PREPARATION INFLUENZA SEASON:

Interim Report

Submitted by: Janvi Amin

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 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

- States with a greater proportion of vulnerable populations (adults over 65 years old) are likely to experience higher mortality rate from Influenza compared to states with a lower proportion of vulnerable populations.

DATA OVERVIEW

Influenza Deaths:

The Centers for Disease Control and Prevention (CDC) offers an external dataset that includes influenza death counts categorized by age groups for each state, spanning the years 2009 to 2017.

US Census Bureau:

The Census Bureau data offers state-wise population information for different age groups, including variables such as state and year, covering the period from 2009 to 2017.

DATA LIMITATION

Influenza Deaths (CDC):

The influenza death dataset lacks detailed information on all age groups of patients and does not account for co-morbidities. Consequently, the underlying causes of death are likely more complex than just influenza, potentially introducing bias into the data.

Population data (US Census Bureau):

As survey data, this information will not represent the entire U.S. population but will only provide estimates for each state. Additionally, the infrequent collection of this data may result in some gaps.

DESCRIPTIVE ANALYSIS

Data Spread		
Variable	**Vulnerable Population**	**Total Death of Vulnerable Population**
Dataset Name	Integrated Data Set	Integrated Data Set
Sample or Population?	Sample	Sample
Normal Distribution?	Left-Skewed	Left-Skewed
Variance	9458520694232	1232029
Standard Deviation	3075471	1110
Mean	3970232	826
Outlier lower bound	-2180710	-1394
outlier upper bound	10121173	3046
Outlier count	16	21
Outlier Percentage	3%	6%

DESCRIPTIVE ANALYSIS I

Correlation		
Variable	Total Vulnerable Population & Total Death of Vulnerable Population	Total Population & Total Death
Proposed Relationship	To test the relationship between total vulnerable population and total death of the vulnerable population due to Influenza.	To test the relationship between total population and total death of population due to Influenza.
Correlation Coefficient	0.37	0.96
Strength of Correlation	weak relationship	Strong relationship
Usefulness / Interpretation	So far it looks like the indanger age group is not as much correlated to number of deaths we may need further analysis to prove that.	As the population grows, so does the number of deaths caused by Influenza. This suggests a correlation between Influenza deaths and population size, regardless of age group. Therefore, further investigation is necessary to determine which age group contributes the most to these deaths.

DESCRIPTIVE ANALYSIS II

- A descriptive analysis was performed to evaluate the data quality and understand the similarities and variances between the variables. Consequently, the mean and standard deviation were calculated for both the number of influenza deaths in individuals over 65 years and the population over 65 years.
- The correlation provides weak support for the hypothesis that individuals aged 65 and above are more prone to severe influenza complications and a higher risk of death.

RESULTS AND INSIGHTS

Research hypothesis	Patient over 65 or more is more likely to die.
Independent variable	Age group
Dependent variable	Influenza death rate
Null hypothesis (H_0)	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 (H_A)	The influenza death rate of patients 65 years or older is greater than the influenza death rate of patients younger than 65 years old.
Two-tailed or one-tailed test	This is a one-tailed test because we are only interested in influenza death rates of one direction, being patients 65 years or older.
Significance level	0.05
p-value (one-tail)	The t-test comparing the influenza death rates between the 0-64 years age group and the 65+ years age group resulted in a one-tailed p-value of 1.0838E-180, which is effectively zero. This p-value is significantly lower than the threshold of 0.05, indicating a statistically significant difference in the mean death rates between these two age groups. Given that the mean death rate for the 65+ years age group is higher (0.09199%) compared to the 0-64 years age group (0.00087%), we can reject the null hypothesis and accept the alternative hypothesis.
Additional analysis	Since the t-test showed a significant statistical difference in influenza death rates between the 65+ years age group and the 0-64 years age group, it may be beneficial to analyze the differences across other 10-year age groups to identify specific vulnerable populations. Additionally, understanding the underlying causes of this influenza death rate difference can help identify risk factors in individuals. Knowing these risk factors enables us to predict which patients are vulnerable and allows hospitals to staff accordingly.

RESULTS AND INSIGHTS

t-Test: Two-Sample Assuming Unequal Variances		
	0-64 yrs	65+ years
Mean	0.00086950%	0.09199440%
Variance	1.04541E-10	1.39089E-07
Observations	423	423
Hypothesized Mean Difference	0	
df	423	
t Stat (t-score)	-50.23403534	
P(T<=t) one-tail	1.0838E-180	
t Critical one-tail	1.648463868	
P(T<=t) two-tail	2.1675E-180	
t Critical two-tail	1.965587999	

REMAINING ANALYSIS AND NEXT STEP

- After confirming that high-risk populations (those aged 65 and older) experience higher influenza deaths, our analysis will focus on identifying states with a higher proportion of populations over 65 years of age.