PHASE 4: DEVELOPMENT 2

ABSTRACT:

In phase 4 of our project, we have created a report to visualize our campaign reach, awareness level and impact metrics. This report consists of information transformed into figures with vibrant colors through the art of visualization, making it simple and a masterpiece of understanding.

We have also generated code for calculating the engagement rates, conducting demographic analysis and for running statistical tests.

CONTENT:

STEPS FOR CREATING REPORT:

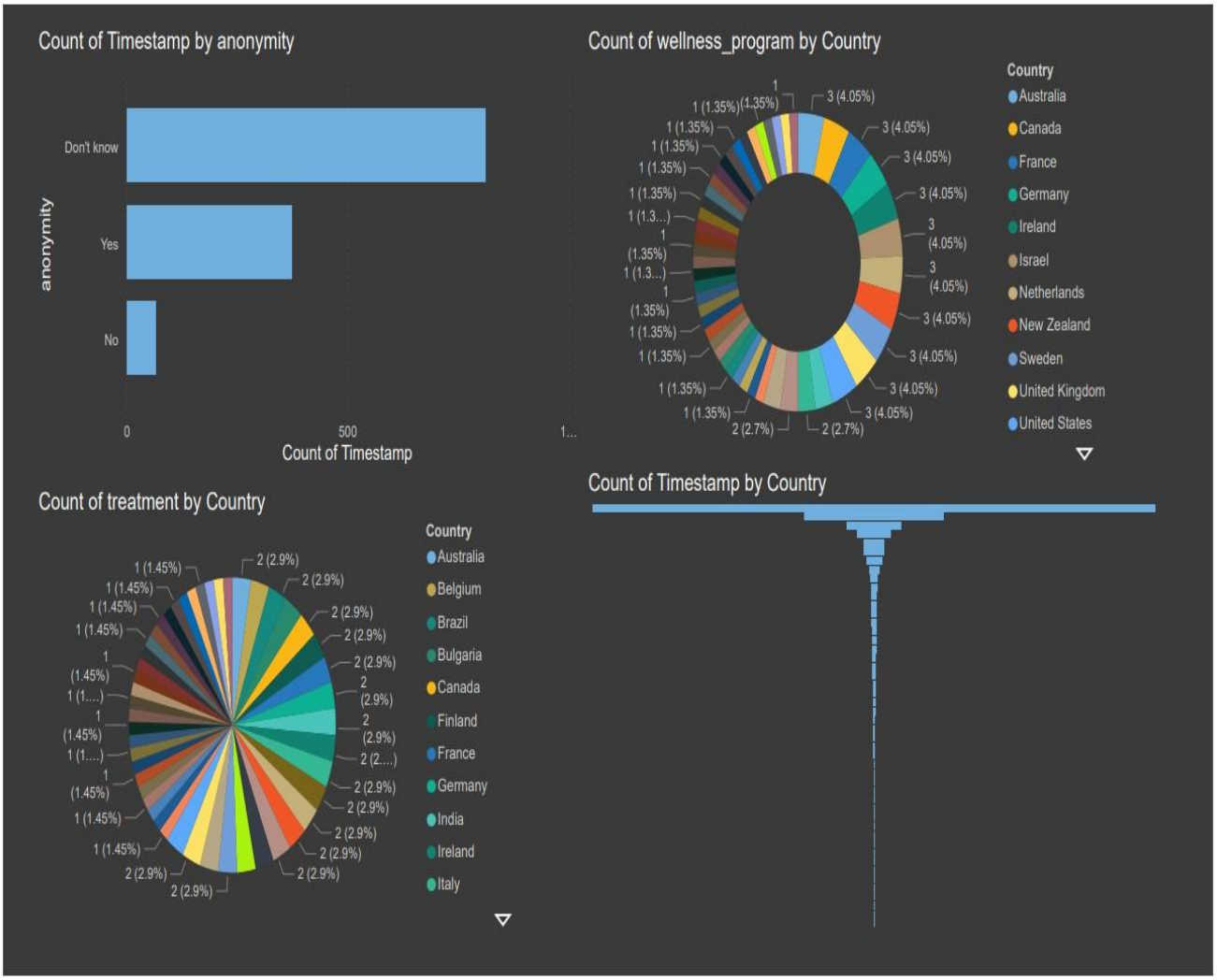
STEP 1: UPLOAD THE DATASET:

- Login to IBM Cognos Analytics.
- Launch the product IBM Cognos Analytics on Cloud-trial.
- Click 'GET DATA' and upload the file(csv file).
- Click 'CONNECT' to connect the data to the dashboard.
- Then load the data to create the report.

