Contents

[Introduction: 2](#__RefHeading___Toc445_3522661246)

[Dataset: 2](#__RefHeading___Toc447_3522661246)

[Research Questions: 2](#__RefHeading___Toc449_3522661246)

[Formulation of Null and Alternative Hypotheses 3](#__RefHeading___Toc451_3522661246)

[Visualization: 3](#__RefHeading___Toc453_3522661246)

[Analysis: 5](#__RefHeading___Toc455_3522661246)

[Formulation of Null and Alternative Hypotheses 5](#__RefHeading___Toc457_3522661246)

[Significance level 6](#__RefHeading___Toc459_3522661246)

[Calculations: 6](#__RefHeading___Toc461_3522661246)

[Conclusion: 6](#__RefHeading___Toc463_3522661246)

# Introduction:

This study is carried out to assess the covid-19 situation across all the states in USA. This study analyses the daily covid cases from diffrent counties in each state and thery by aggregate results at state level to compare the proportion of covid deaths registered for 100000 confirmed cases in each State. This research empowers other studies on social and economic factors in USA.

## Dataset:

The dataset used was taken by website

https://data.world/associatedpress/johns-hopkins-coronavirus-case-tracker

The Following are the variable from dataset:

last\_update

location\_type

state

county\_name

county\_name\_long

fips\_code

lat

lon

NCHS\_urbanization"

total\_population

confirmed

confirmed\_per\_100000 (County Level)

deaths

deaths\_per\_100000 (County level)

The dataset can benefit greatly from additional content. Economics, additional demographics and more.

## Research Questions:

Is there a difference in number of covid deaths per 100000 confirmed cases among all the states of USA?

Here, we want to confirm If all the states in USA have equal proportion of covid deaths.

## Formulation of Null and Alternative Hypotheses

**Null hypothesis**

H0: P0=P1=P2=……….=P51=P52

There is no significant difference in number of covid deaths per 100000 confirmed cases among all the states of USA.

**Alternative Hypothesis:**

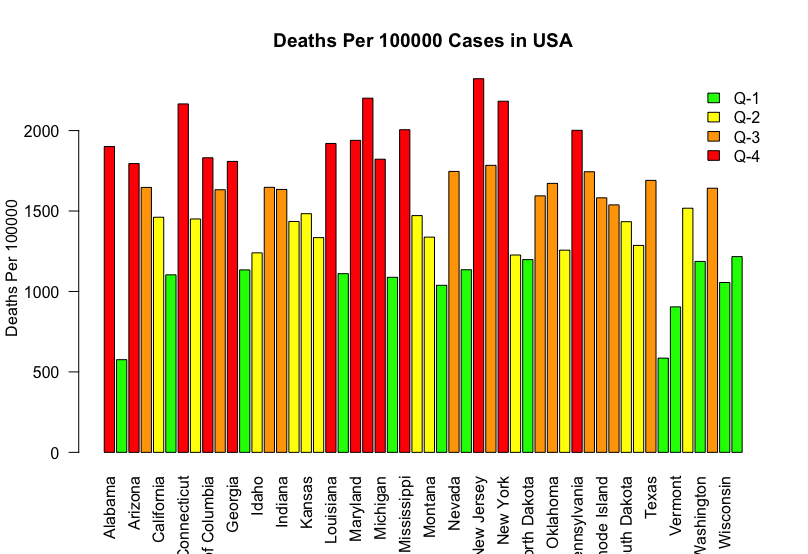
H1: P0≠P1≠P2≠……..≠P51≠P52

There is a significant difference in number of covid deaths per 100000 confirmed among all the states of USA.

# Visualization:

Different visualization was used in this study to assess the true picture of data.

