Identify The Most Significant Factors that cause stress and anxiety in the society

Imalsha Sewwandi Bulathsinhala

Table of Contents

Contents

Introduction 3

Background 3

Importance of Study 3

Scope of the Research 3

Objectives 4

Primary Objective 4

Secondary Objectives 4

Literature Review 5

Existing Studies 5

Gap Analysis 6

Methodology 7

Data Collection 7

Feedback about the Survey data 10

Data Cleaning and Preprocessing 12

Data Analysis 16

Exploratory Data Analysis (EDA) 16

Key Stress Factors 24

Gender Differences in Stress Levels 26

Bias in Dataset and Limitations of Results 27

Conclusion 28

Overview of Findings 28

Summary of Results 28

Introduction

Background

Stress and anxiety are pervasive issues in modern society, affecting individuals across all demographics. With the increasing pace of life, the demands of work, financial pressures, and social expectations, people are experiencing higher levels of stress and anxiety than ever before. These conditions can lead to serious mental health problems, affecting not only individuals but also their families, communities, and workplaces. Understanding the root causes of stress and anxiety is crucial for developing effective interventions and support systems to improve overall well-being.

Importance of Study

Identifying the significant factors causing stress and anxiety is vital for several reasons. First, it allows healthcare providers and policymakers to design targeted interventions that address the most pressing issues. Second, it helps individuals recognize and manage their own stressors more effectively. Lastly, by understanding these factors, employers and educators can create healthier environments that promote mental well-being. Given the widespread impact of stress and anxiety on productivity, health, and quality of life, this study aims to contribute valuable insights that can inform both public health strategies and personal coping mechanisms.

Scope of the Research

This research focuses on identifying the most significant factors contributing to stress and anxiety within a specific population sample in Canada. The study utilizes primary data collected through a survey. The survey provides direct insights from individuals about their experiences and perceptions. The scope of the study is limited to analyzing the data collected from 26 survey respondents and the available public datasets. Due to the limited sample size and scope of data, the findings may not be generalizable to the entire population. However, the research aims to highlight key trends and factors that can be further explored in larger studies.

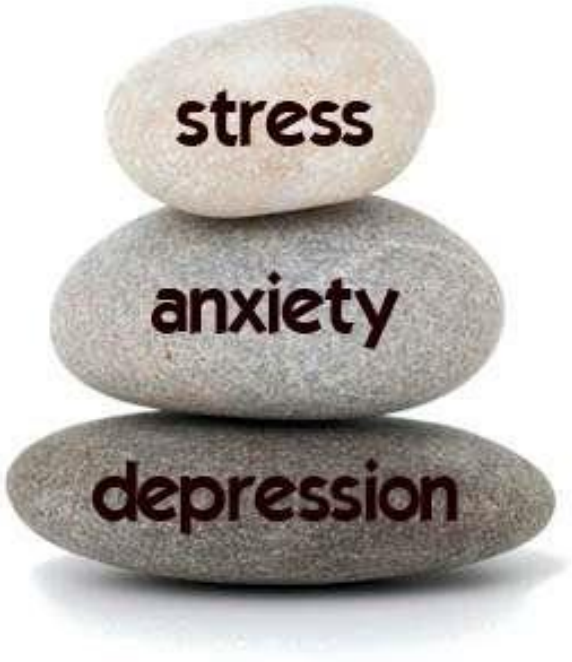
Objectives

Primary Objective

- **Identify the most significant factors contributing to stress and anxiety in Canada.**
    - This involves analyzing data to determine which factors, such as work-related stress, financial pressures, health issues, and other relevant variables, are most strongly associated with increased levels of stress and anxiety among individuals.

Secondary Objectives

- **Analyze the impact of these factors on different demographics.**
    - Investigate how stress and anxiety factors affect various demographic groups, such as age, gender, and location. This will help to understand if certain groups are more vulnerable to specific stressors.
- **Explore trends over time or across regions.**
    - Examine how the significance of stress and anxiety factors may change over time or vary between different geographical regions. This can reveal patterns or shifts in the causes of stress and anxiety, providing insights for more tailored interventions.

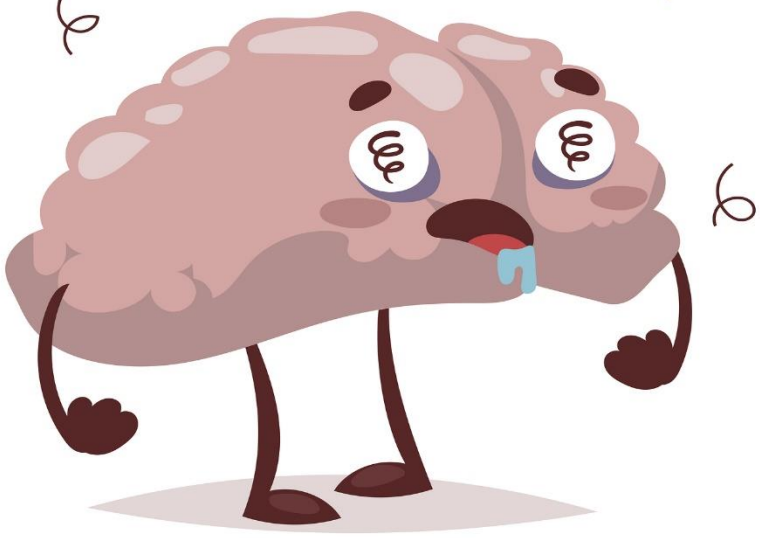


Literature Review

Existing Studies

Numerous studies have examined the causes and effects of stress and anxiety in society. Key findings from this body of research include:

1. **Work-Related Stress**: Studies consistently find that job-related factors such as workload, job insecurity, and lack of work-life balance are significant contributors to stress and anxiety. According to the American Psychological Association (APA), work is one of the top sources of stress for many individuals.
2. **Financial Pressures**: Financial stress, including concerns about income, debt, and economic stability, is a major source of anxiety. Research by the Financial Planning Association (FPA) indicates that financial stress is linked to poorer mental health outcomes.
3. **Health Issues**: Chronic health conditions and concerns about personal or family health are significant stressors. The World Health Organization (WHO) reports that health-related stress is particularly prevalent among older adults and those with pre-existing medical conditions.
4. **Family and Relationship Problems**: Interpersonal relationships, including family conflicts and relationship issues, are common sources of stress. Studies show that supportive relationships can buffer against stress, while strained relationships can exacerbate it.
5. **Social and Environmental Factors**: Social isolation, environmental stressors (such as noise and pollution), and societal pressures (such as cultural expectations and social media) have been identified as contributing factors. Research highlights the growing impact of social media on anxiety levels, particularly among younger individuals.
6. **Educational Stress**: Students face significant stress from academic demands, including exams, assignments, and future career prospects. The National Institute of Mental Health (NIMH) reports high levels of anxiety among college students.



Gap Analysis

Despite the extensive research on stress and anxiety, several gaps remain:

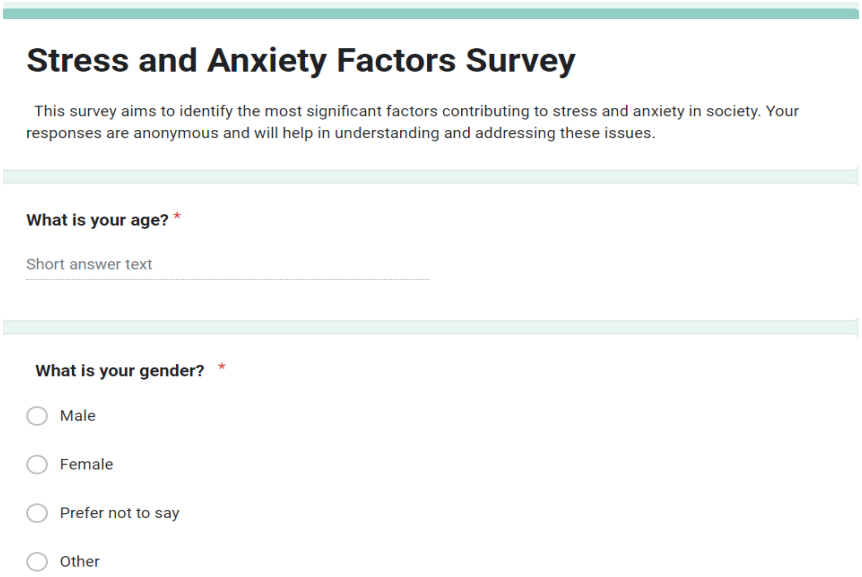
1. **Integration of Multiple Factors**: Many studies focus on single factors in isolation, such as work-related stress or financial pressures. There is a need for comprehensive analyses that integrate multiple stressors to understand their combined impact.
2. **Demographic-Specific Insights**: While some studies address how stress affects different demographic groups, there is limited research that thoroughly examines the intersectionality of factors such as age, gender and location.
3. **Longitudinal Data**: Much of the existing research is cross-sectional, providing a snapshot of stress and anxiety at a single point in time. Longitudinal studies are needed to track changes in stress levels and contributing factors over time.
4. **Regional Variations**: There is a lack of research comparing stress and anxiety factors across different geographical regions. Understanding regional variations can inform more localized and culturally relevant interventions.
5. **Impact of Technology**: With the rapid advancement of technology and the increasing use of social media, there is a need for more research on how digital environments contribute to stress and anxiety, particularly among younger populations.

This study aims to address these gaps by integrating multiple stress factors, analyzing their impact on different demographics, and exploring trends across regions using both primary and secondary data.

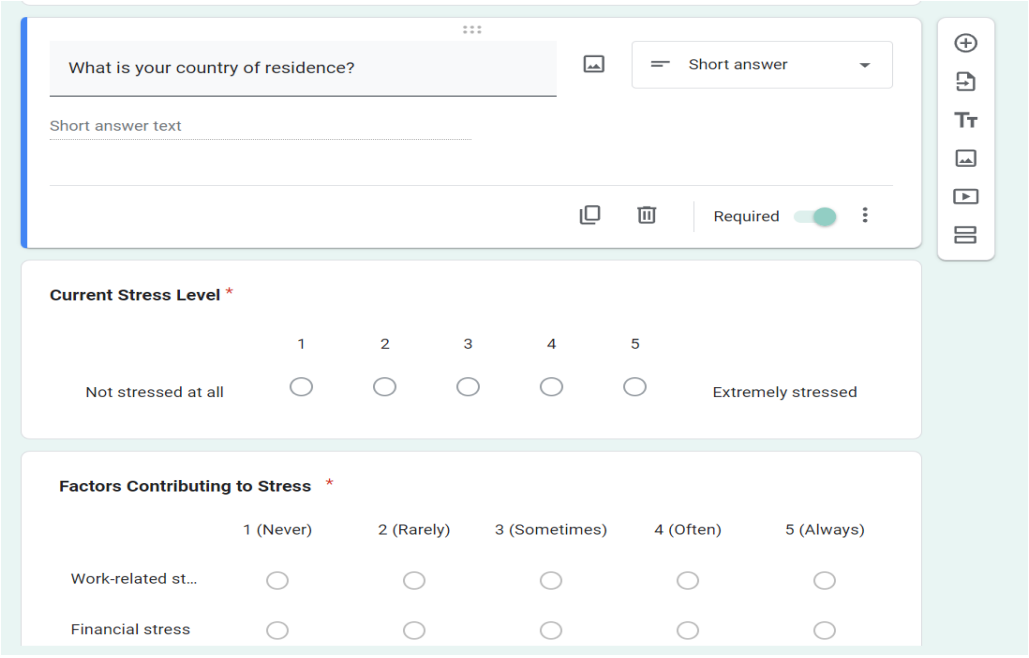
Methodology

Data Collection

Survey Design: The primary data for this study was collected through a structured survey designed to capture various factors contributing to stress and anxiety. The survey included the following sections:

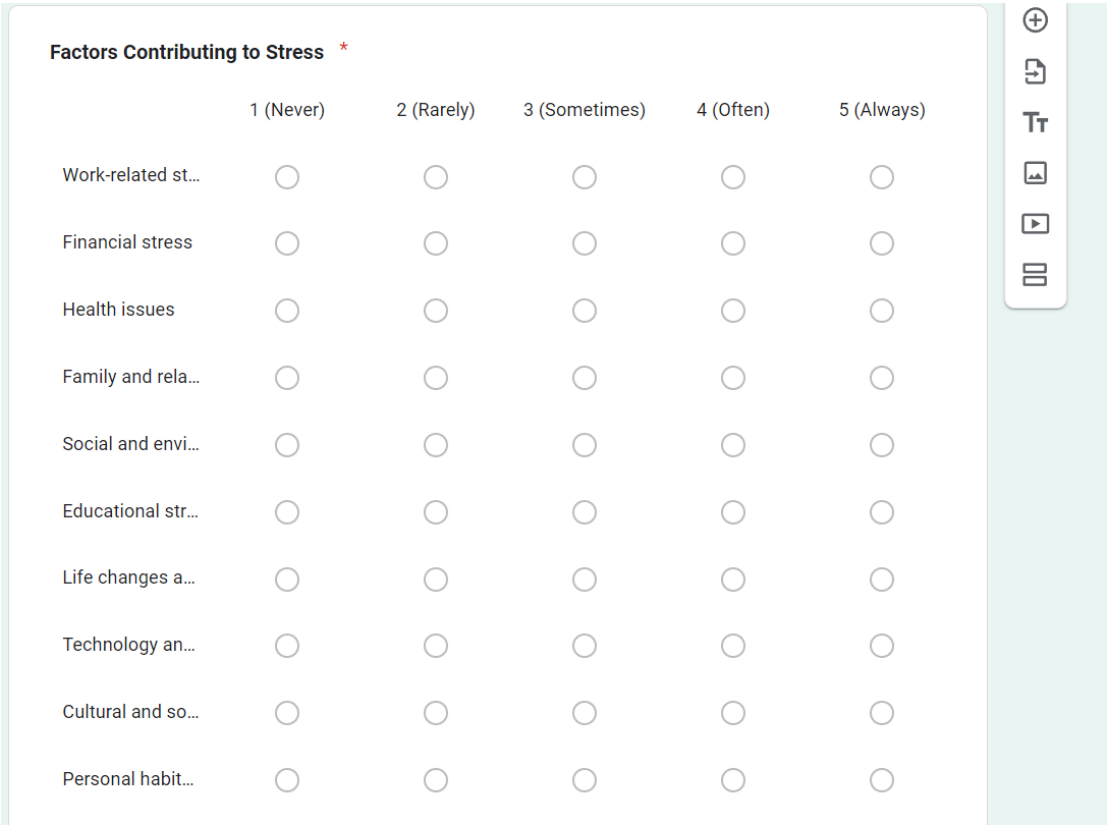


Demographics: Questions about age, gender and location



- **Stress Factors**: Rating scales for different potential stressors, including work, financial, health, family, social, educational, and environmental stress.

- **Stress Levels**: Self-reported stress and anxiety levels on a scale from 1 to 5.

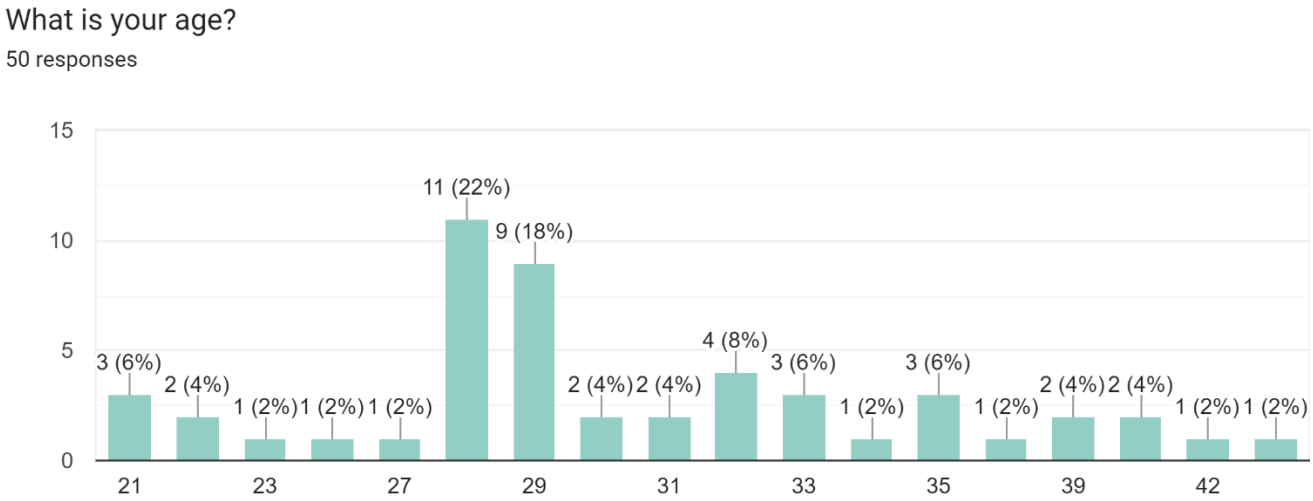


Distribution Methods: The survey was distributed through multiple channels to ensure a diverse respondent pool:

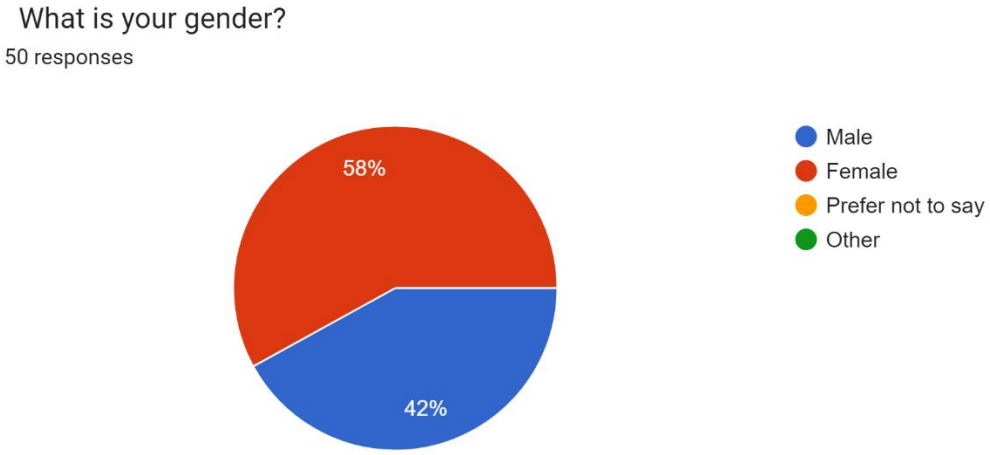
- **Online Platforms**: Shared on social media (Facebook, Twitter, LinkedIn) and through email invitations.
- **Community Groups**: Posted in local community groups and forums.

Demographics of Respondents: The survey received 50 responses. The demographics of the respondents are as follows:

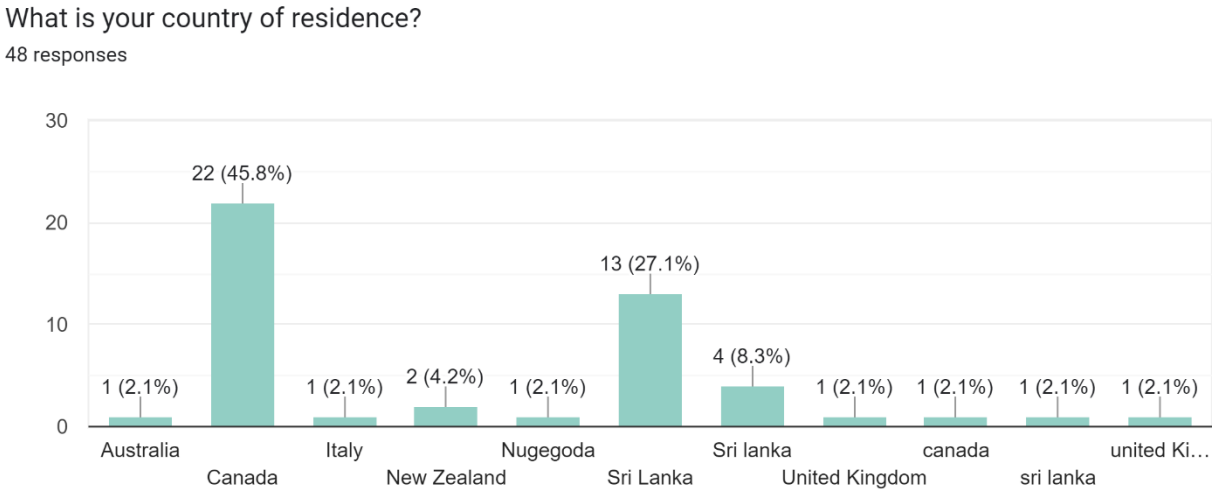
- **Age**: Respondents ranged from 21 to 45 years old, with a median age of 33.



- **Gender**: 58% female, 42% male.



- **Location**: Respondents were from Canada, Sri Lanka, UK, New Zealand and Australia



Feedback about the Survey data

During the data collection process for our study on stress and anxiety factors, I aimed to gather a diverse and representative sample of respondents. However, I encountered some challenges that introduced potential biases into my dataset.

Survey Participant Demographics:

The survey was distributed through various channels, but the majority of responses came from individuals located in Sri Lanka and Canada. Specifically:

- **Sri Lanka**: 37.5% of respondents
- **Canada**: 47.9% of respondents
- **Other Countries**: 14.6% of respondents

Potential Biases:

1. **Geographical Bias**: The overrepresentation of participants from Sri Lanka and Canada may skew the results, as cultural, economic, and social factors unique to these countries could influence the reported stress and anxiety levels. This geographical bias limits the generalizability of our findings to other regions.
2. **Cultural Influence**: Cultural differences in how stress and anxiety are perceived and reported might affect the data. For example, societal norms in Sri Lanka and Canada regarding mental health awareness and stigma could impact respondents' willingness to report certain factors.

3. **Demographic Homogeneity**: If the demographic characteristics (e.g., age, gender, occupation) of the respondents from Sri Lanka and Canada are similar, this homogeneity could further limit the variability and richness of the data.

Impact on Results:

The biases mentioned above could affect the analysis and interpretation of the results. For instance:

- Factors contributing to stress and anxiety might appear more significant or less significant than they are on a global scale.
- Certain stress factors prevalent in Sri Lanka and Canada might dominate the analysis, overshadowing factors more relevant to other regions.

Mitigation Strategies:

To address these biases and improve the representativeness of future studies, I recommend the following strategies:

1. **Diversified Distribution Channels**: Use a broader range of distribution channels to reach a more diverse audience. This could include international social media platforms, collaboration with global organizations, and targeting specific demographic groups.
2. **Stratified Sampling**: Implement stratified sampling techniques to ensure that the sample includes participants from various regions, age groups, and occupations.
3. **Weighted Analysis**: Apply statistical weighting methods to adjust for overrepresented and underrepresented groups in the data, providing a more balanced analysis.

Conclusion:

While my study provides valuable insights into the factors contributing to stress and anxiety, the identified biases highlight the importance of considering the context in which data is collected. By acknowledging these limitations, I aim to provide a transparent and honest evaluation of my findings and suggest improvements for future research. But I have to continue my analysis for the purpose of studying and experience using the data set that I gathered.

