DSC 520 Final Project - Mental Health in the Tech industry

Kausik Chattapadhyay

2022-11-17

Introduction:

All too often we may find ourselves joking around with someone in the tech industry and they may tell you that they must be nuts to be in this job. Even though we joke around about these things, if you suffer from mental health issues know you are not alone. This survey was completed to assess the presence of mental health issues in the tech industry.

Problem Statement:

- Is there a presence of mental health problems in the tech industry and does it impact how they perform their jobs?
- How you addressed this problem statement?
- After cleaning up the survey results, we reviewed a series of questions to determine the impacts.

Analysis.

- Out of the participants, what % are male compared to the % of females?

  Ans: 51% Male, 45% Female and 5% other.
- Out of the male and female participants, what is the average age or age range?

  Ans: Age Median is 35 (Female 34, Male 36, Other 34).
- Out of those surveyed how many have family histories of mental health issues?

  Ans: 66% have a family history of mental health issues.
- Out of those with family history, what % are seeking treatment?

  Ans: Of the 66% with a family history, 100% are seeking treatment. Of the 10% that are not and 24% with an unknown family history, only 2% are not seeking treatment.
- Out of those surveyed how many works in the tech industry?

  Ans: 89% are in the tech industry in their current role, but only 49% was previously.
- In relationship to the tech industry, what are the average and the median age of females and males?

  Ans: The average age is 36. The average age for Females is 35 in the industry and 31 The average age for Males is 36 in the industry and 30 out The average of the other is 34
- What is the likelihood that someone working in the tech industry seeks treatment for mental health issues?

Ans:Currently, 98% of those survey is seeking treatment.

- Comparatively those females and males working in the tech industry that are seeking treatment to those not seeking treatment feel that their mental health interferes with work.

  Ans: 63% of those who are in the tech industry say that they often have interference from their mental health at work while untreated. While only 8% often have interference in their work when receiving treatment.

Implications:

The results show that majority of those working in the tech industry are men. Of the men and women who survived a high percentage has a family history of mental health issues. Those who have this history have a high rate of seeking treatment for these health issues. Those who survived also have a higher rate of having health issues interfere with work, compared to when they are being treated.

Limitations:

- The limits have to do with the data collected and how it was collected. When designed a survey like this it is best to make it multiple choice and limited fill-able areas. Because this form had so many areas that could be skipped or left blank a lot of the information was not usable.
- On a personal side there was a limitation in time and understand to dive deeper into the data.

Concluding Remarks:

- As a person with mental health issues that works in the tech industry, it was nice to conclude that I am not alone and it is more common than I suspected.

Reading and Setting up the Data Process

```
## Set the working directory to the root of your DSC 520 directory
setwd("/Users/kausik/desktop/MS Data Science/DSC 520/dsc520-stats-r-assignments")

library(readr)
library(ggplot2)
library(dplyr)
library(labeling)
library(corpcor)
library(ggpubr)
library(MASS)
library(ppcor)
library(pastecs)
library(psych)
library(foreign)
library(tidyr)
```

Datasets:

I used the dataset from Kaggle for my research. 3 datasets for 2016, 2019 and 2020.

https://www.kaggle.com/code/zakisher/mental-health-in-tech-analysis-prediction/data?select=mental-heath-in-tech-2016_20161114.csv

```
## Set the working directory to the root of your DSC 520 directory
setwd("/Users/kausik/desktop/MS Data Science/DSC 520/dsc520-stats-r-assignments")
OSMI2019 <- read_csv("data/OSMI_2019.csv")
```

```
## New names:
## Rows: 352 Columns: 82
## -- Column specification
## -------------------------------------------------------- Delimiter: "," chr
## (55): How many employees does your company or organization have?, Does y... dbl
## (9): Overall, how much importance does your employer place on physical ... lgl
## (18): *Are you self-employed?*, Is your employer primarily a tech compan...
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `Describe the conversation your coworker had with you about their mental
##   health (please do not use names).` -> `Describe the conversation your
##   coworker had with you about their mental health (please do not use
##   names)....19`
## * `Describe the conversation your coworker had with you about their mental
##   health (please do not use names).` -> `Describe the conversation your
##   coworker had with you about their mental health (please do not use
##   names)....45`
## * `Why or why not?` -> `Why or why not?...61`
## * `Why or why not?` -> `Why or why not?...63`
```

```
str(OSMI2019)
```

```
## spec_tbl_df [352 x 82] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ *Are you self-employed?*
##  $ How many employees does your company or organization have?
##  $ Is your employer primarily a tech company/organization?
##  $ Is your primary role within your company related to tech/IT?
##  $ Does your employer provide mental health benefits as part of healthcare coverage?
##  $ Do you know the options for mental health care available under your employer-provided health cover
##  $ Has your employer ever formally discussed mental health (for example, as part of a wellness campai
##  $ Does your employer offer resources to learn more about mental health disorders and options for see
##  $ Is your anonymity protected if you choose to take advantage of mental health or substance abuse tr
##  $ If a mental health issue prompted you to request a medical leave from work, how easy or difficult
##  $ Would you feel more comfortable talking to your coworkers about your physical health or your menta
##  $ Would you feel comfortable discussing a mental health issue with your direct supervisor(s)?
##  $ Have you ever discussed your mental health with your employer?
##  $ Describe the conversation you had with your employer about your mental health, including their rea
##  $ Would you feel comfortable discussing a mental health issue with your coworkers?
##  $ Have you ever discussed your mental health with coworkers?
##  $ Describe the conversation with coworkers you had about your mental health including their reaction
##  $ Have you ever had a coworker discuss their or another coworker's mental health with you?
##  $ Describe the conversation your coworker had with you about their mental health (please do not use
```

```
##  $ Overall, how much importance does your employer place on physical health?
##  $ Overall, how much importance does your employer place on mental health?
##  $ Do you have medical coverage (private insurance or state-provided) that includes treatment of men
##  $ Do you know local or online resources to seek help for a mental health issue?
##  $ If you have been diagnosed or treated for a mental health disorder, do you ever reveal this to cl
##  $ If you have revealed a mental health disorder to a client or business contact, how has this affec
##  $ If you have been diagnosed or treated for a mental health disorder, do you ever reveal this to co
##  $ If you have revealed a mental health disorder to a coworker or employee, how has this impacted yo
##  $ Do you believe your productivity is ever affected by a mental health issue?
##  $ If yes, what percentage of your work time (time performing primary or secondary job functions) is
##  $ *Do you have previous employers?*
##  $ Was your employer primarily a tech company/organization?
##  $ Have your previous employers provided mental health benefits?
##  $ Were you aware of the options for mental health care provided by your previous employers?
##  $ Did your previous employers ever formally discuss mental health (as part of a wellness campaign o
##  $ Did your previous employers provide resources to learn more about mental health disorders and how
##  $ Was your anonymity protected if you chose to take advantage of mental health or substance abuse t
##  $ Would you have felt more comfortable talking to your previous employer about your physical health
##  $ Would you have been willing to discuss your mental health with your direct supervisor(s)?
##  $ Did you ever discuss your mental health with your previous employer?
##  $ Describe the conversation you had with your previous employer about your mental health, including
##  $ Would you have been willing to discuss your mental health with your coworkers at previous employe
##  $ Did you ever discuss your mental health with a previous coworker(s)?
##  $ Describe the conversation you had with your previous coworkers about your mental health including
##  $ Did you ever have a previous coworker discuss their or another coworker's mental health with you?
##  $ Describe the conversation your coworker had with you about their mental health (please do not use
##  $ Overall, how much importance did your previous employer place on physical health?
##  $ Overall, how much importance did your previous employer place on mental health?
##  $ Do you *currently* have a mental health disorder?
##  $ Have you ever been *diagnosed* with a mental health disorder?
##  $ *What disorder(s) have you been diagnosed with?*
##  $ *If possibly, what disorder(s) do you believe you have?*
##  $ *If so, what disorder(s) were you diagnosed with?*
##  $ Have you had a mental health disorder in the past?
##  $ Have you ever sought treatment for a mental health disorder from a mental health professional?
##  $ Do you have a family history of mental illness?
##  $ If you have a mental health disorder, how often do you feel that it interferes with your work *wh
##  $ If you have a mental health disorder, how often do you feel that it interferes with your work *wh
##  $ Have your observations of how another individual who discussed a mental health issue made you les
##  $ How willing would you be to share with friends and family that you have a mental illness?
##  $ Would you be willing to bring up a physical health issue with a potential employer in an intervie
##  $ Why or why not?...61
##  $ Would you bring up your *mental* health with a potential employer in an interview?
##  $ Why or why not?...63
##  $ Are you openly identified at work as a person with a mental health issue?
##  $ Has being identified as a person with a mental health issue affected your career?
##  $ How has it affected your career?
##  $ If they knew you suffered from a mental health disorder, how do you think that your team members/
##  $ Have you observed or experienced an *unsupportive or badly handled response* to a mental health i
##  $ Describe the circumstances of the badly handled or unsupportive response.
##  $ Have you observed or experienced a *supportive or well handled response* to a mental health issue
##  $ Describe the circumstances of the supportive or well handled response.
##  $ Overall, how well do you think the tech industry supports employees with mental health issues?
##  $ Briefly describe what you think the industry as a whole and/or employers could do to improve ment
```

```
##  $ If there is anything else you would like to tell us that has not been covered by the survey questi
##  $ Would you be willing to talk to one of us more extensively about your experiences with mental heal
##  $ What is your age?
##  $ What is your gender?
##  $ What country do you *live* in?
##  $ What US state or territory do you *live* in?
##  $ What is your race?
##  $ What country do you *work* in?
##  $ What US state or territory do you *work* in?
##  - attr(*, "spec")=
##   .. cols(
##   ..   `*Are you self-employed?*` = col_logical(),
##   ..   `How many employees does your company or organization have?` = col_character(),
##   ..   `Is your employer primarily a tech company/organization?` = col_logical(),
##   ..   `Is your primary role within your company related to tech/IT?` = col_logical(),
##   ..   `Does your employer provide mental health benefits as part of healthcare coverage?` = col_char
##   ..   `Do you know the options for mental health care available under your employer-provided health
##   ..   `Has your employer ever formally discussed mental health (for example, as part of a wellness c
##   ..   `Does your employer offer resources to learn more about mental health disorders and options fo
##   ..   `Is your anonymity protected if you choose to take advantage of mental health or substance abu
##   ..   `If a mental health issue prompted you to request a medical leave from work, how easy or diffi
##   ..   `Would you feel more comfortable talking to your coworkers about your physical health or your
##   ..   `Would you feel comfortable discussing a mental health issue with your direct supervisor(s)?`
##   ..   `Have you ever discussed your mental health with your employer?` = col_logical(),
##   ..   `Describe the conversation you had with your employer about your mental health, including thei
##   ..   `Would you feel comfortable discussing a mental health issue with your coworkers?` = col_chara
##   ..   `Have you ever discussed your mental health with coworkers?` = col_logical(),
##   ..   `Describe the conversation with coworkers you had about your mental health including their rea
##   ..   `Have you ever had a coworker discuss their or another coworker's mental health with you?` = c
##   ..   `Describe the conversation your coworker had with you about their mental health (please do not
##   ..   `Overall, how much importance does your employer place on physical health?` = col_double(),
##   ..   `Overall, how much importance does your employer place on mental health?` = col_double(),
##   ..   `Do you have medical coverage (private insurance or state-provided) that includes treatment of
##   ..   `Do you know local or online resources to seek help for a mental health issue?` = col_characte
##   ..   `If you have been diagnosed or treated for a mental health disorder, do you ever reveal this t
##   ..   `If you have revealed a mental health disorder to a client or business contact, how has this a
##   ..   `If you have been diagnosed or treated for a mental health disorder, do you ever reveal this t
##   ..   `If you have revealed a mental health disorder to a coworker or employee, how has this impacte
##   ..   `Do you believe your productivity is ever affected by a mental health issue?` = col_character
##   ..   `If yes, what percentage of your work time (time performing primary or secondary job functions
##   ..   `*Do you have previous employers?*` = col_logical(),
##   ..   `Was your employer primarily a tech company/organization?` = col_logical(),
##   ..   `Have your previous employers provided mental health benefits?` = col_character(),
##   ..   `Were you aware of the options for mental health care provided by your previous employers?` =
##   ..   `Did your previous employers ever formally discuss mental health (as part of a wellness campai
##   ..   `Did your previous employers provide resources to learn more about mental health disorders and
##   ..   `Was your anonymity protected if you chose to take advantage of mental health or substance abu
##   ..   `Would you have felt more comfortable talking to your previous employer about your physical he
##   ..   `Would you have been willing to discuss your mental health with your direct supervisor(s)?` =
##   ..   `Did you ever discuss your mental health with your previous employer?` = col_logical(),
##   ..   `Describe the conversation you had with your previous employer about your mental health, inclu
##   ..   `Would you have been willing to discuss your mental health with your coworkers at previous emp
##   ..   `Did you ever discuss your mental health with a previous coworker(s)?` = col_logical(),
##   ..   `Describe the conversation you had with your previous coworkers about your mental health inclu
```

```
##   ..   'Did you ever have a previous coworker discuss their or another coworker's mental health with
##   ..   'Describe the conversation your coworker had with you about their mental health (please do no
##   ..   'Overall, how much importance did your previous employer place on physical health?' = col_dou
##   ..   'Overall, how much importance did your previous employer place on mental health?' = col_doubl
##   ..   'Do you *currently* have a mental health disorder?' = col_character(),
##   ..   'Have you ever been *diagnosed* with a mental health disorder?' = col_character(),
##   ..   '*What disorder(s) have you been diagnosed with?*' = col_logical(),
##   ..   '*If possibly, what disorder(s) do you believe you have?*' = col_character(),
##   ..   '*If so, what disorder(s) were you diagnosed with?*' = col_character(),
##   ..   'Have you had a mental health disorder in the past?' = col_character(),
##   ..   'Have you ever sought treatment for a mental health disorder from a mental health professiona
##   ..   'Do you have a family history of mental illness?' = col_character(),
##   ..   'If you have a mental health disorder, how often do you feel that it interferes with your wor
##   ..   'If you have a mental health disorder, how often do you feel that it interferes with your wor
##   ..   'Have your observations of how another individual who discussed a mental health issue made yo
##   ..   'How willing would you be to share with friends and family that you have a mental illness?' =
##   ..   'Would you be willing to bring up a physical health issue with a potential employer in an int
##   ..   'Why or why not?...61' = col_character(),
##   ..   'Would you bring up your *mental* health with a potential employer in an interview?' = col_ch
##   ..   'Why or why not?...63' = col_character(),
##   ..   'Are you openly identified at work as a person with a mental health issue?' = col_logical(),
##   ..   'Has being identified as a person with a mental health issue affected your career?' = col_log
##   ..   'How has it affected your career?' = col_double(),
##   ..   'If they knew you suffered from a mental health disorder, how do you think that your team mem
##   ..   'Have you observed or experienced an *unsupportive or badly handled response* to a mental hea
##   ..   'Describe the circumstances of the badly handled or unsupportive response.' = col_character()
##   ..   'Have you observed or experienced a *supportive or well handled response* to a mental health
##   ..   'Describe the circumstances of the supportive or well handled response.' = col_logical(),
##   ..   'Overall, how well do you think the tech industry supports employees with mental health issue
##   ..   'Briefly describe what you think the industry as a whole and/or employers could do to improve
##   ..   'If there is anything else you would like to tell us that has not been covered by the survey
##   ..   'Would you be willing to talk to one of us more extensively about your experiences with menta
##   ..   'What is your age?' = col_double(),
##   ..   'What is your gender?' = col_character(),
##   ..   'What country do you *live* in?' = col_character(),
##   ..   'What US state or territory do you *live* in?' = col_character(),
##   ..   'What is your race?' = col_character(),
##   ..   'What country do you *work* in?' = col_character(),
##   ..   'What US state or territory do you *work* in?' = col_character()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

First Update in the cleaning of the data

