Final Report Semester Project

# Loading the Data and Necessary Packages

knitr::opts\_chunk$set(comment = NA)  
library(dplyr)  
library(tidyr)  
library(ggplot2)  
library(grid)  
library(gridExtra)  
library(knitr)  
health <- read.csv("mentalsurvey\_clean.csv")

# Introduction

This document contains the fully commented analysis script and outputs, written summary of the project, and summary of contributorship to each part of the project.

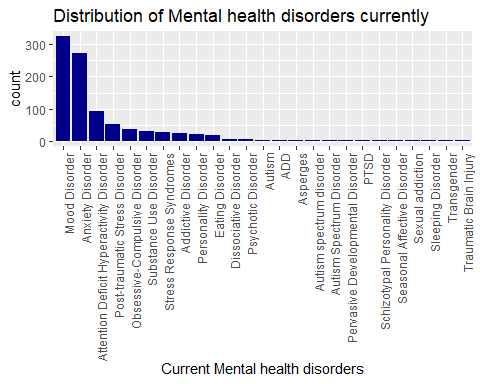
# Section 1: Distribution of mental health disorders based on different demographics

### Distribution of Mental health disorders

#since some people have multiple disorders, we have to split them   
tmp = health %>% separate(mhd,   
 sep = '\\|',  
 c('mhd\_1', 'mhd\_2', 'mhd\_3', 'mhd\_4', 'mhd\_5',  
 'mhd\_6', 'mhd\_7', 'mhd\_8', 'mhd\_9'),  
 fill = 'right')  
  
  
chk = tmp %>% select(matches('mhd\_[1-9]'))  
  
comb = as.vector(t(chk))  
comb1 = comb[comb != ""&!is.na(comb)]  
  
#Shortening the factor levels by splitting  
disorders=sapply(strsplit(comb1, split = " \\("), `[`, 1)  
dataf <- as.data.frame(matrix(disorders))  
colnames(dataf)<- "Disorder"  
#Removing inconsistent response  
subdataf <- subset(dataf, disorders != "I haven\'t been formally diagnosed, so I felt uncomfortable answering, but Social Anxiety and Depression.")  
count\_disorder <-data.frame(count(subdataf, Disorder, sort = TRUE))  
kable(count\_disorder)

|  |  |
| --- | --- |
| Disorder | n |
| Mood Disorder | 325 |
| Anxiety Disorder | 272 |
| Attention Deficit Hyperactivity Disorder | 93 |
| Post-traumatic Stress Disorder | 51 |
| Obsessive-Compulsive Disorder | 36 |
| Substance Use Disorder | 31 |
| Stress Response Syndromes | 28 |
| Addictive Disorder | 24 |
| Personality Disorder | 22 |
| Eating Disorder | 19 |
| Dissociative Disorder | 5 |
| Psychotic Disorder | 4 |
| Autism | 2 |
| ADD | 1 |
| Asperges | 1 |
| Autism spectrum disorder | 1 |
| Autism Spectrum Disorder | 1 |
| Pervasive Developmental Disorder | 1 |
| PTSD | 1 |
| Schizotypal Personality Disorder | 1 |
| Seasonal Affective Disorder | 1 |
| Sexual addiction | 1 |
| Sleeping Disorder | 1 |
| Transgender | 1 |
| Traumatic Brain Injury | 1 |

reorder\_size <- function(x) {  
 factor(x, levels = names(sort(table(x), decreasing = TRUE)))  
}  
#Plot to see the distribution of current mental health illness  
p <- ggplot(data=subdataf, aes(x = reorder\_size(Disorder)))+  
 geom\_bar(fill='dark blue')+  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1))  
  
p+xlab("Current Mental health disorders")+ggtitle("Distribution of Mental health disorders currently")



#### Interpretation of Results

The most common mental health disorder appears to be mood disorder with a count of 325, followed by anxiety disorder with a count of 272.

### Distribution of mental health status among gender

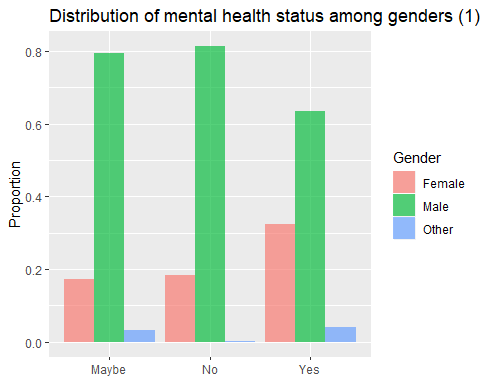
gender\_disorder <- table(health$Do.you.currently.have.a.mental.health.disorder., health$gender)  
gender\_disorder

Female Male Other  
 Maybe 44 202 8  
 No 81 359 1  
 Yes 146 286 19

#Table: % of who may, don't or have mental disorder among genders  
rowprop <- prop.table(gender\_disorder,1)   
rowprop

Female Male Other  
 Maybe 0.173228346 0.795275591 0.031496063  
 No 0.183673469 0.814058957 0.002267574  
 Yes 0.323725055 0.634146341 0.042128603

dr1 <-as.data.frame(rowprop)  
# Distribution of respondents among gender provided they may/have/donot have a mental diorder   
ggplot(data = dr1, aes(x =Var1 , y = Freq, fill = Var2)) +   
 geom\_bar(stat = 'identity', position = 'dodge', alpha = 2/3) +   
 labs(x = NULL, y = 'Proportion', fill = 'Gender')+ggtitle('Distribution of mental health status among genders (1)')



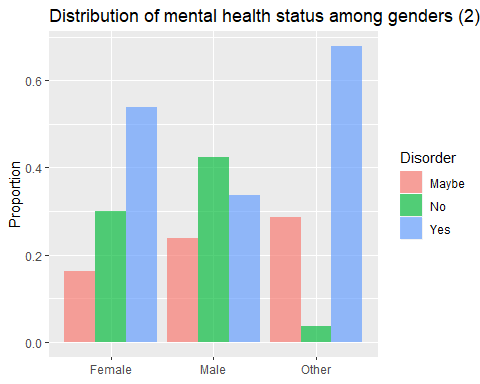
#### Interpretation of results

The table and figure above shows that among the people that said Yes, 32.3% are females, 63.4% are males and 4.2% are otherss. Among the people that said No, 81.4% are males, 18.3% are females and 0.2% are others.Among the people that said Maybe, 79.5% are males, 17.3% are females and 3.1% are others.

#Table : % of each level gender who may, don't or have mental disorder  
colprop <- prop.table(gender\_disorder,2)  
colprop

Female Male Other  
 Maybe 0.16236162 0.23848878 0.28571429  
 No 0.29889299 0.42384888 0.03571429  
 Yes 0.53874539 0.33766234 0.67857143

# Distribution of respondents who may/have/donot have a mental diorder over the gender  
dr2 <-as.data.frame(colprop)  
ggplot(data = dr2, aes(x =Var2 , y = Freq, fill = Var1)) +   
 geom\_bar(stat = 'identity', position = 'dodge', alpha = 2/3) +   
 labs(x = NULL, y = 'Proportion', fill = 'Disorder')+ggtitle('Distribution of mental health status among genders (2)')



#### Interpretation of results

The table and figure above show that the proportion of females who said Yes (0.53) are more than the proportion of females that said no (0.29) amd maybe (0.16). The proportion of males that said No (0.42) are more than the proportion of males who said Yes (0.33) and Maybe (0.23). In the gender Other, the proportion that said Yes (0.67) are more than the proportion that said No (0.03) and Maybe (0.28).

### Distribution of mental health disorder based on age groups

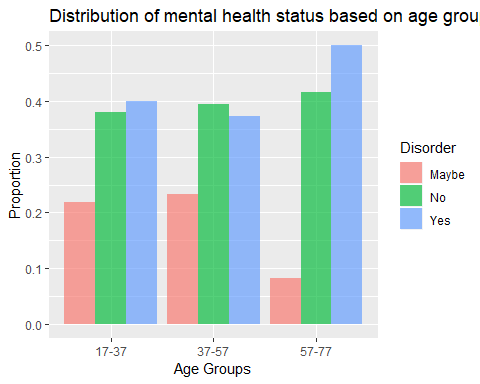
summary(health$age)

Min. 1st Qu. Median Mean 3rd Qu. Max.   
 17.00 28.00 32.00 33.37 38.00 74.00

#Creating labels for age groups  
labs <- c(paste(seq(17, 74, by = 20), seq(37, 80, by = 20),  
 sep = "-"))  
health$AgeGroup <- cut(health$age, breaks = c(seq(17, 74, by = 20), Inf), labels = labs, right = FALSE)  
  
agegr <- table(health$Do.you.currently.have.a.mental.health.disorder.,health$AgeGroup)  
ageprop <- prop.table(agegr,2)  
ageprop

17-37 37-57 57-77  
 Maybe 0.21933086 0.23241590 0.08333333  
 No 0.38042131 0.39449541 0.41666667  
 Yes 0.40024783 0.37308869 0.50000000

dfageprop <- as.data.frame(ageprop)  
  
#Distribution of respondents reporting a mental disorder by age groups  
ggplot(data = dfageprop, aes(x =Var2 , y = Freq, fill = Var1)) +   
 geom\_bar(stat = 'identity', position = 'dodge', alpha = 2/3) +   
 labs(x = "Age Groups", y = 'Proportion', fill = 'Disorder')+ggtitle('Distribution of mental health status based on age groups (1)')



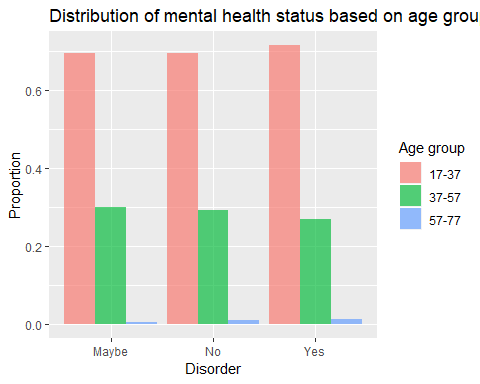
#### Interpretation of results

The table and plot above show that in the age group of 17-37, 40% said they currently have a mental disorder, 38% said they donot have a mental disorder currently and 21.9% said are not sure. In the 37-57 age group, 39.4% said they donot have a mental disorder currently, 37.3% said Yes and 23.2% said maybe or are not sure. In the 57-77 agegroup, 50% said Yes, 41.6% said No and 8.3% said maybe.

ageprop1 <- prop.table(agegr,1)  
ageprop1

17-37 37-57 57-77  
 Maybe 0.696850394 0.299212598 0.003937008  
 No 0.696145125 0.292517007 0.011337868  
 Yes 0.716186253 0.270509978 0.013303769

dfageprop1 <- as.data.frame(ageprop1)  
ggplot(data = dfageprop1, aes(x =Var1 , y = Freq, fill = Var2)) +   
 geom\_bar(stat = 'identity', position = 'dodge', alpha = 2/3) +   
 labs(x = "Disorder", y = 'Proportion', fill = 'Age group')+ggtitle('Distribution of mental health status based on age groups (2)')



#### Interepretion of results

The table and plot above show that among the people that said they have mental disorder currently, 71.6% lie in the 17-37 age group, 27% lie in the 37-57 age group and 1.3% lie in the age group of 57-77. Among the people that said No, 69.6% lie in the 17-37 age group, 29.2% lie in the 37-57 age group and 1.1% lie in the 57-77 group. Among the people that said Maybe, 69.6% liein the 17-37 age group, 29.9% lie in the 37-57 age group and .3% lie in the 57-77 age group.

