Term Project: A Simulation Model of Occupancy Rates in a Nursing Home

Jaiden Angeles, David Eng

Beedie School of Business, Simon Fraser University

BUS 440: Simulation in Management Decision-Making

Prof. Zhe George Zhang

December 16, 2023

A Simulation Model of Occupancy Rates in a Nursing Home

Jaiden Angeles, David Eng
SFU Beedie School of Business, Dr. Zhe George Zhang of Management Sciences;

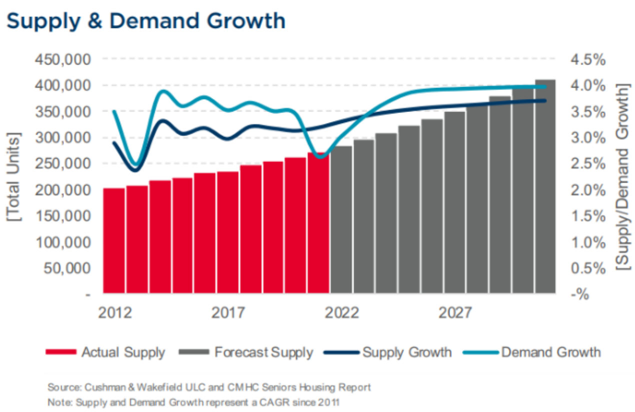
This article presents a comprehensive study conducted in a single senior living facility, focused on developing a predictive model to estimate the move-out times of residents. The primary objective of the study was to optimize room occupancy by minimizing the duration that rooms remain vacant after a resident moves out. This is crucial for operational efficiency and decreasing wait time for waitlisted residents. The study employed a simulation model using Excel, incorporating lognormal distribution to simulate resident move-out patterns and align simulated data with existing sample distributions, providing a foundation for the predictive model. This involved generating random numbers using Excel's RAND function, with the creation of 1000 random move-out times, followed by a detailed analysis of their frequency distribution using Excel's FREQUENCY function. This was paired with an Arena simulation to analyze scenarios at 85.5% and 100% bed occupancy, revealing the relationships between bed availability, room turnover rates, and overall wait times simulated for 10,000 days. The findings of the study offer significant implications for senior living facility management, particularly in planning and resource allocation. By accurately predicting move-out times, facilities can better prepare for incoming residents, thereby enhancing operational efficiency and resident satisfaction. The article concludes with a discussion of the model's potential applications in other senior care settings and recommendations for future research to refine and expand upon the predictive capabilities of the model.

Keywords: nursing home; operational efficiency; simulation

1. Introduction

In Canada, we are currently seeing a trend with a growing population of retirees within the country. According to Statistics Canada, the number of people 65 and older increased by 18.3% to 7 million from 2016 to 2021. Similarly, the number of people aged 85 and older has doubled since 2011 with 861,000 as of 2021 (Government of Canada, 2022). The senior living sector is currently in high demand and is expected to grow along with the ever-increasing number of seniors as seen in **Figure 1**. In 2021, 238,000 people older than the age of 85 live in senior residences and other long-term care facilities (Government of Canada, 2022).

In order to cater to the influx of seniors and maximize their own profits, senior living residences must maximize their available capacity by reducing the down time of rooms. Currently, [redacted]'s capacity is 100% with the goal of maintaining this value in order to help maximize profits. As of 2021, the vacancy rates in seniors' residence are on the rise with all provinces except for Newfoundland and Labrador growing by 7%. Some notable regions are Alberta up to 26.8%, Saskatchewan with 22.4%, and Ontario with 19.6%. Manitoba saw an increase to 8.7% which was the fourth lowest for 2021, with Newfoundland and Labrador being the only region which saw a decrease in vacancy by 17.5% (CMHC, 2021).



The problem we are trying to address is the minimizing of days that a room within [redacted]'s

Figure 1 – Supply & Demand Growth for Retirement Home Units. Source: Cushman & Wakefield Report by McCrorie et al (2021) and CMHC Senior Housing Report (2021).

facility will sit vacant. This will allow for shorter wait times for residents, as well as profit maximization for [redacted], as rooms are always occupied. Currently the smallest room at [redacted] is an independent living studio at around $3,000 per month with the most expensive option being $5,000 per month ([redacted], 2023). An issue that can cause rooms to sit empty is the time required to find another tenant after one has left. The goal of this study is to create a predictive simulation model which can help any manager predict when the next move out period will be and to reach out to prospective tenants ahead of time to create a smooth transition from move-out of existing tenant to move-in of a new one.

Applying Discrete Event Simulation-based Optimization to evaluate retirement home data in modeling the Length of Stay has been utilised in previous studies before to capture the trend of Nursing Home data. However, not many have done this, and it is a field seldom explored, with general hospitals and other types of medical clinics being researched more intensely.

For instance, Griffiths et al (2010) and Ridge et al (1998), among others, explore capacity planning and bed occupancy optimization in ICUs rather than retirement homes.

Another instance is a 2004 study, where Faddy et al modelled the Length of Stay in a Hospital and Other Right Skewed Data -- utilizing Phase-type, Gamma, and Log-normal -- as these distributions seemed to fit best with Hospital data. Although different from the environment of a retirement home, they found that the log-normal fit was a much better fit to the gamma model, but slightly inferior to the phase-type model. They go on to state that problems arise from models with poorer fits which lead to extreme uncertainty, and this paired with assumptions drastically impact these models, so fitting the distribution alongside regression is critical in right-skewed data.

In 1999, Jun et al attempted to lower wait times for potential patients, reduce the Length of Stay, and streamline patient throughput. They did this by using discrete-event simulation modeling on health care clinics and complex healthcare systems like hospitals, outpatient clinics, emergency departments, and pharmacies. Through a wide range of sources, they demonstrated the ability to plan new policies based on their findings to facilitate this.

A study in 2019 by Zhang et al tries to address the limitation that lognormal, Weibull, and gamma distributions do not consistently fit length of stay data for hospitals, so they manipulate and mix distribution models, with lognormal distribution variations being the most used. However, this is associated with various weaknesses like inability to work properly on new data, computational complexity, and parameter estimation difficulties.

Another study in 2012 by Zhang et al attempted to refine the methodologies used for setting long-term care capacity in retirement homes. They integrated demographic and survival analysis, discrete event simulation, and optimization to develop a decision support system. This was applied in case studies in British Columbia, Canada, and at a long-term care facility. Their method was compared to traditional approaches like the fixed ratio and methods from call center literature, showing its superiority in handling long service times and providing policy recommendations.

The most related literature was by Bae et al (2009), where the paper explored modeling patient flow to optimize capacity for nursing homes, analyzing the United States area of Kentucky to better serve retirement home residents. Through this, they fitted Lognormal, Weibull, and Gamma distributions to the data to find predictions of wait times, utilisations, and quality.

The focus of this study is to employ a lognormal distribution model to analyze the move-in and move-out periods of residents at [redacted], with the aim of optimizing these transition periods to maximize bed occupancy rates. The primary goal is to develop a predictive model based on existing resident data that can accurately forecast the most probable timeframes for resident departures.

