

ECONOMETRICS

Socio-Medical Factors Affecting
SUICIDE RATE

The Roles of Mental Illness,
Alcohol Abuse, and Health Expenditure

GROUP 3

Get to know our group



Việt Huy



Hà Phương



Khánh Linh



Danh Dũng



Thu Hiền



Công Huy

I. Introduction

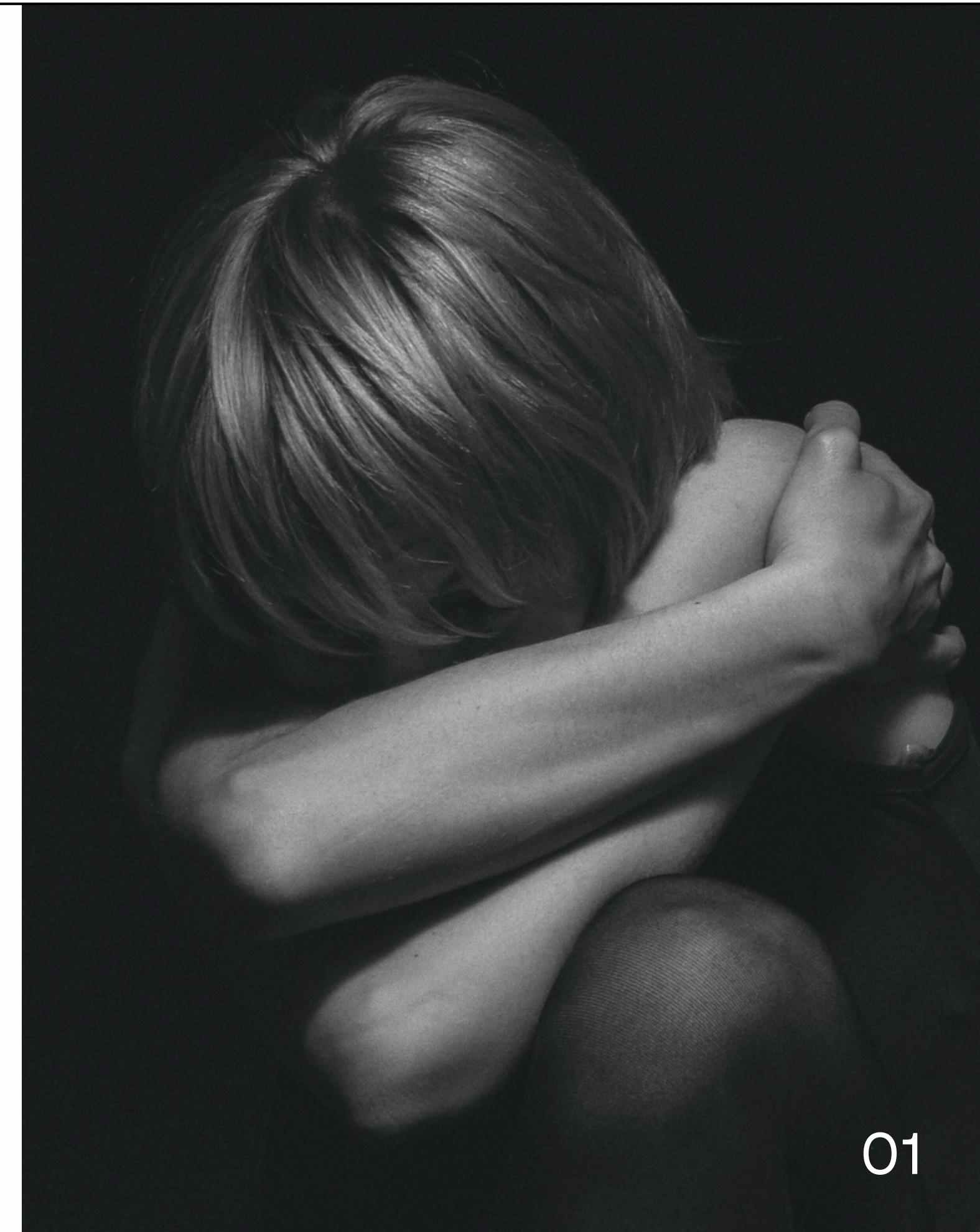
1. Research Motivation

Suicide accounted for 1.1% of global deaths in 2021, with **727,000** deaths annually (WHO).

