Executive Summary

To the Secretary of the U.S. Department of Housing,

Why is it illegal to sleep?

For every one of the 580,000 Americans experiencing homelessness, there are 28 vacant homes in the U.S.[1] Existing programs such as the Title V program of the McKinney–Vento Homeless Assistance Act of 1987 aim to alleviate this striking inequality, but they evidently have not been effective. Such a persistent issue affecting hundreds of thousands of people demands immediate and meaningful action.

Our team has analyzed the state of the housing market and homelessness in the U.S. today and leveraged this analysis to project future trends. Then, using these data, we arrived at a plan to reduce homelessness in the context of rising home prices and homelessness.

To start our model, we developed predictions for the housing market in two U.S. cities: Seattle, Washington, and Albuquerque, New Mexico. We started by taking historical data on three measurements over the past 12 years: total housing units, occupied units, and vacant units. Then, we found the average year-over-year rate of change for each measurement and the standard deviation. Using this and normal distribution curves, we graphed our predictions for each measurement over three intervals: the next 10 years, the next 20 years, and the next 50 years.

Our results showed that these cities are growing rapidly, with Seattle in 50 years potentially nearing 1 million household units and the ability to support 18 million people, and Albuquerque nearly reaching 500,000 household units. External factors support our results: With climate change and rapid global warming, areas further North such as Seattle are seeing increased interest, along with an exponentially expanding population creating a surge in the need for housing.

Then, we developed a model to project homelessness in both of these American cities for the next 10, 20, and 50 years. We utilized a regression technique, ARIMA, to match the data published by the U.S. Department of Housing and Urban Development. We used Matlab Econometrics Toolbox features such as the augmented Dickey-Fuller test, autocorrelation functions and partial autocorrelation functions. This allowed us to approximate parameters for our ARIMA regression in order to project future homelessness numbers for each city. In Seattle, we projected 14,760 homeless people in 10 years, 16,180 homeless in 20 years, and 20,400 homeless in 50 years, building on the current 13,300 homeless individuals at the most recent census. In Albuquerque, we projected 1,330 homeless individuals in 10 years, 1,390 homeless in 20 years, and 1,550 homeless in 50 years, building on the current 1,277 homeless people at the recent census. These predictions highlight how current census data implies that homelessness is on the rise in the future, making it a higher priority concern for city officials.

Finally, we developed a plan to tackle the issue of homelessness in our chosen city, Seattle. We chose to create a plan for Seattle due to the results we found in our first model: Seattle has the potential to grow into an extremely large city, with the ability to support millions of people. That being said, we also predicted a large margin of error, meaning that the total number of housing units could also decrease and Seattle could become a dying city. Thus, we must ensure the homeless in Seattle are housed.

Our methodology consisted of determining potential factors for Seattle's homelessness issues and comparing them to national averages. Each of our 7 factors had a designated solution: quantity/quality of jobs and job creation, economic assistance and increasing housing, lack of healthcare and increasing healthcare, high domestic violence rates and increased domestic violence reduction programs, racial inequality and equitable housing, homes sitting on the market and rapid rehousing programs, and lack of affordable housing and increased low-income housing. We then developed scores for each category based on chi-squared equations to determine which solutions needed to be prioritized. We found that the most important issues Seattle needs to focus on are reducing racial inequality followed by increasing housing affordability.

Thus, we recommend that Seattle focus on reducing racial inequality by increasing equitable housing, followed by increasing housing affordability by increasing the amount of low-income housing available. We hope that our methodology is useful to policymakers in other urban areas as well.

Contents

1 Q1: It Was the Best of Times

1.1 Defining the Problem

We need to model the housing supply in Seattle, Washington, and Albuquerque, New Mexico, over the next 10, 20, and 50 years. We used previous housing market data to make these predictions.

1.2 Variables

Table 1. Variable Definitions

Variable	Definition
R	Rate of housing supply
R_{rv}	Rate of rental vacancy
R_{hv}	Rate of homeowner vacancy
a	Initial quantity in 2010
b	Average Year over Year rate of change
k	Standard deviation adjustment factor
n	Number of standard deviations
σ	Standard deviation
x	Number of years after 2010
$f_1(x)$	Predicted Value
$g_1(x)$	1 Standard Deviation (68% Confidence Interval & Purple Region)
$h_1(x)$	2 Standard Deviations (95% Confidence Interval & Orange Region)
M	The multiplier we used to multiply unit by persons per unit
$M * P/H$	Multiplier * People per house; This tells us the rate at which to multiply the total house projections

1.3 Assumptions

1. **The total rate of housing supply is the sum of the rates of rental and homeowner vacancy ($R = R_{rv} + R_{hv}$).**

   Justification: Most vacant homes are either rentals or owned by a homeowner, and abandoned homes likely are not reasonably inhabitable. This assumption is also convenient because abandoned homes are usually not represented in official data.

2. **Housing supply is measured in three variables: total housing units, occupied units, and vacant units.**

   Justification: Total housing supply defines the overall supply in the specified region, while occupied/vacant units measure the proportion of the total housing supply that is in use, along with the float size.

3. **Each housing unit is a household, and each household is a housing unit.**

Justification: This allows us to use data from the U.S. Census which details the average persons per household: 2.05 in Seattle and 2.32 in Albuquerque.[2][3] Through this assumption, we can make stronger projections.

4. **The proportion of unit types of households is constant within vacant households.**

   Justification: This allows us to use the proportion of distribution of the households with different numbers of units as dictated by the data given by the M3 Challenge. This allows us to apply the proportion to the number of vacant houses and give data that is better understood by the user.

5. **Housing market trends consist of two types: region-specific and non-region-specific.**

   Justification: Housing fluctuations depend on the region and the local market while global recessions affect all levels of the economy, although ground-level effects can vary slightly by region.

6. **Housing market trends consist of two types: region-specific and non-region-specific.**

   Justification: Housing fluctuations depend on the region and the local market while global recessions affect all levels of the economy, although ground-level effects can vary slightly by region.

7. **There is no maximum number of housing units that can be developed.**

   Justification: If there were to be a maximum quantity of housing units in either metropolitan area, our model could be easily adapted by adding an upper bound.

