College of Vocational Studies
University of Delhi

Data analysis and visualization Project file

Topic: Dental Health Analysis

Submitted By

Yash Rana
2k21/CS/114
21013570107

Acknowledgement

I extend my sincere thanks to Mrs. Geetika Vashist, my esteemed professor at the College of Vocational Studies, for her valuable guidance and unwavering support throughout the duration of this project. Her expertise and insights significantly influenced the direction and quality of this research. I also appreciate the College of Vocational Studies for providing an environment conducive to learning and research. The resources and opportunities offered by the college have greatly contributed to the successful completion of this project.

Lastly, I want to acknowledge the hard work and commitment that I put into this project. It has been a challenging yet rewarding journey, and I am proud to have been able to contribute to the field of Computer Science through this research endeavour.

Certificate

This is to certify that the project entitled "Dental Health Analysis" has been successfully completed by Yash Rana under the guidance of Mrs. Geetika Vashist. This project was undertaken as a partial fulfilment of the requirements for Data Analysis and Visualization. The work presented in this project is the original outcome of the research conducted by Yash Rana , and it has not been submitted elsewhere for any academic or professional purpose.

Signature: _____
Mrs. Geetika Vashist

Abstract:

In our contemporary health-conscious society, the importance of maintaining optimal oral hygiene is undeniable. The "Dental Health Analysis" project delves into the intricate interplay between individual habits, choices, and their influence on dental health. Conducted through a comprehensive survey utilizing Google Forms, the project strategically collects and analyzes data, focusing on factors such as toothpaste usage, brushing habits, beverage consumption, sugar intake, and specific dental issues faced by participants.
The primary objectives of the study are to identify prevalent toothpaste preferences, investigate the relationships between brushing habits and dental health, and explore the potential impact of various factors on specific dental issues. The project employs rigorous data analysis techniques, including cross-tabulations and visualizations, to draw meaningful insights from the collected data. Through meticulous analysis, the project aims to contribute essential knowledge to both individuals and dental health professionals. By understanding the intricate relationships between toothpaste choices, brushing practices, and oral diseases, the project seeks to empower individuals to make informed decisions about their oral care routines. Additionally, the findings promise to serve as a valuable resource for dental care providers, public health agencies, and individuals seeking to enhance their dental well-being.

The "Dental Health Analysis" project not only sheds light on current oral hygiene practices but also sets the stage for future research and targeted strategies for promoting and improving overall dental health.

Table Of Content

Introduction:

In an era marked by an increasing emphasis on health and well-being, the significance of maintaining optimal oral hygiene cannot be overstated. The "Dental Health Analysis" project embarks on a comprehensive exploration into the intricate dynamics between individual habits, choices, and their profound impact on dental health. As a crucial aspect of overall well-being, understanding the factors influencing oral health is paramount.

This project employs a meticulous approach to unravel the complexities surrounding dental health, focusing on key determinants such as toothpaste usage, brushing habits, beverage consumption, sugar intake, and the specific dental issues faced by individuals. By utilizing a robust survey methodology facilitated through Google Forms, we aim to capture a diverse and extensive range of data, enabling a nuanced analysis of oral health practices.

The project's overarching objective is to shed light on prevalent toothpaste preferences, examine the relationships between brushing habits and dental health outcomes, and explore how various lifestyle factors contribute to specific dental issues. Through rigorous data analysis techniques, including cross-tabulations and data visualizations, we aspire to derive meaningful insights that go beyond conventional wisdom, providing a deeper understanding of the intricate interplay between oral care practices and dental well-being.

As oral health is a holistic concept that transcends individual choices, this project not only serves to inform individuals about their own oral care routines but also holds the potential to contribute valuable insights to dental care professionals, public health agencies, and policymakers. By bridging the gap in understanding between habits, choices, and dental health outcomes, this project lays the foundation for informed decision-making and the development of targeted strategies to enhance overall dental health.

In the following sections, we delve into the methodology, data collection process, and detailed analysis, aiming to present a comprehensive picture of the intricate relationships shaping our oral health landscape.

Objective:

The "Dental Health Analysis" project aims to achieve the following objectives:

1. Toothpaste Preference:

- Identify and analyse the most commonly used toothpaste brands and types among participants.

2. Brushing Habits:

- Investigate the frequency of brushing habits, including daily routines and the times of day when brushing occurs.
- Explore the number of times individuals brush their teeth each day.

3. Dental Disease Assessment:

- Gather data on the prevalence of oral diseases, including cavities, gum disease, tooth sensitivity, and other related issues.

4. Factors Affecting Dental Health:

- Examine the potential associations between toothpaste preferences, brushing habits, beverage consumption, sugar intake, and specific dental issues faced by participants.

5. Data Collection:

- Utilize Google Forms to collect a comprehensive dataset, ensuring precision and efficiency in gathering a diverse range of information.

6. Data Visualization:

- Create visually appealing and informative graphs and charts to present clear insights into the data, facilitating a deeper understanding of the relationships between variables.

7. Outcome Analysis:

- Identify the most commonly reported dental issues and assess their correlation with toothpaste preferences, brushing habits, and other lifestyle factors.

8. Adherence to Oral Hygiene Practices:

- Understand whether participants adhere to recommended oral hygiene practices and identify areas for improvement.

9. Impact of Beverage Consumption:

- Analyse the impact of hot and cold beverage consumption on dental health outcomes.

The ultimate goal of the project is to contribute valuable insights into the complex relationships between individual habits, choices, and their influence on dental health. By achieving these objectives, the project aims to empower individuals to make informed decisions about their oral care routines and provide valuable information for dental care professionals, public health agencies, and individuals seeking to enhance their overall dental well-being.

Data Collection And Survey Design:

Overview

The survey conducted for the "Dental Health Analysis" project was designed to gather comprehensive information on participants' oral health practices and preferences. The survey focused on key areas such as toothpaste usage, brushing habits, dental issues, and lifestyle factors. The goal was to collect data that would enable a detailed analysis of the relationships between these variables and provide valuable insights into oral health. Done with the help of Google form

Survey Questions and Categories

1. Participant Information

- Name:(Text Response)
- Age:(Text Response)
- Gender:(Dropdown Menu)
- Phone Number:(Text Response)
- Region:(Text Response)

2. Toothpaste Usage

- What brand of toothpaste do you currently use?(Dropdown Menu)
- How satisfied are you with your current toothpaste?(Dropdown Menu)
- Are you satisfied with the toothpaste you use ?(Dropdown Menu)
- Which type of brush are you using ?(Dropdown Menu)

3. Brushing Habits

- How frequently do you brush your teeth in a day?(Dropdown Menu)
- When do you usually brush your teeth?(Checkbox)
- How much time do you spend on average during one brushing session?(Dropdown Menu)

4. Dental Health Assessment

- Do you experience any of the following dental issues currently?(Checkbox)
- Have you ever diagnose with any problem related to teeth ?(Dropdown Menu)

5. Lifestyle Factors

- Do you regularly consume hot or cold beverages?(Radio Buttons)
- Do you drink tea or coffee ?(Radio Buttons)
- How would you rate your sugar consumption ?(Dropdown Menu)

Participation and Consent

Participants were informed about the purpose of the survey and the confidentiality of their responses. They provided voluntary consent to participate in the survey, and their personal information was handled in accordance with privacy guidelines.

1.Data Processing:

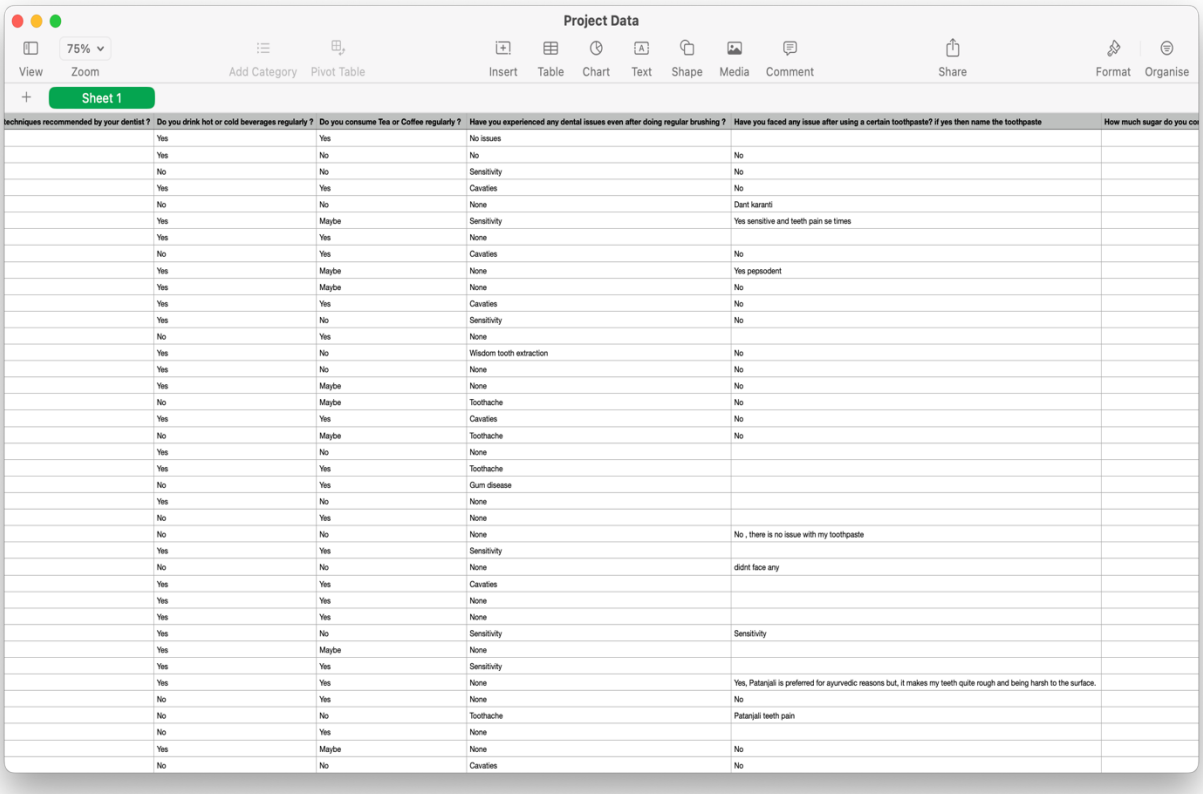
- Import data set.
- Fill the missing values with mean of column in which null values is present.
- Removing any kind of spaces which cause any problem in getting data.
- Removing Duplicates.

CODE:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import csv

data = pd.read_csv("Project Data.csv")
df= pd.DataFrame(data)
print(df)

df=df.drop(columns=['Timestamp'])
df=df.drop(columns=['Is there is any gum and teeth related diseases have been there in your family trade  '])

columns_with_missing_values = df.columns[df.isnull().any()].tolist()

for column in columns_with_missing_values:
    if df[column].dtype == 'float64':
        df[column].fillna(df[column].mean(), inplace=True)

for column in columns_with_missing_values:
    if df[column].dtype == 'object':
        df[column].fillna(df[column].mode().iloc[0], inplace=True)

for column in df.columns:
    if df[column].dtype == 'object':  # Check if the column contains object/string data
        df[column] = df[column].str.upper()

for column in df.columns:
    if df[column].dtype == 'object':  # Check if the column contains object/string data
        df[column] = df[column].str.strip()

