Humana-Mays Healthcare Analytics 2022 Case Competition
Predicting Housing Insecurity

Table of Contents

Executive Summary 2

Background 2

Case Introduction and Business Problem 3

Definition of Metrics / Key Performance Indicators 3

Preliminary Research 4

Housing Insecurity 4

The Data 5

Exploratory Data Analysis 5

Data Cleaning 6

Additional Datasets 8

Modeling 8

Model Creation 8

Final Model Selection and KPI Evaluation 10

Fairness Analysis 12

Feature Evaluation 13

Business Implications 15

Housing insecurity is harmful to individuals 15

Housing insecurity is harmful to Humana 16

Current Approaches to resolving housing insecurity 17

Recommended approaches for Humana 17

Care Coordination Facilitation 18

Case Management 19

Meal Delivery 20

Intervention Summary 21

Intervention Comparison 22

Conclusion 22

Reference List 24

Executive Summary

Housing Insecurity affects approximately 1 in 5 people in the United States each year. Our goal was to use the data provided by Humana to identify which Medicare members are housing insecure using machine learning, and offer proposals to help members avoid prolonged housing insecurity so they can get back to living healthier lives.

Before creating models, we investigated the underlying data to learn more about the included variables. After gaining an understanding of these variables, we then removed null values from the dataset by removing columns, setting null values to zero, or replacing null values with the median depending on what information each column stored. We also converted some numeric columns into categorical columns where appropriate.

Once the data was processed, we constructed a tree-based XGBoost model to predict which members were likely at risk of being housing insecure. The primary metrics we used to evaluate our model were the AUC score and an evaluation of how the model performed across different demographics to ensure fairness in outcome. However, the model is only the first step in implementing need-based policies. We must understand the implications of housing insecurity to affect positive change.

Housing insecurity is harmful to both individuals experiencing housing insecurity and to Humana itself, as housing insecure individuals are more likely to incur large medical costs. Current approaches to mitigating housing insecurity include financial assistance, such as tax credits and affordable housing, as well as social assistance, such as shelters and healthcare provision. Given the complexity of this issue, we see several opportunities for Humana to provide assistance by implementing a three-pronged approach focusing on care coordination facilitation, case management, and meal delivery. We estimate this strategy would result in an annual savings of $1.66 million per 1,000 members flagged as potentially being housing insecure.

Background

In recent years, industry stakeholders have increasingly focused on health related social needs (HRSNs) to improve health outcomes. HSRNs are defined by the Center of Medicare and Medicaid Services as any health-harming conditions such as food insecurity or housing insecurity.[1] Additionally, 39% of Medicare users reported experiencing at least one unmet social

---

[1] Riseborough, P. (2017). *Managing health-related social needs: The prevention imperative in an accountable Health System*

need.[2] Studies have shown that these unmet social needs could increase the risk of developing chronic conditions, reduce the ability of the individual to properly manage these conditions, increase the cost of healthcare, and inevitably lead to avoidable healthcare utilization.[3] Thus, by addressing these needs, healthcare service providers could help at risk individuals decrease healthcare costs and reduce healthcare utilization.

Out of these social needs, housing insecurity consistently ranks as the one with the largest healthcare impact. Among Medicare patients, housing instability ties transportation as the number one indicator of predicting health status.[4] In addition, studies have shown that housing insecurity contributes significantly to hospitalizations and ED visits,[5] and intervention programs addressing housing insecurity lead to high cost savings for the patients and a large return on investment for healthcare providers.[6] [7] Therefore, it is a high priority for Humana to identify the population that is most at risk of housing insecurity.

Case Introduction and Business Problem

The current business problem is to identify populations with housing insecurity utilizing data provided by Humana regarding the Medicare Advantage Prescription Drug member population. With the given datasets, we will create a classification model based on the personal, medical, and regional information that is given to determine whether the Humana member is at risk of housing insecurity, identify the features that are most indicative of housing insecurity, and provide recommendations as opportunities of intervention by Humana.

Definition of Metrics / Key Performance Indicators

The first focus of our model is to output a binary variable which will indicate the housing insecurity status of a member. To ensure accuracy, the first key performance indicator we will use is the Area Under the Curve for the Receiver Operator Characteristic curve (AUC-ROC). The

---

[2] Coe, E. H. et al. (2022). *Understanding the impact of unmet social needs on consumer health and healthcare*

[3] *Accountable Health Communities Model* (no date) *Cms.gov*

[4] Coe, E. H. et al. (2022). *Understanding the impact of unmet social needs on consumer health and healthcare*

[5] Sadowski, L. S. *et al.* (2009) “Effect of a housing and case management program on emergency department visits and hospitalizations among chronically ill homeless adults: a randomized trial: A randomized trial”

[6] Basu, A. *et al.* (2012) “Comparative cost analysis of housing and case management program for chronically ill homeless adults compared to usual care”

[7] *Investing in social services as a core strategy for healthcare organizations: Developing the Business Case* (2018) *Commonwealthfund.org*

AUC was chosen as our primary metric as it indicates the accuracy of our model on how well it predicts whether a certain member is housing insecure or not.

The second focus of our model is our fairness indicator or disparity score. Given that housing insecurity could be closely tied to socioeconomic or demographic factors, we want to ensure our model does not disproportionately target certain members based on these variables. Therefore, by incorporating a fairness analysis into our model, we can establish fairness in our predictions.

Both metrics stated above will be explored in further detail in the modeling section.

Preliminary Research

Housing Insecurity

Before creating a machine learning-based analysis of this problem, we felt it was crucial to fully understand what is meant by the term "Housing Insecurity." Housing Insecurity generally refers to a multitude of housing-related issues that people may confront, including affordability, eviction, security, and condition[8][9]. Housing insecurity is not a new problem in the United States, but the recent COVID-19 pandemic exacerbated many pre-existing housing issues in the country that are still being felt today.

In the United States, rent or mortgage payments that are in excess of 30% of monthly income are considered to be "unaffordable." In 2020, 14% of households in the United States had to pay 50% of their monthly income to cover housing costs, and 20% of households overall were housing insecure. As record inflation affects the United States, rents continue to increase, as national rents increased by an average of 12% from March 2021, to March 2022.[9] With wages for lower-paying jobs stagnant (the federal minimum wage has remained at $7.25 since 2009, or $14,500 a year, which is barely above the poverty line), and the federal eviction moratorium ending at the beginning of October,[10] housing insecurity will likely increase in the near future.

