

People Health Awareness Campaign Analysis

Phase 4: Advanced Analysis and Visualization

In Phase 4, we will perform advanced analyses and create visualizations to gain deeper insights into the dataset. We will leverage IBM Cognos for this purpose.

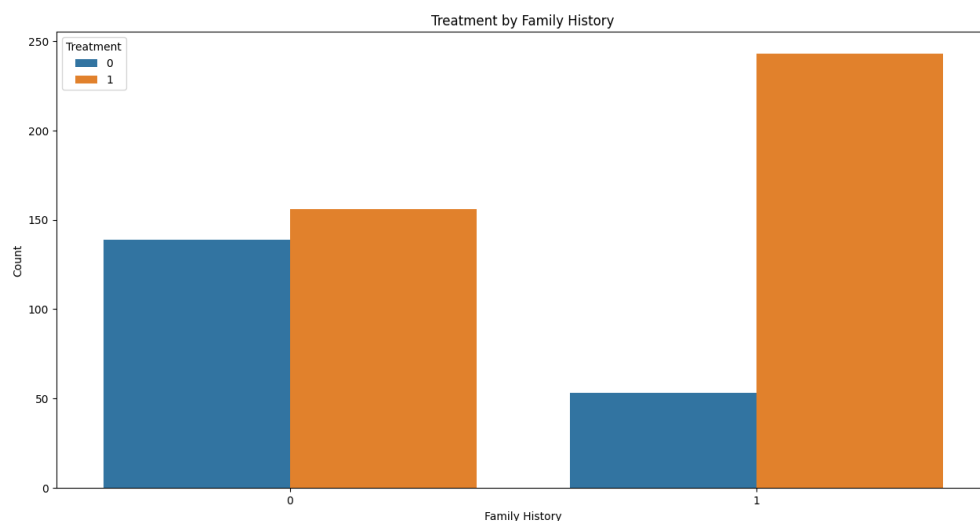
Dataset : <https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey>

Analysis 1: Treatment by Family History

This analysis aims to understand how family history of mental health issues influences the decision to seek treatment.

- Query: `SELECT family_history, COUNT(*) FROM dataset GROUP BY family_history`
- Visualization: Bar chart showing the count of individuals seeking treatment based on family history.

```
plt.figure(figsize=(8, 5))
sns.countplot(data=df, x='family_history', hue='treatment')
plt.title('Treatment by Family History')
plt.xlabel('Family History')
plt.ylabel('Count')
plt.legend(title='Treatment')
plt.show()
```



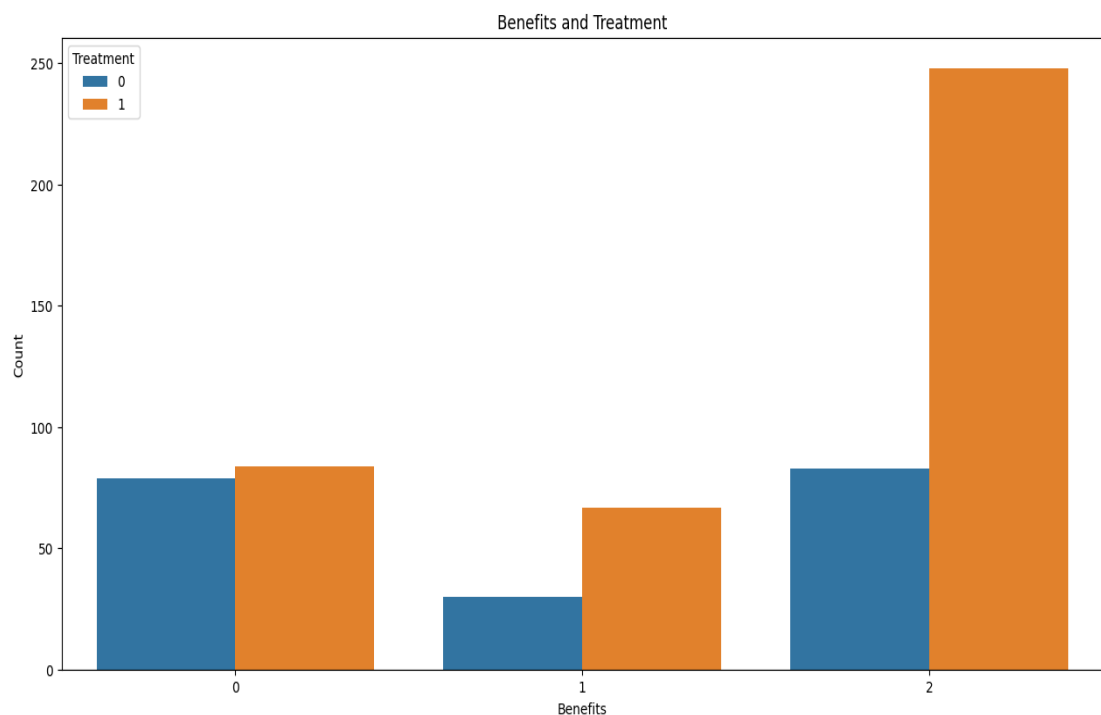
Analysis 2: Work Interference and Treatment

