

# Abstract/Executive Summary

There are many studies conducted on Covid19 pandemic in evaluating the spread of the disease within the population across U.S and other countries. These studies are in terms of quantifying the number of contaminated and deaths. However, little research has been done to explore the impact of Covid19 pandemic on the population demographic. The purpose of this study is to investigate the population demographic facts such as the number of children per household, marital status and life expectancy before(2019) and after (2021/2022) Covid19 pandemic. This research will compare U.S. results to the other most impacted countries by Covid19.

This research will be conducted in academic approach and will go through data collection from official sources (U.N , etc,) and perform data analysis applying explanatory analysis method using R and Python programming languages.The data visualization will be wrapped in apps for easy use.

One of the most popular preventive measures used during Covid19, stay home/ Working From Home(WFH) was found to have a significance on the increase of number of children per household in the U.S (other countries as well?). although this study definively answers questions regarding the correlation between the Covid19 pandemic on the population demographic, further studies are needed to take the analysis at the state level and possibly create a map for a better visualization.

## Literature Review/Research Conducted

Recent studies show that China was already experiencing Covid19 deaths, By the end of 2019. Then, the Covid19 rapidly spread and went viral in many countries. The Covid19 pandemic has been causing death globally for the past 03 years(2019-2022).

## Theory and Hypotheses (if applicable)

H\_0 : there were more children per household during COVID-19 H\_a:

### Research Questions

What is the number of children per household before and after Covid19 in the U.S.? Does the number of children per household increase or decrease during Covid19 in the U.S. ? Are there more divorces during Covid19 or before in the U.S.? How is the number of children per household in the U.S. compared to other countries? Do Women live longer than Men in The U.S.? Do Women in the U.S. live longer than women in other countries? Do Men in the U.S. live longer than men in other countries?

## Data and Methods

### Data Sources

This project relies heavily on population demographics data. Due to the sensitive of the topics being discussed on this project and to remain objective and credible, we explored many official or well-known sources included but not limited: the United Nations tatistics Division data[4], the United States Census Bureau[5], Our World in Data[3], worldometers[6].

## remember to combine data

### Exploratory Data Analysis

let's us understand the dataset ...

- Data available -> [https://github.com/asmozo24/DATA606\\_Project\\_Proposal](https://github.com/asmozo24/DATA606_Project_Proposal)

Using R to acquire data

Population by marital status, age, sex and urban/rural residence

Country.or.Area <chr>	Y... Reference.Date <chr><chr>	Area S... Age <chr> <chr><chr>	Marital.status <chr>		
1United States of America	2012 1-Jul-12	Total Male Total	Total		
2United States of America	2012 1-Jul-12	Total Male Total	Single (never married)		
3United States of America	2012 1-Jul-12	Total Male Total	Married		
4United States of America	2012 1-Jul-12	Total Male Total	Widowed and not remarried		
5United States of America	2012 1-Jul-12	Total Male Total	Divorced and not remarried		
6United States of America	2012 1-Jul-12	Total Male Total	Married but separated		

6 rows | 1-8 of 13 columns

```
##
## Attaching package: 'skimr'
```

```
## The following object is masked from 'package:nanian':
##
##   n_complete
```



Data summary

Name	UN_popM
Number of rows	6800
Number of columns	12
<hr/>	
Column type frequency:	
character	10
numeric	2
<hr/>	
Group variables	None

Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
Country.or.Area	0	1	0	24	1	7	0
Year	0	1	0	124	1	37	0
Reference.Date	0	1	0	8	6	33	0
Area	0	1	0	5	6	4	0
Sex	0	1	0	6	6	3	0
Age	0	1	0	7	6	33	0
Marital.status	0	1	0	26	6	9	0
Record.Type	0	1	0	38	6	4	0
Reliability	0	1	0	22	6	2	0
Value.Footnotes	0	1	0	3	345	5	0