1.4 Model 1: Housing Projections

1.4.1 Developing our Model

We started by attempting to model household projections as our team processed the data regarding total housing units, occupied units, and vacant units for each of Seattle and Albuquerque.[4] To make future housing supply predictions, we found the average rate of change per year for each measurement for each year, and the standard deviation. We then used the average rate of change to create exponential equations that would represent each measurement over time.

1.4.2 Executing our Model

Equations for Total Housing Units (Albuquerque)

$$f_1(x) = 234891(1.006940195)^x$$

$$g_1(x) = 234891(1.006940195 \pm 0.005355323366)^x$$

$$h_1(x) = 234891(1.006940195 \pm 2 \times 0.005355323366)^x$$

Equations for Total Occupied Units (Albuquerque)

$$f_1(x) = 217256(1.008312102)^x$$

$$g_1(x) = 217256(1.008312102 \pm (0.31125)0.01058883438)^x$$

$$h_1(x) = 217256(1.008312102 \pm 2 \times (0.31125)0.01058883438)^x$$

Equations for Total Vacant Units (Albuquerque)

$$f_1(x) = 17635(0.9909189477)^x$$

$$g_1(x) = 17635(0.9909189477 \pm (0.31125)0.06918158605)^x$$

$$h_1(x) = 17635(0.9909189477 \pm 2 \times (0.31125)0.06918158605)^x$$

After completing a similar process for Seattle, we gained the projections for each city.

Table 2. Housing Projections based on Standard Deviations

	10-Year Estimate	20-Year Estimate	50-Year Estimate
Seattle, WA (Total housing units)	459,673 (225,870-650,705)	547,252 (199,994-895,391)	923,431 (138,834-5,816,310)
Seattle, WA (Occupied units)	426,172 (336,746-538,113)	507,346 (363,417-705,985)	855,974 (456,783-1,594,268)
Seattle, WA (Vacant units)	34,697 (15,108-77,496)	41,941 (12,915-130,930)	74,076 (8,069-631,420)
Albuquerque, NM (Total housing units)	277,303 (214,532-357,468)	297,161 (206,579-425,819)	365,690 (184,446-719,763)
Albuquerque, NM (Occupied units)	265,003 (226,407-309,861)	287,873 (230,333-359,265)	369,020 (242,523-559,964)
Albuquerque, NM (Vacant units)	14,167 (4,877-39,329)	12,932 (2,855-54,395)	9,836 (572-149,718)

From the Housing Numbers in 2022, we calculated the proportions of each unit-type. Using that proportion, we could calculate a reliable housing number per unit based on our projections in Table 2. We could then use the Multiplier * Persons per Household as defined by the U.S. Consensus in order to project the number of people that can be supported by housing in 10, 20, and 50 years.

Table 3. Albuquerque Conversion Rates (2022)

	Housing Numbers	Percent	M	M*P/H	10 year	20 year	50 year
1-unit, detached	158,495	62.10%	1	2.32	399,516	428,126	526,857
1-unit, attached	13,809	5.40%	1	2.32	34,741	37,228	45,814
2 units	4,134	1.60%	2	4.64	20,587	22,061	27,149
3 or 4 units	16,177	6.30%	3.5	8.12	141,857	152,016	187,072
5 to 9 units	11,328	4.40%	7	16.24	198,150	212,339	261,307
10 to 19 units	12,168	4.80%	14.5	33.64	48,504	51,977	63,964
20 or more units	29,262	11.50%	20	46.4	1,480,964	1,587,018	1,953,004
Other 1 Unit Households	9805	3.90%	1	2.32	25,090	26,887	33,088

The total numbers of people that can be supported by housing in 10 years in Albuquerque, predicted by our model is 2,349,409 people. In 20 years, it is 2,517,652 people. In 50 years, it is 3,098,255 people.

Table 4. Seattle Conversion Rates (2022)

	Housing Numbers	Percent	M	M*P/H	10 year	20 year	50 year
1-unit, detached	144,289	38.70%	1	2.05	364,682	434,162	732,604
1-unit, attached	20,943	5.60%	1	2.05	52,770	62,825	106,010

2 units	7,946	2.10%	2	4.10	39,578	47,118	79,507
3 or 4 units	12,496	3.40%	3.5	7.175	112,137	133,502	225,270
5 to 9 units	19,914	5.30%	7	14.35	349,604	416,213	702,315
10 to 19 units	26,415	7.10%	14.5	29.725	970,128	1,154,962	1,948,878
20 or more units	139,472	37.50%	20	41	7,067,472	8,414,000	14,197,752
Other 1 Unit Households	961	0.30%	1	2.05	2,870	3,366	5,679

The total numbers of people that can be supported by housing in 10 years in Seattle, predicted by our model is 8,959,241 people. In 20 years, it is 10,666,148 people. In 50 years, it is 17,998,015 people.

