Mod 8 – Capstone Project: Healthpeak Properties, Inc.

Kristine N. Plunkett

Colorado State University – Global Campus

MIS581: Capstone – Business Intelligence and Data Analytics

Dr. Steve Chung

September 6, 2020

Mod 8 – Capstone Project: Healthpeak Properties, Inc.

Abstract

Healthpeak Properties, Inc., a real estate investment trust operating within the U.S., specializes in the diversification of investments between ownership and development of healthcare real estate properties. This specific organization was chosen due to having interviewed with the company in the recent past. Healthpeak Properties has been in business since 1985, currently operates with 204 employees, and is headquartered in Irvine, CA. The aim of this capstone research project is to identify and highlight the specific regions of the organization that will have the most potential for high impact relationships with deaths related to the SARS-CoV-2 pandemic. The goal of this research project is to illuminate these top regions within the senior housing sector for Healthpeak Properties, Inc. so that the business can prove better equipped to strategize for future growth and recovery. Thorough data analysis and testing within this study appear to indicate a statistically significant relationship between Healthpeak Properties units and Covid-19 deaths in the 65+ population.

Introduction

Healthpeak Properties, Inc. operates in life science and medical office sectors, however also maintains a large portfolio in the U.S. senior housing sector. March 2020 quarterly reports indicate that (Healthpeak Properties, Inc., 2019) the organization has an astounding liquidity of $3 billion, $2.5 billion in revolving credit, and roughly $500 million in cash or cash equivalents. The organizations 2019 annual report indicates the total portfolio income as $297 million with total pro-forma portfolio income as $289 million. There are potential outcomes that this research can highlight, to the benefit of Healthpeak Properties, which will enable for the planning and strategy necessary for this U.S. organization to pivot and recover from the inevitable economic

impact this pandemic virus has and will continue to impose on the senior community for the foreseeable future. Studies indicate that older age groups are most at risk of death from Covid-19, which will allow for potential relationships to be witnessed between those pandemic related deaths and the regions where the organization operates.

Research Hypothesis

The formulated primary hypothesis aims to understand the U.S. markets that could have a strong relationship with the high impact areas stemming from Covid-19 deaths within the senior housing sector of Healthpeak Properties, Inc. The null hypothesis (H0) is adequately stated as: the regions that encompass the highest concentrations of Healthpeak Properties senior housing units will prove to have a relationship with the regions of highest novel coronavirus deaths. Conversely, the alternative hypothesis (H1) can be most appropriately stated as: the regions that encompass the highest concentrations of Healthpeak Properties senior housing units will not prove to have a relationship to the regions of highest novel coronavirus deaths.

Objectives

The objective of this research project is to facilitate the ability to identify the high risk senior market housing regions that will be most probable in observing high impact from deaths due to the novel coronavirus pandemic. There are a number of publicly available datasets, specifically from Healthpeak Properties and the U.S. Center for Disease Control, that have been accessed and retrieved to aid in the analysis and testing of the hypothesis. These datasets encompass predominantly quantitative variables, however some categorical variables do exist. The quantitative nature of the datasets provide numeric variables gathered specifically in reference to those affected by the novel coronavirus pandemic. The data represent case counts and deaths by age and region, the information derived from the United States healthcare facilities

and government agencies. According to (O'Leary, 2017) the quantitative traditions are most beneficial to aid in the ability to become intimately familiar with the scientific/hypothetico-deductive methods in research allowing for the investigation and experimentation in population studies. Qualitative research methodologies have a great many benefits, however for this capstone research project the quantitative methodology will be most appropriate.

Overview of Study

The risk of death from the SARS-CoV-2 virus is heightened among the aging population. According to (CDC, 2020) 80% of Covid-19 deaths are reported from the 65 and older population. The objective of this research project is to highlight the relationship between the regions where Healthpeak Properties has concentrations of senior housing, and the total deaths from Covid-19 in those same regions that are taking place heavily within the aging population, the demographic of tenants prevalent at those properties.

