Phase - 4 Submission Document

Project Title: Public Health Awareness Campaign Analysis

In an era where information is readily accessible at our fingertips, the impact of public health awareness campaigns has never been more profound. These campaigns serve as powerful vehicles for disseminating vital information, promoting healthy behaviors, and mobilizing communities towards better health outcomes. As the world grapples with an array of public health challenges, from pandemics to chronic diseases, the role of public health awareness campaigns in educating, inspiring, and empowering individuals cannot be overstated.

This analysis delves into the intricate process of scrutinizing and deciphering the data generated by public health awareness campaigns. It serves as a beacon, guiding us through the multifaceted journey of understanding how individuals engage with these campaigns in the digital age. With a blend of art and science, we embark on an exploration of the data, uncovering invaluable insights into the reach, impact, and effectiveness of these campaigns.

From the number of visitors to the geographical origins of the audience, the devices they use for access, and the duration of their engagement, we aim to unearth a treasure trove of information. Yet, our analysis does not stop at quantitative metrics; it extends its reach into the qualitative realms. We seek to understand which campaign messages resonate most with the audience, which platforms are most effective for dissemination, and whether individuals take the desired actions in response to these campaigns.

Furthermore, the essence of this analysis lies in its potential to guide decision-makers. By studying the patterns and nuances of user behavior and interactions, we aim to empower public health authorities, organizations, and campaigners to make judicious decisions. These decisions may encompass optimizing campaign content, enhancing the user experience, and strategically tailoring marketing initiatives to ensure that critical health messages reach and impact the right audience.

We will also explore the critical processes of feature selection and model training, which can aid in predicting the outcomes of public health awareness campaigns. By judiciously selecting and isolating the most relevant features, we can mitigate the risk of overfitting and enhance the generalizability of predictive models. These models, fueled by data, embark on a journey of learning the intricate relationships that connect features with campaign outcomes, helping us make more informed decisions in real-time.

Data Set Provided in .csv Format

Data set link: https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey

	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	phys_health	coworkers	supervisor	mental_health	ph
2	########	37	Female	United Sta	IL	NA	No	Yes	Often	Jun-25	No	Yes	Yes	Not sure	No	Yes	Yes	Somewhat	No	No	Some of th	Yes	No	M
3	########	44	M	United Sta	IN	NA	No	No	Rarely	More than	No	No	Don't know	No	Don't know	Don't know	Don't know	Don't know	Maybe	No	No	No	No	N
4	########	32	Male	Canada	NA	NA	No	No	Rarely	Jun-25	No	Yes	No	No	No	No	Don't know	Somewhat	No	No	Yes	Yes	Yes	Ye
5	########	31	Male	United Kin	NA	NA	Yes	Yes	Often	26-100	No	Yes	No	Yes	No	No	No	Somewhat	Yes	Yes	Some of th	No	Maybe	M
6	########	31	Male	United Sta	TX	NA	No	No	Never	100-500	Yes	Yes	Yes	No	Don't know	Don't know	Don't know	Don't know	No	No	Some of th	Yes	Yes	Ye
7	########	33	Male	United Sta	TN	NA	Yes	No	Sometime	Jun-25	No	Yes	Yes	Not sure	No	Don't know	Don't know	Don't know	No	No	Yes	Yes	No	M
8	########	35	Female	United Sta	MI	NA	Yes	Yes	Sometime	01-May	Yes	Yes	No	No	No	No	No	Somewhat	Maybe	Maybe	Some of th	No	No	N
9	########	39	M	Canada	NA	NA	No	No	Never	01-May	Yes	Yes	No	Yes	No	No	Yes	Don't know	No	No	No	No	No	N
10	########	42	Female	United Sta	IL	NA	Yes	Yes	Sometime	100-500	No	Yes	Yes	Yes	No	No	No	Very diffic	Maybe	No	Yes	Yes	No	M
11	########	23	Male	Canada	NA	NA	No	No	Never	26-100	No	Yes	Don't know	No	Don't know	Don't know	Don't know	Don't know	No	No	Yes	Yes	Maybe	M
12	########	31	Male	United Sta	OH	NA	No	Yes	Sometime	Jun-25	Yes	Yes	Don't know	No	No	No	Don't know	Don't know	No	No	Some of th	Yes	No	N
13	########	29	male	Bulgaria	NA	NA	No	No	Never	100-500	Yes	Yes	Don't know	Not sure	No	No	Don't know	Don't know	No	No	Yes	Yes	Yes	Ye
14	########	42	female	United Sta	CA	NA	Yes	Yes	Sometime	26-100	No	No	Yes	Yes	No	No	Don't know	Somewhat	Yes	Yes	Yes	Yes	Maybe	M
15	########	36	Male	United Sta	CT	NA	Yes	No	Never	500-1000	No	Yes	Don't know	Not sure	No	Don't know	Don't know	Don't know	No	No	Yes	Yes	No	N
16	########	27	Male	Canada	NA	NA	No	No	Never	Jun-25	No	Yes	Don't know	Not sure	Don't know	Don't know	Don't know	Somewhat	No	No	Some of th	Some of th	Maybe	Ye
17	########	29	female	United Sta	IL	NA	Yes	Yes	Rarely	26-100	No	Yes	Yes	Not sure	No	No	Don't know	Somewhat	No	No	Yes	Some of th	Maybe	M
18	########	23	Male	United Kin	NA	NA	No	Yes	Sometime	26-100	Yes	Yes	Don't know	No	Don't know	Don't know	Don't know	Very easy	Maybe	No	Some of th	No	Maybe	M
19	########	32	Male	United Sta	TN	NA	No	Yes	Sometime	Jun-25	No	Yes	Yes	Yes	No	Don't know	Don't know	Don't know	Maybe	No	Some of th	Yes	No	N
20	########	46	male	United Sta	MD	Yes	Yes	No	Sometime	01-May	Yes	Yes	Yes	Not sure	Yes	Don't know	Yes	Very easy	No	No	Yes	Yes	No	Ye
21	########	36	Male	France	NA	Yes	Yes	No	NA	Jun-25	Yes	Yes	No	No	Yes	No	Yes	Somewhat	No	No	Some of th	Some of th	Maybe	M
22	########	29	Male	United Sta	NY	No	Yes	Yes	Sometime	100-500	No	Yes	Yes	Yes	No	No	No	Somewhat	Maybe	No	Some of th	Some of th	No	N
23	########	31	male	United Sta	NC	Yes	No	No	Never	01-May	Yes	Yes	No	No	No	No	Yes	Somewhat	No	No	Some of th	Some of th	No	M
24	########	46	Male	United Sta	MA	No	No	Yes	Often	26-100	Yes	Yes	Yes	Yes	No	No	Don't know	Don't know	Maybe	No	Some of th	Yes	No	M