[redacted] currently has a capacity of 118 beds, of which we will assume 101 are occupied, resulting in an 85.5% utilization rate mirroring the industry average (See **Figure 2**). Currently, there is a waiting list comprising 12 individuals awaiting admission into the facility. This highlights the potential for improving bed occupancy rates through more effective management of move-in and move-out schedules.

A critical aspect to consider in this study is the mandatory unoccupied period between resident transitions. When a resident vacates a room, a specific time frame is required for the facility staff to clean and prepare the space for the incoming residents. This is crucial for maintaining the facility's standards for new residents. However, this necessary process introduces a variable factor into our model, as the time a room remains unoccupied can vary significantly depending on the reasons for a resident's departure. For instance, the time needed to prepare a room after a resident's death might be longer than the time required following a voluntary departure.

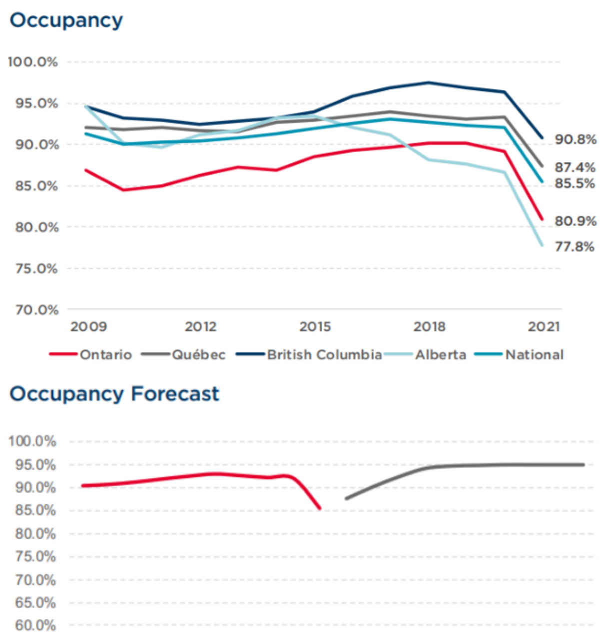


Figure 2 - Canada Occupancy Rates in Retirement Homes. Source: Cushman & Wakefield Report by McCrorie et al (2021) and CMHC Senior Housing Report (2021).

This variability poses a challenge in accurately predicting room turnover times and, consequently, impacts the overall efficiency in managing bed occupancy. Some goals are to improve the utilization rate beyond the current 85.5%, which has the side effect of also reducing wait times for prospective residents, ultimately benefiting both the facility and its clientele.

1.1 Long-term Care Infrastructure and Analysis

Currently, Canada has *3,651* long-term care homes according to the annual CMHC Senior Housing Survey in 2022. Through this, 319,925 units are offered. However, finding the number of medium-sized retirement homes is limited due to only 1,401 facilities submitting data to the Continuing Care Retirement Community (CCRS) or Integrated Retirement and Residential Services (IRRS) out of the 3,651 facilities in Canada (Canadian Institute for Health Information, 2023). Thus, in order to find the **optimal minimum occupancy rate**, we will analyze the United States because they have a more extensive database on the amount of retirement homes due to the reporting of U.S. Centers for Medicare and Medicaid Services on continuing care services.

First, we compare metrics of the Canada and the US's aging populations to ensure they share similar demographic challenges. The US has a median age of 38.9, while Canada's is 41, which are both approximately 30% higher than the global median age at 30 years old during 2022 (United States Census Bureau, 2023; Statista Research Department, 2023; Richie et al, 2019).

Moreover, the U.S. Percentage of population over 65 steadily increased up to 17.3% in 2022 and is forecasted at 20.6% in 2030, while Canada's increased to 19.0% in 2022 and is forecasted to reach 22.5% in 2030 (Statista Research Department, 2023; Eisen, 2022). Looking at **Figure 3**, the population aged 65 and above dramatically increased for both countries, reaching 17% in the US and 19% in Canada (The World Bank Group, 2022a).

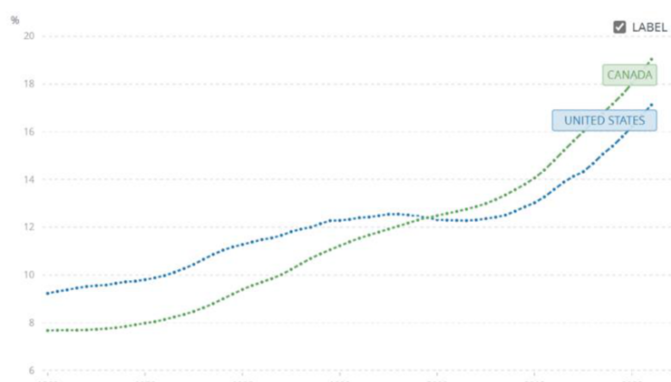


Figure 3 - Population ages 65 and above (% of total population) - United States, Canada. Source: The World Bank Group (2023a)

Finally, the US's and Canada's Age dependency ratio, which is the ratio of dependants (those under 15 or over 64) compared to the working population (15-64) has grown steadily higher over the past 10 years and has reached 54% and 53% respectively as seen in **Figure 4** (The World Bank Group, 2022b). Thus, the U.S. is an appropriate country to analyze their data to extract key information on current trends across North America as a whole.

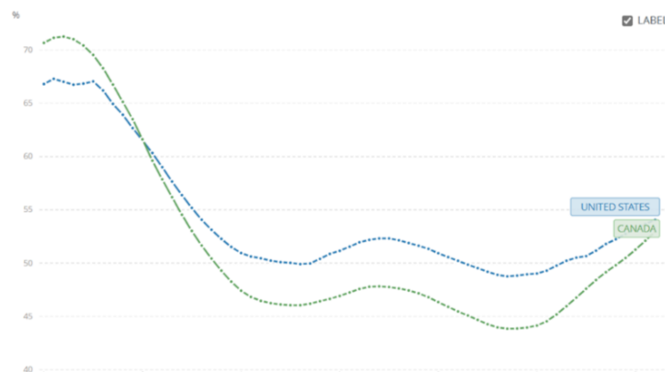


Figure 4 - Canada vs U.S. Age Dependency Ratio (% of working-age population) 1960 – 2022. Source: The World Bank Group (2023b)

According to our analysis conducted using data from the "U.S. Centers for Medicare & Medicaid Services" from October 2023, there are 7,714 medium-sized continuing care and nursing home facilities (having 26 to 99 certified beds) out of a total of 14,495 facilities in the US. This database was critical to assess various performance metrics critical to optimizing occupancy rates and cost-effectiveness in nursing homes, which we examined using Python programming with

Jupyter Notebook to extract insights. Although Altria is considered large at 118 beds, it barely passes the threshold of 99 suites, and we will classify it as medium-sized in this paper.

