



Flu Shot Learning: Predict Seasonal Flu Vaccines

BY

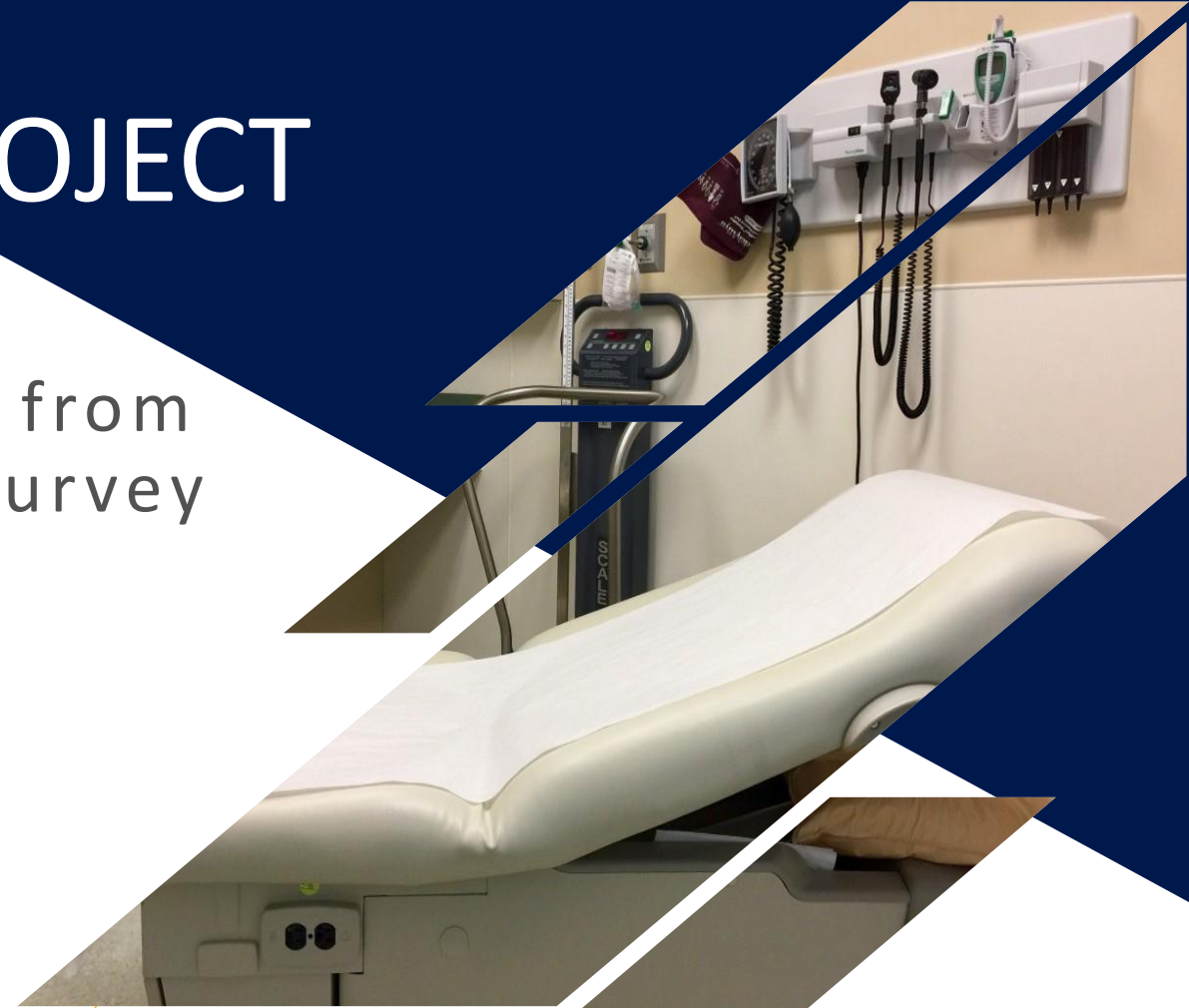
SEGO MICH

[This Photo](#) by Unknown Author is licensed under [CC BY-SA-NC](#)

OVERVIEW OF THE PROJECT

Project Overview

Analysis of Vaccination Patterns from
the National 2009 Seasonal Flu Survey



Business problem

- A vaccine for the H1N1 flu virus became publicly available in October 2009. In late 2009 and early 2010, the United States conducted the National 2009 H1N1 Flu Survey.
- This phone survey asked respondents whether they had received the H1N1 and seasonal flu vaccines, in conjunction with questions about themselves.
- These additional questions covered their social, economic, and demographic background, opinions on risks of illness and vaccine effectiveness, and behaviors towards mitigating transmission.
- A better understanding of how these characteristics are associated with personal vaccination patterns can provide guidance for future public health efforts.

