Is the Flu Vaccine Effective in Reducing the Number of Flu Cases?

# **Christine Hathaway Spring 2020** <https://github.com/chhathaway71/DSC-680>

# Which Domain?

1. Andriana, T., Tomic, I., Dekker, C. L., Maecker, H. T., & Davis, M. M. (2019). The FluPRINT dataset, a multidimensional analysis of the influenza vaccine imprint on the immune system. *BioRxiv*, Preprint https://doi.org/10.1101/564062. FluPRINT is a research project that investigates influenza vaccines and their imprint on the immune system. It contains information regarding the history of the project, as well as information on the data that was collected for the project database.

2. Thomas, J. (2018, November 19). The Flu: Facts, Statistics, and You. Retrieved from Healthline: https://www.healthline.com/health/influenza/facts-and-statistics#1. This site contains information regarding the types of influenza and statistics surrounding flu.

3. Gillespie, C. (2020, February 11). *This Is How Many People Die From the Flu Each Year, According to the CDC*. Retrieved from Health: https://www.health.com/condition/cold-flu-sinus/how-many-people-die-of-the-flu-every-year. This site contains information comparing statistics for the seasonal flu to the corona virus.

4. Vincent Iannelli, M. (2020, February 5). *Annual Flu Deaths Among Adults and Children*. Retrieved from Very Well Health: https://www.verywellhealth.com/deaths-from-flu-2633829. This site contains CDC estimates on flu deaths as well as flu pandemic deaths in history. It attempts to put different strains of flu virus in perspective to each other.

5. DerSarkissian, C. (2017, November 16). *What Are Your Odds of Getting the Flu?* Retrieved from WebMD: https://www.webmd.com/cold-and-flu/flu-statistics This site has statistics around the number of people who get the flu, as well as the incubation times for flus.

# Which Data?

The dataset I will be using is the FluPRINT database. “The FluPRINT is the name of a unified database for a large-scale study exploring novel cellular and molecular underpinnings of successful immunity to influenza vaccines. It contains information on more than 3,000 parameters measured using mass cytometry, flow cytometry, phosphorylation-specific cytometry (phospho-flow), multiplex ELISA, clinical lab tests (hormones and complete blood count), serological profiling with hemagglutination inhibition assay, and virological tests. The dataset represents fully integrated and normalized immunology measurements from 747 individuals from eight clinical studies conducted between 2007 to 2015 at the Human Immune Monitoring Center of Stanford University. The dataset represents a unique source in terms of value and scale, which will broaden our understanding of influenza immunity” (Andriana, Tomic, Dekker, Maecker, & Davis, 2019). The FluPRINT dataset, a multidimensional analysis of the influenza vaccine imprint on the immune system. *BioRxiv*Preprint <https://doi.org/10.1101/564062>.

The data can be downloaded here <https://zenodo.org/record/3222451#.Xml9WKhKhPZ>

# Research Questions? Benefits? Why analyze these data?

I want to examine whether there are any indicators as to who will become infected with influenza, and what impact, if any, flu vaccinations have. I will also research if there are any indicators as to who could be hospitalized with the flu. Knowing whether or not vaccinations help stem the flu could encourage more people to receive one.

# What Method?

First, I will use R and Python to perform exploratory data analysis, as well as clean the data. The data will need to be binned by age group, as well as type of vaccines, whether a person was infected by the flu, and if they were hospitalized as a result. I plan to use logistic regression, as the data is categorical in nature. I will look for correlations between the variables and the outcome of a flu infection and the vaccine.

# Potential Issues?

Some challenges include insufficient data, missing data, or incorrect data. There may not be sufficient data to make conclusions regarding the impact of vaccines on the flu. Or, the dataset may be too big to run on my local computer, which could cause large processing time. One of the biggest obstacles is that the dataset may not contain health information, which could be one of the biggest impacts on whether or not a person is susceptible to getting the flu.

# Concluding Remarks

With the current looming crises of the Corona virus, many questions come to mind regarding viruses in general. Influenza in particular impacts millions of individuals each year, including 5% to 20% of the U.S. population. Some cases can even lead to death, with 58% of these occurring in adults ages 65 and over (Thomas, 2018).Others at high risk include young children, pregnant woman, people with weakened immune systems, people with chronic illnesses, as well as people who are very obese, with a body mass index (BMI) of 40 or higher (Thomas, 2018).

This year, the focus has been on the Corona virus, and rightly so. We have had a number of flu pandemics throughout history, such as Russian flu in 1889 with 1 million deaths, and most recently H1N1 flu in 2009 with 203,000 deaths worldwide (Vincent Iannelli, 2020) Hopefully, we can learn from these historical pandemics. By identifying who is most susceptible to a virus, we can help protect those that are most vulnerable. This project will focus on influenza and vaccination data, and ask questions such as who was infected or hospitalized with influenza? Are any indicators or correlation that will help identify or predict whether a flu vaccination is beneficial in preventing the flu? As more people take precautions to protect themselves, we can limit the spread of infectious viruses around the world.

# References

Andriana, T., Tomic, I., Dekker, C. L., Maecker, H. T., & Davis, M. M. (2019). The FluPRINT dataset, a multidimensional analysis of the influenza vaccine imprint on the immune system. *BioRxiv*, Preprint https://doi.org/10.1101/564062.

Thomas, J. (2018, November 19). *The Flu: Facts, Statistics, and You*. Retrieved from Healthline: https://www.healthline.com/health/influenza/facts-and-statistics#1

Vincent Iannelli, M. (2020, February 5). *Annual Flu Deaths Among Adults and Children*. Retrieved from Very Well Health: https://www.verywellhealth.com/deaths-from-flu-2633829