Data Analytics Immersion
CareerFoundry
Achievement 1: Final Exercise

Preparing for Influenza Season

Interim Report

Poojaben Thummar
May 2024

Index

PROJECT OVERVIEW

- Goal: Enhance preparedness for the influenza season by effectively managing staffing needs at clinics and hospitals served by the medical staffing agency.

- 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 scope of the analysis encompasses examining historical influenza data, including incidence rates, peak seasons, and regional variations, to forecast staffing needs across the country. It involves identifying patterns, seasonal fluctuations, and high-risk periods for influenza outbreaks. Additionally, the analysis will assess the implications of influenza trends on the demand for temporary healthcare workers, considering factors such as clinic size, patient demographics, and geographic location.

- Stakeholder Identification:
  - Medical agency frontline staff (nurses, physician assistants, and doctors)
  - Hospitals and clinics using the staffing agency services
  - Influenza patients
  - Staffing agency administrators

HYPOTHESIS

- If the population data by age are analyzed relating to mortality rates, then the older individuals over 65 years will exhibit a higher mortality rate, which emphasizes the significance to focus on the elderly population during the flu season.

DATA OVERVIEW

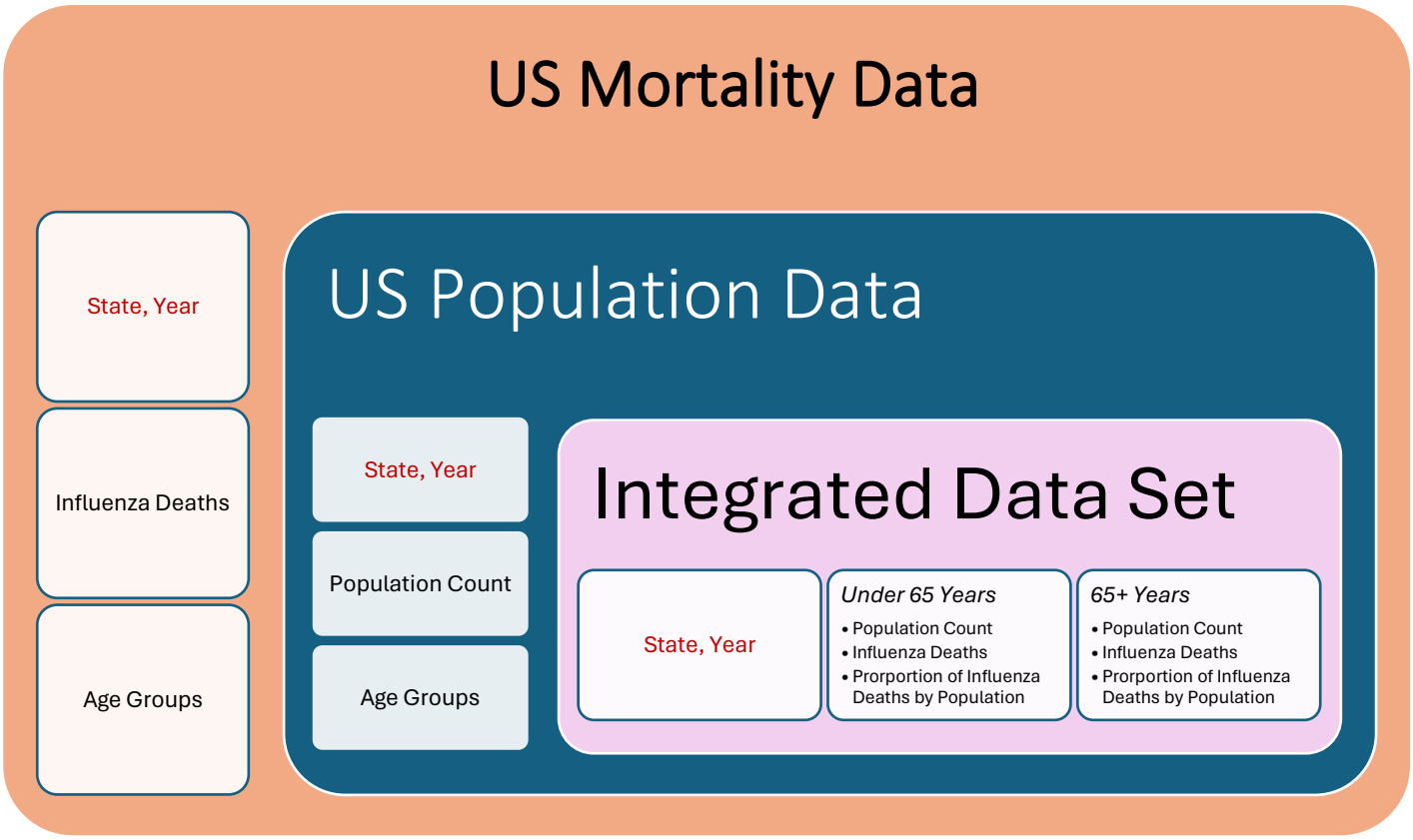
Overview & Limitations

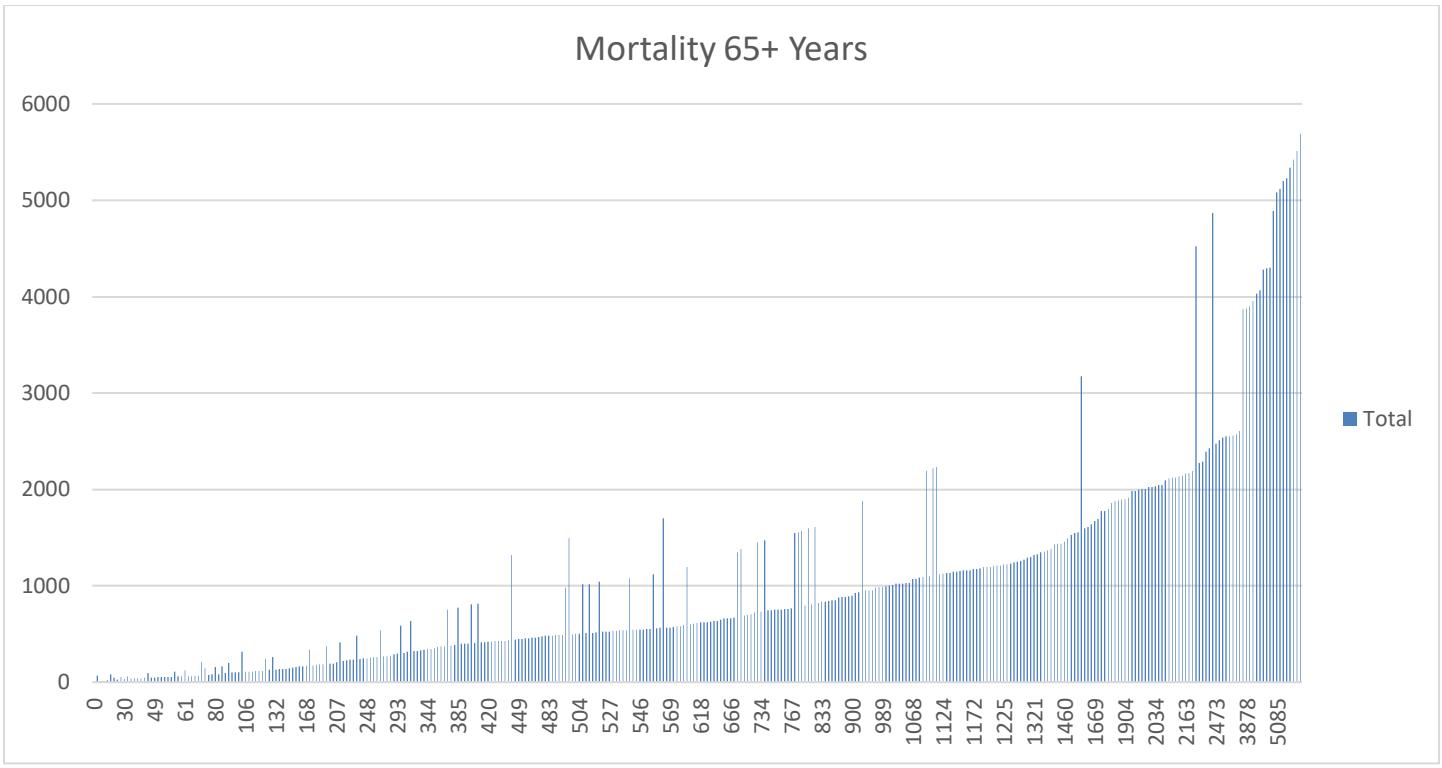
Data set: US Mortality Data	
Summary	The Centers for disease Control and Prevention (CDC)provides this External data set, contains influenza death counts for different age groups in each state for the years 2009-2017.
Limitations	The data source is based on death certificates in which providers report only a single underlying cause of death, which could lead to masking and inaccurate reporting of patients with pre-existing conditions.