### Distribution of mental health disorder based on geography

table(health$What.US.state.or.territory.do.you.live.in. , health$Do.you.currently.have.a.mental.health.disorder.)

Maybe No Yes  
 105 209 123  
 Alabama 0 1 3  
 Alaska 1 1 0  
 Arizona 0 4 0  
 California 22 35 53  
 Colorado 7 6 8  
 Connecticut 0 2 1  
 Delaware 0 0 1  
 District of Columbia 0 0 2  
 Florida 4 7 6  
 Georgia 5 3 3  
 Idaho 0 0 3  
 Illinois 5 18 24  
 Indiana 3 7 11  
 Iowa 0 1 4  
 Kansas 3 7 4  
 Kentucky 1 0 1  
 Louisiana 0 0 2  
 Maine 1 1 3  
 Maryland 4 6 3  
 Massachusetts 5 8 6  
 Michigan 11 15 17  
 Minnesota 6 13 19  
 Missouri 1 6 3  
 Montana 0 1 0  
 Nebraska 1 5 4  
 Nevada 1 1 0  
 New Hampshire 1 2 2  
 New Jersey 1 0 4  
 New Mexico 0 0 1  
 New York 7 9 19  
 North Carolina 4 5 11  
 North Dakota 2 0 1  
 Ohio 6 5 10  
 Oklahoma 1 7 4  
 Oregon 4 11 13  
 Pennsylvania 4 8 18  
 Rhode Island 0 2 1  
 South Carolina 0 0 1  
 South Dakota 1 1 1  
 Tennessee 5 9 13  
 Texas 12 7 18  
 Utah 2 0 4  
 Vermont 0 3 0  
 Virginia 6 3 6  
 Washington 10 11 14  
 West Virginia 1 0 0  
 Wisconsin 1 1 6

#### Interpretation of results

From the above table output, we can see that the US state with the majority of mental health illness cases is California, followed by Illoinois and New York

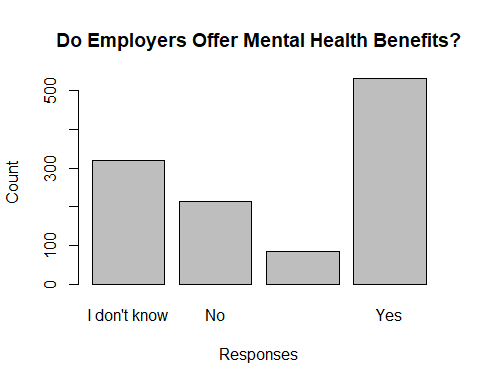
# Section 2: Analysis of distribution of mental health benefits offered by companies in the dataset

### Distribution of employers offering mental healh benefits as part of healthcare coverage

#Creating a table for the variable "Does your employer provide mental health benefits as part of healthcare coverage?""  
  
table(health$Does.your.employer.provide.mental.health.benefits.as.part.of.healthcare.coverage.)

I don't know No   
 319 213   
Not eligible for coverage / N/A Yes   
 83 531

barplot(table(health$Does.your.employer.provide.mental.health.benefits.as.part.of.healthcare.coverage.), xlab = "Responses", ylab = "Count", main = "Do Employers Offer Mental Health Benefits?")



#### Interpretation of results

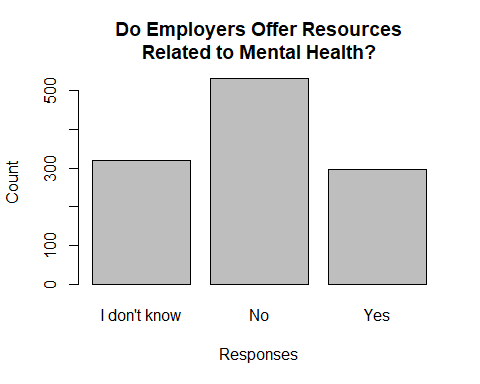
From the above output, we can see that out of a total 1433 companies, only 531 provide mental health benefits as part of healthcare coverage, which is approximately one-third of the responses.

### Distribution of employers offering resources to learn more about mental health concerns and options for seeking help

table(health$Does.your.employer.offer.resources.to.learn.more.about.mental.health.concerns.and.options.for.seeking.help.)

I don't know No Yes   
 320 531 295

barplot(table(health$Does.your.employer.offer.resources.to.learn.more.about.mental.health.concerns.and.options.for.seeking.help.), xlab = "Responses", ylab = "Count", main = "Do Employers Offer Resources\nRelated to Mental Health?")



#### Interpretation of results

From the above output, we can see that a very small number of companies from the dataset of 1433 entries offer resources to employees to learn more about mental health concerns and options for seeking help.i.e., only 295 employers. For the dataset under study, this implies that a very small number of employers are taking steps to aquaint their employees with mental health benefits in the respective companies. Also, we can see that a significant number of employees, i.e., 320 out of 1433, do not know if their employers offer resporces to seek help regading mental health concerns. This also implies lack of initiative in generating awareness about mental health-related policies.

### Do Companies in the survey dataset have formal discussions to Assist workers with Mental and Behavioral Health Policies

For companies with mental health care options available, let us find out what percentage have a formal discussion (as part of a campaign/official communication) with employees to generate awareness about the mental health care policies. Here, we can create subset for companies with mental health care options available. This dataset(subset) will contain only those companies that provide mental health benefits as part of healthcare coverage

Company<-subset(health, Does.your.employer.provide.mental.health.benefits.as.part.of.healthcare.coverage.=="Yes")

#### Percentage of employers who have a formal discussion about mental health with their employees

#Creating a table for for the variable related to employers formally discussing mental health with employees  
table(Company$Has.your.employer.ever.formally.discussed.mental.health..for.example..as.part.of.a.wellness.campaign.or.other.official.communication..)

I don't know No Yes   
 59 331 141

#Computing the percentage  
percent<-table(Company$Has.your.employer.ever.formally.discussed.mental.health..for.example..as.part.of.a.wellness.campaign.or.other.official.communication..)/length(Company$Has.your.employer.ever.formally.discussed.mental.health..for.example..as.part.of.a.wellness.campaign.or.other.official.communication..)  
#The parameter percent gives us the value of the percenntage needed  
percent

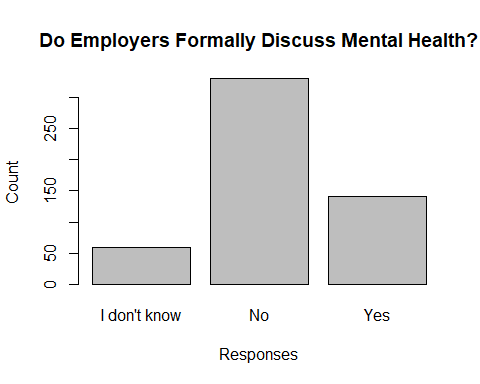
I don't know No Yes   
 0.1111111 0.6233522 0.2655367

#### Interpretation of results

The value of the parameter percent in the above output for the column “Yes” gives us the percentage of employers who have formally discussed mental health with their employees (considering only those companies who provide mental health benefits to their employees). Thus, about 26.5 percent of employers hve a formal discussion about mental health and related policies with their employees. For the dataset under study, this is a very small number. In order to make the employees familiar with the mental health policies and benefits, it is crucial that the employers have a formal discussion with them.

#### Visualizing the proportion of responses for employers formally discussing mental health with theri employees

Company<-droplevels(Company)  
barplot(table(Company$Has.your.employer.ever.formally.discussed.mental.health..for.example..as.part.of.a.wellness.campaign.or.other.official.communication..), xlab = "Responses", ylab = "Count", main = "Do Employers Formally Discuss Mental Health?")



# Section 3: What are Factors that Might Limit the Impact of Behavioral Health Outreach

### Potential Negative Consequences for discussing Health Issues with Employer: Physical vs. Mental

If employees feel that discussing physical or mental health issues would have negative consequences. If so, they may be less likely to reach out for help.

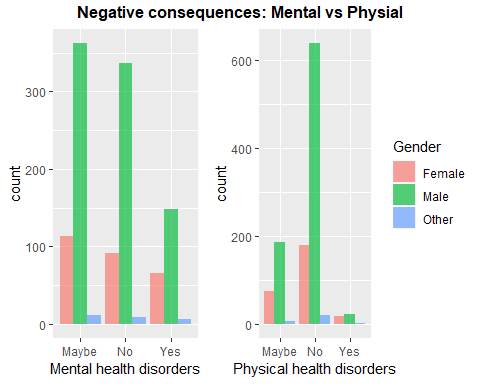
knitr::opts\_chunk$set(echo = TRUE)  
#Creating a table with gender and mental health responses  
tab <- table(health$Do.you.think.that.discussing.a.mental.health.disorder.with.your.employer.would.have.negative.consequences., health$gender)  
tab

Female Male Other  
 Maybe 113 362 12  
 No 92 337 9  
 Yes 66 148 7

#Creating a table with gender and physical health responses  
tab1 <- table(health$Do.you.think.that.discussing.a.physical.health.issue.with.your.employer.would.have.negative.consequences., health$gender)  
tab1

Female Male Other  
 Maybe 75 187 6  
 No 179 638 20  
 Yes 17 22 2

#creating the combined plot  
tabdf <- as.data.frame(tab)  
tab1df <- as.data.frame(tab1)  
  
plot1 <- ggplot(data = tabdf, aes(x = Var1, y =Freq, fill = Var2)) +   
 geom\_bar(stat = 'identity', position = 'dodge', alpha = 2/3) +   
 labs(x = "Mental health disorders", y = 'count')+ theme(legend.position='none')  
  
plot2 <- ggplot(data = tab1df, aes(x = Var1, y =Freq, fill = Var2)) +   
 geom\_bar(stat = 'identity', position = 'dodge', alpha = 2/3) +   
 labs(x = "Physical health disorders", y = 'count', fill = 'Gender')  
title1=textGrob("Negative consequences: Mental vs Physial", gp=gpar(fontface="bold"))  
grid.arrange(plot1, plot2,ncol=2,top = title1, widths = c(3/4,1))



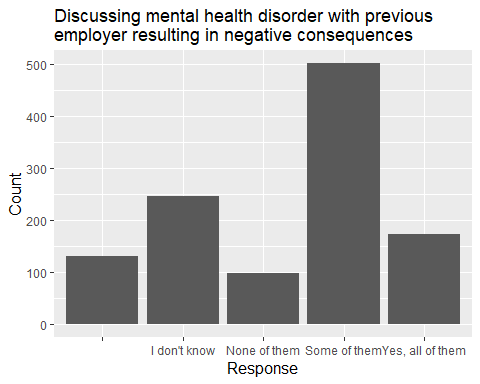
#### Interpretation of Results

The figure above shows that majority of people are certain that discussing physical health issues with their employers will not result in negative consequences. Very few believe that discussing physical health issues will have negative consequences. The responses to whether discussing mental health disorders with employers will result in negative consequences is more divided. Some people believe that discussing mental health disorders will have a negative consequence. This is probably due to the stigma at workplace which may make them believe that they might face negative consequences if their mental issues are brought infront of a employer. This will make them less likely to reach out for help, discuss their issues with their employers, take time off work. However, there are still more people who believe there won’t be any negative consequences. This could be because they have already developed a smooth employer-employee relationship. There are comparatively more number of people who are unsure of whether or not there would be any negative consequences if mental health disorders are discussed with their employers.

