

Anxiety Disorder

Akash Gangadharan (akganga@iu.edu)

Sandhya Iyer (sniyer@iu.edu)

Sakshi Rathi (sakrathi@iu.edu)

Vedika Halwasiya (vhalwasi@iu.edu)

ABSTRACT:

This analysis is about predicting anxiety disorder rates spanning from 1990 to 2019 across 157 countries. Key questions explore how GDP, population age structure, and gender relate to the prevalence of anxiety disorders. Models estimate the relationships and predictive power between potential explanatory factors like GDP, male and female population percentage, and anxiety levels over time. Findings reveal clear links between income status and anxiety, with GDP demonstrating significant explanatory ability in predictive models. Analysis incorporating age, and gender, into enhanced models to deepen understanding of variables influencing national anxiety outcomes and refine future forecasting capabilities.

INTRODUCTION:

Data Description:

Data Source:

The dataset utilized in this analysis was obtained from Kaggle, a platform hosting various public datasets. Specifically, the data is from the "Mental Health" dataset compiled by AmirHosein Mousavian. This dataset provides comprehensive information on the prevalence of anxiety disorders across different countries, along with economic indicators such as GDP per capita and demographic information like population size. The dataset spans a period from 1990 to 2019 and includes a wide range of countries, facilitating a global analysis of mental health trends in relation to economic factors.

Dataset Link: Kaggle - [Mental Health Dataset](#)

WorldBank dataset for income classification - [World Bank data](#)

Scope of Data:

The dataset encompasses a period from 1990 to 2019, providing a comprehensive 30-year overview of mental health trends, specifically focusing on the prevalence of anxiety disorders. Geographically, the dataset covers a global scope, including a diverse range of countries from various continents. However, for the purpose of our analysis, we have selectively focused on eight countries representing different income classifications as per the World Bank's standards: United States and Spain (High income), China and Argentina (Upper middle income), India and

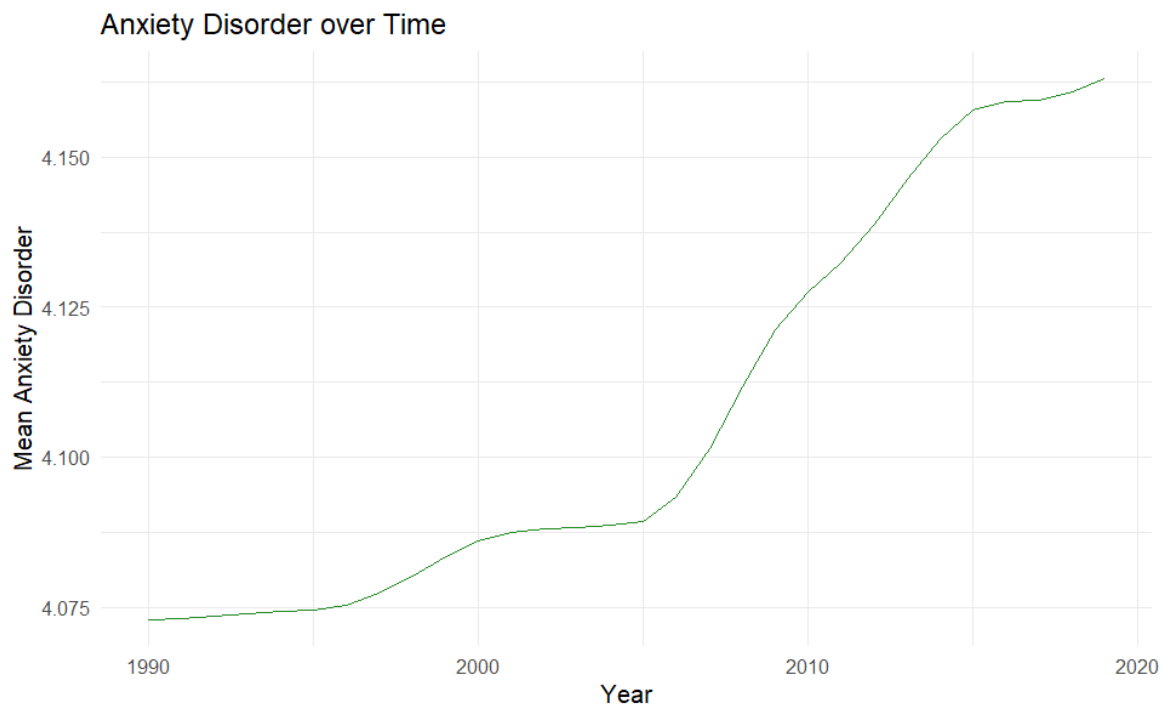
Bhutan (Lower middle income), and Madagascar and Ethiopia (Low income). This selection allows for a nuanced analysis of mental health trends in relation to both economic status and geographic diversity. The dataset includes key variables such as anxiety disorder prevalence, GDP per capita, population size, and the respective continents of the countries.