Data Cleaning and Preprocessing

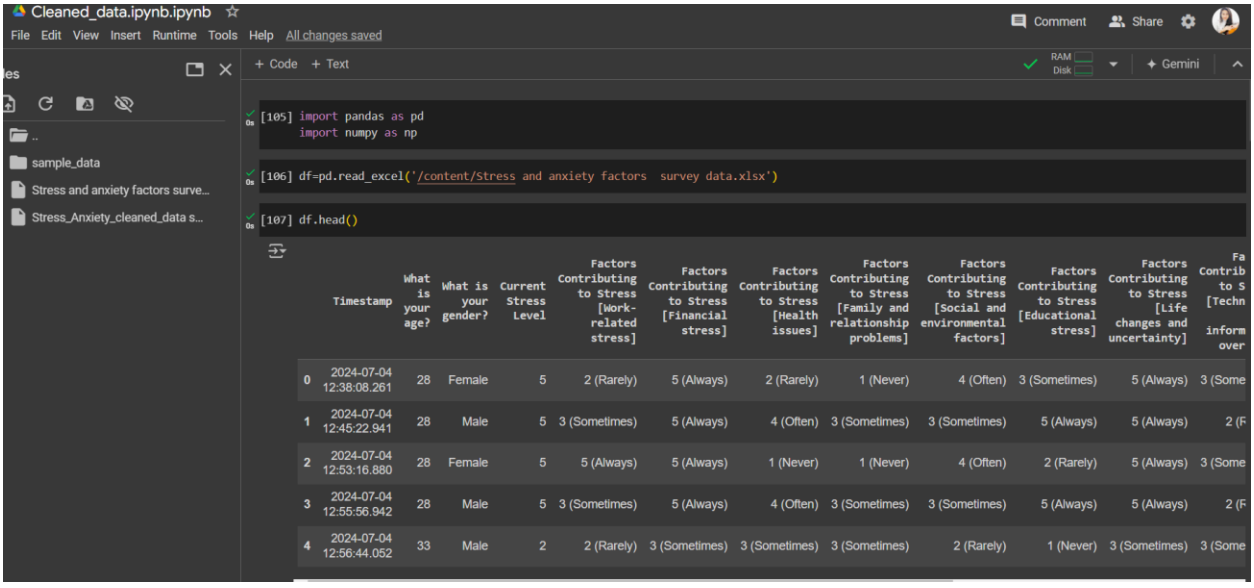
Survey Data Cleaning:

To ensure the quality and accuracy of our survey data, I performed several data cleaning steps using python's pandas libraries:

1. **Removing Duplicates**: We identified and removed any duplicate responses from the dataset, ensuring that each respondent was only represented once.
2. **Handling Missing Values**: We addressed missing values by dropping rows with incomplete data. This ensured that our analysis was based on complete and reliable information.
3. **Correcting Data Types**: We checked and corrected data types for each column to ensure that numerical, categorical, and date/time data were appropriately formatted.
4. **Standardizing Text Data**: We standardized the text data by capitalizing the first letter of each word in the 'Country' column and removing any leading or trailing spaces from all string columns.
5. **Encoding Categorical Variables**: We converted categorical variables into numerical format using one-hot encoding, preparing the data for analysis and modeling.

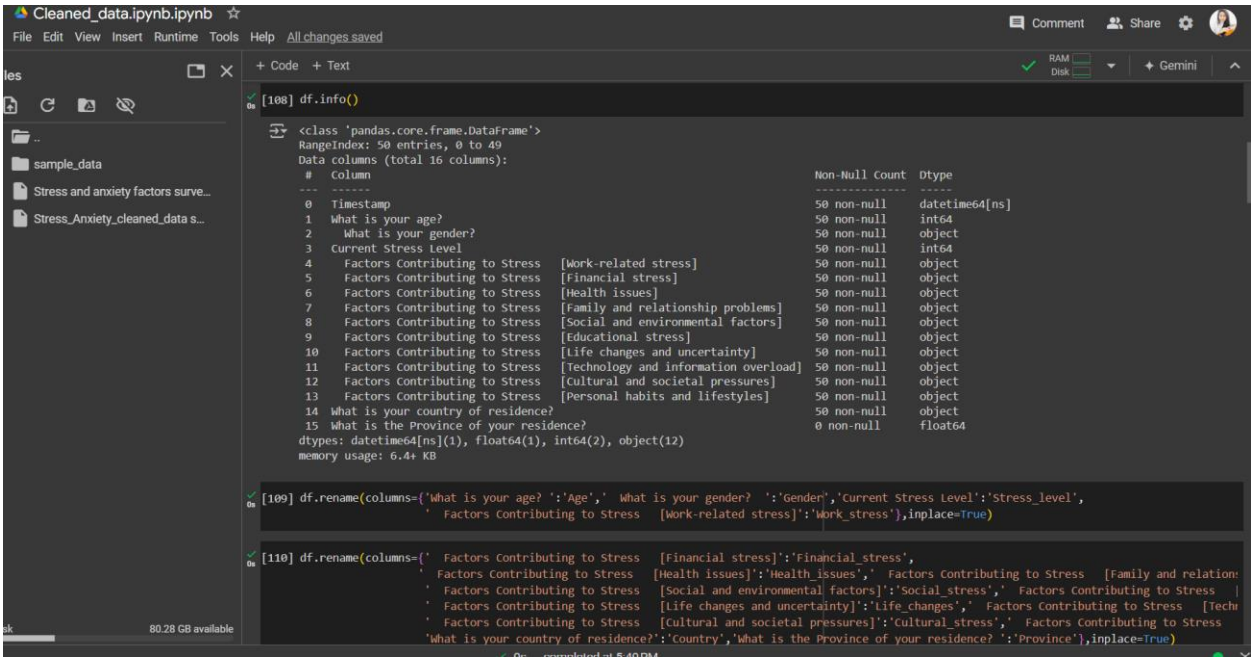
These steps helped us create a clean, well-structured dataset that forms the foundation of our analysis.

#Loading the gathered data into a pandas data frame



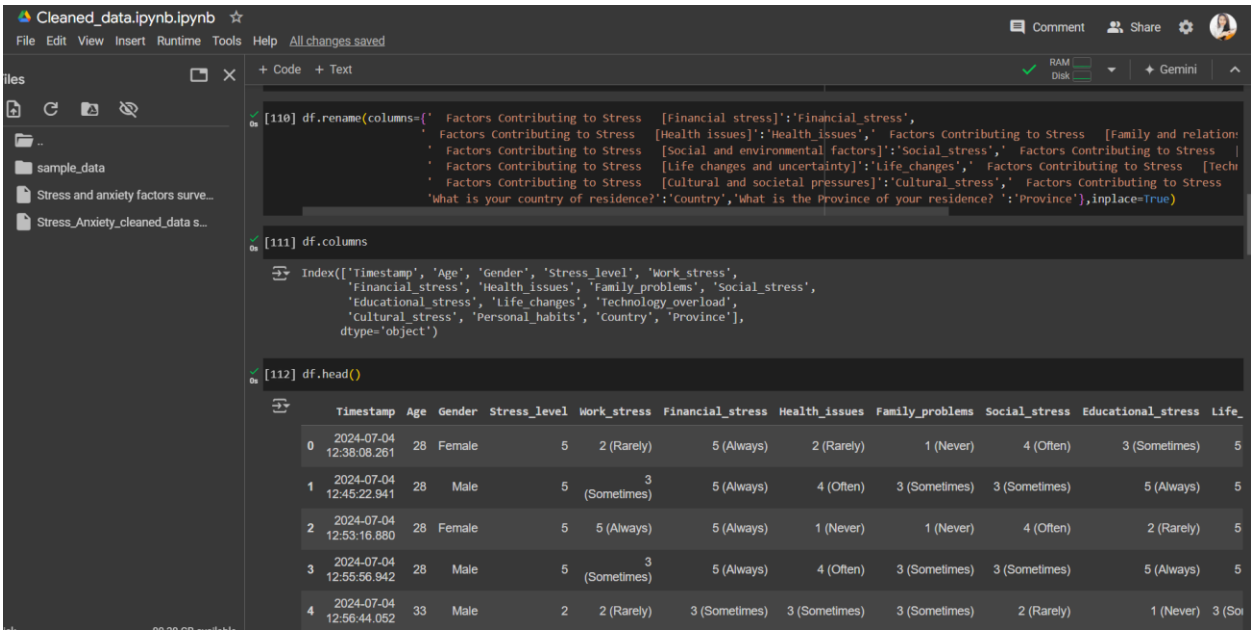
	Timestamp	What is your age?	What is your gender?	Current Stress Level	Factors Contributing to Stress [Work-related stress]	Factors Contributing to Stress [Financial stress]	Factors Contributing to Stress [Health issues]	Factors Contributing to Stress [Family and relationship problems]	Factors Contributing to Stress [Social and environmental factors]	Factors Contributing to Stress [Educational stress]	Factors Contributing to Stress [Life changes and uncertainty]	Fa Contrib to S [Techn inform over
0	2024-07-04 12:38:08.261	28	Female	5	2 (Rarely)	5 (Always)	2 (Rarely)	1 (Never)	4 (Often)	3 (Sometimes)	5 (Always)	3 (Some
1	2024-07-04 12:45:22.941	28	Male	5	3 (Sometimes)	5 (Always)	4 (Often)	3 (Sometimes)	3 (Sometimes)	5 (Always)	5 (Always)	2 (F
2	2024-07-04 12:53:16.880	28	Female	5	5 (Always)	5 (Always)	1 (Never)	1 (Never)	4 (Often)	2 (Rarely)	5 (Always)	3 (Some
3	2024-07-04 12:55:56.942	28	Male	5	3 (Sometimes)	5 (Always)	4 (Often)	3 (Sometimes)	3 (Sometimes)	5 (Always)	5 (Always)	2 (F
4	2024-07-04 12:56:44.052	33	Male	2	2 (Rarely)	3 (Sometimes)	3 (Sometimes)	3 (Sometimes)	2 (Rarely)	1 (Never)	3 (Sometimes)	3 (Some

#Checking for the data types and Null values. Since I made all the survey fields as required fields, there cannot be any Null values. So I moved forward without further digging deep into the null values.



```
[108] df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50 entries, 0 to 49
Data columns (total 16 columns):
 #   Column                                                  Non-Null Count  Dtype
---  ------                                                  --------------  -----
 0   Timestamp                                               50 non-null     datetime64[ns]
 1   What is your age?                                       50 non-null     int64
 2   What is your gender?                                    50 non-null     object
 3   Current Stress Level                                    50 non-null     int64
 4   Factors Contributing to Stress   [Work-related stress]  50 non-null     object
 5   Factors Contributing to Stress   [Financial stress]     50 non-null     object
 6   Factors Contributing to Stress   [Health issues]        50 non-null     object
 7   Factors Contributing to Stress   [Family and relationship problems]  50 non-null  object
 8   Factors Contributing to Stress   [Social and environmental factors]  50 non-null  object
 9   Factors Contributing to Stress   [Educational stress]   50 non-null     object
 10  Factors Contributing to Stress   [Life changes and uncertainty]  50 non-null  object
 11  Factors Contributing to Stress   [Technology and information overload]  50 non-null  object
 12  Factors Contributing to Stress   [Cultural and societal pressures]  50 non-null  object
 13  Factors Contributing to Stress   [Personal habits and lifestyles]  50 non-null  object
 14  What is your country of residence?                      50 non-null     object
 15  What is the Province of your residence?                 0 non-null      float64
dtypes: datetime64[ns](1), float64(1), int64(2), object(12)
memory usage: 6.4+ KB

[109] df.rename(columns={'What is your age? ':'Age','  What is your gender?  ':'Gender','Current Stress Level':'Stress_level',
                         '  Factors Contributing to Stress   [Work-related stress]':'Work_stress'},inplace=True)

[110] df.rename(columns={'  Factors Contributing to Stress   [Financial stress]':'Financial_stress',
                         '  Factors Contributing to Stress   [Health issues]':'Health_issues','  Factors Contributing to Stress   [Family and relation
                         '  Factors Contributing to Stress   [Social and environmental factors]':'Social_stress','  Factors Contributing to Stress
                         '  Factors Contributing to Stress   [Life changes and uncertainty]':'Life_changes','  Factors Contributing to Stress   [Techn
                         '  Factors Contributing to Stress   [Cultural and societal pressures]':'Cultural_stress','  Factors Contributing to Stress
                        'What is your country of residence?':'Country','What is the Province of your residence? ':'Province'},inplace=True)
```

