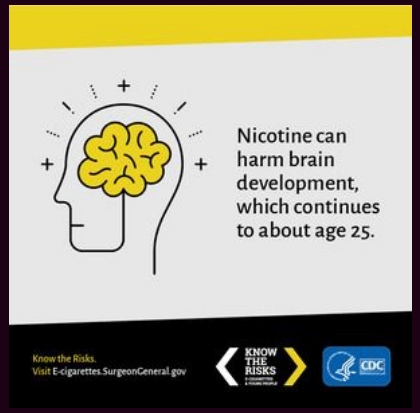
FINAL PRESENTATION:

SMOKING PREVALENCE AND SOCIO-DEMOGRAPHICS OF U.S. COUNTIES IN 2022

Daniel Woodford
Carson Chapman
Anna Kenney-Hynes

INTRODUCTION



RESEARCH QUESTION AND PUZZLE

Is there a significant relationship between county-level socio-demographic factors such as

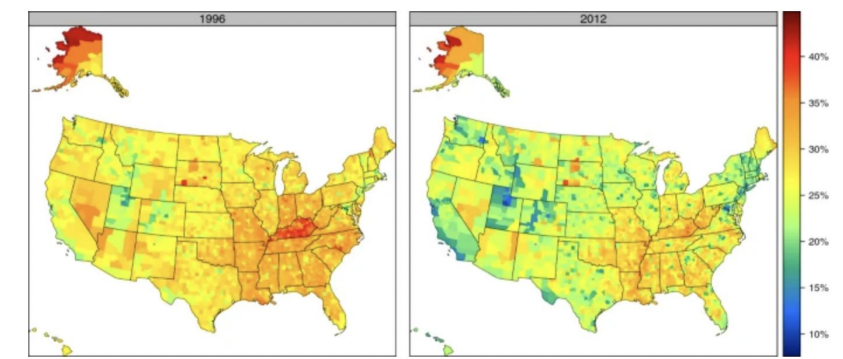
1. median annual household income,
2. educational attainment, and the
3. severity of the COVID-19 virus in 2020,

and county-level smoking prevalence among adults in the United States in 2022?

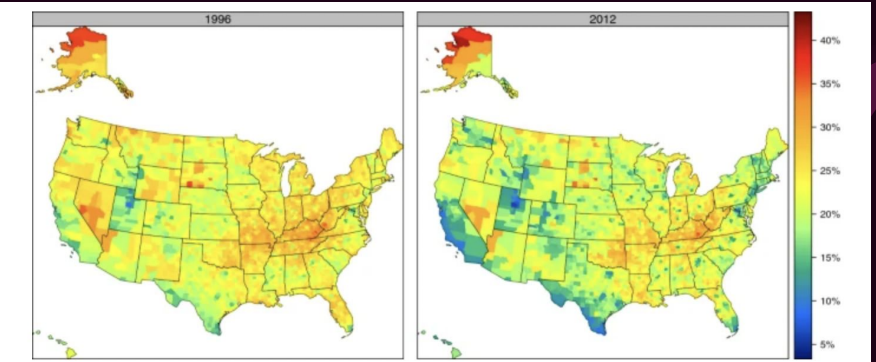
BACKGROUND:

Cigarette smoking prevalence in US counties: 1996-2012 (2014)

Dwyer-Lindgren, L., Mokdad, A. H., Srebotnjak, T., Flaxman, A. D., Hansen, G. M., & Murray, C. J.



Age-standardized total cigarette smoking prevalence, males, 1996 and 2012.



Age-standardized total cigarette smoking prevalence, females, 1996 and 2012.

Highest rates of total cigarette smoking : counties in the South and counties with large Native American populations.

Lowest rates of total cigarette smoking: counties in Utah and other Western states

BACKGROUND:

Cigarette smoking prevalence in US counties: 1996-2012 (2014)

Dwyer-Lindgren, L., Mokdad, A. H., Srebotnjak, T., Flaxman, A. D., Hansen, G. M., & Murray, C. J.

Total cigarette smoking prevalence **varies dramatically** between counties even within states.

Income:

Between 1996 and 2012, **total smoking prevalence** in the United States **decreased**, *however*, **statistically significant declines were concentrated** within a small number of counties and counties in the top quintile in terms of income experienced faster declines in smoking prevalence compared to counties in the bottom quintile.

BACKGROUND:

Momentary smoking context as a mediator of the relationship between SES and smoking (2018)

Jahnel, T., Ferguson, S. G., Shiffman, S., Thrul, J., & Schüz, B.

- SES was operationalized via education

Education:

- lower SES participants were more likely to encounter places where smoking was allowed compared to higher SES participants. **Participants who encountered more smoking-permissive environments smoked more cigarettes per day.**
- There was a significant indirect effect of SES on cigarettes per day via the momentary context, indicating that SES was indirectly associated with smoking behavior through differential exposure to smoking-friendly environments.

