

The relationship between mental health, suicide and the consumption of licit and illicit substances.

Carlos Gomes¹ and Ian Machado¹

¹ University of Minho, Braga 4710-057, Portugal

Motivation and Objectives

Over the last three decades, substance abuse and mental health interconnection has turned out to be an increasing and important concern. The study, therefore, seeks to collate and organize the related data to this problem, enhancing a visual view and understanding. The paper also aims to aid in research, analysis, and public awareness due to an indispensable interrelation with the studied factor and trends of this kind of variable over the years. This is essential comprehensive information in developing effective intervention strategies and public health policies aimed at improving the quality of life for persons affected by these issues.

The study will thus use various advanced tools in database and visualization. The database to be selected is MongoDB, a NoSQL-type database that is seen to be flexible and adequate for handling complex and unstructured data relevant to substance abuse and mental health. This allows storage, retrieval, and management of various data in large volumes. The use of Power BI in the current study was deemed necessary, especially regarding the powerful visualization and data analysis abilities that this software application possesses. This would be through the development of dashboards and interactive reports which communicate insights and trends clearly to stakeholders, policy-makers, and the general public. In combination with these two tools, the study seeks to provide a clear and detailed picture of the evolution between substance abuse and mental health in connection, hence contributing to valuable information designed for the development of intervention and public health policies.

1 Methodology and technical parameters

1.1 Collecting the Data

Firstly, the information considered useful for comparison on the topics under analysis was collected. Using the “Our World in Data” website (<https://ourworldindata.org>), datasets were selected for further analysis. The data sets selected were as follows:

- Burden of disease from each mental illness

(<https://ourworldindata.org/grapher/burden-disease-from-each-mental-illness>)

- Share with drug use disorders
(<https://ourworldindata.org/grapher/share-with-drug-use-disorders>)
- Death rate from suicides
(<https://ourworldindata.org/grapher/death-rate-from-suicides-gho>)
- Total alcohol consumption per capita
(<https://ourworldindata.org/grapher/total-alcohol-consumption-per-capita-litres-of-pure-alcohol>)

1.2 Data storage

After collecting the data, it was decided to use MongoDB to store it. MongoDB's ability to handle large volumes of data and its compatibility with Power BI, which has native connectors for this platform, facilitates data integration for future analyses.

1.3 Data Processing and Transformation

The data was processed and transformed to ensure its quality for analysis. Null and/or missing values were identified and removed, columns and rows that were not relevant to the analysis were eliminated and the data was normalized to facilitate comparison between different data sets.

The first step was to install and import some packages and modules that will be used later.

```
Package installation:
# pip install pyspark
# pip install pandas
# pip install "pymongo[srv]"

Requirement already satisfied: pyspark in c:\users\carla\anaconda3\lib\site-packages (3.5.0)
Requirement already satisfied: py4j==0.10.9.7 in c:\users\carla\anaconda3\lib\site-packages (from pyspark) (0.10.9.7)
Note: you may need to restart the kernel to use updated packages.

Import of the modules:

import time
from functools import reduce
import pandas as pd
import pyspark
from pyspark.sql import SparkSession
from pymongo.mongo_client import MongoClient
from pymongo.server_api import ServerApi
```

Fig. 1. Install and import packages.

After the import, the chosen datasets were associated with a variable that will be used throughout the work.

```
# Reading csv files
file1 = pd.read_csv(url1)
file2 = pd.read_csv(url2)
file3 = pd.read_csv(url3)
file4 = pd.read_csv(url4)
```

Fig. 2. Chosen datasets.

Initially, we tested Pandas to process the data. Time was also used to see how long processing took.

```
# Starts counting the execution time
start = time.time()

# List containing the DataFrames to be combined
Data_Frames = [file1, file2, file3, file4]

# Uses the reduce function to merge all the DataFrames in the Data_Frames list
# The join is made using the columns 'Entity', 'Year' and 'Code' as join keys
combdados = reduce(lambda left, right: pd.merge(left, right, on=['Entity', 'Year', 'Code']), Data_Frames)

# Gets a list of all unique entities/years in the combined DataFrame
países_comuns = combdados['Entity'].unique()
anos_comuns = combdados['Year'].unique()

# Filters the combined DataFrame to include only rows where 'Entity' is in common_countries and 'Year' is in common_years
df_final = combdados[combdados['Entity'].isin(países_comuns) & combdados['Year'].isin(anos_comuns)]

# Ends the execution time count
end = time.time()

# Calculates the total execution time
tempo_pandas = (end - start)

print(tempo_pandas)
df_final
```

Fig. 3. Data processing by Pandas.

Filtering was carried out to ensure that there was only data from countries that appeared at least five times in the 'Entity' column.

```
# Count the number of occurrences of each 'Entity' in the final DataFrame and store in 'z'.
# 'z' will be a boolean series where the value is True if the count is less than 5
z = df_final.Entity.value_counts() < 5

# Checks if there are any entities with less than 5 occurrences, if so, filters the final DataFrame to exclude these entities
if z.any():
    df_final = df_final[~df_final.Entity.isin(z[z].index)]
else:
    pass

# Resets the index of the final DataFrame
df_final = df_final.reset_index()

df_final
# Removes the 'index' column
df_final = df_final.drop(columns=['index'])

df_final

# Counts the number of occurrences of each value in the 'Entity' column of the 'df_final' DataFrame
# and compares whether each count is equal to 5.
df_final.Entity.value_counts() == 5

# Convert the DataFrame 'df_final' to CSV format and save the file with the name 'Data_Frame_Pandas.csv'
df_final_csv = df_final.to_csv('Data_Frame_Pandas.csv')
```

Fig. 4. Filtering for country.

