

Zainab Begum

Prof Rebin Muhammad

DATA 110

30 October 2024

Analysis Report

From the many datasets available, the decision to pick U.S Contagious Diseases was a quite easy one. Starting off with the importance of knowing different popular diseases and their trends over the years is extremely important for the future. Knowing the background of these diseases will increase the likelihood of being able to fight them. Being able to analyze this data sheds some light on the main problem, and when noticing these trends through the data, it will further improve prevention of these diseases.

The dataset includes all of the 50 states and District of Columbia, making it 51, and ranges from the 1930s to early 2000s. The six columns include: disease, state, year, weeks_reporting, count, and population. The columns I decided to focus on were disease, count population, and year because they seem to have a noticeable connection, worth noting down.

For my first graph, I wanted to plot each disease count over the years to see if any disease had an outlier, or a drastic change over the years. The one that stood out the most was Measles in 1935-1940 because there was a huge outlier where the measles count was over 120,000. This was greater than any other disease which spiked an interest for me. I researched how measles drastically went down after the 1960s and it was due to the vaccine finally coming out and people taking it. Projected by CDC, “Two doses of MMR vaccine are about 97% effective at preventing measles” explaining why the sudden drastic 80% decrease in cases (2024). Also, I was able to see this trend because I used a scatterplot, where I could clearly see a point farther

than the others. The scatterplot was the best plot to use rather than the others because with a bar graph it can be hard to pinpoint outliers because all the data is close together.

On to the second graph, I used a bar graph specifically because it is easier to line up with the y-axis to tell the approximate population. I also made the graph a bit smaller because the data is not complicated and very easy to understand, it is just population amount over the years. This graph was made to show how the population is growing at a rapid rate and this just puts into perspective how much we have to be careful in the future with these diseases. Which is why I used the color red to highlight that the population growing rapidly is not a good thing because a lot of people are not very easy listeners when it comes to diseases and viruses. For example, the Covid 19 cases were remarkably high due to our population also being high, since it was extremely contagious. People were told to wear masks and stay inside but not everyone did that which lead to even more cases until finally a decrease occurred after the vaccine was released. But even after the vaccine released, people were very hesitant to take it, and this caused a lot of deaths to happen. With the population only rising, I believe that people will not react well to newer diseases that come out.

The third and final graph includes disease counts over the years. I did not add the specific disease counts because it is not necessary to show the message of how much vaccines help with disease cases. In the graph you can see that the counts drastically decreased meaning that there was a cure eventually and people took it to survive and bring the counts down. I included this specific graph as a line plot because it does not need much extra information other than showing the major decrease in disease counts.

Although it may be comforting knowing that these specific disease counts are exceptionally low now in the 2000s, there are plenty of other diseases to come and it is particularly important to

remember these trends from previous popular diseases that had an outburst and got tamed from a vaccine. Vaccines and prevention techniques are very important for people to learn and implement in their daily lives. Some include, washing hands regularly, wearing a mask when there is a virus in your area, etc. Preventive actions recommended by the CDC should be taken seriously and it needs to be more common since the population will only be going up. So therefore, prevention is better than a cure.

References:

CDC (Measles):

<https://www.cdc.gov/measles/about/index.html#:~:text=Two%20doses%20of%20MMR%20vac,cine%20are%20about%2097%25%20effective%20at,dose%20is%20about%2093%25%20effective.>

Covid-19 (CDC): <https://www.cdc.gov/museum/timeline/covid19.html>