Literature Review

SARS-CoV-2 has only been in existence for roughly eight months, however a body of literature is already accumulating as research into the pandemic virus continues. Studies continue to overwhelmingly suggest that age and prevalence of comorbidities are the leading factors in death from COVID-19. The earliest cases studied in Wuhan, China were all reportedly older patients with comorbidities (Jordan, Adab, and Cheng, 2020). The Cruise ship outbreak in early February 2020, The Diamond Princess, experienced 14 deaths all aged 65 and older (Russel, et. all, 2020). The CDC morbidity and mortality weekly report from March 2020 illustrated 80% of the Covid-19 deaths from February 2020 to March 2020 were patients aged 65 years and older (Covid, C.D.C., & Team, R., 2020). Data studies in July 2020 also have illustrated that more than 50% of Covid deaths in the U.S. are from the aging population, while data from Italy

communicates that those 70 and over make up 85% of deaths (Brannen, et. al., 2020). The body of literature devastatingly communicates that age is a leading factor in death from the Covid-19 disease leading to support the hypothesis that Covid-19 deaths in the U.S. will have a relationship with the senior housing sector of Healthpeak Properties, Inc.

Research Design

Methodology

Data has been derived from viral testing and Covid-19 death confirmations that have been provided to government agencies from healthcare establishments around the country. The CDC dataset adequately communicates this data, albeit within a limited date range and some noted discrepancies. The quantitative nature of the datasets that have been obtained provide numeric variables gathered specifically in reference to those affected by the novel coronavirus pandemic. These datasets represent cumulative case counts and deaths by age and region, the information derived from the United States healthcare facilities and government agencies. According to (O'Leary, 2017) the quantitative traditions are most beneficial to aid in the ability to become intimately familiar with the scientific/hypothetico-deductive methods in research allowing for the investigation in population studies.

Methods and Tools

The specific methods of testing will include descriptive analytics and statistical hypothesis testing, in addition to some additional data visualizations to support and present the statistical findings. Regression analysis testing will be used to test the hypothesis in this study, which will allow for an estimation of the relationships witnessed between the dependent and independent variables available within the dataset. In addition, the descriptive nature of this research project aims to establish associations between the variables contained in these publicly

available datasets in order to provide better understanding of the data. R programming, SAS, and Tableau will be the primary data analytics tools used, as well as MS Excel and SQL sparingly.

Limitations

One area of limitation to note stems from lack of a solid body of research regarding pandemics and the effects they may have on real estate investment trusts. The last large pandemic outbreak took place in 1918-1919, real estate investment trusts not having much if any prevalence at that time. Data accuracy may also prove to be an unfortunate limitation within this study due in part to the U.S. Center for Disease Control and other governmental agencies realizing data discrepancies or reporting data in a different manner. Recently, California realized a glitch in reporting that according to (Colliver, 2020) may have under reported coronavirus cases by up to 300,000 due to a server outage. Regardless, it will still be possible to test the theories within this project and identify if a relationship does exist between a regions senior housing units and novel coronavirus total deaths.

Ethical Concerns

Ethics, when pertaining to research projects include a number of factors that according to (Lund Research, 2012) fall into five categories that include minimizing risk, obtaining informed consent by research participants, addressing and protecting confidentiality and anonymity, always avoiding deceptive practices, and allowing for participants to have the ability to withdraw from the research study. Maintaining the highest level of ethical standards will always remain of strict focus for this research project. Also, the researcher holds strong principles regarding the use of data and knowledge in data analytics and research activities. In addition all efforts have been made to utilize public datasets that will not compromise the integrity of personally identifiable information or privacy.

