

Forecasting of Amarillo, TX and Sacramento, CA Data Sets

Advanced Business Forecasting
CIDM – 6362- 70
Dr. Barthel
Spring 2024

Research and Analysis Completed By:
Miguel M. Sanchez



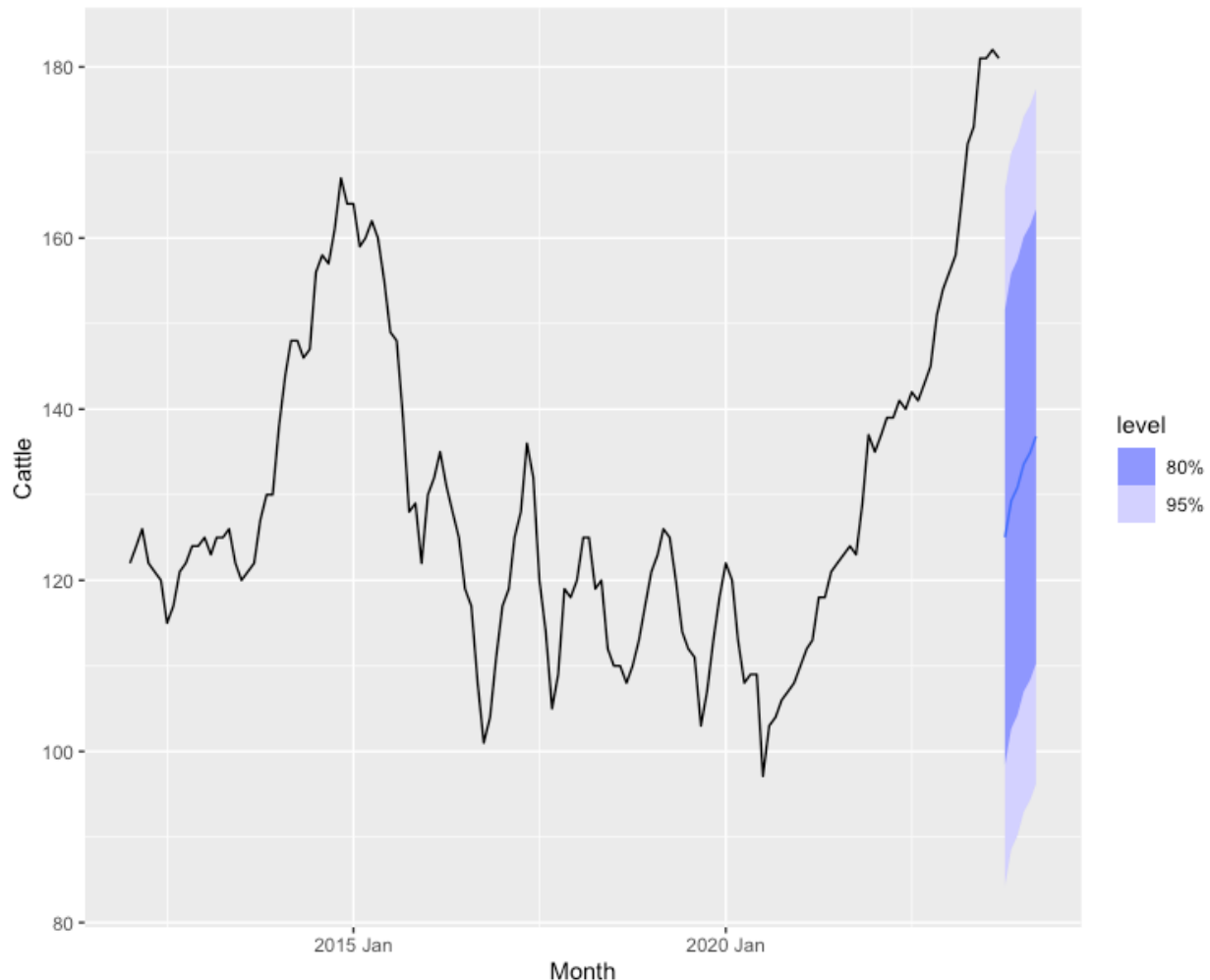
Introduction to Data Sets

This forecasting study will focus on Amarillo, Texas and Sacramento, California. Amarillo is a major cattle hub in Texas. Amarillo has cattle packing operations, ranches, feeding lots, and auctions. Amarillo is a key cattle industry logistics hub and transportation route. Agriculture, oil and gas, and manufacturing thrive in Amarillo. Amarillo is centrally positioned in the Texas Panhandle (25,000 miles) and has many direct transit routes. Texas is the second largest state in the continental US, with diverse landscapes, habitats, and climates. Second, I'll analyze Sacramento, California. California's agriculture capital is Sacramento. Sacramento's economy includes government (being California's capitol), technology, and healthcare. Wine production from Sacramento and California grapes is the city's biggest economic driver. Sacramento has summer, spring, fall, and winter with different temperatures and climates. Sacramento also serves as a regional economic center for California and its neighbors. Amarillo, Texas and Sacramento, California are both economic hubs and industry leaders in their respective states (Cattle for Amarillo and Grape Production for Sacramento).



