M3 Challenge 2024:

A Tale of Two Crises: *The Housing Shortage and Homelessness*

TEAM #17432

March 3rd, 2024

Executive Summary

To the Secretary of the U.S. Department of Housing and Urban Development,

As the housing affordability gap widens, an increasing number of families and individuals are unable to secure stable shelter. Every night, over 650,000 people experience homelessness nationwide[1]. Homelessness exposes people to dangerous health conditions, including contagious diseases, physical attacks, and extreme weather. The life expectancy of a homeless person is just 50 years, compared to 78 years for a non-homeless person[2]. Not only is homelessness a public health issue, but it also negatively impacts the economy[3] and the environment[4]. Although weather, political violence, and interpersonal conflicts can displace individuals, the lack of affordable housing[5] is the primary driver of homelessness. To guide effective policy-making, it is crucial to understand the relationship between the housing crisis and the rise of homelessness.

We forecasted the growth of housing supply in Seattle, Washington, and Albuquerque, New Mexico using a logistic regression model. We determined the carrying capacities of both cities and predicted the housing supply with 95% confidence intervals. Our model estimates 451,164; 489,908; and 529,251 housing units for Seattle and 312,107; 339,504; and 368,391 housing units for Albuquerque in 10, 20, and 50 years respectively.

We then quantified changes in the homeless populations of Seattle and Albuquerque with a vector autoregressive (VAR) model. Our results project an increase in the homeless populations of both Seattle and Albuquerque. We predict that Seattle's homelessness rate will increase by 17.2% in 10 years, 32.6% in 20 years, and 63.2% in 50 years compared to its baseline population in 2022. We forecast drastically different growth rates of 46.8% in 10 years, 52.4% in 20 years, and 54.3% in 50 years for Albuquerque's homelessness rate. Our model incorporates the proportion of income spent on housing, mortgage rates, and population changes, contributing to a comprehensive and robust model.

Finally, to assist city planners, we developed a novel affordable housing optimization model that suggests affordable housing projects (in housing units) based on the goal of a homelessness threshold and a number of years. For example, if the minimum threshold of remaining homeless individuals without affordable housing plans is 800 and we have 20 years to set up our affordable housing plan, then our model finds that we can succeed using only 75% of Seattle's urban development budget. We then adapted the model to consider the risks and impacts of economic turmoil, natural disasters, and migrant influx. Our model showed that at the same threshold and time period, budget loss due to economic recessions, housing loss due to natural disasters, and increased homelessness due to migrant influx have a substantial effect on the minimum budget required for a successful affordable housing development plan.

We believe our results will help policymakers better understand the housing crisis and enable them to design more effective proposals to address homelessness. Our models offer key insights and are a powerful tool for crafting generalizable solutions across the nation.

Table of Contents

Global Assumptions

G-1. **Housing supply refers to the total number of housing units, including occupied units.**

- **Justification:** Because housing is a non-consumable good, every housing unit can theoretically be sold and bought at any time, contributing to the supply. Even when a residence is occupied, a buyer can always request the homeowner to sell their property.

G-2. **For our purposes, the price per person of renting a housing unit is equivalent to the price per person of mortgage for a housing unit.**

- **Justification:** Both mortgages and rent prices are derived monthly from an individual's income. Furthermore, mortgage prices have a direct correlation to a landlord's rent.

G-3. **When discussing the region of a US city, we are referring to the US Census-designated metropolitan area of that city.**

- **Justification:** Different references to cities include varying inclusions of suburban areas. To avoid this discrepancy, we consider the Census-designated boundary of cities as the official metric.

G-4 **Every buyer in a region purchases a house at the same mortgage rate, mortgage term, down payment, and credit score.**

- **Justification:** Mortgage payments depend on mortgage rates, mortgage terms, down payments, and credit scores, which vary among sellers and buyers. To simplify our models and account for variation, we consider the average of these factors in a local area.

G-5 **Data points during and before the 2008-2009 recession do not accurately reflect the modern housing market.**

- **Justification:** The 2008 financial crisis drastically altered the housing market. To ensure accuracy, we only consider data from 2010 and beyond. This is to consider the housing market during the post-2008 ample-reserve monetary policy[20].

G-6 **Home ownership is considered affordable if a buyer spends no more than 30% of their pre-tax income on housing costs, including mortgage payments and utilities.**

- **Justification**: Lenders traditionally recommend allocating no more than 30% of pre-tax income to housing[6]. Because utilities are essential for comfortable living, they are included in the 30% rule.

G-7 **Homeless individuals are unable to purchase vacant housing.**

- **Justification:** By definition, homeless people are unable to gain a permanent home. Because the housing market is a key factor in homelessness, we assume that homeless individuals cannot afford vacant housing.

Q1: It Was the Best of Times

1.1 Defining the Problem

The first problem asks us to predict changes in the housing supply in either the two given U.S. regions or the two given U.K. regions in the next 10, 20, and 50 years. To allow for better generalizability for the following questions, we chose to analyze the U.S. regions of Seattle, Washington, and Albuquerque, New Mexico.

1.2 Assumptions

1-1. **Every region has a carrying capacity for the number of houses built.**

- **Justification:** Because a region is a fixed area of land, it has a limited amount of space for houses. The carrying capacity refers to the maximum number of houses in a region due to its physical constraints.

1-2. **The housing supply in Seattle and Albuquerque grew the fastest in 2021.**

- **Justification:** In a logistic model, the fastest growth rate occurs at the midpoint. To fit our model to the logistic function, the midpoint we use must be included in our data. Since 2021 is the most recent data point we have, we assume 2021 is the midpoint.

1-3. **External factors, such as weather and government policies, do not drastically impact the housing supply on a yearly basis.**

- **Justification:** Previous data on yearly housing supply already accounts for external influences, so it is sufficient for predicting future housing supply values.

1-4. **Seattle and Albuquerque's household unit growth capacity is proportional to its current housing units.**

- **Justification:** Seattle and Albuquerque are both relatively large cities growing steadily, hence, we can conclude that the ratio of the growth capacity of housing units to current housing units is proportional across cities.

1.3 Variables

Symbol	Definition	Units
$S_{s, a}$	Yearly housing supply	Number of Housing Units
t	Time	Year
t_0	Time of logistic function midpoint	Year
k	Growth rate	1/Year
$L_{s, a}$	Carrying capacity of housing supply in Seattle, WA, and Albuquerque, NM	Number of Housing Units
$G_{s, a}$	Growth capacity of housing supply in Seattle, WA, and Albuquerque, NM	Number of Housing Units

Table 1: Variable Definitions for Problem 1

1.4 The Model

1.4.1 Developing the Model

We utilized a logistic regression model to predict changes in the housing supply of Seattle and Albuquerque. Logistic growth often represents variables with resource constraints, making it suitable for housing supply, which is limited by land available for residential development. Furthermore, logistic models are fairly resistant to outliers and can address variations in the housing market resulting from the COVID-19 pandemic.

To predict the housing supply *S* in year *t,* we used the formula for the logistic function:

$$S = \frac{L}{1 + e^{-k(t-t_0)}}.$$

Equation 1: Logistic function

In the equation, t_0 is the year corresponding to the midpoint of the logistic curve, where the rate of change is highest. We assumed that t_0 is 2021. The carrying capacity for housing units is *L*. In the years before the midpoint year, the housing supply is rising rapidly. After the midpoint year, the housing supply begins to flatten out because $t - t_0$ becomes positive. Over time, the denominator becomes close to 1 and the housing supply asymptotically approaches the carrying capacity.

Our team considered using a simple linear regression to predict the growth of the housing supply. However, linear regression is less effective for long-term modeling, because the local linearity of a relationship may not hold for extended periods of time. We only use past data from 12 years, so it is unlikely that linear extrapolation would accurately forecast the housing supply 50 years into the future. Additionally, linear regression assumes a constant rate of change lasting indefinitely, which is problematic when modeling a variable with a clear upper bound.

