Analyzing Financial Resilience in U.S. Nursing Homes: Influential Factors and the Impact of COVID-19

Bach Nguyen, Nicolas Loke, Rajan Shukla - Advisor: Dr. Raju Mainali
Missouri State University
2024 Business Analytics Competition

BACKGROUND

The U.S. nursing home sector is a critical component of the national healthcare system. Eider Lee and Matt Chur are looking for investing options in nursing home.

Geri Attrick and his team will be analyzing financial patterns, key elements, and the effects of Covid-19 on the U.S. nursing homes market to identify possible investment opportunities.

From the data analysis, Elder and Matt will understand the various investment opportunities in nursing homes, including the locations and methods for their investment.

RESEARCH QUESTIONS

FINANCIAL PERFORMANCE: How has the key factors of the financial performance in U.S. nursing homes influenced profitability and sustainability from 2015 to 2021?

INFLUENTIAL FACTORS: Which operational, regulatory, and market factors have had the most significant impact on the financial health of U.S. nursing homes during the study period?

COVID-19 IMPACTS: How has the onset of the COVID-19 pandemic altered the financial situation of U.S. nursing homes, and what implications does this have for future pandemic preparedness?

METHODOLOGY

TWO-STAGE LEAST SQUARES (2SLS) APPROACH: To predict the financial and overall performance of nursing homes, 2SLS regression model was used to address potential endogeneity in explanatory variables.

1. **STAGE ONE - INSTRUMENTATION:** Instrumental variables (IVs) are identified and utilized to predict the endogenous variables. IVs are chosen based on their strong correlation with endogenous predictors and their expected zero correlation with the error term in the outcome equation.
2. **STAGE TWO - ESTIMATION:** The predicted values from the first stage are used as inputs in the second stage, where the model estimates the impact of predictors on nursing home performance. This stage provides the adjusted coefficients, offering insight into the causal relationships.
3. **VALIDATION:** To ensure the robustness of our instruments, we conduct overidentification tests, along with tests for endogeneity. These tests validate the suitability of our IVs and the specification of the 2SLS model.
4. **IMPLEMENTATION:** The 2SLS analysis is conducted using specialized econometric software, ensuring precision in our estimations and interpretations. The final model is assessed for statistical significance and the strength of the predictors.

REFERENCES

- Cuadros, D. F., Branscum, A. J., Mukandavire, Z., Miller, F. D., & MacKinnon, N. (2021, April 22). Dynamics of the covid-19 epidemic in urban and rural areas in the United States. PubMed Central. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8061094/
- Centers for Disease Control and Prevention. (2024, April 7). COVID-19 Nursing Home Data. Centers for Medicare & Medicaid Services Data. https://data.cms.gov/covid-19/covid-19-nursing-home-data
- Economic Research. (2023, December 19). Inflation, consumer prices for the United States. FRED. https://fred.stlouisfed.org/series/FPCPITOTLZGUSA
- Maldonado, R. W., Pradhan, R., Dayama, N., Lord, J., & Gupta, S. (2019, February 9). Sage Journals. Http://journals.sagepub.com/doi/abs/10.1177/0887302x07303626/
- Mather, M., & Scommegna, P. (2024, January 9). Fact sheet: Aging in the United States. PRB. https://www.prb.org/resources/fact-sheet-aging-in-the-united-states/
- Nursing home compare five-star quality rating system. (n.d.). https://www.cms.gov/medicare/provider-enrollment-and-certification/certificationandcompliancdownloads/consumerfactsheet.pdf
- Nursing facilities. KFF. (n.d.-b). https://www.kff.org/state-category/providers-service-use/nursing-facilities/
- The impact of covid-19 on Medicare beneficiaries in ... Centers for Medicare & Medicaid Services. (n.d.-b). https://www.cms.gov/files/-document/medicare-covid-19-nursing-home-analysis.pdf

TOOLS USED

GENERAL INFORMATION

Number of Nursing Home Facilities and Residents (2015-2021)

Year	Resident	Nursing Home Facilities
2015	1367548	15648
2016	1350121	15639
2017	1342876	15670
2018	1307982	15609
2019	1330591	15523
2020	1316950	15397
2021	1098305	15291

- Between 2015 and 2020, the nursing facility resident population decreased by 3.70%, then in 2021, it dropped sharply by 16.60%, mainly because of pandemic-related fatalities.
- The number of nursing facilities remained steady from 2015 to 2019, but over 100 facilities closed annually in 2020 and 2021 (less than 1%/year).

