Humana-Mays Healthcare Analytics
2021 Case Competition

COVID-19 Vaccine Hesitancy Analysis

October 10th, 2021

Table of Contents

1. Executive Summary

After the first COVID-19 case reported in the United States in January 2020, this respiratory disease is still affecting millions of people globally. The virus continues to spread as a form of multiple different variants, making it more difficult to terminate this pandemic. To overcome this crisis, it is more important than ever to follow what citizens can do to reduce the risk of diffusion, and vaccination is one of the key methods. For this year's Humana-Mays Case Competition, the goal of this project is to estimate Humana members' COVID-19 vaccine hesitancy and recommend Humana for ways to reduce vaccine hesitancy among its members. Using various classification models, our team computed the hesitancy score for each member. We wanted to generate actionable insights and proposed an optimal solution to help solve problems as well as meeting Humana's needs.

The first part of this report is focusing on technical analysis. In this part, we first look into the data structure and understand how this dataset is delivering information. Then we explore several characteristics of Humana members in the dataset, such as age, region, financial stability, etc. After these steps, we elaborate in more detail how we prepared the data for the machine learning training and introduce what kind of model that we tried and how was the performance. Then evaluate which model shows the best performance in terms of AUC score. We started with six popular classification models (Logistic regression, Weighted Random Forest, SVM, XGBoost, and neural networks) and chose Logistic Regression as the best predictive model. This prediction model can be a key tool for Humana to estimate vaccine hesitancy of each individual and find better strategies for those members to provide them healthier life and better services.

Secondly, we cover insights that we drove throughout this project. We provide actionable insights with demographic characteristics of members who are most likely to hesitate to get vaccinated. Based on these characteristics, we suggest Humana several action items and their potential benefits for Humana. Our recommendations for Humana are categorized into two groups:

- Mainly targeting elderly people, 65+ years of age
- Focus on low income Black/Hispanic to lower inequity in healthcare

2. Case Background

2.1 Context and Motivation

The COVID-19 pandemic which was first identified in Wuhan, China in late 2019 has had devastating social and economic impacts in the United States and globally.[1] With over 4.82 million deaths worldwide which include nearly 710 thousand confirmed deaths in the United States so far, combating this pandemic effectively with urgency is the utmost priority. [2]After expediting the development of a vaccine, the pressing issue in the United States and other developed nations became whether the public is willing to accept this vaccine. According to Morning Consult, U.S is ranked #2 in Covid-19 vaccine opposition than any country tracked right after Russia.[3] Vaccine rollout in western countries such as the United States have slowed due to lack of acceptance of the vaccine in some. Even though unvaccinated Americans are 11 times more likely to die and 10 times more likely to be hospitalized from Covid-19 than people who are fully vaccinated, a large portion of the population is still resistant to receiving the vaccine even when it is available.[4] Coupled with easing social and public health restrictions, there is a looming threat of greater transmissions, increase in deaths, and potential emergence of new variants with the absence of herd immunity due to vaccine hesitancy.[5]

2.2 The Business Problem

Humana is a top provider of Medicare Advantage benefits to more than 4.8 million members in the United States.[6] As one of the leading health insurance companies in the United States, Humana's role in identifying members who are hesitant about receiving the COVID-19 vaccine remains a priority and is crucial for continuing efforts to vaccinate the U.S. population. The purpose of this study is to help Humana understand the key factors that could identify the members who are likely to be hesitant about the COVID-19 vaccine through a predictive model based on a profile of characteristics.

---

[1] *Novel coronavirus circulated undetected months before first COVID-19 cases in Wuhan, China*. UC Health - UC San Diego. (n.d.). Retrieved October 10, 2021, from https://health.ucsd.edu/news/releases/Pages/2021-03-18-novel-coronavirus-circulated-undetected-months-before-first-covid-19-cases-in-wuhan-china.aspx.

[2] Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., & Roser, M. (2020, March 5). *Coronavirus pandemic (COVID-19) – the data - statistics and Research*. Our World in Data. Retrieved October 10, 2021, from https://ourworldindata.org/coronavirus-data?country=~USA#confirmed-covid-19-deaths-by-country.

[3] *Vaccine hesitancy isn't unique to the U.S., survey shows ...* (n.d.). Retrieved October 10, 2021, from https://www.usnews.com/news/best-countries/articles/2021-09-30/vaccine-hesitancy-isnt-unique-to-the-us-survey-shows.

[4] Richmendezcnbc. (2021, September 10). *New Study finds unvaccinated are 11 times more likely to die from Covid, CDC says*. CNBC. Retrieved October 10, 2021, from https://www.cnbc.com/2021/09/10/new-study-finds-unvaccinated-are-11-times-more-likely-to-die-from-covid-cdc-says-.html.

[5] Ellyatt, H. (2021, September 6). *Covid vaccination rates have slumped in some parts of the world, and experts are worried.* CNBC. Retrieved October 10, 2021, from https://www.cnbc.com/2021/09/06/covid-vaccinations-have-slumped-in-parts-of-the-us-and-europe.html.

6 Humana 2021 Kick-off slide deck

The main barrier to the timely distribution of the vaccine in the U.S. is vaccine hesitancy. In the United States, only 55.6% of the population are fully vaccinated at this point with 90 million Americans (27.3% of the population) eligible for the vaccine that have not gotten it.[7] The American Hospital Association estimates a financial impact of at least $323.1 billion in 2020 attributed to the COVID-19 pandemic and another $53 billion to $122 billion loss for 2021.[8] The pandemic greatly reduced the amount of non-urgent care and procedures with an average decline of 19.5% in inpatient volume and 34.5% in outpatient volume. This affects Humana as a health insurance provider because the healthcare industry is a heavily integrated system that relies on each part of the value chain to function. Delaying vaccination efforts puts tremendous strain on the healthcare system and hampers the recovery of the economy and return to normalcy. According to the Penn Wharton Budget Model, the impact of doubling the number of vaccine doses administered daily to 3 million would create more than 2 million jobs, boost real GDP by an estimated 1%, and prevent about 2 million cases in 2021.[9] These projections and economic recovery rely on the pace of vaccination.

Therefore, the goal of this analysis is to aid in the identification of groups of Humana members that are likely to be hesitant about receiving the vaccine for targeted outreach campaigns. Only through effective action to increase vaccine acceptance can herd immunity be achieved and signal an endpoint of the COVID-19 pandemic. Not only does increasing vaccination rates in the United States benefit public health and have positive social and economic impacts, but Humana is also a stakeholder in the healthcare industry ravaged by the COVID-19 pandemic and has a financial interest in promoting this effort.

3. Data Preparation

3.1 Exploration of Data

The given dataset provided by Humana contained 367 potential variables columns and 974,864 observation rows. The purpose of our data analysis is to identify and determine the probability of a patient's likelihood to reject getting vaccinated. The dependent variable for that is the ‘covid_vaccination’ column which is a categorical data type with a binary indicator (1= positive indicator and 0= negative indicator).

