

Untitled21.ipynb

File Edit View Insert Runtime Tools Help Last edited on 28 January

Comment Share Connect

```
+ Code + Text
```

```
1 import pandas as pd
2 import numpy as np
3 import matplotlib.pyplot as plt
4 import seaborn as sns
5 import missingno as msno
6 import warnings
7
8 %config InlineBackend.rc = {'figure.figsize': (6, 6)}
9 warnings.filterwarnings("ignore")
```

[] 1 df=pd.read_csv("/content/3 jan augmented_new_data - 3 jan augmented_new_data.csv")

[] 1 df.shape

(775, 30)

[] 1 df.head()

	Age	Gender	diagnosed with depression	symptoms of depression frequency	professional help or counseling for depression	feel stressed due to academic workload	engage in stress-relief activities	effectiveness of stress-management strategies	feel pressure to excel academically	coping with academic pressure	following factors importance at education institution	accessible counseling services	Supportive academic environment	Peer support groups	Mental health awareness campaigns	Effectiveness stress-rel techniques activit	
0	22-25	Female	Yes	3.0	Yes	Frequently	No	2.0	No	2.0	...	5.0	Important	Extremely Important	Important	Extremely Important	Friendship and relation
1	22-25	Female	Maybe	3.0	Yes	Constantly	No	3.0	Yes	2.0	...	5.0	Important	Important	Important	Important	Listening to m
2	22-25	Male	Yes	5.0	No	Sometimes	No	4.0	No	4.0	...	4.0	Important	Important	Important	Important	Y
3	22-25	Male	Yes	2.0	Yes	Frequently	Yes	2.0	Yes	3.0	...	3.0	Not Important	Neutral	Neutral	Neutral	Meditation/Mindful Listening to m
4	26-30	Male	Yes	5.0	Yes	Constantly	No	3.0	Yes	5.0	...	5.0	Important	Important	Important	Important	Meditation/Mindful Listening to m

5 rows × 30 columns

[] 1 df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 775 entries, 0 to 774
Data columns (total 30 columns):
 #   Column           Non-Null Count  Dtype  
 ---  -- 
 0   Age              773 non-null    object  
 1   Gender           773 non-null    object  
 2   diagnosed with depression      773 non-null    object  
 3   symptoms of depression frequency 774 non-null    float64 
 4   professional help or counseling for depression 773 non-null    object  
 5   feel stressed due to academic workload 773 non-null    object  
 6   engage in stress-relief activities 773 non-null    object  
 7   effectiveness of stress-management strategies 773 non-null    float64 
 8   feel pressure to excel academically 773 non-null    object  
 9   coping with academic pressure 773 non-null    float64 
 10  academic counseling or support 773 non-null    object  
 11  engaging in self-harming behaviors 773 non-null    object  
 12  professional help or counseling for self-harm 773 non-null    object  
 13  comfort in discussing mental health issues with others 773 non-null    float64 
 14  technology (e.g., NLP and HCI) can positively impact mental wellness 773 non-null    object  
 15  open to use technology based solutions to manage mental health 773 non-null    object  
 16  comfort in discussing mental health with friends or family 773 non-null    object  
 17  facing discrimination related to your mental health issues 773 non-null    object  
 18  interested in participating in workshops or programs 773 non-null    object  
 19  Preference to access mental health resources and information 774 non-null    object  
 20  following factors importance at education institution 773 non-null    float64 
 21  accessible counseling services 773 non-null    object  
 22  Supportive academic environment 773 non-null    object  
 23  Peer support groups 773 non-null    object  
 24  Mental health awareness campaigns 773 non-null    object  
 25  Effectiveness of stress-relief techniques or activities 773 non-null    object  
 26  Effectiveness of talking with someone 773 non-null    float64 
 27  contribution of lack of sleep on stress level 773 non-null    float64 
 28  engaging in activities or hobbies 773 non-null    object  
 29  primary method of coping with academic or personal stress 773 non-null    object  
dtypes: float64(7), object(23)
memory usage: 181.8+ KB
```

[] 1 # df = df.drop(['Timestamp','Name (Optional)',"What is your Educational Institution's name? (Optional)",'What type of mental health support or services do you feel are lacking at your educational institution? (Optional)'],axis=1)

Traceback (most recent call last)
<ipython-input-46-2f9cff53c6e3> in <cell line: 1>()
----> 1 df = df.drop(['Timestamp','Name (Optional)',"What is your Educational Institution's name? (Optional)",'What type of mental health support or services do you feel are lacking at your educational institution? (Optional)'],axis=1)

5 frames
/usr/local/lib/python3.10/dist-packages/pandas/core/indexes/base.py in drop(self, labels, errors)
6932 if mask.any():

