Humana-Mays Healthcare Analytics

2021 Case Competition

COVID Vaccine Hesitation Member Prediction

And

Potential Solutions with Segmentation

October, 2021

Executive Summary

The U.S. Centers for Disease Control and Prevention reports that the covid-19 cases have reached 44 million while the full vaccination rate still lingers around 56% in the US. As a health service provider, Humana strives to provide vaccination opportunities and remove barriers for the most vulnerable and underserved populations. To address this business challenge, we built a predictive XGBOOST model with AUC score of 0.6728 to identify members hesitant to vaccination. Then we drilled down into the key drivers and identified sub-segments with knowledge from secondary research and PCA and K-means quantitative methods. Targeted outreaches to 5 groups of members are proposed and prioritized to improve the vaccination rates of Humana problems and to ensure scalability and reach profitability of over 4 million dollars.

Table of Contents

1 Introduction

1.1 Background

The covid-19 pandemic has been one of the greatest health crises in recorded human history. According to U.S. Centers for Disease Control and Prevention (CDC) (2020), the total covid-19 cases in the US have reached 44 million by October 9th, 2021. Fortunately, the light at the end of the tunnel is getting closer thanks to rapid medical advances. Among all approaches, mass vaccination is the most effective, just as WHO put it, “Equitable access to safe and effective vaccines is critical to ending the COVID-19 pandemic”. Researchers have found that on July 25, by which time the delta variant had become dominant, “infection and hospitalization rates among unvaccinated persons were 4.9 and 29.2 times, respectively, those in fully vaccinated persons” (Griffin, 2021).

As of today, 56% of the US population is fully vaccinated (CDC). Increasing the Covid-19 vaccination rate continues to be a priority both for individual health and for a thriving society. Specifically, it is optimal to identify which of the population are most hesitant or resistant to get the vaccine and dive deep into the reasons hidden behind. Then all healthcare providers can design targeted outreaches to address different problems efficiently and effectively.

1.2 The Humana Analytics Competition

1.2.1 The Business Issue

Humana’s mission is to help people achieve lifelong wellbeing. During pandemic, Humana strives to provide vaccination opportunities for the most vulnerable and underserved populations.

Murphy et al. (2021) found that vaccine hesitant respondents differentiated on sociodemographic and health-related factors but geographic differences existed. According to Forbes report (Hart, 2021), a polling carried out by Kaiser Family Foundation revealed that those who were not willing to get vaccination included 21% of youngsters aged from 18 to 29, 46% of republicans, 37% of agricultural workers, 33% of South Dakota residents, and 22% of people with lower than a college level degree. And the primary reasons for denying vaccinations lied in skepticism, concern on side effects, and distrust. Those information were taken into consideration when we conducted the segmentation.

In this case competition, we created a classification model to predict which members were likely to be hesitant with provided data by Humana and supplementing public data from CDC. We then examined the most important features affecting vaccination status and categorized the underserved population into five major groups and proposed targeted solutions throughout their vaccination journey. Lastly we prioritized the most vulnerable underserved groups and the most effective approach.

1.2.2 Key Performance Indicators

Our main challenge is to build an accurate predictive model and propose actionable business solutions. While the accuracy of the model serves as base for the challenge and has proved effective with ROC-AUC score, our KPIs focus on evaluating the business outcomes.

1. **Vaccination rate.** Vaccination rate is the primary indicator of the effectiveness of the segmentation and the proposed solutions. Increase in vaccination rate denotes the success of the solutions.
2. **Scalability.** After verifying the effectiveness of the solutions, the scalability must be evaluated to ensure the solutions can expand and improve overtime without too much

friction. In a business context, an effective but not scalable solution does not add much value.

3. **Profitability implication.** Covid-19 vaccines are paid with taxpayer dollars and therefore do not increase costs for Humana, but wider vaccination can potentially save significant health care costs.

2 Data Preparation

2.1 Analytical Tools Used

All of our analysis shown in the following parts are coded in Python and implemented by Jupyter Notebook. We used pandas and numpy to explore our data and excluded useless features, Scikit-learn and XGBoost to preprocess our data, build and evaluate our models, and matplotlib to create visualization so as to better understand feature importance and performances of our ML models.

2.2 Preliminary Data Analysis

2.2.1 Data Description

There are 367 columns and 974,842 members in the training dataset. And the columns can be categorized into the following 8 groups:

· Medical Claims Features (132 features)

· Pharmacy Claims Features (70 features)

· Lab Claims Features (2 features)

· Demographics/Consumer Data (110 features)

- Credit Data (16 features)
- Condition Related Features (10 features)
- CMS Features (7 features)
- Others (18 features)

Moreover, these data are either in object type or numeric types. However, there are features that are saved under the wrong data types, which would be further discussed in the following sections.

2.2.2 Dropping Unhelpful Features

In the preliminary phase, we decided to delete features under the following two criteria, standard deviation == 0 and proportion of null value > 30%.

Features would be useless once the standard deviation of that whole column is 0, so we dropped these columns. Moreover, columns with a large proportion of null values would not be helpful in our predictive models, especially columns “lang_spoken_cd” and “mabh_seg”, which contained 74% and 65% of null value respectively. However, we did not want to give up on too many features, as each column itself could be a potentially influential feature. Therefore, we set the threshold as 30% and dropped columns contain more than 30% null values.

2.2.3 Dropping Unhelpful Records

After dropping unhelpful columns, we noticed that there might be some records consisting of certain numbers of null values. Therefore, we used a quantile function in pandas to drop 10% of the records that contained most null values.

2.2.4 Data Cleaning

There were still a large number of null values within the dataset, not only in the format of pd.NA, but also in the format of string "*". We used the Simple Imputer in Sckit-learning to convert all the null values into pd.NA.

Additionally, as we selected XGBoost as our model (we would introduce our reason in the following section), we decided not to deal with NA values in our training set, because XGBoost has its own way of treating missing values.

2.3 Feature Engineering

2.3.1 One Hot Encoding

As categorical columns are inevitable in datasets, how to convert categorical data into numeric values would be a key problem. Fortunately, with the help of One Hot Encoder within Scikit-Learning, we were able to convert categorical features, as well as numerical indices (each numeric index has different meaning) columns, into dummy variables. For example, variable "race_cd" has six values: 0 = Unknown, 1 = White, 2 = Black, 3 = Other, 4 = Asian, 5 = Hispanic, 6 = N. American Native. Therefore, instead of one single column of “race_cd”, there will be 7 new dummy variables such as “race_cd = 0”, “race_cd = 1”, “race_cd = 2”, etc.