There are 202 floats data types and 115 object data types. Features in the dataset were categorized into the following groups:

---

[7] Doheny, K. (2021, August 10). *COVID vaccine hesitancy: 90 million still on the fence*. WebMD. Retrieved October 10, 2021, from https://www.webmd.com/vaccines/covid-19-vaccine/news/20210810/covid-vaccine-hesitancy-90-million.
[8] *Aha Covid-19 Financial Impact Report*. (n.d.). Retrieved October 10, 2021, from https://www.aha.org/system/files/media/file/2020/06/aha-covid19-financial-impact-report.pdf.
[9] Paulson, M. (2021, March 18). *Epidemiological and economic effects of the COVID-19 vaccine in 2021*. Penn Wharton Budget Model. Retrieved October 10, 2021, from https://budgetmodel.wharton.upenn.edu/issues/2021/3/1/epidemiological-and-economic-effects-of-covid-19-vaccine.

Variables	Data type	Columns
ID: Identifier of members	Categorical	2 variables
Basic demographic and health scores	Categorical and Numerical	14 variables
Claims for each of the BETOS codes for Health service	Categorical and Numerical	34 variables
Binary indicator for each of the BH (Disease)	Categorical Binary	7 variables
Binary indicator for CSS code	Categorical Binary	21 variables
Health and insurance information from CMMS (Insurance and Disease)	Categorical and Numerical	13 variables
Binary indicator for each of the CMS level 2 diagnosis categories	Categorical Binary	9 variables
Consumer socio-economics	Categorical	14 variables
Credit data	Numerical	99 variables
HEDIS Disease indicators	Categorical	9 variables
Indicators for participation in a senior exercise program	Categorical Binary	3 variables
Indicator for abnormal lab result	Categorical Binary	8 variables
Non-BH related claims	Numerical	17 variables
Percent Days covered value for drugs	Numerical	8 variables
Physician evaluation and management data	Categorical Binary	3 variables
Using HIPPA provider specialty data	Categorical Binary	13 variables

Revenue code CMS categories data	Categorical binary	7 variables
Prescriptions	Categorical and Numerical	201 variables
Claims for each of the MCC	Categorical and Numerical	326 variables
Combined Med and BH values	Numerical	17 variables

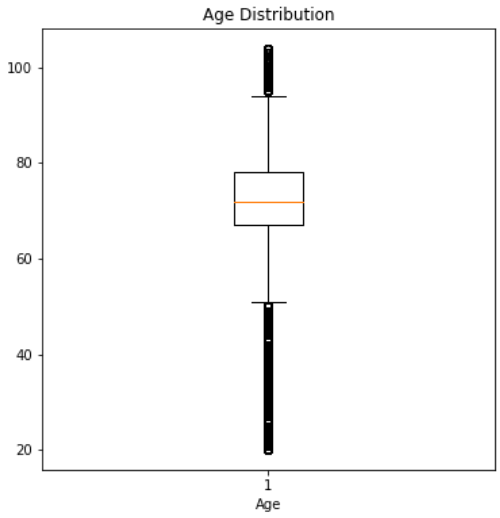
There are many features in the dataset that contained null values and required a thorough data cleaning process. The analysis was performed using Vertex AI and Jupyter Notebooks. We used Pandas, Matplotlib, Seaborn, Numpy and Scikit Learn to clean, analyze, visualize and test various machine learning models in this project. Age- range of 974,842 patients is the most influential among the variables provided for 974842 patients.

3.2 Understanding Variables

We performed exploratory data analysis (EDA) to make more sense of the data and understand the key demographic characteristics of Humana members in the dataset.

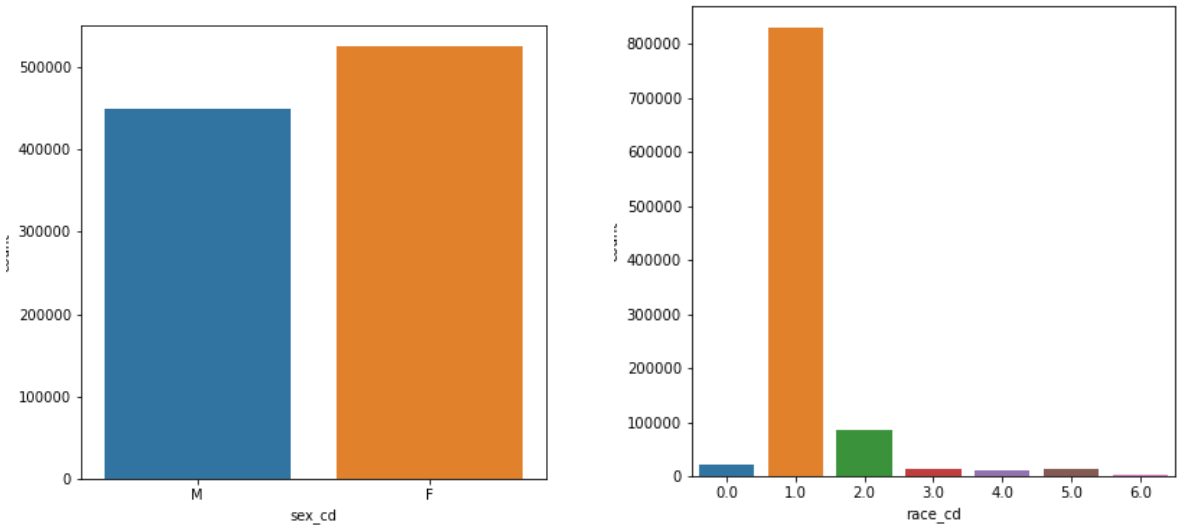
Figure 3.2.1 shows the information about members' age. The median age is 72, and the maximum value of age is 104. Even though the minimum age is 20, the box plot shows that the distribution of age is mainly concentrated in the range between 67 and 78. Therefore, we can conclude that the majority of members in the dataset are elderly.

	est_age
count	974842.000000
mean	71.175321
std	11.063120
min	20.000000
25%	67.000000
50%	72.000000
75%	78.000000
max	104.000000



(Figure 3.2.1) Summary of Members' Age

Figure 3.2.2 shows the information about members' gender and race. Understanding the distribution of these columns is important since these columns can potentially create bias in machine learning models. The bar plot shows that the gender of members is nearly evenly distributed. In terms of race, however, we can see that the data is highly imbalanced. 'Race 1.0', which indicates White ethnicity, shows a notably high number of members. There are 829,212 white members, and this takes about 85% of our total data points.



