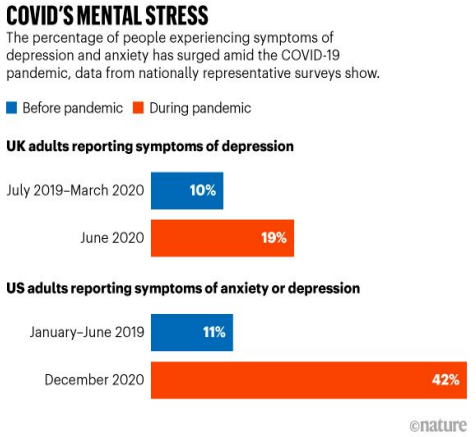
Mental Health Disorder in the Tech Industry

MGT 6203 Team 5

Why did we choose this topic?

- Increase in the number of people with anxiety or depression symptoms
  - Increase from 11% in 2019 to 42% in 2020 (US census Bureau)

- Pandemic causes increase in anxiety and depression
  - True for people in the healthcare industry
  - Systematic research review: ~20% of all healthcare workers experience anxiety and depression (BMJ Open Journal Sept 2021)
  - But for tech industry?



Source: Office for National Statistics (UK data); Centers for Disease Control and Prevention (US data).

Our dataset

- **Mental Health in Tech Survey** collected by Open Source Mental Health (OSMH)
- Years picked: 2019, 2020 and 2021.

Reason for choice

- In consecutive years
- Account for the impact of the pandemic

Research Questions

1. What are the factors that might explain cases of mental health disorders at the workplace?
2. Does the pandemic have an effect on mental health disorders?
3. Does the pandemic have an effect on the attitudes towards mental health disorders?

Approach and Methodology - Data Cleaning

Research Qn 1: Features affecting Mental Health Disorders

- Dependent variable (Outcome): Presence of Mental Health Disorder
- Independent variable (Features):
  - Employment Type→ 1: self-employed, 0: employed
  - Job Scope → 1: Tech, 0: Non-Tech, Missing: Unknown
  - Family History
  - Gender → classify into 4 groups (M, F, Others, Unknown)
  - Country that people live in
  - Country that people work in
    - Picked the top 3 countries
    - Merge the rest of the data into “others”
  - Age
    - Remove unreasonable ages (<18 & >100)
    - Group the rest into 5 different age groups
- Replace missing values with “unknown”

Approach and Methodology - Data Cleaning

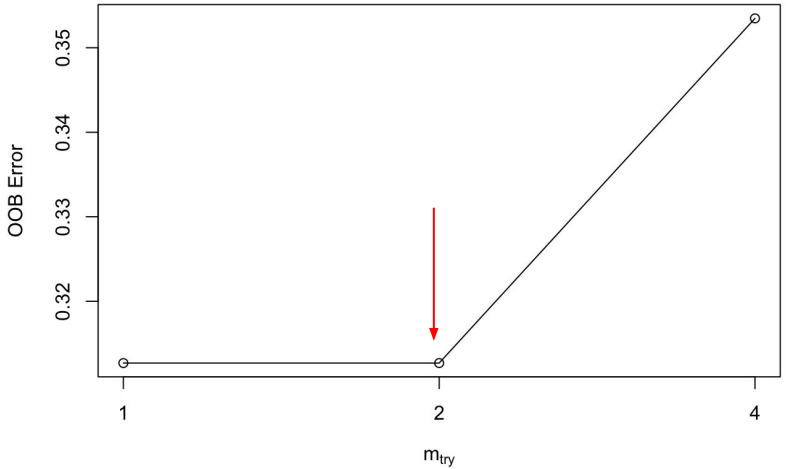
Research Qn 1: Features affecting Mental Health Disorders