After the conversion, we expanded our columns from 334 to 846. Additionally, we did not column “zip_cd” into OneHotEncoder since there are too many unique zip codes.

race_cd
0
1
3
5
4

Unknown	White	Black	Other	Asian	Hispanic	N. American Native
1	0	0	0	0	0	0
0	1	0	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	0	1	0
0	0	0	0	1	0	0

Figure 2.1. An Example of OneHotEncoder Conversion

2.3.2 Adding Features

We believed that whether a client is hesitant of covid vaccine could be influenced by their neighborhoods’ political beliefs. Therefore, we incorporated the US President Election Result in 2020 by each County which was supplied by Havard Dataverse. We merged the most likely political belief of each client’s neighborhood with their own zip codes.

3 Modeling

3.1 Model Selection

We chose XGBOOST over other decision-tree based models to help us model the data because 1) our data is very unbalanced with regards to the ground truth variable as out of a total of 894,340 observations we have, only 159,921(17.8%) are unvaccinated and 736,509(82.2%) are vaccinated. XGBOOST is particularly good at dealing with unbalanced dataset in ways that it gives more preferences and weights to nodes where anomalies take place; 2) in terms of the speed, it fully utilized all computational units available on a machine, which is far more efficient than a typical gradient boosting method; 3) XGBOOST has smart algorithms that decide the way to impute empty values, which is what we need considering the fact that there are many missing values across the whole dataset.

3.2 Model Tuning

We used grid search and cross-validation to find the hyperparameters that yield the best result. There are 3 fixed hyperparameters that we set: objective, nthread, and tree_method. The objective referred to the learning task. In our case, it was set as "binary:logistic" because the prediction label was categorical and had 2 classes. We wanted to have the probability for both classes that each observation belonged to.The "nthread" referred to the number of threads for parallel computing. Since we were using a 6-core server, we set it as 6. We chose "hist" as the value to be fed into the "tree-method" field because we had a larger dataset and this method was faster with a better performance.

There were 6 hyperparameters we were interested in: learning_rate, max_depth, subsample, colsample_by_tree, gamma and n_estimators. The learning rate referred to the step size shrinkage used in learner updating. We chose 0.04. The max_depth referred to the maximum depth of a tree, and we chose 10. The subsample was the portion of sample that XGBOOST chose to grow a tree, which greatly prevented overfitting. We set it at 0.7. The colsample_bytree was the subsample ratio of columns when constructing a tree, which forced the model fitter to learn about weak features, we set it at 0.8. Gamma was the minimum loss reduction required to make a further split, which functioned as overfitting prevention as well, we set it as 0.02. Lastly, for the number of boosting rounds, we set n_estimators to 100, since we observed no large improvements after 100 epochs of training.

3.3 Final model

After having 5 cross validations, we achieved a mean test score of 0.6728 using the AUC metric.

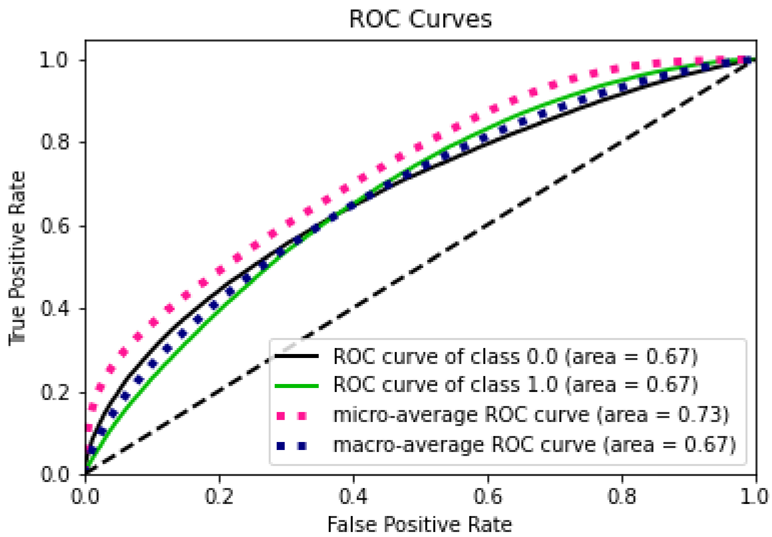

Figure 3.1. AUC of model in Cross Validation set

3.4 Feature Importance with SHAP

3.4.1 Summary analysis of top features

In order to derive actionable business insights, it is crucial to understand the effect and magnitude of how each feature affects our prediction. Therefore, the SHAP method was employed. Based on coalitional game theory, SHAP values could show how much a given feature changed our prediction and was widely used for interpretation. A feature with a higher SHAP value has a greater impact on the prediction.

The 20 most important features were plotted (Figure 3.2).The upper position indicated a higher importance of a feature. And they fell into six categories: demographics/consumer data (8 features), pharmacy claims feature (5 features), CMS feature (3 features), clinical care feature (2 features), and condition related features (1 feature). As shown below, the SHAP value on the horizontal axis and the colored feature value helped us interpret the result. For example, est_age represents a member age and it has a higher SHAP value for older people, indicating that the probability of an older member classified as vaccinated is higher (in other words, older people are less likely to be classified as hesitant to get vaccination). Among the top 10 features, five features are positively related to the probability of being hesitant to vaccination and the others are negatively related. We would further analyze the top 10 features in the next part.