Additionally, despite the worrying signs of increasing housing insecurity, housing insecurity is not distributed equally among America's populace. Per The Center for Economic and Policy Research (CEPR), since the pandemic began, housing insecurity for both Black and Hispanic renters has skyrocketed, as approximately 44% of renters in these groups experience housing

---

[8] *Measuring housing insecurity in the American housing survey* (no date) *Huduser.gov*
[9] *State of the Nation's Housing report* (no date) *Habitat for Humanity*
[10] *Federal Moratorium on Evictions For Nonpayment of Rent* (2021) *Nihc.org*

insecurity.[11] Housing insecurity is also not equally distributed between the 50 states, as Florida, Louisiana, Mississippi, and Texas reported the highest rates of housing insecurity, which mirrored troubling trends of food insecurity in these states.

The Data

Exploratory Data Analysis

First, we performed basic exploratory data analysis (EDA). The training set had 48,300 rows with 881 columns, while the holdout set had 12,220 rows with 880 columns. The missing column in the holdout set was the "hi_flag" that indicates whether the individual was housing insecure, which was withheld from us to score each team's models. After investigating the columns more closely, we removed certain columns that contained too many null values. We also converted some columns into categorical values rather than numeric values to ensure our model assigned weights to different variables correctly. We frequently used this approach for columns storing discrete scores for members where the score indicates the presence of many other factors.

Continuing with our EDA, we found that the training set had approximately 4.39% positive responses in the hi_flag variable. There were 60% females and 40% males in both the training set and the holdout set (shown below). However, the distribution of housing insecurity is not equal between men and women, with men representing 46% of all housing insecure persons in the dataset and having a housing insecurity risk 29% higher than women.

Table 1: Housing insecurity by sex

Sex	Number Housing Insecure	Total in Training Data	Percent of Whole
Male	973	19,200	5.07%
Female	1,145	29,100	3.93%
Total	2118	48,300	4.39%

[11] *Housing Insecurity By Race and Place During the Pandemic* (2021) *Center for Economic and Policy Research*

Additionally, while the exact races of the individuals in the data set were obscured, we wanted to ensure we understood the distribution of races and the percentage of members of each race experiencing housing insecurity (shown below).

Table 2: Housing insecurity by race

Race	Number Housing Insecure	Total in Training Data	Percent of Whole
5	99	1,068	9.27%
6	10	132	7.58%
3	44	759	5.80%
4	22	298	7.38%
2	470	7,706	6.10%
0	43	781	5.51%
1	1,429	37,549	3.81%
Unknown	1	7	14.29%
Total	2,118	48,300	4.39%

It is clear that housing insecurity does not affect each race equally. Every race except for race1 suffers from an above average rate of housing insecurity, especially race5, which has a housing insecurity rate 2.25 times the sample average, and over 2.6 times the rate of race1. No other race is as drastically above the sample average, but every race besides race1 may need additional assistance in terms of finding a solution to housing insecurity.

Data Cleaning

Knowing the importance of clean data in building a predictive model, we evaluated different data cleaning methodologies that we could perform. Our first step in cleaning the data was to investigate which variables had missing data. We observed that 260 of the 881 variables had at least one missing datapoint. We knew that simply dropping those columns was a possible option, but removing 30% of the provided columns seemed irresponsible without investigating the data further. Therefore, we examined each of the columns containing null values to determine what information the columns stored, whether the columns were likely to be relevant, and then we decided which null-handling approach to use.

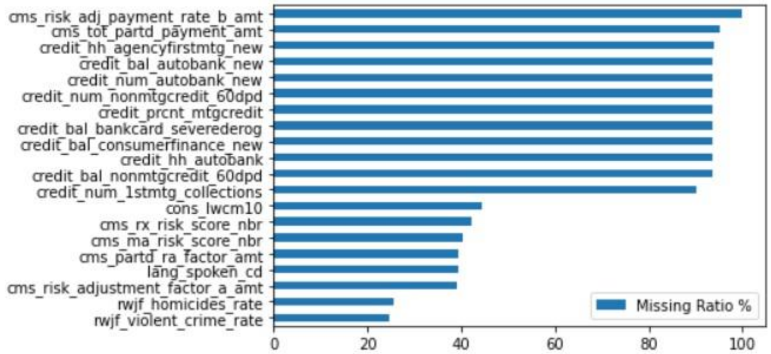

Figure 1: Top 20 columns: Ratio of missing data

Given best practices, we decided to drop any columns with greater than 75% of null values. This included all credit-related columns as well as some variables from the Centers of Medicare and Medicaid Services. For others, we were able to convert many of the null values to the mode, usually 0 or 480, when over 90% of the rows contained the same value. For instance, if the cost per month of prescriptions relating to pain management (rx_hum_77_pmpm_cost) was missing, we felt we could safely assume that that cost was 0. Otherwise, we reasoned, it would likely not be missing, as transactional data is often maintained well. For another set of variables, we chose to add an additional categorical variable column, showing that the variable was missing data. For example, the missing data for “lang_spoken_cd” could be revealing of other underlying demographic information that might be relevant to predicting housing insecurity. Indeed, our added variable, “lang_spoken_cd_nan” was retained by our final model, illustrating that we made the correct choice in adding this supplementary indicator variable. Finally, the null values in many columns with continuous data were able to be replaced by the median of that column as a safe estimate of the missing value.

Table 3: Missing data handling method

Scenario	Handling Method
Over 75% null	Drop column
Over 90% same value	Add mode
Categorical	Add indicator for missing

Continuous data	Add median

After the null values were handled, we needed to convert string variables into separate indicator columns to ensure the values could be fed through our different models (One Hot Encoding). This was challenging due to the seemingly-numeric variables that were actually ordinal in nature, meaning that using them as continuous variables would be unwise and lead to worse performance in our models. This was seen in variables that represented risk scores like “cci_score” and “dcsi_score.”

