~~**Subject Code:**~~	~~(ECON1313)~~
Subject Name:	BASIC ECONOMETRICS
Location & Campus:	SGS
Title of Assignment:	GROUP IMPIRICAL PROJECT
Student Name:	**Nguyen Quoc Thuong- s3914321** **Tran Huu Gia Bao- s3891808** **Le Ngoc Lam- s3976513**
Lecturer Name:	Mrs. Ngoc Hoang
Group Number:	SGS02- Team7
Assignment Due Date:	23 May 2024
Date of Submission:	3 May 2024
Number of Page Including this one:	35 pages
Word counted:	**3713 words** (excluding references, tables of content and figures)

Student ID	First name	Parts contributed	Description of task contributed	Contribution %	Signature
S3914321	Thuong	100	100	100	Thuong
S3891808	Bao	100	100	100	Bao
S3976513	Lam	100	100	100	Lam

Contents

ASSIGNED TOPIC: URBANIZATION AND ENVIRONMENT ISSUES

Part 1

1.1 Introduction

Urbanization and environmental issues is critical because the growth of cities has a significant impact on natural systems. This subject is important because it investigates the relationship between rapid urbanization and the environmental consequences, such as air pollution, carbon emissions, and resource depletion. The report will analyze these issues in depth in 266 countries, as outlined in the World Bank (n.d), providing a comprehensive understanding of global patterns and the various consequences of urbanization. The primary goal of this project is to investigate the relationship between urban population density, industrial employment, government spending on education, and electricity consumption.

1.2 Literature review and research questions

The process of urbanization in developing nations has a profound effect on environmental concerns, which are greatly influenced by factors such as the density of the urban population (UPD), carbon emissions (COE), employment in the industrial sector (EII), government expenditure on education (GEE), and urban electricity consumption (EUP).

High urban population density frequently results in elevated carbon emissions as a result of intensified economic activities and energy consumption. Urbanization will lead to significant growth in cities, intensifying environmental degradation due to increased emissions of greenhouse gases, air pollution, and waste production (Rashed 2023). The shift to industrial employment exacerbates these issues, as industrial activities significantly contribute to environmental pollution. Multiple studies indicate that regions with substantial industrial workforces exhibit high levels of carbon emissions, particularly in India. Mumbai, as a major industrial center, has significantly contributed to high carbon emissions and environmental degradation through its manufacturing and construction sectors (Hayat & Khan 2023). The rapid growth of industries in this area has exerted substantial strain on the surrounding environment, resulting in compromised air quality and exacerbating health issues for urban dwellers.

In discussions about urbanization. Allocating resources towards education and school improvement is the most expeditious method to foster awareness, which have the potential to reduce carbon emissions associated with urban living **(**Bera et al. 2023**)**. Urban areas characterized by substantial electricity consumption frequently depend on non-renewable energy sources, resulting in elevated carbon emissions. New York City is a highly populated urban area in the United States that experiences high levels of energy consumption as a result of residential, commercial, and industrial activities. As per the New York Independent System Operator (NYISO), the electricity consumption of the entire state is attributed to 25% of city-wide infrastructure and fossil fuel plants **(**Lesser 2022**)**. A city's electricity consumption makes a substantial contribution to carbon emissions, particularly during peak periods in the summer when there is a high demand for air conditioning.

Research Question

- ***What is the impact of urban population density on carbon emissions in major urban centers and how does government spending on education contribute to their reduction***?

This question seeks to investigate the explicit relationship between the (UPD) and the release of (COE). Urban areas characterized by many people living in a relatively small space tend to produce higher levels of carbon emissions because of intense economic activities and energy usage (Zhang et al. 2022).

- ***What is the relationship between employment in industry and electricity of urban population and how their collective impact on environmental sustainability in urban areas is being investigated?***

This question examines the correlation between (EII) and (EUP), both of which have a substantial impact on environmental deterioration. Industrial operations frequently require significant energy consumption. It is experiencing rapid industrialization and capable of providing sustainable urban and industrial development strategies.

Part 2

i) Briefly Statistic

The data which was selected and researched from 270 countries according to *the World Bank* (n. d) to analyze the level of Urbanization and effective of environmental Issues. Through this research, finding some indicators can determine how two problems could be linked. These variables are extracted from reliable sources in the World Bank (n.d).

The Data set which included 270 countries in the world, the variable is calculated by Urban population density (UPD), linked with COE (Carbon emission); EII (Employment in Industry); GEE (Government expenditure on education); EUP (Electricity of Urban Population).

	UPD	COE	PM2.5	EII	GEE	EUP
Mean	377.02	3.77	89.38	25.13	13.72	93.9
Median	82.11	2.55	100	25.83	15.14261	99.97
SD	1813.49	4.32	25.56	9.011268	4.29	12.81
Min	0.14	0.03	0	5.37	5.01	14.58
Max	20555.71	31.73	100	52.55	34.24	100

Table 1: Descriptive Statistic of Urbanization and Environmental Issues (Reproduced From R)

The summary statistics reveal significant insights into six key variables such as UPD (Urbanization Population Density), COE (Carbon Emission), PM2.5 (Air Pollution), EII (Employment in Industry), GEE (Government Expenditure on Education), and EUP (Electricity of Urban Population). The average population density (UPD) is 377.02, with a substantial variation between 0.14 and 20,555.71 people per square kilometer, indicating extreme density disparities across regions. COE emissions average 3.77 tons per capita, and PM2.5 pollution levels reach 89.38 µg/m3 on average, highlighting significant environmental challenges (Appendix 3). Employment in industry (EII) averages 25.13% of the population but varies widely, suggesting different levels of industrialization and economic structures (Appendix 4). Government expenditure on education (GEE) stands at an average of 13.72% of total spending, indicating variability in educational priorities (Appendix 5). Meanwhile, electricity access among urban populations (EUP) averages 93.90%, with notable differences across regions. Together, these statistics underscore disparities in urbanization, environmental health, economic development, and access to essential services, offering comprehensive data for understanding global urban challenges and informing targeted policies (Meliciani 2006).

