UNIVERSITY OF WESTMINSTER

INFORMATICS INSTITUTE OF TECHNOLOGY

5DATA006C

Data Visualization

PORTFOLIO

Module Leader:

Fouzul Hassan

Full Name: Sithuli Nayanama Kothalwala

Blackboard Name: Sithuli Kothalawala

IIT ID: 20230336

UOW ID: 20521288

UOW username : w2052128

Table of Contents

Research Question and Data Sourcing

Research Question : "How does workplace culture influence employees seeking treatment for mental health issues?"

Relevance:

Understanding the influence of workplace culture on mental health treatment decision is crucial in promoting supportive work environments. Mental health challenges are common among employees but often remain unaddressed due to workplace barriers, such as unsupportive co-workers, managers or inadequate leave policies. Addressing these factors can help organizations implement policies that encourage mental well-being, improve employee satisfaction, and enhance productivity.

Dataset used :

The analysis utilizes the *"Mental Health in Tech Survey"* dataset provided by Open Sourcing Mental Illness on Kaggle. This dataset contains responses over 1200 individuals, primarily working in technological-relates fields. It includes information on demographics, workplace characteristics, and mental health-related experiences.

The dataset is ideal for exploring this research question as it captures variables related to workplace culture (e.g., support from co-workers and supervisors, leave policies, remote work policies) and mental health treatment-seeking behaviour. For this study, key variables include whether employees sought treatment for mental health (treatment) and workplace-related factors (co-workers, supervisor, leave, benefits, remote work, work interference).

Reference:

Open Sourcing Mental Illness, LTD (2014). *Mental Health in Tech Survey*. [online] Kaggle.com. Available at: https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey [Accessed 30 Dec. 2024].

Data Preparation

Tidiness Evaluation

The dataset is particularly tidy, but improvements are needed.

Observations

1. The dataset has 1260 entries and 27 columns
2. Columns include demographics details, treatment status, physical and mental health interview, and workplace related factors.
3. Some columns have irrelevant data, missing values, and format issues that need to be dealt with.

Variables

Dependant variable - treatment

Independent -

- Demographic information : Age, Gender, Country
- Health Interview: mental_health_interview, seek_help
- Workplace related factors : work_interfere, remote_work, coworkers, supervisor, leave, benefits)

Assessment of Tidiness

Some columns, such as ‘comments’, have excessive missing values, making them less useful for analysis. Since this column is not directly relevant to the research question or the analysis it could be dropped.

Some columns have mixed data types (e.g, categorical values stored as strings)

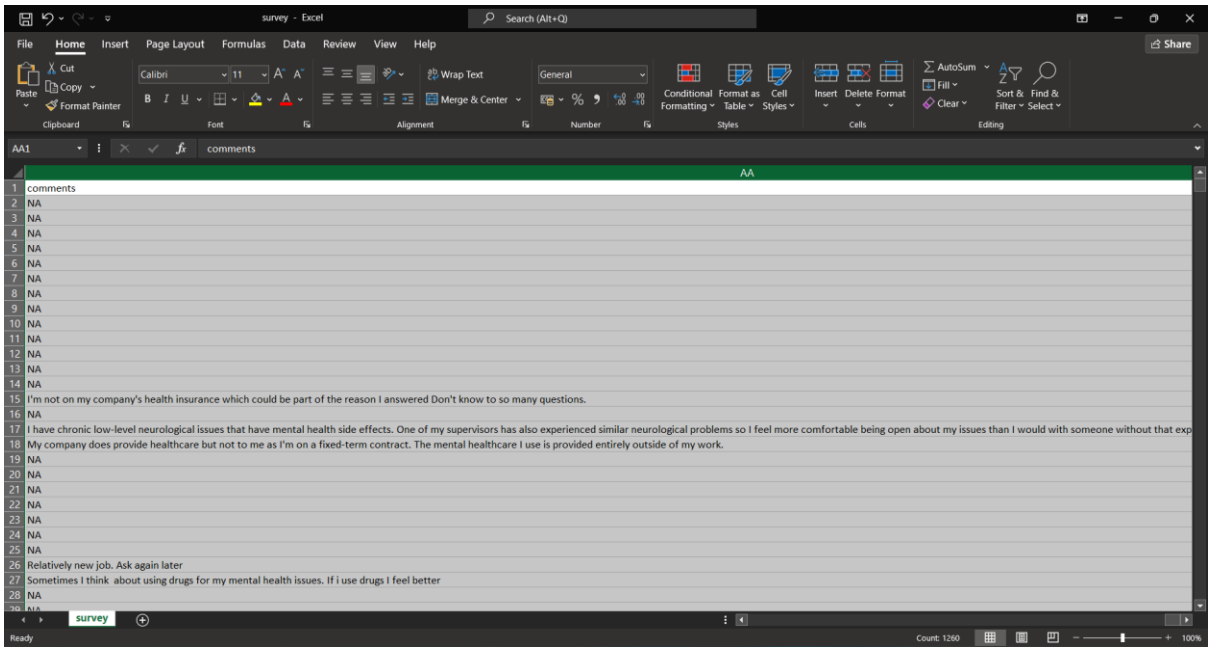
Missing values in columns such as ‘self_employed’ and ‘work_interfere’ need handling.

The same word has been written in different ways which causes errors in analysis (In Gender column Male and Female are written in many different ways and with typos.)

There is one duplicate row that need to be removed.

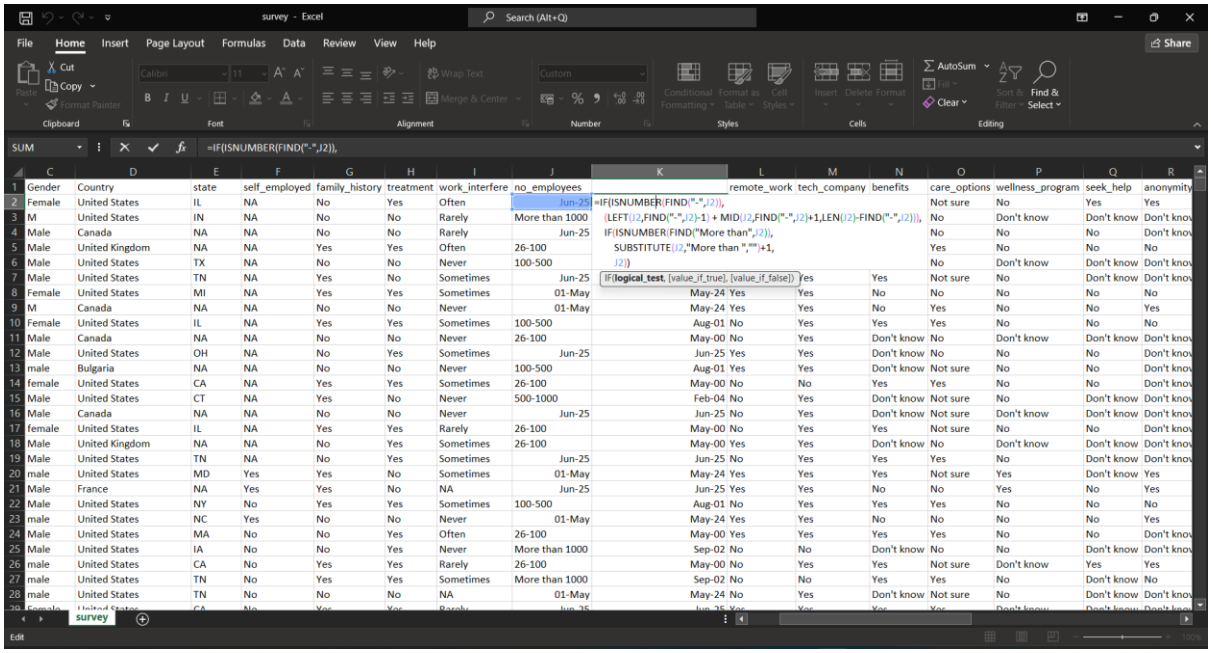
Tidy Dataset Transformation

1. Removing irrelevant columns

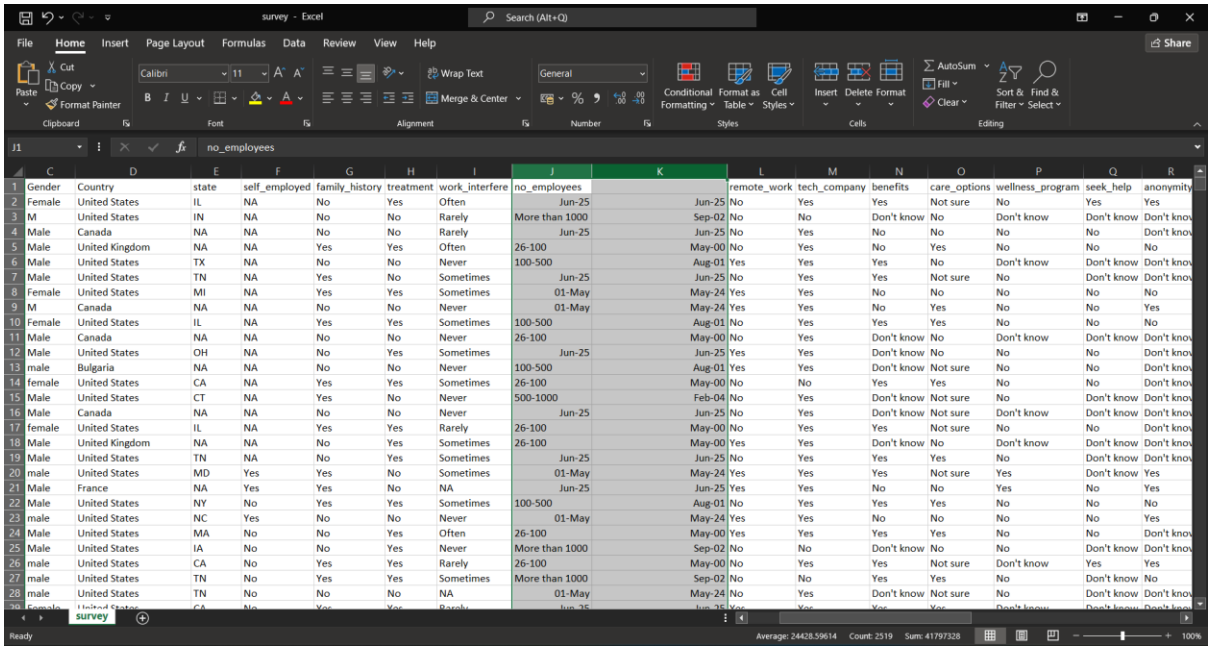


2. Formatting no_employees column

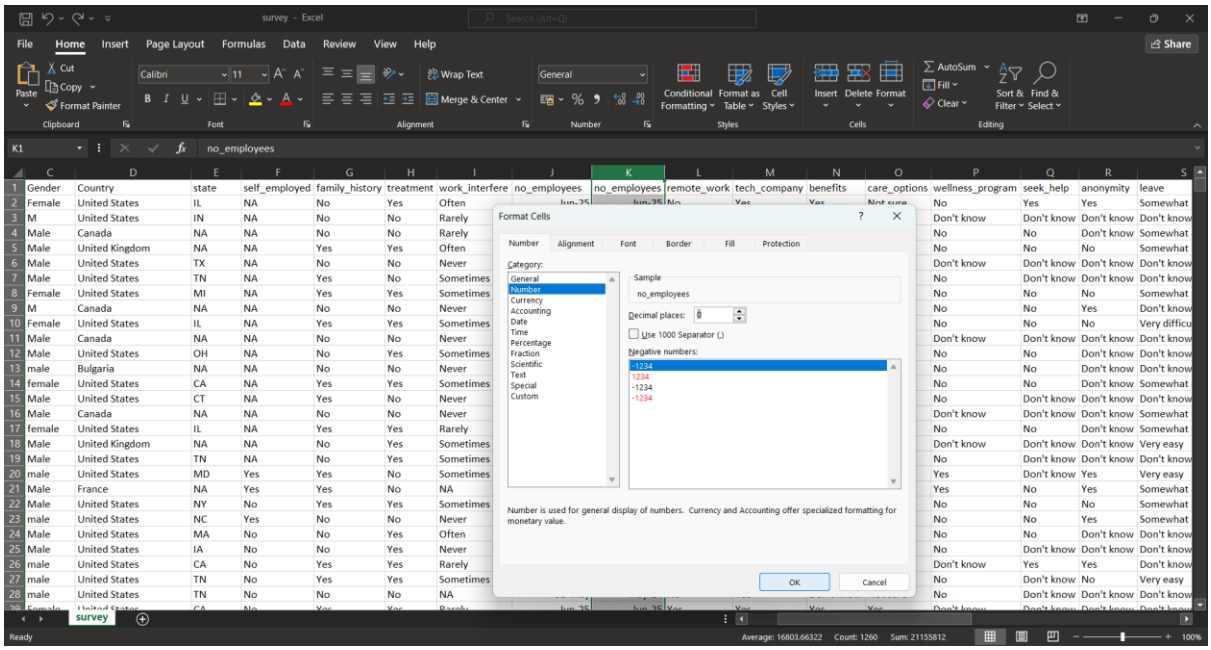
This formular is used to get the mid-point of ranges in the column. Phrases like 'More than 1000' are replaced with a numeric approximation like 1001. Values with neither range nor descriptor remain unchanged.