### Analzing the relationship between perceiving the negative impact of mental illness at previous workplace and the impact at current workplace

It is likely that if employees perceived a negative impact from discussing mental illness at a previous workplace, they may have similar perception at the current workplace as well. We can plot the ggplot for the independent variable “Do you think that discussing a mental health disorder with previous employers would have negative consequences?”

ggplot(health, aes(x = health$Do.you.think.that.discussing.a.mental.health.disorder.with.previous.employers.would.have.negative.consequences.)) + geom\_bar() + ggtitle("Discussing mental health disorder with previous\nemployer resulting in negative consequences") + xlab("Response") + ylab("Count") + theme(axis.title=element\_text(size=12))



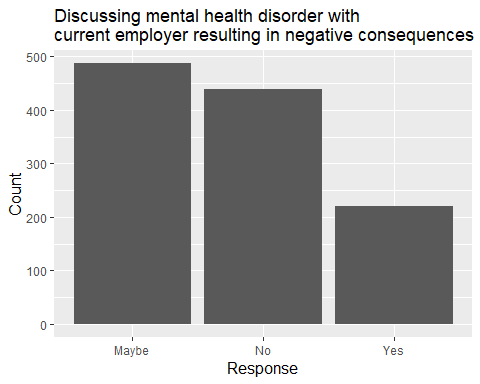
summary(health$Do.you.think.that.discussing.a.mental.health.disorder.with.previous.employers.would.have.negative.consequences.)

I don't know None of them Some of them   
 131 245 97 501   
Yes, all of them   
 172

From the above ggplot, we can observe that a majority of respondents(615) feel that discussing mental health disorders with some of their previous employers resulted in negative consequences for them. While a significant number of respondents responded that they do not know, we can conclude from the graph that the majority did experience negative consequences after discussing mental health disorders with their previous employers

Plotting the ggplot for the dependent variable “Do you think that discussing a mental health disorder with your employer would have negative consequences?”

ggplot(health, aes(x = health$Do.you.think.that.discussing.a.mental.health.disorder.with.your.employer.would.have.negative.consequences.)) + geom\_bar() + ggtitle("Discussing mental health disorder with\ncurrent employer resulting in negative consequences") + xlab("Response") + ylab("Count") + theme(axis.title=element\_text(size=12))



summary(health$Do.you.think.that.discussing.a.mental.health.disorder.with.your.employer.would.have.negative.consequences.)

Maybe No Yes   
 487 438 221

From the above ggplot, we can see that a vast majority of respondents(487) are not sure if discussing mental health disorders with their current employers would have negative consequences; followed by about 438 respondents who do not think that it could lead to negative consequences.

#### Chi-squared test

m1<-table(health$Do.you.think.that.discussing.a.mental.health.disorder.with.previous.employers.would.have.negative.consequences., health$Do.you.think.that.discussing.a.mental.health.disorder.with.your.employer.would.have.negative.consequences.)  
chisq.test(m1)

Pearson's Chi-squared test  
  
data: m1  
X-squared = 158.59, df = 8, p-value < 2.2e-16

#### Interpretation of results

We can see from the above chi-squared test that the p-value is very small, meaning we can reject the hypothesis of independence. Thus there is a relationship between perceiving the negative impact of mental illness at previous workplace and perceiving the same at the current workplace. Employees who have perceived or experienced a negative impact of discussing a mental illness with their previous employer feel the same way about discussing it with their current employer. This type of attitude may be the result of some form of negative impact such as bias, discrimination or lack of oppurtunities that they faced at the previous workplace. This might have influenced their decision to discusstheir mental illnesses with their current employer as well.

## Anonymity Policies

If employees feel that their anonymity is not protected, they may be less likely to reach out for help.

### Anonymity Policies by Company Size

#### Crosstab Table

#Put company size in order  
health$CompanySize = factor(health$How.many.employees.does.your.company.or.organization.have, levels=c("1-5", "6-25", "26-100", "100-500", "500-1000", "More than 1000"))  
#Create Cross-tab Table  
APbCS <- table(health$Is.your.anonymity.protected.if.you.choose.to.take.advantage.of.mental.health.or.substance.abuse.treatment.resources.provided.by.your.employer., health$CompanySize)  
head(APbCS)

1-5 6-25 26-100 100-500 500-1000 More than 1000  
 I don't know 31 133 196 166 56 160  
 No 20 24 18 8 2 12  
 Yes 9 53 78 74 22 84

#### Proportions table

#Create a proportions table by Row  
prop.table(APbCS, 2)

1-5 6-25 26-100 100-500 500-1000  
 I don't know 0.51666667 0.63333333 0.67123288 0.66935484 0.70000000  
 No 0.33333333 0.11428571 0.06164384 0.03225806 0.02500000  
 Yes 0.15000000 0.25238095 0.26712329 0.29838710 0.27500000  
   
 More than 1000  
 I don't know 0.62500000  
 No 0.04687500  
 Yes 0.32812500

#### Chi-Squared Test

#Do a chi-squared test  
chisq.test(APbCS)

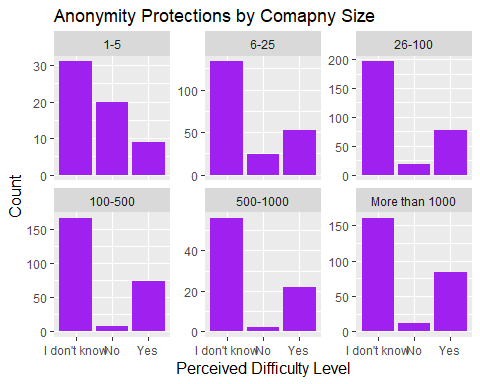
Warning in chisq.test(APbCS): Chi-squared approximation may be incorrect

Pearson's Chi-squared test  
  
data: APbCS  
X-squared = 80.824, df = 10, p-value = 3.461e-13

We get a warning that the approximation may be incorrect, likely because some of the values from the table are small. Therefore, we cannot draw and conclusions for this test about whether these two variables are related.

#### Visualization of Data

t0 <- ggplot(health, aes(x = health$Is.your.anonymity.protected.if.you.choose.to.take.advantage.of.mental.health.or.substance.abuse.treatment.resources.provided.by.your.employer.))+geom\_bar(fill="purple")+facet\_wrap(~health$CompanySize, scales="free\_y") + ggtitle("Anonymity Protections by Comapny Size") + xlab("Perceived Difficulty Level") + ylab("Count")+ theme(axis.title=element\_text(size=12))  
t0



#### Interpretation of Results

The table and chart above show that across all company sizes, most respondents indicate that they do not know what their current company’s anonymity policy is towards those using mental health services. For companies of 1-5 employees, the next most common answer is “No” (anonymity is not protected). However, for companies larger than 5 employees, the results appear fairly consistent, with the majority of respondents (60-70 percent) indicating that they do not know anonymity policies, a sizable minority of respondents saying that anonymity is protected (25-32 percent), and less than 10 percent of respondents in each group saying anonymity is not protected. Based off of the chi-squared test, it is not clear if there is a statistically significant difference in anonymity policies of companies of different sizes. However, the fact that the majority of respondents across company sizes do not know their company’s anonymity policy is something that has to be addressed in any outreach program, as a study by Milne, et al (1994) has found that participants tended to be more confident in a company’s Employee Assistance Program if they felt it was confidential (p. 141). Therefore, any outreach must make the company’s anonymity policy clear.

### Anonymity Policies by Whether it is a Tech Company

#### Cross-tab Table

#Add labels to tech company org  
health$TLabel <- factor(health$Is.your.employer.primarily.a.tech.company.organization.,  
levels = c(0, 1),  
labels = c("Not primarily tech", "Primarily Tech"))

#Create a crosstab table  
APbT2 <- table(health$Is.your.anonymity.protected.if.you.choose.to.take.advantage.of.mental.health.or.substance.abuse.treatment.resources.provided.by.your.employer., health$TLabel)  
head(APbT2)

Not primarily tech Primarily Tech  
 I don't know 169 573  
 No 15 69  
 Yes 79 241

#### Proportions table

#Create a proportions table by Row  
prop.table(APbT2, 2)

Not primarily tech Primarily Tech  
 I don't know 0.64258555 0.64892412  
 No 0.05703422 0.07814270  
 Yes 0.30038023 0.27293318

#### Chi-Squared Test

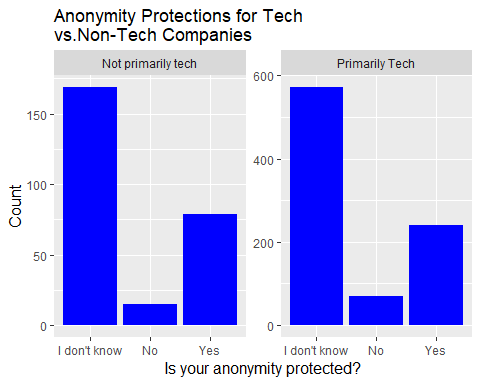
#Do a chi-squared test  
chisq.test(APbT2)

Pearson's Chi-squared test  
  
data: APbT2  
X-squared = 1.7911, df = 2, p-value = 0.4084

This chi-squared test’s p value is very large. Therefore, there is likely not a relationship between these two variables.

#### Visualization of Data

#Create Visualization  
t1 <- ggplot(health, aes(x = health$Is.your.anonymity.protected.if.you.choose.to.take.advantage.of.mental.health.or.substance.abuse.treatment.resources.provided.by.your.employer.))+geom\_bar(fill="blue")+facet\_wrap(health$TLabel ~ ., scales="free\_y") + ggtitle("Anonymity Protections for Tech\nvs.Non-Tech Companies") + xlab("Is your anonymity protected?") + ylab("Count") + theme(axis.title=element\_text(size=12))  
t1



#### Interpretation of Results

The difference between primarily tech and non-primarily tech companies is not statistically significant in regards to anonymity policies for those seeking mental health services. However, this may be influenced by the fact that the majority of those surveyed work for primiarily tech-oriented companies, so we do not have a large sample of those working for non-tech companies. Furthermore, we do not know the industry that these non-tech companies are in, which means we cannot draw any firm conclusions here. Nevertheless, in both cases, the clear majority of those surveyed indicated that they do not know their current company’s anonymity policy is towards those using mental health services. In both cases, the next most common response to the question is that anonymity is protected for those using mental health services. As described above, the fact the majority of respondents do not know their company’s anonymity policy is something that has to be addressed.

## Ease of Asking for Leave For A Mental Health Issue

This is important because if employees feel that asking for time off from work for medical leave, they may be less likely to seek the help they need.