df.to_csv('Project Data2.csv', index=False)
```

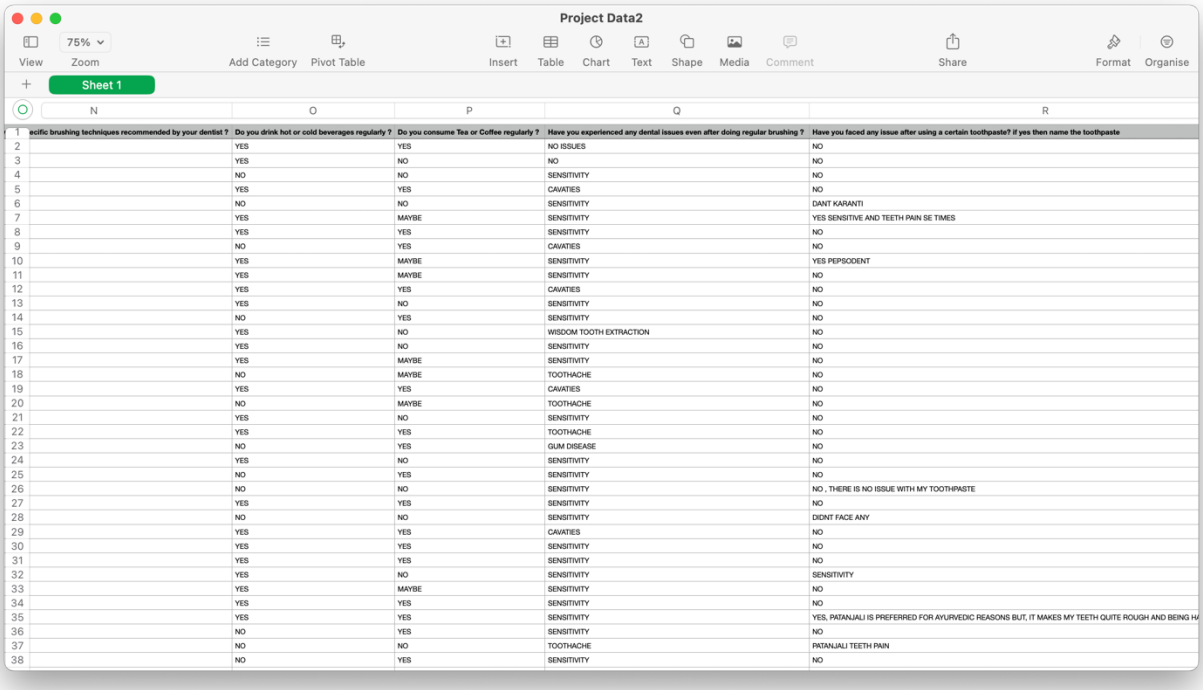
Data Set before and after cleaning :

- **Before**:



echniques recommended by your dentist ?	Do you drink hot or cold beverages regularly ?	Do you consume Tea or Coffee regularly ?	Have you experienced any dental issues even after doing regular brushing ?	Have you faced any issue after using a certain toothpaste? if yes then name the toothpaste	How much sugar do you co
	Yes	Yes	No issues		
	Yes	No	No	No	
	No	No	Sensitivity	No	
	Yes	Yes	Cavaties	No	
	No	No	None	Dant karanti	
	Yes	Maybe	Sensitivity	Yes sensitive and teeth pain se times	
	Yes	Yes	None		
	No	Yes	Cavaties	No	
	Yes	Maybe	None	Yes pepsodent	
	Yes	Maybe	None	No	
	Yes	Yes	Cavaties	No	
	Yes	No	Sensitivity	No	
	No	Yes	None		
	Yes	No	Wisdom tooth extraction	No	
	Yes	No	None	No	
	Yes	Maybe	None	No	
	No	Maybe	Toothache	No	
	Yes	Yes	Cavaties	No	
	No	Maybe	Toothache	No	
	Yes	No	None		
	Yes	Yes	Toothache		
	No	Yes	Gum disease		
	Yes	No	None		
	No	Yes	None		
	No	No	None	No , there is no issue with my toothpaste	
	Yes	Yes	Sensitivity		
	No	No	None	didnt face any	
	Yes	Yes	Cavaties		
	Yes	Yes	None		
	Yes	Yes	None		
	Yes	No	Sensitivity	Sensitivity	
	Yes	Maybe	None		
	Yes	Yes	Sensitivity		
	Yes	Yes	None	Yes, Patanjali is preferred for ayurvedic reasons but, it makes my teeth quite rough and being harsh to the surface.	
	No	Yes	None	No	
	No	No	Toothache	Patanjali teeth pain	
	No	Yes	None		
	Yes	Maybe	None	No	
	No	No	Cavaties	No	

- **After**



	N	O	P	Q	R
1	ecific brushing techniques recommended by your dentist ?	Do you drink hot or cold beverages regularly ?	Do you consume Tea or Coffee regularly ?	Have you experienced any dental issues even after doing regular brushing ?	Have you faced any issue after using a certain toothpaste? if yes then name the toothpaste
2		YES	YES	NO ISSUES	NO
3		YES	NO	NO	NO
4		NO	NO	SENSITIVITY	NO
5		YES	YES	CAVATIES	NO
6		NO	NO	SENSITIVITY	DANT KARANTI
7		YES	MAYBE	SENSITIVITY	YES SENSITIVE AND TEETH PAIN SE TIMES
8		YES	YES	SENSITIVITY	NO
9		NO	YES	CAVATIES	NO
10		YES	MAYBE	SENSITIVITY	YES PEPSODENT
11		YES	MAYBE	SENSITIVITY	NO
12		YES	YES	CAVATIES	NO
13		YES	NO	SENSITIVITY	NO
14		NO	YES	SENSITIVITY	NO
15		YES	NO	WISDOM TOOTH EXTRACTION	NO
16		YES	NO	SENSITIVITY	NO
17		YES	MAYBE	SENSITIVITY	NO
18		NO	MAYBE	TOOTHACHE	NO
19		YES	YES	CAVATIES	NO
20		NO	MAYBE	TOOTHACHE	NO
21		YES	NO	SENSITIVITY	NO
22		YES	YES	TOOTHACHE	NO
23		NO	YES	GUM DISEASE	NO
24		YES	NO	SENSITIVITY	NO
25		NO	YES	SENSITIVITY	NO
26		NO	NO	SENSITIVITY	NO , THERE IS NO ISSUE WITH MY TOOTHPASTE
27		YES	YES	SENSITIVITY	NO
28		NO	NO	SENSITIVITY	DIDNT FACE ANY
29		YES	YES	CAVATIES	NO
30		YES	YES	SENSITIVITY	NO
31		YES	YES	SENSITIVITY	NO
32		YES	NO	SENSITIVITY	SENSITIVITY
33		YES	MAYBE	SENSITIVITY	NO
34		YES	YES	SENSITIVITY	NO
35		YES	YES	SENSITIVITY	YES, PATANJALI IS PREFERRED FOR AYURVEDIC REASONS BUT, IT MAKES MY TEETH QUITE ROUGH AND BEING HA
36		NO	YES	SENSITIVITY	NO
37		NO	NO	TOOTHACHE	PATANJALI TEETH PAIN
38		NO	YES	SENSITIVITY	NO

2.Data Analysis

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import csv

df = pd.read_csv('Project Data2.csv')

columns_of_interest = [
 'Which brand of toothpaste do you commonly use ?',
 'Which factor influence your choice of toothpaste ?',
 'Are you satisfied with the toothpaste you use ?',
 'Have you switched toothpaste brands in the past year? if yes then why?',
 'For how long you brush your teeth ?',
 'Which type of brush are you using ?',
 'How often do you brush your teeth in a typical day ?',
 'Do you use any additional oral care product such as mouthwash or dental floss ?',
 'Do you follow any specific brushing techniques recommended by your dentist ?',
 'Do you drink hot or cold beverages regularly ?',
 'Do you consume Tea or Coffee regularly ? ',
 'Have you experienced any dental issues even after doing regular brushing ? ',
 'Have you faced any issue after using a certain toothpaste? if yes then name the
toothpaste ',
 'How much sugar do you consume daily ?',
 'Have you ever diagnose with any problem related to teeth ?',
 'Do you consume any of these Tobacco, alcohol and smoking ?']


for column_name in columns_of_interest:
    unique_values = df[column_name].unique()
    print(f"Column: {column_name}")
    for value in unique_values:
        value_count = (df[column_name] == value).sum()
        total_values = len(df[column_name])
        percentage = (value_count / total_values) * 100
        print(f"Percentage of '{value}' occurrences: {percentage:.2f}%")
    print()
```

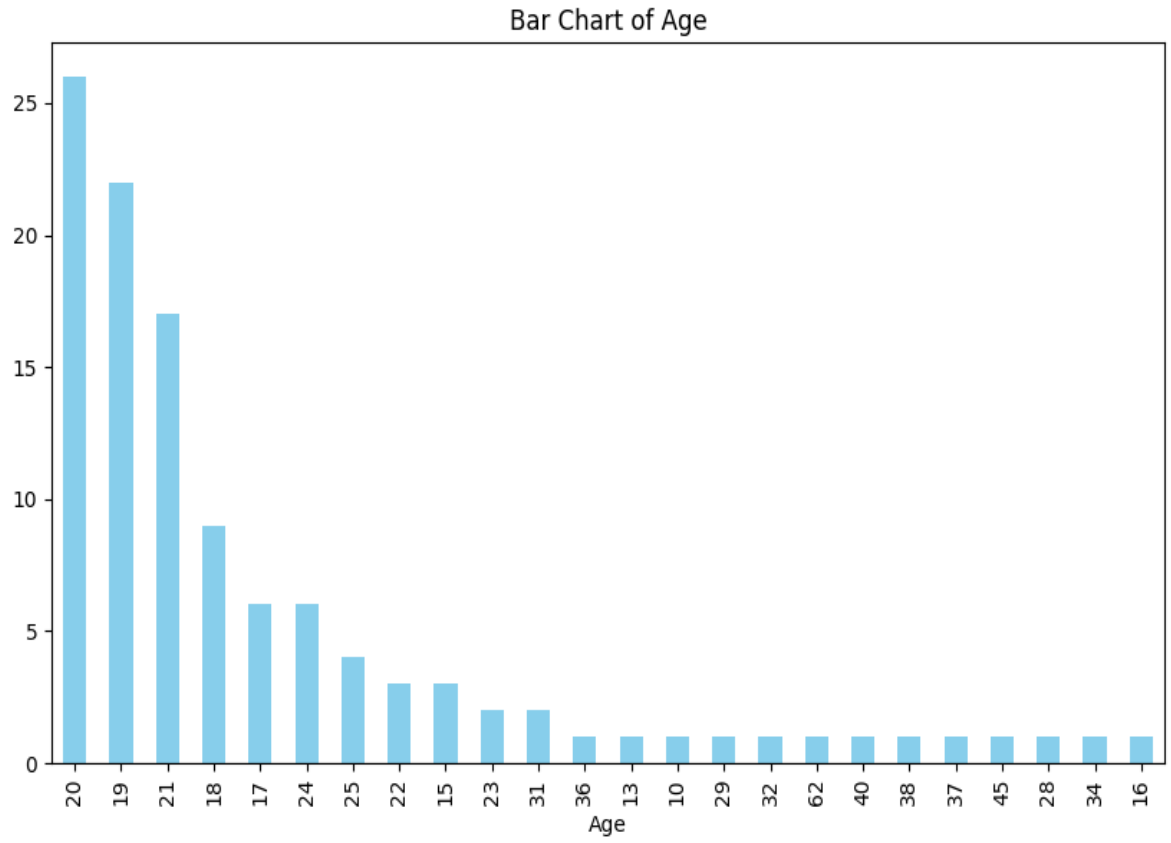
```
Percentage of 'DIDNT FACE ANY' occurrences: 0.88%
Percentage of 'SENSITIVITY' occurrences: 0.88%
Percentage of 'YES, PATANJALI IS PREFERRED FOR AYURVEDIC REASONS BUT, IT MAKES MY TEETH QUITE ROUGH
Percentage of 'PATANJALI TEETH PAIN' occurrences: 0.88%
Percentage of 'NAH' occurrences: 0.88%
Percentage of 'SENSODYNE' occurrences: 0.88%
Percentage of 'NOPE' occurrences: 0.88%
Percentage of 'COLGET' occurrences: 0.88%
Percentage of 'AYURVEDIC.' occurrences: 0.88%
Percentage of 'AFTER USING IT BUT STILL FEELING SENSITIVE' occurrences: 0.88%
Percentage of 'DABUR TASTE BUDS GONE FOR TOO LONG' occurrences: 0.88%
Percentage of 'SOME TIME SENSITIVITY ISSUE' occurrences: 0.88%
Percentage of 'nan' occurrences: 0.00%
Percentage of 'NO BUT THERE'S A CAVITY IN MY TEETH THAT HURTS SOMETIMES' occurrences: 0.88%
Percentage of 'COLGATE' occurrences: 0.88%
Percentage of 'YES BABOOL IS NOT GOOD FOR MY TEETH' occurrences: 0.88%
Percentage of 'SEVERELY CRACKED, DRY LIPS. PATANJALI' occurrences: 0.88%

Column: How much sugar do you consume daily ?
Percentage of 'AVERAGE' occurrences: 92.04%
Percentage of 'ABOVE AVERAGE' occurrences: 7.08%
Percentage of 'TOO MUCH' occurrences: 0.88%

Column: Have you ever diagnose with any problem related to teeth ?
Percentage of 'DON'T KNOW' occurrences: 77.88%
Percentage of 'NO PROBLEM' occurrences: 0.88%
Percentage of 'SWALLOW GUM' occurrences: 3.54%
Percentage of 'SOMETIMES GUM BLEEDING' occurrences: 0.88%
Percentage of 'BLEEDING GUM' occurrences: 7.96%
Percentage of 'CALCIUM DEIFICATE' occurrences: 2.65%
Percentage of 'CAVITY' occurrences: 1.77%
Percentage of 'PROTUDING AND MISALIGNED' occurrences: 0.88%
Percentage of 'DIABETIC' occurrences: 1.77%
Percentage of 'PAIN IN GUMS' occurrences: 0.88%
Percentage of 'NO PROBLEM YET' occurrences: 0.88%

Column: Do you consume any of these Tobacco, alcohol and smoking ?
Percentage of 'NO' occurrences: 92.04%
Percentage of 'YES' occurrences: 5.31%
Percentage of 'PREFER NOT TO SAY' occurrences: 2.65%

yashrana@Ranas-MacBook-Air CVS %
```

```
/usr/local/bin/python3 /Users/yashrana/Desktop/CVS/percent.py
yashrana@Ranas-MacBook-Air CVS % /usr/local/bin/python3 /Users/yashrana/Desktop/CVS/percent.py
Column: Which brand of toothpaste do you commonly use ?
Percentage of 'NUTRAVEDIC' occurrences: 0.88%
Percentage of 'COLGATE' occurrences: 38.94%
Percentage of 'PATANJALI' occurrences: 17.70%
Percentage of 'COLGATE, PEPSODENT, SENSODYNE, DABUR, PATANJALI' occurrences: 0.88%
Percentage of 'COLGATE, CLOSE-UP' occurrences: 1.77%
Percentage of 'CLOSE-UP' occurrences: 1.77%
Percentage of 'COLGATE, PEPSODENT' occurrences: 3.54%
Percentage of 'SENSODYNE' occurrences: 7.96%
Percentage of 'COLGATE, PEPSODENT, SENSODYNE, PATANJALI' occurrences: 0.88%
Percentage of 'COLGATE, PEPSODENT, PATANJALI' occurrences: 0.88%
Percentage of 'VESTIGE DENT ASSURE' occurrences: 0.88%
Percentage of 'COLGATE, SENSODYNE, PATANJALI' occurrences: 1.77%
Percentage of 'COLGATE, SENSODYNE' occurrences: 2.65%
Percentage of 'DABUR' occurrences: 6.19%
Percentage of 'DABUR, PATANJALI' occurrences: 0.88%
Percentage of 'COLGATE, PEPSODENT, SENSODYNE, CLOSE-UP, PATANJALI, HIMALAYA' occurrences: 0.88%
Percentage of 'COLGATE, CLOSE-UP, PATANJALI' occurrences: 1.77%
Percentage of 'COLGATE, PEPSODENT, CLOSE-UP' occurrences: 1.77%
Percentage of 'COLGATE, DABUR' occurrences: 1.77%
Percentage of 'MODI CARE' occurrences: 0.88%
Percentage of 'COLGATE, HIMALAYA,' occurrences: 0.88%
Percentage of 'PEPSODENT' occurrences: 2.65%
Percentage of 'AMWAY GLISTER' occurrences: 0.88%
Percentage of 'COLGATE, DABUR, PATANJALI' occurrences: 0.88%

Column: Which factor influence your choice of toothpaste ?
Percentage of 'PRICE' occurrences: 10.62%
Percentage of 'FAMILY' occurrences: 48.67%
Percentage of 'AYURVEDIC' occurrences: 22.12%
Percentage of 'DENTIST' occurrences: 7.08%
Percentage of 'SENSITIVITY' occurrences: 8.85%
Percentage of 'QUALITY' occurrences: 0.88%
Percentage of 'TASTE' occurrences: 1.77%

Column: Are you satisfied with the toothpaste you use ?
Percentage of 'AVERAGE' occurrences: 30.09%
Percentage of 'VERY SATISFIED' occurrences: 20.35%
Percentage of 'SATISFIED' occurrences: 49.56%
```

