Factors Influencing the Prevalence of Asthma in United States Counties

By: Sadhika Dhanasekar, Aditi Kodali, Sumaja Bandreddi

Introduction

- Asthma is a respiratory condition that causes difficulty breathing
- With the rise of global warming and depletion of the Ozone layer it is important to study the effects air pollution can have on asthma as well as smoking, state region and other factors
- We want to examine which factors most affect the prevalence of asthma in the U.S

Research Questions

1. Do counties with low air quality have a higher asthma population?
2. Do counties with a high smoking rate have a higher asthma population?
3. Do counties with a high proportion of disabilities have a higher asthma population?

Data Summary

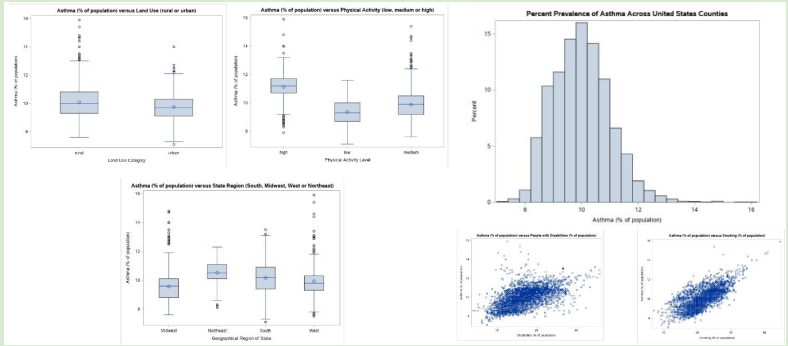
Response Variable:

Variable	Category	Measurement
Asthma	Quantitative (% of population)	Age-adjusted prevalence of current asthma among adults >= 18 (2018)

Explanatory Variables:

Variable	Category	Measurement
Air Quality	Quantitative ($\mu g/m^3$)	PM2.5: Annual average concentration (monitor + modeled data) (2018)
State Region	Qualitative (Northeast, West, Midwest, South)	Census Gov Data classifying State Regions
Smoking	Quantitative (% of population)	Age-adjusted prevalence of current smoking among adults >= 18 (2018)
Land Use	Qualitative (Rural or Urban)	Land Use - Classification of County from Rural to Urban (2013)
Physical Activity	Qualitative (Low, Medium or High)	Age-adjusted prevalence of No Leisure-time Physical activity among adults >= 18 (2018)
People Without Health Insurance	Quantitative (% of population)	Percent of Population without health insurance (2018)
Public Transportation	Quantitative (% of workers)	Percent of Workers 16 and Older who Take Public Transportation to Work (2016-2020)
Disabilities	Quantitative (% of Population)	Percent of Population 5 years or older with a Disability

EDA



Analysis

Stage 1: Quantitative Variables

Initial - asthma = B0 + B1disabilities + B2publictransport + B3smoking + B4airquality + B5healthinsurance

Revised - asthma = B0 + B1disabilities + B2publictransport + B3smoking

Stage 2: Qualitative Variables

Initial - We tested the interactions between State Region and Rural/Urban and Physical activity and Rural/Urban. After refitting the model with state region, the disabilities variable was no longer significant.

Revised - asthma = B0 + B1publictransport + B2smoking + B3DummyWest + B4DummyMidwest + B5DummyNortheast;

Stage 3: Quantitative x Qualitative Interactions

Initial - We tested the interactions between Public Transport and Rural/Urban

Final- We concluded that this interaction is insignificant and kept the same end model as Stage 2

Multicollinearity

- Avg VIF < 3; no individual VIF > 10

Final Model

asthma = 5.64849 + 0.20414*Smoking - 0.51143*DummyMidwest + 0.74773*DummyNortheast + 0.97693*DummyWest

Weight: REWEIGHT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	6.87102	1.71776	49.94	<.0001
Error	93	3.19873	0.03439		
Corrected Total	97	10.06975			

Root MSE	0.18546	R-Square	0.6823
Dependent Mean	9.96159	Adj R-Sq	0.6687
Coeff Var	1.86174		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	5.64849	0.38498	14.67	<.0001
smoking2	1	0.20414	0.01656	12.33	<.0001
dummyMidwest	1	-0.51143	0.14545	-3.52	0.0007
dummyNortheast	1	0.74773	0.22305	3.35	0.0012
dummyWest	1	0.97693	0.19834	4.93	<.0001

Additional Techniques

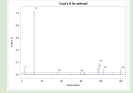
Variable Screening

- Stepwise Regression - SLE = 0.5 SLS = 0.1
- Smoking, Disabilities, and Public Transport were significant

Summary of Stepwise Selection

Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	smoking2		1	0.4599	0.4599	9.4322	83.45	<.0001
2	disabilities2		2	0.0240	0.4839	6.7378	4.52	0.0360
3	publictransport2		3	0.0219	0.5059	4.4587	4.26	0.0417

Removing Influential Outliers



Removed observations 12 and 83 Public Transport was removed

Weighted Squares Regression

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	5.24978	0.38483	13.64	<.0001
publictrans2	1	0.05768	0.01944	2.97	0.0038
smoking2	1	0.21624	0.01646	13.31	<.0001
dummyMidwest	1	-0.45182	0.14649	-3.08	0.0027
dummyNortheast	1	0.81534	0.22480	3.63	0.0005
dummyWest	1	1.08803	0.19753	5.51	<.0001

Conclusion

Since our final p-value (<0.001) is less than our significance level we can conclude that smoking and state region are influential factors in the prevalence of asthma. In the U.S. Although the global F test indicates that all the overall model is useful, the correlation does not necessarily equate causation.

Efficacy: According to our adjusted R-Square the model will predict 66.67% of the variation in the data

Further Research - We want to test more variables such as drinking water and conduct more advanced methods of regression such as Lasso regression. This can potentially be used to distribute medical supplies given the demographics of a county.

References

- EPA, Environmental Protection Agency, https://www.epa.gov/air-trends/.
- "Data Access - Urban Rural Classification Scheme for Counties." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 1 June 2017, https://www.cdc.gov/nchs/data_access/urban_rural.htm.
- "Air Pollution." Asthma & Allergy Foundation of America, 11 Nov. 2022, https://www.aafa.org/air-pollution-smog-asthma/.

Tiotiu, Angelica I, et al. "Impact of Air Pollution on Asthma Outcomes." International Journal of Environmental Research and Public Health, U.S. National Library of Medicine, 27 Aug. 2020, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7503605/.

- "Asthma and Secondhand Smoke." Centers for -Disease Control and Prevention, Centers for Disease Control and Prevention, 5 May 2022, https://www.cdc.gov/tobacco/campaign/tips/diseases/secondhand-smoke-asthma.html.
- Ehrlich, Samantha F., et al. "Patients Diagnosed with Diabetes Are at Increased Risk for Asthma, Chronic Obstructive Pulmonary Disease, Pulmonary Fibrosis, and Pneumonia but Not Lung Cancer." American -Diabetes Association, American Diabetes Association, 6 Oct. 2009,