Research Questions:

1. How do factors like GDP, Population, Age, and Gender of the population affect anxiety disorder levels?
2. Which of these factors are the predictive factors for modeling a country's anxiety disorder level (of its mean share population)?

DATA ANALYSIS:

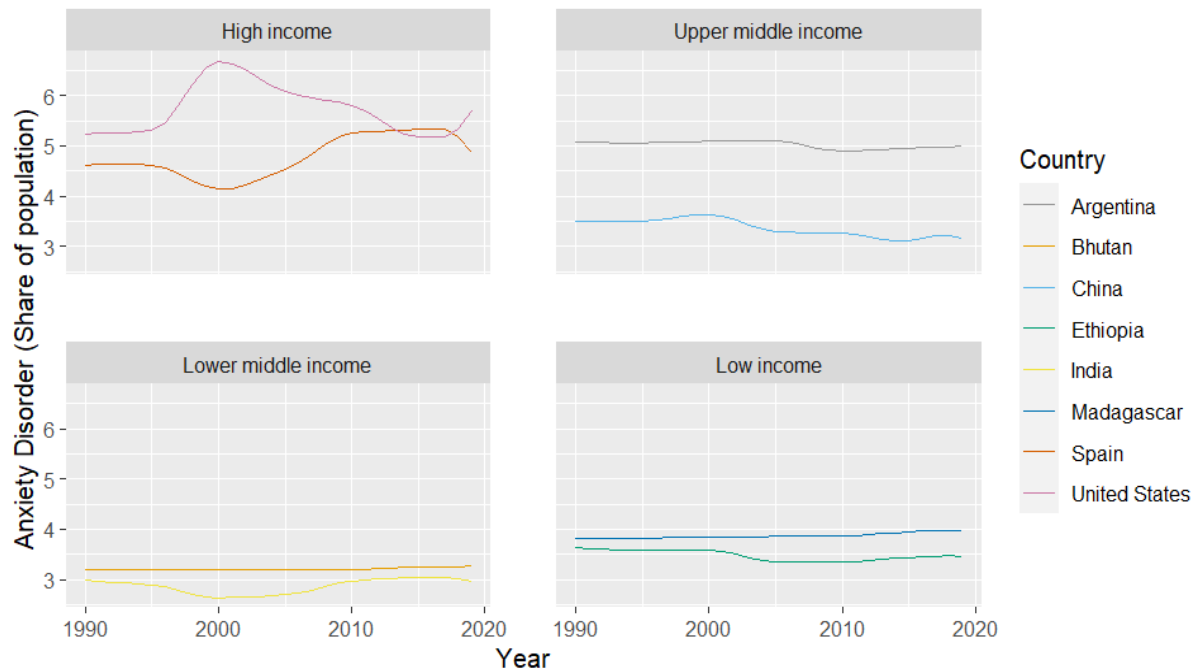
Anxiety Disorder over Time



By observing the 1990 and 2019 trends of log GDP vs Anxiety disorder, the trend seems that with the increase in GDP, there is an increase in anxiety. But the data points are way too scattered. So, let's consider other factors like population, income of a country, gender, and age.

Early on, we mentioned taking into account the income classification, we plot the anxiety variation with time.