3.Visualization of data set using pandas :

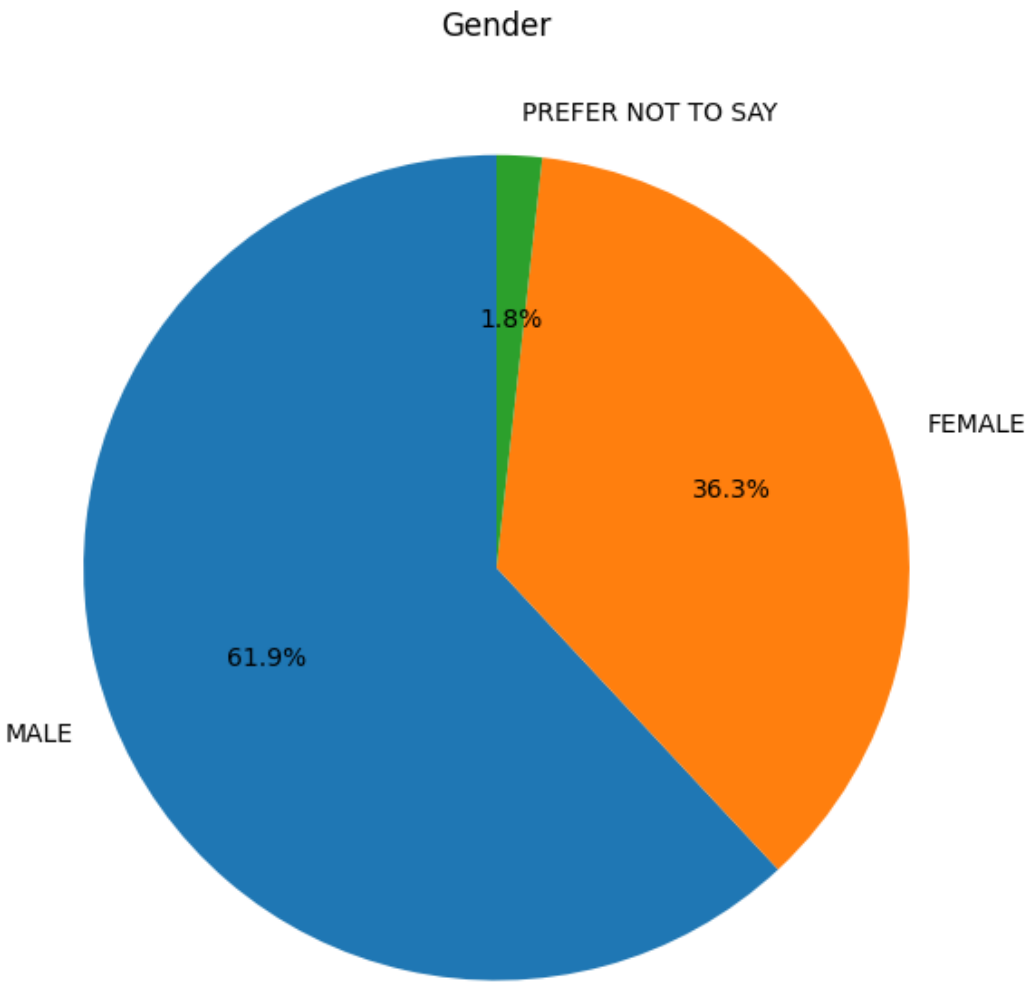
Data Set 1: Age

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import csv
df = pd.read_csv('Project Data2.csv')
####################'Age'#########################

value_counts = df['Age '].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Bar Chart of Age')
plt.xlabel('Age')
plt.ylabel(' ')
plt.show()
```



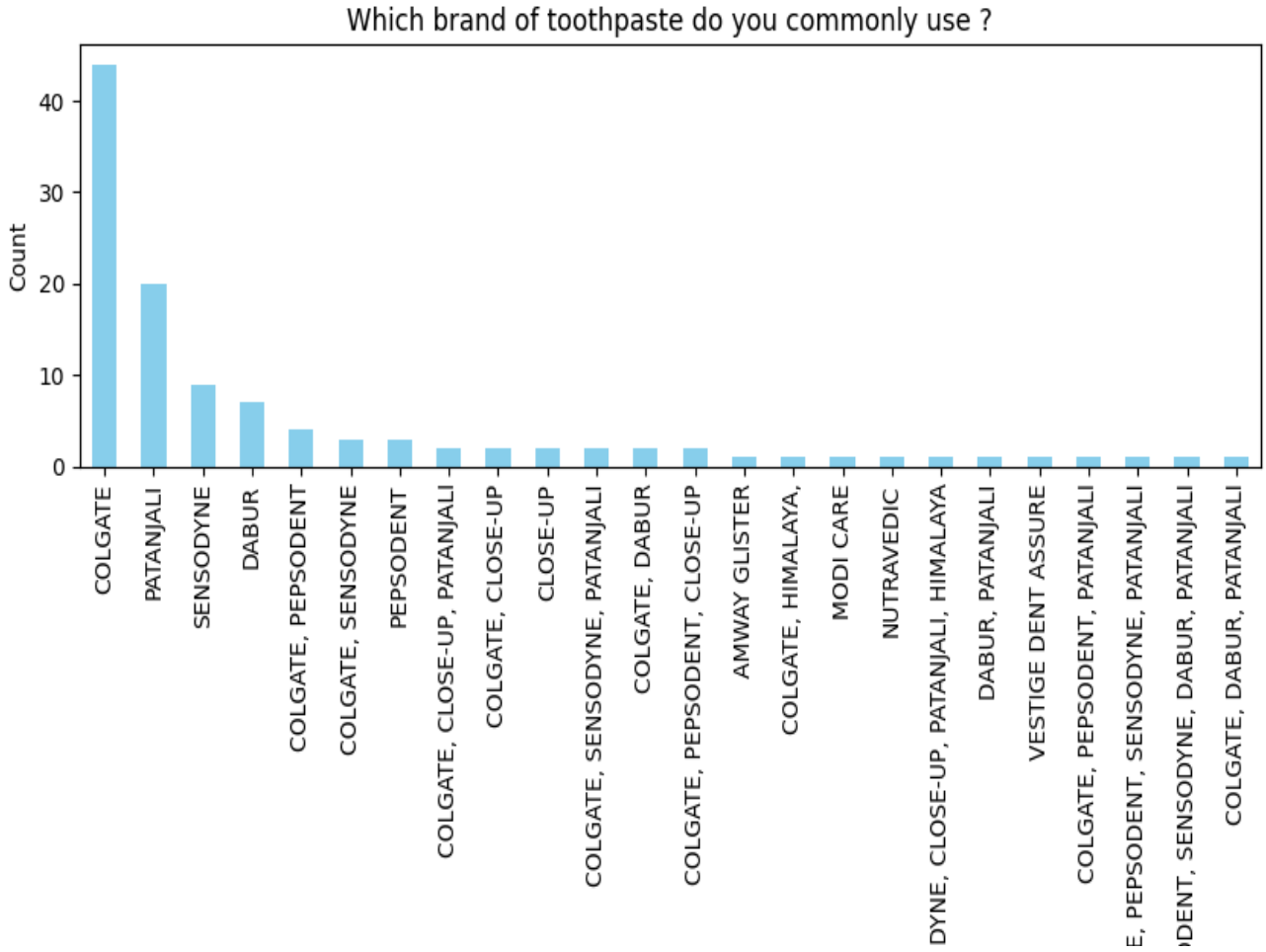
Data Set 2: Gender

```
df = pd.read_csv('Project Data2.csv')
value_counts = df['Gender '].value_counts()
plt.figure(figsize=(8, 8))
plt.pie(value_counts, labels=value_counts.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Gender')
plt.show()
```



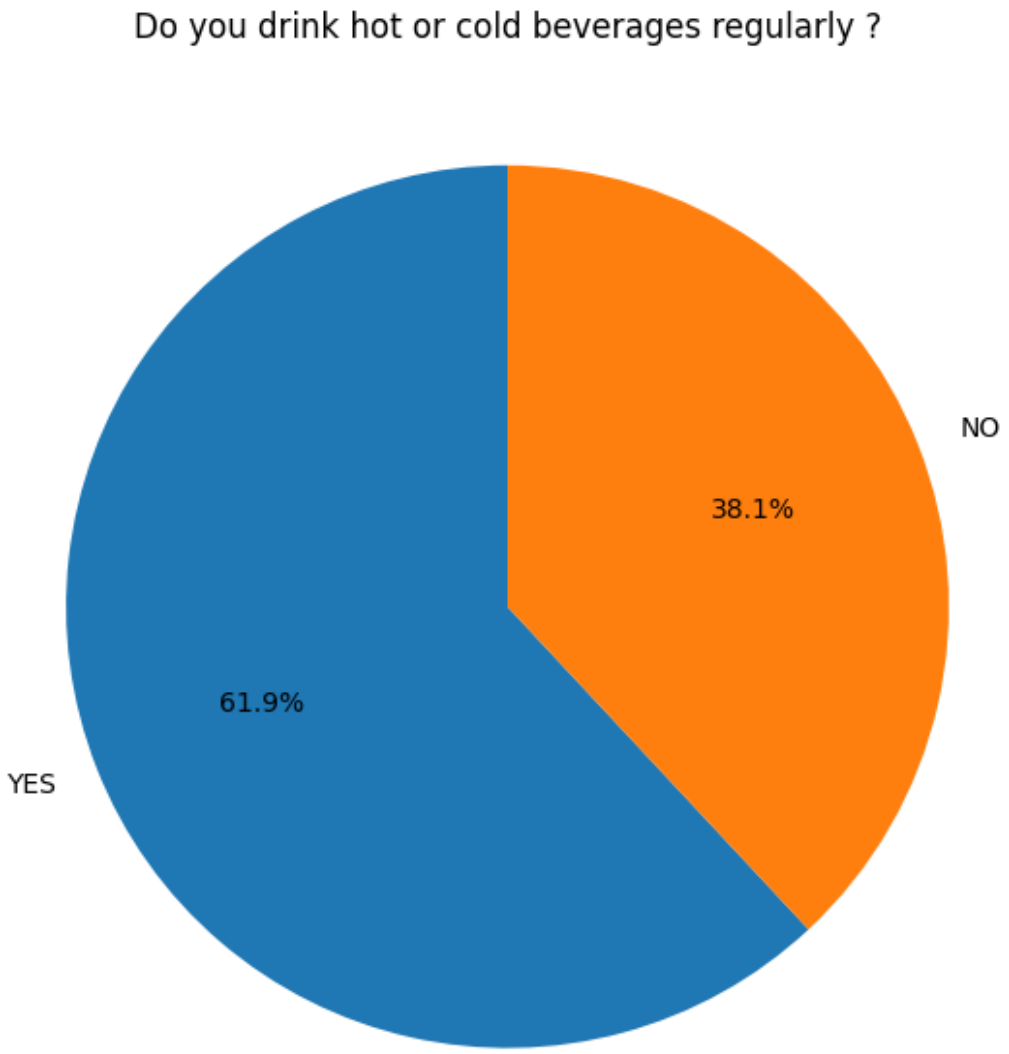
Data Set 3: Which brand of toothpaste do you commonly use?

