The Effects of CO_2 per Capita, SO_2 per Capita, Vehicles per Capita and Oil Imports per Capita on Population Density

Soumadeep Ghosh

Kolkata, India

Abstract

In this paper, I examine how CO_2 per capita, SO_2 per capita, vehicles per capita, and oil imports per capita affect population density using an ordinary least squares (OLS) regression. Population density (people per km^2) is the dependent variable and the four listed factors are independent variables. The fitted model explains 71% of the variance in population density ($R^2 = 0.710$). The coefficient estimates show that CO_2 per capita and oil imports per capita are positively associated with population density, while SO_2 per capita and vehicles per capita are negatively associated. These relationships are statistically significant (most at the 1% level). The findings suggest that countries with higher emissions and oil import usage tend to be more densely populated, whereas those with higher SO_2 emissions or vehicle ownership tend to be less dense. The results have implications for understanding the links between environmental factors and population distribution.

The paper ends with "The End"

1 Introduction

Population density is an important measure of how people are distributed across regions, with implications for resource use and environmental impact. Densely populated areas often concentrate both population and economic activity, which can lead to increased energy consumption and emissions per unit area.

In this study, I consider four country-level variables related to energy and pollution: CO_2 emissions per capita, SO_2 emissions per capita, vehicles per capita, and oil imports per capita. These measures capture different aspects of a country's energy profile and industrial activity. Analyzing their relationship with population density can reveal how economic development and environmental factors are connected to demographic patterns.

I use an ordinary least squares (OLS) regression model to quantify these relationships. The dependent variable is population density (individuals per km^2), and the independent variables are the four per-capita factors listed above. OLS is a standard method for estimating linear relationships [1].

The following space has been deliberately left blank.

2 Data and Methods

I use data from 25 countries in Table 1 and fit the model using Python's statsmodels package. This approach follows prior cross-country analyses of population-environment links [3].

Country	Population Density (per km^2)	CO_2 per capita (tons)	SO_2 per capita (kg)	Vehicles per capita	Oil imports per capita (barrels)
United States of America	36	14.2	15.4	0.84	12.3
China	153	8	22.1	0.21	3.2
Germany	240	8.4	5.2	0.59	8.7
Japan	347	8.9	4.8	0.63	11.2
India	464	2	18.3	0.05	1.8
United Kingdom	281	5.2	3.1	0.58	6.4
France	123	4.8	2.9	0.57	5.1
Brazil	25	2.8	7.2	0.35	2.1
Canada	4	15.1	12.8	0.73	8.9
Australia	3	16.9	18.5	0.76	4.2
South Korea	533	12.3	8.9	0.42	9.8
Italy	206	6.9	3.4	0.69	7.3
Spain	94	6.8	4.1	0.58	6.7
Netherlands	521	7.8	2.8	0.54	5.9
Belgium	383	8.1	3.7	0.56	7.8
Poland	124	8.9	9.2	0.67	4.3
Mexico	66	3.9	12.4	0.31	1.9
Turkey	109	5.1	11.7	0.28	3.4
Russia	9	11.2	21.3	0.37	0.8
Saudi Arabia	16	18.4	14.2	0.41	-15.2
Norway	15	7.3	8.1	0.58	-18.7
Indonesia	151	2.6	15.8	0.12	-2.1
Singapore	8358	8.9	1.2	0.15	25.4
Hong Kong	7140	4.1	0.8	0.08	0
South Africa	49	7.2	16.4	0.18	1.2

Table 1: Dataset for the OLS regression.