Section 1 - Time Series Regression (TSR) Model Selection

Amarillo Cattle Trend Model

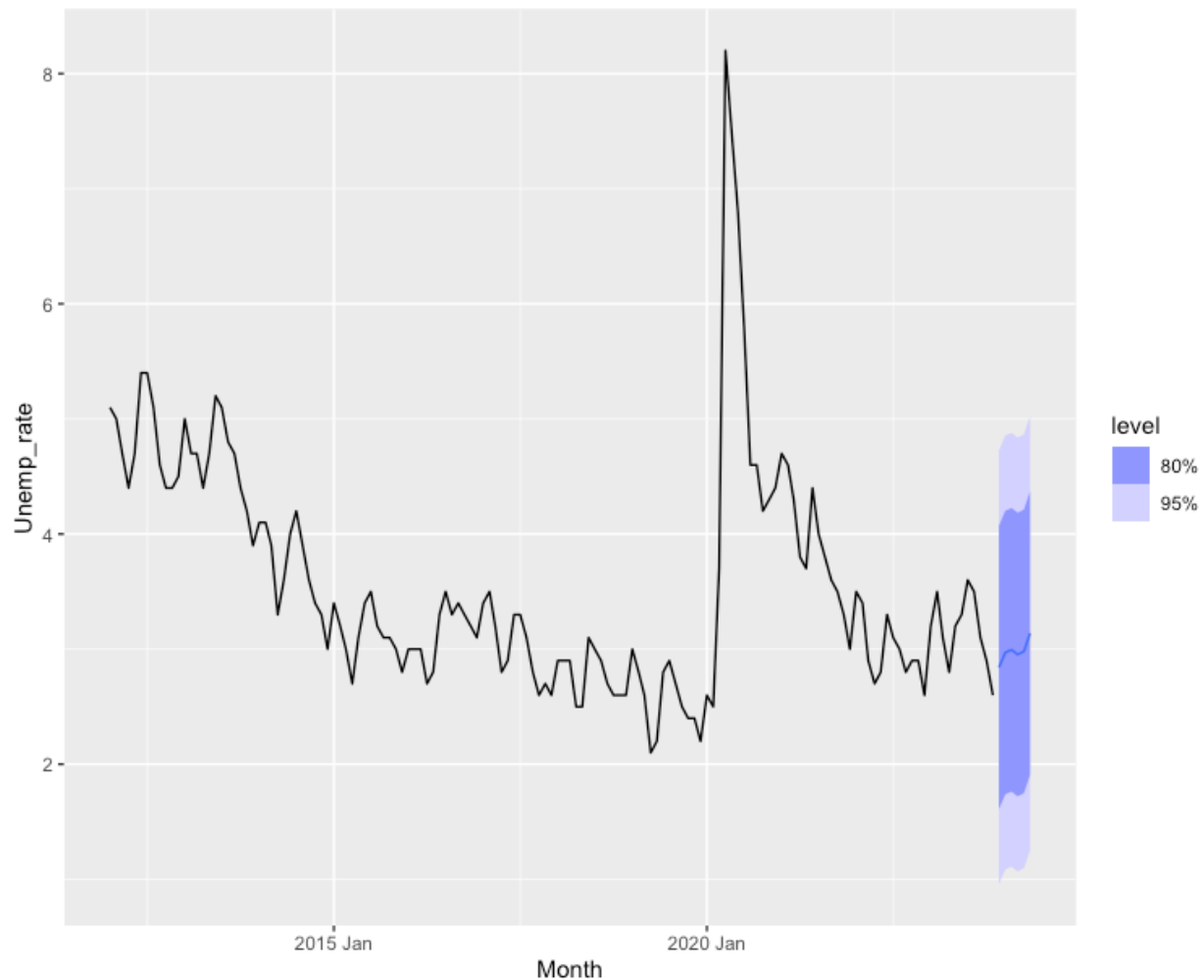


When reviewing the Amarillo Cattle data set, the TSR model selection I chose was the Trend Model. This is due to this specific model having the lowest cross-validation “CV” value across all other models. The CV value for the Trend Model of the Cattle data set When a model has the lowest CV value it essentially means that the predictive performance of this model has less errors than others and will also be able to generalize better when unseen data may be presented. It takes not only historical data but helps to accurately predict the future and the unseen data.