```
tech2019 <- OSMI2019[, -c(1:3,5:21,23:27,33:47, 58:75 ),] # delete unneeded columns down to 24
names(tech2019)
```

```
##  [1] "Is your primary role within your company related to tech/IT?"
##  [2] "Do you have medical coverage (private insurance or state-provided) that includes treatment of
##  [3] "Do you believe your productivity is ever affected by a mental health issue?"
##  [4] "If yes, what percentage of your work time (time performing primary or secondary job functions)
```

```
##  [5] "*Do you have previous employers?*"
##  [6] "Was your employer primarily a tech company/organization?"
##  [7] "Have your previous employers provided mental health benefits?"
##  [8] "Do you *currently* have a mental health disorder?"
##  [9] "Have you ever been *diagnosed* with a mental health disorder?"
## [10] "*What disorder(s) have you been diagnosed with?*"
## [11] "*If possibly, what disorder(s) do you believe you have?*"
## [12] "*If so, what disorder(s) were you diagnosed with?*"
## [13] "Have you had a mental health disorder in the past?"
## [14] "Have you ever sought treatment for a mental health disorder from a mental health professional?
## [15] "Do you have a family history of mental illness?"
## [16] "If you have a mental health disorder, how often do you feel that it interferes with your work
## [17] "If you have a mental health disorder, how often do you feel that it interferes with your work
## [18] "What is your age?"
## [19] "What is your gender?"
## [20] "What country do you *live* in?"
## [21] "What US state or territory do you *live* in?"
## [22] "What is your race?"
## [23] "What country do you *work* in?"
## [24] "What US state or territory do you *work* in?"
```

```
head(tech2019)
```

```
## # A tibble: 6 x 24
##   Is your prim~1 Do yo~2 Do yo~3 If ye~4 *Do y~5 Was y~6 Have ~7 Do yo~8 Have ~9
##   <lgl>          <lgl>   <chr>   <chr>   <lgl>   <lgl>   <chr>   <chr>   <chr>
## 1 TRUE           NA      <NA>    <NA>    TRUE    FALSE   I don'~ Don't ~ <NA>
## 2 TRUE           NA      <NA>    <NA>    TRUE    FALSE   Yes, t~ Possib~ <NA>
## 3 TRUE           NA      <NA>    <NA>    TRUE    TRUE    I don'~ No      <NA>
## 4 TRUE           NA      <NA>    <NA>    TRUE    TRUE    I don'~ No      <NA>
## 5 TRUE           NA      <NA>    <NA>    TRUE    TRUE    I don'~ No      <NA>
## 6 FALSE          NA      <NA>    <NA>    TRUE    TRUE    Yes, t~ Yes     Yes
## # ... with 15 more variables:
## #   `*What disorder(s) have you been diagnosed with?*` <lgl>,
## #   `*If possibly, what disorder(s) do you believe you have?*` <chr>,
## #   `*If so, what disorder(s) were you diagnosed with?*` <chr>,
## #   `Have you had a mental health disorder in the past?` <chr>,
## #   `Have you ever sought treatment for a mental health disorder from a mental health professional?`
## #   `Do you have a family history of mental illness?` <chr>, ...
```

Data Summary

```
summary(tech2019)
```

```
##  Is your primary role within your company related to tech/IT?
##  Mode :logical
##  FALSE:22
##  TRUE :282
##  NA's :48
##
```

```
##
##  Do you have medical coverage (private insurance or state-provided) that includes treatment of mental
##  Mode :logical
##  FALSE:16
##  TRUE :32
##  NA's :304
##
##
##  Do you believe your productivity is ever affected by a mental health issue?
##  Length:352
##  Class :character
##  Mode  :character
##
##
##
##  If yes, what percentage of your work time (time performing primary or secondary job functions) is ar
##  Length:352
##  Class :character
##  Mode  :character
##
##
##
##  *Do you have previous employers?*
##  Mode :logical
##  FALSE:56
##  TRUE :296
##
##
##
##  Was your employer primarily a tech company/organization?
##  Mode :logical
##  FALSE:108
##  TRUE :188
##  NA's :56
##
##
##  Have your previous employers provided mental health benefits?
##  Length:352
##  Class :character
##  Mode  :character
##
##
##
##  Do you *currently* have a mental health disorder?
##  Length:352
##  Class :character
##  Mode  :character
##
##
##
##  Have you ever been *diagnosed* with a mental health disorder?
##  Length:352
##  Class :character
##  Mode  :character
```

```
##
##
##
##  *What disorder(s) have you been diagnosed with?*
##  Mode:logical
##  NA's:352
##
##
##
##
##  *If possibly, what disorder(s) do you believe you have?*
##  Length:352
##  Class :character
##  Mode  :character
##
##
##
##  *If so, what disorder(s) were you diagnosed with?*
##  Length:352
##  Class :character
##  Mode  :character
##
##
##
##  Have you had a mental health disorder in the past?
##  Length:352
##  Class :character
##  Mode  :character
##
##
##
##  Have you ever sought treatment for a mental health disorder from a mental health professional?
##  Mode :logical
##  FALSE:135
##  TRUE :217
##
##
##
##  Do you have a family history of mental illness?
##  Length:352
##  Class :character
##  Mode  :character
##
##
##
##  If you have a mental health disorder, how often do you feel that it interferes with your work *when
##  Length:352
##  Class :character
##  Mode  :character
##
##
##
##  If you have a mental health disorder, how often do you feel that it interferes with your work *when*
##  Length:352
```

```
##  Class :character
##  Mode  :character
##
##
##
##  What is your age? What is your gender? What country do you *live* in?
##  Min.   : 0.00     Length:352           Length:352
##  1st Qu.:28.75     Class :character     Class :character
##  Median :34.00     Mode  :character     Mode  :character
##  Mean   :35.49
##  3rd Qu.:41.00
##  Max.   :64.00
##  What US state or territory do you *live* in? What is your race?
##  Length:352                                   Length:352
##  Class :character                             Class :character
##  Mode  :character                             Mode  :character
##
##
##
##  What country do you *work* in? What US state or territory do you *work* in?
##  Length:352                     Length:352
##  Class :character               Class :character
##  Mode  :character               Mode  :character
##
##
##
```

Second update in Cleaning of Data

```
tech_2019 <- tech2019[, -c(2:4,7,10:12,23,24)] # delete unneeded columns down to 15
names(tech_2019)
```

```
##  [1] "Is your primary role within your company related to tech/IT?"
##  [2] "*Do you have previous employers?*"
##  [3] "Was your employer primarily a tech company/organization?"
##  [4] "Do you *currently* have a mental health disorder?"
##  [5] "Have you ever been *diagnosed* with a mental health disorder?"
##  [6] "Have you had a mental health disorder in the past?"
##  [7] "Have you ever sought treatment for a mental health disorder from a mental health professional?
##  [8] "Do you have a family history of mental illness?"
##  [9] "If you have a mental health disorder, how often do you feel that it interferes with your work
## [10] "If you have a mental health disorder, how often do you feel that it interferes with your work
## [11] "What is your age?"
## [12] "What is your gender?"
## [13] "What country do you *live* in?"
## [14] "What US state or territory do you *live* in?"
## [15] "What is your race?"
```

```
names(tech_2019)[names(tech_2019) == 'Is your primary role within your company related to tech/IT?'] <-
names(tech_2019)[names(tech_2019) == '*Do you have previous employers?*'] <- 'TF_Pre_emp'
names(tech_2019)[names(tech_2019) == 'Was your employer primarily a tech company/organization?'] <- 'P_
```