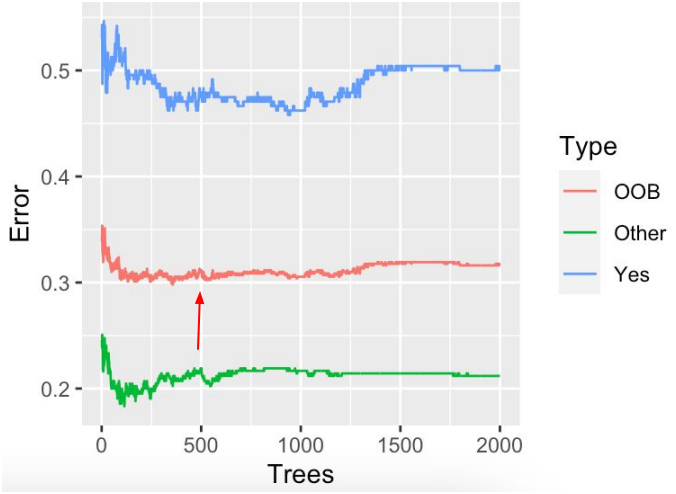
- Random Forest as classifier
  - Dependent variable: Do you currently have a mental health disorder?
  - 4 possible answers - Yes, No, Don’t Know, Possibly
  - High Accuracy Rate
  - Offers feature selection indicator (Gini Importance)

Approach and Methodology - Data Cleaning

Research Qn 1: Features affecting Mental Health Disorders

- Problem Faced: Poor accuracy rate (0.46)
  - Reclassified the answers - Yes and Other (No, Don't Know, Possibly)
  - Use OOB (out-of-bag) error to find the number of features to consider at each split
  - Find the best number of decision trees to use





Results

Research Qn 1: Features affecting Mental Health Disorders

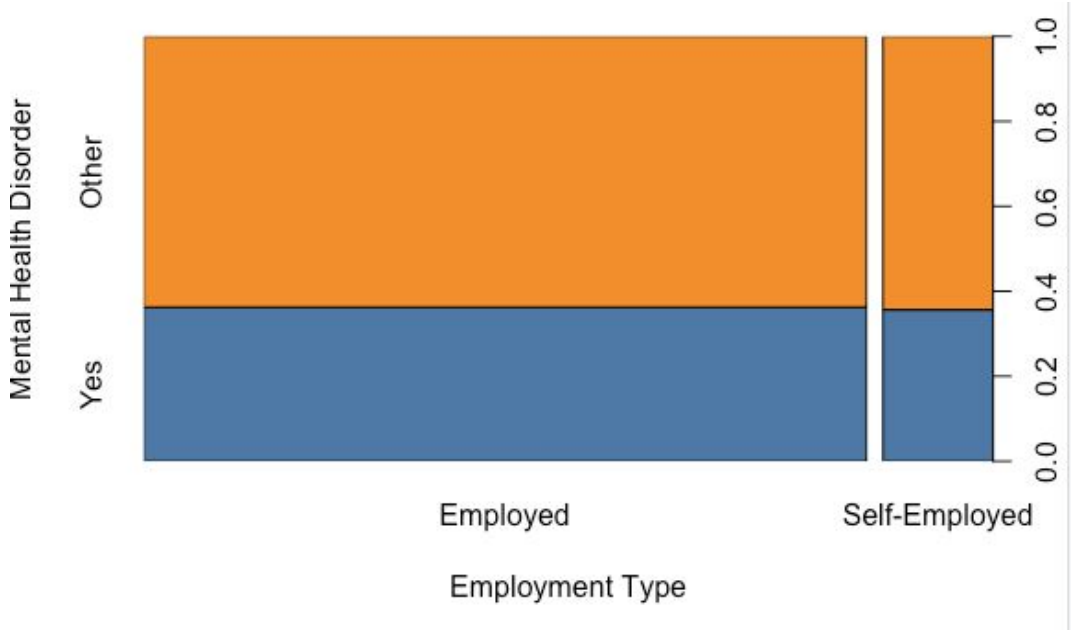
- Feature: Categorical Data , Outcome: Categorical Data
- Mosaic Plot + Chi Square Statistical Test
- Code shown below:

```
Chi_func <- function(v1, v2){
  tb <- table(v1, v2)
  chi <- chisq.test(tb, correct = F)
  chi$p.value
}
```