Objectives

Main Objective

To analyze the demographic characteristics of respondents, including age, education, income, employment, and household composition.

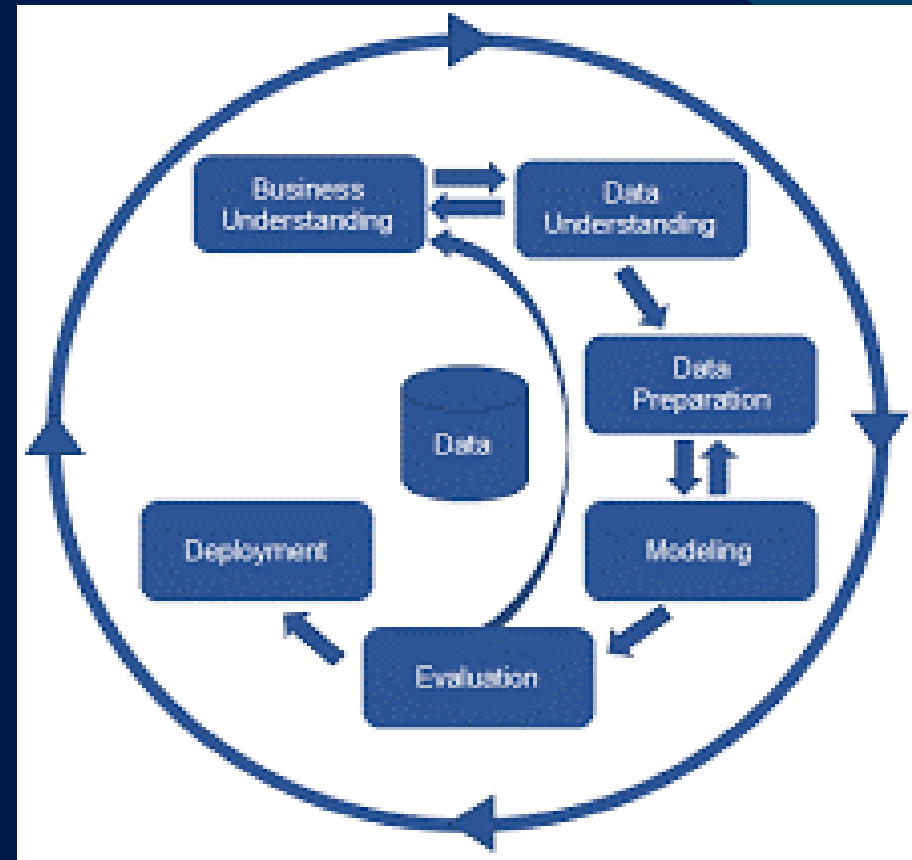
Specific Objectives

- 1.Age Distribution** – Examine the age group distribution among respondents.
- 2.Educational Attainment** – Analyze the levels of education across different respondents.
- 3.Income and Employment Status** – Assess variations in income levels and employment status.
- 4.Household Composition** – Investigate household structure based on marital status, homeownership, and number of adults/children.
- 5.Geographic Demographics** – Identify demographic variations across different regions.