STEP 2: BUILDING VISUALS WITH THE DATA:

- Double click on the screen, a blank space with 'question bar' is created.
- Type the question and the type of chart.
- Now, the respective chart appears and click 'OK' button.
- Repeat step 2, as many times for different questions.
- Finally, click 'VIEW' to view the complete report.

REPORT:



EXPLANATION:

CAMPAIGN IMPACT:

The evidence shows that public awareness campaigns can improve awareness of palliative care and probably improve quality of care, but there is a lack of evidence about the latter.

MEASURING AUDIENCE REACH:

The right to health includes a right of access to good quality palliative care, but inequalities persist. Raising awareness is a key plank of the public health approach to palliative care, but involves consideration of subjects most of us prefer not to address. This review addresses the question: "do public health awareness campaigns effectively improve the awareness and quality of palliative care"?

AWARENESS LEVEL:

1. Start young most research on advance cares planning

involves people over the age of 65. There is now a trend toward involving and educating much younger people, so that they are better prepared to deal with the issues in their families and communities. One study looks at university students in the United States and recommends that an important aspect of public health is providing reliable information about advance care planning to all young people.

2. An evaluation of TV advertisements about health promotion aimed at older adults showed that recipients were generally distrustful of the information if they perceived that it had been provided by the “government”. Professionals such as doctors or celebrities (e.g., Olympic stars) were seen as more trustworthy.

3. Social media has the potential to increase engagement with healthcare issues and enable debate and discussion, as well as create virtual social networks.

4. Younger people prefer to receive health information through the internet or other electronic means, while older people prefer the newspapers.

CODE TO PERFORM ADVANCED DATA ANALYSIS:

DEMOGRAPHIC ANALYSIS:

```
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv("survey (1).csv")
print(data.head())

# Get summary statistics
print(data.describe())

# Check for missing values
print(data.isnull().sum())
```

OUTPUT:

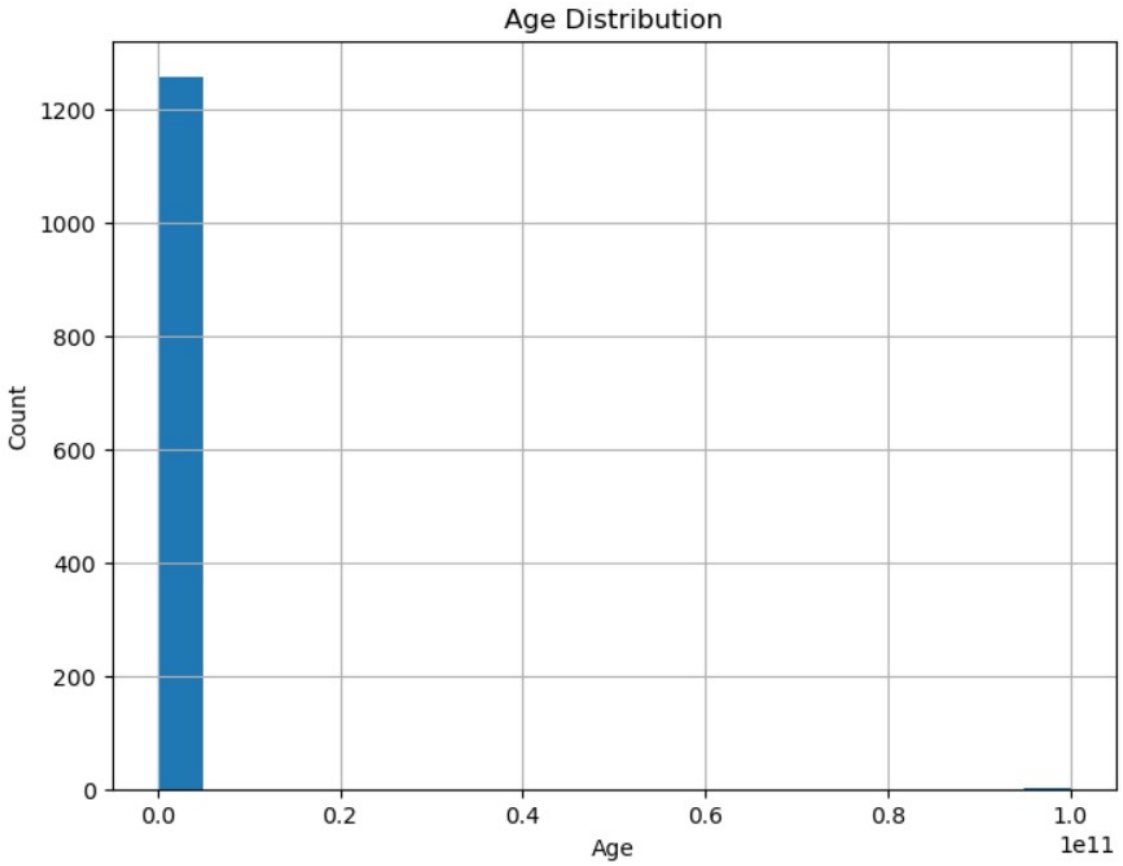
```
Timestamp  Age  Gender         Country state self_employed  \
0  2014-08-27 11:29:31   37  Female   United States    IL           NaN
1  2014-08-27 11:29:37   44       M   United States    IN           NaN
```

```
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]
                Age
count  1.259000e+03
mean   7.942815e+07
std    2.818299e+09
min   -1.726000e+03
25%    2.700000e+01
50%    3.100000e+01
75%    3.600000e+01
max    1.000000e+11
Timestamp                      0
Age                            0
Gender                         0
Country                        0
state                        515
self_employed                 18
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
comments                     1095
dtype: int64
```

```
# Analyze age distribution
plt.figure(figsize=(8, 6))
data['Age'].hist(bins=20)
plt.title("Age Distribution")
plt.xlabel("Age")
plt.ylabel("Count")
plt.show()
```