Amarillo Unemployment Fourier Model



Amarillo Unemployment Rate Fourier Model

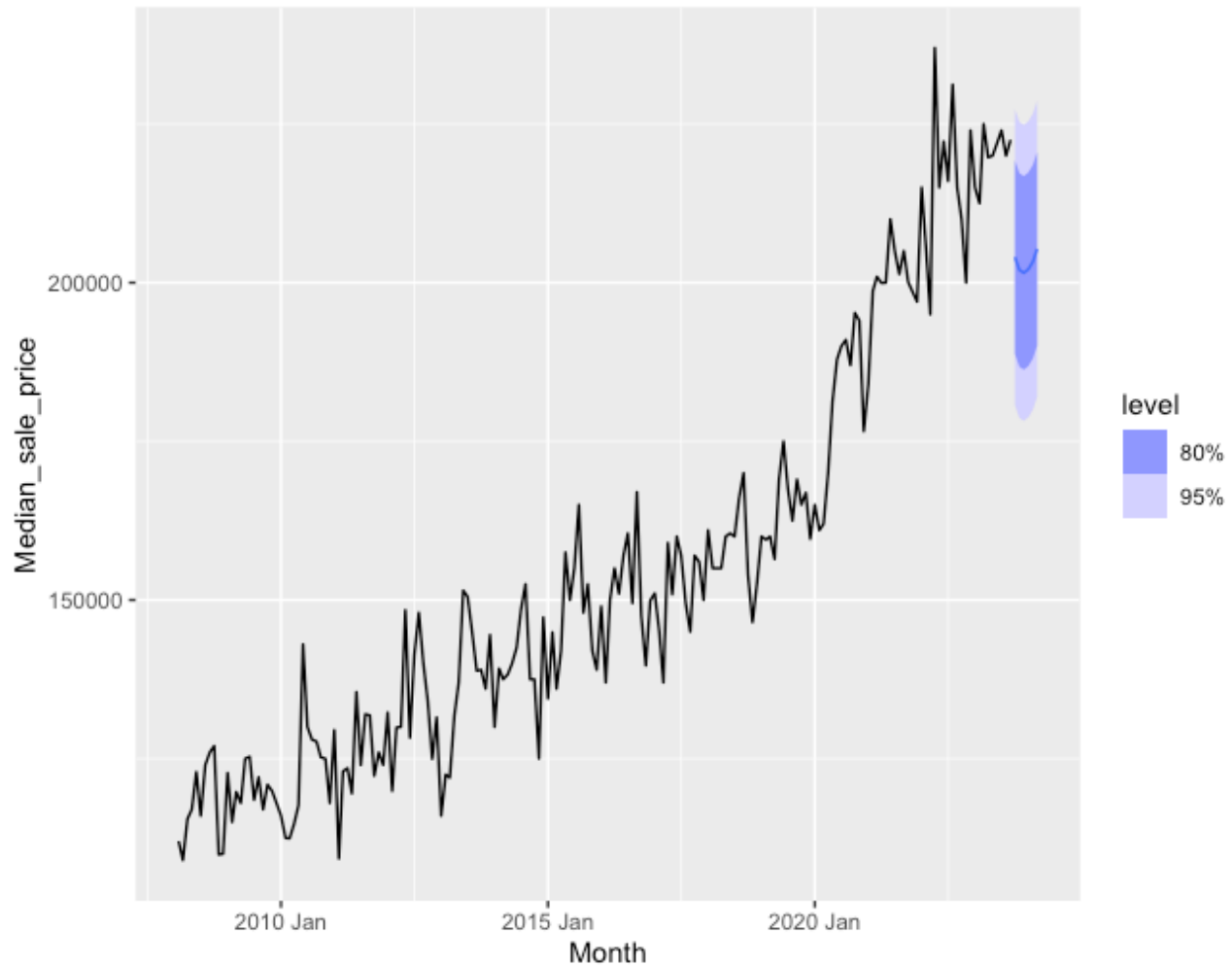
 $K = 2$

AIC = -11.6, AICc = -11.4, CV = 0.906

When reviewing the Amarillo Unemployment data set, the TSR model selection I chose was the Fourier Model. This is due to this specific model having the lowest cross-validation “CV” value across all other models. The CV value of the Fourier Model for the Amarillo Unemployment data set was 0.906 after testing multiple K values and finalizing the model with a K value =2. When a model has the lowest CV value it essentially means that the predictive performance of this model has less errors than others and will also be able to generalize better when unseen data may be presented. It takes not only historical data but helps to accurately predict the future and the unseen data.



