

Index

Sr no	Content	Page No.
1	Introduction and background	2
2	Dataset	4
3	Data Story	7
4	Dashboard	14
5	Dataset Limitation	15
6	Summary and Conclusion	15
7	References	16

Introduction And Background

Nowadays, a lot of people have mental health problems like anxiety and depression as a result of several lifestyle-related causes. Because it has an unseen impact on people, it appears that the transition from the traditional to the modern era is not as seamless as we might have thought. Compare today's issues with those that people 50 years ago faced. People are too privileged today, as a result of lifestyle and knowledge growth, to respond even to pointless conflict. Thinking, mood, and/or behaviour are all impacted by certain mental diseases. Additionally, it hinders a person's functioning and quality of life.

The variety of mental disorders and diseases covered in this project includes the following:

- Schizophrenia
- Bipolar disorder
- Eating disorders
- Anxiety disorders
- Drug use disorders
- Depression
- Alcohol-consumption disorders

I learnt that mental health is one of the most neglected areas of health globally. The neglect of mental health has occurred for a number of reasons. Associative stigma is the first. The second is that mental health issues are seen as a "luxury good" as opposed to true illnesses. The fragmented and out-of-date service model is one of the extra main factors. Providing mental health services mostly in psychiatric institutions, the acute lack of preventative mental health care, the slow implementation of policy changes, and a lack of human resources are a few of these.

From the analysis I performed, I accomplished the main reasons behind the increment in the number of mental illnesses in which the few of the factors are anxiety and depression which showed

the highest proportions among the others. The other main reason that notably was the cause for the mental disorders was due to alcohol consumption. From the beginning of the year 1990 to 2017, both anxiety and depression displayed the highest prevailing factors among the victims of mental illnesses and this comes to flash on practices that may reduce mental stress, anxiety, and depression.

In order to offer potential solutions to anyone who is concerned about the situations and take proactive steps to resolve them, my ultimate goal is to analyse the dataset and provide a clear visualisation of the main factors that may be the ultimate cause of such a problem among the people residing in different countries.

The following are a set of questions which are going to be answered and visualised in order to achieve the goal: -

1. What is the most common reason for mental disorders?
2. What are the average mental health disorders in each country?
3. Over the years, what can be observed in the prevalence among the disorders?
4. Are there any relationships between mental disorders?
5. Where does it tend to be more common? Woman or Man?
6. Is there any correlation between depression and anxiety and are these factors related to suicide rates?
7. What is the tendency of Suicide over the years?
8. How many people in each nation experience mental health issues?
9. In which age groups is depression more common?

Dataset:

The dataset is extracted from the Kaggle website: [Global Trends in Mental Health Disorder | Kaggle](#). The dataset contains the data regarding mental disorders with respect to different countries.

The dataset contains 11 columns and the data description is as below:

Column name	Description
Entity	Unique identifier for each country or region included in the data set. (String)
Code	Unique code associated with an Entity/Country or region included in the data set. (String)
Year	Year that the data about that particular Entity/Country was collected. (Integer)
Schizophrenia (%)	Percentage of people with schizophrenia in that country/region during that year. (Float)
Bipolar disorder (%)	Percentage of people with bipolar disorder in that country/region during that year. (Float)
Eating disorders (%)	Percentage of people with eating disorders in that country/region during that year. (Float)
Anxiety disorders (%)	Percentage of people with anxiety disorders in that country/region during that year. (Float)
Drug use disorders (%)	Percentage of people with drug use disorders in that country/region during that year. (Float)
Depression (%)	Percentage of people with depression in that country/region during that year. (Float)
Alcohol use disorders (%)	Percentage of people with alcohol use disorders in that country/region during that year. (Float)