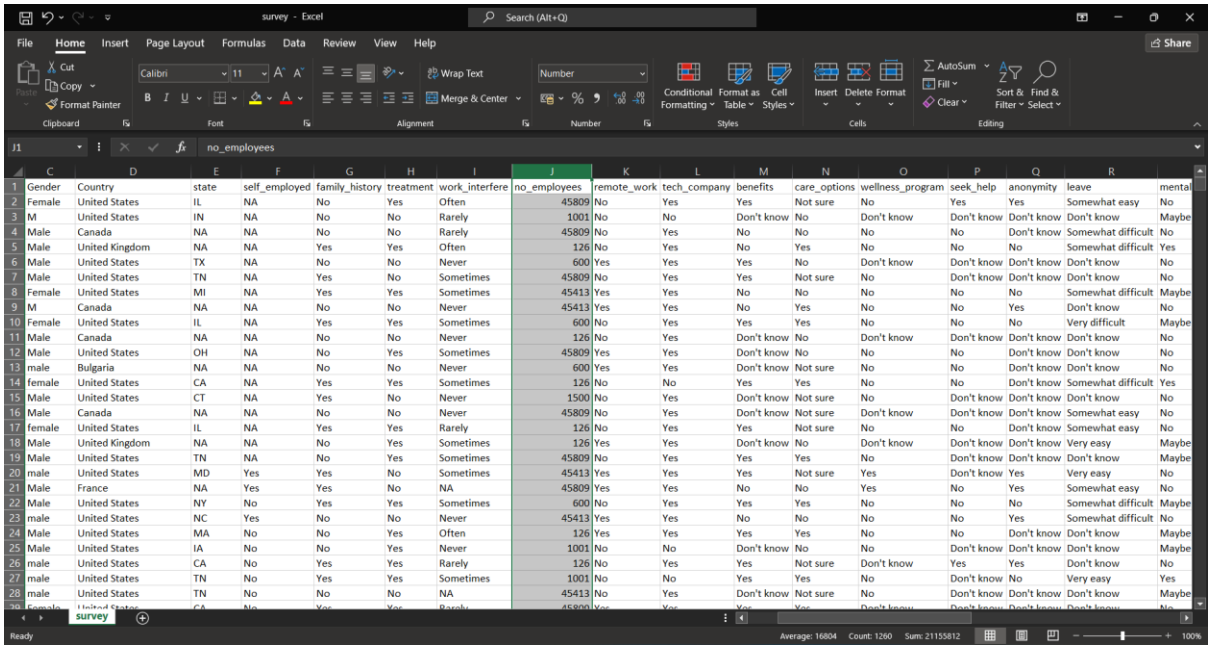


Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	remote_work	tech_company	benefits	care_options	wellness_program	seek_help	anonymity
Female	United States	IL	NA	No	Yes	Often	Jun-25				Not sure	No	Yes	Yes
M	United States	IN	NA	No	No	Rarely	More than 1000				No	Don't know	Don't know	Don't know
Male	Canada	NA	NA	No	No	Rarely	Jun-25				No	No	No	Don't know
Male	United Kingdom	NA	NA	Yes	Yes	Often	26-100				Yes	No	No	No
Male	United States	TX	NA	No	No	Never	100-500				No	Don't know	Don't know	Don't know
Male	United States	TN	NA	Yes	No	Sometimes	Jun-25		Yes	Yes	Not sure	No	Don't know	Don't know
Female	United States	MI	NA	Yes	Yes	Sometimes	01-May	May-24	Yes	No	No	No	No	No
M	Canada	NA	NA	No	No	Never	01-May	May-24	Yes	No	Yes	No	No	Yes
Female	United States	IL	NA	Yes	Yes	Sometimes	100-500	Aug-01	No	Yes	Yes	No	No	No
Male	Canada	NA	NA	No	No	Never	26-100	May-00	No	Yes	Don't know	No	Don't know	Don't know
Male	United States	OH	NA	No	Yes	Sometimes	Jun-25	Jun-25	Yes	Yes	Don't know	No	No	Don't know
male	Bulgaria	NA	NA	No	No	Never	100-500	Aug-01	Yes	Yes	Don't know	Not sure	No	Don't know
female	United States	CA	NA	Yes	Yes	Sometimes	26-100	May-00	No	No	Yes	Yes	No	Don't know
Male	United States	CT	NA	Yes	No	Never	500-1000	Feb-04	No	Yes	Don't know	Not sure	No	Don't know
Male	Canada	NA	NA	No	No	Never	Jun-25	Jun-25	No	Yes	Don't know	Not sure	Don't know	Don't know
female	United States	IL	NA	Yes	Yes	Rarely	26-100	May-00	No	Yes	Yes	Not sure	No	Don't know
Male	United Kingdom	NA	NA	No	Yes	Sometimes	26-100	May-00	Yes	Yes	Don't know	No	Don't know	Don't know
Male	United States	TN	NA	No	Yes	Sometimes	Jun-25	Jun-25	No	Yes	Yes	Yes	No	Don't know
male	United States	MD	Yes	Yes	No	Sometimes	01-May	May-24	Yes	Yes	Not sure	Yes	Don't know	Yes
Male	France	NA	Yes	Yes	No	NA	Jun-25	Jun-25	Yes	Yes	No	No	Yes	No
Male	United States	NY	No	Yes	Yes	Sometimes	100-500	Aug-01	No	Yes	Yes	Yes	No	No
male	United States	NC	Yes	No	No	Never	01-May	May-24	Yes	Yes	No	No	No	No
Male	United States	MA	No	No	Yes	Often	26-100	May-00	Yes	Yes	Yes	Yes	No	No
Male	United States	IA	No	No	Yes	Never	More than 1000	Sep-02	No	No	Don't know	No	No	Don't know
male	United States	CA	No	Yes	Yes	Rarely	26-100	May-00	No	Yes	Yes	Not sure	Don't know	Yes
male	United States	TN	No	Yes	Yes	Sometimes	More than 1000	Sep-02	No	No	Yes	Yes	No	Don't know
male	United States	TN	No	No	No	NA	01-May	May-24	No	Yes	Don't know	Not sure	No	Don't know



Then the values in the column are formatted into Numeric values as follows.



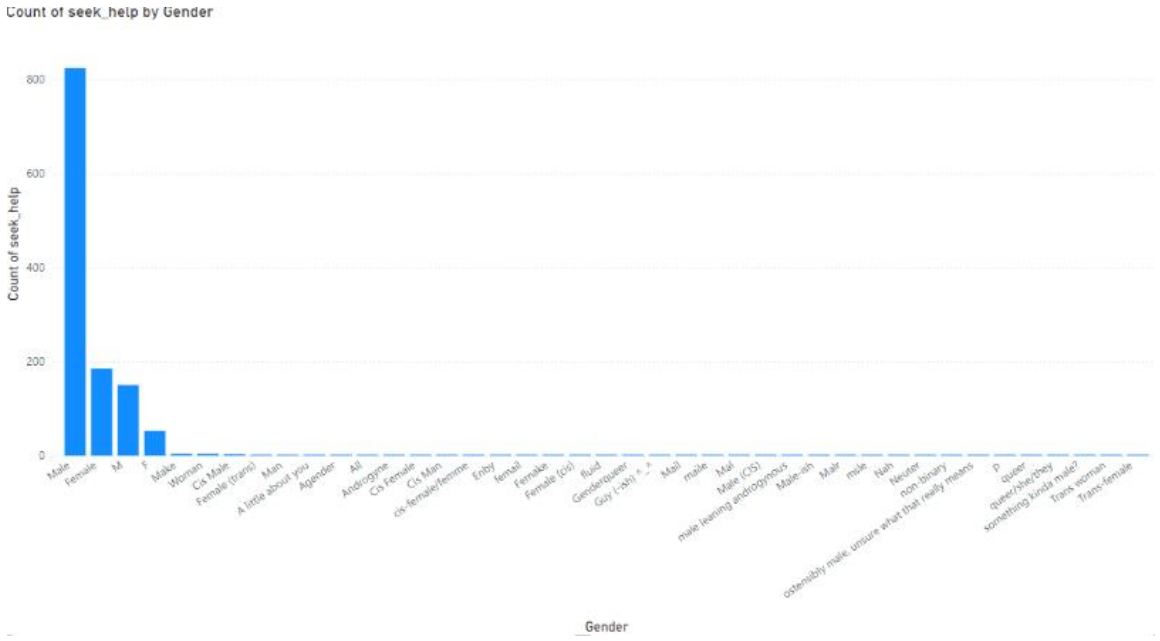


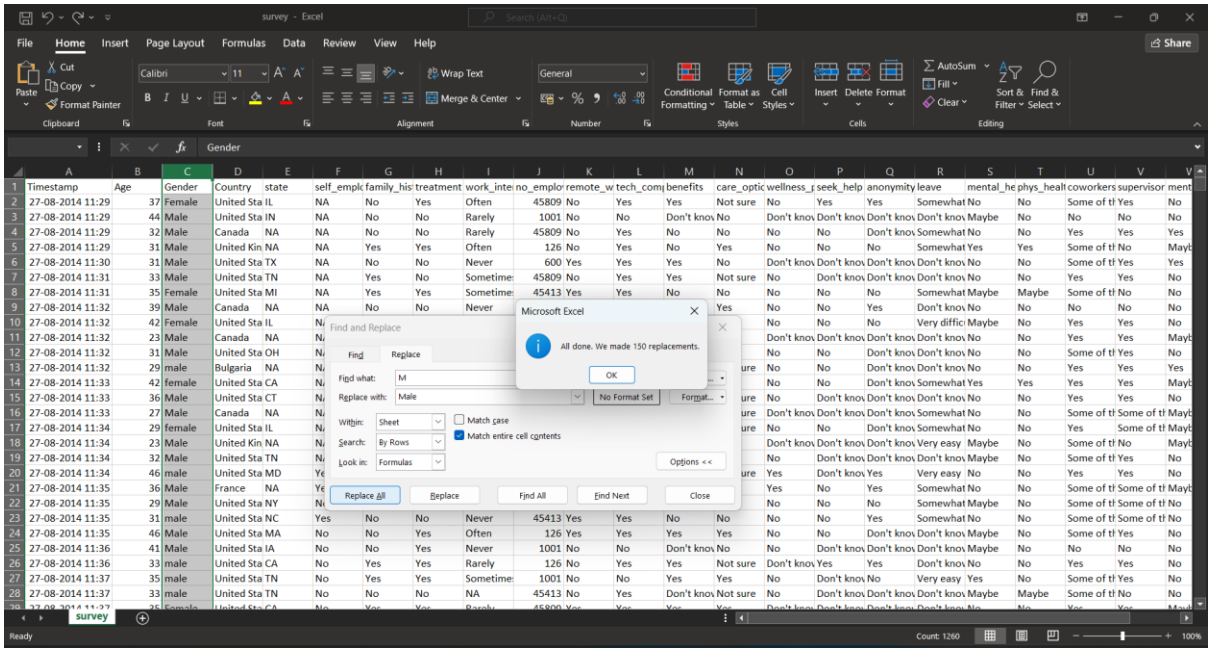
Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	remote_work	tech_company	benefits	care_options	wellness_program	seek_help	anonymity	leave	mental
Female	United States	IL	NA	No	Yes	Often	45809	No	Yes	Yes	Not sure	No	Yes	Yes	Somewhat easy	No
M	United States	IN	NA	No	No	Rarely	1001	No	No	Don't know	No	Don't know	Don't know	Don't know	Don't know	Maybe
Male	Canada	NA	NA	No	No	Rarely	45809	No	Yes	No	No	No	No	Don't know	Somewhat difficult	No
Male	United Kingdom	NA	NA	Yes	Yes	Often	126	No	Yes	No	Yes	No	No	No	Somewhat difficult	Yes
Male	United States	TX	NA	No	No	Never	600	Yes	Yes	Yes	No	Don't know	Don't know	Don't know	Don't know	No
Male	United States	TN	NA	Yes	No	Sometimes	45809	No	Yes	Yes	Not sure	No	Don't know	Don't know	Don't know	No
Female	United States	MI	NA	Yes	Yes	Sometimes	45413	Yes	Yes	No	No	No	No	No	Somewhat difficult	Maybe
M	Canada	NA	NA	No	No	Never	45413	Yes	Yes	No	Yes	No	No	Yes	Don't know	No
Female	United States	IL	NA	Yes	Yes	Sometimes	600	No	Yes	Yes	Yes	No	No	No	Very difficult	Maybe
Male	Canada	NA	NA	No	No	Never	126	No	Yes	Don't know	No	Don't know	Don't know	Don't know	Don't know	No
Male	United States	OH	NA	No	Yes	Sometimes	45809	Yes	Yes	Don't know	No	No	No	Don't know	Don't know	No
male	Bulgaria	NA	NA	No	No	Never	600	Yes	Yes	Don't know	Not sure	No	No	Don't know	Don't know	No
female	United States	CA	NA	Yes	Yes	Sometimes	126	No	No	Yes	Yes	No	No	Don't know	Somewhat difficult	Yes
Male	United States	CT	NA	Yes	No	Never	1500	No	Yes	Don't know	Not sure	No	Don't know	Don't know	Don't know	No
Male	Canada	NA	NA	No	No	Never	45809	No	Yes	Don't know	Not sure	Don't know	Don't know	Don't know	Somewhat easy	No
female	United States	IL	NA	Yes	Yes	Rarely	126	No	Yes	Yes	Not sure	No	No	Don't know	Somewhat easy	No
Male	United Kingdom	NA	NA	No	Yes	Sometimes	126	Yes	Yes	Don't know	No	Don't know	Don't know	Don't know	Very easy	Maybe
Male	United States	TN	NA	No	Yes	Sometimes	45809	No	Yes	Yes	Yes	No	Don't know	Don't know	Don't know	Maybe
male	United States	MD	Yes	Yes	No	Sometimes	45413	Yes	Yes	Yes	Not sure	Yes	Don't know	Yes	Very easy	No
Male	France	NA	Yes	Yes	No	NA	45809	Yes	Yes	No	No	Yes	No	Yes	Somewhat easy	No
Male	United States	NY	No	Yes	Yes	Sometimes	600	No	Yes	Yes	Yes	No	No	No	Somewhat difficult	Maybe
male	United States	NC	Yes	No	No	Never	45413	Yes	Yes	No	No	No	No	Yes	Somewhat difficult	No
Male	United States	MA	No	No	Yes	Often	126	Yes	Yes	Yes	Yes	No	No	Don't know	Don't know	Maybe
Male	United States	IA	No	No	Yes	Never	1001	No	No	Don't know	No	No	No	Don't know	Don't know	Maybe
male	United States	CA	No	Yes	Yes	Rarely	126	No	Yes	Yes	Not sure	Don't know	Yes	Yes	Don't know	No
male	United States	TN	No	Yes	Yes	Sometimes	1001	No	No	Yes	Yes	No	Don't know	No	Very easy	Yes
male	United States	TN	No	No	No	NA	45413	No	Yes	Don't know	Not sure	No	Don't know	Don't know	Don't know	Maybe

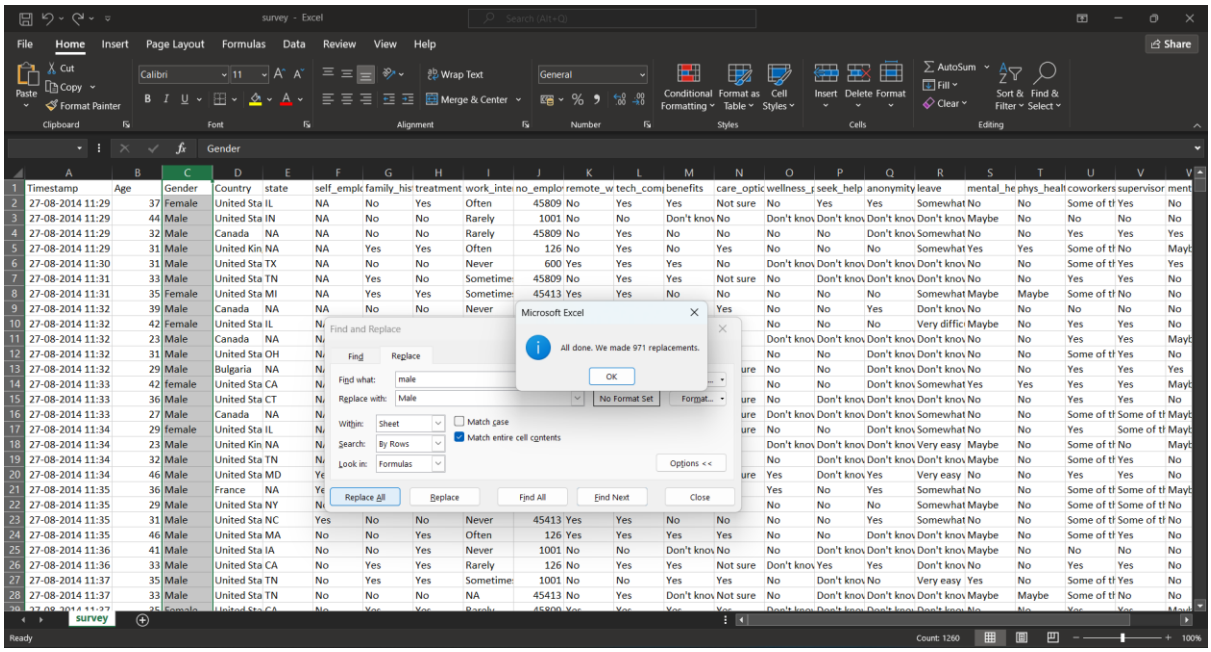
3. Cleaning the Gender column

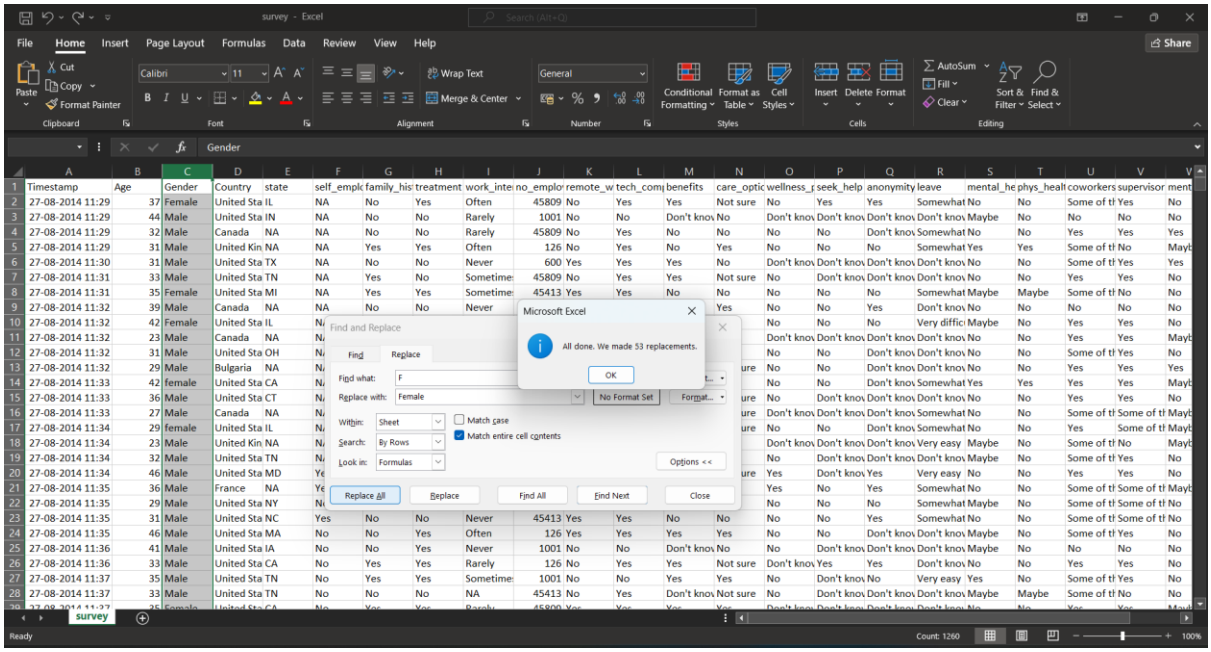
In the Gender column Male was written in many different ways such as M, male and even with typos like make and Mal. Female was also written in many different ways such as female and femake. These were cleaned using Find & Replace dialogue box (Ctrl + H).