Anxiety Disorder for countries classified based on income 1990-2019



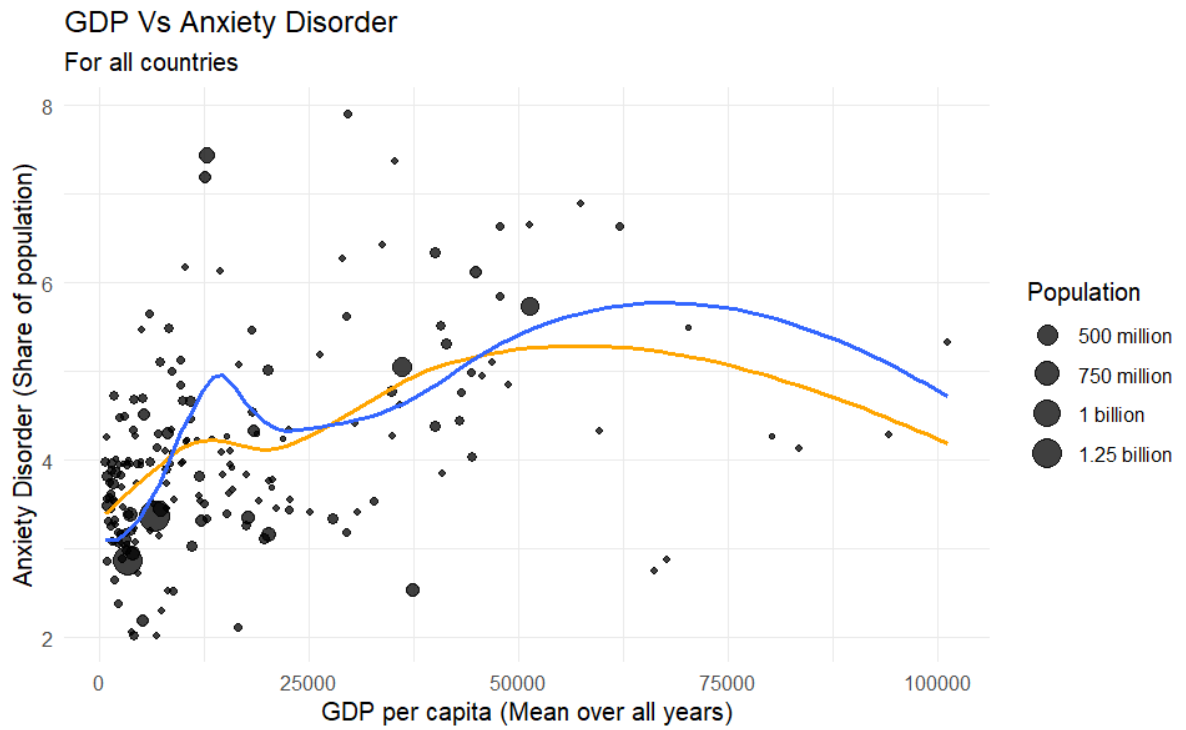
High-income and upper-middle-income categories have a higher anxiety level as compared to lower-middle-income and low-income.

In the high-income category, US (high population) has higher anxiety disorder levels as compared to Spain (low population). We can say that a higher population with high income contributes to high anxiety levels.

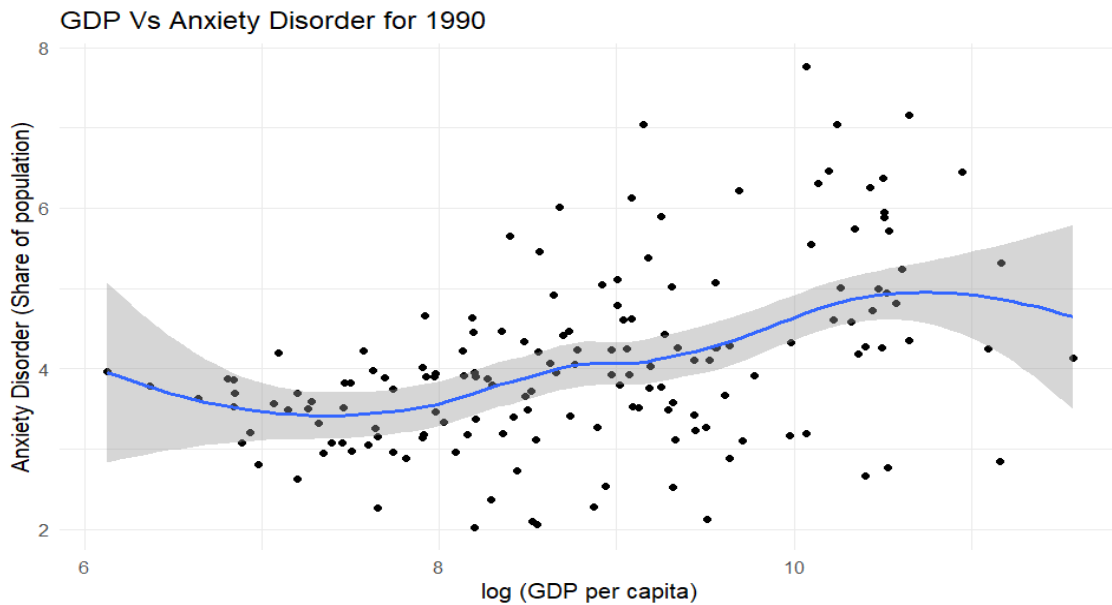
Whereas, in the remaining three categories, it is observed that among pairs of countries, the one with a smaller population experiences greater levels of anxiety in comparison to its counterpart with a higher population.