```
    if errors != 'ignore':
-> 6934      raise KeyError(f"{'list(labels[mask])' not found in axis}")
6935      indexer = indexer[~mask]
6936  return self.delete(indexer)
```

	1	# df = df.drop(['If you answered "Yes" to the previous question, please briefly describe the reasons or triggers for self-harm.', 'Do you experience other mental health issues?'])	Meditation/Mindfulness/Yoga, Listening, Breathing, Relaxation, and Cognitive Behavioral Therapy																	
	2	df.head(20)	7	26-30	Female	Yes	5.0	Yes	Frequently	No	2.0	Yes	3.0	...	5.0	Extremely Important	Extremely Important	Important	Extremely Important	Meditation/Mindfulness/Yoga, Listening, Breathing, Relaxation, and Cognitive Behavioral Therapy
8	26-30	Male	Yes	4.0	No	Frequently	No	2.0	No	3.0	...	2.0	Important	Extremely Important	Important	Important	Meditation/Mindfulness/Yoga, Listening, Breathing, Relaxation, and Cognitive Behavioral Therapy			
9	22-25	Male	No	1.0	No	Sometimes	Yes	2.0	Yes	2.0	...	2.0	Not Important	Not Important	Somewhat Important	Neutral	Listening			
10	26-30	Male	Maybe	3.0	No	Sometimes	Yes	2.0	Yes	3.0	...	3.0	Not Important	Important	Neutral	Important	Breathing			
11	22-25	Male	No	3.0	No	Sometimes	No	2.0	Yes	2.0	...	5.0	Extremely Important	Extremely Important	Neutral	Important	Meditation/Mindfulness/Yoga, Listening, Breathing, Relaxation, and Cognitive Behavioral Therapy			
12	22-25	Male	No	3.0	No	Sometimes	Yes	3.0	Yes	3.0	...	3.0	Important	Important	Extremely Important	Extremely Important	Breathing, Listening			
13	26-30	Male	Maybe	3.0	No	Sometimes	Yes	3.0	Yes	3.0	...	3.0	Not Important	Somewhat Important	Somewhat Important	Neutral	Listening			
14	22-25	Female	Yes	5.0	Yes	Frequently	Yes	3.0	Yes	3.0	...	3.0	Important	Neutral	Neutral	Neutral	Meditation/Mindfulness/Yoga, Listening, Breathing, Relaxation, and Cognitive Behavioral Therapy			
15	22-25	Female	Yes	5.0	No	Frequently	No	1.0	Yes	4.0	...	3.0	Not Important	Not Important	Not Important	Extremely Important	Listening			
16	22-25	Female	Yes	4.0	No	Sometimes	Yes	3.0	Yes	3.0	...	5.0	Important	Extremely Important	Important	Important	Meditation/Mindfulness/Yoga, Listening, Breathing, Relaxation, and Cognitive Behavioral Therapy			
17	22-25	Female	Maybe	2.0	No	Frequently	No	2.0	Yes	3.0	...	5.0	Important	Important	Extremely Important	Important	Listening			
18	18-21	Female	Yes	5.0	Yes	Sometimes	Yes	3.0	Yes	3.0	...	5.0	Extremely Important	Extremely Important	Extremely Important	Extremely Important	Meditation/Mindfulness/Yoga, Listening, Breathing, Relaxation, and Cognitive Behavioral Therapy			
19	22-25	Female	No	4.0	No	Frequently	Yes	4.0	Yes	4.0	...	4.0	Important	Important	Extremely Important	Important	Meditation/Mindfulness/Yoga, Listening, Breathing, Relaxation, and Cognitive Behavioral Therapy			

20 rows × 30 columns