(Figure 3.2.2) Gender and Race count

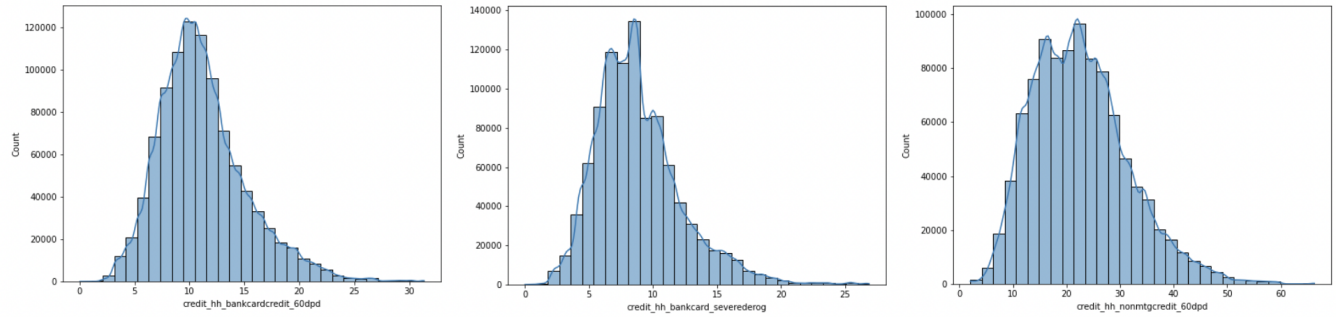
Figure 3.2.3 below shows the information about members' regions. Only East Central and Great lakes/Central North show a proportion higher than 10%. The States of Texas and Florida were treated as a region by themselves, but the rest of the states were aggregated with others.

	hum_region	Proportion
EAST CENTRAL	155468	15.95
GREAT LAKES/CENTRAL NORTH	130896	13.43
MID-ATLANTIC/NORTH CAROLINA	92716	9.51
EAST	69441	7.12
TEXAS	68980	7.08
SOUTHEAST	67117	6.88
FLORIDA	65608	6.73
CENTRAL	61266	6.28
MID-SOUTH	61079	6.27
NORTHEAST	48535	4.98
GULF STATES	47137	4.84
CALIFORNIA/NEVADA	43909	4.50
CENTRAL WEST	29656	3.04
INTERMOUNTAIN	27280	2.80
PR	4631	0.48
PACIFIC	1123	0.12

(Figure 3.2.3) The number of members in specific region and proportion

Finally, Figure 3.2.4 shows the inference about members' financial stability. We used three finance-related columns, credit_hh_bankcard_severederog, credit_hh_bankcardcredit_60dpd, credit_num_nonmtgcredit_60dpd. These columns are not written at the individual level, but still provide much insight into Humana members. The forest histogram shows the percentage of Severe Derogatory Accounts. This is slightly right-skewed, 7~10% got the highest responses. The second histogram shows the percentage of bankcard accounts that have passed more than 60 days due. It is also slightly right-skewed, but more members have accounts past its due date as you can see the maximum of y-axis labels is higher than the previous histogram. Lastly, the third histogram shows the number of Non-Mortgage Loan Accounts which have passed more than 60 days due. This histogram is more dispersed than others, but it shows fewer members in all ranges of numbers. This indicates that people take meeting loan dues more seriously than

meeting bankcard due and managing derogatory accounts.



(Figure 3.2.4) Histograms to understand the financial stability

3.3 Data Cleaning

To prepare the dataset for modeling, we performed the following data formatting and feature reduction sequences:

For numeric features:

- We filled the missing values in 258 numeric columns with the median value of the observed values to mitigate the effect of outliers.

For categorical features:

- We first identified 18 variables that were categorical data in numeric data types. These features were cast into category columns.
- We imputed the missing values in 93 categorical columns with (*) and then replaced (*) with ‘Na’ before filling the missing values with the most occurring value in the column.
- To use categorical features in a machine learning model, we turned all categorical features into dummy variables.
- We extracted the first digit in zip code and cast it as a category datatype to cluster by location. The first digit in the American zip code system indicates a broad national area grouping from 0 (Northeast) to 9 (West Coast).

For feature reduction:

- We removed columns with more than 70% of missing data. From our investigations, the only feature with too many missing values according to this filter was ‘lang_spoken_cd’, so it was dropped from the dataset.

- We removed columns with a variance of zero because they do not influence the target variable ‘covid_vaccination_vacc’. This included 36 columns that have “0” across all rows in numeric type and another 19 columns that have zero variance in mixed data types.

4. Technical Analysis

4.1 Model selection

To achieve our goal to predict vaccine hesitancy among Humana members, we examined 6 different classification models: Logistic regression, Support Vector Machine(SVM), Random Forests, Naive Bayes, Gradient Tree Boosting, and Binary Classification with Neural Network. In this procedure, we trained all our models on an 80-20 train-test split.

Each model was evaluated by comparing the AUC score with the initial model built without hyperparameter tuning. SVM model and Sequential model achieved specifically lower AUC scores and also encountered computationally inefficiency. The naive Bayes model was good in terms of computation issues but achieved low AUC. We, therefore, chose the Logistic regression, Weighted Random Forest and the Neural Network for the further tuning process.

4.2 Model Training

Logistic regression model

The expected response for this model is a value of “1” represents a patient who is vaccinated, and “0” represents a patient who is not vaccinated. We use the logistic regression classifier on our train dataset to predict the outcome of each patient and the latter calculates the probability of the patient likelihood of getting vaccinated. We also performed repeated cross-validation for 10 folds and 3 times. We used the Logistic Regression classifier and set its hyperparameter to 'liblinear'.

Weighted Random Forest

Random Forest is an algorithm made to overcome the overfitting of the decision tree. This model builds multiple decision trees using a random subset of features each time it splits nodes. It is a key characteristic that makes the model produce generalizable algorithms. Finally, the model takes a vote within the results of individual trees. Another key advantage of the Random Forest is that it can handle thousands of variables efficiently compared to other models.

For this project, we used the Weighted Random Forest Classifier as our train dataset was highly imbalanced in the target variable. These are the hyperparameters showing the highest AUC scores:

- Class weight = balanced subsample
- Max_depth = 10
- Max_features = 27
- Min_samples_leaf = 2
- Min_samples_split = 20
- N_estimator = 2000

Binary Classification with Neural Network

Our team used Keras to effectively build, design, and train neural networks and deep learning. Since the expected prediction takes a value of either “0” and “1”, we used sigmoid as the activation function and relu as hidden layer function. These are the hyperparameters showing the highest AUC scores.

- Drop out = 0.3
- First hidden layer = “tanh”
- Second layer = ‘relu’
- Activation function = ‘sigmoid’
- Loss = ‘binary_crossentropy
- Optimizer = Adam
- Learning rate = 0.005
- Hidden units = 500, 100, 1

4.3 Feature Selection

Logistic Regression

There are 741 features including dummies variables in the dataset. After we did data cleaning, we included all the variables into the logistic regression then split it into a train and test set. Since we are working with a lot of predictors and observations, we decided to use the StandardScaler() function from SciKit Learn modules to scale all the x values to optimize the overall performance of our model later.