This analysis explores the impact of work interference on the likelihood of seeking mental health treatment.

- Query: `SELECT work_interfere, COUNT(*) FROM dataset GROUP BY work_interfere`

- Visualization: Stacked bar chart representing the count of individuals seeking treatment, categorized by work interference level.

```
plt.figure(figsize=(8, 5))
sns.countplot(data=df, x='work_interfere', hue='treatment')
plt.title('Work Interference and Treatment')
plt.xlabel('Work Interference')
plt.ylabel('Count')
plt.legend(title="Treatment")
plt.show()
```



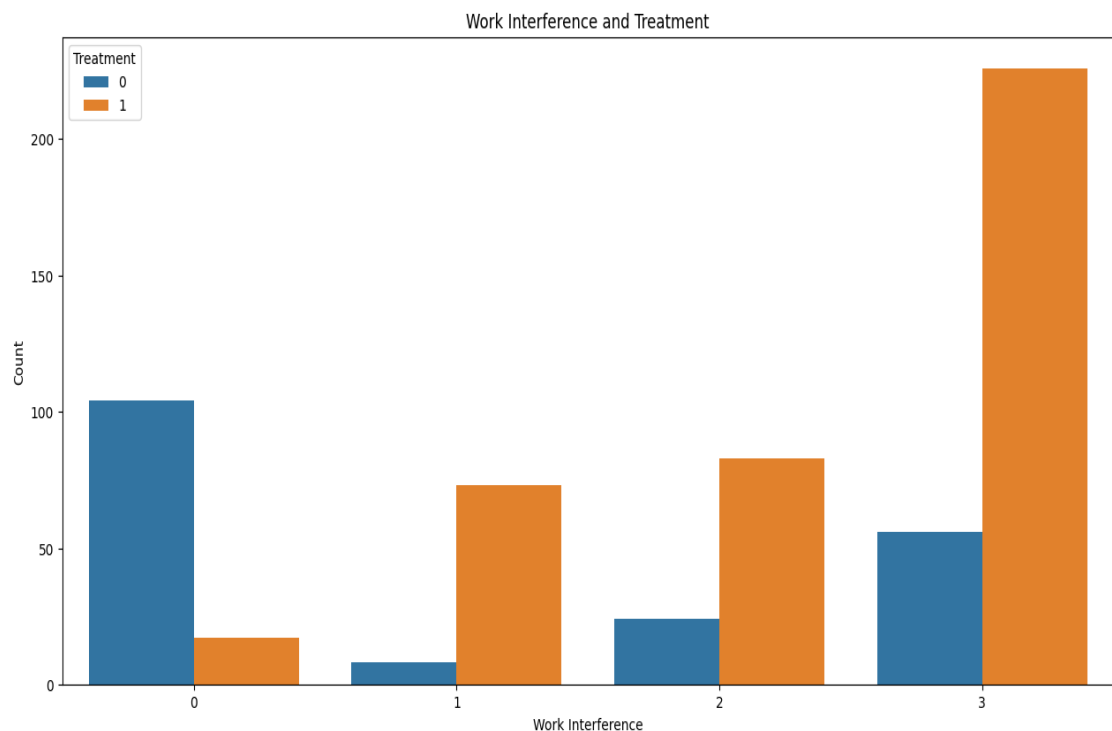
Analysis 3: Benefits and Treatment

This analysis examines the relationship between the availability of mental health benefits and the decision to seek treatment.

