Ann Mariya Francis

Exercise 1.10

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.

If the individual is aged 65 or older, then the individual has a higher risk of developing flu complications that could lead to increasing the number in deaths in the state.

Data overview

- **Influenza deaths by geography:** this data set comes from the CDC (Centers for Disease Control and Prevention) and contains the information about flu-related death counts between 2009 and 2017. The data was collected monthly per each US state and is divided into ten-year age groups.
- **Population data by geography, time, age, and gender:** this data set is provided by US Census Bureau and comprises of US population numbers. The data were collected in years 2009-2017 per counties in each state. Besides the total population counts, the numbers are organized across genders and five-year age groups.

Data limitations

The data available for this analysis was collected from year 2009 until 2017, therefore it is not the most up to date. It would be beneficial to have more recent data from the years 2018-2023, as the trends in infection rates and deaths may vary. In the mortality data set most death records falls into the category ‘Suppressed’.

Descriptive Analysis

Children under 5 (vulnerable group) data is omitted as the sample data of deaths were indicated “Suppressed”. This could most likely be due to privacy concerns and lack of consent from parents to share that information publicly. As well, the government data sources do not provide more information regarding pregnancy or chronic medical information (ex: HIV/AIDs) of individual status. Therefore, this project’s analysis will primarily focus on population over 65 as it has the most delivered information.

	Total Influenza Death	Total death 65+
Data Set	Integrated Census data & Flu data	
Sample or Population	Sample	Sample
Normal Description	Right Skewed	Right Skewed
Variance	1370743.21	1053020.57
STD.Deviation	1170.79	1026.17
Mean	982.08	896.61
Outlier Lower Bound	-1359.5	-1155.73
Outlier Upper Bound	3323.65	2948.95
Outlier Count	18	18
Outlier Percentage	4%	4%

Dependent Variable	Total Flu Death
Independent Variable	Population 65+
Null Hypothesis	The risk of dying from the flu either equal or higher for individual aged less than 65 compared to those aged 65 and older
Alternative Hypothesis	If state has a high number of people 65+,then it will have more flu related death overall
One tail	Examining if there's a higher mortality rate in state with higher 65+ proportion or not
Result	It appears that the percent of the population over 65 doesn't have a significant effect on the total number of flu death

Statistical Analysis

	0-64 years old	65+ years old
Mean	0.000809457%	0.086474735%
Variance	1.00510680496592E-10	1.785623159328 69E-07
Observations	450	450
Hypothesized Mean Difference	0	
df	449	
t Stat	-42.9926273815464	
P(T<=t) one-tail	1.42776992102307E-161	
T Critical one-tail	1.64825431724876	
P(T<=t) two-tail	2.85553984204614E-161	
t Critical two-tail	1.96526138811176	

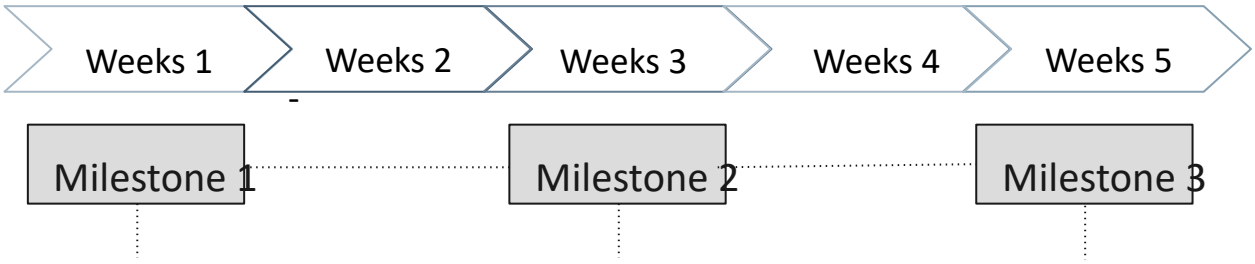
Results & Insights

t Stat	-42.9926273815464
P(T<=t) one-tail	1.42776992102307E-161
T Critical one-tail	1.64825431724876
P(T<=t) two-tail	2.85553984204614E-161
t Critical two-tail	1.96526138811176

Remaining analysis and next steps

Further analysis will include comparison of different age groups mortality and determining age mortality relationship across years and geographical regions (states).
This leads to some other essential questions that need to be answered:

- Does influenza occur seasonally or throughout the entire year and is it the same in every state?
- Which states should be prioritized in staff allocation based on size of the vulnerable populations?



Deliverables

- Gain domain knowledge of influenza records and seasons
- Data cleansing, analysis, provide limitations from appropriate sources
- Integrate 2 data sources into 1 cohesive view and conduct statistical analyses to find any correlation

Deliverables

- Provide biweekly updates of project progress status

- Provide 1 interim report to confirm/ disconfirm hypotheses
- Receive feedback, enhance data further, and preparation of data visualizations (bar chart, treemap, time forecast, spatial analysis)

Deliverables

- Schedule meeting to communicate insights with key stakeholders (staffing agency)
- Presentation of influenza trends and provide recommendations of how many and when to send frontline staffing workforce to hospital/clinics (50 states)