Economics & Emissions

Obesity

Earth Girls: Amanda Huang, Diyang Lu, Nan Nie, Siqi Ma

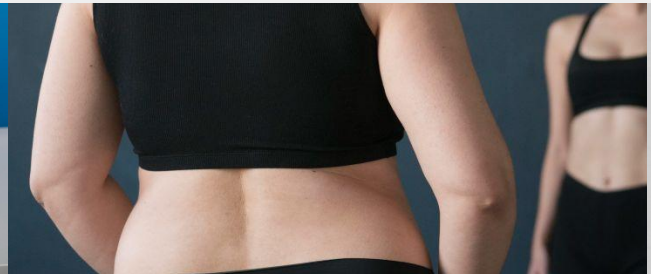


TABLE OF CONTENTS

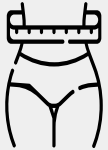
Background

01

Background

Economic

Environment



Public Health

Previous Researches

Hammond

There is no direct cost estimate assigned to greenhouse gas emissions due to obesity, but it acknowledges the intertwining of obesity, economic factors, and health risks

Flechtner-Mors

Economic growth in the United States was correlated with increasing CO2 emissions, which have indirect implications for obesity through lifestyle changes

Tomiyama

The relationship between economic growth and obesity has been explored, highlighting implications for both health and environmental impacts such as carbon emissions

Bryant

There is evidence to suggest that rising obesity rates may contribute to greenhouse gas emissions both directly through increased food production and indirectly through lifestyle changes

Hypothesis

Null Hypothesis (H0)

There is no significant impact of the ten variables (MLN_USD, USD_CAP, AVWAGE, FERTILITY, EMP, UR, HRWKD, GGDEBT, YNGPOP, MtCO2) on adult obesity rates.

Alternative Hypothesis (H1)

At least one of the ten variables (MLN_USD, USD_CAP, AVWAGE, FERTILITY, EMP, UR, HRWKD, GGDEBT, YNGPOP, MtCO2) significantly impacts adult obesity rates.

Dataset

02

Source

Country Economic Indicators

Source: OECD
Indicators: MLN_USD, USD_CAP, AVWAGE, FERTILITY, EMP, UR, HRWKD, GGDEBT, YNGPOP

Emissions by Country

Source: GCP
Indicator: MtCO2

Obesity among Adults by Country

Obesity among Adults by Country
Source: World Health Organization
Indicator: Obesity prevalence

Data Preparation and Cleaning Overview

01

Obesity Data Prep

·Import and refine dataset
·Focus on overall rates
·Reformat 'Obesity (%)'

02

Country Code Standardization

·Import ISO codes
·Merge with obesity data

03

Economic Indicators

·Import from OECD
·Select relevant columns
·Data reshaping & cleaning

04

Emissions Data Integration

·Import CO2 emissions data
·Standardize country names

05

Data Merging

·Merge based on 'Country_Code' & 'Year'
·Inner join for data completeness

Key Variables Overview

	NO.	Variable Name	Description
Country Economic Indicators Dataset	1	MLN_USD	Total economic output of the country in millions of U.S. dollars
	2	USD_CAP	Economic output per capita in U.S. dollars
	3	AVWAGE	Average annual wages in the country
	4	FERTILITY	Fertility rate, representing the average number of children a woman will have
	5	EMP	Employment rate, the percentage of the working-age population that is employed
	6	UR	Unemployment rate, the percentage of the labor force that is unemployed

Table continued on next slide

	NO.	Variable Name	Description
Country Economic Indicators Dataset	7	HRWKD	Average hours worked per week
	8	GGDEBT	Gross government debt as a percentage of GDP
	9	YNGPOP	Percentage of the population that is young (usually defined as under 15 or 18 years)
Emissions by Country Dataset	10	MtCO2	Emissions measured in metric tons of CO2
Obesity among Adults by Country Dataset	11	Obesity	Prevalence of obesity among adults, often expressed as a percentage of the population

Focus on 2016 Data

Reason

- Pre-COVID period for consistent comparison
- Data completeness and reliability
- Representative of global trends

