# Analysis of Employee Mental Wellness and Workplace Support Systems

## Project Goal:

The goal of this project is to analyze the factors influencing employee mental wellness and evaluate the effectiveness of workplace support systems. By identifying key trends and areas of improvement, we aim to provide actionable insights for organizations to foster a healthier work environment.

### Why This Analysis Matters:

Employee mental wellness is critical for organizational productivity, employee satisfaction, and overall workplace harmony. Understanding how workplace support systems impact mental health helps organizations implement strategies to enhance employee well-being.

### Dataset Information

This dataset is from a 2014 survey that measures attitudes towards mental health and frequency of mental health disorders in the tech workplace.

***MOUNTING DRIVE***

In [ ]:

```
from google.colab import drive
drive.mount('/content/drive', force_remount=True)
```

```
Mounted at /content/drive
```

***IMPORTING NECESSARY LIBRARIES***

In [ ]:

```
import warnings
warnings.filterwarnings('ignore')

#basic libraries
import numpy as np
import pandas as pd

#visualization libraries
import missingno as msno
import matplotlib
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
import plotly.graph_objects as go
%matplotlib inline
```

***IMPORTING SURVEY.CSV DATA FROM DRIVE AND LOAD THE DATA***

In [ ]:

```
# Load the dataset
data = pd.read_csv('/content/drive/My Drive/Analysis of Employee Mental Wellness and Workplace Support Systems/survey.csv')

# Preview the data
print(data.head())
```

```
             Timestamp  Age  Gender         Country state self_employed  \
0  2014-08-27 11:29:31   37  Female   United States    IL           NaN   
1  2014-08-27 11:29:37   44       M   United States    IN           NaN   
2  2014-08-27 11:29:44   32    Male          Canada   NaN           NaN   
3  2014-08-27 11:29:46   31    Male  United Kingdom   NaN           NaN   
4  2014-08-27 11:30:22   31    Male   United States    TX           NaN   

  family_history treatment work_interfere    no_employees  ...  \
0             No       Yes          Often            6-25  ...   
1             No        No         Rarely  More than 1000  ...   
2             No        No         Rarely            6-25  ...   
3            Yes       Yes          Often          26-100  ...   
4             No        No          Never         100-500  ...   

                leave mental_health_consequence phys_health_consequence  \
0       Somewhat easy                        No                      No   
1          Don't know                     Maybe                      No   
2  Somewhat difficult                        No                      No   
3  Somewhat difficult                       Yes                     Yes   
4          Don't know                        No                      No   

      coworkers supervisor mental_health_interview phys_health_interview  \
0  Some of them        Yes                      No                 Maybe   
1            No         No                      No                    No   
2           Yes        Yes                     Yes                   Yes   
3  Some of them         No                   Maybe                 Maybe   
4  Some of them        Yes                     Yes                   Yes   

  mental_vs_physical obs_consequence comments  
0                Yes              No      NaN  
1         Don't know              No      NaN  
2                 No              No      NaN  
3                 No             Yes      NaN  
4         Don't know              No      NaN  

[5 rows x 27 columns]
```

In [ ]:

```
# Assuming 'data' is your DataFrame
print(data.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1259 entries, 0 to 1258
Data columns (total 27 columns):
 #   Column                     Non-Null Count  Dtype 
---  ------                     --------------  ----- 
 0   Timestamp                  1259 non-null   object
 1   Age                        1259 non-null   int64 
 2   Gender                     1259 non-null   object
 3   Country                    1259 non-null   object
 4   state                      744 non-null    object
 5   self_employed              1241 non-null   object
 6   family_history             1259 non-null   object
 7   treatment                  1259 non-null   object
 8   work_interfere             995 non-null    object
 9   no_employees               1259 non-null   object
 10  remote_work                1259 non-null   object
 11  tech_company               1259 non-null   object
 12  benefits                   1259 non-null   object
 13  care_options               1259 non-null   object
 14  wellness_program           1259 non-null   object
 15  seek_help                  1259 non-null   object
 16  anonymity                  1259 non-null   object
 17  leave                      1259 non-null   object
 18  mental_health_consequence  1259 non-null   object
 19  phys_health_consequence    1259 non-null   object
 20  coworkers                  1259 non-null   object
 21  supervisor                 1259 non-null   object
 22  mental_health_interview    1259 non-null   object
 23  phys_health_interview      1259 non-null   object
 24  mental_vs_physical         1259 non-null   object
 25  obs_consequence            1259 non-null   object
 26  comments                   164 non-null    object
dtypes: int64(1), object(26)
memory usage: 265.7+ KB
None
```

In [ ]:

```
# Assuming 'data' is your DataFrame
print(data.columns)
```

```
Index(['Timestamp', 'Age', '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_health_consequence', 'phys_health_consequence', 'coworkers',
       'supervisor', 'mental_health_interview', 'phys_health_interview',
       'mental_vs_physical', 'obs_consequence', 'comments'],
      dtype='object')
```

## This code cleans and normalizes the 'Gender' column by:

1. Defining a function (`clean_gender`) to categorize gender into 'Male', 'Female', or 'Non-binary/Other' while handling inconsistencies.
2. Applying the function to create a new cleaned column ('Gender_Cleaned').
3. Displaying unique cleaned values for verification.
4. Saving the updated dataset with the cleaned column to a new Excel file.

In [ ]:

```
import pandas as pd
import numpy as np

# Define a mapping function to clean the 'Gender' column
def clean_gender(gender):
    if pd.isna(gender):
        return np.nan
    gender = gender.strip().lower()  # Normalize case and trim spaces
    if gender in ['male', 'm', 'man', 'male-ish', 'maile', 'cis male', 'male (cis)', 'msle', 'malr', 'mal', 'mail', 'make', 'g
        return 'Male'
    elif gender in ['female', 'f', 'woman', 'cis female', 'femake', 'female (cis)', 'female (trans)', 'trans-female', 'femail'
        return 'Female'
    elif gender in ['non-binary', 'genderqueer', 'androgyne', 'enby', 'fluid', 'agender', 'queer', 'queer/she/they', 'male lea
        return 'Non-binary/Other'
    else:
        return np.nan  # For unclear or invalid entries

# Apply the cleaning function to the 'Gender' column
data['Gender_Cleaned'] = data['Gender'].apply(clean_gender)

# Display the unique cleaned values
print(data['Gender_Cleaned'].unique())

# Save the cleaned dataset to a new Excel file
data.to_excel('/content/drive/My Drive/Analysis of Employee Mental Wellness and Workplace Support Systems/survey.xlsx', index=
```

```
['Female' 'Male' 'Non-binary/Other' nan]
```

## This code cleans the 'Age' column by:

1. Applying a condition to filter out unrealistic values (keeping ages between 15 and 70) and setting others to `NaN`.
2. Creating a new column ('Age_Cleaned') to store the cleaned age values.
3. Displaying the original and cleaned age values for verification.
4. Saving the updated dataset with the cleaned column to a new Excel file.