The evaluation of the top 100 medium-sized residency care centers, specifically focusing on the lowest "Percentage of short-stay residents who were re-hospitalized after a nursing home admission" and lowest "Number of hospitalizations per 1000 long-stay resident days", revealed an occupancy of **94.65%.** The average percentage of short-stay residents who were re-hospitalized in these top-performing facilities was found to be 6.35%, whereas the overall average for this metric stood at 22.45%. Similarly, the average number of hospitalizations per 1000 long-stay resident days for these top facilities was 0.28, in contrast to the overall average of 1.77. These findings highlight the correlation of targeted strategies in maximizing capacity utilization with maintaining high care standards in continuing care facilities. Thus, we will aim for an optimization of at least **94.65%** as facilities with this value show indications of efficient operational management and superior patient care, serving as a benchmark for [redacted]'s operational goals.

1.2 Nursing Home Operations

The focus of this study will be on the independent living and assisted living sectors of [redacted]. For the purpose of this study, we assume the operational time are primarily driven by room maintenance and utilities, aligned with industry standards.

We assume a standard occupancy rate of rooms based on industry averages. A study by the Cushman & Wakefield Senior Housing Industry Overview of Canada (2021) suggests that the average occupancy in senior living residences is approximately **85.5%,** primarily due to health-related move-outs or COVID-19 and it is anticipated to increase (See **Figure 4**).

Our study aims to refine these operational costs by improving room turnover efficiency and maintaining high occupancy rates. By optimizing these aspects, we anticipate a potential reduction in annual operational costs by **5-10%**, enhancing [redacted]'s profitability while ensuring the delivery of high-quality care to its residents.

1.3 Data

[redacted] in [redacted], part of [redacted], is a key player in the Canadian senior living sector. This award-winning facility, with a [redacted], has 118 beds with an occupancy rate of 100% per day (which as mentioned we will assume to be 85.5%) and a waiting list of 12 people. It is part of a network of [redacted] retirement communities across [redacted] provinces, making [redacted] the [redacted] largest retirement home company in

Canada ([redacted]; Mordor Intelligence, 2023).

Data collected with move-out rates from December 22, 2022, to December 2, 2023 – approximately 2 years – offers insights into occupancy trends and resident demographics on [redacted] location in [redacted]. The dataset includes information on resident identifiers, gender, move-in and move-out dates, rates per room, room assignments, payor type, service and care levels, and move-out reasons ([redacted]). This analysis aims to optimize room vacancy management and streamline resident transitions to enhance occupancy rates. The initial dataset contained over **200 entries** which included residents that are currently still staying at [redacted]'s facilities. As we are only interested in the residents that have moved out, we removed any data related to residents currently residing at [redacted]. This left us with the dataset used in this analysis containing the data for **131 anonymous patients**, all of which have moved in after October 27, 2013, and moved out due to a variety of common reasons.

2. Methods

2.1 Summary of the Model & Parameter Estimation

We decided to build a simulation model using Microsoft Excel. Excel was chosen for a few reasons. Firstly, since our dataset from [redacted] was provided to us in Excel file format, doing the simulation with Excel meant that we did not need to move our dataset around. Excel also provides us with easy data visualization tools and built-in functions for manipulating data such as graphs, charts, data tables and random number generation. Finally, we chose Excel so that we could easily share our findings back to [redacted] as Excel is a familiar tool to the management team. Starting with the sample data that was provided, we found 131 data points which represented individuals moving out of the facility. Given the move-in and move-out date, we were able to calculate the length of stay. From the length of stay data given, we were able to find the sample mean and standard deviation using the built in Excel functions MEAN and STDEV which were **830.94 days** and **881.692 days respectively**. From this we also calculated the Min and Max values along with dividing our data into 42 bins with a width of 76. After graphing our distribution on a histogram, we can see that the data we have is rightly skewed.

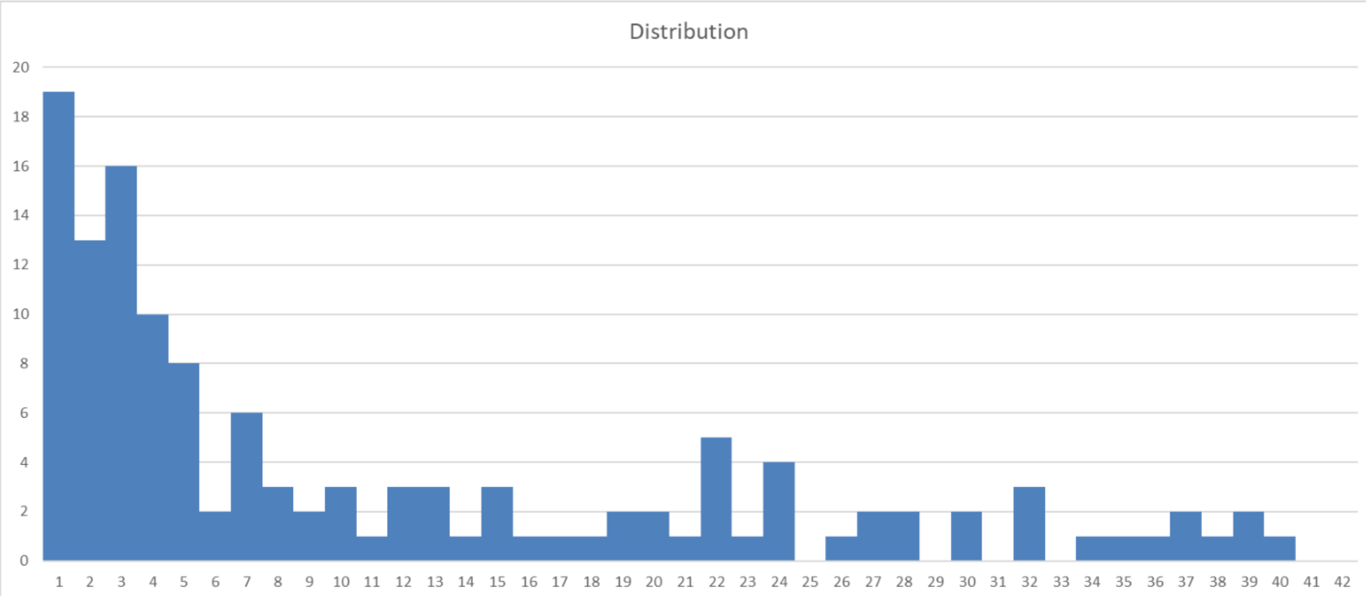


Figure 5 - [redacted] 2021-2022 Length of Stay Distribution.

As we can see, there are only a handful of distributions that follow the trend in **Figure 5**. This estimation is useful for minimizing the duration a room remains unoccupied, therefore optimizing occupancy rates and enhancing revenue. To achieve this, we concluded to use a simulation model that uses a lognormal distribution after conducting a Kolmogorov-Smirnov Test (Calculated in Section 2.2) in combination with the visual aspect of **Figure 5**. Our approach involved the creation of 1000 replications of a random number, each uniformly distributed between 0 and 1. This was accomplished using Excel's RAND function, a tool that generates random numbers within the specified range. The RAND function plays an important role in the simulation, as it serves as the basis for our predictive model. To do a lognormal distribution, we first had to create a new column in Excel with would calculate the normal log of our length of stay data. Afterwards, we calculated the mean and standard deviation which was approximately **5.97** and **1.43** respectively. To do the lognormal distribution we used the formula LOGNORM.INV with the random numbers previously generated. We started at 100 replications and ended up doing 1000 replications. After generating 1000 random move-out times, we proceeded to analyze the frequency distribution of these times. Utilizing Excel's FREQUENCY function, we were able to categorize these times into 42 distinct bins. The resulting histogram, as shown below, provides a visual representation of this frequency distribution. Notably, the simulated distribution closely follows the general shape of the sample distribution.

