Low-Income Jobs Lost to COVID-19

Week 5 Presentation

By: Yarra Abozaed, Zach Johansen, Janelle Acob, Tim Chen
DAT 4500

TABLE OF CONTENTS

RECAP

Last Week:

- Research Question: How do different states respond to the covid pandemic and how does that affect the workforce?
- Mathematical measure of covid response on a state level and the impact it had on the workforce in said state
- Looking at other datasets mentioned
  - Hypothesis testing for Rate of Vaccinations, Average Salary, etc.
  - Metric that correlates to low-income jobs
- Continue seeking explanations to relationships we find (start creating a wholistic picture for what we're seeing)

Metrics to Measure COVID-19 Regulations

- Rate of vaccinations
- Length of Vaccine Availability
- Length of Masking Requirement
- Gathering Restrictions (Max Length)
- Income tax
- Average Salary Per State
- Health Ranking by State
- Political Standing
- Rate of Jobs that Went Remote

OUR HYPOTHESIS TESTING

Hypothesis Testing 1: Rate of Vaccination

The relationship between the proportion of low-income jobs lost per state (by population) and rates of vaccinations →

H_0 : There is no correlation between the proportion of low-income jobs lost per state (by population) and the rate of vaccinations in the U.S
H_a : There is a correlation between the proportion of low-income jobs lost per state (by population) and the rate of vaccinations in the U.S

Results: With a p-value of 1.497×10^{-5} we reject the null hypothesis and conclude that there is a moderate positive correlation between the proportion of low-income jobs lost per state (by population) and the rate of vaccinations (by state). Furthermore, the 95% confidence interval of (0.3441461, 0.7280574) does not include 0, which further supports the conclusion that there is a relationship between these variables.

Hypothesis Testing 2: Length of Vaccine Availability (by state)

The relationship between the proportion of low-income jobs lost per state (by population) and length of Vaccine Availability (by each state) →

H_0 : There is no correlation between the proportion of low-income jobs lost per state (by population) and the length of vaccine availability in the U.S
H_a : There is a correlation between the proportion of low-income jobs lost per state (by population) and the length of vaccine availability in the U.S

Results: With a p-value of 6.24×10^{-6} we reject the null hypothesis and conclude that there is a moderate positive correlation between the proportion of low-income jobs lost per state (by population) and the length of vaccine availability in the U.S (by each state). Furthermore, the 95% confidence interval of (0.3702306, 0.7418059) does not include 0, which further supports the conclusion that there is a relationship between these variables.

Hypothesis Testing 3: Length of Masking Requirement

The relationship between the proportion of low-income jobs lost per state (by population) and length of mask mandates →

H_0 : There is no correlation between the proportion of low-income jobs lost per state (by population) and length of mask mandates
H_a: There is a correlation between the proportion of low-income jobs lost per state (by population) and length of mask mandates

Results: With a p-value of 1.929×10^{-7}, we reject the null hypothesis and conclude that there is a positive, strong correlation between low-income jobs lost per state (by population) and length of mask mandates. Furthermore, the 95% confidence interval of (0.4665281, 0.7922343) does not include 0, which further supports the conclusion that there is a relationship between these variables.

Hypothesis Testing 4: Gathering Restrictions (Max Length)

The relationship between the proportion of low-income jobs lost per state (by population) and length of gathering bans per state →

H_0: There is no correlation between the proportion of low-income jobs lost per state (by population) and length of gathering bans per state
H_a: There is a correlation between the proportion of low-income jobs lost per state (by population) and length of gathering bans per state

Results: With a p-value of 0.0009606, we reject the null hypothesis and conclude that there is a positive, moderate correlation between the proportion of low-income jobs lost per state (by population) and length of gathering bans per state. Furthermore, the 95% confidence interval of (0.20, 0.64) does not include 0, which further supports the conclusion that there is a relationship between these variables.

Hypothesis Testing 5: Income Tax

The relationship between the proportion of low-income jobs lost per state (by population) and the average income tax per state →

H_0: There is no correlation between the proportion of low-income jobs lost per state (by population) and the average income tax per state
H_a: There is a correlation between the proportion of low-income jobs lost per state (by population) and the average income tax per state

Results: With a p-value of 0.03623, (and a cor. coef. of 0.294) we reject the null hypothesis and conclude that there is a positive, weak correlation between the proportion of low-income jobs lost per state (by population) and the average income tax per state. Furthermore, the 95% confidence interval of (0.02007688 , 0.52692053) does not include 0, which further supports the conclusion that there is a relationship between these variables.

Hypothesis Testing 6: Average Salary Per State

The relationship between the proportion of low-income jobs lost per state (by population) and the average salary per state →

H_0: There is no correlation between the proportion of low-income jobs lost per state (by population) and the average salary per state
H_a: There is a correlation between the proportion of low-income jobs lost per state (by population) and the average salary per state

Results: With a p-value of 0.00113, (and a cor. coef. of 0.447) we reject the null hypothesis and conclude that there is a positive, moderate correlation between the proportion of low-income jobs lost per state (by population) and the average salary per state. Furthermore, the 95% confidence interval of (0.193, 0.645) does not include 0, which further supports the conclusion that there is a relationship between these variables.