### Answers by Gender

#Modify Labels (ideal if we can get it to fit)  
health$pd2 <- factor(health$If.a.mental.health.issue.prompted.you.to.request.a.medical.leave.from.work..asking.for.that.leave.would.be.,  
levels = c("Very easy", "Somewhat easy", "Neither easy nor difficult","Somewhat difficult", "Very difficult", "I don't know"),  
labels = c("Very easy (VE)", "Somewhat Easy (SE)", "Neutral (N)", "Somewhat Hard (SH)", "Very Hard (VH)", "Don't Know (DK)"))  
#Modify Labels (if health$pd2 does not fit, use this as an alternate)  
health$pd3 <- factor(health$If.a.mental.health.issue.prompted.you.to.request.a.medical.leave.from.work..asking.for.that.leave.would.be.,  
levels = c("Very easy", "Somewhat easy", "Neither easy nor difficult","Somewhat difficult", "Very difficult", "I don't know"),  
labels = c("VE", "SE", "N", "SH", "VH", "DK"))

#Crosstab table  
ELDbG <- table(health$pd2, health$gender)  
ELDbG

Female Male Other  
 Very easy (VE) 48 168 4  
 Somewhat Easy (SE) 60 213 8  
 Neutral (N) 38 137 3  
 Somewhat Hard (SH) 58 136 5  
 Very Hard (VH) 39 76 3  
 Don't Know (DK) 28 117 5

prop.table(ELDbG, 2)

Female Male Other  
 Very easy (VE) 0.17712177 0.19834711 0.14285714  
 Somewhat Easy (SE) 0.22140221 0.25147580 0.28571429  
 Neutral (N) 0.14022140 0.16174734 0.10714286  
 Somewhat Hard (SH) 0.21402214 0.16056671 0.17857143  
 Very Hard (VH) 0.14391144 0.08972845 0.10714286  
 Don't Know (DK) 0.10332103 0.13813459 0.17857143

#### Chi-Squared

We get a warning that the approximation may be incorrect, likely because some of the values from the table are small. However, it potentially indicates that that the differences between genders may not be statistically significant (or we do not have enough information)

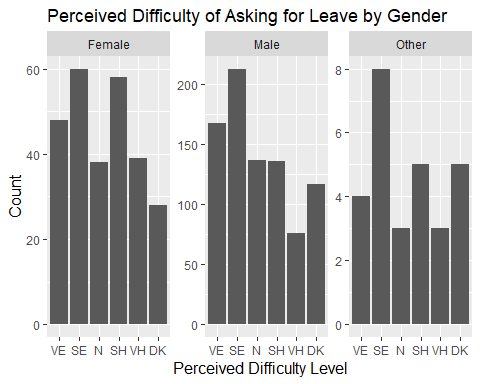
chisq.test(ELDbG)

Warning in chisq.test(ELDbG): Chi-squared approximation may be incorrect

Pearson's Chi-squared test  
  
data: ELDbG  
X-squared = 14.474, df = 10, p-value = 0.1524

#### Visualization of Data

#Create Visualization  
t2 <- ggplot(health, aes(x = health$pd3))+geom\_bar()+facet\_wrap(health$gender ~ ., scales="free\_y") + ggtitle("Perceived Difficulty of Asking for Leave by Gender") + xlab("Perceived Difficulty Level") + ylab("Count") + theme(axis.title=element\_text(size=12))  
t2

 NOTE: VE = Very Easy, SE = Somewhat Easy, N = Neutral (or Neither Easy nor Difficult), SH = Somewhat Hard, VH = Very Hard, DK = Don’t Know.

#### Interpretation of Results

Across all genders, the most common answer is “Somewhat Easy”. Interestingly, for women the second most common answer is “Somewhat Difficult”, while for men the second most common answer is “Very Easy.” For those indicating their gender falls into another category, an equal amount indicated that they felt asking for leave was very hard or provided a neutral response (“neither easy or hard”). Across all genders there is a significant group of people indicating they find it somewhat to very hard to ask for leave. That being said, there are fewer women and those from other genders participating in this survey (compared to men), which may impact our results. Furthermore, the chi-squared test indicated that the approximation may be incorrect, likely because some of the values from the table are small.

### Answers by Age Group

#Do the AgeGroup Variable  
labs <- c(paste(seq(17, 74, by = 20), seq(37, 80, by = 20),  
 sep = "-"))  
health$AgeGroup <- cut(health$age, breaks = c(seq(17, 74, by = 20), Inf), labels = labs, right = FALSE)

#Crosstab table  
ELDbA <- table(health$pd2, health$AgeGroup)  
ELDbA

17-37 37-57 57-77  
 Very easy (VE) 157 62 1  
 Somewhat Easy (SE) 203 77 1  
 Neutral (N) 126 48 4  
 Somewhat Hard (SH) 134 63 2  
 Very Hard (VH) 82 33 3  
 Don't Know (DK) 105 44 1

#Proportions table  
prop.table(ELDbA)

17-37 37-57 57-77  
 Very easy (VE) 0.1369982548 0.0541012216 0.0008726003  
 Somewhat Easy (SE) 0.1771378709 0.0671902269 0.0008726003  
 Neutral (N) 0.1099476440 0.0418848168 0.0034904014  
 Somewhat Hard (SH) 0.1169284468 0.0549738220 0.0017452007  
 Very Hard (VH) 0.0715532286 0.0287958115 0.0026178010  
 Don't Know (DK) 0.0916230366 0.0383944154 0.0008726003

#### Chi-Squared

chisq.test(ELDbA)

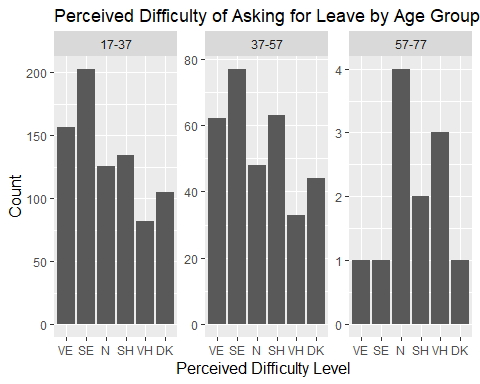
Warning in chisq.test(ELDbA): Chi-squared approximation may be incorrect

Pearson's Chi-squared test  
  
data: ELDbA  
X-squared = 8.6657, df = 10, p-value = 0.5641

We get a warning that the approximation may be incorrect, likely because some of the values from the table are small. However, it potentially indicates that that the differences between age groups may not be statistically significant (or we do not have enough information)

#### Visualization of Data

#Create Visualization  
t3 <- ggplot(health, aes(x = health$pd3))+geom\_bar()+facet\_wrap(health$AgeGroup ~ ., scales="free\_y") + ggtitle("Perceived Difficulty of Asking for Leave by Age Group") + xlab("Perceived Difficulty Level") + ylab("Count") + theme(axis.title=element\_text(size=12))  
t3

 NOTE: VE = Very Easy, SE = Somewhat Easy, N = Neutral (or Neither Easy nor Difficult), SH = Somewhat Hard, VH = Very Hard, DK = Don’t Know.

#### Interpretation of Results

The most common response for those in the 17-27 and 37-57 age ranges is that asking for leave for a mental health issue was “Somewhat Easy” and the second most common isthat asking for leave “Very Easy”. However, there is a significant number in both age groups who provided a neutral answer (e.g., neither easy or difficult) or indicated that they felt asking for leave was hard/very hard. That being said, there are few people who are 57-77 participating in this survey and there are fewer people in the 37-57 age group than the 17-27 age group. Furthermore, the chi-squared test indicated that the approximation may be incorrect, likely because some of the values from the table are small (since few people fell into the 57-77 age range).

### Answers by Company Size

#Crosstab table  
ELDbCS <- table(health$pd2, health$CompanySize)  
ELDbCS

1-5 6-25 26-100 100-500 500-1000 More than 1000  
 Very easy (VE) 16 42 59 48 15 40  
 Somewhat Easy (SE) 12 41 74 62 30 62  
 Neutral (N) 6 40 43 30 14 45  
 Somewhat Hard (SH) 11 45 46 48 6 43  
 Very Hard (VH) 13 16 36 15 5 33  
 Don't Know (DK) 2 26 34 45 10 33

prop.table(ELDbCS, 2)

1-5 6-25 26-100 100-500  
 Very easy (VE) 0.26666667 0.20000000 0.20205479 0.19354839  
 Somewhat Easy (SE) 0.20000000 0.19523810 0.25342466 0.25000000  
 Neutral (N) 0.10000000 0.19047619 0.14726027 0.12096774  
 Somewhat Hard (SH) 0.18333333 0.21428571 0.15753425 0.19354839  
 Very Hard (VH) 0.21666667 0.07619048 0.12328767 0.06048387  
 Don't Know (DK) 0.03333333 0.12380952 0.11643836 0.18145161  
   
 500-1000 More than 1000  
 Very easy (VE) 0.18750000 0.15625000  
 Somewhat Easy (SE) 0.37500000 0.24218750  
 Neutral (N) 0.17500000 0.17578125  
 Somewhat Hard (SH) 0.07500000 0.16796875  
 Very Hard (VH) 0.06250000 0.12890625  
 Don't Know (DK) 0.12500000 0.12890625

#### Chi-Squared

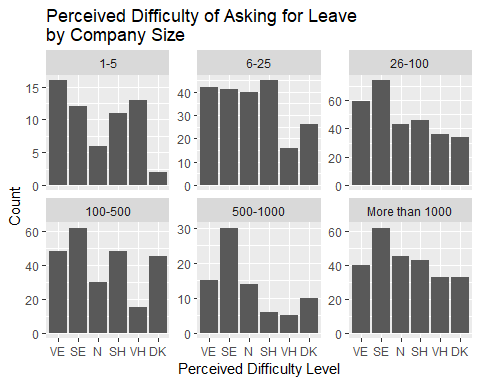
chisq.test(ELDbCS)

Pearson's Chi-squared test  
  
data: ELDbCS  
X-squared = 52.482, df = 25, p-value = 0.001041

Because the p value is less than .01 we can reject the null and say that we think that there is a relationship between these two variables.

#### Visualization of Data

#Create Visualization  
t4 <- ggplot(health, aes(x = health$pd3))+geom\_bar()+facet\_wrap(health$CompanySize~ ., scales="free\_y") + ggtitle("Perceived Difficulty of Asking for Leave\nby Company Size") + xlab("Perceived Difficulty Level") + ylab("Count")  
t4

 NOTE: VE = Very Easy, SE = Somewhat Easy, N = Neutral (or Neither Easy nor Difficult), SH = Somewhat Hard, VH = Very Hard, DK = Don’t Know.

#### Interpretation of Results

For very small companies (1-5 employees), there seems to be a fairly even split between those who perceive taking leave for mental health to be easy and those who perceive it to be hard. For other small companies (6-25 employees), approximately the same number of people say that it is easy/somewhat easy or provide a neutral response. However, in this case, the most common answer is that they would perceive it as somehwat hard to ask for leave. For larger companies, the most common answer is that they perceive it to be somewhat easy to request leave for a mental illness. Therefore it appears that as the company size gets larger, it may be perceived as easier in most companies to request leave related to a mental illness. However, there is still sizable minority of employees who work for larger companies (over 25 employees) who do perceive it as hard or very hard to ask for leave. Therefore, a lot may vary by the company.