In [ ]:

```
# Clean the Age column by applying the condition
data['Age_Cleaned'] = data['Age'].apply(lambda x: x if 15 < x < 70 else np.nan)

# Display the cleaned Age column
print(data[['Age', 'Age_Cleaned']].head())

# Save the cleaned dataset to a new Excel file
data.to_excel('/content/drive/My Drive/Analysis of Employee Mental Wellness and Workplace Support Systems/survey.xlsx', index=
```

```
   Age  Age_Cleaned
0   37         37.0
1   44         44.0
2   32         32.0
3   31         31.0
4   31         31.0
```

## This code removes unnecessary columns from the dataset by:

1. Dropping the 'Timestamp', 'state', and 'comments' columns using the `drop()` method.
2. Displaying the updated dataset to ensure the columns have been removed.
3. Saving the cleaned dataset to the existing Excel file.

In [ ]:

```
# Drop the specified columns from the dataset
data = data.drop(columns=['Timestamp', 'state', 'comments'])

# Display the updated DataFrame to verify
print(data.head())

# Save the cleaned dataset to the existing Excel file
data.to_excel('/content/drive/My Drive/Analysis of Employee Mental Wellness and Workplace Support Systems/survey.xlsx', index=
   Age  Gender         Country self_employed family_history treatment  \
0   37  Female   United States           NaN             No       Yes   
1   44       M   United States           NaN             No        No   
2   32    Male          Canada           NaN             No        No   
3   31    Male  United Kingdom           NaN            Yes       Yes   
4   31    Male   United States           NaN             No        No   

  work_interfere    no_employees remote_work tech_company  ...  \
0          Often            6-25          No          Yes  ...   
1         Rarely  More than 1000          No           No  ...   
2         Rarely            6-25          No          Yes  ...   
3          Often          26-100          No          Yes  ...   
4          Never         100-500         Yes          Yes  ...   

  phys_health_consequence     coworkers supervisor mental_health_interview  \
0                      No  Some of them        Yes                      No   
1                      No            No         No                      No   
2                      No           Yes        Yes                     Yes   
3                     Yes  Some of them         No                   Maybe   
4                      No  Some of them        Yes                     Yes   

  phys_health_interview mental_vs_physical obs_consequence Gender_Cleaned  \
0                 Maybe                Yes              No         Female   
1                    No         Don't know              No           Male   
2                   Yes                 No              No           Male   
3                 Maybe                 No             Yes           Male   
4                   Yes         Don't know              No           Male   

  Age_Cleaned Age_Group  
0        37.0     36-45  
1        44.0     36-45  
2        32.0     26-35  
3        31.0     26-35  
4        31.0     26-35  

[5 rows x 27 columns]
```

## Description of Gender and Age Group Distribution Graphs

The following graphs provide insights into the distribution of participants based on Gender and Age Group.

**Gender Representation (Pie Chart):** This pie chart visualizes the proportion of individuals across different genders in the dataset. The chart divides participants into three categories: Male, Female, and Other. The color-coded segments represent the gender distribution, with Male shown in skyblue, Female in pink, and Other in green. The percentage labels on the chart allow for quick insight into the gender makeup of the population.

**Age Group Distribution (Pie Chart):** The second pie chart illustrates the distribution of participants across various age groups. The groups are divided into 18-25, 26-35, 36-45, and 46+. Each age group is color-coded: 18-25 is displayed in light blue, 26-35 in orange, 36-45 in green, and 46+ in red. The chart offers a clear overview of the age demographics within the dataset, highlighting which age group constitutes the largest portion of participants.

Both charts are presented side by side for comparative analysis, allowing viewers to easily observe the relationship between gender and age group in the dataset.

In [ ]:

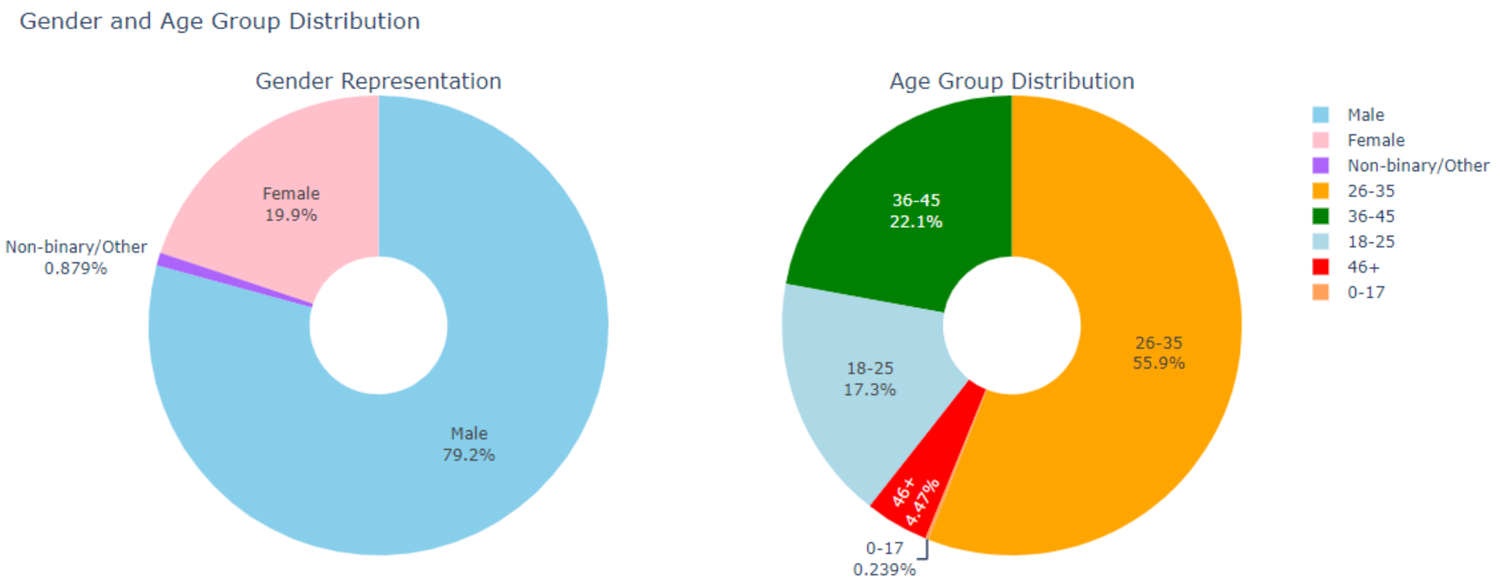
```
import pandas as pd

# Define bins and labels for Age Groups
bins = [0, 17, 25, 35, 45, 100]  # Age ranges (0-17, 18-25, 26-35, 36-45, 46+)
labels = ['0-17', '18-25', '26-35', '36-45', '46+']

# Create the 'Age_Group' column by binning the 'Age' column
data['Age_Group'] = pd.cut(data['Age'], bins=bins, labels=labels, right=True)

# Verify the new column
print(data[['Age', 'Age_Group']].head())

# Save the cleaned dataset to the existing Excel file
data.to_excel('/content/drive/My Drive/Analysis of Employee Mental Wellness and Workplace Support Systems/survey.xlsx', index=
```

```
   Age Age_Group
0   37     36-45
1   44     36-45
2   32     26-35
3   31     26-35
4   31     26-35
```

