DATA ANALYST PROJECT CASE STUDY

by: Padmasini Venugopal

INTRODUCTION

The main goal of this project is to help a medical staffing agency that provides temporary workers to clinics and hospitals on an as-needed basis. The analysis will help plan for influenza season, a time when additional staff are in high demand. The final results will examine trends in influenza and how they can be used to proactively plan for staffing needs across the country.

DATA AND PROJECT

- The influenza death data is collected from the CDC
  Link: source
  The US Census data
  Link: source
  Project Brief
  Link: source

DURATION

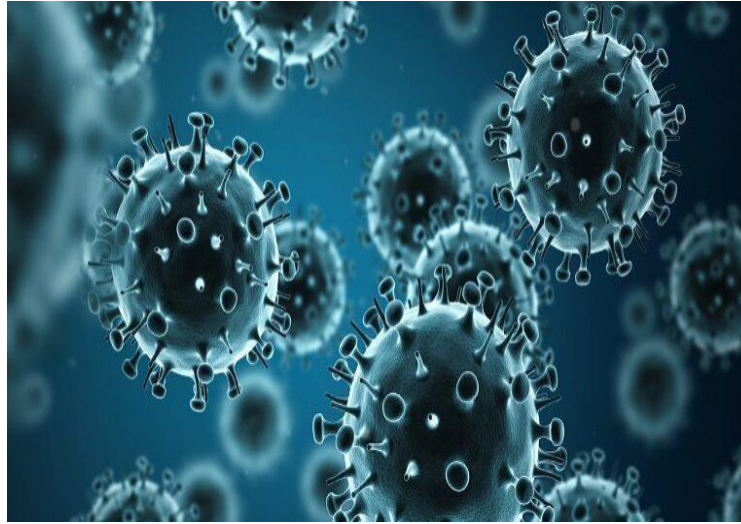It took almost 6 weeks for me to complete the entire analysis.
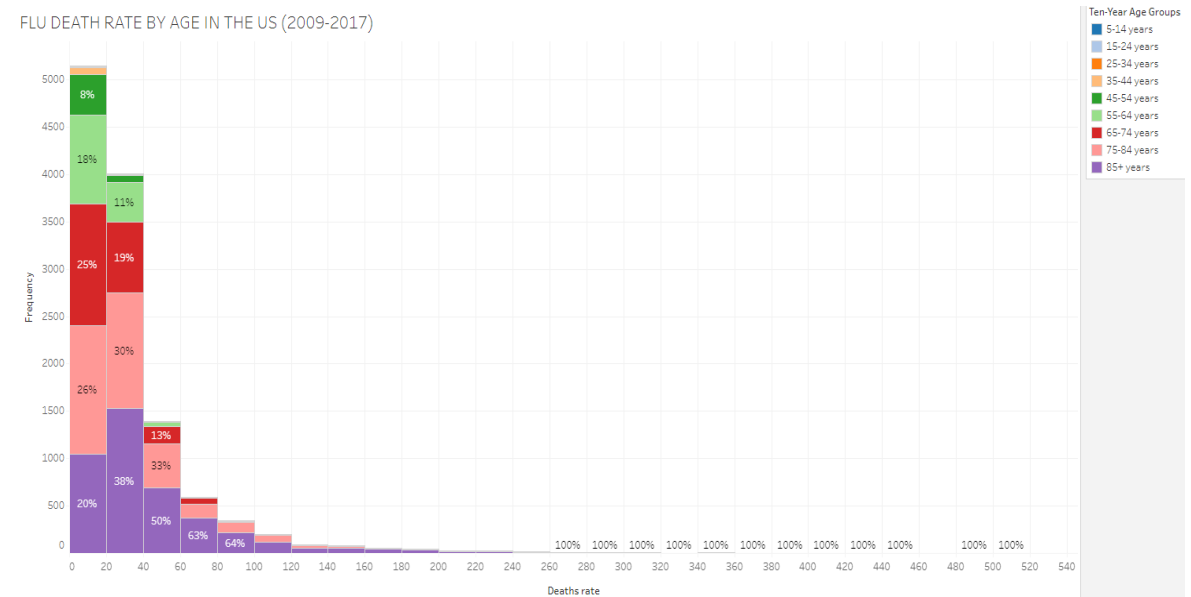
TOOLS USED

Microsoft Excel

Tableau

TECHNIQUES APPLIED

- Data cleaning
- Data quality and Integrity checks.
- Data integration & transformation
- Data grouping and Summarizing
- Statistical analysis
- Hypothesis testing
- Forecasting
- Visualizations using Tableau.