ii) Visualize Data

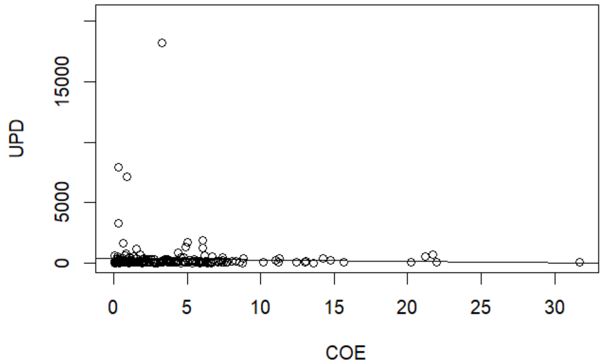


Figure 1: UPD and COE relationship (Copy clipboard from R)

The coloration between urbanization population and Carbon emissions shows that UPD concentrates in the first point.

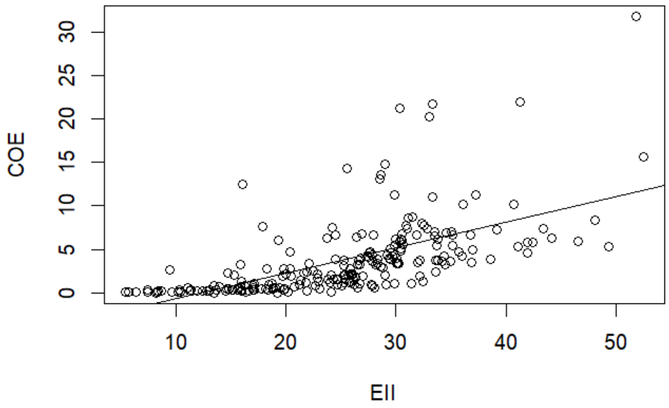


Figure 2: The connection of COE and EII (Clipbroad R)

The picture 2 illustrates that Carbon emissions tend to grow sharply and increasing in EII.

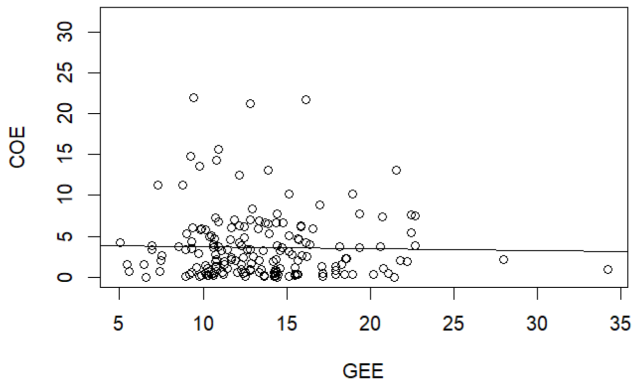


Figure 3: COE (Carbon Oxide Emission) & GEE (Clipbroad R)

Figure 3 which is show the fluctuation between COE and GEE cofficient decreased

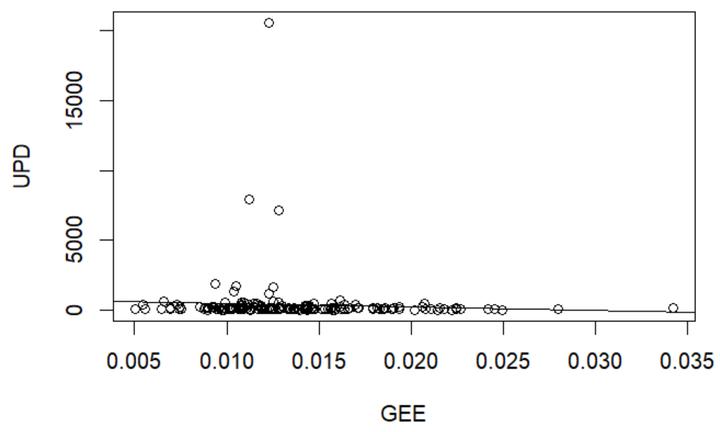


Figure 4: UPD (urbanization population density) & GEE (Government expenditure on edu)

The picture indicating that the significant grow in UPD compared with Government expenditure on education.

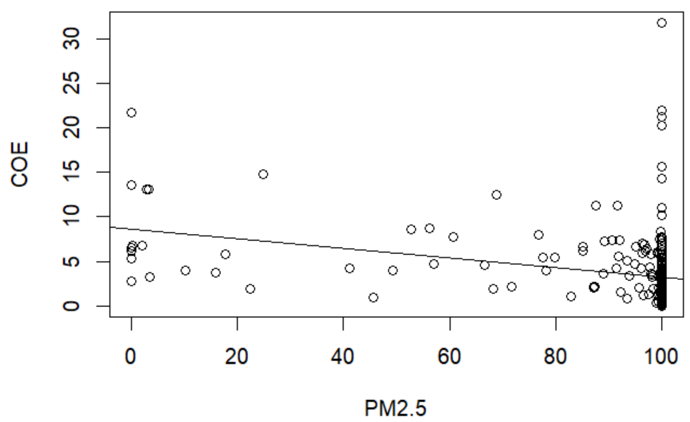


Figure 5: COE and PM2.5 coefficient

The picture exposed an increase in PM2.5 coefficient compared with COE. At this time, they peaked sharply at approximately 100p.

iii) Missing Value

In data available which exposed that 23 countries missing value which is Armenia, Belize, Switzerland, Cuba, Curacao, France, Ghana, Grenada, Guatemala, India, Latvia, Macao, Morocco, Mongolia, Namibia, Paraguay, El Salvador, Eswatini, Syrian Arab Republic, Turkiye, Venezuela, Samoa, Zimbabwe (*R stats)* in Carbon Emission Variable.