It is **the third leading cause of death** among 15-29-year-olds in 2021.

Suicide affects families, communities, and countries, with **long-lasting impacts**.

Mental health disorders and Economic factors are key determinants contributing to suicidal behaviors.



I. Introduction

2. Research objectives

- 01** To examine the link between mental illness, health expenditure, economics factors and suicide rates.
- 02** To provide insights for designing effective prevention strategies and policies.
- 03** To understand the role of mental health in suicide prevention.



II. Data Overview

1. Data Sources

**900 cross-
sectional
observations**

**5 distinct time
points between
2000 and 2016**

**1 dependent variable
& 9 independent
variables**

**1. World Bank Open Data
2. Global Health Data
Analysis 1990–2019 dataset
(Kaggle)**

2. Variables description

Country name	Official name of the country.
Country code	Alphanumeric code for merging and consistency.
Continent	Denotes the continent of each country, controlling region-specific effects.
Year	The year of each observation (2000, 2004, 2008, 2012, 2016).
Suicide mortality rate	Suicide deaths per 100,000 individuals.
Bipolar disorder	Prevalence of bipolar disorder, per 100 individuals.
Anxiety disorder	Prevalence of anxiety disorder, per 100 individuals.

Depression	Prevalence of depression, per 100 individuals.
Eating disorders	Prevalence of eating disorders, per 100 individuals.
Alcohol disorder	Prevalence of alcohol-related disorders, per 100 individuals.
Current health expenditure per capita	Health expenditure per capita, as a proxy for mental health risks.
GDP per capita	GDP per midyear population, reflecting economic development.
Inflation	Yearly change in consumer prices, indicating economic stability.
Unemployment	Share of the labor force that is unemployed.

3. Why focus on these mental health disorders?

01

Mental Health Disorders and Suicide:

Conditions like depression, anxiety, bipolar disorder, eating disorders, and substance abuse are major contributors to suicide risk.

02

Impact on Vulnerable Groups:

These disorders often lead to hopelessness and isolation, especially affecting vulnerable populations like adolescents and women.

03

Barriers to Care:

Many individuals with mental health disorders are untreated, with over 50% of suicide victims not diagnosed, highlighting the need for better access to care.

4. Data Cleaning and Preprocessing

```
<class 'pandas.core.frame.DataFrame'>
Index: 588 entries, 3 to 899
Data columns (total 14 columns):
 #   Column
 ---  -----
 0   country_name
 1   country_code
 2   year
 3   suicide_mortality_rate
 4   log_prevalence_of_bipolar_disorder
 5   log_prevalence_of_anxiety_disorder
 6   log_prevalence_of_depression
 7   log_alcohol_use_disorders
 8   log_prevalence_of_eating_disorders
 9   continent
 10  log_current_health_expenditure_per_capita
 11  log_gdp_per_capita
 12  inflation
 13  unemployment
dtypes: float64(10), int64(1), object(3)
memory usage: 68.9+ KB
```

	Non-Null Count	Dtype
0	588 non-null	object
1	588 non-null	object
2	588 non-null	int64
3	588 non-null	float64
4	588 non-null	float64
5	588 non-null	float64
6	588 non-null	float64
7	588 non-null	float64
8	588 non-null	float64
9	588 non-null	object
10	588 non-null	float64
11	588 non-null	float64
12	588 non-null	float64
13	588 non-null	float64

After removing the outliers, the boxplots show that:

- Most variables have an even distribution, with no significant outliers
- Except for the unemployment rate, whose outliers were retained to preserve important information.