Then we did the same procedure for PySpark.

Initially, we tested PySpark to process the data. Time was also used to see how long processing took.

```
# Create a new SparkSession with the application name 'Project' or get an existing session
spark = SparkSession.builder.appName('Projeto').getOrCreate()

# Import the dataframes
# The command 'option(header, true)' is important to define the title of the columns, which by default would be c_0, c_1, etc...
# inferSchema=True is important for transforming data that is assumed to be str into int

df_file1 = spark.read.option('header', 'true').csv("C:\\Users\\carla\\OneDrive\\Ambiente de Trabalho\\Universidade\\Mestrado\\18 Au
df_file2 = spark.read.option('header', 'true').csv("C:\\Users\\carla\\OneDrive\\Ambiente de Trabalho\\Universidade\\Mestrado\\18 Au
df_file3 = spark.read.option('header', 'true').csv("C:\\Users\\carla\\OneDrive\\Ambiente de Trabalho\\Universidade\\Mestrado\\18 Au
df_file4 = spark.read.option('header', 'true').csv("C:\\Users\\carla\\OneDrive\\Ambiente de Trabalho\\Universidade\\Mestrado\\18 Au

# Prints the schema (structure of columns and their types) of the DataFrame df_file1
# If we don't specify, the data type will be assumed to be a string
df_file1.printSchema()
# Shows the type of data present in DataFrame df_file1
df_file1

# Show the first 5 lines of the DataFrame
df_file1.show(5)
```

Fig. 5. Data processing by Pyspark.

The datasets were joined according to the columns 'Entity', 'Code', 'Year' and then the processing time was analyzed.

```
# Starts counting the execution time
start = time.time()

# Performs internal joins of DataFrames df_file1, df_file2, df_file3 and df_file4 using the columns 'Entity', 'Code' and 'Year'
df_joined = df_file1.join(df_file2, ['Entity', 'Code', 'Year'], "inner").join(df_file3, ['Entity', 'Code', 'Year'], "inner").join(df_file4, ['Entity', 'Code', 'Year'], "inner")

# Ends the execution time count
end = time.time()

# Calculates the total execution time
tempo_spark = (end - start)

# Prints the execution time
print(tempo_spark)
```

Fig. 6. Merging the data.

Then we decided to compare the time required by each of the libraries

```
print(f'Using PySpark it takes {tempo_spark} seconds!')
print(f'Using Pandas it takes {tempo_pandas} seconds!')
```

Using PySpark it takes 0.041808128356933594 seconds!
Using Pandas it takes 0.02780604362487793 seconds!

Fig. 7. Comparison of times between libraries.

Once the data has been processed, we place it in our MongoDB database.

```
# Connection String provided by MongoDB - Replace <password> with the real password
uri = "mongodb+srv://carlosfgomes0:<password>@bigdatadb.qzzcrlk.mongodb.net/?retryWrites=true&w=majority&appName=BigDataDb"

# Create a new client and connect to the MongoDB server
client = MongoClient(uri, server_api=ServerApi('1'))

# Confirm that the connection was successful
try:
    client.admin.command('ping')
    print("Pinged your deployment. You successfully connected to MongoDB!")
except Exception as e:
    print(e)

# Accessing the database
db = client.get_database("BigDataDb")

# Access the collection
collection = db.get_collection("BigData")

# Read the csv created by Pandas and assign it to a variable
data = pd.read_csv('Data_Frame_Pandas.csv')

# Insert the data into the MongoDB collection
collection.insert_many(data_dict)

# Close the connection
client.close()
```

Fig. 8. Data migration to MongoDB.

1.4 Visualization and Analysis

After processing the data, Power BI resources were used to build and visualize charts that contributed to data analysis. These visualizations were essential for comparisons, identifying patterns, trends and relationships between datasets.

2 Results

This chapter will illustrate potential interactions between several variables in the dataset through some charts. These charts will be given globally, across 178 countries, for the period between 2000 and 2018. However, this is a wide analysis and may not be very accurate because external forces have an effect on the results, such as wars, economic problems, and natural disasters.

As this 'noise' cannot be eliminated completely, the following measures have been put in place to avoid interference of these variables to the highest possible degree, which could change the results of the parameters being studied. Three countries have been chosen for an individual detailed examination as and when needed. These three countries are: Nigeria, Morocco, North Macedonia.

These countries have been chosen to enhance a proper understanding of the data, as they fall in different regions and their socio-economic backgrounds are quite different. The study will take place with a more specific focus on these nations in order to cancel out the extraneous variables and determine a clearer perspective on the relationship between substance abuse and mental health across different contexts. The specific analysis will, therefore, support the global one, which will be more comprehensive and give a wider perspective on the interaction of these variables across different environments and conditions.

2.1 Alcohol consumption

It is widely known that excessive alcohol consumption is theoretically connected with momentary problems that affect the individual. That is why we have decided to analyze other parameters, which may also be related to these problems, in order to find out if there is the possibility of existence of correlations between the following. The correlations are: Alcohol Consumption and Number of Suicides; Alcohol Intake and Depression Cases.

These correlations will be investigated with statistical studies to determine if there is a connection between alcohol consumption and the increase in suicide and depression rates. The analysis will seek to understand if the increase in alcohol consumption is associated with a rise in mental health problems, thus contributing to a better understanding of the impact of excessive alcohol consumption on the mental health of individuals.