1.4.2 Executing the Model

We utilized the data provided by the Mathworks Math Modeling Challenge on total housing units from 2010 to 2021[21]. We omitted data from 2022 to test the accuracy of our model. After importing our datasets into a Python notebook, we employed the curve fit function from the SciPy library. This approach enables us to automatically fit a logistic curve to our data while allowing us to input our own carrying capacity value.

Using a report produced by Seattle's Office of Planning and Community Development, we found that the growth capacity for housing units in Seattle is 172,440 units. Due to the lack of data from Albuquerque's local government, we assumed that the ratio between the growth capacity of the housing supply and the current housing supply is the same for Albuquerque and Seattle. We calculated that the growth capacity of Albuquerque is 120,212 housing units. Adding this value to the current number of housing units in Albuquerque, we get the carrying capacity for Albuquerque.

$$\frac{G_s}{S_s} = \frac{G_a}{S_a}$$

Equation 2: Growth capacity proportionality assumption

To find the growth rate *k*, we iterated through values between 0.001 and 0.999 and identified the value minimizing the mean squared error (MSE) for both Seattle and Albuquerque.

Variable	Value for Seattle	Value for Albuquerque
L	535,249	373,136
k	0.064	0.068
t_0	2021	2021

Table 2: Variables for the logistic function

1.5 Results

Using the values in Table 2, we predicted the housing supply in 10 years, 20 years, and 50 years with the logistic function. We calculated and plotted the values with Python. We also included a 95% confidence interval to represent the inherent uncertainty of long-term predictions.

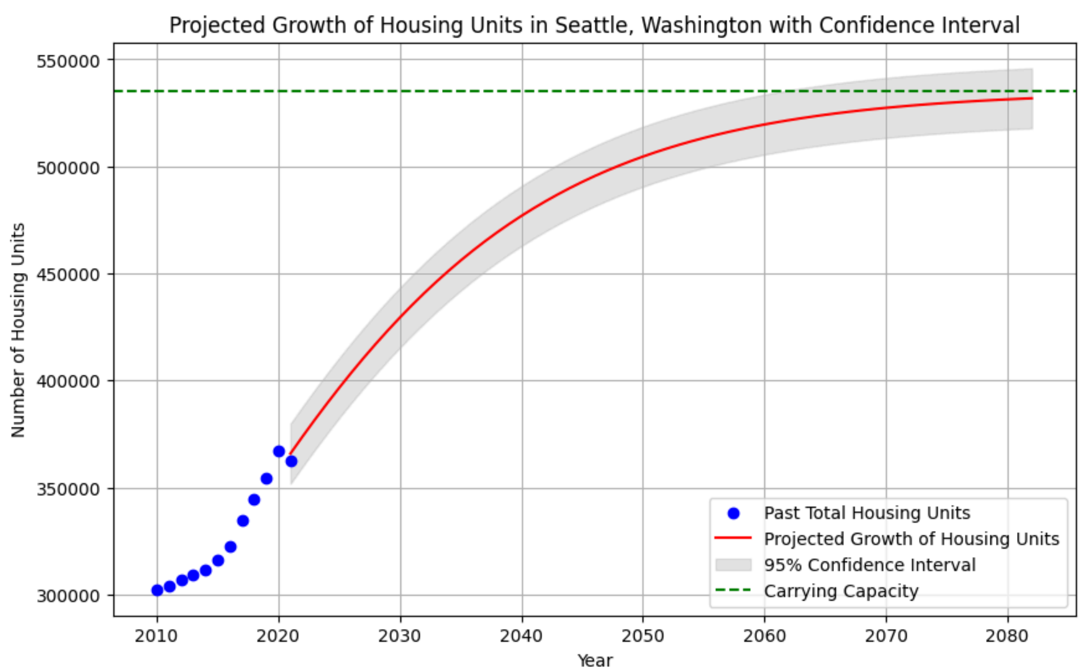

Figure 1: Graph of past and predicted housing supply in Seattle through 2074

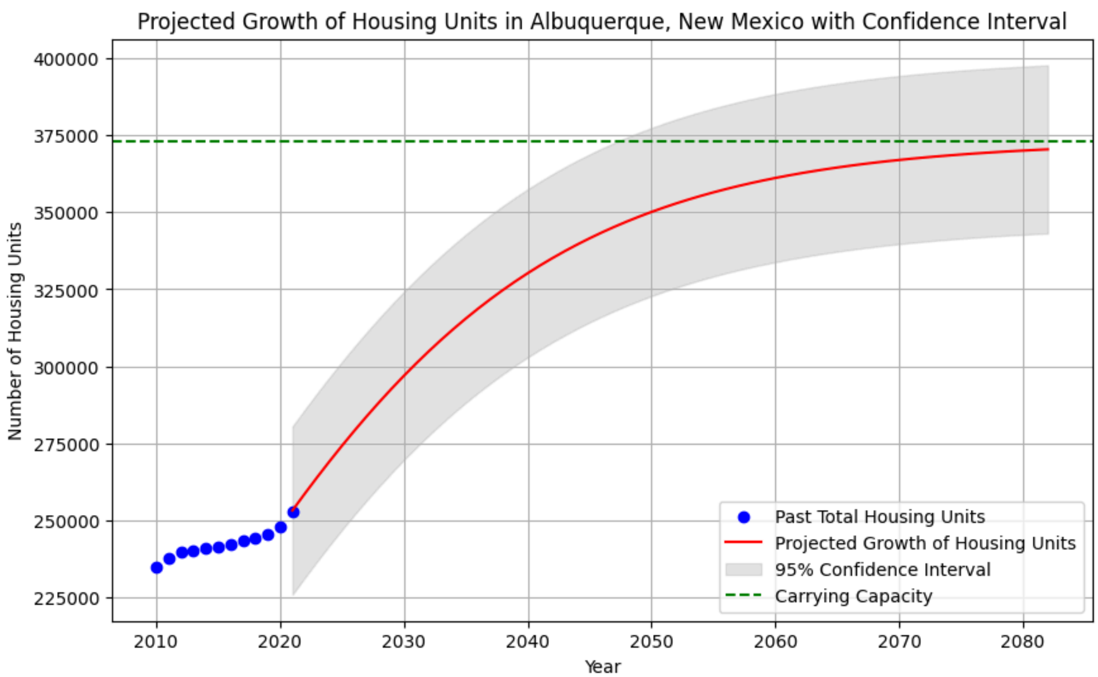

Figure 2: Graph of past and predicted housing supply in Albuquerque through 2074

Year	Predicted Housing Supply (Seattle)	Percent Growth Since 2021
2034	451,164	24.353%
2044	489,908	35.032%
2074	529,251	45.876%

Table 3: Predicted housing supply in Seattle

Year	Predicted Housing Supply (Albuquerque)	Percent Growth Since 2021
2034	312,107	22.310%
2044	339,504	33.046%
2074	368,391	44.366%

Table 4: Predicted housing supply in Albuquerque

1.6 Discussion

Our logistic function predicts that Seattle will have a housing supply of 451,164 in 2034, 489,908 in 2044, and 529,251 in 2074. In Albuquerque, our model forecasts a housing supply of 312,107 in 2034, 339,504 in 2044, and 368,391 in 2074. As expected, the housing supply increases slower in later years for both cities. The housing supply grows at similar rates for

Seattle and Albuquerque, which makes sense because the growth rate values *k* were very similar. However, the confidence interval for Albuquerque is larger than the confidence interval for Seattle, which may be because Albuquerque is in the earlier stages of urban development.

$$P.E = \frac{predicted - actual}{actual} \times 100\%$$

Equation 3: Percent error formula

As mentioned above, we omitted the 2022 data while fitting the logistic regression curve so that we can use it to test our model's accuracy without bias. In 2022, we predicted the housing supply of Seattle to be 373,853, while the actual housing supply in 2022 was 372,436. Using the percent error formula, we get a **0.38%** error. For Albuquerque in 2022, we predicted the housing supply to be 258,727, while the actual housing supply was 255,178. This prediction yields a **1.39%** error. The higher percent error for Albuquerque aligns with its wider confidence interval for predicting the housing supply. The low percent errors of both Seattle and Albuquerque highlight the accuracy of our model.