III. Modelling

1. Initial Model

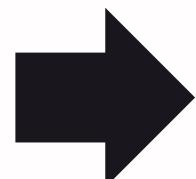
suicide_mortality_rate = $\beta_0 + \beta_1 * \log_{\text{prevalence_of_bipolar_disorder}}$
+ $\beta_2 * \log_{\text{prevalence_of_anxiety_disorder}}$
+ $\beta_3 * \log_{\text{prevalence_of_depression}}$
+ $\beta_4 * \log_{\text{alcohol_use_disorders}}$
+ $\beta_5 * \log_{\text{prevalence_of_eating_disorders}}$
+ $\beta_6 * \log_{\text{current_health_expenditure_per_capita}}$
+ $\beta_7 * \log_{\text{gdp_per_capita}}$
+ $\beta_8 * \text{inflation} + \beta_9 * \text{unemployment} + \varepsilon$

III. Modelling

2. Variable Selection

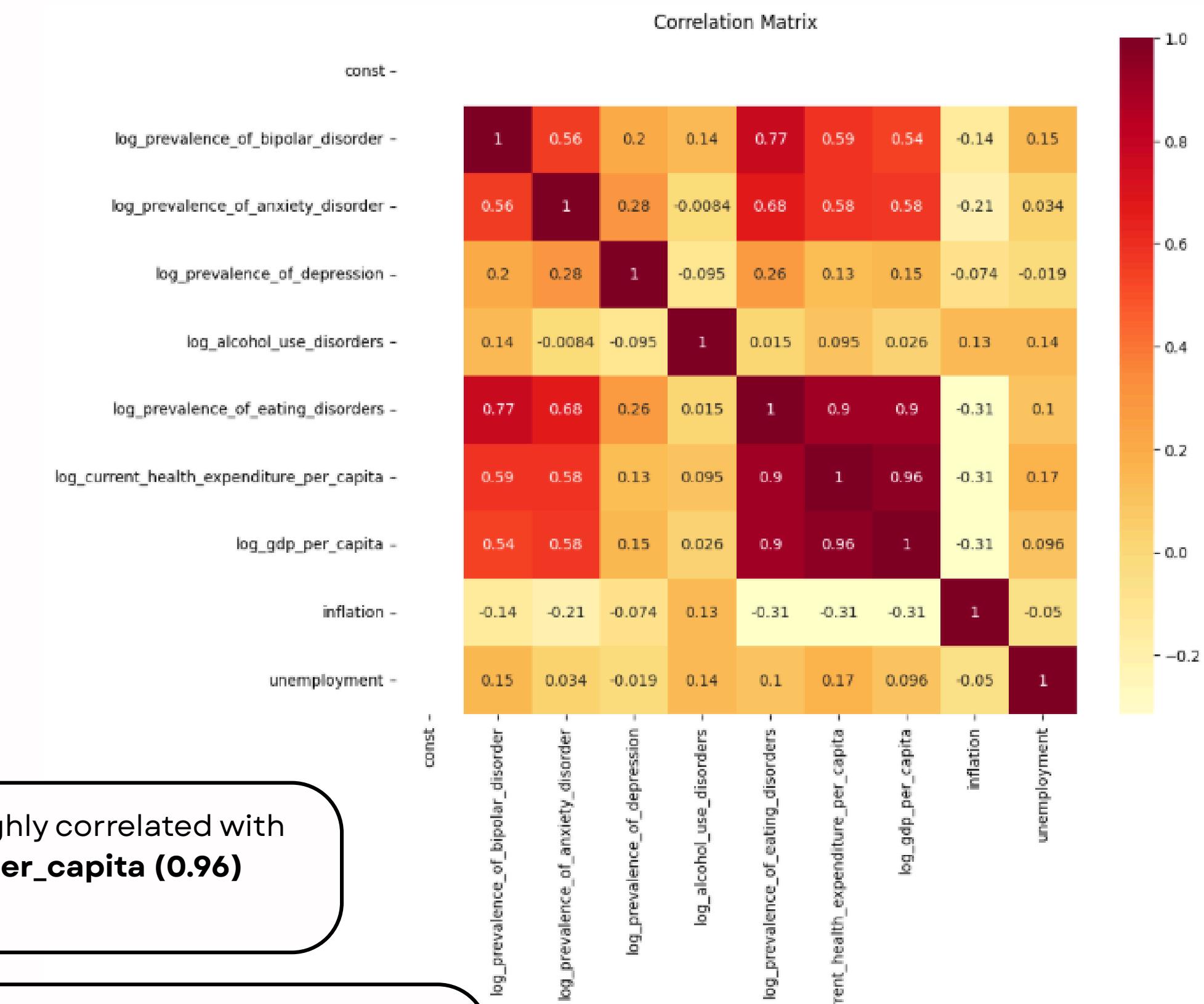
Variable	VIF
const	3413.808579
log_prevalence_of_bipolar_disorder	4.033381
log_prevalence_of_anxiety_disorder	1.936941
log_prevalence_of_depression	1.188261
log_alcohol_use_disorders	1.175772
log_prevalence_of_eating_disorder	17.818187
log_current_health_expenditure_per_capita	15.964839
log_gdp_per_capital	17.521044
inflation	1.167006
unemployment	1.119337

