Analysis of Influenza-Related Medical Staffing Needs in the USA for 2018: Interim Report

Project Overview

Motivation:
The United States has an influenza season where more people than usual suffer from the flu. Some people, particularly those in vulnerable populations, develop serious complications and end up in the hospital. Hospitals and clinics need additional staff to adequately treat these extra patients. The medical staffing agency provides this temporary staff.

Objective:
Determine when to send staff, and how many, to each state.

Scope:
The agency covers all hospitals in each of the 50 states of the United States, and the project will plan for the upcoming influenza season.

Research Hypothesis

If a state has a high percentage of their population in what the CDC class as 'vulnerable' age groups, then it should expect a high rate of patients going to hospital with Influenza Like Infections (ILI).

Data Overview

U.S. Census Bureau Population:
Includes county, state, year, total population, male and female population, and age categories starting with under age 5 through 85 years and older.

CDC Influenza Deaths:
Includes state, year, month, ten-year age groups, amount of deaths.

CDC Influenza Visits:
Includes state, year, week number, % weighted ILI, % unweighted ILI, age categories starting with under age 5 through 65 years and older, ILI total, num of providers, total patients.

Data Limitations

Timeliness:
The only consistent years over all 3 data sets is 2011 – 2017. Census data was collected in 2010, and all subsequent years up to 2017 is an estimate based on 2010s data. CDCs Influenza Visits data is collected weekly, and CDC Influenza Deaths is monthly while the Census Data is yearly.

U.S. Census Bureau Population:
There were errors and duplicates, 11% in total, in the data set which, when removed left some gaps in some states. These gaps were not deemed to be a big issue to the overall data set.

CDC Influenza Deaths:
The data is collected on a voluntary basis, from a limited collection of hospitals and other providers. This means that it is very likely there were more deaths due to Influenza Like Infections, but we cannot say what percentage is missing.
'Not Stated' is one of the options for the Ten-Year Age groups, making those particular entries of no use to us, as we cannot deduce if they are in the vulnerable age groups.
81% of the entries are 'Suppressed', meaning there was less than 10 deaths in that state/month/age group, but with no way to know the right number.

CDC Influenza Visits:
The data is collected on a voluntary basis, from a limited collection of hospitals and other providers. This means that it is very likely there were more visits due to Influenza Like Infections, but we cannot say what percentage is missing.
The Age Groups have no actual data, only the ILI Total. Therefore, we don't know the amount of people from the vulnerable groups that visited.
Florida State did not provide any data in this data set.

Descriptive Analysis

Columns are calculated by State and Year

	% of Deaths, due to ILI, of people aged 65+	% of Population aged over 65 years	% of Visits, due to ILI *
Average	95.23%	14.09%	1.68%
Standard Deviation	5.18%	1.65%	0.91%
Outliers	0	23	12
% that are Outliers	0.00%	7.01%	3.74%

*Florida has not given us any data in relation to any visits.
Column 1 shows that, on average, 95% of deaths due to Influenza Like Infections are in people aged over 65.
Column 2 shows that, on average, States have an average of 14.09% of their population over 65. The max is 19% and the minimum is 9%.

Column 3 shows that, on average, 1.68% of visits per year to our medical partners, that have given us data, is due to Influenza Like Infections.

Results and Insights

Null Hypothesis:
The number of deaths in the 65+ age group, has no correlation to the population aged over 65 in a state.

Alternative Hypothesis:
If there is a high population of people aged 65+ in a state, then there will be more deaths due to Influenza Like Infections.

Result:
I ran a one-tail test, with a significance of 0.5, comparing the number of Deaths, from Influenza Like Infections, of people over 65, with the Census Population of people aged over 65. The results, the P-Value was below the Significance level, allowed me to reject the Null Hypothesis. We now have statistical evidence that states with more people over the age of 65 will have more deaths due to Influenza Like Infections. Therefore, we should provide more staff to states that have a higher proportion of citizens aged 65 and over.

Remaining Analysis and Next Steps

- Assess when influenza season occurs and investigate any shifts or patterns in recent years and/or by state including visualization.
- Analyze health care provider visits with Influenza Like Infections in most recent years by US state including visualization.
- Provide forecast for staffing plan during upcoming influenza season by highlighting large vulnerable populations (categorize states by level of need).
- Create a presentation with visualizations and recommendations.

Appendix

Project Overview:
With our current datasets we should be able to work out the data by month, and state, allowing us to see when the Influenza season will be in 2018. We will also be able to confirm which states will need more staff during that time.

Data Overview:
Due to CDC Influenza Deaths having no information for deaths of children under 5, and CDC Influenza Visits having no data for each specific age group, we will have to focus on the 65+ age group, in order to keep with our hypothesis surrounding 'vulnerable populations'.