PREPARING FOR INFLUENZA SEASON: INTERIM REPORT

PROJECT OVERVIEW

The analysis sought to determine which populations had the highest risk of flu-related hospitalizations and deaths, so more staff could be sent to areas where more of these vulnerable populations reside. Following the hypothesis that those younger than 5 years and older than 65 are more likely to die of flu-related complications, it was confirmed that there is a greater death rate among this vulnerable age group than for all other age groups. Going forward, this knowledge will be used to determine the number of staff allocated to each state.

- **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

There will be higher flu-related mortality rates in regions where there are more people in vulnerable age groups.

If an individual is older than 65 years of age or less than 5 years of age, then that individual has a greater chance of hospitalization or death because of inflection with the flu. If more staff is sent to the areas with more vulnerable people in these age brackets, then hospital and staff shortages can be lessened or avoided.

DATA OVERVIEW

- **US Census Data**
  This US government data source contains population data by state, year, and county for each 5-year age category. Data was gathered for the years 2009-2017.

- **CDC Influenza Deaths Data**
  This US government data source contains monthly death counts for influenza-related deaths in the United States from 2009 to 2017 and are categorized by 10-year age groups. Death records come from death certificates, in which a doctor codes the primary cause of death as "Influenza" or "Pneumonia."

DATA LIMITATIONS

This data does not consider the COVID-19 pandemic over the last two years and how it may have impacted the number of those suffering from the flu or increased the overall need for staff.

- **US Census Data**
  There are numerous missing county population records in various years, accounting for 11.45% of records. Otherwise, this is still a reliable government data set. Calculations requiring population data used state population data, not county, which helped lessen the effect of the missing data.

- **CDC Influenza Deaths Data**
  There was a significant amount of suppressed death data in most age groups, with the exception of 65+ years age groups. This data is suppressed when there are less than 10 deaths per record. This is done for privacy reasons. Random values under 10 were substituted to calculate reasonably accurate estimates of death rates. Additionally, there was a "not stated" age category that held entirely suppressed death records, for each state and year. It was not included because it was missing vital age-related information.

DESCRIPTIVE ANALYSIS

VARIABLE	AVERAGE (MEAN)	STANDARD DEVIATION
% OF POPULATION 65 OR OLDER BY STATE/YEAR	14%	0.018
NUMBER OF DEATHS OF THOSE 65+ BY STATE/YEAR	889	975

There is a strong positive correlation between the number of people 65+ who reside in a given state and year and the total number of deaths in that area and time frame, meaning that the larger the elderly population, the larger the number of deaths in a given area and time frame.

RESULTS & INSIGHTS
- **Null hypothesis:** The death rate in populations of people older than 65 years and younger than 5 years is equal to or less than the number of deaths in all other age groups of the population.
- **Alternative hypothesis:** There is a higher death rate in vulnerable age groups, 65+ and younger than 5 years, than for other age groups in a population.

At an alpha of 0.05, or confidence level of 95 percent, there is a significant difference between the flu death rate of those 65+ and children younger than 5 years and the flu death rates of all other age groups.

REMAINING ANALYSIS & NEXT STEPS
- At this midpoint in the analysis, there will be updates for key stakeholders as outlined in the project management plan:
  - A call with upper management stakeholders, such as hospital administrators and staffing agency administrators for a brief progress update and to answer any questions.
  - E-mail message to all key stakeholders, including management, frontline healthcare workers, and patients.
- A brief, focused statistical evaluation of 65+ age group, instead of 65+ and under 5 years combined to see if more staffing resources should be prioritized for states with higher elderly populations as opposed to young or young & old combinations.
- Acquiring additional data sources that may provide a more complete picture of the situation, such as more comprehensive death or infection data on the younger age groups or vaccination rates.
- Create visualizations of analysis for key stakeholders.

APPENDIX

GOAL
The goal of this project is to help a medical staffing agency that provides temporary workers to hospitals and clinics on an as-needed basis. This analysis will help plan for influenza season, a time when additional staff are in high demand, with the goal of mitigating potential staffing shortages.

- Project Brief
- Project Management Plan

SUCCESS FACTORS
- A staffing plan that utilizes all available agency staff per state requirements, without necessitating additional resources
- Minimal instances of understaffing and overstaffing across states (a state can be considered understaffed if the staff-to-patient ratio is lower than 90% of the required ratio and overstaffed if greater than 110%)

DATA SETS
- Flu Deaths, CDC
- Population Census, US Census

STATISTICAL HYPOTHESIS TESTING PERFORMED

T-TEST: TWO-SAMPLE ASSUMING UNEQUAL VARIANCES

	Flu Death Rate Under 5/65+	Flu Death Rate Other Ages
MEAN	0.001109225	0.000197893
VARIANCE	2.97578E-07	4.02823E-08
OBSERVATIONS	459	459
HYPOTHESIZED MEAN DIFFERENCE	0	
DF	580	
T STAT	33.59035711	
P(T<=T) ONE-TAIL	1.81E-138	
T CRITICAL ONE-TAIL	1.647485042	
P(T<=T) TWO-TAIL	3.6254E-138	
T CRITICAL TWO-TAIL	1.964062511	

From this analysis, we reject the null that the death rate for the vulnerable age groups is equal or less than the death rate for other age groups. This means that the death rate for the vulnerable population is greater than the death rate for other age groups, with a p-value of 1.81 E-138, which is in scientific notation. This p-value is much smaller than our chosen alpha of 0.05, so we reject the null.