Simone van der Graaf
18 December 2023
Task 2.10 Data Immersion

Project Data Limitations and Metrics

Were there any limitations that prevented you from conducting an analysis?

The data from the population data set is from the US Census Bureau and because it is only done every 10 years, it has a risk that it is not up to date and therefore has limitations. For a future project a data set that is more up to date has less of these risks.

The data shows that for some counties one or more years are missing. Many duplicated have been found and removed, because the data is collected manually there is always a risk of duplicates and/or errors.

When there are 1 to 9 deaths counted per state and year in different age groups In the data set of influenza deaths, it's described as "Suppressed" instead of actual numbers given. Therefore, some of the data is not accurate and has been replaced with random numbers between 1 and 9. In the age groups of 65 years and older this almost did not occur and therefore the data on the most vulnerable age groups is accurate.

Did your data have any limitations that may have affected your results?

Using random numbers between 1 and 9 has affected the number of deaths for the other age groups as there is no certainty these random numbers are correct. In the future we need data to be more accurate.

How might you monitor the impact of the staffing changes you recommended?

Compare the influenza deaths of season 2018 with earlier seasons. If the states that received additional staff show a decrease in influenza deaths, this might be because of them.

Is there a metric that could be used for monitoring this impact?

In order to know whether it is the additional staff that is responsible for a decrease in influenza patients and influenza deaths, there should also be data available about how much staff there is available for influenza in each state, preferably in each hospital and clinic, so that influenza deaths and the number of staff can be compared. Also extra data on influenza deaths is necessary, such as the duration of having the flu for patients that become really ill and for the patients that die from it. Lastly more data should be available about medication. Is it the additional staff that is the solution, specific medication or the combination of the two?

Tableau storyboard:
https://public.tableau.com/app/profile/simone.van.der.graaf/viz/ProjectInfluenzaSeason-CareerFoundry/InfluenzaAnalysis

Oral presentation: https://www.youtube.com/watch?v=Or7NlAYdZ-s