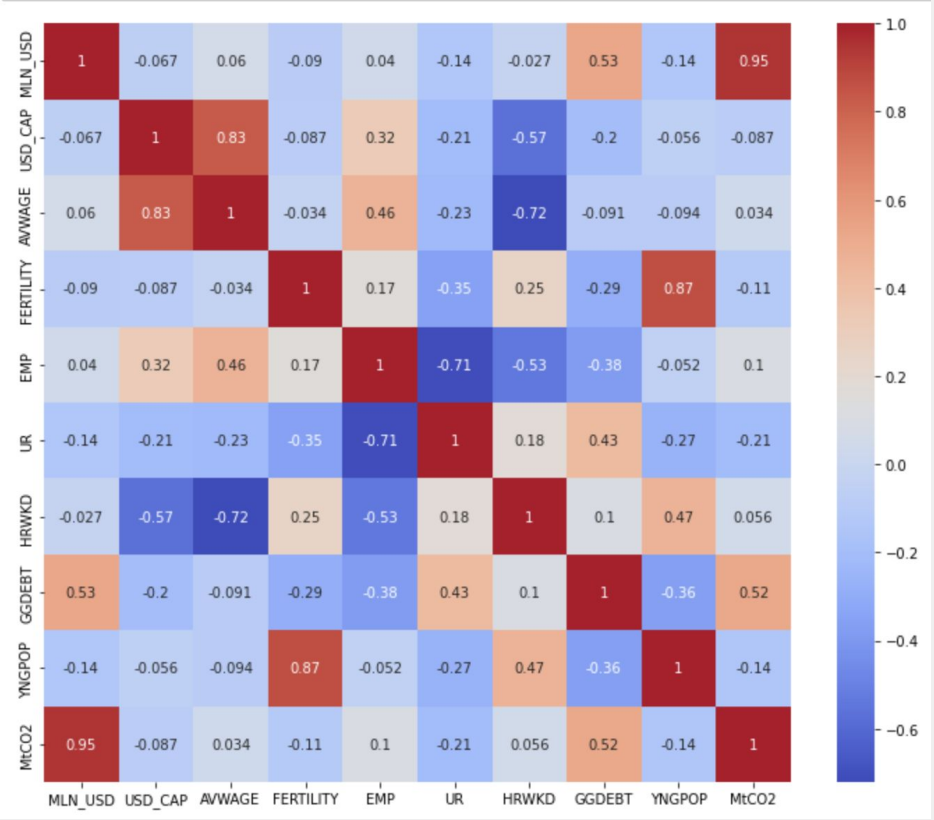
Extraction

- Creation of df_2016 dataframe
- 30 observations representing key indicators

03

Methodology

Correlation



Positive correlation between total economic output and emissions (0.95)

Young population and Fertility (0.87)

Average wage and economic output per capita(0.83)

Assumption 1: Collinearity

After deleting Obesity

feature	VIF
const	1422.263135
MLN_USD	15.752480
USD_CAP	3.915619
AVWAGE	6.461392
FERTILITY	7.719707
EMP	4.432994
UR	3.277995
HRWKD	6.475678
YNGPOP	11.715303
GGDEBT	3.124238
MtCO2	18.685135

- MtCO2 - important independent variable
- MLN_USD and USD_CAP are all economic indicators
- Second biggest

Delete MLN_USD

Assumption 1: Collinearity

After deleting MLN_USD

	feature	VIF
0	const	1147.247128
1	USD_CAP	3.915079
2	AVWAGE	6.305317
3	FERTILITY	7.692634
4	EMP	4.065157
5	UR	3.241069
6	HRWKD	4.983642
7	YNGPOP	11.544099
8	GGDEBT	3.121381
9	MtCO2	2.194338

Attempt 1

Delete AVWAGE and YNGPOP directly

Attempt 2

Try to combine some of them

Assumption 1: Collinearity

Attempt 1

feature	VIF
const	1132.203242
USD_CAP	1.573431
FERTILITY	1.383803
EMP	3.251171
UR	3.076783
HRWKD	2.466367
GGDEBT	2.424678
MtCO2	2.114772