BACKGROUND:

Socioeconomic Differences in Cigarette Smoking Among Sociodemographic Groups (2019)

Garrett, B. E., Martell, B. N., Caraballo, R. S., & King, B. A.

Table 1. Current Cigarette Use Among US Adults Aged 25 or Older, by Sociodemographic Characteristic and Education Level, National Survey on Drug Use and Health, 2011–2014

Characteristic	Less Than High School Diploma [Reference], % (95% CI)	High School Diploma, % (95% CI) [*P* Value]	Some College (No Degree), % (95% CI) [*P* Value]	College Graduate, % (95% CI) [*P* Value]
Overall[b]	31.6 (30.6–32.7)	27.5 (26.8–28.2) [<.001]	25.1 (24.4–25.8) [<.001]	10.8 (1.4–11.3) [<.001]
US Census region[c]				
Northeast	31.2 (28.6–34.0)	27.0 (25.4–28.6) [.008]	25.4 (23.8–27.2) [<.001]	11.1 (10.1–12.1) [<.001]
Midwest	37.4 (35.2–39.8)	30.6 (29.3–31.9) [<.001]	27.7 (26.4–29.0) [<.001]	11.6 (10.8–12.4) [<.001]
South	33.4 (31.7–35.1)	28.0 (26.8–29.2) [<.001]	25.9 (24.7–27.1) [<.001]	11.7 (10.9–12.5) [<.001]
West	24.5 (22.6–26.6)	23.2 (21.7–24.8) [.29]	21.3 (19.8–22.8) [.009]	8.9 (8.0–9.9) [<.001]

[a] Source: Substance Abuse and Mental Health Services Administration, Center or Behavioral Health Statistics and Quality, National Survey on Drug Use and Health; surveys for 2011–2014 (8–11).
[b] Overall row includes data on respondents who reported being of more than one racial/ethnic group although these data were excluded from numbers in race/ethnicity categories.
[c] Northeast: Connecticut, Maine, Massachusetts, New Jersey, New Hampshire, New York, Pennsylvania, Rhode Island, and Vermont; Midwest: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; South: Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia; West: Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

BACKGROUND:

Socioeconomic Differences in Cigarette Smoking Among Sociodemographic Groups (2019)

Garrett, B. E., Martell, B. N., Caraballo, R. S., & King, B. A.

Income:

- According to Garrett et al., (2019) men and women who lived **at or above the poverty line had a significantly lower smoking prevalence** than those who lived below the federal poverty level.

Education:

- Also suggests **a significant inverse relationship between smoking status and highest education level attained**, finding a higher smoking prevalence among populations with lower educational attainment.

BACKGROUND:

The impact of COVID-19 pandemic on tobacco use: A population-based study (2023)

Alomari, M. A., Khabour, O. F., Alzoubi, K. H., & Maikano, A. B.

COVID-19:

- Changes in smoking habits during the COVID-19 pandemic were reported by **about 50%** of survey participants of adults in Jordan who smoke tobacco, with **more participants reporting a decrease than increase in use.**

BACKGROUND:

Impact of COVID-19 lockdown on smoking consumption in a large representative sample of Italian adults (2021)

Carreras, G., Lugo, A., Stival, C., Amerio, A., Odone, A., Pacifici, R., Gallus, S., & Gorini, G.

COVID-19:

- found that the lockdown **increased cigarette consumption by 9.1%** and emphasized the association of increased cigarette consumption with increased mental distress among the study's participants.

LITERATURE ON INFLUENCE OF INCOME LEVEL:

- According to Dwyer-Lindgren et al., (2014), between 1996 and 2012, **total smoking prevalence** in the United States **decreased**, *however*, **statistically significant declines were concentrated** within a small number of counties and counties in the top quintile in terms of income experienced faster declines in smoking prevalence compared to counties in the bottom quintile.
- According to Garrett et al., (2019) men and women who lived **at or above the poverty line had a significantly lower smoking prevalence** than those who lived below the federal poverty level.

THEORY:

Counties with higher median incomes:

- better access to smoking cessation programs and support
- better access to more diverse, recreational and stress-relief opportunities

Counties with lower median incomes:

- individuals may perceive smoking as a more affordable stress-relief option
- individuals may be more susceptible to tobacco companies' marketing and pricing strategies

Hypothesis: Counties with lower median annual household income in 2022 will have a higher smoking prevalence.

LITERATURE ON INFLUENCE OF EDUCATION LEVEL:

- According to Jahnel et al. (2018), lower **SES participants were more likely to encounter places where smoking was allowed compared to higher SES participants**. Participants who encountered more smoking-permissive environments smoked more cigarettes per day.
- Garrett et al., (2019) suggests a significant inverse relationship between smoking status and highest education level attained, finding a **higher smoking prevalence among populations with lower educational attainment.**

THEORY:

Counties with higher educational attainment level::

- more informed about the health risks associated with smoking
- likely to have better work and economic conditions

Counties with lower educational attainment level:

- Individuals may be more likely to encounter places where smoking is allowed, tolerated or even encouraged

Hypothesis: Counties with populations with greater educational attainment levels in 2022 will have a lower smoking prevalence.

