Data Mining Exercise

**Data mining is** the process of sorting through large amounts of data to identify patterns and establish relationships in order to solve problems through data analysis.

1. You will research the data that is available on the COVID 19 outbreaks in at least 4 different countries.
2. Using Google Docs or MS Word, use the data to answer these questions:
3. Where are the largest clusters of COVID 19 cases in each country?
4. Why do you think that is happening in those places?
5. Separate and count deaths in each country by age group (0-20yrs; 21-55 yrs; 55-65 yrs; 65+ yrs) If you see some different patterns, suggest those as well.
6. Compare the deaths in each age group using a bar graph to illustrate. (For each country, meaning 4 bar graphs)
7. Now, compare the bar graphs. Do they look similar or different?
8. Why do you think this is or is not?
9. Look at the timeline for the COVID 19 spread in these countries and illustrate using a line graph that shows the date on the x-axis (horizontal) and number of cases on the y-axis (vertical).
10. Do the 4 graphs look similar or different and give an explanation.