Variable type: numeric

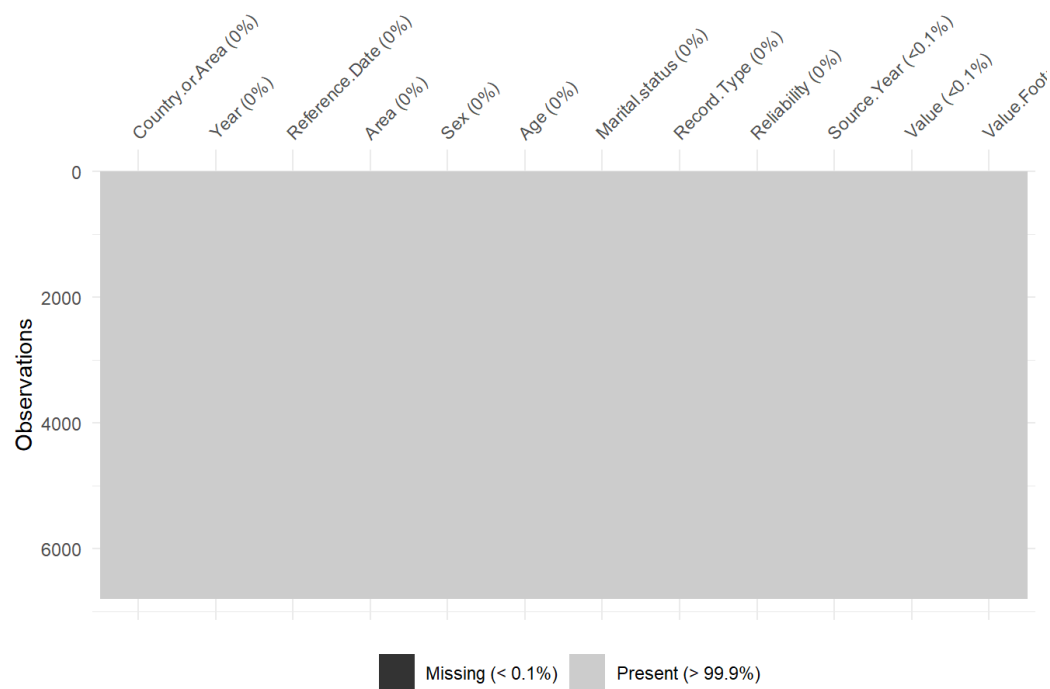
skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
Source.Year	6	1	1986.58	9.28	1976	1979	1984.0	1994	2014	
Value	6	1	4990783.24	14157671.38	0	22000	418705.5	4765312	153327000	

```
## Loading required package: usethis
```

```
##
## Attaching package: 'devtools'
```

```
## The following object is masked from 'package:rsconnect':  
##  
## lint
```

```
## Warning: `gather_()` was deprecated in tidyr 1.2.0.  
## Please use `gather()` instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was generated.
```



```
##  
##  
## processing file: report.rmd
```

```

##
|
|
|
|..
## inline R code fragments
##
##
|
|...
## label: global_options (with options)
## List of 1
## $ include: logi FALSE
##
##
|
|.....
## ordinary text without R code
##
##
|
|.....
## label: introduce
##
|
|.....
## ordinary text without R code
##
##
|
|.....
## label: plot_intro

```

```

##
|
|.....
## ordinary text without R code
##
##
|
|.....
## label: data_structure
##
|
|.....
## ordinary text without R code
##
##
|
|.....
## label: missing_profile

```

```

##
|
|.....| 26%
## ordinary text without R code
##
##
|
|.....| 29%
## label: univariate_distribution_header
##
|
|.....| 31%
## ordinary text without R code
##
##
|
|.....| 33%
## label: plot_histogram

```

```

##
|
|.....| 36%
## ordinary text without R code
##
##
|
|.....| 38%
## label: plot_density
##
|
|.....| 40%
## ordinary text without R code
##
##
|
|.....| 43%
## label: plot_frequency_bar

```

```

##
|
|.....| 45%
## ordinary text without R code
##
##
|
|.....| 48%
## label: plot_response_bar
##
|
|.....| 50%
## ordinary text without R code
##
##
|
|.....| 52%
## label: plot_with_bar
##
|
|.....| 55%
## ordinary text without R code
##
##
|
|.....| 57%
## label: plot_normal_qq