```
names(tech_2019)[names(tech_2019) == 'Do you *currently* have a mental health disorder?'] <- 'C_MHealth
names(tech_2019)[names(tech_2019) == 'Have you ever been *diagnosed* with a mental health disorder?'] <-
names(tech_2019)[names(tech_2019) == 'Have you had a mental health disorder in the past?'] <- 'P_MHealth
names(tech_2019)[names(tech_2019) == 'Have you ever sought treatment for a mental health disorder from a
names(tech_2019)[names(tech_2019) == 'Do you have a family history of mental illness?'] <- 'FamHist'
names(tech_2019)[names(tech_2019) == 'If you have a mental health disorder, how often do you feel that
names(tech_2019)[names(tech_2019) == 'If you have a mental health disorder, how often do you feel that
names(tech_2019)[names(tech_2019) == 'What is your age?'] <- 'Age'
names(tech_2019)[names(tech_2019) == 'What is your gender?'] <- 'Gender_IDT'
names(tech_2019)[names(tech_2019) == 'What country do you *live* in?'] <- 'Country'
names(tech_2019)[names(tech_2019) == 'What US state or territory do you *live* in?'] <- 'StateTerr'
names(tech_2019)[names(tech_2019) == 'What is your race?'] <- 'race'
summary(tech_2019)
```

```
##    C_role        TF_Pre_emp        P_role         C_MHealth
##  Mode :logical   Mode :logical   Mode :logical   Length:352
##  FALSE:22        FALSE:56        FALSE:108       Class :character
##  TRUE :282       TRUE :296       TRUE :188       Mode  :character
##  NA's :48                        NA's :56
##
##
##   Diagnosed          P_MHealth         Treatment         FamHist
##  Length:352         Length:352         Mode :logical   Length:352
##  Class :character   Class :character   FALSE:135       Class :character
##  Mode  :character   Mode  :character   TRUE :217       Mode  :character
##
##
##
##    Interfer         Interfer_NT             Age         Gender_IDT
##  Length:352         Length:352         Min.   : 0.00   Length:352
##  Class :character   Class :character   1st Qu.:28.75   Class :character
##  Mode  :character   Mode  :character   Median :34.00   Mode  :character
##                                        Mean   :35.49
##                                        3rd Qu.:41.00
##                                        Max.   :64.00
##    Country           StateTerr            race
##  Length:352         Length:352         Length:352
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
```

Current Data Structure

```
str(tech_2019)
```

```
## tibble [352 x 15] (S3: tbl_df/tbl/data.frame)
##  $ C_role     : logi [1:352] TRUE TRUE TRUE TRUE TRUE FALSE ...
##  $ TF_Pre_emp : logi [1:352] TRUE TRUE TRUE TRUE TRUE TRUE ...
```

```
##  $ P_role     : logi [1:352] FALSE FALSE TRUE TRUE TRUE TRUE ...
##  $ C_MHealth  : chr [1:352] "Don't Know" "Possibly" "No" "No" ...
##  $ Diagnosed  : chr [1:352] NA NA NA NA ...
##  $ P_MHealth  : chr [1:352] "No" "Possibly" "No" "No" ...
##  $ Treatment  : logi [1:352] FALSE FALSE FALSE FALSE FALSE TRUE ...
##  $ FamHist    : chr [1:352] "No" "Yes" "I don't know" "Yes" ...
##  $ Interfer   : chr [1:352] "Not applicable to me" "Sometimes" "Not applicable to me" "Not applicable
##  $ Interfer_NT: chr [1:352] "Not applicable to me" "Often" "Not applicable to me" "Not applicable to
##  $ Age        : num [1:352] 25 51 27 37 46 36 39 35 49 45 ...
##  $ Gender_IDT : chr [1:352] "Male" "male" "Male" "male" ...
##  $ Country    : chr [1:352] "United States of America" "United States of America" "United States of
##  $ StateTerr  : chr [1:352] "Nebraska" "Nebraska" "Illinois" "Nebraska" ...
##  $ race       : chr [1:352] "White" "White" "White" "White" ...
```

Updating objectives

#Yes to True, No to False #Updating Gender to three types: Male, Female and Other

```
tech_2019[tech_2019 == "Yes"] <- "TRUE"
tech_2019[tech_2019 == "No"] <- "FALSE"
tech_2019[tech_2019 == "m"] <- "Male"
tech_2019[tech_2019 == "M"] <- "Male"
tech_2019[tech_2019 == "male"] <- "Male"
tech_2019[tech_2019 == "Cishet male"] <- "Male"
tech_2019[tech_2019 == "cis male"] <- "Male"
tech_2019[tech_2019 == "Cis Male"] <- "Male"
tech_2019[tech_2019 == "CIS Male"] <- "Male"
tech_2019[tech_2019 == "I have a penis"] <- "Male"
tech_2019[tech_2019 == "Identify as male"] <- "Male"
tech_2019[tech_2019 == "Masculine"] <- "Male"
tech_2019[tech_2019 == "masculino"] <- "Male"
tech_2019[tech_2019 == "Trans man"] <- "Male"
tech_2019[tech_2019 == "man"] <- "Male"
tech_2019[tech_2019 == "Man"] <- "Male"
tech_2019[tech_2019 == "Make"] <- "Male"
tech_2019[tech_2019 == "make"] <- "Male"
tech_2019[tech_2019 == "cis woman"] <- "Female"
tech_2019[tech_2019 == "Agender trans woman"] <- "Female"
tech_2019[tech_2019 == "Female-identified"] <- "Female"
tech_2019[tech_2019 == "Female (cis)"] <- "Female"
tech_2019[tech_2019 == "Femile"] <- "Female"
tech_2019[tech_2019 == "femmina"] <- "Female"
tech_2019[tech_2019 == "woman"] <- "Female"
tech_2019[tech_2019 == "Woman"] <- "Female"
tech_2019[tech_2019 == "F"] <- "Female"
tech_2019[tech_2019 == "f"] <- "Female"
tech_2019[tech_2019 == "female"] <- "Female"
tech_2019[tech_2019 == "agender"] <- "Other"
tech_2019[tech_2019 == "I am a Wookie"] <- "Other"
tech_2019[tech_2019 == "Non-binary"] <- "Other"
tech_2019[tech_2019 == "Non-binary and gender fluid"] <- "Other"
tech_2019[tech_2019 == "Non binary"] <- "Other"
```

```
tech_2019[tech_2019 == "Nonbinary"] <- "Other"
tech_2019[tech_2019 == "None"] <- "Other"
tech_2019[tech_2019 == "Questioning"] <- "Other"
tech_2019[tech_2019 == "rr"] <- "Other"
tech_2019[tech_2019 == "Trans non-binary/genderfluid"] <- "Other"
tech_2019 <- tech_2019 %>%
  mutate_at(c('Gender_IDT'), ~replace_na(.,"Other"))
tech_2019$Gender_IDT[tech_2019$Gender_IDT == "NA"] <- "Other"
tech_2019$Gender_IDT[tech_2019$Gender_IDT == "43"] <- "Other"
tech_2019$Gender_IDT[startsWith(tech_2019$Gender_IDT, "Let's keep it simple")] <- "Male"
tech_2019$Gender_IDT[tech_2019$Gender_IDT == "Na"] <- "Other"
tech_2019$Gender_IDT[tech_2019$Gender_IDT == "na"] <- "Other"
tech_2019$Gender_IDT[tech_2019$Gender_IDT == ""] <- "Other"
tech_2019[tech_2019 == "White"] <- "Caucasian"
tech_2019[tech_2019 == "Hispanic"] <- "Caucasian"
tech_2019[tech_2019 == "White Hispanic"] <- "Caucasian"
tech_2019[tech_2019 == "Black or African American"] <- "Black"
tech_2019[tech_2019 == "I prefer not to answer"] <- "Undisclosed"
tech_2019[tech_2019 == "More than one of the above"] <- "Other"
tech_2019[tech_2019 == "I don't know"] <- "Unknown"
tech_2019[tech_2019 == "European American"] <- "Other"
tech_2019$StateTerr[tech_2019$StateTerr == ""] <- "Undisclosed/Non-Us"
tech_2019A <- tech_2019

cc = is.na(tech_2019$C_role)
m = which(cc == c("TRUE"))
tech_2019 = tech_2019[-m,]
tech_2019B <- tech_2019
summary(tech_2019)
```

```
##    C_role        TF_Pre_emp        P_role         C_MHealth
##  Mode :logical   Mode :logical   Mode :logical   Length:304
##  FALSE:22        FALSE:51        FALSE:98        Class :character
##  TRUE :282       TRUE :253       TRUE :155       Mode  :character
##                                  NA's :51
##
##
##   Diagnosed          P_MHealth         Treatment         FamHist
##  Length:304         Length:304         Mode :logical   Length:304
##  Class :character   Class :character   FALSE:112       Class :character
##  Mode  :character   Mode  :character   TRUE :192       Mode  :character
##
##
##
##    Interfer         Interfer_NT             Age         Gender_IDT
##  Length:304         Length:304         Min.   : 0.00   Length:304
##  Class :character   Class :character   1st Qu.:28.00   Class :character
##  Mode  :character   Mode  :character   Median :34.00   Mode  :character
##                                        Mean   :34.86
##                                        3rd Qu.:40.00
##                                        Max.   :63.00
##    Country           StateTerr            race
##  Length:304         Length:304         Length:304
```

```
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
```

```
tech_2019 <- na.omit(tech_2019)
summary(tech_2019)
```

```
##    C_role        TF_Pre_emp       P_role         C_MHealth
##  Mode :logical   Mode:logical   Mode :logical   Length:84
##  FALSE:9         TRUE:84        FALSE:42        Class :character
##  TRUE :75                       TRUE :42        Mode  :character
##
##
##
##   Diagnosed          P_MHealth         Treatment         FamHist
##  Length:84          Length:84          Mode :logical   Length:84
##  Class :character   Class :character   FALSE:2         Class :character
##  Mode  :character   Mode  :character   TRUE :82        Mode  :character
##
##
##
##    Interfer         Interfer_NT             Age         Gender_IDT
##  Length:84          Length:84          Min.   :19.00   Length:84
##  Class :character   Class :character   1st Qu.:28.75   Class :character
##  Mode  :character   Mode  :character   Median :35.50   Mode  :character
##                                        Mean   :35.56
##                                        3rd Qu.:41.00
##                                        Max.   :54.00
##    Country           StateTerr            race
##  Length:84          Length:84          Length:84
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
```

```
str(tech_2019)
```

```
## tibble [84 x 15] (S3: tbl_df/tbl/data.frame)
##  $ C_role     : logi [1:84] FALSE TRUE TRUE TRUE FALSE TRUE ...
##  $ TF_Pre_emp : logi [1:84] TRUE TRUE TRUE TRUE TRUE TRUE ...
##  $ P_role     : logi [1:84] TRUE FALSE TRUE TRUE FALSE TRUE ...
##  $ C_MHealth  : chr [1:84] "TRUE" "TRUE" "TRUE" "TRUE" ...
##  $ Diagnosed  : chr [1:84] "TRUE" "TRUE" "TRUE" "TRUE" ...
##  $ P_MHealth  : chr [1:84] "TRUE" "TRUE" "Don't Know" "Possibly" ...
##  $ Treatment  : logi [1:84] TRUE TRUE TRUE TRUE TRUE TRUE ...
##  $ FamHist    : chr [1:84] "TRUE" "TRUE" "FALSE" "Unknown" ...
##  $ Interfer   : chr [1:84] "Sometimes" "Rarely" "Sometimes" "Sometimes" ...
##  $ Interfer_NT: chr [1:84] "Often" "Sometimes" "Often" "Often" ...
##  $ Age        : num [1:84] 36 39 26 37 40 37 37 45 28 41 ...
##  $ Gender_IDT : chr [1:84] "Female" "Female" "Female" "Male" ...
```

```
##  $ Country    : chr [1:84] "United States of America" "United States of America" "United States of Am
##  $ StateTerr  : chr [1:84] "Nebraska" "Nebraska" "Pennsylvania" "Wisconsin" ...
##  $ race       : chr [1:84] "Caucasian" "Caucasian" "Caucasian" "Caucasian" ...
##  - attr(*, "na.action")= 'omit' Named int [1:220] 1 2 3 4 5 8 9 10 11 12 ...
##   ..- attr(*, "names")= chr [1:220] "1" "2" "3" "4" ...
```

Updated Factors in filtered Data Frame

```
tech_2019C <- tech_2019
tech_2019 <- as.data.frame(unclass(tech_2019), stringsAsFactors = TRUE)
#tech_2019 <- stringsAsFactors(tech_2019)
str(tech_2019)
```