Entity	Code	Year	# Schizophr...	# Bipolar dis...	# Eating dis...	# Anxiety di...	# Drug use ...	# Depressio...
Afghanistan	AFG	1990	0.16056	0.697779	0.101855	4.82883	1.677082	4.071831
Afghanistan	AFG	1991	0.160312	0.697961	0.099313	4.82974	1.684746	4.079531
Afghanistan	AFG	1992	0.160135	0.698107	0.096692	4.831108	1.694334	4.088358
Afghanistan	AFG	1993	0.160037	0.698257	0.094336	4.830864	1.70532	4.09619
Afghanistan	AFG	1994	0.160022	0.698469	0.092439	4.829423	1.716069	4.099582
Afghanistan	AFG	1995	0.160076	0.698695	0.09098	4.828337	1.728112	4.104207
Afghanistan	AFG	1996	0.160249	0.698914	0.089709	4.828083	1.737643	4.1075
Afghanistan	AFG	1997	0.160554	0.699127	0.088372	4.827726	1.746891	4.110834
Afghanistan	AFG	1998	0.160931	0.699372	0.08733	4.826971	1.756963	4.114438
Afghanistan	AFG	1999	0.161311	0.699674	0.086267	4.826413	1.770791	4.117633
Afghanistan	AFG	2000	0.161621	0.700015	0.086021	4.827047	1.788395	4.11861
Afghanistan	AFG	2001	0.161957	0.700499	0.086517	4.831409	1.839123	4.121381
Afghanistan	AFG	2002	0.162414	0.701141	0.087023	4.838318	1.934326	4.124928
Afghanistan	AFG	2003	0.162916	0.70186	0.087189	4.845538	2.051106	4.12523
Afghanistan	AFG	2004	0.163377	0.702556	0.088158	4.851512	2.163044	4.126384
Afghanistan	AFG	2005	0.163706	0.703078	0.088933	4.854684	2.247443	4.126908
Afghanistan	AFG	2006	0.163977	0.703517	0.090054	4.856685	2.32102	4.128638
Afghanistan	AFG	2007	0.164202	0.703908	0.091600	4.858061	2.405564	4.130790

I collected the data from Kaggle platform from the above-mentioned link. This data set, which examines the prevalence of mental health conditions such schizophrenia, bipolar disorder, eating disorders, anxiety disorders, drug use disorders, depression, and alcohol use disorders, was gathered from nations all around the world. You can obtain insight into how these problems are affecting people's life by viewing this data in an easily understandable style, which enables a deeper comprehension of these conditions and their ramifications.

The timeline of this dataset is from the year 1990 to 2017. The dataset totally has 108553 entries. This dataset contains information about the prevalence of mental health disorders in countries across the globe

- Entity: Unique identifier for each country or region included in the data set. (String)
- Code: Unique code associated with an Entity/Country or region included in the data set. (String)
- Year: Year that the data about that Entity/Country was collected. (Integer)
- Schizophrenia (%): Percentage of people with schizophrenia in that country/region during that year. (Float)

- Bipolar disorder (%): Percentage of people with bipolar disorder in that country/region during that year. (Float)
- Eating disorders (%): Percentage of people with eating disorders in that country/region during that year. (Float)
- Anxiety disorders (%): Percentage of people with anxiety disorders in that country/region during that year. (Float)
- Drug use disorders (%): Percentage of people with drug use disorders in that country/region during that year. (Float)
- Depression (%): Percentage of people with depression in that country/region during that year. (Float)
- Alcohol use disorders (%): Percentage of people with alcohol use disorders in that country/region during that year. (Float)

Data Story:

First, we need to clean the data and eliminate the unrequired and empty data from the dataset. After then analysed the data set for duplicate data and removed them from the dataset using python.

```
import pandas as pd
import numpy as np

df = pd.read_csv('Mental health Depression disorder Data.csv', low_memory=False) #Low memory to hide the error of different datatypes.
df.head()
```

[57] ✓ 0.4s

index	Entity Code	Year	Schizophrenia (%)	Bipolar disorder (%)	\
0	Afghanistan	AFG	1990	0.16056	0.697779
1	Afghanistan	AFG	1991	0.160312	0.697961
2	Afghanistan	AFG	1992	0.160135	0.698107
3	Afghanistan	AFG	1993	0.160037	0.698257
4	Afghanistan	AFG	1994	0.160022	0.698469

	Eating disorders (%)	Anxiety disorders (%)	Drug use disorders (%)	\
0	0.101855	4.828830	1.677082	
1	0.099313	4.829740	1.684746	
2	0.096692	4.831108	1.694334	
3	0.094336	4.830864	1.705320	
4	0.092439	4.829423	1.716069	

