Economic Inequality and the Global Mental health Burden

Alona Sychevska, 2769408, Tutorial lecturer's name: Chantal Schouwenaar, Jack Fitzgerald

June 5, 2025

1 Problem Motivation & Literature

Why It's a Recognized Social Problem

Mental health issues like depression and anxiety are rising worldwide—not only in high-income countries like the Netherlands or UK but also in low- and middle-income countries (LMICs). The WHO lists mental disorders among the top causes of disability globally, with burdens expected to increase (World Health Organization, 2023).

Research shows **economic inequality**, more than poverty or income alone, strongly predicts mental health problems. Higher income inequality correlates with greater psychological distress even after accounting for GDP per capita (Praeg, Mills, & Wittek, 2014).

A review focusing mainly on high-income countries found income inequality increases depression risk by 19% (Patel et al., 2018). LMIC studies confirm poverty and inequality also predict common mental disorders (Lund et al., 2010). Cross-continental research reinforces that inequality worsens anxiety and depression across all economies, emphasizing the need for global analysis (Shahbazi, Shahbazi, & Poorolajal, 2022).

Theoretical Framing & Mechanisms

Economic inequality affects mental health through:

- **Relative deprivation**: Social comparison breeds chronic stress and lowers self-worth (Smith et al., 2012).
- **Social capital erosion**: Inequality weakens trust and cohesion, increasing loneliness and mental illness risk (Kawachi & Berkman, 2000).
- **Poverty–mental health loop**: Mental illness and poverty reinforce each other, especially in LMICs (Lund et al., 2010).

Gap in the Literature

Most studies link inequality (via Gini) to mental health in high-income settings, treating GNI or GDP per capita as controls. Few examine Gini, GNI, and GDP per capita together, especially in LMICs where mental health burdens are rapidly growing.

This project fills that gap by analyzing how these economic factors jointly influence mental health across diverse countries.

2 Data Sourcing & Description

Data Sources and Credibility

This project integrates datasets: the Global Burden of Disease (GBD) and multi-sourced data compiled from the World Bank (WDI), OECD, and other internationally reputable institutions that includes socioeconomic variables.

The GBD Study is produced by the Institute for Health Metrics and Evaluation (IHME) at the University of Washington. It is considered the gold standard for global health statistics and is widely cited by academic institutions and policymakers worldwide (Global Burden of Disease Collaborative Network, 2024).

The compilation of datasets drawing from sources like the World Bank's World Development Indicators (WDI), and World Population Review. These are recognized institutions known for robust methodologies and transparency. WDI, in particular, is frequently used in cross-country empirical studies due to its consistency and broad temporal coverage.

Metadata Overview

Go to ReadMe.md

Complementarity of the Datasets

GBD offers high-quality health burden metrics (prevalence in numbers), crucial for understanding the impact of mental health disorders, while the additional datasets offer explanatory variables on inequality and economic performance. These datasets are complementary in structure and purpose: the GBD gives outcome measures, while the additional datasets give potential predictors. Their integration allows for robust modeling of social determinants of mental health at the national level.

While I initially collected more than two datasets, the final dataset has complementary strengths and ability to support the causality-oriented focus of this project: exploring how inequality and structural factors relate to mental health burdens across countries.

Relevance to the Topic

The project's focus is on the causal relationship between inequality and mental health in low-, middle-, and high-income countries. The selected datasets are ideal because: The GBD dataset provides mental health burden metrics (e.g., Prevalence of depression or anxiety), which are essential for measuring the societal impact of mental disorders across countries. The merged dataset allows for correlational and regression analyses of how inequality indicators (Gini index), and economic performance relate to mental health prevalence. Datasets cover a wide range of countries, including those outside the high-income bracket, which aligns with the study's comparative angle.

Limitations of the Data

Despite their strengths, datasets have limitations: GBD data, though standardized, aggregate country-level estimates, potentially masking subnational disparities and cultural variation in diagnosis/reporting. The multi-sourced data suffers from missing data for certain indicators in low-income countries, which could bias regression results or reduce sample size. Differences in data collection years (some indicators are 2021, others 2022 or 2023) may introduce temporal misalignment. Mental health prevalence estimates from sources like World Population Review may lack the methodological rigor of epidemiological surveys.