After deleting YNG_POP and AVWAGE

Attempt 2

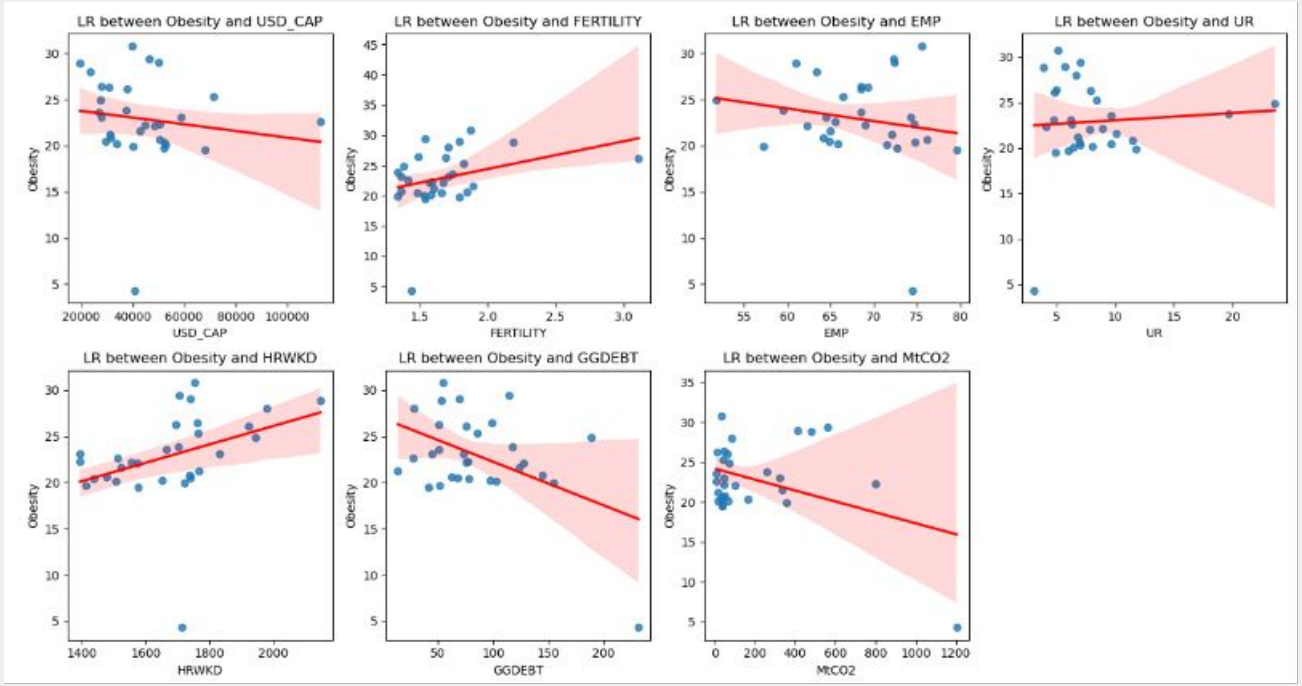
	feature	VIF
0	const	358.839870
1	USD_CAP	2.781762
2	YNGxFer	1.674867
3	WagexEmployment	4.786740
4	UR	2.108346
5	HRWKD	3.197445
6	GGDEBT	2.353892
7	MtCO2	2.159999

YNG x FER : the extent to which this group contributes to the total fertility rate

