University of Findlay

Demographic, Socio-Economic, and Health Factors Affecting COVID-19 Infection Rate in Ohio

Indumathi Balasubramanian; William Doyle Ph.D., Capstone Advisor.
Department of Environmental, Safety, and Health Management, The University of Findlay, Ohio.

ABSTRACT

Some of the demographic, socio-economic, and health factors are chosen for analyzing their effect on the COVID-19 Infection Rate. A literature review is conducted with the journals to understand some key factors causing severe disease and mortality with coronavirus A Multiple Regression Analysis is done with the Statistical Package for Social Sciences (SPSS) software, with the Correlation and Collinearity Analysis. The factors are detected to have a 74% prediction on the Infection Rate, with significant coefficient values and p-values.

INTRODUCTION

The state of Ohio State in the United States with its 88 counties had its first case on March 9th, 2020 and has recorded 115,651 COVID-19 cases on the 24th of August,2020 occupying the 16th position within the 50 states of the country. The vast data collected by the various authorities like the Centers for Disease Control and Prevention (CDC), United States Census Bureau, United States Department of Agriculture (USDA), Ohio Department of Health and others can be utilized for understanding the factors affecting the COVID-19 outbreak in the counties of Ohio. The data analysis done by the previous research studies has helped solve some mystery that prevailed as to why some countries and people are more vulnerable than others and succumb to the infection. The present study can help find the relevance of the past studies in the various counties of Ohio and thereby give us insights in dealing with the infection and protecting us from further outbreaks.

CASE DESCRIPTION

The standard data of COVID-19 cases across counties in the State of Ohio from March 2020 to August 2020 are collected from standard databases like the Ohio Department of Health (Ohio.gov), and the Centers of Disease Prevention and Control (CDC). The CDC brings certain data in partnership with educational and technical organizations. The US Census Bureau had dedicated a resource page for the coronavirus situation as the data can help further research and devise economic and health policies at the government level. This secondary research study started with getting the approval of the Institutional Review Board (IRB) and studying the guidelines stated by my Capstone Advisor. The period of data from March 2020 till August 2020 will be collected, analyzed, and interpreted to understand the implications of the various factors on the infection rates of COVID-19. The data does not reveal the personal identity of anyone and is the public data published in the standard databases. The demographic, socio-economic, and health factors are taken up as independent variables and correlated with the infection rate of coronavirus cases.

• The dependent variable being the infection rate of COVID-19 cases in the various counties of the State of Ohio, United States, is log transformed for statistical analytic purposes. The research design of the present project is a retrospective, quantitative, cross sectional study involving data collected from standard databases. No intervention was made to the people as it is a data-oriented secondary research study. The demographic and economic studies can help the prospective interventions in helping the community protect itself through the insights made. The problem statement, hypothesis and research questions are carefully chosen to identify the relationship between the various demographic and economic variables in influencing the incident rate of the COVID-19 cases.

The demographic, socio-economic, and health factors in the general population are collected from the 2020 County Health Rankings of Ohio published by the University of Wisconsin Population Health Institute and funded by the Robert Wood Johnson Foundation Program, the U.S. Census Bureau COVID-19 site and United States Department of Agriculture (USDA) economic research service.

INTERVENTIONS

A Multiple Linear Regression, Descriptive Statistics, Correlation Analysis, and Collinearity Diagnostics were done to analyze the impact of the independent variables on the dependent variable. The analyses had led to significant results which will be discussed here. The Multiple Regression Analysis gave out the chosen outputs in the form of tables and graphs.

OUTCOMES

The literature review has cited many research studies and standard databases explaining the importance of the demographic, socio-economic, and health factors impacting the COVID-19 infection rate throughout the world and the various counties of Ohio. The Regression Analysis shows a good Model Fit with a R-Value of 0.860, R-Squared Value of 0.740, and a Significant p-value of 0.000 (<0.05). The data analysis though multiple linear regression analysis in SPSS software has proved a good model fit through the Model Summary, ANOVA summary and Regression Coefficient Analysis. A 74% (0.740 R-Squared value) prediction proved to be possible for the infection rate with the independent variables taken up for the study. A significant p-value of 0.000 as a model summary signifies a significant study on the effect of the predictor variables on the outcome variable. The Standardized coefficients (beta) and the significant p-values of the various predictor variables show wither a positive or negative effect on the coronavirus infection rate. Though many factors are included in the present study, there are other innumerable factors like climate, absolute humidity, political ideologies, media influences etc., which are known to affect the infection rate. For want of data and time, only the present factors are considered for our present study.

To show the variations in Infection Rate, I removed the Median HH income % of State Total 2018 variable. The R-Squared value changed from 0.740 to 0.711, and Adjusted R-Squared value changed from 0.570 to 0.531. This shows that Infection Rate varies with adding or removing one or more variables (Morgan et al.,2019). The Model Summary below shows the regression analysis summary after removing the Median Household (HH) Income % of State Total 2018 variable. The Sig. F-Change before and after removing the variable has a value 0.000 which is lesser than 0.05 and therefore, the variable does not affect the significance of the regression analysis and can be included in the study.

DISCUSSION

The Coefficient Statistics showed that almost all the predictor variables influenced the infection rate in the various counties of Ohio. Either through positive or negative direction, the infection rate is influenced by the predictors. Variables like % Males in Total Population, % Median age 65 years and over Male, % Black, % Asian, % Native Hawaiian/Other Pacific Islander, % Hispanic or Latino (Any race), % Not Proficient in English, % Rural, % Unemployed, %Internet-at-home, % With Access to Exercise Opportunities, Residential segregation - non-White/White (Segregation index), % Physically Inactive, % Frequent Physical Distress, and % Adults with Obesity increase by 1 unit, the log infection rate increases by the mentioned standardized coefficient value (beta). For example, % Male in Total Population increases by 1 unit, the log infection rate increases by 0.503. The independent variables like Population Density, % 65 years age and over, % American Indian/ Alaska Native, % Multiracial, Median Income % of State Total 2018, % Severe Housing problems, % Uninsured, % Food Insecure, % Limited Access to Healthy Foods, % of adults with a high school diploma only, 2014-18, % Some College, Life Expectancy, % Fair or Poor Health, % Frequent Mental distress, % Insufficient Sleep, % Adults with Diabetes, % Smokers, and % Excessive Drinking have an impact but in the negative direction

CONCLUSIONS

With all the eight assumptions of the Multiple Regression Analysis being met, we can infer the conduct and validity of the current research. The results have a high R-Squared value and significant p-value. Health care quality and interventions, political and social constructs, and global connectedness are some key factors which need due consideration in predicting a pandemic like COVID-19.

Health care quality and interventions, political and social constructs, and global connectedness are some key factors which need due consideration in predicting a pandemic like COVID-19.

Health care quality and interventions, political and social constructs, and global connectedness are some key factors which need due consideration in predicting a pandemic like COVID-19.

REFERENCES

• Centers for Disease Control and Prevention,2020. Explore Human mobility and COVID-19 transmission in your local area. https://covid.cdc.gov/covid-data-tracker/?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fcases-updates%2Fcases-in-us.html#mobility
• Centers for Disease Control and Prevention,2020. United States Laboratory testing. https://covid.cdc.gov/covid-data-tracker/#testing