In both the confidence interval and the percent error analysis, Albuquerque was more variant and error-prone than Seattle. This is because Albuquerque has less housing development than Seattle, and thus is in its nascent stages. Data from this stage is harder to extrapolate and thus produces more error-prone results.

1.7 Sensitivity Analysis

To perform a sensitivity analysis on our model, we chose to analyze the impact of changing the carrying capacity of the housing supply in Seattle (L_s) on our predictions. We increased the carrying capacity by 5% from 535,249 houses to 562,011 houses.

Year	Predicted Housing Supply (Seattle)	Predicted Housing Supply with 5% increased Carrying Capacity	Percent Change
2034	451,164	449,445	-0.381%
2044	489,908	492,703	+0.571%
2074	529,251	548,344	+3.608%

Table 5: Effects of modifying carrying capacity for Seattle

By increasing the carrying capacity, the initial predicted values decreased slightly and the later predicted values increased. This phenomenon occurs because the model assumes that by having a larger carrying capacity, Seattle is in an early stage of housing development. Thus, growth will be slower initially and will speed up to reach the increased carrying capacity afterward. Our new results correlate with expected intuition and therefore establish the robustness of our model.

1.8 Strengths & Weaknesses

Since the housing supply is constrained by available land, it is natural to incorporate the carrying capacity of the housing supply into a housing growth model. By implementing a logistic growth model we incorporated the carrying capacity of housing units and showed the intuitive evolution of housing growth rate through our graph. Through percent error calculations, we showed that our model successfully forecasted Seattle's housing supply in 2022, demonstrating its potential for future predictions. Additionally, our model uncovered and highlighted the differences between Seattle and Albuquerque through their different percent errors and confidence intervals. Thus, we can gain key insights from our model to understand the prediction domain. Our sensitivity analysis further reveals the robustness of our model, even in the long term. Unlike other models, logistic models handle outliers well, so they can account for fluctuations in the housing market. Moreover, our model is practical and can easily find values for other years by expanding the time scale.

However, our model may not be entirely accurate due to the unreliability of long-term forecasting. We only had 12 years of data to predict the housing supply in 50 years, which is a large extrapolation. Our results highlight this uncertainty, as our confidence intervals were fairly wide for both Seattle and Albuquerque. We also assumed that the housing supply will grow the fastest (inflection point) in 2021, which may not reflect real-world trends. Furthermore, the logistic function is univariate and does not consider economic factors or demographic changes impacting the housing supply. Given more time, we would have tested a multivariate approach for predicting the housing supply.

Q2: It Was the Worst of Times

2.1 Defining the Problem

The second problem asks us to forecast changes in the homeless population in the next 10, 20, and 50 years in the two cities we chose in question one: Seattle and Albuquerque.

2.2 Assumptions

2-1 **Changes in the housing market are the largest factor behind changes in homelessness.**

- **Justification:** As housing prices increase, more families and individuals are unable to afford their mortgage costs, resulting in higher levels of homelessness.

2-2 **Natural disasters, economic crashes, and increased migration do not significantly impact rates of homelessness.**

- **Justification:** Although extreme events would displace communities and increase homelessness, it is impossible to predict such events. Thus, we did not include them in our model.

2-3 **The ratio of income required for housing, mortgage rate, and population are correlated with each other.**

- **Justification:** Fluctuations in mortgage rates are expected to affect housing prices, while population has a more long-term impact[17]. Although causation does not imply a correlation, we assume correlation for the sake of time and our model.

2.3 Variables

Symbol	**Definition**	**Units**
n	Number of endogenous variables	Number
$\mathbf{Y_t}$	*n*-dimensional vector of endogenous variables	Number
$\mathbf{X_p}$	*n* x *n* coefficient matrices representing the *p*-th lagged relationship between variables	Number
$\mathbf{u_t}$	Vector of white noise residuals at time t	Number
M	Median household income	USD per year
P	Median mortgage payment per year	USD per year
r	Ratio of the income proportion to finance a house compared to the 30% recommended ratio	Number

Table 6: Variables for Problem 2

2.4 The Model

2.4.1 Developing the Model

We chose a vector autoregressive (VAR) model to predict changes in the homeless population of Seattle and Albuquerque. We considered univariate methods, including simple regression, but determined that they would not account for the multifaceted nature of the housing market. Homelessness is a complex issue with no single cause, so it is important to examine different factors driving homelessness.

We utilized the VAR model because of its flexibility and ability to analyze multivariate time series data[8]. The VAR model has widespread applications in econometrics, making it suitable for predicting the rate of homelessness, which depends on financial factors in the housing market[9]. In the VAR model, each variable is a linear function of its previous values and the previous values of other variables. The VAR model employs endogenous variables, which are variables assumed to be correlated to each other. Mathematically, the *p*-th generated value from the VAR model (for *p*-lagged observations) can be expressed as:

$$VAR(p) = Y_t = \sum_{i=0}^{p} \left(X_i Y_{t-i} + u_t \right)$$

Equation 4: Mathematical representation of VAR

For example, a VAR(1) model for two variables Y_1 and Y_2 is represented as:

$$VAR(1) = Y_t = X_1 Y_{t-1} + u_t$$

Equation 5: Mathematical representation of VAR(1)

VAR(1) would be calculated as follows, where a_{ij} is the impact of $Y_{j,\,t-1}$ on $Y_{i,\,t}$:

$$\begin{pmatrix} Y_{1,t} \\ Y_{2,t} \end{pmatrix} = X_1 \begin{pmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}, X_1 = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$$

Equation 6: A more detailed matrix representation of the Equation 5

2.4.2 Executing the Model

The VAR model uses the endogenous variables of mortgage rate, the average ratio of income used to finance housing and population, forecasting the homeless population in Seattle and Albuquerque in 10, 20, and 50 years. Based on our assumption that the housing market would be a major influence on homelessness, these endogenous variables were chosen to explicitly represent the changes in the housing market itself, as well as its collateral.

We utilized the population data provided by the Mathworks Math Modeling Challenge from the years 2010 to 2022[21]. Additional data on mortgage rates was obtained from reputable sources, including the Federal Housing Finance Agency[10], [11] and Zillow[12], [13], as no single source contained every value from 20010 to 2022. We found the monthly mortgage payment from the mortgage rate, mortgage term, down payment, and credit score using an online calculator[14]. The most popular mortgage term is 30 years, and the traditional down payment is 15%[15]. Using historical data on average credit scores by state, we considered the credit score range of 700 to 719 for Washington and 680 to 699 for New Mexico. We multiplied the monthly mortgage payment by 12 to find the yearly mortgage payment *M*. Because we assume that a household should only be spending 30% of their income on housing, we compared the proportion of annual income used on housing to the threshold 30% value through a ratio *r*. We obtained the median income and median listing price from the Mathworks Math Modeling Challenge[21].

$$r = P(x)/(0.3 * M)$$

Equation 7: Formula for the ratio of income needed to afford housing-related finances

We could not find the full mortgage rate data from 2010 to 2022 for Albuquerque, so we initially replaced the missing values with the national mortgage rate. However, using the VAR on the values yielded sporadic and dramatic trends, which makes sense as a single city may not accurately reflect national trends. Instead, we omitted the mortgage ratio for Albuquerque.

The VAR model requires stationary variables, so we tested for stationarity using the Augmented Dickey-Fuller (ADF) test with a significance level of 0.05. However, even after implementing common transformation methods such as differencing and log transformations, some of the variables remained non-stationary. Instead, we tested for cointegration of the two non-stationary variables, which shows whether the combination of non-stationary variables yields a stationary (and therefore, stable) relationship in the long term.