	Depression (%)	Alcohol use disorders (%)
0	4.071831	0.672404
1	4.079531	0.671768
2	4.088358	0.670644
3	4.096190	0.669738
4	4.099582	0.669260

After importing the data into the python file, I performed analysis to eliminate redundant data from the dataset so that further analysis produces accurate results.

```
print(df.shape) #shape of data before dropping rows that are not countries
df = df.loc[~df['entity'].str.match('^(Australia|South America|Caribbean|Central African Republic|Central Asia|Central Europe|Central Europe, Eastern Europe, and Central Asia|Central Latin America|Central Sub-Saharan Africa|East Asia|Eastern Europe|Eastern Sub-Saharan Africa|High Income Asia Pacific|High middle SDI|Latin America and Caribbean|Low middle SDI|Lower middle SDI|Middle East and North Africa|North America|South Asia|South East Asia, East Asia, and Oceania|Southern Latin America|Southern Sub-Saharan Africa|Tropical Latin America|Western Europe|Western Sub-Saharan Africa)')] #select everything except this rows that are not countries
print(df.shape)
df = df.rename(columns={'entity': 'country'})

df[5705:5720]

df = df.drop(columns=['index']).reset_index()

df.loc[df['year'] == 'year']

df1 = df[15712] #first table of dataset
df2 = df[15712:15718] #second table of dataset
df3 = df[15718:15720] #third table of dataset
df4 = df[15720:] #fourth table of dataset

df2, df2.columns = df2[1:], df2.iloc[0] #df2[1:] to setting row from 1 to last and df2.iloc[0] to lock the first row as a header.
df2 = df2.iloc[1:]
df2 = df2.rename(columns={'entity': 'country'}) #changing column entity to country as well
df2.head()
```

I got rid of the unrequired data from the dataset such as the continents were mixed-up with the countries so we needed to filter the data and remove the unwanted data and also eliminated the index column as it was of no purpose in the further analysis.


```

df2, df2.columns = df2[1:] , df2.loc[0] df2[1:] to setting row from 1 to last and df2.loc[0] to lock the first row as a header.
df2 = df2.iloc[1, 3:]
df2 = df2.rename(columns= {'entity' : 'country'}) #changing column entity to country as well
df2.head()

df1, df1.columns = df1[1:] , df1.loc[0] df1[1:] to setting row from 1 to last and df1.loc[0] to lock the first row as a header.
df1 = df1.iloc[1, 3:]
df1 = df1.rename(columns= {'entity' : 'country'}) #changing column entity to country as well
df1.head()

df4, df4.columns = df4[1:] , df4.loc[0] df4[1:] to setting row from 1 to last and df4.loc[0] to lock the first row as a header.
df4 = df4.iloc[1, 3:]
df4 = df4.rename(columns= {'entity' : 'country'}) #changing column entity to country as well
df4.head()

df1 = pd.merge(df1, df2, how='left', on=['country', 'code', 'year']) #we use left join concatenate them with these three columns that have in common.
df4 = pd.merge(df4, df2, how='left', on=['country', 'code', 'year'])
c_data = pd.merge(df1, df4, how='left', on=['country', 'code', 'year'])
c_data.head()

c_data = c_data.drop(columns=['population_y', 'index']).rename(columns={'population_x' : 'population'}).dropna()

c_data.info()

```

It came to my notice that the dataset had 4 different tables which were at the following locations:

```

df1 = df[:5712] #First table of dataset
df2 = df[5712:52478] #Second table of dataset
df3 = df[52478:99244] #Third table of dataset
df4 = df[99244:] #Fourth table of dataset

```

We finally mapped the countries to their respective regions by merging the disorders data with the continents2.csv dataset and created a new column named as region and sub-region:

The data was mapped through the 'code' column in both the datasets and finally saved this data to a new csv file which I imported into tableau for analysis.

```

c_data = c_data.astype({'year': 'int', 'Schizophrenia (%)': 'float', 'Bipolar disorder (%)': 'float', 'Eating disorders (%)': 'float', 'Prevalence in males (%)': 'float',
'Prevalence in females (%)': 'float', 'Population': 'float', 'Suicide rate (deaths per 100,000 individuals)': 'float', 'Depressive disorder rates (number suffering per
100,000)': 'float', 'Prevalence - Depressive disorders - Sex: Both - Age: All Ages (number) (people suffering from depression)': 'float' })
c_data.head()

c_data = c_data.loc[c_data['country'] != 'World'] #any row different that world

#add country mapping dataframe
continent_df = pd.read_csv('continents2.csv')
continent_df.head()

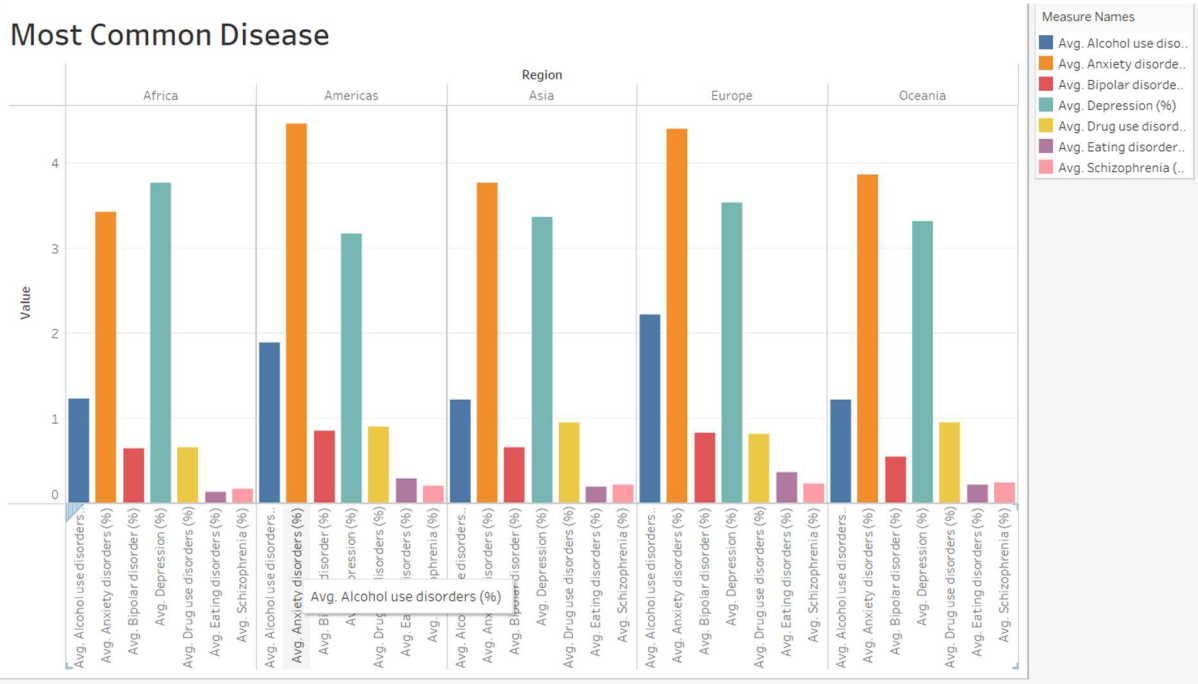
#keep necessary columns only and rename
continent_df = continent_df[['alpha-3', 'region', 'sub-region']]
continent_df = continent_df.rename(columns = {'alpha-3' : 'code', 'sub-region': 'sub-region'})
continent_df.head()

merge_df = pd.merge(c_data, continent_df, how='inner', left_on='code', right_on='code')
#check null value again
merge_df.isna().sum()

merge_df.info()

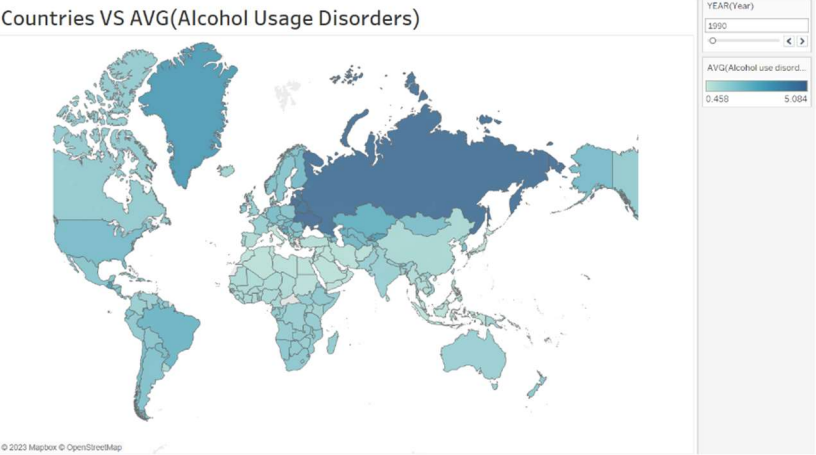
```