- Query: `SELECT benefits, COUNT(*) FROM dataset GROUP BY benefits`

- Visualization: Pie chart displaying the distribution of individuals seeking treatment based on benefits availability.

```
plt.figure(figsize=(8, 5))
sns.countplot(data=df, x='benefits', hue='treatment')
plt.title('Benefits and Treatment')
plt.xlabel('Benefits')
plt.ylabel('Count')
plt.legend(title="Treatment")
plt.show()
```

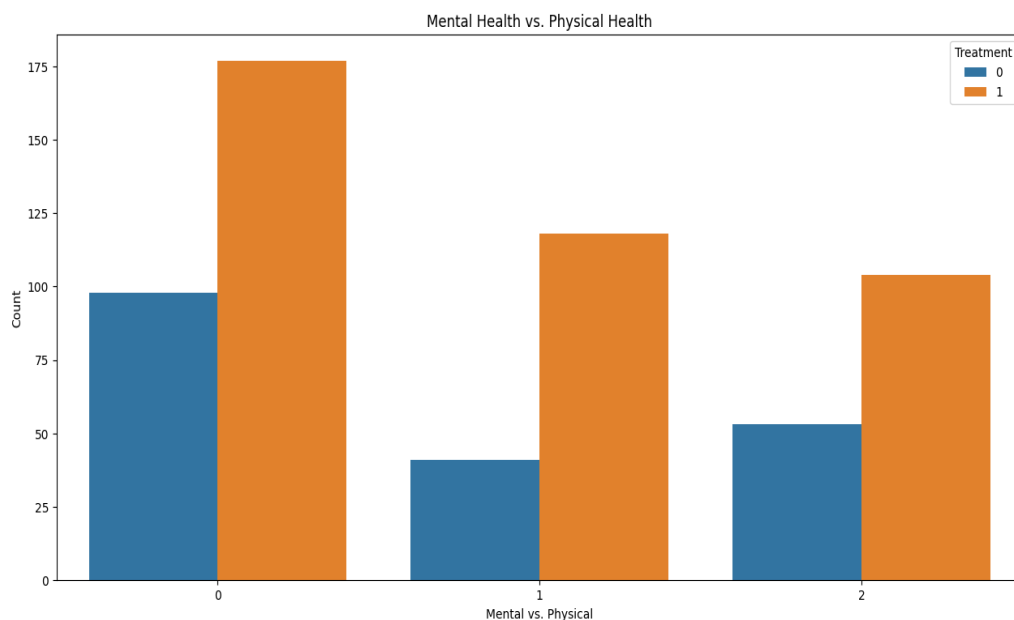


Analysis 4: Mental Health vs. Physical Health

This analysis compares the preference for seeking treatment for mental health issues versus physical health issues.

- Query: `SELECT mental_vs_physical, COUNT(*) FROM dataset GROUP BY mental_vs_physical`
- Visualization: Donut chart showing the proportion of individuals preferring mental health treatment over physical health.

```
plt.figure(figsize=(8, 5))
sns.countplot(data=df, x='mental_vs_physical', hue='treatment')
plt.title('Mental Health vs. Physical Health')
plt.xlabel('Mental vs. Physical')
plt.ylabel('Count')
plt.legend(title="Treatment")
plt.show()
```