```
## 'data.frame':    84 obs. of  15 variables:
##  $ C_role     : logi  FALSE TRUE TRUE TRUE FALSE TRUE ...
##  $ TF_Pre_emp : logi  TRUE TRUE TRUE TRUE TRUE TRUE ...
##  $ P_role     : logi  TRUE FALSE TRUE TRUE FALSE TRUE ...
##  $ C_MHealth  : Factor w/ 1 level "TRUE": 1 1 1 1 1 1 1 1 1 1 ...
##  $ Diagnosed  : Factor w/ 1 level "TRUE": 1 1 1 1 1 1 1 1 1 1 ...
##  $ P_MHealth  : Factor w/ 4 levels "Don't Know","FALSE",..: 4 4 1 3 4 4 4 4 4 4 ...
##  $ Treatment  : logi  TRUE TRUE TRUE TRUE TRUE TRUE ...
##  $ FamHist    : Factor w/ 3 levels "FALSE","TRUE",..: 2 2 1 3 2 3 2 2 2 2 ...
##  $ Interfer   : Factor w/ 5 levels "Never","Not applicable to me",..: 5 4 5 5 4 5 4 4 4 1 ...
##  $ Interfer_NT: Factor w/ 3 levels "Often","Rarely",..: 1 3 1 1 1 3 3 3 1 3 ...
##  $ Age        : num  36 39 26 37 40 37 37 45 28 41 ...
##  $ Gender_IDT : Factor w/ 3 levels "Female","Male",..: 1 1 1 2 1 2 2 1 2 2 ...
##  $ Country    : Factor w/ 1 level "United States of America": 1 1 1 1 1 1 1 1 1 1 ...
##  $ StateTerr  : Factor w/ 28 levels "California","Colorado",..: 14 14 20 27 16 16 16 27 2 1 ...
##  $ race       : Factor w/ 5 levels "Asian","Black",..: 3 3 3 3 5 3 3 3 3 3 ...
```

New Summary

```
summary(tech_2019)
```

```
##    C_role        TF_Pre_emp       P_role        C_MHealth Diagnosed
##  Mode :logical   Mode:logical   Mode :logical   TRUE:84   TRUE:84
##  FALSE:9         TRUE:84        FALSE:42
##  TRUE :75                       TRUE :42
##
##
##
##
##       P_MHealth  Treatment          FamHist                   Interfer
##  Don't Know: 4   Mode :logical   FALSE  : 8   Never                : 4
##  FALSE     : 2   FALSE:2         TRUE   :56   Not applicable to me: 3
##  Possibly  : 9   TRUE :82        Unknown:20   Often                : 8
##  TRUE      :69                                Rarely               :31
##                                               Sometimes            :38
```

```
##
##
##     Interfer_NT      Age          Gender_IDT                      Country
##  Often    :61   Min.   :19.00   Female:37   United States of America:84
##  Rarely   : 2   1st Qu.:28.75   Male  :42
##  Sometimes:21   Median :35.50   Other : 5
##                 Mean   :35.56
##                 3rd Qu.:41.00
##                 Max.   :54.00
##
##       StateTerr           race
##  Nebraska  : 9   Asian      : 1
##  Washington: 9   Black      : 2
##  California: 6   Caucasian  :76
##  Michigan  : 6   Other      : 2
##  New York  : 6   Undisclosed: 3
##  Ohio      : 6
##  (Other)   :42
```

Updated Factors Cleaned Role Data Frame

```
tech_2019D <- tech_2019B
tech_2019D <- as.data.frame(unclass(tech_2019D), stringsAsFactors = TRUE)
str(tech_2019D)
```

```
## 'data.frame':    304 obs. of  15 variables:
##  $ C_role     : logi  TRUE TRUE TRUE TRUE TRUE FALSE ...
##  $ TF_Pre_emp : logi  TRUE TRUE TRUE TRUE TRUE TRUE ...
##  $ P_role     : logi  FALSE FALSE TRUE TRUE TRUE TRUE ...
##  $ C_MHealth  : Factor w/ 4 levels "Don't Know","FALSE",..: 1 3 2 2 2 4 4 2 3 3 ...
##  $ Diagnosed  : Factor w/ 2 levels "FALSE","TRUE": NA NA NA NA NA 2 2 NA NA NA ...
##  $ P_MHealth  : Factor w/ 4 levels "Don't Know","FALSE",..: 2 3 2 2 2 4 4 4 1 3 ...
##  $ Treatment  : logi  FALSE FALSE FALSE FALSE FALSE TRUE ...
##  $ FamHist    : Factor w/ 3 levels "FALSE","TRUE",..: 1 2 3 2 1 2 2 1 3 1 ...
##  $ Interfer   : Factor w/ 5 levels "Never","Not applicable to me",..: 2 5 2 2 2 5 4 1 5 2 ...
##  $ Interfer_NT: Factor w/ 5 levels "Never","Not applicable to me",..: 2 3 2 2 2 3 5 2 5 5 ...
##  $ Age        : num  25 51 27 37 46 36 39 49 45 40 ...
##  $ Gender_IDT : Factor w/ 3 levels "Female","Male",..: 2 2 2 2 2 1 1 3 2 2 ...
##  $ Country    : Factor w/ 27 levels "Brazil","Canada",..: 27 27 27 27 27 27 27 26 26 26 ...
##  $ StateTerr  : Factor w/ 38 levels "Alabama","Alaska",..: 22 22 11 22 22 22 22 NA NA NA ...
##  $ race       : Factor w/ 5 levels "Asian","Black",..: 3 3 3 3 3 3 3 NA NA NA ...
```

New Summary Only Role cleaned

```
summary(tech_2019D)
```

```
##    C_role        TF_Pre_emp        P_role             C_MHealth   Diagnosed
##  Mode :logical   Mode :logical   Mode :logical   Don't Know: 21   FALSE:  3
```

```
##  FALSE:22        FALSE:51        FALSE:98        FALSE     : 88   TRUE :128
##  TRUE :282       TRUE :253       TRUE :155       Possibly  : 64   NA's :173
##                                  NA's :51        TRUE      :131
##
##
##
##       P_MHealth   Treatment          FamHist                    Interfer
##  Don't Know: 20   Mode :logical   FALSE  : 88   Never                :  8
##  FALSE     : 91   FALSE:112       TRUE   :141   Not applicable to me:110
##  Possibly  : 61   TRUE :192       Unknown: 75   Often                : 23
##  TRUE      :129                                 Rarely               : 67
##  NA's      :  3                                 Sometimes            : 96
##
##
##                Interfer_NT       Age         Gender_IDT
##  Never               :  3   Min.   : 0.00   Female: 87
##  Not applicable to me: 94   1st Qu.:28.00   Male  :200
##  Often               :130   Median :34.00   Other : 17
##  Rarely              : 12   Mean   :34.86
##  Sometimes           : 65   3rd Qu.:40.00
##                             Max.   :63.00
##
##                      Country         StateTerr            race
##  United States of America:184   California: 18   Asian      :  9
##  United Kingdom          : 26   Nebraska  : 18   Black      :  2
##  Portugal                : 17   New York  : 14   Caucasian  :160
##  Brazil                  : 13   Tennessee : 13   Other      :  8
##  Canada                  : 12   Washington: 13   Undisclosed:  5
##  India                   : 12   (Other)   :108   NA's       :120
##  (Other)                 : 40   NA's      :120
```

Linear Model for age and gender

```
LMGender_Age <- lm(Age ~ Gender_IDT, data = tech_2019)
summary(LMGender_Age)
```

```
##
## Call:
## lm(formula = Age ~ Gender_IDT, data = tech_2019)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.6667  -6.6667   0.3333   4.4054  19.6216
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      34.3784     1.3827  24.863   <2e-16 ***
## Gender_IDTMale    2.2883     1.8964   1.207    0.231
## Gender_IDTOther   0.6216     4.0075   0.155    0.877
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 8.411 on 81 degrees of freedom
## Multiple R-squared:  0.01794,    Adjusted R-squared:  -0.00631
## F-statistic: 0.7398 on 2 and 81 DF,  p-value: 0.4804
```

LinearModel for age, current role and gender

```
LMGender_Role <- lm(Age ~ C_role + Gender_IDT, data = tech_2019)
summary(LMGender_Role)
```

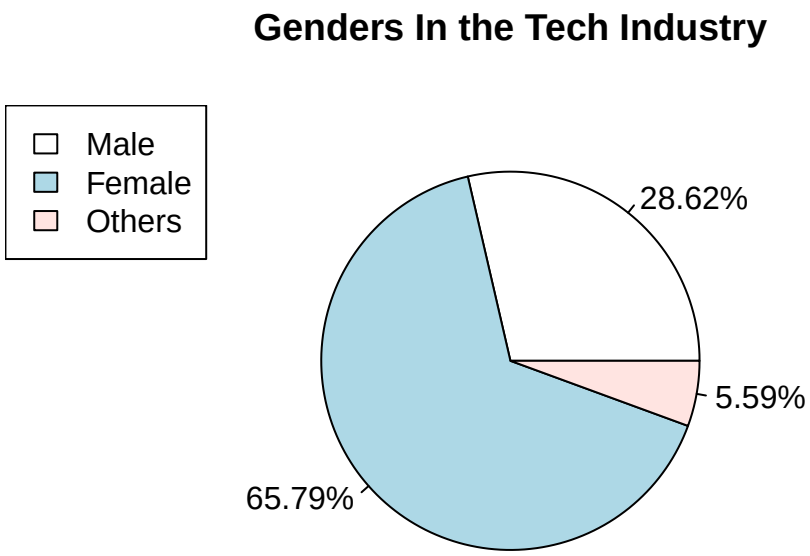
```
##
## Call:
## lm(formula = Age ~ C_role + Gender_IDT, data = tech_2019)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.1226  -6.2821   0.2397   4.3826  18.8774
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      31.1887     2.8440  10.966   <2e-16 ***
## C_roleTRUE        3.9340     3.0689   1.282    0.204
## Gender_IDTMale    1.6377     1.9559   0.837    0.405
## Gender_IDTOther   0.6642     3.9918   0.166    0.868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.378 on 80 degrees of freedom
## Multiple R-squared:  0.0377, Adjusted R-squared:  0.001618
## F-statistic: 1.045 on 3 and 80 DF,  p-value: 0.3774
```

Percentage of genders Full Survey

```
f <- table(tech_2019B$Gender_IDT)
f
```

```
##
## Female   Male  Other
##     87    200     17
```

```
pie_labels <- paste0(round(100 * f/sum(f), 2), "%")
pie(f,main = "Genders In the Tech Industry", labels = pie_labels)
legend("topleft", legend = c("Male", "Female", "Others"),
       fill =  c("white", "lightblue", "mistyrose"))
```



LinearModel for age, current role and gender

```
LMGender_Role <- lm(Age ~ C_role + Gender_IDT, data = tech_2019)
summary(LMGender_Role)
```

```
##
## Call:
## lm(formula = Age ~ C_role + Gender_IDT, data = tech_2019)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.1226  -6.2821   0.2397   4.3826  18.8774
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      31.1887     2.8440  10.966   <2e-16 ***
## C_roleTRUE        3.9340     3.0689   1.282    0.204
## Gender_IDTMale    1.6377     1.9559   0.837    0.405
## Gender_IDTOther   0.6642     3.9918   0.166    0.868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.378 on 80 degrees of freedom
```

```
## Multiple R-squared:  0.0377, Adjusted R-squared:  0.001618
## F-statistic: 1.045 on 3 and 80 DF,  p-value: 0.3774
```

LinearModel for age and family history

```
LMFamilyHist <- lm(Age ~ FamHist, data = tech_2019)
summary(LMFamilyHist)
```