Results

Research Qn 1: Features affecting Mental Health Disorders

- Employment Type vs. Presence of Mental Health Disorders



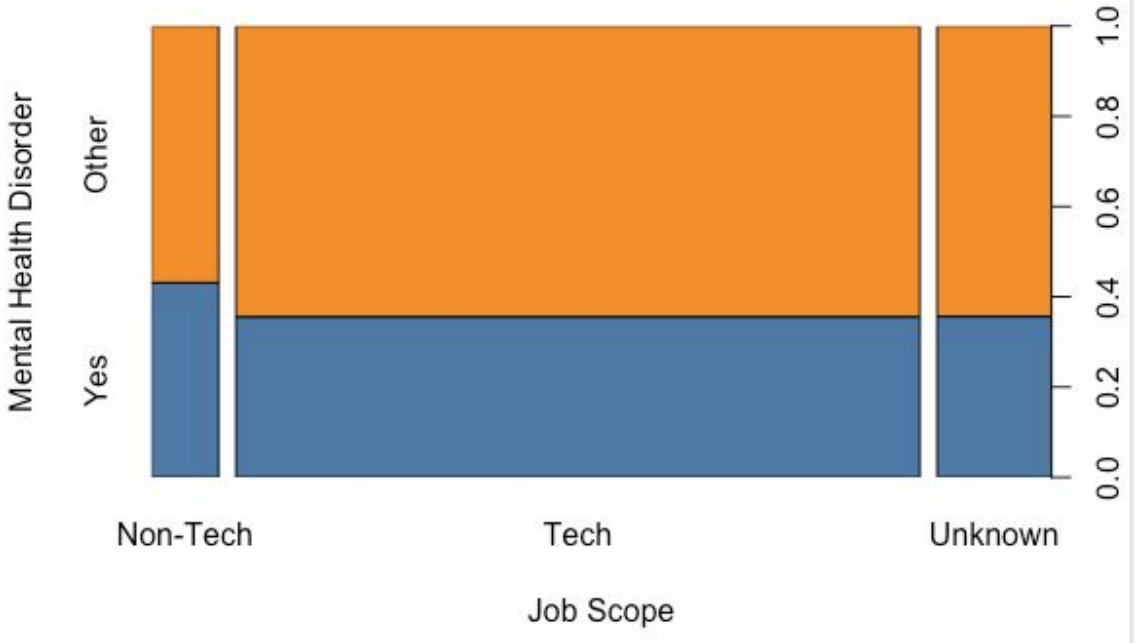

```
Chi_func(dat$EmploymentType,
dat$CurrentlyMentalHealthDisorder)
# [1] 0.9107304
```

- No difference in likelihood of getting mental health disorders
- P > 0.05: Fail to reject null hypothesis