Besides that, missing value of Urbanization Population in Kosovo, in Korea which missing GEE (Government expenditure on education) and EUP.

El Salvador only exposed variable of GEE, EUP, it should be deleted from data collection.

Outliners:

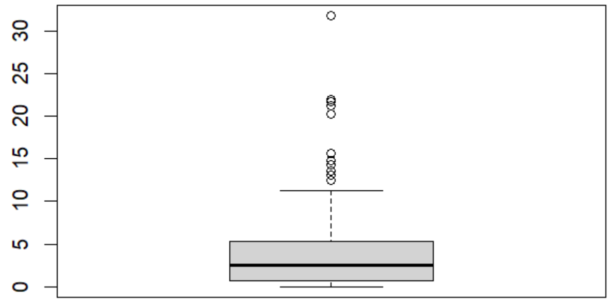


Figure 6: Boxplot of COE (Clipboard from R)

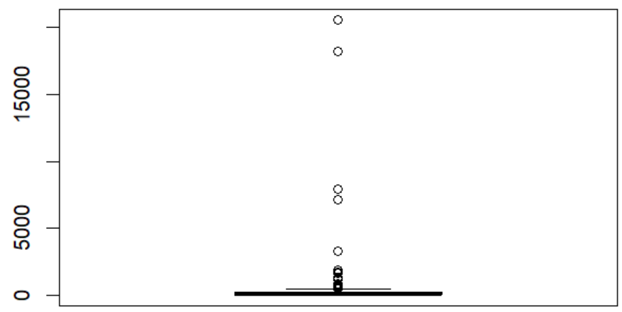


Figure 7: UPD Boxplot

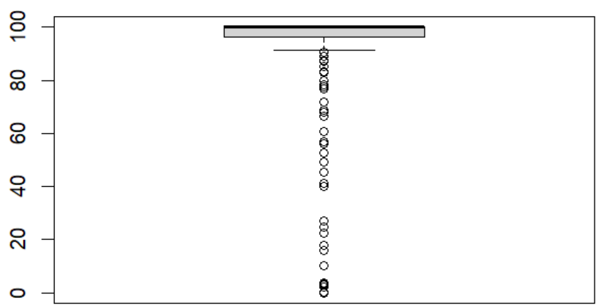


Figure 8: Plot of PM2.5

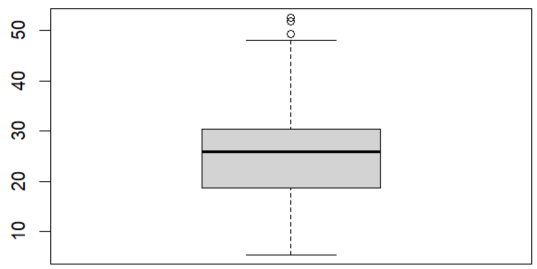


Figure 9: Boxplot of Employment in Industry

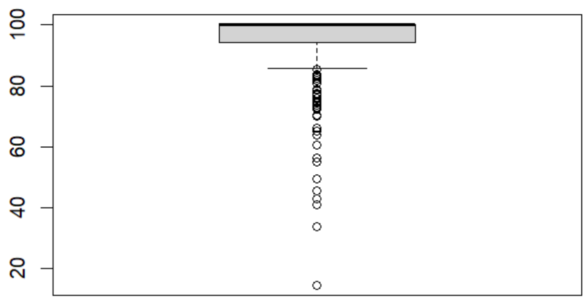


Figure 10: GEE variable of Plot

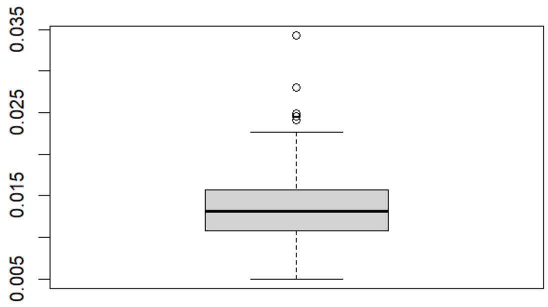


Figure11: Boxplot of EUP (Urbanization population)

According to Boxplot, UPD (Urbanization Population Density) are the most outlier. Besides that, positive effectiveness which exposed in EII, EUP. The Boxplot of EII and EUP did not have extreme value. We can handle this by using the median instead of the mean for central tendency or robust regression methods for modeling. If the outliers do not provide valuable information, they can be removed. However, this should be done cautiously, as removing data can lead to biased analyses.

iv) Transform Variable

A logarithmic transformation can be applied to a highly skewed variable to convert it into a more symmetric data set (Benoit 2011). When one or more variables are converted to their logarithmic forms instead of being kept in their original state, the relationship becomes non-linear (Olivie et al. 2008).

The histogram for UPD, COE, and EII shows a notable rightward skew (Appendix 3,4,5). Taking the logarithm of the variables may lead to a better model fit by transforming the distribution of the features into a bell curve. The logarithm of a negative number is undefined because it's impossible to raise a negative number to the power of a positive base. Similarly, negative numbers cannot be multiplied by a positive base to define their logarithm (Packard 2013).

In this situation, despite the inflation variable being significantly right-skewed, it includes a negative value of -2.0 in the dataset. Therefore, it cannot be transformed using a logarithmic function.

Part 3
3.1 Model Specification

i) Dependence variable

The UPD coefficient, also referred to as the UPD index, is the most widely used measure for assessing inequality in a country's population, income, and spending (Polinesi 2020). When the coefficient is 0, it indicates that all members of the population have equal status (Rotos 2020). Conversely, a coefficient of 100% represents complete population, where all the income is earned by one individual while no one else earns anything (Recchioni 2020).

ii) Independence variable

A list independent variable that affected UPD which is COE, GEE, EII. These variables are selected because they are often incorporated in models of population issues. They represent key environment and social factors that influence the distribution of Urbanization (Zhang et al. 2020).