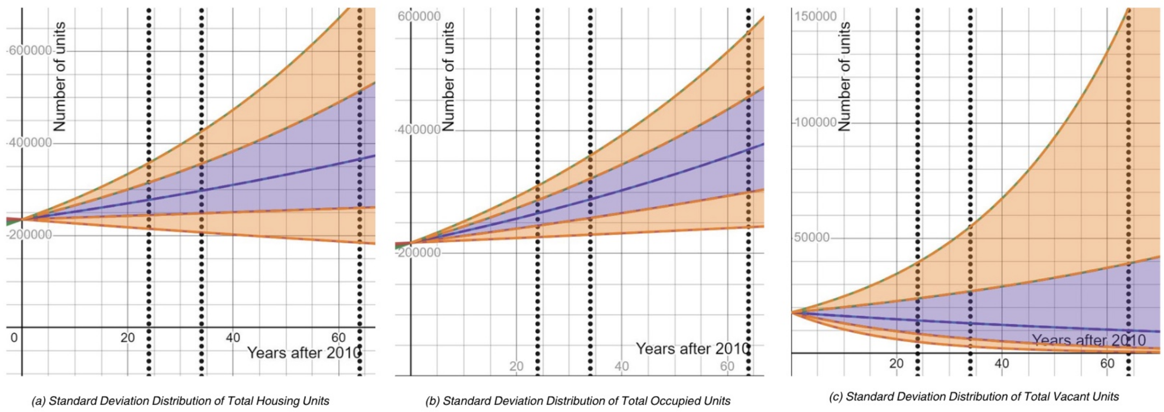


Figure 1: Standard Deviation Distributions in Albuquerque

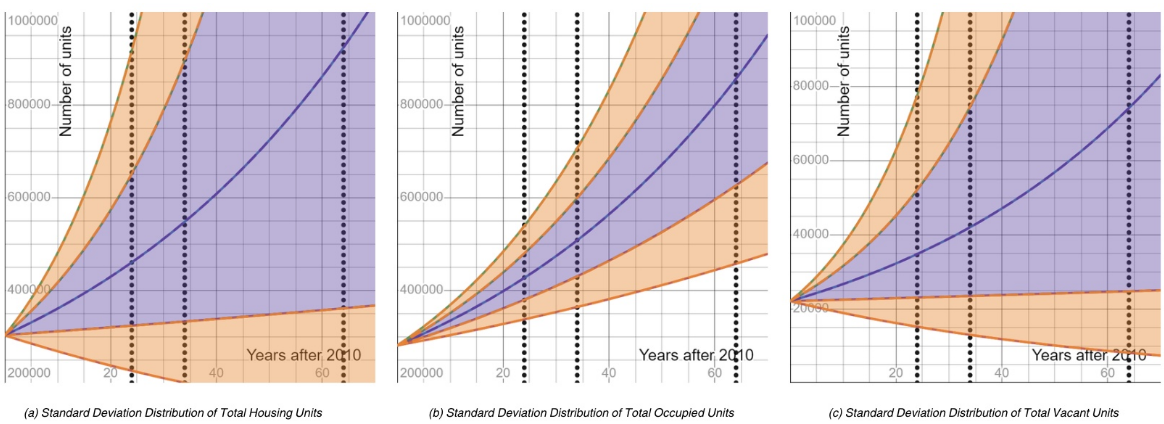


Figure 2: Standard Deviation Distributions in Seattle

1.5 Discussions

Our results show a stark difference between Seattle's projections and Albuquerque's projections: despite both trending towards a greater number of units of housing and the number of those units occupied, Seattle's growth rate is much greater than Albuquerque's and is also more volatile. Albuquerque is also trending toward a lower number of vacancies, while Seattle's is trending upward.

Seattle, Washington has a high housing supply growth rate, but also a high margin of error due to its rapid ascent. This matches recent feedback on Seattle's 2035 Plan, which originally projected growth of 70,000 units by 2035 from 2015, but recent years showed has nearly been surpassed. The majority of Seattle's newly constructed housing units are projected to be occupied, with a marginal year-over-year increase in vacancies as well. Over the next 50 years, we expect Seattle to potentially near 1 million housing units.

Albuquerque, New Mexico, has a slower average year-over-year growth rate, at around 1.007, and our confidence interval is tighter than for Seattle. Its total vacant units are also decreasing each year. Albuquerque has the potential for its measurements to both grow and shrink, making policy changes in the city especially important.

External factors support our results: With climate change and rapid global warming, areas further North such as Seattle are seeing increased interest, while Albuquerque, which lies in the Southwest, is facing extreme heat waves and water shortage.

Table 5. Results of Model 1

	Housing Units (10 years)	Housing Units (20 years)	Housing Units (50 years)	Supportable Population (10 years)	Supportable Population (20 years)	Supportable Population (50 years)
Seattle	459,673	547,252	923,431	8,959,241	10,666,148	17,998,015
Albuquerque	277,303	297,161	365,690	2,349,409	2,517,652	3,098,255

1.6 Sensitivity Analysis

Although we reached numbers up to 17,998,015 people that can be supported for our 50-year projection of Seattle in our Model Execution, it's important to recognize that this is an estimation based on our people per unit value. This is inherently limited by a lack of knowledge of how many people reside in each unit household, thus without further data (that is accessible), we can't project a more solidified number.

Looking at our house supply projections in Table 2 offers a more holistic view of our model. For example, the 10-year estimate of housing units in Seattle, WA is 459,673 compared to the total housing units in 2022, which is 372,436. This is a much more reasonable contrast to the drastic increase in population that can be supported.

To back-test our model, we inputted our data into our models and found that our data lined up with our models' projections. All of our data was contained within our error margins, verifying our predictions.

1.6.1 Strengths and Weaknesses

Our model on the large scale fits qualitative analysis. Seattle is seeing a high growth rate at the moment, so its future is more difficult to predict compared to Albuquerque. We also project Seattle over the next 10 years to gain about 80,000 units, which is in line with the past 10 years.

Over longer periods, especially around the 50-year time frame, our model lacks precision. This aligns with the measurements we are attempting to model because a wide range of events could occur over the next 50 years. This makes a high-growth city like Seattle have outcomes where it becomes larger than modern-day New York City, and outcomes where its size collapses below 150,000 family units.

Another limitation of our model is that our projections do not add up properly. The sum of the occupied units and vacant units should equal the total housing units projection. However, because we analyzed each variable independently, the average growth rates and standard deviations varied, creating data that differed from the projected total. Other than for the 50-year totals, however, the predictions were in similar ranges. Thus, taking our prior limitation into analysis, our error here is not significant because our data conveys the same message for the 50-year time period: there is a wide range of predictions for half a century from now.

2 Q2: It Was the Worst of Times

2.1 Defining the Problem

We need to project homeless populations in the same cities of Seattle and Albuquerque over the next 10, 20, and 50 years. To do this, we looked at the data provided by the U.S. Department of Housing and Urban Development's data on Homelessness in Seattle and Albuquerque. We utilized a regressive technique, ARIMA, described below, to forecast future homelessness numbers in these cities.

2.2 Variables

Table 6. Variable Definitions

Variable	**Definition**
p	AutoRegressive order (AR)
d	Differencing order
q	Moving Average Order (MA)

2.3 Assumptions

1. **Our data for the populations of homeless people are based on a point-in-time analysis.**

   Justification: Homeless populations fluctuate greatly when attempting to consider every person who has experienced homelessness at any point throughout a year. Thus, to simplify, we only consider the number of homeless people at a specific moment in time for future projections.

2. **Future projections can be calculated using previous data**

   Justification: It allows us to project the rate of homelessness in 10 years, 20 years, and 50 years without having to take into account internal and external factors that may affect homelessness rates. Thus, we can use past homelessness rates to project future ones.