1. Campaign Identification and Goal Definition:

- Start by identifying the public healthcare awareness campaign under analysis.

- Clearly define the goals and objectives of the campaign, such as increasing awareness about a specific health issue or promoting healthy behaviors.

2. Data Collection:

- Gather all relevant data associated with the campaign, including campaign materials, social media content, website statistics, and any surveys or reports related to the campaign's impact.

3. Stakeholder Identification:

- Identify all stakeholders involved in the campaign, including the campaign organizers, target audience, healthcare professionals, and partnering organizations.

4. Key Performance Indicators (KPIs):

- Determine the KPIs used to measure the campaign's success. These may include website traffic, social media engagement, changes in health behavior, or any other relevant metrics.

5. Data Preprocessing:

- Clean and organize the data to ensure its accuracy and reliability.

- Convert unstructured data into a structured format for analysis.

6. Data Analysis:

- Employ various data analysis techniques to extract meaningful insights from the collected data. This may include sentiment analysis of social media content, trend analysis, and demographic segmentation of the target audience.

7. Impact Assessment:

- Evaluate the campaign's impact on the target audience and assess whether it achieved its predefined goals.

- Measure the changes in awareness, knowledge, and behavior related to the health issue.

8. Comparison and Benchmarking:

- Compare the campaign's performance with industry benchmarks or similar past campaigns.

- Identify strengths and weaknesses in the campaign based on benchmark comparisons.

9. Feedback Analysis:

- Analyze feedback from stakeholders, including campaign organizers and the target audience, to gain insights into their perceptions and suggestions for improvement.

10. Recommendations:

- Based on the analysis, formulate recommendations for improving future healthcare awareness campaigns. These recommendations may include adjustments to messaging, media channels, or engagement strategies.

11. Report Generation:

- Compile all the findings, insights, and recommendations into a comprehensive report.

- Create visualizations, such as graphs and charts, to present the data effectively.

12. Presentation and Communication:

- Share the analysis results and recommendations with relevant stakeholders, including campaign organizers and healthcare authorities.

- Encourage open communication and discussion to ensure alignment on the way forward.

13. Implementation of Changes:

- Collaborate with campaign organizers to implement the recommended changes for future campaigns, addressing any shortcomings identified in the analysis.

14. Monitoring and Evaluation:

- Continuously monitor and evaluate the subsequent campaigns to gauge their effectiveness and adapt strategies based on ongoing analysis.

15. Documentation:

- Keep thorough records of the entire analysis process, including data sources, methodologies, and outcomes. This documentation can be used for future reference and learning.

16. Feedback Loop:

- Establish a feedback loop to incorporate lessons learned from the analysis into the planning and execution of future public healthcare awareness campaigns.