Wage x Employment: the overall wage income level

Assumption 2: linearity

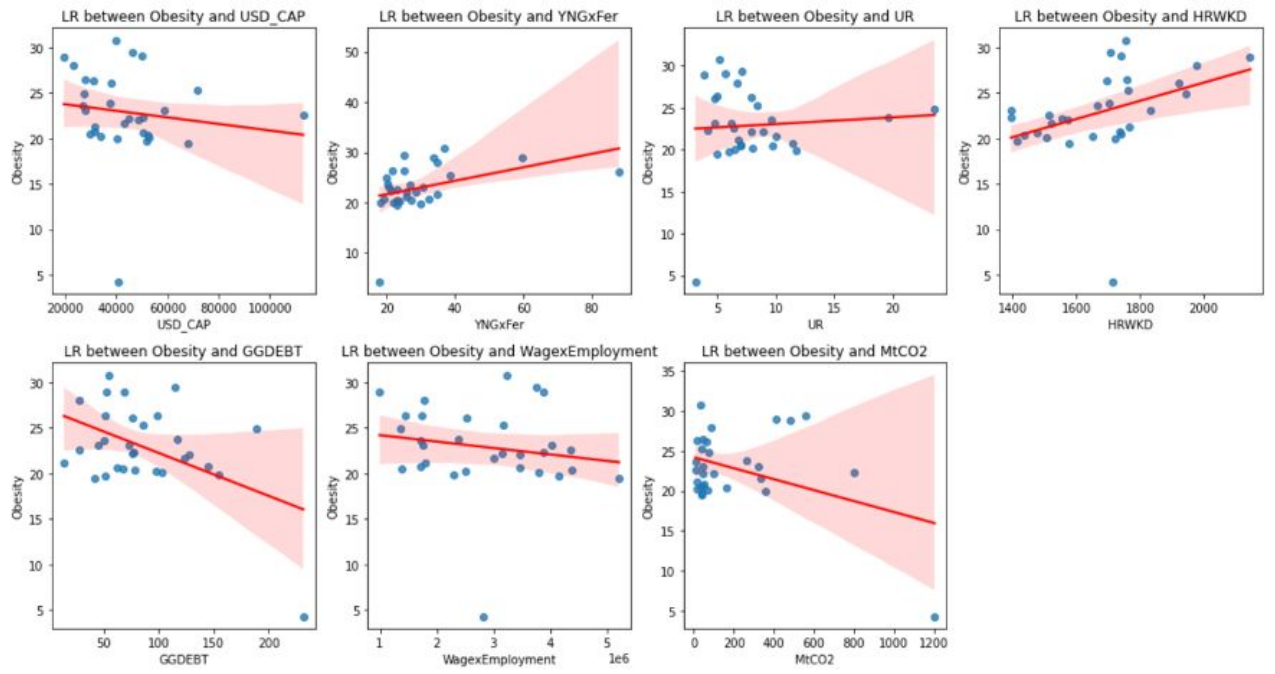
Attempt 1: deleting features



The remaining features have some degree of linearity, although pretty weak

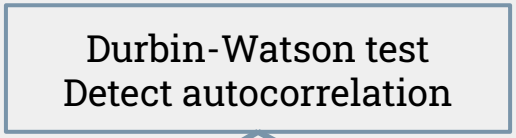
Assumption 2: linearity

Attempt 2: combining features



There exists linearity between features and the Obesity, although some are weak

Assumption 3: Independent Residuals




```
from statsmodels.stats.stattools import durbin_watson

#perform Durbin-Watson test
durbin_watson(model.resid)
```

```
1.4177984017671879
```

Attempt 1
Delete features

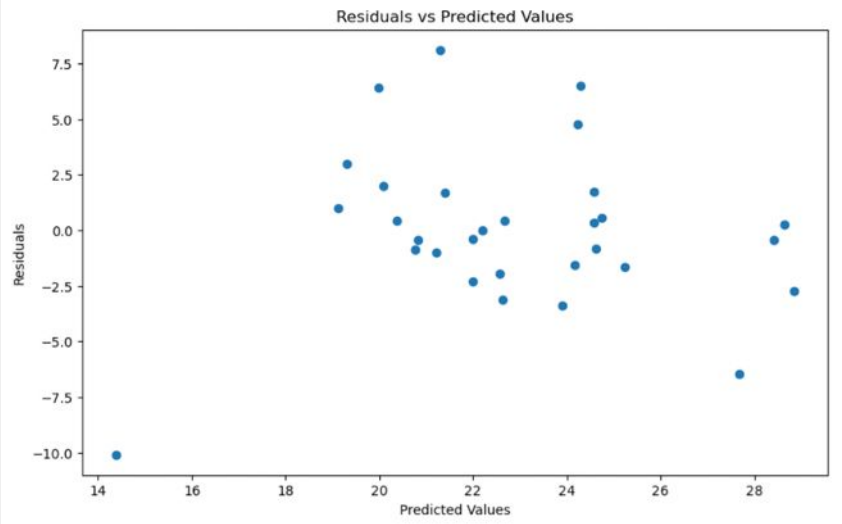
```
from statsmodels.stats.stattools import durbin_watson

#perform Durbin-Watson test
durbin_watson(model.resid)
```

```
1.4655253404708712
```