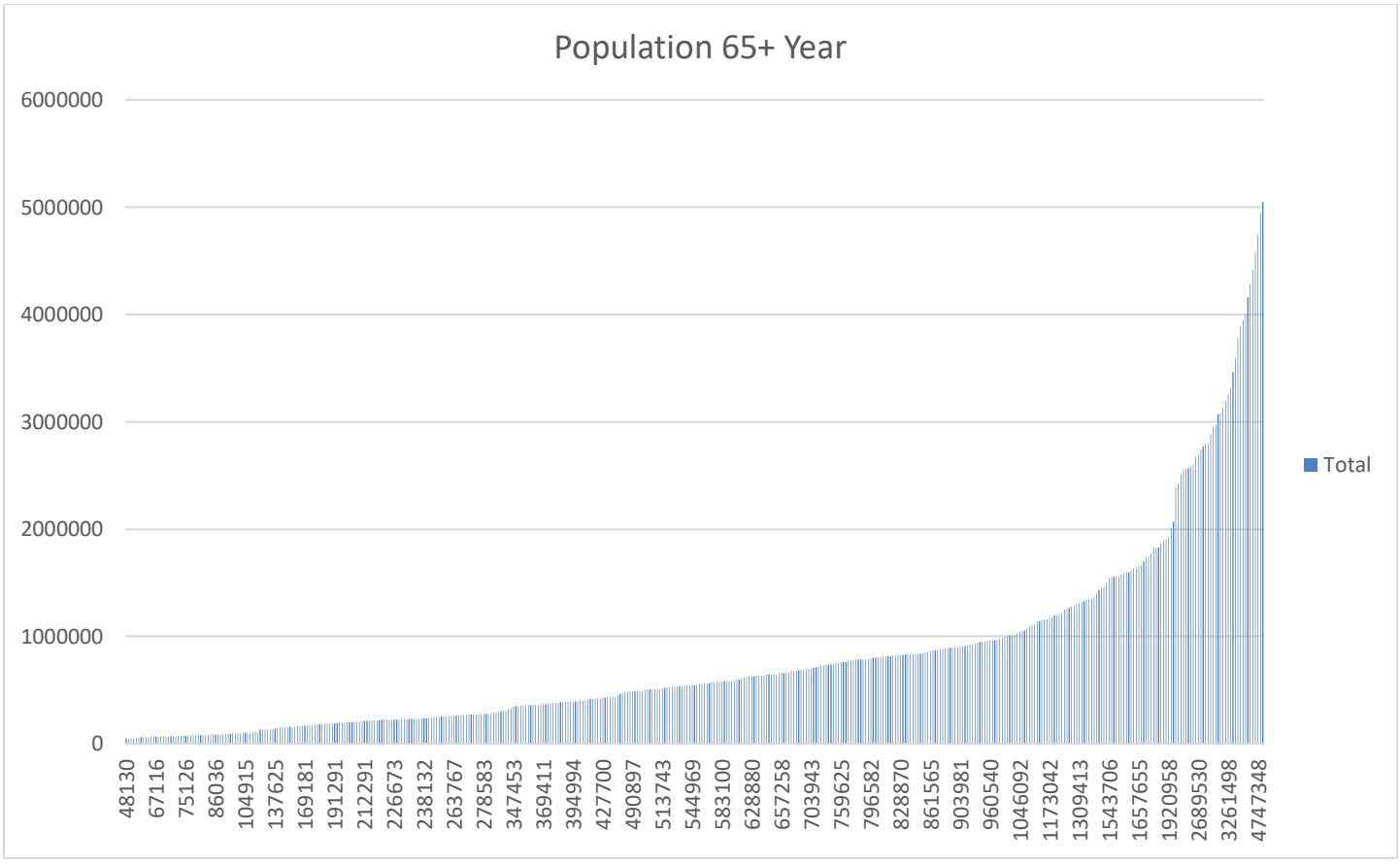
Data set: US Population Data	
Summary	The US Census provides this survey data. It contains demographic data for states including gender and age groups in the US from 2009 to 2017.
Limitations	As survey data this will not be representative of the entire US population and will only serve as estimates of US population in each state. Another limitation is that this data is not collected frequently which may create some gaps in the data.