LITERATURE ON INFLUENCE OF COVID-19 PANDEMIC:

- According to Alomari et al. (2023), changes in smoking habits during the COVID-19 pandemic were reported by about 50% of survey participants of adults in Jordan who smoke tobacco, **with more participants reporting a decrease than increase in use.**
- Carreras et al. (2021) found that the lockdown **increased cigarette consumption by 9.1%** and emphasized the association of increased cigarette consumption with **increased mental distress among the study's participants.**

THEORY:

Counties with higher COVID-19 Severity:

- Individuals may undergo higher levels of stress due to health concerns, economic uncertainties, and disruptions to daily life, and may use smoking as a coping mechanism

Counties with lower COVID-19 Severity:

- Lower COVID-19 severity may be indicative of communities that have successfully implemented proactive public health measures, emphasizing health consciousness which may extend to lower acceptance of smoking

Hypothesis: Counties that experienced greater levels of severity of the COVID-19 virus will have higher smoking prevalence rate in 2022 then counties that experienced lower levels of severity.

HYPOTHESIS:

H_1: Counties with *lower* **median annual household income** in 2022 will have a *higher* **smoking prevalence.**

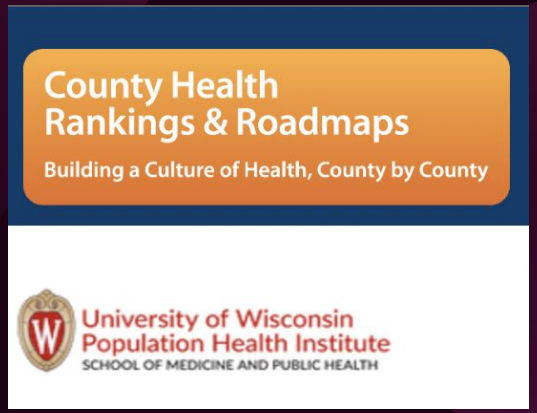
H_2: Counties with populations with *greater* **educational attainment levels** in 2022 will have a *lower* **smoking prevalence**.

H_3: Counties that experienced *greater* **levels of severity of the COVID-19 virus** will have *higher* **smoking prevalence** rate in 2022 then counties that experienced lower levels of severity.

DATA:

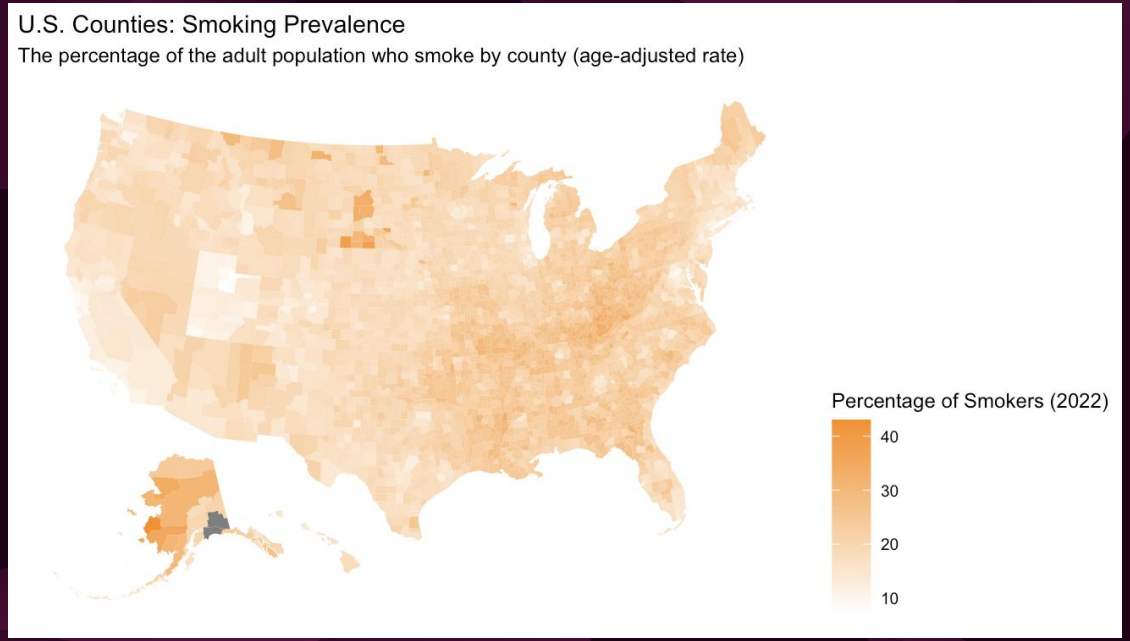
Smoking Prevalence (Dependent Variable)