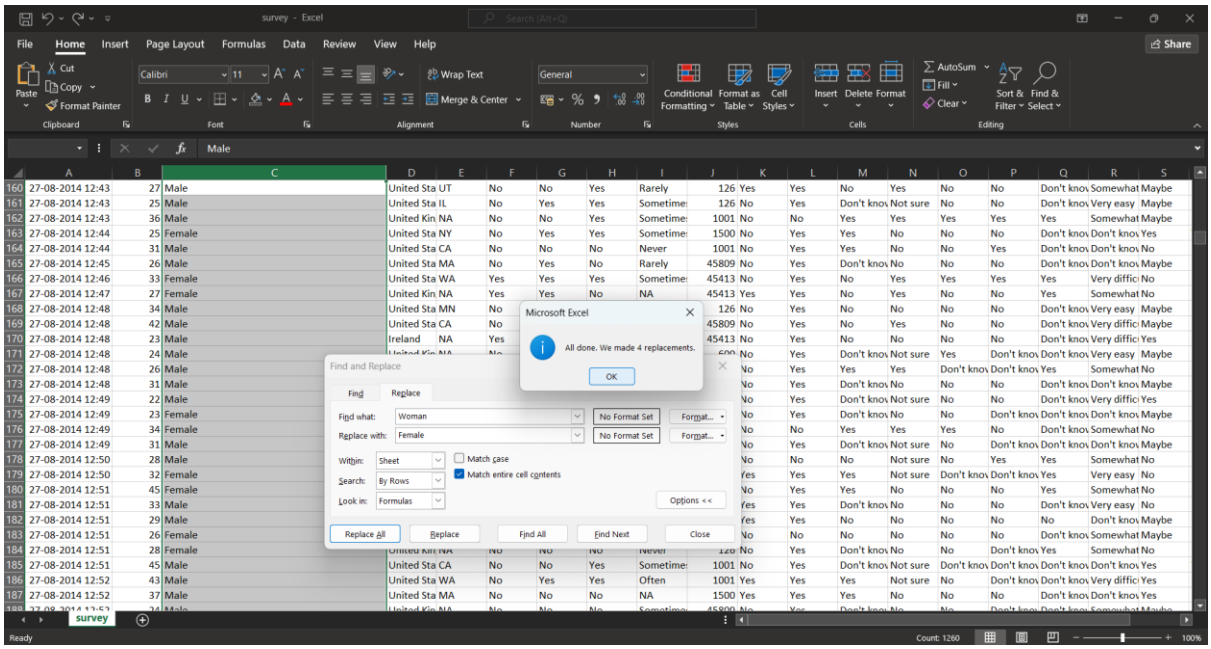
Other than that, there were invalid values which were manually replaced to closest value or assigned as ‘Other’.

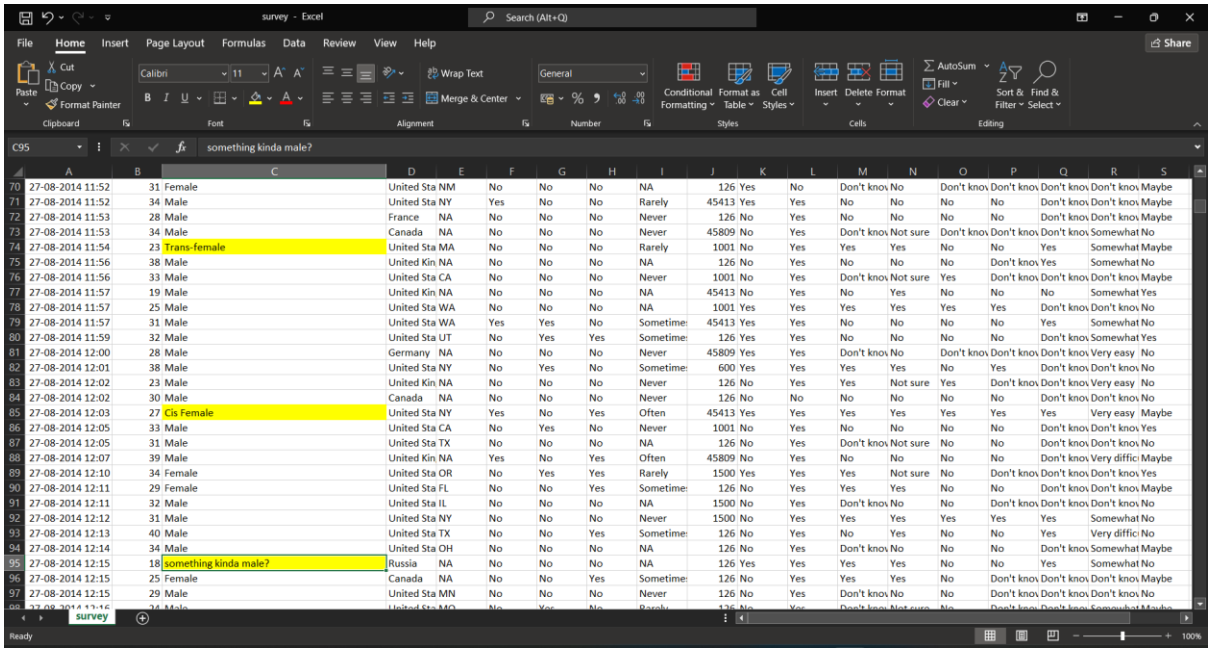


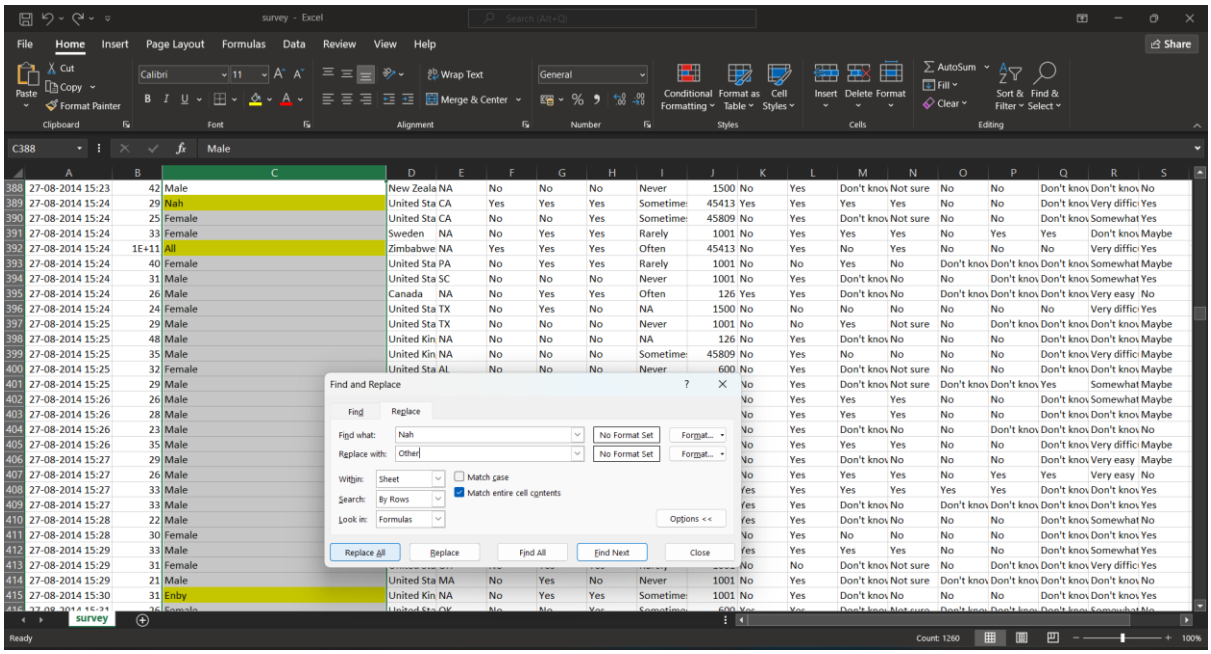

survey - Excel
Microsoft Excel
All done. We made 150 replacements.
OK
Find and Replace
Find what: M
Replace with: Male
Replace All
Replace
Find All
Find Next
Close

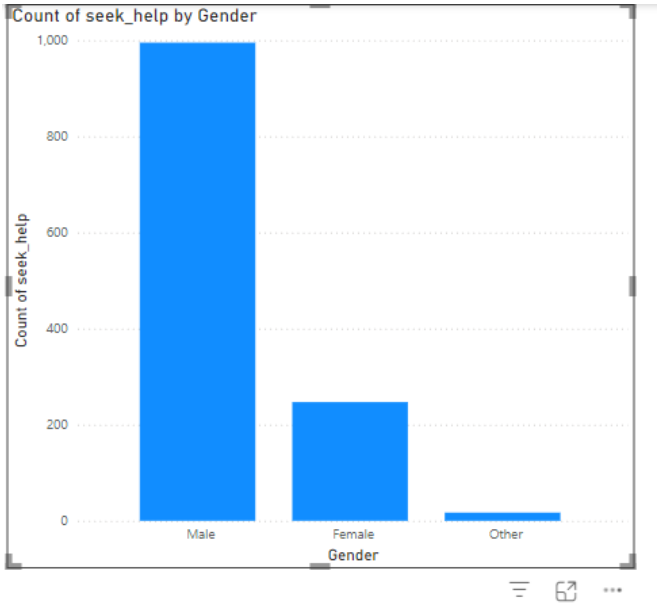

survey - Excel
Microsoft Excel
All done. We made 971 replacements.
OK
Find and Replace
Find what: male
Replace with: Male
Replace All
Replace
Find All
Find Next
Close


survey - Excel
Microsoft Excel
All done. We made 53 replacements.
OK
Find and Replace
Find
Replace
Find what: F
Replace with: Female
No Format Set
Format...
Within: Sheet
Search: By Rows
Look in: Formulas
Match case
Match entire cell contents
Options <<
Replace All
Replace
Find All
Find Next
Close
survey
Ready


survey - Excel
Microsoft Excel
All done. We made 4 replacements.
OK
Find and Replace
Find
Replace
Find what: Woman
Replace with: Female
No Format Set
Format...
Within: Sheet
Search: By Rows
Look in: Formulas
Match case
Match entire cell contents
Options <<
Replace All
Replace
Find All
Find Next
Close
survey
Ready


survey - Excel
Search (Alt+Q)
File Home Insert Page Layout Formulas Data Review View Help
Share
Paste Cut Copy Format Painter
Calibri 11
Wrap Text
Merge & Center
General
Conditional Formatting
Format as Table
Cell Styles
Insert Delete Format
AutoSum Fill Clear
Sort & Filter
Find & Select
Clipboard Font Alignment Number Styles Cells Editing
C95
something kinda male?
survey
Ready

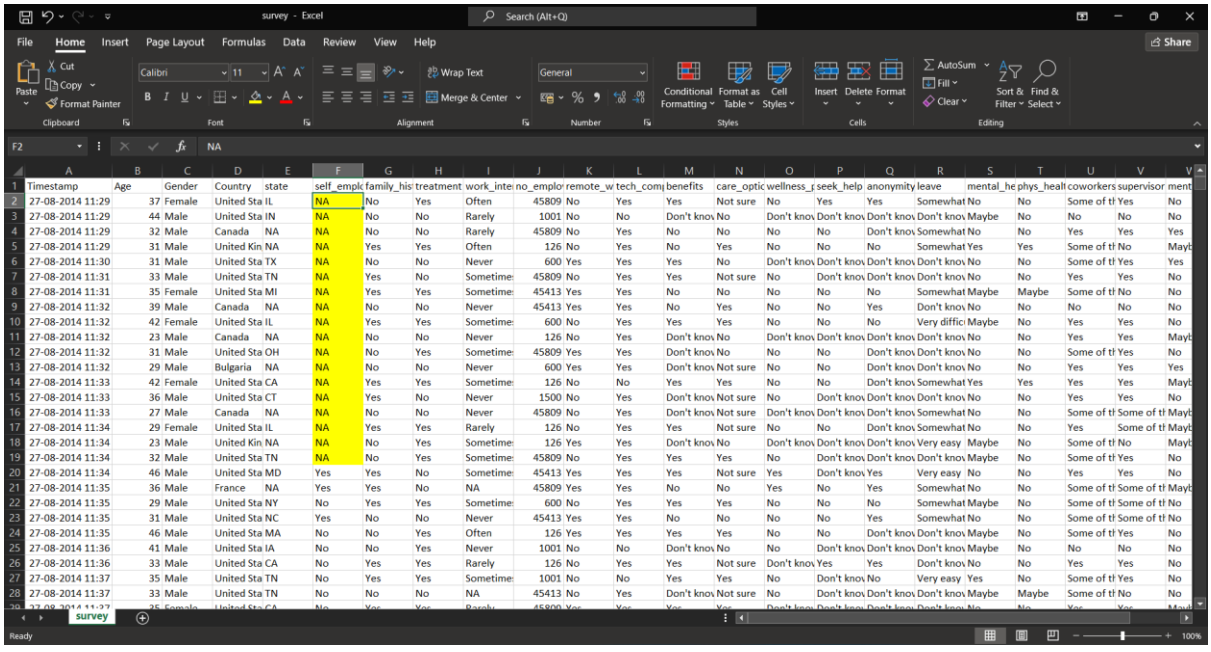

survey - Excel
Search (Alt+Q)
File Home Insert Page Layout Formulas Data Review View Help
Share
Paste Cut Copy Format Painter
Calibri 11
Wrap Text
Merge & Center
General
Conditional Formatting
Format as Table
Cell Styles
Insert Delete Format
AutoSum Fill Clear
Sort & Filter
Find & Select
Clipboard Font Alignment Number Styles Cells Editing
C388
Male
Find and Replace
Find
Replace
Find what: Nah
No Format Set
Format...
Replace with: Other
No Format Set
Format...
Within: Sheet
Match case
Search: By Rows
Match entire cell contents
Look in: Formulas
Options <<
Replace All
Replace
Find All
Find Next
Close
survey
Ready
Count: 1260