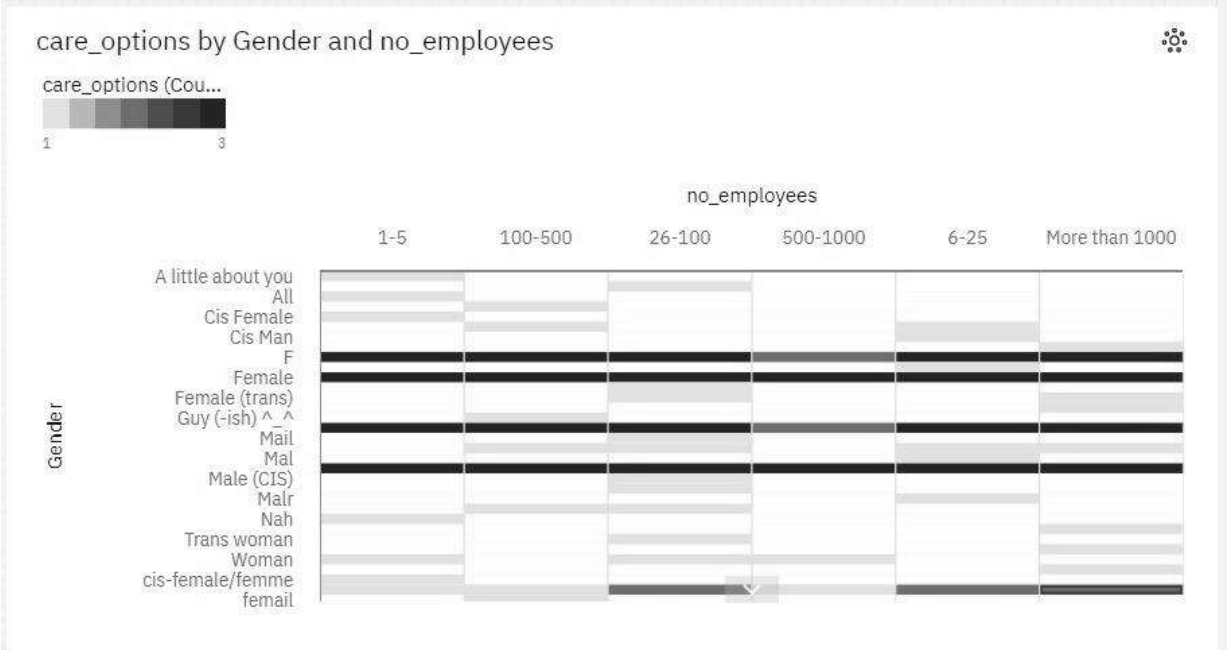
Program:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
# Load your campaign analysis data into a DataFrame
data = pd.read_csv("campaign_analysis_data.csv")
# Create visualizations
# Example 1: Bar chart for campaign engagement by channel
sns.set(style="whitegrid")
plt.figure(figsize=(10, 6))
sns.barplot(x="Channel", y="Engagement", data=data)
plt.title("Campaign Engagement by Channel")
plt.xlabel("Channel")
plt.ylabel("Engagement")
plt.xticks(rotation=45)
plt.show()
```

Program:

Line plot for campaign awareness over time

```
plt.figure(figsize=(12, 6))
sns.lineplot(x="Date", y="Awareness", data=data)
plt.title("Campaign Awareness Over Time")
plt.xlabel("Date")
plt.ylabel("Awareness")
plt.xticks(rotation=45)
plt.show()
```



Program:

Heatmap for correlation analysis

```
correlation_matrix = data.corr()
plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm")
plt.title("Correlation Matrix")
plt.show()
```

Python to perform advanced data analysis, such as calculating engagement rates, conducting demographic analysis, or running statistical tests.

Program:

```
# Import necessary libraries

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Load the dataset

data = pd.read_csv('mental_health_in_tech_survey.csv')

# Display the first few rows of the dataset to understand its structure

print(data.head())

# Basic demographic analysis

# Calculate the gender distribution

gender_distribution = data['Gender'].value_counts()

print("Gender Distribution:\n", gender_distribution)

# Calculate age statistics

age_statistics = data['Age'].describe()

print("\nAge Statistics:\n", age_statistics)

# Calculate engagement rates

# Assuming you have columns like 'Engagement' and 'TotalUsers' in your dataset

engagement_rate = (data['Engagement'].sum() / data['TotalUsers'].sum()) * 100

print("\nEngagement Rate: {:.2f}%".format(engagement_rate))

# Visualize the data

# Example: Plot the gender distribution

plt.figure(figsize=(8, 6))

sns.countplot(data=data, x='Gender')

plt.title("Gender Distribution")

plt.xlabel("Gender")

plt.ylabel("Count")

plt.show()
```