For answering (Q1), I have used a bar chart, which displays the most common reason for mental disorder across regions of the globe.



The bar graph illustrates the main reason being anxiety followed by depression as the leading factors for the mental illnesses and disorders.

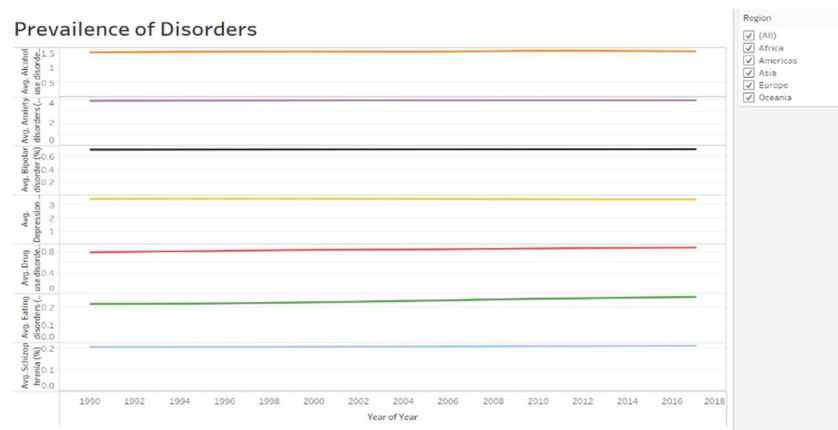
For answering (Q2), I have made various map charts for each of the factors;





Similarly plotted the maps for other prevailing factors.

For answering (Q3), I proposed a line chart that displays the prevalence of the mental disorders between the years 1990 and 2017.



We can observe that all the factors that may lead to mental disorders have a similar trend which indicates rising in the future but the difference is in the mean of people influenced by factor leading to disorders.

For (Q5), I have used the table to display the gender possessing most of the disorders by taking three data from the data set; regions, prevalence among males, prevalence among females.

The data can also be filtered according to a particular year. From this, I came to the conclusion that women are more victims of mental disorders than the men.

Total Mental disorders Prevalance Gender wise in Regions

Region	Prevalence in females	Prevalence in males
Africa	6,515	4,608
Americas	4,032	2,693
Europe	4,824	3,008
Oceania	1,497	1,106

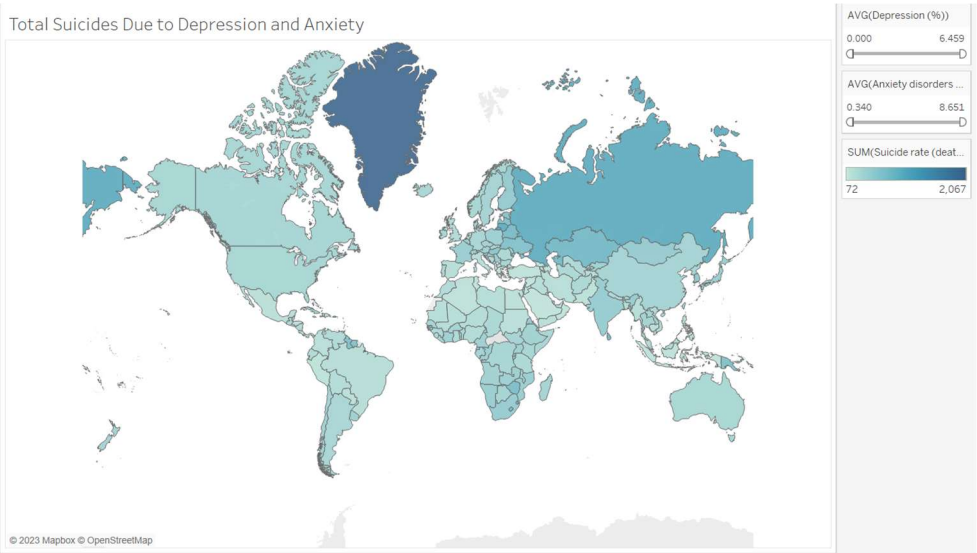
Next, I created a table which described the prevalence of mental disorders among the males and females in different regions, sub-regions and countries.