```
value_counts = df['Which brand of toothpaste do you commonly use
?'].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Which brand of toothpaste do you commonly use ?')
plt.xlabel('Brand')
plt.ylabel('Count')
plt.show()
```



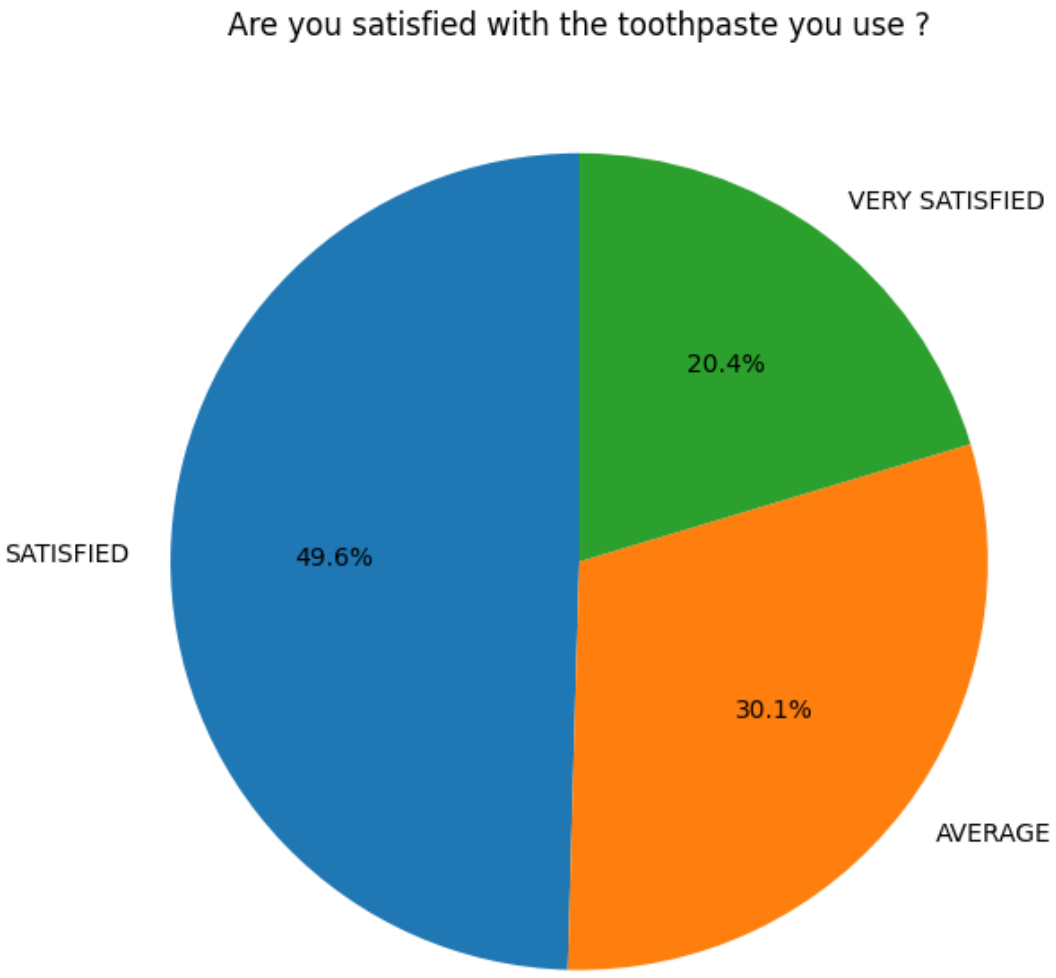
Data Set 4: Do you drink hot or cold beverages regularly?

```
value_counts = df['Do you drink hot or cold beverages regularly ?'].value_counts()
plt.figure(figsize=(8, 8))
plt.pie(value_counts, labels=value_counts.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Do you drink hot or cold beverages regularly ?')
plt.show()
```



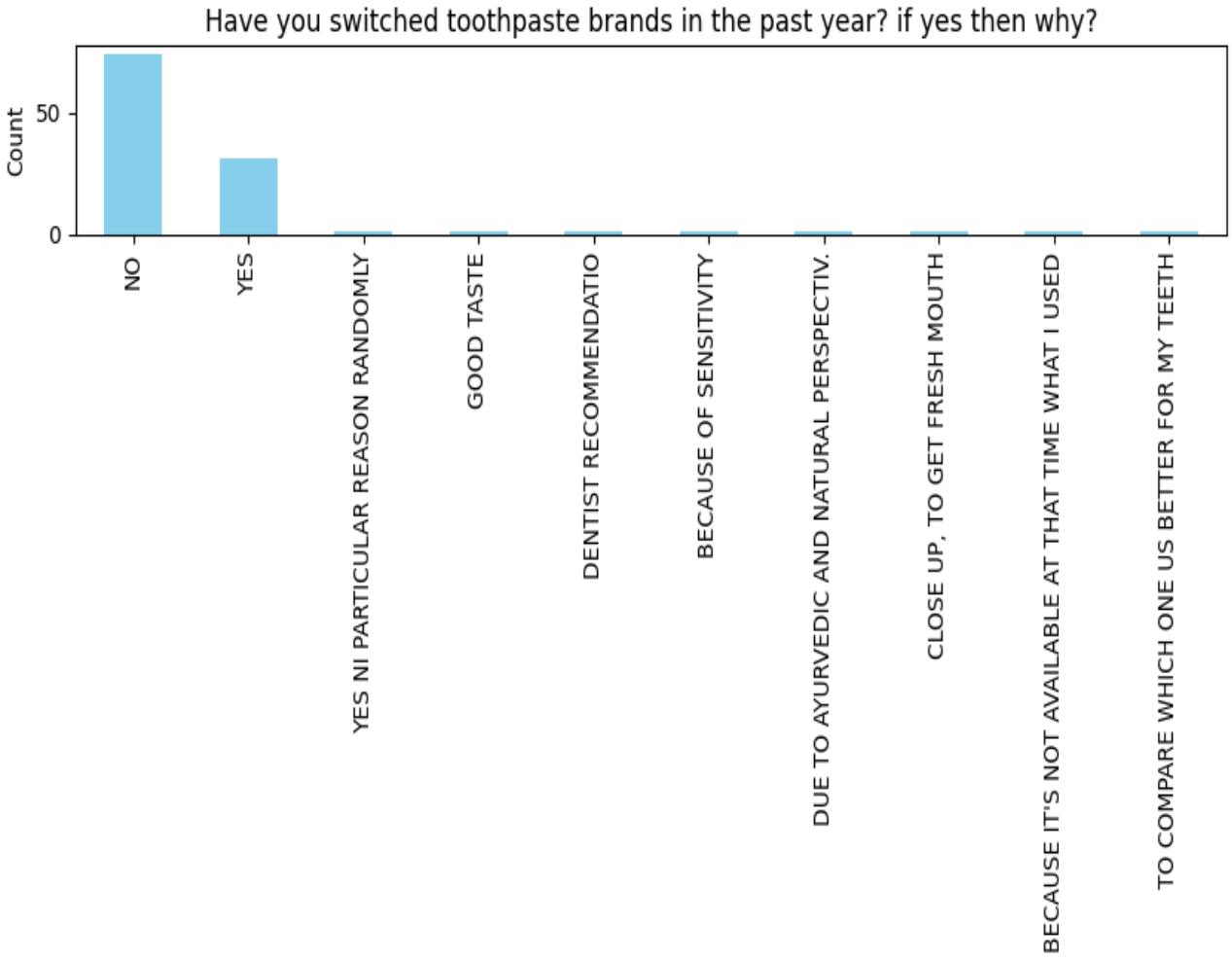
Data Set 5: 'Are you satisfied with the toothpaste you use ?

```
value_counts = df['Are you satisfied with the toothpaste you use ?'].value_counts()
plt.figure(figsize=(8, 8))
plt.pie(value_counts, labels=value_counts.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Are you satisfied with the toothpaste you use ?')
plt.show()
```



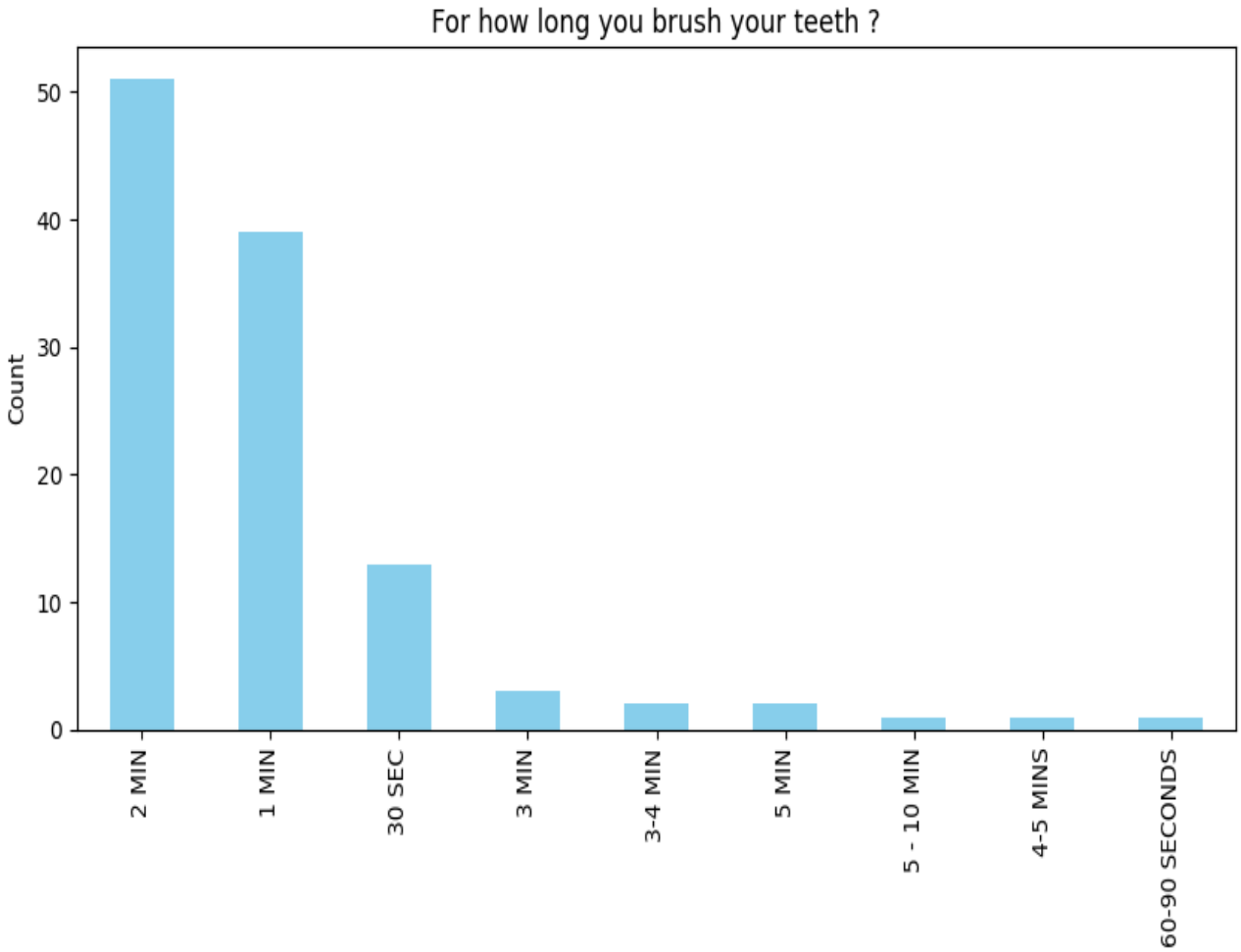
Data Set 6: Have you switched toothpaste brands in the past year? if yes then why?

