <chr>, coworkers <chr>, supervisor <chr>,  
## # mental\_health\_interview <chr>, phys\_health\_interview <chr>,  
## # mental\_vs\_physical <chr>, obs\_consequence <chr>, comments <chr>

head(data)

## # A tibble: 6 × 27  
## Timestamp Age Gender Country state self\_employed family\_history treatment  
## <dbl> <dbl> <chr> <chr> <chr> <chr> <chr> <chr>   
## 1 41879. -1726 male United Ki… NA No No Yes   
## 2 41879. -29 Male United St… MN No No No   
## 3 41882. -1 p United St… AL Yes Yes Yes   
## 4 41878. 0 m Canada NA Yes No No   
## 5 41879. 0 Male United St… NY No No Yes   
## 6 41879. 5 Male United St… OH No No No   
## # … with 19 more variables: work\_interfere <chr>, no\_employees <chr>,  
## # remote\_work <chr>, tech\_company <chr>, benefits <chr>, care\_options <chr>,  
## # wellness\_program <chr>, seek\_help <chr>, anonymity <chr>, leave <chr>,  
## # mental\_health\_consequence <chr>, phys\_health\_consequence <chr>,  
## # coworkers <chr>, supervisor <chr>, mental\_health\_interview <chr>,  
## # phys\_health\_interview <chr>, mental\_vs\_physical <chr>,  
## # obs\_consequence <chr>, comments <chr>

#Visualizing structure of the dataset

library(DataExplorer)  
plot\_intro(data)



#Exploratory Analysis

str(data)

## tibble [1,190 × 27] (S3: tbl\_df/tbl/data.frame)  
## $ Timestamp : num [1:1190] 41879 41879 41882 41878 41879 ...  
## $ Age : num [1:1190] -1726 -29 -1 0 0 ...  
## $ Gender : chr [1:1190] "male" "Male" "p" "m" ...  
## $ Country : chr [1:1190] "United Kingdom" "United States" "United States" "Canada" ...  
## $ state : chr [1:1190] "NA" "MN" "AL" "NA" ...  
## $ self\_employed : chr [1:1190] "No" "No" "Yes" "Yes" ...  
## $ family\_history : chr [1:1190] "No" "No" "Yes" "No" ...  
## $ treatment : chr [1:1190] "Yes" "No" "Yes" "No" ...  
## $ work\_interfere : chr [1:1190] "Sometimes" "NA" "Often" "Sometimes" ...  
## $ no\_employees : chr [1:1190] "26-100" "More than 1000" "44201" "44201" ...  
## $ remote\_work : chr [1:1190] "2" "1" "Yes" "No" ...  
## $ tech\_company : chr [1:1190] "No" "No" "Yes" "Yes" ...  
## $ benefits : chr [1:1190] "No" "Yes" "Yes" "No" ...  
## $ care\_options : chr [1:1190] "No" "No" "Yes" "Yes" ...  
## $ wellness\_program : chr [1:1190] "No" "Don't know" "Yes" "Yes" ...  
## $ seek\_help : chr [1:1190] "No" "Yes" "Yes" "No" ...  
## $ anonymity : chr [1:1190] "Don't know" "Don't know" "Yes" "Don't know" ...  
## $ leave : chr [1:1190] "Somewhat difficult" "Don't know" "Very easy" "Don't know" ...  
## $ mental\_health\_consequence: chr [1:1190] "Yes" "No" "Yes" "No" ...  
## $ phys\_health\_consequence : chr [1:1190] "No" "No" "Yes" "No" ...  
## $ coworkers : chr [1:1190] "No" "Some of them" "Yes" "Yes" ...  
## $ supervisor : chr [1:1190] "No" "Yes" "Yes" "Yes" ...  
## $ mental\_health\_interview : chr [1:1190] "No" "No" "Yes" "No" ...  
## $ phys\_health\_interview : chr [1:1190] "Maybe" "No" "Yes" "No" ...  
## $ mental\_vs\_physical : chr [1:1190] "Don't know" "Don't know" "Yes" "Yes" ...  
## $ obs\_consequence : chr [1:1190] "No" "No" "Yes" "No" ...  
## $ comments : chr [1:1190] "NA" "NA" "password: testered" "NA" ...

summary(data)

## Timestamp Age Gender Country   
## Min. : 2014 Min. :-1.726e+03 Length:1190 Length:1190   
## 1st Qu.:41879 1st Qu.: 2.700e+01 Class :character Class :character   
## Median :41879 Median : 3.100e+01 Mode :character Mode :character   
## Mean :41779 Mean : 8.410e+07   
## 3rd Qu.:41880 3rd Qu.: 3.600e+01   
## Max. :41988 Max. : 1.000e+11   
## NA's :1 NA's :1   
## state self\_employed family\_history treatment   
## Length:1190 Length:1190 Length:1190 Length:1190   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## work\_interfere no\_employees remote\_work tech\_company   
## Length:1190 Length:1190 Length:1190 Length:1190   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## benefits care\_options wellness\_program seek\_help   
## Length:1190 Length:1190 Length:1190 Length:1190   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## anonymity leave mental\_health\_consequence  
## Length:1190 Length:1190 Length:1190   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## phys\_health\_consequence coworkers supervisor   
## Length:1190 Length:1190 Length:1190   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## mental\_health\_interview phys\_health\_interview mental\_vs\_physical  
## Length:1190 Length:1190 Length:1190   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## obs\_consequence comments   
## Length:1190 Length:1190   
## Class :character Class :character   
## Mode :character Mode :character   
##   
##   
##   
##

sapply(data, **function**(x) sum(is.na(x)))

## Timestamp Age Gender   
## 1 1 1   
## Country state self\_employed   
## 1 1 5   
## family\_history treatment work\_interfere   
## 5 5 5   
## no\_employees remote\_work tech\_company   
## 5 6 5   
## benefits care\_options wellness\_program   
## 5 5 5   
## seek\_help anonymity leave   
## 5 5 5   
## mental\_health\_consequence phys\_health\_consequence coworkers   
## 5 5 5   
## supervisor mental\_health\_interview phys\_health\_interview   
## 5 5 5   
## mental\_vs\_physical obs\_consequence comments   
## 5 5 6