FLU DEATH RATE BY AGE IN THE US (2009-2017)

THE PROCESS

1. THE DATA PREPARATION PHASE

To evaluate the Influenza death rate for the upcoming year, I conducted a thorough scan for duplicates, missing numbers, and other anomalies. I subsequently eliminated extraneous columns and duplicates from the data set.

2. THE ANALYSIS PHASE

After the management team answered all my queries, I found it easier to proceed with my additional investigation. I created separate pivot tables for each instance, performed a deeper analysis, and merged both datasets using Vlookup. Finally, I used Tableau for visualizations and further analysis.
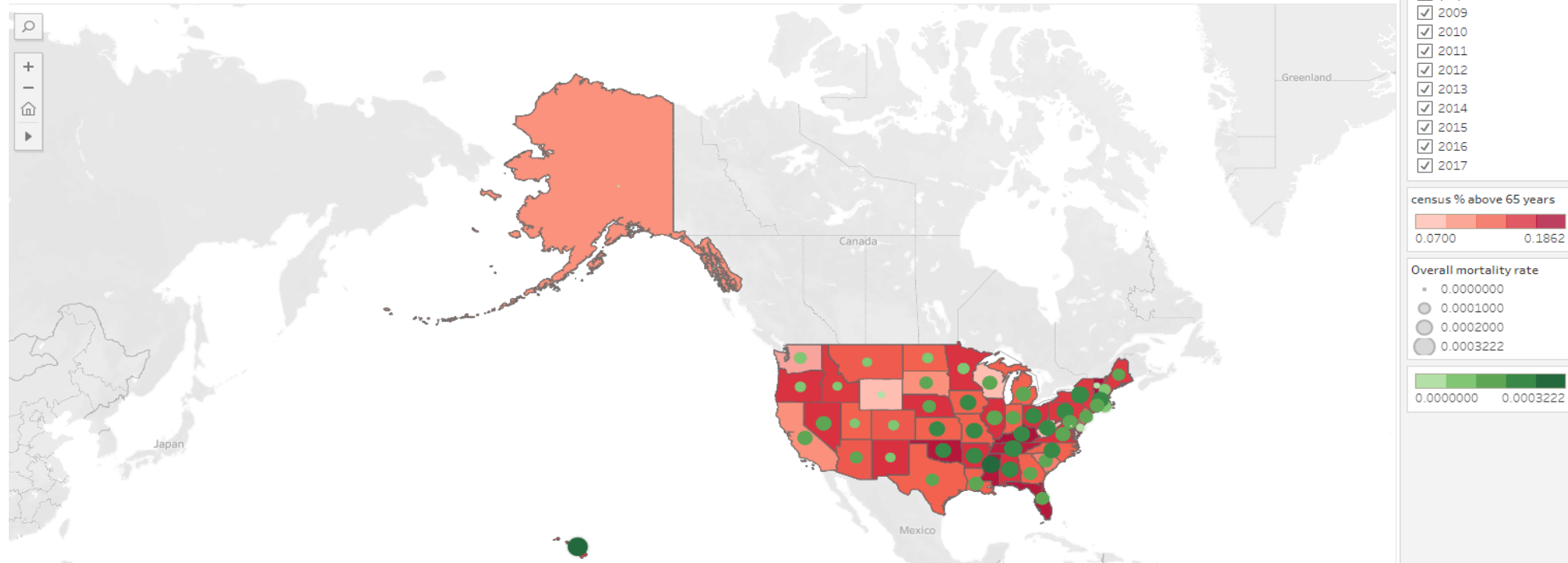
3. THE RESULTS