Results

Research Qn 1: Features affecting Mental Health Disorders

- Job Scope vs. Presence of Mental Health Disorders



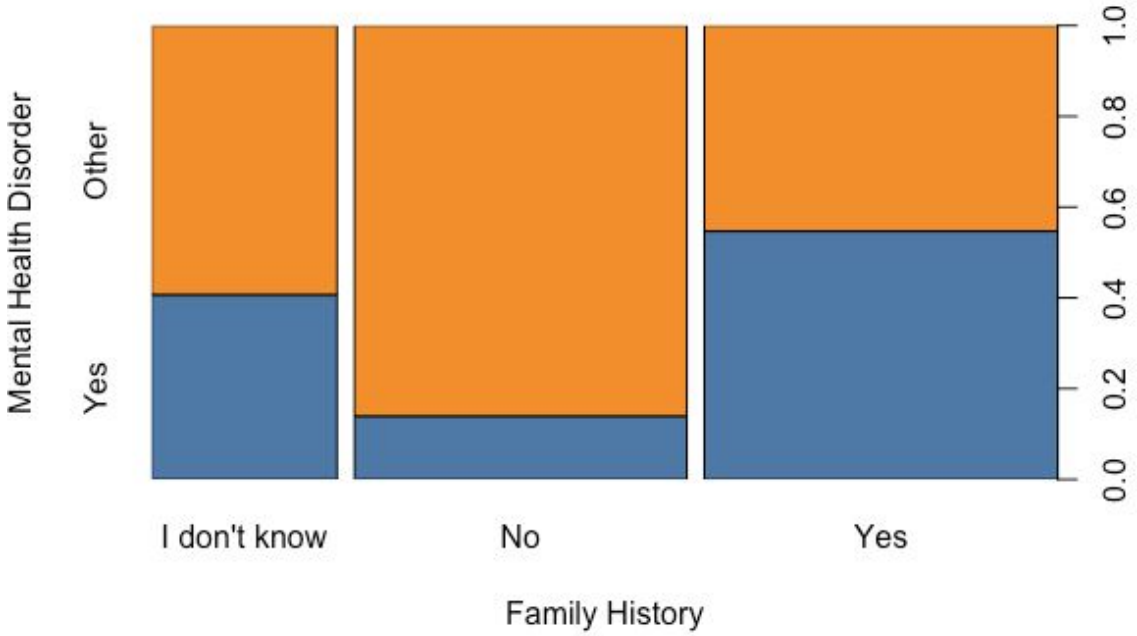

```
Chi_func(dat$JobScope,
dat$CurrentlyMentalHealthDisorder)
# [1] 0.5592186
```

- No difference in likelihood of getting mental health disorders
- P > 0.05: Fail to reject null hypothesis

Results

Research Qn 1: Features affecting Mental Health Disorders

- Family History vs. Presence of Mental Health Disorders




```
# FamilyHistory ***
Chi_func(dat$FamilyHistory,
dat$CurrentlyMentalHealthDisorder)
# [1] 2.888205e-21
```

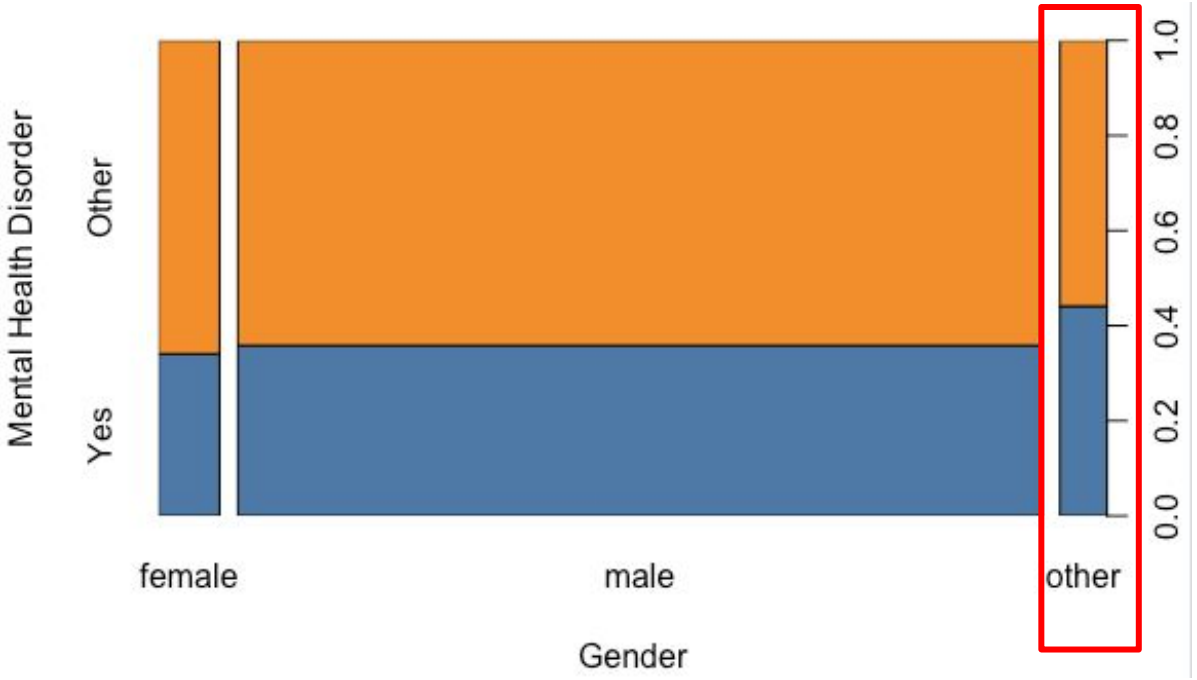
- People who has family history of mental health disorders are more likely to get mental health disorders