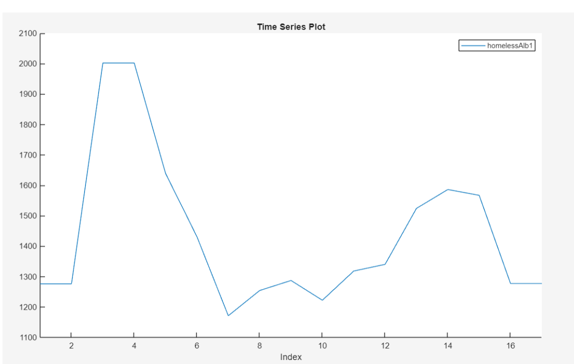
2.4 Model 2: Homeless Projection Model

We started by forecasting time series models using autoregressive integrated moving average (ARIMA) models in order to calculate what would happen in the future. The autoregressive (AR) component models the relationship between observations and the number of lagged observations. The integrated (I) component represents the differencing of raw observations to make the time series stationary. We need to make sure that the time series is stationary, as the forecasting model assumes that the underlying data is stationary. If the data were not stationary, we would get misleading results, which would lead to poor forecasting. The moving average (MA) component models the relationship between an observation and a residual error from a moving average model that is being applied to lagged observations. Because the provided data only gave us around seventeen data points, the predictions were somewhat flawed in accuracy, and more data points would be preferable.

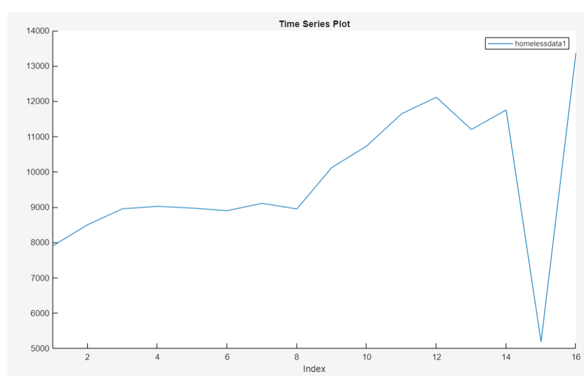
Table 7. Seattle and Albuquerque Homelessness Data

Year	Albuquerque	Seattle
2007	1,276	7,902
2008	1,276	8,501
2009	2,002	8,952
2010	2,002	9,022
2011	1,639	8,972
2012	1,431	8,899
2013	1,171	9,106
2014	1,254	8,949
2015	1,287	10,122
2016	1,222	10,730
2017	1,318	11,643
2018	1,340	12,112
2019	1,524	11,199
2020	1,586	11,751
2021	1,567	5,183
2022	1,277	13,368

As mentioned, we used Matlab's Econometric Toolbox to implement the ARIMA model. First, we imported this data into Matlab.



(a) Albuquerque Homelessness



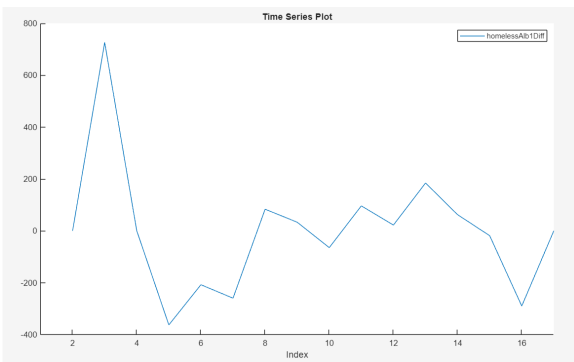
(b) Seattle Homelessness

Figure 3: City Homelessness Data

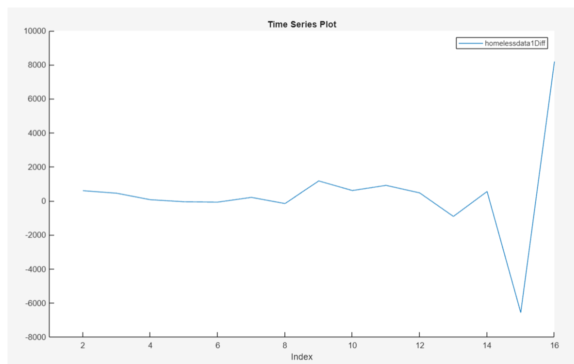
ARIMA models are only useful on stationary data, which means that the variance, or probability distribution, of the data does not change over time. This means that the data can then be effectively modeled as a stochastic, or probabilistic, process. To assess if the current data is stationary or not, we used the augmented Dickey-Fuller test.

The augmented Dickey-Fuller test is used to determine if there will be a unit root in the time series data. The unit root in a time series suggests if the series has a stochastic trend does not mean-revert to some constant level over a period of time. It shows the tendency to drift away from its mean level rather than fluctuating around some fixed value. If there is a unit root in the time series, then the null hypothesis applies. This would mean that the time series is non-stationary. If the time series does not contain the unit root, then the alternate hypothesis applies, which would mean that the time series is stationary.

We applied the Dickey-Fuller test to both data sets and found that they were both non-stationary. So, we transformed the non-stationary data by taking the difference; essentially, we took the discrete derivative of the data points by taking consecutive differences between adjacent data points. This changes whether a data set is stationary or not. Once we checked again with the Dicker-Fuller test, both transformed data sets were stationary.