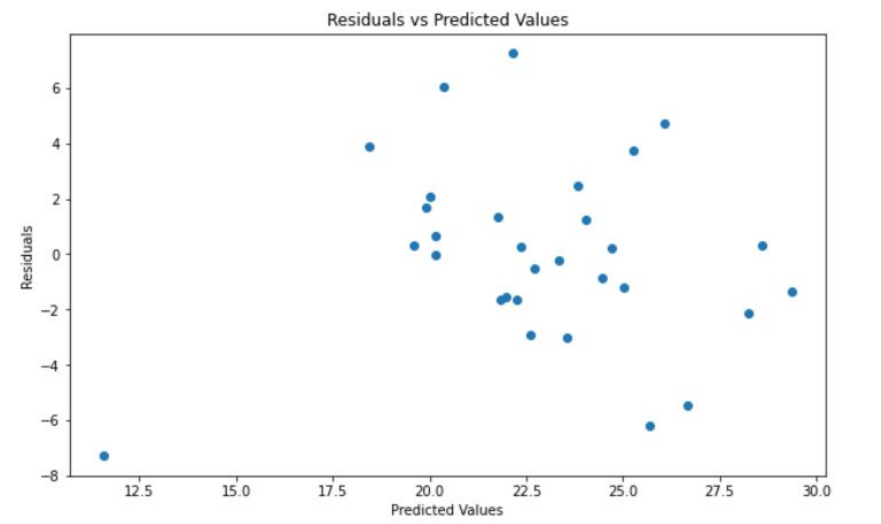
Attempt 2
Combine features

Assumption 4: Homoscedasticity

Attempt 1
Delete features



Attempt 2
Combine features



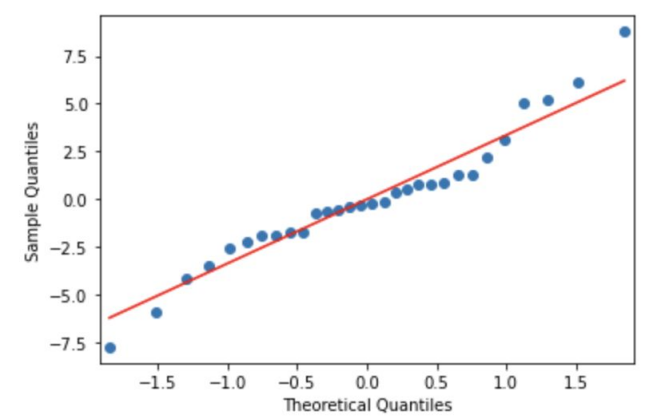
The points are randomly distributed around the horizontal axis with no clear pattern

Homoscedasticity

Assumption 5: Residuals Normality

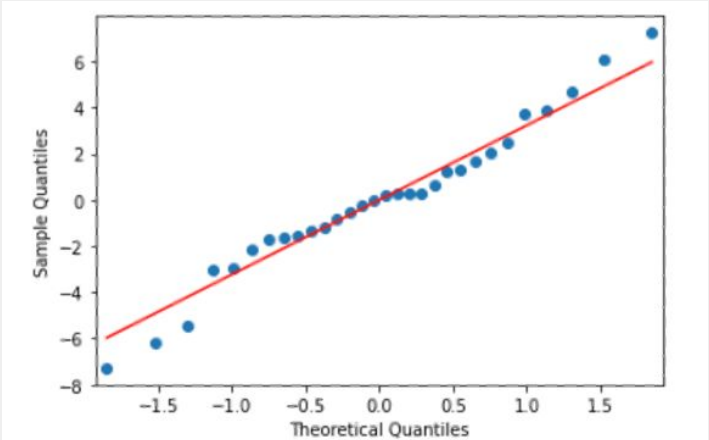
Attempt 1
Delete features

Q-Q Plot




```
Statistics=0.962, p=0.352
Sample looks Gaussian (fail to reject H0)
```

