Preparing for Influenza Season:

Interim report

"Don't let algorithms decide,

what news you see"

Mother Jones Magazine

Zoya Matsiy

Careerfoundry Course 2023

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. 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.

HYPOTHESIS

Influenza infections are associated with thousands of deaths every year in the United States, with the majority of deaths from seasonal influenza occurring among adults aged ≥65 years [1]. Other research gives the mean age at death, as 72.7 years [2]. If the population of any region has a larger group of people 65+, then the influenza mortality will be higher and there will be more additional stuff necessary.

DATA OVERVIEW

The analysis is based on 2 datasets: the CDC report on influenza deaths and the CDC Census Population report.
The CDC report on influenza deaths includes data on influenza-related deaths specified by geography, time, age, and gender for the period 2009-2017. The data is external and public and it is trustworthy as automatically and frequently collected administrative data. It's relevant even though some not significant errors and biases are presented.
The Population Census data set represents data of population by country, by year (2009-2017), by age groups (11 groups), and by gender. The data is external and public and is automatically collected. There are 11,43% missing data presented. The missing data were imputed duplicating the nearest available by region/year/age group data. Therefore, the dataset is trustworthy and relevant.

DATA LIMITATIONS

However, a diagnosis of influenza virus infection often is not confirmed with sensitive and specific laboratory diagnostics, particularly among older persons, and even when identified is rarely recorded on death certificates. Many deaths associated with influenza infections occur from secondary infections such as bacterial pneumonia or complications of chronic conditions such as congestive heart failure and chronic obstructive pulmonary disease. Therefore, estimates using underlying respiratory and circulatory mortality data (which include pneumonia and influenza causes) can provide an upper bound for influenza-associated deaths [1].
Furthermore, the data reflect actual changes in demographic characteristics, healthcare access, patient management, and increased influenza vaccination coverage in populations of the last 10 years. The average data we use for stuff distribution planning by region can be more or less than the forecast based on the newest data only.

DESCRIPTIVE ANALYSIS

The available dataset has important characteristics to take into consideration. Some of them are shown in Table 1.
The Population mean represents the average population in certain ages over the years and states. The Death mean represents the average number of influenza deaths for certain age groups over the years and states. The data shows the Population mean grows by half and the Deaths mean decreases by half by moving to the older group.
Life acceptance in the USA is about 79 years in under review period [3] and we can expect a higher number of Deaths in the correspondent age group. But the data on influenza Deaths don't confirm this

assumption (see **Table 1**): shows linear progression in Deaths/Population and the level of outliners occurrence is mostly near the allowed level of 5%.

Table 1. DATASPREAD

	Population 65-74	***Death 65-74***	***Population 75-84***	***Death 75-84***	***Population 85+***	***Death 85+***
Variance **D**	233,927,030,308	27912	80,018,449,159	92838	14,884,069,343	301474
Standard Deviation **σ**	483,660	167	282,875	305	122,000	549
Mean **μ**	454,347	114	261,785	242	110,217	455
Outlier Percentage	8%	6%	10%	5%	10%	5%

According to data test results (see **Table 2)**, influenza Deaths have a very strong (94%) correlation to the population age group size. That is the data confirm that influenza-related Deaths are the age-dependent parameter in groups over 65.

Table2. Data Correlation

Variables	Population 65-74	Death 65-74	Population 75-84	Death 75-84	Population 85+	Death 85+
Proposed Relationship	linear		linear		linear	
Correlation Coefficient	0.94		0.94		0.94	
Strength of Correlation	strong		strong		strong	

RESULTS AND INSIGHT

The goal of the current data analysis is to create some directions for a staffing plan that utilizes all available agency staff per state requirements, without necessitating additional resources.

It would be rational to distribute the stuff in proportion to the number of population groups over 65 years of age by state. To do this it's necessary to find out whether the hypothesis of the highest number of Influenza Deaths by groups over 65 is statistically confirmed. This analysis was carried out by conducting of t-test for 2 groups of the population: the age group 0-64 and the age group over 65.