Top 5 States With The Most Residents/Facility Ratio (2021)

State	R/F Ratio
NY	142.28
DC	114.00
NJ	97.62
FL	91.15
PA	89.46

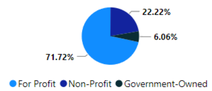
- 4 out of the 5 leading states with the highest ratios are located in the Northeast region.

Top 5 States With The Most Residents Density (2021)

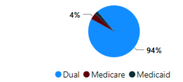
State	Density
DC	31.77
RI	5.84
NJ	4.75
MA	3.85
CT	3.70

- D.C. stands out as its density surpasses that of the second most dense state by over 5.44 times.

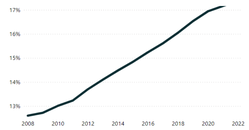
Distribution of Certified Nursing Facilities by Ownership Type



Distribution of Certified Nursing Facilities by Certification Type

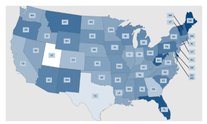


The 65+ Age Group Share in the U.S. (2008-2022)



- The 65+ age group's share of the total population rose 4.79% from 2008 to 2022.
- According to PRB, it is projected to increase 47% from 58 million in 2022 to 82 million by 2050, and the share of the total population is projected to rise from 17% to 23%.

The 65+ Age Group Share In Each State (07/2021)



- The age group share in most Southeast and Northeast states exceeds the US average.
- The Southern states' age group share that is generally lower than the Northern states.

Nursing Home Facilities Distribution (07/2021)

Nursing Home Facility Residents Distribution (07/2021)

- Both maps share the similar distribution structure. Most of the Nursing Home Facilities are located in the Midwest and Northeast region.
- On the remaining part of the US, California, Texas, and Florida placed 2nd, 3rd, and 4th states in the number of nursing home residents.

Average of Percent Vaccinated Residents by States (2021)

Residents Distribution - Residents Vaccinated Comparison:
In both maps, there is a contrasting trend where states with smaller populations tend to have higher vaccination rates. However, in some states like California, Nevada, and Arizona, there is a correlation between the distribution of vaccines and the percentage of vaccinated residents.

Distribution - Popular Density Comparison:
Although California, Texas, and Florida are among the states with the highest populations, they do not rank in the top five for population density, owing to their large geographical sizes.

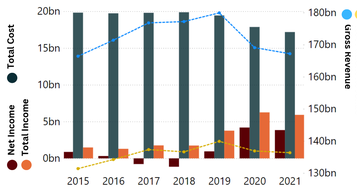
Age Group Share - Distribution Comparison:
The Age Group Share map shows an inverse relationship to the Distribution maps, where states with many facilities and residents have a lower percentage of people aged 65 and above, except for Florida, where the trend does not apply.

DISCUSSION AND RESULTS

FINANCIAL PERFORMANCE & INFLUENTIAL FACTORS

WHAT INSIGHTS CAN BE DERIVED FROM THE AVAILABLE DATA REGARDING THE FINANCIAL PERFORMANCE OF NURSING HOME?

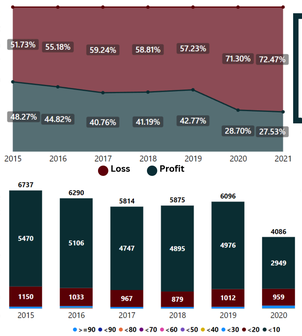
FINANCIAL INDICATORS: Revenue, Income, and Costs



- **Gross Revenue** and **Net Patient Revenue** experienced an increase from 2015 to 2019, followed by a decline during the pandemic period.
- **Total Costs** were stable from 2015 to 2018, dipped in 2019, and significantly decreased in 2020 and 2021. As a result, **Net Income** and **Total Income** are low from 2015 to 2018 but increased from 2019 to 2021.

Note: **Net Patient Revenue** is calculated by deducting **Contractual Allowances and Discounts** from **Gross Revenue**. Using Gross Revenue alone for financial ratios is inaccurate due to significant allowances and discounts. Therefore, it's essential to adjust Gross Revenue to Net Patient Revenue to avoid errors like facilities reporting a Net Income higher than Gross Revenue.

FINANCIAL INDICATORS: Operating Margin at Loss and Distribution by Year

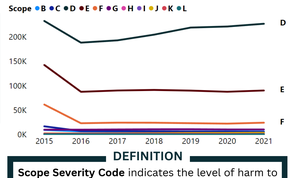


DEFINITION

The 100% Stacked Area Chart shows the percentage of nursing home facilities with loss and profit operating margin. **The bar chart** illustrates how the positive operating margin was distributed over the period.