Attempt 2
Combine features



Q-Q Plot

```
Statistics=0.977, p=0.750
Sample looks Gaussian (fail to reject H0)
```

H0: data are normally distributed

Shapiro-Wilk test

- Statistics close to 1
- p-value greater than 0.05

Fail to reject H0
Residuals are normally distributed

04

Results

Results

Attempt 1
Delete features

OLS Regression Results

Dep. Variable:	Obesity	R-squared:	0.496
Model:	OLS	Adj. R-squared:	0.335
Method:	Least Squares	F-statistic:	3.087
Date:	Wed, 13 Dec 2023	Prob (F-statistic):	0.0200
Time:	10:27:25	Log-Likelihood:	-78.869
No. Observations:	30	AIC:	173.7
Df Residuals:	22	BIC:	184.9
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	22.8900	0.715	32.015	0.000	21.407	24.373
USD_CAP	0.1242	0.897	0.138	0.891	-1.736	1.984
FERTILITY	0.8733	0.841	1.038	0.310	-0.871	2.618
EMP	-0.2876	1.289	-0.223	0.826	-2.961	2.386
HRWKD	1.5192	1.123	1.353	0.190	-0.809	3.848
UR	1.3822	1.254	1.102	0.282	-1.219	3.983
GGDEBT	-2.8016	1.113	-2.516	0.020	-5.110	-0.493
MtCO2	-0.0487	1.040	-0.047	0.963	-2.205	2.108

Omnibus:	2.552	Durbin-Watson:	1.418
Prob(Omnibus):	0.279	Jarque-Bera (JB):	1.265
Skew:	0.340	Prob(JB):	0.531
Kurtosis:	3.742	Cond. No.	3.85

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

R-squared: 0.496

Attempt 2
Combine features

OLS Regression Results

Dep. Variable:	Obesity	R-squared:	0.532
Model:	OLS	Adj. R-squared:	0.383
Method:	Least Squares	F-statistic:	3.567
Date:	Wed, 13 Dec 2023	Prob (F-statistic):	0.0103
Time:	02:22:11	Log-Likelihood:	-77.756
No. Observations:	30	AIC:	171.5
Df Residuals:	22	BIC:	182.7
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	22.8900	0.689	33.224	0.000	21.461	24.319
USD_CAP	-0.9413	1.149	-0.819	0.421	-3.324	1.442
YNGxFer	0.6379	0.892	0.715	0.482	-1.211	2.487
WagexEmployment	1.9268	1.507	1.278	0.214	-1.199	5.053
HRWKD	2.4029	1.232	1.950	0.064	-0.152	4.958
UR	1.6522	1.000	1.652	0.113	-0.422	3.727
GGDEBT	-2.7060	1.057	-2.560	0.018	-4.898	-0.514
MtCO2	-0.3464	1.013	-0.342	0.736	-2.446	1.754

Omnibus:	0.588	Durbin-Watson:	1.466
Prob(Omnibus):	0.745	Jarque-Bera (JB):	0.050
Skew:	-0.018	Prob(JB):	0.975
Kurtosis:	3.196	Cond. No.	4.59

