Influenza Mortality and Vulnerability Analysis : Interim Report

1.Project Overview

Motivation:

The United States experiences an annual influenza season, increasing the demand for hospital staffing. Vulnerable populations, such as the elderly and young children, require additional care, placing a strain on hospitals and clinics. This project aims to help a medical staffing agency effectively plan for the upcoming influenza season.

Objective:

To determine when and how many staff members should be sent to each state to manage seasonal influenza demands.

Scope:

This project covers staffing needs for hospitals across all 50 U.S. states during influenza season.

Success Factors:

- Effective staffing with minimal instances of understaffing (<90%) or overstaffing (>110%).
- Full utilization of available staffing resources without additional hires.

2. Research Hypothesis

Hypothesis Statement:

States with larger vulnerable populations will experience higher influenza-related mortality rates, requiring increased medical staffing.

3. Data Overview

Dataset 1: Influenza Deaths by Geography

- **Source:** CDC
- **Key Variables:** State, year, number of deaths.
- **Purpose:** Highlights historical trends in influenza-related fatalities to identify states with greater mortality risks.
- **Summary:**
  - Mean deaths per state: **34.38**

- Standard deviation: **37.13**
- States with highest death totals: **California, New York, Texas.**

Dataset 2: Population Data by Age and Gender

- **Source:** [U.S. Census Bureau]
- **Key Variables:** State, age group (<5 years, >65 years), gender, population count.
- **Purpose:** Provides insights into vulnerable populations by geography to inform staffing prioritization.
- **Summary:**
    - Mean population per state: **59,540,484.05**
    - Standard deviation: **220,140,503.70**
    - States with largest vulnerable populations: **Texas, Georgia, Missouri.**

Integrated Data:

- **Description:** Combines CDC Influenza Mortality data and Census Population data for 2009–2017.
- **Normalization:**
    - Death rates per 100 population were normalized for comparability across states and age groups.
    - Age groups were reconciled to align CDC and Census data formats.
- **Key Insights:**
    - Senior Citizens (65+) exhibit higher normalized mortality rates than Adults (<65).
    - Smaller states often show disproportionately high normalized mortality rates, highlighting healthcare disparities.

4. Data Limitations

1. **Influenza Deaths by Geography (CDC Data):**
    - **Historical Data:**
      Data covers 2009–2017 and may not reflect the dynamics of more recent influenza seasons, including any changes in patterns due to COVID-19 or other factors.
    - **Reporting Standards:**
      Differences in state reporting methods (e.g., variations in how deaths are attributed to influenza) could lead to inconsistencies in the dataset.
2. **Population Data by Age and Gender (US Census Data):**

- **Estimates:**
  Census population data relies on projections, which may not perfectly reflect actual populations. Variability in estimation accuracy between states may also impact analysis.

- **Granularity:**
  Data is limited to age and gender categories, without details about other health conditions, healthcare access, or socioeconomic factors, which may impact vulnerability.

3. **Integrated Data Challenges:**

   - **Missing Data:**
     Imputation was performed for missing entries (e.g. Washington D.C. 2011), introducing potential uncertainty in mortality estimates.

   - **Age Alignment:**
     Reconciliation of mismatched age groups between datasets (for instance combining “Under 1 year” and “1–4 years” into “Under 5 years”) introduces approximations that may affect accuracy in age-specific analysis.

5. Descriptive Analysis

Core Metrics:

- **Influenza Deaths:**
  - Mean deaths per state: **34.38**
  - Standard deviation: **37.13**
  - States with highest death totals: **California, New York, Texas.**
- **Vulnerable Populations:**
  - Mean population per state: **59,540,484.05**
  - Standard deviation: **220,140,503.70**
  - States with largest vulnerable populations: **Texas, Georgia, Missouri.**

Correlation Analysis:

- A **strong positive correlation** exists between the size of the elderly population in a state and influenza-related deaths.
- **Key Insight:** States with larger elderly populations require more targeted staffing and resources to mitigate mortality risks.