Amarillo Housing Market Fourier Model



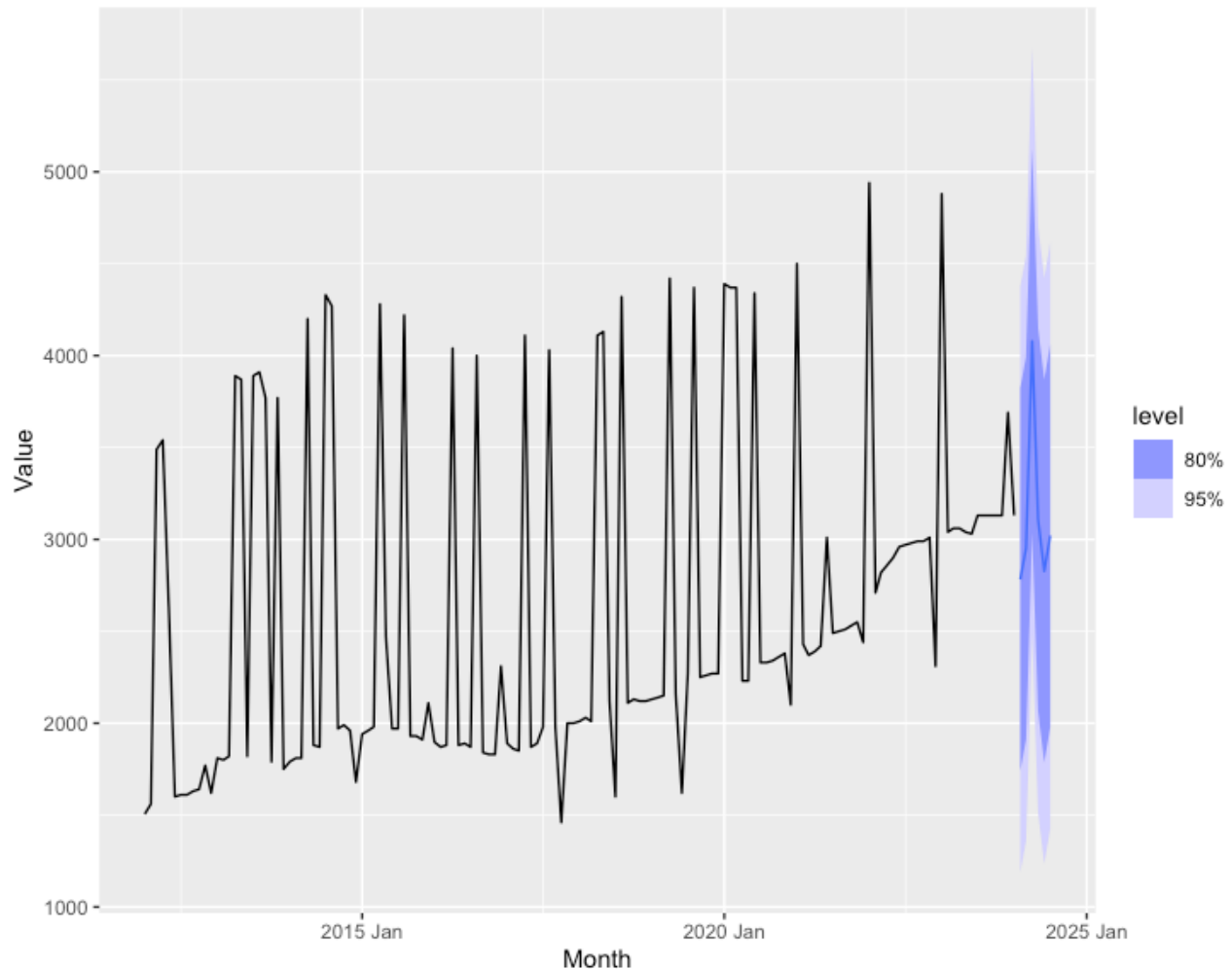
$K = 2$

$AIC = 3528, AICc = 3529, CV = 140365860$

When reviewing the Amarillo Housing Market data set, the TSR model selection I chose was the Fourier Model. This is due to this specific model having the lowest cross-validation “CV” value across all other models. The CV value of the Fourier Model for the Amarillo Unemployment data set was 140365860, with a $K = 2$ value. When a model has the lowest CV value it essentially means that the predictive performance of this model has less errors than others and will also be able to generalize better when unseen data may be presented. It takes not only historical data but helps to accurately predict the future and the unseen data.



