KGiSL Institute of Technology

NAAN MUDHALVAN

PROJECT TITLE :

Public Health Awareness

TEAM MEMBERS:

- Mersheena.F
- Dhiviya Shree.K
- Navin.c
- Surya.S P

PROJECT DESCRIPTION :

PHASE 4 : Development On The Public Health Awareness DataSet

OBJECTIVE :

STEPS :

IN GOOGLE COLAB NOTEBOOK:

- Mount the google drive
- Load the DataSet to the Google ColabNotebook

DESCRIPTIVE ANALYSIS :

- Frequency
- Proposition
- Summary Statistics & Reports
- Corelation
- Graph Representation

DESCRIPTIVE ANALYSIS :

FREQUENCY:

code :

```
# Frequency of treatement :

response_counts = id['treatment'].value_counts()
print(response_counts)
```

output:

PROPOSITION:

code:

```
#Proposition :

response_proportions = id['treatment'].value_counts(normalize=True)
print(response_proportions)
```

output:

SUMMARY STATISTICS :

code :

```
#SUMMARY STATISTICS FOR TREATMENT :
```

```
# summary statistics for "Yes" responses


yes_stats = id[id['treatment'] == 'Yes'].describe()
print(yes_stats)
```

output :

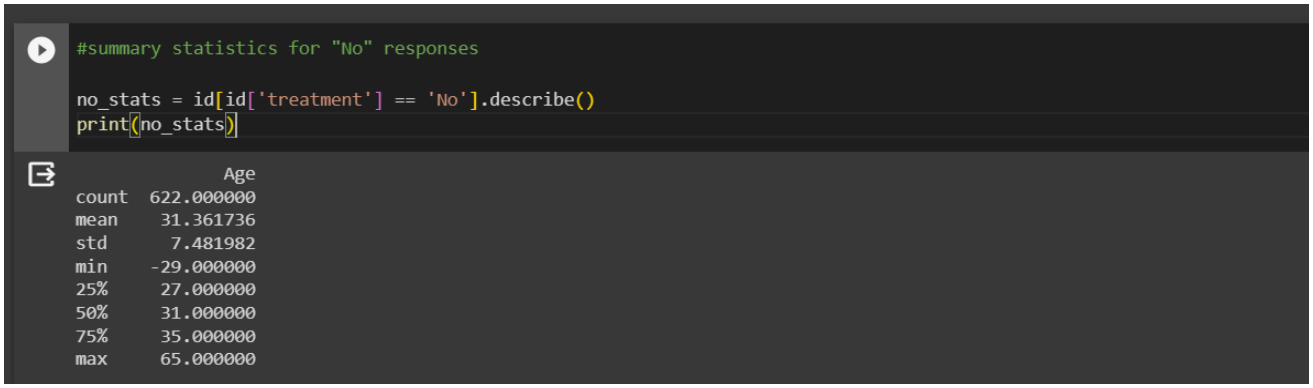
```
#SUMMARY STATISTICS FOR TREATMENT :

# summary statistics for "Yes" responses

yes_stats = id[id['treatment'] == 'Yes'].describe()
print(yes_stats)
```

```
                Age
count  6.370000e+02
mean   1.569859e+08
std    3.962144e+09
min   -1.726000e+03
25%    2.700000e+01
50%    3.200000e+01
75%    3.700000e+01
max    1.000000e+11
```

code:

```
#summary statistics for "No" responses

no_stats = id[id['treatment'] == 'No'].describe()
print(no_stats)
```

output:



```
#summary statistics for "No" responses

no_stats = id[id['treatment'] == 'No'].describe()
print(no_stats)
```

```
              Age
count  622.000000
mean    31.361736
std      7.481982
min    -29.000000
25%     27.000000
50%     31.000000
75%     35.000000
max     65.000000
```

GRAPH REPRESENTATION :

code:

```
#BAR GRAPH :

import matplotlib.pyplot as plt

plt.bar(response_counts.index,response_counts.values,color='skyblue',ed
gecolor='k')
plt.xlabel('Treatment')
plt.ylabel('Count')
plt.title('Distribution of Satisfaction Responses')
plt.show()
```