Urbanization often leads to the growth and expansion of urban areas, changes in land use, and the development of infrastructure to support the increased population (Zhu 2020).

Environmental Quality which exposed high levels of carbon emissions can degrade the air quality in urban areas, making cities less attractive places to live. Poor environmental conditions may deter people from moving to urban areas or encourage them to move to less polluted regions (Maloney & McCormick 2017). Investment in education can drive innovation and technological advancements, often centered in urban areas. This can create dynamic urban environments that attract talent and promote further urban growth (Psacharopoulos 2018).

High carbon emissions may hinder urbanization by affecting environmental quality and health. Increased government expenditure on education can promote urbanization by improving human capital, economic opportunities, and social mobility (Patrinos 2018).

iii) Population model

$$UPD=\beta 0+\beta 1logCOE+\beta 2\ (GEE)+\beta 3log\ (PM2.5)\ +\beta 4log\ (EII)+\beta 5EUP$$

3.2 Estimate and interpretation

i) Model 1

$$UPD=\square 1.17\square 0.062COE\ +0.36log\ (PM2.5)+0.006log\ (EII)\square 0.015log\ (GEE)\ -\ 0.0003log\ (EUP)$$

(n=85; R^2= 0.02095; Adjusted R^2= 0.009457)

Decision ruled =n-k-1=85-5-1=79, α=0.05

Critical t-value =1.955(two-sided alternatives)

In significant level of 10%, the p-value of log (COE) is 0.6325, which is below the threshold of α = 0.1 (Appendix =7). Therefore, we reject the null hypothesis (H0), indicating that there is a statistically significant relationship between government spending and the UPD coefficient, while holding other factors constant, at the 10% significance level (Appendix 7).

At the level of 5%, the p-values for log (GEE) and EII are 2.02e-04 and 0.004, respectively, both of which are less than 0.05. Thus, we reject the null hypothesis (H0), indicating that log (EII) and urban significantly affect the UPD coefficient, after controlling for all other factors, at the 5% significance level.

Interpret coefficients.

The coefficient for (PM2.5) is 0.36, which indicates a positive relationship between openness and the UPD coefficient. With all other variables held constant, an increase of 1% in openness is expected to result in a 0.005 unit rise in the UPD coefficient.

The coefficient for log (GEE) is approximately -0.015, indicating that government expenditure has a negative effect on the UPD coefficient. Thus, a 1% increase in government spending is predicted to reduce the UPD coefficient by 0.0001 units, provided that all other factors are constant (Zhao et al. 2013).

Urban is roughly 0.003, demonstrating a positive correlation between urbanization and the UPD coefficient. It is anticipated that for every unit increase in urbanization, the UPD coefficient will rise by about 0.0003 units, assuming other variables remain constant.

ii) MLR-assumption

COE

H0: $\beta_{\text{COE}} = 0$ No effect Emission on Urbanization

H1: $\beta_{\text{COE}} \neq 0$ Emissions has a statistically significant impact on UPD.

The t-test statistic for the coefficient of Urbanization is t- Coe = -0.209. Since the absolute value |tCOE| = |-0.209| is less than the critical t-value (0.645 < 1.955), we cannot reject the null hypothesis (H0). This indicates that the effect of Urbanization is not significantly different from zero at the 5% significance level, meaning the Urbanization variable does not significantly affect the UPD coefficient after controlling for all other factors.

PM2.5

H0: $\beta_{PM2.5}$ =0 There is no effect PM2.5 on UPD

H1: $\beta_{PM2.5} \neq 0$ PM2.5 Emission effected on Urbanization.

The t-test statistic for the coefficient of PM2.5 is t-PM2.5 = 0.337. Since the absolute value |tPM2.5| = |0.337| is greater than the critical t-value (0.337 > 1.955), we reject the null hypothesis (H0). This implies that, at the 5% significance level, there is a statistically significant relationship between PM2.5 and the UPD coefficient.

EII

H0: β_{EII} =0 Wasn't not affected to UPD

H1: $\beta_{EII} \neq 0$ Employment in Industry effected on Urbanization.

The t-test statistic for the coefficient of log (EII) is t-EII = -0.433. Since the absolute value |tlEII| = |-0.433| is less than the critical t-value (0.433 < 1.955), we cannot reject the null hypothesis (H0). This indicates that the effect of Employment in Industry is not significantly different from zero at the 5% significance level, meaning the log (EII) variable does not significantly affect the Human Development Index (HDI).

GEE

H0: β_{GEE} =0 No result in UPD

H1: $\beta_{GEE} \neq 0$ Significantly effected in GEE.

The p-value for log (GEE) is -1.139, which is lower than 0.1. Therefore, we reject the null hypothesis (H0), indicating that government spending has a statistically significant impact on the UPD coefficient.

iii) Multicollinearity

The variance inflation factors (VIFs) for Urbanization, COE, and GEE, EII indicate that the variables are moderately correlated. However, since all VIF values in the model are less than 5, there is no substantial multicollinearity. Nonetheless, it is important to recognize that VIFs are just one measure of multicollinearity and may not always be a fully reliable indicator.

To address multicollinearity in Model 1, one approach is to remove several key independent variables from the model. Multicollinearity can be mitigated by excluding variables with high VIFs. Depending on the context, it may be necessary to remove factors such as globalization or openness.

iv) Breusch-Pagan test

Using the Breusch-Pagan test, the p-value is 0.8394, which is lower than $\alpha = 0.1$ (Appendix 9). Therefore, we reject the null hypothesis, indicating evidence of heteroskedasticity in the equation at the 10% significance level (Appendix 9 & 11)

Result:

Because of the non-constant variance of residuals, OLS estimates do not retain their minimal variance attribute. The presence of heteroscedasticity results in larger standard errors for coefficients and diminishes their accuracy, reducing the efficiency of OLS estimators (Bun & Harrison 2019). The OLS method does not accommodate heteroscedasticity, which can lead to anomalously low p-values. As a result, OLS might misestimate the actual variance during the computation of t-values and F-values (Rosopa 2013). This could falsely indicate statistical significance for a model term when there is none.

v) Model 2

Removing COE variable from Model 1 so we have Model 2:

$$UPD = \square 1.17 + 0.36\log(PM2.5) + 0.006\log(EII) \square 0.015\log(GEE) - 0.0003\log(EUP)$$

(n=85; R^2= 0.0215; Adjusted R^2= 0.002074)

Model 2 has an adjusted R^2 of 0.002074, which is an improvement over Model 1's adjusted R^2 of 0.009457. This increase suggests that the independent variables in Model 2 collectively account for a greater portion of the variance in the dependent variable, UPD (Appendix 12). This improvement implies that removing the globe variable from Model 2 enhances its explanatory power regarding UPD.