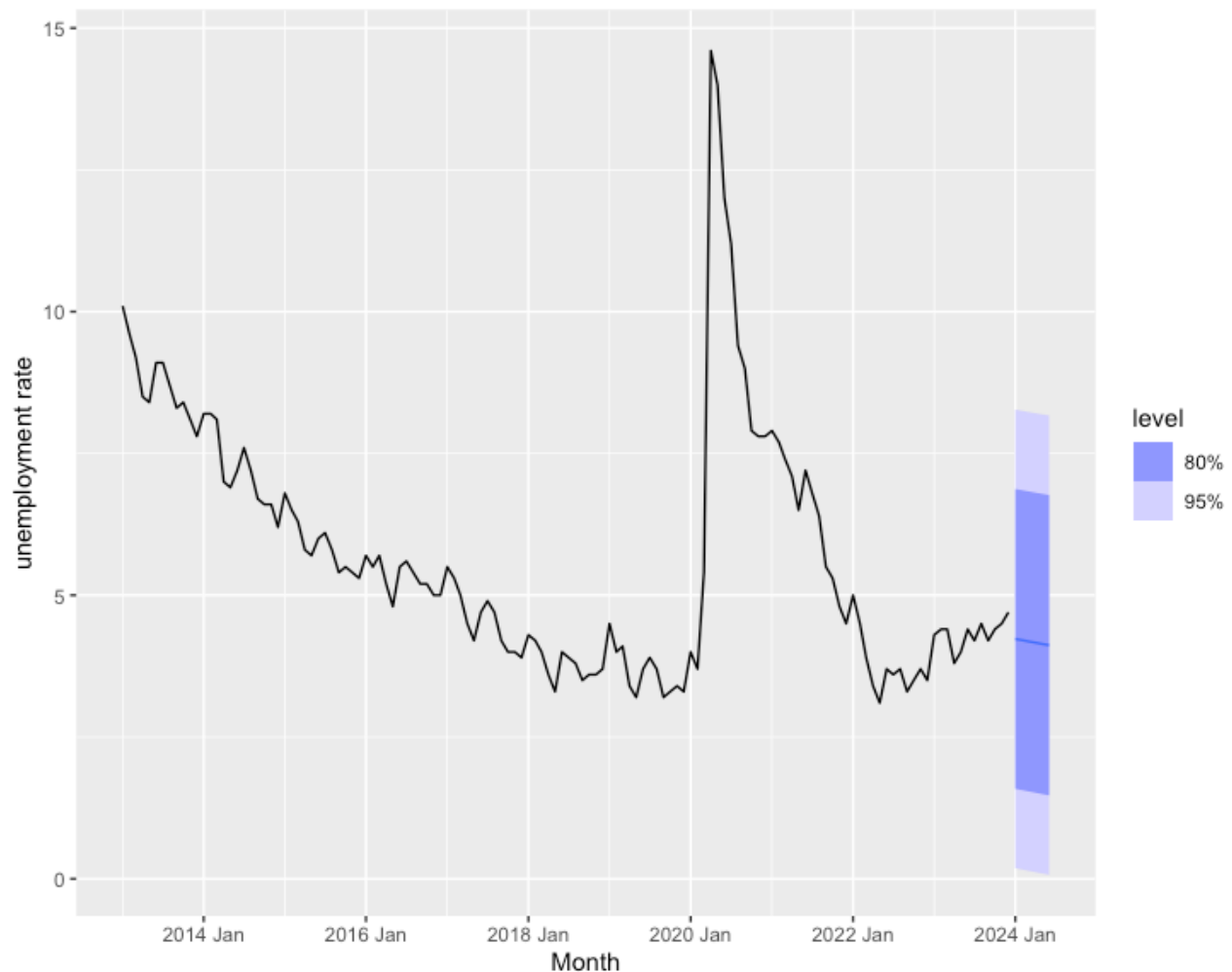
Sacramento Grapes Trend + Season Model



When reviewing the Sacramento Grapes data set, the TSR model selection I chose was the Trend + Season Model. This is due to this specific model having the lowest cross-validation “CV” value across all other models. The CV value of the Trend + Season Model for the Sacramento Grapes data set was 654026. When a model has the lowest CV value it essentially means that the predictive performance of this model has less errors than others and will also be able to generalize better when unseen data may be presented. It takes not only historical data but helps to accurately predict the future and the unseen data.