table 2. VIF results of model 1



The correlation matrix shows that **log_gdp_per_capita** is highly correlated with key variables, especially **log_current_health_expenditure_per_capita (0.96)** and **log_prevalence_of_eating_disorders (0.90)**.

log_gdp_per_capita is not directly related to suicide rates, but **log_current_health_expenditure_per_capita** and **log_prevalence_of_eating_disorders** provide more relevant insights into mental health and healthcare systems.



III. Modelling

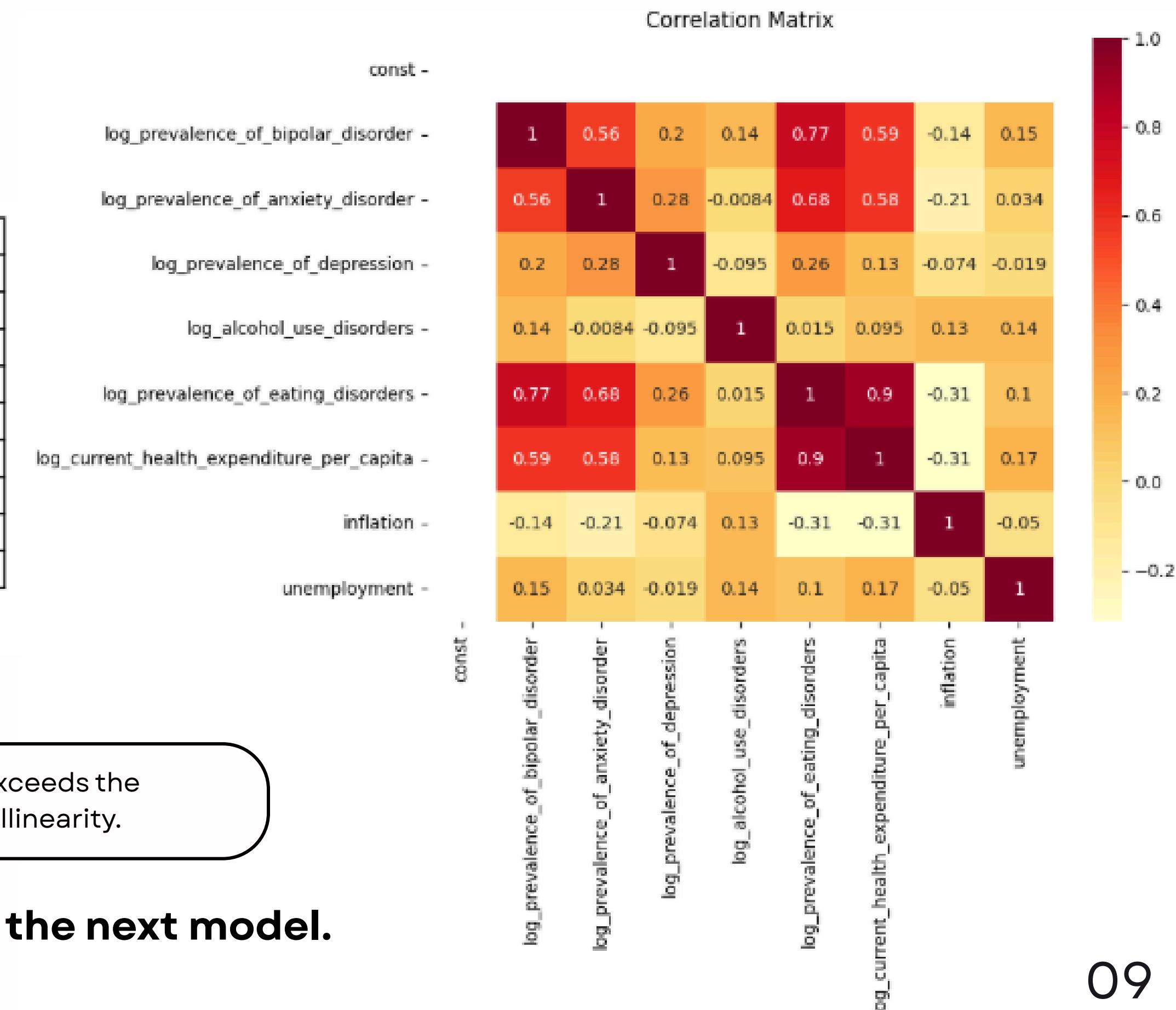
2. Variable Selection

Variable	VIF
const	3231.134275
log_prevalence_of_bipolar_disorder	3.411409
log_prevalence_of_anxiety_disorder	1.931323
log_prevalence_of_depression	1.185238
log_alcohol_use_disorders	1.171068
log_prevalence_of_eating_disorder	14.089693
log_current_health_expenditure_per_capita	7.930084
inflation	1.163050
unemployment	1.095427

table 4. VIF results of model 2

The VIF of log_prevalence_of_eating_disorders is 14.09, which exceeds the commonly accepted threshold of 10, indicating severe multicollinearity.

→ We choose to remove this variable in the next model.



III. Modelling

2. Variable Selection

OLS Regression Results

	coef	std err	t	P> t	[0.025	0.975]