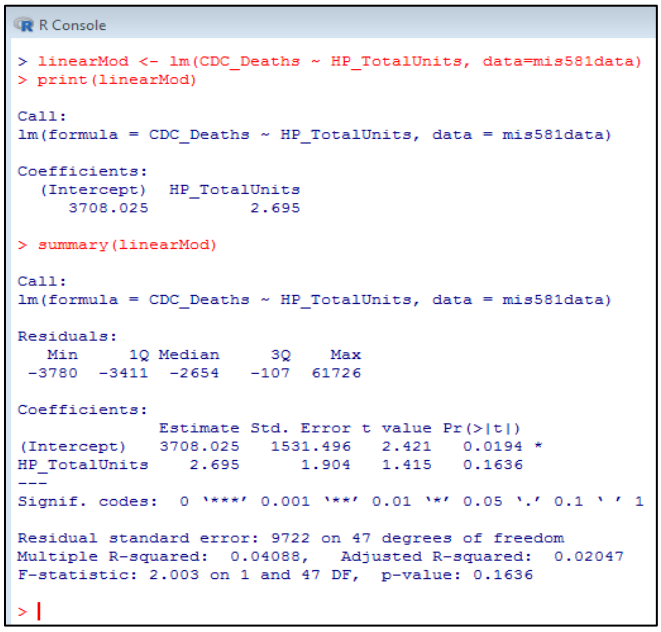
Data Analysis Results

R Programming

Linear regression hypothesis testing has taken place using R programming. Given the output results, it is safe to assume that there is substantial support for the null hypothesis. The linear regression was run against the standard variables from the dataset, units and deaths, which resulted in a p-value of 0.1636. The large p-value, in comparison to a standard alpha significance level of 0.05, indicates there is a lack of strong evidence to reject the null hypothesis. The residuals distribution indicates that outlier data can be expected within the dataset, and the intercept estimate shows that for every 2.695 units we can expect an intercept of 3708.025 deaths. A residual standard error indicates that deaths can deviate by 9722 with 47 degrees of freedom.

Figure 1

Linear Modelling Code and Output Results



Note. The above screenshot depicts the linear regression modeling code and output for the CDC_Deaths and HP_TotalUnits variables.

In order to find further evidence of a significant relationship between units and deaths a second linear regression test was performed. This test was performed against a dataset where SQL was used to transform the counts of deaths and units into binary variables. Interestingly, the linear regression output results from the dataset having binary variables illustrated an even greater relationship between Healthpeak units and Covid-19 deaths. This linear regression was run against the binary variables from the dataset and resulted in a larger p-value of 0.292 in comparison to the same standard alpha significance level of 0.05, indicating even stronger supporting evidence for the null hypothesis.

Figure 2

Linear Modelling Code and Output Results – Binary Variables

```
> linearMod <- lm(binCDC_Deaths ~ binHP_TotalUnits, data=mis581databin)
> print(linearMod)

Call:
lm(formula = binCDC_Deaths ~ binHP_TotalUnits, data = mis581databin)

Coefficients:
     (Intercept)  binHP_TotalUnits
         0.95652           0.04348

> summary(linearMod)

Call:
lm(formula = binCDC_Deaths ~ binHP_TotalUnits, data = mis581databin)

Residuals:
     Min       1Q   Median       3Q      Max
-0.95652  0.00000  0.00000  0.04348  0.04348

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.95652    0.02975  32.156   <2e-16 ***
binHP_TotalUnits 0.04348    0.04084   1.065    0.292
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1427 on 47 degrees of freedom
Multiple R-squared:  0.02355,   Adjusted R-squared:  0.002775
F-statistic: 1.134 on 1 and 47 DF,  p-value: 0.2925

>
```

Note. The above screenshot depicts the linear regression modeling code and output for the binCDC_Deaths and binHP_TotalUnits binary variables.