6. Results and Insights

Statistical Hypothesis

- **Objective:** To determine if the normalized mortality rate for Senior Citizens (65+) is statistically higher than that for Adults (<65).
- **Hypotheses:**
  - H0(Null Hypothesis): Senior Citizens (65+) normalized mortality rate ≤ Adults (<65) normalized mortality rate.
  - HA (Alternative Hypothesis): Senior Citizens (65+) normalized mortality rate > Adults (<65) normalized mortality rate.

Statistical Test:

- **Test Type:** One-tailed t-test (α=0.05).
- **Results:**
  - Mean (Adults): **0.343**
  - Mean (Senior Citizens): **5.035**
  - Variance (Adults): **3.622**
  - Variance (Senior Citizens): **1489.121**
  - T Statistic: **-2.602**
  - p-value: **0.0048**

Conclusion:

- **Reject H0:** Senior Citizens (65+) have statistically higher normalized mortality rates compared to Adults (<65).

Interpretation:

- **Key Insight:** Senior Citizens (65+) are at a significantly greater risk, underscoring the need for targeted interventions such as vaccination campaigns and resource prioritization.

7. Appendix Materials

1. Data Profiles

- **Dataset 1:** **Influenza Deaths by Geography (CDC):**
  - Time range: 2009–2017.
  - Key variables: State, year, deaths.
  - Summary: Highlights historical mortality trends to identify high-risk states.

- **Dataset 2: Population Data by Age and Gender (U.S. Census Bureau):**
  - Key variables: State, age group (<5, >65), population count.
  - Summary: Used to identify vulnerable populations by geography.

2. Statistical Analysis Outputs

- **Descriptive Statistics:**
  - Means and standard deviations for:
    - Influenza deaths by state (mean: 34.38, SD: 37.13).
    - Vulnerable populations (mean: 59,540,484.05, SD: 220,140,503.70).
- **Hypothesis Testing Summary:**
  - Hypotheses, t-statistic, critical value, p-value, and conclusion (Senior Citizens have higher normalized mortality rates than Adults).

3. Integration Process

- **Key Adjustments:**
  - Age groups were aligned between datasets (e.g., combining "Under 1 year" and "1–4 years" into "Under 5 years").
  - Missing data (e.g., Washington D.C. deaths in 2011) was imputed using averages from adjacent years.

4. References

- **Data Sources:**
  - CDC Influenza Mortality Data.
  - U.S. Census Bureau Population Data.

8. Remaining Analysis and Next Steps

Remaining Analyses

1. **Spatial Visualizations:**
   - Create heatmaps to illustrate:
     - Influenza mortality rates by state.
     - Normalized death rates per 100 population across regions.
2. **Temporal Trends Analysis:**
   - Analyze yearly trends in influenza-related mortality to identify peak seasons or anomalies (e.g. severe flu seasons like 2009 with H1N1).

- Compare normalized death rates over time to assess the effectiveness of interventions (e.g., vaccination campaigns).

3. **Detailed Age Group Analysis:**
   - Further investigate mortality patterns within specific age groups (e.g. "85+ years" or Infants).
   - Correlate age-specific normalized death rates with healthcare access or vaccination coverage.
4. **Socioeconomic Correlations**
   - Explore socioeconomic variables (e.g., poverty rates, healthcare access) to explain variations in mortality rates between states.

Next Steps

1. **Stakeholder Presentation:**
   - Develop a stakeholder-friendly presentation focusing on:
     - Key findings (for instance Senior Citizens have higher mortality rates).
     - Recommendations for resource allocation and targeted interventions.
     - Visual summaries of high-risk areas and trends.
2. **Recommendations Development:**
   - Use the insights from analyses to develop actionable recommendations:
     - Increase vaccination drives for high-risk groups (Senior Citizens, specific states).
     - Allocate resources to states with high normalized death rates and vulnerable populations.
     - Propose funding for states with healthcare disparities.