- In 2015, 51.73% of nursing homes were operating at a loss, of which 33.60% were nonprofit or government-owned facilities.
- By 2019, this figure had risen slightly to 57.23%, with 33.81% being nonprofit or government facilities.
- In 2015, operating margins of 10% or less were reported by 81.19% of profit nursing homes, with a slight increase to 81.63% by 2019.

HEALTH DEFICIENCIES: Scope Severity Code



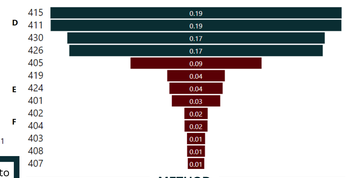
DEFINITION

Scope Severity Code indicates the level of harm to the resident(s) involved within the nursing home, ranging in 4 different levels and 3 different types.

- In 2016, the number of codes declined and has since stabilized, except for D-level codes which are trending upwards, nearing their 2015 peak.
- Nursing homes report the highest number of codes at level 2 (which includes D, E, F codes), with D (Isolated type) being the most frequent, occurring twice as often as E (Pattern type), the second most common.

Note: The definitions of scope levels and types can be viewed at https://anha.org/uploads/ScopeSeverity2018.pdf.

QUALITY MEASURES: Weighted Average Technique



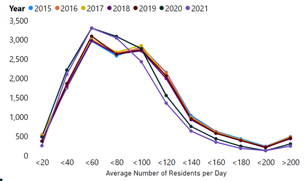
METHOD

From 2015 to 2021, quality is measured annually using Measure Codes (MCs) with varying weights. **The Weighted Average Technique** calculated weighted values for Q1, Q2, Q3, Q4, and the annual average.

- Every year, the highest weighted percentages (>10%) are attributed to the percentage of **long-stay** and **short-stay** accessed and appro-priately given the **seasonal influenza vaccine** and **pneumococcal vaccine** (MCs: 415, 411, 430, 426).

Note: Quality is measured from 2015 to 2021. The chart above shows the score in 2015.

PROVIDER INFO: Number of Residents per Day



DEFINITION

Average Number of Residents per Day comes from the MDS daily census. Due to variations in data collection, the data is grouped for clarity. For example, "<40" covers a range from 20 to just under 40.

- From 2020 to 2021, there was a reduction in the number of facilities where the average daily resident count was over 80, whereas facilities with an average daily resident count between 20 and 80 experienced growth. This shift is believed to be a result of COVID-19 restrictions.

KEY TAKEAWAY

The COVID-19 pandemic has impacted nursing homes, leading to a decrease in revenue. However, there has also been a reduction in costs and a sharp increase in income.

During the COVID-19 pandemic, the proportion of nursing homes with a negative operating margin rose from about 58% to approximately 72%. Meanwhile, about 81% of nursing homes with a positive operating margin reported margins of 10% or less.

Some Health Deficiency Scope Codes and Quality Measure Codes are disproportionately higher than others and should be closely examined in the operation of Nursing Homes

During the pandemic, the daily number of residents has decreased, showing a shift towards lower numbers (right skew).

COVID-19 IMPACTS

HOW HAS THE COVID-19 IMPACTED THE NURSING HOME OVERALL AND FINANCIAL PERFORMANCE?

OVERVIEW COVID-19

29.9%
National Percent of Residents Up to Date with Vaccines per Facility

2,022,998
Total Resident COVID-19 Confirmed Cases

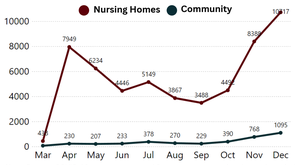
170,099
Total Resident COVID-19 Deaths

1,897,763
Total Staff COVID-19 Confirmed Cases

Note: Up-to-Date Data as of the week ending 04/14/2024 from data.cms.gov

2020 OVERALL PERFORMANCE: Nursing Home vs Community

Type	Count of COVID-19 Cases	Adjusted Cumulative Incidence	Average Monthly Incidence
Community	2377347	3875	388
Nursing Home	654583	54214	5421
Total	**3031930**	**58089**	**5809**



Note: Data on COVID-19 cases was sourced from cms.gov, covering the study period from March 2020 to December 2020.