The Null Hypothesis: Number of Deaths under 65 =Number of Deaths over65

The Alternative Hypothesis: Number of Deaths under 65 < Number of Deaths over 65

Figure: 1. T-test

t-Test: Two-Sample Assuming Unequal Variances

	Death >65	*Death < 65*
Mean	810.3974359	77.25
Variance	1021567.769	22579.77248
Observations	468	468
Hypothesized Mean Difference	0	
df	488	
t Stat	15.5214865	
P(T<=t) one-tail	9.46581E-45	
t Critical one-tail	1.647982077	
P(T<=t) two-tail	1.89316E-44	
t Critical two-tail	1.96483707	

The result is P =0,0% (one-tailed test) which is much less than the significance level of 5%. At a confidence level of 5 percent, it is found a significant difference in the number of Deaths (dependent variable) between 2 Age Groups in the USA population: under 65 and over 65. Therefore, the working Hypothesis is statistically confirmed.

REMAINING ANALYSIS AND NEXT STEPS

There are the parameters to include in the analysis as the according datasets will be available: data about the different influenza virus types on the population age group's Mortality, about the more vulnerable parts inside of the over 65 age population group, and their participation in flu shorts events, data about the flu seasons over the years.

To include is also the data about the available medical staff and historical data about staff illness during the influenza seasons.

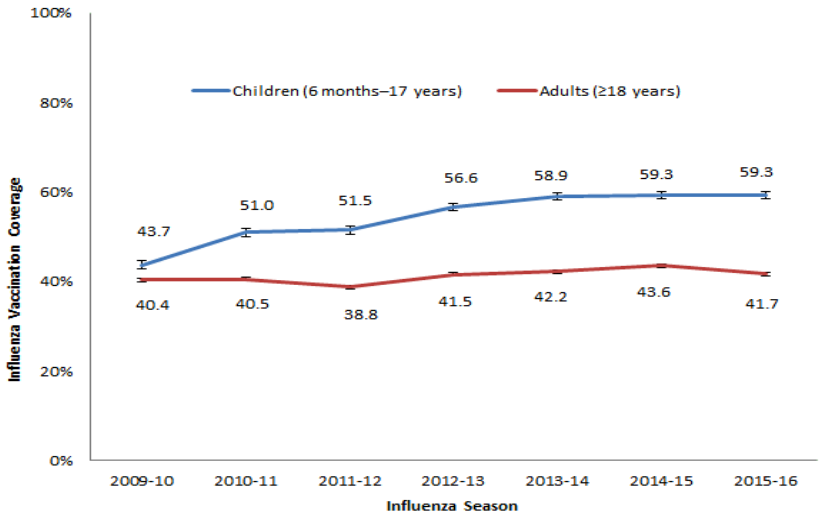
The future step of the current analysis should be a list of regions to their vulnerability and according to the country map, tableau storyboard featuring visualizations of analysis, and actual PowerPoint Presentation.

APPENDIX

The dataset presents population data on illness and mortality tendencies which are very dangerous and they are valid not only for the patients but also for the staff of the agency. For both categories, vaccination became of big importance.

The vaccination level is only 50%-60%. It is presented in Figure 3. [7]

Figure 1. Seasonal Flu Vaccination Coverage by Age Group and Season, United States, 2009–2016



We need to conduct another analysis to see if understaffing and overstuffing in some regions emerged because of influenza illness by staff.

ADDITIONAL DATA SOURCES

[1] CDCP "Morbidity and Mortality Weekly Report"
https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5933a1.htm?s_cid=mm5933a1_e%0d%0a#t
[2] "Mortality Associated with Influenza and Respiratory Syncytial Virus in the US, 1999-2018" Chelsea L. Hansen, MPH; Sandra S. Chaves, MD, MSc; Clarisse Demont, Ph.D.; Cécile Viboud, Ph.D.
https://jamanetwork.com/journals/jamanetworkopen/article-abstract/2789446
[3] U.S. Life Expectancy 1950-2023
https://www.macrotrends.net/countries/USA/united-states/life-expectancy

[4] CDC “People at Higher Risk of Flu Complications”
https://www.cdc.gov/flu/highrisk/index.htm
[5] CDC “Past Flu Seasons” https://www.cdc.gov/flu/season/past-flu-seasons.htm
[6] United Nation, WPP “United States of America, Demographic Profiles”
https://population.un.org/wpp/Graphs/DemographicProfiles/Line/840
[7] CNBC “People of color face higher risk of hospitalization in the US”
https://www.cnbc.com/2022/10/18/people-of-color-face-higher-risk-of-flu-hospitalization-as-us-faces-severe-season.html
[8]CDC Flu Vaccination Coverage, United States, 2015-16 Influenza Seas
https://www.cdc.gov/flu/fluvaxview/coverage-1516estimates.htm