```

```
##
|
|.....| 60%
## ordinary text without R code
##
##
|
|.....| 62%
## label: plot_response_qq
##
|
|.....| 64%
## ordinary text without R code
##
##
|
|.....| 67%
## label: plot_by_qq
##
|
|.....| 69%
## ordinary text without R code
##
##
|
|.....| 71%
## label: correlation_analysis
```

```
##
|
|.....| 74%
## ordinary text without R code
##
##
|
|.....| 76%
## label: principal_component_analysis
```

```

##
|
| ..... | 79%
## ordinary text without R code
##
##
|
| ..... | 81%
## label: bivariate_distribution_header
##
|
| ..... | 83%
## ordinary text without R code
##
##
|
| ..... | 86%
## label: plot_response_boxplot
##
|
| ..... | 88%
## ordinary text without R code
##
##
|
| ..... | 90%
## label: plot_by_boxplot
##
|
| ..... | 93%
## ordinary text without R code
##
##
|
| ..... | 95%
## label: plot_response_scatterplot
##
|
| ..... | 98%
## ordinary text without R code
##
##
|
| ..... | 100%
## label: plot_by_scatterplot

```

```
## output file: C:/Users/owner/OneDrive/Documents/R/Master_Research/report.knit.md
```

```

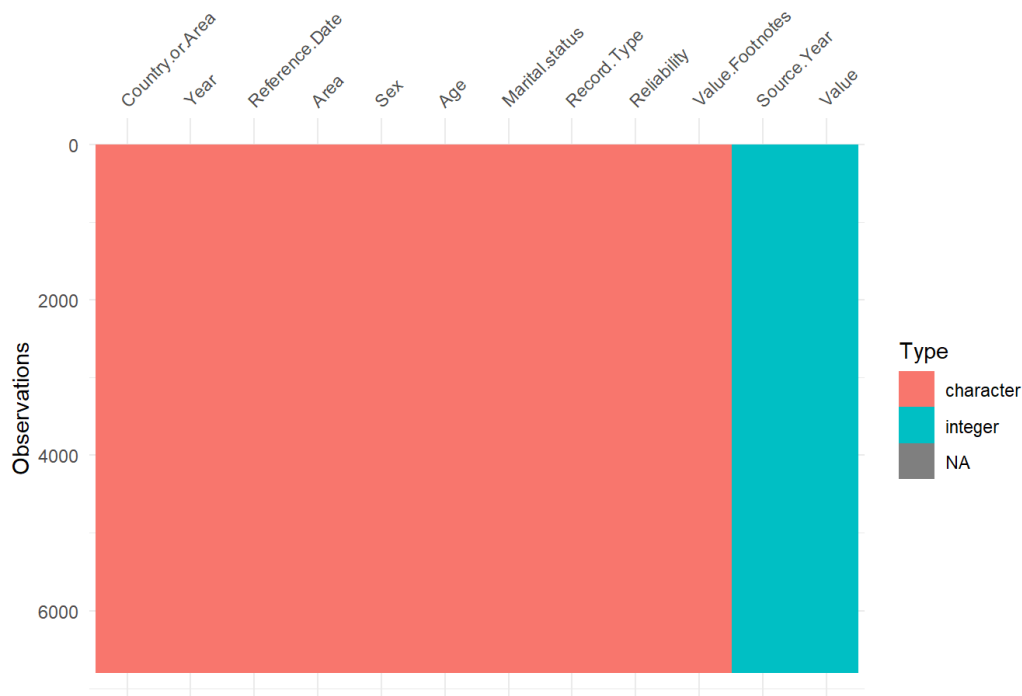
## "C:/Program Files/R/RStudio/bin/quarto/bin/pandoc" +RTS -K512m -RTS "C:/Users/owner/OneDrive/Documents/R/Master_Research/report.knit.md" --to html4 --from markdown+autolink_bare_uris+tex_math_single_backslash --output pandoc8c041b9b6c.html --lua-filter "C:/Users/owner/OneDrive/Documents/R/win-library/4.1/rmarkdown/rmarkdown/latex-div.lua" --self-contained --variable bs3=TRUE --standalone --section-divs --table-of-contents --toc-depth 6 --template "C:/Users/owner/OneDrive/Documents/R/win-library/4.1/rmarkdown/rmd/h/default.html" --no-highlight --variable highlightjs=1 --variable theme=yeti --mathjax --variable "mathjax-url=https://mathjax.rstudio.com/latest/MathJax.js?config=TeX-AMS-MML_HTMLorMML" --include-in-header "C:/Users/owner/AppData/Local/Temp/RtmpMDmvNb/rmarkdown-str8c044bfb46e6.html"