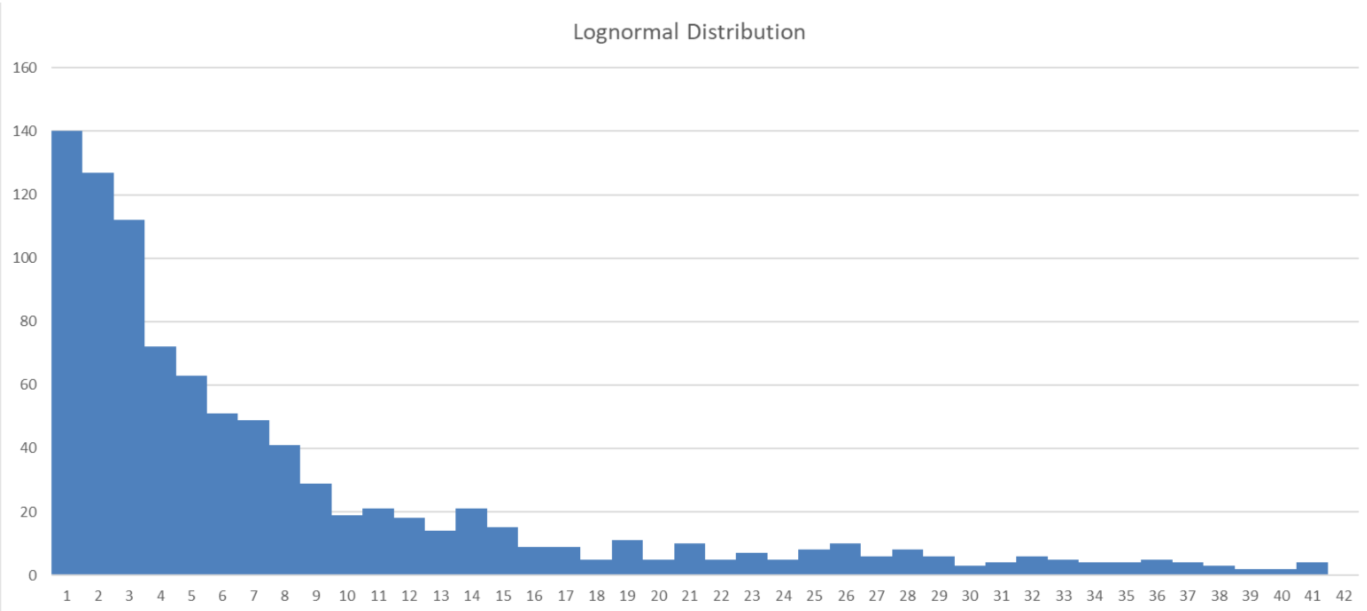


Figure 6 - Simulated Lognormal Distribution [redacted] *Length of Stay Distribution.*

As noticeable in **Figure 6**, this resemblance is a strong indicator of the model's reliability and its potential effectiveness in predicting resident move-out times for [redacted]. The outputs of this Excel simulation can then be used for further analysis within Arena to demonstrate the turnover time of the waiting line inventory.

2.2 Distribution Analysis for "Length of Stay"

In the pursuit of optimizing wait times for residents on the waiting list of a retirement home, our analysis first involved the application of the Chi-Square test to identify the best-suited distribution for the "Length of Stay" data. However, this approach encountered limitations due to the small dataset size of 131 records and significant right-skewness in the data. The Chi-Square test's dependency on adequate expected frequencies in each bin presented a challenge, as several bins had expected frequencies below 5, leading to inaccurate and inconclusive results. This limitation highlighted the need for a method suited to our data characteristics.

Subsequently, we turned to the Kolmogorov-Smirnov (K-S) test for a goodness-of-fit assessment. The advantages of the K-S test in our context were numerous: it is non-parametric and distribution-free, making no assumptions about the data's distribution. This is particularly beneficial for our dataset, which did not adhere to common distributions like normal or exponential. Additionally, the K-S test's sensitivity to differences in both the location and shape of empirical cumulative distribution functions made it an excellent tool for evaluating goodness of fit, especially for the "Length of Stay" data which is inherently continuous. Another key advantage of the K-S test is its ability to work directly with raw data, eliminating the need for data binning, thus avoiding potential information loss. This feature was crucial, given the challenges we faced with the Chi-square test's binning requirements. Furthermore, the K-S test's effectiveness even with small sample sizes and its straightforward interpretability (through the maximum distance between the cumulative distribution functions) made it an ideal choice for our study.

We tested six different types of appropriate distributions using Python within a Jupyter Notebook environment based on the histogram distribution see in **Figure 5**: Exponential, Log-Normal, Weibull, Gamma, Pareto, and Gumbel. Furthermore, we constructed a table, as seen in **Table 1**, showing all the crucial statistical information for the distribution types we tested. Among these, the Log-Normal distribution emerged as the best fit for our data, as indicated by its K-S Statistic (0.08434325015587263) and p-value (0.29234909312094925). This led us to formulate our null hypothesis (H0) that the "Length of Stay" data follows a log-normal distribution, with the alternative hypothesis (H1) being that it does not. The relatively high p-value for the Log-Normal distribution provided less evidence against H0, suggesting

Table 1 – Summary Measures for different distributions of Length of Stay from Kolmogorov-Smirnov (K-S) Test

Distribution Type	Gumbel	Pareto	Gamma	Weibull	Log-Normal	Exponential
Hypotheses (H0/H1)	*H0*: Follows Gumbel *H1*: Does not follow	*H0*: Follows Pareto *H1*: Does not follow	*H0*: Follows Gamma *H1*: Does not follow	*H0*: Follows Weibull *H1*: Does not follow	*H0*: Follows Log-Normal *H1*: Does not follow	*H0*: Follows Exponential *H1*: Does not follow
K-S Statistic	0.1891	0.3528	0.1009	0.0920	0.0843	0.1483
P-Value	0.0001	<0.0001	0.1292	0.2044	0.2923	0.0056
Sample Size (N)	131	131	131	131	131	131
Mean	830.94	830.94	830.94	830.94	830.94	830.94
Std. Dev.	878.32	878.32	878.32	878.32	878.32	878.32
Comments	Poor fit	Very poor fit	Moderate fit	Moderate fit	Good fit	Poor fit

a better fit to our data compared to the other distributions tested.

These findings align with previous research that has often identified the log-normal distribution as a good fit for length of stay data, likely due to its ability to accommodate data with long tails. When we paired the results from the Kolmogorov-Smirnov Test with the initial distribution in **Figure 5**, it conclusively suggests that it follows a log-normal distribution. This characteristic is particularly relevant in the context of retirement home residency, where stay durations can vary significantly. The adoption of the log-normal distribution in our model reflects these unique aspects of the "Length of Stay" data, allowing for a more accurate and representative analysis of wait times and potential strategies for their reduction.