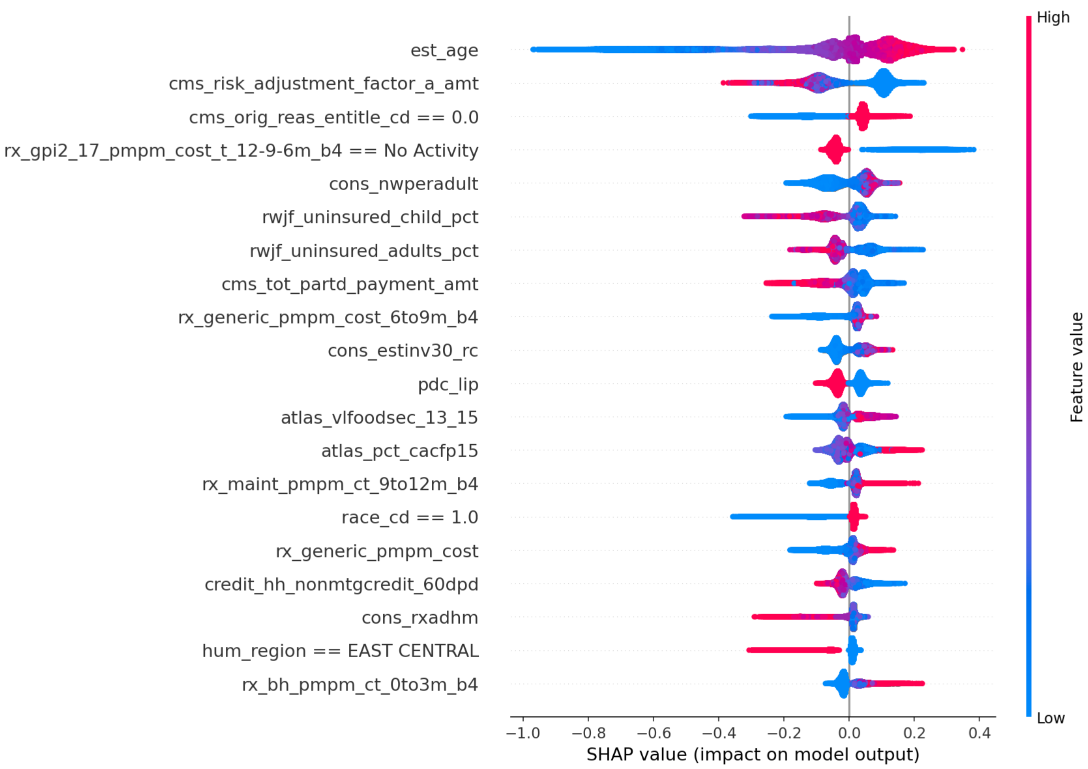

Figure 3.2. Top 20 most important features with SHAP value

Table 3.1. Top 20 most important features and descriptions

Feature	Description
est_age	Member age {calculated using est_bday, relative to score/index date}
cms_risk_adjustment_factor_a_amt	Risk Adjustment Factor A Amount
cms_orig_reas_entitle_cd == 0.0	Old Age Survivors Insurance (OASI) is the original reason for entry into Medicare.
rx_gpi2_17_pmpm_cost_t_12-9-6m_b4 == No Activity	trend of cost per month of prescriptions related to VACCINES drugs in the past sixth to ninth month versus ninth to twelfth month prior to the score date {Based on GPI2 grouping}
cons_nwperadult	Net Worth Per Adult
rwjf_uninsured_child_pct	Clinical Care - Percentage of children under age 19 without health insurance
rwjf_uninsured_adults_pct	Clinical Care - Percentage of adults under age 65 without health insurance
cms_tot_partd_payment_amt	Total Part D Payment Amount
rx_generic_pmpm_cost_6to9m_b4	cost per month of prescriptions related to generic drugs in the past sixth to ninth month prior to the score date
cons_estinv30_rc	Estimated Household Investable Assets Recoded
pdc_lip	proportion of days covered for prescriptions related to hyperlipidemia in the past one year
atlas_vlfoodsec_13_15	Household very low food security (%, three-year average), 2013-15
atlas_pct_cacfp15	Child & Adult Care (% pop)
rx_maint_pmpm_ct_9to12m_b4	count per month of prescriptions related to maintenance drugs in the past ninth to twelfth month prior to the score date
Race_cd == 1	Code indicating a member's race {0 = Unknown, 1 = White, 2 = Black, 3 = Other, 4 = Asian, 5 = Hispanic, 6 = N. American Native}
rx_generic_pmpm_cost	cost per month of prescriptions related to generic drugs in the past one year
credit_hh_nonmtgcredit_60dpd	% HH Non-Mortgage Loan Accts 60+ Days Past Due
cons_rxadhm	RX Adherence - Maintenance
Hum_region == EAST CENTRAL	Member geographic information - Humana Region
rx_bh_pmpm_ct_0to3m_b4	count per month of prescriptions related to behavioral health drugs in the past three months prior to score date

3.4.2 Top 10 Features: Individual Dependency Analysis

Demographic/consumer--age:

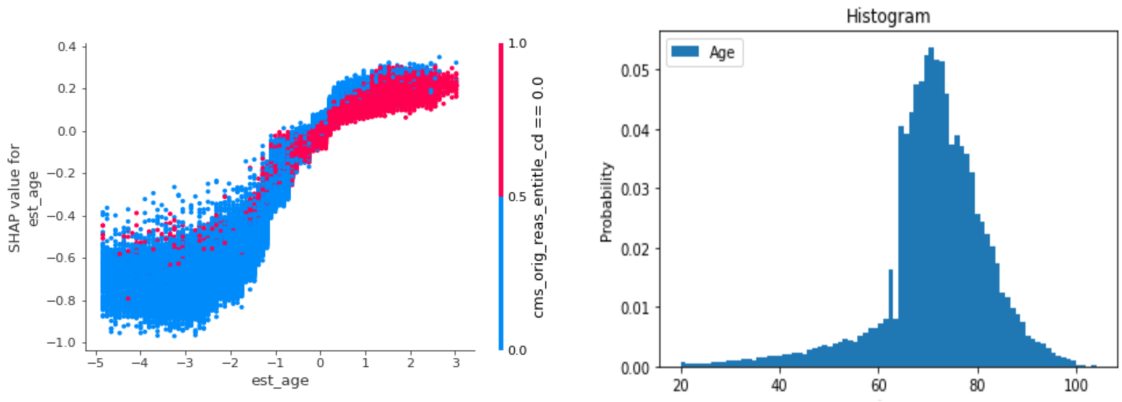

Figure 3.3. Dependency figures of demographic/consumer--age with SHAP

In terms of age, the overall trend is that with the increase in age, the probability of getting vaccinated is higher. And age interacts most with feature Cms_orig_reas_entitle_cd. Cms_orig_reas_entitle_cd = 0 means that Old Age Survivors Insurance (OASI) is the original reason for entry into Medicare. Specifically, when the age exceeds the average age of 72, having an OASI increases the probability of vaccination hesitation.

Demographic/consumer--financial:

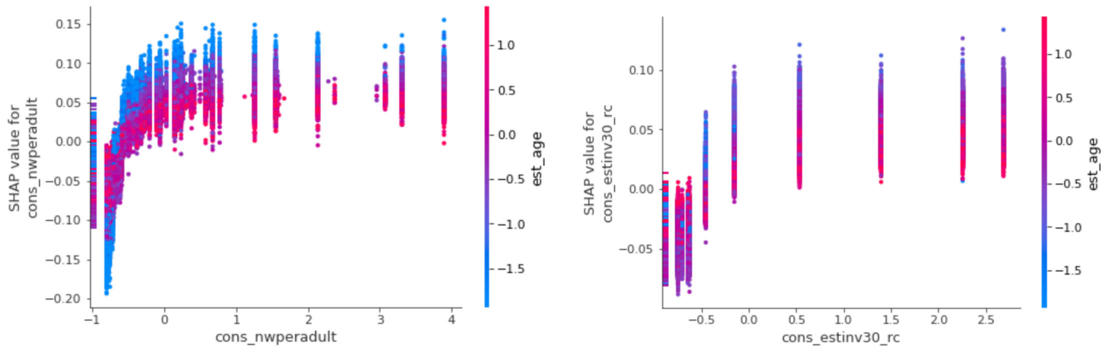

Figure 3.4. Dependency figures of demographic/consumer--financial with SHAP

Net worth per person(cons_nwperadult) and household investable assets are both key indicators of one's financial status. A higher net worth and a higher household investable assets is associated with a higher probability of being vaccinated.