- P<0.05: Reject null hypothesis

Results

Research Qn 1: Features affecting Mental Health Disorders

- Gender vs. Presence of Mental Health Disorders



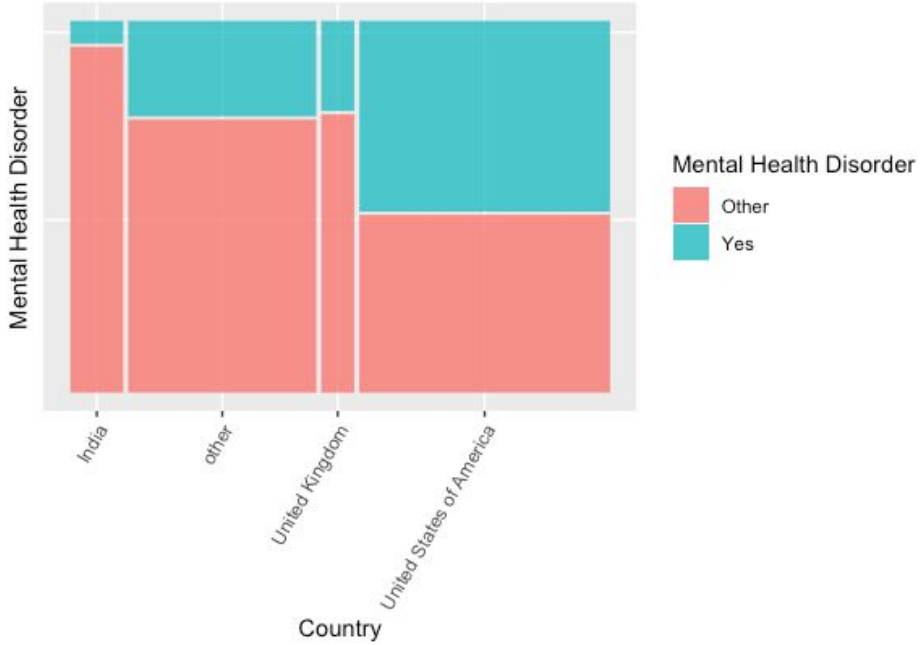

```
Chi_func(dat$Gender,
dat$CurrentlyMentalHealthDisorder)
# [1] 0.5955878
```

- Non-Cisgender are more likely to get mental health disorders
- Limitation: Small data sample for non-cisgnder
- P > 0.05: fail to reject null hypothesis

