



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.

Research Hypothesis:

- Elderly people are part of the vulnerable population. If you are over 65 then the influenza mortality rate will increase.

Data Overview:

Population Death Data:

- This data shows the US population from 2009 to 2017 broken down by county, gender, and age groups by every 5 years.

Influenza Deaths Data:

- The data contains monthly death counts for influenza-related deaths in the United States from 2009 to 2017. Counts are broken into two categories: state and age.

Data Limitations:

Population Death Data:

- **Manual errors:** Only one person per household fills out the census by hand. This can lead to manual errors and interpreting question differently than how they were written.
- **Timeliness:** Data is only from 2009-2017. The data does not include the effects of Covid-19 after 2020. This effect would decrease the population especially in the vulnerable population of under 5 and over 65 years old.

Influenza Death Data:

- **Insufficient Data:** The deaths on a death certificate only list one cause of death. This could cause issues to the elderly and vulnerable population as the cause of death may be related to another health issue such as AIDs, diabetes, or asthma. The decline in health may have been initiated by influenza.
- **Bias:** Only being able to list one cause of death on a death certificate can lead to bias from the providers. Providers may have incentive to put influenza as the cause of death to gain awareness of influenza or get research funding. If the government is concerned about the increase of influenza deaths, there could be pressure on providers to either provide an increase or decrease in those numbers.

Descriptive Analysis:

Variable	Mean	Stand Deviation
under 65 population	5,278,879	5,973,747
over 65 population	829,430	892,630

under 65 deaths
over 65 deaths

79	151
826	1,014

The correlation of the variables between the 65 population and over 65 deaths is 0.94. This is a strong positive relationship between the two variables. This correlation supports the hypothesis that as age increases death also increases.

Results and Insights:

Null Hypothesis:

- Flu deaths in population under 65 is equal to or greater than the population of 65+

Alternative Hypothesis:

- Flu deaths in population under 65 is less than the population of 65+

At an alpha of 0.05 or a confidence level of 95 percent, there's a significant difference in flu deaths and the population of over 65.

Analysis and Next Steps:

Discuss with stakeholders the significant difference between the groups of populations between people under 65 and over 65. We can use this analysis to see where to send staff in the upcoming influenza season. We can focus on the populations that have high elderly populations as they are most vulnerable.

Visuals- Create comparison charts, time forecast, statistical visualizations, and spatial visuals

Presentations- Create a visual storyboard and a video presentation to stakeholders

Appendix:

Project Brief

Hypothesis Development:

- **Questions asked:** When is flu season? Is flu season the same in every state? Is flu season only once a year? Which states are most affected by influenza? How long is flu season? Why are certain states more affected by influenza?
- **Article context:** Flu shots can prevent people from getting influenza, and having an adequate medical staff can prevent deaths during the influenza season. Each state is affected by influenza differently because of the different population composition. Some states have a highly vulnerable population compared to other states. The flu season all year around, however it is normally in the fall/ winter. The peak season is between December and February.

Profiles of data:

- **Population Death Data:**
 - Data source is external as it comes from the government and is available to the public. The US Census Bureau owns the data. Since this is run by the government this is trustworthy data. This data is administrative data. The US census is collected every 10 years, therefore there is a significant time lag in the data. A limitation to the data is that only one person per household fills out the census by hand. This can lead to manual errors and interpreting question differently than how they were written. This data shows the US population from 2009 to 2017 broken down by county, gender, and age groups by every 5 years. This data is relevant to the project as the objective is to know where to send the staff. This data will help to determine which states have a high or low population that needs more staff. The hypothesis also is broken down to look at the elderly population, which the data is grouped by age.
- **Influenza Death Data:**
 - The data source is external as it comes from the government agency of the CDC. The CDC owns the data. Since this is run by the government this is trustworthy. This data is administrative data. Each US state is required to record all birth and deaths records. Death records come from the providers that code the cause of death as influenza. The deaths on a death certificate only list one cause of death. This could cause issues to the elderly and vulnerable population as the cause of death may be related to another health issue such as AIDs, diabetes, or asthma. The decline in health may have been initiated by influenza. Only being able to list one cause of death on a death certificate can lead to bias from the providers. Providers may have incentive to put influenza as the cause of death to

gain awareness of influenza or get research funding. If the government is concerned about the increase of influenza deaths, there could be pressure on providers to either provide an increase or decrease in those numbers. The data contains monthly death counts for influenza-related deaths in the United States from 2009 to 2017. Counts are broken into two categories: state and age. This data will show the deaths data by age needed to test our hypothesis if age is correlated with deaths.

Results & Insights:

Results of the two-sample t-test with a one-tailed test:

	<i>Over 65 Deaths</i>	<i>Under 65 Deaths</i>
Mean	826.2875817	78.76470588
Variance	1028483.747	22903.91395
Observations	459	459
Hypothesized Mean Difference	0	
df	478	
t Stat	15.61886217	
P(T<=t) one-tail	4.96E-45	
t Critical one-tail	1.648047653	
P(T<=t) two-tail	9.91009E-45	
t Critical two-tail	1.964939272	