Preparing for Influenza Season: Interim Report

Submitted by Ryan Wick

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.

Hypothesis

- If a person's age increases, then their risk of influenza-related mortality will also increase.

Data Overview

- **Influenza deaths by geography**
  - **Source:** Centers for Disease Control and Prevention (CDC)
  - **Ownership:** U.S. Government
  - **Trustworthiness:** Reliable, given the CDC's reputation as a reputable government agency. However, data quality may vary due to factors like reporting accuracy and completeness.
  - **Data Collection Method:** Administrative data collected automatically through reporting systems and databases. Potential time lag between death occurrence and data inclusion.
  - **Data Contents:** State, year, month, age group, and number of influenza-related deaths. Lacks population demographic data.
  - **Limitations:** Potential reporting errors, incomplete data, missing demographic information.
- **Population data by geography, time, age, and gender**
  - **Source:** U.S. Census Bureau ([Link](hQps://[invalid URL removed]) - **Note:** This link might be restricted)
  - **Content:** County-level population estimates for the United States from 2009 to 2017.

- **Details:**
  - **Age Groups**: 5-year increments from 0 to 85+
  - **Gender**: Male and Female
  - **Geography**: Counties within each U.S. state
- **Note:** This dataset provides population estimates, and the sum of individual age groups might not perfectly match the total population due to rounding or estimation methods.

Data Limitations

- This analysis relies on two primary datasets:
  - ### Influenza Deaths by Geography:
    - **Limitations:**
      - **Data Quality Issues:** Missing values, inconsistencies, and potential errors due to manual data entry.
      - **Data Suppression:** Some states suppressed mortality data, limiting the scope of analysis.
      - **Age Group Limitations:** Many states did not report deaths for age groups younger than 65.
  - ### Population Data by Geography, Time, Age, and Gender:
    - **Limitations:**
      - **Time Lag:** Potential time lag between actual population counts and data availability.
      - **Data Accuracy:** Manual reporting by citizens might introduce errors and inaccuracies.
      - **Data Age:** The dataset covers the period from 2009 to 2017, which may not accurately reflect current population trends and the impact of recent health advancements.

Descriptive analysis

	65+ Mortality Rate	15-44 Mortality Rate	65+ Mortality Rate (Formatted)	15-44 Mortality Rate (Formatted)
Mean	0.001316527	0.000211332	0.0013	0.0002
Variance	2.74417E-07	4.92678E-08	0.0000	0.0000
Observations	459	459	459.0000	459.0000
Hypothesized Mean Difference	0		0.0000	
df	617		617.0000	
t Stat	41.61828981		41.6183	
P(T<=t) one-tail	1.4203E-181		0.0000	
t Critical one-tail	1.647327004		1.6473	
P(T<=t) two-tail	2.8407E-181		0.0000	
t Critical two-tail	1.963816258		1.9638	

- **P-value:** 1.42035E-181 (0.0000)
- **Significance Level:** 0.05

- **Assessment of p-value to Significance Level**
  - Since our p-value is so low it is still formatted as being zero four decimal places out. This then shows us that our p-value is significantly smaller than our .05 (alpha) meaning we a 95%+ certainty that in fact as your age increases you will have a higher mortality rate when it comes to influenza.

- **Correlation Result:**

Variable	**85+ Flu Deaths and Total Flu Deaths**	Total Population & Total Death
Proposed Relationship	This should demestrate that as 85+ deaths increase the overall total deaths will increase.	To demostrate that the larger the population the higher number of flu related deaths should exist.
Correlation Coefficient	0.99	0.95
Strength of Correlation	Strong relationship (>0.7)	Strong relationship (>0.7)
Usefulness / Interpretation	A correlation coefficient of 0.99 between age group 85+ and total flu deaths indicates a very strong positive relationship, suggesting that the elderly population is significantly more vulnerable to flu-related deaths. This information can inform public health strategies, healthcare resource allocation, and policy decisions to protect the elderly.	A correlation coefficient of 0.95 between total flu deaths and total population indicates a very strong positive relationship. This suggests that a larger population size is associated with a higher number of flu deaths. This information can help public health officials anticipate and prepare for flu seasons in regions with larger populations.

Results and insight

- **Null Hypothesis**
  - There is no significant relationship between age and the risk of influenza-related mortality.
- **Alternative Hypothesis**
  - As age increases, the risk of influenza-related mortality also increases.
- **Hypothesis Testing**
  - Our analysis aimed to determine if there is a significant relationship between age and the risk of influenza-related mortality. The null hypothesis stated that there is no significant relationship, while the alternative hypothesis suggested a positive correlation between age and mortality risk.