- **Nursing home residents** account for about 2% of the Medicare population, but about 22% of all COVID-19 cases. Nursing home residents were 14 times more likely to be diagnosed with COVID-19 compared to beneficiaries in the community.
- **Nursing home** residents consistently faced a higher risk of COVID-19 diagnosis compared to those in the **community**, being at least 11 times more likely even when community spread was low.

NET INCOME PERFORMANCE: Before and After COVID

Variables	Before Covid	After Covid
Intercept	−0.73 ***	−0.43 **
OVERALL_RATING	100.27 ***	92.07 ***
AVG Number of residents per day (RESTOT)	2.00 ***	3.81 ***
Medicare and Medicaid/Medicaid (1,0)	266.98 ***	220.54 *
Rural/Urban (1,0)	114.09 ***	161.43 ***
Nonprofit	−210.29 ***	−452.84 ***

Note: $***p < 0.001$; $**p < 0.010$; $*p < 0.050$

Regression Output for Net Income in Millions Before and After COVID

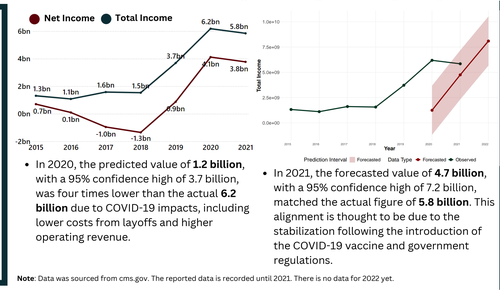
METHOD

A regression model was performed to see the change in Net Income performance in the period from 2015 to 2021, especially the change during the pandemic. From this model, we run the model 2 times, one of them it before the pandemic and one after the pandemic.

NOTE: The provided data contains over 150 variables. To select the appropriate variables for our model, we consider the correlation between each variable. The independent variables selected are not correlated with each other, but they do correlate with the response variable. 5 variables was applied to the model across two periods: Before Covid and After Covid.

- Initially, the model primarily focused on **Nonprofit** and **Certification (Medicare/Medicaid)**. However, the COVID-19 pandemic significantly elevated the importance of **Rural/Urban** and **Nonprofit** factors in managing case distribution. The National Library of Medicine highlighted that in 2020, states with higher rural infection and mortality rates became epicentric centers. Meanwhile, the influence of **Overall Rating** and **Certification** decreased slightly.
- Additionally, the model showed a shift in the **Intercept**, suggesting higher **Net Income** under the same conditions in the second model, indicating potential profit increases for Nursing Homes.

TOTAL INCOME FORECASTING: Autoregressive Integrated Moving Average Model



- In 2020, the predicted value of **1.2 billion**, with a 95% confidence high of 3.7 billion, was four times lower than the actual **6.2 billion** due to COVID-19 impacts, including lower costs from layoffs and higher operating revenue.
- In 2021, the forecasted value of **4.7 billion**, with a 95% confidence high of 7.2 billion, matched the actual figure of **5.8 billion**. This alignment is thought to be due to the stabilization following the introduction of the COVID-19 vaccine and government regulations.

Note: Data was sourced from cms.gov. The reported data is recorded until 2021. There is no data for 2022 yet.

METHOD

ARIMA is employed to forecast and estimate **Total Income** over time, utilizing both current and projected data, with its accuracy enhanced by adjusting for inflation. The data was used to predict is from 2015 to 2019. 2020 and 2021 is the predicted value which will be compared with the actual data.
Income, which includes **Net Income** and **Total Income**, is favored over **Revenue** due to the uncertainties or inconsistencies from Contractual Allowances and Discounts. Total Income is determined by summing Net Income and Total Other Income.

KEY TAKEAWAY

The mortality rate among residents is approximately 8.4%.

Nursing homes perform significantly poorer in handling COVID-19 compared to the community.

Net income has improved during the pandemic. Given similar conditions, a nursing home facility will experience an increase in net income.

Forecasting for nursing homes indicate a potential increase in total income, which is expected to continue growing strongly.

INVESTMENT SOLUTION

SHOULD ELDER LEE AND MATT CHUR INVEST IN NURSING HOME? IF SO, HOW AND WHERE?