```

```

##
## Output created: report.html

```



# Results

# Discussion

# Conclusion

# References

<http://data.un.org/Data.aspx?d=GenderStat&f=inID:37&c=1,2,3,4,5,6&s=crEngName:asc,sgvEngName:asc,timeEngName:desc&v=1#GenderStat>

<https://unstats.un.org/unsd/demographic-social/products/dyb/dybcensusdata.cshtml>

<http://data.un.org/Data.aspx?d=POP&f=tableCode:323>

<http://data.un.org/Data.aspx?d=POP&f=tableCode:40>

<http://data.un.org/Data.aspx?d=POP&f=tableCode:330>

<http://data.un.org/Data.aspx?d=POP&f=tableCode:41>

<http://data.un.org/Data.aspx?d=POP&f=tableCode%3a325>

<http://data.un.org/Data.aspx?d=POP&f=tableCode%3a23>

<https://app.datacamp.com/learn/projects/166?open-modal=project-upgrade-modal>

# R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

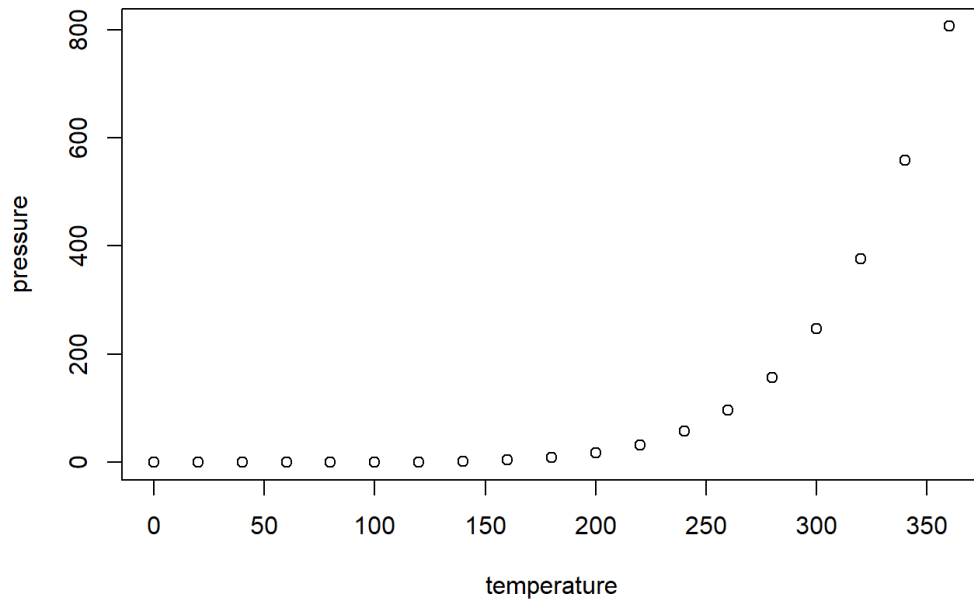
```
#We have created a framework application that you can use to test out different RMarkdown functions. Simply run the following code within the introDS package by using either

# RMarkdown Web
#introDS::runShiny('rmd')

# RMarkdown Mobile
#introDS::runShiny('rmd_mini')
```



You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

## References