Finally, after all the analysis the final presentation was created using Tableau. The stakeholders were advised based on the results of the analysis with a clear disclaimer regarding the limitations of the data.
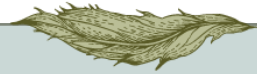
ANALYSIS
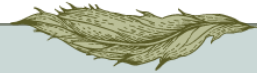
Relationship between % of Census data over 65 years and Total Mortality Rate(2009-2017)
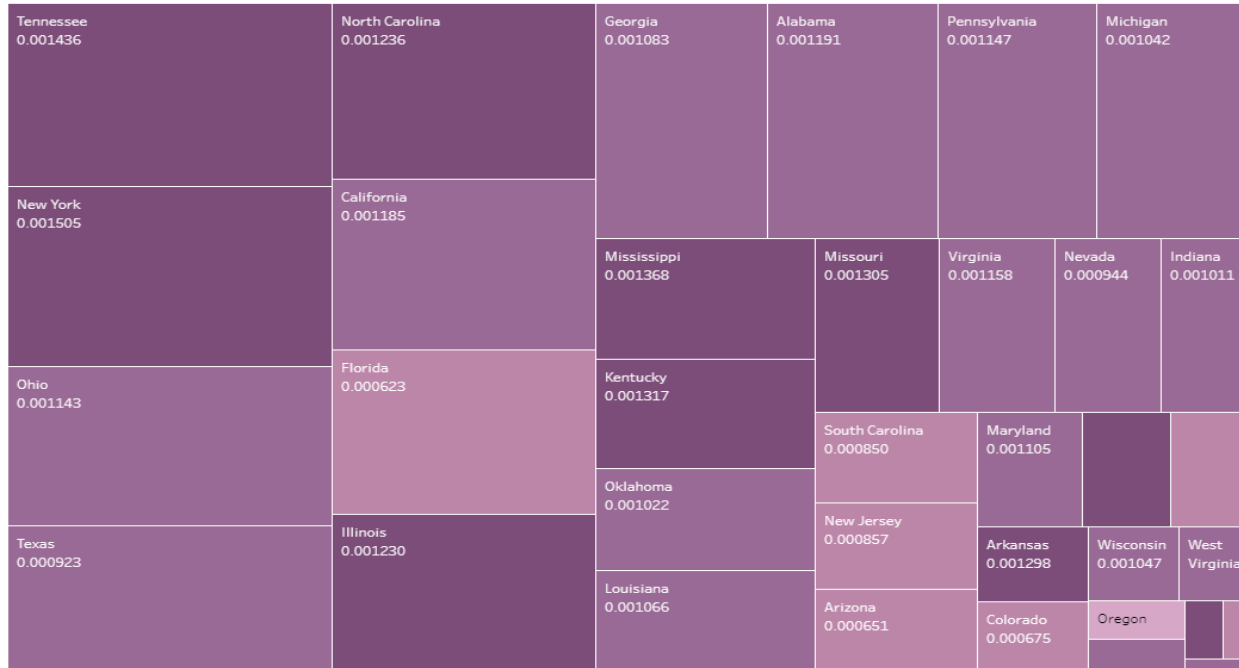
Hawaii state has the highest mortality rate followed by it is New York, Tennessee, and Mississippi this shows that the state with a more vulnerable population has the most death rates.
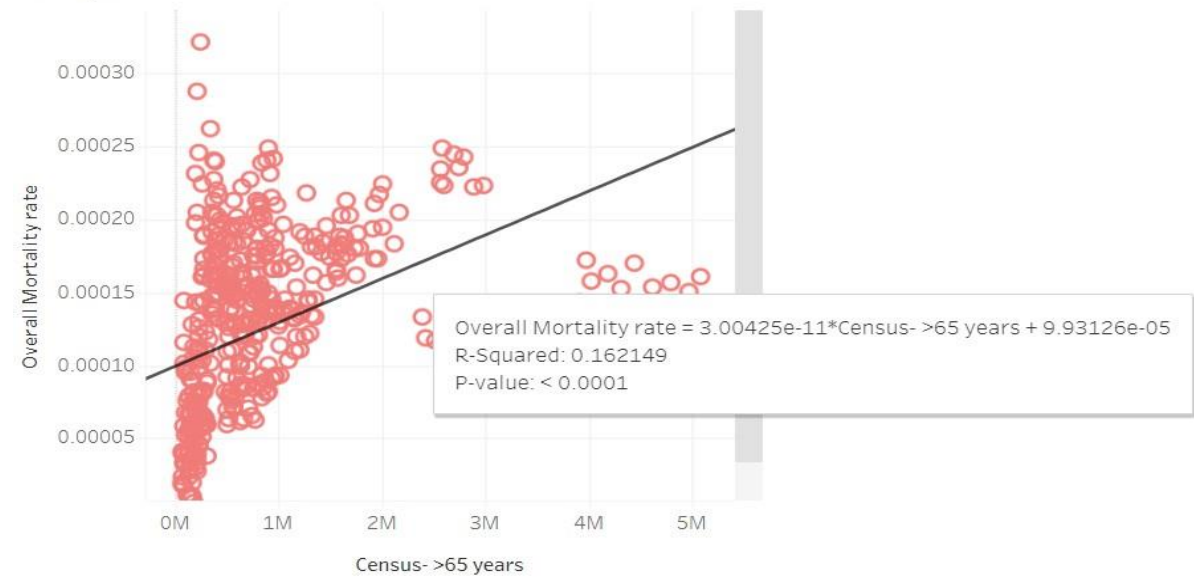
ANALYSIS