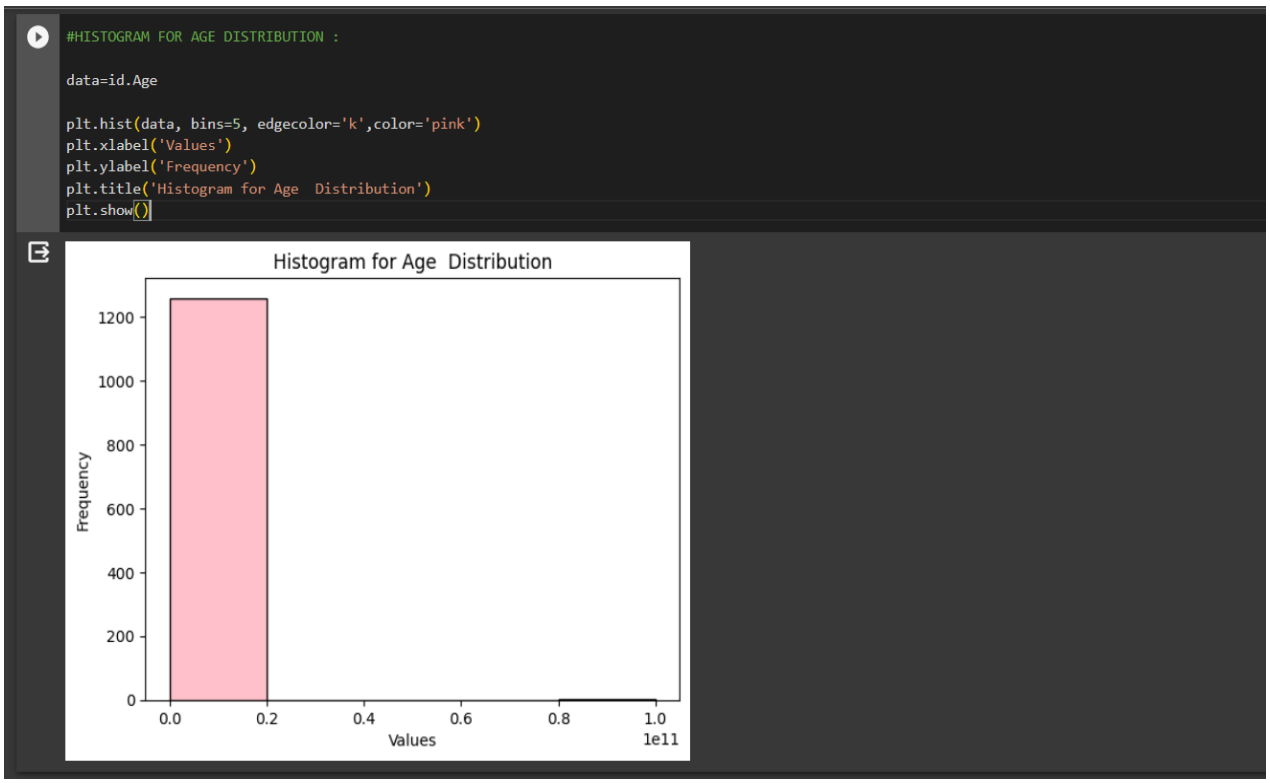
output :

code :

```
#HISTOGRAM FOR AGE DISTRIBUTION :
```

```
data=id.Age

plt.hist(data, bins=5, edgecolor='k',color='pink')
plt.xlabel('Values')
plt.ylabel('Frequency')
plt.title('Histogram for Age  Distribution')
plt.show()
```

output :



