Influenza season-Interim Report.

1. Project Overview

- **Motivation**: Help the agency allocate staff as accurately as possible in states where flu mortality is the highest.
- **Objective:** Find out in which states additional staff are most needed.
- **Scope:** The agency recruits staff for hospitals across the country. The project will cover the upcoming flu season.

2. Research Hypothesis

If the influenza program is successful for the target age, the death rate for that age group will be less. We can identify how many of them are vulnerable and which age group is affected most by the population.

3. Data Overview

- **Us Census Population**: This is an official survey of the population of the United States that is carried out to find out how many people live there and to obtain details of such things as people's ages.
- **CDC Influenza Deaths:** Contains data on mortality from influenza by age groups over the years 2009-2017 for each state.

4. Data Limitation

- **Us Census Population** Data is collected every 10 years. In addition to omissions, erroneous enumerations, and duplications, enumerations in the wrong location can also affect the accuracy of census counts. ... For example, if someone's age is misreported on the census form, this adds one tally in error to the count for one age group and subtracts one tally in error for another. Some people may not tell truth, hide something.
- **CDC Influenza Deaths:** Some parts of data are collected via survey. There might be a typo error and those cases might not be added to the total number of Influenza cases. About 80% of data regarding the number of deaths for each age group in each state is suppressed. During the pandemic, Influenza cases might be counted as Covid-19 and vice-versa. The data presented each week are preliminary and may change as more data are received.

5. Descriptive analysis:

As I do not have data on the vaccination rate among the most vulnerable people, I would like to show a relationship between states with a higher number of vulnerable people and the mortality rate. People aged up to 5 and over 65 are the most vulnerable.

Mortality in vulnerable groups in the entire population

	Total Population of vulnerable	The death rate in vulnerable groups
Standard Deviation	6 799 372	972
Mean	5 973 847	1017
Outliners	0,22%	0%

The number of deaths caused by the flu has a strong relationship with age. The correlation is at 0,94 , so it is high. Therefore, it will be necessary to focus on this group of patients.

6. Summary

The number of patients in the groups most at risk of dying from influenza was checked and compared with the mortality rate among the most at risk. Age is strongly related to mortality. The older you get, the more likely you are to die from the flu.

- **Null Hypothesis**

  The death rate in non-vulnerable groups is higher or equal to the death rate in vulnerable groups.

- **Alternative Hypothesis**

The death rate in vulnerable groups is higher. The T-test showed that the mortality rate among those most at risk in terms of age is higher than among those at the lower risk. This proved that the null hypothesis had not been confirmed.

t-Test: Two-Sample Assuming Unequal Variances

	Variable 1	*Variable 2*
Mean	1 016,80	416,63
Variance	944 307,02	14 183,16
Observations	459,00	459,00
Hypothesized Mean Difference	0,00	
df	472,00	
t Stat	13,13	
P(T<=t) one-tail	0,00	
t Critical one-tail	1,65	
P(T<=t) two-tail	0,00	
t Critical two-tail	1,97	

7. Next Step

- Find more vaccination data among the groups most at risk of dying from flu.
- Take a closer look at the relationship between age and size of the population checking for any deviations and if it does, investigate what could be the cause.
- A visual presentation will be performed when the data collection process is complete.

8. Appendix

- **CDC influenza data Sources:** The Influenza Division at CDC collects, compiles, and analyses information on influenza activity year-round in the United States. Everything I've been working on so far is in this file.
- **US Census Bureau Population Data**: The United States census is a population census, which is distinct from the U.S. Census of Agriculture, which is no longer the responsibility of the Census Bureau. It is also distinct from local censuses conducted by some states or local jurisdictions**.**

If you want to look at what I have been working on so far, please click on the link below.

Excel file I'm working on