We employed the Engle-Granger Two-Step Method to test for cointegration, which first regresses one variable on another and then checks the residuals for stationarity. Using the ADF test again, we found that the residuals were stationary. Thus, we can safely proceed with our VAR model.

City	Endogenous Variable	Stationary?	Cointegrated?
Seattle	Ratio of Income	No	Yes
	Mortgage Rate	No	Yes
Albuquerque	Ratio of Income	No	Yes
	Total Population	Yes	Yes

Table 7. Variables in the VAR models for Seattle and Albuquerque

2.5 Results

We used Python and libraries such as pandas, numpy, and statsmodels to carry out and implement the VAR model. Below are the graphs generated by matplotlib for the VAR predictions for homeless population growth for the next 50 years.

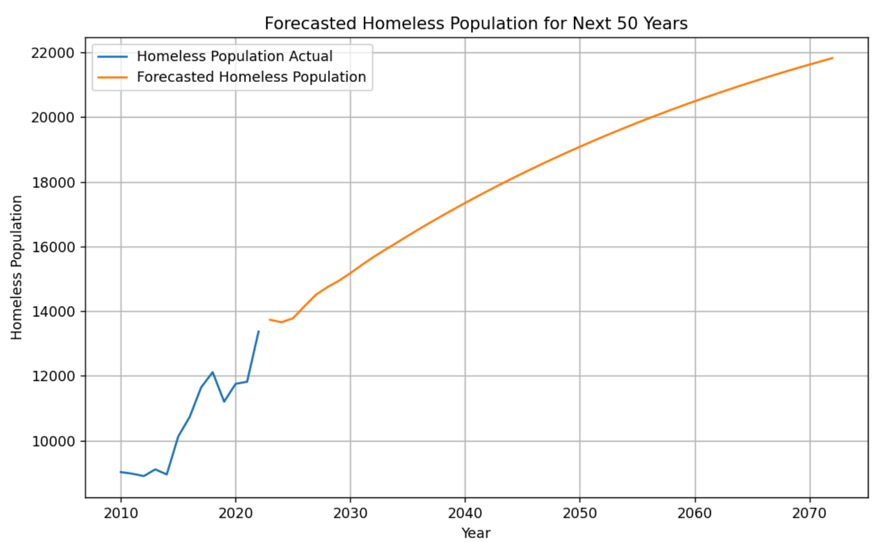

Figure 3. Graph of predicted homeless population in Seattle through 2072

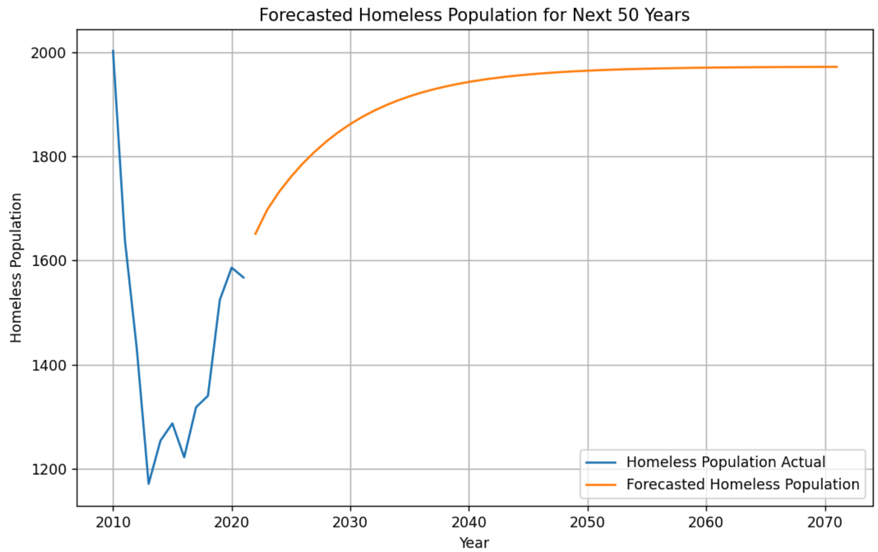

Figure 4. Graph of predicted homeless population in Albuquerque through 2072

As the differing variables of mortgage rate and population may result in changes in the model, we also considered using population instead of mortgage rate in the VAR model for Seattle.

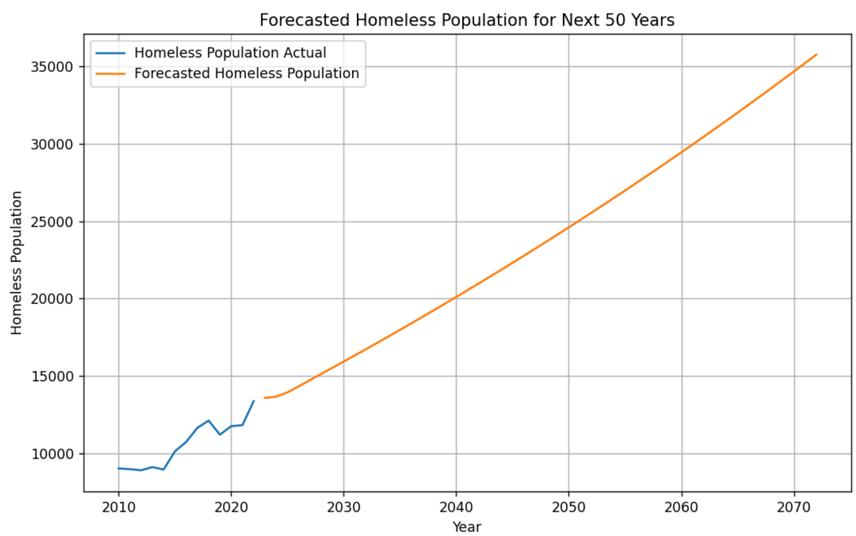

Figure 5. VAR model predictions for Seattle's homeless population given that population replaces mortgage rate as a predictor

We ran the VAR model separately for steps of 10, 20, and 50 years to project the homeless population in Seattle and Albuquerque.

City	Year	Homeless Population	% Change from 2022
Seattle	2022	13,368	0
	2032	15,668	+17.2
	2042	17,722	+32.6
	2072	21,822	+63.2
Albuquerque	2022	1277	0
	2032	1875	+46.8
	2042	1946	+52.4
	2072	1971	+54.3

Table 8: Predicted changes in the homeless populations of Seattle and Albuquerque

2.6 Discussion

The vector autoregressive model finds that the homeless population is expected to increase in both Seattle and Albuquerque, which aligns with the current trend of high cost housing. The graphs reveal that growth in homeless populations is logarithmic for both cities. Seattle's homeless population seems to increase more rapidly and to a greater extent. This rate is especially rapid when Seattle's homeless population is predicted with population instead of mortgage rate, appearing almost linear. Meanwhile, with the same variables, Albuquerque's homeless population growth remains logarithmic, approaching an asymptotic limit past the 50-year mark. The trends are expected to differ, as Seattle and Albuquerque, though both populous cities have vastly different environments. It also reflects upon the existing data on homeless populations - compared to Albuquerque, Seattle has a substantially greater ratio of homeless people to its total population.

2.7 Strengths & Weaknesses

As a multivariate time series model, vector autoregression models are well suited to modeling the movements of various time-series data simultaneously while also taking into account feedback between variables in the model[16]. VAR models are known to be effective in modeling and forecasting the dynamics of macroeconomic time series and are regarded as efficient alternatives to the traditional system involving several equations[17]. Within its prediction, VAR models can also provide insight into the causal relationships between variables aiding in understanding the strength and direction of interactions[18].

However, the utility of VAR becomes limited with higher dimensionality. This is due to the increase in the number of parameters in the VAR model, which can lead to increased uncertainty. Additionally, as with any time series model, VAR models require stationarity and further assume a correlation between the endogenous variables. In these cases, further steps would be needed to bypass these restrictions, potentially leading to situations more complicated than needed. For the purposes of this model, we also made the critical assumption that the variables used were correlated, while we only had evidence of causation. VAR was also extremely sensitive to what parameters are used, so given the limited parameters that we chose to investigate, a more accurate representation can likely be obtained with a more diverse set of variables. Given more time, we would have confirmed the correlation of the variables used, as well as explored other variables that could've been integrated. Outside of time series, we may have also considered econometric models that can provide better insight into the economic trends used.