const	-55.1756	9.779	-5.642	0.000	-74.382	-35.970
log_prevalence_of_bipolar_disorder	-3.5215	0.597	-5.895	0.000	-4.695	-2.348
log_prevalence_of_anxiety_disorder	-5.4350	0.861	-6.309	0.000	-7.127	-3.743
log_prevalence_of_depression	10.7547	1.033	10.416	0.000	8.727	12.783
log_alcohol_use_disorders	4.5938	0.415	11.067	0.000	3.779	5.409
log_current_health_expenditure_per_capita	2.0155	0.176	11.459	0.000	1.670	2.361
inflation	-0.0551	0.047	-1.166	0.244	-0.148	0.038
unemployment	0.0104	0.035	0.294	0.769	-0.059	0.079
Omnibus:	51.217	Durbin-Watson:		0.684		
Prob(Omnibus):	0.000	Jarque-Bera (JB):		62.593		
Skew:	0.752	Prob(JB):		2.56e-14		
Kurtosis:	3.542	Cond. No.		1.01e+03		

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.01e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Variable	VIF
const	2896.598764
log_prevalence_of_bipolar_disorder	1.783510
log_prevalence_of_anxiety_disorder	1.834156
log_prevalence_of_depression	1.110304
log_alcohol_use_disorders	1.088232
log_current_health_expenditure_per_capital	1.953321
inflation	1.147681
unemployment	1.059134

table 6. VIF results of model 3

3. Model diagnostics

Ramsey's test for linear model

```
Ramsey RESET Test: <F test: F=0.015245417955895837, p=0.9017759272935483, df_denom=579, df_num=1>
```