Additional Datasets

Though we were permitted to add additional data to the model, we elected against using outside data for three reasons. First, due to the anonymity of the data, finding a data set that could be accurately merged with the Humana-provided dataset would prove difficult. Though we could potentially merge based on certain demographics, too much would be estimated and imprecise for us to feel confident in using outside data. Second, we wanted our model to be usable immediately or in the very near future. Utilizing an outside dataset to make predictions would rely on that dataset always being available, recently updated, and of guaranteed quality – three requirements that seemed unrealistic and would introduce more risk than we were willing to accept. Finally, this competition takes place in a relatively short time period. We decided it was a better use of our time to commit to creating a top-tier model with a thorough report than to spend time searching for datasets that would meet our standards. Thus, we chose not to utilize any 3rd party data when constructing our model. Given a longer time frame or data with slightly less anonymization, we would have allocated more time searching for datasets that could be incorporated into our model.

Modeling

Model Creation

Once the data was finally cleaned, we began modeling. Knowing the data was not of identical scale, we scaled the data to a normal distribution to ensure equal weights be given to each of the variables. Though some of the methods we employed were not affected by the scale of the data (specifically, tree-based models), any regularization performed on regression models such as LASSO, Ridge, or ElasticNet would cause uneven weights to be given to variables of differing scales. Therefore, we scaled the data around a mean of 0 with a unit variance for each column in the dataset. Though this was unnecessary for our binary columns, performing this operation does no harm to them.

Next, we randomly split our data into two groups: a training set and a validation set, using a training set size of 80% of the provided data, and a validation set size of 20%.

Once we had our data ready, we proceeded to implement and train four different models: Random Forests, XGBoost, Cross-Validated LASSO Logistic Regression, and LightGBM. Due to the large numbers of variables in the dataset, we focused solely on models that included some form of feature-selection and regularization. This allowed us to only include relevant variables in the model, while simultaneously ensuring the model was not overfitting the training data.

Once the models were trained, we predicted values on the validation set, and calculated the Area Under the Receiver Operating Characteristic (ROC) Curve to investigate which of the models had the best preliminary performance. AUC was chosen as a primary validation metric for two reasons. First, AUC performs better as a validation metric when the dataset is unbalanced, which ours was. Second, as the AUC is essentially measuring multiple thresholds of a binary classification, it allows us to pick an appropriate cut-off ourselves that is best suited to the situation at hand.

The results of the initial models are shown in the table below.

Table 4: AUC of initial classification models

Model	Training AUC	Validation AUC
LightGBMT	0.9372	0.7507
XGBoost	0.8633	0.7455
Cross-Validated LASSO Logistic Regression	0.7535	0.7331
Random Forest	1.0	0.7000

As is apparent from the above table, the XGBoost and LightGBMT models performed better than any of the other models we attempted on our validation set, so we began hyper-tuning the parameters of these two models to optimize them and get the best results possible.

Despite the initial highest score from the LightGBM model, we were better able to tune the hyperparameters of the XGBoost model to get a higher AUC score on the holdout dataset. We utilized several iterations of a cross-validated grid search to tune the hyperparameters of our XGBoost model. The specific parameters we tuned and their purpose are shown here:

- Learning Rate (eta): The Learning Rate is a parameter that helps prevent overfitting. After every boosting step in the algorithm, the learning rate shrinks the weights of the features. The larger the Learning Rate, the more conservative the model.
- Number of Estimators: This refers to the number of boosting rounds, or the number of trees that are built when constructing the model. The higher this value, the longer the algorithm takes to run.
- Min Child Weight: The minimum child weight is the smallest permitted value for the sum of instance weight in a node (leaf). A higher value for min_child_weight makes the model more conservative.
- Colsample_bytree: This is a parameter that deals with subsampling of variables. Specifically, Colsample_bytree refers to the fraction of columns that are used for building each tree in the model. There are other subsampling methods (such as subsampling by level and subsampling by node, but we found that they did not improve model fit, so they were excluded from our tuning process).
- Max Depth: This is the maximum depth of each tree in the model. A higher value often leads to overfitting, and we found that the default value, 6, was the best for our model.

Final Model Selection and KPI Evaluation

After all the hyper parameters were tuned, we achieved an AUC of 0.755 on our test set and an AUC 0.759 on the holdout data set, per the Official Case Competition Leaderboard.

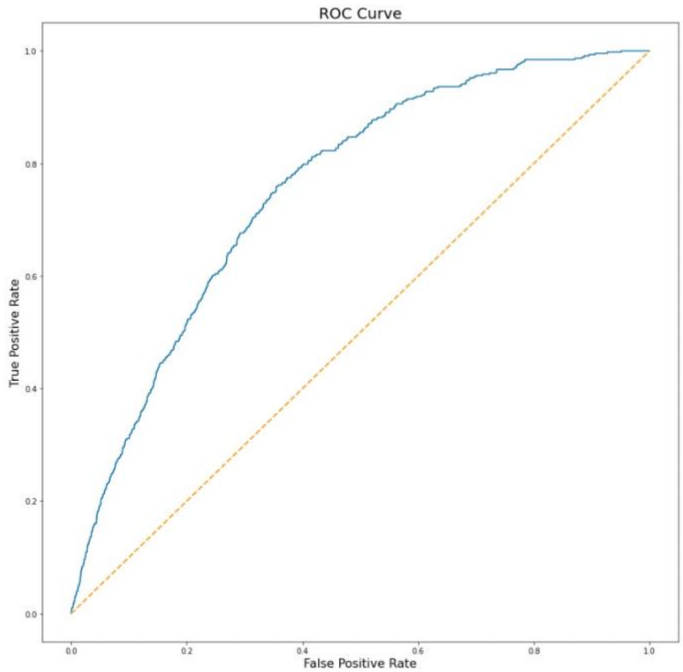

Figure 2: ROC curve of Final XGBoost model

The consequences of incorrectly identifying someone as at risk of experiencing housing insecurity are less significant than the impact of failing to identify someone who is experiencing housing insecurity. Therefore, we set an initial threshold that would capture 50% of all people experiencing housing insecurity. In capturing such a large percentage of people experiencing housing insecurity (true positives), we captured an even larger number of people who are not experiencing housing insecurity (false positives). The confusion matrix at this threshold is below:

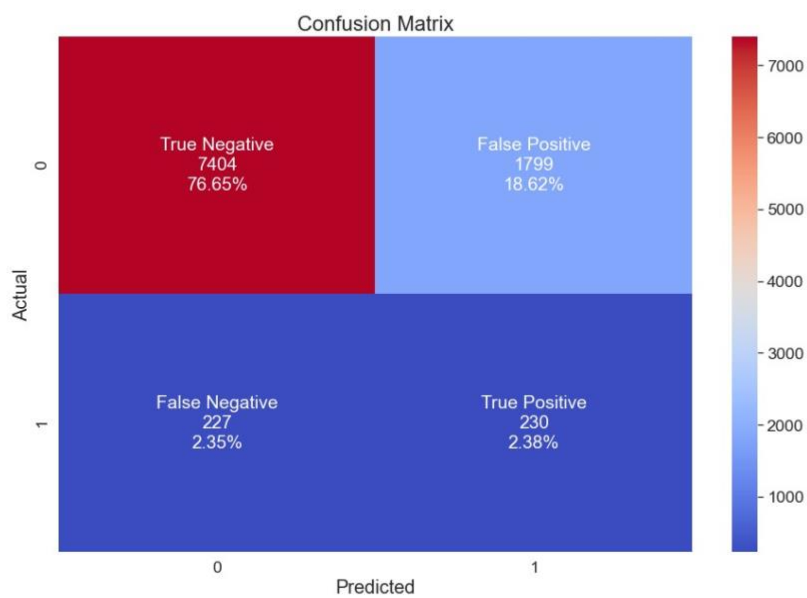

Figure 3: Confusion matrix of predictions from XGBoost model

Fairness Analysis

Next, we wanted to check the demographic breakdown of our model to ensure fairness in our predictions. In the training data, men were 29% more likely to be housing insecure than women. Looking at the model predictions, the model maintains this distribution, predicting 30% more men to be housing insecure than women.

Looking at the performance of our model across each race, we see the model generally performed well in predicting the number of people experiencing housing insecurity, though there does seem to be a small penalty for race0 in our model. Future iterations of this model should focus on improving outcomes for race0. The table below shows the likelihood multiplier of being housing insecure for each race. We set race1 to have a likelihood multiplier of 1 since this race has the lowest housing insecurity rate in the training data. The middle column stores the difference in probability of being housing insecure in our training set (for example, race0 is 1.45x more likely to be housing insecure than race1 and race 5 is 2.43x more likely to be housing insecure than race1). The rightmost column stores the likelihood multiplier for a person from each race being housing insecure from our trained model. We can see our likelihood multipliers align reasonably well with the underlying differences in housing insecurity across each race in the training data apart from race0.

Table 5: Fairness evaluation by race

Race	**Training Set** Housing Insecurity Likelihood Multiplier	**Model** Housing Insecurity Likelihood Multiplier
1	1	1
0	1.45	.94
2	1.60	1.9
3	1.52	1.6
4	1.94	2.09
5	2.43	2.67
6	1.99	2.86

Feature Evaluation

In tree-based models, the Feature Importances are used to describe the relative importances of different explanatory variables in a model. However, they lack the ability to show how each variable affects the prediction. So instead we use SHAP values (**Sh**apely **A**dditive Ex**p**lanations) to give us an idea of how the value of each variable affects the final prediction, similar to the coefficients in a linear regression model. First, we can take a look at the feature importances to get a sense of which features are affecting the model's predictions the most. The top 20 most important features are shown below.

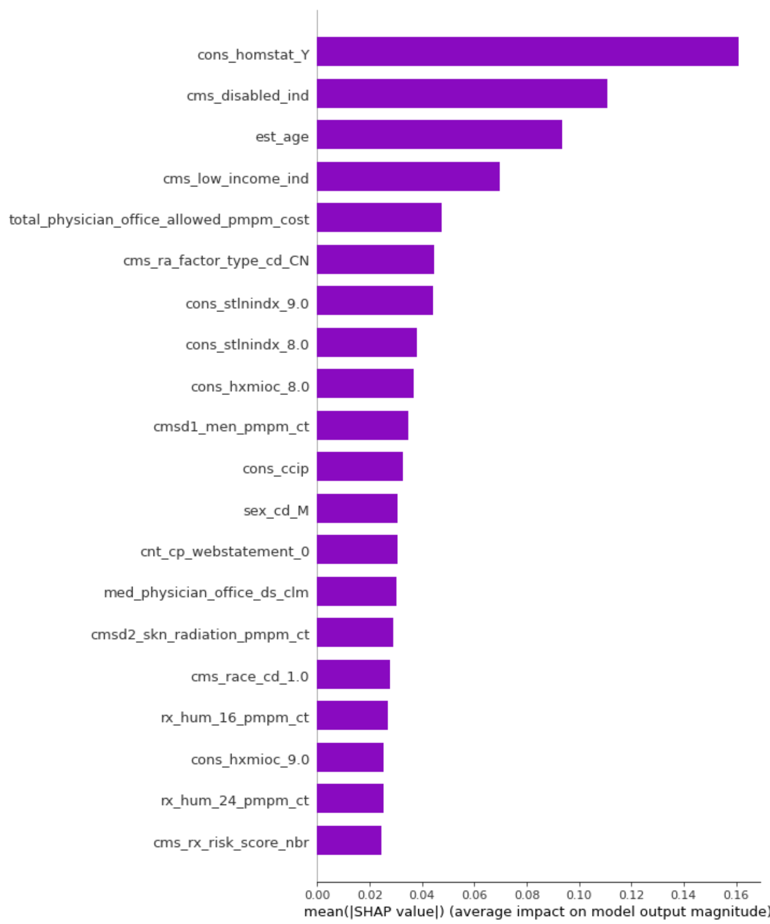

Figure 4: Feature selection from SHAP values

Grouping the above features into 8 categories reveals which types of features are most important in determining housing insecurity. Clearly, demographic information is the most important type of variable when determining whether a member is housing insecure or not. However, it is also important to note that variables relating to battling chronic diseases, or managing a complex health situation are also quite important in the model's performance.