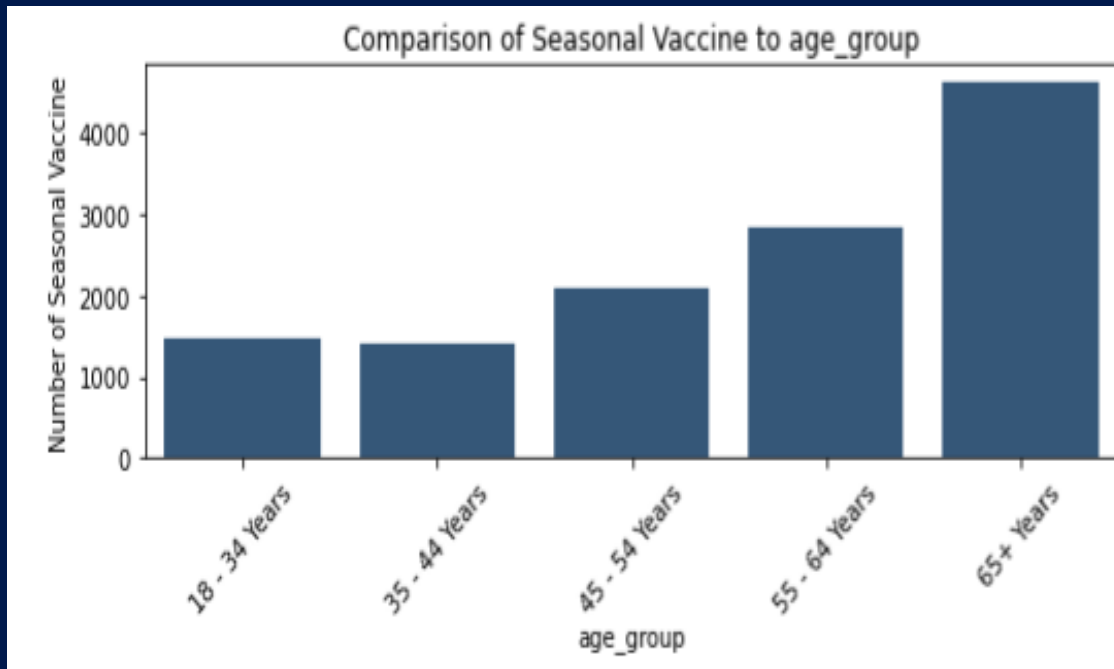
Methodology

CRISP-DM

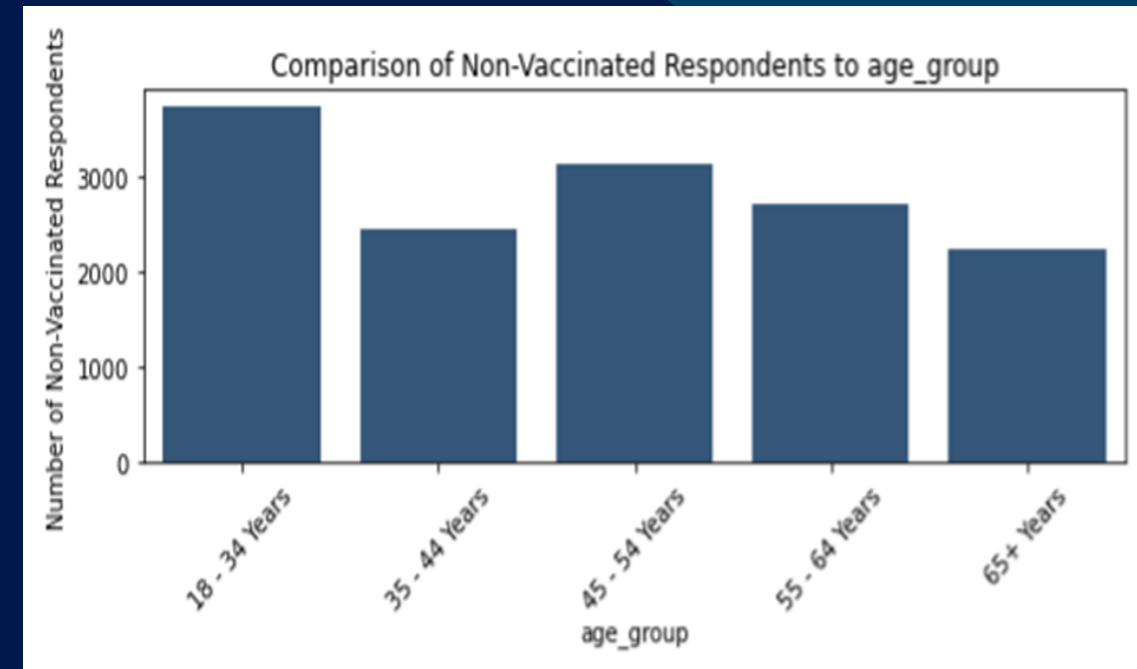
- ✓ Business Understanding
- ✓ Data Understanding
- ✓ Data Preparation
- ✓ Modeling
- ✓ Evaluation