STEP 1: Variables Chosen and Descriptive Statistics

Variable	2015	2016	2017	2018	2019	2020	2021
Number of Observations	13913	13969	14202	14218	14157	14139	14086
Total Income (in Ten Thousand $)							
Mean	9.51	7.95	11.38	11.01	26.30	43.77	41.51
SD	169.86	387.32	1188.16	1247.62	1295.22	1114.07	1208.75
Net Income (in Ten Thousand $)							
Mean	5.17	0.93	−6.43	−9.22	6.43	29.35	28.93
SD	159.53	186.36	313.68	173.60	217.17	257.26	326.42
Ownership							
Government	708 (5%)	750 (5%)	742 (5%)	725 (5%)	723 (5%)	687 (5%)	692 (5%)
Non-profit	2923 (21%)	2921 (21%)	2989 (21%)	2993 (21%)	3034 (21%)	3029 (21%)	2984 (21%)
Profit	10282 (71%)	10298 (71%)	10476 (71%)	10500 (71%)	10400 (71%)	10423 (71%)	10410 (71%)
Rural vs Urban							
Rural	3972 (29%)	3892 (28%)	3885 (27%)	3857 (27%)	3779 (27%)	3751 (27%)	3722 (26%)
Urban	9941 (71%)	10077 (72%)	10322 (73%)	10361 (73%)	10378 (73%)	10388 (73%)	10364 (74%)
Certification							
Medicare and Medicaid	13370 (96%)	13420 (96%)	13654 (96%)	13656 (96%)	13623 (96%)	13611 (96%)	13588 (96%)
Medicare	543 (4%)	549 (4%)	553 (4%)	562 (4%)	534 (4%)	528 (4%)	498 (4%)
Overall Rating							
*	2184	2367	1863	1842	2021	2078	2259
**	2819	2707	2707	2887	2768	2735	2835
***	2645	2578	2435	2357	2537	2521	2663
****	3210	3183	3115	3159	3082	3041	3064
*****	2947	3027	3681	3774	2990	3568	3056
No Rating	108	107	406	199	159	196	209
Avg Number of Beds	117.47	117.91	975.99	122.78	120.79	122.70	117.77
Avg Residents Per Day	89.80	88.75	88.18	88.41	88.26	79.22	76.15

Nursing Home Descriptive Statistics for 2015-2021

The provided data contains over 150 variables. To select the appropriate variables for our model, we consider the correlation between each variable. The independent variables selected are not correlated with each other, but they do correlate with the response variable.

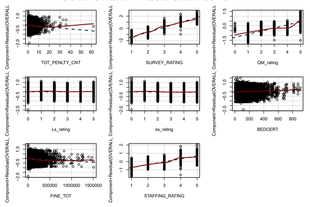
The two-stage least squares approach mitigates the influence of hidden biases, yielding more accurate insights into the determinants of financial and overall performance in nursing homes.

STEP 2: 2SLS Stage I - Assessing Overall Rating

Variable	1st Stage Model
Intercept***	4.24***
COVID(1,0)	-0.033***
Avg Residents Per Day	-0.003**
Nonprofit	0.59***
Medicare and Medicaid/Medicaid	0.99**
Rural vs Urban(1,0)	-0.10***
Avg Residents Per Day:Medicare and Medicaid/Medicaid(1,0)	0.001***
Observations	96991
F-Statistic	1165
p-value	$< 2e-16$***
R-squared	0.06725

1st Stage Model Estimate for Overall Rating

COMPONENT + RESIDUAL PLOTS



First Stage (Predicting Overall Rating with RU as IV)

$$\text{Overall\_Rating}_{it,t} = \beta_0 + \beta_1 \times \text{Covid}_{i,t}(1,0) + \beta_2 \times \text{Average Residents Per Day}_{i,t} + \beta_3 \times \text{Nonprofit}_{i,t} + \beta_4 \times \text{Medicare and Medicaid/Medicaid}_{i,t} + \beta_5 \times \text{Rural vs Urban}_{i,t} + \beta_6 \times (\text{Average Residents Per Day}_{i,t} \times \text{Medicare and Medicaid/Medicaid}_{i,t}) + \varepsilon_{i,t}$$

We're focusing on what influences a nursing home's **Overall Rating**. Some key factors here are:

- **COVID(1,0):** A coefficient of -0.033 suggests that an increased impact from COVID-19 lowers the Overall Rating.
- **Nonprofit:** Nonprofit status positively correlates with higher ratings in nursing homes, as indicated by a coefficient of +0.5884.
- **Rural vs Urban (RU):** As an instrumental variable, its negative coefficient of -0.1035 indicates that nursing homes in rural areas, or those with certain characteristics represented by this variable, generally receive lower ratings than urban ones.
- **RESTOT:** Representing total residents per day, the negative coefficient of -0.0039 indicates that higher numbers of residents are associated with a lower Overall Rating.