Alcohol Consumption - Number of Suicides. We started by looking at alcohol consumption and the number of suicides on a global scale, as shown in Figure 9:

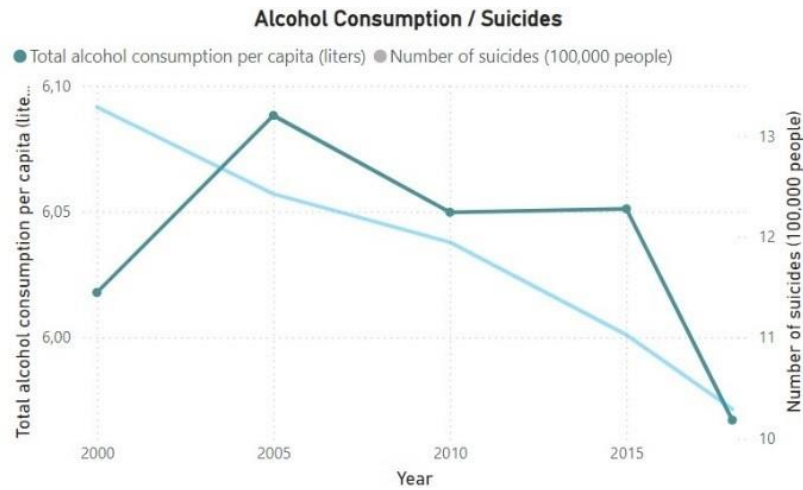


Fig. 9. A worldwide analysis of alcohol consumption and number of suicides between 2000 and 2018.

The chart reveals a slight increase in alcohol consumption between 2000 and 2005, which is not aligned with the figures for the number of suicides, which decreased. However, after 2005, the rates showed a similar decline, getting closer and closer over the years.

To investigate further, it was decided to analyze the three chosen countries individually. This approach will allow for a more detailed understanding of the specific trends in each country regarding alcohol consumption and suicide rates. By examining the data from each nation separately, it will be possible to identify more precise patterns and correlations, taking into account the particularities of each socio-economic and cultural context.

The comparison for Nigeria can be seen in Figure 10.

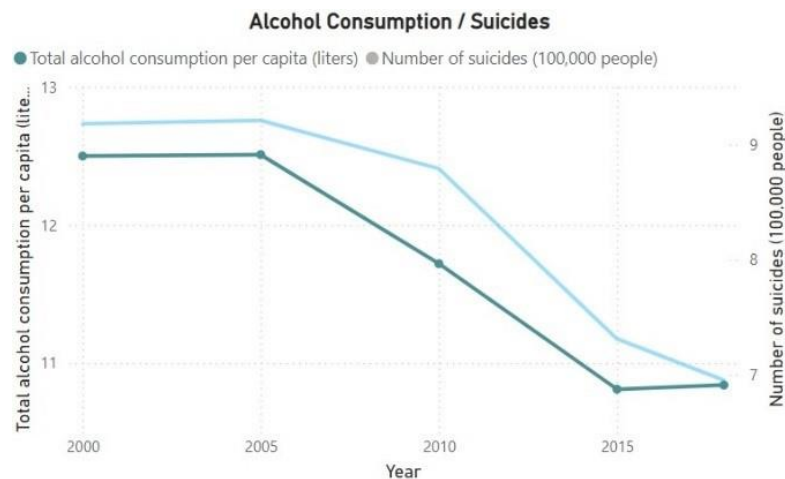


Fig. 10. An analysis of alcohol consumption and the number of suicides in Nigeria between 2000 and 2018.

Looking at the chart above, you can see that the data for both variables follow the same trend over the years. This could be a sign that the parameters are closely related.

We then analyzed the chart for Morocco:

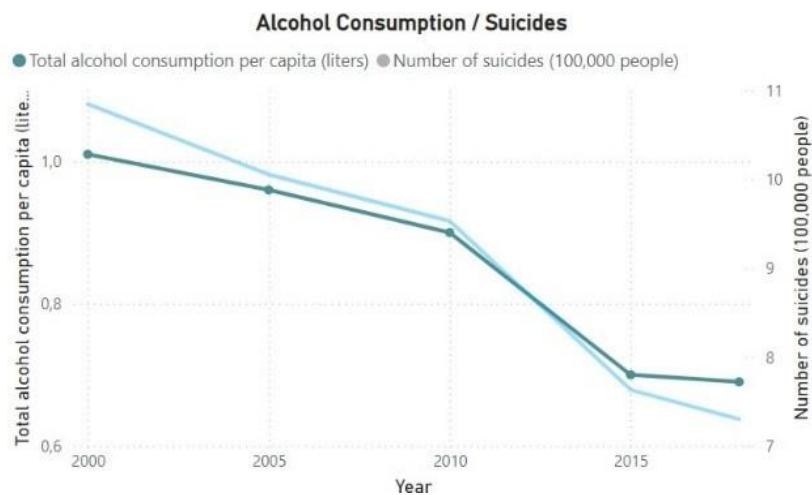


Fig. 11. An analysis of alcohol consumption and the number of suicides in Marrocco between 2000 and 2018.

Looking at the chart, we can see an enormous similarity in the evolution of the values of the variables over the years. Once again, this may indicate that suicide and alcohol consumption are strongly related.