In [ ]:

```
import plotly.graph_objects as go
import plotly.express as px
from plotly.subplots import make_subplots

# Assuming 'data' is your main DataFrame
# Count Gender distribution
by_gen = data['Gender_Cleaned'].value_counts().reset_index(name='count')
by_gen.columns = ['Gender_Cleaned', 'count']
```

```
# Group the data by Age Group to get counts
by_age = data['Age_Group'].value_counts().reset_index(name='count')
by_age.columns = ['Age_Group', 'count']  # Renaming columns for better readability

# Create the first pie chart for Gender distribution
gender_fig = px.pie(by_gen, names='Gender_Cleaned', values='count', title='Gender Representation',
                     color='Gender_Cleaned', color_discrete_map={'Male': 'skyblue', 'Female': 'pink', 'Other': 'green'})
gender_fig.update_traces(textinfo='percent+label', hole=0.3)

# Create the second pie chart for Age group distribution
age_fig = px.pie(by_age, names='Age_Group', values='count', title='Age Group Distribution',
                  color='Age_Group', color_discrete_map={'18-25': 'lightblue', '26-35': 'orange', '36-45': 'green', '46+': 're
age_fig.update_traces(textinfo='percent+label', hole=0.3)

# Create subplots with two pie charts side by side
fig = make_subplots(rows=1, cols=2, subplot_titles=('Gender Representation', 'Age Group Distribution'),
                    specs=[[{'type': 'pie'}, {'type': 'pie'}]])

# Add the gender pie chart to the first subplot
fig.add_trace(gender_fig.data[0], row=1, col=1)

# Add the age group pie chart to the second subplot
fig.add_trace(age_fig.data[0], row=1, col=2)

# Update the layout to improve the spacing
fig.update_layout(title_text='Gender and Age Group Distribution', showlegend=True)

# Show the final combined figure
fig.show()
```

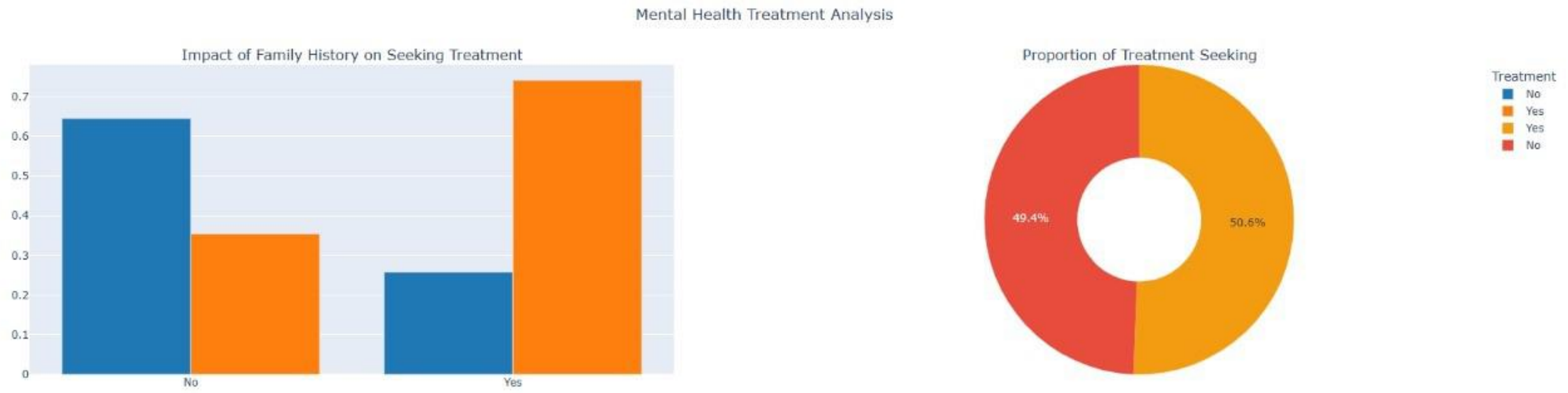


## EXPLORATORY DATA ANALYSIS

Family History and Treatment Proportion This analysis explores the relationship between family history and the likelihood of seeking mental health treatment, as well as the overall proportion of individuals seeking treatment. The data is visualized through two distinct charts: a bar chart that highlights the impact of family history on treatment decisions, and a pie chart illustrating the distribution of different treatment types chosen by individuals.

In [ ]:

```
from plotly.subplots import make_subplots
import plotly.graph_objects as go

# Calculate data for the first bar chart (Family History Impact)
family_treatment = data.groupby('family_history')['treatment'].value_counts(normalize=True).unstack().fillna(0)

# Calculate data for the second pie chart (Proportion of Treatment)
treatment_proportion = data['treatment'].value_counts(normalize=True) * 100

# Create a subplot layout with 1 row and 2 columns
fig = make_subplots(
    rows=1, cols=2,
    specs=[[{'type': 'bar'}, {'type': 'domain'}]],  # 'domain' is used for pie charts
    subplot_titles=("Impact of Family History on Seeking Treatment", "Proportion of Treatment Seeking")
)

# Color scheme for the bar chart
bar_colors = ['#1f77b4', '#ff7f0e', '#2ca02c', '#d62728', '#9467bd']  # Customize as needed

# Add bar chart for Family History Impact
for i, treatment in enumerate(family_treatment.columns):
    fig.add_trace(
        go.Bar(
            x=family_treatment.index,
            y=family_treatment[treatment],
            name=treatment,
            marker=dict(color=bar_colors[i % len(bar_colors)])  # Apply the color scheme
        ),
        row=1, col=1
    )

# Color scheme for the pie chart
pie_colors = ['#f39c12', '#e74c3c', '#3498db', '#2ecc71', '#9b59b6']  # Customize as needed

# Add pie chart for Proportion of Treatment
fig.add_trace(
    go.Pie(
        labels=treatment_proportion.index,
        values=treatment_proportion,
        hole=0.4,
        name="Treatment Proportion",
        marker=dict(colors=pie_colors)  # Apply the color scheme
    ),
    row=1, col=2
)

# Update layout for better aesthetics and interactivity
fig.update_layout(
    title="Mental Health Treatment Analysis",
    title_x=0.5,  # Center align the title
    barmode='group',  # Group bars in the bar chart
    legend_title="Treatment",
    font=dict(size=12),
    hovermode="x unified"  # Interactive hover
)

# Show the figure
fig.show()
```



## Key Findings:

1. Impact of Family History on Seeking Treatment: The bar chart reveals that individuals with a family history of mental health issues are more likely to seek treatment compared to those without such a history. However, the treatment decision varies by other factors, and the data shows a nuanced picture of how family history influences mental health care choices.

2. Proportion of Treatment: The pie chart provides insights into the proportion of individuals who opt for various types of mental health treatment. This breakdown helps to understand the preferences in treatment choices, whether it is medication, therapy, or other forms of mental health care.