```
##
## Call:
## lm(formula = Age ~ FamHist, data = tech_2019)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.0357  -6.2393   0.0571   4.9643  17.9643
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     36.5000     2.9805  12.246   <2e-16 ***
## FamHistTRUE     -0.4643     3.1863  -0.146    0.885
## FamHistUnknown  -2.6500     3.5266  -0.751    0.455
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.43 on 81 degrees of freedom
## Multiple R-squared:  0.01341,    Adjusted R-squared:  -0.01095
## F-statistic: 0.5503 on 2 and 81 DF,  p-value: 0.5789
```

LinearModel for age, family history and treatment

```
LMFamHistTreat <- lm(Age ~ FamHist + Treatment, data = tech_2019)
summary(LMFamHistTreat)
```

```
##
## Call:
## lm(formula = Age ~ FamHist + Treatment, data = tech_2019)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.0357  -6.0962  -0.0357   4.7827  17.9643
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     42.2222     6.9619   6.065 4.15e-08 ***
## FamHistTRUE     -0.4643     3.1897  -0.146    0.885
## FamHistUnknown  -3.2222     3.5859  -0.899    0.372
## TreatmentTRUE   -5.7222     6.2902  -0.910    0.366
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.439 on 80 degrees of freedom
## Multiple R-squared:  0.02351,    Adjusted R-squared:  -0.01311
## F-statistic: 0.642 on 3 and 80 DF,  p-value: 0.5903
```

LinearModel for age, current role and treatment

```
LMRoleTreatment <- lm(Age ~ Treatment +  C_role, data = tech_2019)
summary(LMRoleTreatment)
```

```
##
## Call:
## lm(formula = Age ~ Treatment + C_role, data = tech_2019)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.9726  -5.9726   0.5274   5.0274  18.0274
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     34.472      6.599   5.223 1.33e-06 ***
## TreatmentTRUE   -3.027      5.984  -0.506    0.614
## C_roleTRUE       4.528      2.950   1.535    0.129
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.349 on 81 degrees of freedom
## Multiple R-squared:  0.03231,    Adjusted R-squared:  0.008419
## F-statistic: 1.352 on 2 and 81 DF,  p-value: 0.2644
```

LinearModel for age, gender, current role, treatment and job interference

```
LMInterfer <- lm(Age ~ Treatment + C_role + Gender_IDT + Interfer + Interfer_NT, data = tech_2019)
summary(LMInterfer)
```

```
##
## Call:
## lm(formula = Age ~ Treatment + C_role + Gender_IDT + Interfer +
##     Interfer_NT, data = tech_2019)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.3169  -6.1694  -0.0779   4.2500  19.7686
##
## Coefficients:
```

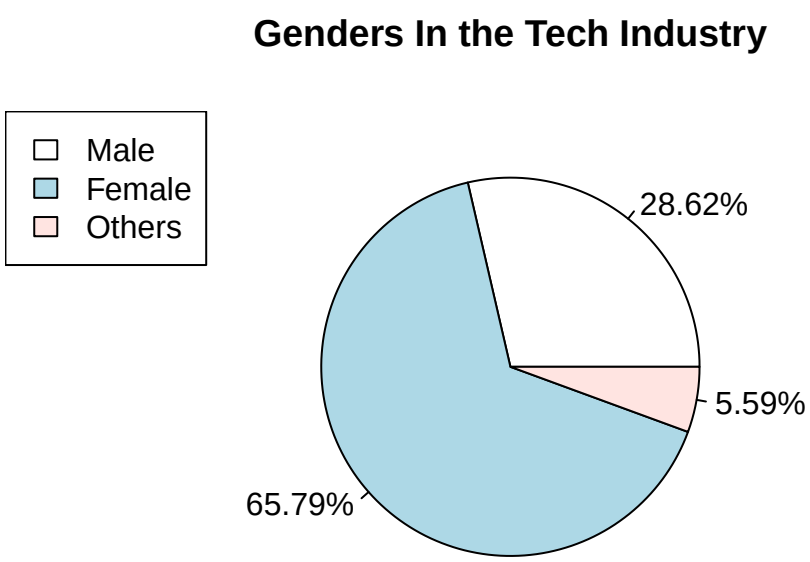
```
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     34.2578     7.4842   4.577 1.89e-05 ***
## TreatmentTRUE                    3.7797     6.8900   0.549   0.5850
## C_roleTRUE                       3.4713     3.1885   1.089   0.2799
## Gender_IDTMale                   1.9008     2.0389   0.932   0.3543
## Gender_IDTOther                  0.6412     4.0559   0.158   0.8748
## InterferNot applicable to me -10.4097     7.1057  -1.465   0.1472
## InterferOften                 -3.4084     5.4846  -0.621   0.5362
## InterferRarely                -7.3404     5.2163  -1.407   0.1636
## InterferSometimes             -7.2774     5.2554  -1.385   0.1703
## Interfer_NTRarely            -13.2551     6.8203  -1.943   0.0558 .
## Interfer_NTSometimes           2.1485     2.5007   0.859   0.3931
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.407 on 73 degrees of freedom
## Multiple R-squared:  0.1158, Adjusted R-squared:  -0.005321
## F-statistic: 0.9561 on 10 and 73 DF,  p-value: 0.4886
```

Percentage of genders Cleaned Primary Role

```
f <- table(tech_2019D$Gender_IDT)
f
```

```
##
## Female   Male  Other
##     87    200     17
```

```
pie_labels <- paste0(round(100 * f/sum(f), 2), "%")
pie(f,main = "Genders In the Tech Industry", labels = pie_labels)
legend("topleft", legend = c("Male", "Female", "Others"),
       fill =  c("white", "lightblue", "mistyrose"))
```

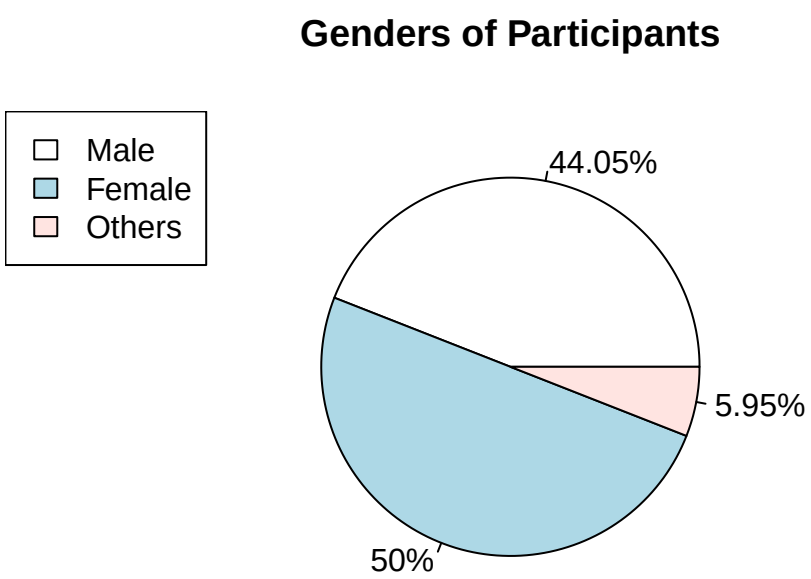


Survey

```
f <- table(tech_2019$Gender_IDT)
f
```

```
##
## Female   Male  Other
##     37     42      5
```

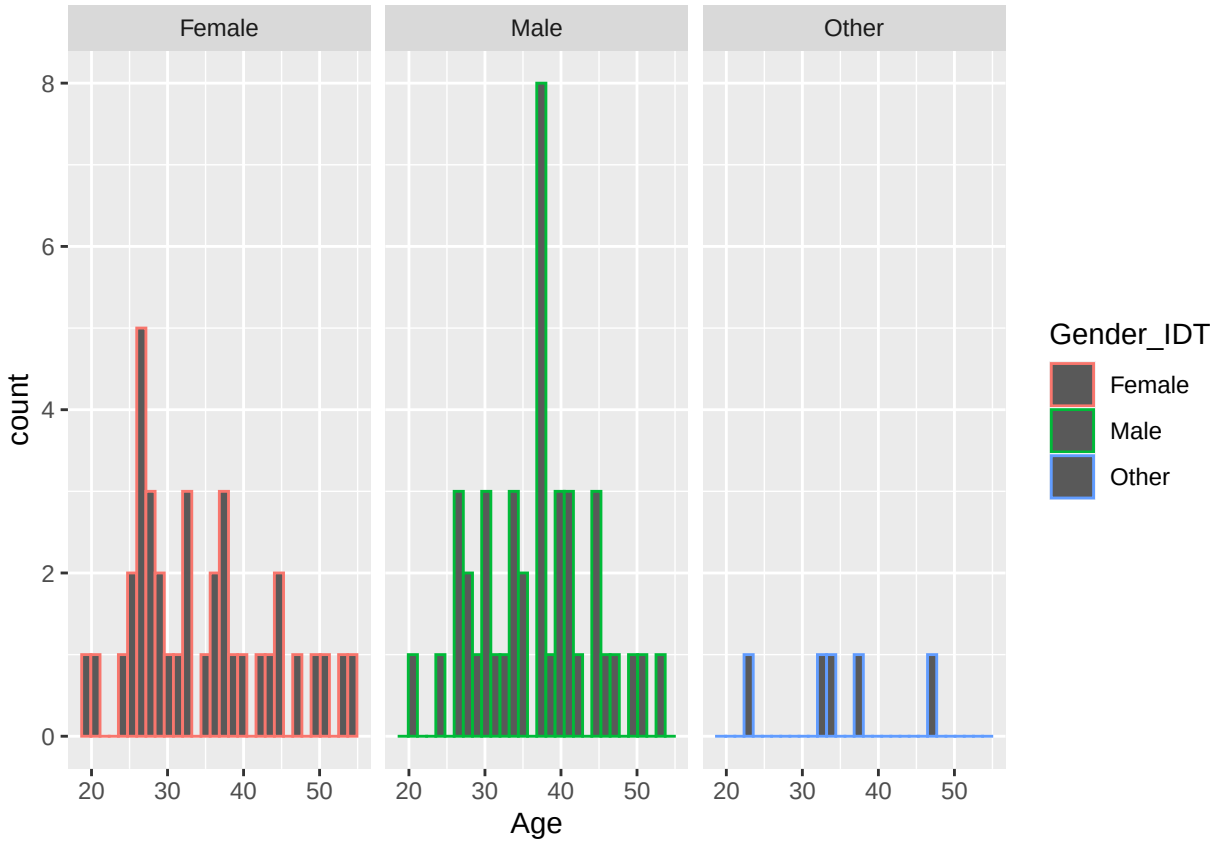
```
pie_labels <- paste0(round(100 * f/sum(f), 2), "%")
pie(f,main = "Genders of Participants", labels = pie_labels)
legend("topleft", legend = c("Male", "Female", "Others"),
       fill =  c("white", "lightblue", "mistyrose"))
```



Age and Gender

```
ggplot(tech_2019, aes(x=Age, color=Gender_IDT)) + geom_histogram() + facet_wrap(~Gender_IDT)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



% of those surveyed with a family history of mental health issues.

```
f <- table(tech_2019$FamHist)
f
```

```
##
##   FALSE    TRUE Unknown
##       8      56      20
```