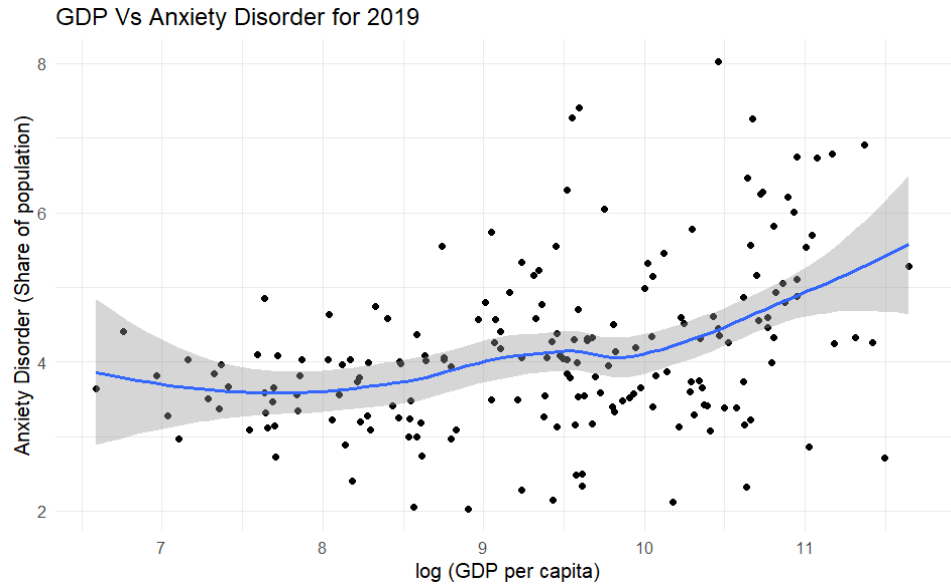
1. GDP

We observe the Anxiety levels with the variation in GDP (or log GDP). First, we plot the mean GDP from the years 1990-2019 with respect to the anxiety level for the mean share of the population. We observe that the curve tends to dip when the GDP is really high. This can be explained by the fact there are a few countries it is trying to fit, which could be the outliers. Also, an extremely high GDP could mean extremely wealthy countries but they might not have a very high population.



We plot the GDP with respect to anxiety level for the years 1990 and 2019 to observe the trend.





2. Age Category

We plot the anxiety disorder over the different age categories, faceted by the countries' classification by income. (Considering 8 countries for the given income classification).

Anxiety Disorder Prevalence by Age in High-Prevalence Countries
1990-2019



In the high-income category, all age groups exhibit high anxiety levels, but the working population tends to have even higher anxiety compared to other age groups. This might be due

to increased workplace stress, career expectations, or work-related pressures prevalent in high-income societies.