Results

Research Qn 1: Features affecting Mental Health Disorders

- Country people live in vs. Presence of Mental Health Disorders



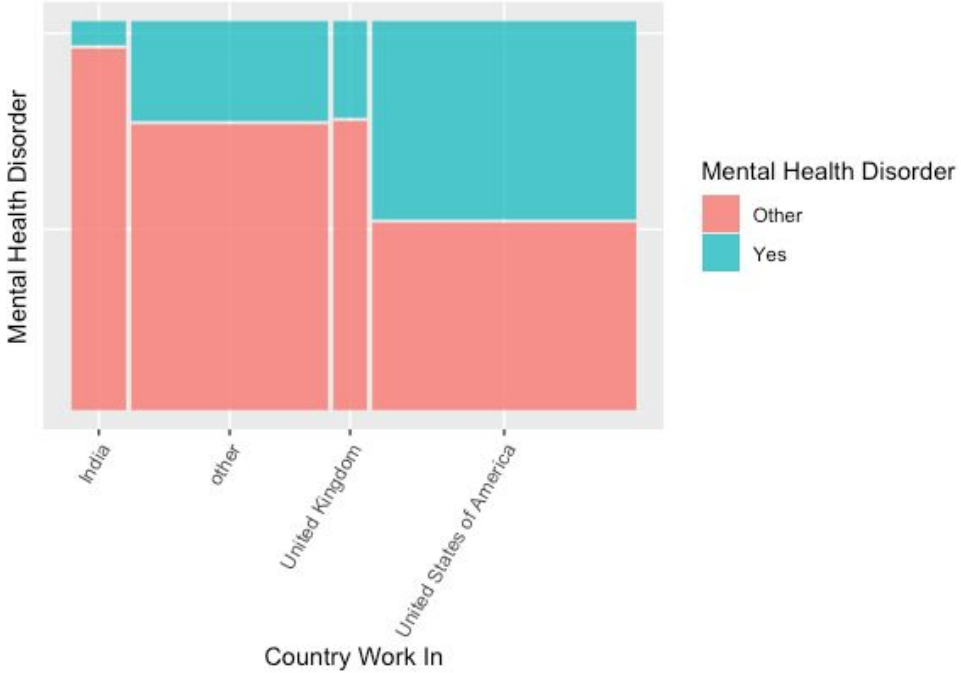

```
Chi_func(dat$CountryLiveIn,
dat$CurrentlyMentalHealthDisorder)
# [1] 1.323527e-15
```

- People living in USA are more likely to get mental health disorders
- Limitation: Lack of sample size from other countries
- $P < 0.05$: Reject null hypothesis

Results

Research Qn 1: Features affecting Mental Health Disorders

- Country people work in vs. Presence of Mental Health Disorders



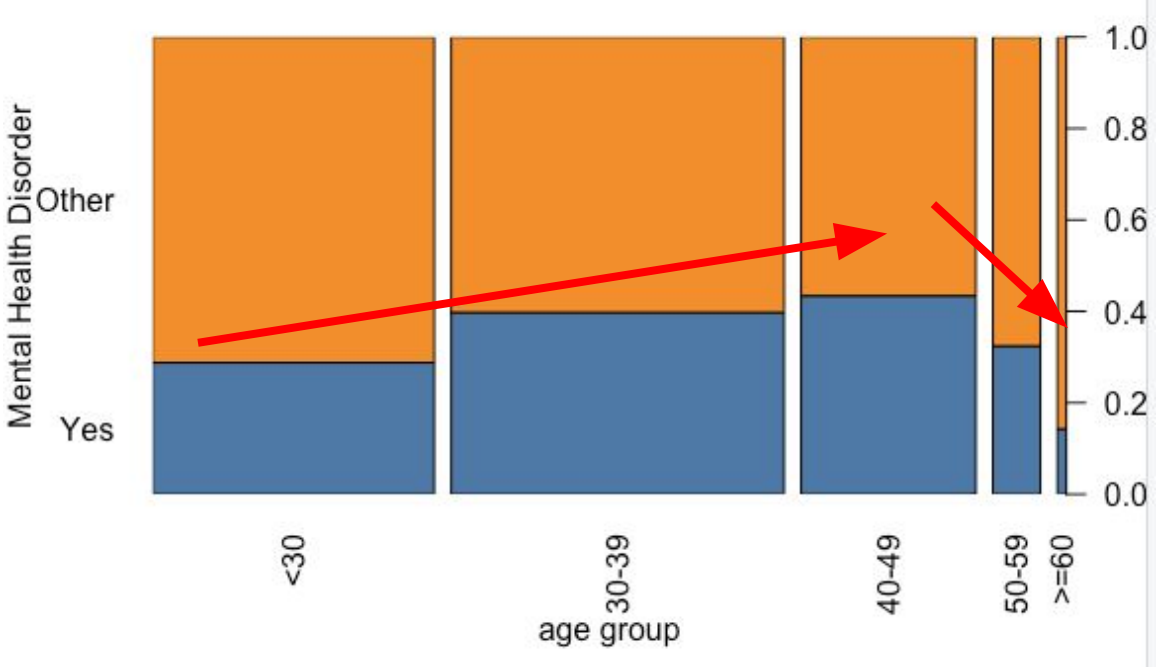

```
# CountryWorkIn ***
Chi_func(dat$CountryWorkIn,
dat$CurrentlyMentalHealthDisorder)
# [1] 3.592461e-15
```