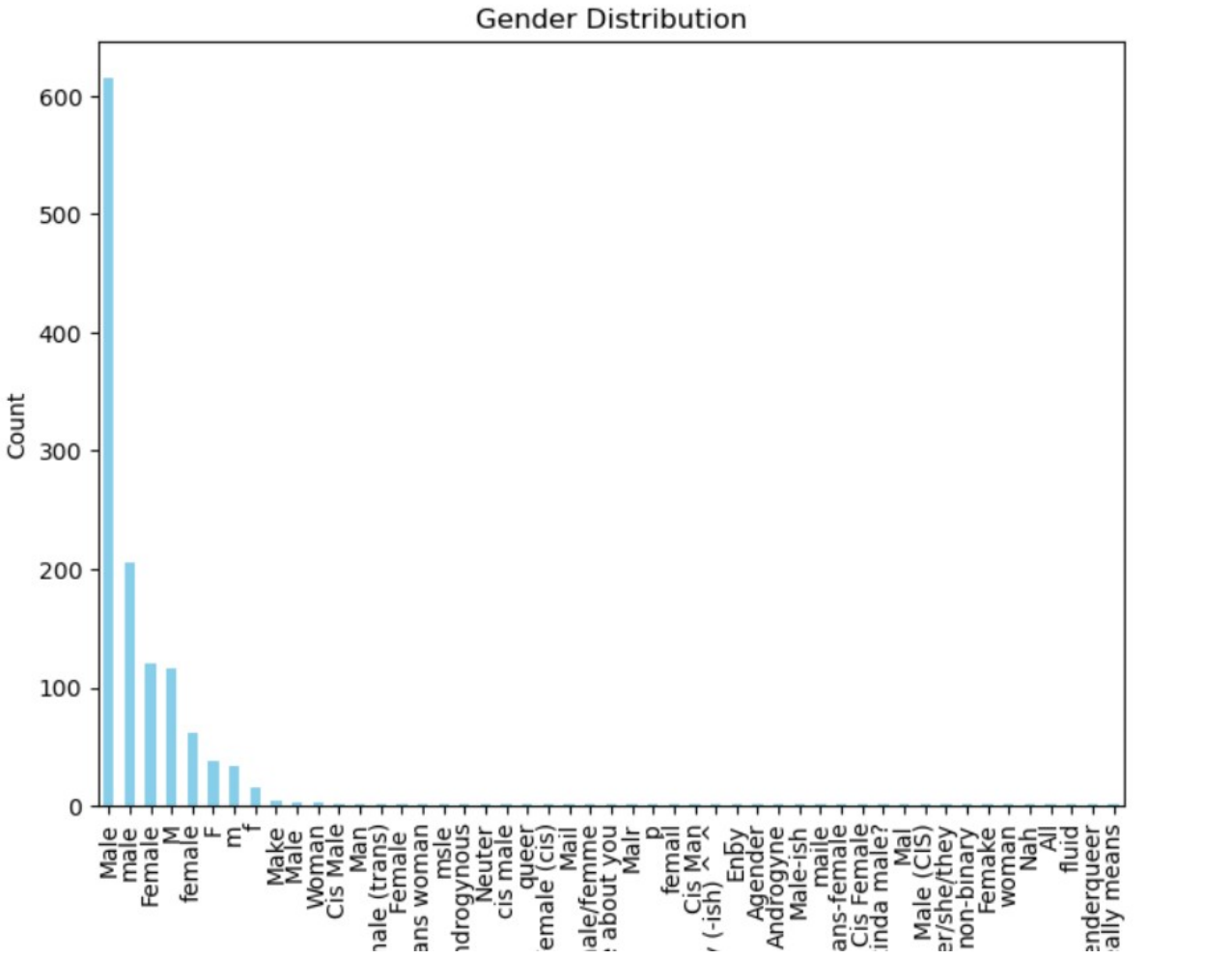
OUTPUT:




```
# Analyze gender distribution
gender_counts = data['Gender'].value_counts()
plt.figure(figsize=(8, 6))
gender_counts.plot(kind='bar', color='skyblue')
plt.title("Gender Distribution")
plt.xlabel("Gender")
plt.ylabel("Count")
```