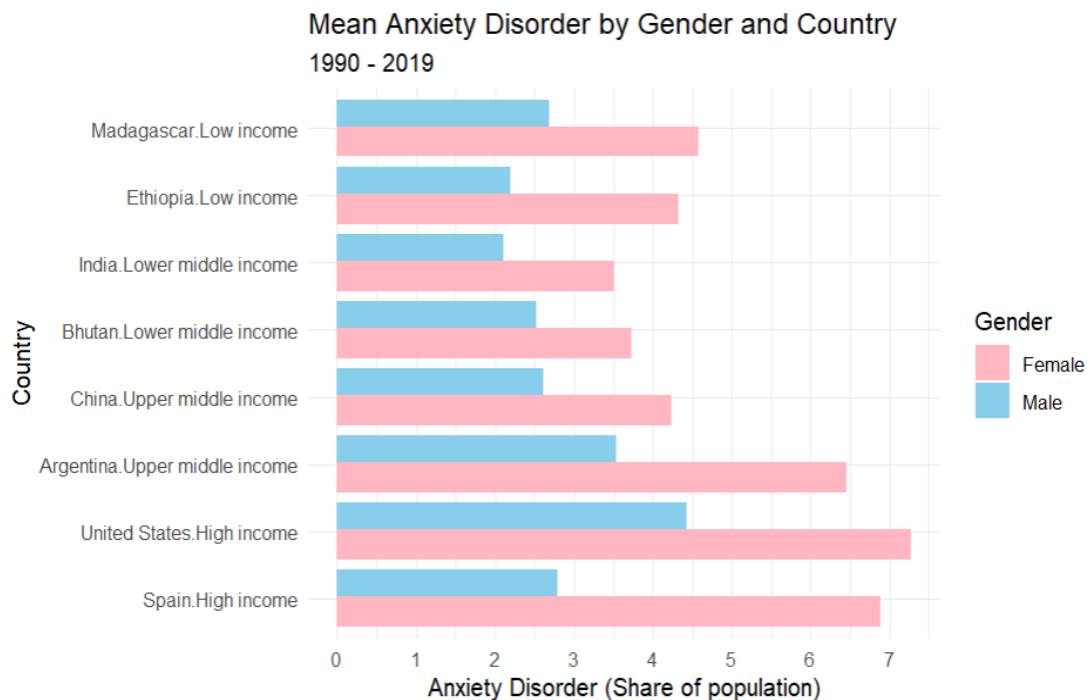
For upper-middle-income countries, the anxiety levels are pretty similar to the high-income category. Here, all the age groups have similar anxiety levels.

For the lower middle and low-income countries, anxiety levels are notably lower compared to higher-income categories. However, among these lower-income countries, the working population tends to have higher anxiety levels, similar to observations in higher-income countries. This could be due to economic uncertainties, job insecurities, and workplace stress. Surprisingly, the older population also experiences a comparable level of anxiety in these lower-income countries. This might be attributed to various factors such as health concerns, financial worries, or social support systems available to the elderly.

Across all income categories, the working population seems to be the most affected by anxiety. This trend might be linked to occupational stress, job insecurities, career advancement pressures, and the demanding nature of work across different economic landscapes.

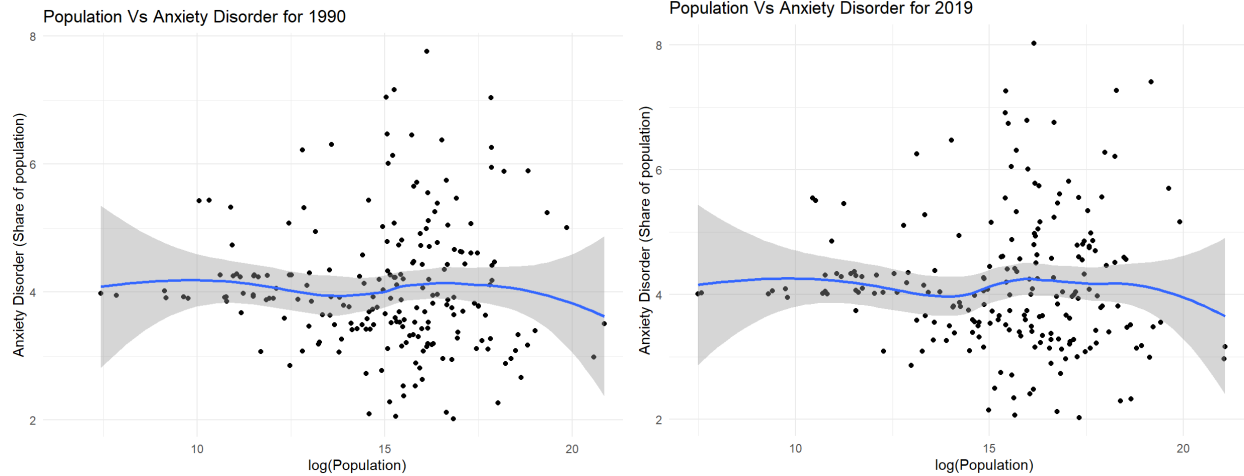
3. Gender

Considering the mean of anxiety for the years 1990-2019, for all the categories of countries based on income, we observe that the anxiety disorder levels are significantly higher in females as compared to males.