The p-values for EII and urban in Model 2 are -0.665 and -1.193 respectively, as noted in Appendix 13. These values are below the 0.05 threshold at the 5% significance level, leading to the rejection of the null hypothesis (H0). This indicates that both UPD and urban significantly affect UPD at the 5% level, with EII also being significant at the 10% level.

While the coefficient for the urban variable remains the same in both models, there is a notable difference in the coefficients for log (EII) and log (GEE). The coefficient for log (EII) is lower in Model 2 compared to Model 1, whereas the coefficient for log (GEE) is higher in Model 2 than in Model 1.

vi) Binary Variable

Model 3 which a binary variable for African countries in the model allows for assessing the wealth disparity in nations compared to other regions globally. This addition helps to specifically evaluate how wealth distribution differs within countries relative to the rest of the world.

$$UPD = \square 1.35 + 0.357 \log(PM2.5) + 0.0077 \log(EII) \square 0.0215 \log(GEE) — 0.0002 \log(EUP) + Angola$$

(N=85; R^2= 0.01459; Adjusted R^2=-0.003116)

The p-value for the variable representing Angola countries is 0.00582 (Appendix 13) which is below the 0.05 threshold at a 5% significance level. Therefore, we reject the null hypothesis (H0), indicating that being an Angola country has a statistically significant effect on the UPD coefficient when other variables are held constant.

The slope coefficient for the variable representing Angola countries is 0.061. This means that, with all other factors constant, the UPD coefficient for Angola countries is higher by 0.061 units compared to non-countries.

Part 4: Conclusion and policy Implications

4.1 summaries and findings

To analyze the relationship between economic factors and pressing environmental issues based on data from 270 countries collected from the World Bank database. Research factors include UPD (Urban population density), linked with COE (Carbon emission); EII (Employment in Industry); GEE (Government expenditure on education); EUP (Electricity of Urban Population) and PM2.5 (Air Pollution). First, the statistical data show large gaps in urbanization, environmental issues, economic growth, health services, and relative data collection between developing and developed countries. The remaining factors are considered independent factors to impact Urban population density.

Air pollution has a positive relationship with the Urban population density coefficient. A 1% increase in openness is expected to lead to an increase in the Urban population density coefficient of 0.005 units; A 1% increase in government spending is predicted to reduce the Urban population density coefficient by 0.0001 units; For each unit of urbanization increased, the UPD coefficient will increase by about 0.0003 units.

4.2 Policy Recommendations

An effective solution to increase the value of government spending is to shift the value of investment in the education and training system to reduce urban population density; the government needs to invest in increased education in rural and underdeveloped areas; instead of focusing on high population density urban areas. The government needs to invest in expanding and upgrading educational facilities; improve the quality of lecturers and basic training systems; Expand the quality of educational infrastructure in rural and economically underdeveloped areas (Junaid et al. 2023). This provides more suitable educational training and work opportunities for rural people; help these individuals not have to leave their homeland in search of better employment opportunities or educational facilities (Tran 2021). At the same time, the government also needs to promote policies to support special education. These provide enormous opportunities for the local workforce and reduce urban population pressure (Hajebi et al. 2023).

To reduce air pollution, local authorities need to promote the expansion of satellite urban areas and surrounding residential areas. This strategy helps reduce population in concentrated urban areas and reduce pressure on transportation infrastructure and public services in big cities. These satellite residential areas need to be planned and regulated by the government with modern public transport infrastructure; and environmentally friendly such as investing in green landscapes and using resources effectively (Castells et al. 2021). Besides, promoting the use of renewable energy needs to be increased; includes public facilities and residential areas such as in buildings and industrial zones. The government needs to encourage people to increase the application of renewable energy such as solar, wind and other clean energy sources; thereby reducing environmental impact and emissions causing air pollution (Isaifan et al. 2021).

For traveling between regions, green and energy-efficient transportation services can be used. In addition, local governments need to promote green areas around urban areas such as green spaces, parks and trees; thereby serving as a foundation to improve air quality and create a healthier living environment for residents (Castells et al. 2021). By combining these measures, local governments can reduce urban population density; as well as reducing environmental pollution and building a more sustainable living area for everyone (Rezaei and Millard 2023).

4.3 Limitation and Future research

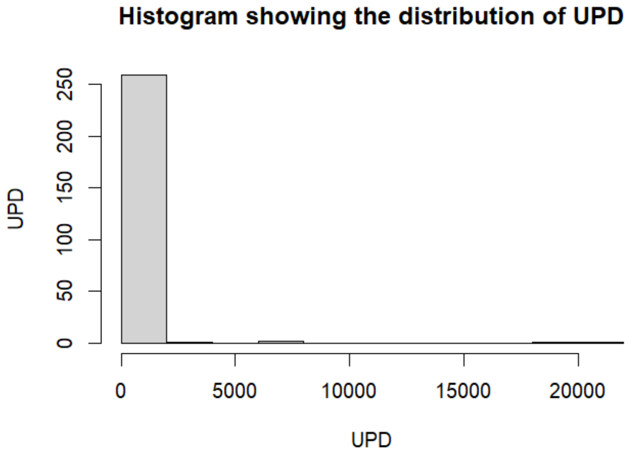
The main limitation is that this report is so broad that the results may not be relevant to a given region or country; Therefore, future research can focus on a certain area to find suitable solutions. Second, this study only focuses on evaluating factors such as Urban population density, Carbon emission, Employment in Industry, Government expenditure on education and Electricity of Urban Population; But in reality there are many other factors. Therefore future research may add some other factors.