Weighted Random Forest

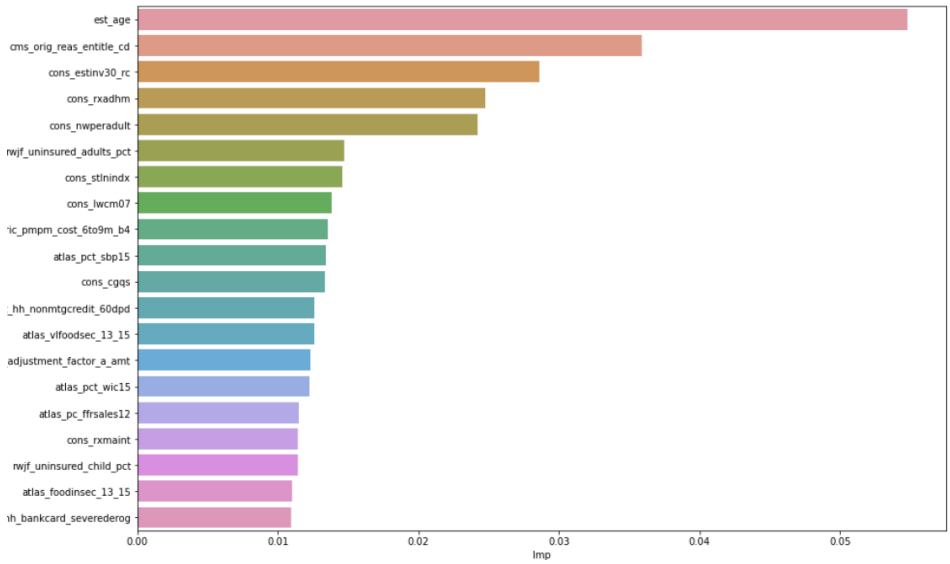
We provided all features in cleaned_data but limited the maximum number of features used for each node to 27 to prevent overfitting.

Binary Classification with Neural Network

Before we trained the model, we had to create dummy variables with the columns in pandas categorical dtype, since the sequential model does not use that data type. Then we provided all features to the model and let the model choose meaningful ones itself.

4.4 Feature Importance

One practical feature provided via sklearn ensemble tree model is feature importance. This is very helpful to understand the effect of different features on the probability of vaccine hesitancy. We used the feature_importance_ attribute of the random forest model. Here is the plot of the Top 20 features in terms of importance.

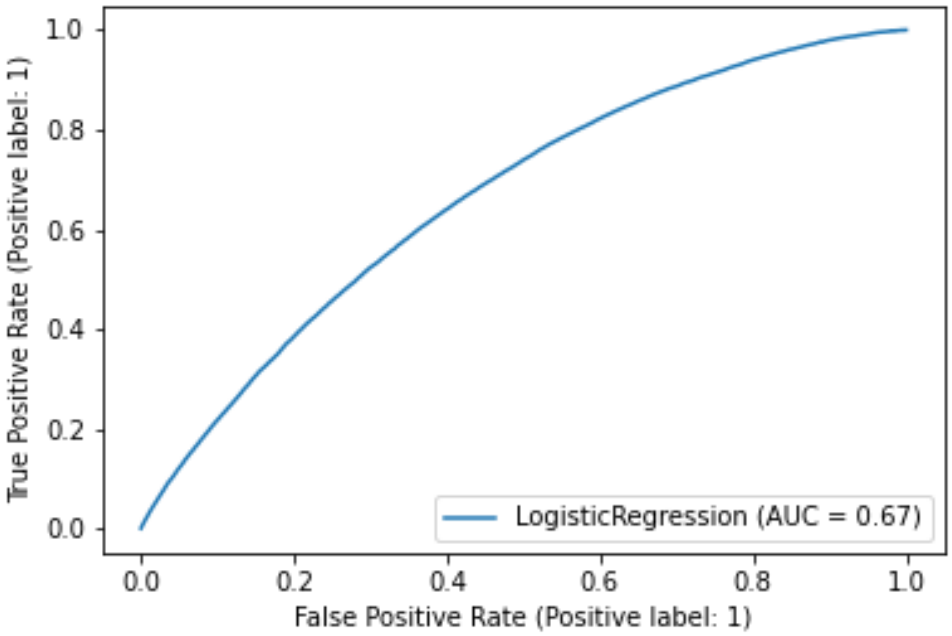


Some important features learned by our model are - Age of the member, Original reason for entry into Medicare, Estimated Household Investable Assets and RX adherence. The age shows noticeably high importance itself, and the reasons for entry into medicare is also related to age.

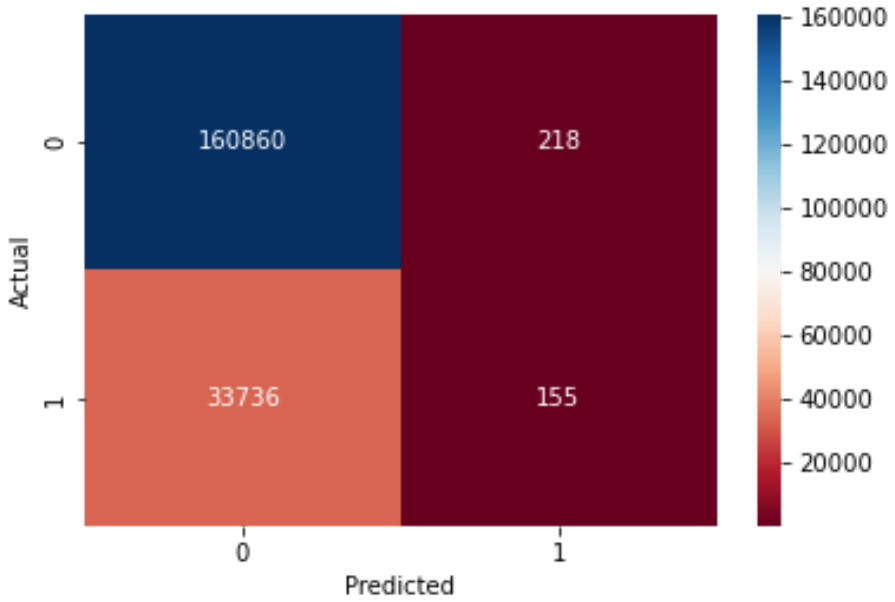
4.5 Model Evaluation

Logistic Regression

For this model, we identified a ROC AUC metric of 0.66580718365 as a baseline score to compare with other models. The model will be compared and selected based on the area under the curve for the receiver operator characteristic curve (AUC-ROC). When performing cross-validation, the average k-fold score is 0.8245 standard error is 0.002.



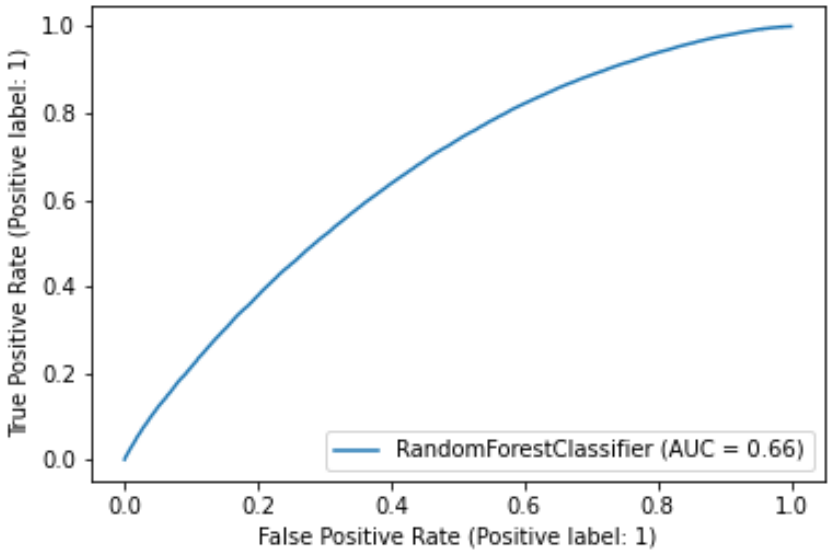
(Figure 4.5.1) ROC Curve of Logistic Regression Model