```
[ ] 1 new_columns = ['Age','Gender','diagnosed with depression','symptoms of depression frequency','professional help or counseling for depression','feel stressed or hopeless frequently']
2 df.columns=new_columns
```

```
[ ]    1 df.isnull().sum()
```

Age	2
Gender	2
diagnosed with depression	2
symptoms of depression frequency	1
professional help or counseling for depression	2
feel stressed due to academic workload	2
engage in stress-relief activities	2
effectiveness of stress-management strategies	2
feel pressure to excel academically	2
coping with academic pressure	2
academic counseling or support	2
engaging in self-harming behaviors	2
professional help or counseling for self-harm	2
comfort in discussing mental health issues with others	2
technology (e.g., NLP and HCI) can positively impact mental wellness	2
open to use technology based solutions to manage mental health	2
comfort in discussing mental health with friends or family	2
facing discrimination related to your mental health issues	2
interested in participating in workshops or programs	2
Preference to access mental health resources and information	1
following factors importance at education institution	2
accessible counseling services	2
Supportive academic environment	2
Peer support groups	2
Mental health awareness campaigns	2
effectiveness of stress-relief techniques or activities	2
effectiveness of talking with someone	2
contribution of lack of sleep on stress level	2
engaging in activities or hobbies	2
primary method of coping with academic or personal stress	2
dtype: int64	2

```
[ ]    1 df.dropna(inplace=True)  
      2 df.isnull().sum()
```

Age	0
Gender	0
diagnosed with depression	0
symptoms of depression frequency	0
professional help or counseling for depression	0
feel stressed due to academic workload	0

```
engage in stress-relief activities          0
effectiveness of stress-management strategies 0
feel pressure to excel academically        0
coping with academic pressure              0
academic counseling or support           0
engaging in self-harming behaviors        0
professional help or counseling for self-harm 0
comfort in discussing mental health issues with others 0
technology (e.g., NLP and HCI) can positively impact mental wellness 0
open to use technology based solutions to manage mental health 0
comfort in discussing mental health with friends or family 0
facing discrimination related to your mental health issues 0
interested in participating in workshops or programs 0
Preference to access mental health resources and information 0
following factors importance at education institution 0
accessible counseling services           0
Supportive academic environment          0
Peer support groups                      0
Mental health awareness campaigns        0
effectiveness of stress-relief techniques or activities 0
effectiveness of talking with someone    0
contribution of lack of sleep on stress level 0
engaging in activities or hobbies        0
primary method of coping with academic or personal stress 0
dtype: int64
```

```
[ ] 1 df.duplicated().sum()
```

```
→ 301
```

```
[ ] 1 # Identify and print duplicate rows
2 duplicate_mask = df.duplicated(keep=False)
3 duplicate_rows = df[duplicate_mask]
4
5 # Display the duplicate rows
6 print("Duplicate Rows:")
7 print(duplicate_rows)
8
9 # Drop duplicates from the original DataFrame
10 df_no_duplicates = df.drop_duplicates()
11
12 # Display the DataFrame without duplicates
13 print("\nDataFrame after dropping duplicates:")
14 print(df_no_duplicates)
15
16 df.drop_duplicates(inplace=True)
17 print("DataFrame after dropping duplicates in place:")
18 print(df)
```

```
→ Duplicate Rows:
   Age      Gender diagnosed with depression \
2  22-25       Male                  Yes
3  22-25       Male                  Yes
7  26-30     Female                 Yes
9  22-25       Male                  No
10 26-30      Male                Maybe
.. ...
764 22-25      Male                  Yes
765 22-25      Male                  No
767 22-25      Male                  No
770 18-21     Female                 Yes
772 Under 18 Prefer not to say      Yes
```

```
      symptoms of depression frequency \
2                           5.0
3                           2.0
7                           5.0
9                           1.0
10                          3.0
.. ...
764                          3.0
765                          3.0
767                          2.0
770                          5.0
772                          4.0
```