No error, no omitted variable

The model is satisfied linearity assumption

t-test for zero mean

```
T-Statistic: 2.5585628505742805e-13, P-Value: 0.9999999999997959
```

VIF for multicollinear

	feature	VIF
0	const	2896.598764
1	log_prevalence_of_bipolar_disorder	1.783510
2	log_prevalence_of_anxiety_disorder	1.834156
3	log_prevalence_of_depression	1.110304
4	log_alcohol_use_disorders	1.088232
5	log_current_health_expenditure_per_capita	1.953321
6	inflation	1.147681
7	unemployment	1.059134

BP-test for Heteroskedasticity

→ Breusch-Pagan Test Results:

```
Lagrange multiplier statistic: 3.4713e+01
p-value: 1.2663e-05
F-statistic: 5.1984e+00
F p-value: 9.2271e-06
```

Homoskedasticity is violated.

```
1.2662710917134855e-05
```

Durbin-Watson test for autocorrelation

Durbin-Watson test for autocorrelation:

Independent variable: log_prevalence_of_bipolar_disorder
Durbin-Watson statistic: 0.5775

Independent variable: log_prevalence_of_anxiety_disorder
Durbin-Watson statistic: 0.5721

Independent variable: log_prevalence_of_depression
Durbin-Watson statistic: 0.5916

Independent variable: log_alcohol_use_disorders
Durbin-Watson statistic: 0.5488

Independent variable: log_current_health_expenditure_per_capita
Durbin-Watson statistic: 0.6355

Independent variable: inflation
Durbin-Watson statistic: 0.5998

Independent variable: unemployment
Durbin-Watson statistic: 0.5733

3. Model diagnostics

We employed a Generalized Least Squares (GLS) model to improve estimation efficiency and inference validity.

GLS + HAC:							
GLS Regression Results							
Dep. Variable:	suicide_mortality_rate	R-squared:	0.344				
Model:	GLS	Adj. R-squared:	0.336				
Method:	Least Squares	F-statistic:	29.46				
Date:	Sun, 20 Apr 2025	Prob (F-statistic):	7.63e-35				
Time:	15:44:37	Log-Likelihood:	-1687.6				
No. Observations:	588	AIC:	3391.				
Df Residuals:	580	BIC:	3426.				
Df Model:	7						
Covariance Type:	HAC						