Hypothesis Testing 7: Health Ranking by State

The relationship between the proportion of low-income jobs lost per state (by population) and the nationwide state's health ranking →

H_0: There is no correlation between the proportion of low-income jobs lost per state (by population) and the nationwide state's health ranking
H_a: There is a correlation between the proportion of low-income jobs lost per state (by population) and the nationwide state's health ranking

Results: With a p-value of 0.001125, (and a cor. coef. of –0.447) we reject the null hypothesis and conclude that there is a negative, moderate correlation between the proportion of low-income jobs lost per state (by population) and the nationwide state's health ranking. Furthermore, the 95% confidence interval of (-0.645, -0.193) does not include 0, which further supports the conclusion that there is a relationship between these variables.

Hypothesis Testing 8: Political Standing

The relationship between the proportion of low-income jobs lost per state (by population) and the state political standing →

H_0 : There is no correlation between the proportion of low-income jobs lost per state (by population) and the state political standing
H_a : There is a correlation between the proportion of low-income jobs lost per state (by population) and the state political standing

Results: With a p-value of 0.259, (and a cor. coef. of -0.163) we fail to reject the null hypothesis and conclude that there is a negative, weak correlation between the proportion of low-income jobs lost per state (by population) and the state political standing. Furthermore, the 95% confidence interval of (-0.422, 0.121) does include 0, which does not support the conclusion that there is a relationship between these variables.

Hypothesis Testing 9: Rate of Jobs that went Remote

The relationship between the proportion of low-income jobs lost per state (by population) and the Proportion of jobs that went remote due to covid-19 (in each state) →

H_0 : There is no correlation between the proportion of low-income jobs lost per state (by population) and the Proportion of jobs that went remote due to covid-19
H_a : There is a correlation between the proportion of low-income jobs lost per state (by population) and the Proportion of jobs that went remote due to covid-19

Results: With a p-value of 0.04745 we reject the null hypothesis and conclude that there is a weak but positive correlation between the proportion of low-income jobs lost per state (by population) and the Proportion of jobs that went remote due to covid-19 in the U.S (by each state). Furthermore, the 95% confidence interval of (0.003690818, 0.519365939) does not include 0, which further supports the conclusion that there is a relationship between these variables.

Summary

Variable	P-Value	Correlation Coefficient	Relationship With PLIJs Lost
Rate of Vaccination	1.497×10^{-5}	0.57	Moderate & Positive (PLIJs Lost ↑, Rate of Vaccination ↑)
Length of Vaccine Availability	6.24×10^{-6}	0.60	Moderate & Positive (PLIJs Lost ↑, Length of Vaccine Availability ↑)
Length of Masking Requirement	1.929×10^{-7}	0.70	Strong & Positive (PLIJs Lost ↑, Length of Requirements ↑)
Gathering Restrictions (Max Length)	0.0009606	0.45	Moderate & Positive (PLIJs Lost ↑, Length of Gathering Restrictions ↑)
Income Tax	0.03623	0.294	Weak & Positive (Higher Income Tax, More LIJ Lost)
Average Salary (per state, 2021)	0.00113	0.45	Moderate & Positive (PLIJs Lost ↑, Average Salary ↑)
Healthcare Rank (2021)	0.001125	-0.50	Moderate & Negative (PLIJs Lost ↑, Healthcare rank ↓)
Political Standing	0.259	-0.163	No Association
Rate of Jobs that went Remote	0.04745	0.29	Weak & Positive (PLIJs Lost ↑, Rate of Jobs that went Remote ↑)

What are we doing with all these datasets?

- Perform ANOVA testing.
  - *We want to see if all variables have equal pull, if they do not, we then dive into what that means in regards to our response variable.*
- Use all variables but political standing to model the metric.
  - Will require some experimentation.

OUR CHALLENGES

Dataset Findings

- Finding the right external dataset for our metrics to do hypothesis/ correlation & a lot of data wrangling

Metrics

- Figuring out what Metrics to Use to measure covid-19 regulations (and making sure they are not confounding).

Our Goal

- Figuring out how to use all these variables & combining them into one measure

OUR NEXT STEPS

Visualizing the Results

Create visuals that get the story we're trying to tell across well.

Putting the Pieces Together

Start testing linear regression models with multiple variables

External Research

Can we find external resources that will help us tell a story through the data and results we have? (articles, research papers, journals...)

LIST OF REFERENCES

Variable	Dataset
Length of masking requirement	https://ballotpedia.org/State-level_mask_requirements_in_response_to_the_coronavirus_(COVID-19)_pandemic,_2020-2022
Healthcare rank (2021)	https://www.usnews.com/news/best-states/rankings/health-care
Gathering Restrictions (Max Length)	https://healthdata.gov/dataset/U-S-State-and-Territorial-Gathering-Bans-March-11-/8tfm-md2h/data?no_mobile=true
Rate of Vaccination	https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-total-admin-count-total
Length of Vaccine Availability	https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-total-admin-count-total
Rate of Jobs that went Remote	https://www.teamflowhq.com/blog/states-where-the-most-people-worked-remote-because-of-covid-19
Political Standing	https://ballotpedia.org/Election_results,_2020:_State_trifectas_and_the_2020_presidential_vote
Average Salary (per state, 2021)	https://www.statista.com/statistics/243850/private-industry-wages-per-employee-in-the-us-by-state/
Income Tax	https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-2020-zip-code-data-soi

QUESTIONS?