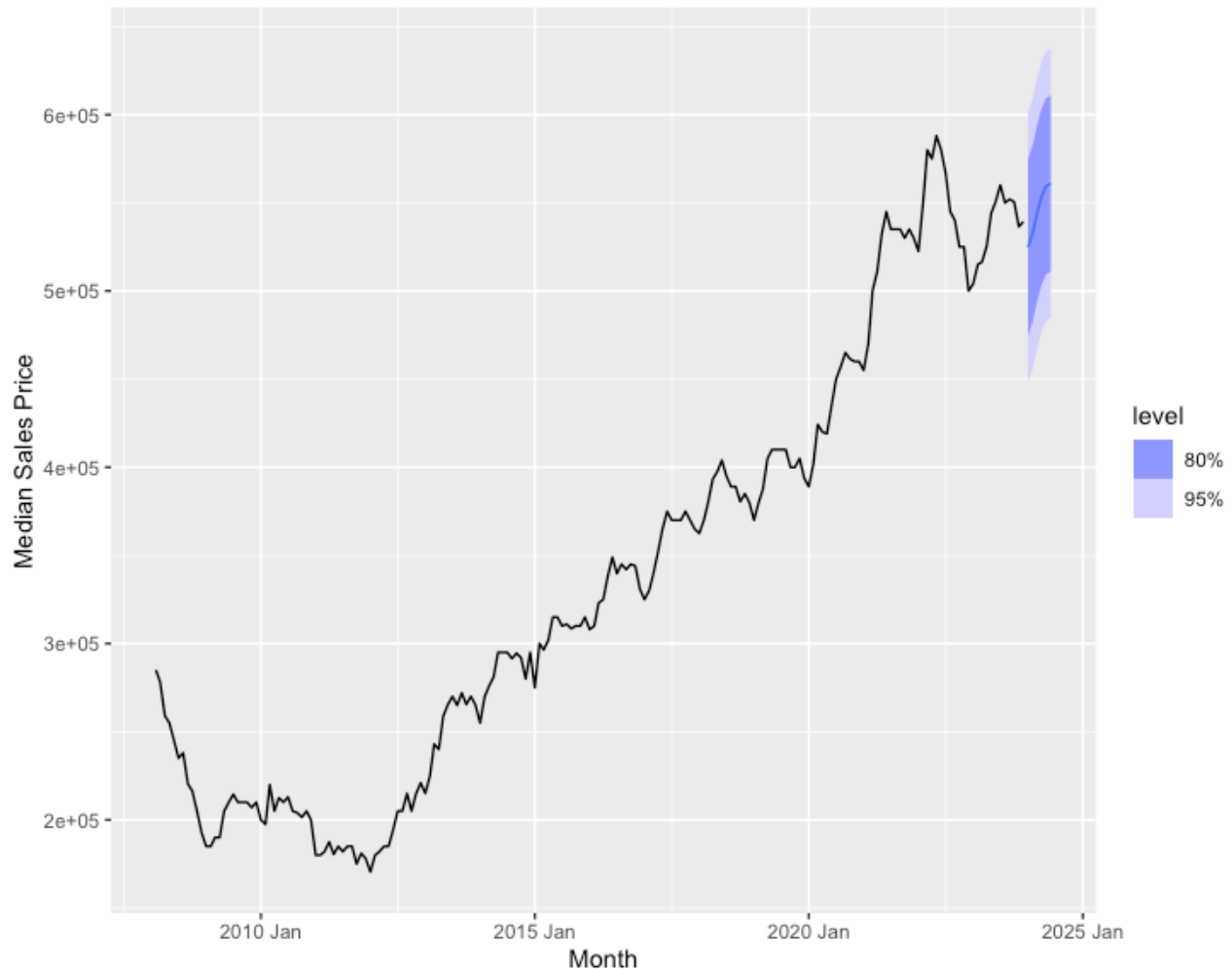
Sacramento Unemployment Trend Model



When reviewing the Sacramento Unemployment data set, the TSR model selection I chose was the Trend Model. This is due to this specific model having the lowest cross-validation “CV” value across all other models. The CV value for the Trend Model of the Sacramento Unemployment data set is 4.16. When a model has the lowest CV value it essentially means that the predictive performance of this model has less errors than others and will also be able to generalize better when unseen data may be presented. It takes not only historical data but helps to accurately predict the future and the unseen data.