The Integrated Data Set

- The influenza Mortality and US population data sets were combined using the concatenated variable State-Year as a key id [Combined key].
- The age groups were combined into 2 groups:
    - o High Risk (65+ Years of Age)
    - o Low Risk (Under 65 Year)




Mortality 65+ Years
6000
5000
4000
3000
2000
1000
0
0
30
49
61
80
106
132
168
207
248
293
344
385
420
449
483
504
527
546
569
618
666
734
767
833
900
989
1068
1124
1172
1225
1321
1460
1669
1904
2034
2163
2473
3878
5085
Total


Population 65+ Year
6000000
5000000
4000000
3000000
2000000
1000000
0
48130
67116
75126
86036
104915
137625
169181
191291
212291
226673
238132
263767
278583
347453
369411
394994
427700
490897
513743
544969
583100
628880
657258
703943
759625
796582
828870
861565
903981
960540
1046092
1173042
1309413
1543706
1657655
1920958
2689530
3261498
4747348
Total

DESCRIPTIVE ANALYSIS

- The research hypothesis proposes that a higher population in the 65+ age group will result in an overall higher number of influenza deaths in the same age group.
- A descriptive analysis was conducted to assess the quality of the data and to understand the similarity and variance between the variables. Therefore, the mean and standard deviation were calculated for both the number of influenza deaths over 65 years and the population over 65 years.

	Mortality 65+ Years	Population 65+ years
Mean	826	806566
Standard Deviation	1014	880783

	Mortality 65+ Years	Population 65+ Years
Correlation Coefficient	0.94	0.29
Strength of Correlation	Strong relationship	Weak relationship

The correlation coefficient weakly supports the hypothesis that individuals aged 65 and above are more prone to severe influenza complications and a higher risk of death. Instead, factors like vaccination rates, access to healthcare, general health, and various demographic and environmental elements seem to influence influenza deaths in this age group.

RESULTS AND INSIGHTS

To obtain further insight from the sample data sets, inferential statistics were used to translate the research hypothesis into conclusions about the population as a whole.

- Null Hypothesis: The influenza death rate of patients 65 years or older is less than or equal to the influenza death rate of patients younger than 65 years old.

- Alternative Hypothesis: The influenza death rate of patients 65 years or older is greater than the influenza death rate of patients younger than 65 years old.

t-Test: Two-Sample Assuming Unequal Variances		
	0-64	**65+**
Mean	0.00079739%	0.0847813%
Variance	1.01026E-10	1.91276E-07
Observations	459	459
Hypothesized Mean Difference	0	
df	458	
t Stat	-41.12990387	
P(T<=t) one-tail	3.525E-156	
t Critical one-tail	1.648187415	
P(T<=t) two-tail	7.0493E-156	
t Critical two-tail	1.965157098	

REMAINING ANALYSIS AND NEXT STEPS

- After confirming that high risk populations (65+ years of age) do experience higher influenza deaths, our analysis will begin to look for states where populations older than 65 years of age is higher.
- The analysis should also begin to look at influenza patterns found in other data sets to predict how influenza impacts different states at different times of the year to accurately predict and allocate medical resources as needed.