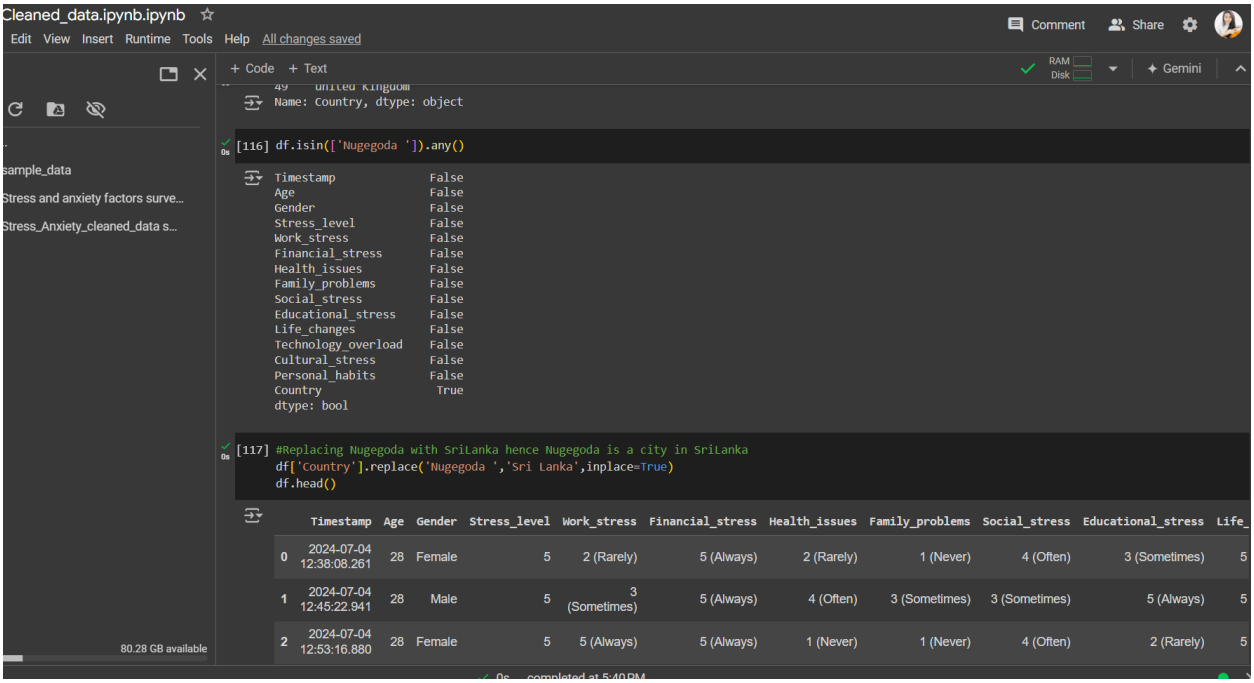
#Changing the column names

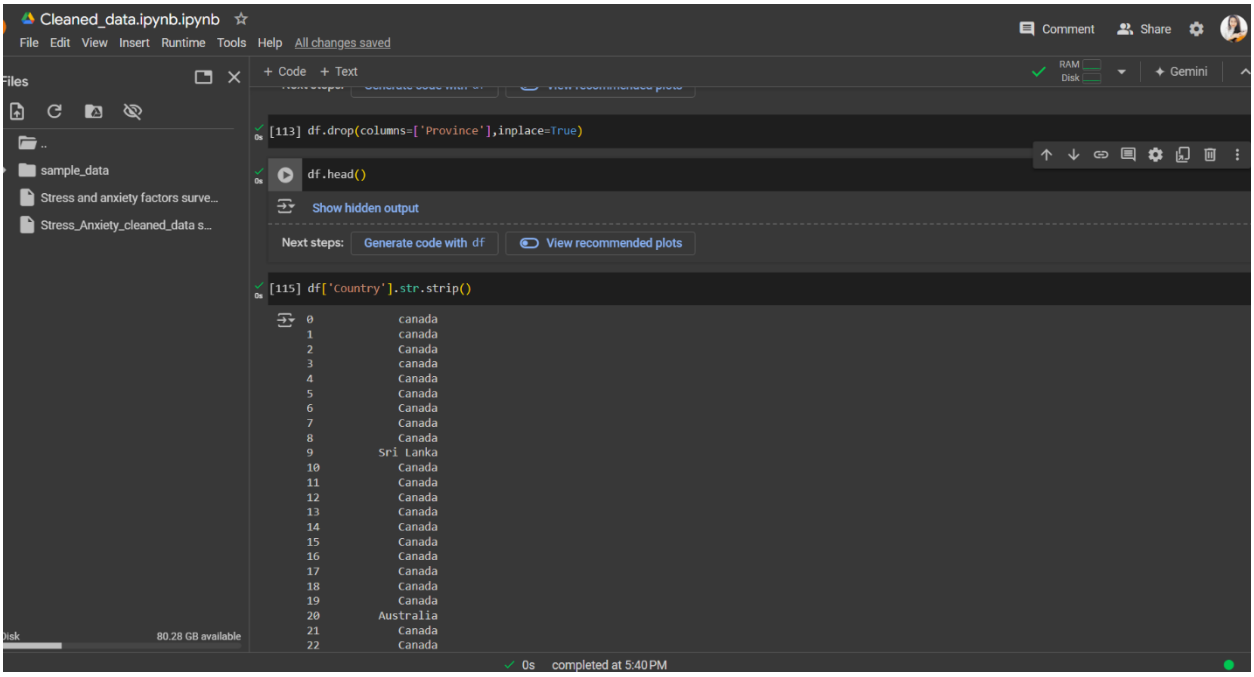


```
[110] df.rename(columns={'  Factors Contributing to Stress   [Financial stress]':'Financial_stress',
                         '  Factors Contributing to Stress   [Health issues]':'Health_issues','  Factors Contributing to Stress   [Family and relation
                         '  Factors Contributing to Stress   [Social and environmental factors]':'Social_stress','  Factors Contributing to Stress
                         '  Factors Contributing to Stress   [Life changes and uncertainty]':'Life_changes','  Factors Contributing to Stress   [Techn
                         '  Factors Contributing to Stress   [Cultural and societal pressures]':'Cultural_stress','  Factors Contributing to Stress
                        'What is your country of residence?':'Country','What is the Province of your residence? ':'Province'},inplace=True)

[111] df.columns

Index(['Timestamp', 'Age', 'Gender', 'Stress_level', 'Work_stress',
       'Financial_stress', 'Health_issues', 'Family_problems', 'Social_stress',
       'Educational_stress', 'Life_changes', 'Technology_overload',
       'Cultural_stress', 'Personal_habits', 'Country', 'Province'],
      dtype='object')

[112] df.head()
```

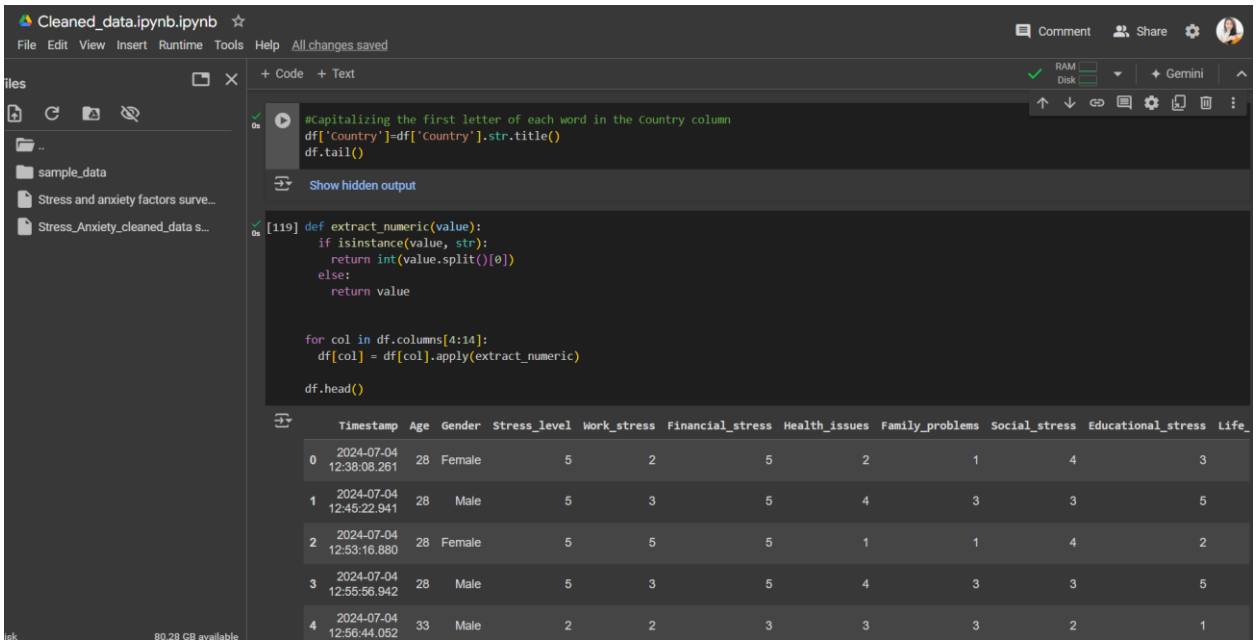
	Timestamp	Age	Gender	Stress_level	Work_stress	Financial_stress	Health_issues	Family_problems	Social_stress	Educational_stress	Life_
0	2024-07-04 12:38:08.261	28	Female	5	2 (Rarely)	5 (Always)	2 (Rarely)	1 (Never)	4 (Often)	3 (Sometimes)	5
1	2024-07-04 12:45:22.941	28	Male	5	3 (Sometimes)	5 (Always)	4 (Often)	3 (Sometimes)	3 (Sometimes)	5 (Always)	5
2	2024-07-04 12:53:16.880	28	Female	5	5 (Always)	5 (Always)	1 (Never)	1 (Never)	4 (Often)	2 (Rarely)	5
3	2024-07-04 12:55:56.942	28	Male	5	3 (Sometimes)	5 (Always)	4 (Often)	3 (Sometimes)	3 (Sometimes)	5 (Always)	5
4	2024-07-04 12:56:44.052	33	Male	2	2 (Rarely)	3 (Sometimes)	3 (Sometimes)	3 (Sometimes)	2 (Rarely)	1 (Never)	3 (So

#Standardizing data





#Extracting Numeric values and removing text from the stress levels



Cleaned_data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share Gemini

+ Code + Text

```
#Capitalizing the first letter of each word in the Country column
df['Country']=df['Country'].str.title()
df.tail()
```

Show hidden output

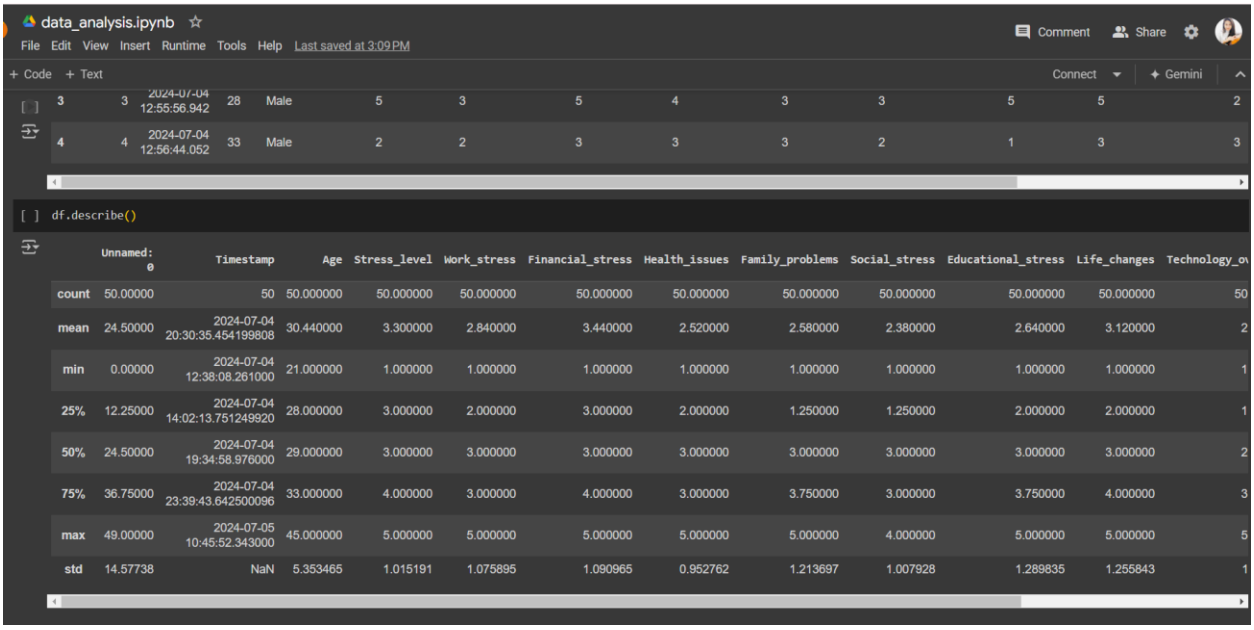
```
[119] def extract_numeric(value):
        if isinstance(value, str):
          return int(value.split()[0])
        else:
          return value


      for col in df.columns[4:14]:
        df[col] = df[col].apply(extract_numeric)

      df.head()
```