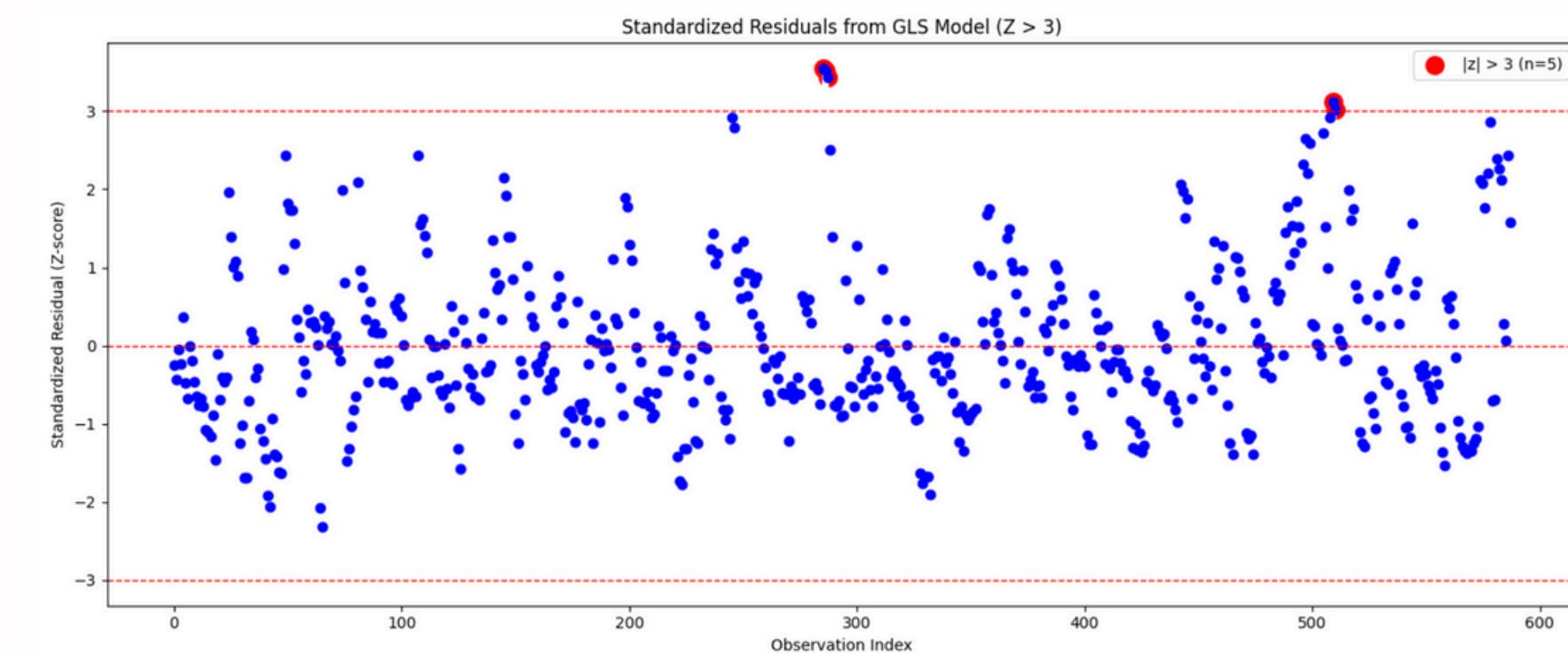
		coef	std err	t	P> t	[0.025	0.975]
const		-65.8462	13.985	-4.708	0.000	-93.313	-38.380
log_prevalence_of_bipolar_disorder		-3.7441	0.796	-4.703	0.000	-5.308	-2.180
log_prevalence_of_anxiety_disorder		-4.5182	1.266	-3.569	0.000	-7.005	-2.032
log_prevalence_of_depression		11.3976	1.113	10.242	0.000	9.212	13.583
log_alcohol_use_disorders		4.7672	0.500	9.543	0.000	3.786	5.748
log_current_health_expenditure_per_capita		1.7376	0.257	6.772	0.000	1.234	2.241
inflation		-0.0859	0.047	-1.832	0.067	-0.178	0.006
unemployment		-0.0182	0.036	-0.506	0.613	-0.089	0.053

Omnibus:	65.530	Durbin-Watson:	0.673				
Prob(Omnibus):	0.000	Jarque-Bera (JB):	85.635				
Skew:	0.865	Prob(JB):	2.54e-19				
Kurtosis:	3.710	Cond. No.	1.07e+03				

4. Residual Diagnostics and Influential Observations in the GLS Model

Observations with absolute Z-scores greater than 3 were flagged as statistical outliers.

A GLS-specific influence score was computed using a leave-one-out (LOO) approach.



	country_name	year	suicide_mortality_rate	gls_outlier_z>3	gls_influence_score
410	Japan	2000	23.9	True	3.679987
411	Japan	2004	24.1	True	3.357421
565	Myanmar	2000	4.7	False	3.172833
412	Japan	2008	24.4	True	2.945951
155	Central African Republic	2000	19.0	False	2.568986
566	Myanmar	2004	4.2	False	1.873788
413	Japan	2012	21.6	False	1.777083
157	Central African Republic	2008	15.1	False	1.609349
111	Brazil	2004	4.4	False	1.493545
156	Central African Republic	2004	15.7	False	1.483687

5. Model comparison

	R-squared	Adj. R-squared	AIC	BIC
Model 1	0.396	0.387	3414.	3458.
Model 2	0.392	0.383	3416.	3455.
Model 3	0.385	0.378	3421.	3456.
GLS Model	0.344	0.336	3391.	3426.
Cleaned GLS Model	0.363	0.356	3300.	3335.