Finally, the chart for North Macedonia was analyzed:

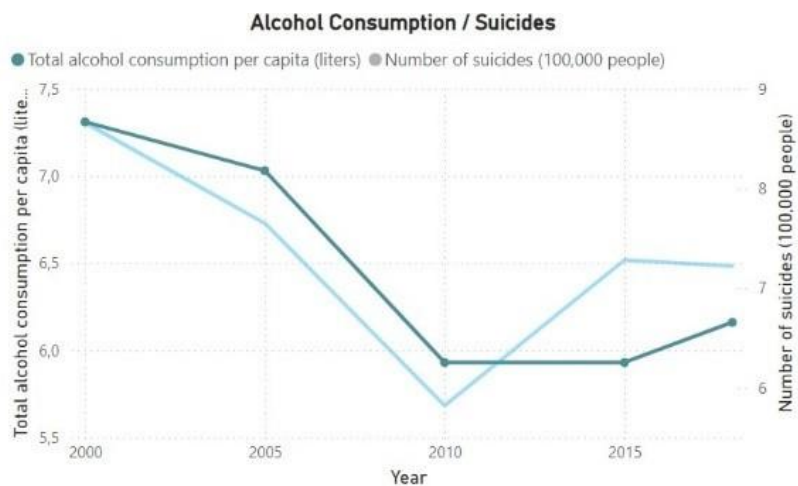


Fig. 12. An analysis of alcohol consumption and the number of suicides in North Macedonia between 2000 and 2018.

This chart also shows a similar trend in the values for the two variables, with only slight deviations from 2010 onwards. However, due to the major similarity, this chart can also be an insight into the relationship between suicide and alcohol consumption.

Depression - Alcohol consumption. After the previous analyses, we tried to find a connection between alcohol consumption and depression in the data. Figure 13 shows the evolution of alcohol consumption per capita and the number of cases of depression per 100,000 inhabitants worldwide.

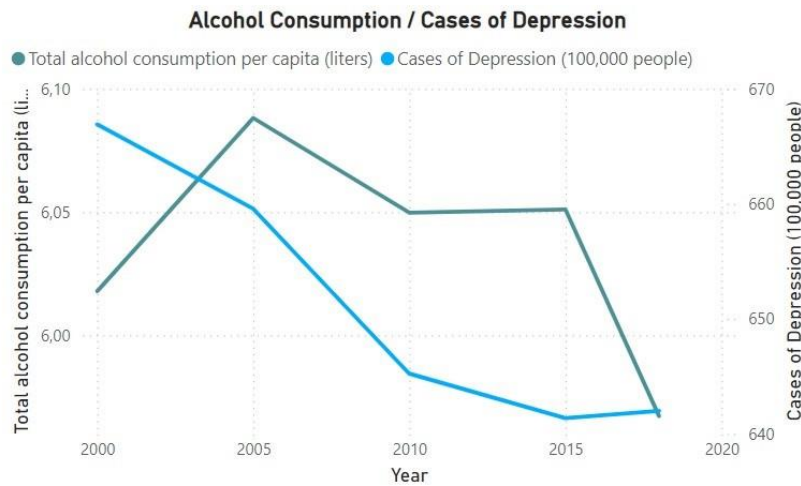


Fig. 13. A worldwide analysis of alcohol consumption and cases of depression between 2000 and 2018.

In the Figure, although there is a convergence in values between 2010 and 2018, there is no clear interaction between the two parameters under analysis. It was therefore decided to carry out an individual analysis of the countries chosen upstream. We started with the analysis of Nigeria.

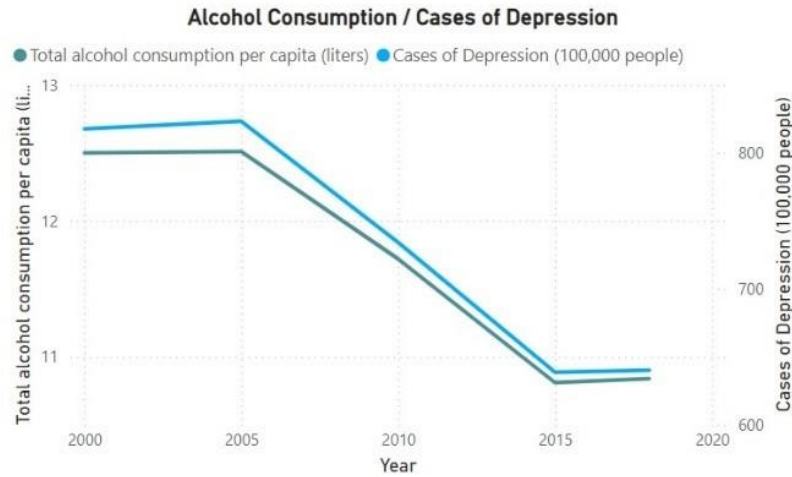


Fig. 14. An analysis of alcohol consumption and cases of depression in Nigeria between 2000 and 2018.

In the chart above, you can see that the lines corresponding to the variables almost overlap. This shows that there is probably a strong correlation between the variables under study. Moving on to the analysis of the figures for North Macedonia.