	Timestamp	Age	Gender	Stress_level	Work_stress	Financial_stress	Health_issues	Family_problems	Social_stress	Educational_stress	Life_
0	2024-07-04 12:38:08.261	28	Female	5	2	5	2	1	4	3	
1	2024-07-04 12:45:22.941	28	Male	5	3	5	4	3	3	5	
2	2024-07-04 12:53:16.880	28	Female	5	5	5	1	1	4	2	
3	2024-07-04 12:55:56.942	28	Male	5	3	5	4	3	3	5	
4	2024-07-04 12:56:44.052	33	Male	2	2	3	3	3	2	1	

Data Analysis

Exploratory Data Analysis (EDA)

In this section, I will conduct exploratory data analysis (EDA) to understand the underlying patterns and distributions in the survey data. This includes calculating descriptive statistics and creating visualizations to illustrate the data.



data_analysis.ipynb

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

Comment Share

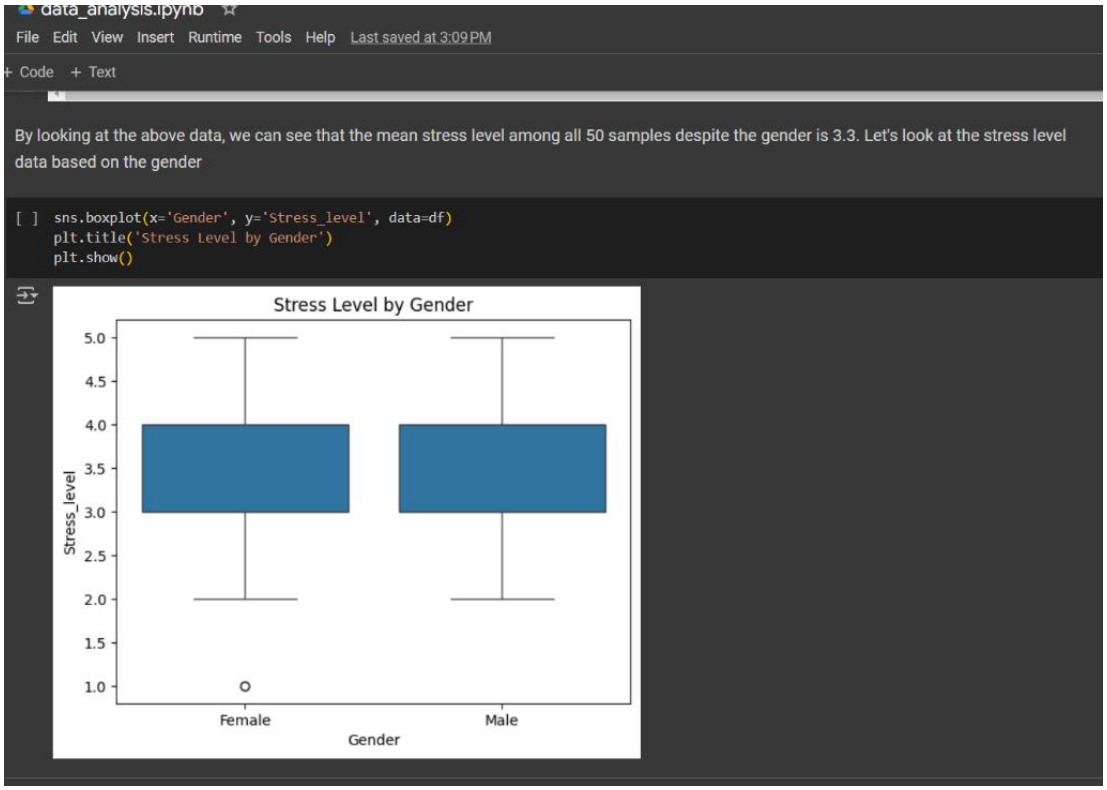
+ Code + Text

Connect ✦ Gemini

3	3	2024-07-04 12:55:56.942	28	Male	5	3	5	4	3	3	5	5		2
4	4	2024-07-04 12:56:44.052	33	Male	2	2	3	3	3	2	1	3		3

```
[ ] df.describe()
```

	Unnamed: 0	Timestamp	Age	Stress_level	Work_stress	Financial_stress	Health_issues	Family_problems	Social_stress	Educational_stress	Life_changes	Technology_ov
count	50.00000	50	50.000000	50.000000	50.000000	50.000000	50.000000	50.000000	50.000000	50.000000	50.000000	50
mean	24.50000	2024-07-04 20:30:35.454199808	30.440000	3.300000	2.840000	3.440000	2.520000	2.580000	2.380000	2.640000	3.120000	2
min	0.00000	2024-07-04 12:38:08.261000	21.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1
25%	12.25000	2024-07-04 14:02:13.751249920	28.000000	3.000000	2.000000	3.000000	2.000000	1.250000	1.250000	2.000000	2.000000	1
50%	24.50000	2024-07-04 19:34:58.976000	29.000000	3.000000	3.000000	3.000000	3.000000	3.000000	3.000000	3.000000	3.000000	2
75%	36.75000	2024-07-04 23:39:43.642500096	33.000000	4.000000	3.000000	4.000000	3.000000	3.750000	3.000000	3.750000	4.000000	3
max	49.00000	2024-07-05 10:45:52.343000	45.000000	5.000000	5.000000	5.000000	5.000000	5.000000	4.000000	5.000000	5.000000	5
std	14.57738	NaN	5.353465	1.015191	1.075895	1.090965	0.952762	1.213697	1.007928	1.289835	1.255843	1



data_analysis.ipynb

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

+ Code + Text

By looking at the above data, we can see that the mean stress level among all 50 samples despite the gender is 3.3. Let's look at the stress level data based on the gender

```
[ ] sns.boxplot(x='Gender', y='Stress_level', data=df)
    plt.title('Stress Level by Gender')
    plt.show()
```

data_analysis.ipynb

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

Comment Share Connect Gemini

+ Code + Text

Female Male

Gender

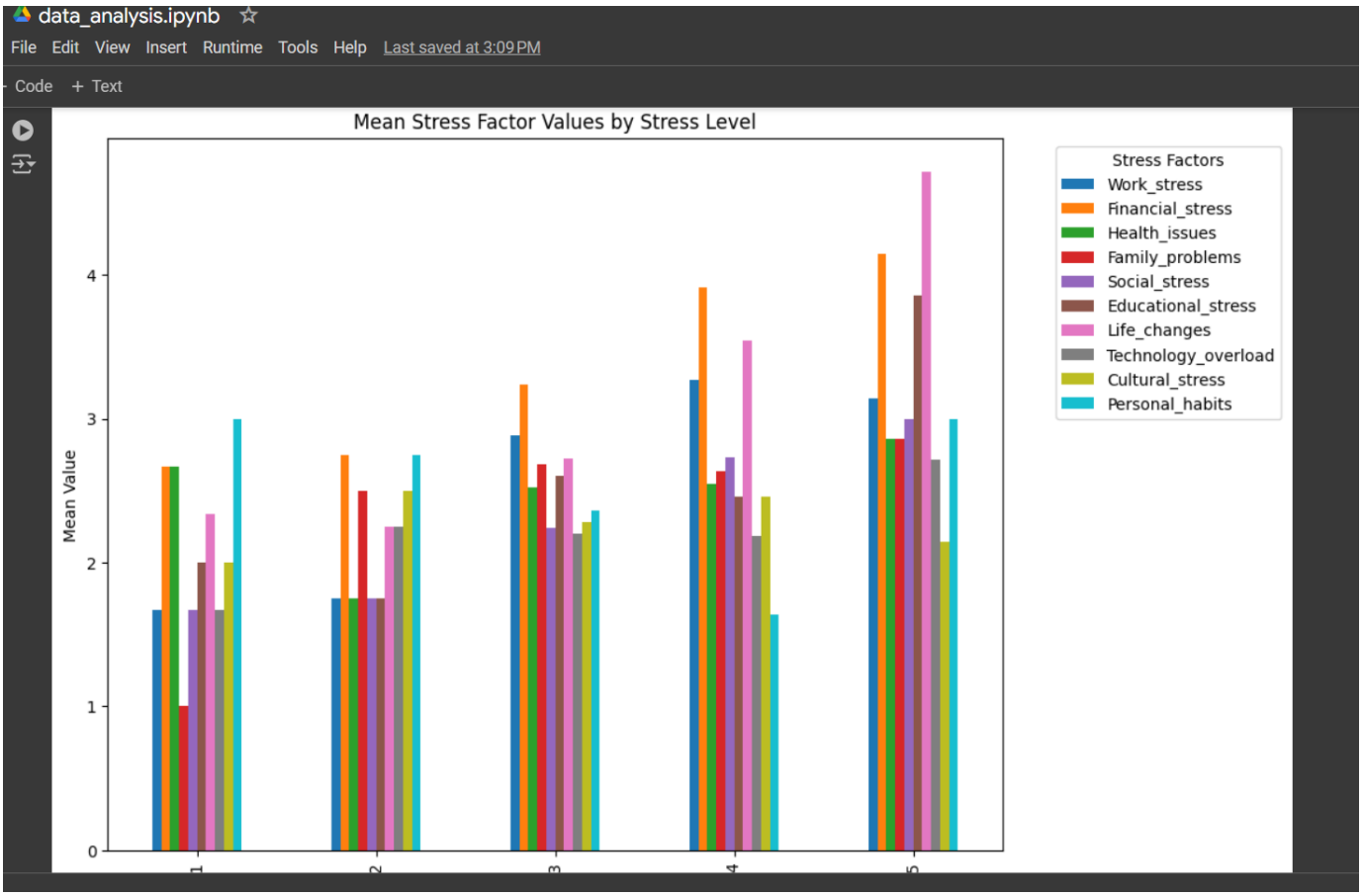
```
[ ] df.columns
```

```
Index(['Unnamed: 0', 'Timestamp', 'Age', 'Gender', 'Stress_level',
       'Work_stress', 'Financial_stress', 'Health_issues', 'Family_problems',
       'Social_stress', 'Educational_stress', 'Life_changes',
       'Technology_overload', 'Cultural_stress', 'Personal_habits', 'Country'],
      dtype='object')
```

```
[ ] factors =df[['Work_stress', 'Financial_stress', 'Health_issues', 'Family_problems',
       'Social_stress', 'Educational_stress', 'Life_changes',
       'Technology_overload', 'Cultural_stress', 'Personal_habits','Stress_level']]
```

Stress_level	Work_stress	Financial_stress	Health_issues	Family_problems	Social_stress	Educational_stress	Life_changes	Technology_overload	Cultural_stress	Personal_habits
1	1.666667	2.666667	2.666667	1.000000	1.666667	2.000000	2.333333	1.666667	2.000000	3.000000
2	1.750000	2.750000	1.750000	2.500000	1.750000	1.750000	2.250000	2.250000	2.500000	2.750000
3	2.880000	3.240000	2.520000	2.680000	2.240000	2.600000	2.720000	2.200000	2.280000	2.360000
4	3.272727	3.909091	2.545455	2.636364	2.727273	2.454545	3.545455	2.181818	2.454545	1.636364
5	3.142857	4.142857	2.857143	2.857143	3.000000	3.857143	4.714286	2.714286	2.142857	3.000000

```
[ ] stress_group.plot(kind='bar', figsize=(10, 8))
    plt.title('Mean Stress Factor Values by Stress Level')
    plt.xlabel('Stress Level')
    plt.ylabel('Mean Value')
    plt.legend(title='Stress Factors', bbox_to_anchor=(1.05, 1), loc='upper left')
```