(a) Transformed Albuquerque Data

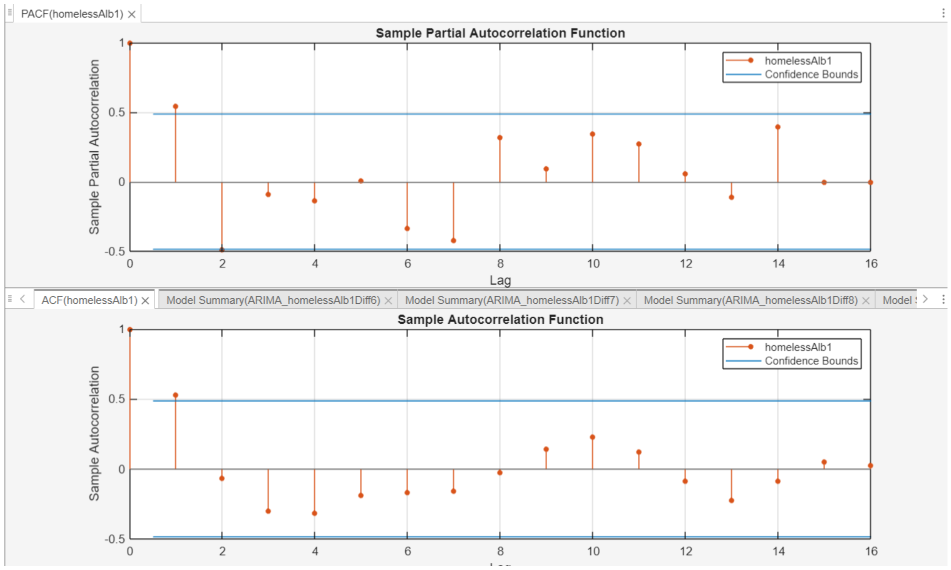


(b) Transformed Seattle Data

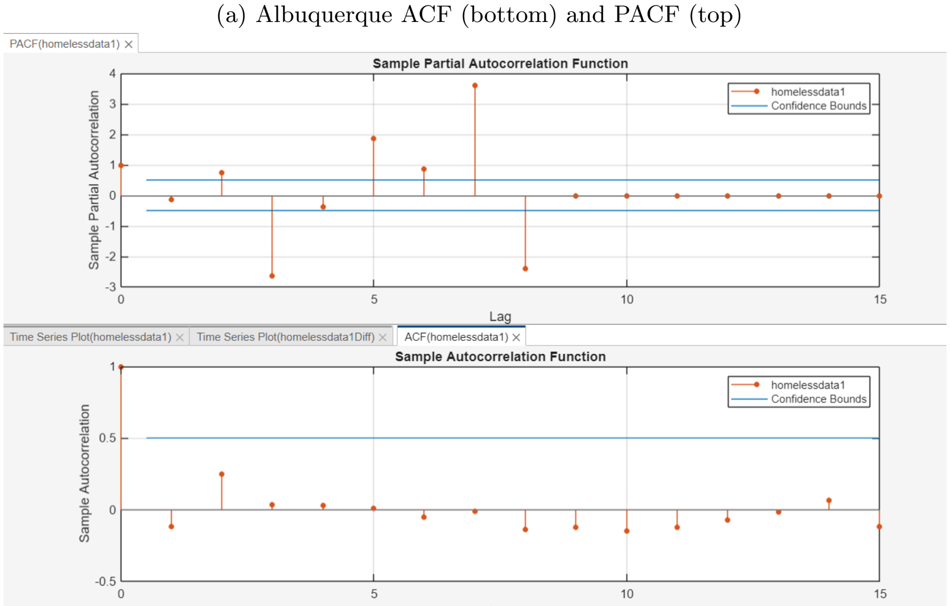
Figure 4: Taking the discrete derivative Homelessness Data

Once we obtained stationary data, we could now fit it with an ARIMA model. An ARIMA model is described by three parameters: p, d, and q. These are autoregressive order, differencing order, and the moving average order, respectively. The parameter d is determined by how many differences had to be taken to reach a stationary data set. Then, to determine p and q, we analyzed the autocorrelation function (ACF) and the partial autocorrelation function (PACF). The ACF measures the correlation between a series and its lagged values. The correlation coefficients calculated at various lags is used to determine how strongly the current observation compares with older correlations. From the ACF plot, we can obtain the q value by first analyzing the significant peaks beyond the confidence interval. The lag value at the first significant spike after any initial significant spike represents the moving average order, q. The PACF measures the correlation between a series and its lagged values while controlling for the effects of other lags in between. PACFs are useful in determining the direct relationship between observations at different lags without the impact of intermediary lags. To obtain the p value, we can first look at the PACF plot to identify the significant spike beyond the confidence interval. The lag value at the first significant spike after any initial significant spike is the auto-regressive order, p. This process is essentially the same as for finding the q value.

From the ACF and PACF graphs, we successfully determined the parameters for each ARIMA model. For Seattle data, we took $p = 2$, $d = 1$, $q = 2$, and for Albuquerque data we took $p = 2$, $d = 1$, and $q = 1$. This got us models that fit the data that we could use to project future trends.



(a) Albuquerque ACF (bottom) and PACF (top)

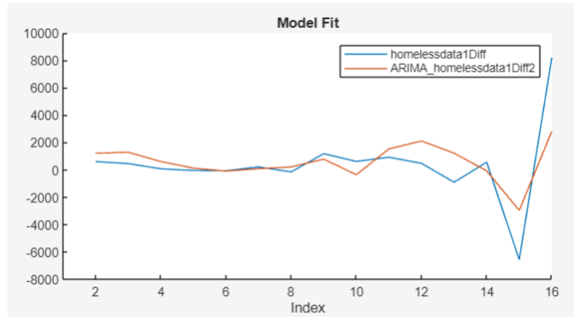


(b) Seattle ACF (bottom) and PACF (top)

Figure 5: Autocorrelation and Partial Autocorrelation Functions

Though our model doesn't fit the given data exactly, this isn't terrible since if it fits exactly then our model could be subject to overfitting, in which it would inaccurately predict future data. So, we projected homelessness rates using this data, and wrote MatLab code that was used to backtrack (or integrate) from the differenced data to homeless numbers in each city. This code is provided in Appendix B. Table 8 provides our final results for each city and each time interval, and graphs are provided showing what our model predicts over time.

(a) Albuquerque Model Fit

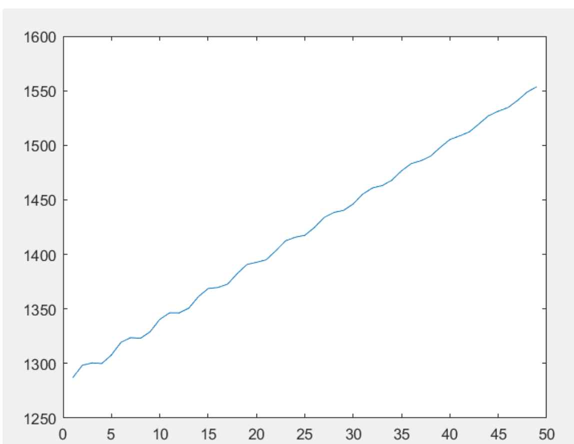


(b) Seattle Model Fit

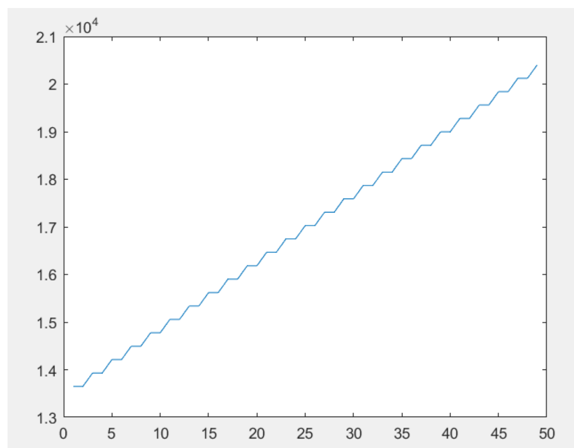
Figure 6: Fitting ARIMA Model

Table 8. Seattle and Albuquerque Homelessness Projections (Number of People)