Pharmacy Claims Features:

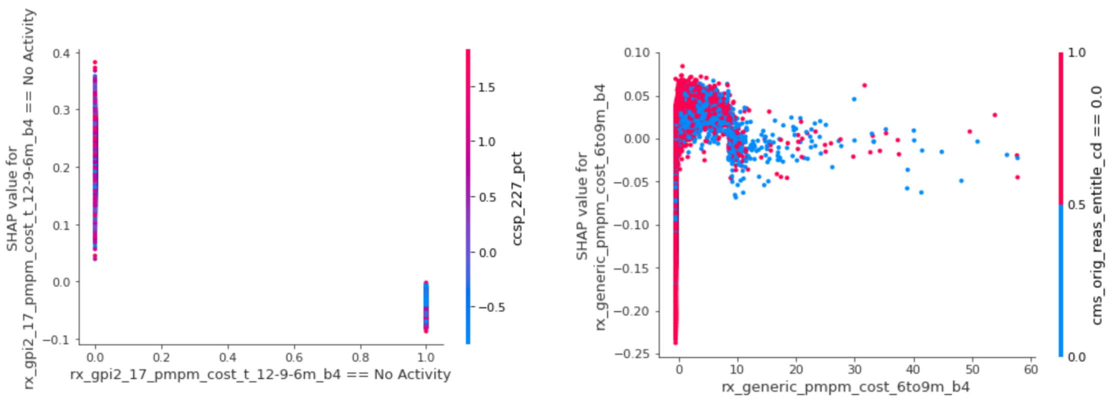

Figure 3.5. Dependency figures of pharmacy claims features with SHAP

This factor means that if a person does not incur any prescription cost associated with vaccine drugs, he would most likely be not vaccinated. This is likely due to the reason that a person without being vaccinated would not need to get any vaccine side effect drugs. Besides, having a higher cost of prescriptions on generic drugs is less associated with being a vaccine hesitant.

CMS features:

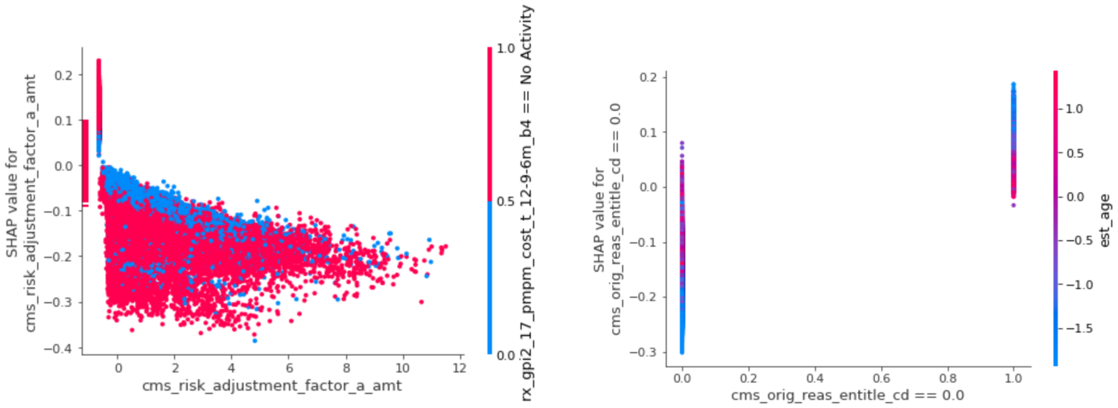

Figure 3.6. Dependency figures of CMS features-I with SHAP

Risk adjustment factor A amount is a score to evaluate a person's health status, but its relationship with the target is not clear because of the sparse distribution. When this factor is more than 8 standard deviations away from the mean value, it converges and indicates a lower probability of getting vaccinated. The effect of OASI is explained in age.

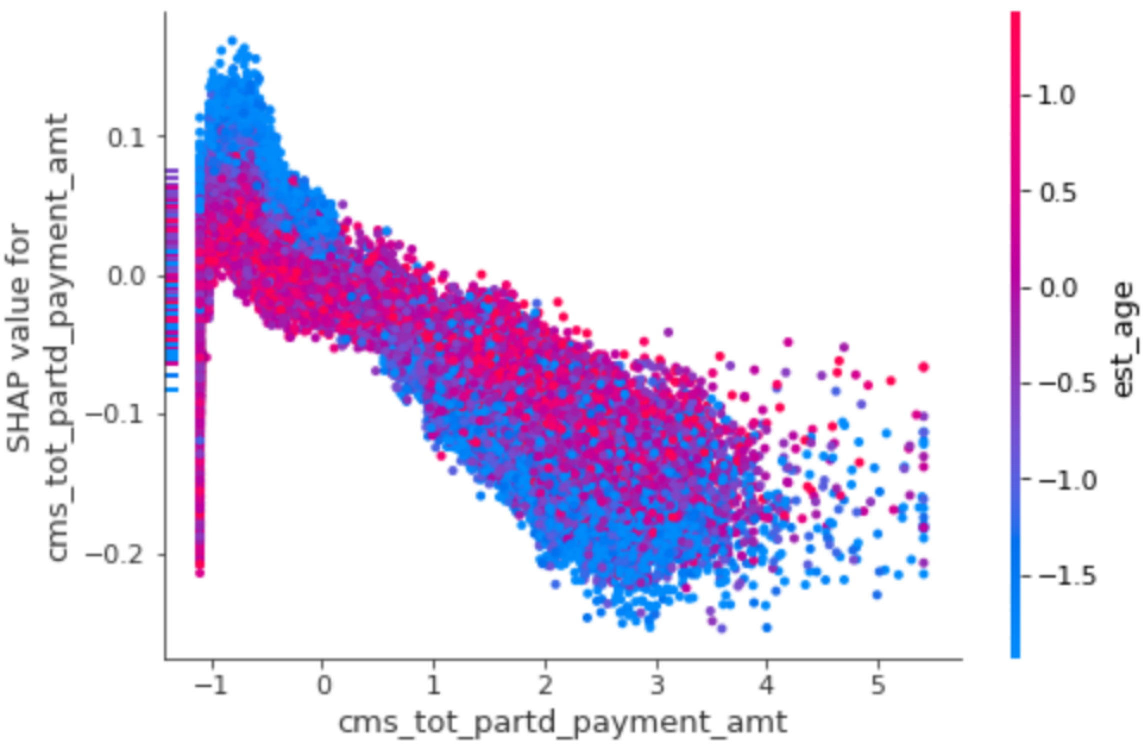

Figure 3.7. Dependency figures of CMS features-II with SHAP

A higher total part D payment is associated with a higher probability of being vaccine hesitant. Similar to OASI, this feature interacts most with age. Below the average payment of 137, older members are more likely to be vaccine hesitant; above the average payment, older members are less likely to be vaccine hesitant.

