**REJANE BERINGER**

What question did you investigate and explain how it pertains to the project’s goals?

How the Covid 19 hit the different countries’ population differently

What problems did you have while exploring & analyzing your data, how did you overcome them?

The fields that were blank and not having exactly the kind of data I was looking for

What did you find along the way that you didn’t expect?

I’ve got some unexpected but interesting results regarding to

COVID Deaths per million people and COVID Total Cases per Km2

What were your primary findings on the question you investigated?

The charts show clearly

Can any general conclusions be drawn from your analysis?

‘COVID Deaths per million people’ were represented clearly on a chart bar showing the discrepancy between countries, probably a reflex of their policy adopted against the pandemic. Sweden shows a high concentration of deaths per million, although it had only 5933 deaths this country has a total population of 10,099,270.

‘COVID Total Cases per Km2’ shows an unexpected concentration of people per area, once the virus is highly airborne transmissible, which had un effect especially in UK and Italy, smaller countries compared to US and Brazil.

Scatter plots - ‘COVID Cases vs. Median Age and ‘COVID Cases vs. Aged 65 Older’ both showed no relationship between age/number of deaths and age/number of cases.

Linear Regression - Total deaths per median age and aged 65 older. Low negative correlation and Insufficient evidence against Null hypothesis

Most Interesting Figure(s)?

Chart, bar chart

Description automatically generatedChart, histogram

Description automatically generatedChart, line chart, scatter chart

Description automatically generatedChart, line chart, scatter chart

Description automatically generated

If you had two more weeks, what would you have liked to expand upon or look deeper into?

I’d like to obtain more and specific data regarding to age ranges affected by pandemic

-------------------------------------------------------------------------------------------------------

Next: Combining Jupyter NoteBooks

Hypothesis: Countries with stricter mandates were more successful in limiting the spread and lethality of Covid-19.

Question 1:  Did countries who implemented mandates earlier on have greater success in curbing the spread and mortality rate of Covid-19?

Question 2:

Question 3:

Question 4:

**--------------------------------:Questions & Data:-------------------------------**

-   Elaborate on your questions, describing what kinds of data you need to answer them, and where you found it.

Question 1:  How the pandemic affected a country’s population in numbers?

Question 2: Is the population density relevant during a highly transmissible virus?

Question 3: Is there any relation between the number of cases and population’s age (median age or 65 older age)?

Question 4: Is there any relation between the number of deaths and population’s age (median age or 65 older age)?

**------------------:Data Cleanup & Exploration:----------------**

-   Describe the exploration and cleanup process

Starting from the file owid-covid-data.csv, I kept only the columns of interest(Location, Date, Total Cases, Total Deaths, Population, Population Density, Median Age (years), Aged 65 Older (%)), then the countries of interest were kept ("China", "Japan", "New Zealand", "United Kingdom", "Italy", "Sweden", "United States", "Brazil") and then only the last row of each with the most updated data. I saved to a clean csv file owid\_data.csv.

-   Discuss insights you had while exploring the data that you didn’t anticipate

-   Discuss any problems that arose after exploring the data, and how you resolved them

-   *Present and discuss interesting figures developed during exploration, w/ JupytNotebook*

**-----------------------------------:Data Analysis:-----------------------------------**

-   Discuss the steps you took to analyze the data and answer each question you asked in your proposal.

Outside of the dataframe, I calculated the mortality rate per million people and rate of cases per km2 of area, for every country. They were plotted bar and scatter charts and I also calculated the linear regression for Total deaths per median age and aged 65 older.

-   *Present and Discuss interesting figures developed during analysis, w/ Jupyt NoteBook*

**----------------------------------:Discussion:--------------------------------------**

-   Discuss your findings. Did you find what you expected to find? If not, why not? What inferences or general conclusions can you draw from your analysis?

(answered above under general conclusions)

**---------------------------:Post Mortem:--------------------------**

**-   Discuss any difficulties that arose, and how you dealt with them**

**-   Discuss any additional questions that came up, but didn’t have time to answer; What would you research next, if you had two more weeks?**