Beginning

With each flu season in the United States, the number of cases of sick people arriving at the hospital increases considerably.

This generates a great need for medical personnel (doctors, nurses, nursing technicians) to meet considerably higher demands for care.

The number of patients treated varies according to state and due to this, a great deal of prior planning is necessary so that all people are properly cared for and in the best possible way.

The motivation of this analysis is helping the Medical Staffing Agency plan for the influenza season in the year of 2018.

The main objective is determining **when** and **how many** staff is needed for each state.

Middle

The datasets from the American population Census for the period 2009-2017 (US Census Bureau) and the count of deaths from influenza by geographic region and age (CDC) for the same period were cleaned and integrated.

Both Datasets were integrated, and the final table was used as the main source for all analysis.

A line chart was also made with the number of deaths per month to see at which time of the year the influenza related deaths mostly occur.

Statistical Analysis

After this first step, and hypothesis was created to conduct the research. That was:

“When a State has a larger number of elderly people, then the Influenza mortality rate will be higher”.

This hypothesis was chosen because this question is crucial for the medical staff distribution. The states with higher mortality rate and high influenza infections value will need more assistance.

Based on these data, the existence of a correlation between age group and mortality was established (are the elderly age groups more likely to die from influenza?). There was a positive weak correlation. A scatter plot chart was made to facilitate visualization.

Then, T test between the mortality rate over 65 and under 65 to see if there was a difference between then. There was a difference between the groups ($p<0,05$).

With these results, as a complementary analysis, the states were divided into 2 groups: the 10 states with the largest proportion of elderly people and the 10 states with the smallest proportion of elderly people for the year 2017 (most recent data). Then a T test was made with the mortality rates from those 2 groups to see if the states with more elderly people have a higher mortality rate. Here there was no difference found ($p>0,05$).

After the results, an extra analysis was made with a third Dataset that brought information about average household income per State in the year of 2017 (source: U.S. Department of Commerce, Census Bureau 2017), to establish if there was a correlation between income and mortality rate. There was a weak negative correlation. A scatterplot was also made for visualization.

Visualizations

-Bar chart to show the number of deaths per state in 2017.

-A line chart to visualize deaths per month

-1 scatterplots to visualize correlations.

-Spatial Analysis: Choropleth Maps and Point Maps combined.

End

Conclusions

With the statistical results presented, it was not possible to confirm that states with a greater number of elderly people have a higher mortality rate due to influenza.

Although there is a correlation between age and mortality, and there is a statistically significant difference between mortality in people over 65 and under 65, it is not possible to state that only the proportion of people over 65 in a state would be responsible for a high mortality rate (There was no statistical difference between the states with the highest and lowest proportions).

This can be explained by the fact that only data from the year 2017 were used, which makes the N number very low.

In addition, other factors can also influence mortality, such as poverty rates (the income scatterplot is an example), distribution of hospitals, presence of health insurance, etc. This data can be included in the next stage of the study.

Even so, it would be possible to initially use population data for a first analysis of the distribution of medical personnel by state, as more populous states will need more personnel regardless of the elderly population present. And because of that, this will be used on the presentation for the stakeholders.

Another important factor to consider is the Influenza season, that is, in which period of the year the number of deaths increases. With the analysis it was possible to see in the Line Chart that in winter there are more deaths caused by influenza. So, this is the time of the year that the states will need more medical staff.