### Answers by Whether Current Company Provides Mental Health Benefits

#Modify Labels for mental health benefit provision  
health$mhb2 <- factor(health$Does.your.employer.provide.mental.health.benefits.as.part.of.healthcare.coverage.,  
levels = c("I don't know", "No", "Not eligible for coverage / N/A","Yes"),  
labels = c("Don't Know", "No", "Ineligible/NA", "Yes"))

#Crosstab Table  
ELDbMHB <- table(health$pd2, health$mhb2)  
ELDbMHB

Don't Know No Ineligible/NA Yes  
 Very easy (VE) 55 36 18 111  
 Somewhat Easy (SE) 81 46 17 137  
 Neutral (N) 52 29 16 81  
 Somewhat Hard (SH) 51 49 17 82  
 Very Hard (VH) 25 31 10 52  
 Don't Know (DK) 55 22 5 68

#### Chi-Squared

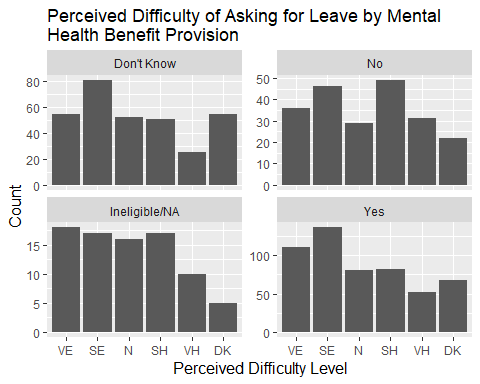
chisq.test(ELDbMHB)

Pearson's Chi-squared test  
  
data: ELDbMHB  
X-squared = 25.93, df = 15, p-value = 0.03877

Because the p value is not less than .01 we cannot reject the null. The null hypothesis is that there is no relationship between these two variables.

#### Create Visualization

t5 <- ggplot(health, aes(x = health$pd3))+geom\_bar()+facet\_wrap(~ health$mhb2, scales="free\_y") + ggtitle("Perceived Difficulty of Asking for Leave by Mental\nHealth Benefit Provision") + xlab("Perceived Difficulty Level") + ylab("Count") + theme(axis.title=element\_text(size=12))  
t5

 NOTE: VE = Very Easy, SE = Somewhat Easy, N = Neutral (or Neither Easy nor Difficult), SH = Somewhat Hard, VH = Very Hard, DK = Don’t Know.

#### Interpretation of Results

For those who do not get mental health benefits, more participants responded that they perceive it as somewhat difficult to request leave for a mental health issue than any other response (with somewhat easy being the second most common). Among those who do get mental health benefits, most indicated they perceive it as very easy or somwhat easy to request leave for a mental health issue. Most of those who did not know if mental health benefits are offered inidicated they perceived it to be somewhat easy to request leave for a mental health issue. Those who are ineligible (or for whom it was not applicable) were evently split among the responses, although relatively few people selected this response compared to the others. However, the difference between the two variables does not appear to be statistically significant.

### Answers by Whether Current Company Has Ever Formally Discussed Mental Health

ELDbEDM <- table(health$pd2, health$Has.your.employer.ever.formally.discussed.mental.health..for.example..as.part.of.a.wellness.campaign.or.other.official.communication..)  
ELDbEDM

I don't know No Yes  
 Very easy (VE) 23 129 68  
 Somewhat Easy (SE) 25 183 73  
 Neutral (N) 20 126 32  
 Somewhat Hard (SH) 13 161 25  
 Very Hard (VH) 3 105 10  
 Don't Know (DK) 19 109 22

#### Chi-Squared

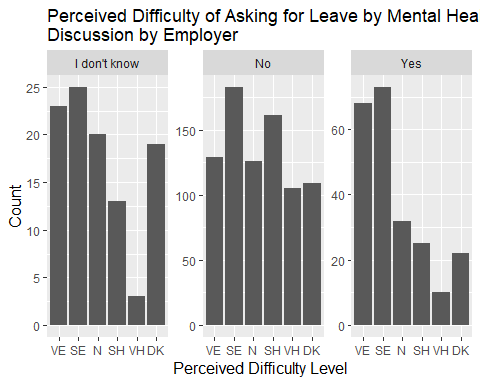
chisq.test(ELDbEDM)

Pearson's Chi-squared test  
  
data: ELDbEDM  
X-squared = 58.706, df = 10, p-value = 6.364e-09

Because the p value is less than .01 we can reject the null and say that we think that there is a relationship between these two variables.

#### Create Visualization

t6 <- ggplot(health, aes(x = health$pd3))+geom\_bar()+facet\_wrap( health$Has.your.employer.ever.formally.discussed.mental.health..for.example..as.part.of.a.wellness.campaign.or.other.official.communication.. ~., scales="free\_y") + ggtitle("Perceived Difficulty of Asking for Leave by Mental Health\nDiscussion by Employer") + xlab("Perceived Difficulty Level") + ylab("Count") + theme(axis.title=element\_text(size=12))  
t6

 NOTE: VE = Very Easy, SE = Somewhat Easy, N = Neutral (or Neither Easy nor Difficult), SH = Somewhat Hard, VH = Very Hard, DK = Don’t Know.

#### Interpretation of Results

Across all categories (e.g., yes they have discussed, no they haven’t, don’t know), the most common response is that most recipients think it would be somewhat easy to request leave. Interestingly, for those who did not know if there has been a discussion, a large contingent also indicated they don’t know how hard it would be to request leave. For those where there has been a discussion, far fewer people say it would be hard to request leave to handle a mental illness than those who say it would be easy to very easy. For those where there has not been a discussion (which is the majority of respondents), the second most common answer is that it would be somewhat hard to request leave to hanlde a mental illness.

### Answers by Whether Current Company Offers Other Resources

ELDbEOR <- table(health$pd2, health$Does.your.employer.offer.resources.to.learn.more.about.mental.health.concerns.and.options.for.seeking.help.)  
ELDbEOR

I don't know No Yes  
 Very easy (VE) 59 84 77  
 Somewhat Easy (SE) 80 123 78  
 Neutral (N) 49 80 49  
 Somewhat Hard (SH) 52 110 37  
 Very Hard (VH) 30 74 14  
 Don't Know (DK) 50 60 40

#### Chi-Squared

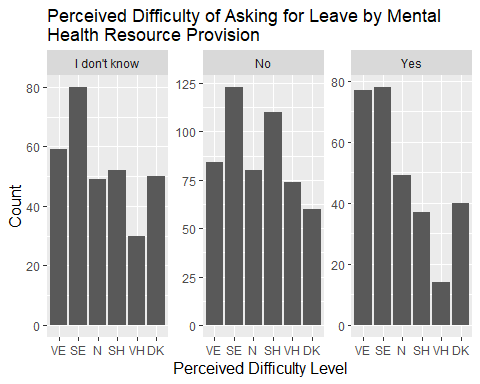
Because the p value is less than .01 we can reject the null and say that we think that there are differences between categories.

chisq.test(ELDbEOR)

Pearson's Chi-squared test  
  
data: ELDbEOR  
X-squared = 38.212, df = 10, p-value = 3.486e-05

#### Create Visualization

t7 <- ggplot(health, aes(x = health$pd3))+geom\_bar()+facet\_wrap( health$Does.your.employer.offer.resources.to.learn.more.about.mental.health.concerns.and.options.for.seeking.help. ~., scales="free\_y") + ggtitle("Perceived Difficulty of Asking for Leave by Mental\nHealth Resource Provision") + xlab("Perceived Difficulty Level") + ylab("Count") + theme(axis.title=element\_text(size=12))  
t7

 NOTE: VE = Very Easy, SE = Somewhat Easy, N = Neutral (or Neither Easy nor Difficult), SH = Somewhat Hard, VH = Very Hard, DK = Don’t Know.

#### Interpretation of Results

Across all categories (e.g., yes there are resources, no they aren’t, don’t know), the most common response is that most recipients think it would be somewhat easy to request leave. Similarly to the results provided above, in companies where there are other resources provided, fewer people say it would be hard to request leave to handle a mental illness than those who say it would be easy to very easy. For those where there are not additional resources offered related to mental health (the majority of respondents), the second most common answer is that it would be somewhat hard to request leave to hanlde a mental illness.

### Answers by If They Ever Saw A Poor Response to Mental Health Issue

ELDbPR <- table(health$pd2, health$Have.you.observed.or.experienced.an.unsupportive.or.badly.handled.response.to.a.mental.health.issue.in.your.current.or.previous.workplace.)  
ELDbPR

Maybe/Not sure N/A No Yes, I experienced  
 Very easy (VE) 34 7 136 17  
 Somewhat Easy (SE) 58 14 126 35  
 Neutral (N) 46 10 70 18  
 Somewhat Hard (SH) 60 10 66 32  
 Very Hard (VH) 38 3 24 20  
 Don't Know (DK) 42 8 69 10  
   
 Yes, I observed  
 Very easy (VE) 26  
 Somewhat Easy (SE) 48  
 Neutral (N) 34  
 Somewhat Hard (SH) 31  
 Very Hard (VH) 33  
 Don't Know (DK) 21

prop.table(ELDbPR, 2)

Maybe/Not sure N/A No  
 Very easy (VE) 0.12230216 0.13461538 0.27698574  
 Somewhat Easy (SE) 0.20863309 0.26923077 0.25661914  
 Neutral (N) 0.16546763 0.19230769 0.14256619  
 Somewhat Hard (SH) 0.21582734 0.19230769 0.13441955  
 Very Hard (VH) 0.13669065 0.05769231 0.04887984  
 Don't Know (DK) 0.15107914 0.15384615 0.14052953  
   
 Yes, I experienced Yes, I observed  
 Very easy (VE) 0.12878788 0.13471503  
 Somewhat Easy (SE) 0.26515152 0.24870466  
 Neutral (N) 0.13636364 0.17616580  
 Somewhat Hard (SH) 0.24242424 0.16062176  
 Very Hard (VH) 0.15151515 0.17098446  
 Don't Know (DK) 0.07575758 0.10880829

# Revise Labels

#Modify Labels for unsupportive response to make them shorter  
health$epr <- factor(health$Have.you.observed.or.experienced.an.unsupportive.or.badly.handled.response.to.a.mental.health.issue.in.your.current.or.previous.workplace.,  
levels = c("Maybe/Not sure", "N/A", "No","Yes, I experienced", "Yes, I observed"),  
labels = c("Don't Know", "N/A", "No", "Yes, experienced", "Yes, observed"))  
table(health$pd2, health$epr)

Don't Know N/A No Yes, experienced Yes, observed  
 Very easy (VE) 34 7 136 17 26  
 Somewhat Easy (SE) 58 14 126 35 48  
 Neutral (N) 46 10 70 18 34  
 Somewhat Hard (SH) 60 10 66 32 31  
 Very Hard (VH) 38 3 24 20 33  
 Don't Know (DK) 42 8 69 10 21

#### Chi-Squared

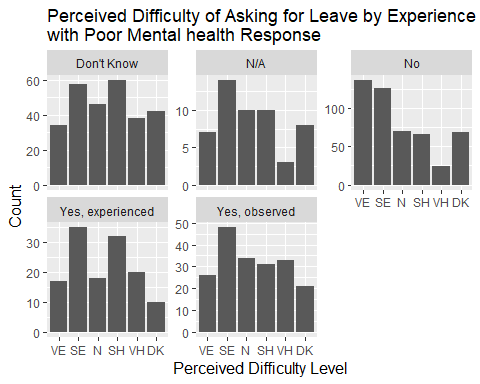
chisq.test(ELDbPR)

Pearson's Chi-squared test  
  
data: ELDbPR  
X-squared = 82.616, df = 20, p-value = 1.406e-09

Because the p value is less than .01 we can reject the null and say that we think that there are differences between categories.