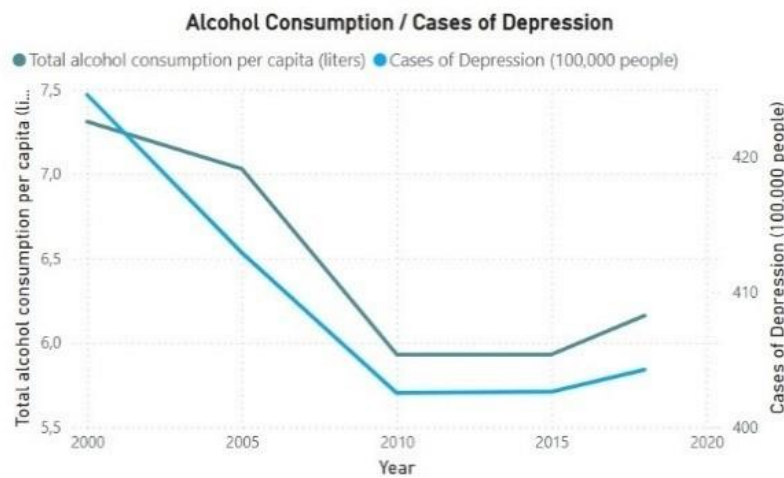


Fig. 15. An analysis of alcohol consumption and cases of depression in North Macedonia between 2000 and 2018.

As in the case of Nigeria, the values corresponding to the variables show a very similar trend, with a decrease until 2010, followed by a slight increase until 2018. So, we moved on to the visualization of the chart for Morocco.

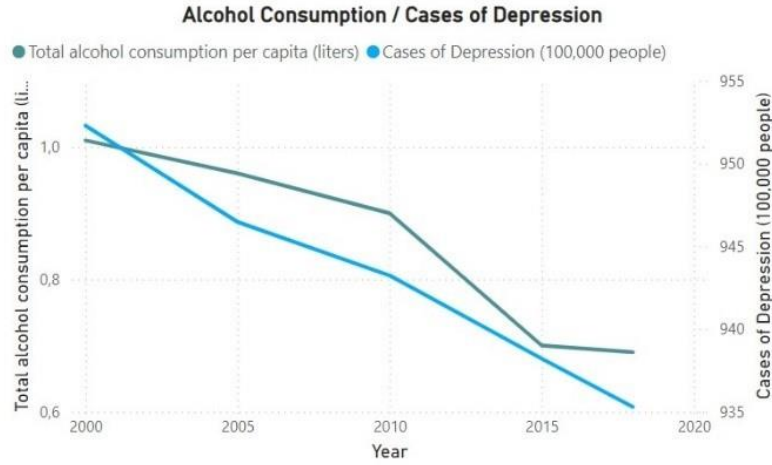


Fig. 16. An analysis of alcohol consumption and cases of depression in Morocco between 2000 and 2018.

Finally, looking at Morocco, although it is not as noticeable as in the other countries analyzed individually, it is possible to correlate the variables since they show a similar drop in the amount of alcohol consumption and cases of depression.

2.2 Suicide

In this subchapter, we have chosen to analyze the relationship between two variables that are closely related in the literature: depression and suicide. We defined the number of suicides per 100,000 inhabitants and the percentage of depressive disorders per 100,000 inhabitants for the years 2000-2018.

First, we undertook an analysis of 178 countries to understand global patterns in a comprehensive manner with a time frame. The wide approach of this nature enabled us to observe general trends and differences in the relationship between depression and suicide rates across diverse regions and populations.

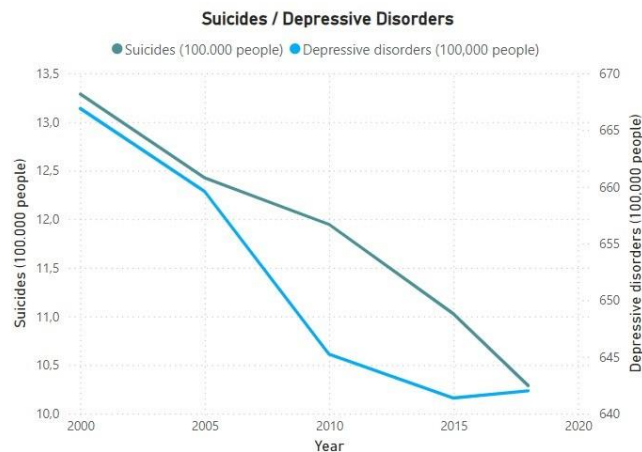


Fig. 17. A worldwide analysis of the number of suicides and depressions between 2000 and 2018.

The chart in Figure 17 shows the similarity in the trends of the data over the years. From the year 2000 to 2018, there was a general observed decrease from the world data on the number of suicides and depression prevalence. Additionally, the trajectory of the patterns was the same, almost pointing to a correlation between the two parameters. Despite these similarities, individual country analyses were decided upon. Therefore, we began the examination by disaggregating the data for Nigeria.

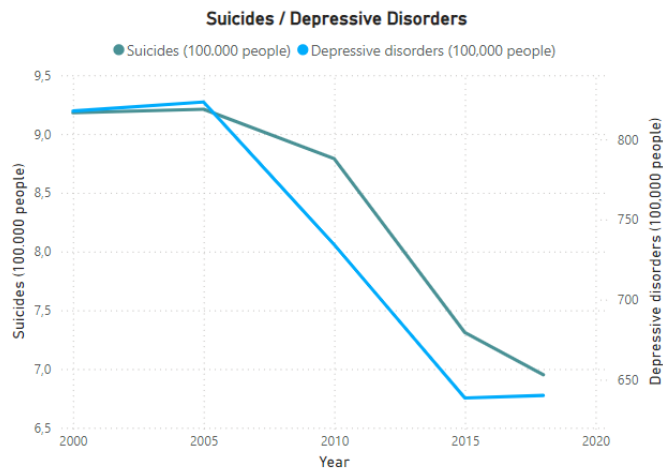


Fig. 18. An analysis of the number of suicides and depressions in Nigeria between 2000 and 2018.