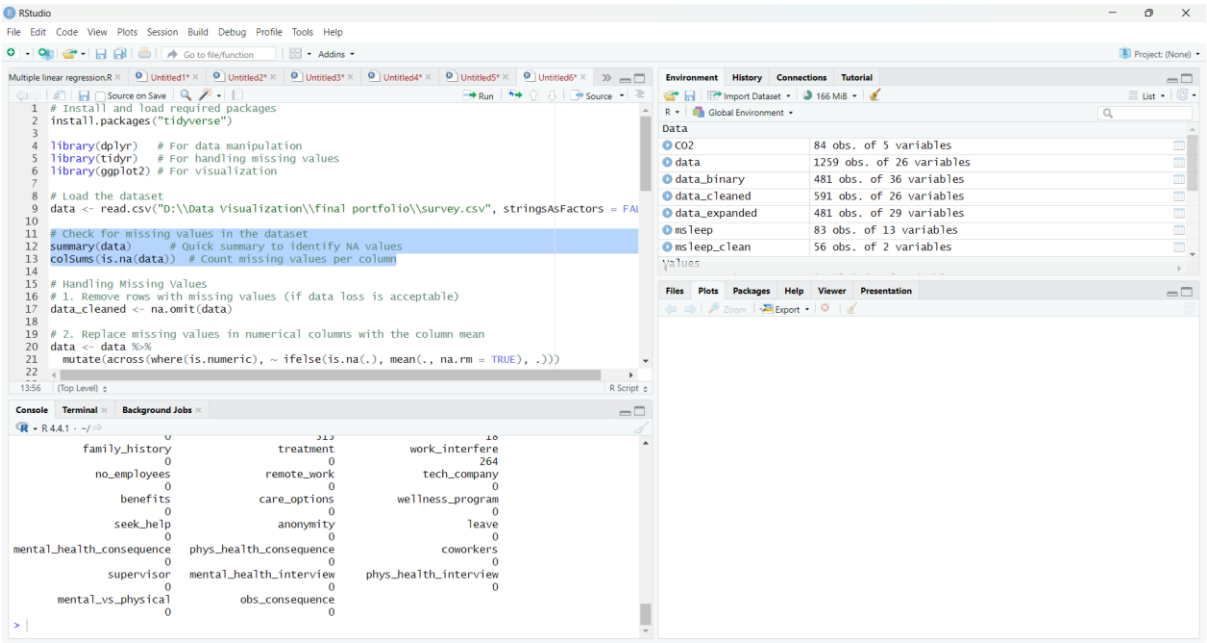


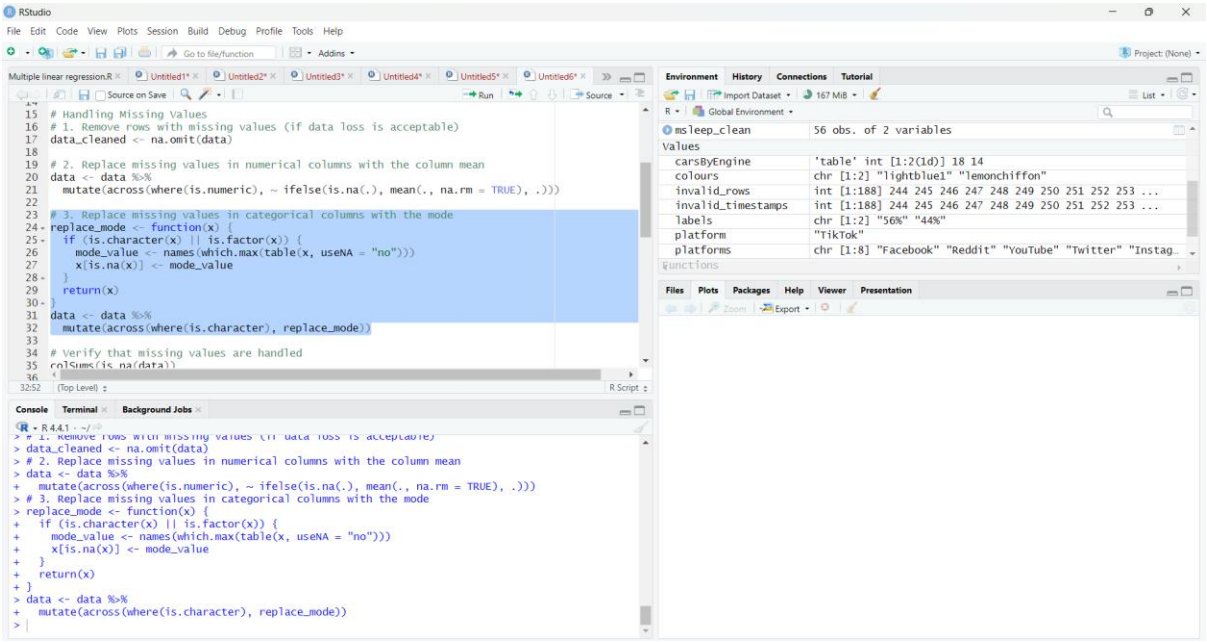
4. Handling Missing Values and outliers

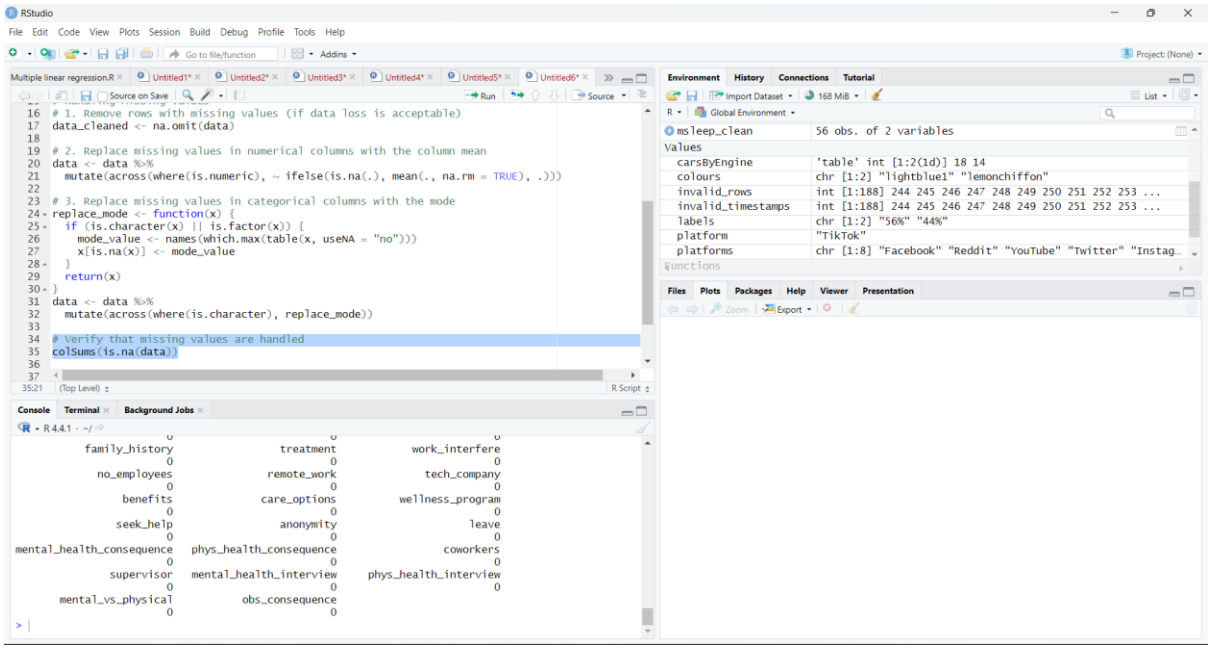
Handling missing values

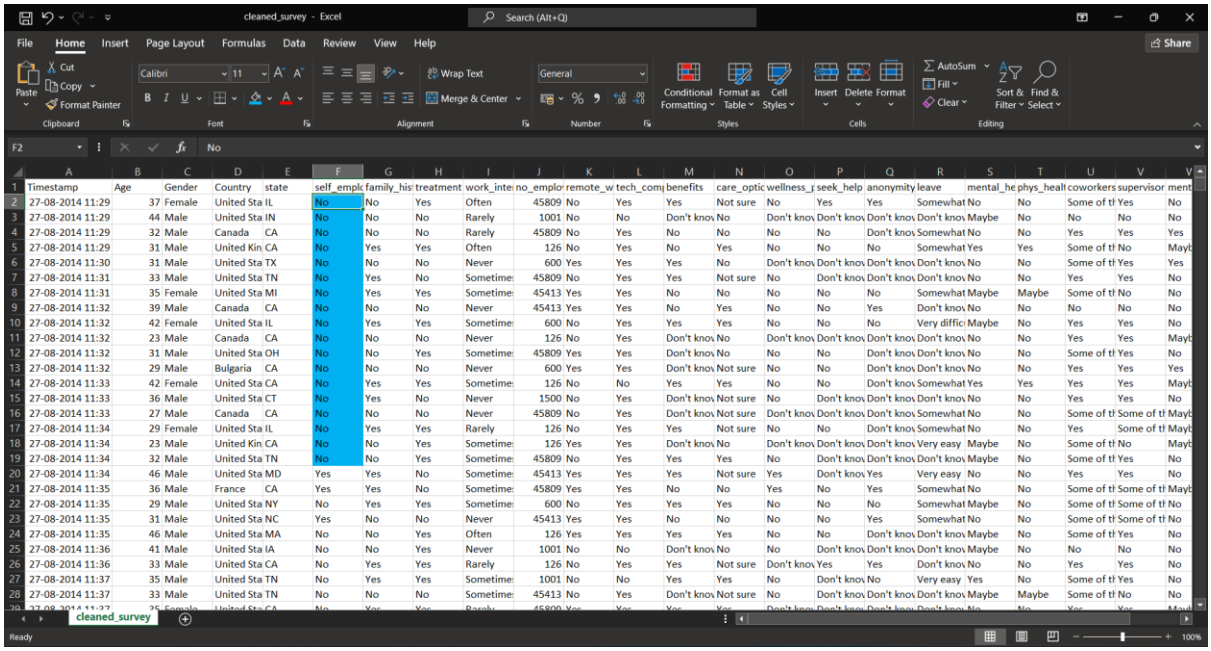
The missing values in the columns, self_employed and work_interfere are replaced with mode which is 'No' in this case and the cleaned dataset is saved into a new .csv file using an R code.




RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Multiple linear regression.R
Untitled1*
Untitled2*
Untitled3*
Untitled4*
Untitled5*
Untitled6*
# Install and load required packages
install.packages("tidyverse")
library(dplyr) # For data manipulation
library(tidyr) # For handling missing values
library(ggplot2) # For visualization
# Load the dataset
data <- read.csv("D:\\Data Visualization\\final portfolio\\survey.csv", stringsAsFactors = FA
# Check for missing values in the dataset
summary(data) # Quick summary to identify NA values
colSums(is.na(data)) # Count missing values per column
# Handling Missing Values
# 1. Remove rows with missing values (if data loss is acceptable)
data_cleaned <- na.omit(data)
# 2. Replace missing values in numerical columns with the column mean
data <- data %>%
mutate(across(where(is.numeric), ~ ifelse(is.na(.), mean(., na.rm = TRUE), .)))
Console
Terminal
Background Jobs
family_history 0
treatment 0
work_interfere 264
no_employees 0
remote_work 0
tech_company 0
benefits 0
care_options 0
wellness_program 0
seek_help 0
anonymity 0
leave 0
mental_health_consequence 0
phys_health_consequence 0
coworkers 0
supervisor 0
mental_health_interview 0
phys_health_interview 0
mental_vs_physical 0
obs_consequence 0
Environment
History
Connections
Tutorial
Import Dataset
166 MiB
Global Environment
Data
CO2 84 obs. of 5 variables
data 1259 obs. of 26 variables
data_binary 481 obs. of 36 variables
data_cleaned 591 obs. of 26 variables
data_expanded 481 obs. of 29 variables
msleep 83 obs. of 13 variables
msleep_clean 56 obs. of 2 variables
Values
Files
Plots
Packages
Help
Viewer
Presentation
Export

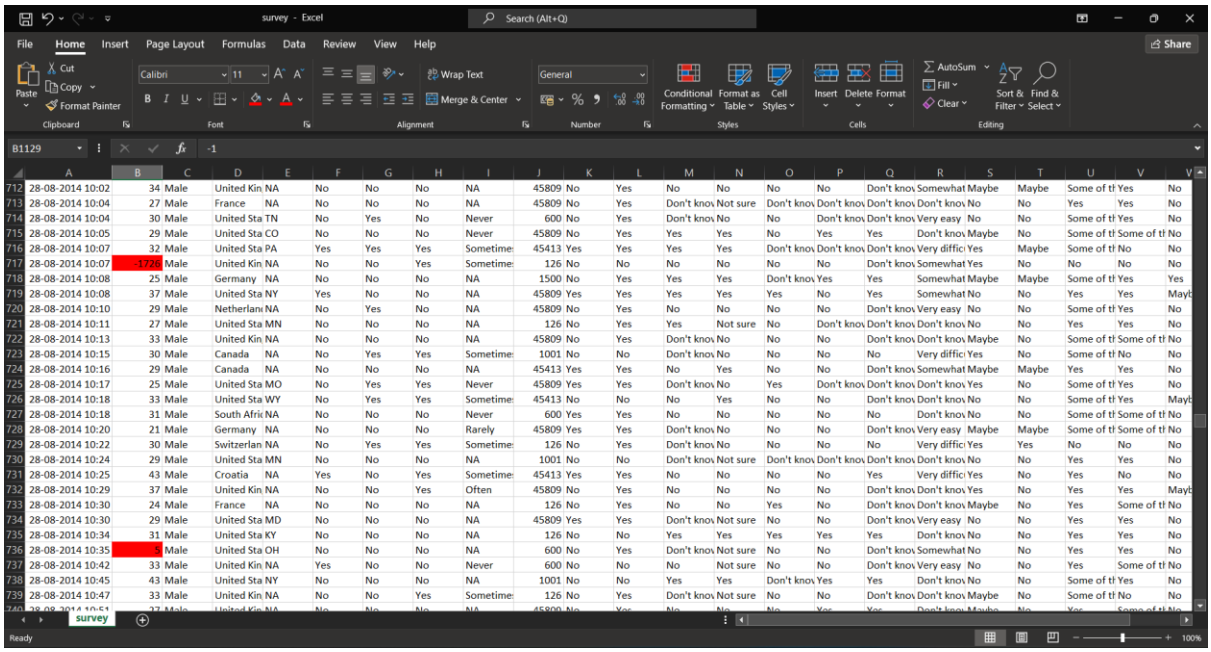

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Multiple linear regression.R
# Handling Missing Values
# 1. Remove rows with missing values (if data loss is acceptable)
data_cleaned <- na.omit(data)
# 2. Replace missing values in numerical columns with the column mean
data <- data %>%
mutate(across(where(is.numeric), ~ ifelse(is.na(.), mean(., na.rm = TRUE), .)))
# 3. Replace missing values in categorical columns with the mode
replace_mode <- function(x) {
if (is.character(x) || is.factor(x)) {
mode_value <- names(which.max(table(x, useNA = "no")))
x[is.na(x)] <- mode_value
}
return(x)
}
data <- data %>%
mutate(across(where(is.character), replace_mode))
# Verify that missing values are handled
colSums(is.na(data))
Console
Terminal
Background Jobs
> # 1. Remove rows with missing values (if data loss is acceptable)
> data_cleaned <- na.omit(data)
> # 2. Replace missing values in numerical columns with the column mean
> data <- data %>%
+ mutate(across(where(is.numeric), ~ ifelse(is.na(.), mean(., na.rm = TRUE), .)))
> # 3. Replace missing values in categorical columns with the mode
> replace_mode <- function(x) {
+ if (is.character(x) || is.factor(x)) {
+ mode_value <- names(which.max(table(x, useNA = "no")))
+ x[is.na(x)] <- mode_value
+ }
+ return(x)
+ }
> data <- data %>%
+ mutate(across(where(is.character), replace_mode))
Environment
History
Connections
Tutorial
Import Dataset
167 MiB
Global Environment
msleep_clean 56 obs. of 2 variables
Values
carsByEngine 'table' int [1:2(1d)] 18 14
colours chr [1:2] "lightblue1" "lemonchiffon"
invalid_rows int [1:188] 244 245 246 247 248 249 250 251 252 253 ...
invalid_timestamps int [1:188] 244 245 246 247 248 249 250 251 252 253 ...
labels chr [1:2] "56%" "44%"
platform "TikTok"
platforms chr [1:8] "Facebook" "Reddit" "YouTube" "Twitter" "Instag...
Functions
Files
Plots
Packages
Help
Viewer
Presentation
Export