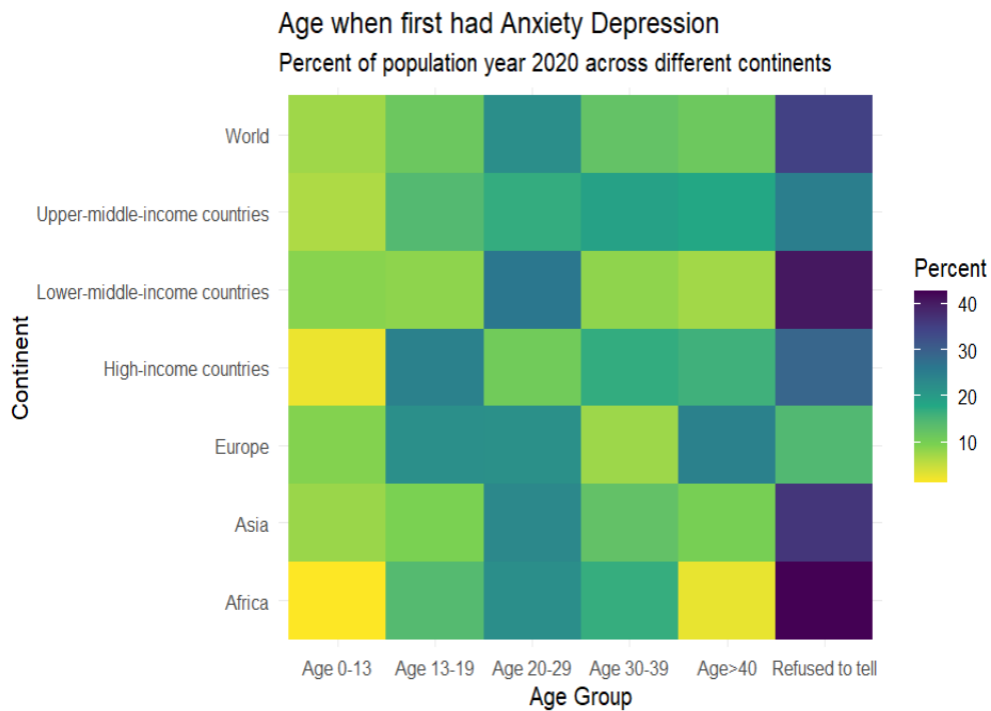


4. Population

From the plots below, log population does not seem to have a major effect on the variation of anxiety levels.