2.3 Wait Time Analysis & Other Calculations

Although we discovered the distribution and parameters of the length of stay in a retirement home, one thing we need to create an effective simulation was the same information for "wait times". In our examination of the wait times for a retirement home, we've conducted a thorough analysis to understand and potentially optimize the rate at which beds become available and, consequently, the rate at which individuals can join the waitlist.

The original situation presented us with a facility, [redacted], operating at full capacity with an occupancy rate of 100% and a mean length of stay for residents at approximately 830.9 days, following a log-normal distribution. The standard deviation of the stay length was quite high at 881.7 days, indicating significant variation in individual resident stays. To manage the transition between residents, a **3-day** sanitization and preparation period was factored in before a new resident could occupy a vacated bed.

Upon deeper inspection, it was necessary to calculate the **departure rate** – the frequency with which residents leave – thereby opening up beds. This rate, denoted as λ, was initially given as 0.00120346 departures per day. This figure was derived under the assumption of full occupancy (100%) where the mean length of stay was approximately **831 days** and we translated that into λ by the formula: **λ = 1 / mean length of stay**. However, to reflect a more realistic scenario and to potentially reduce wait times, we decided to adjust our calculations to an **85.5% occupancy rate**. This percentage is not arbitrary but is instead based on industry averages, providing a more pragmatic and attainable target for the facility.

To determine the adjusted departure rate at this occupancy level, we utilized the original λ but scaled it to the new occupancy rate. The rationale behind this is that the same number of departures would now be spread over fewer occupied beds, effectively increasing the departure rate. Mathematically, this was achieved by dividing the original λ by the occupancy rate, which in this case is **0.855**. This calculation resulted in an adjusted λ of approximately **0.0014076** departures per day.

With this new departure rate in hand, we then calculated the effective time between departures, which also includes the aforementioned 3-day waiting period for bed turnover. To clarify, the waiting period is essential for ensuring that the facility maintains high standards of cleanliness and safety – a non-negotiable aspect in the care of elderly residents. This aligns with standard practices for cleaning, sanitization, and necessary repairs or inspections that ensure a room is suitable for a new occupant. Thus, the total time from one resident's departure to the next resident's admission is the sum of the reciprocal of the adjusted λ and the waiting period. Through this calculation, we determined that it would take around **713 days** for a new individual to join the waitlist, given the adjusted parameters.

Furthermore, it is important to clarify that the dynamics of the waiting list are assumed to be exponential in nature, as suggested by the graph illustrating the steady incline in the population aged 75+ in Canada (See **Figure 7**). Unlike the length of stay distribution, which is log-normal, the waiting list is expected to grow exponentially. This means that the rate at which people are added to the list accelerates over time, in line with the increasing elderly population.

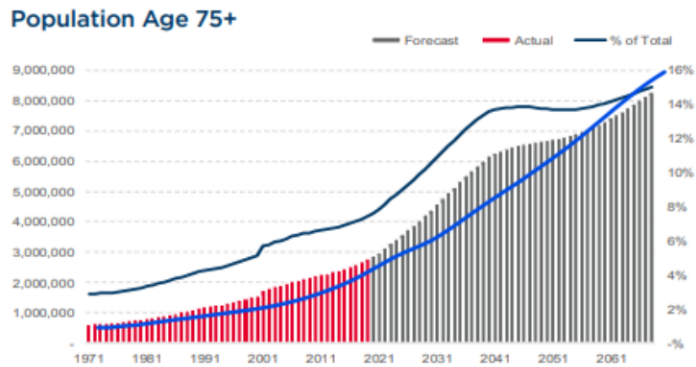


Figure 7 - Trend of Population Aged 75+ in Canada. Source: Cushman & Wakefield Report by McCrorie et al (2021) and StatsCan Tables 17-10-0005-01 and 17-10-0057-01 Projection scenario M4: medium-growth

In conclusion, the recalibration to an 85.5% occupancy rate, coupled with a steadfast 3-day preparation window, yields a more refined perspective on the waitlist dynamics. It suggests a more expedited turnover than a facility operating at full capacity, thus potentially reducing wait times for future residents.

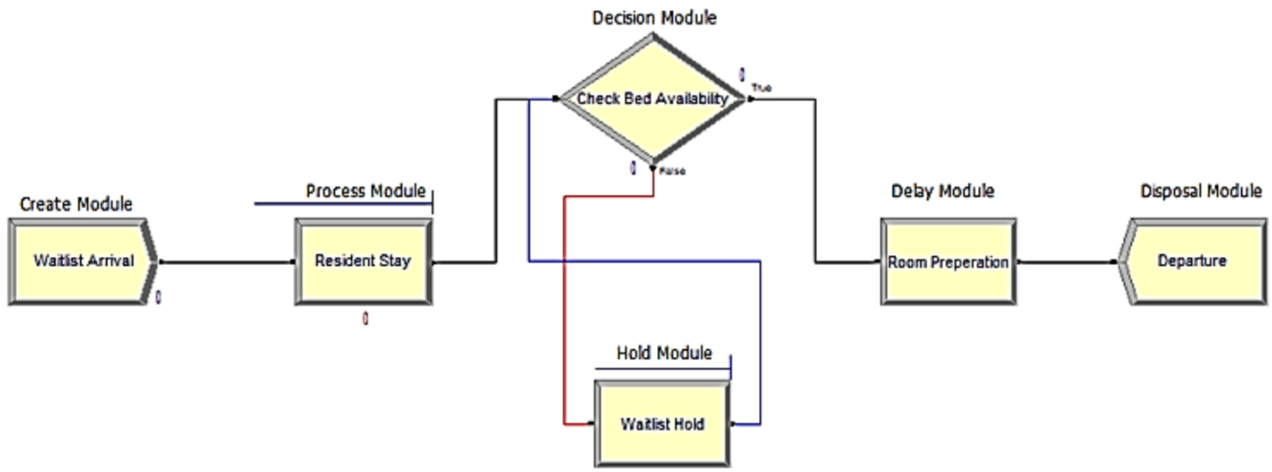


Figure 8 - Arena Model of Waiting List Queue Simulation into Retirement Home.

2.4 Verification and Validation

In an extensive simulation study of a retirement home with a current waitlist of 12 potential residents and a capacity of 118 beds, we employed **Arena simulation software** to evaluate the dynamics of bed occupancy and waitlist management. The simulation ran over a period of 10,000 days with 100 replications to ensure statistical reliability. This can be seen in **Figure 8**. Two scenarios were assessed: one with the facility operating at 85.5% bed occupancy (101 beds), reflecting the industry standard, and another at full capacity with 100% bed occupancy (118 beds). The mean length of stay for a current resident was modeled at 830.94 days with a log-normal distribution, and the waitlist was assumed to grow exponentially with an average of 713 days for a new individual to join the list.

The initial setup involved creating a 'Waitlist Arrival' module to simulate incoming residents, configuring a 'Process' module for their stay, and establishing a 'Dispose' module to represent departures. Resources were carefully adjusted to reflect bed availability, queues were defined to manage the waitlist, and delays were incorporated to simulate room preparation time. A 'Decide' module was necessary in checking bed availability, directing the flow based on occupancy.