Note: *** p<0.001

The plots shows that higher **survey and staffing ratings** correlate with better nursing home quality perceptions, while the impact of **financial penalties** on ratings is less clear, suggesting they don't directly influence quality ratings. Additionally, the **number of certified beds** has little effect on ratings, indicating that a nursing home's size doesn't necessarily determine its quality. These insights are useful for evaluating nursing home performance and the factors affecting their ratings.

STEP 3: 2SLS Stage II and OLS Model: Predicting Net Income

Variable	OLS Model	IV Model
Intercept***	0.14***	0.15***
COVID(1,0)	0.32***	0.31***
Avg Residents Per Day	-0.02***	-0.02***
Nonprofit	-0.28**	-0.27***
Medicare and Medicaid/Medicaid	-0.65**	-0.59***
OVERALL_RATING	0.10***	0.10***
Avg Residents Per Day:Medicare and Medicaid/Medicaid(1,0)	0.02***	0.02***
Rural vs Urban(1,0)(Instrument Variable)	0.14***	
Observations	96990	96991
F-Statistic	179.8	200.4
p-value	$< 2e-16$***	$< 2e-16$***
R-squared	0.01821	0.01225

2nd Stage and OLS Model Estimates for Net Income (In Millions)

Second Stage (Main Model: Predicting Net Income)

$$\text{Net Income (in millions)}_{i,t} = \delta_0 + \delta_1 \times \text{Average Residents Per Day}_{i,t} + \delta_2 \times \text{Nonprofit}_{i,t} + \delta_3 \times \text{Covid}_{i,t}(1,0) + \delta_4 \times (\text{Average Residents Per Day}_{i,t} \times \text{Medicare and Medicaid/Medicaid}_{i,t}) + \delta_5 \times \text{Medicare and Medicaid/Medicaid}_{i,t} + \delta_6 \times \text{Overall\_Rating}_{i,t} + \mu_{i,t}$$

The initial prediction of the **Overall Rating** is incorporated into the main model to examine its impact on a nursing home's financial performance. It is hypothesized that a **higher Overall Rating**, indicative of better quality, will lead to **greater Net Income**.

In this analysis, the **Rural vs Urban classification** is used as an **instrumental variable**. This approach helps to isolate the effect of daily resident counts on quality ratings, minimizing the direct impact of location. This method strengthens our confidence that the observed relationships accurately represent the true influence of these variables on a nursing home's rating and its financial outcomes.

For the 2nd model, we still consider the same variables, in addition to Overall Rating, to have better insights in finding the solution for the investors.

- **COVID(1,0)**: The positive coefficient (+0.3122) suggests that COVID-19 is linked to higher scaled income, possibly due to additional funding received by nursing homes facing the pandemic.
- **Nonprofit**: A negative coefficient (-0.2729) shows that nonprofit status correlates with lower financial performance, likely due to different objectives compared to for-profits.
- **Overall Rating**: The positive coefficient (+0.0964) indicates that higher ratings correlate with better financial performance, suggesting that well-rated homes may attract more residents or funding.
- **RESTOT**: Again, similar to model 1, in model 2, the average resident per day with the negative coefficient of -0.018 indicates that higher numbers of residents are associated with a lower Net Income

SOLUTION RECOMMENDATION

Given the data from the **General Information** section and Keys Takeaway from **Financial Performance & Influential Factors** and **COVID-19 Impacts** sections, we have brought the following conclusion:

Elder Lee and Matt Chur should invest in Nursing Home

Recommendations:
- **Location:** Midwest and Northeast region, or California, Texas, and Florida
- **Certification**: Both Medicare and Medicaid
- **Location Type**: Urban
- **Size**: Concentrate on numerous medium-sized facilities (60-100 residents per day) rather than a few larger ones.

Considerations:
- Taking care of Health Deficiency Scope Codes (especially level 2) and Quality Measure Codes.
- Focusing on optimizing operational costs.
- Study on the effectiveness of community management in enhancing COVID-19 protocols.

Thank You

- **Dr. Raju Mainali, Dr. Xiang Guo,**
- **Department of Information Technology and Cybersecurity, College of Business, Missouri State University**

The Nursing Home Project

began in February 2024, and was funded by Department of Information Technology and Cybersecurity (MSU), is focused on analyzing U.S. nursing home market to deliver investment solutions to investors.