Smoking prevalence **operationalized** by:
the percentage of the adult population who smoke by county (age-adjusted rate)



DATA:

Smoking Prevalence (Dependent Variable)



DATA:

Median Annual Household Income
(H_1: Independent Variable)

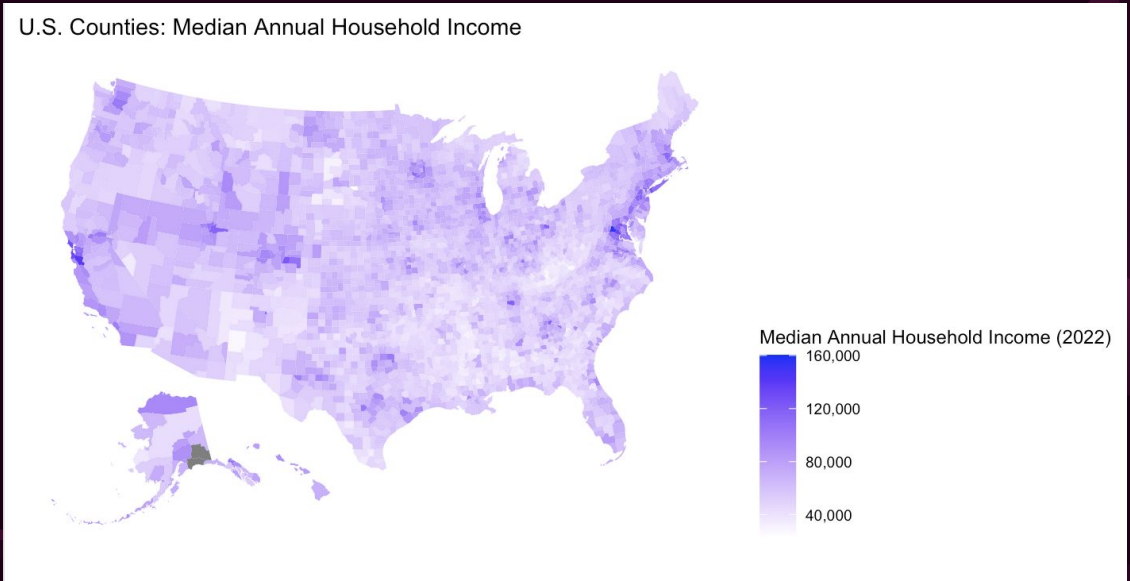
United States® Census Bureau

Small Area Income and Poverty Estimates (SAIPE) Program

The U.S. Census Bureau's Small Area Income and Poverty Estimates program produces single-year estimates of income and poverty for all U.S. states and counties

DATA:

Median Annual Household Income
(H_1: Independent Variable)



DATA:

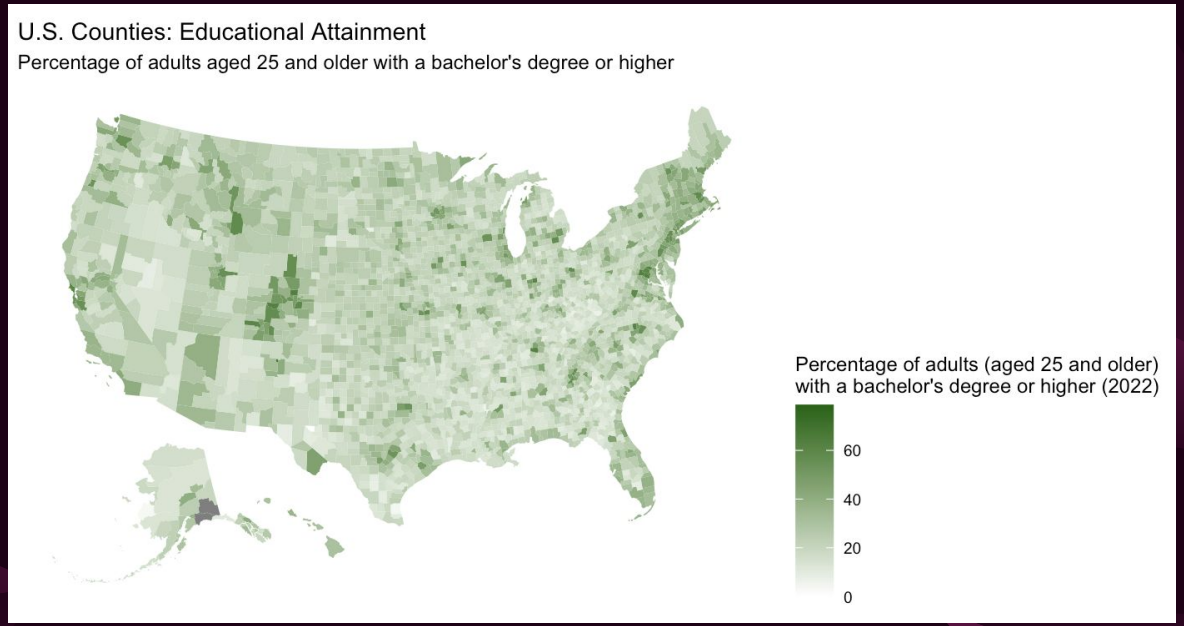
Educational attainment
(H_2: Independent Variable)