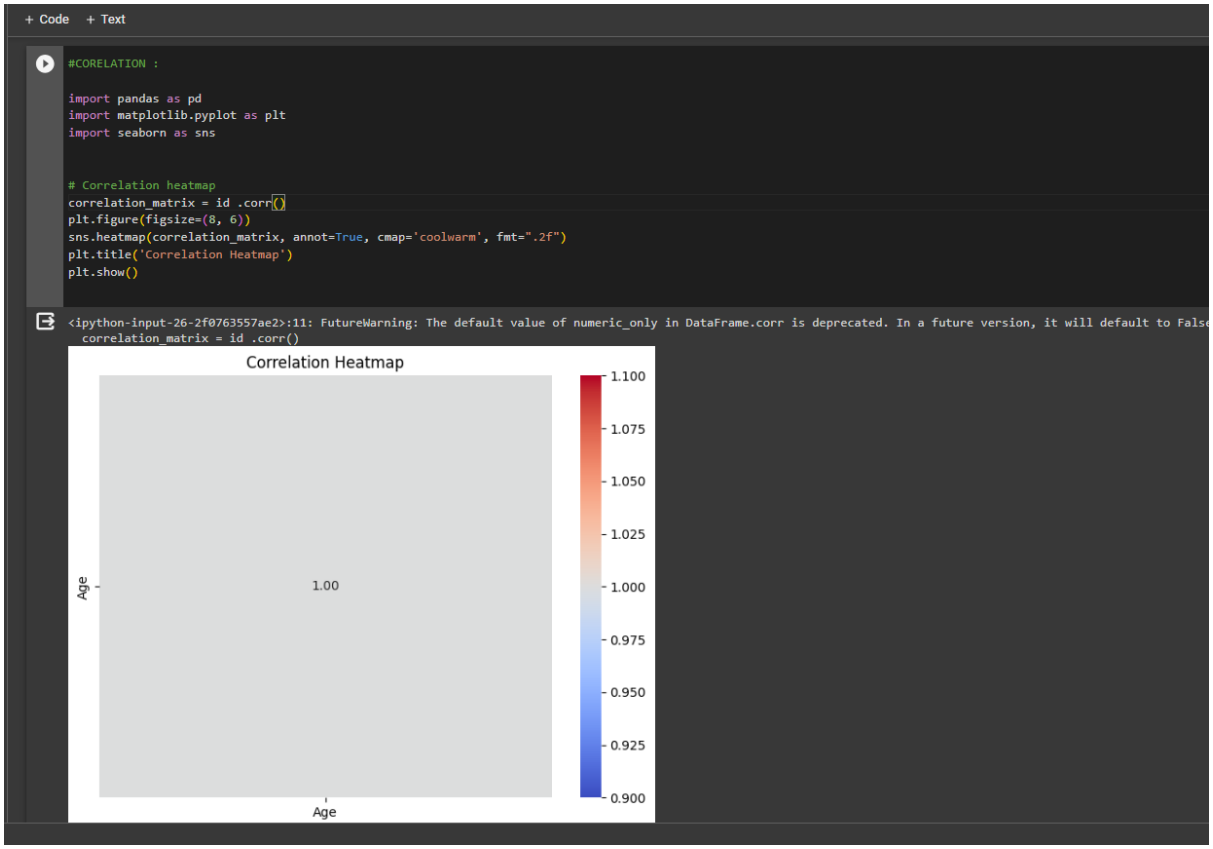
CORELATION :

code:

```
#CORELATION :

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
# Correlation heatmap
correlation_matrix = id .corr()
plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt=".2f")
plt.title('Correlation Heatmap')
plt.show()
```

output :