```
      professional help or counseling for depression \
2                           No
3                         Yes
7                         Yes
9                         No
10                        No
.. ...
764                        No
765                        Yes
767                        Yes
770                        Yes
772                        Yes
```

```
      feel stressed due to academic workload engage in stress-relief activities \
2             Sometimes                  No
3             Frequently                 Yes
7             Frequently                 No
9             Sometimes                 Yes
10            Sometimes                 Yes
.. ...
764            Frequently                 Yes
765            Frequently                 Yes
767            Occasionally               Yes
770            Constantly                 Yes
772            Frequently                 Yes
```

```
      effectiveness of stress-management strategies \
2                           4.0
3                           2.0
7                           2.0
9                           2.0
```

```
[ ] 1 df.duplicated().sum()
```

```
→ 0
```

```
[ ] 1 df.info()
```

```

[2]: <class 'pandas.core.frame.DataFrame'>
Int64Index: 472 entries, 0 to 771
Data columns (total 30 columns):
 #   Column           Non-Null Count Dtype  
--- 
 0   Age              472 non-null    object  
 1   Gender            472 non-null    object  
 2   diagnosed with depression 472 non-null    object  
 3   symptoms of depression frequency 472 non-null    float64 
 4   professional help or counseling for depression 472 non-null    object  
 5   feel stressed due to academic workload 472 non-null    object  
 6   engage in stress-relief activities 472 non-null    object  
 7   effectiveness of stress-management strategies 472 non-null    float64 
 8   feel pressure to excel academically 472 non-null    object  
 9   coping with academic pressure 472 non-null    float64 
 10  academic counseling or support 472 non-null    object  
 11  engaging in self-harming behaviors 472 non-null    object  
 12  professional help or counseling for self-harm 472 non-null    object  
 13  comfort in discussing mental health issues with others 472 non-null    float64 
 14  technology (e.g., NLP and HCI) can positively impact mental wellness 472 non-null    object  
 15  open to use technology based solutions to manage mental health 472 non-null    object  
 16  comfort in discussing mental health with friends or family 472 non-null    object  
 17  facing discrimination related to your mental health issues 472 non-null    object  
 18  interested in participating in workshops or programs 472 non-null    object  
 19  Preference to access mental health resources and information 472 non-null    object  
 20  following factors importance at education institution 472 non-null    float64 
 21  accessible counseling services 472 non-null    object  
 22  Supportive academic environment 472 non-null    object  
 23  Peer support groups 472 non-null    object  
 24  Mental health awareness campaigns 472 non-null    object  
 25  effectiveness of stress-relief techniques or activities 472 non-null    object  
 26  effectiveness of talking with someone 472 non-null    float64 
 27  contibution of lack of sleep on stress level 472 non-null    float64 
 28  engaging in activities or hobbies 472 non-null    object  
 29  primary method of coping with academic or personal stress 472 non-null    object  
dtypes: float64(7), object(23)
memory usage: 114.3+ KB

```