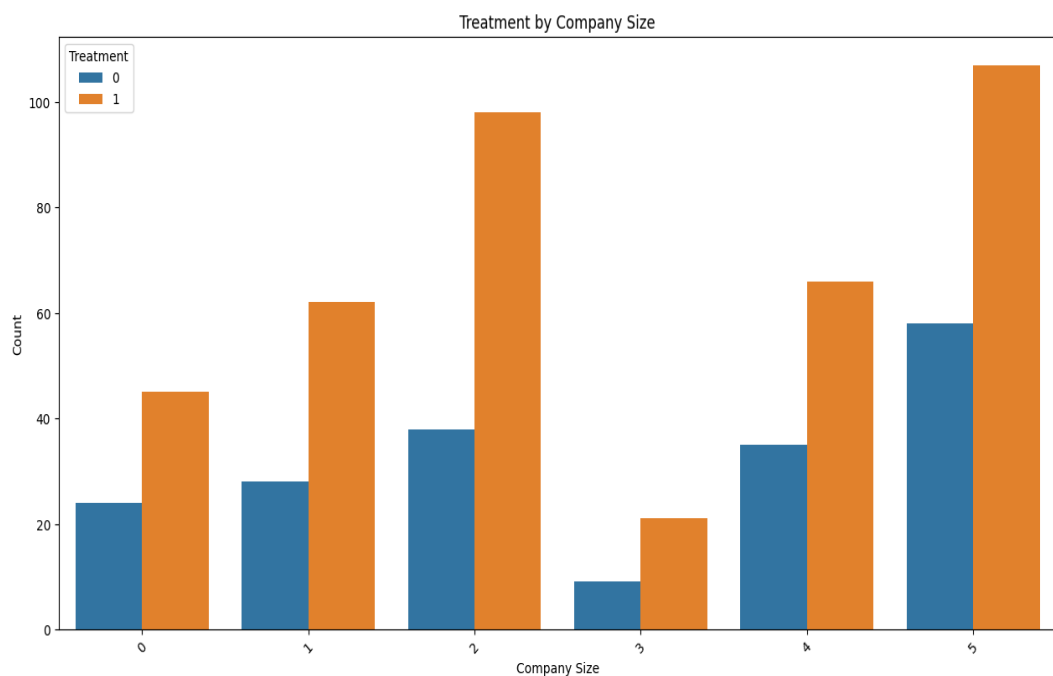


Analysis 5: Treatment by Company Size

This analysis investigates how the size of the company influences the likelihood of seeking mental health treatment.

- Query: `SELECT no_employees, COUNT(*) FROM dataset GROUP BY no_employees`
- Visualization: Horizontal bar chart depicting the count of individuals seeking treatment categorized by company size.

```
plt.figure(figsize=(8, 5))
sns.countplot(data=df, x='no_employees', hue='treatment')
plt.title("Treatment by Company Size")
plt.xlabel('Company Size')
plt.ylabel('Count')
plt.legend(title="Treatment")
plt.xticks(rotation=45)
plt.show()
```