Total Mental disorders Prevalance Gender wise in Regions

Region	Sub Region	Country	Prevalence in females	Prevalence in males
Africa	Northern Africa	Algeria	121.2	83.4
		Egypt	110.3	68.5
		Libya	123.7	88.4
		Morocco	184.7	128.1
		Sudan	120.6	91.6
		Tunisia	127.1	91.2
	Sub-Saharan Africa	Angola	130.8	102.5
		Benin	118.9	80.3
		Botswana	123.7	97.3
		Burkina Faso	115.7	84.1
		Burundi	128.7	92.3
		Cameroon	123.8	83.0
		Cape Verde	122.0	100.6
		Chad	132.7	86.4
		Comoros	109.9	75.3
		Congo	132.5	99.9
		Cote d'Ivoire	109.5	75.7
		Democratic Republic of Co.	124.4	96.7
		Djibouti	117.3	85.0
		Equatorial Guinea	130.0	102.2
		Eritrea	123.5	90.7
		Ethiopia	121.8	99.1
		Gabon	125.3	104.6
		Gambia	161.1	72.9
		Ghana	108.8	79.0
		Guinea	120.5	74.8
		Guinea-Bissau	122.9	76.6
		Kenya	116.9	84.9
		Lesotho	173.6	124.7

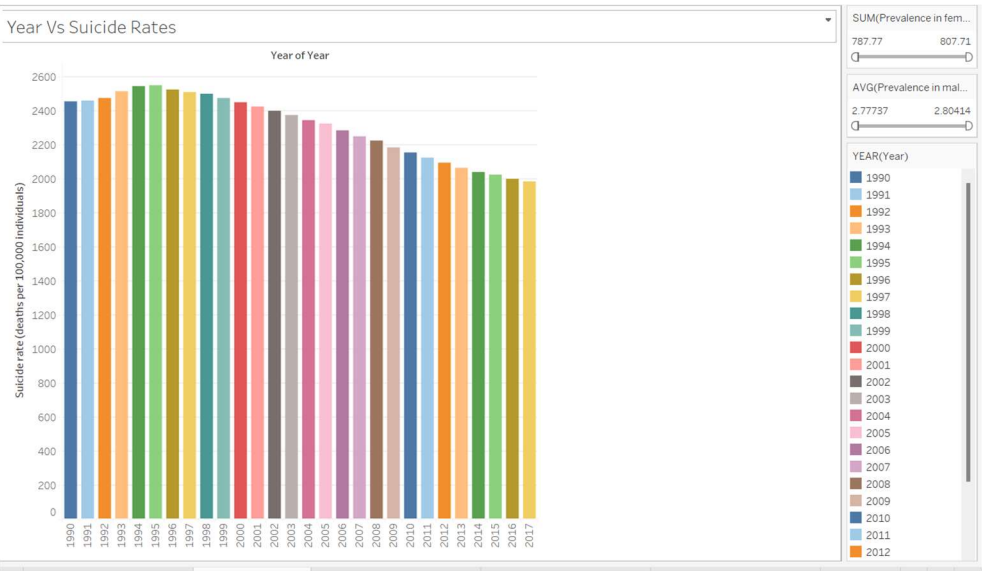
For question(Q6), I plotted the world map regarding suicide rates and to check the correlation between suicide rates and depression and anxiety, I filtered the data using the average of depression

and anxiety and found out that depression and anxiety widely affect the suicide rates as they may be the main reasons for suicide.