% OF FLU DEATH RATE 65+ BY STATES

|			
Tennessee 0.001436	North Carolina 0.001236	Georgia 0.001083 / Alabama 0.001191	Pennsylvania 0.001147 / Michigan 0.001042
New York 0.001505	California 0.001185	Mississippi 0.001368 / Missouri 0.001305 / Virginia 0.001158 / Nevada 0.000944 / Indiana 0.001011	
Ohio 0.001143	Florida 0.000623	Kentucky 0.001317	
		South Carolina 0.000850 / Maryland 0.001105	
		Oklahoma 0.001022	
Texas 0.000923	Illinois 0.001230	New Jersey 0.000857 / Arkansas 0.001298 / Wisconsin 0.001047 / West Virginia	
	Louisiana 0.001066	Arizona 0.000651 / Colorado 0.000675 / Oregon	

Despite California and Texas having the highest 65+ population count the states that have the highest flu mortality rate are New York, Tennessee, and Hawaii.

Relationship between the Census above 65 years and the overall mortality rate(2009-2017)

Overall Mortality rate = 3.00425e-11*Census- >65 years + 9.93126e-05
R-Squared: 0.162149
P-value: < 0.0001

Above 65 years pop. has the highest influenza mortality rate and they are a vulnerable population.

ANALYSIS
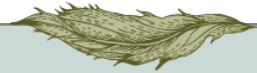
States from the southern region are going to have high flu deaths next year so it's best to send the faculties there in advance.
Next is the West (California) will have a high death rate

TOTAL FLU DEATH RATES(2009-2017) & FORECASTING FOR 2018
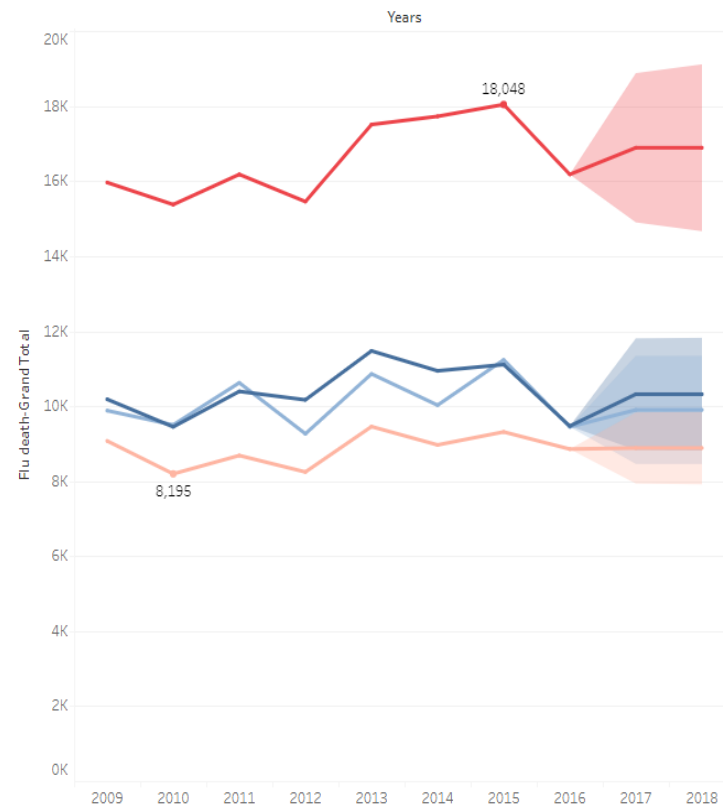
RETROSPECTIVE

WHAT WENT WELL?

I was surprised to discover the extensive range of functions that Excel offers to make tasks easier and more efficient. Upon learning that it is the most powerful tool for data analysis, I delved deeper into it. Learning Tableau for data visualization was exciting and impressive as it enhances analysis. And I personally Loved using Tableau and creating innovative storyboard presentations in it.

CHALLENGES FACED?

When I was learning Excel, one of the biggest hurdles I encountered was practicing and memorizing the complex formulas. With so many functions to keep in mind, it was difficult to remember them all at once. Additionally, I struggled with pivot tables at first, but with practice, I was able to become more proficient with them.

RECOMMENDATIONS

- ✓ To deploy more staff to the highest mortality states to prevent more flu deaths.
- ✓ Should prepare before the season- so sending adequate staff in late November to all the states(especially to those 5 states) will help us to be prepared for the season.
- ✓ To create health awareness by publishing more about flu spread and its prevention in newspapers and social media platforms.

Thank you

The Link to the Final Presentation

Click here

For further queries kindly contact
drpadmasini.v@gmail.com