Clinical care:

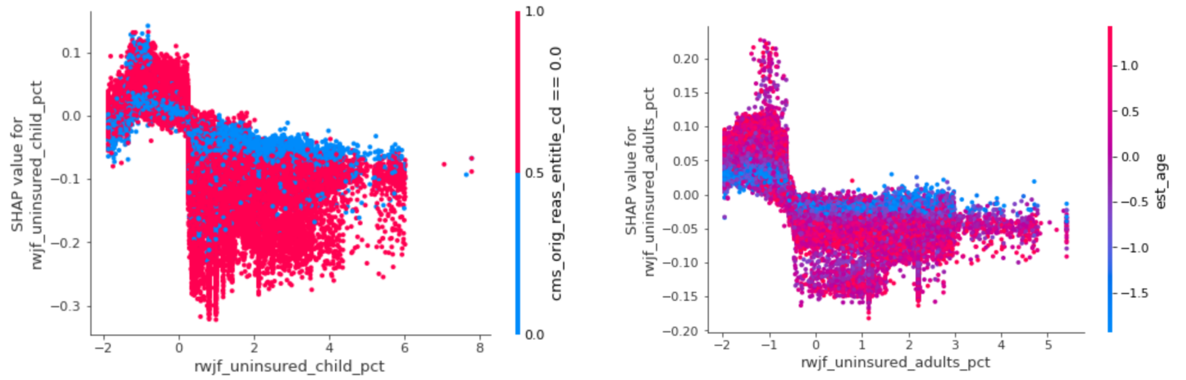

Figure 3.8. Dependency Figures of linical care with SHAP

These two features fall into the clinical care category and reflect regional level. They show that if a member comes from a county with a higher percentage of children under age 19 without health insurance and of adults under age 65 without health insurance, then he is more likely to be vaccine-hesitant.

4 Business Implications

4.1 Analysis Summary

By understanding key contributors to the possibility of vaccination from predictive modelling and SHAP analysis, we would further investigate the reasons behind the COVID vaccine hesitation, and propose differentiated strategies for sub-segments to drive more vaccinations, especially for the most vulnerable and underserved one.

Our analysis process can be summarized with the following flow chart.

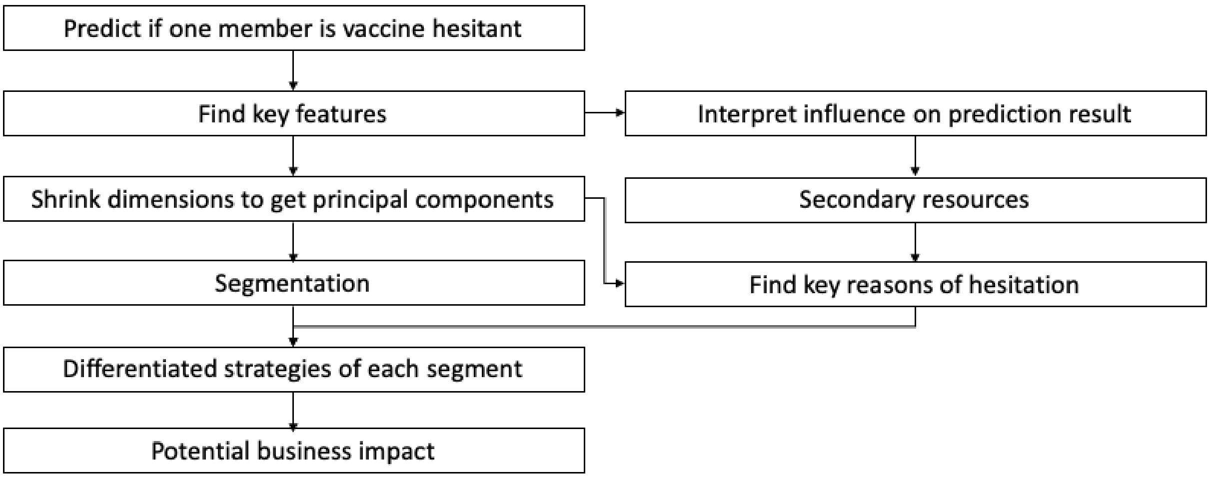

Figure 4.1. Business analysis flowchart

4.2 Key Contributors to Vaccine Hesitation

To identify key factors that drive COVID vaccine hesitation, we applied Principal Component Analysis (PCA) on the 20 most important features (see Appendix A). Our analysis shows that the features can be categorized into 5 factors with 54.05% variance explained. And these factors orient reasons of hesitation combining with segmentation that will be discussed later.

Component 1: Regional factors

The features contributing to this component include: census data of percentage of children under age 19 and adults under age 65 who are without health insurance, demographic features of race and geographic region, and credit records. According to previous SHAP analysis results, if a member comes from a low insurance rate county and from “East Central” region, the member would be less likely to be vaccinated. This indicates the impact of regions: the higher the score on this component, the member is more likely coming from an underserved region with lower vaccination rate.

Component 2: Health condition factors

The cost and count of different kinds of prescriptions (generic, maintenance, and behavioral health) and part D payment amount mainly contributes to the second component. The SHAP analysis indicates that all the 4 prescription cost features are negatively related to the vaccine hesitation. That is, if a member has a higher score on this component, the member is generally in a better health condition, and is more likely to be hesitant on vaccination.

Component 3: Social welfare factors

This component consists of hyperlipidemia prescriptions coverage, food security status of the region, and child & adult care coverage (CACFP) of the region. The Child and Adult Care Food Program (CACFP) provides nutritious meals and snacks to the children and older adults or chronically impaired persons with disabilities in their care (Child and Adult Care Food Program (CACFP), n.d.). Food security and CACFP coverage reflect the social economic environment a member is living in. With a higher score on this component, the member would suffer worse socio-economic environment, and is less likely to get vaccination.