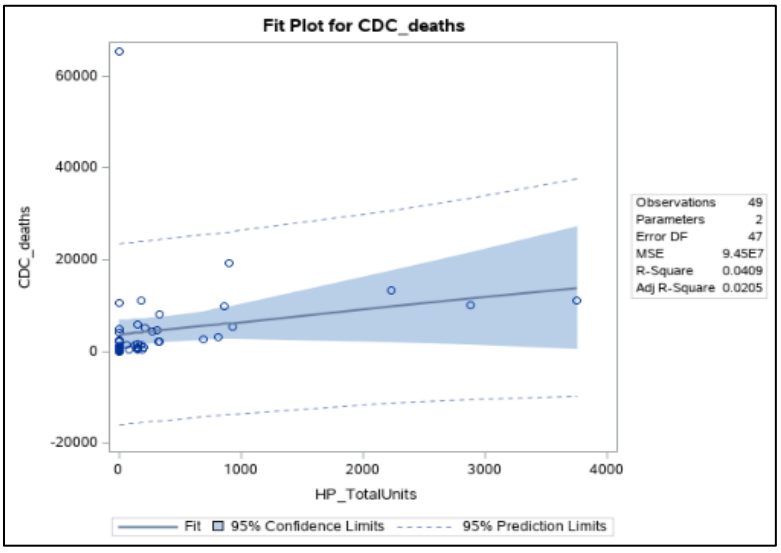
SAS - Regression Analysis

The SAS regression procedure output results match what was uncovered using R programming, indicating that the p-value is greater than the small significance level of 0.05, or

Pr > F = 0.1636 with alpha set to 0.05. This again illustrates a positive statistical relationship granting evidence that the null hypothesis can be supported. To further support the statistical signifance of the relationship, exploring the regression fit plot can also highlight the relationship. The below fit plot suggests the linear relationship to be increasing gradually, however it is also highlighting the outliers in the data. These outliers represent high concentrations of deaths or units in specific regions and will be further analyzed.

Figure 3

Fit Plot for HP_Total Units



Note. The Fit Plot output results for HP_TotalUnits and CDC_Deaths created using the SAS regression procedure.

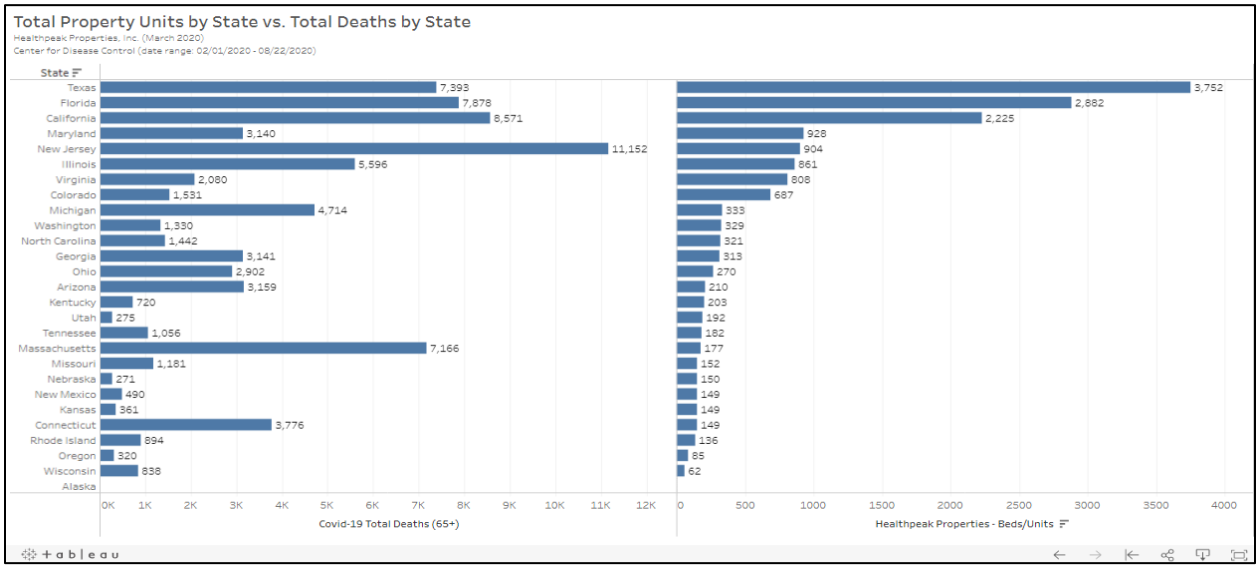
Tableau Data Analysis and Visualization

Data visualization provides an image to highlight the states with highest Covid-19 deaths in the 65+ population in addition to the states with highest concentrations of units that belong to Healthpeak Properties, Inc. There are discrepancies witnessed where we have high concentrations of Covid-19 deaths among the 65+ populations in states where there are few Healthpeak units and properties. Alternatively, we also can witness some states where there are

high levels of properties and units for Healthpeak, yet lower incidents of deaths from Covid-19. Although there are some irregularities these results further support the regression analysis outcome leading to support of the null hypothesis. As a result, it appears that the regions that encompass the highest concentrations of Healthpeak Properties senior housing units have a statistically significant relationship to the regions of highest novel coronavirus deaths.

