Interim Report

Project Overview

Motivation

The United States has an influenza season where more people than usual suffer form 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 person is older than 65, they have a higher risk of death due to influenza.

Data Overview

US Census Population

This data set indicates the U.S. population, by county, from 2009 to 2017, and breaks it down by gender and age group.

CDC Influenza Deaths

This data set provides detailed information on influenza-related deaths across all U.S. states, from 2009 to 2017, and breaks it down by month and age group.

Data Limitations

US Census Population

A decennial census is performed in years ending on zero, and is done either online, by mail, in person, or by phone. To account for changes in the demographic landscape, these numbers are updated annually with estimates. Errors can occur during collection and processing, and it is possible that incorrect assumptions might be used when estimating annual updates.

CDC Influenza Deaths

The data is collected by way of surveillance system, updated weekly, and historic information is continuously being revised. Deaths included are those classified by way of ICD-10 code as associated with either influenza or pneumonia, and counts less than 10 are suppressed to protect the identify of individuals.

The data is subject to data entry errors, and the weekly updates and revision of reported figures mean that accurate and complete data may not be readily available. It is also difficult to determine the true cause of death, since many patients that contract influenza, ultimately die from pneumonia and a death certificate can only list one cause of death. This does not mean however that all pneumonia patients contracted influenza so the data set includes deaths unrelated to influenza that will distort the analysis. Results are further distorted by suppressed counts, which could otherwise have an impact in aggregate.

Descriptive Analysis

Variable	Mean	Standard Deviation
Population 65+ Years of Age	793553	802355
Influenza Deaths 65+ Years of Age	890	975

The number of influenza deaths has a strong correlation (0.9) with age, meaning the older a person is, the more likely they are to die from influenza.

Results & Insights

Null Hypothesis

The influenza death rate for people older than 65 will be less than, or equal to the influenza death rate of people younger than 65.

Alternative Hypothesis

The influenza death rate for people older than 65 will be higher than the influenza death rate of people younger than 65.

At an significance level (alpha) of 0.05, or confidence level of 95 percent, there's a significant difference in death ratio between people younger than 65, and those aged 65 and older.

Remaining Analysis & Next Steps

Next steps to support the timing and spatial distribution of the staffing plan:

1. Analysing influenza trends to determine when flu season occurs in a state.

2. Doing a spatial analysis to determine which states have the most vulnerable populations (older than 65) that would require more support staff.

These results will be included in the final report and presented at the final meeting.

Appendix

Business Requirements Document / Project Brief

Data Profile - US Census Population

Variables	Time-variant / -invariant	Structured / Unstructured	Qualitative / Quantitative	Qualitative: Nominal / Ordinal Quantitative: Discrete / Continuous
County	Time-invariant	Structured	Qualitative	Nominal
Year	Time-invariant	Structured	Qualitative	Ordinal
Total Population	Time-variant	Structured	Qualitative	Discreet
Male Total Population	Time-variant	Structured	Quantitative	Discreet
Female Total Population	Time-variant	Structured	Quantitative	Discreet
Age Groups	Time-variant	Structured	Quantitative	Discreet

Data Profile - CDC Influenza Deaths

Variables	Time-variant / -invariant	Structured / Unstructured	Qualitative / Quantitative	Qualitative: Nominal / Ordinal Quantitative: Discrete / Continuous
State	Time-invariant	Structured	Qualitative	Nominal
State Code	Time-invariant	Structured	Qualitative	Nominal
Year	Time-invariant	Structured	Qualitative	Ordinal
Month	Time-invariant	Structured	Qualitative	Ordinal
Month Code	Time-invariant	Structured	Qualitative	Ordinal
Ten-Year Age Groups	Time-invariant	Structured	Qualitative	Ordinal
Ten-Year Age Groups Code	Time-invariant	Structured	Qualitative	Ordinal
Deaths	Time-variant	Structured	Quantitative	Discreet