Component 4: Financial factors

Risk Adjustment Factor A amount, net worth per adult, and household investable assets are the key compositions of the fourth component. They evaluate the health status and wealth of a member. While the relationship of vaccination and RIF is not clear as suggested in SHAP analysis, the relationship of vaccination and member's worth and assets is positive. The higher the score on this component, the poorer the member would be, and the more likely to be vaccine hesitant.

Component 5: Age factors

This component is summarized as age, since age is dominant with much higher feature importance than other factors and it coordinates with others, such as "entering medicare by OASI" – the program is only eligible after retirement (Kagan, 2021). Therefore, a higher score on this component informs a younger member who's hesitant to be vaccinated.

4.3 Segmentation and Hesitation Reasoning

After discovering key contributors to hesitation, we would like to identify different groups in Humana's members that are unvaccinated. We therefore completed K-means clustering based on the members' score on the 5 components explained above. The approach would help us understand their different reasons in different groups and propose focused and efficient strategies. We found members falling into 5 main clusters with the highest Calinski-Harabasz Score at 165,738. The score indicates the segmentation performs the best with clusters that are dense and well separated.

With the help of Radar Chart, it can be easily observed that the clusters are differentiated on the 5 dimensions from PCA. For example, cluster 1 is extremely higher than other clusters on the component of age.

Figure 4.2. Radar chart of principal components contribution in each group

We summarized the key features of every cluster in the following table. The top 5 features are components extracted from PCA analysis while the others are general statistics of clusters (demographic and health insurance related).

Table 4.1. Summary information of each group

Feature	Group 1	Group 2	Group 3	Group 4	Group 5
Regional factors	-1.66	0.67	1.64	-1.26	-0.24
Health condition factors	-3.75	-1.44	-0.08	1.67	0.31
Social welfare factors	0.11	-0.09	0.20	0.02	-0.11
Financial factors	-0.21	-0.81	-0.18	-0.98	0.75
Age factors	-0.93	2.05	-1.05	-0.37	0.25
members(#, %)	130,371, 17.7%	158,406, 21.5%	313,402, 42.6%	47,355, 6.4%	86,975, 11.8%
Average age**	74	73	74	67	52
Primary regions***	Great Lakes/Central North(20%) East Central(18%)	Texas(28%) South East(13%)	East Central(23%) Great Lakes/Central North(15%)	East Central(17%) Great Lakes/Central North(11%)	East Central(17%) Great Lakes/Central North(12%) Mid-Atlantic/North Carolina(12%)
Health risk****	0.435	0.745	0.657	1.460	0.876

* Identified by the centroid of each cluster. For readability, some numbers are reversed to ensure the larger the number, the higher possibility of vaccine hesitation

** Identified by the average of the “est_age” feature

***	Identified by the proportion members in each region of the “hum_region” feature
****	Identified by the average of the “cms_risk_adjustment_factor_a_amt” and “cms_partd_ra_factor_amt” feature

In addition, with the idea of Customer Journey Map and combining with secondary resources, we mapped the potential reasons of hesitation of each group into steps that a member would experience to get vaccinated: know, understand, be willing, access, complete. We believe it would help us better understand the pain points of Humana’s unvaccinated members and support outreach strategies. We will further explain the reasoning by each group.

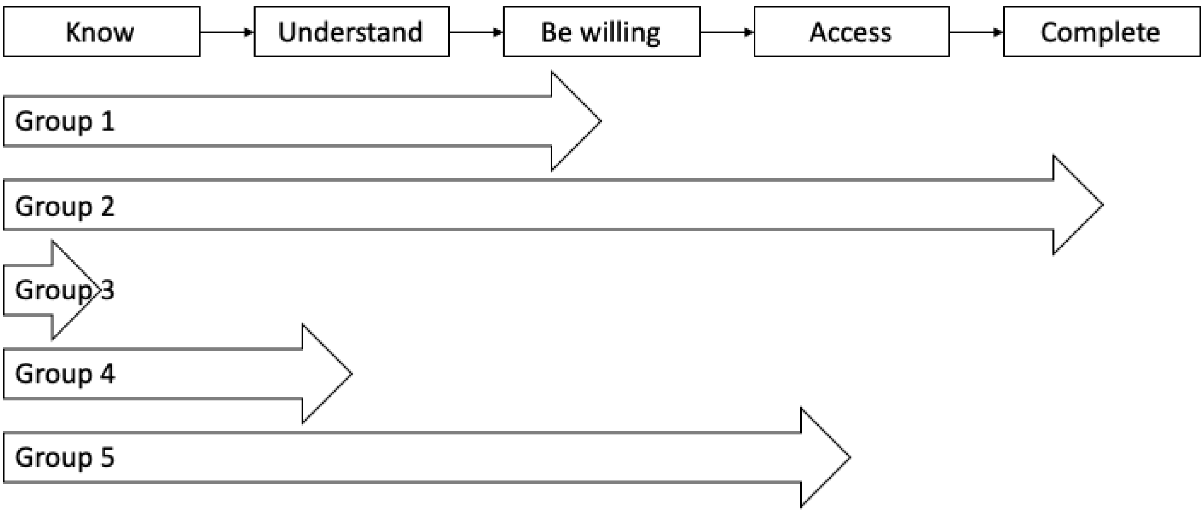

Figure 4.3. Customer Journey Map with each group’s situation

Group 1: Seniors with health issues

This group consists of 17.7% members of Humana with the highest average age at 74. Their problems are the most severe on social welfare (0.11) – their life is not secured and they easily feel unsafe. Although they have the worst health condition (-3.75), they are still not willing to be vaccinated. According to the report from Forbes (Hart, 2021), 36% of adults on the vaccine fence are mostly worried about side effects. Especially for those with chronic health problems, they are resistant to vaccines with the fear of iatrogenic effects and presence of individuals for whom vaccines are medically contraindicated (Murphy, 2021). Therefore, this group of less secured, worse health condition seniors, their willingness becomes the most severe reason to vaccine hesitation.

Group 2: Texans

The second group consists of 21.5% of Humana members with 73 average age. They have problems with regional factors (0.67) and age factors (2.05). They are the group most geographically separated from other groups – 28% of them live in Texas and 13% in the South East, compared to other groups which are basically distributed in Central North and East Central. According to a report in March (Agnew & Menchaca, 2021) that is corresponding to the timing of the dataset, Texas was facing vaccine challenges that left many eligible people unable to get a shot because of unequal distribution and lack of supply. The hardness to complete vaccination primarily prevents this geographically isolated group from vaccination.