plt.show()

OUTPUT:



Analyze location distribution

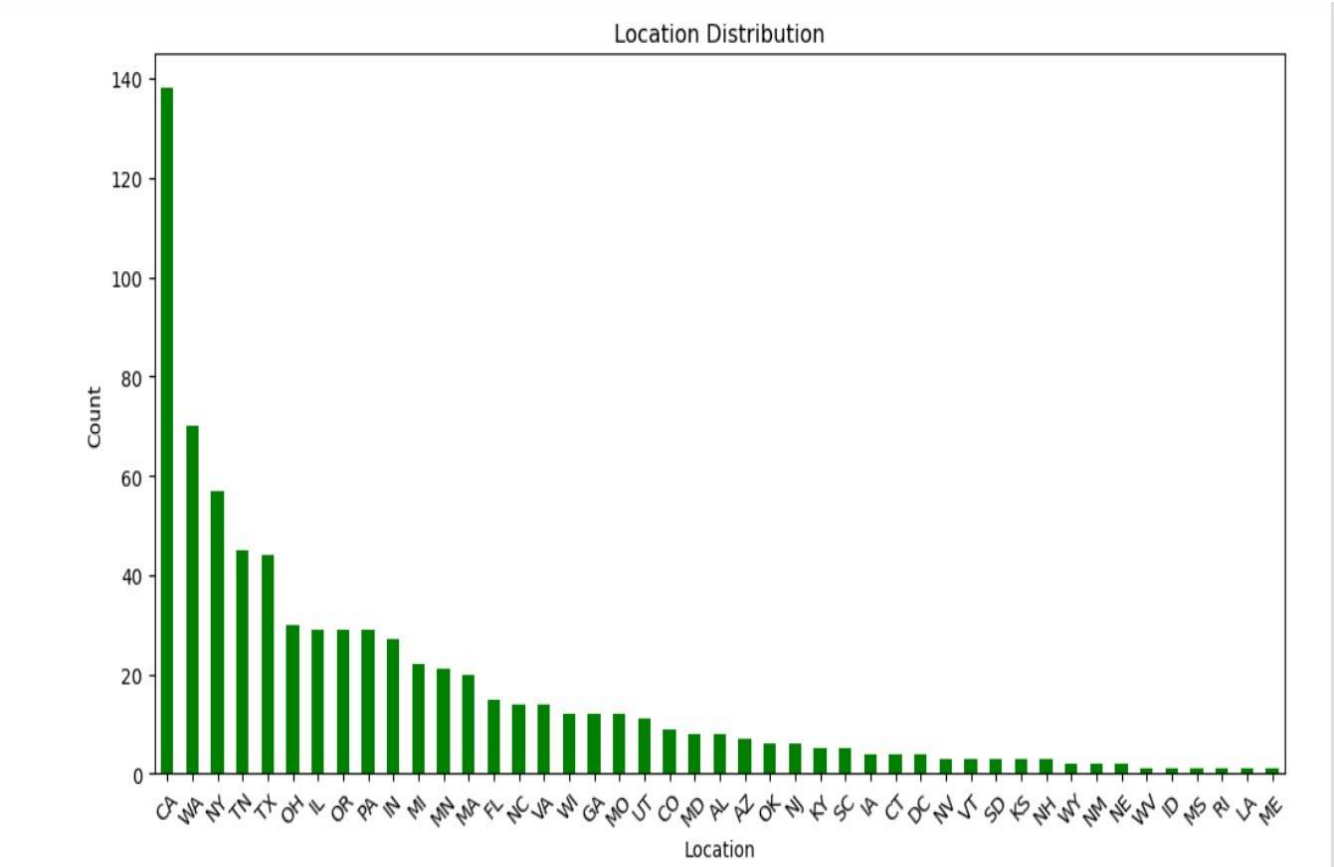
location_counts = data['Location'].value_counts()

