Influenza-Interim Report

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:

If the vulnerable population (patient is older than 65 years old) increases, then the rate of mortality increases.

Data Overview:

Influenza deaths by geography:

- 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.

Population data by geography, time, age, and gender:

- The US Census data includes United States population data by geography, age, and gender from 2009 to 2017. It provides county-wise counts of males and females, along with the population segregated by age.

Data Limitations

Influenza deaths by geography:

- Limitations of influenza death data by geography include potential underreporting due to variations in reporting practices among regions, discrepancies in diagnostic criteria, and differences in healthcare infrastructure.

Population data by geography, time, age, and gender:

- The US Census data has a temporal lag as it is typically conducted every ten years, which may lead to outdated information.In the surveys, some individuals or groups may be missed, leading to an undercount. Lack of Health Information Data.

Descriptive Analysis

Conducted Descriptive Analysis to find Mean, Standard Deviation, Outlier Percentage of different Variables.

Variables	Total Population	Total Deaths	65+ Years Population	65+ Years Deaths
Standard Deviation	6806791	1084	887017	972
Mean	5973849	1435	806989	897
Outlier Percentage	5%	4%	6%	4%

Correlation

Variables	0- 64 Years Population and 0- 64 Years Influenza Deaths	65+ Years Population and 65+ Years influenza Deaths
Correlation Coefficient	0.91	0.94
Strength of Correlation	Strong Correlation	Strong Correlation

There is a Strong Correlation between 65+ Years Population and Influenza Deaths.

Results and Insights

Inferential Statistics used to Test our Hypothesis. Statistical analyses are employed to draw conclusions about populations based on sample data.

Null Hypothesis(H0)

The influenza mortality rate is higher in people younger than 65 years old.

Alternative Hypothesis(HA)

The mortality rate due to influenza is higher among individuals aged 65 and older.

t-Test: Two-Sample Assuming Unequal Variances

	0-64 Years	65+ Years
Mittelwert	538.0021786	896.581699
Varianz	15228.6441	945092.956
Beobachtungen	459	459
Hypothetische Differenz der Mittelwerte	0	
Freiheitsgrade (df)	473	
t-Statistik	-7.839411768	
P(T<=t) einseitig	1.51202E-14	
Kritischer t-Wert bei einseitigem t-Test	1.648081483	
P(T<=t) zweiseitig	3.02405E-14	
Kritischer t-Wert bei zweiseitigem t-Test	1.964991997	

- The p-values for both one-sided (1.51202E-14) and two-sided (3.02405E-14) tests are extremely small. This suggests a high level of significance and provides further evidence against the null hypothesis.
- The small p-values and the t-statistic falling beyond the critical values, we would likely reject the null hypothesis. This implies that there is a statistically significant difference in influenza mortality rates between the two age groups.
- The p-values are extremely smaller than alpha, for this reason we can reject the Null hypothesis.
- **With a confidence level of 95 percent, we found that Influenza Mortality is more in 65+ years old.**

Remaining analysis and next steps

- Spatial and temporal visualizations.
- Conclusions, recommendations, and proposed next steps in the final presentation.
- The audience when determining which analysis components to include in the final presentation.

Appendix

Datasets
1. Influenza deaths by geography Source:
- [CDC](#)
- Type of Data : Administrative Data
- Trustworthy: The data is administrative data collected as part of the National Vital Statistics Cooperative Program. Each of the U.S. states and territories is required to record all births, deaths, marriages, and divorces within their jurisdiction. Death records come from death certificates, in which a doctor codes the primary cause of death as "Influenza" or "Pneumonia" (ICD-10 codes J09-J18).

2. Population data by geography, time, age, and gender Source:
- [US Census Bureau Dataset](#)
- Type of Data Source: External
- Data Ownership : US Census Bureau

- Trustworthy: US Census Bureau is a reputable government agency for collecting data. So the data Source is Trustworthy. Data Collection The Medical Staffing Agency does not have this data. It is collected from the US Census Bureau. The US census bureau collects data by conducting Surveys every 10 years.

Hypothesis

Clarifying Questions.
- Which city needs more Healthcare workers in Peak Influenza Season?
- Which month has the peak number of Influenza Cases?
- How do healthcare facilities respond to rising demand during the Influenza Season?
- Where is the more number of cases occurring(Rural area or Urban Area)?

Funneling Questions
- Are there any particular months or weeks within the flu season with the highest demand of Temporary staff?
- Are there any specific types of healthcare positions that are in higher demand during flu season?
- Are there the same number of healthcare workers needed in Rural and Urban areas?

Privacy and Ethics Questions
- What measures will be taken to ensure that patient information remains confidential?
- What kind of data will be collected and shared with third parties, and how will individuals give consent for this sharing?
- How will the staffing agency ensure that temporary staff have proper Qualification?
- How will staffing agencies allocate temporary staff? What factors will be considered for the allocation of temporary staff