For the parameters analyzed in Nigeria between 2000 and 2018, it is possible to observe a great similarity in the decrease of their values. This could be an indicator of the direct

correlation between cases of depression and the number of suicides. Then we move onto the next country.

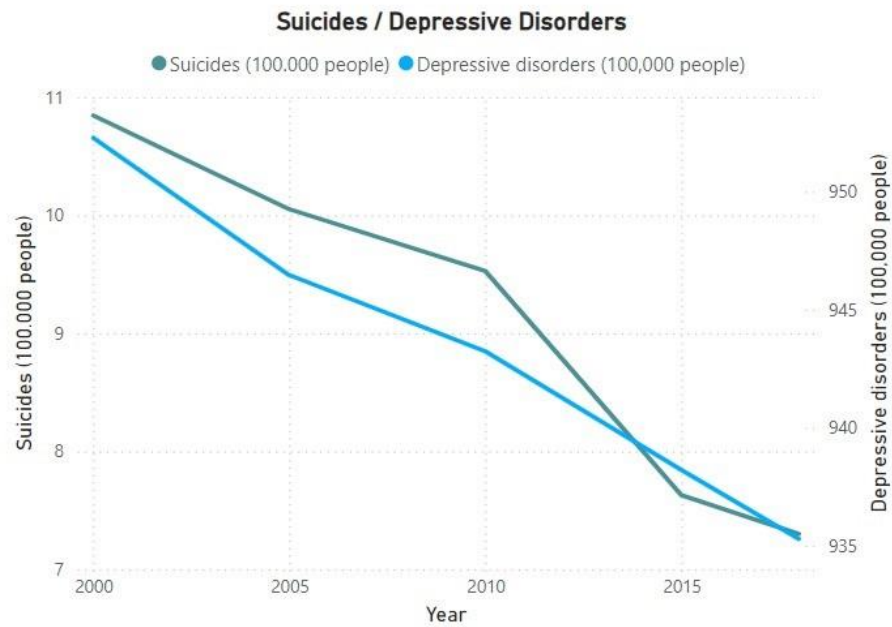


Fig. 19. An analysis of the number of suicides and depressions in Morocco between 2000 and 2018.

With regard to the parameters under analysis in Morocco between 2000 and 2018, it is also possible to observe a similarity in the decrease in their values. Once again, the presence of similarities in the evolution of the two variables may be an indicator of the direct correlation between cases of depression and the number of suicides.

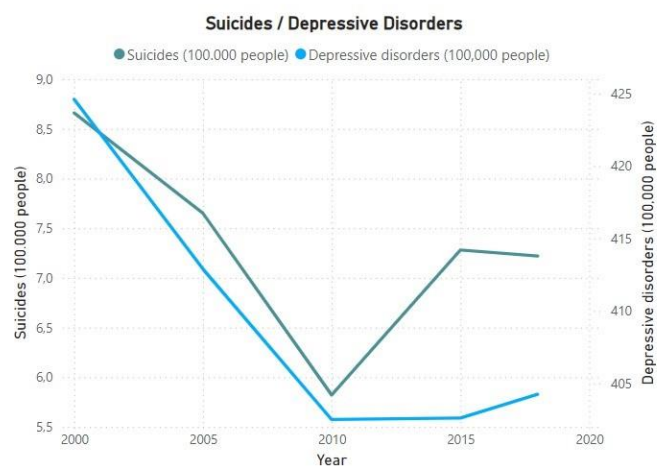


Fig. 20. An analysis of the number of suicides and depressions in North Macedonia 2000 and 2018.

Finally, in the parameters under analysis in North Macedonia between 2000 and 2018, it is also possible to observe a parity in the decrease of their values until 2010, and then there is some discrepancy in their behavior. That's because between 2010 and 2015, there was a large increase in the number of suicides, while the number of depressions remained steady.

2.3 Drugs

This subsection focuses on the analysis of the relationship between the use of illegal psychoactive substances and the prevalence of mental disorders in the population. Specifically, in this analysis, the charts for the following mental disorders are examined: Bipolar Disorder, Schizophrenia, Eating Disorders.

It should be noted that under the eating disorders category, there is no distinction made between the various types of eating disorders, such as anorexia nervosa, bulimia nervosa, binge eating disorder, or restrictive and avoidant food intake disorders. Thus, all of these disorders are considered in aggregate under the eating disorders category.

Whereas the above is the case, after the introduction, we examined the charts in detail to attempt to determine if any underlying correlations or patterns could be drawn between the use of illegal psychoactive substances and the mental disorders under discussion.

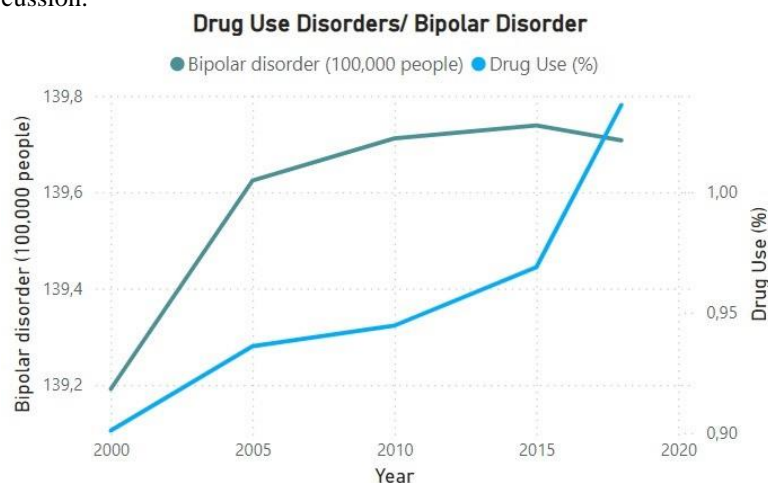


Fig. 21. A worldwide analysis of the rate of drug use disorders and the number of bipolar patients between 2000 and 2018.