Comparison of seasonal vaccine per age group

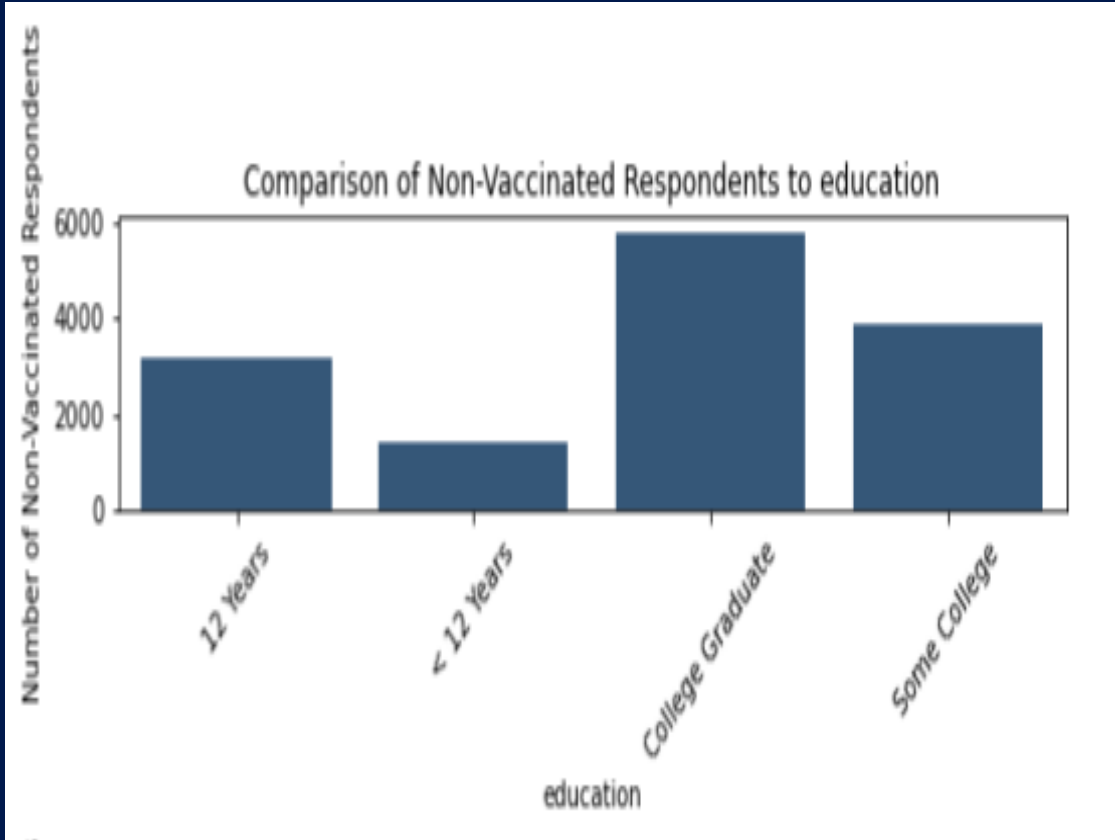


65+ Years are most vaccinated

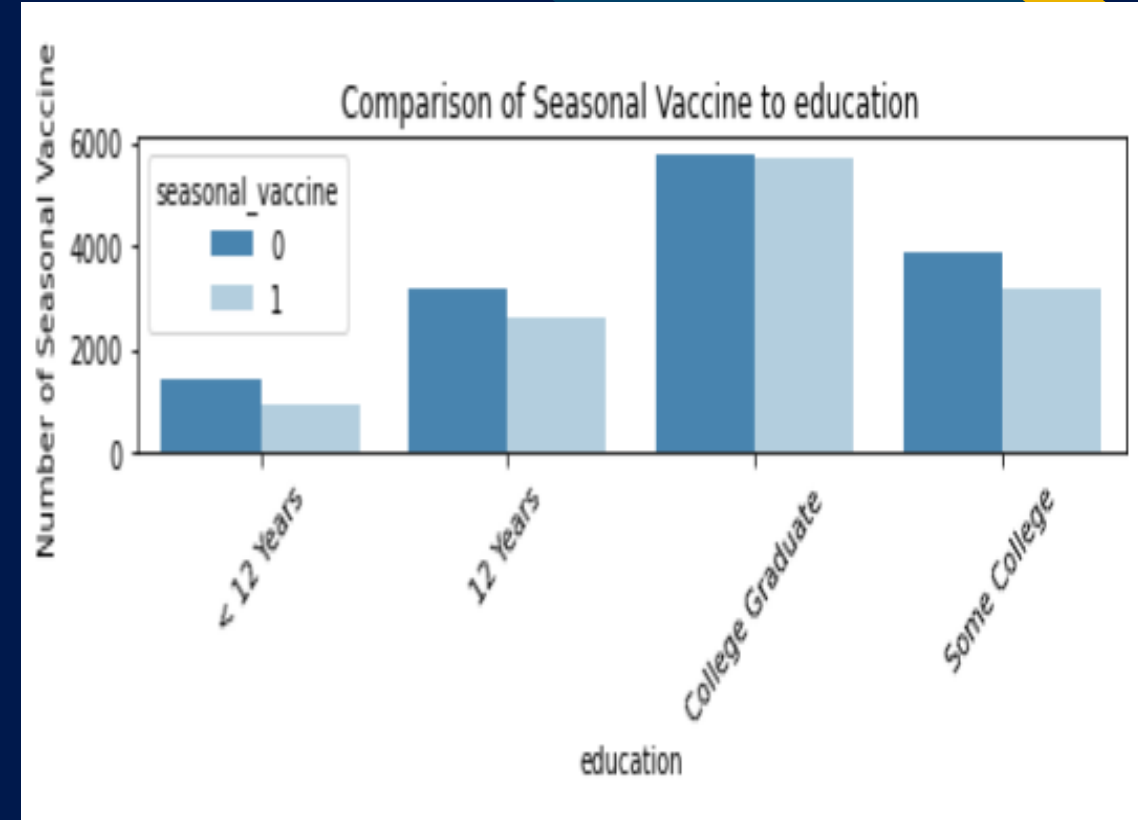


18 – 34 Years are most non vaccinated