---

## Description of the analysation related to workplace mental health benefits, wellness programs, and employee support for mental health treatment

This analysis explores several aspects related to workplace mental health benefits, wellness programs, and employee support for mental health treatment. The data focuses on respondents' access to mental health benefits, their encouragement to seek treatment, the effectiveness of wellness programs, and the ease of taking mental health leave.

**Proportion of Respondents with Mental Health Benefits:**

The first part analyzes the availability of mental health benefits in the workplace, categorizing responses as "Yes" or "No." This data is represented through an interactive pie chart, providing an overview of the proportion of employees who have access to mental health benefits.

**Does the Presence of Benefits Encourage Seeking Treatment?:**

The second part investigates whether having mental health benefits in place encourages employees to seek treatment. This is displayed in a grouped bar chart, breaking down responses based on the presence of mental health benefits and employees' willingness to seek help.

**Wellness Programs and Mental Health Care Options:**

The third part examines the effectiveness of wellness programs and mental health care options in workplaces. It looks at how the availability of these options impacts employee engagement with mental health care services. This data is shown in a grouped bar chart, which compares wellness programs to care options.

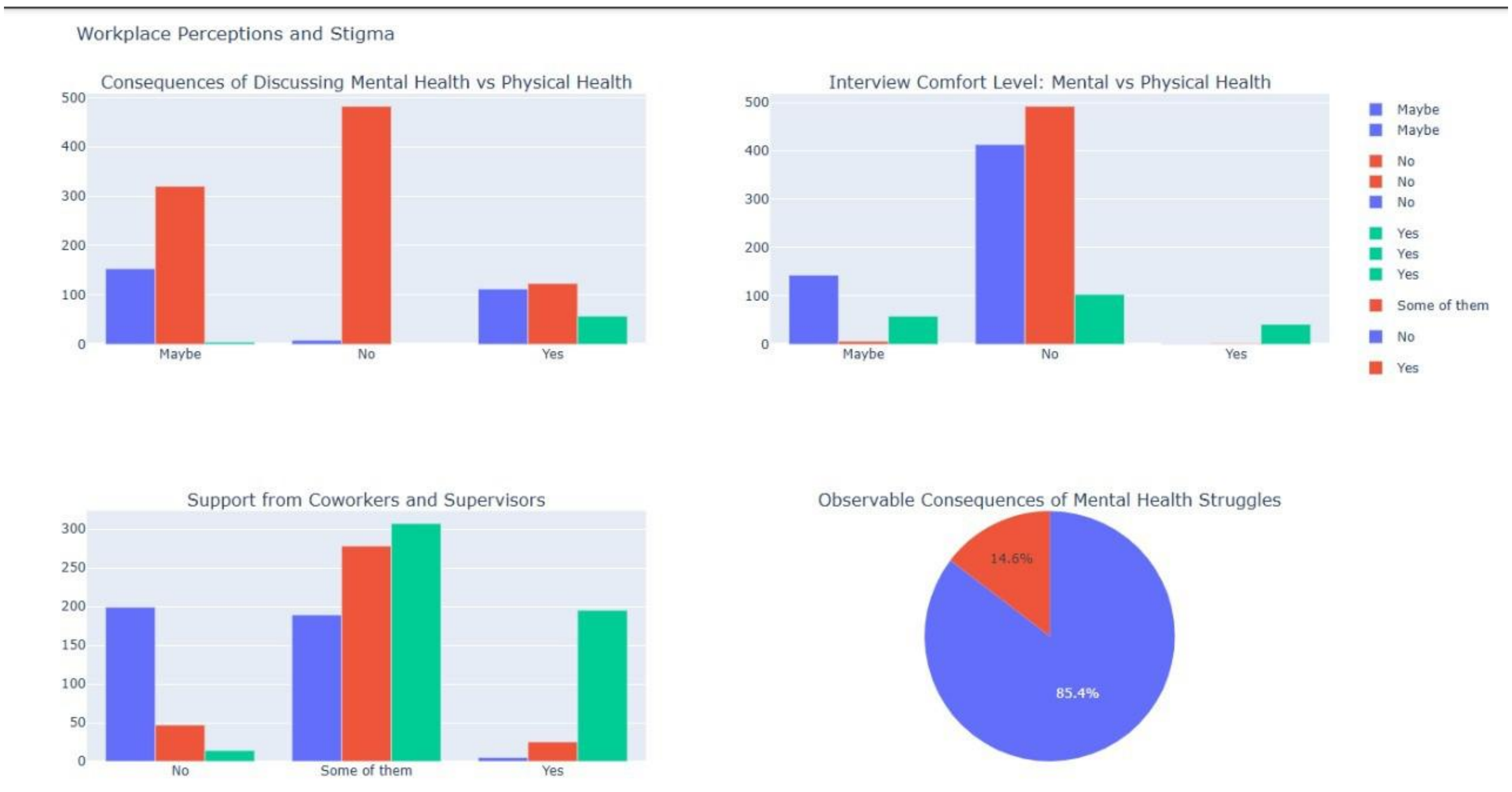
**Ease of Taking Mental Health Leave:**

The final part evaluates how comfortable employees are with taking mental health leave. An interactive pie chart illustrates the ease or difficulty of taking mental health leave based on employee responses.