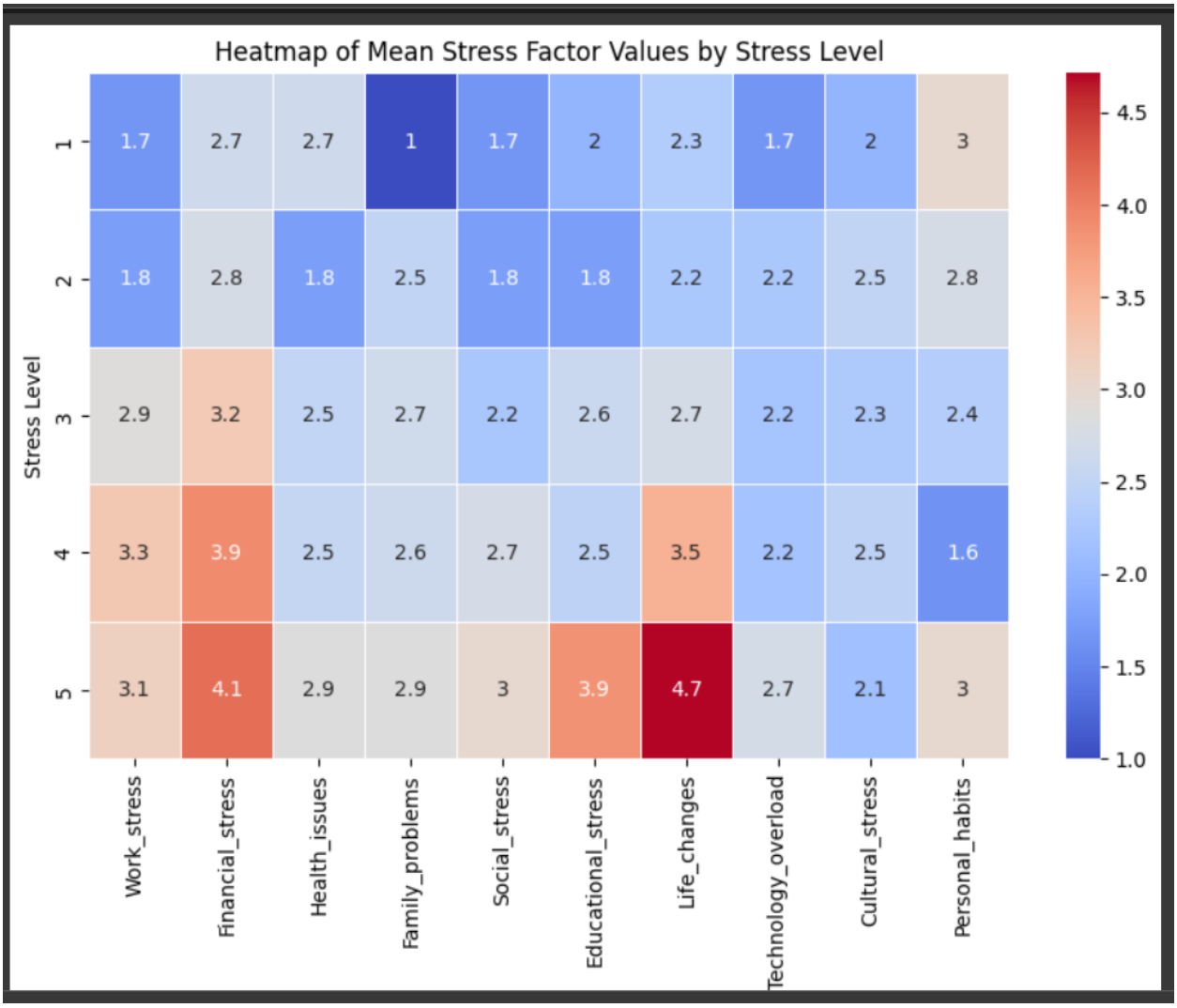


data_analysis.ipynb ☆

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

Code + Text

```
# Plot heatmap of the mean values
plt.figure(figsize=(10, 6))
sns.heatmap(stress_group, annot=True, cmap='coolwarm', linewidths=0.5)
plt.title('Heatmap of Mean Stress Factor Values by Stress Level')
plt.xlabel('Stress Factors')
plt.ylabel('Stress Level')
plt.show()
```



data_analysis.ipynb ☆

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

+ Code + Text

Box Plot Analysis

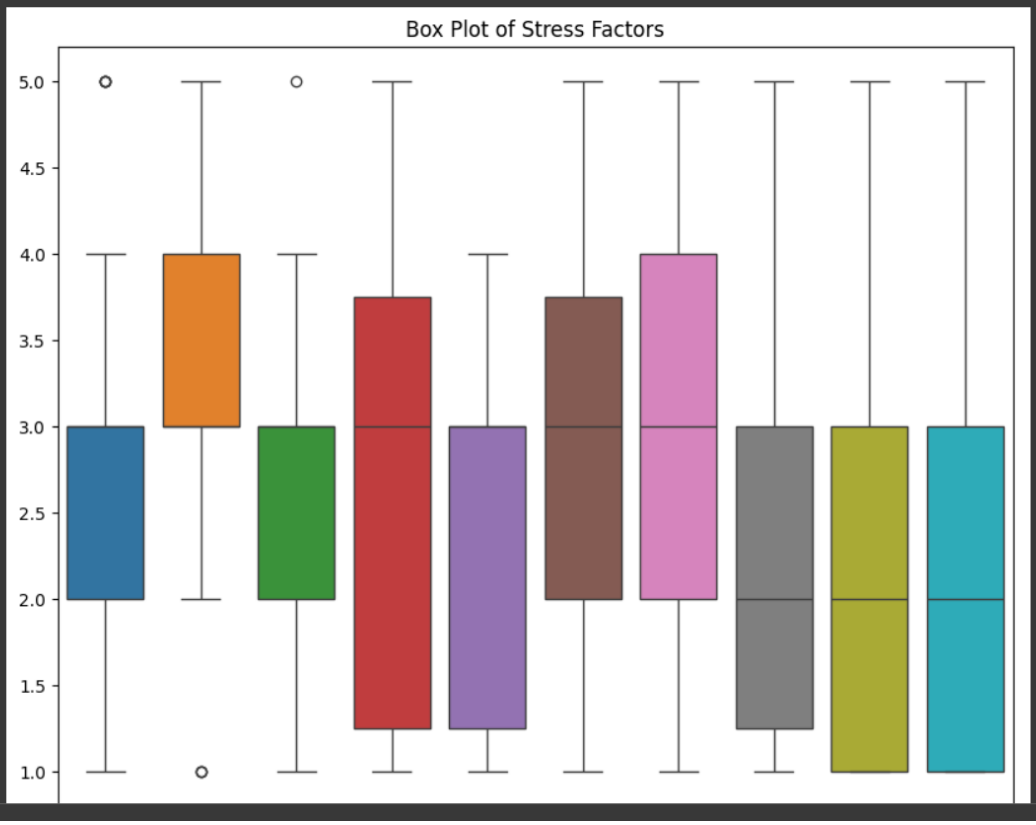
```
stress_factors =df[['Work_stress', 'Financial_stress', 'Health_issues', 'Family_problems',
       'Social_stress', 'Educational_stress', 'Life_changes',
       'Technology_overload', 'Cultural_stress', 'Personal_habits']]
```

```
[ ] # Plot box plots for each stress factor
    plt.figure(figsize=(10, 8))
    sns.boxplot(data=stress_factors)
    plt.title('Box Plot of Stress Factors')
    plt.xticks(rotation=45)
    plt.show()
```

data_analysis.ipynb ☆

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

Code + Text



data_analysis.ipynb ☆

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

Code + Text

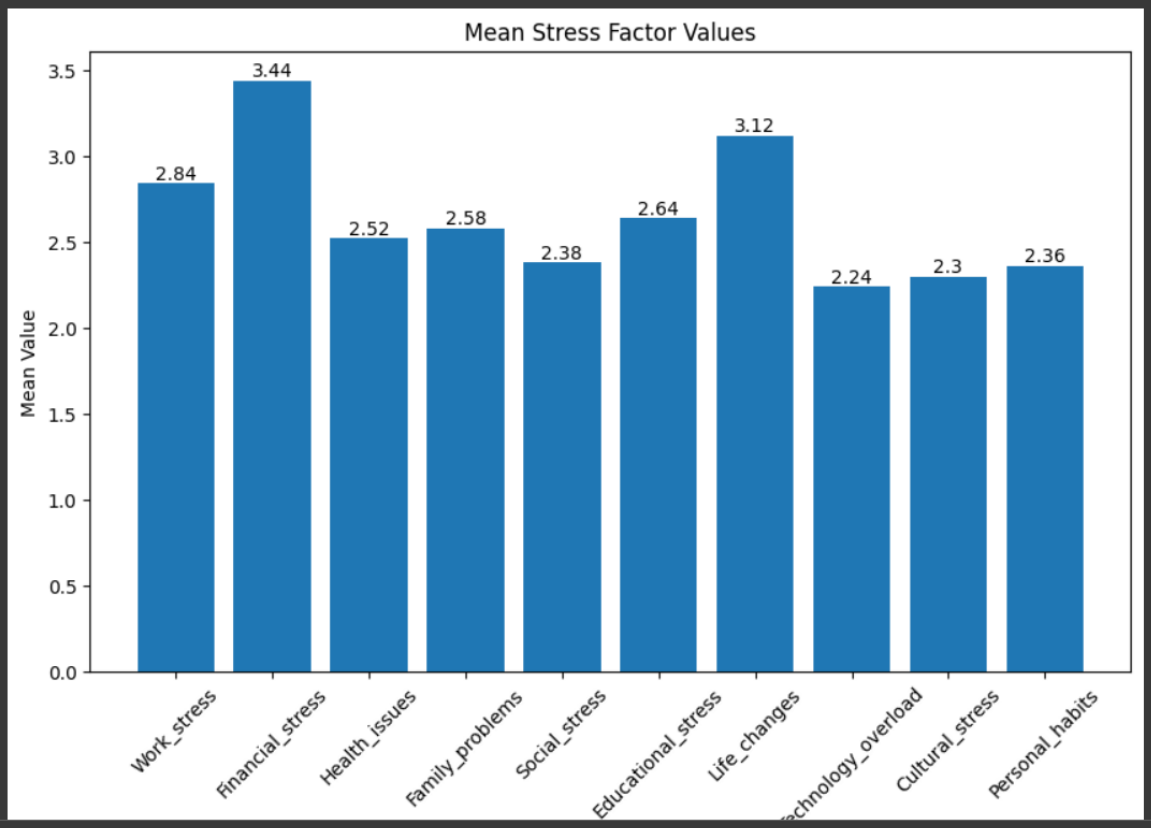
```
from inspect import Attribute
y = stress_factors.mean()
x = stress_factors.columns
plt.figure(figsize=(10, 6))
plt.bar(x, y, width=0.8)
plt.title('Mean Stress Factor Values')
plt.xlabel('Stress Factors')
plt.ylabel('Mean Value')
plt.xticks(rotation=45)

for index, value in enumerate(y):
    plt.text(index, value, str(round(value, 3)), ha='center', va='bottom')

plt.show()
```

