

Stat 512 Project Summary Page

1. **Names:** Daniel Furry, Alex Scarcelli, Cavan McFarland
2. **Section number:** STAT 51200 - 010 (10:30-11:45)
3. **Project title:** COVID Mortality in Indiana – A state with mixed vaccination rates
4. **Project YouTube link:** <https://youtu.be/r8PdE0ci3VA>
5. **Project background introduction (why this is an important question, what has been done on the question, what are your major research questions in the project, etc.):**

General Research Question: What characteristics primarily affect the death rate from Covid-19 in the state of Indiana by county.

We are testing this research question to help find a group of individuals that are more prone to contract and die from COVID-19. Over the past year, there has been extensive testing and analysis on this disease to discover more about it and how we can be better prepared to not contract the disease and therefore decreasing the mortality rates. This is important in order for the population to be more aware of their chances of dying from this disease if they wish to take extra precautions knowing that they are apart of a group that has a higher mortality rate.

Question 1: Does infection rate and the vaccine rate have a significant linear impact on the Covid-19 death rate?

Question 2: Do demographic factors such as age and gender have a significant linear effect on the COVID death rate/population of each county?

Question 3: Does the amount of high-risk people living in a county have a significant linear impact on the COVID death rate?

6. **Project result highlight (what are the major findings of your project, what do you consider the most contribution of this project):**

In our research we found that to best predict the number of deaths in the state of Indiana we should use only 3 predictors out of our 5 chosen originally. The most significant predictors were vaccine rate, percentage of the population at high risk, and the average death age due to COVID-19.

The biggest contribution to this project was transforming our Y variable because we needed to transform Y for all of our hypothesis tests. This shows that we had flaws in the normality of our data and once we transformed Y, this fixed our issues with the Total_deaths variable. After this transformation we were able to run the appropriate tests we needed on our data to drive results and draw conclusions.

7. **Project data introduction (the exact data resource, a table summarizes variable notation and definition, such as the one on the first page in the homework).**

Data background

The data set provides selected county demographic information for the 92 counties in Indiana. Each line of the data set has an identification number with a county name and provides information on 12 variables for a single county. The 14 variables are:

Variable number	Variable name	Description
1	County	County name
2	population	County population
3	full_vaccinated	Number of fully vaccinated people by county
4	vaccine_100k	(Number of vaccinated people / county population) * 100,000
5	cases_100k	(Number of cases / county population) * 100,000
6	pop_65_plus	Population over 65 (at risk)
7	high_risk_100k	(Population over 65 / county population) * 100,000
8	reg_deaths	Deaths under 65
9	at_risk_deaths	Deaths over 65
10	avg_death_age	Average death age
11	M_deaths	Number of male deaths
12	F_deaths	Number of female deaths
13	M_F_diff_100k	(Male deaths - Female deaths) * 100,000
14	Total_deaths	Total deaths

8. Project reference (attach the research paper if applicable):

Data Resources:

General: <https://www.nytimes.com/interactive/2021/us/marion-indiana-covid-cases.html?auth=link-dismiss-google1tap>

Vaccination Rate: <https://data.rgi.com/covid-19-vaccine-tracker/indiana/18/>

COVID-19 Deaths by Demographics (Age & Gender):

https://hub.mph.in.gov/dataset/covid-19-case-demographics-by-county/resource/9ae4b185-b81d-40d5-ae2-f0e30405c162?inner_span=True

High Risk (Age 65+ by county): <https://www.covid19atrisk.org/at-risk.html?state=IN>

Infection Rate:

<https://hub.mph.in.gov/dataset/covid-19-county-categorization-scores/resource/3db61861-f7dc-4b5d-a917-cb5d1a845afd>

Dict:

<https://hub.mph.in.gov/dataset/covid-19-county-categorization-scores/resource/3f08c302-468c-4d45-9cdc-5952d605f5ff>