- People working in USA are more likely to get mental health disorders
- Limitation: Lack of sample size from other countries
- P < 0.05: Reject null hypothesis

Results

Research Qn 1: Features affecting Mental Health Disorders

- Age vs Presence of Mental Health Disorders




```
chisq.test(dat$AgeGroup,
dat$CurrentlyMentalHealthDisorder,
simulate.p.value = TRUE)

# data:  dat$AgeGroup and
dat$CurrentlyMentalHealthDisorder
# X-squared = 11.391, df = NA, p-value =
0.02149
```

- The likelihood of one having a mental health disorder increases gradually and eventually peaks at 40-49. Beyond that, the likelihood of developing the mental health disorder decreases.

- As P-value < 0.05, reject the null hypothesis

Results

Research Qn 1: Features affecting Mental Health Disorders

Confusion Matrix and Statistics

Prediction	Reference: Yes	Reference: Other
Yes	144	77
Other	94	343

Accuracy : 0.7401

95% CI : (0.7048, 0.7733)

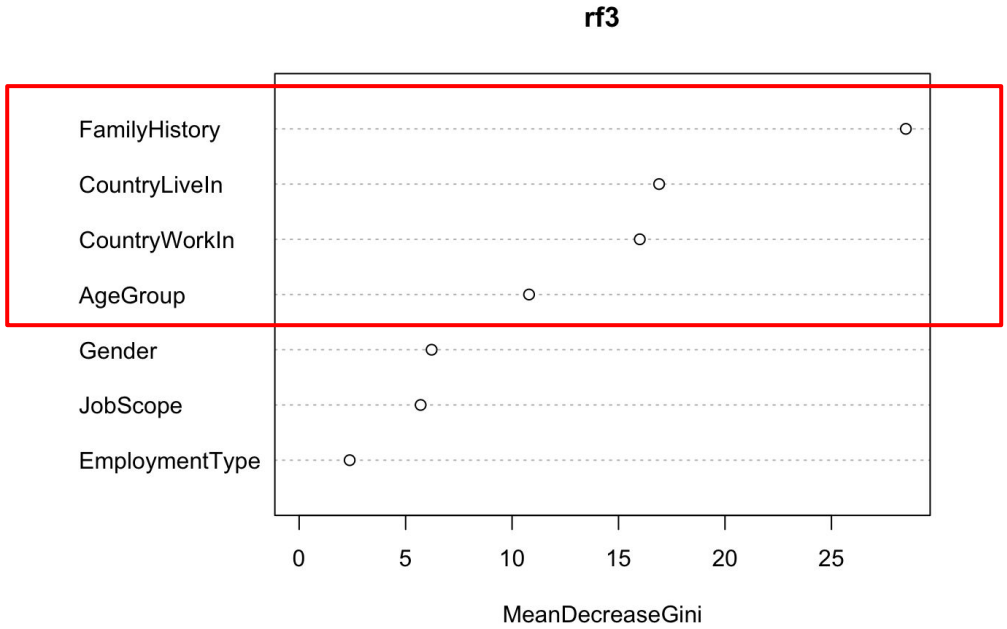
No Information Rate : 0.6383

P-Value [Acc > NIR] : 1.613e-08

- Our accuracy rate improved to 0.74 from 0.46

Results

Research Qn 1: Features affecting Mental Health Disorders



Approach and Methodology - Data Cleaning

Research Qn 2: Does the pandemic has an effect on Mental Health Disorders?