Part 5: Guest Speaker

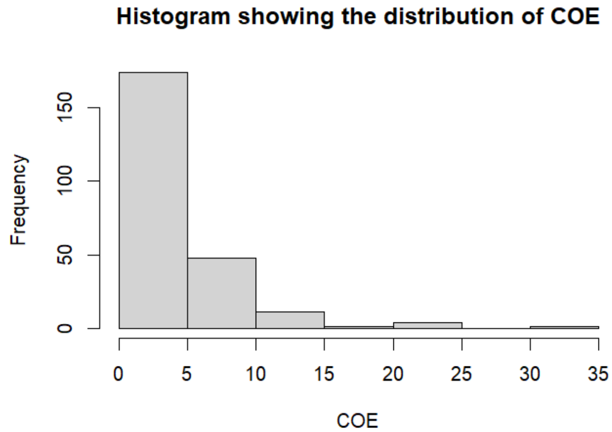
CZ's products and services include providing supply chain solutions for a number of industries such as food, energy and agriculture; In addition, CZ also provides professional consulting services on market analysis and sustainable development. CZ has driven many integrated activities into a number of initiatives that positively impact the environment. CZ has provided certain services such as data analysis in the global supply chain related to the Container river terminal in Vietnam; and focus assessment on improving the efficiency of the goods supply chain. CZ's data research and analytics services drive more efficient terminal operations; such as increasing mining efficiency, reducing carbon emissions or increasing market share.

CZ faces several data analysis challenges such as No learning, scattered and limited data, poor communication and interpretation of insights, Data illiteracy, Data bias and Abundance of data but lacking quality; includes various data and information related to the market, customers or operations management. CZ also needs to screen and evaluate data quality to serve various sales and business tasks. Finally, it involves interaction and data sharing with customers and between departments. CZ can build a centralized integrated data management system; with a comprehensive database stored from many different sources and organized by Data Warehouse; Apply the Extract, Transform, Load process to ensure data consistency. CZ can invest in additional technology infrastructure and artificial intelligence applications such as Cloud Computing or Scalable Storage Solutions. In particular, the focus of CZ group is promoting the value of Data Visualization. CZ group often organizes training courses and participates in workshops on the importance of Data Visualization in business decision making; to promote and analyze the role of Data Visualization in decision making and driving operational efficiency.