Group 3: Seniors from insured regions

The third group of members are also elders of 74 years old on average. They are the largest proportion of members at 42.6%. They prominently face regional problems (1.64), as in their area of living, less people are insured. Since health insurance companies would put much effort on advocating vaccinations in the regions with mass members for the best efficiency, a region that is not well covered by insurance would be overlooked and thus, the member would be less likely to know about vaccinations from local reach and peers. They are excluded from the beginning of the vaccination journey for not being well informed.

Group 4: Healthy but riskful

Group 4 has 6.4% members who currently have good health conditions (1.67) and seldomly spend money on drugs. However, they also have the highest health risk (1.460). They seem healthy so they would be overconfident and be skeptical to COVID and vaccines. A survey (Dohr, 2021) reveals that not believing vaccines are effective and not needing them are among the top 3 reasons for hesitation. And according to a Forbes report (Hart, 2021), 90% of skeptical rejecters claim they are not worried about being sick from COVID. It raises attention that this group of seemingly healthy members may bring high risk for Humana if they are not vaccinated and are infected, despite their relatively low number of members. Education on understanding the seriousness of COVID and the effectiveness of vaccines are in need for this group.

Group 5: Young and low-income

The last group of 11.8% members are the youngest with average 52 years old. But they typically have financial problems (0.75). This phenomenon is supported by the article from Nature Communications (Murphy et al., 2021) that vaccine resistance is associated with lower income. A health policy professor at Boston University states that a lot of low-income workers are working hard to make ends meet and it becomes hard for them to take unpaid time off to get vaccinated (Herman, 2021). On the other hand, some locations offer vaccinations through drive-thru service that excludes many low-income people without their own vehicle. For other locations, people still need to have access to transportation to get to their appointments. (Agnew & Menchaca, 2021) Therefore, managing convenient access to vaccinations is the top one task to do.

4.4 Recommendations and Solutions

In this session, we propose solutions for each group accordingly. The pain point for vaccination and focus are given first, followed by solutions which are proposed by scalability from high to low.

Solution for Group 1: Seniors with health issues

The pain point for vaccination in this cluster is no willingness despite their older age, insecure life and poor physical health. The focus is to provide easy-accessed, trustworthy information to encourage vaccination. First, make full use of existing solutions. Humana currently provides a covid information website which can be accessed through the banner on top of its official website, this information should be delivered to this cluster via mail, e-mail, or even in-door visit. The network telehealth visits should be emphasized as professional doctors can give this group a feeling of trustworthiness to help them overcome the fear and bias towards vaccination that is common in older-aged groups. Secondly, collaborate with more parties to expand current influence. Humana can partner with government, food banks and other local organizations to provide objective information and educational resources, video modules, communication toolkits, and outreach to help guide through the roll-out of the COVID-19 vaccine.

Solution for Group 2: Texans

The pain point for vaccination in this cluster is in completion which is associated with low supply of vaccination and inequality. The focus is to support vaccine supply and eliminate geographic inequality. First of all, it is highly suggested that Humana perform a deep analysis on which communities are most vulnerable. The current dataset has errors on the zip code information and therefore does not allow for more drilldown. Humana should get the whole picture of geographic inequality by first focusing on identifying members who are vulnerable to covid-19 and reside in areas where the vaccination rates are among the most inequitable with more accurate granular data and advanced analytics. Next, Humana can consider building a community-based strategy to address potential inequality related to covid-19 vaccine access. For example, Humana can give members recommendations on location of most convenient available vaccination sites and send notifications once vaccines are available. Lastly, Humana can partner with pharmacies to open more vaccination sites on where vaccinations are most needed based on Humana's analysis, and work with the state to expand the efficient and equitable distribution of vaccines.

Solution for Group 3: Seniors from insured regions

The pain point for vaccination in this cluster is in awareness which is possibly due to the reason that they live in an out-of-business-focus area where their neighbors may not buy insurance or get vaccinations. The focus is to provide sufficient objective information in the most acceptable and persuasive way. Humana currently provides short interview videos of members sharing how they walked from hesitant to getting vaccination to happy to enjoy and share the protections of vaccination. Those in-person stories are the most powerful to deliver both informational value and emotional echo. Humana should engage these audiences by making a presence and impact on their favorite platforms or social media. Podcasting can be considered as it gains heat in both older and younger generations especially during the pandemic. Though it may be difficult to serve this group extensively in the short term, Humana can collaborate with local organizations to reach out and provide resources to under-served communities. Last but not the least, once members in this group change their minds and get vaccinations, Humana should invite them as a trusted voice to share their personal vaccine story.

Solution for Group 4: Healthy but riskful

The pain point for vaccination in this cluster is in not understanding the necessity for vaccination. The focus is to help them understand the necessity and take precautions early even when they are in good health for now. Humana can create a series of video interviews with community leaders to emphasize the health benefits and importance of receiving vaccines, and release the videos on Youtube and other popular channels that these group of people would browse. Since this group of people are not worried about the side effects of vaccination but simply do not take covid and vaccination seriously. Humana can further carry out a public health ad with a serious tone and simple message to help those people understand vaccination is not just for the good of themselves, but for their family and everyone else.

Solution for Group 5: Young and low-income

The pain point for vaccination in this cluster is in restricted access associated with poor financial status. The focus is to remove barriers in time availability and transportation. Humana can coordinate travel for them by providing language assistance, flexible appointment scheduling, and arranging transportation. It can sponsor Bluebikes rides to and from vaccination sites. Moreover, Humana can arrange for some special vaccination events in major cities during which members can just show up and get vaccinated without appointments. Another creative approach is that it can purchase a van and transform it into a "mobile vaccination truck". Then it can act as a mobile clinic to this underserved community.

4.5 Potential Business Impact and Priority Analysis

According to research conducted by Kaiser Family Foundation (KFF), preventable covid-19 hospitalizations among unvaccinated adults cost $5.7 billion from June to August 2021, and over 280,000 covid-19 hospitalizations could have been prevented by vaccination during this period. The unvaccinated would potentially cost Humana a large amount of money. In this part, we would like to discuss how much vaccination is going to influence the probability of being infected with covid-19 and the probability of being hospitalized, and also the potential amount of money Humana could save.

4.5.1 Vaccinated are less likely to be infected with Covid-19

According to a study published in the CDC's Morbidity and Mortality Weekly Report, unvaccinated people were nearly five times more likely to be infected with Covid-19 than people who got the shots. Moreover, the study also indicates that unvaccinated people are about 29 times more likely to be hospitalized with Covid-19 than those who are fully vaccinated. In another word, the risk of infection with Covid-19 is much higher among the unvaccinated, and the symptom would be much more severe for those who are unvaccinated. While the average hospitalization cost for covid-19 patients is roughly $20,000, the cost is much more likely to add up for the unvaccinated, as they are more likely to have severe symptoms because of a bigger chance of being hospitalized. Therefore, convincing the hesitant to get their vaccines would certainly help lower the chance of any potential copays.