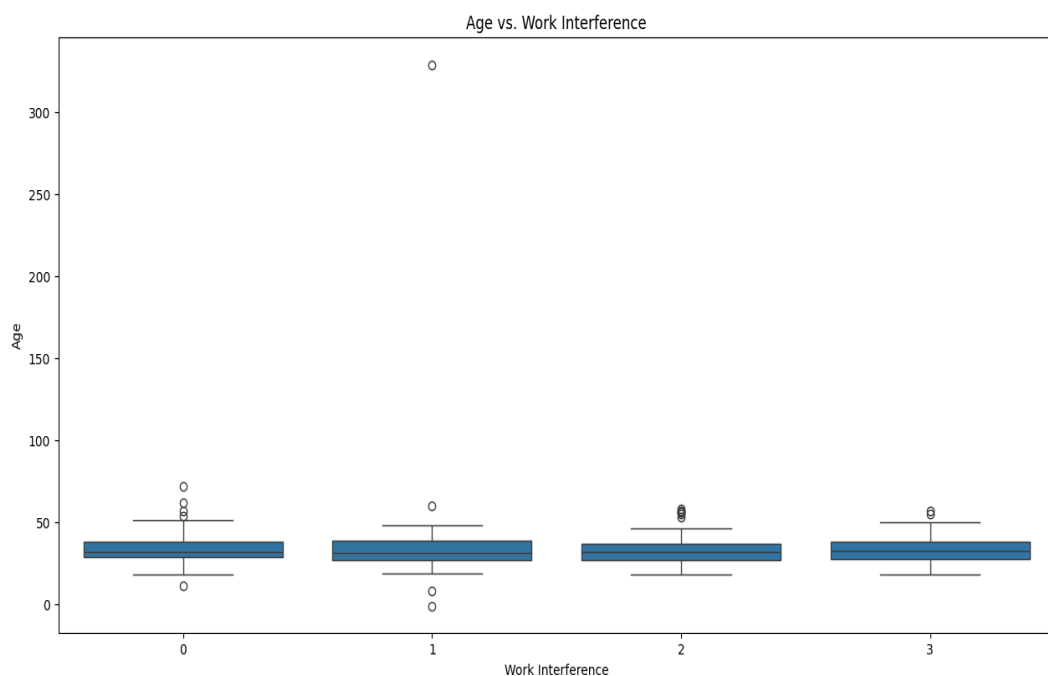


Analysis 6: Age vs. Work Interference

This analysis explores the relationship between age and the level of work interference due to mental health issues.

- Query: `SELECT Age, AVG(work_interfere) FROM dataset GROUP BY Age`
- Visualization: Line plot representing the average work interference level for each age group.

```
plt.figure(figsize=(10, 6))
sns.boxplot(data=df, x='work_interfere', y='Age')
plt.title('Age vs. Work Interference')
plt.xlabel('Work Interference')
plt.ylabel('Age')
plt.show()
```



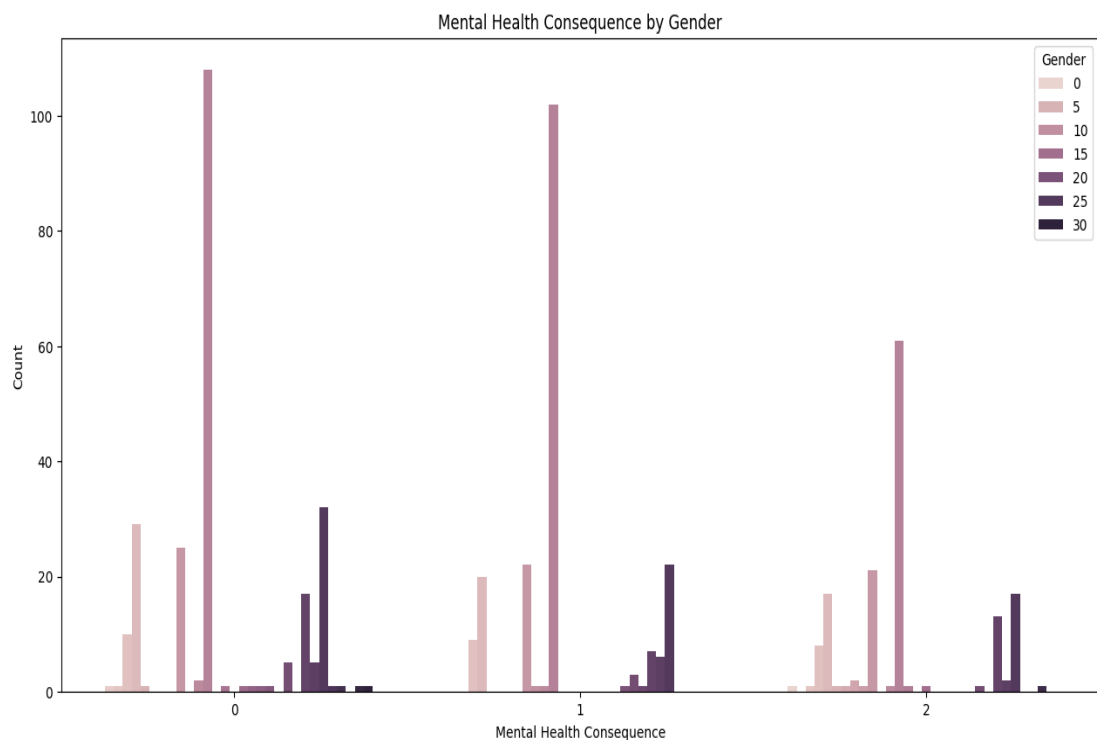
Analysis 7: Mental Health Consequence by Gender

This analysis aims to understand the consequences of mental health issues based on gender.