2.1 Load in the data

2.2 Checking the structure

```
head(GBD %>% select(measure_name, location_name, cause_name, metric_name, year, val), 3)
```

```
## # A tibble: 3 x 6
##   measure_name location_name cause_name       metric_name  year       val
##   <chr>        <chr>         <chr>            <chr>       <dbl>     <dbl>
## 1 Deaths       Israel        Mental disorders Number       2010 0.616
## 2 Deaths       Israel        Mental disorders Percent      2010 0.0000156
## 3 Deaths       Israel        Mental disorders Rate         2010 0.00786
```

2.3 Quantifying

Data cleaning

Creating new variable: gini_bucket

```
economic_inequality_gini_index <- economic_inequality_gini_index %>%
  filter(Year >= 2012) %>%
  rename("Gini_index" =
           "Gini coefficient (2017 prices) - Income or consumption consolidated") %>%
  select(-Code, -`1039568-annotations`)
# Creating gini buckets (new var)
gini_cats <- economic_inequality_gini_index %>%
  mutate(gini_bucket = case_when(
    Gini_index < 0.25                 ~ "Low",
    Gini_index >= 0.25 & Gini_index < 0.35 ~ "Moderate",
    Gini_index >= 0.35                ~ "High",
    TRUE                            ~ NA_character_
  ))
```

To facilitate group-wise comparisons and enhance interpretability, a new categorical variable—gini_bucket—was created based on the continuous Gini index. Countries were grouped into three inequality categories: Low (< 0.25), Moderate (0.25–0.35), and High (> 0.35). This transformation enables easier comparison across inequality levels, particularly in visualizations and descriptive analyses. Although some granularity is lost by converting continuous values into bins, the approach improves clarity and supports the identification of broader patterns in mental health outcomes across varying levels of income inequality.

```
gni_by_country_year <- gni_by_country_year %>%
  rename("2016" = "...31", "2017" = "...32", "2018" = "...33", "2019" = "...34") %>%
  pivot_longer(
    cols = matches("^\\d{4}$"),
    names_to = "year",
    values_to = "income_group"
  ) %>%
  mutate(year = as.integer(year)) %>%
  select(Country, year, income_group)
gni_by_country_year <- gni_by_country_year %>%
```

```r
  mutate(income_group = ifelse(income_group == "..", NA, income_group)) %>%
  filter(year >= 2012)
```

All datasets were filtered to include only data from 2012 onward, and unnecessary columns or duplicates were removed. Variables were renamed for clarity, and formats were standardized (e.g., reshaping wide GNI data to long format, converting ".." to missing values). These steps ensured consistency across datasets and prepared the data for merging and analysis.

Creating new variables: percent and rate of depression by country

```r
# rate per 100,000 and percent
GBD <- GBD %>%
  filter(metric_name == "Number", measure_name == "Prevalence") %>%
  mutate(
    val_m = round((val / pop_est) * 100000, 2),
    val_rate = format(val_m, big.mark = ",", scientific = FALSE),
    val_percent = (val / pop_est) * 100
  ) %>%
  select(Country, cause, Year, val_rate, val_percent, pop_est) %>%
  filter(Year >= 2012)
```

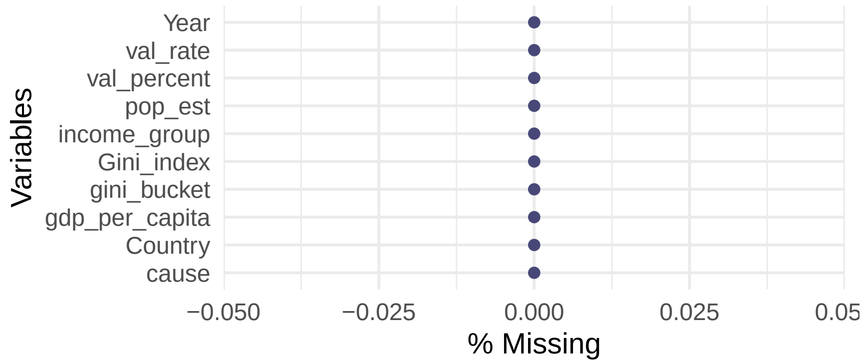
To standardize mental health prevalence data across countries with varying population sizes, two new variables were created: val_rate, representing the number of cases per 100,000 population, and val_percent, indicating the proportion of the population affected by a disorder as a percentage. These transformations ensure comparability across countries and time periods, regardless of absolute population size. The per-100,000 rate (val_rate) is commonly used in public health reporting, while the percentage (val_percent) aids in interpreting the relative burden of disorders within a country's population.

2.4 Datasets merging

```r
# Mental + gini_cats data
merged_data <- GBD %>%
  inner_join(gini_cats, by = c("Country" = "Entity", "Year" = "Year"))
```

To ensure consistency and minimize missing values, datasets were merged using inner_join. This method retains only rows with complete matches across key variables (Country and Year), which helps maintain data integrity for comparative analysis. While this approach may reduce the overall number of observations, it avoids issues introduced by incomplete records and ensures that all included data points are reliable and consistent across mental health, inequality, GDP, and income classification data.

```r
gg_miss_var(merged_data, show_pct = TRUE)
```