**Total Covid deaths in states:**

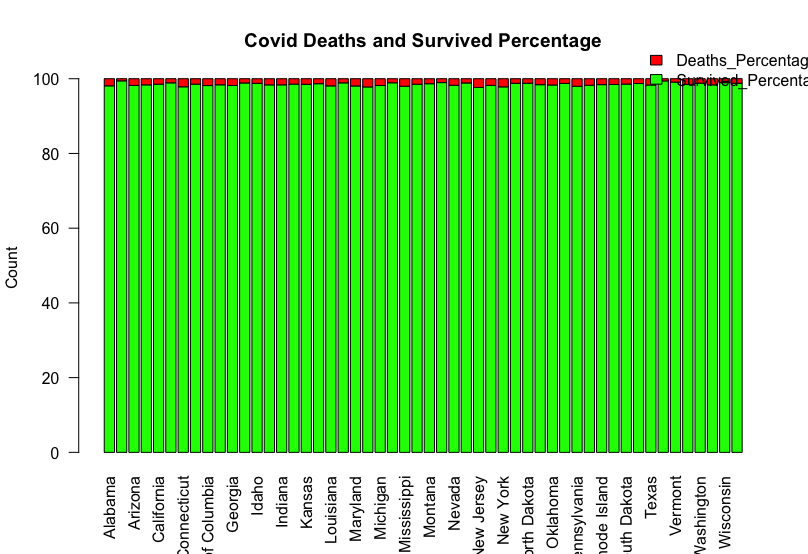
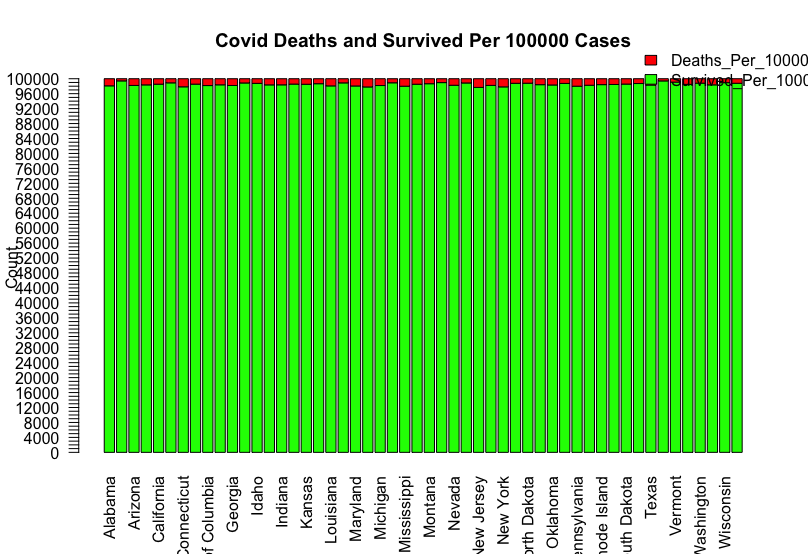


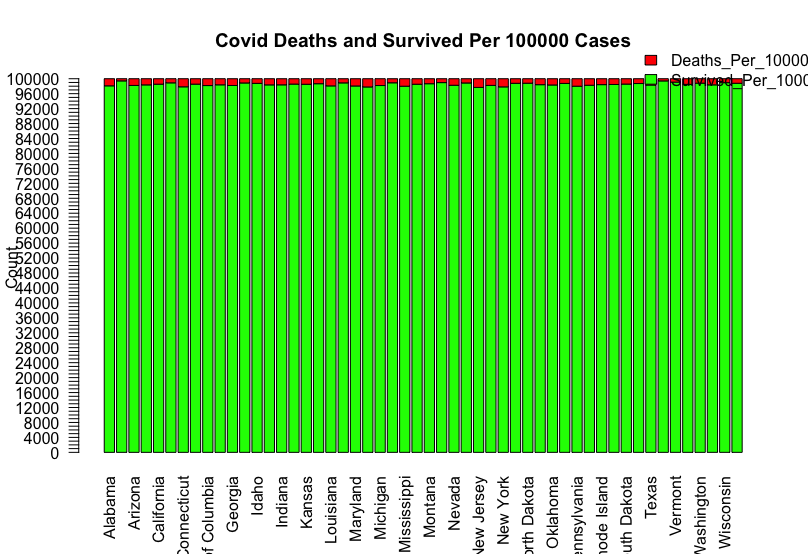
The above bar graph shows total number of covid deaths in each state and each state is categorized in to four buckets based the below quantiles.

* Covid deaths < 25 quantile: green
* Covid deaths < 50 and Covid deaths > 25: yellow
* Covid deaths< 75 and Covid deaths > 50: orange
* Covid deaths > 75: red

We can observe all the states are not categorized to same bucket.

**Covid Death Proportions per 100000 cases:**





From the above two diagrams we can observe there is a slight difference in the deaths proportion for each state. To get the better understanding of data we need to perform the statistical analysis and there by can conclude with better results.

# Analysis:

We have used 52-sample Z test for equality of proportions.

We used this test to address our research question that is number of covid deaths per 100000 confirmed cases among all the states of USA is different or not?

The steps of testing are given as

## Formulation of Null and Alternative Hypotheses

**Null hypothesis**

H0: P0=P1=P2=……….=P51=P52

There is no significant difference in number of covid deaths per 100000 confirmed cases among all the states of USA.

**Alternative Hypothesis:**

H1: P0≠P1≠P2≠……..≠P51≠P52

There is a significant difference in number of covid deaths per 100000 confirmed among all the states of USA.

## Significance level

ɑ=0.05

## 

## Calculations:

We calculated Z-statistic and p-value using R.

data: state\_level\_deaths$Deaths\_Per\_100000 out of population

X-squared = 5143.5

df = 51

p-value < 2.2e-16

# Conclusion:

The p-value of the test is lesser than significance level alpha=0.05, We can conclude that the of number covid deaths per 100000 confirmed cases in different states of USA is not equal. There by we can accept the alternate hypothesis and reject the null hypothesis.