- Query: `SELECT Gender, mental_health_consequence, COUNT(*) FROM dataset GROUP BY Gender, mental_health_consequence`

- Visualization: Clustered bar chart showing the count of individuals facing consequences, segmented by gender.

```
plt.figure(figsize=(8, 5))
sns.countplot(data=df, x='mental_health_consequence', hue='Gender')
plt.title('Mental Health Consequence by Gender')
plt.xlabel('Mental Health Consequence')
plt.ylabel('Count')
plt.legend(title='Gender')
plt.show()
```



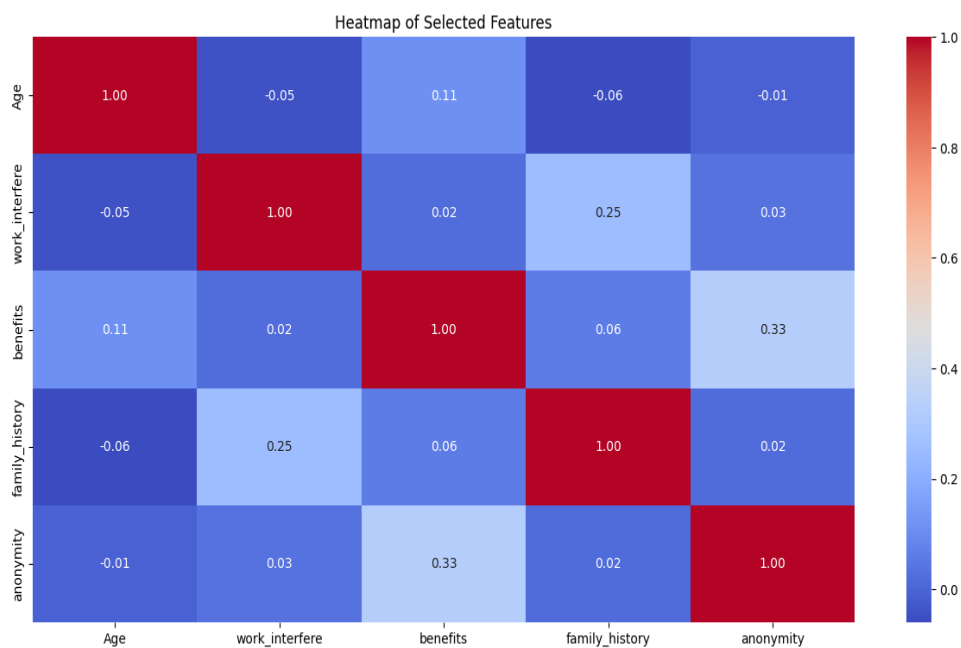
Analysis 8: Correlation Heatmap

This visualization displays the correlation between selected features, providing insights into potential patterns.

- Query: N/A

- Visualization: Heatmap depicting the correlation matrix of selected features.

```
plt.figure(figsize=(10, 8))
selected_features = ['Age', 'work_interfere', 'benefits', 'family_history', 'anonymity']
sns.heatmap(df[selected_features].corr(), annot=True, cmap='coolwarm', fmt=".2f")
plt.title('Heatmap of Selected Features')
plt.show()
```



Analysis 9: Tech Company and Treatment

This stacked bar chart visualizes the relationship between working in a tech company and the decision to seek mental health treatment.

- Query: `'SELECT tech_company, COUNT(*) FROM dataset GROUP BY tech_company'`
- Visualization: Stacked bar chart representing the count of individuals seeking treatment, categorized by tech company employment.

```
plt.figure(figsize=(10, 6))
tech_treatment = df.groupby('tech_company')['treatment'].value_counts().unstack().fillna(0)
tech_treatment.plot(kind='bar', stacked=True)
plt.title('Tech Company and Treatment')
plt.xlabel('Tech Company')
plt.ylabel('Count')
plt.legend(title='Treatment', loc='upper right')
plt.show()
```

