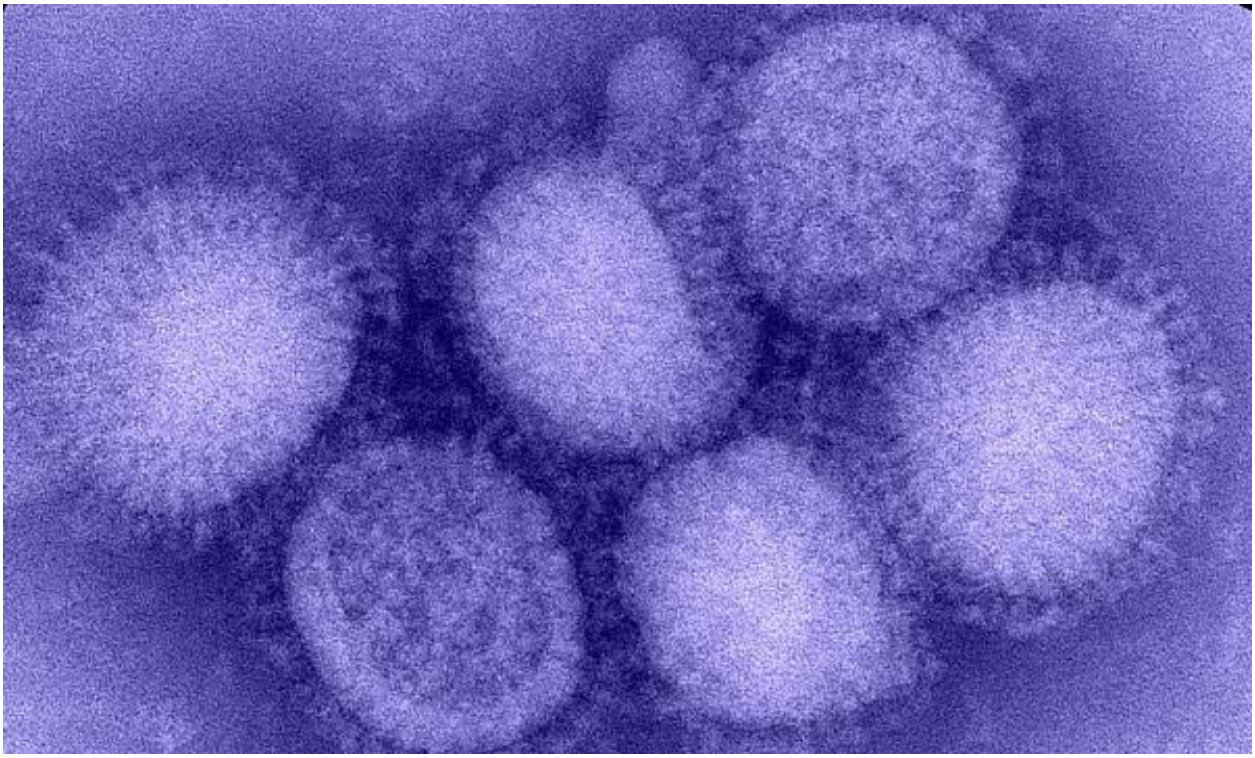

STAFFING FOR INFLUENZA

INTERIM REPORT BY MIGUEL OLIVEIRA

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 you are part of the vulnerable population, which includes any person 65 years old or over, you are more likely to die from influenza, upon infection.

DATA OVERVIEW

CDC Influenza Deaths: This data set contains the number of influenza deaths from the year 2009 until 2017, organized by US state and age group.

US Census Population: This data set contains the population number from each county in each US state, organized by year and age group.

DATA LIMITATIONS

CDC Influenza Deaths

- **Missing data:** Certain mortality values were listed as “Suppressed”, and some age groups were not specified (shown as NS).

- **Bias:** Influenza may not have been the only, or the leading, cause for the death of certain individuals. This means certain assumptions about the data were made to proceed with the analysis.

US Census Population

- **Missing data:** Certain counties were missing information for all 9 years.
- **Manual errors:** There is possibility for errors or incomplete data, since the census was filled out manually.
- **Timeliness:** Census data is only collected every 10 years, which means that counts may have been affected by a multitude of factors (moving, births, deaths, etc.), thus influencing the accuracy of the data.

DESCRIPTIVE ANALYSIS

	Mean	Standard Deviation
Influenza Deaths: 0 to 64 years old	**536.63**	**119.09**
Influenza Deaths: 65 years and older	**896.80**	**971.75**

The Correlation Coefficient between Influenza Deaths for people under 65 and Influenza Deaths for people 65 years or older is 0.91. This conveys a strong relationship between the two variables, which validates the hypothesis that people over the age of 65 are more likely to die from influenza than those below the age of 65.