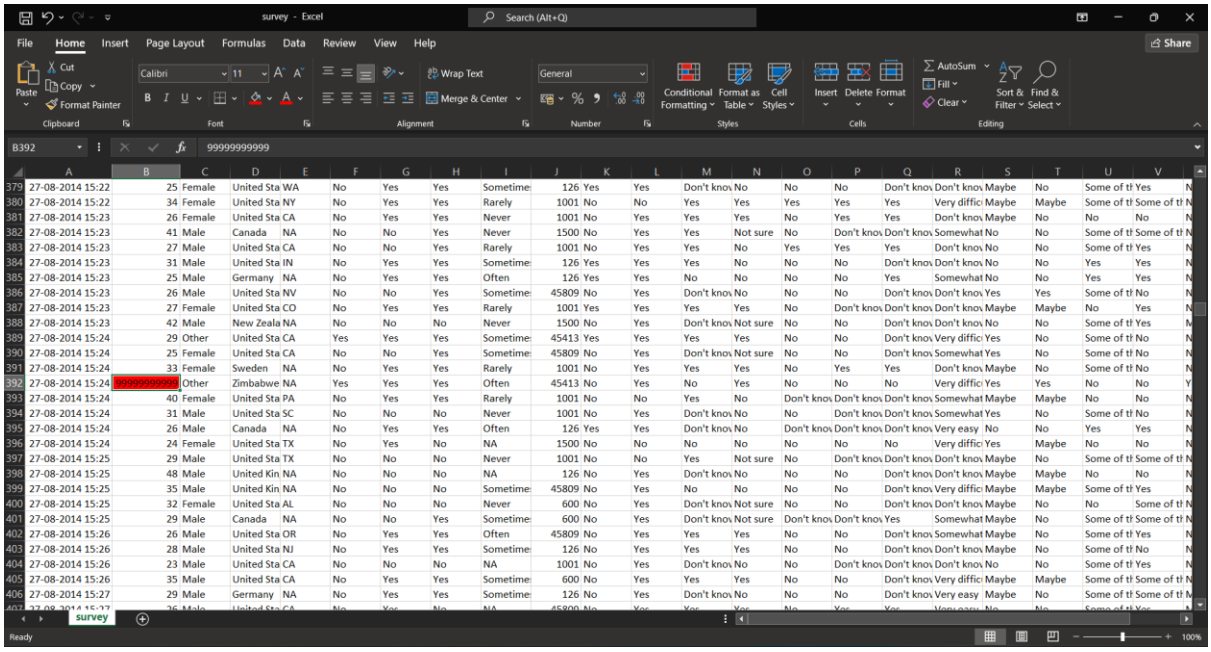

RStudio
# 1. Remove rows with missing values (if data loss is acceptable)
data_cleaned <- na.omit(data)
# 2. Replace missing values in numerical columns with the column mean
data <- data %>%
mutate(across(where(is.numeric), ~ ifelse(is.na(.), mean(., na.rm = TRUE), .)))
# 3. Replace missing values in categorical columns with the mode
replace_mode <- function(x) {
if (is.character(x) || is.factor(x)) {
mode_value <- names(which.max(table(x, useNA = "no")))
x[is.na(x)] <- mode_value
}
return(x)
}
data <- data %>%
mutate(across(where(is.character), replace_mode))
# Verify that missing values are handled
colSums(is.na(data))
family_history 0
treatment 0
work_interfere 0
no_employees 0
remote_work 0
tech_company 0
benefits 0
care_options 0
wellness_program 0
seek_help 0
anonymity 0
leave 0
mental_health_consequence 0
phys_health_consequence 0
coworkers 0
supervisor 0
mental_health_interview 0
phys_health_interview 0
mental_vs_physical 0
obs_consequence 0

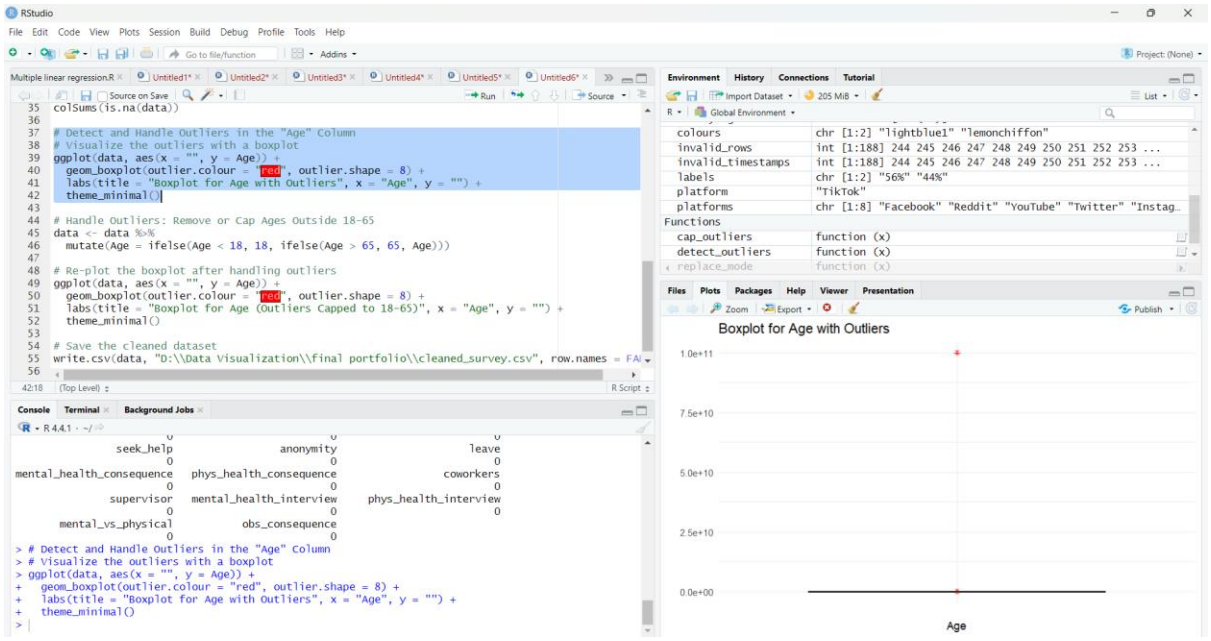

cleaned_survey - Excel
cleaned_survey

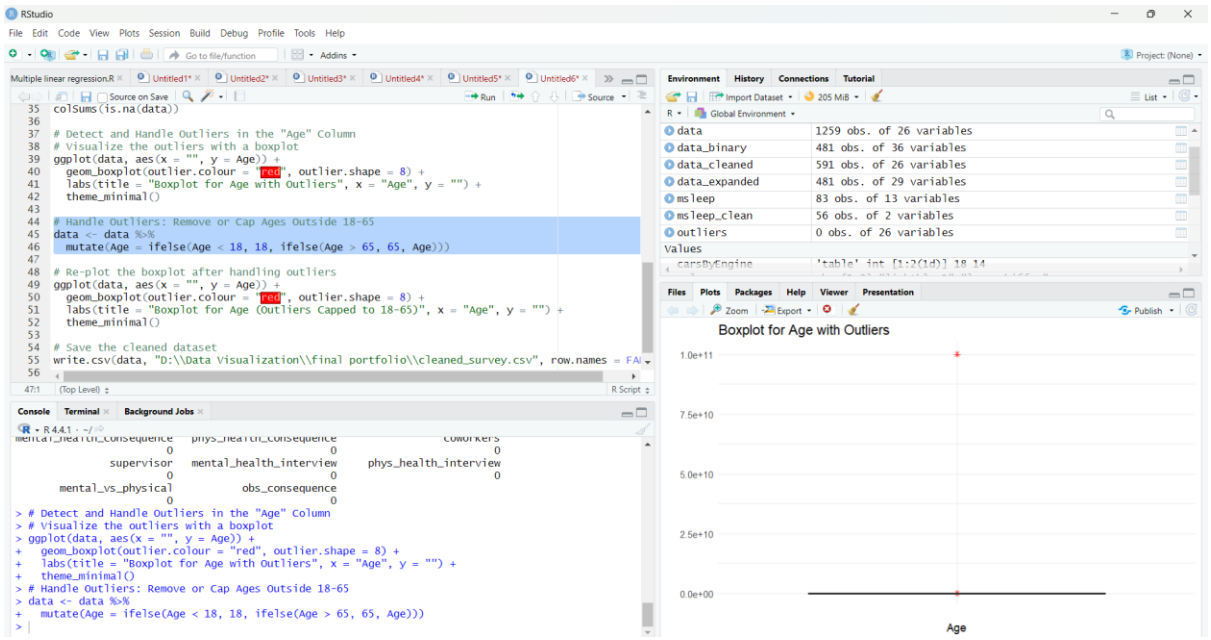
Handling outliers

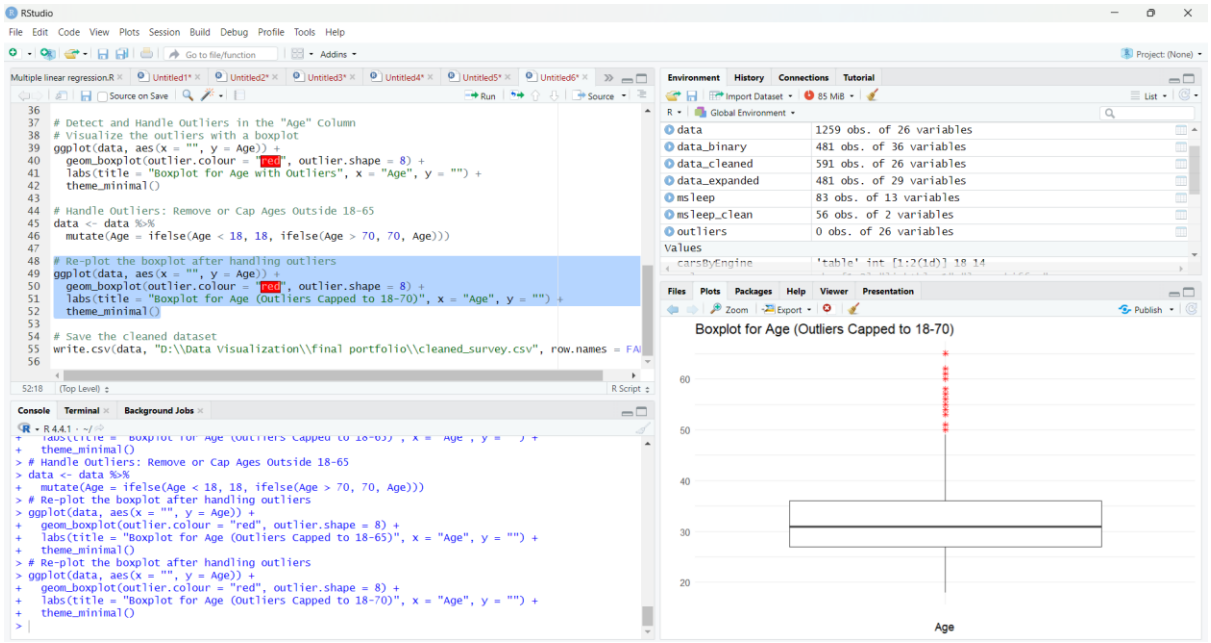
In the Age column of the dataset there are a couple of outliers as follows. There are extremely unrealistic values and even negative values. Other than that, since this is about a work environment ages like 5 are also invalid. Therefore, ages above 65 and below 18 are considered outliers and are removed.





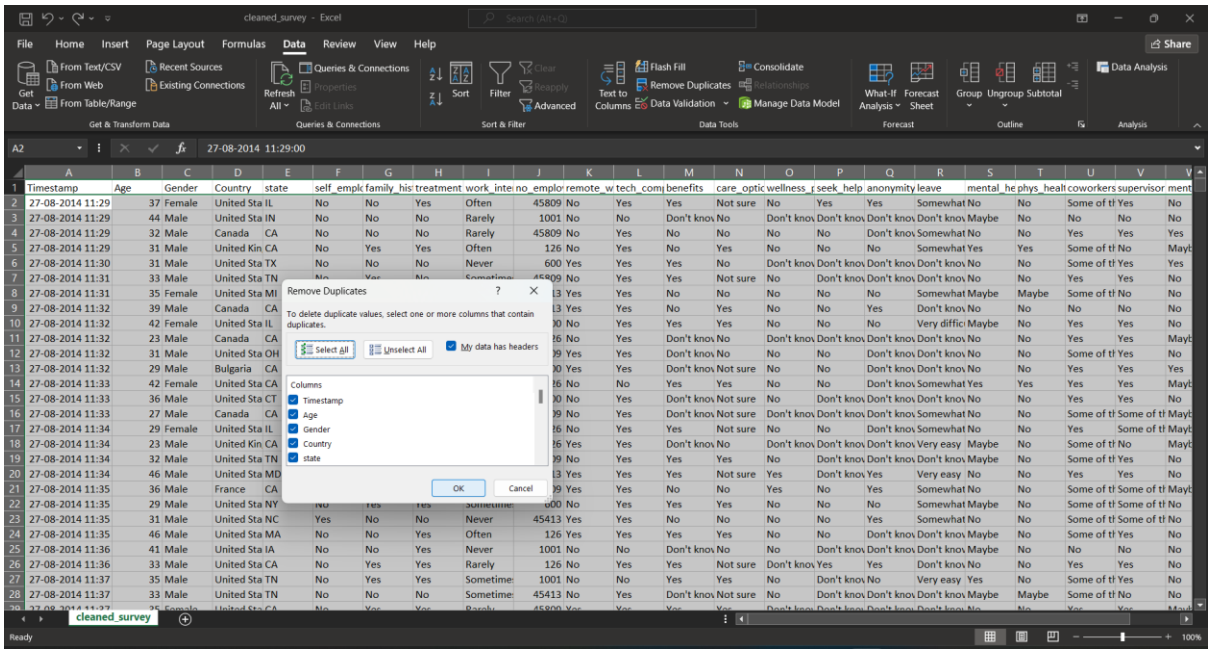

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
colSums(is.na(data))
# Detect and Handle Outliers in the "Age" Column
# Visualize the outliers with a boxplot
ggplot(data, aes(x = "", y = Age)) +
geom_boxplot(outlier.colour = "red", outlier.shape = 8) +
labs(title = "Boxplot for Age with Outliers", x = "Age", y = "") +
theme_minimal()
# Handle Outliers: Remove or Cap Ages Outside 18-65
data <- data %>%
mutate(Age = ifelse(Age < 18, 18, ifelse(Age > 65, 65, Age)))
# Re-plot the boxplot after handling outliers
ggplot(data, aes(x = "", y = Age)) +
geom_boxplot(outlier.colour = "red", outlier.shape = 8) +
labs(title = "Boxplot for Age (Outliers Capped to 18-65)", x = "Age", y = "") +
theme_minimal()
# Save the cleaned dataset
write.csv(data, "D:\\Data Visualization\\final portfolio\\cleaned_survey.csv", row.names = FA
Console Terminal Background Jobs
seek_help anonymity leave
0 0 0
mental_health_consequence phys_health_consequence coworkers
0 0 0
supervisor mental_health_interview phys_health_interview
0 0 0
mental_vs_physical obs_consequence
0 0
> # Detect and Handle Outliers in the "Age" Column
> # Visualize the outliers with a boxplot
> ggplot(data, aes(x = "", y = Age)) +
+ geom_boxplot(outlier.colour = "red", outlier.shape = 8) +
+ labs(title = "Boxplot for Age with Outliers", x = "Age", y = "") +
+ theme_minimal()
Environment History Connections Tutorial
colours chr [1:2] "lightblue1" "lemonchiffon"
invalid_rows int [1:188] 244 245 246 247 248 249 250 251 252 253 ...
invalid_timestamps int [1:188] 244 245 246 247 248 249 250 251 252 253 ...
labels chr [1:2] "56%" "44%"
platform "TikTok"
platforms chr [1:8] "Facebook" "Reddit" "YouTube" "Twitter" "Instag...
Functions
cap_outliers function (x)
detect_outliers function (x)
replace_mode function (x)
Files Plots Packages Help Viewer Presentation
Boxplot for Age with Outliers
1.0e+11
7.5e+10
5.0e+10
2.5e+10
0.0e+00
Age