data_analysis.ipynb ☆

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

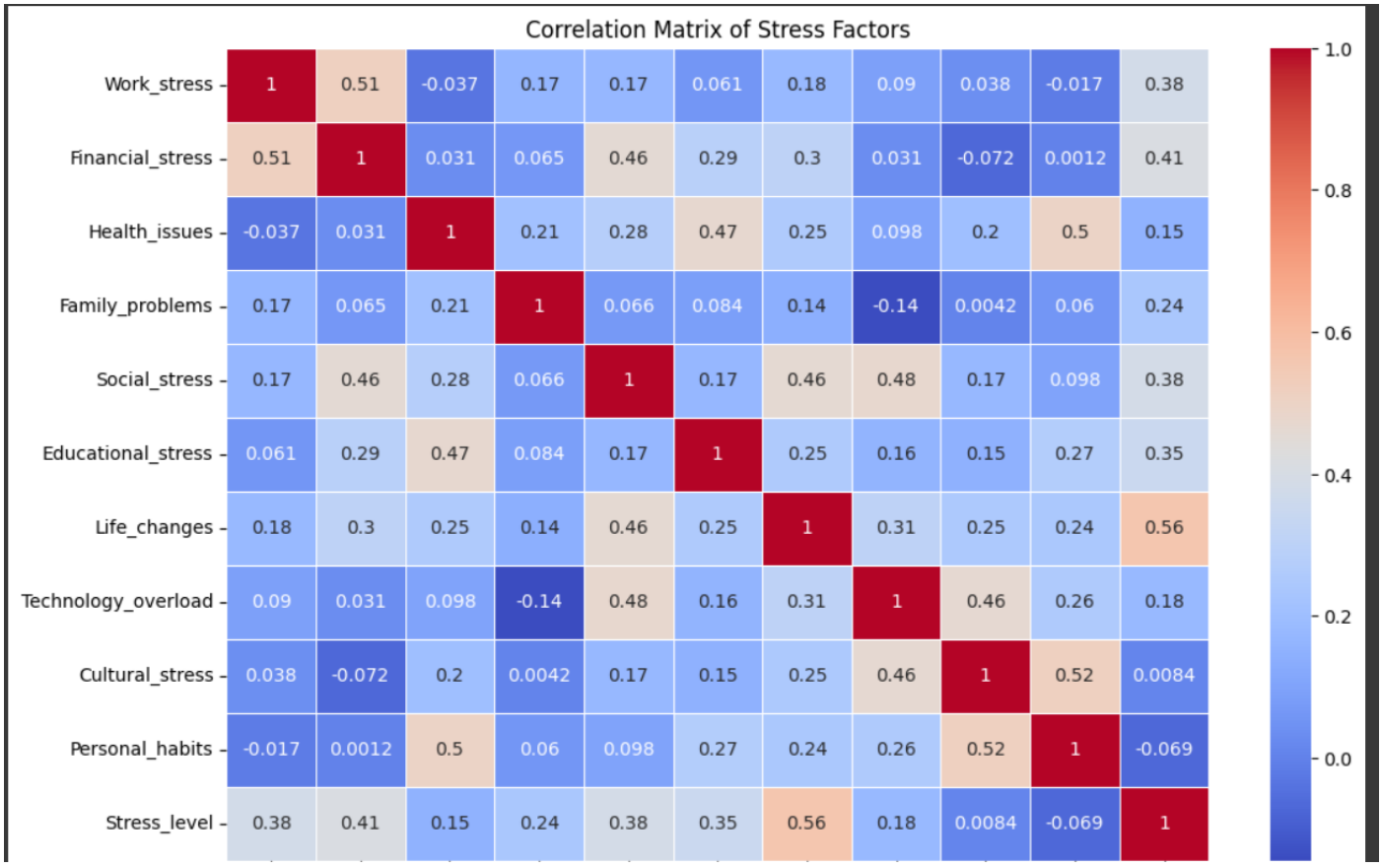
Code + Text

```
plt.show()
```



Correlation Analysis

```
[ ]  # Calculate correlation matrix
     correlation_matrix = stress_factors.corr()

     # Plot heatmap of the correlation matrix
     plt.figure(figsize=(12, 8))
     sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=0.5)
     plt.title('Correlation Matrix of Stress Factors')
     plt.show()
```



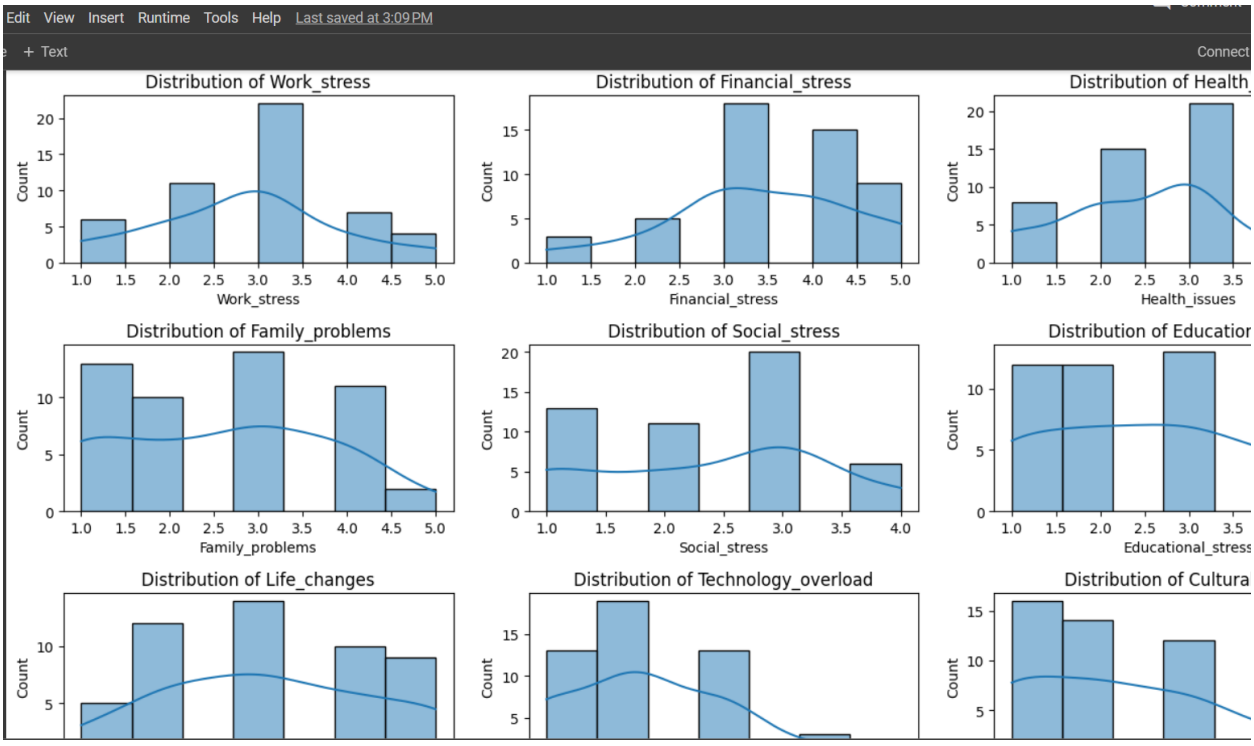
data_analysis.ipynb

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

+ Code + Text

Distribution Analysis

```
[ ]  # Plot distribution for each stress factor
     plt.figure(figsize=(14, 10))
     for i, column in enumerate(stress_factors.columns, 1):
         plt.subplot(4, 3, i)
         sns.histplot(stress_factors[column], kde=True)
         plt.title(f'Distribution of {column}')
     plt.tight_layout()
     plt.show()
```



data_analysis.ipynb

File Edit View Insert Runtime Tools Help Last saved at 3:09 PM

Comment Share

+ Code + Text

Connect Gemini

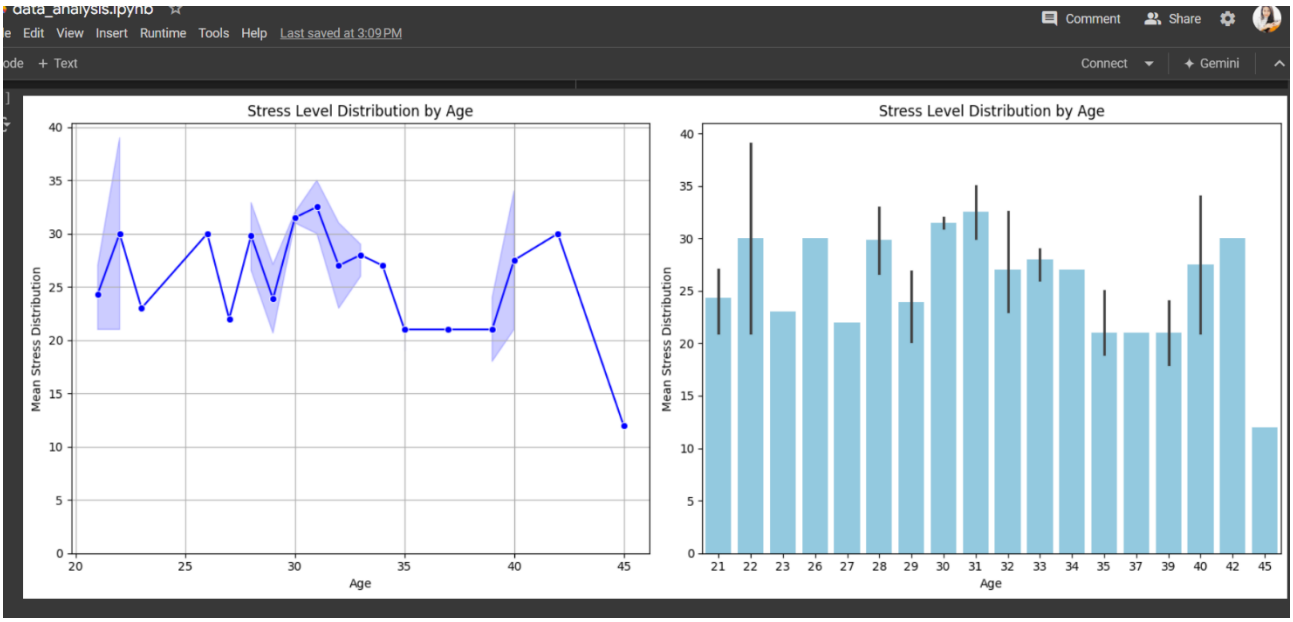
Let's find out the most stressed age of the sample I have collected

```
[ ] #calculating a value for to collect insights about the overall stress level of the individuals

    for col in stress_factors.columns:
        Toatl_stress = stress_factors.sum(axis=1)
        df['Total_stress value'] = Toatl_stress
```

```
[ ] df.head()
```