Figure 4

Total Property Units by State vs. Total Deaths by State



Note: Tableau visualization illustrating a comparison of values for Healthpeak Properties units and Covid-19 65+ deaths per state.

Findings

Statistical testing has provided plentiful evidence indicating that there is a strong relationship between Healthpeak Properties units and Covid-19 deaths. It would seem entirely reasonable that properties within these states will experience the highest impact stemming from coronavirus deaths. Texas, California, and Florida should be regions of specific interest and concern as they boast the highest potential for impact, having the greatest amount of Covid-19 deaths and Healthpeak Properties units.

Conclusion

In conclusion, there is a large body of research and literature regarding the SARS-CoV-2 disease, highlighting the fact that seniors are at highest risk of death from the virus. Independent studies from China, Japan, Italy, and the United States all overwhelmingly indicate this to be true. It stands to good reason to believe that the senior housing sector of Healthpeak Properties is uniquely situated to feel an economic impact from these events being that they have an interest in the senior housing sector. Thorough analysis and testing was clear in statistically proving a relationship and supports the null hypothesis: the regions that encompass the highest concentrations of Healthpeak Properties senior housing units will be related to the regions of highest novel coronavirus deaths.

Recommendations

Recommendations for Healthpeak Properties includes addressing the top regions that are proving to be potential areas of high impact to business stemming from novel coronavirus deaths. Those regions include but are not limited to California, Texas, and Florida. Real-time data from each property regarding status of outbreak clusters and deaths will prove incredibly useful in further tracking and analysis. Ultimately, one of the best recommendations that can be made for Healthpeak Properties, Inc. is to invest in preventative measures within each of their locations to protect the senior housing tenants from contracting the SARS-CoV-2 pandemic virus. All SAS and R code utilized for this project can be accessed within GitHub at the following URL: https://github.com/kplunkett13/MIS581-Capstone

References

Brannen, S., Ahmed, H., & Newton, H. (2020). Covid-19 reshapes the future. *Center for Strategic and International Studies (CSIS)*. Doi: 10.2307/resrep25198

CDC. (2020). Coronavirus disease 2019 (COVID-19): older adults. Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/older-adults.html

Colliver, V. (2020). California sorts through coronavirus data as glitch stymies policy decisions. Retrieved from https://www.politico.com/states/california/story/2020/08/07/california-sorts-through-coronavirus-data-as-glitch-stymies-policy-decisions-1305972

Covid, C. D. C., & Team, R. (2020). Severe outcomes among patients with coronavirus disease 2019 (COVID-19)—United States, February 12–March 16, 2020. *MMWR Morb Mortal Wkly Rep, 69*(12), 343-346.

Healthpeak Properties. (2019). *Healthpeak Properties: About us*. Retrieved from https://www.healthpeak.com/about-us/

Jordan, R. E., Adab, P., & Cheng, K. K. (2020). Covid-19: risk factors for severe disease and death. *BMJ 2020*;368:m1198 doi: 10.1136/bmj.m1198

Lund Research. (2012). Principles of research ethics. Retrieved from http://dissertation.laerd.com/principles-of-research-ethics.php

O'Leary, Z. (2017). The essential guide to doing your research project. (3rd ed.). Thousand Oaks, CA: SAGE Publishing.

Russell, W., Hellewell, J., van Zandvoort, K., Abbott, S, Ratnayake, R., Flasche, S., Eggo, R., Edmunds, J., Kucharski, A. (2020). Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the

Diamond Princess Cruise ship. *Euro Surveill. 2020;25*(12):pii=2000256.

https://doi.org/10.2807/1560-7917.ES.2020.25.12.2000256