```
[3]: df.info()
```

```

[3]: <class 'pandas.core.frame.DataFrame'>
Int64Index: 472 entries, 0 to 771
Data columns (total 30 columns):
 #   Column           Non-Null Count Dtype  
--- 
 0   Age              472 non-null    object  
 1   Gender            472 non-null    object  
 2   diagnosed with depression 472 non-null    object  
 3   symptoms of depression frequency 472 non-null    float64 
 4   professional help or counseling for depression 472 non-null    object  
 5   feel stressed due to academic workload 472 non-null    object  
 6   engage in stress-relief activities 472 non-null    object  
 7   effectiveness of stress-management strategies 472 non-null    float64 
 8   feel pressure to excel academically 472 non-null    object  
 9   coping with academic pressure 472 non-null    float64 
 10  academic counseling or support 472 non-null    object  
 11  engaging in self-harming behaviors 472 non-null    object  
 12  professional help or counseling for self-harm 472 non-null    object  
 13  comfort in discussing mental health issues with others 472 non-null    float64 
 14  technology (e.g., NLP and HCI) can positively impact mental wellness 472 non-null    object  
 15  open to use technology based solutions to manage mental health 472 non-null    object  
 16  comfort in discussing mental health with friends or family 472 non-null    object  
 17  facing discrimination related to your mental health issues 472 non-null    object  
 18  interested in participating in workshops or programs 472 non-null    object  
 19  Preference to access mental health resources and information 472 non-null    object  
 20  following factors importance at education institution 472 non-null    float64 
 21  accessible counseling services 472 non-null    object  
 22  Supportive academic environment 472 non-null    object  
 23  Peer support groups 472 non-null    object  
 24  Mental health awareness campaigns 472 non-null    object  
 25  effectiveness of stress-relief techniques or activities 472 non-null    object  
 26  effectiveness of talking with someone 472 non-null    float64 
 27  contibution of lack of sleep on stress level 472 non-null    float64 
 28  engaging in activities or hobbies 472 non-null    object  
 29  primary method of coping with academic or personal stress 472 non-null    object  
dtypes: float64(7), object(23)
memory usage: 114.3+ KB

```

```

[4]: 1 df['symptoms of depression frequency'] = df['symptoms of depression frequency'].astype(int)
 2 df['effectiveness of stress-management strategies'] = df['effectiveness of stress-management strategies'].astype(int)
 3 df['coping with academic pressure'] = df['coping with academic pressure'].astype(int)
 4 df['comfort in discussing mental health issues with others'] = df['comfort in discussing mental health issues with others'].astype(int)
 5 df['following factors importance at education institution'] = df['following factors importance at education institution'].astype(int)
 6 df['effectiveness of talking with someone'] = df['effectiveness of talking with someone'].astype(int)
 7 df['contibution of lack of sleep on stress level'] = df['contibution of lack of sleep on stress level'].astype(int)

```

```
[5]: df.info()
```

```

[5]: <class 'pandas.core.frame.DataFrame'>
Int64Index: 472 entries, 0 to 771
Data columns (total 30 columns):
 #   Column           Non-Null Count Dtype  
--- 
 0   Age              472 non-null    object  
 1   Gender            472 non-null    object  
 2   diagnosed with depression 472 non-null    object  
 3   symptoms of depression frequency 472 non-null    int64  
 4   professional help or counseling for depression 472 non-null    object  
 5   feel stressed due to academic workload 472 non-null    object  
 6   engage in stress-relief activities 472 non-null    object  
 7   effectiveness of stress-management strategies 472 non-null    int64  
 8   feel pressure to excel academically 472 non-null    object  
 9   coping with academic pressure 472 non-null    int64  
 10  academic counseling or support 472 non-null    object  
 11  engaging in self-harming behaviors 472 non-null    object  
 12  professional help or counseling for self-harm 472 non-null    object  
 13  comfort in discussing mental health issues with others 472 non-null    int64  
 14  technology (e.g., NLP and HCI) can positively impact mental wellness 472 non-null    object  
 15  open to use technology based solutions to manage mental health 472 non-null    object  
 16  comfort in discussing mental health with friends or family 472 non-null    object  
 17  facing discrimination related to your mental health issues 472 non-null    object  
 18  interested in participating in workshops or programs 472 non-null    object  
 19  Preference to access mental health resources and information 472 non-null    object  
 20  following factors importance at education institution 472 non-null    int64  
 21  accessible counseling services 472 non-null    object  

```

```
22 Supportive academic environment          472 non-null  object
23 Peer support groups                    472 non-null  object
24 Mental health awareness campaigns      472 non-null  object
25 effectiveness of stress-relief techniques or activities 472 non-null  object
26 effectiveness of talking with someone  472 non-null  int64
27 contribution of lack of sleep on stress level    472 non-null  int64
28 engaging in activities or hobbies       472 non-null  object
29 primary method of coping with academic or personal stress 472 non-null  object
dtypes: int64(7), object(23)
memory usage: 114.3+ KB
```

```
[ ] 1 df.to_csv('new data.csv', index=False)
```

Colab paid products - Cancel contracts here