RESULTS AND INSIGHTS

Null Hypothesis: People 65 or older have a mortality rate smaller than or equal to people under the age of 65.

Alternative Hypothesis: People under the age of 65 have a mortality rate smaller than people aged 65 or older.

Interpretation: With 95% certainty, people over 65 are significantly more likely to die from Influenza relative to people under the age of 65.

REMAINING ANALYSIS & NEXT STEPS

Remaining Analysis

- Based on the data, determine when the influenza season is, to aid in the effective assignment of medical staff.
- Further examine both data sets to determine which states are in most need of staffing and when.

Next Steps

- Create a data visualization design checklist.
- Create Tableau storyboard to aid in final presentation to all stakeholders.

APPENDIX

Project Overview

https://images.careerfoundry.com/public/courses/data-immersion/A1-A2_Influenza_Project/A1-A2-data-immersion-project-brief.pdf

Hypothesis Development

The hypothesis was developed by asking the following questions:

- When is the influenza season? How long is it?
- What percentage of the US population is considered at risk? How is that determined?
- How many nurses, physicians and doctors are there in the US? How many are needed per state?
- How does population size affect the number of cases and transmission rate for influenza in each area/city/state?
- What other health factors contribute to patients' "at risk" status? (For example: obesity, diabetes, heart disease, lifestyle, etc.)

Data Overview

CDC Influenza Deaths

Variables: State, State Code, Year, Month, Month Code, Ten-Year Age Groups, Ten-Year Age Groups Code, Deaths

Count for each variable: 66096

	Minimum	***Maximum***	***Mean***
Deaths	0	512	6.29

US Census Data

Variables: County, Year, Total population, Male Total population, Female Total population, Under 5 years, 5 to 9 years, 10 to 14 years, 15 to 19 years, 20 to 24 years, 25 to 29 years, 30 to 34 years, 35 to 39 years, 40 to 44 years, 45 to 49 years, 50 to 54 years, 55 to 59 years, 60 to 64 years, 65 to 69 years, 70 to 74 years, 75 to 79 years, 80 to 84 years, 85 years and over

Count for each variable: 28985

	Minimum	***Maximum***	***Mean***
Total Population	41	10105722	97842.18
Male Population	23	4979641	48121.57
Female Population	15	5126081	49720.61
Under 5 YO	0	733897.13	6309.26
5 - 9 YO	0	665400.83	6377.7
10 - 14 YO	0	724111.83	6496.23
15 - 19 YO	0	753656.52	6804.06
20 - 24 YO	0	777987.83	6908.61
25 - 29 YO	0	831276	6714.16
30 - 34 YO	0	762619	6436.46
35 - 39 YO	0	753467.72	6344
40 - 44 YO	1.001	733897.13	6613.69
45 - 49 YO	0	704717.78	6900.9
50 - 54 YO	0	683886.54	6921.97
55 - 59 YO	2.98	628513	6292.77
60 - 64 YO	0	535357	5416.44
65 - 69 YO	0	415243	4189.2
70 - 74 YO	0	295420	3128.96
75 - 79 YO	0	215181	2394.11
80 - 84 YO	0	161647	1821.32
85+ YO	0	177493	1774.31

Integrated Data

Variables: State & Year Key, Influenza Deaths (0 to 85+), US Census Population (0 to 85+), Mortality Rate (0 to 85+)

Count for each variable: 459

Integration process:

- Created a combined key for State and Year
- Matched data to State & Year Key for both the Influenza Deaths and US Census Data
- Created “0 to 64” and “65+” columns for both Influenza Deaths and US Census Data
- Created Mortality Rate section by dividing number of deaths with population.

Results and Insights

Results of the two sample, one-tailed t-test:

t-Test: Two-Sample Assuming Unequal Variances

	Mortality Rate Under 65	***Mortality Rate 65 or Older***
Mean	0.000269513	0.001316858
Variance	7.6196E-08	2.74967E-07
Observations	458	458
Hypothesized Mean Difference	0	
df	692	
t Stat	-37.82405035	
P(T<=t) one-tail	6.954E-171	
t Critical one-tail	1.647058574	
P(T<=t) two-tail	1.3908E-170	
t Critical two-tail	1.963398024	

The t-test indicates that people aged 65 and over are significantly more likely to die from the flu compared to those under the age of 65. The p value is smaller than the significance level and for this reason we can reject the null hypothesis. This validates the belief that if you are 65 years old or older, you are more likely to die from influenza.