Figure 21 analyzes the relationship between drug use (in %) and the presence of bipolar disorder (per 100,000 people). It can be seen that over the years there has been a tendency for both parameters to increase, with a slight similarity in the evolution of their values.

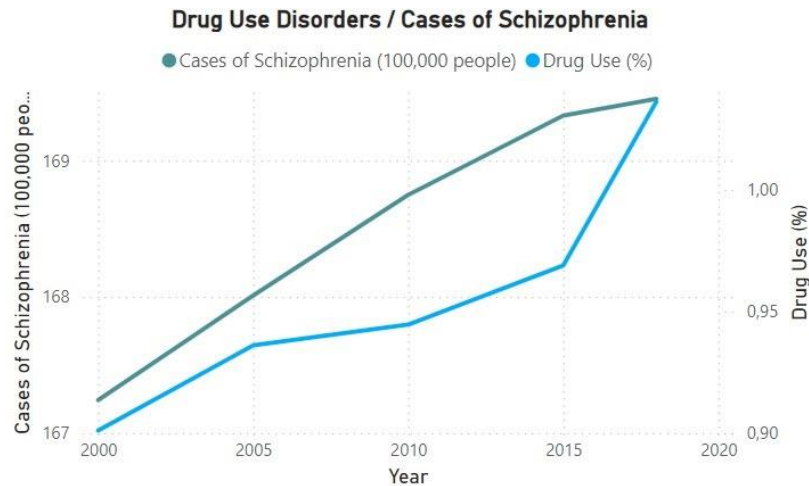


Fig. 22. A worldwide analysis of the number of drug use disorders and the number of cases of schizophrenia between 2000 and 2018.

The chart shows the relationship between drug consumption (in %) and the number of schizophrenia cases (per 100,000 people). It can be seen that over the years, both parameters follow an upward trend, and there is also a similarity in the way their values evolve, with similar trend lines.

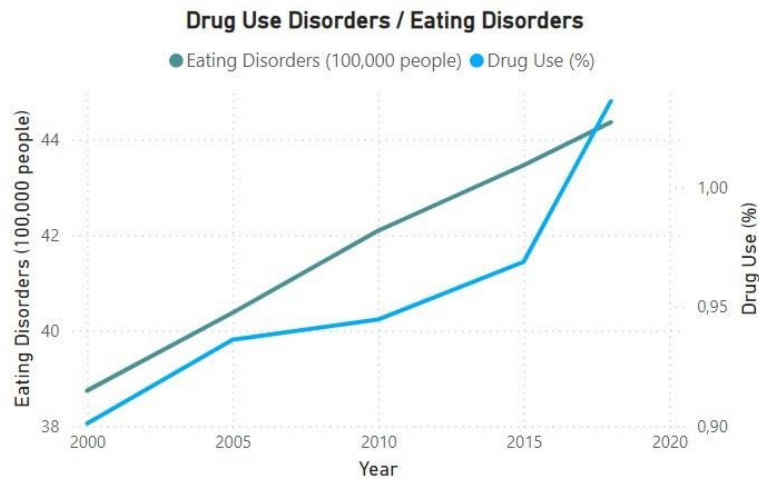


Fig. 23. A worldwide analysis of the number of drug use disorders and the number of eating disorders between 2000 and 2018.

Finally, Figure 23 depicts the relationship between the percentage of drug use and the incidence of eating disorders per 100,000 people. As in the previous charts (Fig. 21 and

Fig. 22), the parameters show parallel behavior over the years and in the wider context of global data averages.

In addition, since the beginning of the early 2000s, the growth of both overall drug use in society and the incidence of mental disorders in society has been synchronously positive. This fact suggests the presence of a certain relationship between the studied parameters.

3 Discuss

3.1 Alcohol consumption and suicide

Literature underscores the linkage between drinking alcohol and committing suicide. An extensive study [1], covering 183 countries found that the evidence was consistent across countries with different levels of income for both males and females. Suicides were also reported to be significantly higher among males, especially those falling into the middle to older age groups. A further study [2], on suicidal behavior as it connects with a variety of factors discovered the role of aggressive, impulsive, and alcohol-dependent behaviors in suicidal tendencies. In another large-scale study comprising over 10 million individuals from different strata of populations, the results indicated that 94% of suicide mortality risks were accounted for by alcohol consumption [3] Further analysis of the same data indicated that the risk is more pronounced among young populations and those with relatively high alcohol consumption.

3.2 Alcohol consumption and depression

Psychological studies have shed light on the association between alcohol consumption and depression. For example, one study looked into this multifaceted association, considering bidirectionality in three periods and observing over 10,000 samples. The findings revealed that in the cases where alcohol was consumed in harmful levels, there was a significantly higher risk of depression. Further, the analysis considered not only the quantity of alcohol consumed but the pattern of consumption and showed that the situation worsened with higher levels of consumption [4]. On the analysis of the Canadian population in 1994, involving approximately 18,000 participants, it was inferred that those who had consumed alcohol in the year prior to the analysis were more likely to present a depressive episode compared to non-consumers. Besides, heavy episodic drinking (more than 10 drinks in one occasion) was associated with an elevated risk of depression [5].

