Preparing for Influenza Season: 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.

Hypothesis

People over age 65+ years have a much higher risk of complications and fatality than normal If the patient is 65+ years then, there is a much higher risk of complications and fatality.

Data Overview

US Census Data: The data was collected from the US Government website, so its external data. the US Census Bureau owns the data. since the data is collected from Government website Its trustworthy. Data is collected both manually and automatically (from the census survey and by the data integration from other government records.). This data set contains total population, male/female populations, and populations of several age ranges and the data is broken down by county. The data spans from 2009 to 2017.

CDC Influenza Deaths: The CDC Influenza Deaths data was collected from 100 public health providers and over 300 clinical laboratories located throughout the United States and its territories. Its external data owned by the CDC. It contains the number of Deaths by state, Year and age group.

Limitations

US Census Data: Data is collected both manually and automatically (from the census survey and by the data integration from other government records.) so there can be error in manually collected data. As the data is collected after ten years, there will be a time lag, and this can lead to wrong analysis.

CDC Influenza Deaths: There are so many supressed values (more than 5%) in deaths column, and this may affect the analysis.

Descriptive Aanalysis

The mean, variance, standard deviation and correlation coefficient values were calculated to understand and the data and draw correlation between the variables as below.

	Population - 65Years and over	Deaths - 65 Years and over
Variance	786688471011	621674
Standard Deviation	886954.6048	788.4629381
Mean	809089.7379	1305.947598
Two standard deviation range	-964819.4718	-270.9782779
	2582998.947	2882.873474
Count of Outliers	29	18
%of outliers	6%	4%

Correlation Coefficient	0.850012812

The correlation coefficient is o.85, which shows a strong relationship between variables Population and Deaths. As the proportion of 65+ Years age population over total over total increasing, deaths are increasing. % outliers for population over 65 years and over age group and Deaths over 65 years and over are 6% and 4% respectively.

statistical hypothesis and interpretation

Null Hypothesis: Death rate among age group under 65 ≥ death rate among population over 65+ year age group.

Alternative hypothesis: Death rate among age group under 65 < death rate among population over 65+ year age group.

t-Test: Two-Sample Assuming Unequal Variances

	Deaths-Under 65 Years	*Daeths-65 Years and above*
Mean	3192.08952	1305.947598
Variance	9515.801597	621673.8047
Observations	458	458
Hypothesized Mean Difference	0	
df	471	
t Stat	50.80742531	
P(T<=t) one-tail	1.46E-193	
t Critical one-tail	1.648095217	
P(T<=t) two-tail	2.9172E-193	
t Critical two-tail	1.965013401	

With a p-value of 1.46E-193, which is less than the significance level of 0.05, we can reject the null hypothesis. This means that the alternative hypothesis is true, and we can conclude that the risk of death during the Influenza season is higher for people who are 65 years or older than those who are less than 65 years old.

Next Steps

- Determine the states and counties having higher number of 65+ years aged people.
- Aallocate staff for clinics and hospitals in each state based on the vulnerable (over 65 years) population of people.
- Continue doing analysis for other vulnerable group (under age 5 years)

- Prepare and present final presentation to stakeholders.

Appendix

CDC
CDC(Fluview)