Educational attainment **operationalized** by:
Percentage of adults aged 25 and older with a bachelor's degree or higher

Estimates are derived from survey data collected over a 5-year period from 2017 to 2021

DATA:

Educational Attainment
(H_2: Independent Variable)



DATA:

Severity of the COVID-19 Virus (H_3: Independent Variable)

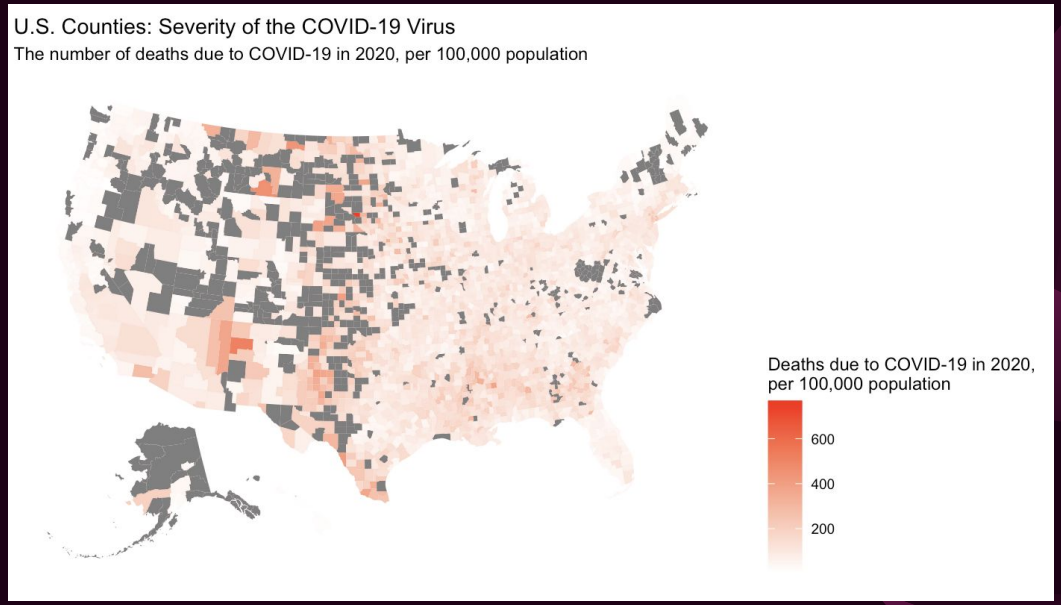
Severity of the COVID-19 virus **operationalized** by:
the number of deaths due to COVID-19 in 2020, per 100,000 population.

CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™
National Center for Health Statistics

NVSS
National Vital Statistics System

DATA:

Severity of the COVID-19 Virus
(H_3: Independent Variable)