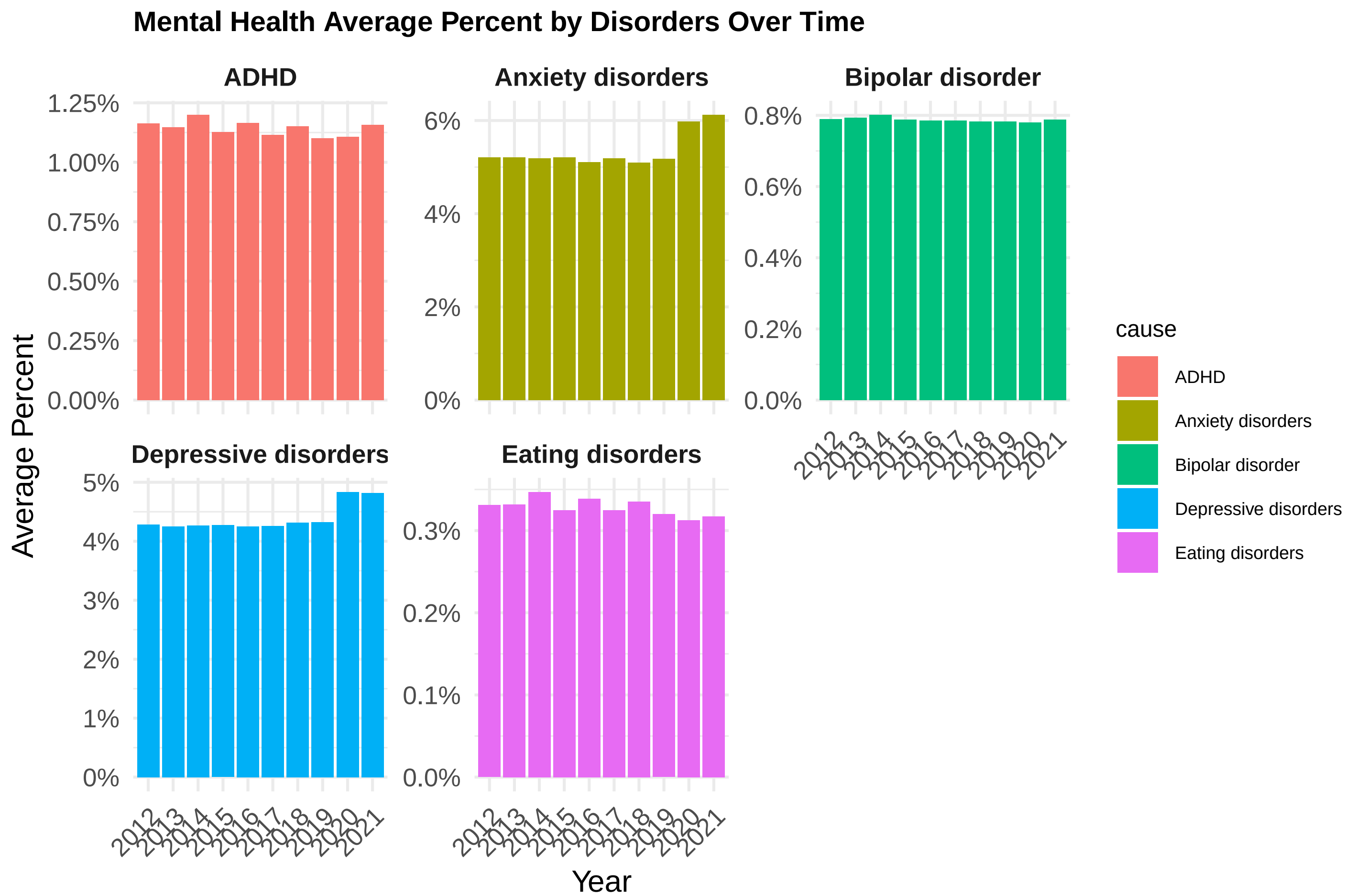
2.5 Describe the type of variables included

```
describe_variables(merged_data)
```

```
## - Country: categorical (character), levels: China, Mongolia, Belarus, Kazakhstan, Armenia ...
## - cause: categorical (character), levels: Schizophrenia, Attention-deficit/hyperactivity disorder, I
## - Year: continuous (numeric), mean = 2016.37, sd = 2.82
## - val_rate: categorical (character), levels:    366.55,    366.01,    365.29,    364.93,    365.07 .
## - val_percent: continuous (numeric), mean = 2.78, sd = 4.25
## - pop_est: continuous (numeric), mean = 104628265.19, sd = 315184396.73
## - Gini_index: continuous (numeric), mean = 0.34, sd = 0.07
## - gini_bucket: categorical (character), levels: High, Moderate, Low
## - income_group: categorical (character), levels: Upper-middle income, Lower-middle income, High inco
## - gdp_per_capita: continuous (numeric), mean = 30698.76, sd = 24331.04
```

3.1 Visualizations

3.1.1 Temporal Variations



The figure shows the average percentage prevalence of five major mental health disorders — Depressive disorders, Anxiety disorders, Bipolar disorder, Eating disorders, and ADHD — over several years. Each disorder is presented in a separate facet for clearer comparison, with the y-axis scaled independently to reflect differences in prevalence magnitude.