#### Create Visualization

t8 <- ggplot(health, aes(x = health$pd3))+geom\_bar()+facet\_wrap(health$epr ~., scales="free\_y") + ggtitle("Perceived Difficulty of Asking for Leave by Experience\nwith Poor Mental health Response") + xlab("Perceived Difficulty Level") + ylab("Count") + theme(axis.title=element\_text(size=12))  
t8

 NOTE: E = Very Easy, S = Somewhat Easy, N = Neutral (or Neither Easy nor Difficult), SH = Somewhat Hard, H = Very Hard, D = Don’t Know.

#### Interpretation of Results

Most people surveyed indicate that they have never had an experience with an unsupportive or badly handled response to mental health in the workplace. Among those who selected this response, the vast majority indicate they think that it would be very or somewhat easy to request leave to handle a mental illness. For those who have observed a badly handled response, the most common response is that they would find it somewhat easy to request leave, but the second most common is that it they would find it somewhat difficult. For those who have experienced a poorly handled response, the most common response is that they would find it somewhat easy to request leave, but a significant amount said they would find it hard or very hard to request leave (or provided a neutral response saying it would be neither easy or difficult). Therefore, it is possible that experience with a poorly handled response may have an impact on perceived ease of requesting leave.

## Comfort with talking to Supervisor About Mental Health

This is important because if an employee has a mental health issue, they should ideally be comfortable discussing it with a supervisor in case they need additional assistance. In addition, we may want to consider training supervisors on how to handle instances where employees reach out to them about mental health.

### Answers by Gender

MHbG <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s.., health$gender)  
MHbG

Female Male Other  
 Maybe 87 286 9  
 No 84 245 7  
 Yes 100 316 12

prop.table(MHbG, 2)

Female Male Other  
 Maybe 0.3210332 0.3376623 0.3214286  
 No 0.3099631 0.2892562 0.2500000  
 Yes 0.3690037 0.3730815 0.4285714

#### Chi-Squared

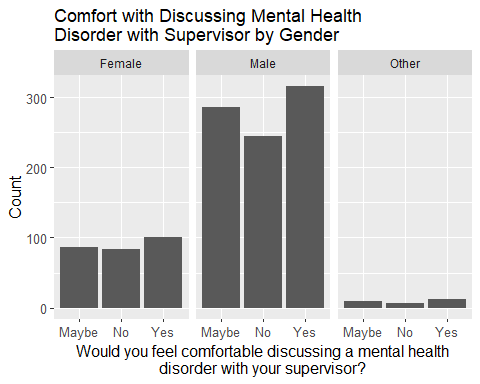
chisq.test(MHbG)

Pearson's Chi-squared test  
  
data: MHbG  
X-squared = 0.90786, df = 4, p-value = 0.9234

The p value is very large. Therefore, we cannot reject the null hypothesis. The null hypothesis is that these two variables are indepdendent (no relationship between them)

#### Create Visualization

t9 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s..))+geom\_bar()+facet\_grid(~ health$gender) + ggtitle("Comfort with Discussing Mental Health\nDisorder with Supervisor by Gender") + xlab("Would you feel comfortable discussing a mental health\ndisorder with your supervisor?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t9



#### Interpretation of Results

Most respondents across genders indicate that they would be comfortable discussing a mental health disorder with their supervisors. With male employees, a lower proportion indicate that they would not be comfortable discussing a mental health disorder with their supervisors. Meanwhile, with female and employees of other genders, the results are slightly more evenly split among the three options. However, the difference is not statistically significant.

### Answers by Age Group

MHbA <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s.., health$AgeGroup)  
MHbA

17-37 37-57 57-77  
 Maybe 265 114 3  
 No 236 96 4  
 Yes 306 117 5

prop.table(MHbA, 2)

17-37 37-57 57-77  
 Maybe 0.3283767 0.3486239 0.2500000  
 No 0.2924411 0.2935780 0.3333333  
 Yes 0.3791822 0.3577982 0.4166667

#### Chi-Squared

chisq.test(MHbA)

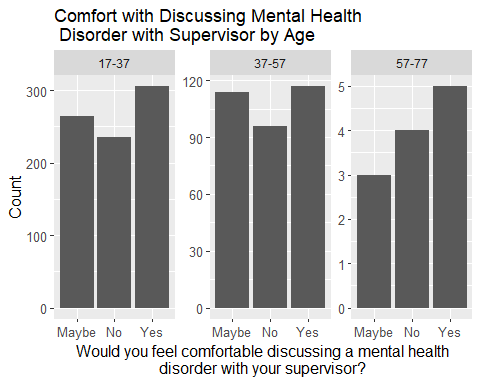
Warning in chisq.test(MHbA): Chi-squared approximation may be incorrect

Pearson's Chi-squared test  
  
data: MHbA  
X-squared = 0.95201, df = 4, p-value = 0.917

We get a warning that the approximation may be incorrect, likely because some of the values from the table are small. However, it potentially indicates that that the differences between age groups may not be statistically significant (or we do not have enough information).

#### Create Visualization

t10 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s..))+geom\_bar()+facet\_wrap(health$AgeGroup ~., scales="free\_y") + ggtitle("Comfort with Discussing Mental Health\n Disorder with Supervisor by Age") + xlab("Would you feel comfortable discussing a mental health\ndisorder with your supervisor?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t10



#### Interpretation of Results

The majority of respondents in the 17-37 and 37-57 age groups responded that they would feel comfortable discussing a mental health disorder with their supervisor. The next most common response is that they might, and no received the least amount of responses from the 17-37 and 37-57 age groups. However, in both cases over a quarter of participants in both age groups answered that they would not feel comfortable. The 57-77 range has a low number of responses, although in that case “No” is the second most common reponse. Given the chi-squared test results and the visualization, it does not appear that age has much of an impact on comfort level in discussing a mental health disorder with a supervisor. However, the fact that there are fewer responses from the 37-57 and especially the 57-77 age groups may impact our results.

### Answers by Company Size

MHbCS <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s.., health$CompanySize)  
MHbCS

1-5 6-25 26-100 100-500 500-1000 More than 1000  
 Maybe 14 69 105 87 32 75  
 No 21 61 74 66 17 97  
 Yes 25 80 113 95 31 84

prop.table(MHbCS, 2)

1-5 6-25 26-100 100-500 500-1000 More than 1000  
 Maybe 0.2333333 0.3285714 0.3595890 0.3508065 0.4000000 0.2929688  
 No 0.3500000 0.2904762 0.2534247 0.2661290 0.2125000 0.3789062  
 Yes 0.4166667 0.3809524 0.3869863 0.3830645 0.3875000 0.3281250

#### Chi-Squared

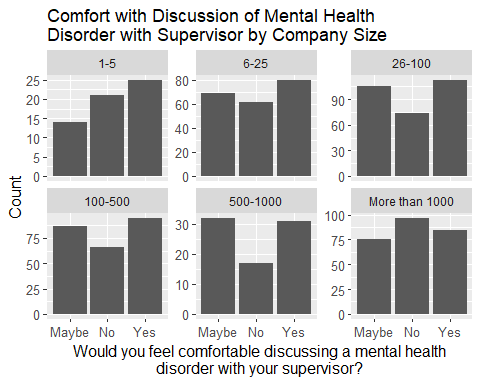
chisq.test(MHbCS)

Pearson's Chi-squared test  
  
data: MHbCS  
X-squared = 18.002, df = 10, p-value = 0.05493

Because the p value is not less than .01 we cannot reject the null. The null hypothesis is that there is no relationship between these two variables.

#### Create Visualization

t11 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s..))+geom\_bar()+facet\_wrap(health$CompanySize ~., scales="free\_y") + ggtitle("Comfort with Discussion of Mental Health\nDisorder with Supervisor by Company Size") + xlab("Would you feel comfortable discussing a mental health\ndisorder with your supervisor?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t11



#### Interpretation of Results

The pattern that we see in the data visualization is interesting. For companies in the 26-100, 100-500, and 500-1000 groups (and to a lesser extent in the 6-25 group), fewer respondents indicate they would not feel comfortable discussing a mental health disorder with their supervisor than selecting the “Maybe” or “Yes” options. However, in the “More than 1000” group, most respondents indicated they would not feel comfortable (although a significant amount answered “Maybe” or “Yes”). It is possible that for larger companies people may feel less comfortable discussing these issues with supervisors because it the organization is more bureaucratic or there are employee/supervisor relationships that are less personal. However, it is also possible that these results indicate that the results are depedendent on the company as well as other factors. Because the null is larger than .01, we fail to reject the null that there is a significant relationship between the two variables.

### Answers by If They Ever Saw A Poor Response to Mental Health Issue

MHbPR <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s.., health$Have.you.observed.or.experienced.an.unsupportive.or.badly.handled.response.to.a.mental.health.issue.in.your.current.or.previous.workplace.)  
MHbPR

Maybe/Not sure N/A No Yes, I experienced Yes, I observed  
 Maybe 97 12 161 48 64  
 No 101 24 109 35 67  
 Yes 80 16 221 49 62

prop.table(MHbPR, 2)

Maybe/Not sure N/A No Yes, I experienced  
 Maybe 0.3489209 0.2307692 0.3279022 0.3636364  
 No 0.3633094 0.4615385 0.2219959 0.2651515  
 Yes 0.2877698 0.3076923 0.4501018 0.3712121  
   
 Yes, I observed  
 Maybe 0.3316062  
 No 0.3471503  
 Yes 0.3212435

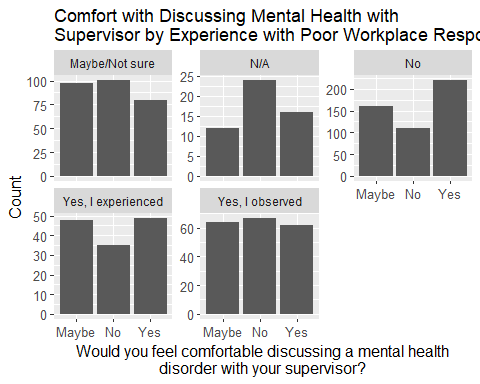
#### Chi-Squared

chisq.test(MHbPR)

Pearson's Chi-squared test  
  
data: MHbPR  
X-squared = 37.901, df = 8, p-value = 7.852e-06

#### Create Visualization

t17 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s..))+geom\_bar()+facet\_wrap(health$Have.you.observed.or.experienced.an.unsupportive.or.badly.handled.response.to.a.mental.health.issue.in.your.current.or.previous.workplace. ~., scales="free\_y") + ggtitle("Comfort with Discussing Mental Health with\nSupervisor by Experience with Poor Workplace Response") + xlab("Would you feel comfortable discussing a mental health\ndisorder with your supervisor?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t17



#### Interpret Results

When asked about experience with a poorly handled response to mental health, respondents who were not sure if they expereinced it were evenly split between answering that they maybe comfortable discussing a mental health disorder with supervisors or not comfortable (there was also a large group that said they would be comfortable). For those who answered that they did not have experience with a poor response, they also answered yes more often than maybe and no when also asked about their comfort with discussing a mental health disorder with a supervisor. Interestingly, those who stated “Yes, I observed” a poor response to mental health in the workplace were evenly split selected maybe the most often when also asked if they would feel comfortable discussing a mental health disorder with coworkers (and “yes” was the second most common response). However, those who stated that they experienced a poorly handled response were evenly split between the yes, no, and maybe answers when asked about comfort level discussing mental health with a supervisor.