```
value_counts = df['Have you switched toothpaste brands in the past year? if yes then why?'].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Have you switched toothpaste brands in the past year? if yes then why?')
plt.xlabel('')
plt.ylabel('Count')
plt.show()
```



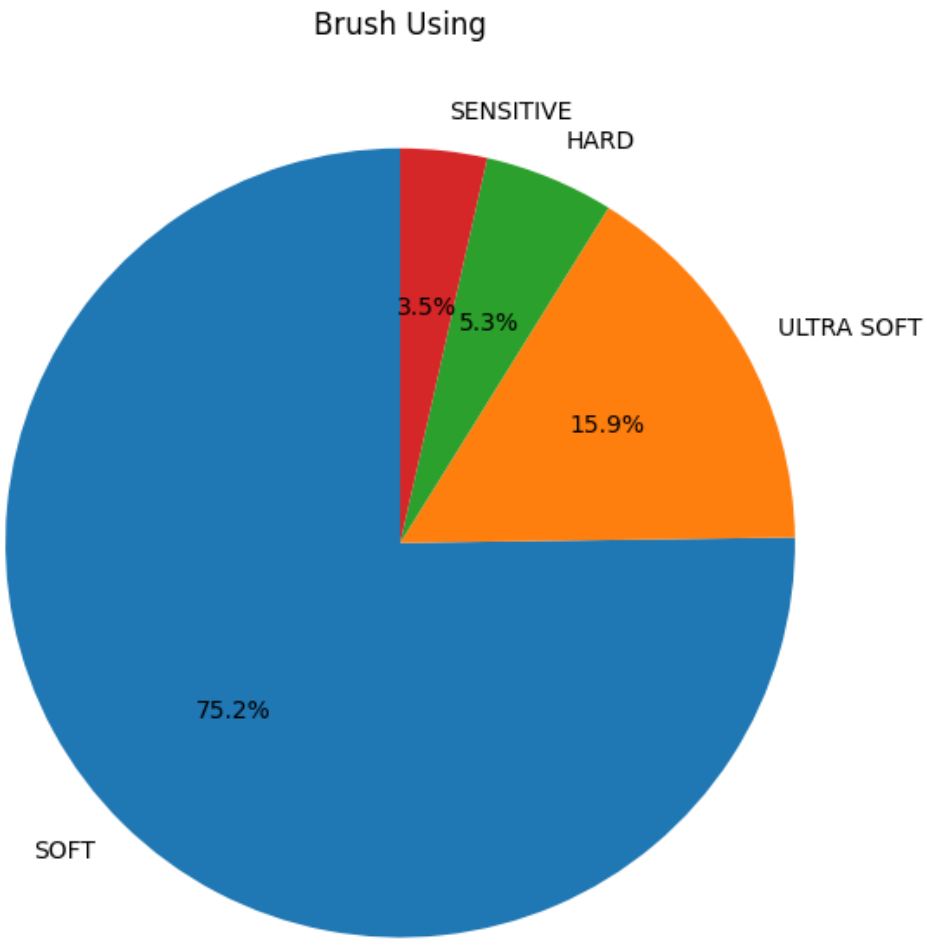
Data Set 7: For how long you brush your teeth?

```
value_counts = df['For how long you brush your teeth ?'].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'For how long you brush your teeth ?')
plt.xlabel('')
plt.ylabel('Count')
plt.show()
```



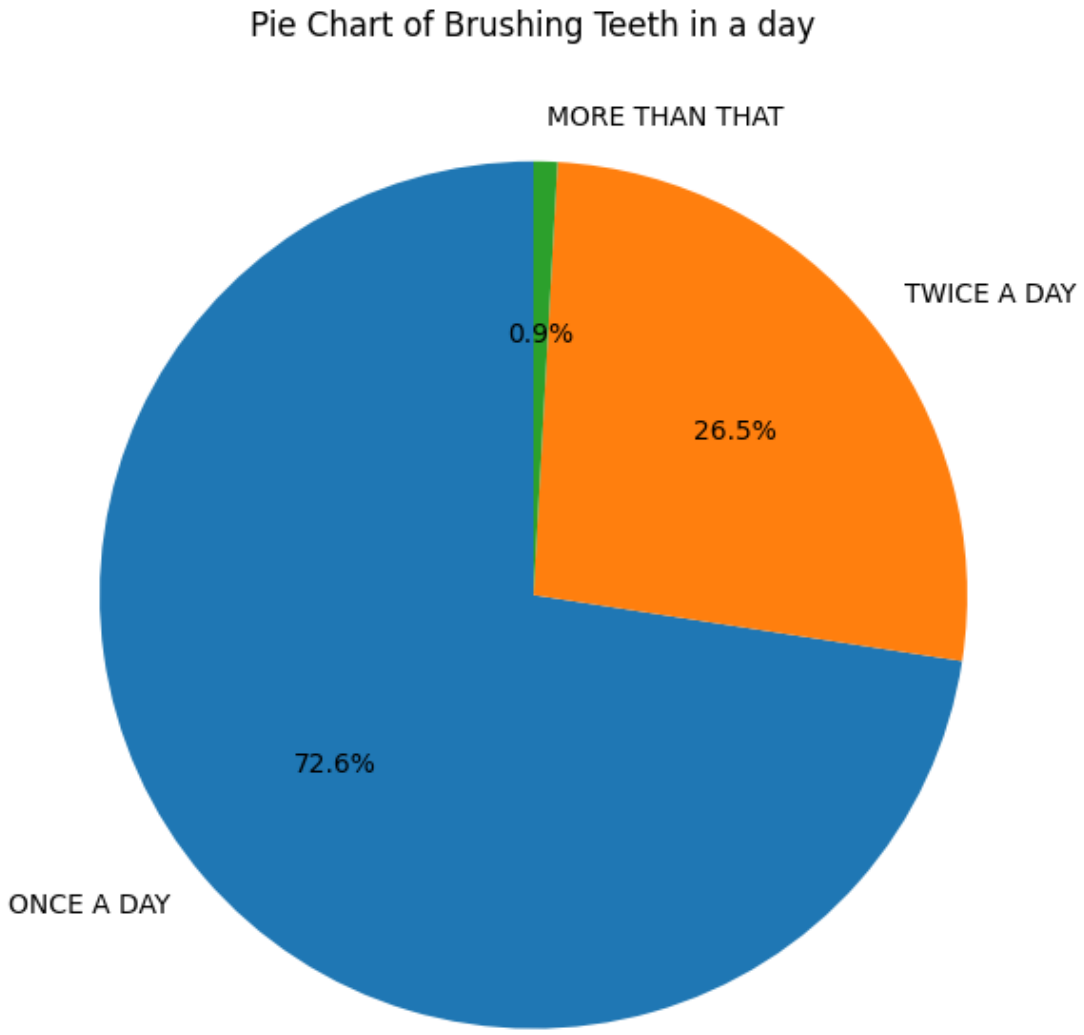
Data Set 8: 'Which type of brush are you using?

```
value_counts = df['Which type of brush are you using ?'].value_counts()
plt.figure(figsize=(8, 8))
plt.pie(value_counts, labels=value_counts.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Brush Using')
plt.show()
```



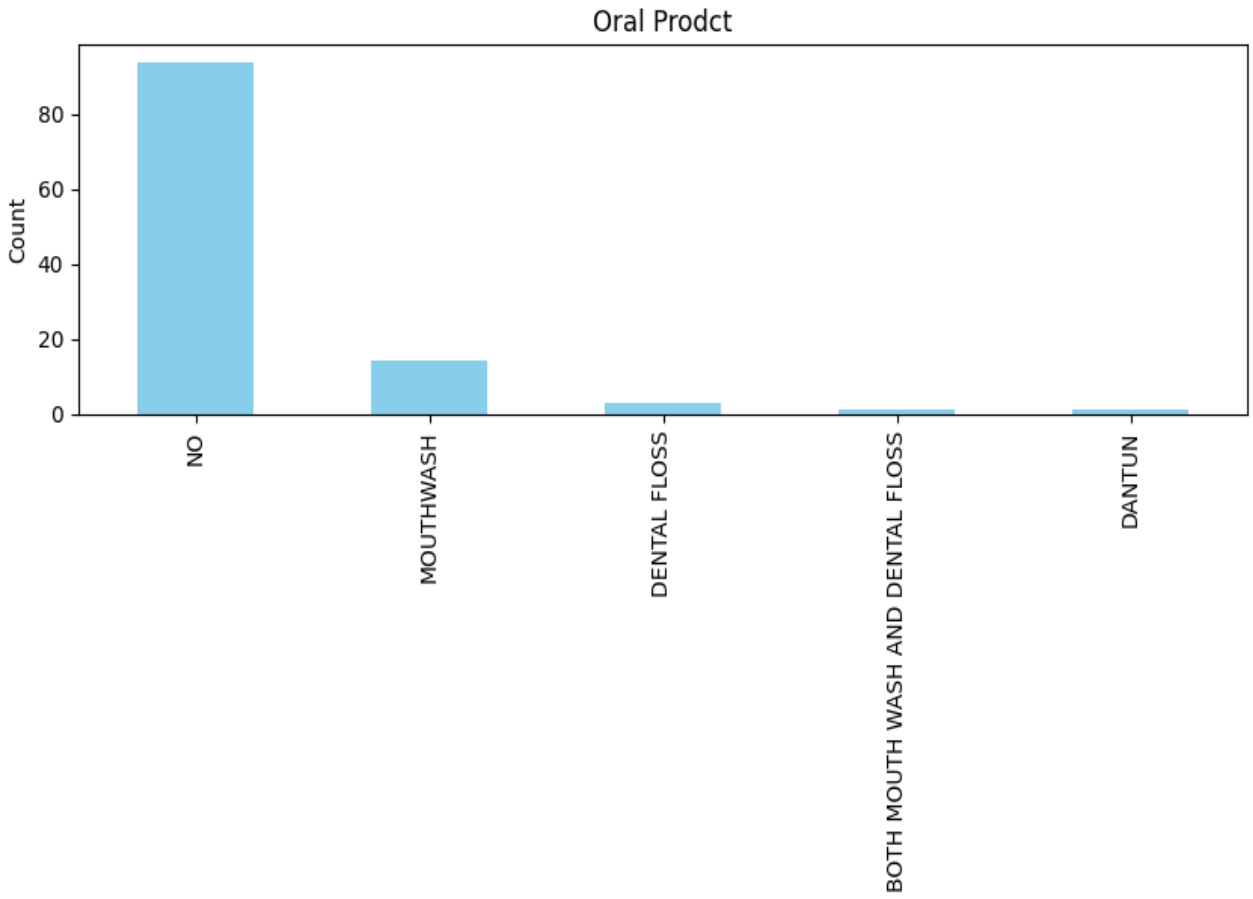
Data Set 9: How often do you brush your teeth in a typical day?

```
value_counts = df['How often do you brush your teeth in a typical day
?'].value_counts()
plt.figure(figsize=(8, 8))
plt.pie(value_counts, labels=value_counts.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Pie Chart of Brushing Teeth in a day')
plt.show()
```



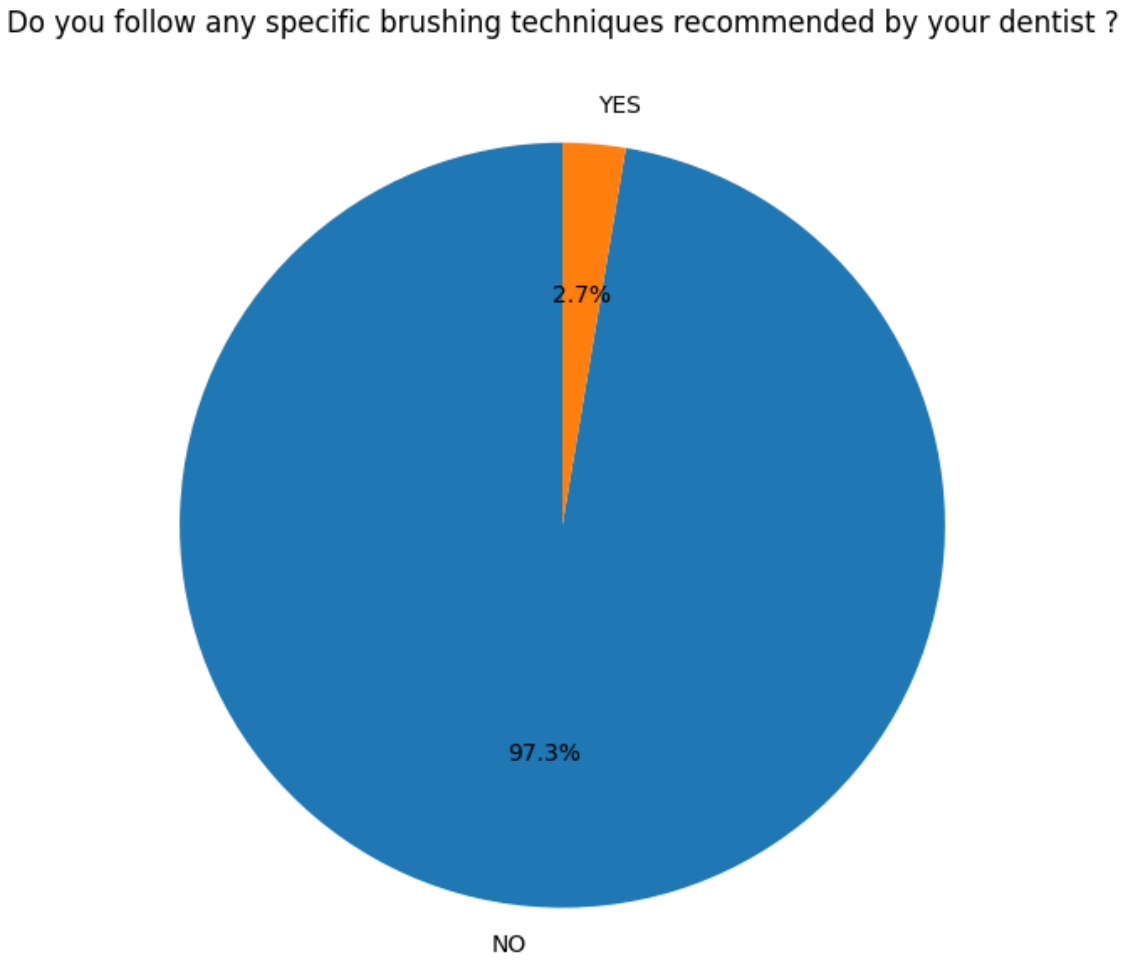
Data set 10: 'Do you use any additional oral care product such as mouthwash or dental floss?