Q3: Rising from This Abyss

3.1 Defining the Problem

The third problem asks us to create a model that helps a city determine a long-term plan to address homelessness. Our previous results exhibit a higher reliability for Seattle, so we will be focusing on Seattle.

3.2 Assumptions

3-1 **Migrants purchase vacant housing when moving into their city. Those who migrate, but are unable to attain housing, will be considered as part of the homeless population.**

- **Justification**: This is an extension of G-7, as migrants unable to purchase housing in their new area lack a residence.

3-2 **The price to build one unit of affordable housing is constant across a given region.**

- **Justification**: This assumption is created to address volatility within the construction industry and variation of development costs across Seattle.

3-3 **Sheltered and unsheltered homeless individuals are both considered homeless, as they neither own nor rent a property.**

- **Justification**: The U.S. Department of Housing and Urban Development regards both sheltered and unsheltered homeless individuals as homeless[19].

3-4 **We assume a constant price of $373,285 to build one affordable housing unit in Seattle.**

- **Justification**: This price is based on data from a new study in 2018. We adjusted the production cost to account for inflation and the 2024 purchasing power[20].

3-5 **During economic recessions, Seattle's budget and government spending will significantly decrease for affordable housing projects.**

- **Justification**: While government spending has historically increased during economic turmoil to support citizens, most funds were taken from specific programs and spent towards overall well-being (i.e. stimulus packages during COVID-19).

3-6 **During natural disasters, Seattle's affordable housing units and total housing units will decrease due to damages.**

- **Justification**: Natural disasters will render a proportion of housing units unviable for safe living. Thus, we will consider this proportion of housing units lost during a natural disaster.

3.3 Variables

Symbol	Definition	Units
B	Final Budget	USD
$\mathbf{B_g}$	Given Budget = $339,000,000	USD
D	Admissible homeless population	People
T	Timeline of suggested development plan	Years
$\mathbf{A_i}$	Affordable housing units at index *i*	Housing units
$\mathbf{A_{i,adj}}$	Adjusted number of affordable housing units at index *i* due to natural disasters	Housing units
$\mathbf{P_j}$	Pricing increase at index *j*	Percentage
I	Average income of a homeless person in Seattle	USD
ER_loss	Percent of budget lost due to economic recession	Percentage
ND_loss	Percent of budget lost due to natural disaster	Percentage
n	Length of the list of affordable housing units	Units
m	Length of the list of pricing increases	Units
$\mathbf{N_i}$	Number of new migrants	People
$\mathbf{V_u}$	Vacant housing units	Housing units

Table 9: Variables for Problem 2

3.4 The Model

Our model uses several factors to determine the smallest percentage of Seattle's housing budget needed to build new affordable housing units each year so that all homeless people above a given threshold can gain affordable housing in a given timeline.

$$B = \left(B_{\mathrm{g}} + \sum_{i=0}^{n-1} (A_i \cdot P_{\min(i,m-1)} \cdot I) \right) \times (1 - \mathrm{ER\_loss})$$

Equation 8. Iterative budget formula

Using Equation 8, we determined the budget available to build affordable houses for each year. The budget increases due to payments by new tenants on their affordable housing plans and a generally stronger economy due to reduced homelessness[23]. The tenants will start by paying 5% of their income and will progressively increase their payment by 5% until it reaches 25% of their income. Afterward, their payment increases by 1% until it reaches 30% of their income and stabilizes. This is our affordable housing pricing model. These additional rent payments are added to the given budget for the next year.

The budget will also decrease if there is an economic recession, which illustrates our model's adaptability to economic changes.

$$\mathrm{homeless\_immigrants} = (N_i - V_u) \times (V_u < N_i)$$

Equation 9. Homeless immigrants formula

$$A_{i,\mathrm{adj}} = \mathrm{round}(A_i \times (1 - \mathrm{ND\_loss}))$$

Equation 10. Natural Disaster Adjustment formula for Affordable Housing Supply

As showcased in Equation 9, our model accounts for migration influx by providing migrants with all available vacant housing. If there are more migrants than vacant housing units, the remaining migrants are added to the homeless population. Furthermore, Equation 10 demonstrates how our model considers natural disasters by decreasing the number of affordable houses and total houses built using ND_loss.

This method repeatedly iterates through each year, finding the number of affordable houses that can be built from the designated budget each year. Using this process, we can determine if Seattle is able to stay within the minimum threshold of homeless individuals within our timeline budget. If we are successful, we can reduce our budget allocation percentage and/or our timeline duration and repeat the process to optimize our model suggestions.

3.5 Results

To simulate our novel affordable housing optimization model, we tested our model. We considered 1,000 admissible homeless individuals across an ideal timeline of 15 years. The model produced the following plan for the next 15 years.

Year	New Affordable Housing Units	Estimated Development Costs (in millions)
2023	908	$338.94
2024	913	$340.81
2025	923	$344.54
2026	939	$350.51
2027	960	$358.35
2028	986	$368.06
2029	1014	$378.51
2030	1044	$389.71
2031	1075	$401.28
2032	1109	$413.97
2033	1144	$427.04
2034	1180	$440.48
2035	1217	$454.29
2036	1256	$468.85
2037	1295	$483.40

Table 10: Estimated development costs over the next 15 years

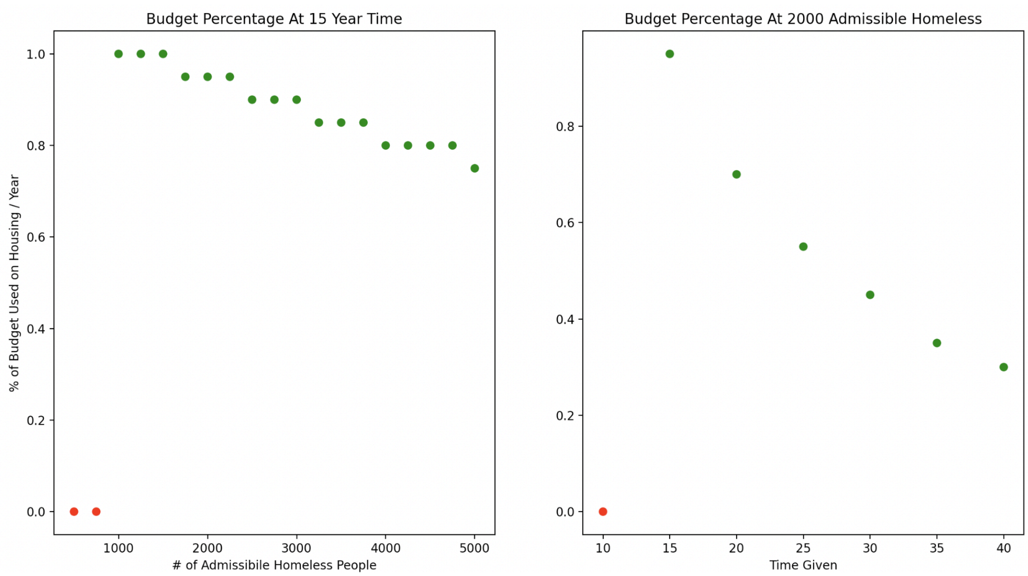

Figure 6. Effect of Admissible Homeless People and Timeline Length on minimum budget necessary to meet affordable housing goals.

3.6 Discussion

The number of houses our model is suggesting to build each year increases because of the income gained from affordable housing tenants and holistic economic benefits drawn from reduced homelessness[23]. The total number of affordable houses we were able to build using 100% of our available budget was just sufficient to meet our minimum threshold. When we tried to use 95% of our budget, we were unable to meet the minimum threshold of remaining homeless individuals without affordable housing plans in the given 15-year timeline. Therefore, with the given initial conditions, we must use 100% of the budget.