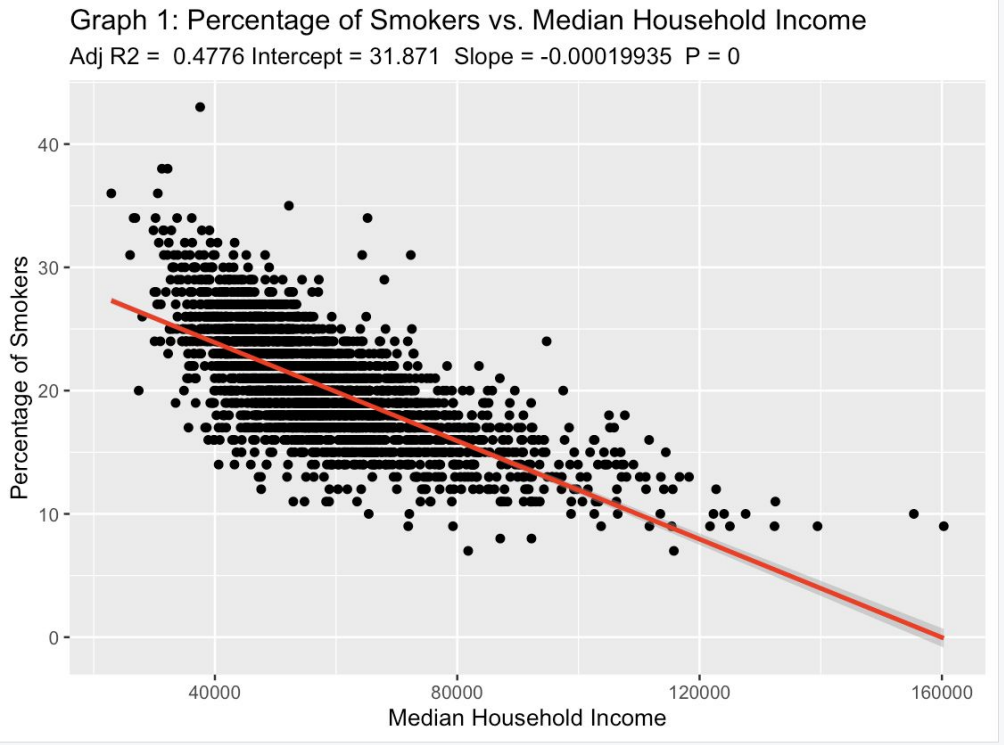
RESULTS: CORRELATION COEFFICIENT

Table 1: Correlation Coefficients (Pearson's) and Significance Values

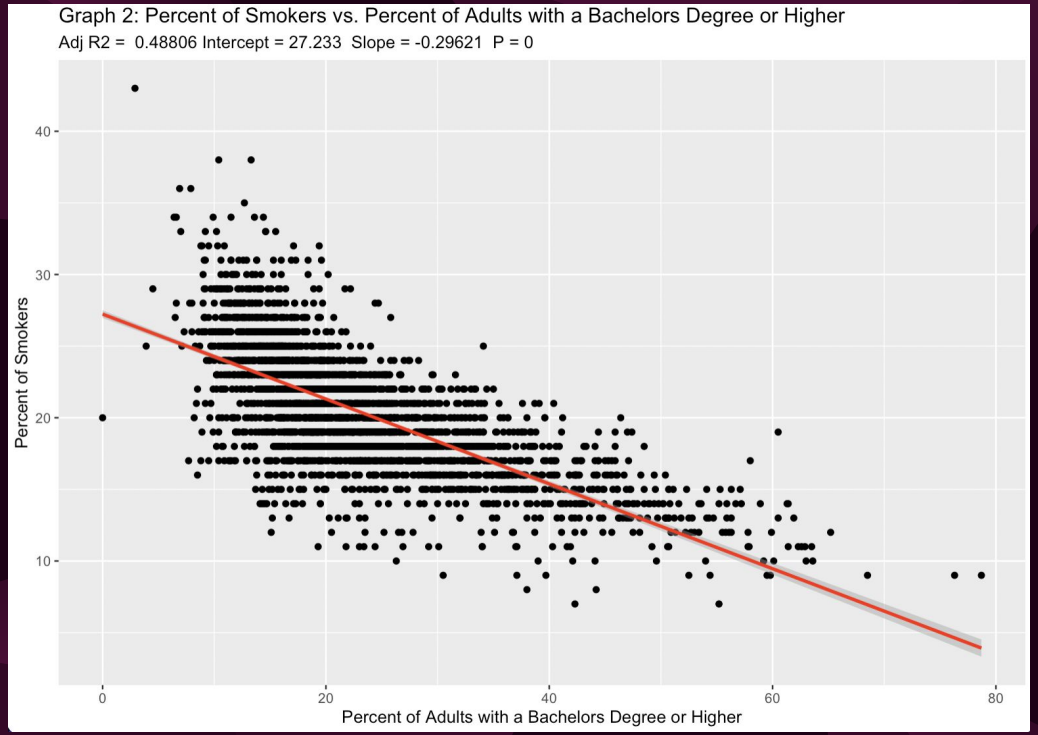
	Percentage of Smokers
Median Household Income	-0.6912086***
Percent of Adults with a Bachelor's Degree or Higher	-0.6987278***
COVID-19 Death Rate	0.2453013***