In [ ]:

```python
import pandas as pd
import plotly.express as px
from plotly.subplots import make_subplots
import plotly.graph_objects as go

# Create a subplot grid (2 rows, 2 columns) and specify the pie chart type for (1,1) and (2,2)
fig = make_subplots(
    rows=2, cols=2,
    subplot_titles=[
        'Availability of Mental Health Benefits',
        'Does the Presence of Mental Health Benefits Encourage Seeking Help?',
        'Effectiveness of Wellness Programs and Mental Health Care Options',
        'Ease of Taking Mental Health Leave'
    ],
    specs=[[{'type': 'pie'}, {'type': 'xy'}], [{'type': 'xy'}, {'type': 'pie'}]]
)

# --------- 1. Proportion of Respondents with Mental Health Benefits ---------
if 'benefits' in data.columns:
    # Count the responses for mental health benefits
    benefits_counts = data['benefits'].value_counts().reset_index(name='count')
    benefits_counts.columns = ['benefits', 'Count']

    # Create a pie chart for the availability of mental health benefits
    pie_fig = px.pie(benefits_counts, names='benefits', values='Count', color='benefits',
                     color_discrete_map={'Yes': 'lightgreen', 'No': 'lightcoral'})
    pie_fig.update_traces(textinfo='percent+label', hole=0.3)

    # Add pie chart to the subplot (row 1, col 1)
    for trace in pie_fig.data:
        fig.add_trace(trace, row=1, col=1)

# --------- 2. Does the Presence of Benefits Encourage Seeking Treatment ---------
if 'benefits' in data.columns and 'seek_help' in data.columns:
    # Count the responses for seeking help based on benefits
    benefits_seek_help_counts = data.groupby(['benefits', 'seek_help']).size().reset_index(name='count')

    # Create a bar chart for encouragement to seek help based on mental health benefits
    bar_fig = px.bar(benefits_seek_help_counts, x='benefits', y='count', color='seek_help',
                     title='Does the Presence of Mental Health Benefits Encourage Seeking Help?',
                     labels={'benefits': 'Mental Health Benefits', 'seek_help': 'Encourages Seeking Help', 'count': 'Number of
                     barmode='group')

    # Add bar chart to the subplot (row 1, col 2)
    for trace in bar_fig.data:
        fig.add_trace(trace, row=1, col=2)
```

```
# --------- 3. Wellness Programs and Care Options ---------
if 'wellness_program' in data.columns and 'care_options' in data.columns:
    # Count the responses for wellness programs and care options
    wellness_care_counts = data.groupby(['wellness_program', 'care_options']).size().reset_index(name='count')

    # Create a bar chart for the effectiveness of wellness programs and care options
    bar_fig = px.bar(wellness_care_counts, x='wellness_program', y='count', color='care_options',
                     title='Effectiveness of Wellness Programs and Mental Health Care Options',
                     labels={'wellness_program': 'Wellness Program', 'care_options': 'Mental Health Care Options', 'count': 'N
                     barmode='group')

    # Add bar chart to the subplot (row 2, col 1)
    for trace in bar_fig.data:
        fig.add_trace(trace, row=2, col=1)

# --------- 4. Ease of Taking Mental Health Leave ---------
if 'leave' in data.columns:
    # Count the responses for the ease of taking mental health leave
    leave_counts = data['leave'].value_counts().reset_index(name='count')
    leave_counts.columns = ['Ease of Taking Leave', 'Count']

    # Create an interactive pie chart for the ease of taking mental health leave
    pie_fig = px.pie(leave_counts, names='Ease of Taking Leave', values='Count', title='Ease of Taking Mental Health Leave')
    pie_fig.update_traces(textinfo='percent+label', hole=0.3)  # To add percentage and label in the pie chart

    # Add pie chart to the subplot (row 2, col 2)
    for trace in pie_fig.data:
        fig.add_trace(trace, row=2, col=2)

# Update layout and show the final grid
fig.update_layout(
    title_text="Mental Health Survey Results",
    showlegend=False,  # Optionally hide the legend if it's not needed
    height=800,  # Adjust height to fit all graphs
    width=1400,  # Adjust width for a better layout
)

fig.show()
```



# Workplace Perceptions and Stigma: Insights on Mental Health

#### Consequences of Discussing Mental Health vs Physical Health

This graph compares the perceived consequences of discussing mental health and physical health in the workplace. It reveals whether employees believe they face greater consequences when discussing mental health compared to physical health issues, highlighting the stigma and societal challenges surrounding mental health at work.

#### Interview Comfort Level: Mental vs Physical Health

The second graph explores how comfortable candidates feel discussing mental health versus physical health during job interviews. It sheds light on whether stigma discourages candidates from addressing mental health concerns during the hiring process, offering valuable insights into potential barriers to open communication in professional settings.

#### Support from Coworkers and Supervisors

This stacked bar chart illustrates the levels of support employees receive from coworkers and supervisors concerning mental health issues. By visualizing the combination of support from both groups, it helps identify areas where workplace culture may need improvement to foster a more supportive environment for mental health discussions and accommodations.

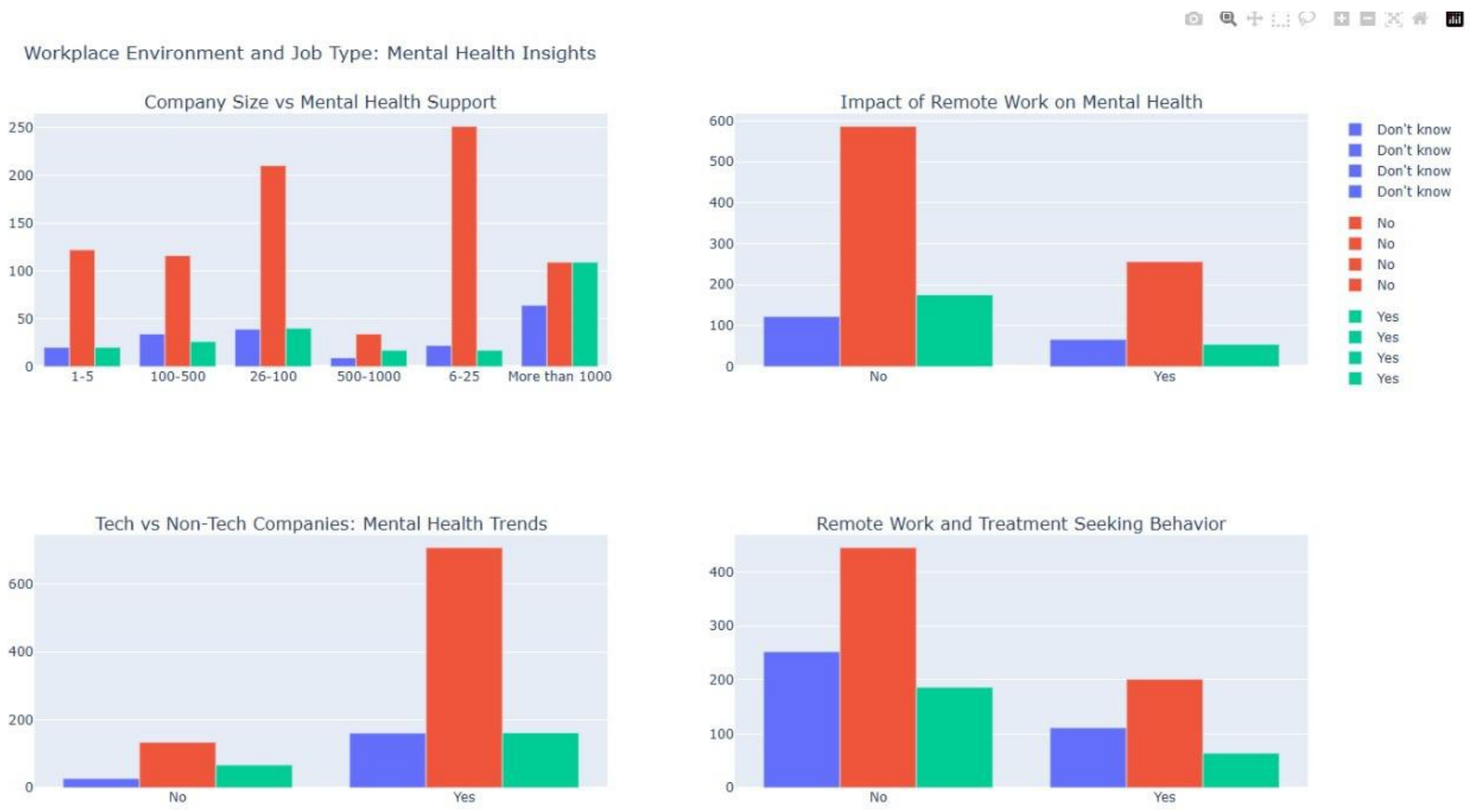
#### Observable Consequences of Mental Health Struggles

The pie chart highlights the observable consequences of mental health struggles as reported by employees. This graph helps identify how frequently mental health struggles manifest visibly in the workplace, providing a deeper understanding of how these challenges are perceived and recognized by others.