3.3 Drugs and mental illness

Several studies have explained some of the relationships between substance abuse and mental health disorders. To exemplify, eating disorders such as bulimia, predominantly in women because of the societal pressures, usually come with the use of cigarettes,

diuretics, emetics, and illicit drugs to suppress their appetite, manage their weight, or even initiate vomiting. These substances that are used are basically meant to \"treat\" symptoms related to eating disorders [6]. Other studies conducted on binge eating disorder using the Yale Food Addiction Scale 2.0 found that patients with one or more addictive behaviors were likely to report greater clinical severity [7].

In the case of schizophrenia, research suggests that patients are more predisposed to the use of substances, mainly because genetic factors have a role in determining the risk for psychosis and addiction. This predisposition might become apparent earlier than the manifestations of the disorder itself and therefore increase the risk of a substance use disorder [8]. Reports from North America also show that patients with schizophrenia are at a heightened risk of violent behavior when there is co-morbid substance abuse, highlighting their vulnerability, which can make substance abuse have antisocial and at times violent behavior as an outcome [9].

The relationship between bipolar disorders and substance abuse has also been widely studied. Literature shows that the relationship between bipolar disorder and the use of drugs is quite prominent. Patients often present for in-patient treatment early in the course of the disorder and with higher levels of depression and suicide attempts [10]. Furthermore, bipolar patients with substance abuse also demonstrate poorer adherence to treatment, resulting in more hospitalizations, and this trend is rather concerning as it calls for specific interventions to be made [11].

4 Conclusion

Therefore, after several analyses, it is possible to confirm the intrinsic link between these parameters. Depression often contributes to the use of psychoactive substances and alcohol, in search of momentary relief which, over time, can worsen depressive symptoms, making it difficult for the patient to recover.

In addition, excessive alcohol use carries a higher risk of suicidal behaviour, because according to some studies, people who have some kind of dependency, along with depressive disorders, have significantly higher suicide rates, and so interventions that highlight all the problems derived from alcohol are vital.

Finally, this project contributes to raising public awareness and calling for the creation of new public health policies. The creation of safe and supportive environments, for instance, is essential to break this harmful cycle. Effective intervention strategies are vital to promote mental health, reduce substance abuse and prevent suicide, contributing to a healthier and more resilient society.

References

1. Kim, A.M.: Alcohol consumption and suicide rate: A cross-sectional analysis in 183 countries. *Psychiatry Res.* 295, 113553 (2021). <https://doi.org/10.1016/j.psychres.2020.113553>.

2. Sher, L.: Alcohol consumption and suicide. *QJM: An International Journal of Medicine*. 99, 57–61 (2006). <https://doi.org/10.1093/qjmed/hci146>.
3. Isaacs, J.Y., Smith, M.M., Sherry, S.B., Seno, M., Moore, M.L., Stewart, S.H.: Alcohol use and death by suicide: A meta-analysis of 33 studies. *Suicide Life Threat Behav*. 52, 600–614 (2022). <https://doi.org/10.1111/sltb.12846>.
4. Gémes, K., Forsell, Y., Janszky, I., László, K.D., Lundin, A., Ponce De Leon, A., Mukamal, K.J., Moller, J.: Moderate alcohol consumption and depression – a longitudinal population-based study in Sweden. *Acta Psychiatr Scand*. 139, 526–535 (2019). <https://doi.org/10.1111/acps.13034>.
5. Patten, S.B., Charney, D.A.: Alcohol Consumption and Major Depression in the Canadian Population. *The Canadian Journal of Psychiatry*. 43, 502–506 (1998). <https://doi.org/10.1177/070674379804300509>.
6. Bulik, C.M.: Abuse of drugs associated with eating disorders. *J Subst Abuse*. 4, 69–90 (1992). [https://doi.org/10.1016/0899-3289\(92\)90029-W](https://doi.org/10.1016/0899-3289(92)90029-W).
7. Miranda-Olivos, R., Agüera, Z., Granero, R., Vergeer, R.R., Dieguez, C., Jiménez-Murcia, S., Gearhardt, A.N., Fernández-Aranda, F.: Food addiction and lifetime alcohol and illicit drugs use in specific eating disorders. *J Behav Addict*. 11, 102–115 (2022). <https://doi.org/10.1556/2006.2021.00087>.
8. Khokhar, J.Y., Dwiel, L.L., Henricks, A.M., Doucette, W.T., Green, A.I.: The link between schizophrenia and substance use disorder: A unifying hypothesis. *Schizophr Res*. 194, 78–85 (2018). <https://doi.org/10.1016/j.schres.2017.04.016>.
9. Smith, J., Hucker, S.: Schizophrenia and Substance Abuse. *British Journal of Psychiatry*. 165, 13–21 (1994). <https://doi.org/10.1192/bjp.165.1.13>.
10. Ortiz, Ó.M.: Comorbilidad por abuso de sustancias en el trastorno bipolar*. *Rev Colomb Psiquiatr*. 41, 371–383 (2012). [https://doi.org/10.1016/S0034-7450\(14\)60011-1](https://doi.org/10.1016/S0034-7450(14)60011-1).
11. Sherwood Brown, E., Suppes, T., Adinoff, B., Rajan Thomas, N.: Drug abuse and bipolar disorder: comorbidity or misdiagnosis? *J Affect Disord*. 65, 105–115 (2001). [https://doi.org/10.1016/S0165-0327\(00\)00169-5](https://doi.org/10.1016/S0165-0327(00)00169-5).