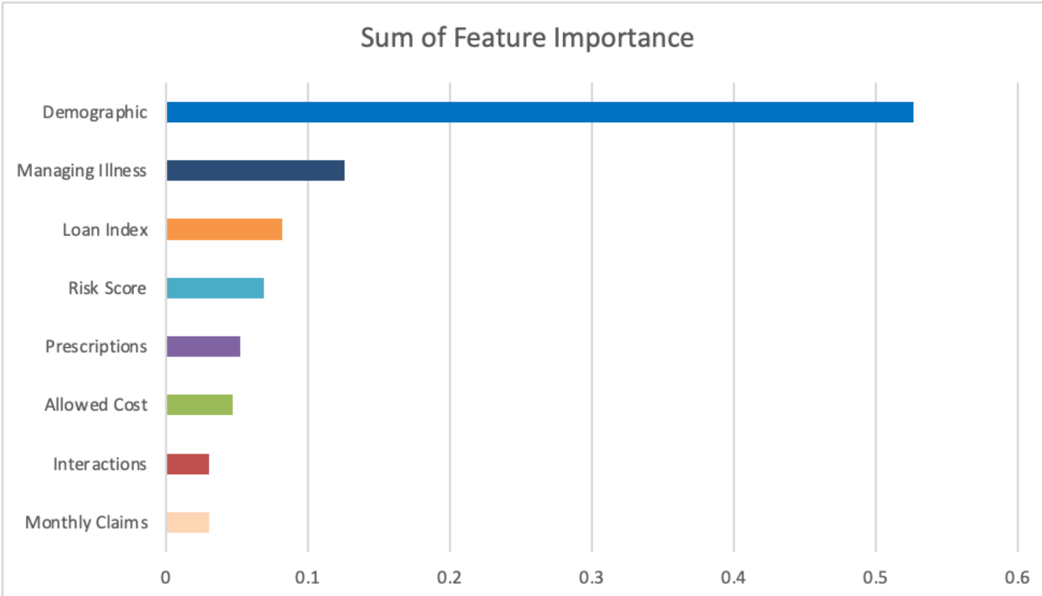

Figure 5: Categorical feature importance from XGBoost model

Next, we can analyze the top-20 shap values to evaluate how each important variable affects the end model result.

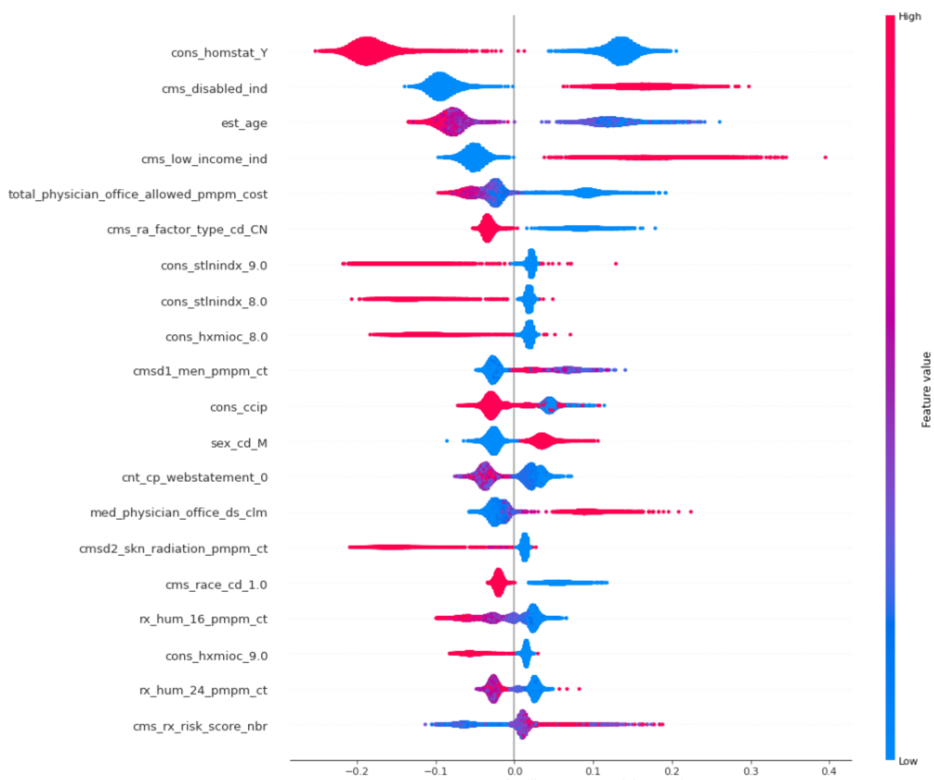

Figure 6: Effect magnitude of features

This plot can be interpreted as how a high/low value for a variable affects the likelihood of a person being housing insecure. For example, the first variable, cons_homstat_Y indicates whether a person is a homeowner or not. Homeowners are far less likely to be housing insecure than non-homeowners. Further, cms_disabled_ind is a binary variable for whether a Medicare Supplement member is under 65. From the beeswarm plot, we can see that being under 65 leads to a higher likelihood of being housing insecure than being over 65. This is reiterated by the est_age variable, which is a member's relative age, and we see that the younger a person is, the more likely they are housing insecure.

Business Implications

Housing insecurity is harmful to individuals

Individuals who experience housing insecurity are likely to suffer many additional hardships related to their health and quality of life, both directly and indirectly. People who are homeless tend to have higher rates of serious chronic diseases such as diabetes, and hypertension, in addition to being more likely to suffer from substance abuse disorders. For these reasons and other reasons related to housing instability, the mortality rate is significantly higher for people experiencing homelessness than for the general population.[12] This relationship holds even after controlling for previous health. Among the housing insecure, negative health outcomes are especially pronounced in middle aged people, as well as Black people.[13] Cancer survivors are also likely to experience increased levels of housing insecurity, which can pose additional health burdens for people who have already experienced serious illnesses.[14] Depression and substance use among mothers seems to increase the risk of experiencing housing insecurity for families that are already vulnerable.[15] The evidence suggests that housing insecurity tends to be compounded with health issues.