Comparison of seasonal vaccine per education level

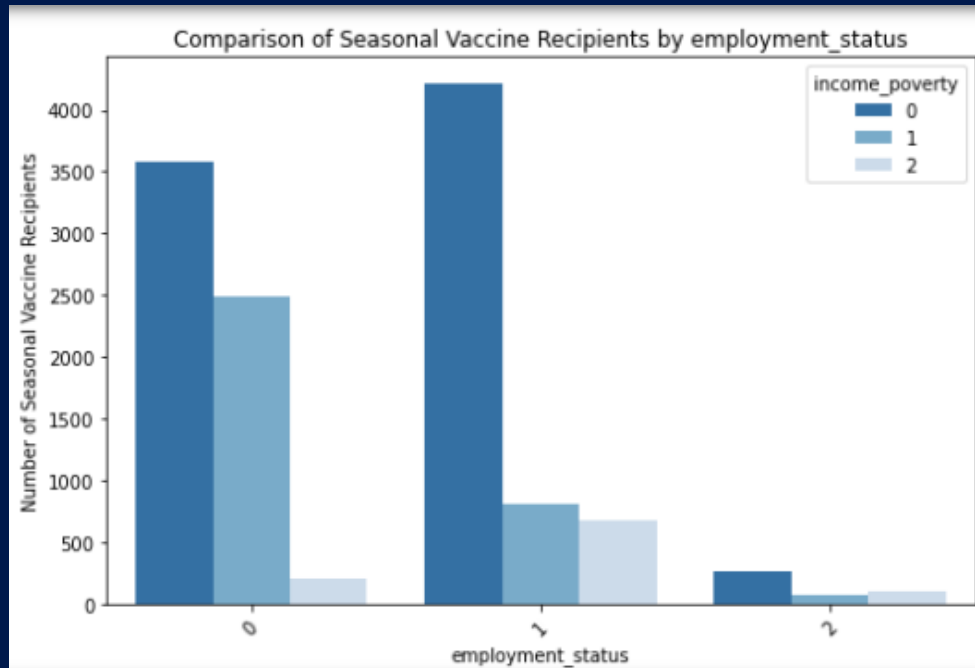


In the survey College Graduate are most vaccinated

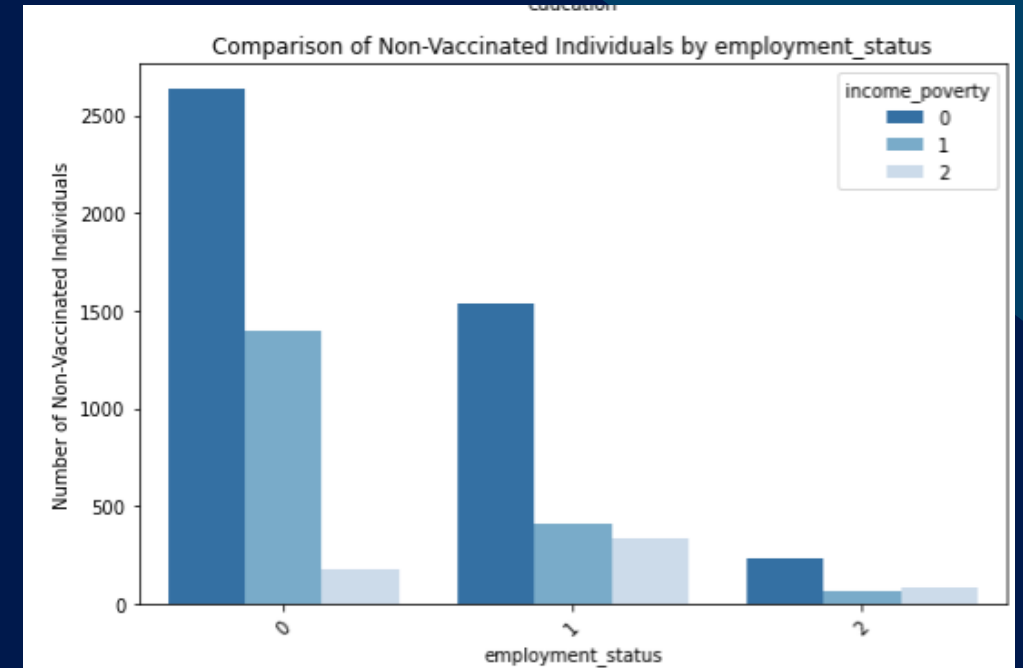


All education levels are more non vaccinated than vaccinated.