These insights collectively offer a comprehensive view of the challenges employees face regarding mental health in the workplace, the stigma they encounter, and the level of support they receive, guiding potential interventions and improvements in workplace culture.

In [ ]:

```python
from plotly.subplots import make_subplots
import plotly.graph_objects as go
import plotly.express as px

# Create a subplot grid (2 rows, 2 columns)
fig = make_subplots(
    rows=2, cols=2,
    subplot_titles=[
        'Consequences of Discussing Mental Health vs Physical Health',
        'Interview Comfort Level: Mental vs Physical Health',
        'Support from Coworkers and Supervisors',
        'Observable Consequences of Mental Health Struggles'
    ],
    specs=[[{'type': 'xy'}, {'type': 'xy'}], [{'type': 'xy'}, {'type': 'pie'}]]
)

# 1. Consequences of Mental Health vs Physical Health
if 'mental_health_consequence' in data.columns and 'phys_health_consequence' in data.columns:
    consequence_counts = data.groupby(['mental_health_consequence', 'phys_health_consequence']).size().reset_index(name='count
    fig_1 = px.bar(consequence_counts, x='mental_health_consequence', y='count', color='phys_health_consequence',
                   labels={'mental_health_consequence': 'Mental Health Consequence', 'phys_health_consequence': 'Physical Heal
    for trace in fig_1.data:
        fig.add_trace(trace, row=1, col=1)

# 2. Interview Comfort Level
if 'mental_health_interview' in data.columns and 'phys_health_interview' in data.columns:
    interview_comfort_counts = data.groupby(['mental_health_interview', 'phys_health_interview']).size().reset_index(name='cou
    fig_2 = px.bar(interview_comfort_counts, x='mental_health_interview', y='count', color='phys_health_interview', barmode='s
    for trace in fig_2.data:
        fig.add_trace(trace, row=1, col=2)

# 3. Support from Coworkers and Supervisors (Replaced Heatmap with Stacked Bar Chart)
if 'coworkers' in data.columns and 'supervisor' in data.columns:
    support_counts = data.groupby(['coworkers', 'supervisor']).size().reset_index(name='count')
    fig_3 = px.bar(support_counts, x='coworkers', y='count', color='supervisor',
                   title='Support from Coworkers and Supervisors for Mental Health Issues',
                   labels={'coworkers': 'Support from Coworkers', 'supervisor': 'Support from Supervisors', 'count': 'Number o
                   barmode='stack')  # Stacked bar chart to visualize support levels
    for trace in fig_3.data:
        fig.add_trace(trace, row=2, col=1)

# 4. Observable Consequences
if 'obs_consequence' in data.columns:
    obs_consequence_counts = data['obs_consequence'].value_counts().reset_index(name='count')
    obs_consequence_counts.columns = ['Observable Consequence', 'Count']
    fig_4 = px.pie(obs_consequence_counts, names='Observable Consequence', values='Count')
    for trace in fig_4.data:
        fig.add_trace(trace, row=2, col=2)

# Update layout and show the final grid
fig.update_layout(
    title_text="Workplace Perceptions and Stigma",
    height=800,
    width=1400
)

fig.show()
```



---

# Workplace Environment and Job Type: Mental Health Insights

This dashboard presents key insights into how workplace characteristics influence mental health support, wellness programs, and treatment-seeking behavior. The visualizations offer a clear comparison across company size, remote work arrangements, and industry type.

## 1. Company Size vs Mental Health Support

- **Description**: This chart examines the availability of wellness programs based on company size.
- **Insight**: Larger companies tend to offer more wellness programs compared to smaller organizations, showcasing a potential correlation between company resources and mental health support.

## 2. Impact of Remote Work on Mental Health Support

- **Description**: This visualization highlights the relationship between remote work arrangements and wellness programs.
- **Insight**: Companies with remote work policies may show varying levels of wellness program adoption. It explores whether remote work promotes or limits access to mental health support.

## 3. Tech vs Non-Tech Companies: Mental Health Trends

- **Description**: A comparative analysis of mental health support availability in tech and non-tech industries.
- **Insight**: Tech companies might display unique trends in adopting wellness programs, which could reflect their culture, employee demands, or resources dedicated to mental health initiatives.

## 4. Remote Work and Treatment Seeking Behavior

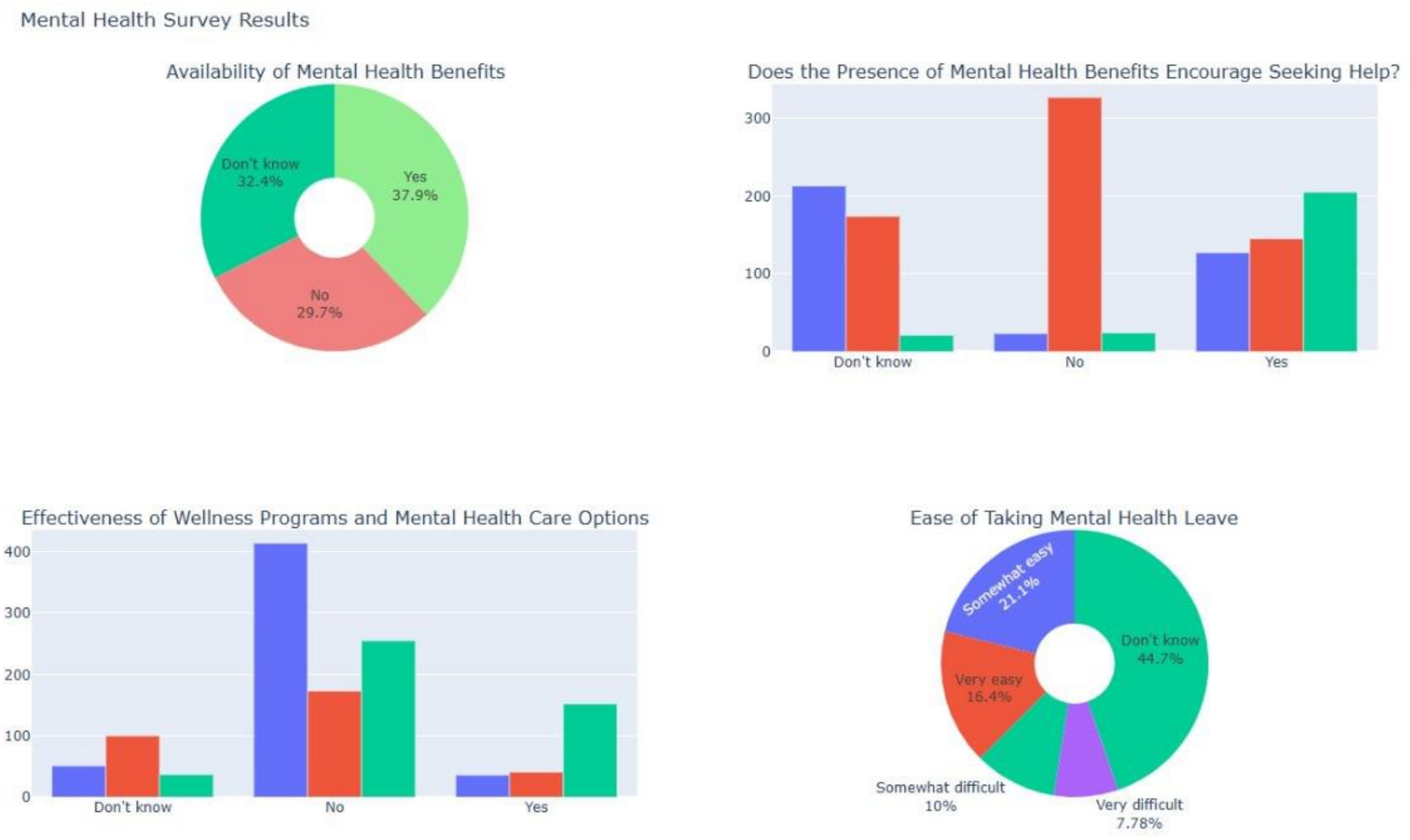
- **Description**: This chart focuses on whether remote work influences employees' willingness to seek mental health treatment.
- **Insight**: Employees working remotely may exhibit different treatment-seeking behaviors, providing insight into how work arrangements impact mental health awareness and help-seeking culture.