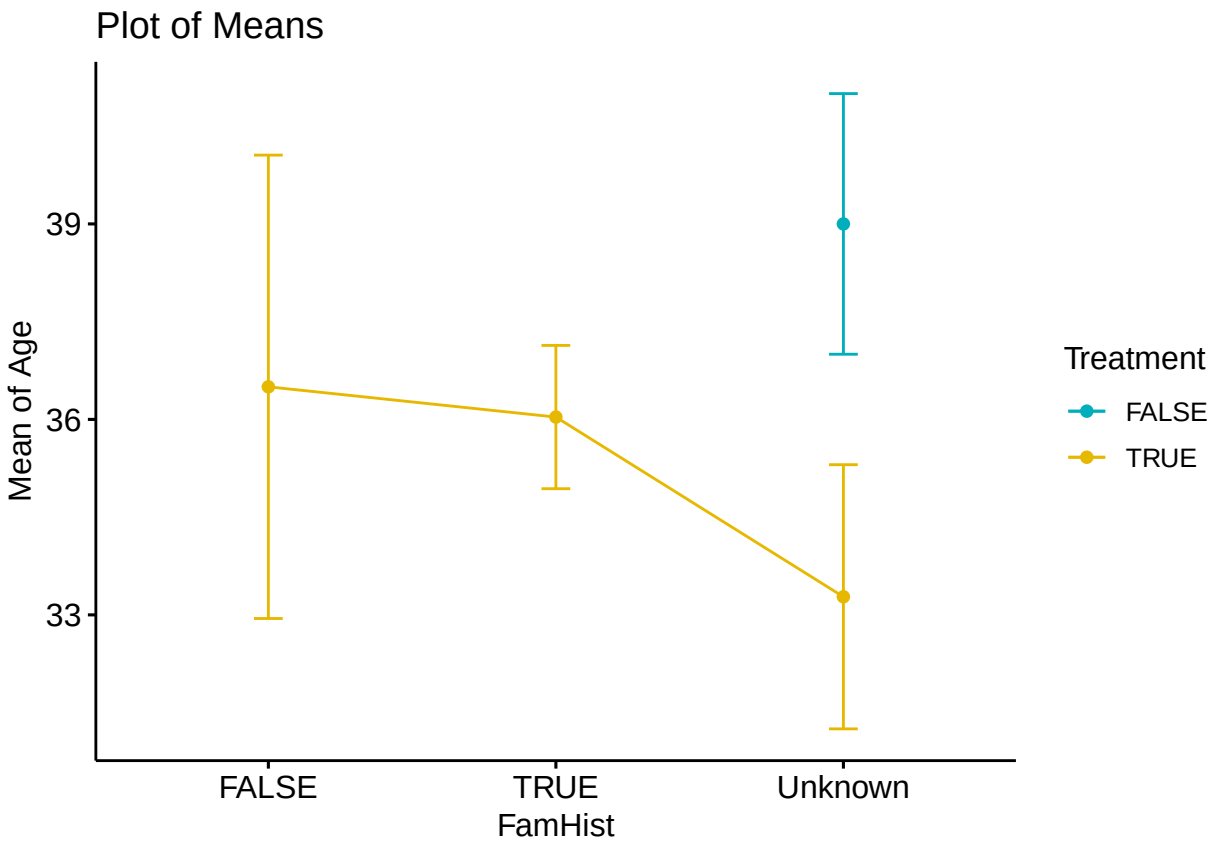
```
pie_labels <- paste0(round(100 * f/sum(f), 2), "%")
```

Treatment

```
library(gplots)
```

```
##
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':
##
##     lowess
```

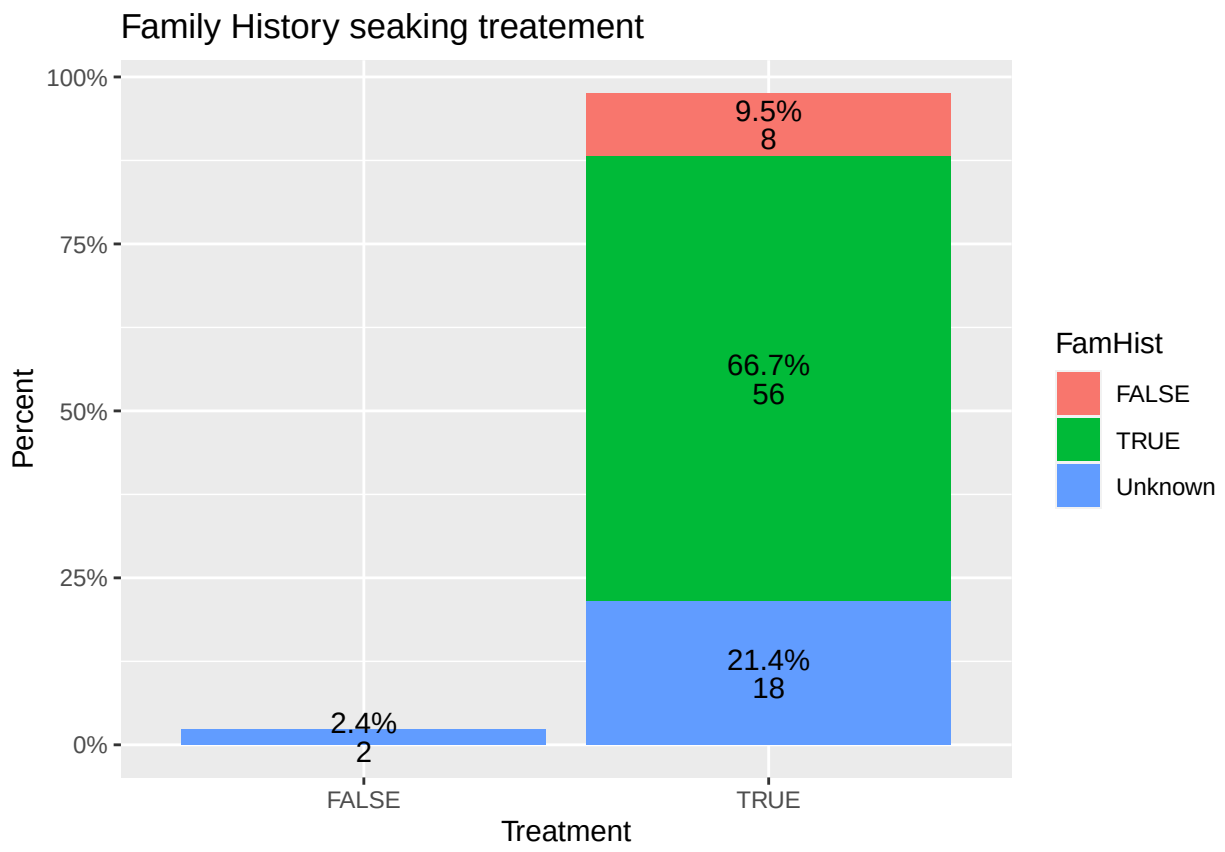
```
library(ggpubr)

ggline(tech_2019, x = "FamHist", y = "Age", color = "Treatment",
 add = "mean_se", palette = c("#00AFBB", "#E7B800"),
 main="Plot of Means",
 xlab="FamHist",
 ylab="Mean of Age",
 legend="right")
```



Percent of those with a Family History seaking treatement

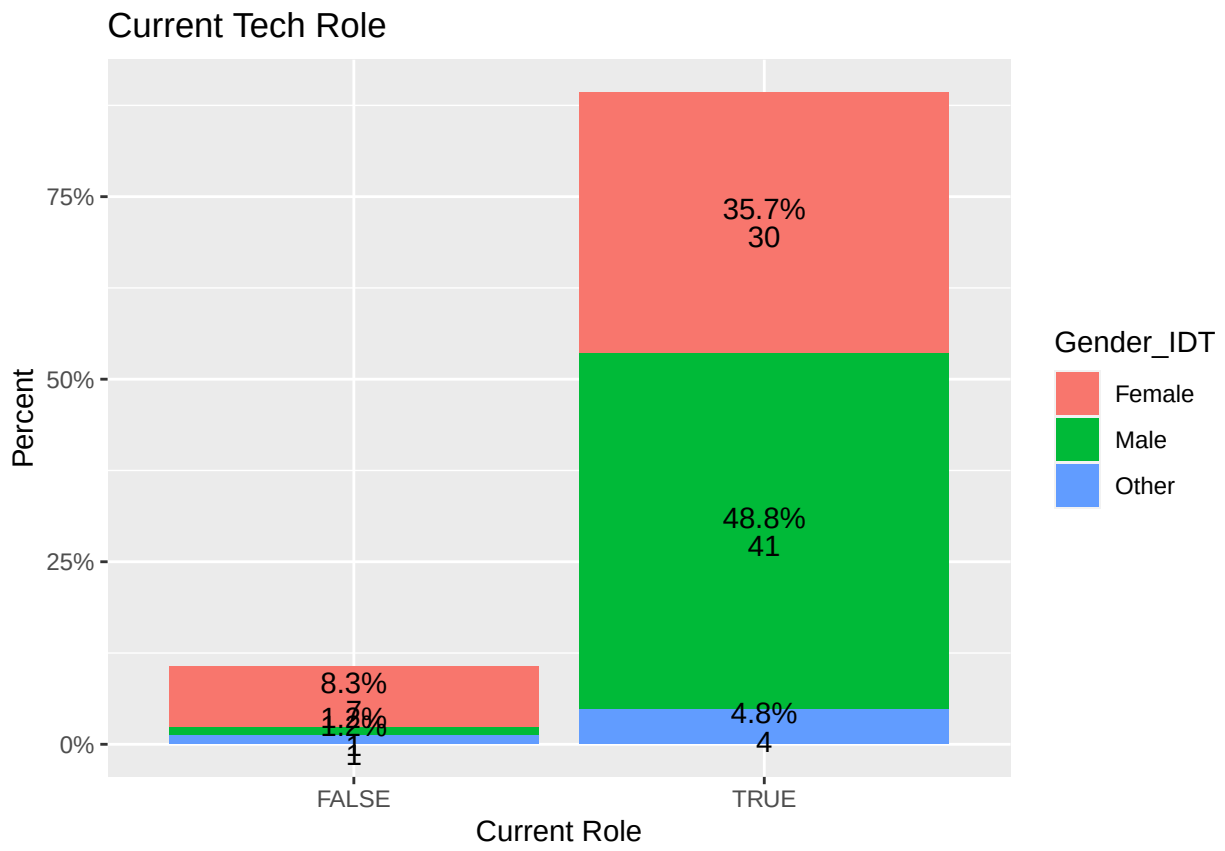
```
library(ggpubr)
library(dplyr)
tech_df <- tech_2019 %>%
  count(Treatment, FamHist) %>%
  mutate(pct = n / sum(n),
         pct_label = scales::percent(pct))

ggplot(tech_df, aes(x= Treatment, fill = FamHist, y = pct)) +
  geom_col() +
  geom_text(aes(label = paste(pct_label, n, sep = "\n")),
                lineheight = 0.8,
```

```
              position = position_stack(vjust = 0.5)) +
  scale_y_continuous(labels = scales::percent) +
   labs(title="Family History seaking treatement", x= "Treatment", y = "Percent")
```



Count in Current Tech role

```
library(ggpubr)
library(dplyr)
tech_df <- tech_2019 %>%
  count(C_role, Gender_IDT) %>%
  mutate(pct = n / sum(n),
         pct_label = scales::percent(pct))

ggplot(tech_df, aes(x= C_role, fill = Gender_IDT, y = pct)) +
  geom_col() +
  geom_text(aes(label = paste(pct_label, n, sep = "\n")),
                lineheight = 0.8,
                position = position_stack(vjust = 0.5)) +
  scale_y_continuous(labels = scales::percent) +
   labs(title="Current Tech Role", x= "Current Role", y = "Percent")
```

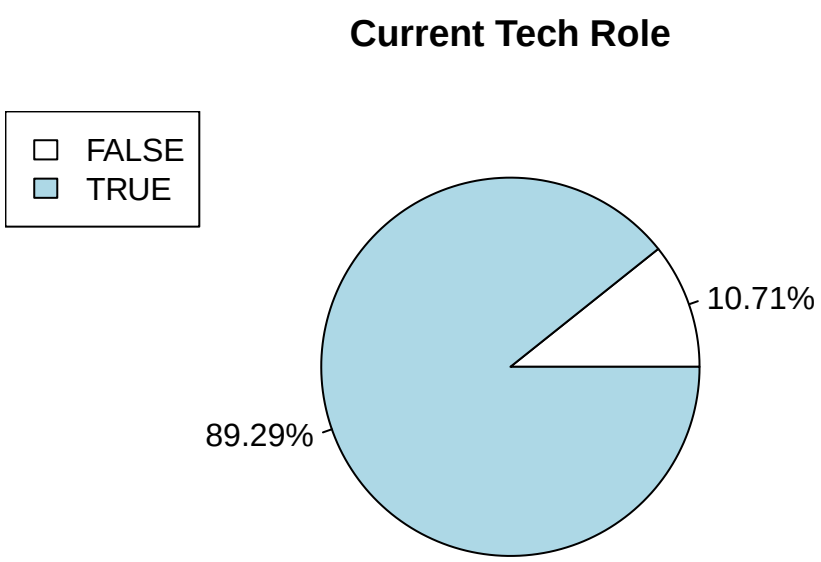


Current Tech Role

```
f <- table(tech_2019$C_role)
f
```

```
##
## FALSE  TRUE
##     9    75
```

```
pie_labels <- paste0(round(100 * f/sum(f), 2), "%")
pie(f,main = "Current Tech Role", labels = pie_labels)
legend("topleft", legend = c("FALSE", "TRUE"),
       fill =  c("white", "lightblue"))
```