Key observations include:

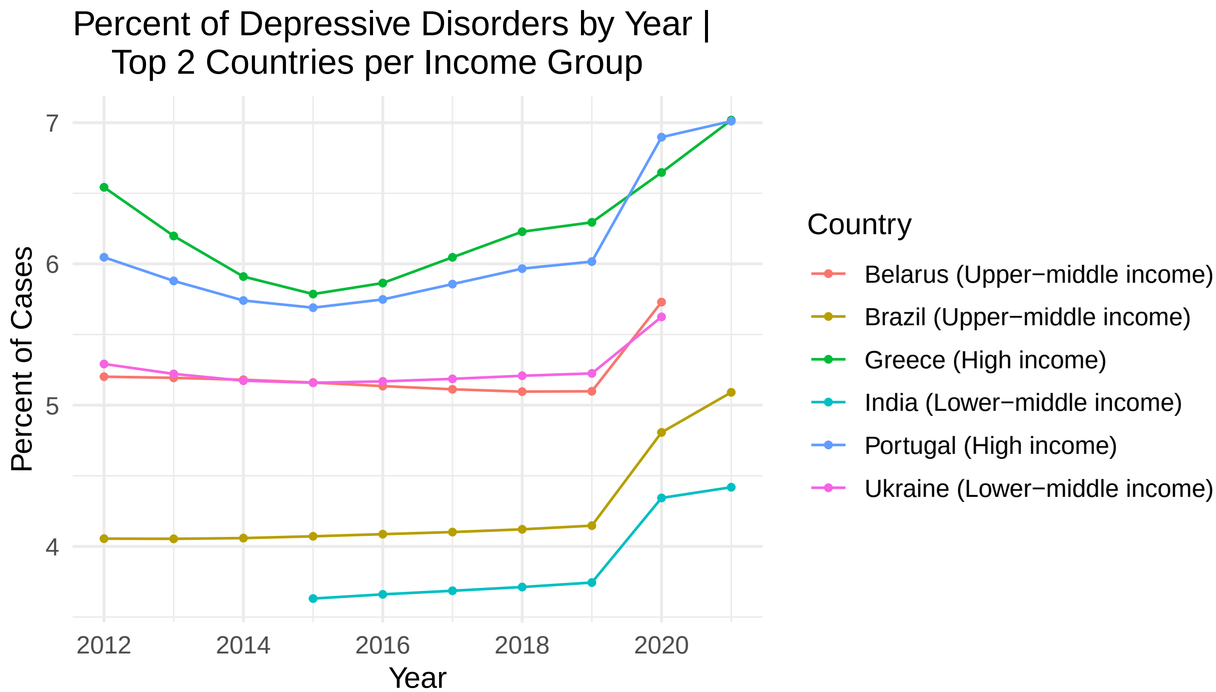
Depressive disorders consistently show the highest average prevalence, remaining relatively stable between 4.2% and 4.4%, with a slight upward trend over time.

Anxiety disorders exhibit a comparable pattern but at slightly higher rates, ranging from 5% to 5.3%, with mild fluctuations and a noticeable increase in 2020.

Bipolar disorder and ADHD exhibit much lower prevalence values, remaining relatively stable over time without clear upward or downward trends.

Eating disorders show the lowest prevalence overall but display a slight increase in recent years.

Overall, the data suggest that while some disorders like depression and anxiety remain widespread and possibly increasing, others remain steady at lower prevalence levels. This temporal analysis highlights the importance of continuous monitoring to understand evolving mental health trends.



This line plot shows the percentage of depressive disorder cases over several years for the top two countries within each income group (Low, Lower-middle, Upper-middle, High). Each country is labeled along with its income group in the legend for clarity.

Key points:

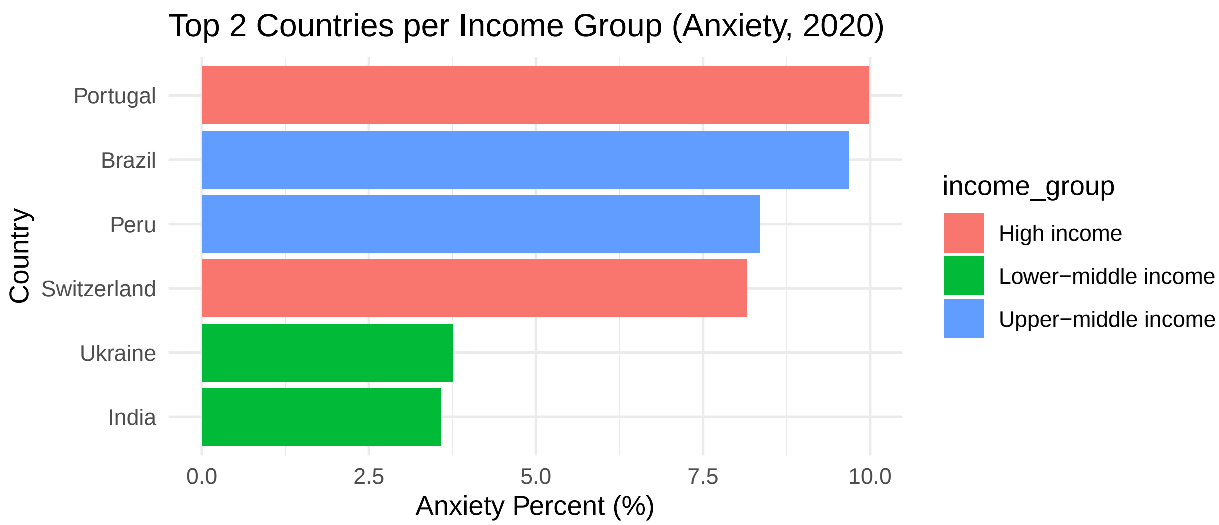
- High-income countries (e.g., Portugal, Greece) consistently report higher prevalence percent, generally between 5.6% and 7%, with some fluctuations.
- Upper-middle-income countries (e.g., Brazil) exhibit moderate prevalence around 4% to 5+, with mild variation over time, and a jump during the COVID time.
- Lower-middle- and low-income countries (e.g., India) tend to have lower prevalence, though some show slight upward or downward trends.