## Overall Takeaway

The analysis sheds light on how workplace size, remote work policies, and industry type influence mental health initiatives and behaviors. These findings can help organizations better tailor mental health support systems to meet employees' needs and improve overall workplace well-being.

In [ ]:

```
from plotly.subplots import make_subplots
import plotly.graph_objects as go
import plotly.express as px

# Create a subplot grid (2 rows, 2 columns)
fig = make_subplots(
    rows=2, cols=2,
    subplot_titles=[
        'Company Size vs Mental Health Support',
        'Impact of Remote Work on Mental Health',
        'Tech vs Non-Tech Companies: Mental Health Trends',
        'Remote Work and Treatment Seeking Behavior'
    ],
    specs=[[{'type': 'xy'}, {'type': 'xy'}], [{'type': 'xy'}, {'type': 'xy'}]]  # Adjusted specs for bar charts
)

# 1. Company Size vs Mental Health Support
if 'no_employees' in data.columns and 'wellness_program' in data.columns:
    size_support_counts = data.groupby(['no_employees', 'wellness_program']).size().reset_index(name='count')
    fig_1 = px.bar(size_support_counts, x='no_employees', y='count', color='wellness_program',
                   labels={'no_employees': 'Company Size', 'wellness_program': 'Wellness Program'})
    for trace in fig_1.data:
        fig.add_trace(trace, row=1, col=1)

# 2. Impact of Remote Work on Mental Health
if 'remote_work' in data.columns and 'wellness_program' in data.columns:
    remote_support_counts = data.groupby(['remote_work', 'wellness_program']).size().reset_index(name='count')
    fig_2 = px.bar(remote_support_counts, x='remote_work', y='count', color='wellness_program',
                   labels={'remote_work': 'Remote Work', 'wellness_program': 'Wellness Program'})
    for trace in fig_2.data:
        fig.add_trace(trace, row=1, col=2)

# 3. Tech vs Non-Tech Companies: Mental Health Trends
if 'tech_company' in data.columns and 'wellness_program' in data.columns:
    tech_support_counts = data.groupby(['tech_company', 'wellness_program']).size().reset_index(name='count')
    fig_3 = px.bar(tech_support_counts, x='tech_company', y='count', color='wellness_program',
                   labels={'tech_company': 'Tech Company', 'wellness_program': 'Wellness Program'})
    for trace in fig_3.data:
        fig.add_trace(trace, row=2, col=1)

# 4. Remote Work and Treatment Seeking Behavior
if 'remote_work' in data.columns and 'seek_help' in data.columns:
    remote_seek_help_counts = data.groupby(['remote_work', 'seek_help']).size().reset_index(name='count')
    fig_4 = px.bar(remote_seek_help_counts, x='remote_work', y='count', color='seek_help',
                   labels={'remote_work': 'Remote Work', 'seek_help': 'Seeking Treatment'})
    for trace in fig_4.data:
        fig.add_trace(trace, row=2, col=2)

# Update layout and show the final grid
fig.update_layout(
    title_text="Workplace Environment and Job Type: Mental Health Insights",
    height=800,
    width=1400
)

fig.show()
```



## Combined Mental Health Insights Visualization

This project analyzes survey data related to mental health and workplace support systems. The dataset includes attributes such as gender, age, workplace programs, mental health support, family history, and treatment-seeking behavior.

### Objective:

To combine and visualize key factors that impact mental health in the workplace, answering questions like:

### Gender and Workplace Support:

Do men and women perceive workplace support differently?

### Age and Company Benefits:

Are younger or older employees more likely to work in companies offering mental health benefits?

### Family History and Work Interference:

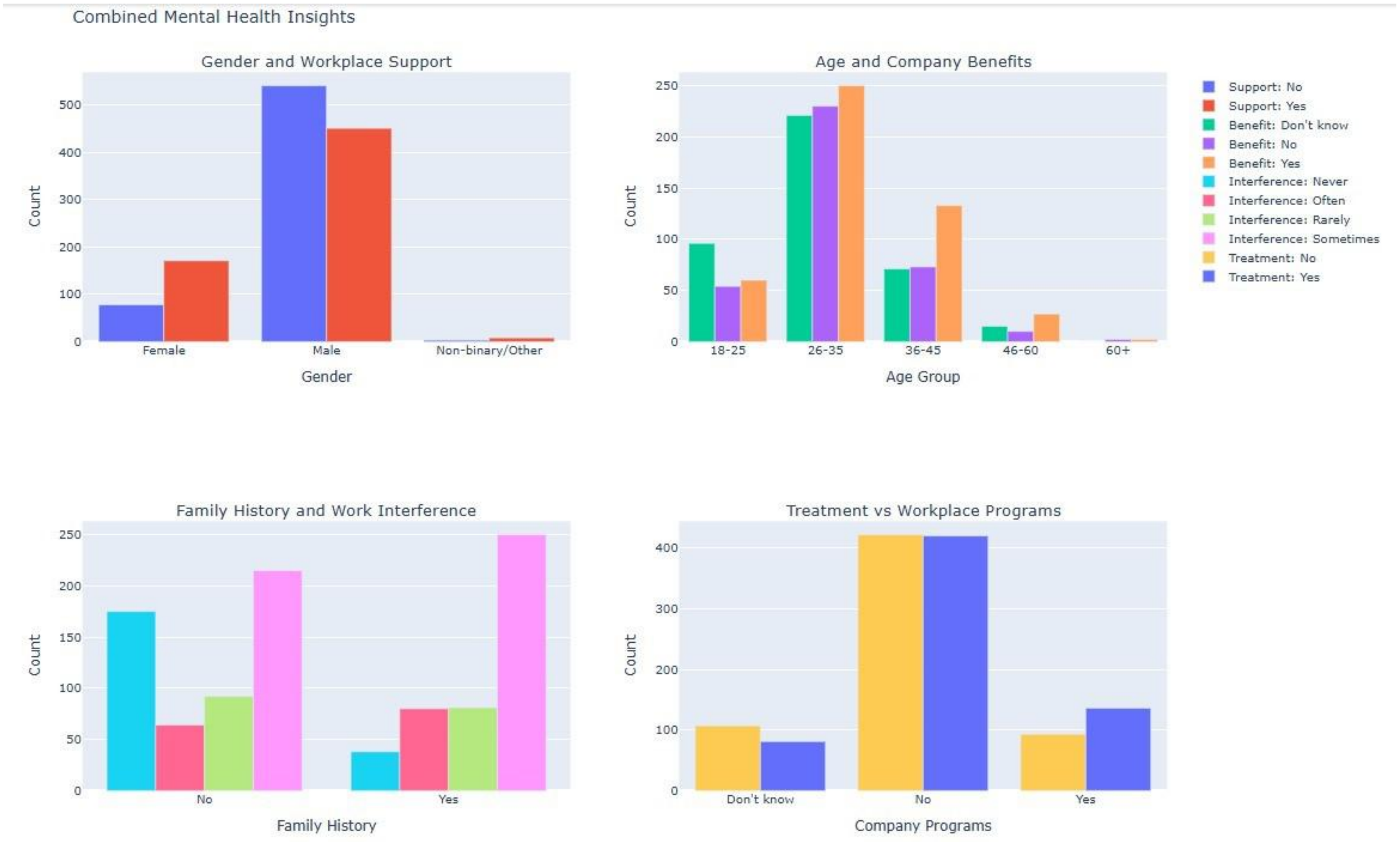
Does having a family history of mental illness correlate with higher work interference?