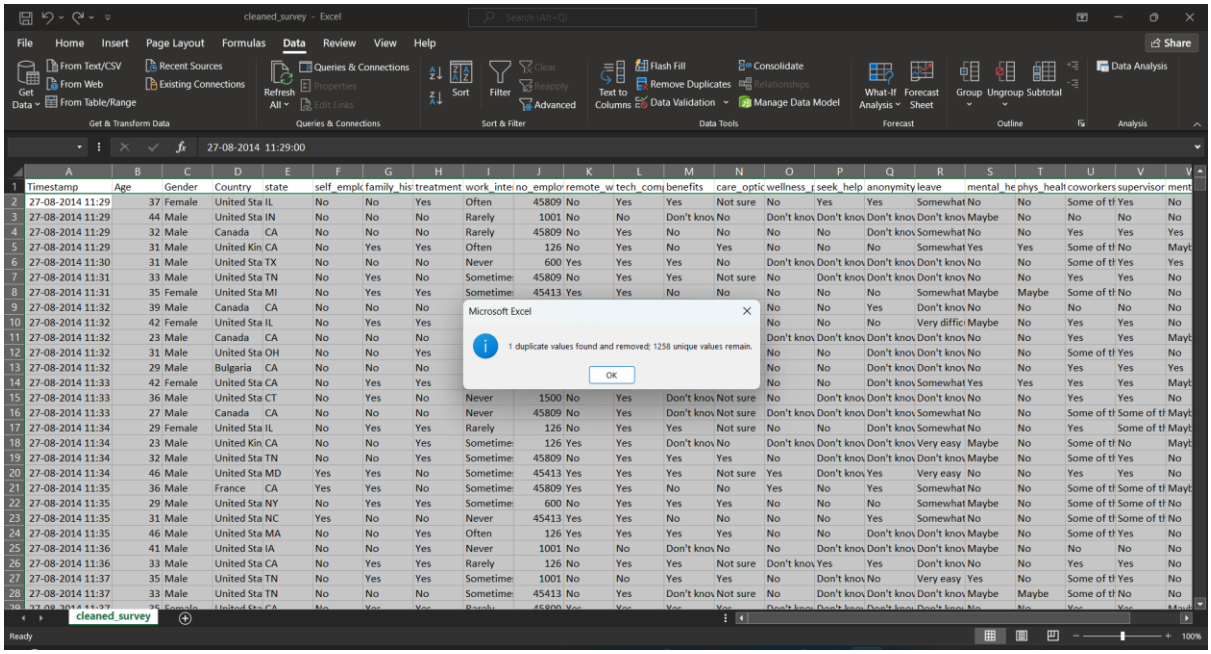

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
colSums(is.na(data))
# Detect and Handle Outliers in the "Age" Column
# Visualize the outliers with a boxplot
ggplot(data, aes(x = "", y = Age)) +
geom_boxplot(outlier.colour = "red", outlier.shape = 8) +
labs(title = "Boxplot for Age with Outliers", x = "Age", y = "") +
theme_minimal()
# Handle Outliers: Remove or Cap Ages Outside 18-65
data <- data %>%
mutate(Age = ifelse(Age < 18, 18, ifelse(Age > 65, 65, Age)))
# Re-plot the boxplot after handling outliers
ggplot(data, aes(x = "", y = Age)) +
geom_boxplot(outlier.colour = "red", outlier.shape = 8) +
labs(title = "Boxplot for Age (Outliers Capped to 18-65)", x = "Age", y = "") +
theme_minimal()
# Save the cleaned dataset
write.csv(data, "D:\\Data Visualization\\final portfolio\\cleaned_survey.csv", row.names = FA
Console Terminal Background Jobs
0 0 0
supervisor mental_health_interview phys_health_interview
0 0 0
mental_vs_physical obs_consequence
0 0
> # Detect and Handle Outliers in the "Age" Column
> # Visualize the outliers with a boxplot
> ggplot(data, aes(x = "", y = Age)) +
+ geom_boxplot(outlier.colour = "red", outlier.shape = 8) +
+ labs(title = "Boxplot for Age with Outliers", x = "Age", y = "") +
+ theme_minimal()
> # Handle Outliers: Remove or Cap Ages Outside 18-65
> data <- data %>%
+ mutate(Age = ifelse(Age < 18, 18, ifelse(Age > 65, 65, Age)))
Environment History Connections Tutorial
data 1259 obs. of 26 variables
data_binary 481 obs. of 36 variables
data_cleaned 591 obs. of 26 variables
data_expanded 481 obs. of 29 variables
msleep 83 obs. of 13 variables
msleep_clean 56 obs. of 2 variables
outliers 0 obs. of 26 variables
Values
carsByEngine 'table' int [1:2(1d)] 18 14
Files Plots Packages Help Viewer Presentation
Boxplot for Age with Outliers
1.0e+11
7.5e+10
5.0e+10
2.5e+10
0.0e+00
Age



5. Removing duplicates

Row number 861 and 862 are duplicated with exact same values. Therefore, one column is removed using 'Remove duplicates' option.



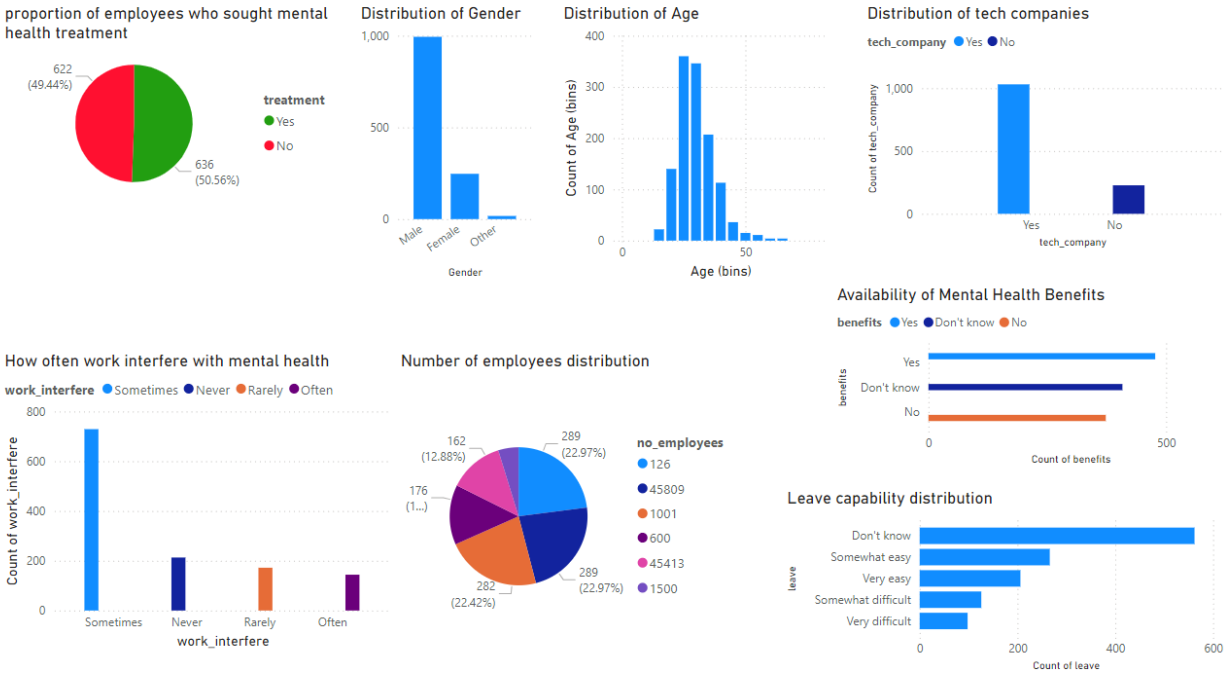

cleaned_survey - Excel
Microsoft Excel
1 duplicate values found and removed; 1258 unique values remain.
OK
cleaned_survey
Ready

Exploratory Data Analysis (EDA)

Univariate Analysis

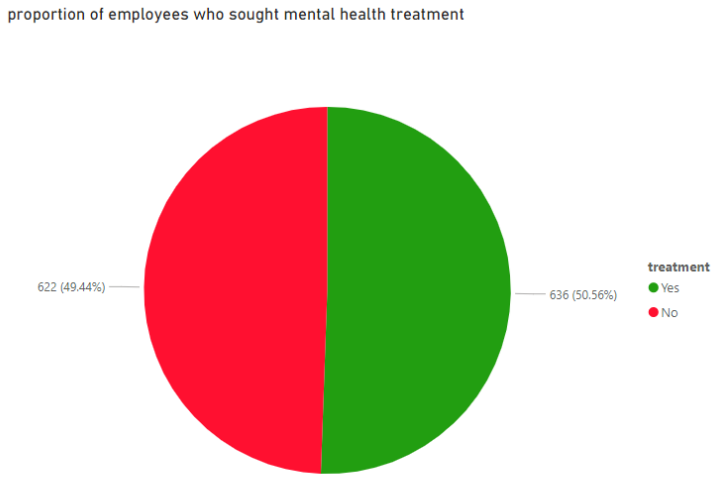
Univariate analysis examines the distribution of individual variables to understand their central tendency, dispersion and shape.

The variables in this dataset include categorical (e.g., Gender, Country, treatment, work_interfere) and numerical variables (eg., no_employees)



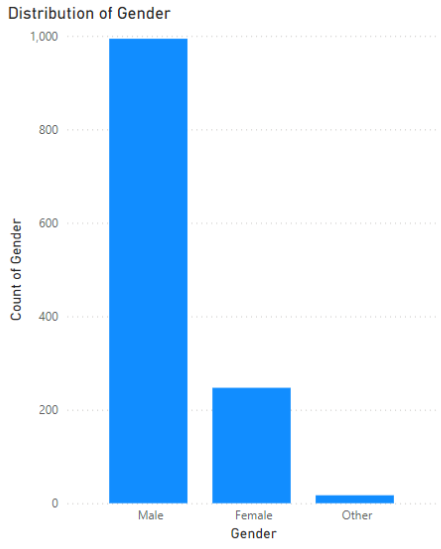
Insights of Univariate Analysis

I. Proportion of Employees who sought Mental Health Treatment (treatment)



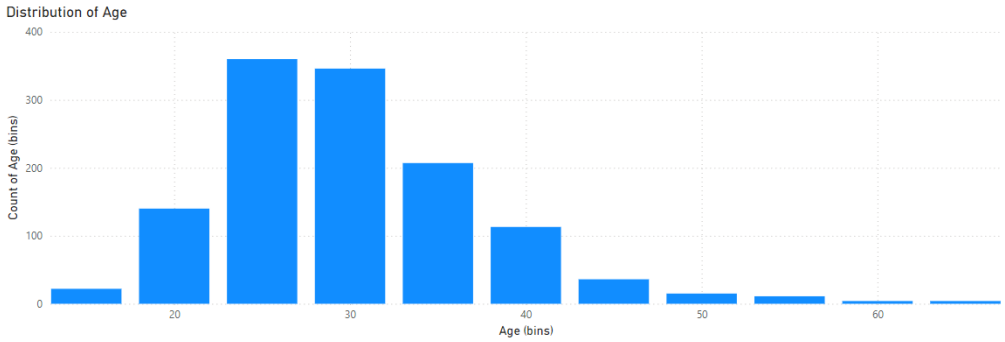
Insights : 49.44% of respondents sought mental health treatment while 50.65 did not. This indicates a fairly even split, highlighting that mental health issues are a significant concern among employees.

II. Distribution of Gender



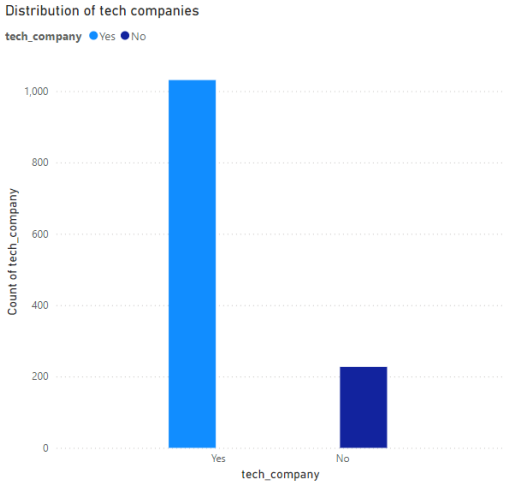
Insights : The majority of respondents identify as male. This distribution suggests a gender imbalance, which may influence how mental health issues are perceived in the workplace

III. Distribution of Age



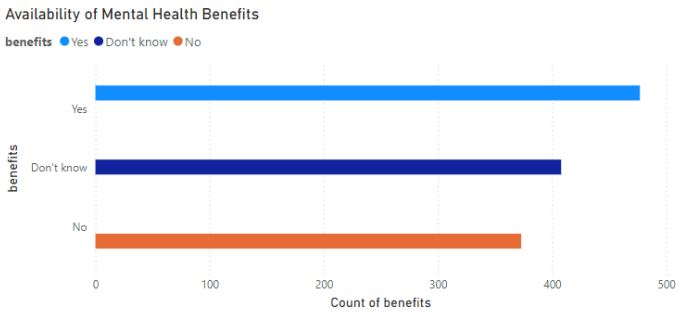
Insights : The average age of respondents is approximately 32 with most employees falling between the ages of 27 and 36 indicating that mental health issues are prevalent among younger professionals.

IV. Distribution of Tech Companies (tech_company)



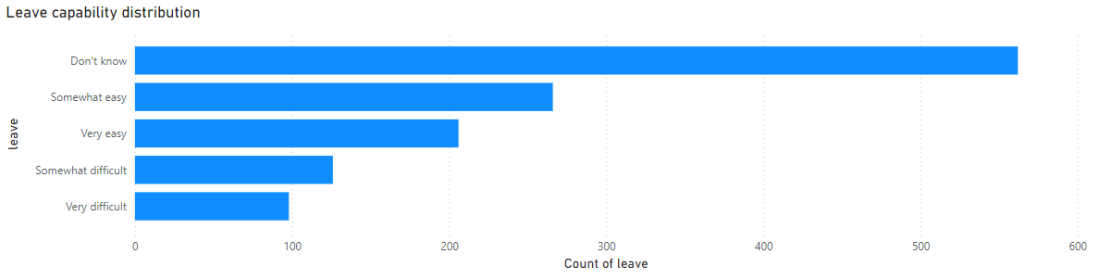
Insights : A significant portion of respondents work in tech companies aligning with the research focus.

V. Availability of Mental Health Benefits (benefits)



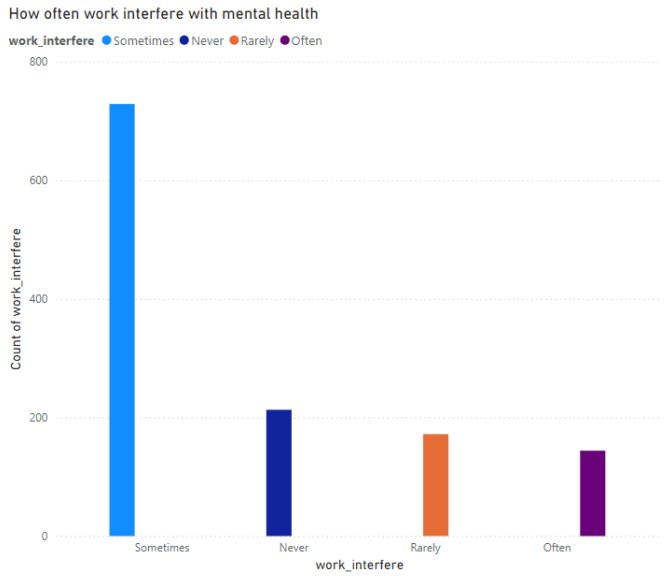
Insights : 47.7% respondents are provided with mental health benefits while a notable portion do not. This suggests that awareness and availability of mental health support could be improved.