The data reveal a clear pattern where higher-income countries report higher rates of depressive disorders, while lower-income groups show comparatively lower prevalence, emphasizing economic context's role in mental health.

This graph highlights important disparities in depressive disorder prevalence across economic groups, underscoring the need for income-sensitive mental health strategies.

```
dep_anx %>%
  filter(cause == "Anxiety disorders", Year == 2020) %>%
  group_by(income_group) %>%
  slice_max(order_by = val_percent, n = 2, with_ties = FALSE) %>%
  ungroup() %>%
  ggplot(aes(x = reorder(Country, val_percent), y = val_percent, fill = income_group)) +
  geom_col() +
  coord_flip() +
  labs(
    title = "Top 2 Countries per Income Group (Anxiety, 2020)",
    x = "Country",
```

```
  y = "Anxiety Percent (%)"
) +
theme_minimal()
```



Top Countries per Income Group: Anxiety Prevalence in 2020

This chart shows the top two countries with the highest anxiety rates in 2020 within each World Bank income group: High, Upper-middle, and Lower-middle income.

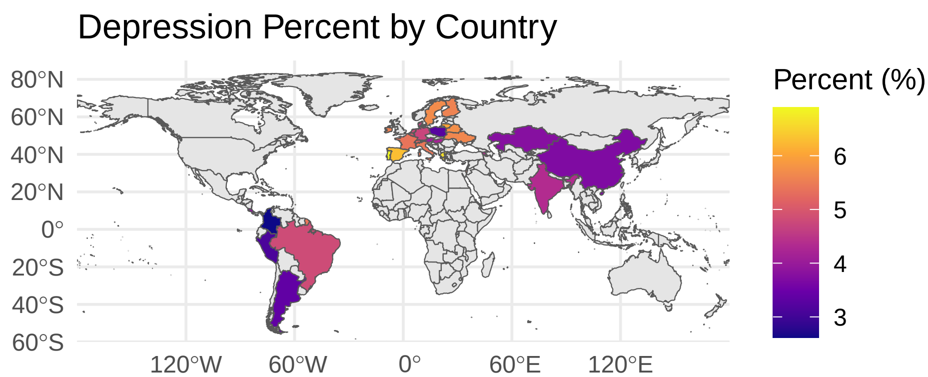
Findings

- High-Income: Portugal and Switzerland lead in anxiety prevalence. Portugal's rates may reflect lingering effects of the Eurozone debt crisis, youth unemployment, and austerity. Switzerland's anxiety may relate to social pressure, isolation, and greater mental health awareness.
- Upper-Middle-Income: Peru and Brazil have the highest rates. Brazil faces urban violence, inequality, and healthcare strain. Peru's challenges include political instability and economic volatility.
- Lower-Middle-Income: Ukraine and India top this group. Ukraine's high anxiety is linked to ongoing geopolitical conflict and economic insecurity. India's rates may be driven by rapid urbanization, population density, limited mental health infrastructure, and improved diagnostic reporting.

This comparison shows that anxiety is widespread across all income levels. National income alone doesn't predict anxiety prevalence; instead, socio-political factors, healthcare access, stigma, and stress exposures play major roles. The data highlights the need for tailored mental health strategies across all economic contexts.

```
ggplot() +
  geom_sf(data = world, fill = "grey80", color = "white") +
  geom_sf(data = world_dep, aes(fill = val_percent)) +
  scale_fill_viridis_c(option = "plasma", na.value = "grey90") +
   coord_sf(xlim = c(-180, 180), ylim = c(-60, 90), expand = FALSE)+
  theme_minimal() +
  labs(title = "Depression Percent by Country", fill = "Percent (%)")
```

3.1.2 Spatial Analysis (Maps)



Global Distribution of Depression Percent (2020)

A choropleth map was created to visualize country-level depression prevalence in 2020, highlighting regional patterns and geographic disparities beyond what tables or bar charts show.