Seattle				Albuquerque			
Current	10 Years	20 Years	50 Years	Current	10 Years	20 Years	50 Years
13,368	14,760	16,180	20,400	1,277	1,330	1,390	1,550



(a) Albuquerque # Homeless Projections



(b) Seattle # Homeless Projections

Figure 7: Final Model Predictions

2.5 Discussions

2.6 Sensitivity Analysis

One important component of our model is the data that we used to predict the homelessness rates. In Table 7, we see that in year 2021, Seattle has 5,183 homeless people; however, this isn't representative of the data, as it excludes the unsheltered total. While this is an important conceptual consideration, its effect in practice is essentially negligible, as it doesn't affect our model's computation.

2.6.1 Strengths and Weaknesses

A main strength of ARIMA is how extendable it is to different datasets, so this method can be applied to any other city in the world with enough data. It also utilizes stochastic elements, meaning it isn't assuming that rates are completely determined by a limited set of factors when, in the real world, the future is unpredictable. It is also useful on smaller data sets, such as the one we used in this model, to readily predict homeless numbers. With Matlab, ARIMA is also extendable to seasonal data and is powerful enough for more complicated sets of data. Bouncing off of this, although ARIMA is still inherently a regression model, it is much more accurate than other regression models for intricate data. Almost all real world data cannot be approximated by a line, polynomial, or logarithmic curve if one intends to extrapolate data. ARIMA takes advantage of stochastic processes and moving averages to make sure the data does not go out of control and stays within reason.

However, the power of ARIMA does have some setbacks. As seen in how our model fits data, though the ARIMA model fits the data fairly well, on a smaller scale it isn't quite accurate at all. If you were to, say, look exactly 100 months into the future, our model could be predicting something wildly different from what may happen, but on a more general time period, such as 10 years into the future, it is able to give a good estimate of the data. Also, ARIMA models make it difficult to accurately predict data given that you have to take the discrete derivative more than once. If your data is really not stationary, you'll have to take the difference multiple times to get stationary data (if you are even able to). This is the inherent assumption of ARIMA, so without it, the model is essentially useless. It is also difficult to predict data for each level taking the difference is performed, since then many more initial values are needed to climb back up the ladder to backtrack to the predicted data number.

3 Q3: Rising from This Abyss

3.1 Defining the Problem

We are tasked with determining a plan that the city of Seattle should follow to address homelessness. Our approach consists of comparing statistics of Seattle residents to those of the general U.S. population in subjects we identify as important factors in determining rates of homelessness. The goal is to identify areas in which the residents of Seattle need extra assistance that may currently be contributing to their status as the third-highest population in the country.[5]

3.2 Assumptions

1. **There are a set amount of plans that can be taken to alleviate homelessness**

   Justification: This allows us to limit suggested solutions from our model to those from the National Alliance to End Homelessness.

2. **Plans can be prioritized in a specified order**

   Justification: This allows us to suggest specific plans for different environments. Based on the statistics of each factor in the location, we can suggest one plan that will help target the specific issues in the region and effectively reduce homelessness.

3. **The statistics of the criteria can be correlated to the solutions**

   Justification: This allows us to draw direct correlations between factors and homelessness so that we know how to compare analyzed data to different plans.

3.3 Model 3: Homeless Projections

3.3.1 Developing our Model

We started our process by searching for potential solutions for homelessness online, and the problems that each of them solved. We found 7 problems that each had a potential solution: quantity/quality of jobs and job creation, economic assistance and increasing housing, lack of healthcare and increasing healthcare, high domestic violence rates and increased domestic violence reduction programs, racial inequality and equitable housing, homes sitting on the market and rapid rehousing programs, and unaffordable housing and increased low-income housing. For each of these solutions, to determine their impact on Seattle, we compared Seattle's corresponding metric that the solution would solve to the national average. If Seattle's score was worse than the national average, the corresponding priority became more important.

The equation that we used to develop scores was similar to a chi-squared test:

$$\frac{|\text{Expected} - \text{Observed}|}{\text{Expected}}$$

We then changed the sign depending on whether the observed value (Seattle's value) was better or worse than the expected value (national value).

3.3.2 Executing our Model

Table 9. Quantity/Quality of jobs

	National	Seattle
Proportion Receiving Support for Disabilities[6]	0.037	0.032
High School Graduation Rate[7][8]	0.87	0.86
College Graduation Rate[9][10]	0.64	0.628
Proportion Receiving Social Security Benefits[11]	0.19	0.18

Overall Score: 0.05450274174

Table 10. Housing Assistance

	National	Seattle
Proportion Receiving Housing Assistance[12][13]	0.0420711974	0.05489381334

Overall Score: 0.3047837174

Table 11. Healthcare

	National	Seattle
Proportion of People With Disability[14][2]	0.0420711974	0.05489381334
Proportion of People Without Insurance[14][2]	0.098	0.05

Overall Score: 0.3628754873

Table 12. Domestic Violence

	National	Seattle
Domestic Violence Experience Rate[15][16]	0.335	0.365

Overall Score: -0.1351351351

Table 13. Racial inequality

	National	Seattle
Percentage of White People in Poverty[17][18]	8.6	5.7
Percentage of Black People in Poverty[17][18]	17.1	18.9
Racial Poverty Disparity	8.5	13.2

Overall Score: -0.5529411765

Table 14. Rapid Rehousing

	National	Seattle
Median time on market[19][20]	69	49

Overall Score: 0.2898550725

Table 15. Low-Income Housing

	National	Seattle
Median Cost[21]	348,079	650,949
Median Income	74,580	116,068
Median Income[22]	5	5.608337746

Overall Score: -0.2016520073

3.4 Discussions

We discovered that the most important issues Seattle needs to focus on are reducing racial inequality followed by increasing housing affordability. Our full score order is below:

Priority Number	Plan	Score
1	Racial Equality Programs	-0.5529411765
2	Low-Income Housing	-0.2016520073
3	Domestic Violence	-0.1351351351
4	Quantity/quality of jobs	0.05450274174
5	Median time on market	0.2898550725
6	Housing Assistance	0.3047837174
7	Healthcare	0.3628754873

Thus, we recommend that Seattle focus on reducing racial inequality by increasing equitable housing, followed by increasing housing affordability by increasing the amount of low-income housing available. We hope that our methodology is useful to policymakers in other urban areas as well and that Americans across the United States all find affordable, quality housing.