VI. Leave Capability Distribution



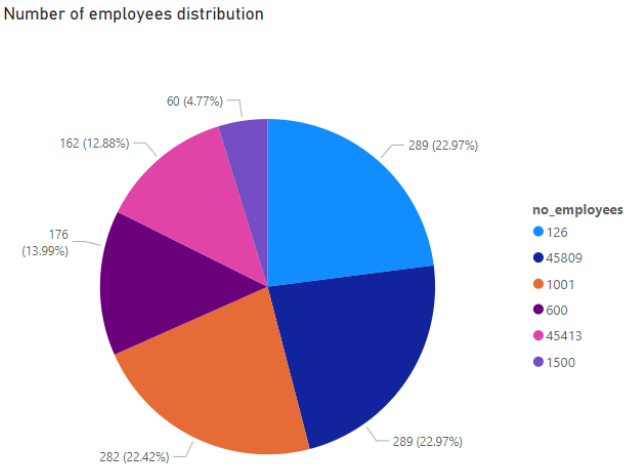
Insights : Many respondents indicated that taking leave for mental health reasons as " Don't know " and " Somewhat difficult". This points to a potential stigma around taking a leave.

VII. How often Work Interferes with Mental Health (work_interfere)



Insights : The majority reported that work "Sometimes" interfere with their mental health highlighting the need for workplace policies to be better accommodate mental well-being.

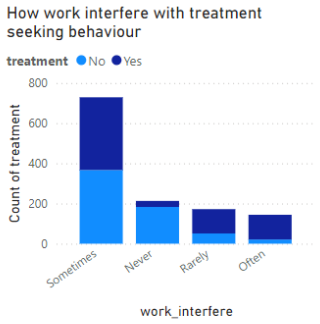
VIII. Number of employees Distribution (no_employees)

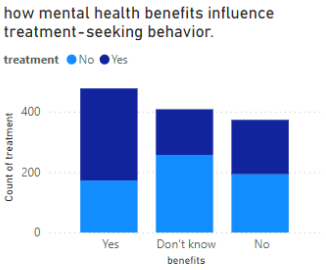


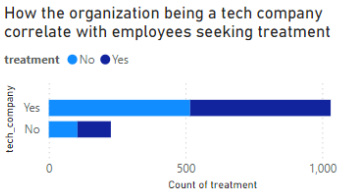
Insights : Respondents work across a range of company sizes with a significant proportion in smaller companies with around 120 to 1000 employees.

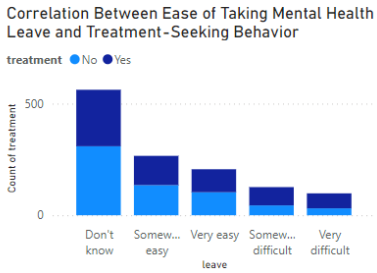
Bivariate Analysis

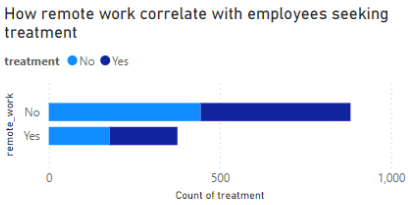
The relationship between variables like work_interfere, benefits, leave and remote_work with the dependant variable treatment are analyzed.



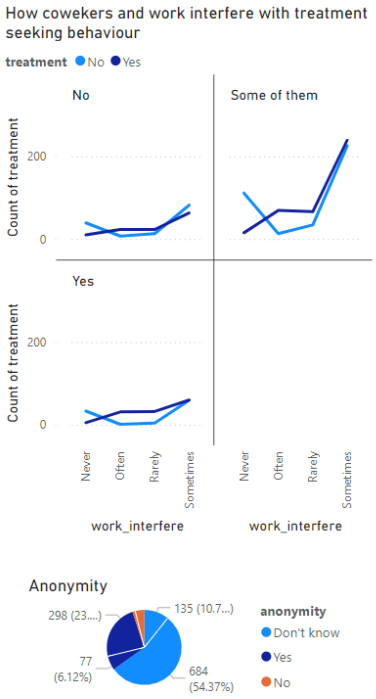


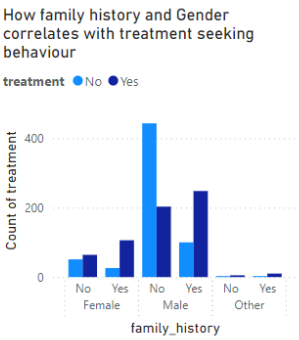


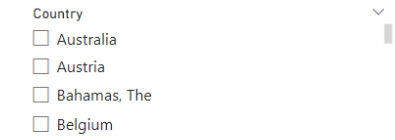


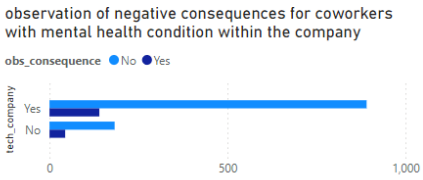


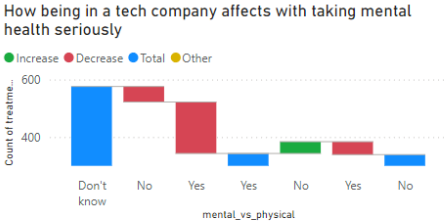
Multivariate Analysis

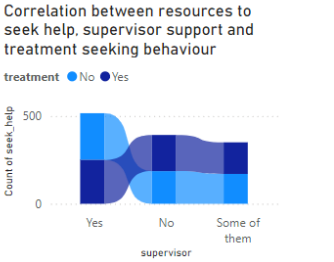


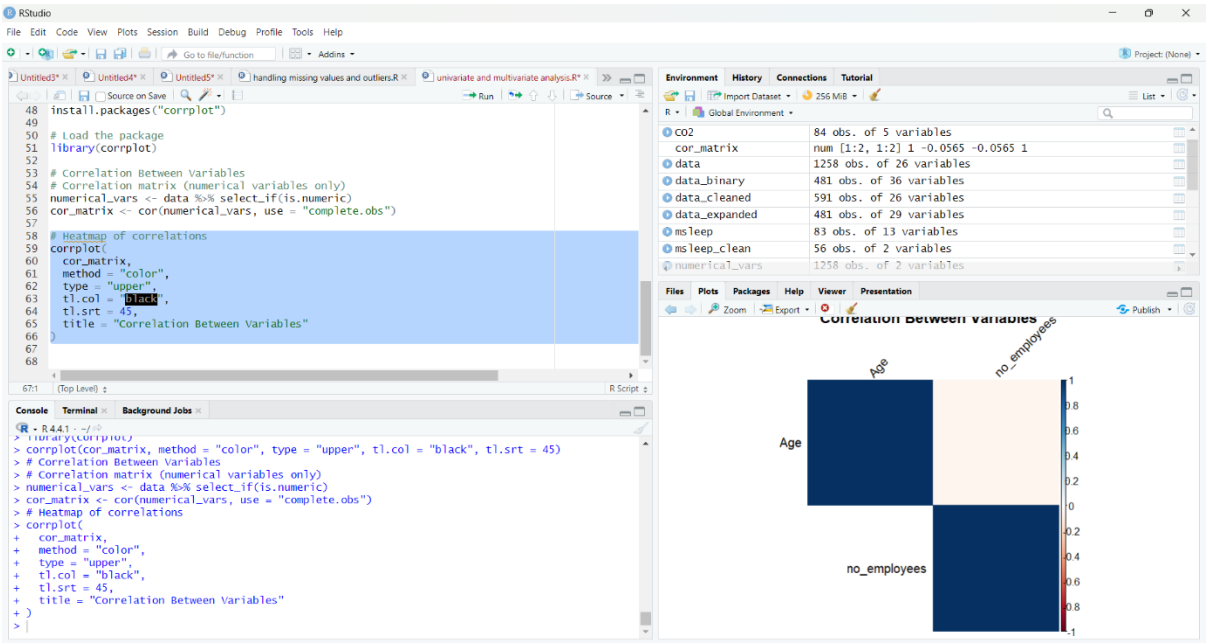






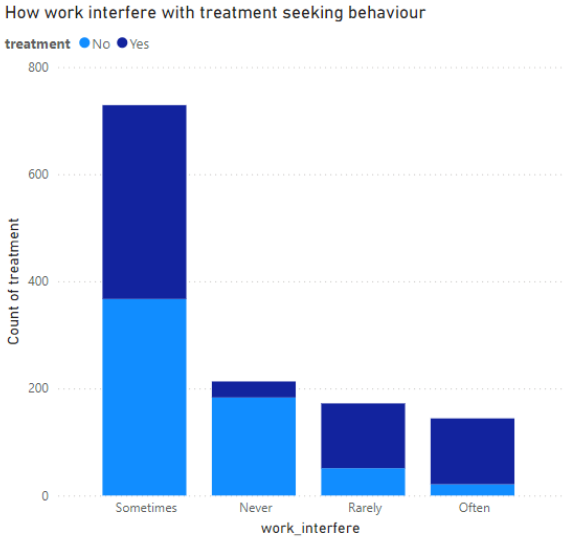






Insights of Bivariate and Multivariate Analysis

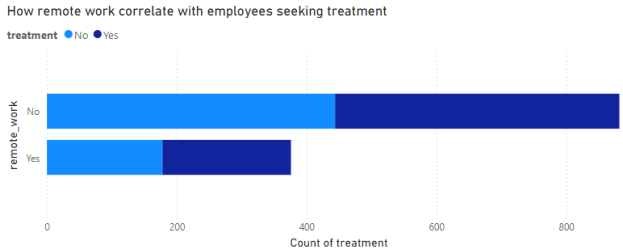
1. Work Interference vs Treatment-seeking behaviour



Employees experiencing higher level of mental health interference with work are most likely to seek treatment.

Insight : Work interference could be a significant stressor, pushing employees to seek help.

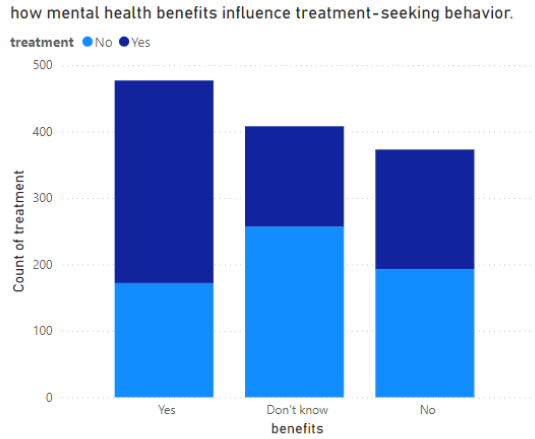
2. Remote work vs Treatment-seeking behaviour



Employees who do not work remotely tend to seek treatment more often.

Insight : Remote work may reduce workplace stressors potentially lowering treatment needs

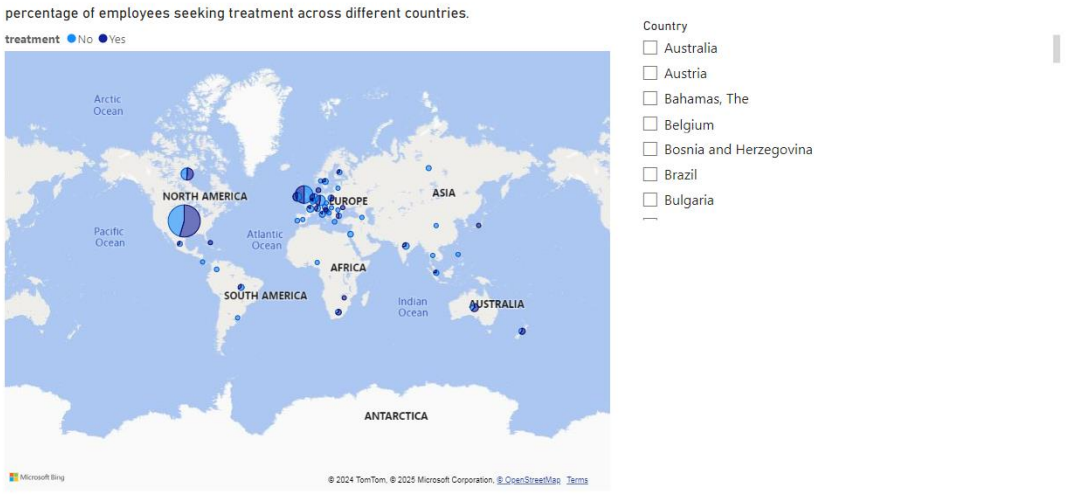
3. Mental Health benefits vs Treatment-seeking behaviour



Employees who know their company provides mental health benefits are more likely to seek treatment than those who are unsure or not know.

Insight : Clear communication and provision of mental health benefits promote a culture of seeking help when needed.

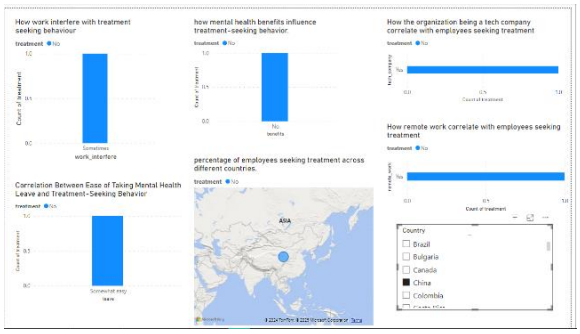
4. Country-wise Distribution

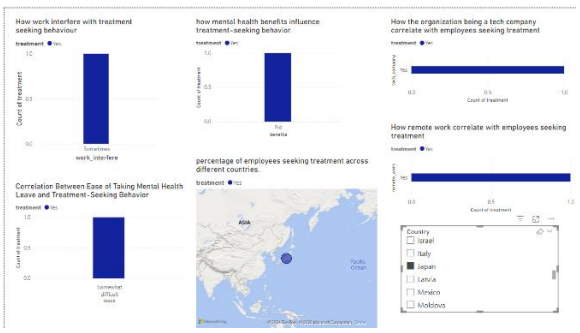


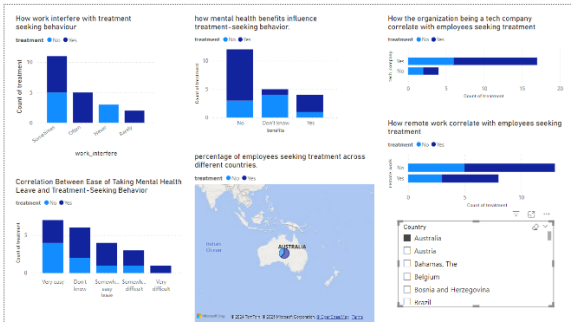
The map visualization highlights geographical variations in treatment-seeking behaviour. North America and Europe have higher concentration of employees seeking help.

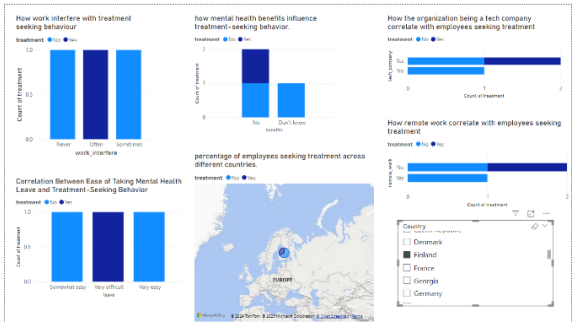
Insight : Cultural and regional attitudes towards mental health likely influence these trends.

Below are examples of how this differs from Country to Country.