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

R-squared: 0.532

Only government debt has a significant effect on the obesity rate

Notable Attempts

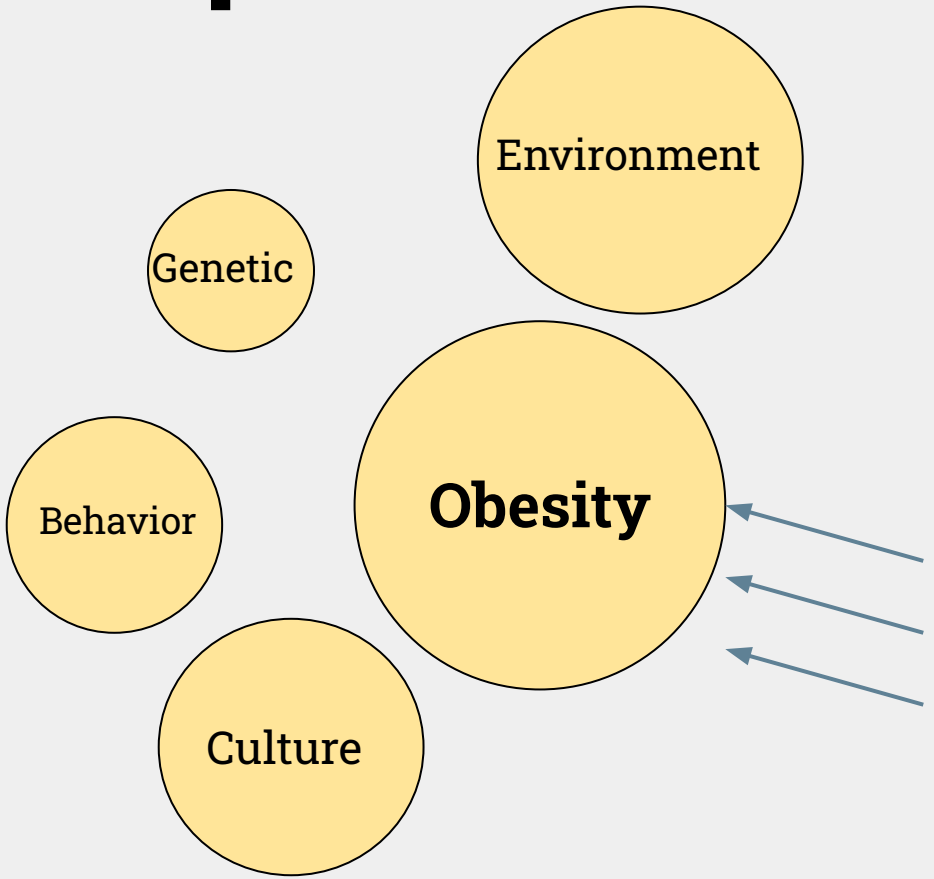
Models	RMSE	R2 score
AdaBoost	6.3971	0.23
Gradient Boosting	6.2353	0.25
Random Forest	6.3660	0.22

- Do not perform well
- Too little data with too complicated models
- Cannot keep loss down

05

Discussion

Surprised...



Complex socioeconomic dynamics

Higher government debt may signal broader economic pressures, which may affect

- Reduce access to healthy food
- Increased pressure on the population
- Changes in health-related policies/funding

Application

Scholars and researchers

Further explore the interplay between economics and public health

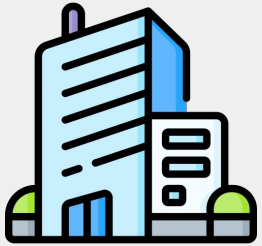
Government

Assess the impact of economic policies on public health

Public health advocates

Promote economic reforms that may also have positive health outcomes

Businesses



Adjust strategies based on the impact of economic factors on obesity

Encourage students like us to investigate problem in daily life

06

Conclusion

Conclusion

- *Rejection of the Null Hypothesis (H0)*
- *Combine features vs delete features*

At least one economic or environmental factor (in this case, government debt) significantly influences adult obesity rates.

References

Liu, Z., Deng, Z., Davis, S. et al. Monitoring global carbon emissions in 2022. Nat Rev Earth Environ 4, 205–206 (2023). https://doi.org/10.1038/s43017-023-00406-z

Hammond, R. A. (2010). The economic impact of obesity in the United States. Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy, 3, 285–295. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3047996/

Flechtner-Mors, M., Thamm, M., Wiegand, S., Reinehr, T., Schwab, K. O., Kiess, W., ... & Holl, R. W. (2015). Overweight and Obesity Based on Four Reference Populations in Pediatric Type 1 Diabetes Subjects. Scientifica, 2015. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4466469/

Tomiyama, A. J. (2019). Stress and Obesity. Annual Review of Psychology, 70, 703–718. https://www.annualreviews.org/doi/10.1146/annurev-psych-010418-102936

Bryant, C. J. (2022). Plant-based animal product alternatives are healthier and more sustainable than animal products. Food and Nutrition Sciences, 3(1). https://www.sciencedirect.com/science/article/pii/S2666833522000612