Additionally, people who experience homelessness and other forms of housing insecurity tend to have weaker social support networks. This can manifest itself through smaller networks, which can result in longer periods of unemployment, and reduced access to support systems,

---

[12] *Housing instability* (no date) *Health.gov*

[13] Bhat, A. C. *et al.* (2022) "A longitudinal analysis of the relationship between housing insecurity and physical health among midlife and aging adults in the United States"

[14] Coughlin, S. S. and Datta, B. (2022) "Housing insecurity among cancer survivors: Results from the 2017 behavioral risk factor surveillance system survey"

[15] Marçal, K. E. (2021) "Perceived instrumental support as a mediator between maternal mental health and housing insecurity"

which can alleviate hardships.[16] Housing insecurity can also negatively affect employment outcomes even during employment. While employment offers access to income, the jobs available to people experiencing housing insecurity may not be near places that are affordable for these same people. As a result, people may have to spend significant periods of time every day commuting to and from work, which can impose additional hardships. Oftentimes wages are not high enough to afford rent.[17] In these situations, there are limited avenues for people to improve their housing situation without access to additional support.

Isolating causation from correlation would be a valuable area for future research, but the model developed here to identify people who are potentially experiencing housing insecurity does not need to separate causation from correlation. The presence of correlation at a minimum is sufficient for targeting individuals for potential interventions.

Housing insecurity is harmful to Humana

The impact of housing insecurity on people's lives is very serious and deserves significant attention, but there are spillover effects to organizations such as Humana as well. As we mentioned above, there tend to be serious health consequences for people experiencing housing insecurity. The average cost for an ED visit was over $2,000 in 2019 according to United Health Group.[18] This represents the average, however, and people experiencing housing insecurity are more likely to have more serious chronic conditions that push this number even higher. The prevalence of serious conditions may also increase the probability of needing to visit the ED, which increases the number of visits to the ED relative to people who are not experiencing housing insecurity.

There are a variety of reasons why people experiencing housing insecurity may tend to have worse health conditions. Some research has indicated that people experiencing housing insecurity who are not homeless are making tradeoffs between purchasing food and paying for rent.[19] This decision can contribute to additional health problems down that line that may result in additional costs for Humana and worse outcomes for its members. Members who are cost-burdened or severely cost-burdened are making additional tradeoffs as well, such as choosing between rent and medication, or between rent and preventative treatment. These decisions are of course logical, but they may contribute to deteriorating health conditions over time that

---

[16] Kim, H., Burgard, S. A. and Seefeldt, K. S. (2017) "Housing assistance and housing insecurity: A study of renters in southeastern Michigan in the wake of the great recession"

[17] Jones, K. *et al.* (2020) "Working and homeless: Exploring the interaction of housing and labour market insecurity"

[18] *The high cost of avoidable hospital emergency department visits* (2021) *Unitedhealthgroup.com*. [19] Leopold, J. *et al.* (no date) *Improving measures of housing insecurity: A path forward*

result in more emergency/treatment-oriented care, which is far less cost-effective than preventative care and general wellbeing.

Current Approaches to resolving housing insecurity

The primary methods for addressing housing insecurity today can be bucketed into two categories: financial and social. Some of the major interventions pursued in each category include:

Financial assistance:

- Tax Credits such as the Earned Income Tax Credit (EITC)
- Low-interest loans offered through USDA Rural Development programs
- Renters assistance such as through the Emergency Rental Assistance Program (ERA)
- Emergency Housing Vouchers through the American Recovery Plan

Social assistance

- Shelters that provide temporary to permanent housing
- Education/training programs to inform people of available resources and to build skills to improve employment outcomes
- Food assistance to provide meals and groceries
- Healthcare assistance to provide basic preventative care and examinations for underserved populations

While many stakeholders support and participate in programs across each area, the large financial assistance programs tend to be sponsored by the government at the federal, state, and local levels. Nonprofits and other local support groups tend to focus on social programs, though the government is still heavily involved with these programs. Such social programs have varying degrees of success, but housing insecurity is a lingering issue, so there is ample opportunity to provide additional assistance to reducing housing insecurity.

Recommended approaches for Humana

We recommend Humana consider a three-pronged approach for tackling housing insecurity that focuses on Care Coordination Facilitation, Case Management, and Meal Delivery, as these three areas align closely with Humana's goals and interests as stated in Humana's June 2020 Housing Brief. The primary goal of the Care Coordination Facilitation prong is to encourage members to take advantage of the health resources available to them. The Case Management prong focuses on initiating housing-related conversations with Medicare members and determining if these members would benefit from receiving housing assistance. The Meal Delivery prong offers the opportunity to assist members with an area of their life that may fall to the wayside when members are more concerned about paying rent and for medical bills. Each prong requires different levels of involvement from Humana, but all should be considered as potential cost-saving and health-improving initiatives. The largest risk to Humana with each of these initiatives is that Humana will need to rely some amount on existing networks and support structures that have their own carrying capacity. While they could likely provide support to additional people, Humana would need to conduct research about how many people can be supported by these programs and ensure the model is set up to flag the appropriate number of people.

Care Coordination Facilitation

Due to Humana's awareness of each member's points of contact within the medical system, Humana could facilitate conversations between members and care coordinators to ensure patients have access to the health and wellness resources available to them in their community. This could include providing access to medical homes and social services offering resources such as housing, food, and transportation. Initiating contacts with members to share targeted information and available resources is a relatively low-lift intervention for Humana and faces little risk of excess resource consumption because few members who would not benefit from these services would likely take advantage of them.

The Center for Health Care Strategies evaluated an intervention similar to this for a population that was at risk of having unmet social needs.[18] This patient population could be similar to the patients flagged by our model as potentially experiencing housing insecurity. Patients enrolled in this study experienced a 26% reduction in EMS trips along with a savings of $17,562 per avoided inpatient admission and $1,387 per avoided ED visit. Participants in this study also experienced an increase in stable housing. Among the top 1,000 scored patients in our test set for our model, there were 24 total monthly ED visits and 10 monthly acute inpatient admissions, meaning 288 annual ED visits and 120 annual inpatient admissions. If 50% of these trips took place in an ambulance and this proposed intervention results in a 26% reduction in those visits, that results in 37 fewer annual ED visits and 16 fewer annual inpatient admissions. Using the estimated costs from above, that translates to an annual savings of $332,000 for the first 1,000 scored patients in our model. Implementing this at scale would of course require coordination with local resources and this would have costs for Humana. However, if this model were rolled out to a larger patient set, Humana could provide this service to many of its members.