- **Statistical Analysis**
    - A statistical test was conducted to assess the validity of the alternative hypothesis. The p-value obtained was significantly smaller than the alpha level ($0.0000 < 0.05$), allowing us to reject the null hypothesis.
- **Interpretation**
    - This finding strongly supports the notion that older individuals are at a significantly higher risk of influenza-related mortality compared to younger populations. This insight underscores the importance of targeted public health interventions, particularly for older adults, to mitigate the impact of influenza outbreaks.

Remaining Analysis and Next Steps

- To further our understanding of the impact of influenza, we will conduct the following analyses:
    - **Geographic Analysis:** A more detailed geographic analysis will be performed to identify regions with higher mortality rates and potential contributing factors, such as socioeconomic disparities, healthcare access, and climate conditions.
    - **Age-Specific Analysis:** We will delve deeper into age-specific mortality rates to understand the impact of influenza on different age groups, particularly vulnerable populations like the elderly and young children.
    - **Visual Analysis:** Interactive visualizations will be created to communicate our findings effectively. These visualizations will include maps, charts, and dashboards to highlight trends, patterns, and correlations.
- In addition to the analysis, the following next steps will be taken:
    - **Final Deliverable:** A comprehensive final report will be prepared, summarizing our key findings, insights, and recommendations. This report will be shared with relevant stakeholders, including policymakers, healthcare providers, and public health officials.
    - **Presentation:** A presentation will be developed to communicate our findings to a broader audience. This presentation will include a clear and concise overview of our analysis, key findings, and actionable recommendations.
- By conducting these analyses and taking the proposed next steps, we aim to provide valuable insights that can inform public health policies, resource allocation, and healthcare interventions to mitigate the impact of influenza.

Appendix

- **Stakeholder communication.**
  - Medical agency frontline staff:
    - Regular updates on project progress and findings
  - Hospitals and clinics:
    - Bi-weekly reports on staffing predictions
  - Staffing agency administrators:
    - Weekly progress meetings and final presentation
- **Schedule and milestones**
  - Achievement 1: Preparing & Analyzing Data
    - **Week 1-2:** Data sourcing and profiling
    - **Week 3-4**: Data cleaning and integration
    - **Week 5-6**: Statistical analysis and hypothesis testing
    - **Week 7**: Interim report preparation
  - Achievement 2: Data Visualization & Storytelling
    - **Week 8-9:** Data visualization creation
    - **Week 10-11:** Narrative development and presentation design
    - **Week 12**: Final presentation and Tableau storyboard creation
- **Project deliverables**
  - Achievement 1:
    - Data research project design
    - Data profiles for each dataset
    - Integrated dataset
    - Statistical analysis results
    - Interim report
  - Achievement 2:
    - Data visualizations
    - Final presentation
    - Tableau storyboard
- **Audience definition.**
  - Primary audience: Staffing agency administrators
    - Decision-makers responsible for allocating medical staff
    - Interested in data-driven insights for efficient staffing
    - Require clear, actionable recommendations
  - Secondary audience: Hospital administrators and medical staff
    - End-users of the staffing recommendations
    - Need to understand the rationale behind staffing decisions
    - Interested in how the analysis accounts for local factors

- **4 proposed hypotheses before selecting Hypothesis 1.**
  - Hypothesis 1 (Age and Influenza Mortality):
    - If a person's age increases, then their risk of influenza-related mortality will also increase.
  - Hypothesis 2:
    - If a state has a higher percentage of vulnerable populations (elderly, children, pregnant women), then it will have a higher rate of influenza-related hospitalizations.
  - Hypothesis 3:
    - If a state has a higher flu vaccination rate among children, then it will have a lower number of influenza cases in the pediatric population.
  - Hypothesis 4:
    - If a state has a higher population density, then it will experience a more rapid spread of influenza during the peak season.
- **Data Wishlist should stakeholders have possible access or contacts.**
  - Age-specific influenza mortality rates:
    - Detailed data on influenza-related deaths categorized by age groups (e.g., 0-4, 5-17, 18-49, 50-64, 65+)
  - Total population data by age group:
    - To calculate mortality rates and compare across different age demographics
  - Influenza case data by age group:
    - To understand the relationship between infection rates and mortality rates across age groups
  - Vaccination rates by age group:
    - To account for potential confounding factors in mortality rates
  - Comorbidity data by age group:
    - Information on pre-existing health conditions that may influence influenza mortality, categorized by age
  - Historical influenza season data:
    - To analyze trends and patterns in age-related mortality over time
  - Geographic data:
    - To account for potential regional variations in age-related influenza mortality
  - Healthcare access data by age group:
    - Information on healthcare utilization and availability, which may impact mortality rates