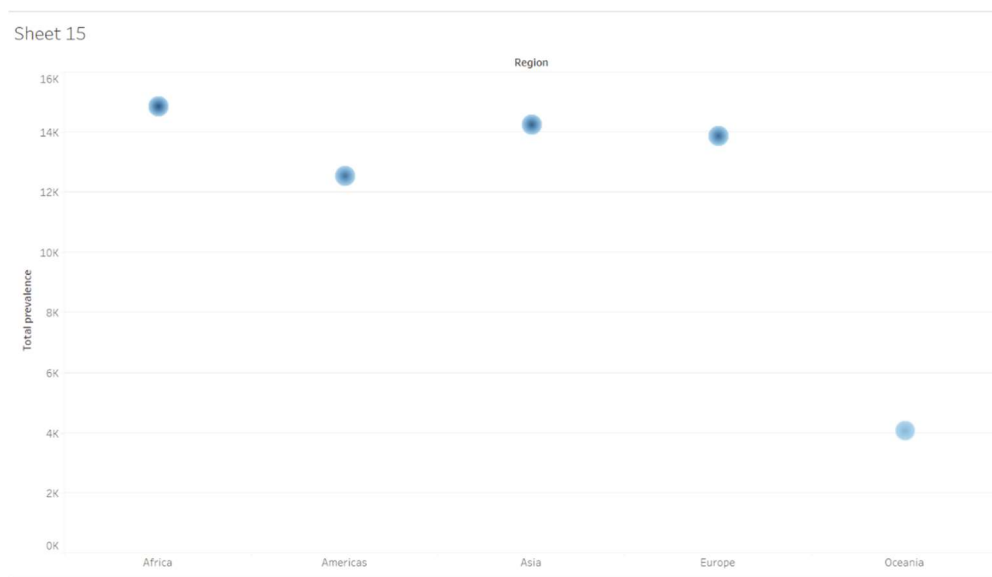


Q7 requires the tendency of the suicide rates over the years so here I have used a bar chart to depict the tendency over the years starting from 1990-2017.

I have placed the years' data into columns and the suicide rates into the rows. We can clearly observe that over the years there is a downfall in the number of suicide rates



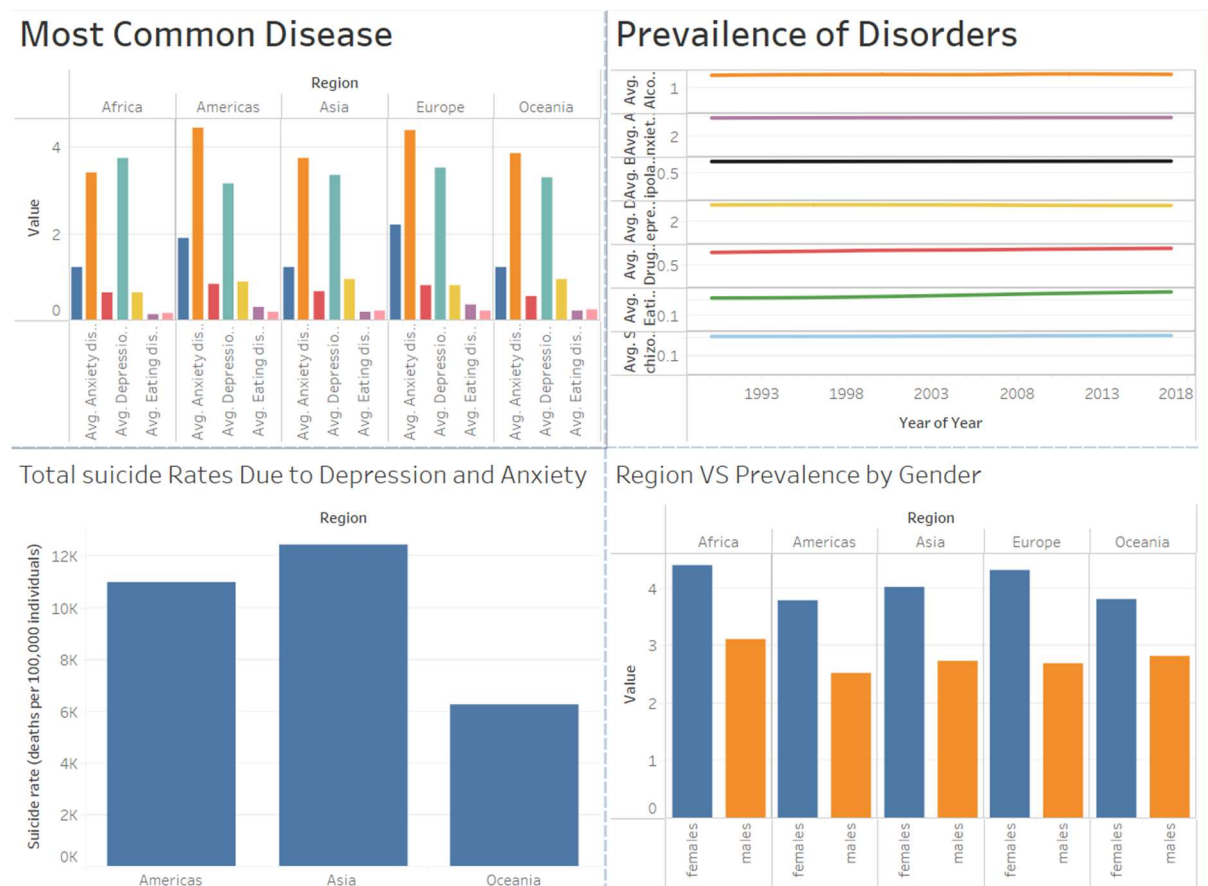
Q8 describes the total number of people experiencing the mental disorders over different regions for which I have plotted a density chart showing bubbles depicting the total number of people with disorders. Here we can conclude that Africa has the highest number of population with mental disorders.