Figure 6 shows that as we increase the number of admissible homeless people in our model and increase the time given we can allocate less of our available yearly affordable housing budget to building new affordable housing units. This correlates with our market intuition.

3.7 Sensitivity Analysis

To test if our model could account for economic recessions, natural disasters, and migrant influx, we set up temporary datasets at an 800-homeless-person threshold and a 20-year development period. The dataset without perturbation succeeded with a 75% budget usage. Our first dataset had a natural disaster in the year 2028 causing a 40% loss to the number of available affordable housing units and the number of total housing units in Seattle. Our model found that we needed an 85% budget usage per year to successfully implement our plan. We then did the same with an economic recession in 2028, causing a 40% loss to our annual budget, resulting in an 80% budget usage for success. Our last dataset had a 30,000 migrant influx in 2028, and our model found that we needed to use 95% of our budget to implement the affordable housing plan. This analysis showed that our optimization model can holistically account for unforeseen events.

3.8 Strengths & Weaknesses

One of the strongest qualities of this model is its inherent reproducibility across several cities given the proper data. Our model is reproducible with a myriad of starting conditions, making it incredibly robust and adaptable to various housing scenarios, and will help housing and urban development planners weigh the different approaches to their policy agenda. Finding specific metrics that would be affected by external factors allowed our model to make conservative assessments that prevent the overallocation of the city's budget.

However, we only considered three unforeseen circumstances on affordable housing development (economic recessions, natural disasters, migrant influx) which leaves out several other factors that may influence affordable housing development. On the other hand, given the simplicity of our model, policymakers can streamline the process to incorporate more variables such as interest rate changes and specific housing types (apartments, condos, townhouses, etc.), making our model incredibly flexible.

Conclusion

In the first question, we used logistic regression to predict changes in the housing supply of Seattle and Albuquerque. Over 10, 20, and 50 years, our model estimates 451,164; 489,908; and 529,251 housing units in Seattle and 312,107; 339,504; and 368,391 housing units in Albuquerque, respectively. Meanwhile, in the second question, we quantified changes in the homeless populations of Seattle and Albuquerque with a vector autoregressive model. Using mortgage rates, the ratio of income needed for housing finances, and total population, the model projects an increase in the homeless population for both Seattle and Albuquerque. Compared to their respective homeless populations in 2022, Seattle's homeless population is forecasted to increase 1cbn7.2% in 10 years, 32.6% in 20 years, and 63.2% in 50 years, while Albuquerque's homeless population is expected to increase 46.8% in 10 years, 52.4% in 20 years, and 54.3% in 50 years. Finally, to assist city planners in planning policy to address homelessness, we developed a novel affordable housing optimization model that can address the homeless crisis and housing shortages together, providing valuable insights into budgetary allocation for specific policy choices to curb the homelessness problem at its root. We found that a 20-year affordable housing program with a minimum threshold of 800 homeless remaining can be done using only 75% of the annual urban development budget and would still be easily implemented after major economic recessions, natural disasters, and migrant influx.

As the housing crisis continues to escalate, so too does the issue of homelessness persist. Our findings show that this issue has the potential to rise to unprecedented heights in the near future - yet at the same time, so too can guided policy and decision-making help create benefits that may last for decades. Truly, in the worst of times, the best of the times may soon follow.

References

[1] Homelessness in America reaches record level amid rising rents and end of COVID aid
[2] Not One More: Honoring Those Who Died Homeless
[3] The Impact of Homelessness on Economic Competitiveness
[4] How Climate Change Impacts Homelessness
[5] How Housing Costs Drive Levels of Homelessness
[6] What is the 30/30/3 Rule for Buying a Home in Uncertain Times?
[7] The influence of mortgage rates on property prices
[8] VARX-L: Structured regularization for large vector autoregressions with exogenous variables
[9] Vector Autoregressions
[10] https://smartasset.com/mortgage/washington-mortgage-rates
[11] https://smartasset.com/mortgage/new-mexico-mortgage-rates
[12] https://www.zillow.com/mortgage-rates/wa/
[13] https://www.zillow.com/mortgage-rates/nm/
[14] https://www.bankrate.com/mortgages/mortgage-calculator/
[15] What's the Average Down Payment on a House?
[16] Vector Autoregression (VAR) Models
[17] Chapter 3: Vector Autoregressive Methods¶
[18] Vector Autoregressive with Exogenous Variable Model and its Application in Modeling and Forecasting Energy Data: Case Study of PTBA and HRUM Energy
[19] A Guide to Counting Unsheltered Homeless People
[20]Building affordable housing in Seattle isn't cheap
[21] A Tale of Two Crises, MathWorks Math Modeling Challenge 2024, curated data, https://m3challenge.siam.org/kdfrldh/
[22] The Ample Reserves Framework and Balance Sheet Reduction: Perspective from the Open Market Desk
[23]The Business Case for Ending Homelessness

Code Appendix

Q1: It Was the Best of Times

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import curve_fit

def logistic_function(x, L, k, x0):
    return L / (1 + np.exp(-k * (x - x0)))

"""With Confidence Interval

Seattle Graph
"""

# Past Number of Housing Units in Seattle
years = np.arange(2010, 2022)  # X axis data from 2010 to 2021. Not
including 2022 in input data so we can use it to test our model.
num_houses_list = [302465, 304164, 306694, 309205, 311286, 315950, 322795,
334739, 344503, 354475, 367337, 362809]
carrying_capacity = 535249

# Curve fitting
popt, pcov = curve_fit(logistic_function, years, num_houses_list,
bounds=([0, 0, years[0]], [carrying_capacity, 0.07, years[-1]]))

# Projection
generations = 60
future_years = np.arange(2021, 2023 + generations)
fitted_population = logistic_function(future_years, *popt)

# Confidence intervals
allowed_error = 0.025
se = np.sqrt(np.diag(pcov)) #standard error array
margin_of_error = allowed_error * se
lower_bound = fitted_population - margin_of_error[0]
upper_bound = fitted_population + margin_of_error[0]

# Plotting
plt.figure(figsize=(10, 6))
```

```
plt.plot(years, num_houses_list, 'bo', label='Past Total Housing Units')
plt.plot(future_years, fitted_population, 'r-', label='Projected Growth of
Housing Units')
plt.fill_between(future_years, lower_bound, upper_bound, color='gray',
alpha=0.2, label = "95% Confidence Interval")
plt.axhline(y=carrying_capacity, color='g', linestyle='--',
label='Carrying Capacity')
plt.title('Projected Growth of Housing Units in Seattle, Washington with
Confidence Interval')
plt.xlabel('Year')
plt.ylabel('Number of Housing Units')
plt.legend()
plt.grid(True)
plt.show()

#Sensitivity Analysis
predicted_2022_value = fitted_population[1]
actual_2022_value = 372436

percent_error =
(predicted_2022_value-actual_2022_value)*100/actual_2022_value
print(predicted_2022_value, actual_2022_value)
print("Percent Error of Seattle 2022 Housing Unit Prediction:",
percent_error, "%")

predicted_2034_value = fitted_population[13]
predicted_2044_value = fitted_population[23]
predicted_2074_value = fitted_population[53]
print("Predicted Total Housing Units in 2034:", predicted_2034_value)
print(lower_bound[13], upper_bound[13])
print("Predicted Total Housing Units in 2044:", predicted_2044_value)
print(lower_bound[23], upper_bound[23])
print("Predicted Total Housing Units in 2074:", predicted_2074_value)
print(lower_bound[53], upper_bound[53])

"""Albuquerque"""

#Albuquerque, NM
# Provided data points
```

```
years = np.arange(2010, 2022)  # Assuming the first year is 2010 and goes
up to 2022
num_houses_list = [234891, 237735, 239718, 240277, 240961, 241326, 242070,
243402, 244382, 245476, 247926, 252924]
# carrying_capacity = 252924 + 120212
carrying_capacity = 373136