Methodology
Using R's `ggplot2` and `sf` packages, depression data for "Depressive disorders" from the merged dataset was joined with global shapefiles by country name. The `val_percent` variable, indicating the population percentage affected, was mapped with a plasma viridis color scale for clarity and accessibility. Countries lacking data appear in grey.

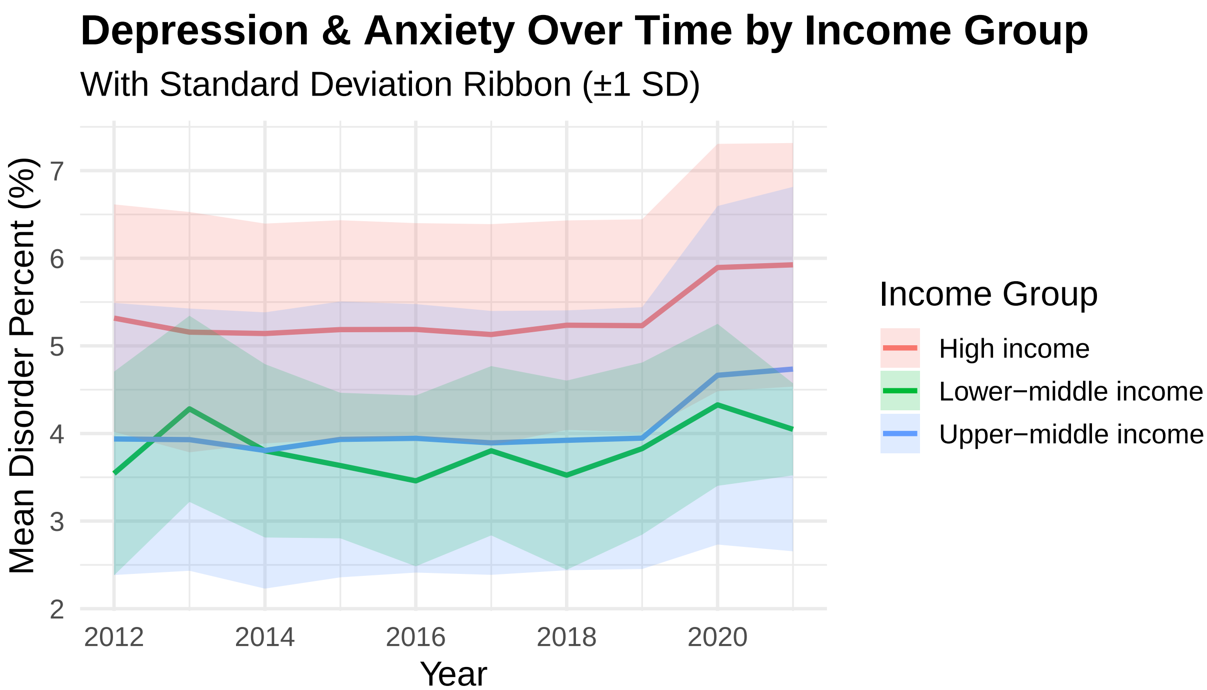
Findings
- Highest rates (6–7%) are seen in Southern and Eastern Europe (e.g., Portugal, Greece, Ukraine), possibly linked to economic instability, post-conflict stress, and limited mental health care.
- South Asia (India) and parts of South America (Brazil) show moderate to high rates, indicating significant mental health challenges in emerging economies.
- Lower rates (3–4%) appear in much of South America, Southeast Asia, though this may reflect underdiagnosis or cultural stigma rather than lower prevalence.
- Data gaps (grey areas) exist in regions like Africa and Central Asia, signaling a need for improved mental health data collection.

Implications
Depression affects all regions unevenly. Higher-income countries report more cases, likely due to better healthcare access and awareness, while lower-income countries may underreport due to stigma and systemic barriers. Mental health policies should consider regional economic and cultural contexts, and global surveillance needs strengthening to guide effective interventions and funding.