Note: *p<0.1; **p<0.05; ***p<0.01

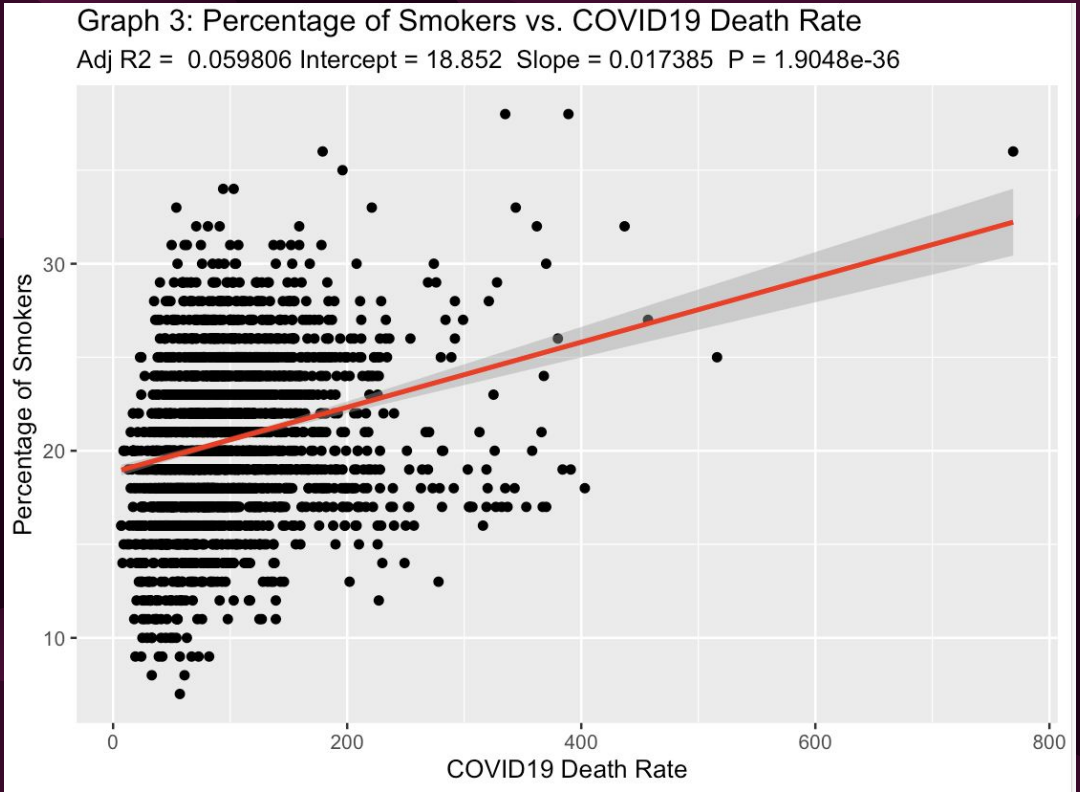
RESULTS: SIMPLE LINEAR REGRESSION



RESULTS: SIMPLE LINEAR REGRESSION



RESULTS: SIMPLE LINEAR REGRESSION



RESULTS: SIMPLE LINEAR REGRESSION

Table 1: Simple Linear Regression

	Dependent variable:		
	Percentage of Smokers		
	(1)	(2)	(3)
Median Household Income	-0.0002***		
	(0.00000)		
Percent of Adults with a Bachelors Degree or Higher		-0.296***	
		(0.005)	
COVID19 Death Rate			0.017***
			(0.001)
Constant	31.871***	27.233***	18.852***
	(0.221)	(0.136)	(0.157)
Observations	3,140	3,141	2,564
R^2	0.478	0.488	0.060
Adjusted R^2	0.478	0.488	0.060
Residual Std. Error	3.040 (df = 3138)	3.010 (df = 3139)	4.035 (df = 2562)
F Statistic	2,870.839*** (df = 1; 3138)	2,994.501*** (df = 1; 3139)	164.033*** (df = 1; 2562)
Note:			*p<0.1; **p<0.05; ***p<0.01

$$y_{Percentage\ of\ Smokers} = \beta_0 + \beta_1 x_{Median\ Household\ Income} + \varepsilon$$

$$y_{Percentage\ of\ Smokers} = \beta_0 + \beta_1 x_{Percent\ of\ population\ with\ Bachelor's\ degree\ or\ higher} + \varepsilon$$

$$y_{Percentage\ of\ Smokers} = \beta_0 + \beta_1 x_{Number\ of\ deaths\ due\ to\ COVID-19\ in\ 2020,\ per\ 100{,}000\ population} + \varepsilon$$

RESULTS: MULTIPLE LINEAR REGRESSION

Table 2: Multiple Linear Regression

	Dependent variable:
	Percentage of Smokers
Median Household Income	-0.0001***
	(0.00001)
Percent of Adults with a Bachelors Degree or Higher	-0.173***
	(0.008)
COVID19 Death Rate	-0.003***
	(0.001)
Constant	31.780***
	(0.260)
Observations	2,564
R^2	0.609
Adjusted R^2	0.608
Residual Std. Error	2.604 (df = 2560)
F Statistic	1,328.520*** (df = 3; 2560)
Note:	*p<0.1; **p<0.05; ***p<0.01

$$y_{Percentage\ of\ Smokers} = \beta_0 + \beta_1 x_{Median\ Household\ Income} + \beta_2 x_{Percent\ of\ population\ with\ Bachelor's\ degree\ or\ higher} + \beta_3 x_{number\ of\ deaths\ due\ to\ COVID-19\ in\ 2020,\ per\ 100{,}000\ population} + \varepsilon$$

RESULTS: ACCOUNTING FOR HETEROSKEDASTICITY

Table 3: Multiple Linear Regression with Robust Standard Errors

	Coefficients	Robust_SE	t_value	p_value
(Intercept)	31.7798934880868	0.350280114655214	90.7270842918622	0.0000
Median.Household.Income	-0.000120089118623953	5.91028840022666e-06	-20.3186562976093	0.0000
Percent.of.adults.with.a.bachelors.degree.or.higher	-0.172848584918991	0.00770948448293894	-22.4202520027771	0.0000
COVID19.death.rate	-0.00272108063075576	0.00163430497924222	-1.66497726270004	0.0960

CONCLUSION:

Smoking research is critical to continue improving public health.