Comparison of seasonal vaccine per income poverty and employment status

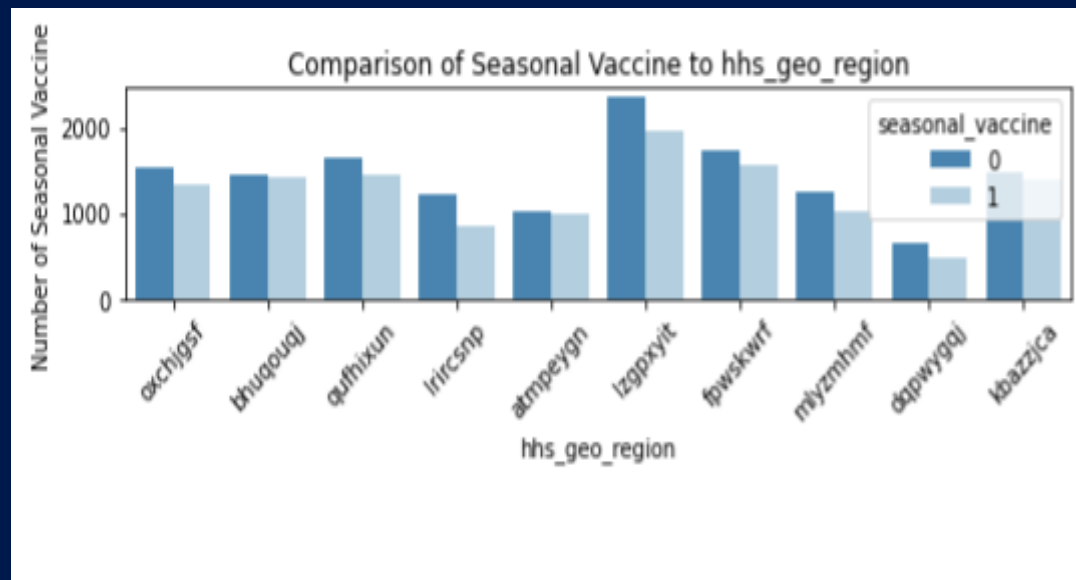


In the survey College Graduate are most vaccinated

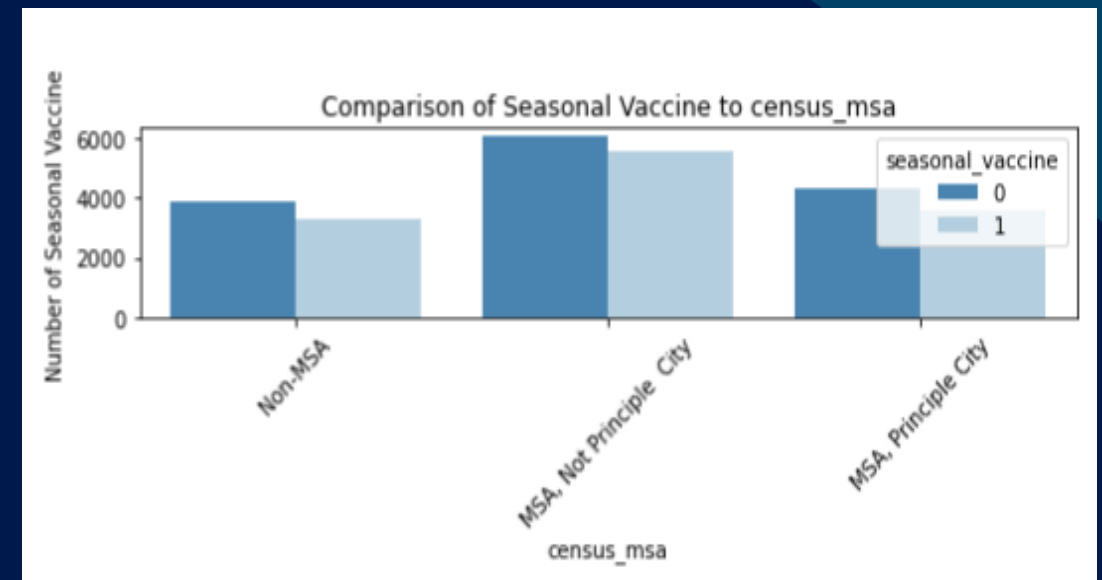