### Answers by If They Think Discussing a Mental Health Disorder with Employer Would Have Negative Consequences

MHbVN <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s.., health$Do.you.think.that.discussing.a.mental.health.disorder.with.your.employer.would.have.negative.consequences.)  
MHbVN

Maybe No Yes  
 Maybe 218 115 49  
 No 164 24 148  
 Yes 105 299 24

#### Chi-Squared

chisq.test(MHbVN)

Pearson's Chi-squared test  
  
data: MHbVN  
X-squared = 414.95, df = 4, p-value < 2.2e-16

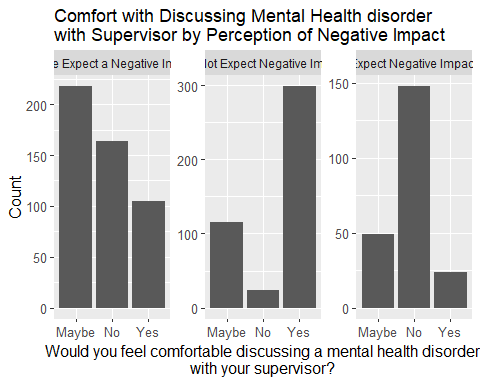
Because the p value is less than .01 we can reject the null and say that we think that there is a statistically significant relationship between these two categories.

#### Create Visualization

#Modify Labels for perception of mental health disclosure impact  
health$negimp <- factor(health$Do.you.think.that.discussing.a.mental.health.disorder.with.your.employer.would.have.negative.consequences.,  
levels = c("Maybe", "No", "Yes"),  
labels = c("Maybe Expect a Negative Impact", "Do Not Expect Negative Impact", "Expect Negative Impact"))  
#Check Results  
table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s.., health$negimp)

Maybe Expect a Negative Impact Do Not Expect Negative Impact  
 Maybe 218 115  
 No 164 24  
 Yes 105 299  
   
 Expect Negative Impact  
 Maybe 49  
 No 148  
 Yes 24

t12 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.direct.supervisor.s..))+geom\_bar()+facet\_wrap(health$negimp ~., scales="free\_y") + ggtitle("Comfort with Discussing Mental Health disorder\nwith Supervisor by Perception of Negative Impact") + xlab("Would you feel comfortable discussing a mental health disorder\nwith your supervisor?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t12



#### Interpret Results

The results here are fairly clear. Most people who responded that they feel that discussing a mental health disorder would have a negative impact also are not comfortable discussing a mental health issue with a supervisor. The reverse is also true, most people who responded that they do not feel that discussing a mental health disorder would have a negative impact also indicate that they would be comfortable discussing a mental health issue with their supervisor. For those who responded that there maybe a negative impact, most also answered that they may be comfortable discussing a mental health disorder with their supervisor (with “no” being the second most common response for that group). Therefore, any outreach that we do or program that we plan needs to keep this in mind.

## Comfort with taking to Coworkers About Mental Health

This is important because if an employee has a mental health issue, they should ideally be comfortable discussing it with a coworkers in case they need additional assistance. In addition, we may want to consider training employees on how to handle instances where coworkers reach out to them about mental health.

### Answers by Gender

MHCbG <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers., health$gender)  
MHCbG

Female Male Other  
 Maybe 107 360 12  
 No 99 287 6  
 Yes 65 200 10

prop.table(MHCbG, 2)

Female Male Other  
 Maybe 0.3948339 0.4250295 0.4285714  
 No 0.3653137 0.3388430 0.2142857  
 Yes 0.2398524 0.2361275 0.3571429

#### Chi-Squared

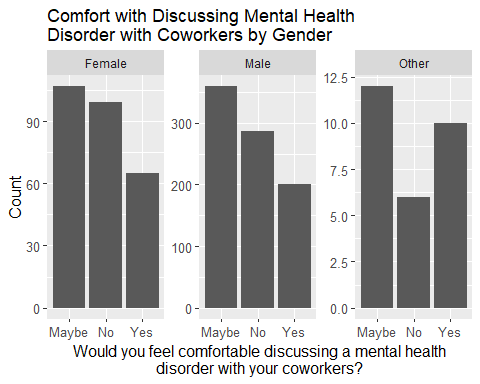
chisq.test(MHCbG)

Pearson's Chi-squared test  
  
data: MHCbG  
X-squared = 3.9002, df = 4, p-value = 0.4197

The chi-squared is very large. Therefore, we cannot reject the null hypothesis that these categories are independent.

#### Create Visualization

t13 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers.))+geom\_bar()+facet\_wrap(health$gender ~., scales="free\_y") + ggtitle("Comfort with Discussing Mental Health\nDisorder with Coworkers by Gender") + xlab("Would you feel comfortable discussing a mental health\ndisorder with your coworkers?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t13



#### Interpret Results

The chi-squared test appears to indicate that there is not a significant relationship between gender and comfort with discussing mental health with coworkers. However, these results may be impacted by the fact that there are fewer respondents who are female or another gender than men. However, a higher proportion of female respondents indicate that they do not feel comfortable discussing a mental health disorder with coworkers. However, across all genders the most common response is “maybe.” This indicates that a lot may depend on their relationships with individual coworkers (e.g., they may feel comfortable with discussing with a specific coworker). Interestingly, as stated above, most respondents across genders indicated that they would be comfortable discussing a mental health disorder with their supervisors.

### Answers by Age Group

MHCbA <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers., health$AgeGroup)  
MHCbA

17-37 37-57 57-77  
 Maybe 338 139 2  
 No 272 116 4  
 Yes 197 72 6

#Proportions (as column percentages)  
prop.table(MHCbA, 2)

17-37 37-57 57-77  
 Maybe 0.4188352 0.4250765 0.1666667  
 No 0.3370508 0.3547401 0.3333333  
 Yes 0.2441140 0.2201835 0.5000000

#### Chi-Squared

chisq.test(MHCbA)

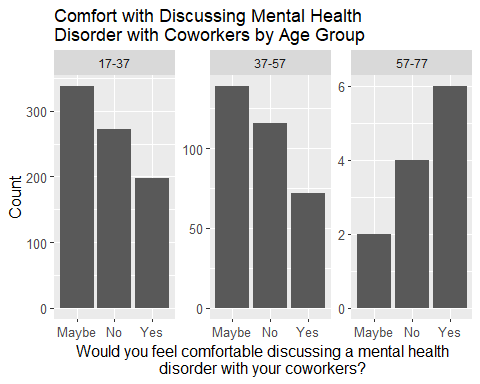
Warning in chisq.test(MHCbA): Chi-squared approximation may be incorrect

Pearson's Chi-squared test  
  
data: MHCbA  
X-squared = 6.0422, df = 4, p-value = 0.196

We get a warning that the approximation may be incorrect, likely because some of the values from the table are small. However, it potentially indicates that that the differences between age groups is not statistically significant.

#### Create Visualization

t14 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers.))+geom\_bar()+facet\_wrap(health$AgeGroup ~., scales="free\_y") + ggtitle("Comfort with Discussing Mental Health\nDisorder with Coworkers by Age Group") + xlab("Would you feel comfortable discussing a mental health\ndisorder with your coworkers?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t14



#### Interpret Results

Members of the 17-37 and 37-57 age group have a similar pattern. The majority of respondents indicated that they might feel comfortable discussing a mental health disorder with coworkers, with the second most common response being that they would not. When I looked at the comfort level of discussing a mental health disorder with supervisors, the results were different. Interestingly, as stated above, the majority of respondents in the 17-37 and 37-57 age groups responded that they would feel comfortable discussing a mental health disorder with their supervisor. Therefore, there may be a higher comfort level discussing mental health disorders with supervisors than other coworkers (or respondents’ comfort level may vary by the coworker).

### Answers by Company Size

MHCbCS <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers., health$CompanySize)  
MHCbCS

1-5 6-25 26-100 100-500 500-1000 More than 1000  
 Maybe 20 83 140 110 32 94  
 No 19 75 84 79 21 114  
 Yes 21 52 68 59 27 48

#Proportions (as column percentages)  
prop.table(MHCbCS, 2)

1-5 6-25 26-100 100-500 500-1000 More than 1000  
 Maybe 0.3333333 0.3952381 0.4794521 0.4435484 0.4000000 0.3671875  
 No 0.3166667 0.3571429 0.2876712 0.3185484 0.2625000 0.4453125  
 Yes 0.3500000 0.2476190 0.2328767 0.2379032 0.3375000 0.1875000

#### Chi-Squared

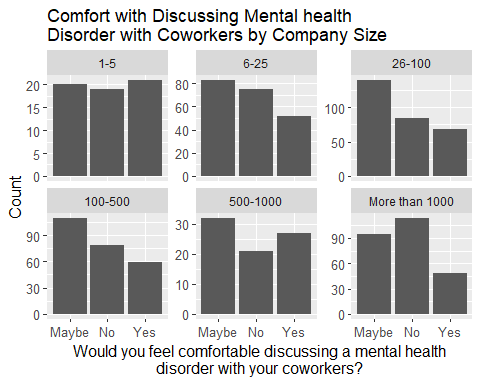
#Chi-Squared (the p value is small)  
chisq.test(MHCbCS)

Pearson's Chi-squared test  
  
data: MHCbCS  
X-squared = 27.848, df = 10, p-value = 0.001909

The chi-squared test shows a p value that is less than 0.01. Therefore, we can reject the null hypothesis (that there is no relationship between these variables). This indicates that there may be a relationship between company size and comfort level with talking to coworkers about a mental health issue.

#### Create Visualization

t15 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers.))+geom\_bar()+facet\_wrap(health$CompanySize ~., scales="free\_y") + ggtitle("Comfort with Discussing Mental health\nDisorder with Coworkers by Company Size") + xlab("Would you feel comfortable discussing a mental health\ndisorder with your coworkers?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t15



#### Interpret Results

Respondents from companies with 1-5 employees are roughly evenly split among the three responses (yes, no, maybe) when asked if they would feel comfortable discussing a mental health disorder with coworkers. Results are similar with companies for 6-25 employees, but a lower proportion of employees in that category indicate they would be comfortable. For the 26-100 and 100-500 employee companies, most indicate that they may be comfortable with “no” being the second most common response. Respondents from companies with 500-1000 employees also select maybe the most often, but “yes” is the second most common response. Finally, respondents from companies with more than 1000 employees select the “no” answer the most often, with maybe being the second most common. The results indicate that there may be a relationship between company size and willingness to talk to coworkers about mental illness. However, the variation in responses also indicates that the reasons for the comfort, discomfort, or uncertainty may vary by company.