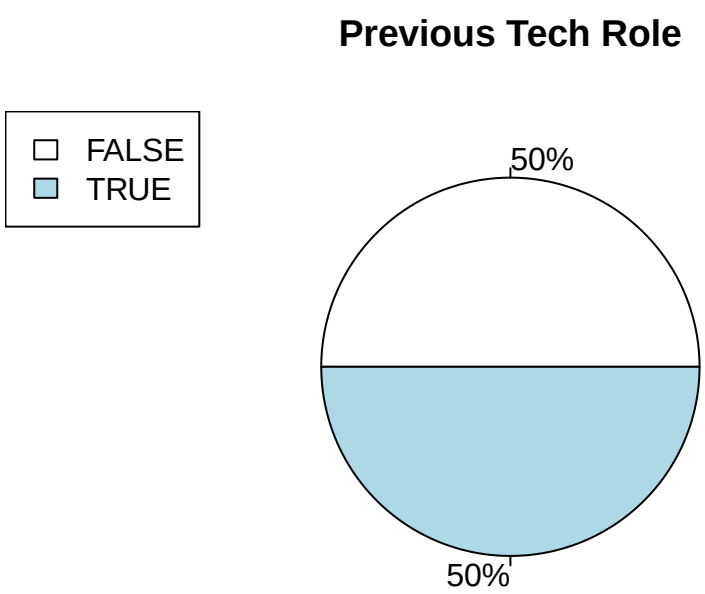


Previous Tech Role

```
f <- table(tech_2019$P_role)
f
```

```
##
## FALSE  TRUE
##    42    42
```

```
pie_labels <- paste0(round(100 * f/sum(f), 2), "%")
pie(f,main = "Previous Tech Role", labels = pie_labels)
legend("topleft", legend = c("FALSE", "TRUE"),
       fill =  c("white", "lightblue"))
```

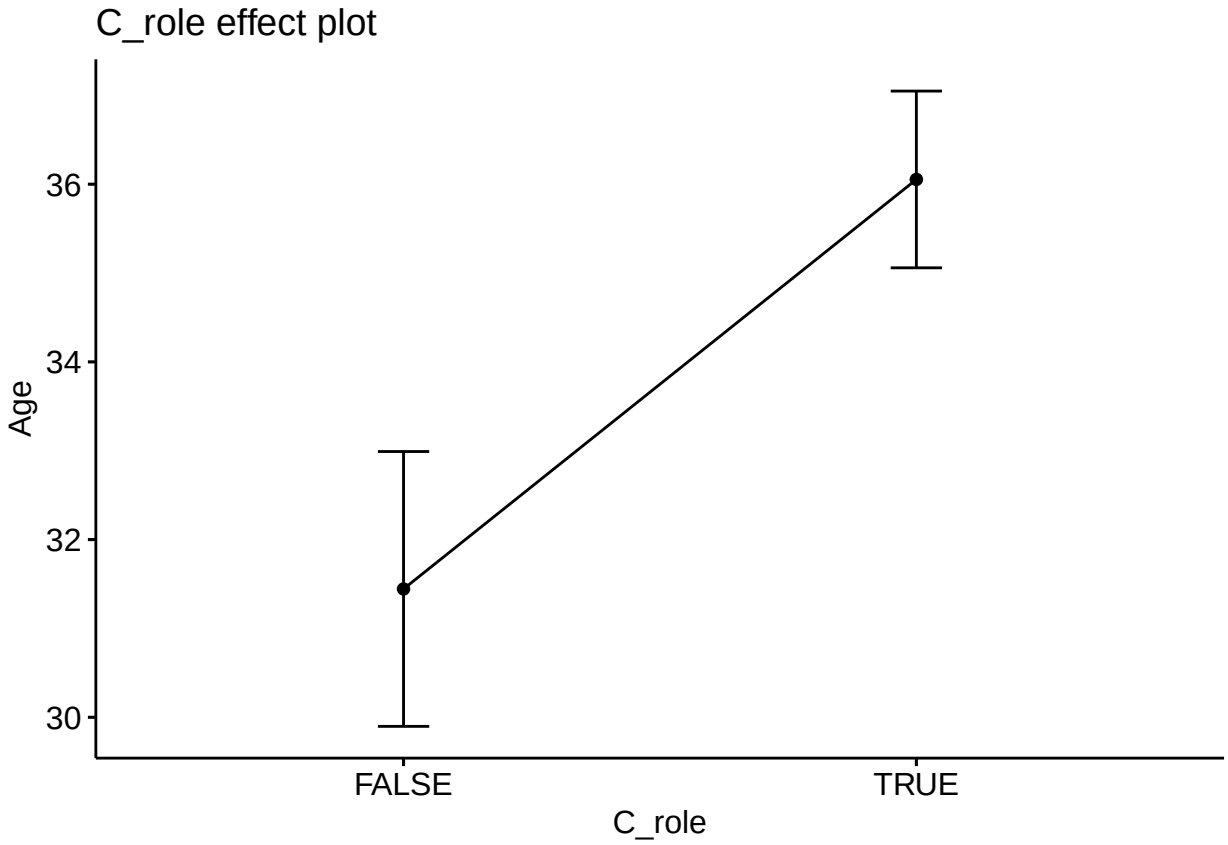


median age of those in the tech industry

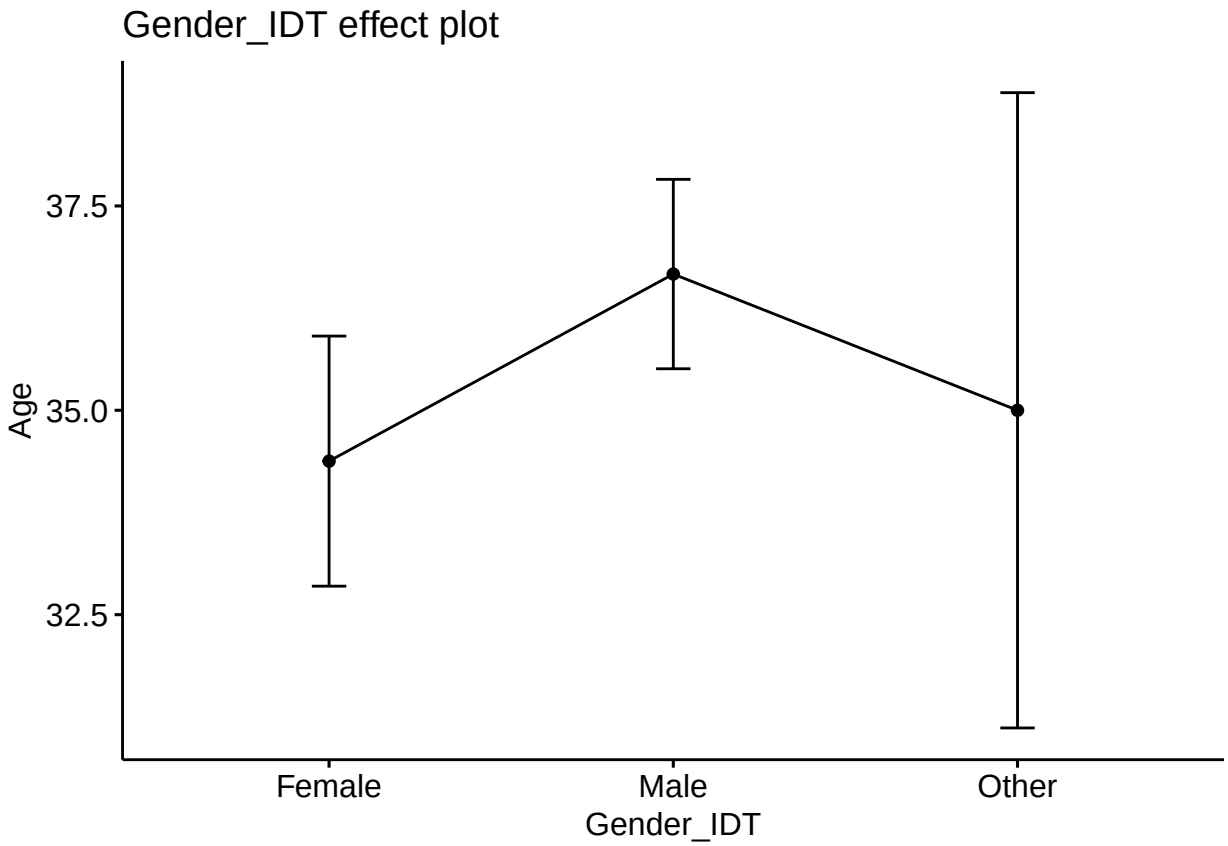
```
t.test(Age~C_role, alternative = 'two.sided', conf.level = .95, var.equal = FALSE, data = tech_2019)
```

```
##
##  Welch Two Sample t-test
##
## data:  Age by C_role
## t = -2.5064, df = 15.698, p-value = 0.02361
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.5132584 -0.7045194
## sample estimates:
## mean in group FALSE  mean in group TRUE
##            31.44444            36.05333
```

```
library(ggpubr)

ggline(tech_2019, x = "C_role", y = "Age",
 add = "mean_se", palette = c("#00AFBB", "#E7B800"),
 main="C_role effect plot",
 xlab="C_role",
 ylab="Age",
 legend="right")
```



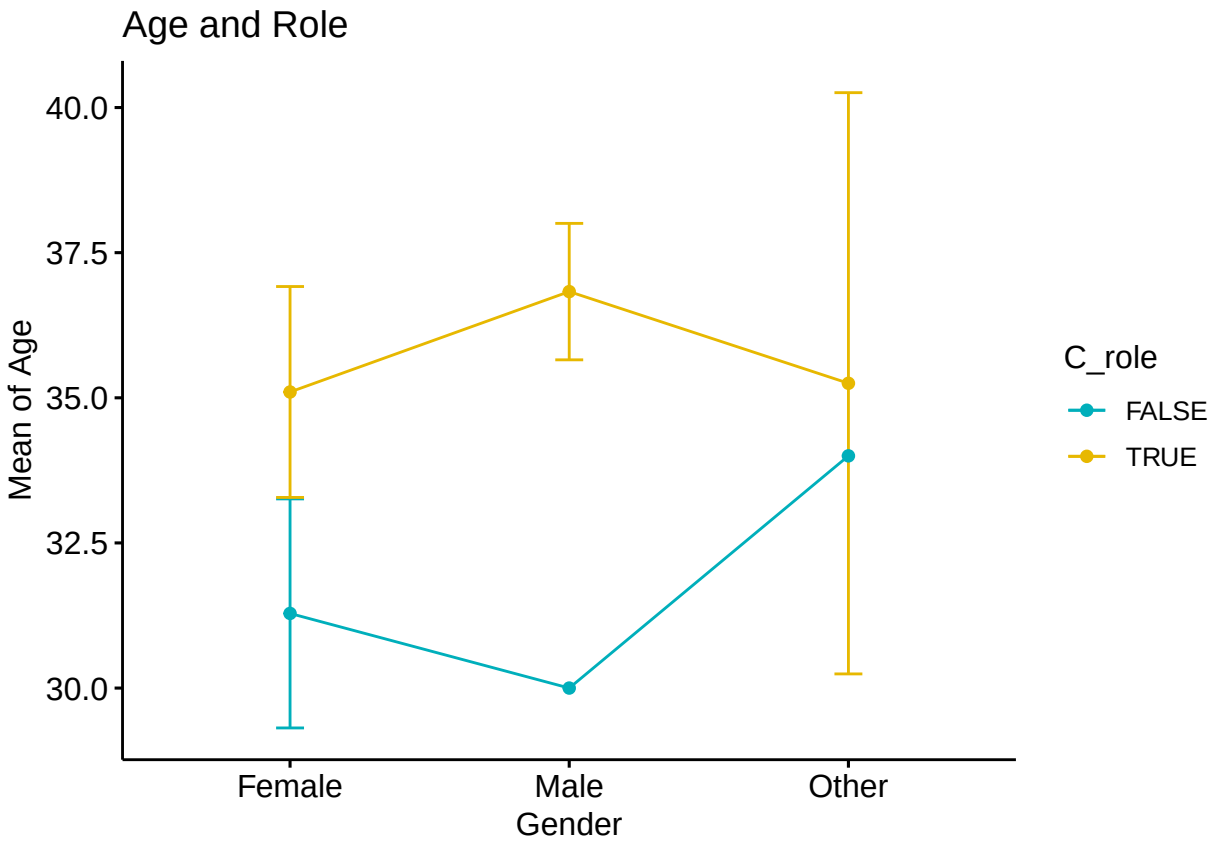

```
ggline(tech_2019, x = "Gender_IDT", y = "Age",
 add = "mean_se", palette = c("#00AFBB", "#E7B800"),
 main="Gender_IDT effect plot",
 xlab="Gender_IDT",
 ylab="Age",
 legend="right")
```




```
ggline(tech_2019, x = "Gender_IDT", y = "Age", color = "C_role",
 add = "mean_se", palette = c("#00AFBB", "#E7B800"),
 main="Age and Role",
 xlab="Gender",
 ylab="Mean of Age",
 legend="right")
```

```
## Warning in stats::qt(ci/2 + 0.5, data_sum$length - 1): NaNs produced
```

```
## Warning in stats::qt(ci/2 + 0.5, data_sum$length - 1): NaNs produced
```