3.5 Sensitivity Analysis

Our results are rather respected, currently, racial inequality in the United States is high, thus it fits that one of the main issues that Seattle must solve is racial housing disparities. Additionally, urban areas are expensive to live in, thus logically housing affordability would be another priority for Seattle.

3.5.1 Strengths and Weaknesses

Our model is strong in determining which problems matter most for Seattle to solve, and in identifying plans to solve those problems. What our problem relies on, however, is that these plans work, which we assume, and if they do not, another measure must be taken to improve the measurement deficiencies we noticed with our models.

4 Conclusion

In the first question, we modeled the future housing supply in Seattle and Albuquerque by utilizing historical data to predict growth and standard deviation change over the next 10, 20 and 50 years. Our first model predicted that in 50 years, Seattle could support up to 18 million people, and Albuquerque up to 3 million people. Of course, the population of these cities don't come close to these numbers now, but that just goes to show how many homes that we will have available to these cities to house people. Our model incorporated various factors including vacancy rates and demographic trends, providing insights into potential housing availability and market dynamics. For question two, we leveraged data from the U.S. Department of Housing and Urban Development to create an ARIMA model. The ARIMA model helped us forecast changes in homelessness in both Seattle and Albuquerque in the coming 10, 20, and 50 years. In both of these cities homelessness is on the rise, with over 20,000 homeless in Seattle and 1,500 in Albuquerque expected in 50 years. This helped to identify the potential scale of homelessness and shows the need for targeted policy interventions. For question three, we focused on developing a plan for city-specific homelessness for Seattle. We addressed the homelessness by comparing city residents' statistics with national averages in key areas influencing homelessness rates. This approach was designed to identify the gaps in support and resources which offers a blueprint for areas to focus on for long-term sustainable solutions to reduce homelessness. Each phase of our work contributed to a holistic understanding of the challenges and opportunities within the housing market and homelessness. This will provide a foundation for informed decision-making and policy development.

5 References

1. https://unitedwaynca.org/blog/vacant-homes-vs-homelessness-by-city/

2. https://www.census.gov/quickfacts/fact/table/seattlecitywashington/DIS010222#DIS010222

3. https://www.census.gov/quickfacts/fact/table/albuquerquecitynewmexico,US/POP815222

4. A Tale of Two Crises, MathWorks Math Modeling Challenge 2024, curated data, https://m3challenge.siam.org/kdfrldh/.

5. https://www.seattletimes.com/seattle-news/homeless/white-house-announces-initiative-to-reduce-homelessness-in-seattle/

6. https://www.dshs.wa.gov/sites/default/files/dvr/pdf/2017%20Disability%20%26%20DVR%20Statistics%20Report.pdf

7. https://nces.ed.gov/fastfacts/display.asp?id=805#:~:text=The%20U.S.%20Department%20of%20Education,the%20ACGR%20in%202010%E2%80%9311.&text=The%20U.S.%20average%20ACGR%20for,87%20percent%20in%202019%E2%80%9320.&text=On%20average%2C%20the%20ACGR%20increased,%E2%80%9319%20to%202019%E2%80%9320.

8. https://www.seattletimes.com/education-lab/wa-high-school-graduation-rates-are-up-who-saw-the-biggest-gains/#:~:text=And%20in%20Seattle%2C%20students%20did,up%20from%20pre%2Dpandemic%20years.

9. https://nces.ed.gov/fastfacts/display.asp?id=40

10. https://www.opendatanetwork.com/entity/1600000US5363000/Seattle_WA/education.graduation_rates.percent_bachelors_degree_or_higher?ref=entity-question&year=2018

11. https://www.seattletimes.com/business/economy/compare-social-security-benefits-and-inflation-in-wa/

12. https://www.seattlehousing.org/about-us#:~:text=In%20a%20variety%20of%20housing,375%20sites%20throughout%20the%20city.

13. https://nlihc.org/sites/default/files/HousingSpotlight2-2.pdf

14. https://www.census.gov/quickfacts/fact/table/US/POP815222

15. https://www.cdc.gov/violenceprevention/intimatepartnerviolence/fastfact.html#:~:text=Data%20from%20CDC%27s%20National%20Intimate,related%20impact%20during%20their%20lifetime.

16. https://wibm.us/domestic-violence-washington-state-male-victims-female-victims-statistics/#:~:text=%E2%80%9CIn%20Washington%20State%2C%2041%25,domestic%20violence%20throughout%20their%20lifetimes.%E2%80%9D&text=(The%20term%20domestic%20violence%20in,sexual%20assault%2C%20or%20stalking.)

17. https://www.statista.com/statistics/200476/us-poverty-rate-by-ethnic-group/#:~:text=U.S.%20poverty%20rate%20in%20the%20United%20States%202022%2C%20by%20race%20and%20ethnicity&text=In%202022%2C%2017.1%20percent%20of,and%20ethnicities%20was%2011.5%20percent.

18. https://www.seattle.gov/rsji/racial-equity-research/poverty

19. https://fred.stlouisfed.org/series/MEDDAYONMARUS

20. https://www.realtor.com/realestateandhomes-search/Seattle_WA/overview#:~:text=Median%20days%20on%20market%3A%2049%20Days&text=Range%3A%200%20to%2080.&text=days%2080%20days-,End%20of%20interactive%20chart.,slightly%20down%20since%20last%20year.