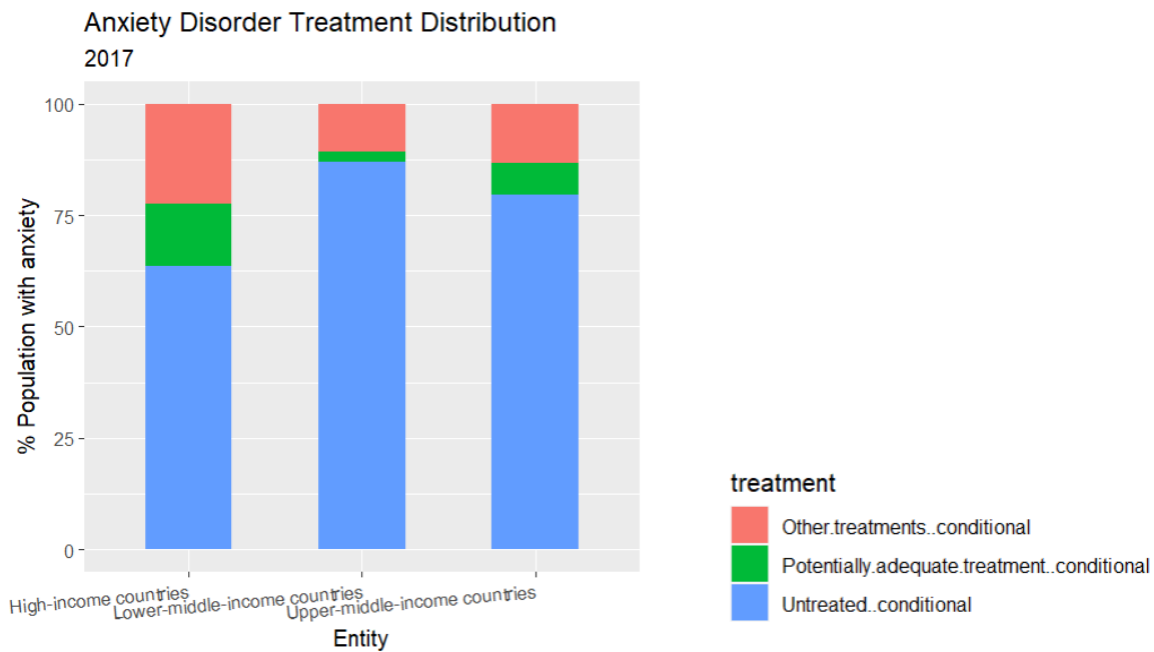


5. Age when they first had anxiety.



A major chunk of the population refuses to tell when they first experienced anxiety. Across the teenage years, high-income countries have a high percentage of the population first experiencing anxiety. As we progress towards the 20-29 age group Lower-middle income countries, Europe, Asia, and Africa see a significant rise too. We can say that a major chunk of the population first experiences anxiety before the age of 30, although many people have refused to reveal their age.

6. Anxiety Disorder Treatment Distribution

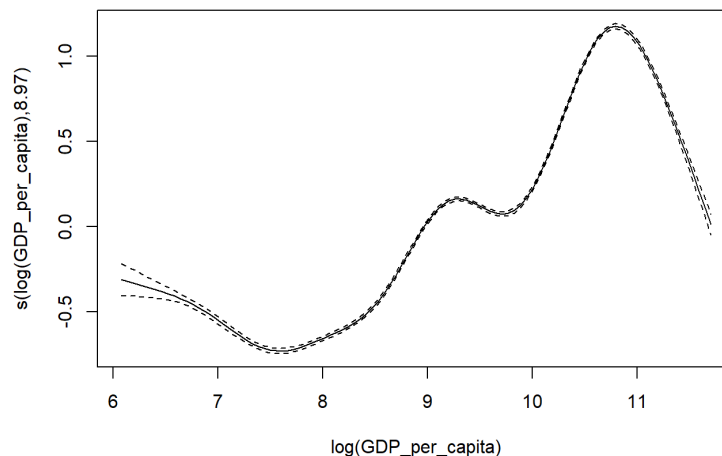


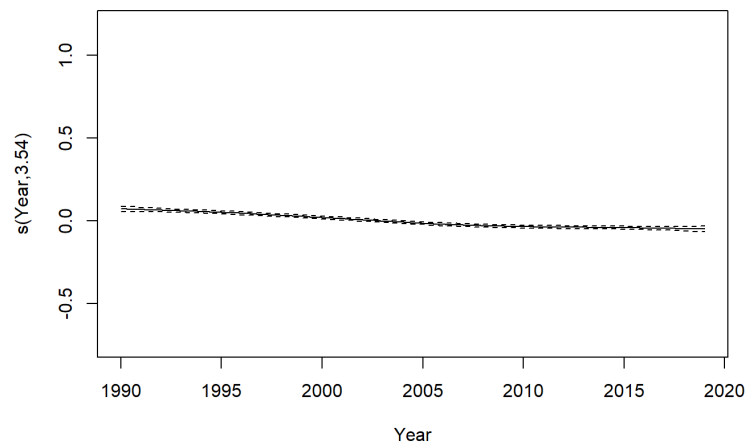
Across all the categories, we can see that the majority of the population is untreated, this is a great cause of concern. For the high-income countries, the share of the population receiving treatments is higher as compared to other middle-low-income countries. (it has better facilities).

DATA MODELING:

We fit a GAM to predict Anxiety_disorders based on the smoothed relationships with the logarithm of GDP per capita ($\log(\text{GDP_per_capita})$) and the year (Year). The use of smooth terms allows for flexibility in modeling potentially non-linear relationships between the predictors and the response variable.