```
value_counts = df['Do you use any additional oral care product such as mouthwash or dental floss ?'].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Oral Prodct')
plt.xlabel('')
plt.ylabel('Count')
plt.show()
```



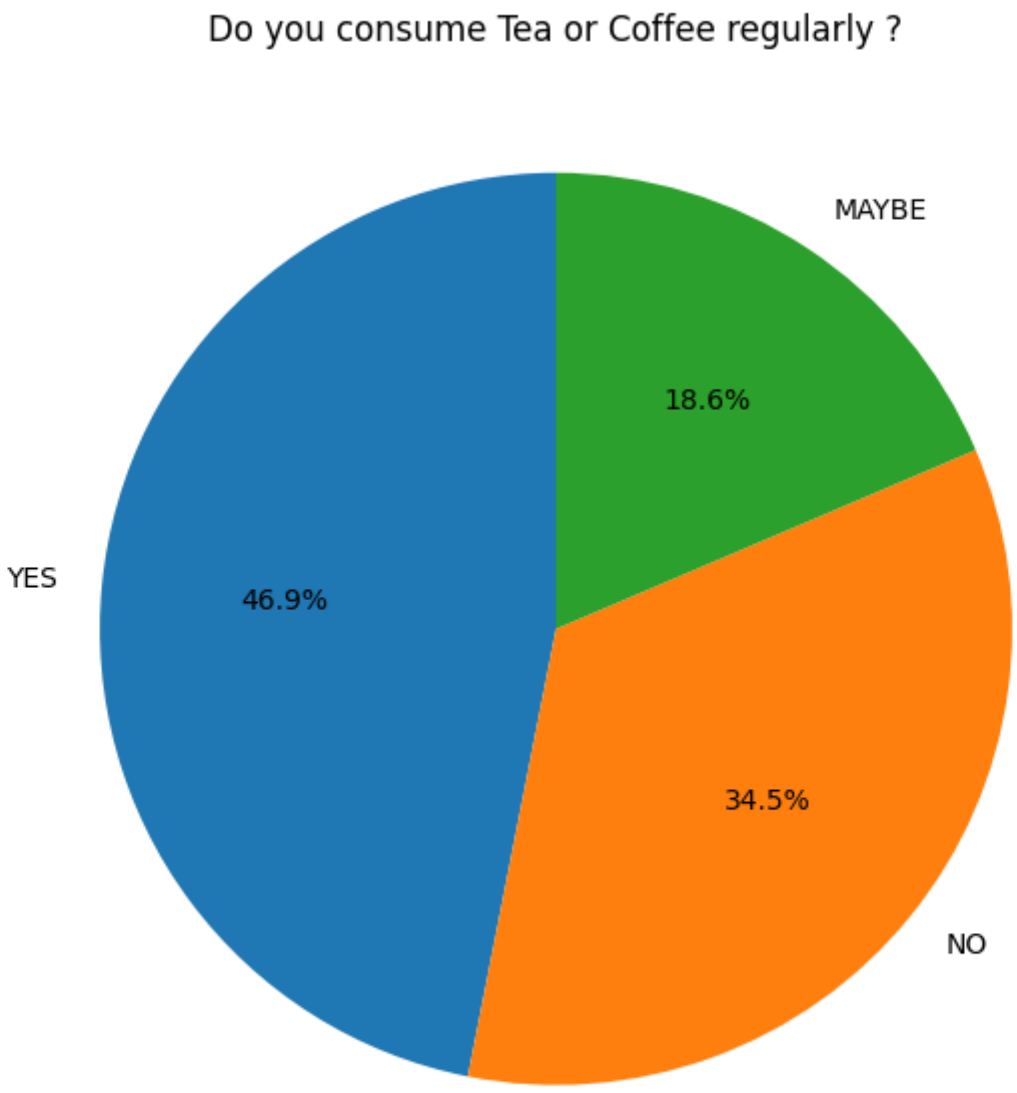
Data Set 11: Do you follow any specific brushing techniques recommended by your dentist?

```
value_counts = df['Do you follow any specific brushing techniques recommended by your dentist ?'].value_counts()
plt.figure(figsize=(8, 8))
plt.pie(value_counts, labels=value_counts.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Do you follow any specific brushing techniques recommended by your dentist ?')
plt.show()
```



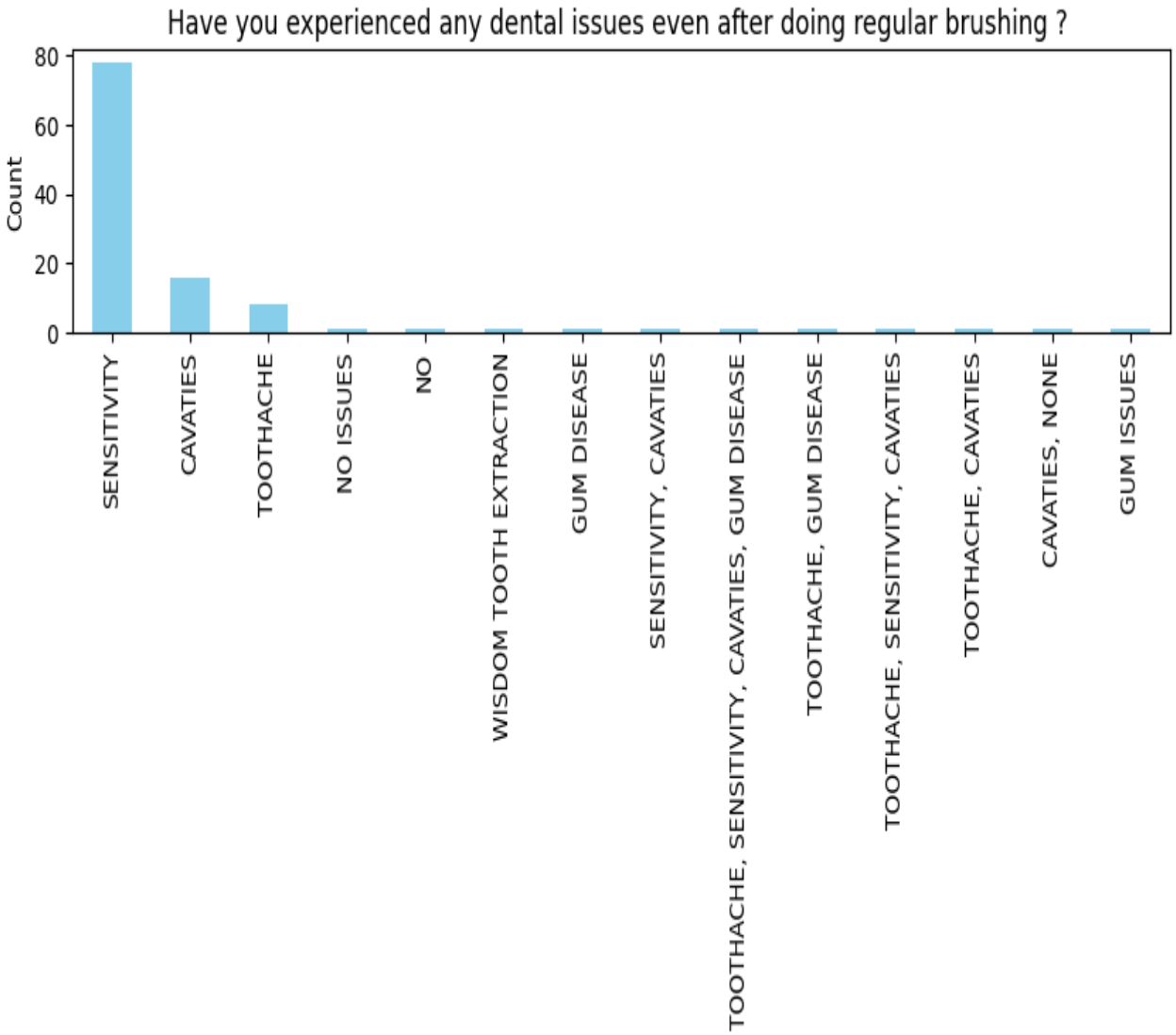
Data set 12: Do you consume Tea or Coffee regularly?

```
value_counts = df['Do you consume Tea or Coffee regularly ? '].value_counts()
plt.figure(figsize=(8, 8))
plt.pie(value_counts, labels=value_counts.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Do you consume Tea or Coffee regularly ?')
plt.show()
```



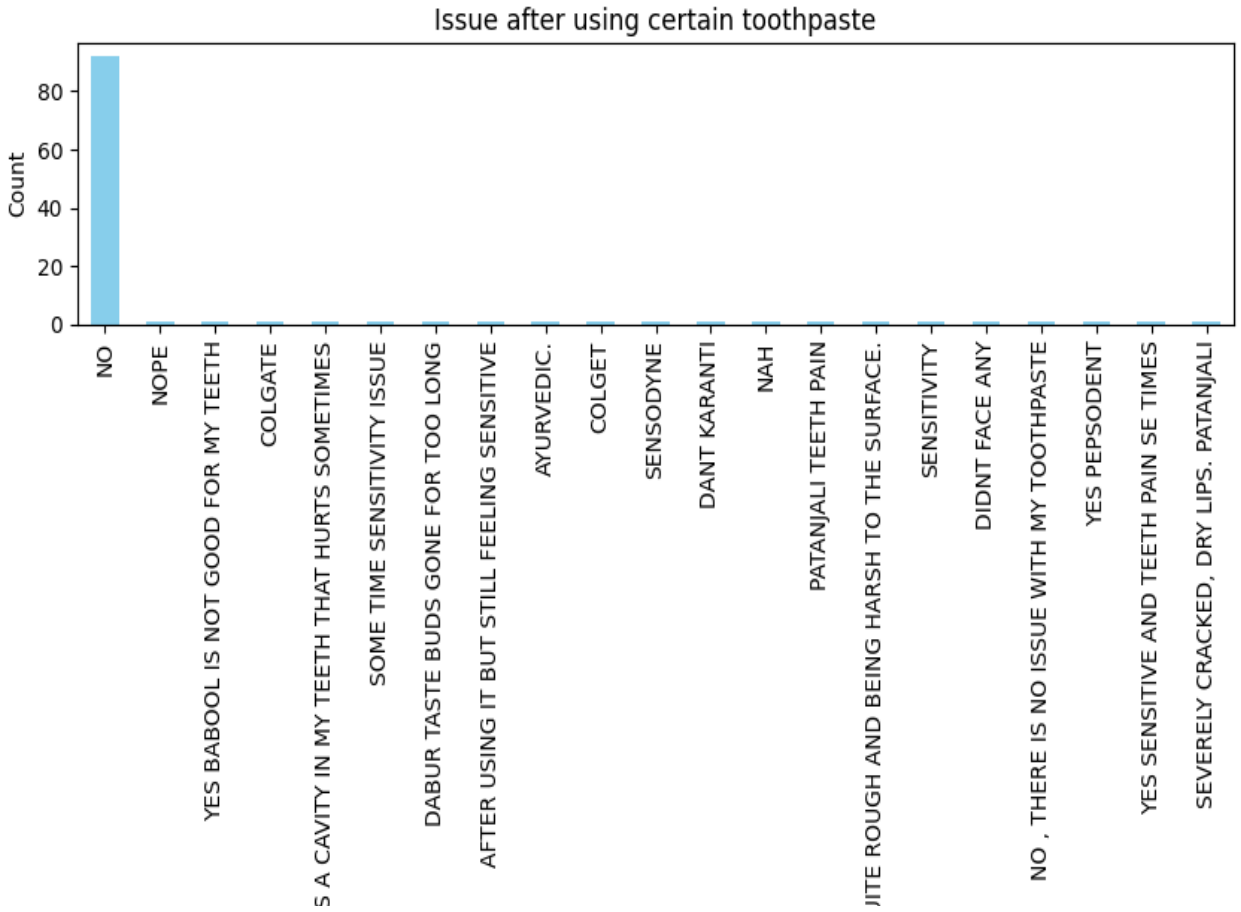
Data Set 13: Have you experienced any dental issues even after doing regular brushing?