21. https://www.thezebra.com/resources/home/average-home-price-in-us-2022/#:~:text=The%20average%20home%20price%20in,a%20house%20is%20West%20Virginia.

22. https://www.census.gov/library/publications/2023/demo/p60-279.html

6 Appendix

6.1 A. Developing our unused Model for Q1

We started by attempting to forecast time series models for our data in order to calculate what would happen in the future. We started off by trying to use the ARIMA (AutoRegressive Integrated Moving Average) model, though for this question we ended up not using it. Because the provided data only gave us thirteen data points, we decided to split the table into quarterly points and apply the SARIMA, which is the ARIMA model with the Seasonal (S) component. The Seasonal component captures repeating patterns that occur at fixed intervals. We got data from Quarterly Vacancy and Homeownership Rates by State and MSA taken from the U.S. Census shown in the following table.

Table 2. Seattle Homeowner Vacancy Rate

Year	Q1	Q2	Q3	Q4
2022	0.5%	0.4%	1.0%	1.1%
2021	1.1%	0.7%	0.6%	0.5%
2020	0.8%	0.3%	0.8%	0.6%
2019	0.4%	0.7%	1.2%	1.6%
2018	0.1%	1.0%	0.6%	1.4%
2017	0.5%	0.2%	0.5%	0.8%
2016	0.6%	0.9%	0.9%	1.0%
2015	0.6%	0.6%	0.9%	2.2%
2014	1.7%	1.5%	0.9%	0.9%
2013	1.3%	2.4%	1.6%	1.5%
2012	2.9%	2.9%	2.0%	1.2%
2011	3.4%	2.6%	2.0%	2.3%
2010	4.1%	2.1%	3.0%	3.7%
2009	2.0%	2.4%	3.6%	3.2%
2008	1.9%	1.5%	2.3%	1.7%
2007	2.1%	1.4%	1.5%	2.1%
2006	0.6%	0.3%	1.6%	1.2%
2005	1.6%	0.8%	0.3%	1.1%

To calculate the value at each quarter for every year we first used the vacant units in Seattle data and divided each year into four quarters.[2] We assumed that the increase between any two year data points was linear so we calculated the amounts at Q1, Q2, Q3, and Q4. Then, we added up the percentages from the previous table for each year and calculated the ratio of the percentage for that quarter versus the the total percentage. We then multiplied this value by the difference between the ladder year and the earlier year. Finally, to obtain the final quarter values, we added this amount to the previously calculated quarter values. The table below shows the final results of this calculation.

Table 2. Seattle Homeowner Vacancy Rate

Year	Q1	Q2	Q3	Q4
2022	22,012	21,908	21,854	21,778
2021	21,684	21,530	21,413	21,324
2020	21,218	21,072	20,927	20,826
2019	20,766	20,708	20,602	20,531
2018	20,464	20,074	19,730	19,523
2017	19,317	19,222	19,128	18,985
2016	18,638	18,859	19,190	19,521
2015	19,889	20,181	20,298	20,590
2014	21,057	21,108	21,618	21,925
2013	22,639	22,646	22,658	22,680
2012	22,708	23,585	23,914	24,790
2011	25,448	26,109	26,529	26,890
2010	Null	Null	Null	27,190

In the end, this approach did not work, since the data got out of control with our models. Some of our more accurate models got close enough to the current Model 1's solution, so we decided to go with that instead since using ARIMA on question 1 was taking too much time.

6.2 B. Model 2 Matlab Code

6.2.1 Seattle

```
function ARIMA_homelessdata1Diff = modelTimeSeries(homelessdata)
%%Time Series Modeling Using the Econometric Modeler
% This code recreates the estimated model produced in the Econometric Modeler app. Use the code below to estimate the same model, or estimate a model with the same parameters on a new set of data.
%
%Input: A numeric matrix with the same number of columns as the data imported into the app (homelessdata)
%
%Output: The model containing estimated parameters (ARIMA_homelessdata1Diff)
%
homelessdata1 = readmatrix('whatever3.csv');

%% First Order Difference
% Perform a first order difference on time series homelessdata1
homelessdata1Diff = [NaN; diff(homelessdata1)];

%% Autoregressive Integrated Moving Average Model
%Estimate an ARIMA Model of homelessdata1Diff
ARIMA_homelessdata1Diff = arima('Constant',NaN,'ARLags',1,'D',1,'MALags',1:2,'Distribution','Gaussian');
g = estimate(ARIMA_homelessdata1Diff,homelessdata1,'Display','off');
num_periods = 50; % Number of periods to forecast
forecast_data = forecast(g, num_periods);
plot(forecast_data);
y = 13368;
m = length(forecast_data);
arr = [];
for val = 1:(m-1)
    arr = [arr, y+forecast_data(val)];
end
plot(arr);
disp(arr(length(arr - 1)));
end
```

6.2.2 Albuquerque

```
function ARIMA_homelessAlb1Diff9 = modelTimeSeries(homelessAlb)
%%Time Series Modeling Using the Econometric Modeler
% This code recreates the estimated model produced in the Econometric Modeler app. Use the code below to estimate the same model, or estimate a model with the same parameters on a new set of data.
%
%Input: A numeric matrix with the same number of columns as the data imported into the app (homelessAlb)
%
%Output: The model containing estimated parameters (ARIMA_homelessAlb1Diff9)
%
%Auto-generated by MATLAB (R2023b) and Econometrics Toolbox Version 23.2 on 03-Mar-2024 18:59:42
homelessAlb1 =  readmatrix('homelessA.csv');

%% First Order Difference
% Perform a first order difference on time series homelessAlb1
homelessAlb1Diff = [NaN; diff(homelessAlb1)];

%% Autoregressive Integrated Moving Average Model
%Estimate an ARIMA Model of homelessAlb1Diff
ARIMA_homelessAlb1Diff9 = arima('Constant',NaN,'ARLags',1:2,'D',1,'MALags',1,'Distribution','Gaussian');
g = estimate(ARIMA_homelessAlb1Diff9,homelessAlb1Diff,'Display','off');
```

```
num_periods = 50; % Number of periods to forecast
forecast_data = forecast(g, num_periods);
plot(forecast_data);
y = 1277;
m = length(forecast_data);
arr = [];
for val = 1:(m-1)
    arr = [arr, y+forecast_data(val)];
end
plot(arr);
disp(arr(length(arr - 1)));
end
```