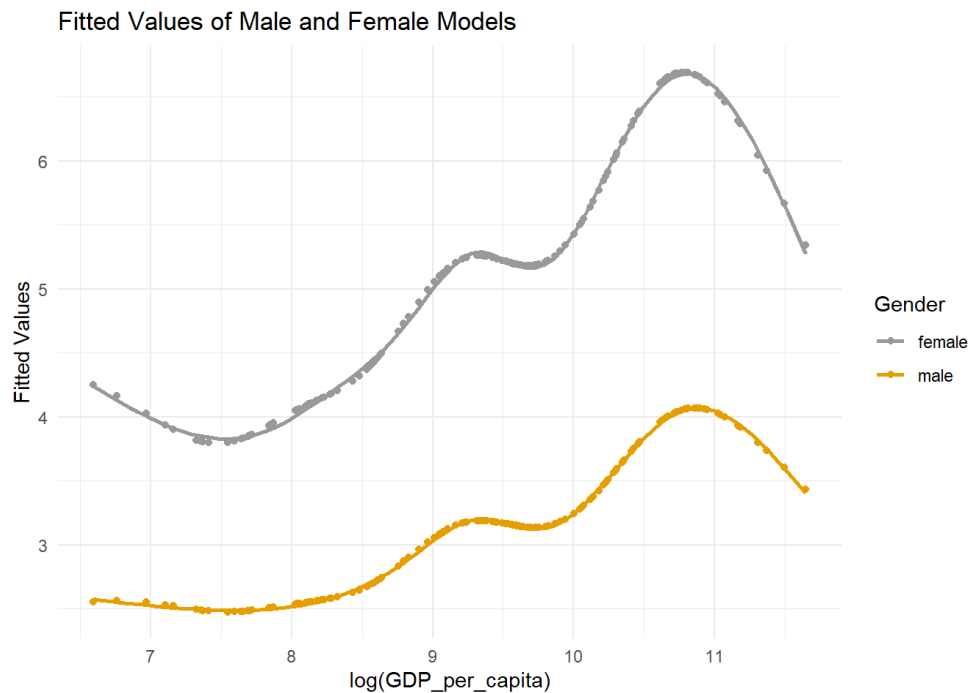
The model has a R^2 of 0.26. may seem low, but considering the variability in the predictability factors of anxiety, it seems pretty decent.





The plots indicate that the increase in log GDP has an increasing effect on anxiety, though with a really high GDP, we see a dip in anxiety level. This can be explained by some outliers, discussed in the data analysis section.

We then model two separate gender anxiety, and then we would compare if the fitted values of females and males.



The data analysis before indicates that female anxiety is significantly larger than males. When we model the male and female anxieties with log GDP and year. The predicted values of the male and female anxiety yet again have a significant difference. So, modeling log GDP along with other factors could prove to be a better fit for predicting anxiety levels.

CONCLUSION:

This exploratory analysis of global mental health data reveals several key insights into factors influencing anxiety disorder prevalence across countries and income levels. Clear patterns emerge linking higher national income status with greater overall anxiety rates. However, rising income also correlates with more treatment availability and lower untreated incidence. GDP demonstrates reasonably strong predictive capabilities as a single variable, explaining over a quarter of the variance in country-level anxiety outcomes. Yet substantial gains in model performance are attainable by incorporating further economic, demographic, healthcare, and other national traits. Population size exhibits little consistent predictive power. However, age and gender factors show promise for enhancing models. Working-age groups tend to endure heightened anxiety worldwide, especially in high-income countries. And women face markedly higher anxiety levels across all income brackets.

In conclusion, income and GDP function as useful starting points for modeling anxiety but fail to capture intricate societal drivers fully. Expanding feature sets to better encapsulate occupational pressures, gender gaps, age dynamics, and care access holds the potential to significantly improve the explanatory prowess of models while offering more granular insights to policymakers.

LIMITATIONS:

The model relies solely on GDP and year variables, failing to account for complex societal drivers of anxiety. Data constraints prevented the inclusion of potentially meaningful factors like unemployment, workplace environments, and stress levels. The limited country sample risks bias and reduced generalizability of global trends.

FUTURE WORK:

1. **Incorporate additional macroeconomic predictors:** The model would benefit greatly from expanding the features set with unemployment rates, income inequality metrics, education access, and other national economic traits that may relate to anxiety levels.
2. **Implement advanced ML modeling approaches:** Techniques like random forests and neural networks can capture highly complex multivariate relationships missed by GAM models.