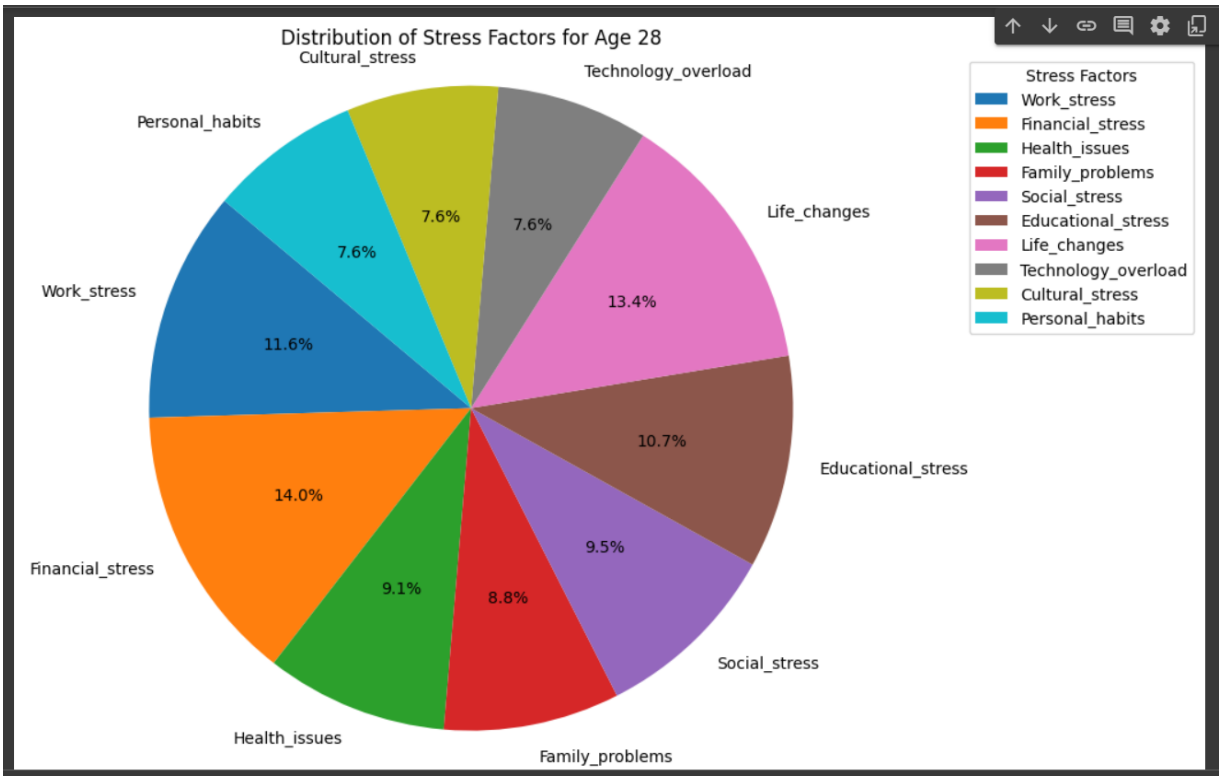
	Unnamed: 0	Timestamp	Age	Gender	Stress_level	Work_stress	Financial_stress	Health_issues	Family_problems	Social_stress	Educational_stress	Life_changes	Technology_overload
0	0	2024-07-04 12:38:08.261	28	Female	5	2	5	2	1	4	3	5	3
1	1	2024-07-04 12:45:22.941	28	Male	5	3	5	4	3	3	5	5	2
2	2	2024-07-04 12:53:16.880	28	Female	5	5	5	1	1	4	2	5	3
3	3	2024-07-04 12:55:56.942	28	Male	5	3	5	4	3	3	5	5	2
4	4	2024-07-04 12:56:44.052	33	Male	2	2	3	3	3	2	1	3	3

```
[ ] fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(15, 6))

    #plotting the line plot(left subplot)
    sns.lineplot(x='Age', y='Total_stress value', data=df, marker='o', color='b', ax=ax1)
    ax1.set_xlabel('Age')
    ax1.set_ylabel('Mean Stress Distribution')
    ax1.set_title('Stress Level Distribution by Age')
    ax1.set_ylim(0, None)  # Set the lower limit of y-axis to 0
    ax1.grid(True)

    # Plotting bar plot (right subplot)
    sns.barplot(x='Age', y='Total_stress value', data=df, ax=ax2, color='skyblue')
    ax2.set_xlabel('Age')
    ax2.set_ylabel('Mean Stress Distribution')
    ax2.set_title('Stress Level Distribution by Age')


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



The analysis of the survey data revealed that the most stressed community in the sample was individuals aged 28. This finding is significant as it sheds light on the particular challenges faced by this age group, which can contribute to higher levels of stress and anxiety. Understanding the reasons behind this elevated stress can help in developing targeted interventions to support this demographic. Characteristics of the Age 28 Community

Key Stress Factors



Individuals aged 28 often find themselves in the midst of their careers, potentially starting families, and managing increased financial responsibilities. This age typically marks a period of significant personal and professional development, bringing with it a unique set of challenges and stressors.

Key Stress Factors

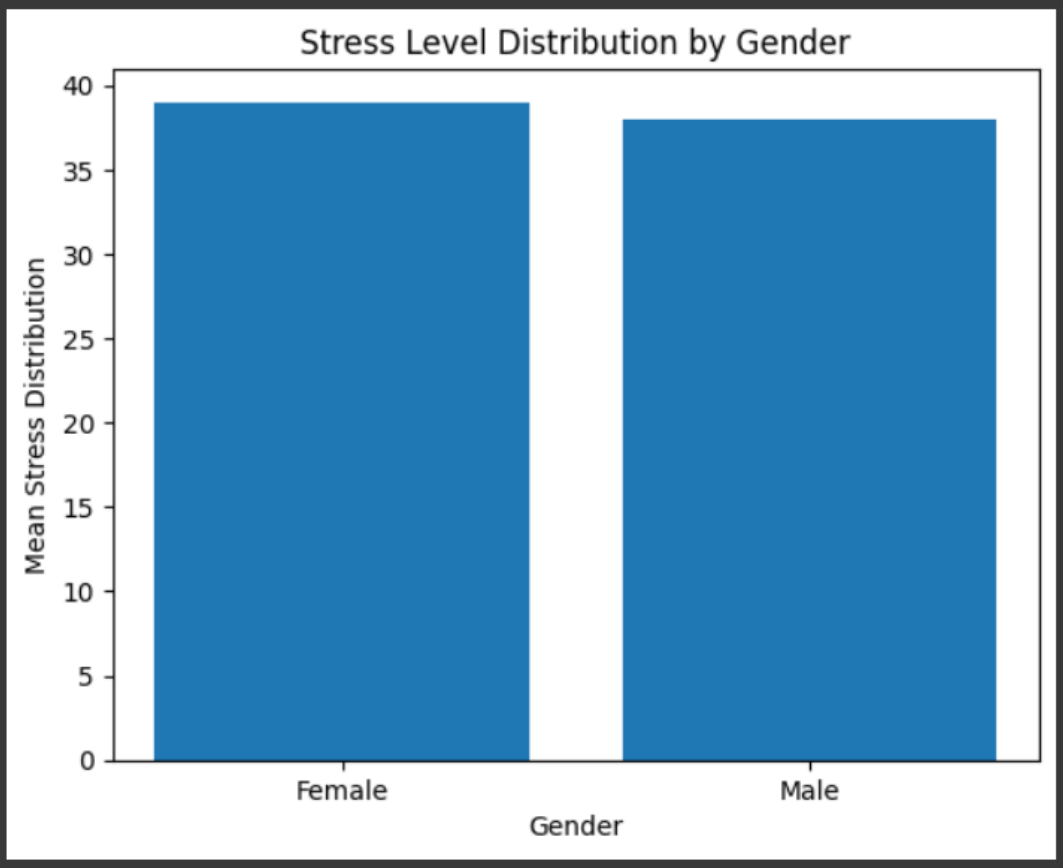
Based on the survey data and the analysis conducted, the following factors were identified as the most significant contributors to stress among 28-year-olds:

1. **Financial Stress**: This was identified as the most significant stress factor for individuals aged 28. Balancing mortgages, loans, bills, and other financial commitments can create substantial stress.
2. **Work Stress**: Career advancement, job stability, and maintaining work-life balance can be particularly challenging at this stage.

3. **Health Issues**: Both physical and mental health concerns start becoming more prominent and can add to the overall stress.
4. **Family Problems**: Managing relationships with partners, children, and extended family members can contribute to stress.
5. **Social Stress**: Navigating social relationships and maintaining social connections can be sources of stress.
6. **Educational Stress**: For those pursuing further education or professional certifications, the pressure to perform can be significant.
7. **Life Changes**: Significant life events such as marriage, parenthood, and moving can add to the stress levels.
8. **Technology Overload**: The constant connectivity and information overload from technology can also contribute to stress.
9. **Cultural Stress**: Adapting to societal expectations and cultural norms can be stressful.
10. **Personal Habits**: Unhealthy personal habits and lifestyle choices can further exacerbate stress levels.

Gender Differences in Stress Levels

```
plt.bar(df['Gender'], df['Total_stress value'])
plt.xlabel('Gender')
plt.ylabel('Mean Stress Distribution')
plt.title('Stress Level Distribution by Gender')
plt.show()
```



The analysis also examined the potential differences in stress levels based on gender. The findings indicate that there was only a slight difference in the stress levels between genders. This suggests that while the sources of stress may vary slightly, both male and female individuals aged 22 experience similar overall stress levels. This highlights the universal nature of the stressors affecting this age group, regardless of gender.

Bias in Dataset and Limitations of Results

Introduction

While analyzing the survey data, it became evident that the dataset contained a significant bias, particularly among respondents aged 20-30. This bias impacts the accuracy and generalizability of the results, leading to certain limitations in the study.

Bias in Age Distribution

The survey showed an overrepresentation of respondents aged 20-25, with a notable concentration of participants being 22 years old. This skew in age distribution can influence the findings, as the stress factors prevalent in this specific age group may not accurately reflect the broader population.

Impact on Results

1. **Age-Related Stress Factors**: The overrepresentation of younger individuals means that the stress factors identified are heavily influenced by the experiences and challenges unique to this age group. For instance, work stress and educational stress were prominent in the findings, which may be less relevant for older age groups who have settled into their careers or completed their education.
2. **Generalizability**: The results may not be applicable to older demographics or other age groups that were underrepresented in the survey. Stress factors such as health issues or family responsibilities, which may be more significant for older age groups, might not have been adequately captured.
3. **Potential Misinterpretation**: Policymakers, educators, and mental health professionals could potentially misinterpret the findings if they are assumed to be representative of the entire population. This could lead to the implementation of interventions that do not address the needs of other age groups.

Recommendations for Future Research

To mitigate the impact of bias and improve the accuracy and reliability of future studies, the following recommendations are suggested:

1. **Diverse Sampling**: Ensure that future surveys include a more diverse age range to provide a comprehensive understanding of stress factors across different demographics.
2. **Larger Sample Size**: Increasing the sample size can help in achieving a more balanced representation of various age groups, reducing the skewness in data.
3. **Stratified Sampling**: Implementing stratified sampling techniques can help ensure that specific age groups are proportionally represented in the sample, enhancing the generalizability of the findings.

4. **Continuous Monitoring**: Regularly monitor and adjust the sampling process to address any emerging biases, ensuring that the data remains representative of the target population.

Conclusion

Overview of Findings

This research aimed to identify the most significant factors contributing to stress and anxiety in society, utilizing both primary survey data and secondary data sources. The analysis revealed that individuals aged 28 experience notably high levels of stress, with financial stress emerging as the most prominent stress factor. The survey also highlighted the influence of educational, financial, social, and life changes on stress levels.

Demographic Bias

While the study provides valuable insights, it is important to acknowledge the demographic biases present in the dataset, which impact the overall accuracy and generalizability of the findings:

1. **Age Bias**: The survey data exhibited a significant bias towards respondents aged 20-30, particularly age 28. This age group was overrepresented, influencing the identification of stress factors that may be unique to younger individuals transitioning from education to the workforce. As a result, the findings may not accurately reflect the stressors affecting older age groups.
2. **Country Bias**: The majority of survey participants were from Sri Lanka and Canada, introducing a geographic bias. Cultural, economic, and societal differences between countries can significantly affect the nature and intensity of stress factors. Thus, the results may not be fully applicable to populations outside these regions.
3. **Occupation Bias**: The dataset also showed a bias towards specific occupations, particularly those related to early career stages. Occupation-related stress can vary widely across different professions and career levels, which means the findings may not comprehensively cover the stress factors faced by individuals in varied occupations and job roles.

Summary of Results

Recommendations for Future Research

To enhance the accuracy and applicability of future studies, the following recommendations are proposed:

1. **Diversified Sampling**: Ensure a more balanced representation of different age groups, countries, and occupations to provide a comprehensive understanding of stress factors across demographics.
2. **Larger Sample Size**: Increase the sample size to reduce the impact of demographic biases and enhance the reliability of the findings.
3. **Stratified Sampling Techniques**: Implement stratified sampling to ensure proportional representation of various demographic groups within the sample.

Implications for Practice

The findings underscore the need for targeted interventions to address the specific stressors faced by different demographic groups. Mental health professionals, policymakers, and educators should consider the unique challenges highlighted in this study when designing support systems and resources. By acknowledging and addressing demographic biases, future research can build on these insights to develop more effective strategies for mitigating stress and promoting mental well-being across diverse populations.

In conclusion, this research has shed light on the significant factors contributing to stress and anxiety, particularly among younger individuals. While the results are influenced by demographic biases, they offer a valuable starting point for further exploration and intervention. By refining research methodologies and expanding the scope of future studies, we can gain a deeper understanding of the complex dynamics of stress and develop more tailored and impactful solutions.

Thank You