4.5.2 The money term

In this part, we would estimate the money that Humana might be able to save based on the data presented by Yuping Tsai, Tara M. Vogt, and Fangjun Zhou, and the CDC study mentioned above. They have conducted a study on the demographics and average cost associated with covid-19–related medical care.

Demographics of the patient

Below is a table that shows the percent of hospitalized covid-19 patients receiving different levels of medical treatment. We can see that there is still a large proportion of patients who need special care.

Table 4.2. The percentage of hospitalized covid-19 patients receiving different levels of medical care

Age	Excluding Death or Ventilator	Ventilator	Death
65-74 y	83%	10%	14%
75-84 y	79%	9%	19%
>85 y	75%	4%	24%

Mean Cost for Different Segments of Patients

Below is a table that indicates the average costs for different segments of patients receiving different levels of medical care. Note that the mean cost per hospitalization was $21,752, the mean cost for those who died in the hospital was $32,015, and the mean cost for those who needed ventilator support during hospitalization was $49,441.

Table 4.3. The mean cost for hospitalized covid-19 patients receiving different levels of medical treatment

Age	Excluding Death or Ventilator	Ventilator	Death
65-74 y	$ 19,405	$ 53,641	$ 42,475
75-84 y	$ 18,424	$ 47,361	$ 32,613
>85 y	$ 17,078	$ 40,706	$ 22,794

The Estimated Amount of Money Humana Could Save

According to CDC's study posted in August, 2021, the age-adjusted hospitalization rate in unvaccinated persons (29.4 per 100,000 population) was 29.2 times the rate in fully vaccinated persons (1.0 per 100,000 population). And also, according to Humana's current dataset, there is a total unvaccinated population of 736,509, including 48.1% hesitant aged 65 – 74, 26.2% hesitant aged 75 – 84, and 6.4% hesitant aged over 85. With these pieces of information, we are able to estimate that Humana has a potential to save over 4 million dollars, if they are able to convince their client to get vaccinated.

Table 4.4. Potential Savings for Humana

Age	Difference in number of hospitalized population	Expected Savings per Person	Total Expected Savings
65-74 y	101	$ 27,500	$ 2,766,192
75-84 y	55	$ 24,827	$ 1,362,254
>85 y	13	$ 19,916	$ 268,328
TOTAL			$ 4,396,774

4.5.3 Suggesting Priority Measures

Based on the PCA analysis and saving estimation above, we are going to offer the following advice ranked by priority.

- Focus on the 65 – 74 year-old people in Group 3, as the reason that they are still unvaccinated is lack of exposure to vaccination information, and also, patients in the age group of 65 – 74 are expected to cost more.
- Do more promotion on the importance of vaccination to people in Group 1 & 4, make sure they understand the benefits of vaccination. Humana could also assign agents to reach out to these clients.
- Focus on the rest of the clients aged from 65 to 74 in other groups. Then, Focus on the clients above 75 years old.

References

Agnew, D., & Menchaca, M. (2021, March 20). *Texas' decentralized, internet-reliant system for vaccine appointments leaves many eligible people unable to access a shot*. The Texas Tribune. https://www.texastribune.org/2021/03/20/texas-coronavirus-vaccine-difficult/

CDC. (2020, March 28). *COVID Data Tracker*. Centers for Disease Control and Prevention. https://covid.cdc.gov/covid-data-tracker/#cases_casesper100k

Child and Adult Care Food Program (CACFP). (n.d.). Benefits.Gov. Retrieved October 11, 2021, from https://www.benefits.gov/benefit/5871

County Presidential Election Returns 2000-2020. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/VOQCHQ

Dohr, B. (2021, April 9). *13% of adults say they will "definitely not" get the COVID-19 vaccine. Here are their reasons.*Milwaukee Journal Sentinel. https://eu.jsonline.com/story/communities/2021/04/09/covid-vaccine-not-everyone-wants-one-why-side-effects-effectiveness-fear-of-needles/4841405001/

Griffin, J. B. (2021). SARS-CoV-2 Infections and Hospitalizations Among Persons Aged ≥16 Years, by Vaccination Status — Los Angeles County, California, May 1–July 25, 2021. MMWR. Morbidity and Mortality Weekly Report, 70. https://doi.org/10.15585/mmwr.mm7034e5

Hart, R. (n.d.). By The Numbers: Who's Refusing Covid Vaccinations—And Why. Forbes. Retrieved October 10, 2021.https://www.forbes.com/sites/roberthart/2021/09/05/by-the-numbers-whos-refusing-covid-vaccinations-and-why/?sh=747a443052ea

Herman, B. (2021, July 12). *Most unvaccinated people have low incomes*. Axios. https://www.axios.com/covid-vaccines-low-income-poor-workers-58698275-0451-4158-a967-37189dbf673c.html

Kagan, J. (2021, October 10). *What Is the Old-Age and Survivors Insurance (OASI) Trust Fund?* Investopedia.https://www.investopedia.com/terms/o/old-age-and-survivors-insurance-trust-fund.asp

Krutika Amin, Cynthia Cox. Unvaccinated COVID-19 hospitalizations cost billions of dollars.https://www.healthsystemtracker.org/brief/unvaccinated-covid-patients-cost-the-u-s-health-system-billions-of-dollars/

Murphy, J., Vallières, F., Bentall, R.P. et al. Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. Nat Commun 12, 29 (2021). https://doi.org/10.1038/s41467-020-20226-9

Yuping Tsai, Tara M. Vogt, Fangjun Zhou. Patient Characteristics and Costs Associated With COVID-19–Related Medical Care Among Medicare Fee-for-Service Beneficiaries. Annals of Internal Medicine, Vol. 174, No. 8, August 2021

Appendix

A. PCA Results

Component	Feature
Regional factors	rwjf_uninsured_child_pct
	rwjf_uninsured_adults_pct
	race_cd
	credit_hh_nonmtgcredit_60dpd
	hum_region
Health condition factors	cms_tot_partd_payment_amt
	rx_generic_pmpm_cost_6to9m_b4
	rx_maint_pmpm_ct_9to12m_b4
	rx_generic_pmpm_cost
	rx_bh_pmpm_ct_0to3m_b4
Social welfare factors	pdc_lip
	atlas_vlfoodsec_13_15
	atlas_pct_cacfp15
Financial factors	cms_risk_adjustment_factor_a_amt
	cons_nwperadult
	cons_estinv30_rc
Age factors	est_age
	cms_orig_reas_entitle_cd
	rx_gpi2_17_pmpm_cost_t_12-9-6m_b4
	cons_rxadhm