### Answers by If They Ever Saw A Poor Response to Mental Health Issue

MHCbPR <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers., health$Have.you.observed.or.experienced.an.unsupportive.or.badly.handled.response.to.a.mental.health.issue.in.your.current.or.previous.workplace.)  
MHCbPR

Maybe/Not sure N/A No Yes, I experienced Yes, I observed  
 Maybe 132 19 195 58 75  
 No 105 22 150 39 76  
 Yes 41 11 146 35 42

#### Chi-Squared

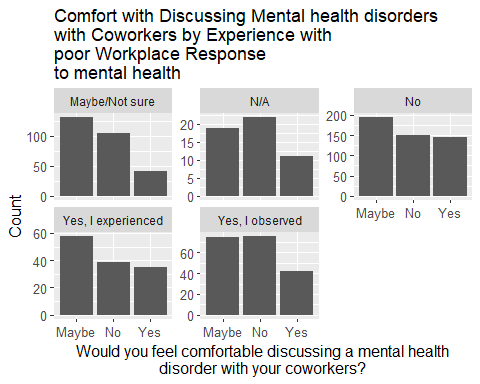
chisq.test(MHCbPR)

Pearson's Chi-squared test  
  
data: MHCbPR  
X-squared = 27.416, df = 8, p-value = 0.0005991

The chi-squared test shows a p value that is less than 0.01. Therefore, we can reject the null hypothesis (that there is no relationship between these variables). This indicates that there may be a relationship between experience with a poor response to a mental health issue with talking to coworkers about a mental health issue.

#### Create Visualization

t16 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers.))+geom\_bar()+facet\_wrap(health$Have.you.observed.or.experienced.an.unsupportive.or.badly.handled.response.to.a.mental.health.issue.in.your.current.or.previous.workplace. ~., scales="free\_y") + ggtitle("Comfort with Discussing Mental health disorders\nwith Coworkers by Experience with\npoor Workplace Response\nto mental health") + xlab("Would you feel comfortable discussing a mental health\ndisorder with your coworkers?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t16



#### Interpret Results

When asked about experience with a poorly handled response to mental health, respondents across all categories except for “N/A” (which had relatively few respondents), also answered that they maybe comfortable discussing a mental health disorder with coworkers. Interestingly, those who stated “Yes, I observed” a poor response to mental health in the workplace were evenly split between the maybe and no answers when also asked if they would feel comfortable discussing a mental health disorder with coworkers. However, those who stated that they experienced a poorly handled response selected “Maybe” most often when also asked if they would feel comfortable discussing a mental health disorder with coworkers. The remainder were about evenly split between the “no” and “yes” answers. This may indicate that observing a poorly handled response to a mental health issue in the workplace may have a greater negative effect on willingness to discuss with coworkers than those experiencing it.

### Answers by If They Think Coworkers Would View Them Negatively If They Knew You Suffered From a Mental Health Issue

MHCbCN <- table(health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers., health$Do.you.think.that.team.members.co.workers.would.view.you.more.negatively.if.they.knew.you.suffered.from.a.mental.health.issue.)  
MHCbCN

Maybe No, I don't think they would No, they do not  
 Maybe 234 137 7  
 No 127 51 1  
 Yes 100 115 36  
   
 Yes, I think they would Yes, they do  
 Maybe 87 14  
 No 201 12  
 Yes 20 4

#### Chi-Squared

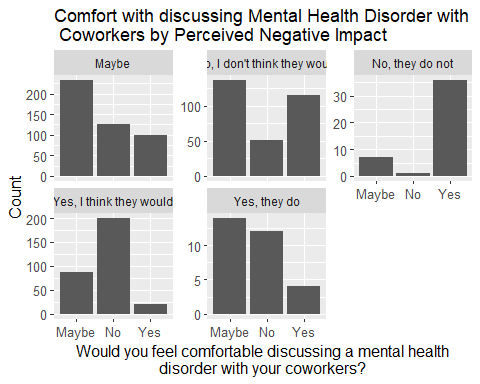
chisq.test(MHCbCN)

Pearson's Chi-squared test  
  
data: MHCbCN  
X-squared = 291.08, df = 8, p-value < 2.2e-16

The chi-squared is less than .01. Therefore, we can reject the null and say that there is a relationship between these two variables.

#### Creating the Visualization

t17 <- ggplot(health, aes(x = health$Would.you.feel.comfortable.discussing.a.mental.health.disorder.with.your.coworkers.))+geom\_bar()+facet\_wrap(health$Do.you.think.that.team.members.co.workers.would.view.you.more.negatively.if.they.knew.you.suffered.from.a.mental.health.issue. ~., scales="free\_y") + ggtitle("Comfort with discussing Mental Health Disorder with\n Coworkers by Perceived Negative Impact") + xlab("Would you feel comfortable discussing a mental health\ndisorder with your coworkers?") + ylab("Count") + theme(axis.title=element\_text(size=12), axis.text=element\_text(size=10))  
t17



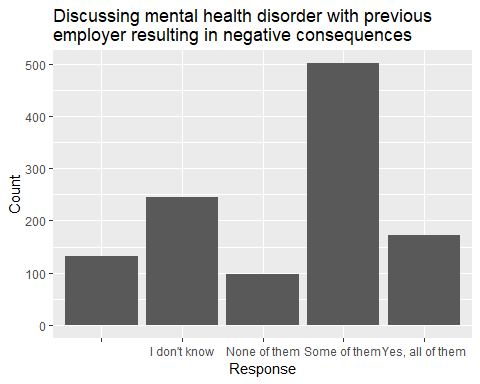
#### Interpretation of Results

In this visualization we see that perceived impact of discussing a mental health disorder with coworkers has a potentially significant impact on comfort level with discussing a mental health disorder with coworkers. For those who feel that they would be viewed negatively by coworkers if they knew the respondent had a mental health issue, they frequently answer that they would not feel comfortable discussing a mental health disorder with coworkers. For those who feel like they are negatively viewed by coworkers because of a mental health issue, appear to be evenly split between maybe being comfortable and not being comfortable. Meanwhile, for those who feel that they are not viewed more negatively because of a mental health disorder, they most frequently say they would be comfortable discussing mental health with coworkers. For those who did not think they would be viewed negatively, they most frequently select that they may be comfortable discussing a mental health disorder with other coworkers, with yes being the second most common response from that group. This indicates that we may want to consider providing training to employees about mental illness in the workplace.

### Analzing the relationship between perceiving the negative impact of mental illness at previous workplace and the impact at current workplace

#### Plotting the ggplot for the independent variable “Do you think that discussing a mental health disorder with previous employers would have negative consequences?”

ggplot(health, aes(x = health$Do.you.think.that.discussing.a.mental.health.disorder.with.previous.employers.would.have.negative.consequences.)) + geom\_bar() + ggtitle("Discussing mental health disorder with previous\nemployer resulting in negative consequences") + xlab("Response") + ylab("Count")



summary(health$Do.you.think.that.discussing.a.mental.health.disorder.with.previous.employers.would.have.negative.consequences.)

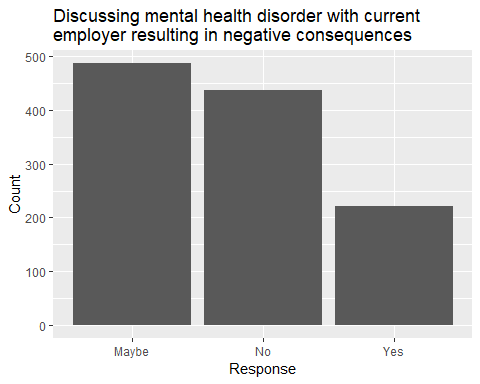
I don't know None of them Some of them   
 131 245 97 501   
Yes, all of them   
 172

#### Interpreting the results

From the above ggplot, we can observe that a majority of respondents(615) feel that discussing mental health disorders with some of their previous employers resulted in negative consequences for them. While a significant number of respondents responded that they do not know, we can conclude from the graph that the majority did experience negative consequences after discussing mental health disorders with their previous employers

### Plotting the ggplot for the dependent variable “Do you think that discussing a mental health disorder with your employer would have negative consequences?”

ggplot(health, aes(x = health$Do.you.think.that.discussing.a.mental.health.disorder.with.your.employer.would.have.negative.consequences.)) + geom\_bar() + ggtitle("Discussing mental health disorder with current\nemployer resulting in negative consequences") + xlab("Response") + ylab("Count")



summary(health$Do.you.think.that.discussing.a.mental.health.disorder.with.your.employer.would.have.negative.consequences.)

Maybe No Yes   
 487 438 221

#### Interpretation of results

From the above ggplot, we can see that a vast majority of respondents(487) are not sure if discussing mental health disorders with their current employers would have negative consequences; followed by about 438 respondents who do not think that it could lead to negative consequences

#### Chi-squared test

t1<-table(health$Do.you.think.that.discussing.a.mental.health.disorder.with.previous.employers.would.have.negative.consequences., health$Do.you.think.that.discussing.a.mental.health.disorder.with.your.employer.would.have.negative.consequences.)  
chisq.test(t1)

Pearson's Chi-squared test  
  
data: t1  
X-squared = 158.59, df = 8, p-value < 2.2e-16

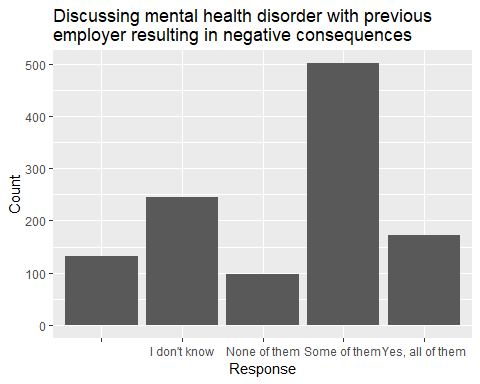
#### Interpret results

We can see that the p-value of the Chi-squared test is very small, meaning we can reject the hypothesis of independence. Thus there is a relationship between perceiving the negative impact of mental illness at previous workplace and the impact at current workplace

### Analzing the relationship between perceiving the negative impact of mental illness at previous workplace and the impact at current workplace

#### Plotting the ggplot for the independent variable “Do you think that discussing a mental health disorder with previous employers would have negative consequences?”

ggplot(health, aes(x = health$Do.you.think.that.discussing.a.mental.health.disorder.with.previous.employers.would.have.negative.consequences.)) + geom\_bar() + ggtitle("Discussing mental health disorder with previous\nemployer resulting in negative consequences") + xlab("Response") + ylab("Count")



summary(health$Do.you.think.that.discussing.a.mental.health.disorder.with.previous.employers.would.have.negative.consequences.)

I don't know None of them Some of them   
 131 245 97 501   
Yes, all of them   
 172

# Summary

This is a placeholder for the summary.

# Who did what in the project

Laura - For the work plan, I worked with the group to draft the sections on dataset used, audience, and research questions. I also reviewed a list of studies on mental health in the workplace and drafted summaries that can be used as a reference by the group. For the data cleaning documentation, I examined nine variables and noted if there were any issues. I worked with other group members to determine how to handle the variables. For the R Script, I did frequency tables related to my analyses and the variables I examined for data cleaning. For the R Plot Draft, I created visualizations and interpretations for my analyses (in section 3, the portions on anonymity policies, ease of taking leave, comfort discussing with coworkers and supervisors). For the presentation, I created a template for the presentation that included what was needed and placeholder slides for the analyses the group did. I also did slides related to my analysis (anonymity policies, ease of taking leave, comfort discussing with coworkers and supervisors) and used the findings to draft some recommendations.For the final package, I did the first draft of the summary.

# Works Cited

Milne, S. H., Blum, T. C., & Roman, P. M. (1994). Factors Influencing Employees??? Propensity to Use an Employee Assistance Program. Personnel Psychology, 47(1), 123???145. Retrieved October 24, 2018 from <http://search.ebscohost.com.proxy-um.researchport.umd.edu/login.aspx?direct=true&db=bth&AN=9411113184&site=ehost-live>