[18] *2-1-1 San Diego: Connecting partners through the community information exchange* (no date) *Chcs.org*

Case Management

Humana could take efforts to ensure people have adequate housing solutions after being discharged from care by performing member outreach as members are flagged by the model. The types of housing that could be offered to members can range from temporary recovery housing to more stable, affordable long term housing. Humana could partner with local organizations and providers around the country to help place members struggling with housing insecurity into more stable situations as appropriate based on the members' needs. This could be included as a post-treatment follow up from Humana when processing claims. If pursued as an opt-in strategy for members with active outreach from Humana, Humana would have limited exposure to excess resource consumption by its members. People who are not housing insecure are unlikely to accept offers of help to find different housing situations because they will not want to leave their own housing. There is a small risk of people who are housing insecure not wanting to leave their own house, but that is an inherent risk when offering to place people in new housing.

Results from one study suggest that providing case management related to procuring housing for people post-hospital discharge resulted in annual cost savings for the treatment group of $6,037 per person, with the largest benefits concentrated among people who are the most housing insecure.[19] Another similar study found that case management efforts to secure housing after hospital discharge resulted in 49 fewer hospitalizations, 270 fewer hospital days, and 116 fewer ED visits for the next year per 100 homeless adults offered this program.[20] We will use the observed $6,037 in savings to estimate the potential benefit to Humana from this approach.

Suppose that Humana implements our model and sets a threshold for outreach that correctly identifies 50% of all members who are experiencing housing insecurity. This would have resulted in 227 people correctly being identified as housing insecure from our test set. If 50% of

[19] Basu, A. *et al.* (2012) "Comparative cost analysis of housing and case management program for chronically ill homeless adults compared to usual care"

[20] Sadowski, L. S. *et al.* (2009) "Effect of a housing and case management program on emergency department visits and hospitalizations among chronically ill homeless adults: a randomized trial: A randomized trial"

the people who are housing insecure take advantage of this program, then 114 people would have accepted help. With an annual savings of $6,037 per person (which is a conservative estimate given that the people most likely to accept help will tend to be the people in more dire situations), this corresponds to an annual savings of $688,000. To identify 50% of all members experiencing housing insecurity in our test set, we had 1,799 false positives (people who are flagged, but not actually housing insecure). This implies a savings of $339,000 if we divide the initial savings estimate by 2.026 to get the estimated savings per 1,000 flagged members. Note that we can assume no members incorrectly flagged as being housing insecure would accept new housing because people are very unlikely to want to leave their home unnecessarily. This program would have significant startup costs and ongoing resource utilization to contact members, but the startup costs will be relatively constant regardless of the number of members considered for this outreach program. It is worth noting that Humana should not set its threshold to flag all people as being housing insecure since there will be limited resources available and Humana will want to ensure that those most likely to be housing insecure are offered new housing first.

Meal Delivery

Many studies have produced strong evidence that medically-tailored meal delivery programs can greatly reduce annual medical costs for people. This is likely to be especially true for people experiencing housing insecurity. In addition to being more prone to suffer from chronic illnesses requiring more careful attention to nutrition, people experiencing housing insecurity may have to choose between purchasing food and paying rent. Further, housing insecurity is likely to be more prevalent in neighborhoods that lack access to affordable healthy food choices. Humana could sponsor or aid local programs that construct and deliver medically-tailored meals to people. This would involve partnering with providers to figure out what meals are appropriate for people and then providing funding or physical resources to existing programs to purchase food, prepare meals, and then deliver them to members. Partnering with local programs offers the added benefit of supporting strong local communities. This could result in positive externalities that put downward pressure on overall housing insecurity.

While this intervention would be more likely to be taken advantage of by people who are not housing insecure than providing housing solutions, Humana should not be too concerned about this because people consuming healthy food results in lower annual costs for Humana and a healthier member population overall. One potential drawback to this intervention is it may not be feasible in rural areas. This could create additional disparities for those living in rural communities.

To predict the return on investment for such a program, we will assume that the ROI is $220 per month for medically tailored meals based on a study from Berkowitz et al. in 2018.[21] Notably, this study was not specific to people at high risk of housing insecurity. Suppose that for every 1,000 members flagged as potentially being housing insecure, only 50% of members are eligible for a meal delivery program due to their location and suppose 75% of all eligible and flagged members accept assistance from a meal delivery service. This would result in 375 members receiving meals per 1,000 flagged members. Not all of these flagged members would actually experience housing insecurity because we would have incorrectly flagged some people as being housing insecure who are not actually housing insecure, but the exact ratio depends on how strict Humana chooses to be with flagging members. Since the $220 monthly savings per person estimate is not specific to people who are housing insecure, however, we can apply this estimate to all people flagged by our model regardless of their true status. This is appropriate since some of the most heavily weighted predictors in our model correspond to serious and/or chronic health conditions. With 375 members per 1,000 flagged members receiving meals at a benefit of $220 per month per person, that results in an annual cost savings of $990,000 per year per 1,000 flagged members.

Intervention Summary

Between Care Coordination Facilitation, Case Management, and Meal Delivery, we estimate Humana could save $1.66 million per year per 1,000 patients. This cost savings estimate does not account for the startup and operational costs required to implement these programs, but the startup costs would only be required once and the operational costs would likely be relatively minimal and not increase rapidly as the program scales. Additionally, there is likely a gradual diminishing marginal return for these initiatives. Further research would be needed to determine how quickly the cost-effectiveness diminishes. The major decision for Humana related to implementing these prongs would be determining:

[21] Berkowitz et al., "Meal Delivery Programs Reduce The Use Of Costly Health Care In Dually Eligible Medicare and Medicaid Beneficiaries"

1. How many resources can be allocated to these initiatives?

2. Once that answer is known, how many members can safely be flagged each year without overwhelming Humana's available resources

The answers to these two questions can help Humana determine where to set its thresholds for intervention using our model.