The simulation results with the facility at 85.5% occupancy and a 3-day delay for room preparation showed an average total time in the system for potential residents at 859.15 days, including an average waiting time of 849.51 days. The bed resource utilization was minimal, with an average instant utilization of 0.01. The Work-In-Process (WIP) for potential residents averaged at 1.0031, suggesting that on average, there was approximately one resident either occupying a bed or waiting for one at any given time.

Upon adjusting the bed availability to full capacity and reducing room preparation delay to 1 day, the average total time in the system for potential residents showed a slight reduction to 857.15 days. However, the average waiting time remained unchanged at 849.51 days, indicating that increasing bed capacity alone did not significantly impact the wait times under the given conditions. This suggests that the bottleneck in the system may lie elsewhere, such as the rate at which beds become available rather than the total number of beds.

3. Output Analysis and Discussion:

3.1 Experiment Design:

The primary objective of this research was to develop an effective predictive model for [redacted], focusing on accurately estimating the move-out times of its residents. This task is important as it assists in minimizing the vacancy period of rooms, thereby maximizing occupancy rates, and consequently increasing revenue. The foundation of this model lies in the construction of a simulation model that leverages the lognormal distribution. This choice was driven by the distribution's ability to accurately model the positively skewed nature of move-out times.

During the design phase of our experiment, we opted for a simulation model primarily based on generating random numbers uniformly distributed between 0 and 1. To facilitate this, we employed Excel's RAND function. This function is integral to our simulation, as it generates random numbers that form the foundation of our predictive model. The process involved creating replications of these random numbers, starting at 100 replications, and incrementally increasing by 100 until the variability from one run to another was minimal. This ensured a robust and comprehensive dataset for our simulation. The choice of 1000 replications were made to strike a balance between computational feasibility and the need for a large enough sample size to capture the variability inherent in move-out times.

We tested several different distributions to ensure that the findings from our K-S goodness of fit test were accurate. We conducted simulations using the exponential distribution, Weibull distribution, and lognormal distribution. After thorough analysis, we concluded that the shape of the lognormal distribution best fit the shape of our data.

Our decision not to use a VLOOKUP table for our data simulation was based on observations made during our preliminary data analysis and creation of the distribution. Notably, three of the bins in our data contained zero values. Utilizing a lookup table in such a scenario would imply a 0% probability for resident move-outs within specific time frames (1900 to 1976 days, 2204 to 2280 days, and 2508 to 2584 days). Such a model would inaccurately represent these periods as having an impossible move-out rate. While the probability of move-outs during these periods is small, it is not nonexistent. However, due to the limitations of our dataset, the best method for extrapolating our data was to find a model that best fit our existing distribution.

The lognormal distribution was selected because of its suitability for modeling data that is not symmetrically distributed. It is particularly effective in situations where the data is positively skewed, as is the case with [redacted] resident move-out times. This distribution can handle the wide range of variability in move-out times, from shorter to significantly longer periods, without losing accuracy in the tails of the distribution.

Our simulation model played a crucial role in refining our predictive capabilities. By incrementally increasing the number of replications and closely monitoring the results, we were able to identify and minimize any inconsistencies or anomalies in the data. This iterative process was instrumental in fine-tuning the model to better reflect the real-world scenario at [redacted].

The reliability of our model was further validated through repeated simulations. Each run of the model produced results that were consistent with the previous ones, demonstrating the model's stability and dependability. This consistency is important for predictive models.

3.2 Statistical analysis:

The statistical analysis phase began with the transformation of our raw data into a format suitable for lognormal distribution analysis. This involved creating a new column in Excel to calculate the natural logarithm of the length of stay data for each resident. This step was crucial as the lognormal distribution models the distribution of the logarithm of a variable, rather than the variable itself.

Once we completed natural logging of our data, the next step was to compute key statistical parameters: the mean and standard deviation of our newly logged values. For our dataset, the mean and standard deviation were calculated to be approximately **5.97** and **1.43**, respectively. These parameters are important as lognormal distribution relies on the mean and standard deviation.

To simulate move-out times that follow a lognormal distribution, we applied the LOGNORM.INV function in Excel. This function requires the mean and standard deviation of the log-transformed data, along with the uniformly distributed random numbers generated earlier.

To represent the 1000 simulated times on a histogram, we used the same 42 bins with a bin width of 76 as our initial histogram. Using the FREQUENCY function in Excel, we were able to find how frequently a generated value would fall within the bounds of each bin. To ensure that we had accounted for all the generated values, we did a sum of the frequency column, double checking that the count was equal to 1000. To calculate the probability, we divided the frequency column by the total count of 1000. This way, we can identify the probability of each of the different bins.

The analysis of the Arena simulation output was approached using descriptive statistics to capture the central tendencies and distribution measures of the system's performance. The statistical examination focused on key metrics, including the average total time within the system for prospective residents, the average wait time before bed allocation, and the bed resource utilization rates. A comparative analysis was conducted between two operational scenarios to identify any statistically significant differences. These analyses provided a basic view of the system's functionality, enabling a quantitative assessment of operational efficiency against the varied bed occupancy rates.

3.3 Simulation results explanation

When looking at the results of the Excel simulation in **Table 2**, we can see that we have generated a frequency and probability column. Interpreting the data in the frequency column, the frequency number corresponds with the given upper limit of the bin. For example, the first bin of 76 has a frequency of 114. This means that out of the 1000 simulated values, 114 of them are less than 76. The corresponding probability indicates the percentage that a simulated value will fall within a certain bin. For example, when looking at the third bin (228) the

Table 2 - Summary of Lognormal Distributed values

Bins (UL)	Frequency	Probability
76	114	0.114
152	118	0.118
228	87	0.087
304	83	0.083
380	66	0.066
456	50	0.05
532	38	0.038
608	37	0.037
684	37	0.037
760	29	0.029
836	25	0.025
912	20	0.02
988	20	0.02
1064	15	0.015
1140	20	0.02
1216	8	0.008
1292	18	0.018
1368	12	0.012
1444	9	0.009
1520	10	0.01
1596	9	0.009
1672	6	0.006
1748	12	0.012
1824	7	0.007
1900	10	0.01
1976	10	0.01
2052	4	0.004
2128	6	0.006
2204	3	0.003
2280	12	0.012
2356	6	0.006
2432	5	0.005
2508	8	0.008
2584	4	0.004
2660	6	0.006
2736	3	0.003
2812	2	0.002
2888	3	0.003
2964	2	0.002
3040	4	0.004
3116	1	0.001
3192	2	0.002

probability indicates that there is an 8.7% chance that a resident will leave the facility between 153 to 228 days.