```
value_counts = df['Have you experienced any dental issues even after doing regular brushing ? '].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Have you experienced any dental issues even after doing regular brushing ? ')
plt.xlabel('')
plt.ylabel('Count')
plt.show()
```



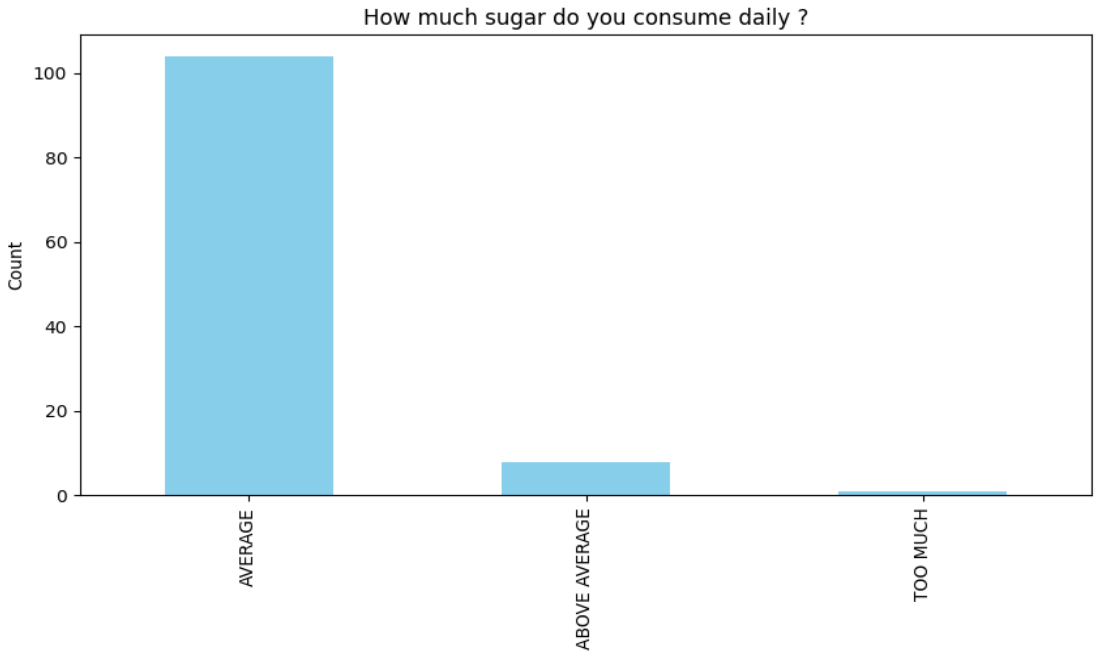
Data Set 14: Have you faced any issue after using a certain toothpaste? if yes then name the toothpaste

```
value_counts = df['Have you faced any issue after using a certain toothpaste? if yes then name the toothpaste '].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Issue after using certain toothpaste')
plt.xlabel(' ')
plt.ylabel('Count')
plt.show()
```



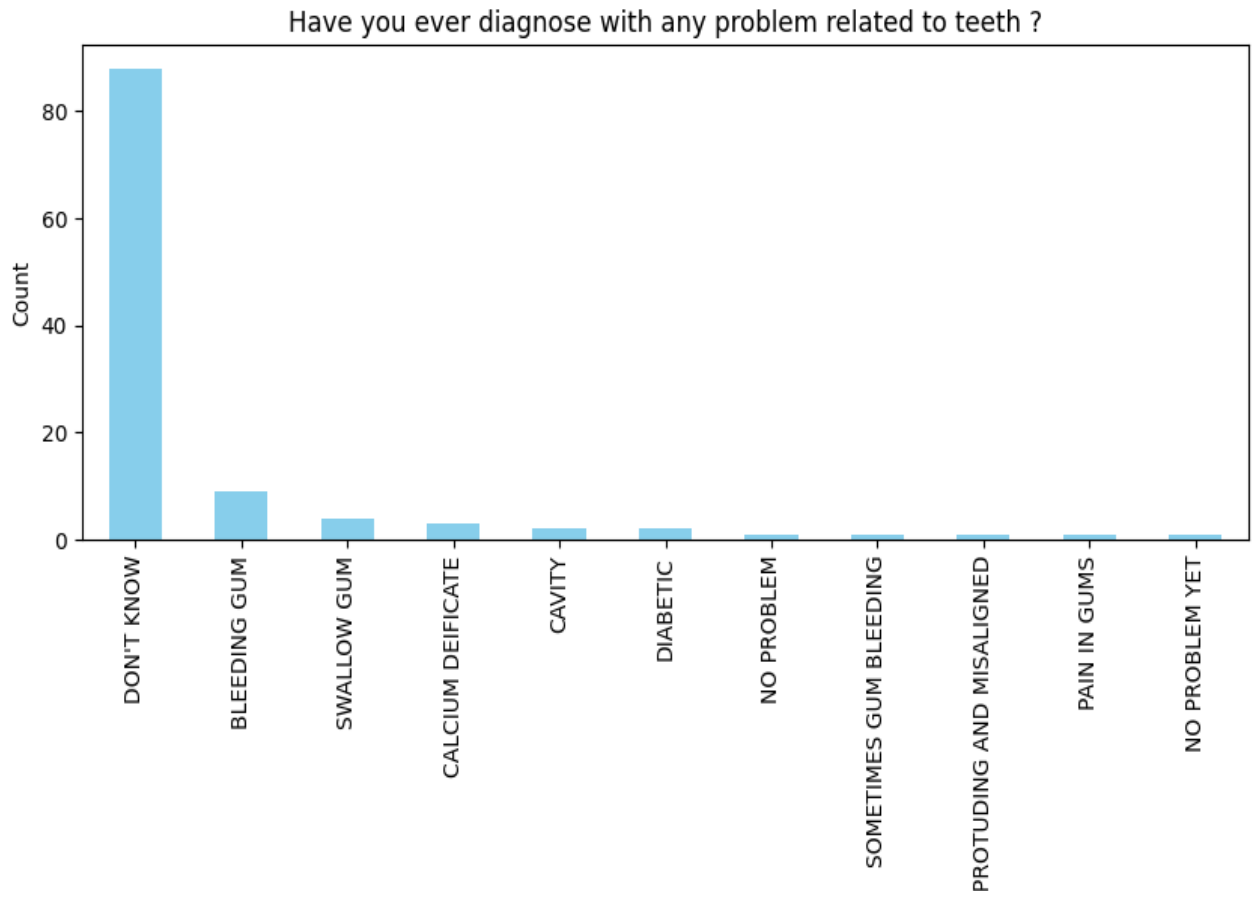
Data Set 15: How much sugar do you consume daily?

```
value_counts = df['How much sugar do you consume daily ?'].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'How much sugar do you consume daily ?')
plt.xlabel('')
plt.ylabel('Count')
plt.show()
```



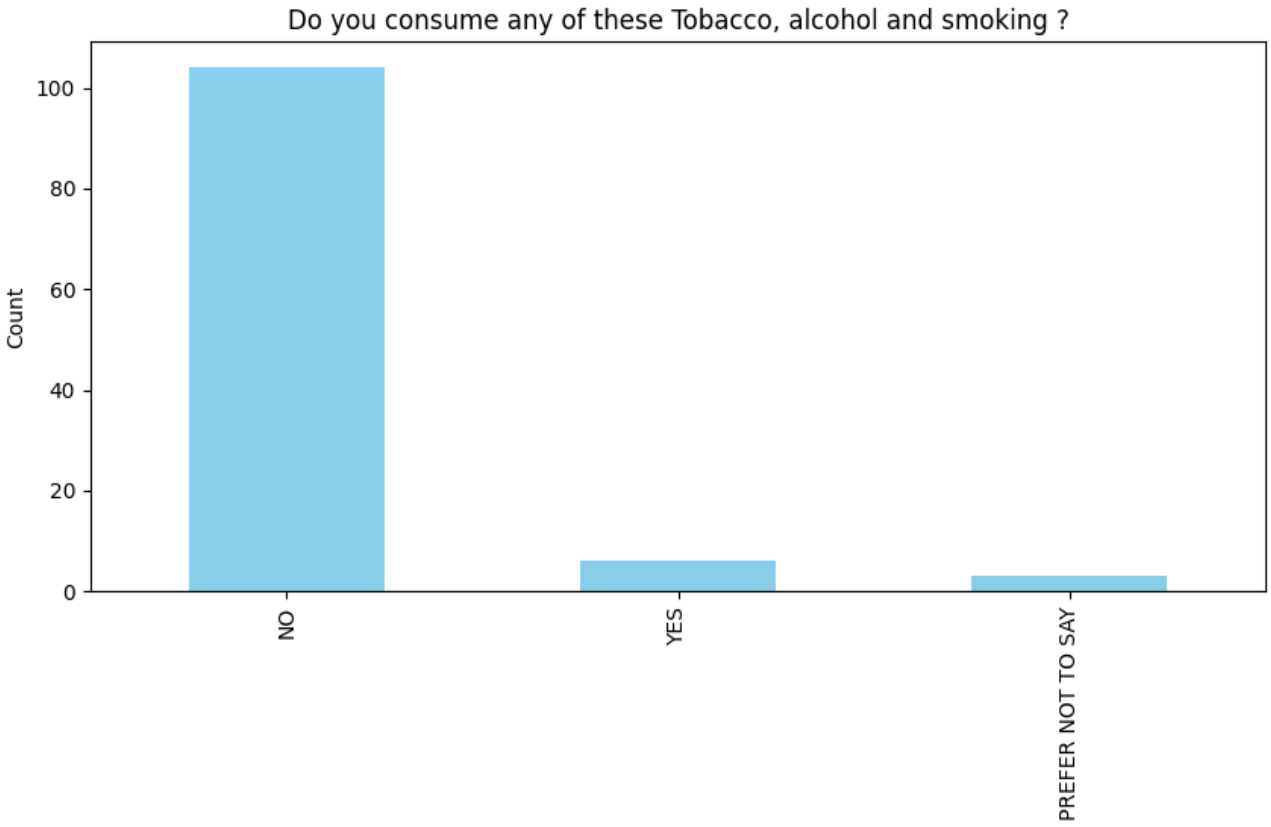
Data Set 16: Have you ever diagnose with any problem related to teeth?

```
value_counts = df['Have you ever diagnose with any problem related to teeth
?'].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Have you ever diagnose with any problem related to teeth ?')
plt.xlabel(' ')
plt.ylabel('Count')
plt.show()
```



Data Set 17: Do you consume any of these Tobacco, alcohol and smoking?

```
value_counts = df['Do you consume any of these Tobacco, alcohol and smoking
?'].value_counts()
plt.figure(figsize=(10, 6))
value_counts.plot(kind='bar', color='skyblue')
plt.title(f'Do you consume any of these Tobacco, alcohol and smoking ?')
plt.xlabel(' ')
plt.ylabel('Count')
plt.show()
```

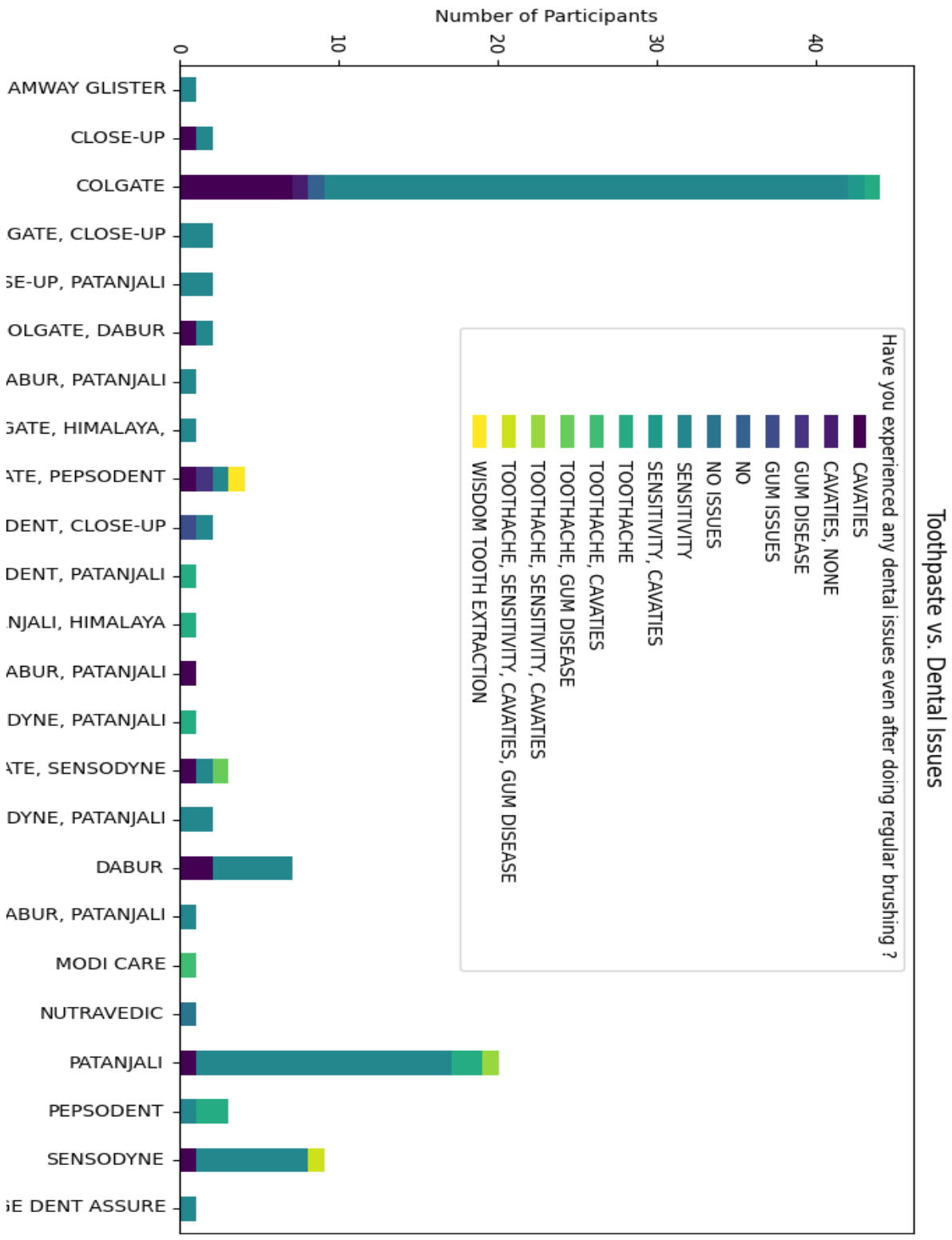


Compression between toothpaste use and dental issue after using toothpaste

This show if certain toothpaste doing any problem to people teeth

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import csv


df = pd.read_csv('Project Data2.csv')


toothpaste_brand = df['Which brand of toothpaste do you commonly use ?']
brushing_frequency = df['How often do you brush your teeth in a typical day ?']
brushing_duration = df['For how long you brush your teeth ?']
hot_cold_beverage = df['Do you drink hot or cold beverages regularly ?']
tea_coffee_consumption = df['Do you consume Tea or Coffee regularly ? ']
sugar_consumption = df['How much sugar do you consume daily ?']
dental_issues = df['Have you experienced any dental issues even after doing regular
brushing ? ']


hot_cold_beverage = np.where(hot_cold_beverage == 'Yes', 1, 0)
tea_coffee_consumption = np.where(tea_coffee_consumption == 'Yes', 1, 0)


toothpaste_dental_issues = pd.crosstab(toothpaste_brand, dental_issues)


brushing_dental_issues = pd.crosstab(brushing_frequency, dental_issues)
duration_dental_issues = pd.crosstab(brushing_duration, dental_issues)


plt.figure(figsize=(12, 8))


plt.subplot(2, 2, 1)
toothpaste_dental_issues.plot(kind='bar', stacked=True, colormap='viridis')
plt.title('Toothpaste vs. Dental Issues')
plt.xlabel('Toothpaste Brand')
plt.ylabel('Number of Participants')
```

```
plt.tight_layout()
plt.show()
```