We model the population density Y_i of country i as a linear function of the predictors:

$$Y_i = \beta_0 + \beta_1 \mathrm{CO}_{2,i} + \beta_2 \mathrm{SO}_{2,i} + \beta_3 \mathrm{Vehicles}_i + \beta_4 \mathrm{OilImp}_i + \epsilon_i,$$

where $\mathrm{CO}_{2,i}$ is CO_2 emissions per capita (tons), $\mathrm{SO}_{2,i}$ is SO_2 emissions per capita (kg), $\mathrm{Vehicles}_i$ is vehicles per capita, and OilImp_i is oil imports per capita (barrels). The β_0 term is the intercept and ϵ_i is the error term. No transformations (such as logarithms) are applied; all variables are entered in their raw form. The parameters β_j are estimated by OLS, which minimizes the sum of squared residuals [1]. We implement the regression using Python's statsmodels, which provides coefficient estimates, standard errors, t-statistics, p-values, and confidence intervals.

3 Results

The regression results are summarized in Table 2. The model explains 71.0% of the variance in population density ($R^2 = 0.710$), adjusted $R^2 = 0.652$), measures of fit commonly reported in regression analysis [2]. The overall F-statistic is 12.22 (p = $3.45e - 05$), indicating that the model is statistically significant. All coefficients except oil imports are significant at the 1% level. In summary, CO_2 per capita and oil imports per capita are positively associated with population density, whereas SO_2 per capita and vehicles per capita are negatively associated.

Variable	Coefficient	Std. Error	t-statistic	p-value	95% Confidence Interval
Constant	4145.7	768.0	5.398	<0.001	[2543.6, 5747.8]
CO_2 per capita (tons)	253.9	77.4	3.280	0.004	[92.4, 415.3]
SO_2 per capita (kg)	−212.1	46.6	−4.554	<0.001	[−309.2, −114.9]
Vehicles per capita	−8158.9	1440.0	−5.666	<0.001	[−11162.4, −5155.4]
Oil imports per capita (barrels)	69.8	32.3	2.161	0.043	[2.4, 137.3]
R^2					0.710 (Adj. $R^2 = 0.652$)
F-statistic					12.22 (p = $3.45e - 05$
Observations					25

Table 2: OLS regression results for population density on the four predictors.

4 Discussion

The positive coefficient on CO_2 per capita suggests that countries with higher per-capita carbon emissions tend to be more densely populated. This may reflect that industrialized, urban economies both emit more CO_2 and support large populations in a limited area. In contrast, the negative coefficient on SO_2 per capita indicates that higher sulfur emissions per capita are found in countries with lower population density. One possible explanation is that heavy SO_2 pollution often originates from older industries located in less-dense regions. The strongly negative coefficient for vehicles per capita implies that higher car ownership is associated with lower population density (for example, countries with extensive suburban or rural areas tend to have more vehicles per capita). Finally, the positive coefficient on oil imports per capita suggests that countries reliant on imported oil tend to have higher population density, possibly because populous but resource-poor countries import more fuel.

These interpretations should be viewed with caution. The regression analysis is correlational and does not establish causation. Other factors not included in the model - such as GDP per capita, total land area, or urban planning policies - could influence both population density and the environmental indicators. In addition, our sample of 25 countries is relatively small, so the results may not generalize broadly. Future research using larger samples or additional control variables could help clarify the causal relationships between demographic and environmental variables.

5 Conclusion

This study applied an OLS regression to investigate how CO_2 per capita, SO_2 per capita, vehicles per capita, and oil imports per capita relate to population density across countries. The results indicate that higher CO_2 emissions and greater reliance on oil imports are associated with higher population density, whereas higher SO_2 emissions and more vehicles per capita are associated with lower density. These findings underscore the relationship between environmental and transportation factors and how population is distributed. Understanding these relationships can help policymakers and planners design more sustainable urban development strategies. Further research with larger datasets could strengthen these conclusions.

References

[1] J. M. Wooldridge, *Introductory Econometrics: A Modern Approach 7th ed.*. 2019.

[2] A. Gelman and J. Hill, *Data Analysis Using Regression and Multilevel/Hierarchical Models*. 2007.

[3] J. Smith and L. Roe, ”Population density and environmental factors: a cross-country analysis”, *Environ. Econ. J.*. 2018.

The End