The Arena simulation's results were suggestive and gave a guide on waitlist management in the retirement home setting. In the first scenario, with an 85.5% occupancy rate, the system revealed an average total time of 859.15 days within the facility for potential residents, dominated by a wait time of 849.51 days. Despite the wait, bed resource utilization remained very low, suggesting inefficiencies in the transition process from waitlist to bed occupancy. When the bed capacity was elevated to full occupancy, there was a negligible decrease in the total system time to 857.15 days, and no improvement in the average waiting time. This persistent wait time, resistant to changes in bed capacity and room turnover rates, suggested operational constraints. Notably, the bed resource utilization remained minimal, with an instantaneous utilization averaging 0.01, and the Work-In-Process (WIP) for potential residents held steady at around 1, suggesting that at any given moment, a bed was either occupied or on the brink of being filled. The simulation thereby showed that the critical factor in reducing wait times was not simply a function of bed numbers but rather a more complex mix of operational processes, suggesting that enhancing the rate at which beds become available could potentially alleviate the bottleneck. This interpretation guides us toward targeted operational adjustments rather than capacity expansion as a solution to the waiting time problem.

4. Recommendations and Limitations:

4.1 Recommendation 1: Enhancing Room Turnover Efficiency

A critical factor in the extended wait times appears to be the turnover rate of rooms. The current three-day delay for room preparation contributes significantly to the overall wait time. By scrutinizing each step within the room preparation process, from cleaning to maintenance, we could identify inefficiencies. For instance, the employment of lean management principles or the introduction of a specialized rapid response team could expedite the turnover without compromising the quality of service. This modification is projected to have a substantial impact on reducing the system's total time, thereby enhancing throughput and resident satisfaction.

Future simulations should introduce variables that reflect optimized room turnover strategies. These will include scenarios with staggered cleaning crews working in shifts to ensure immediate room availability postdeparture and the implementation of predictive maintenance schedules to minimize room downtime. This would include monitoring the effects of these adjustments on the average wait time and the number of residents served.

4.2 Recommendation 2: Revising Departure Rate Protocols

The overall departure rate, which currently stands as a function of residents' length of stay, may be artificially inflating the waitlist backlog. We propose a review of the discharge process to determine if there are opportunities to safely accelerate resident transitions, either to home care or alternative facilities, without compromising care quality. This could involve a reassessment of care plans and discharge criteria, potentially leading to a more responsive departure rate.

We must explore the effects of a more fluid departure rate. This might involve modeling scenarios where residents have variable lengths of stay based on their health improvements or the availability of post-retirement home care options. The simulation will help us understand the relationship between departure rate modifications and waitlist management effectiveness.

4.3 Recommendation 3: Implementing Priority Rules

Given the diversity of needs among waitlisted individuals, it is smart to use a priority system rather than a "first-in first-out" method. This system would evaluate potential residents based on urgency of need, readiness to move in, and other critical factors. By categorizing waitlisted individuals and potentially offering expedited access to those with immediate needs, the home could better manage resident inflow and enhance the care provided to the most vulnerable.

To simulate the impact of a priority system, introducing attributes for potential residents that reflect their priority status. Different queue structures will be tested, such as priority queues that allow higher-need residents to access beds more rapidly. This will help measure how prioritization affects overall wait times and resource utilization.

4.4 Recommendation 4: Adjusting Marketing Strategies

Our analysis suggests the potential to adjust marketing strategies to influence the rate of incoming residents. By targeting marketing efforts during periods of low occupancy, the home could stabilize the occupancy rate throughout the year. This could help flatten variations in demand, leading to a more consistent waitlist and occupancy management.

Future simulations will incorporate marketing impact by modeling increased arrival rates during traditionally low occupancy periods. This strategy's effectiveness will be measured by the change in average waitlist times and the retirement home's ability to maintain optimal occupancy levels.

4.5 Limitations

While our study developed a comprehensive predictive model for [redacted]'s resident move-out times, we were faced with several limitations.

One significant constraint was the absence of data concerning the service time required to prepare a room for a new resident. [redacted] currently does not track the duration it takes for their employees to clean and prepare a room for the next occupant. This gap in data presents a challenge as it omits a crucial factor in accurately estimating the room turnover rate, directly impacting the model's overall efficacy.

Another notable limitation pertains to the variability in reasons for resident departure, such as medical issues, death, or dissatisfaction. The turnover time can fluctuate considerably based on these factors, especially in scenarios where employees must clean out a room following a resident's death, which typically requires more time and resources compared to rooms left by dissatisfied residents.

Additionally, our study faced constraints due to the limited scope of data regarding departure rates across different regions in Canada. Our primary data source was the [redacted] location in [redacted] Manitoba. The absence of comparative data from other [redacted] locations within the province or across the country means that our simulation model's applicability might be confined to the specific circumstances at [redacted]. As our earlier examination in the study highlighted, vacancy rates and waiting times can differ from one province to another, thereby limiting the applicability of our findings.

A further limitation in our study was the lack of comprehensive data on waiting line abandonment. [redacted]'s current method of managing their waiting list, which involves a receptionist recording customer details on a spreadsheet and removing names as individuals are admitted, lacks the sophistication necessary for detailed tracking. This rudimentary system lacks any ability to collect data on the duration prospective residents spend on the waiting list and at what point they decide to abandon the queue. The absence of this data represents a significant gap, as understanding waiting line/queuing system dynamics is crucial for accurately forecasting room availability and optimizing occupancy rates.

Another limitation is that Length of Stay distributions are quite difficult to fit to data from different datasets. One facility varies significantly against another, especially with different operations and multiple units. Thus, lognormal distribution may not work across all length of stay data in different retirement homes.

Looking at the limitations in Arena, we assumed a constant bed turnover and exponential waitlist arrival rates, which may not accurately reflect real-world variability. These rates were based on averages, thereby overlooking potential fluctuations in resident departures and new admissions. This assumption could mask peak periods of high demand or lower occupancy, leading to an oversimplified understanding of the retirement home's operational dynamics. Consequently, this limitation could affect the generalizability of our findings to actual scenarios, where random events and seasonal variations play a significant role in occupancy rates and waitlist movements.

Lastly, we used a student version of Arena that would not allow us to replicate past 150 entities (i.e. residents), thus we had to be conservative in how much people would join the waiting list by creating an inflated mean of 713 with an exponential distribution. With more data on the waiting list entries and distribution, we could have more realistically discovered when people join the waiting list.

5. Conclusion

The study employed a lognormal distribution model, selected for its ability to capture the positively skewed nature of resident move-out times. This approach was chosen after a thorough review of relevant literature, which predominantly focused on hospital and ICU settings, indicating a gap in specific research for retirement homes. Previous studies have shown mixed success with various distribution models in different healthcare contexts, but our analysis using the Kolmogorov-Smirnov Test suggested that the lognormal distribution was the most fitting for our data.

A critical aspect of our model was its consideration of the mandatory unoccupied period between resident transitions, mandated by the cleaning and preparation processes. This factor introduced significant variability into the model, as the time for room preparation could vary based on the reasons for a resident's departure.

Our methodology involved data collection from [redacted], with an emphasis on understanding the length of stay and wait times for residents. We utilized Microsoft

Excel for simulation, due to its familiarity and data visualization capabilities. The model began with 100 replications, eventually expanding to 1000 to ensure a reliable dataset. Our analysis used Excel's RAND function for generating random numbers and the LOGNORM.INV function for the lognormal distribution. The resulting data was then categorized into bins using the FREQUENCY function, providing insights into the probability of move-out times within specific periods.