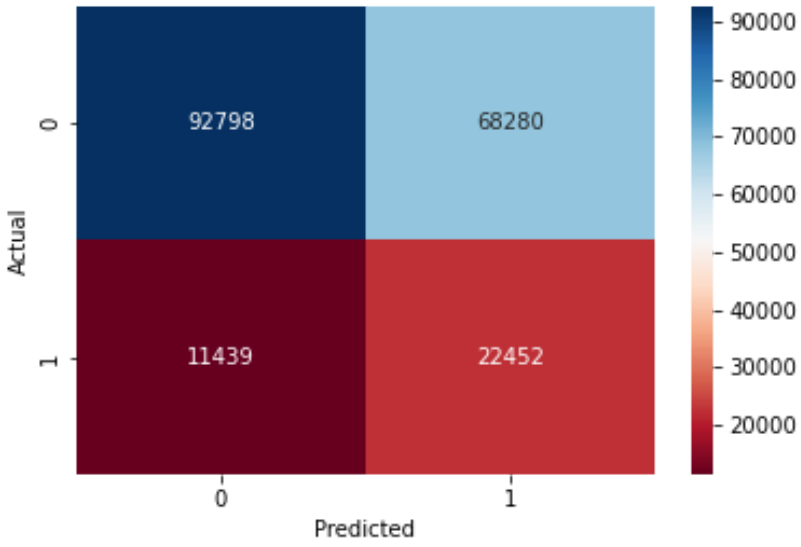
(Figure 4.5.2) Confusion Matrix of Logistic Regression Model

Weighted Random Forest

After training our model, we made predictions with the test data set, computed AUC scores, and generated a confusion matrix. The highest AUC score from the weighted random forest is 0.6636981029230024.



(Figure 4.5.3) ROC Curve of Weighted Random Forest Model



(Figure 4.5.4) Confusion Matrix of Weighted Random Forest Model

Binary Classification with Neural Network

After running the model with 80 epochs, the accuracy came out to be 0.8223. Even though Binary Classification with Neural Network is a widely used technique in the industry, this model did not show good improvement in our project. It showed only 0.56864 of AUC scores.

4.6 Final Model Result

The table is the predicted probability of a patient's likelihood of getting vaccinated. We consider the probability of getting vaccination as willingness, and the probability of not getting vaccinated as hesitancy. The weighted average scores for the precision, recall, and F-1 scores were 0.74, 0.82, and 0.75.

	Prob of not vacinated (0)	Prob of vacinnated (1)
0	0.763321	0.236679
1	0.661320	0.338680
2	0.502458	0.497542
3	0.814794	0.185206
4	0.849689	0.150311
5	0.861284	0.138716
6	0.852402	0.147598
7	0.839800	0.160200
8	0.867855	0.132145
9	0.981560	0.018440

4.7 Fairness in AI

There has been an ongoing discussion of whether AI models display a trait of bias towards specific features, especially towards gender, ethnicity, or race. AI has yet to perfectly replicate human minds. In other words, AI, in theory, should be fair in all categories and should not carry any bias. In practice, however, some groups of people assert there is undeniable discrimination happening behind the scenes. Who is right and what should we do about it? Before criticizing the correctness of a specific model, we need to closely examine the balance of data. For instance, if 99% of observations are favoring one specific feature, the remaining 1% would carry very minimal significance. Nevertheless, we cannot classify this as a bias because the data itself provided the groundwork to compute the significance of each feature.

The fundamental goal of this project is to build a model which provides the highest probability to classify correctly. Therefore, unfairness/bias is most likely due to the misled interpretation of the model that people created instead of the model intentionally disregarding a certain group. We do, however, have a means of resolving the imbalance. When building a model, we should use bootstrap aggregation that allows the model to iterate numerous times. This way the model can train the dataset in a fairway.

5. Recommendations

5.1 Actionable Insights

From our analysis of patterns in characteristics of members of Humana's health insurance program, we constructed a profile using data from the top 20% ranked members who are most likely to be vaccine hesitant. The profile of the most likely to be hesitant member based on our predictive model is a white male who is about 65 years old and resides in the East Central/Florida area.

However, further analysis into the demographic of the top 20% of vaccine hesitant individuals show the proportion of individuals in the most likely to be vaccine hesitant to the total observed data points by race tell a different story. This is because the dataset is overwhelmingly composed of people who identify as white (about 85%). Looking at the proportion of number of individuals in the top 20% group to the total number of individuals in the data by race, Black and Hispanic ethnicities are much more likely to not accept the vaccine than white people at 19.96% and 24.38% respectively compared to 9.49% for white people.

Thus, the demographics that our recommendations are tailored to are older age groups and Black/Hispanic populations. Our solutions are strategies to target this demographic in hopes of making the greatest impact on raising the percentage of vaccinated members and dissemination of information about the facts of the COVID-19 vaccine with a focus on vaccine equity.

5.2 Proposed Solutions

Targeting Adults over 65 Years Old

Members of Humana's Medicare Advantage Prescription Drug (MAPD) program are individuals who are age 65+ or younger individuals with certain disabilities. According to the CDC, these two groups are especially vulnerable and most likely need hospitalization or even die due to COVID-19. More than 80% of COVID-19 deaths occur in people over the age of 65 and adults

with medical conditions are at increased risk.[10] Because statistics show that the specific population that are recipients of Humana's Medicare benefits are much more likely to require hospitalization, intensive care, and possibly even a ventilator, it makes it of even more importance that this group of people are targeted in Humana's strategy to increases vaccination rates for public welfare as well as for Humana's financial interest due to the high medical costs associated with treatment for serious cases of COVID-19.

What are some of the factors that affect vaccine hesitancy in individuals over 65 years old? An AARP indicated that older adults are less likely to accept the vaccine due to these top factors[11]:

- Worries about the side effects of the vaccine (59% of respondents)
- Concern with the effectiveness of the vaccine (29% of respondents)
- Perceived risk of the vaccine (52% of respondents)
- Distrust in the government (47% of respondents)

In general, vaccine hesitancy is due to a lack of accurate information, lack of trust, or a combination of both. In response to factors that could help build trust, disseminating information about the recent FDA (Food and Drug Administration) approval of the COVID-19 vaccine. While the FDA had issued an EUA (Emergency Use Authorization) order earlier for vaccine distribution developed by reputable companies, they officially approved the Pfizer-BioNTech COVID-19 Vaccine which will now be marketed as Comirnaty on August 23, 2021. This recent news can have a potential impact on reinvigorating the slump in vaccination rates. According to a study, an overall 34% of respondents who are still hesitant about the vaccine said they would get the vaccine if it were FDA approved.[12] In light of FDA approval, the situation is still unfurling on the impact it will have on decreasing vaccine hesitancy. Dr. Rutherfod, UCSF Professor of Epidemiology asserts this approval is a necessary step in addressing safety concerns amongst unvaccinated skeptics.[13] Another strategy Humana can approach to tackle this issue amongst individuals in this age group is to partner with medical care providers who are involved in their expansive network of relationships in the healthcare industry to disseminate accurate information and build trust surrounding the vaccine. This is also an important step to getting the information out there about the FDA approval on the Pfizer vaccine just under two months ago. This strategy can help mitigate the issue of misinformation and distrust in two-fold.