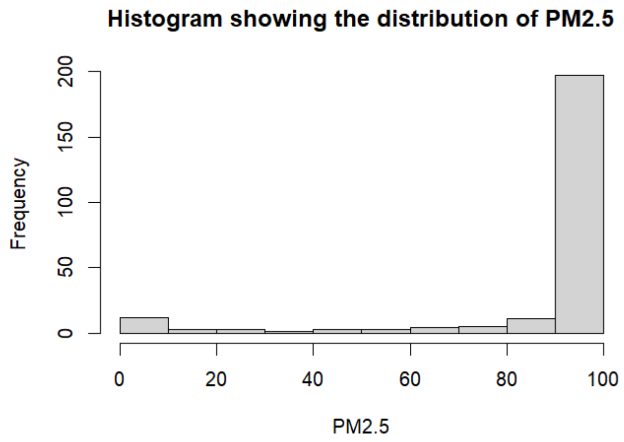
Appendix



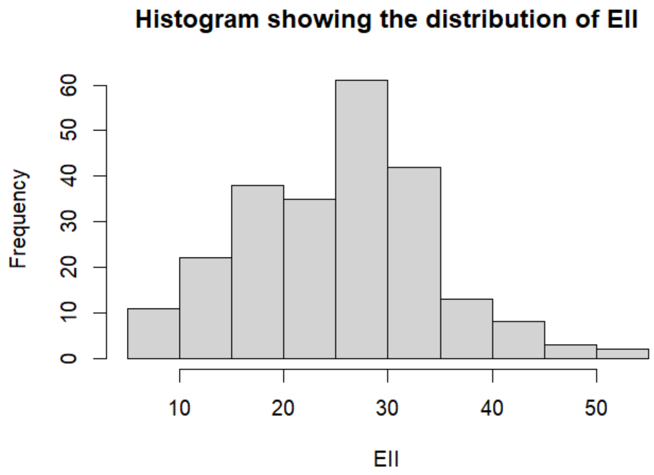
Appendix 1: Histograms of UPD



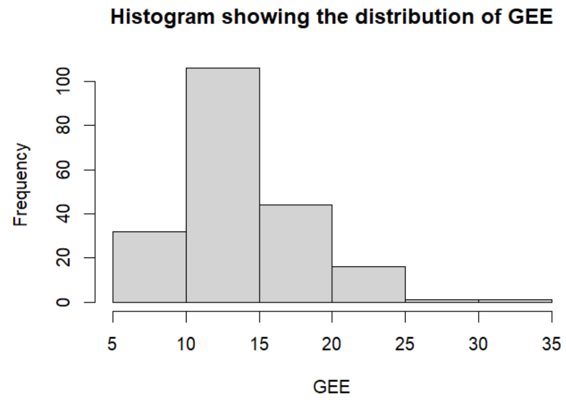
Appendix 2: COE Histograms



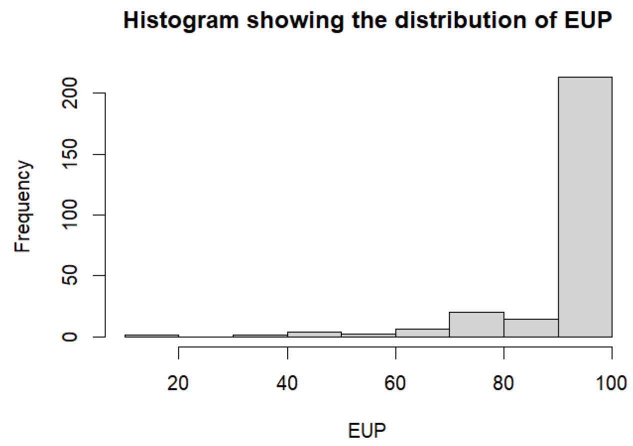
Appendix 3: Histograms of PM2.5



Appendix 4: Show distribution of EII



Appendix 5: GEE Distribution



Appendix 6: EUP Distribution

```
Residuals:
   Min     1Q Median     3Q    Max
-386.0 -240.7 -154.3 -23.2 7533.4

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 15.012    623.439   0.024    0.981
COE         -4.343     20.816  -0.209    0.835
PM2.5        1.009      2.994   0.337    0.737
EII         -3.867      8.926  -0.433    0.665
GEE        -17.794     15.622  -1.139    0.256
EUP          5.363      5.031   1.066    0.288

Residual standard error: 855.6 on 161 degrees of freedom
  (99 observations deleted due to missingness)
Multiple R-squared:  0.02095,	Adjusted R-squared:  -0.009457
F-statistic: 0.689 on 5 and 161 DF,  p-value: 0.6325
```

Appendix 7: The output of Model 1 (linear regression)

```
> vif(Model1)
     COE    PM2.5      EII      GEE      EUP
1.693984 1.195318 1.482271 1.013129 1.019567
```

Appendix 8: The VIF Model Test

```
Bp test(model1)
studentized Breusch-Pagan test
data:  Model1
BP = 2.0698, df = 5, p-value = 0.8394
```

Appendix 9: the Breusch-Pagan test

```
> coeftest (Model1, vcov = vcovHC (Model1, type="HC0"))

t test of coefficients:

            Estimate Std. Error t value Pr(>|t|)
(Intercept) 15.01192 165.79673 0.0905 0.92797
COE          -4.34330   12.60105 -0.3447 0.73079
PM2.5         1.00916    0.77326 1.3051 0.19373
EII          -3.86686    5.24285 -0.7376 0.46186
GEE         -17.79374    7.56297 -2.3527 0.01984 *
EUP           5.36301    2.38621 2.2475 0.02597 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘’ 1
```

Appendix 10: coeftest

	Dependence Variable-----	
	UPD	
	OLS	coefficient test
	(1)	(2)
Constant	15.012	15.012
	(623.439)	(165.797)
COE	-4.343	-4.343
	(20.816)	(12.601)
PM2.5	1.009	1.009
	(2.994)	(0.773)
EII	-3.867	-3.867
	(8.926)	(5.243)
GEE	-17.794	-17.794**
	(15.622)	(7.563)
EUP	5.363	5.363**
	(5.031)	(2.386)
Observations	167	
R2	0.021	
Adjusted R2	-0.009	
Residual Std. Error	855.576 (df = 161)	
F Statistic	0.689 (df = 5; 161)	
Note:	*p<0.1; **p<0.05; ***p<0.01	

Appendix 11: Model 1 before and after using BP test.

```
lm(formula = UPD ~ PM2.5 + EII + GEE + EUP, data = data2)

Residuals:
   Min     1Q Median     3Q    Max 
-382.6 -233.9 -137.1  -26.4 7538.1 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.2255   591.2337   0.000    1.000
PM2.5         1.2951     2.6058   0.497    0.620
EII          -4.8402     7.2820  -0.665    0.507
GEE         -17.8919    14.9944  -1.193    0.234
EUP           5.3409     4.9298   1.083    0.280

Residual standard error: 842.7 on 166 degrees of freedom
  (95 observations deleted due to missingness)
Multiple R-squared:  0.0215,	Adjusted R-squared:  -0.002074 
F-statistic: 0.9121 on 4 and 166 DF,  p-value: 0.4583
```

Appendix 12: Model 2 after removing COE.

```
Residuals:
   Min     1Q Median     3Q    Max 
-455.0 -230.2 -151.8  -28.0 7570.7 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  487.446    384.012   1.269    0.206
PM2.5          1.513      2.599   0.582    0.561
EII           -4.214      7.263  -0.580    0.563
GEE          -19.516     14.927  -1.307    0.193

Residual standard error: 843.2 on 167 degrees of freedom
```

(95 observations deleted due to missingness)
Multiple R-squared: 0.01459, Adjusted R-squared: -0.003116
F-statistic: 0.824 on 3 and 167 DF, p-value: 0.4824

Appendix 13: Model 3 after adding Angola.

References

The World Bank (n.d), 'PM2.5 air pollution, population exposed to levels exceeding WHO guideline value (% of total)', WorldBank website, accessed 17 May 2024. https://data.worldbank.org/indicator/EN.ATM.PM25.MC.ZS?view=chart

The World Bank (n.d),' Population density (people per sq. km of land area)', The World Bank website, accessed 17 May 2024. https://data.worldbank.org/indicator/EN.POP.DNST?view=chart

The World Bank(n.d),'CO2 emissions (metric tons per capita)', World Bank website, accessed in 17 May 2024. https://data.worldbank.org/indicator/EN.ATM.CO2E.PC?view=chart

The World Bank (n.d.),'Employment in industry, male (% of male employment) (modeled ILO estimate), World Bank website, accessed in 17 MAy 2024. https://data.worldbank.org/indicator/SL.IND.EMPL.MA.ZS?view=chart

The World Bank (n.d),' Government expenditure on education, total (% of government expenditure)', World Bank website, accessed in 17 May 2024, https://data.worldbank.org/indicator/SE.XPD.TOTL.GB.ZS?view=chart

The World Bank (n.d.),'Access to electricity, urban (% of urban population)', World Bank website, accessed in 17 May 2024. https://data.worldbank.org/indicator/EG.ELC.ACCS.UR.ZS

Benoit, K., (2011). 'Linear regression models with logarithmic transformations'. London School of Economics, London, 22(1):23-36.