Intervention Comparison

Prong	Intervention Summary	Pros	Cons
Care Coordination Facilitation	Facilitate conversations between members and care coordinators	- Focuses on health rather than other SDOH - Similar approaches have improved housing stability in addition to health - Initiating conversations requires relatively little effort from Humana	- Sharing the information does not guarantee member action - The initial study had few participants.
Case Management	Ensuring members have adequate housing after receiving care	- Can be added on as follow-up after care - Low risk of more people consuming resources than necessary - High-touch channel from Humana can offer opportunities to recommend additional care to patients.	- Requires active involvement from Humana - People who are housing insecure may not use offered resources
Meal Delivery	Deliver medically tailored meals to people	- Less invasive than providing housing options - Can build sustainable habits among members - Helps support local communities	- May be utilized by people who are not housing insecure - May not be feasible for members in rural areas

Conclusion

Housing insecurity is a complex issue that is caused by, and can cause, a variety of issues. Being able to identify people who are at risk of experiencing housing insecurity offers Humana the opportunity to improve the health of Medicare members and build stronger local communities to support those Medicare members while also generating net savings for Humana. We propose implementing an XGBoost model to identify Medicare members that are most likely to be experiencing housing insecurity. Our final model had an AUC of 0.755 on our validation set and an AUC of 0.759 on the official leaderboard, which indicates that our model performs well and can be implemented with confidence. Once the model is implemented, we propose a three-pronged approach for Humana that emphasizes Care Coordination Facilitation, Case Management, and Meal Delivery. Collectively, these three initiatives target multiple aspects of life that may be challenging for people experiencing housing insecurity. While startup and ongoing operational costs need to be researched by the Humana team, we estimate the savings from implementing our proposed prongs would be around $1.66 million per year per 1,000 members flagged as potentially being housing insecure.

Reference List

2-1-1 San Diego: Connecting partners through the community information exchange (no date) *Chcs.org*. Available at: https://www.chcs.org/media/2-1-1-San-Diego-Case-Study_080918.pdf (Accessed: October 16, 2022).

2022 State of the Nation's Housing report (no date) *Habitat for Humanity*. Available at: https://www.habitat.org/costofhome/2022-state-nations-housing-report-lack-affordable-housin g (Accessed: October 16, 2022).

Accountable Health Communities Model (no date) *Cms.gov*. Available at: https://innovation.cms.gov/innovation-models/ahcm (Accessed: October 16, 2022).

Basu, A. *et al.* (2012) "Comparative cost analysis of housing and case management program for chronically ill homeless adults compared to usual care," *Health services research*, 47(1 Pt 2), pp. 523–543. doi: 10.1111/j.1475-6773.2011.01350.x.

Berkowitz, S. A. *et al.* (2018) "Meal delivery programs reduce the use of costly health care in dually eligible Medicare and Medicaid beneficiaries," *Health affairs (Project Hope)*, 37(4), pp. 535–542. doi: 10.1377/hlthaff.2017.0999.

Bhat, A. C. *et al.* (2022) "A longitudinal analysis of the relationship between housing insecurity and physical health among midlife and aging adults in the United States," *SSM - population health*, 18(101128), p. 101128. doi: 10.1016/j.ssmph.2022.101128.

Coe, E. H. *et al.* (2020) *Understanding the impact of unmet social needs on consumer health and healthcare*, *Mckinsey.com*. McKinsey & Company. Available at: https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/understan ding-the-impact-of-unmet-social-needs-on-consumer-health-and-healthcare (Accessed: October 16, 2022).

Coughlin, S. S. and Datta, B. (2022) "Housing insecurity among cancer survivors: Results from the 2017 behavioral risk factor surveillance system survey," *Journal of cancer policy*, 31(100320), p. 100320. doi: 10.1016/j.jcpo.2021.100320.

Federal Moratorium on Evictions for Nonpayment of Rent (2021) *Nlihc.org*. Available at: https://nlihc.org/sites/default/files/Overview-of-National-Eviction-Moratorium.pdf (Accessed: October 16, 2022).

Housing insecurity by race and place during the pandemic (2021) *Center for Economic and Policy Research*. Available at:

https://cepr.net/report/housing-insecurity-by-race-and-place-during-the-pandemic/ (Accessed: October 16, 2022).
Housing instability (no date) *Health.gov*. Available at: https://health.gov/healthypeople/priority-areas/social-determinants-health/literature-summaries/housing-instability (Accessed: October 16, 2022).

Investing in social services as a core strategy for healthcare organizations: Developing the business case (2018) *Commonwealthfund.org*. Available at: https://www.commonwealthfund.org/publications/other-publication/2018/mar/investing-social-services-core-strategy-healthcare?redirect_source=/publications/publication/2018/mar/investing-social-services-core-strategy-healthcare-organizations (Accessed: October 16, 2022).

Jones, K. *et al.* (2020) "Working and homeless: Exploring the interaction of housing and labour market insecurity," *Social policy and society: a journal of the Social Policy Association*, 19(1), pp. 121–132. doi: 10.1017/s1474746419000332.

Kim, H., Burgard, S. A. and Seefeldt, K. S. (2017) "Housing assistance and housing insecurity: A study of renters in southeastern Michigan in the wake of the great recession," *The social service review*, 91(1), pp. 41–70. doi: 10.1086/690681.

Leopold, J. *et al.* (no date) *Improving measures of housing insecurity: A path forward*, *Urban.org*. Available at: https://www.urban.org/sites/default/files/publication/101608/improving_measures_of_housing_insecurity.pdf (Accessed: October 16, 2022).

Marçal, K. E. (2021) "Perceived instrumental support as a mediator between maternal mental health and housing insecurity," *Journal of child and family studies*, 30(12), pp. 3070–3079. doi: 10.1007/s10826-021-02132-w.

Measuring housing insecurity in the American housing survey (no date) *Huduser.gov*. Available at: https://www.huduser.gov/portal/pdredge/pdr-edge-frm-asst-sec-111918.html (Accessed: October 16, 2022).

Riseborough, P. (2017) *Managing health-related social needs: The prevention imperative in an accountable health system*, *JSI*. Available at: https://www.jsi.com/managing-health-related-social-needs-the-prevention-imperative-in-an-ac countable-health-system/ (Accessed: October 16, 2022).

Sadowski, L. S. *et al.* (2009) "Effect of a housing and case management program on emergency department visits and hospitalizations among chronically ill homeless adults: a randomized trial: A randomized trial," *JAMA: the journal of the American Medical Association*, 301(17), pp. 1771–1778. doi: 10.1001/jama.2009.561.

The high cost of avoidable hospital emergency department visits (2021) *Unitedhealthgroup.com*. UnitedHealth Group. Available at: https://www.unitedhealthgroup.com/newsroom/posts/2019-07-22-high-cost-emergency-department-visits.html (Accessed: October 16, 2022).