→ Cleaned GLS Model is the optimal choice

IV. Final model & results

1. Final regression equation

GLS + HAC:

GLS Regression Results

	coef	std err	t	P> t	[0.025	0.975]
const	-77.5773	11.765	-6.594	0.000	-100.686	-54.469
log_prevalence_of_bipolar_disorder	-3.9368	0.777	-5.067	0.000	-5.463	-2.411
log_prevalence_of_anxiety_disorder	-3.6492	1.047	-3.486	0.001	-5.706	-1.593
log_prevalence_of_depression	11.8557	0.985	12.031	0.000	9.920	13.791
log_alcohol_use_disorders	5.0860	0.454	11.204	0.000	4.194	5.978
log_current_health_expenditure_per_capita	1.6412	0.224	7.336	0.000	1.202	2.081
inflation	-0.0691	0.044	-1.562	0.119	-0.156	0.018
unemployment	-0.0044	0.035	-0.128	0.898	-0.072	0.064
Omnibus:	41.966	Durbin-Watson:		0.706		
Prob(Omnibus):	0.000	Jarque-Bera (JB):		49.203		
Skew:	0.701	Prob(JB):		2.07e-11		
Kurtosis:	3.248	Cond. No.		1.09e+03		

$$\begin{aligned}
 \text{suicide_mortality_rate} = & -77.5773 + (-3.9368) * \log_{\text{prevalence_of_bipolar_disorder}} \\
 & + (-3.6492) * \log_{\text{prevalence_of_anxiety_disorder}} \\
 & + 11.8557 * \log_{\text{prevalence_of_depression}} \\
 & + 5.0860 * \log_{\text{alcohol_use_disorders}} \\
 & + 1.6412 * \log_{\text{current_health_expenditure_per_capita}} \\
 & + (-0.0691) * \text{inflation} + (-0.0044) * \text{unemployment} + \varepsilon
 \end{aligned}$$

IV. Final model & results

2. Estimation results and interpretation

Significant Predictors

- log_prevalence_of_depression: Strong positive relationship with suicide mortality ($\beta = 11.86$, $p < 0.01$).
- log_alcohol_use_disorders: Positive association with suicide mortality ($\beta = 5.086$, $p < 0.01$).
- log_current_health_expenditure_per_capita: Positive effect on suicide rates ($\beta = 1.641$, $p < 0.01$).

Signs and Magnitudes of Coefficients

- log_prevalence_of_bipolar_disorder and log_prevalence_of_anxiety_disorder: Negative relationship with suicide rates ($\beta = -3.937$, $p < 0.01$; $\beta = -3.649$, $p < 0.01$).
- Economic factors: Inflation has a small negative effect ($\beta = -0.0691$, $p = 0.088$); unemployment shows no significant impact ($\beta = -0.0044$, $p = 0.888$).

V. Conclusion & Recommendations

1. Conclusions

The study highlights that depression and alcohol abuse are key factors influencing suicide rates, with depression showing a strong positive correlation. However, an increase in healthcare spending appears to be paradoxically linked to rising suicide rates, potentially indicating better detection rather than effective prevention. Additionally, mental disorders like anxiety and bipolar disorder show an inverse relationship with suicide, raising questions about the effectiveness of treatment.

V. Conclusion & Recommendations

2. Recommendations

- 01 Invest in Mental Health:** Focus on improving mental health care and therapeutic approaches to address suicide's root causes.
- 02 Regulate Alcohol Consumption:** Implement stricter controls on alcohol to reduce its impact on mental and physical health.
- 03 Review Healthcare Spending:** Investigate the relationship between healthcare expenditure and suicide rates to ensure effective resource allocation.
- 04 Conduct Further Research:** Explore the inverse relationship between mental disorders and suicide to improve treatment outcomes.

*"Together against suicide,
Bringing hope for the future."*

THE END