College graduate are most non vaccinated also.

Comparison of seasonal vaccine per household composition

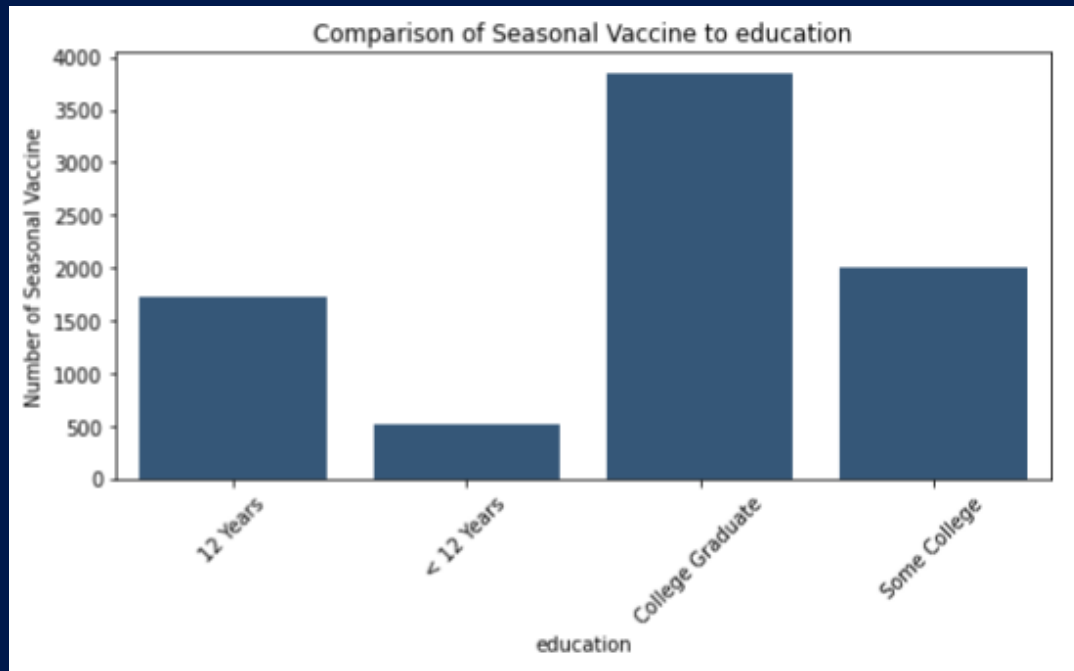


In the survey all regions most respondents do not take most vaccines as those who do.