Olivier, J., Johnson, W.D. and Marshall, G.D., (2008). 'The logarithmic transformation and the geometric mean in reporting experimental IgE results: what are they and when and why to use them?'. Annals of Allergy, Asthma & Immunology, 100(4):333-337.

Packard, G.C., (2013). 'Is logarithmic transformation necessary in allometry?'. Biological Journal of the Linnean Society, 109(2):476-486.

Rosopa, P.J., Schaffer, M.M. and Schroeder, A.N., (2013). 'Managing heteroscedasticity in general linear models'. Psychological methods, 18(3):335.

Polinesi, G., Recchioni, M.C., Turco, R., Salvati, L., Rontos, K., Rodrigo-Comino, J. and Benassi, F., (2020). 'Population trends and urbanization: Simulating density effects using a local regression approach'. ISPRS International Journal of Geo-Information, 9(7):454.

Zhang, Q., Wu, Z., Yu, H., Zhu, X. and Shen, Z., (2020). 'Variable urbanization warming effects across metropolitans of China and relevant driving factors'. Remote Sensing, 12(9):1500.

Maloney, M.T. and McCormick, R.E., (2017). 'A positive theory of environmental quality regulation'. In Distributional Effects of Environmental and Energy Policy ,185-209

Psacharopoulos, G. and Patrinos, H.A., (2018). 'Returns to investment in education: a decennial review of the global literature'. Education Economics, 26(5):445-458. doi.org/10.1080/09645292.2018.1484426

Zhao, J., Chen, S., Jiang, B., Ren, Y., Wang, H., Vause, J. and Yu, H., (2013). 'Temporal trend of green space coverage in China and its relationship with urbanization over the last two decades'. Science of the Total Environment, 442:455-465.

Bun, M.J. and Harrison, T.D., (2019). 'OLS and IV estimation of regression models including endogenous interaction terms'. Econometric Reviews, 38(7):814-827.

Dwesini, N.F., (2019). 'Causes and prevention of high employee turnover within the hospitality industry: A literature review'. African Journal of Hospitality, Tourism and Leisure, 8(3):1-15.

Mitton, T. and Vorkink, K., (2007). 'Equilibrium underdiversification and the preference for skewness'. The Review of Financial Studies, 20(4), pp.1255-1288. doi.org/10.1093/revfin/hhm011

Stanners, D., Dom, A., Gee, D., Martin, J., Ribeiro, T., Rickard, L. and Weber, J.L., (2007). Frameworks for policy integration indicators, for sustainable development, and for evaluating complex scientific evidence. Sustainability Indicators, 145-162.

Sarstedt, M., Mooi, E., Sarstedt, M. and Mooi, E., (2019). 'Regression analysis. A concise guide to market research': The process, data, and methods using IBM SPSS Statistics, 209-256.

The World Bank (n.d) World Development Indicators, The World Bank website, accessed on 21 May 2024 https://databank.worldbank.org/reports.aspx?source=2&series=EN.ATM.PM25.MC.ZS&country=

Rashed A. (2023) The Impacts of Unsustainable Urbanization on the Environment, Sustainable Regional Planning. IntechOpen, doi:10.5772/intechopen.110089

The Impacts of Unsustainable Urbanization on the Environment | IntechOpen

Hayat K. and Khan I. (2023) The effect of technological innovations, urbanization and economic growth on environmental quality: does governance matter? Frontiers in Environmental Science, 11, doi: 10.3389/fenvs.2023.1239288

Bera B. et al. (2023). Urbanization and Its Impact on Environmental Sustainability: A Comprehensive Review. Harbin Gongcheng Daxue Xuebao/Journal of Harbin Engineering University, 44: 1310-1318

Lesser J. (2022) Energy Poverty in New York: The Adverse Impacts of the State's Green Power Mandates, The Manhattan Institute website, accessed on 20 May 2024 Energy Poverty in New York: The Adverse Impacts of the State's Green Power Mandates | Manhattan Institute (manhattan-institute.org)

Zhang Y. et al. (2022) Spatiotemporal Differentiation and Driving Mechanism of Coupling Coordination between New-Type Urbanization and Ecological Environment in China, Urbanization and Environmental Sustainability, 14(18): 11780, doi: 10.3390/su141811780

Castells D, Dienesch E and Krause M (2021) 'Air pollution in an urban world: A global view on density, cities and emissions', *Ecological Economics*, 189. 107153. doi: 10.1016/j.ecolecon.2021.107153.

Hajebi E, Billing C and Hajebi M (2023) 'The Effect of Government Expenditure on Education on the Enrollment Rate of Different Educational Levels in Selected OECD Countries', *International Journal of Scientific Research and Management*, 11. 2783-2795. doi: 10.18535/ijsrm/v11i05.el03.

Isaifan R, Alsalama T and Koç M (2021) 'Mitigation of urban air pollution with green vegetation for sustainable cities: a review', *International Journal of Global Warming*, 25. 498. doi: 10.1504/IJGW.2021.10042689.

Junaid A, Iqbal A, Siddiqui A and Hassan M (2023) 'Government Spending in Education Division, Urbanization and Economic Escalation in Under-Developed and Developing Countries', *Voyage Journal of Educational Studies*, 3. 207-219. doi: 10.58622/vjes.v3i1.37.

Rezaei N and Millard A (2023) 'Urban form and its impacts on air pollution and access to green space: A global analysis of 462 cities', *PLOS ONE*, 18. e0278265. doi: 10.1371/journal.pone.0278265.

Tran T (2021) 'Government Education Expenditure and Economic Growth Nexus: Empirical Evidence from Vietnam', 8. 413-421. doi: 10.13106/JAFEB.2021.VOL8.NO7.0413.