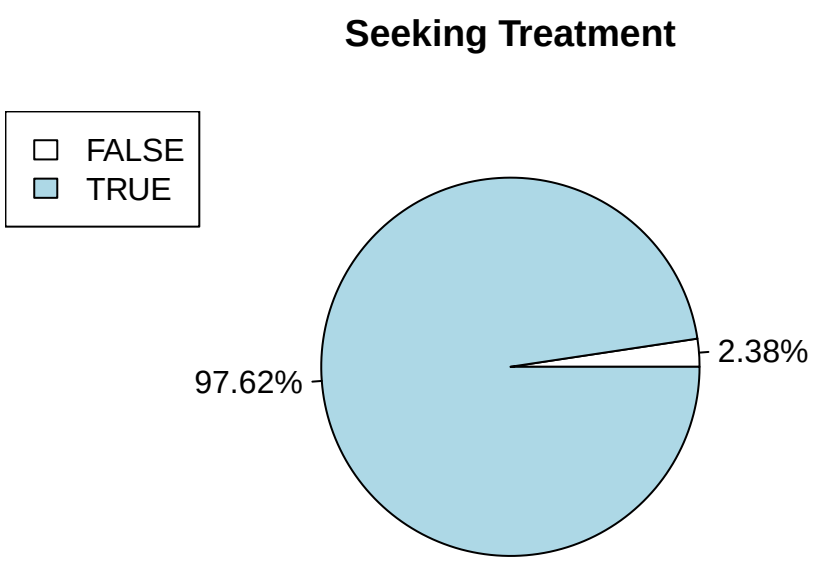


Percent seeking treatement

```
f <- table(tech_2019$Treatment)
f
```

```
##
## FALSE  TRUE
##     2    82
```

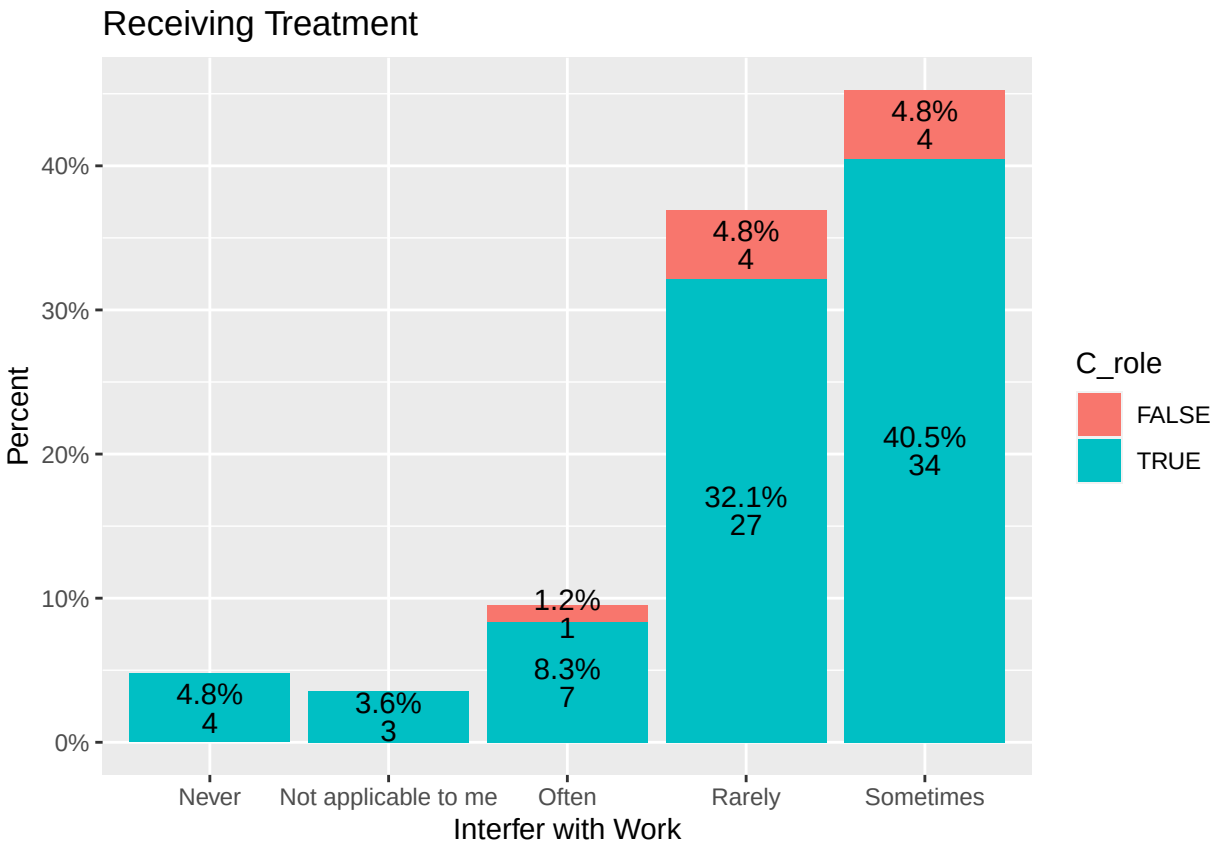
```
pie_labels <- paste0(round(100 * f/sum(f), 2), "%")
pie(f,main = "Seeking Treatment", labels = pie_labels)
legend("topleft", legend = c("FALSE", "TRUE"),
       fill =  c("white", "lightblue"))
```



Mental health interfers in work while being treated

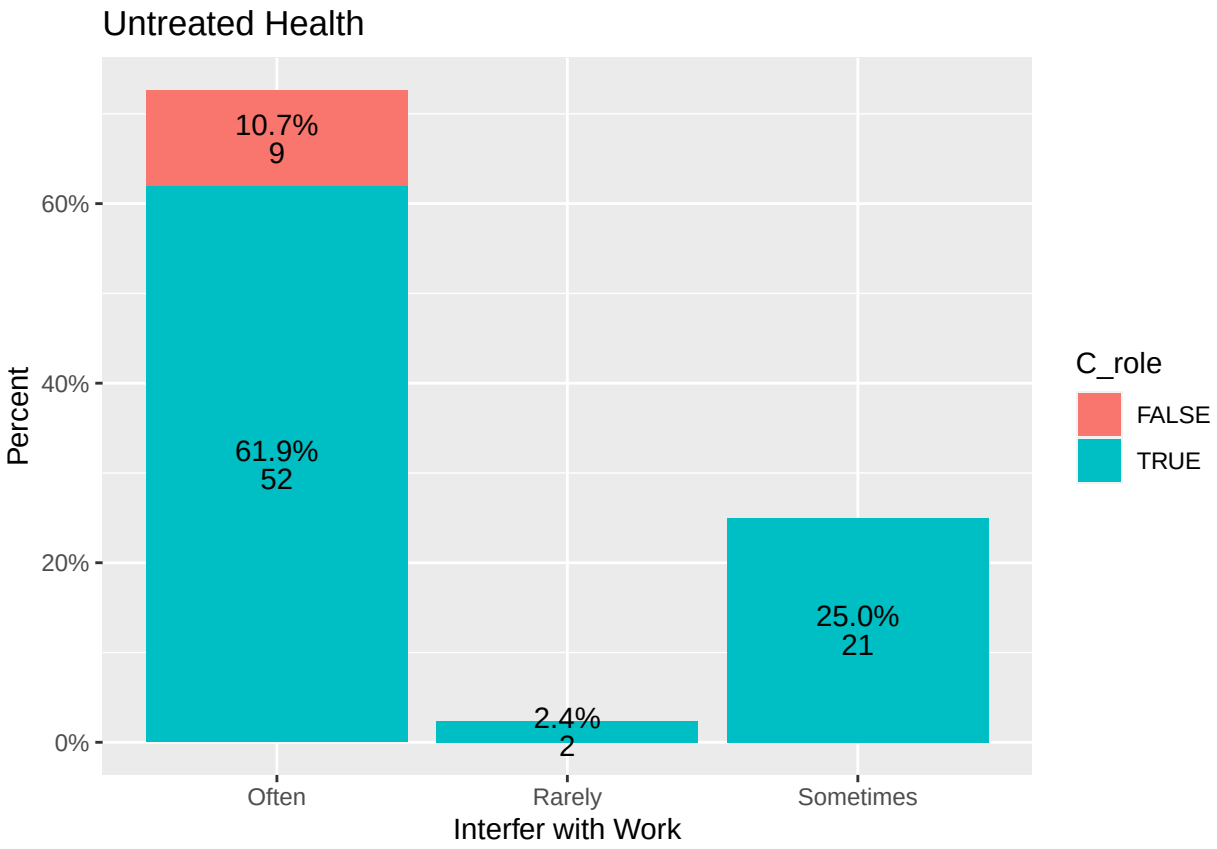
Receiving Treatment

```
library(dplyr)
tech_df <- tech_2019 %>%
  count(Interfer, C_role) %>%
  mutate(pct = n / sum(n),
         pct_label = scales::percent(pct))

ggplot(tech_df, aes(x= Interfer, fill = C_role, y = pct)) +
  geom_col() +
  geom_text(aes(label = paste(pct_label, n, sep = "\n")),
                lineheight = 0.8,
                position = position_stack(vjust = 0.5)) +
  scale_y_continuous(labels = scales::percent) +
   labs(title="Receiving Treatment", x= "Interfer with Work", y = "Percent")
```



while not recieving treatment

```
library(dplyr)
tech_df <- tech_2019 %>%
  count(Interfer_NT, C_role) %>%
  mutate(pct = n / sum(n),
         pct_label = scales::percent(pct))

ggplot(tech_df, aes(x= Interfer_NT, fill = C_role, y = pct)) +
  geom_col() +
  geom_text(aes(label = paste(pct_label, n, sep = "\n")),
                lineheight = 0.8,
                position = position_stack(vjust = 0.5)) +
  scale_y_continuous(labels = scales::percent) +
  labs(title="Untreated Health", x= "Interfer with Work", y = "Percent")
```


Untreated Health
10.7%
9
61.9%
52
2.4%
2
25.0%
21
60%
40%
20%
0%
Percent
Often
Rarely
Sometimes
Interfer with Work
C_role
FALSE
TRUE