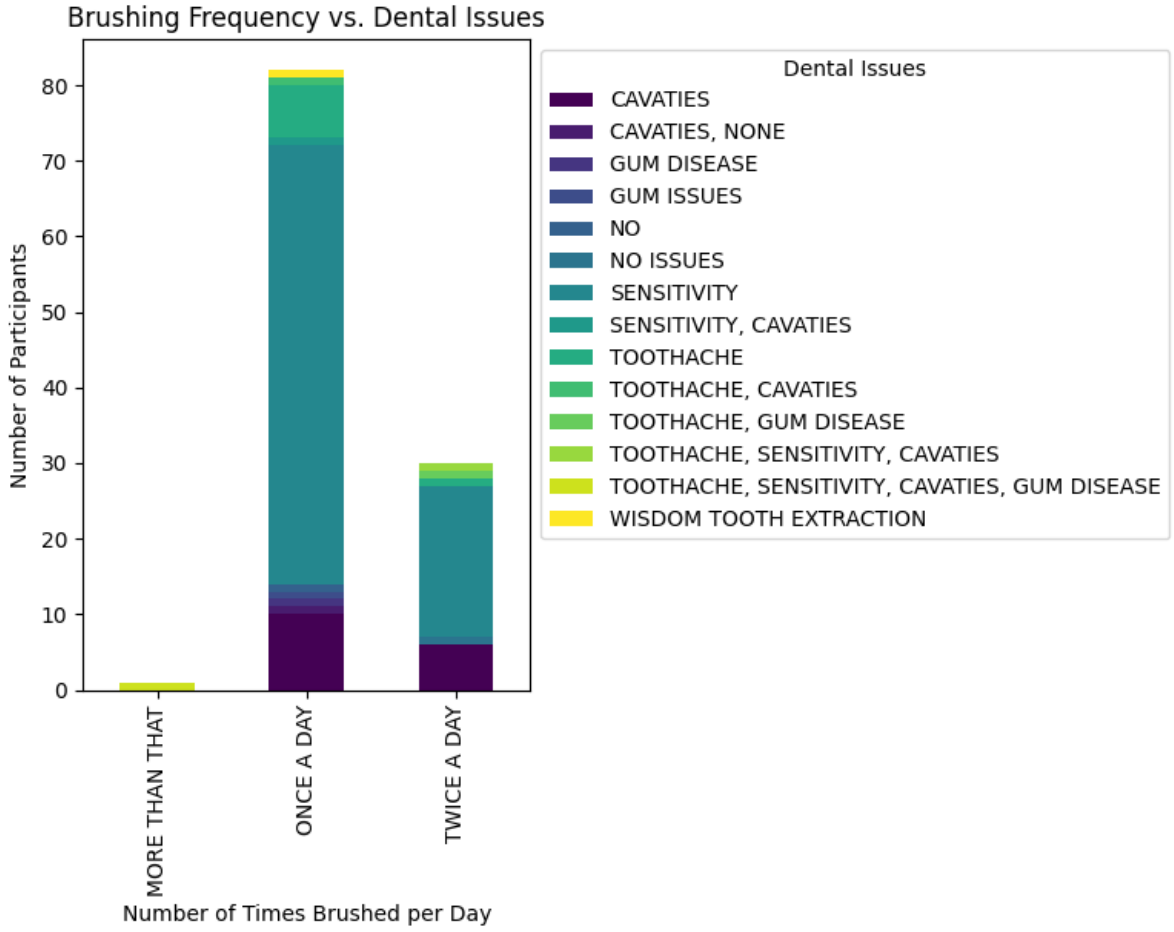


Compression between frequency of brush vs dental issue

This compression show if how many times you doing brush in a day but still getting dental issue

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt


df = pd.read_csv('Project Data2.csv')

toothpaste_brand = df['Which brand of toothpaste do you commonly use ?']
brushing_frequency = df['How often do you brush your teeth in a typical day ?']
brushing_duration = df['For how long you brush your teeth ?']
hot_cold_beverage = df['Do you drink hot or cold beverages regularly ?']
tea_coffee_consumption = df['Do you consume Tea or Coffee regularly ? ']
sugar_consumption = df['How much sugar do you consume daily ?']
dental_issues = df['Have you experienced any dental issues even after doing regular brushing ? ']

hot_cold_beverage = np.where(hot_cold_beverage == 'Yes', 1, 0)
tea_coffee_consumption = np.where(tea_coffee_consumption == 'Yes', 1, 0)


toothpaste_dental_issues = pd.crosstab(toothpaste_brand, dental_issues)


brushing_dental_issues = pd.crosstab(brushing_frequency, dental_issues)

duration_dental_issues = pd.crosstab(brushing_duration, dental_issues)


plt.figure(figsize=(12, 8))

plt.subplot(2, 2, 2)
brushing_dental_issues.plot(kind='bar', stacked=True, colormap='viridis')
plt.title('Brushing Frequency vs. Dental Issues')
plt.xlabel('Number of Times Brushed per Day')
plt.ylabel('Number of Participants')
plt.legend(title='Dental Issues', bbox_to_anchor=(1, 1))
```

```
plt.tight_layout()
plt.show()
```

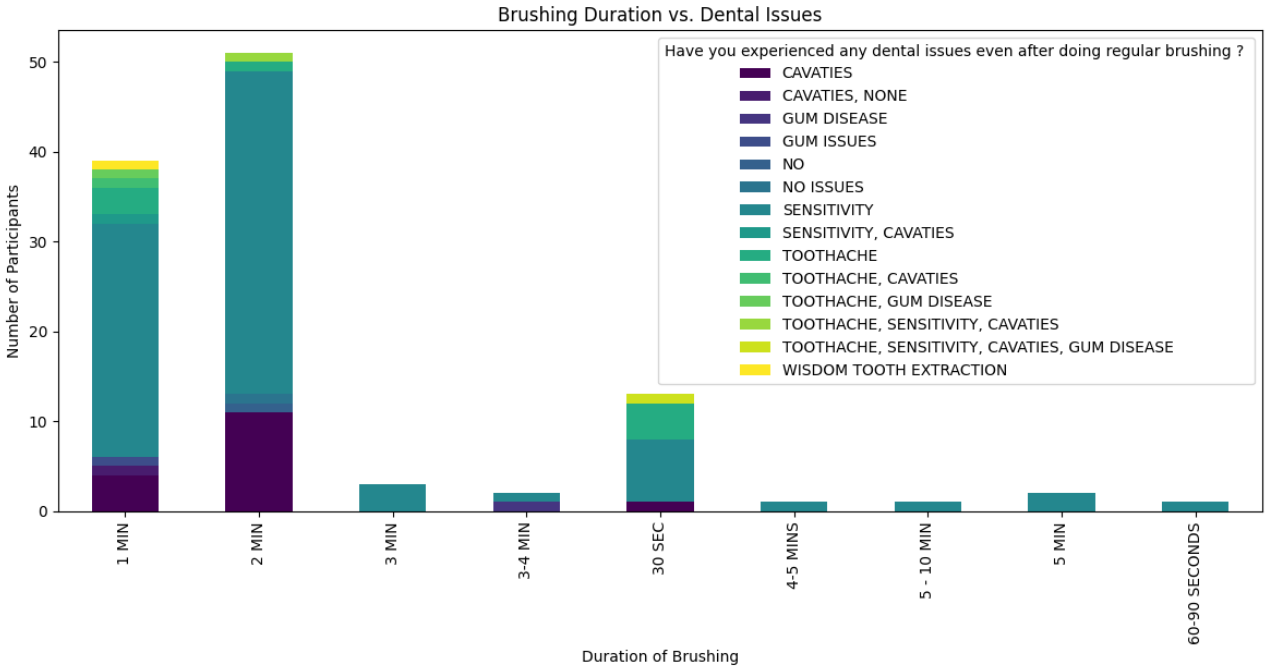


Compression between Brushing Duration vs Dental issue

This compression show people doing regular and sufficient time for brushing but still get dental issue

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import csv


df = pd.read_csv('Project Data2.csv')


toothpaste_brand = df['Which brand of toothpaste do you commonly use ?']
brushing_frequency = df['How often do you brush your teeth in a typical day ?']
brushing_duration = df['For how long you brush your teeth ?']
hot_cold_beverage = df['Do you drink hot or cold beverages regularly ?']
tea_coffee_consumption = df['Do you consume Tea or Coffee regularly ? ']
sugar_consumption = df['How much sugar do you consume daily ?']
dental_issues = df['Have you experienced any dental issues even after doing regular brushing ? ']


hot_cold_beverage = np.where(hot_cold_beverage == 'Yes', 1, 0)
tea_coffee_consumption = np.where(tea_coffee_consumption == 'Yes', 1, 0)

toothpaste_dental_issues = pd.crosstab(toothpaste_brand, dental_issues)

brushing_dental_issues = pd.crosstab(brushing_frequency, dental_issues)
duration_dental_issues = pd.crosstab(brushing_duration, dental_issues)

plt.figure(figsize=(12, 8))

plt.subplot(2, 2, 3)
duration_dental_issues.plot(kind='bar', stacked=True, colormap='viridis')
plt.title('Brushing Duration vs. Dental Issues')
plt.xlabel('Duration of Brushing')
plt.ylabel('Number of Participants')
plt.tight_layout()
```

```
plt.show()
```


Brushing Duration vs. Dental Issues
Have you experienced any dental issues even after doing regular brushing ?
CAVATIES
CAVATIES, NONE
GUM DISEASE
GUM ISSUES
NO
NO ISSUES
SENSITIVITY
SENSITIVITY, CAVATIES
TOOTHACHE
TOOTHACHE, CAVATIES
TOOTHACHE, GUM DISEASE
TOOTHACHE, SENSITIVITY, CAVATIES
TOOTHACHE, SENSITIVITY, CAVATIES, GUM DISEASE
WISDOM TOOTH EXTRACTION
Number of Participants
0
10
20
30
40
50
1 MIN
2 MIN
3 MIN
3-4 MIN
30 SEC
4-5 MINS
5 - 10 MIN
5 MIN
60-90 SECONDS
Duration of Brushing

Outcomes:

- Identification of the most commonly used toothpaste brands among participants.
- Insights into how dental issues may be influenced by toothpaste preferences, brushing habits, and beverage consumption.
- An understanding of whether participants adhere to recommended oral hygiene practices.
- An assessment of the impact of hot and cold beverage consumption on dental health. Insights into dental check-up behaviour and its association with dental health.

Reference

- McKiney, W. (2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython. 2nd edition. O'Reilly Media.
- O'Neil, C., & Schutt, R. (2013). Doing Data Science: Straight Talk from the Frontline O'Reilly Media
- https://www.w3schools.com/python/
- https://www.geeksforgeeks.org/pandas-tutorial/
- https://www.javatpoint.com/python-pandas