---

[10] Centers for Disease Control and Prevention. (n.d.). *People with certain medical conditions*. Centers for Disease Control and Prevention. Retrieved October 10, 2021, from https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html.
[11] McSpadden, J. (2021, February 26). *Vaccine hesitancy among older adults, with implications for covid-19 vaccination and beyond*. AARP. Retrieved October 10, 2021, from https://www.aarp.org/ppi/info-2021/vaccine-hesitancy-among-older-adults.html.
[12] Liz Hamel Follow @lizhamel on Twitter, A. K. F. @A. K. on T., & 2020, D. (2020, December 22). *KFF COVID-19 Vaccine Monitor: December 2020*. KFF. Retrieved October 10, 2021, from https://www.kff.org/coronavirus-covid-19/report/kff-covid-19-vaccine-monitor-december-2020/.
[13] Pena, L. (2021, August 24). *Vaccine hesitancy: Here's what experts are saying about the impact of Pfizer's FDA approval*. ABC7 San Francisco. Retrieved October 10, 2021, from https://abc7news.com/fda-approves-pfizer-vaccine-approval-hesitancy-mandates/10970572/.

Firstly, people in this age group are known to have a lesser comfort level of accessing information through the internet. To give perspective, although trends of internet usage by adults over 65 show an increase, only an estimated 67% have internet access with only about 25% of seniors saying they feel confident navigating the internet and using technology.[14] This is a very big barrier to accessing information because this is how information is searched for and found most easily– just a few keystrokes and information is presented. Due to this constraint, adults older than 65 more likely need a trusted, face-to-face source that is more easily accessible.

Secondly, people are more likely to trust medical professionals and experts over any other prominent figures or sources. In a Pew Research study, more people indicated that they are more trusting towards medical professionals over elected officials, business leaders, and news outlets with 74% holding a favorable view for medical authorities compared to 35%, 46%, and 47% respectively for the other listed sources.[15] Distrust in the vaccine can be most effectively alleviated through talks with primary care physicians who are highly perceived to have their patients' best interests in mind and have an established relationship with the individuals. Considering that Humana MAPD members are likely to sign up for this program because of the added benefit of prescription drugs in addition to typical Medicare benefits, pharmacists are also a specific group of medical professionals suited to execute this strategy. Pharmacists can chat with members who have to routinely stop by to acquire prescription drugs and reduce uncertainty over time.

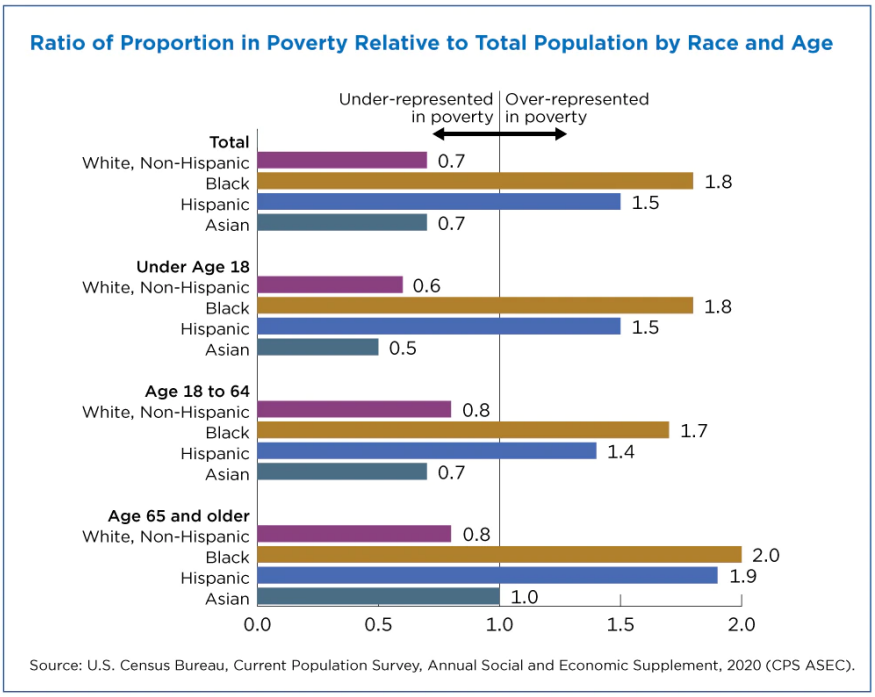
Targeting Low-Income Black/Hispanic Populations

Vaccine equity in distribution especially in underserved populations is a priority to Humana as well as should be for the United States. Due to various reasons, low-income minorities or marginalized groups are more likely to not have received the vaccine due to characteristics that are more pronounced in these populations or be more hesitant towards the vaccine. Racial health inequities exist in the American healthcare system as a result of social and it shows in our prediction that indicate Black and Hispanic populations have a higher rate of vaccine hesitancy than white, Asian, or other ethnicities. This can be examined in several ways to develop a solution that would target increasing vaccination rates in these populations.

---

[14] Poon , L., & Holder, S. (n.d.). *The 'New Normal' for Many Older Adults Is on the Internet*. Bloomberg.com. Retrieved October 10, 2021, from https://www.bloomberg.com/news/features/2020-05-06/in-lockdown-seniors-are-becoming-more-tech-savvy.

[15] Funk, C., & Gramlich, J. (2020, August 27). *Amid coronavirus threat, Americans generally have a high level of trust in medical doctors*. Pew Research Center. Retrieved October 10, 2021, from https://www.pewresearch.org/fact-tank/2020/03/13/amid-coronavirus-threat-americans-generally-have-a-high-level-of-trust-in-medical-doctors/.



Black and Hispanics in poverty are 1.8 and 1.5 times greater overall than the proportion to the population respectively especially in adults age 65 or older while whites and Asians are underrepresented.[16]

With a greater proportion of Black and Hispanic individuals in poverty, their concerns are more centered around lack of accessible information and inability to access healthcare even if they are not opposed to the COVID-19 vaccination. Here are some key statistics about this matter:

- About two-thirds of low-income essential workers are worried about serious side-effects with 53% stating concern about missing work as a result[17]
- Needing to take unpaid time off to get to vaccinated[18]
- Lesser access to reliable internet to make online appointments or information about the vaccine through a trusted provider[19]

Other necessities are more prioritized over making time to get the vaccine for low-income families. Concerns with the vaccine directly impacting these priorities are also another major

[16] Bureau, U. S. C. (2021, October 8). *Inequalities persist despite decline in poverty for all major race and Hispanic origin groups*. Census.gov. Retrieved October 10, 2021, from https://www.census.gov/library/stories/2020/09/poverty-rates-for-blacks-and-hispanics-reached-historic-lows-in-2019.html.
[17] Published: Apr 23, 2021. (2021, April 22). *Essential workers employed outside health care are less enthusiastic about getting a COVID-19 vaccine than other adults*. KFF. Retrieved October 10, 2021, from https://www.kff.org/coronavirus-covid-19/press-release/essential-workers-employed-outside-health-care-are-less-enthusiastic-about-getting-a-covid-19-vaccine-than-other-adults/.
[18] Herman, B. (2021, July 12). *Most unvaccinated people have low incomes*. Axios. Retrieved October 10, 2021, from https://www.axios.com/covid-vaccines-low-income-poor-workers-58698275-0451-4158-a967-37189dbf673c.html.
[19] *Pandemic's racial disparities persist in vaccine rollout ...* (n.d.). Retrieved October 10, 2021, from https://www.nytimes.com/interactive/2021/03/05/us/vaccine-racial-disparities.html.