# Curve fitting
popt, pcov = curve_fit(logistic_function, years, num_houses_list,
bounds=([0, 0, years[0]], [carrying_capacity, 0.068, years[-1]]))

# Projection
generations = 60
future_years = np.arange(2021, 2023 + generations)
fitted_population = logistic_function(future_years, *popt)

# Confidence intervals
allowed_error = 0.025
se = np.sqrt(np.diag(pcov)) #standard error array
margin_of_error = allowed_error * se
lower_bound = fitted_population - margin_of_error[0]
upper_bound = fitted_population + margin_of_error[0]

# Plotting
plt.figure(figsize=(10, 6))
plt.plot(years, num_houses_list, 'bo', label='Past Total Housing Units')
plt.plot(future_years, fitted_population, 'r-', label='Projected Growth of
Housing Units')
plt.fill_between(future_years, lower_bound, upper_bound, color='gray',
alpha=0.2, label = "95% Confidence Interval")
plt.axhline(y=carrying_capacity, color='g', linestyle='--',
label='Carrying Capacity')
plt.title('Projected Growth of Housing Units in Albuquerque, New Mexico
with Confidence Interval')
plt.xlabel('Year')
plt.ylabel('Number of Housing Units')
plt.legend()
plt.grid(True)
plt.show()
```

```
#Sensitivity Analysis
predicted_2022_value = fitted_population[1]
actual_2022_value = 255178

percent_error =
(predicted_2022_value-actual_2022_value)*100/actual_2022_value
print(predicted_2022_value, actual_2022_value)
print("Percent Error of Albuquerque 2022 Housing Unit Prediction:",
percent_error, "%")

predicted_2034_value = fitted_population[13]
predicted_2044_value = fitted_population[23]
predicted_2074_value = fitted_population[53]
print("Predicted Total Housing Units in 2034:", predicted_2034_value)
print(lower_bound[13], upper_bound[13])
print("Predicted Total Housing Units in 2044:", predicted_2044_value)
print(lower_bound[23], upper_bound[23])
print("Predicted Total Housing Units in 2074:", predicted_2074_value)
print(lower_bound[53], upper_bound[53])
```

Q2: It Was the Worst of Times

Forecasting for Seattle:

```
#Import libraries
import pandas as pd
from statsmodels.tsa.api import VAR
import numpy as np
import matplotlib.pyplot as plt

#Read in all neccesary datasets
home_seattle = pd.read_csv('homelessness_seattle.csv')
price_seattle = pd.read_csv('median_price_seattle.csv')
house_seattle = pd.read_csv('housing_seattle2.csv')
pop_seattle = pd.read_csv('population_seattle1.csv')

df = pd.DataFrame(pop_seattle["Year"])
df["Total Homeless"] = list(map(float,home_seattle["Homeless Total"]))
df["Ratio of Required Income for House Financing to Actual Income"] =
price_seattle["Compare to 30%"]
df['Mortgage Rate'] = price_seattle["Mortgage Rate"]
```

```
df["Population"] = pop_seattle["Total Population"]

print(df)

#endogenous variables used for the VAR model; 'Mortgage Rate' can be
interchanged with 'Population'
endog = df[["Total Homeless",'Ratio of Required Income for House Financing
to Actual Income','Mortgage Rate']]

#Test for stationarity with ADF
from statsmodels.tsa.stattools import adfuller

def adf_test(series, name=''):
    result = adfuller(series, autolag='AIC')
    print(f'ADF Test for {name}:')
    print('Test Statistic:', result[0])
    print('p-value:', result[1])
    print('Critical Values:', result[4])
    print('Is the time series stationary?', 'No (reject null hypothesis)'
if result[1] < 0.05 else 'Yes (fail to reject null hypothesis)')
    print('\n')

for column in endog.columns:
    adf_test(endog[column], name=column)

#Test for cointegration with Engle-Granger
import statsmodels.api as sm

y = df['Total Homeless']
X = df[['Ratio of Required Income for House Financing to Actual
Income','Mortgage Rate']]
X = sm.add_constant(X)

model = sm.OLS(y, X)
results = model.fit()

residuals = results.resid
adf_test(residuals, name='Residuals')

#Implement the VAR
```

```
model = VAR(endog)
model_fit = model.fit()

lag_order = model_fit.k_ar
pred = model_fit.forecast(endog.values[-lag_order:], 50)
model_fit.plot_forecast(50)
#print(f"Predicted: {pred}")

forecasted_pop = pred[:, 0]
print(f"Forecasted : {forecasted_pop[-1]}")
forecasted_years = np.arange(df["Year"].max()+1, df["Year"].max() + 51)

#Plot the original homeless population along with the forecasted values
plt.figure(figsize=(10, 6))
plt.plot(df["Year"], df["Total Homeless"], label="Homeless Population
Actual")
plt.plot(forecasted_years, forecasted_pop, label="Forecasted Homeless
Population")
plt.xlabel("Year")
plt.ylabel("Homeless Population")
plt.title("Forecasted Homeless Population for Next 50 Years")
plt.legend()
plt.grid(True)
plt.show()

#Save forecasted data into csv
forecast_df = pd.DataFrame({
    "Year": forecasted_years,
    "Total Homeless": forecasted_pop,
    'Ratio of Required Income for House Financing to Actual Income':
pred[:, 1],
    'Mortgage Rate': pred[:, 2]
})

forecast_df["Total Homeless"] = forecast_df["Total Homeless"].astype(int)

#print(forecast_df)
forecast_df.to_csv('forecasted_data.csv', index=False)
```

Forecasting for Albuquerque:

```
import pandas as pd
from statsmodels.tsa.api import VAR
import numpy as np
import matplotlib.pyplot as plt

mortgage_raw = pd.read_csv('mortgage_rate.csv')
mortgage_adj = mortgage_raw.iloc[::12].reset_index()

home_alb = pd.read_csv('homelessness_alb.csv')
price_alb = pd.read_csv('price_alb.csv')
pop_alb = pd.read_csv('population_alb.csv')

df = pd.DataFrame(home_alb["Year"])
df["Total Homeless"] = list(map(float,home_alb["Homeless Total"]))
df["Ratio of Required Income for House Financing to Actual Income"] =
price_alb["Compare to 30%"]
df['Mortgage Rate'] = price_alb["Mortgage Rate"]    #Note: was faulty, so
it is not used in the final version
df['Population'] = pop_alb['Total Population']

print(df)


endog = df[["Total Homeless",'Ratio of Required Income for House Finacing
to Actual Income','Population']]

#test for stationarity
from statsmodels.tsa.stattools import adfuller

def adf_test(series, name=''):
    result = adfuller(series, autolag='AIC')
    print(f'ADF Test for {name}:')
    print('Test Statistic:', result[0])
    print('p-value:', result[1])
    print('Critical Values:', result[4])
    print('Is the time series stationary?', 'No (reject null hypothesis)'
if result[1] < 0.05 else 'Yes (fail to reject null hypothesis)')
    print('\n')
```

```
for column in endog.columns:
    adf_test(endog[column], name=column)

#Test for cointegration
import statsmodels.api as sm

y = df['Total Homeless']
X = df[['Ratio of Required Income for House Financing to Actual
Income','Population']]
X = sm.add_constant(X)

model = sm.OLS(y, X)
results = model.fit()

residuals = results.resid
adf_test(residuals, name='Residuals')

#Implement the VAR model
model = VAR(endog)
model_fit = model.fit()

lag_order = model_fit.k_ar
pred = model_fit.forecast(endog.values[-lag_order:], 50)
model_fit.plot_forecast(50)

print(f"Predicted: {pred}")

forecasted_pop = pred[:, 0]  # Assuming housing units is the first
variable in your endogenous variables
print(f"Forecasted : {forecasted_pop[-1]}")
forecasted_years = np.arange(df["Year"].max()+1, df["Year"].max() + 51)