3.1.3 Sub-population analysis

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
```

```
## generated.
```



This graph depicts the average prevalence of depressive and anxiety disorders over time, broken down by income groups (Low, Lower-Middle, Upper-Middle, High). The colored lines show mean disorder rates, while the shaded ribbons represent ±1 standard deviation, highlighting variability within each group.

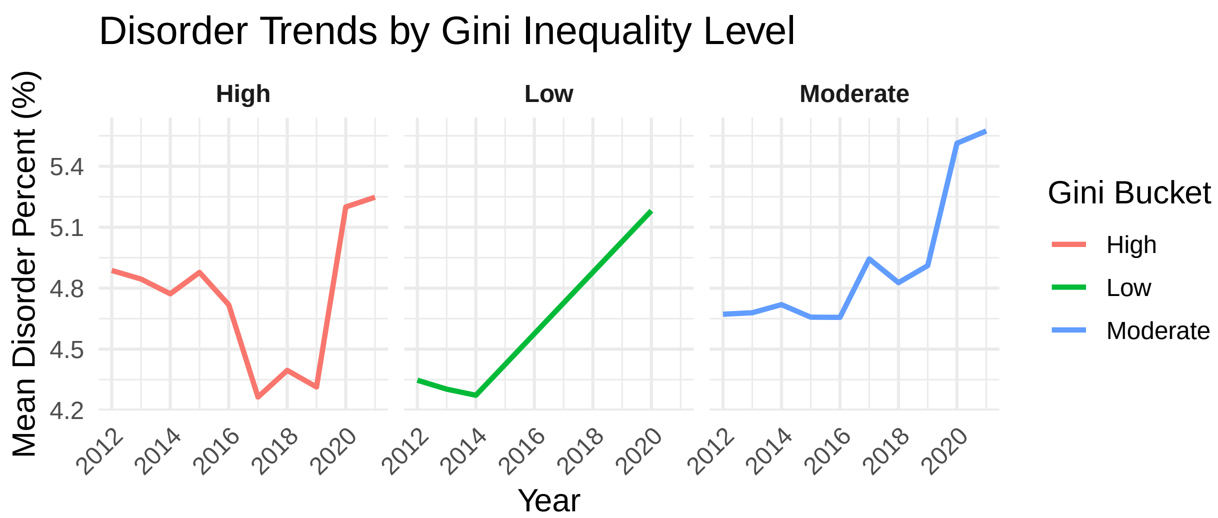
Key observations:

- High-income countries consistently show higher average disorder rates compared to lower-income groups.
- All income groups exhibit a gradual upward trend over the years.

The spread (ribbon width) indicates some variability within income groups but a generally steady increase across the board. This suggests that income level is associated with differences in mental health disorder prevalence, and the gradual increase points to a growing mental health challenge globally.

```r
dep_anx %>%
  filter(!is.na(gini_bucket)) %>%
  filter(cause %in% c("Depressive disorders", "Anxiety disorders")) %>%
  group_by(Year, gini_bucket) %>%
  summarise(mean_disorder = mean(val_percent, na.rm = TRUE), .groups = "drop") %>%
  ggplot(aes(x = Year, y = mean_disorder, color = gini_bucket)) +
  geom_line(size = 1) +
  facet_wrap(~ gini_bucket) +
  labs(
    title = "Disorder Trends by Gini Inequality Level",
    x = "Year",
    y = "Mean Disorder Percent (%)",
    color = "Gini Bucket"
  ) +
```

```r
  theme_minimal(base_size = 13) +
  theme(
    strip.text = element_text(face = "bold"),
    axis.text.x = element_text(angle = 45, hjust = 1)
  )
```

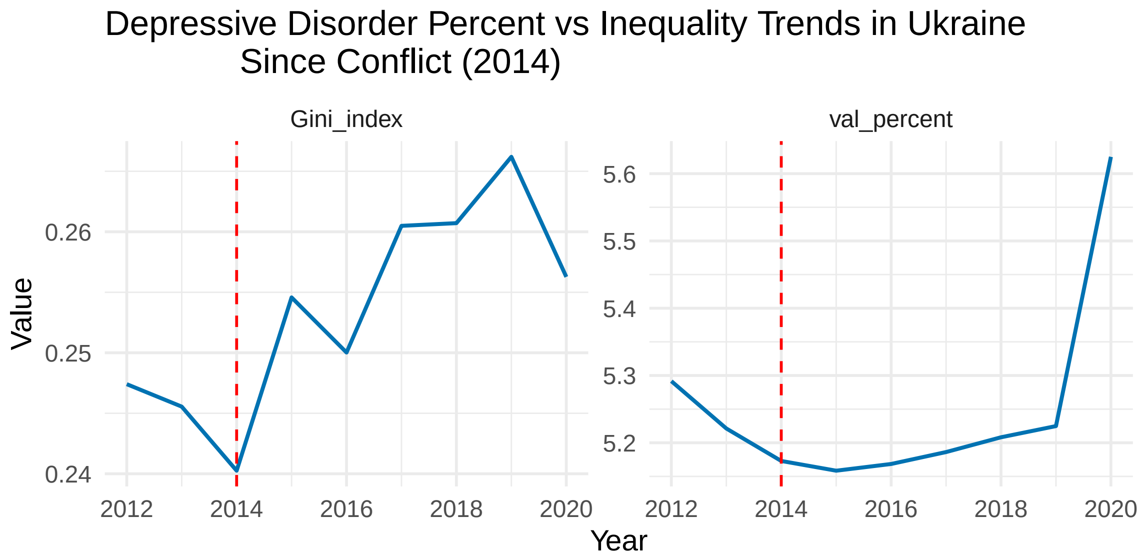


The graph shows the average prevalence of depressive and anxiety disorders from 2012 to 2021, categorized by low, moderate, and high income inequality (Gini index).

Countries with high inequality consistently exhibit higher disorder rates, while those with low inequality maintain lower rates. All groups show an upward trend over time, with a widening gap suggesting that greater economic inequality is linked to increased mental health burdens.

3.2 Event analysis

