

# Preparing For Influenza Season: Interim Report.

By Oluwole Oni

## 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

If a state has a higher population of vulnerable individuals, it's likely to see more influenza deaths.

*Vulnerable - refers to population of 65 years and older.*

*Non-vulnerable - refers to population of 64 years and younger.*

## Data Overview

- **Population Data:** This dataset provides a detailed breakdown of the population count across defined age group, as well as by states and counties in the United State, covering the period from 2009 to 2017. The data is sourced from the US Census Bureau. As government-provided information, it can be considered as a verified Trustworthy data source.
- **Influenza Mortality Data:** This dataset provides a detailed count of influenza-related deaths by state and age group in the United State, spanning the years 2009 to 2017. This data is an external source from the Centers for Disease Control and Prevention (CDC), as government data it can be verified as a trustworthy data source.

## Data Limitations

### Population Data:

- **Data completeness:** Many data values were missing, especially within certain years and counties. No adjustments/imputation were made to the missing values to avoid influencing the overall results of my analysis.
- **Duplicates:** numerous duplicate values needed to be removed
- **Human error:** one of the counties was misspelled and had to be corrected.

## Influenza Mortality Data:

- **Data completeness:** There are suppressed values in the data represented as 0 - 9 mortalities.
- **Human error:** Due to manual imputation, several errors were identified in the data and were corrected.

## Descriptive Analysis

### Population Data

| AGE GROUP | STANDARD DEVIATION | AVERAGE |
|-----------|--------------------|---------|
| <5-14     | 1372775            | 1171978 |
| 15-24     | 975366             | 837688  |
| 25-34     | 962738             | 805351  |
| 35-44     | 931734             | 792875  |
| 45-54     | 947575             | 844762  |
| 55-64     | 772218             | 712847  |
| 65+       | 887070             | 807370  |

### Influenza Mortality Data

| AGE GROUP | STANDARD DEVIATION | AVERAGE |
|-----------|--------------------|---------|
| <5-14     | 17                 | 164     |
| 15-24     | 9                  | 53      |
| 25-34     | 11                 | 55      |
| 35-44     | 13                 | 56      |
| 45-54     | 33                 | 67      |
| 55-64     | 86                 | 98      |
| 65+       | 976                | 889     |

### Correlation

| Row Label               | <5-14 | 15-24 | 25-34 | 35-44    | 45-54  | 55-64  | 65+    |
|-------------------------|-------|-------|-------|----------|--------|--------|--------|
| Correlation Coefficient | -0.02 | 0.03  | 0.13  | 0.40     | 0.81   | 0.94   | 0.94   |
| Strength of Correlation | No    | No    | Weak  | Moderate | Strong | Strong | Strong |

### Interpretation

The lack of significant relationship between the age groups of under 5 and 34 indicates that influenza is not a major cause of mortality in these groups.

The strong positive correlation of 0.94 suggests a significant relationship between states with larger populations of individuals aged 65 and older and the total number of influenza deaths in those states.

## **Results And Insights**

### Statistical Hypotheses

- **Null Hypothesis:** The mortality rate from influenza of vulnerable population is equal to or lower than non-vulnerable population.
- **Alternative Hypothesis:** The mortality rate from influenza of vulnerable population is greater than non-vulnerable population.

## **Interpretation**

The p-value is significantly below the 0.05 significance level, leading to the rejection of the null hypothesis. This result indicates a 95% confidence that the population aged 65 and older has a significant impact on the number of influenza deaths in each state. Based on this finding, I can proceed with using the research hypothesis to predict influenza trends for the next season and make informed decisions regarding the allocation of medical staff to each state.

## **Remaining Analysis And Next Steps**

### Remaining Analysis

- Identify the states with the highest population of individuals aged 65 and older
- Create charts, scatterplots and heatmap that are easily understood

### Next Steps

- Provide recommendations, outline the proposed steps of the staffing plan in the final presentation.
- Prepare a final presentation summarizing the findings and the staffing plan.

### Presentation Approach

- Using Tableau for visual storytelling.
- Using graphs & statistical results to explain findings.
- Record a video presentation for stakeholders