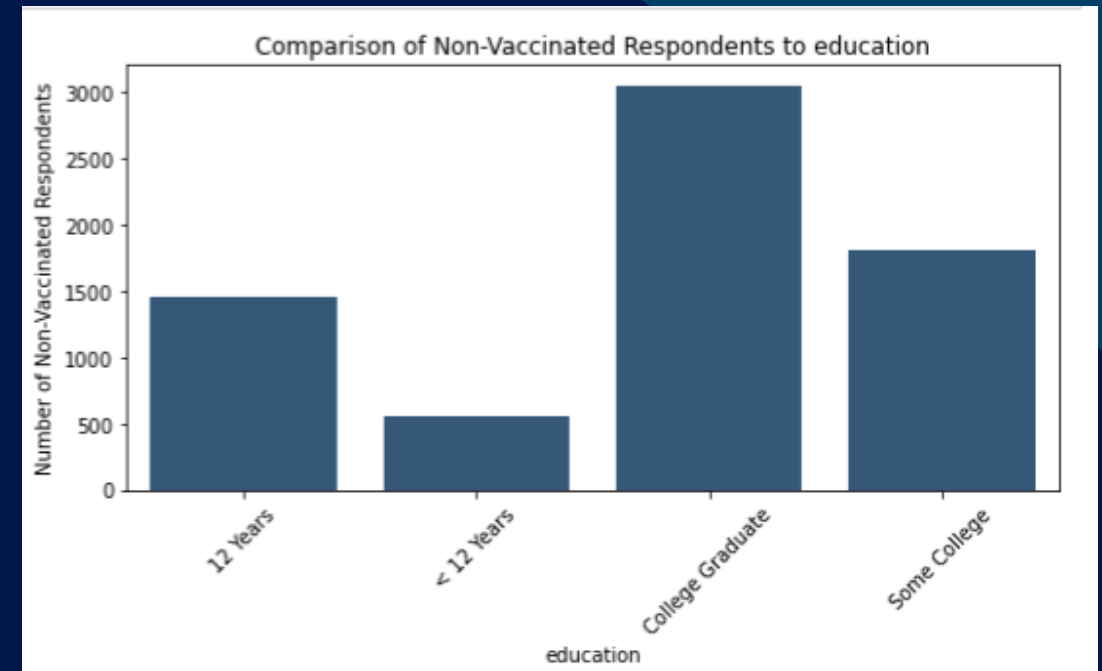


In the survey all regions most respondents do not take most vaccines as those who do.

Comparison of seasonal vaccine per geographic composition



In the survey College Graduate are most vaccinated



College graduate are most non vaccinated also.

Tools and Libraries

MATHEMATICAL	VISUALIZATION	PREPROCESSING	MODELLING	METRICS
Pandas	Seaborn	LabelEncoder	Logistic Regression	Accuracy
Numpy	Matplotlib		Decision Tree Classifier	
			Random Forest Classifier	
			Support Vector Classifier	
			Naive Bayes	

Conclusion

- **Logistic Regression:** 0.62 Accuracy
- **Decision Tree Classifier** 0.59 Accuracy
- **Random Forest:** 0.60 Accuracy
- **Support Vector Machine (SVM):** 0.62 Accuracy
- **Naive Bayes:** 0.62 Accuracy

Logistic Regression, Naive Bayes and SVM achieved the highest accuracy at **0.63**, making it the best-performing model. Decision Tree Classifier had the lowest accuracy at **0.59**. Random Forest performed similarly with accuracies of **0.62**, respectively.

[This Photo](#) by Unknown Author is licensed under [CC BY](#)



ML

FLU SHOT LEARNING

Thank You.



Sego Mich



2547XXXXXX



mich.sego@student.moringaschool.com



https://github.com/Sego-Mich/mich_phase_3_final_project