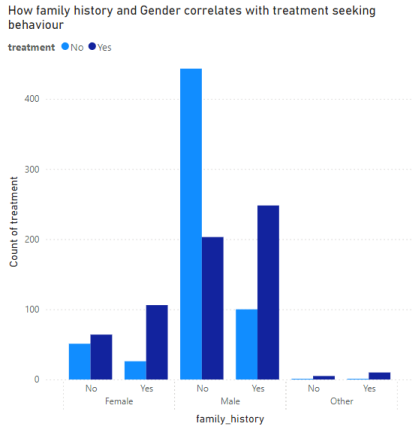








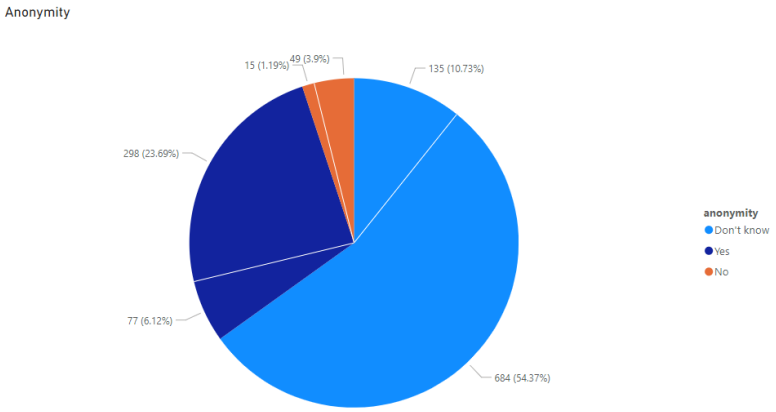
5. Family History vs Gender Correlation with treatment



Employees with a family history of mental health issues show a higher tendency to seek treatment

Insight : Genetic tendency and gender-specific openness to mental health issues effect treatment-seeking behaviour

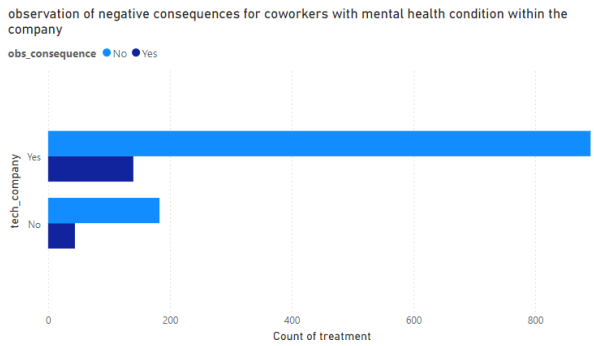
6. Anonymity Concerns



A significant proportion of employees prioritize anonymity when seeking treatment

Insight : Ensuring confidentiality could encourage more employees to seek mental health support.

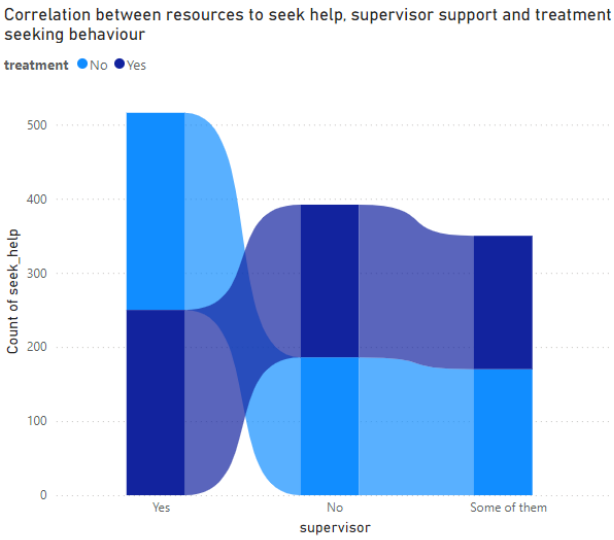
7. Organizational Culture and Consequences



Employees in tech companies and those observing negative consequences for co-workers with mental health issues are less likely to seek treatment.

Insight : Stigma and workplace culture in tech companies may discourage employees from accessing mental health support. Awareness may help reduce stigma.

8. Supervisor support and Resources



Employees who feel supported by supervisors or have access to resources are more likely to seek treatment.

Insight : Supervisor training and resource allocation play a vital role in addressing mental health needs.

Data Storytelling

The Silent Struggle of Emma : A Story of Workplace Mental Health

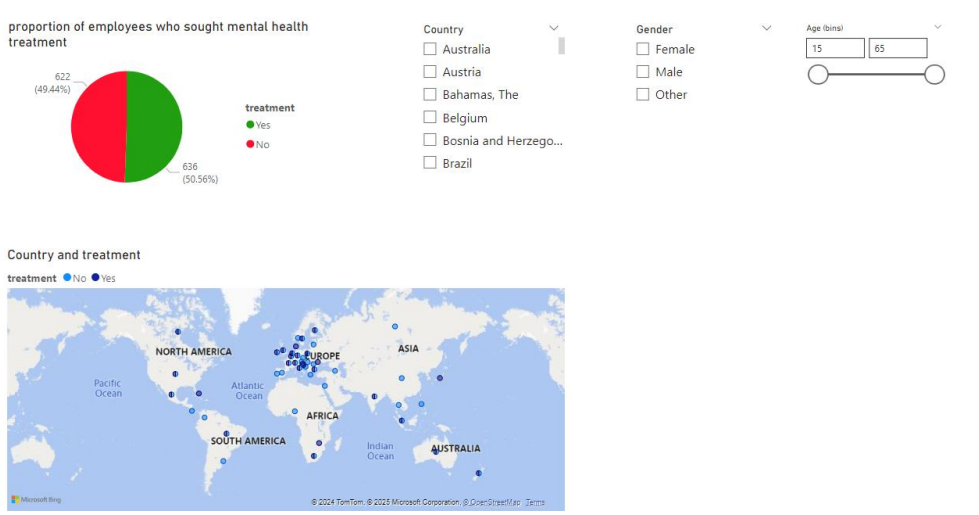
Emma, a 32 year-old marketing manager at a mid-sized tech company, was the kind of employee every organization hopes for. She was hardworking, innovative, and always the first to volunteer for new projects. On the surface, Emma seemed to have it all together – a stable career, a supportive family, and a bright future ahead. But beneath the surface, she was fighting an invisible battle that no one in her workplace knew about.

For months, Emma had been struggling with anxiety and overwhelming stress. Work had started to interfere with her mental health in ways she couldn't have imagined. She used to love the challenge of managing complex projects but now deadlines felt suffocating, meetings drained her energy, and even small tasks felt overwhelming. It didn't help that her company, though progressive in many ways, didn't openly discuss mental health. There was stigma around employees who admitted that they were struggling, and Emma feared being labelled as "weak" or "incapable".

Emma's struggle was a microcosm of a larger issue plaguing in workplaces worldwide. As businesses push for innovation and productivity, the invisible cost, employee mental health, often goes unnoticed. In Emma's case, her tech company's restless pace and lack of supportive policies left her feeling isolated and afraid to seek help. Emma's story is just one among many, but it sheds light on the challenges that employees face and the systemic change needed to create healthier workplaces

The Landscape of Mental Health in Workplace

To understand Emma's struggle, we turn to data. A global survey revealed that 49.44% of employees sought mental health treatment while 50.56% did not.



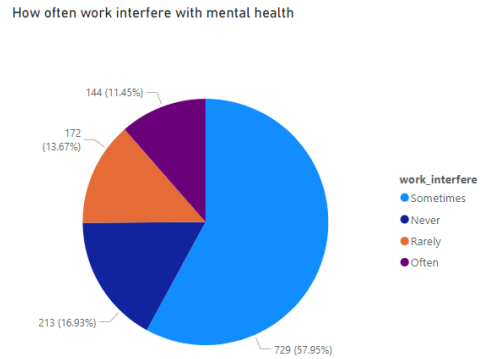
This near-even split underscores how widespread mental health challenges are among working professionals.

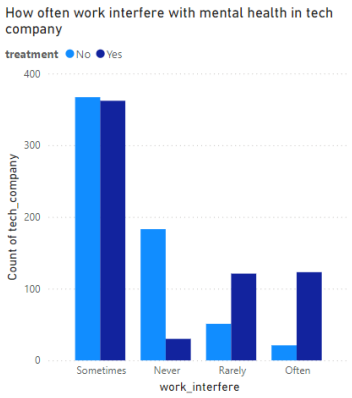
Emma's tech industry background isn't surprising as 82% of respondents in the survey worked in tech companies.

Tech's high-pressure culture often aggravates mental health issues, with employees reporting that work *"sometimes"* interfere with their well-being.

Work Interference and its Impact

For Emma, work interference was a daily struggle. She wasn't alone, 58% of respondents reported that work sometimes interferes with their mental health.

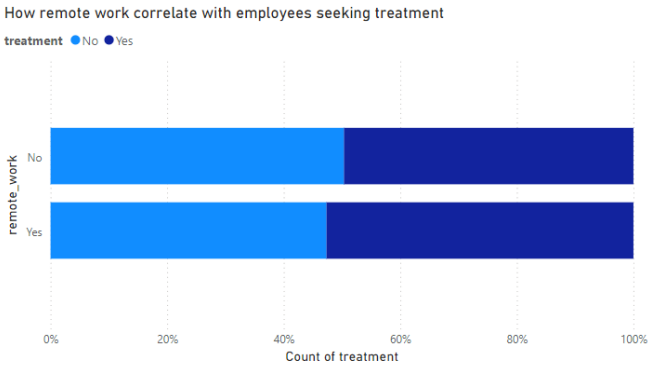




Data also revealed a significant insight : Employees who experience work interfering their mental health are most likely to seek treatment.

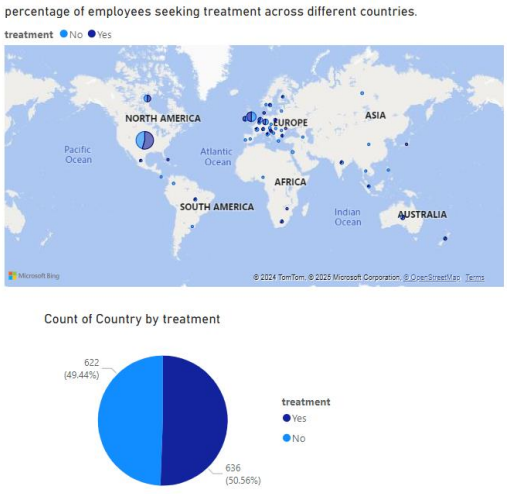
Companies must address the root causes of such interference, from unrealistic deadlines to insufficient support systems.

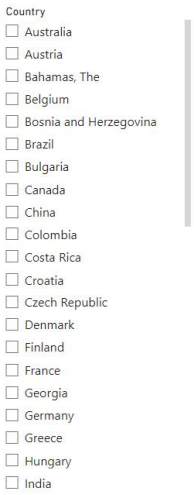
Emma often wondered if working remotely might have helped her manage stress. The data suggests that employees who do not work remotely are more likely to seek treatment.



Cultural and Regional Influence

Emma often found solace in reading about global workplace trends. She was fascinated to learn that mental health attitudes vary widely by region. In countries like United States and parts of Europe, employees are more likely to seek help, reflecting cultural openness towards mental health.

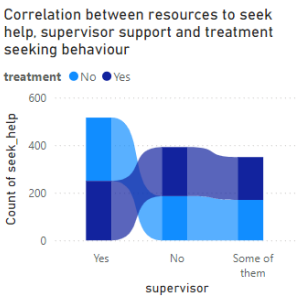


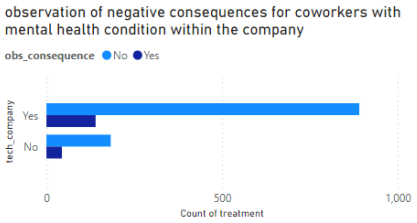


Breaking Stigma

One of Emma's biggest fears was the stigma associated with seeking help. The data revealed a troubling trend : employees in tech companies and those who observed negative consequences for seeking help.

These insights reinforced Emma's believes that the change has to start with leadership. Supervisors must create an environment where employees feel safe discussing mental health without fear of judgement.





The Path Forward

Emma's Journey reached a turning point when she finally reached out to a trusted co-worker. To her surprise, they shared similar struggles and pointed her towards resources outside the company. While it wasn't the ideal solution, it was a step in the right direction.

For companies to truly support employees like Emma, the data offers a clear guidance :

- **Ensure anonymity** : Many employees prioritize confidentiality when seeking help.
- **Provide accessible resources** : Employees with supervisor support and mental health resources are more likely to seek treatment.
- **Train managers** : Supervisors play a significant role in fostering a supportive culture

Emma's story is a call to action. It is a reminder that behind every statistic, is a real person. By listening to the data and implementing changes, workplaces can become spaces where employees thrive – not just professionally, but personally.

Why this research was important? (Gap analysis)

Despite significant strides in understanding workplace mental health, critical gaps remain. Existing researches highlights the importance of leadership, workplace benefits and stigma in shaping employees' mental health outcomes. For instance, *(Smith et al., 2022)* found out that organizations offering mental health support 25% increase in treatment-seeking behaviour. However, factors like work interference, flexibility policies and remote work remain unexplored. Additionally, small-to-medium enterprises and non-Western workplaces are underrepresented, limiting the scope of current findings. Most studies also focus broadly on mental health resources without examining specific triggers for treatment-seeking behaviour.

This research addresses these gaps by leveraging a comprehensive global dataset, particularly from the tech industry, to analyse treatment-seeking behaviour. It examines underexplored factors like work interference, company size, remote work policies while considering diverse cultural contexts. By combining quantitative methods with actionable insights, this study contributes to a deeper understanding of how workplace culture influence mental health treatment, surfacing the way for inclusive, evidence-based interventions.

References

Open Sourcing Mental Illness, LTD (2014). *Mental Health in Tech Survey*. [online] Kaggle.com. Available at: https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey [Accessed 30 Dec. 2024].

Harvey, P.D., Depp, C.A., Rizzo, A.A., Strauss, G.P., Spelber, D., Carpenter, L.L., Kalin, N.H., Krystal, J.H., McDonald, W.M., Nemeroff, C.B., Rodriguez, C.I., Widge, A.S. and Torous, J. (2022). Technology and Mental Health: State of the Art for Assessment and Treatment. *American Journal of Psychiatry*, [online] 179(12), pp.897–914. doi:https://doi.org/10.1176/appi.ajp.21121254.

Nyblade, L., Stockton, M.A., Giger, K., Bond, V., Ekstrand, M.L., Lean, R.M., Ellen, Nelson, R.E., Sapag, J.C., Taweesap Siraprapasiri, Turan, J. and Wouters, E. (2019). Stigma in health facilities: why it matters and how we can change it. *BMC Medicine*, [online] 17(1). doi:https://doi.org/10.1186/s12916-019-1256-2.

Ackerman, N. (2024). *How to support employee wellbeing at work*. [online] Thetimes.com. Available at: https://www.thetimes.com/business-money/entrepreneurs/article/how-to-support-employee-wellbeing-at-work-enterprise-network-p2kf29566? [Accessed 5 Jan. 2025].

Smith, J.M., Smith, J., McLuckie, A., Szeto, A.C.H., Choate, P., Birks, L.K., Burns, V.F. and Bright, K.S. (2022). Exploring Mental Health and Well-Being Among University Faculty Members: A Qualitative Study. *Journal of Psychosocial Nursing and Mental Health Services*, 60(11), pp.17–25. doi:https://doi.org/10.3928/02793695-20220523-01.

Vonderlin, R., Schmidt, B., Müller, G., Biermann, M., Kleindienst, N., Bohus, M. and Lyssenko, L. (2021). Health-Oriented Leadership and Mental Health From Supervisor and Employee Perspectives: A Multilevel and Multisource Approach. *Frontiers in Psychology*, [online] 11. doi:https://doi.org/10.3389/fpsyg.2020.614803.

Elraz, H. (2018). Identity, mental health and work: How employees with mental health conditions recount stigma and the pejorative discourse of mental illness. *Human Relations*, [online] 71(5), pp.722–741. doi:https://doi.org/10.1177/0018726717716752.

Wu, A., Roemer, E.C., Kent, K.B., Ballard, D.W. and Goetzel, R.Z. (2021). Organizational Best Practices Supporting Mental Health in the Workplace. *Journal of Occupational and Environmental Medicine*, [online] 63(12), pp.e925–e931. doi:https://doi.org/10.1097/jom.0000000000002407.

Song, Z. and Baicker, K. (2019). Effect of a Workplace Wellness Program on Employee Health and Economic Outcomes. *JAMA*, [online] 321(15), pp.1491–1491. doi:https://doi.org/10.1001/jama.2019.3307.