```r
ggplot(ukraine_long, aes(x = Year, y = value)) +
  geom_line(color = "#0072B2", size = 0.7) +
  geom_vline(xintercept = 2014, linetype = "dashed", color = "red") +
  facet_wrap(~ variable, scales = "free_y") +
  labs(
    title = "Depressive Disorder Percent vs Inequality Trends in Ukraine
             Since Conflict (2014)",
    x = "Year",
    y = "Value"
  ) +
  theme_minimal()
```



Event Analysis: Depression and Inequality in Ukraine since 2014

This analysis examines Ukraine's income inequality (Gini index) and depression prevalence from 2012 to 2020, focusing on changes after the 2014 conflict.

Data and Visualization
Trends in inequality and depression are shown side-by-side, with 2014 marked as the conflict start.

Findings
- **Income Inequality:** The Gini index rose from ~0.24 to ~0.26 after 2014, reflecting economic disruption from political instability and territorial losses.
- **Depression:** Percentages were stable before 2014 but increased significantly after 2015, peaking around 2020, likely due to trauma, insecurity, and reduced healthcare access.

Interpretation
Rising inequality and depression post-2014 suggest conflict worsens both economic and mental health challenges, with a delayed mental health impact. These results highlight the need for coordinated economic and mental health policies in post-conflict recovery.

4 Discussion

4.1 Discuss your findings

Analysis of temporal, spatial, and subgroup patterns reveals significant disparities in mental health disorder prevalence across time, income levels, and geography.

Temporal Trends: Depression and anxiety consistently show the highest global burden, with slight upward trends over time. While ADHD, bipolar, and eating disorders remain less prevalent, their relative stability may mask localized spikes or diagnostic gaps.

Cross-Country and Income Comparisons: High-income countries like Portugal and Greece exhibit the highest rates of depression, while low- and lower-middle-income countries show lower averages. This could reflect better diagnosis and awareness in wealthier nations, or underreporting and stigma in resource-limited settings. Notably, Ukraine's data shows high rates, likely driven by socio-political instability and rapid societal change.

Spatial Patterns: Choropleth maps highlight regional clusters—Southern and Eastern Europe, parts of South Asia, and South America—with above-average depression rates. In contrast, lower figures in Southeast Asia may indicate underdiagnosis rather than lower burden.

Subpopulation Insights: Income group analysis confirms that wealthier nations report higher average disorder rates, though all groups show a gradual increase over time. Similarly, a positive association is visible between inequality (Gini index) and disorder prevalence, reinforcing the connection between socioeconomic disparities and mental health outcomes.

5 Reproducibility

5.1 Github repository link

https://github.com/qgelena/Quantifying_a_social_problem/tree/main

5.2 Reference List

Global Burden of Disease Collaborative Network. (2024). *Global Burden of Disease Study 2021 (GBD 2021).* Institute for Health Metrics and Evaluation. https://www.healthdata.org/research-analysis/gbd

Kawachi, I., & Berkman, L. F. (2000). Social cohesion, social capital, and health. In *Social Epidemiology* (pp. 174–190).

Lund, C., Breen, A., Flisher, A. J., Kakuma, R., Corrigall, J., Joska, J. A., … & Patel, V. (2010). Poverty and common mental disorders in low and middle income countries: A systematic review. *Social Science & Medicine, 71*(3), 517–528. https://doi.org/10.1016/j.socscimed.2010.04.027

Patel, V., Burns, J. K., Dhingra, M., Tarver, L., Kohrt, B. A., & Lund, C. (2018). Income inequality and depression: A systematic review and meta-analysis of the association and a scoping review of mechanisms. *World Psychiatry, 17*(1), 76–89. https://doi.org/10.1002/wps.20492

Praeg, P., Mills, M., & Wittek, R. (2014). Income and income inequality as social determinants of health: Do social comparisons play a role? *European Sociological Review, 30*(2), 218–229. https://doi.org/10.1093/esr/jct035

Shahbazi, F., Shahbazi, M., & Poorolajal, J. (2022). Association between socioeconomic inequality and the global prevalence of anxiety and depressive disorders: An ecological study. *General Psychiatry, 35*(3), e100735. https://doi.org/10.1136/gpsych-2021-100735

Smith, K. E., Bambra, C., Hill, S. E., & Watt, R. G. (2012). Health inequalities and the social determinants of health: What works? *Journal of Public Health, 34*(4), 523–529. https://doi.org/10.1093/pubmed/fds052

World Health Organization. (2023). *Mental health.* https://www.who.int/health-topics/mental-health