reason for vaccine hesitancy and a barrier to increasing vaccination rates. Low-income workers tend to be too busy worrying about providing food and shelter, but it is exactly because low-income workers make up a large sector of non-health related essential jobs that increase their risk of exposure to COVID-19. A secondary concern outside of work-related impacts is lack of information. According to Pew Research, 63% of Blacks say lack of reliable high-speed internet puts people at a major disadvantage in connecting with medical professionals compared to 53% of Hispanics responding in the same way.[20]

Our suggested solutions to these issues is to mitigate concerns with employers regarding vaccination affecting their work and provide better access to information, vaccine appointment availability, and transportation to vaccination sites. To prioritize more flexibility and access to the vaccine are the main goals to combat the major causes of vaccine hesitancy in this group.

Humana can work with vaccination site providers to help those with issues with scheduling online to ensure that there are accessible appointments that are outside of typical work hours so it does not interfere with their jobs. Humana also has a transportation service, Logisticare, readily available at no-extra cost for purposes such as medical appointments, routine check-ups, etc. and can be extended to vaccine appointments.[21] However, their current hours of operation are from Monday-Friday between 8 AM and 5 PM which are traditional working hours. Offering more flexibility in this area is crucial to removing this barrier along with providing internet access and conversation to inform and assure Black and Hispanic people about the safety and availability of the COVID-19 vaccine.

5.3 Calculated Potential Impact

Financial Cost of COVID-19 Hospitalization to Humana

1. From June to August 2021 (3 months span), an estimated 456,000 COVID-19 hospitalizations were unvaccinated individuals and about 287,000 could have been prevented if they were vaccinated[22]
2. A study consisting of 10% of the U.S. population by the CDC indicated 42.14% of COVID-19 related hospitalizations were amongst people age 65+[23]

---

[20]Atske, S., & Perrin, A. (2021, September 10). *Home Broadband Adoption, computer ownership vary by race, ethnicity in the U.S.* Pew Research Center. Retrieved October 10, 2021, from https://www.pewresearch.org/fact-tank/2021/07/16/home-broadband-adoption-computer-ownership-vary-by-race-ethnicity-in-the-u-s/.

[21] *Medical appointments transportation - tmgipa.com*. (n.d.). Retrieved October 10, 2021, from https://www.tmgipa.com/transport.pdf.

[22] Krutika Amin Twitter and Cynthia Cox Twitter KFF. (2021, September 14). *Unvaccinated covid-19 hospitalizations cost billions of dollars*. Peterson-KFF Health System Tracker. Retrieved October 10, 2021, from https://www.healthsystemtracker.org/brief/unvaccinated-covid-patients-cost-the-u-s-health-system-billions-of-dollars/.

[23] Centers for Disease Control and Prevention. (n.d.). *Covid-19 hospitalizations*. Centers for Disease Control and Prevention. Retrieved October 10, 2021, from https://gis.cdc.gov/grasp/covidnet/COVID19_5.html.

3. *About 120,942 cases were individuals 65+ (287,000 * 42.14%) that could have been prevented
4. About 96% of the U.S. population has Medicare benefits[24]
5. *Estimated 116,104 preventable COVID-19 hospitalizations amongst Medicare members
6. Humana has over 4.8 million members[25]
7. About 16% of Medicare beneficiaries are under 65 years of age[26]
8. *Proportionally, about 4.03 million of Humana members are 65 years or older (4.8 million* (1 - 0.16))
9. About 62.6 million people in the U.S. have Medicare[27]
10. *Humana members represents 7.67% of Medicare members(4.8 million/62.6 million)
11. Estimated 8,905 preventable deaths amongst Humana members (116,104 * 0.0767)
12. Center of Medicare and Medicaid Services reported that Medicare fees for COVID-19 hospitalizations averages $24,033[28]

Estimated preventable cost to Humana due to vaccine hesitancy in 3 months = $214 billion

Assuming even a 1% change in vaccination rate per month due to these strategies, Humana can recoup costs in hospitalization prevention of an estimated $710,000 per month

*Calculated from available statistics

6. Conclusion

Given the Humana competition training dataset of 974,842 observations and 367 columns, we constructed various predictive models and built the best model with a Logistic Regression model that predicted COVID-19 Vaccination hesitancy with an AUC of 0.67. Key features we found through modeling correspond with our actionable insights and recommendations to reduce vaccine hesitancy among Humama members. Our main insights directed us to look into recommendations that target the most vulnerable and most likely to be hesitant about the vaccine populations which are people over 65 years of age and low-income Black or Hispanic individuals. Adults over age 65 are the primary members of Humana's Medicare Advantage

[24] Institute of Medicine (US) Committee to Design a Strategy for Quality Review and Assurance in Medicare. (1990, January 1). *The elderly population*. Medicare: A Strategy for Quality Assurance: Volume 1. Retrieved October 10, 2021, from https://www.ncbi.nlm.nih.gov/books/NBK235450/.

25 Humana 2021 Kick-off slide deck

[26] Juliette Cubanski Follow @jcubanski on Twitter, T. N. F. @tricia_neuman on T. (2016, August 16). *Medicare's role for people under age 65 with disabilities*. KFF. Retrieved October 10, 2021, from https://www.kff.org/medicare/issue-brief/medicares-role-for-people-under-age-65-with-disabilities/

[27] Mikulic, M. (n.d.). *Topic: Medicare*. Statista. Retrieved October 10, 2021, from https://www.statista.com/topics/1167/medicare/

[28] Krutika Amin Twitter and Cynthia Cox Twitter KFF. (2021, September 14). *Unvaccinated covid-19 hospitalizations cost billions of dollars*. Peterson-KFF Health System Tracker. Retrieved October 10, 2021, from https://www.healthsystemtracker.org/brief/unvaccinated-covid-patients-cost-the-u-s-health-system-billions-of-dollars/.

Prescription Plan (MAPD) and also the most at risk for hospitalization which is a major cost to Humana. In addition, healthcare inequality by racial groups is also exhibited in this analysis as the group most likely to be hesitant to the vaccine are Black and Hispanic individuals. Vaccine hesitancy boils down to major barriers that include misinformation, distrust, lack of access to vaccination, and/or interference with life. Implementing strategies to combat vaccine hesitancy in these targeted populations is a priority and essential to the overall welfare of the United States and for the interests of Humana.