# Plot the original homeless population data along with the forecasted
values
plt.figure(figsize=(10, 6))
plt.plot(df["Year"], df["Total Homeless"], label="Homeless Population
Actual")
plt.plot(forecasted_years, forecasted_pop, label="Forecasted Homeless
Population")
plt.xlabel("Year")
```

```
plt.ylabel("Homeless Population")
plt.title("Forecasted Homeless Population for Next 50 Years")
plt.legend()
plt.grid(True)
plt.show()
```

Q3: Rising from This Abyss

```
import pandas as pd
import numpy as np
import math
import matplotlib.pyplot as plt

pricing_increase = [0.05,0.1,0.15,0.2,0.25,0.26,0.27,0.28,0.29,0.3]
carrying_capacity = 535249

def
one_year(affordable_housing,current_homeless,built_houses,vacant_units, #
MODEL INPUT VALUES
            ND_loss,ER_loss,new_immigrants, # STRENGTH TESTING
            const_price=373285,given_budget=339000000,income=42635): #
DEFAULT VALUES
   budget = (given_budget +
sum(affordable_housing[i]*pricing_increase[min(i,len(pricing_increase)-1)]
*income for i in range(len(affordable_housing)))) * (1-ER_loss)

   temp_housing = [round(i * (1-ND_loss)) for i in affordable_housing]
   affordable_housing = temp_housing
   built_houses *= (1-ND_loss)
   homeless_immigrants = (new_immigrants-vacant_units) * (vacant_units <
new_immigrants)
   new_housing = budget//const_price
   if(new_housing > carrying_capacity - built_houses):
       return [],0,False
   affordable_housing.insert(0,new_housing)
   return affordable_housing,homeless_immigrants,True


def simulate_housing(ideal_homeless,ideal_time, # TOOL SPECIFIC ANALYSIS
                    homeless_data,built_houses_data,vacant_units_data, #
MODEL INPUT VALUES
                    ND_loss_data,ER_loss_data,new_immigrants_data, #
STRENGTH TESTING
```

```
const_price=373285,given_budget=339000000,income=42635): # DEFAULT VALUES
   all_housing = []
   homeless_immigrants = 0
   for i in range(0,ideal_time):
       all_housing,new_homeless_immigrants,success =
one_year(all_housing,homeless_data[i]-sum(all_housing)+homeless_immigrants
,built_houses_data[i],

vacant_units_data[i],ND_loss_data[i],ER_loss_data[i],new_immigrants_data[i
],

const_price,given_budget,income)
       homeless_immigrants += new_homeless_immigrants
       if(not success):
           return [],0,False
       if(sum(all_housing) + ideal_homeless >
homeless_data[i]+homeless_immigrants):
           return
all_housing,homeless_data[i]+homeless_immigrants-sum(all_housing),True
   return [],0,False


def
optimal_housing(ideal_homeless,ideal_time,homeless_data,built_houses_data,
vacant_units_data, ND_loss_data,ER_loss_data,new_immigrants_data):
   print()
   print("Admissible # of Homeless People:",ideal_homeless)
   print("Ideal Time:",ideal_time)
   print()
   print("Calculating...")
   print()
   optimal_plan = []
   curr_homeless = 0
   budget_percent = 1.05
   success = True
```

```
   while(success):
       budget_percent -= 0.05
       previous_plan = optimal_plan
       prev_homeless = curr_homeless
       optimal_plan,curr_homeless,success =
simulate_housing(ideal_homeless,ideal_time,
                    homeless_data,built_houses_data,vacant_units_data,

ND_loss_data,ER_loss_data,new_immigrants_data,given_budget=339000000*budge
t_percent)
   if(not previous_plan):
       print("Not possible in the given amount of time!")
       print("To Fix: Increase Homeless Threshold OR Increase Time")
       return 0
   print("Success!")
   print("You can spend",str(math.ceil((budget_percent+0.05)*100))+"% of
your urban development budget on affordable housing")
   print()
   for i in range(1,len(previous_plan)+1):
       print(str(2022+i)+":",int(previous_plan[-i])," new affordable
housing units")
   print()
   print("People in Affordable Housing:",int(sum(previous_plan)))
   print("Homeless after",ideal_time,"years:",int(prev_homeless),"people")
   print()
   return budget_percent+0.05


homeless_data = [13731, 13656, 13778, 14149, 14505, 14741, 14938, 15172,
15427, 15668, 15889, 16105, 16323, 16538, 16747, 16949, 17148, 17343,
17535, 17722, 17905, 18084, 18259, 18431, 18599, 18763, 18924, 19082,
19236, 19388, 19535, 19680, 19822, 19961, 20097, 20229, 20360, 20487,
20612, 20734, 20853, 20970, 21085, 21197, 21307, 21414, 21519, 21622,
21723, 21822]
```

```
built_houses_data = [381633, 389184, 396499, 403572, 410397, 416973,
423297, 429368, 435188, 440759, 446083, 451164, 456007, 460617, 465001,
469163, 473113, 476855, 480398, 483750, 486917, 489908, 492730, 495390,
497897, 500257, 502478, 504566, 506529, 508373, 510105, 511730, 513255,
514685, 516025, 517281, 518458, 519560, 520591, 521557, 522460, 523306,
524096, 524836, 525527, 526173, 526777, 527341, 527868, 528361, 528821,
529251, 529652, 530027, 530377, 530703, 531008, 531293, 531558, 531806]
vacant_units_data = [23414, 23761, 24108, 24455, 24802, 25149, 25496,
25843, 26190, 26537, 26884, 27231, 27578, 27925, 28272, 28619, 28966,
29313, 29660, 30007, 30354, 30701, 31048, 31395, 31742, 32089, 32436,
32783, 33130, 33477, 33824, 34171, 34518, 34865, 35212, 35559, 35906,
36253, 36600, 36947, 37294, 37641, 37988, 38335, 38682, 39029, 39376,
39723, 40070, 40417, 40764, 41111, 41458, 41805, 42152, 42499, 42846,
43193, 43540, 43887]

print()
print("No Changes")
optimal_housing(800,20,homeless_data,built_houses_data,vacant_units_data,[
0]*60,[0]*60,[0]*60)
print()
print("Natural Disaster (40% loss) after 5 years")
optimal_housing(800,20,homeless_data,built_houses_data,vacant_units_data,[
0]*5+[0.4]+[0]*50,[0]*60,[0]*60)
print()
print("Economic Recession (40% loss) after 5 years")
optimal_housing(800,20,homeless_data,built_houses_data,vacant_units_data,[
0]*60,[0]*5+[0.4]+[0]*50,[0]*60)
print()
print("Migrant Influx (30000 people) after 5 years")
optimal_housing(800,20,homeless_data,built_houses_data,vacant_units_data,[
0]*60,[0]*60,[0]*5+[30000]+[0]*50)

x1 = []
y1 = []
c1 = []
x2 = []
```

```
y2 = []
c2 = []
for i in range(500,5250,250):
   x1.append(i)

y1.append(optimal_housing(i,15,homeless_data,built_houses_data,vacant_unit
s_data,[0]*60,[0]*60,[0]*60))
   if(y1[-1]):
       c1.append('g')
   else:
       c1.append('r')

for i in range(10,45,5):
   x2.append(i)

y2.append(optimal_housing(2000,i,homeless_data,built_houses_data,vacant_un
its_data,[0]*60,[0]*60,[0]*60))
   if(y2[-1]):
       c2.append('g')
   else:
       c2.append('r')

figure, axis = plt.subplots(1, 2)
axis[0].scatter(x1,y1,c=c1)
axis[0].set_title("Budget Percentage At 15 Year Time")
axis[0].set_ylabel("% of Budget Used on Housing / Year")
axis[0].set_xlabel("# of Admissibile Homeless People")
axis[1].scatter(x2,y2,c=c2)
axis[1].set_title("Budget Percentage At 2000 Admissible Homeless")
axis[1].set_xlabel("Time Given")
plt.show()
```