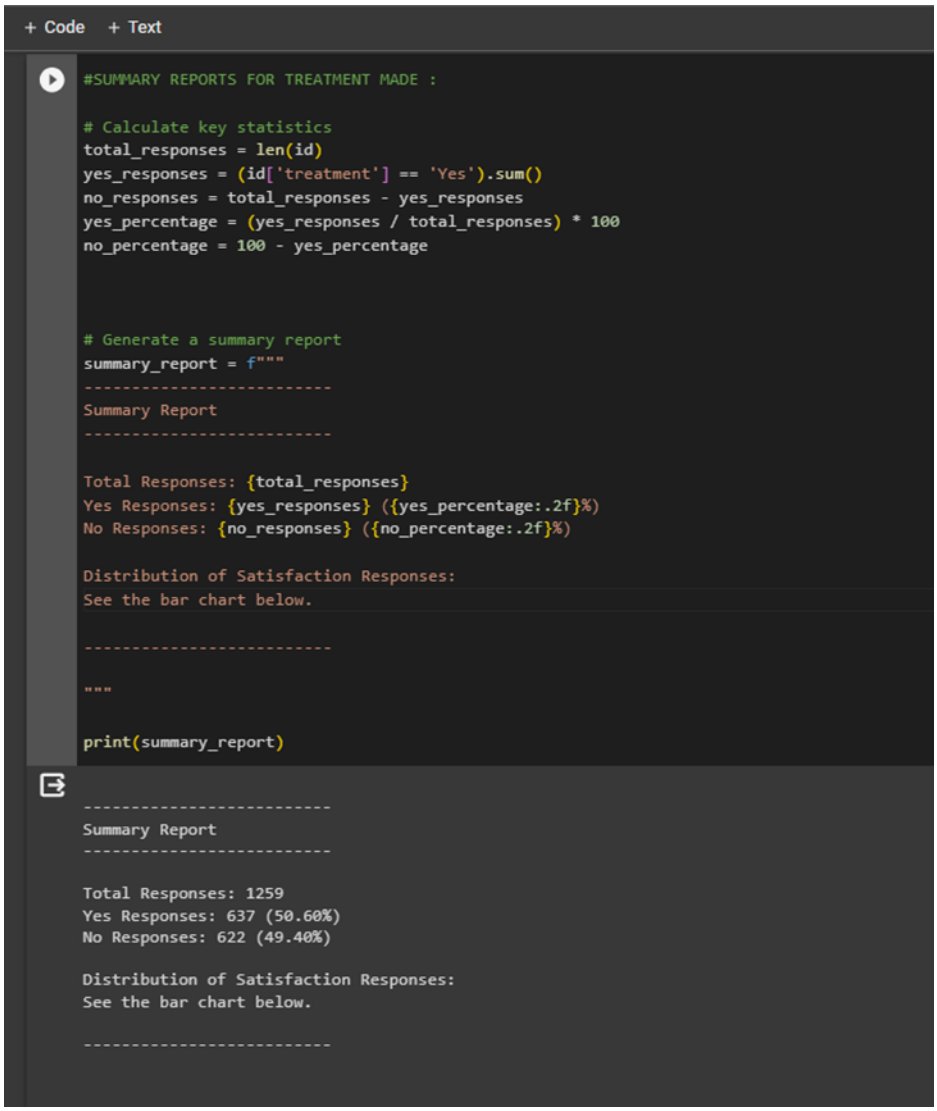


SUMMARY REPORTS:

code:

```
#SUMMARY REPORTS FOR TREATMENT MADE :

# Calculate key statistics
total_responses = len(id)
yes_responses = (id['treatment'] == 'Yes').sum()
no_responses = total_responses - yes_responses
yes_percentage = (yes_responses / total_responses) * 100
no_percentage = 100 - yes_percentage



# Generate a summary report
summary_report = f"""
--------------------------
Summary Report
--------------------------

Total Responses: {total_responses}
Yes Responses: {yes_responses} ({yes_percentage:.2f}%)
No Responses: {no_responses} ({no_percentage:.2f}%)

Distribution of Satisfaction Responses:
See the bar chart below.

--------------------------

"""

print(summary_report)
```

output :



```
+ Code  + Text

#SUMMARY REPORTS FOR TREATMENT MADE :

# Calculate key statistics
total_responses = len(id)
yes_responses = (id['treatment'] == 'Yes').sum()
no_responses = total_responses - yes_responses
yes_percentage = (yes_responses / total_responses) * 100
no_percentage = 100 - yes_percentage



# Generate a summary report
summary_report = f"""
-------------------------
Summary Report
-------------------------

Total Responses: {total_responses}
Yes Responses: {yes_responses} ({yes_percentage:.2f}%)
No Responses: {no_responses} ({no_percentage:.2f}%)

Distribution of Satisfaction Responses:
See the bar chart below.

-------------------------

"""

print(summary_report)
```

```
-------------------------
Summary Report
-------------------------

Total Responses: 1259
Yes Responses: 637 (50.60%)
No Responses: 622 (49.40%)

Distribution of Satisfaction Responses:
See the bar chart below.

-------------------------
```