7. References

1. LaFee, S. (n.d.). Novel coronavirus circulated undetected months before first COVID-19 cases in Wuhan, China. UC Health - UC San Diego. Retrieved October 10, 2021, from https://health.ucsd.edu/news/releases/Pages/2021-03-18-novel-coronavirus-circulated-undetected-months-before-first-covid-19-cases-in-wuhan-china.aspx

2. Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., & Roser, M. (2020, March 5). Coronavirus pandemic (COVID-19) – the data - statistics and Research. Our World in Data. https://ourworldindata.org/coronavirus-data?country=~USA#confirmed-covid-19-deaths-by-country

3. Wazir, Z. (n.d.). Vaccine hesitancy isn't unique to the U.S., survey shows ... Retrieved October 10, 2021, from https://www.usnews.com/news/best-countries/articles/2021-09-30/vaccine-hesitancy-isnt-unique-to-the-us-survey-shows

4. Mendez, R. (2021, September 10). New Study finds unvaccinated are 11 times more likely to die from Covid, CDC says. CNBC. Retrieved October 10, 2021, from https://www.cnbc.com/2021/09/10/new-study-finds-unvaccinated-are-11-times-more-likely-to-die-from-covid-cdc-says-.html

5. Ellyatt, H. (2021, September 6). Covid vaccination rates have slumped in some parts of the world, and experts are worried. CNBC. Retrieved October 10, 2021, from https://www.cnbc.com/2021/09/06/covid-vaccinations-have-slumped-in-parts-of-the-us-and-europe.html

6. Humana 2021 Kick-off slide deck

7. Doheny, K. (2021, August 10). COVID vaccine hesitancy: 90 million still on the fence. WebMD. Retrieved October 10, 2021, from https://www.webmd.com/vaccines/covid-19-vaccine/news/20210810/covid-vaccine-hesitancy-90-million

8. American Hospitals Association. (2020). Hospitals and Health Systems Face Unprecedented Financial Pressures Due to COVID-19. Chicago, Illinois; American Hospitals Association. Retrieved October 10, 2021, from https://www.aha.org/system/files/media/file/2020/06/aha-covid19-financial-impact-report.pdf

9. Arnon, A., Ricco, J., Prisinzano, R., & Paulson, M. (2021, March 18). Epidemiological and economic effects of the COVID-19 vaccine in 2021. Penn Wharton Budget Model. Retrieved October 10, 2021, from https://budgetmodel.wharton.upenn.edu/issues/2021/3/1/epidemiological-and-economic-effects-of-covid-19-vaccine

10.Centers for Disease Control and Prevention. (2021, August 20). Covid 19 People with certain medical conditions. Centers for Disease Control and Prevention. Retrieved October 10, 2021, from https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html

11. McSpadden, J. (2021, February 26). Vaccine hesitancy among older adults, with implications for covid-19 vaccination and beyond. AARP. Retrieved October 10, 2021, from

https://www.aarp.org/ppi/info-2021/vaccine-hesitancy-among-older-adults.html
12. Hamel, L. F. @A. K. on T., Kirzinger, A., Muñana , C., & Brodie, M. (2020, December 15). KFF COVID-19 Vaccine Monitor: December 2020. KFF. Retrieved October 10, 2021, from https://www.kff.org/coronavirus-covid-19/report/kff-covid-19-vaccine-monitor-december-2020/
13. Pena, L. (2021, August 23). Vaccine hesitancy: Here's what experts are saying about impact of Pfizer's FDA approval. ABC7 San Francisco. Retrieved October 10, 2021, from https://abc7news.com/fda-approves-pfizer-vaccine-approval-hesitancy-mandates/10970572/
14. Poon, L., & Holder, S. (2020, May 6). The 'New Normal' for Many Older Adults Is on the Internet. Bloomberg.com. Retrieved October 10, 2021, from https://www.bloomberg.com/news/features/2020-05-06/in-lockdown-seniors-are-becoming-more-tech-savvy
15.Funk, C., & Gramlich, J. (2020, March 13). Amid coronavirus threat, Americans generally have a high level of trust in medical doctors. Pew Research Center. Retrieved October 10, 2021, from https://www.pewresearch.org/fact-tank/2020/03/13/amid-coronavirus-threat-americans-generally-have-a-high-level-of-trust-in-medical-doctors/
16. Creamer, J. (2021, September 15). Inequalities persist despite decline in poverty for all major race and Hispanic origin groups. Census.gov. Retrieved October 10, 2021, from https://www.census.gov/library/stories/2020/09/poverty-rates-for-blacks-and-hispanics-reached-historic-lows-in-2019.html
17.Palosky, C. (2021, April 22). Essential workers employed outside health care are less enthusiastic about getting a COVID-19 vaccine than other adults. KFF. Retrieved October 10, 2021, from https://www.kff.org/coronavirus-covid-19/press-release/essential-workers-employed-outside-health-care-are-less-enthusiastic-about-getting-a-covid-19-vaccine-than-other-adults/
18.Amin, K., & Cox, C. (2021, September 14). Unvaccinated covid-19 hospitalizations cost billions of dollars. Peterson-KFF Health System Tracker. Retrieved October 10, 2021, from https://www.healthsystemtracker.org/brief/unvaccinated-covid-patients-cost-the-u-s-health-system-billions-of-dollars/
19. Centers for Disease Control and Prevention. (n.d.). Laboratory-Confirmed Covid-19 hospitalizations. Centers for Disease Control and Prevention. Retrieved October 10, 2021, from https://gis.cdc.gov/grasp/covidnet/COVID19_5.html
20. Institute of Medicine (US) Committee to Design a Strategy for Quality Review and Assurance in Medicare; (1990, January 1). Medicare: A Strategy for Quality Assurance: Volume 1. Medicare: A Strategy for Quality Assurance: Volume 1. Retrieved October 10, 2021, from https://www.ncbi.nlm.nih.gov/books/NBK235450/
21. Humana 2021 Kick-off slide deck
22. Cubanski , J. F. @tricia_neuman on T., Neuman, T., & Damico, A. (2016, August 16). Medicare's role for people under age 65 with disabilities. KFF. Retrieved October 10, 2021, from https://www.kff.org/medicare/issue-brief/medicares-role-for-people-under-age-65-with-disabilities/
23. Mikulic, M. (2021, October 4). Medicare - Statistics & Facts. Statista. Retrieved October 11, 2021, from

https://www.statista.com/topics/1167/medicare/
24. Amin, K., & Cox, C. (2021, September 14). Unvaccinated covid-19 hospitalizations cost billions of dollars. Peterson-KFF Health System Tracker. Retrieved October 10, 2021, from https://www.healthsystemtracker.org/brief/unvaccinated-covid-patients-cost-the-u-s-health-system-billions-of-dollar/
25 Humana 2021 Kick-off slide deck
26 Cubanski , J. F. @tricia_neuman on T., Neuman, T., & Damico, A. (2016, August 12). Medicare's role for people under age 65 with disabilities. KFF. Retrieved October 10, 2021, from https://www.kff.org/medicare/issue-brief/medicares-role-for-people-under-age-65-with-disabilities/
27 Mikulic, M. (n.d.). *Topic: Medicare*. Statista. Retrieved October 10, 2021, from https://www.statista.com/topics/1167/medicare/.
28 Krutika Amin and Cynthia Cox KFF. (2021, September 14). *Unvaccinated covid-19 hospitalizations cost billions of dollars*. Peterson-KFF Health System Tracker. Retrieved October 10, 2021, from https://www.healthsystemtracker.org/brief/unvaccinated-covid-patients-cost-the-u-s-health-sys