Forecasting US Flu Season Staffing Requirements: Interim Report

Prepared by: Jay Haygood

October 28th, 2021

Project Overview

Motivation: The United States has an influenza season where more people than usual suffer from the flu. Some people develop serious complications and end up in the hospital. Hospitals and clinics need additional staff to adequately treat these extra patients.

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 given state has an increasing rate of flu deaths each year, relative to the population, then these states would see greater challenges with understaffing during the peak of flu season.” ~ In accordance with the project requirements, the hypothesis will also require further analysis to determine which states have higher ratios of vulnerable population, specifically those aged 55 and older.

Data Overview

CDC Influenza Deaths: Shows annual flu mortality for each US state, as provided by the CDC. Integrated with US Census data, variables include:

- US State & Year (2009-2017)
- Age groups, in 10-year increments
- Total flu deaths

US Census Population: Provides estimated counts of US population by state. Since launch of the American Community Survey, collection methods for the Census have evolved into a combination of survey and administrative data. The existing data set, integrated with Influenza Deaths above, currently includes:

- US State & Year (2009-2017)
- Total population counts
- Population per gender (Male, Female)
- Age groups, in 10-year increments

CDC Influenza Visits: Tracks hospital visits for influenza-like illnesses. Data is sourced by the CDC and related providers (Sentinel, FluView), through multiple US healthcare providers, and clinical laboratories. Included variables with applicable data:

- Region (State)
- Year (2010-2019)
- Patient #'s w/Illness
- Total Patients

Data Limitations

CDC Influenza Deaths: This data was collected from Jan, 2009, through Dec, 2017. Ideally, to plan for the upcoming flu season, the most current data would be needed. However, the available data is applicable to identify historical trends and plan for subsequent flu seasons.

US Census Population: Census surveys for the population records were likely collected in 2009 or 2010, with subsequent annual estimates, the most recent being 2017. While census estimates are generally accurate based on past records, census data collected in 2020 would be more ideal. However, using flu death data only through 2017, the existing census data should be timely for the project.

CDC Influenza Visits: Possible bias at the reporting level, and while the most recent records are more current as of 2019, there is no available data to define specific age groups.

Descriptive Analysis

Initial analyses performed establishes predicted correlation with greater population counts over age 55, and higher flu mortality rate.

Data Variable:	Flu Deaths, 55-64 years	Census, Total population
Standard Deviation:	101	6743429
Mean (Average):	57	5924386
Outliers:	67	41
Outlier Percentage:	14%	9%

This indicates that with higher populations counts, also comes an increase in flu deaths. I interpret the correlation as one of potentially many examples which will support the stated hypothesis.

Results and Insights

Statistical hypothesis: If a given state shows a growing population count of residents 55 and older each year, then these states would also see a higher annual flu mortality rate.

Null hypothesis: Flu mortality rates are either unaffected or reduced with increasing 55+ population.

Alternative hypothesis: States with an increasing population age 55+ will also have higher flu mortality.

With p-value result 3.0012, null hypothesis is rejected, as significance level is notably lower (alpha = 0.05.) In simple terms, this result indicates a definite correlation between higher flu mortality, and larger population counts age 55 and older.

Remaining Analysis and Next Steps

- Further analysis to identify which states, if any have a greater flu mortality to population ratio, and any correlation to hospital understaffing.
- Research needed to determine which states have faced previous challenges with understaffing, during past flu seasons.
- Perform more in-depth analyses, which will further support the research hypothesis and project requirements, including ***Data Visualization, Design Basics & Tableau, Charts, Visualizations & Forecasting, Histograms & Box Plots.***
- Additional visual analyses with ***Scatter Plots & Bubble Charts, Spatial & Textual Analysis.***
- Prepare final video presentation, and ***Present Findings to Stakeholders.***

Appendix

Regarding Census data:
"Census data is collected at regular intervals using methodologies such as total counts, sample surveys, and administrative records. After it is collected or generated, census data is summarized to represent counts or estimates of groups of people for different geographic areas." What is Census Data? - Social Science Space

There is a potential time lag factor:
"The CDC/ATSDR SVI is updated every two years based on U.S. Census Bureau data releases. The Census releases American Community Survey data in December of the year following the Survey. Thus, there is a time differential in when we can produce and disseminate updates to the CDC/ATSDR SVI." CDC/ATSDR SVI Frequently Asked Questions (FAQ) | Place and Health | ATSDR

Regarding CDC data, and possible bias at the reporting level:
"Many published reports on test performance, often funded by diagnostics companies, are prone to bias… For the 96 studies analyzed, compliance was <25% for 14/34 STARD 2015 standards, and 3/7 QUADAS-2 domains showed a high risk of bias. All reports lacked reporting of at least one criterion. These biases should be considered in the interpretation of study results." Risk of bias and limits of reporting in diagnostic accuracy studies for commercial point-of-care tests for respiratory pathogens - PubMed (nih.gov)

National, Regional, and State Level Outpatient Illness and Viral Surveillance (cdc.gov)