An Arena simulation was utilized to analyze bed occupancy and waitlist management at [redacted], a retirement home with 118 beds. The simulation, spanning 10,000 days with 100 replications, examined two scenarios: one at an 85.5% bed occupancy rate and another at full capacity. Key elements of the simulation included modules for waitlist arrival, resident stay process, and departures, along with resource management for bed availability and room preparation time. The results indicated that at 85.5% occupancy, the average total system time for a resident was 859.15 days, including a significant waiting period. Increasing the bed capacity to full did not notably reduce the waiting time, suggesting that the bottleneck in reducing wait times was more complex than just bed numbers. These insights pointed towards the need for operational adjustments, particularly in speeding up bed turnover, to effectively manage the waitlist and improve occupancy rates.

The reliability of our model was verified through repeated simulations, demonstrating its stability and potential effectiveness in forecasting move-out times. This predictive capability is crucial for [redacted]'s operational efficiency, as it could significantly reduce the waiting time for new residents, thereby benefiting both the facility and its clientele.

However, the study faced several limitations, including the lack of data on room preparation time and the variability in reasons for resident departure. Furthermore, the study's findings were based on data from a single location, potentially limiting its applicability to other regions or facilities with different operational dynamics. The rudimentary nature of [redacted]'s waiting list management system also presented challenges in tracking and analyzing waitlist dynamics.

In conclusion, while the study provides valuable insights and a robust predictive model for [redacted], addressing these limitations could further enhance its accuracy and applicability. This research contributes significantly to the operational management of retirement homes, offering a model that can adapt to the dynamic nature of resident turnover and improve overall efficiency and profitability.

References

[redacted]

Bae, K. H., Jones, M., Evans, G., & Antimisiaris, D. (2017). *Simulation modelling of patient flow and capacity planning for regional long-term care needs: a case study*. Health systems (Basingstoke, England), 8(1), 1–16. https://doi.org/10.1080/20476965.2017.1405873

Canadian Institute for Health Information. (2021, June 10). *Long-term care homes in Canada: How many and who owns them?* CIHI. https://www.cihi.ca/en/long-term-care-homes-in-canada-how-many-and-who-owns-them

Canadian Institute for Health Information. (2023). *Profile of Residents in Residential and Hospital-Based Continuing Care, 2022–2023*. CIHI. https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.cihi.ca%2Fsites%2Fdefault%2Ffiles%2Fdocument%2Fresidential-hospital-based-continuing-care-ltc-quick-stats-2022-2023-data-tables-en.xlsx&wdOrigin=BROWSELINK

CMHC. (2021, July 6). *Seniors housing survey data*. CMHC SCHL. https://www.cmhc-schl.gc.ca/professionals/housing-markets-data-and-research/housing-data/data-tables/rental-market/seniors-housing-survey-data-tables

CMS. (2023a, December 6). *Medicare Claims Quality Measures*. Centers for Medicare & Medicaid Services. https://data.cms.gov/provider-data/dataset/ijh5-nb2v?fbclid=IwAR3QxJz4PpyVAxURbkzWwSYKQInFDl6-F6DbpkouPA3ln4_FlnSdOi4y9A0

CMS. (2023b, December 6). *Provider Information*. Centers for Medicare & Medicaid Services. https://data.cms.gov/provider-data/dataset/4pq5-n9py?fbclid=IwAR0FdrwJwiQ6CoHMU1VYpijQ3-SU8I1QxIA8fIGb1OVSNaWlPiTA5uY1ylI

Eisen, B. (2022, July 5). *Canada's aging population—what does it mean for government finances?* Fraser Institute. https://www.fraserinstitute.org/blogs/canadas-aging-population-what-does-it-mean-for-government-finances

Faddy, M., Graves, N., & Pettitt, A. (2009). *Modeling length of stay in hospital and other right skewed data: Comparison of phase-type, gamma and log-normal distributions*. Value in Health, 12(2), 309–314. https://doi.org/10.1111/j.1524-4733.2008.00421.x

Government of Canada, S. C. (2022, April 27). *In the midst of high job vacancies and historically low unemployment, Canada faces record retirements from an aging labour force: Number of seniors aged 65 and older grows six times faster than children 0-14*. The Daily. https://www150.statcan.gc.ca/n1/daily-quotidien/220427/dq220427a-eng.htm

Griffiths, J. D., Jones, M., Read, M. S., & Williams, J. E. (2017). *A simulation model of bed-occupancy in a Critical Care Unit*. Journal of Simulation, 4(1), 52–59. https://doi.org/10.1057/jos.2009.22

Jun, J. B., Jacobson, S. H., & Swisher, J. R. (1999). *Application of discrete-event simulation in Health Care Clinics: A survey*. Journal of the Operational Research Society, 50(2), 109–123. https://doi.org/10.1057/palgrave.jors.2600669

McCrorie, S. (2022, January). *Canadian Seniors Housing Industry Overview: Insights: Canada*. Cushman & Wakefield. https://www.cushmanwakefield.com/en/canada

/insights/canadian-seniors-housing-industry-overview

Mordor Intelligence. (2023). *Canada Senior Living Companies*. https://www.mordorintelligence.com/industry-reports/canada-senior-living-market-analysis/companies

Ridge, J. C., Jones, S. K., Nielsen, M. S., & Shahani, A. K. (1998). *Capacity planning for Intensive Care Units*. European Journal of Operational Research, 105(2), 346–355. https://doi.org/10.1016/s0377-2217(97)00240-3

Ritchie, H., & Roser, M. (2019, September 20). *Age Structure*. Our World in Data. https://ourworldindata.org/age-structure

Statista Research Department. (2023, October 10). *U.S. - Seniors as a percentage of the population 1950-2050.* Statista. https://www.statista.com/statistics/457822/share-of-old-age-population-in-the-total-us-population/

Statista Research Department. (2023, October 17). *Canada: Median age of the resident population 2000-2022*. Statista. https://www.statista.com/statistics/444844/canada-median-age-of-resident-population/

The World Bank Group. (2022b). *Age dependency ratio (% of working-age population) - United States, Canada.* World Bank Group. https://data.worldbank.org/indicator/SP.POP.DPND?locations=US-CA

The World Bank Group. (2022a). *Population ages 65 and above (% of total population) - United States, Canada.* World Bank Group. https://data.worldbank.org/indicator/SP.POP.DPND?locations=US-CA

United States Census Bureau. (2023, June 22). *America Is Getting Older*. https://www.census.gov/newsroom/press-releases/2023/population-estimates-characteristics.html

Zhang, X., Barnes, S., Golden, B., Myers, M., & Smith, P. (2019). *Lognormal-based mixture models for robust fitting of hospital length of stay distributions*. Operations Research for Health Care, 22, 100184. https://doi.org/10.1016/j.orhc.2019.04.002

Zhang, Y., Puterman, M. L., Nelson, M., & Atkins, D. (2012). *A simulation optimization approach to long-term care capacity planning*. Operations Research, 60(2), 249–261. https://doi.org/10.1287/opre.1110.1026