### Treatment vs Workplace Programs:

Are employees in workplaces with mental health programs more likely to seek treatment?

```
In [ ]: import pandas as pd
        import plotly.graph_objects as go
        from plotly.subplots import make_subplots
```

```python
# Create age groups if not already categorized
data['Age_Group'] = pd.cut(data['Age'], bins=[18, 25, 35, 45, 60, 100],
                           labels=["18-25", "26-35", "36-45", "46-60", "60+"])

# Initialize Subplots in a Grid Layout (2x2)
fig = make_subplots(
    rows=2, cols=2,
    subplot_titles=(
        "Gender and Workplace Support",
        "Age and Company Benefits",
        "Family History and Work Interference",
        "Treatment vs Workplace Programs"
    )
)

# ----------------- Plot 1: Gender and Workplace Support -----------------
gender_support = data.groupby(["Gender_Cleaned", "treatment"]).size().reset_index(name="Count")

for support in gender_support['treatment'].unique():
    subset = gender_support[gender_support["treatment"] == support]
    fig.add_trace(
        go.Bar(x=subset["Gender_Cleaned"], y=subset["Count"], name=f"Support: {support}"),
        row=1, col=1
    )

# ----------------- Plot 2: Age and Company Benefits -----------------
age_benefits = data.groupby(["Age_Group", "benefits"]).size().reset_index(name="Count")

for benefit in age_benefits['benefits'].unique():
    subset = age_benefits[age_benefits["benefits"] == benefit]
    fig.add_trace(
        go.Bar(x=subset["Age_Group"], y=subset["Count"], name=f"Benefit: {benefit}"),
        row=1, col=2
    )

# ----------------- Plot 3: Family History and Work Interference -----------------
family_interference = data.groupby(["family_history", "work_interfere"]).size().reset_index(name="Count")

for interference in family_interference['work_interfere'].unique():
    subset = family_interference[family_interference["work_interfere"] == interference]
    fig.add_trace(
        go.Bar(x=subset["family_history"], y=subset["Count"], name=f"Interference: {interference}"),
        row=2, col=1
    )

# ----------------- Plot 4: Treatment vs Workplace Programs -----------------
treatment_programs = data.groupby(["wellness_program", "treatment"]).size().reset_index(name="Count")

for treatment in treatment_programs['treatment'].unique():
    subset = treatment_programs[treatment_programs["treatment"] == treatment]
    fig.add_trace(
        go.Bar(x=subset["wellness_program"], y=subset["Count"], name=f"Treatment: {treatment}"),
        row=2, col=2
    )

# ----------------- Update Layout -----------------
fig.update_layout(
    title_text="Combined Mental Health Insights",
    height=900,
    width=1400,
    showlegend=True
)

# Update axes for better readability
fig.update_xaxes(title_text="Gender", row=1, col=1)
fig.update_xaxes(title_text="Age Group", row=1, col=2)
fig.update_xaxes(title_text="Family History", row=2, col=1)
fig.update_xaxes(title_text="Company Programs", row=2, col=2)

fig.update_yaxes(title_text="Count", row=1, col=1)
fig.update_yaxes(title_text="Count", row=1, col=2)
fig.update_yaxes(title_text="Count", row=2, col=1)
fig.update_yaxes(title_text="Count", row=2, col=2)

# Show the figure
fig.show()
```



# Conclusions from the above Data Analysis

1. **Treatment Needs by Age Group:**

Employees aged 46-60 have the highest likelihood (59.6%) of requiring mental health treatment.

Younger employees (18-25 and 26-35) show a similar trend, with around 48.5% seeking treatment, indicating mental health issues span all age groups.

2. **Work Interference:**

46.7% of employees reported "Sometimes" experiencing work interference due to mental health issues, the highest among categories.

This highlights the significant impact of mental health on productivity.

3. **Supervisor Support:**

Employees without supportive supervisors (52.7%) are more likely to seek treatment compared to those with supportive supervisors (48.4%).

Supervisor attitudes significantly influence mental health treatment decisions.

4. **Mental Health Interviews:**

80% of employees reported their workplace does not include mental health discussions during interviews, showing a lack of prioritization for mental health in organizational policies.

# Steps to Improve Workplace Mental Health

1. **Raise Awareness and Normalize Mental Health:**

Conduct regular mental health training and workshops for employees and management to destigmatize mental health discussions.

2. **Incorporate Mental Health in Workplace Policies:**

Introduce mandatory mental health check-ins and include mental health topics in onboarding or periodic employee interviews.

3. **Enhance Supervisor Training:**

Train supervisors to recognize signs of mental health struggles and provide them with tools to support employees effectively.

4. **Flexible Work Arrangements:**

Implement policies for remote work or flexible hours to reduce stress for employees reporting frequent work interference.

5. **Create Peer Support Networks:**

Facilitate peer-to-peer programs where employees can share experiences and seek informal support within a safe environment.

6. **Invest in Mental Health Resources:**

Provide access to professional counseling services or subsidize therapy costs for employees needing mental health support.

---

# Why These Steps Are Necessary

**Improved Productivity:** Addressing mental health issues reduces absenteeism and boosts overall efficiency.

**Employee Retention:** A supportive workplace culture increases employee loyalty and reduces turnover.

**Legal and Ethical Responsibility:** Many regions now require organizations to address employee well-being, aligning with ethical best practices.

**Competitive Advantage:** Organizations prioritizing mental health stand out in recruitment, attracting top talent.

```
In [ ]:
```