- We mapped certain demographics and confirmed the relationship between smoking and low household income and educational attainment.
- By determining which areas are most affected by smoking, future efforts in public health can focus on those areas by county, and on the vulnerable demographics.
- The after-affects of COVID-19 have many more areas to research.
- This research has possibilities for improvement and for continued development.

REFERENCES

Alomari, M. A., Khabour, O. F., Alzoubi, K. H., & Maikano, A. B. (n.d.). *The impact of covid-19 pandemic on tobacco use: A population-based study*. PLOS ONE. https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0287375

Carreras, G., Lugo, A., Stival, C., Amerio, A., Odone, A., Pacifici, R., Gallus, S., & Gorini, G. (2021). Impact of covid-19 lockdown on smoking consumption in a large representative sample of Italian adults. *Tobacco Control*, *31*(5), 615–622. https://doi.org/10.1136/tobaccocontrol-2020-056440

Correlation (Pearson, Kendall, Spearman). Statistics Solutions. (2021a, August 10). https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/correlation-pearson-kendall-spearman/

Dwyer-Lindgren, L., Mokdad, A. H., Srebotnjak, T., Flaxman, A. D., Hansen, G. M., & Murray, C. J. (2014). Cigarette smoking prevalence in US counties: 1996-2012. *Population Health Metrics*, *12*(1). https://doi.org/10.1186/1478-7954-12-5

Garrett, B. E., Martell, B. N., Caraballo, R. S., & King, B. A. (2019). Socioeconomic differences in cigarette smoking among sociodemographic groups. *Preventing Chronic Disease*, *16*. https://doi.org/10.5888/pcd16.180553

REFERENCES

Homoscedasticity. Statistics Solutions. (2021b, August 3). https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/homoscedasticity/

How to interpret regression analysis results: P-values and coefficients. Minitab Blog. (n.d.). https://blog.minitab.com/en/adventures-in-statistics-2/how-to-interpret-regression-analysis-results-p-values-and-coefficients

Introduction to R and rstudio. Introduction to R - ARCHIVED. (2017, September 8). https://hbctraining.github.io/Intro-to-R/lessons/01_introR-R-and-RStudio.html

Jahnel, T., Ferguson, S. G., Shiffman, S., Thrul, J., & Schüz, B. (2018). Momentary smoking context as a mediator of the relationship between SES and smoking. *Addictive Behaviors*, *83*, 136–141. https://doi.org/10.1016/j.addbeh.2017.12.014

Johnston, S. (2015, April 23). *A quick and easy function to plot LM() results with GGPLOT2 in R*. Johnston Lab. https://sejohnston.com/2012/08/09/a-quick-and-easy-function-to-plot-lm-results-in-r/

REFERENCES

Kassambara, Leo, Sara, Mbugua, F., Kassambara, & Visitor. (2018, March 10). *Multiple linear regression in R*. STHDA. http://www.sthda.com/english/articles/40-regression-analysis/168-multiple-linear-regression-in-r/#:~:text=Multiple%20regression%20analysis%20allows%20researchers,of%20other%20predictors%20statistically%20eliminated.

Miller, S. (2021, June 25). *What is R used for? exploring the R programming language*. Codecademy Blog. https://www.codecademy.com/resources/blog/what-is-r-used-for/

Multiple regression analysis. Multiple Regression Analysis - an overview | ScienceDirect Topics. (n.d.). https://www.sciencedirect.com/topics/economics-econometrics-and-finance/multiple-regression-analysis#

O'Donnell, R., Eadie, D., Stead, M., Dobson, R., & Semple, S. (2021). 'I was smoking a lot more during lockdown because I can': A qualitative study of how UK smokers responded to the COVID-19 lockdown. *International Journal of Environmental Research and Public Health*, *18*(11), 5816. https://doi.org/10.3390/ijerph18115816

REFERENCES

Porras, E. M. (2022, December 5). *R linear regression tutorial: LM function in R with code examples*. DataCamp. https://www.datacamp.com/tutorial/linear-regression-R

Rogés, J., Bosque-Prous, M., Colom, J., Folch, C., Barón-Garcia, T., González-Casals, H., Fernández, E., & Espelt, A. (2021). Consumption of alcohol, cannabis, and tobacco in a cohort of adolescents before and during COVID-19 confinement. *International Journal of Environmental Research and Public Health*, *18*(15), 7849. https://doi.org/10.3390/ijerph18157849

RStudio desktop. Posit. (2023, September 22). https://posit.co/download/rstudio-desktop/

What is R?. R. (n.d.). https://www.r-project.org/about.html