- X (Feature): Before/After the pandemic
- Y (Outcome): Presence of Mental Health Disorders
- Split the dataset into before and after the pandemic
  - Rename 2019 data as “before”
  - Rename 2020 and 2021 data as “after”
- Run the chi-square test
  - Both categorical data

Results

Research Qn 2: Does the pandemic has an effect on Mental Health Disorders?

```
chi <- chisq.test(tb, correct = F)
chi$p.value
# > chi$p.value
# [1] 0.002057285
```

- As p-value < 0.05, reject the null hypothesis
- In short, pandemic most likely has an effect on the presence of Mental Health Disorders

Approach and Methodology - Data Cleaning

Research Qn 3: Does the pandemic have an effect on the attitude towards Mental Health DIsorders?

- Select the questions from the survey that are related from the 3 dataset
  - Would you feel comfortable discussing a mental health issue with your direct supervisor?
  - Would you feel comfortable discussing a mental health issue with your coworker?
  - Both qns are the features (Independent variable)
- Combine all the data into 1 dataset
- Split the dataset into before and after the pandemic
  - Rename 2019 data as “before”
  - Rename 2020 and 2021 data as “after”
- Both independent and dependent variable are categorical data→ chi-square test

Results

Research Qn 3: Does the pandemic have an effect on the attitude towards Mental Health DIsorders?

- Discuss with direct supervisors
- P-value <0.05: pandemic most likely has an effect on the discussion of mental health disorders with direct supervisor

```
chi <- chisq.test(tb, correct = F)
chi$p.value
# [1] 0.04401081
```

- Discuss with coworkers
- P-value >0.05: pandemic most likely does not has an effect on the discussion of mental health disorders with coworkers

```
chi <- chisq.test(tb, correct = F)
chi$p.value
# [1] 0.5038958
```

Results

Research Qn 3: Does the pandemic have an effect on the attitude towards Mental Health Disorders?

- Possible reason: people are more comfortable sharing their personal issues with coworkers but not direct supervisors.

- However, the pandemic changed this perception.

Conclusion

1. What are the factors that might explain cases of mental health disorders at the workplace?
    - Top factors: Family History, Country that people live and work in, age group
    - Family History of mental health disorder, Age 40-49, USA
2. Does the pandemic have an effect on mental health disorders?
    - Yes!
3. Does the pandemic have an effect on the attitudes towards mental health disorders?
    - Discussion of mental health disorders with supervisors? Yes
    - Discussion of mental health disorders with coworkers? No

Future Research Ideas

- Is there an Increase or decrease in mental health disorders due to the pandemic?

- Are more or less people willing to share their mental health disorders with their direct supervisor?

- Find other models with better accuracy rate, e.g. logistic regression

References

- Open Sourcing Mental Health. (2014). *Research*. Open Sourcing Mental Illness - Changing how we talk about mental health in the tech community - Stronger Than Fear. Retrieved April 23, 2022, from https://osmhhelp.org/research
- Fernandez , R., Sikhosana , N., Green, H., Halcomb , E. J., Middleton, R., Alananzeh, I., Trakis , S., & Moxham , L. (2021, September 21). *Anxiety and depression among healthcare workers during the COVID-19 pandemic: A Systematic Umbrella Review of the global evidence*. BMJ open. Retrieved April 23, 2022, from https://pubmed.ncbi.nlm.nih.gov/34548373/