plt.figure(figsize=(12, 6))

location_counts.plot(kind='bar', color='green')

```
plt.title("Location Distribution")
plt.xlabel("Location")
plt.ylabel("Count")
plt.xticks(rotation=45)
plt.show()
```

OUTPUT:



Statistical Tests:

```
import pandas as pd
import numpy as np
from scipy import stats
data = pd.read_csv("survey (1).csv")
```

Hypothesis Testing:

One-sample T-Test:

```
sample_data = data['self_employed']  # Replace with your metric of interest
population_mean = 100  # Replace with your known population mean


t_statistic, p_value = stats.ttest_1samp(sample_data, population_mean)
if p_value < 0.05:
    print("The sample mean is significantly different from the population mean.")
else:
    print("There is no significant difference between the sample mean and the population mean.")
```

OUTPUT:

```
There is no significant difference between the sample mean and the populat
ion mean.
```

TWO SAMPLE T-TEST:

```
group1_data = data[data['work_interfere'] == 'Group 1']['work_interfere']  # Replace with your data

group2_data = data[data['self_employed'] == 'Group 2']['self_employed']  # Replace with your data


t_statistic, p_value = stats.ttest_ind(group1_data, group2_data)

if p_value < 0.05:

    print("There is a significant difference between the two groups.")

else:

    print("There is no significant difference between the two groups.")
```

OUTPUT:

```
There is no significant difference between the two groups.
```

CHI-SQUARED TEST:

```
contingency_table = pd.crosstab(data['Country'], data['self_employed'])
chi2, p, dof, expected = stats.chi2_contingency(contingency_table)
if p < 0.05:
    print("The variables are dependent.")
else:
    print("The variables are independent.")
```

OUTPUT:

```
The variables are dependent.
```

```
contingency_table = pd.crosstab(data['state'], data['work_interfere'])
chi2, p, dof, expected = stats.chi2_contingency(contingency_table)
if p < 0.05:
    print("The variables are dependent.")
else:
    print("The variables are independent.")
```

OUTPUT:

```
The variables are independent.
```

ANOVA (Analysis of Variance):

```
group_data = [data[data['Timestamp'] == group]['Age'] for group in data['Timestamp'].unique()]

f_statistic, p_value = stats.f_oneway(*group_data)

if p_value < 0.05:

    print("There is a significant difference between the groups.")

else:

    print("There is no significant difference between the groups.")
```

OUTPUT:

```
There is no significant difference between the groups.
```

Engagement Rates:

```
engagements = 500

total_reach = 10000

# Calculate engagement rate
```

```
engagement_rate = (engagements / total_reach) * 100


# Print the result
print(f"The engagement rate is:
{engagement_rate:.2f}%")
```

OUTPUT:

```
The engagement rate is: 5.00%
```

CONCLUSION:

In conclusion, the visual journey through this report, powered by IBM Cognos, has unraveled insights that transcend numbers. In the realm of data, visualization isn't just a tool, it's the alchemy that turns information into actionable intelligence. Let these visual insights be the compass for future decisions, and may the patterns uncovered here guide us to new heights of understanding and success.