Dashboard:

The dashboard is an elevated view of business data from which the reports are generated. With the help of innovative and real-time visualisation, dashboarding in Tableau offers excellent performance and enables non-technical users to access and utilise the data. The dashboard, which is made up of worksheets, enables simultaneous analysis of numerous points of view.

Here I have imported 4 worksheets into the dashboard, Most common disease, prevalence of disorders over the years, total suicide rates due to depression and anxiety, prevalence among the gender in various regions.



Dataset Limitation:

- Data outdated: I have not found the newest updated dataset so use the data from 2017. It is not showing the current status, I will update it when having a suitable dataset
- Many shortcomings: It had many outliers or unrequired fields so the analysis was limited.

Summary and Conclusion:

From the above performed analysis we can conclude that depression and anxiety are the main factors that led to the mental disorders among the people and so we can take steps to reduce the depression and anxiety among the people. I have also found that the main victim of this are the females so we need to focus more on the measures of reducing the mental stress and anxieties among the women.

We have found the total number of the population suffering from mental disorders and the highest can be observed in Africa. So, by this analysis, It may help the WHO as well as the nations' government to take necessary steps in order to decrease the proportion of increasing mental illnesses and disorders among the citizens

References:

Beach MC, Gary TL, Price EG, et al. Improving health care quality for racial/ethnic minorities: a systematic review of the best evidence regarding provider and organization interventions. *BMC Public Health*. 2006;6:104. [[PMC free article](#)] [[PubMed](#)]

Bebbington PE, Dean C, Der G, et al. Gender, parity and the prevalence of minor affective disorder. *The British Journal of Psychiatry*. 1991;158:40–45. [[PubMed](#)]

Beck AT. BDI-II: Beck Depression Inventory Manual. 2nd edn. Boston, MA: Harcourt Brace; 1996.

Bee PE, Bower P, Lovell K, et al. Psychotherapy mediated by remote communication technologies: a meta-analytic review. *BMC Psychiatry*. 2008;8:60. [[PMC free article](#)] [[PubMed](#)]

Cooper S, Smiley E, Morrison J, et al. Mental ill-health in adults with intellectual disabilities: prevalence and associated factors. *The British Journal of Psychiatry*. 2007;190:27–35.

Jones, P. B. (2013). Adult mental health disorders and their age at onset. *British Journal of Psychiatry*, 202(s54), s5–s10. <https://doi.org/10.1192/bjp.bp.112.119164>