Sacramento Housing Market Fourier Model



$K = 2$

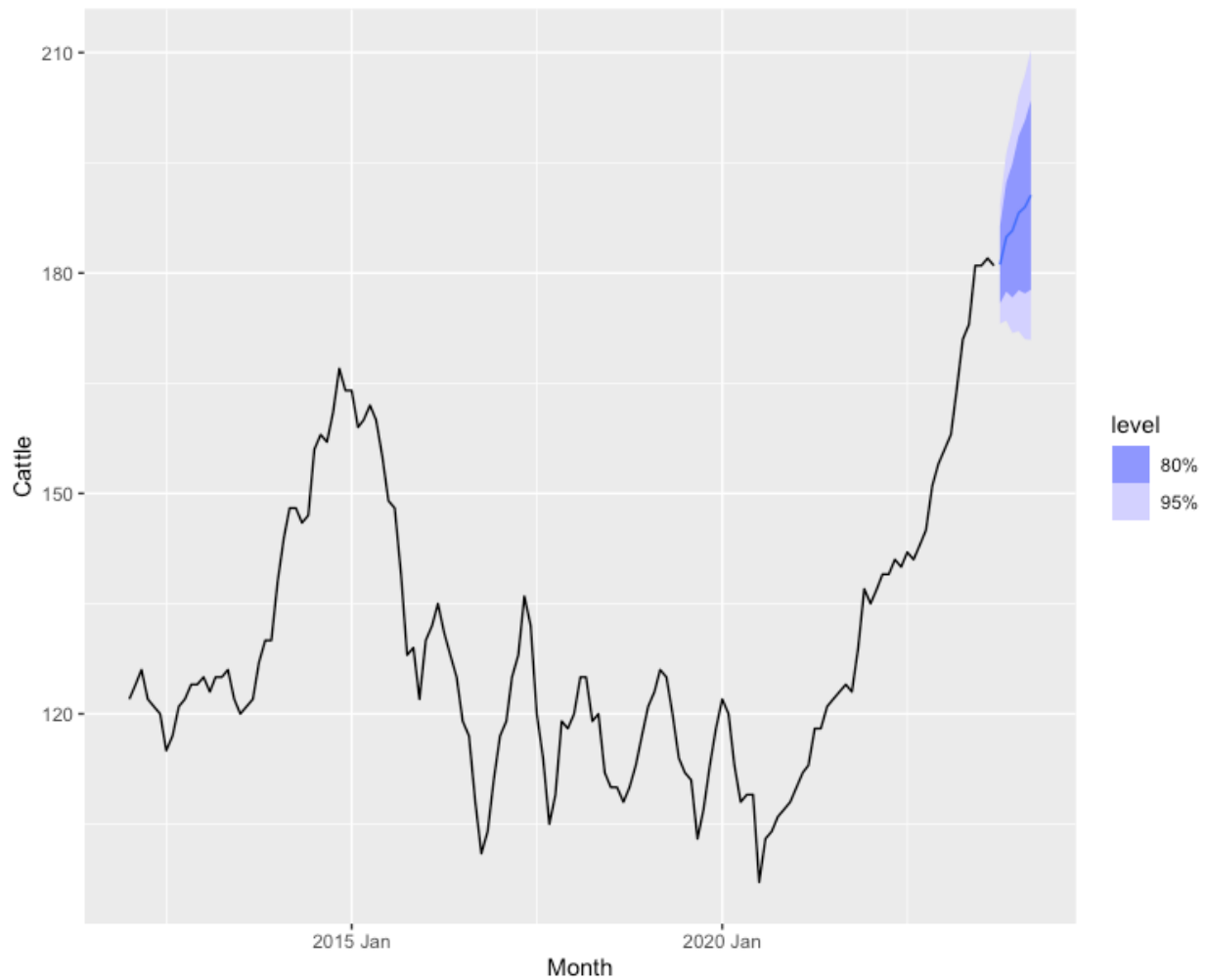
AIC = 4038, AICc = 4038, CV = 1514327476

When reviewing the Sacramento Housing Market data set, the TSR model selection I chose was the Fourier Model. This is due to this specific model having the lowest cross-validation “CV” value across all other models. The CV value of the Fourier Model for the Sacramento Housing Market data set was 1514327476, with a $K = 2$ value. When a model has the lowest CV value it essentially means that the predictive performance of this model has less errors than others and will also be able to generalize better when unseen data may be presented. It takes not only historical data but helps to accurately predict the future and the unseen data.

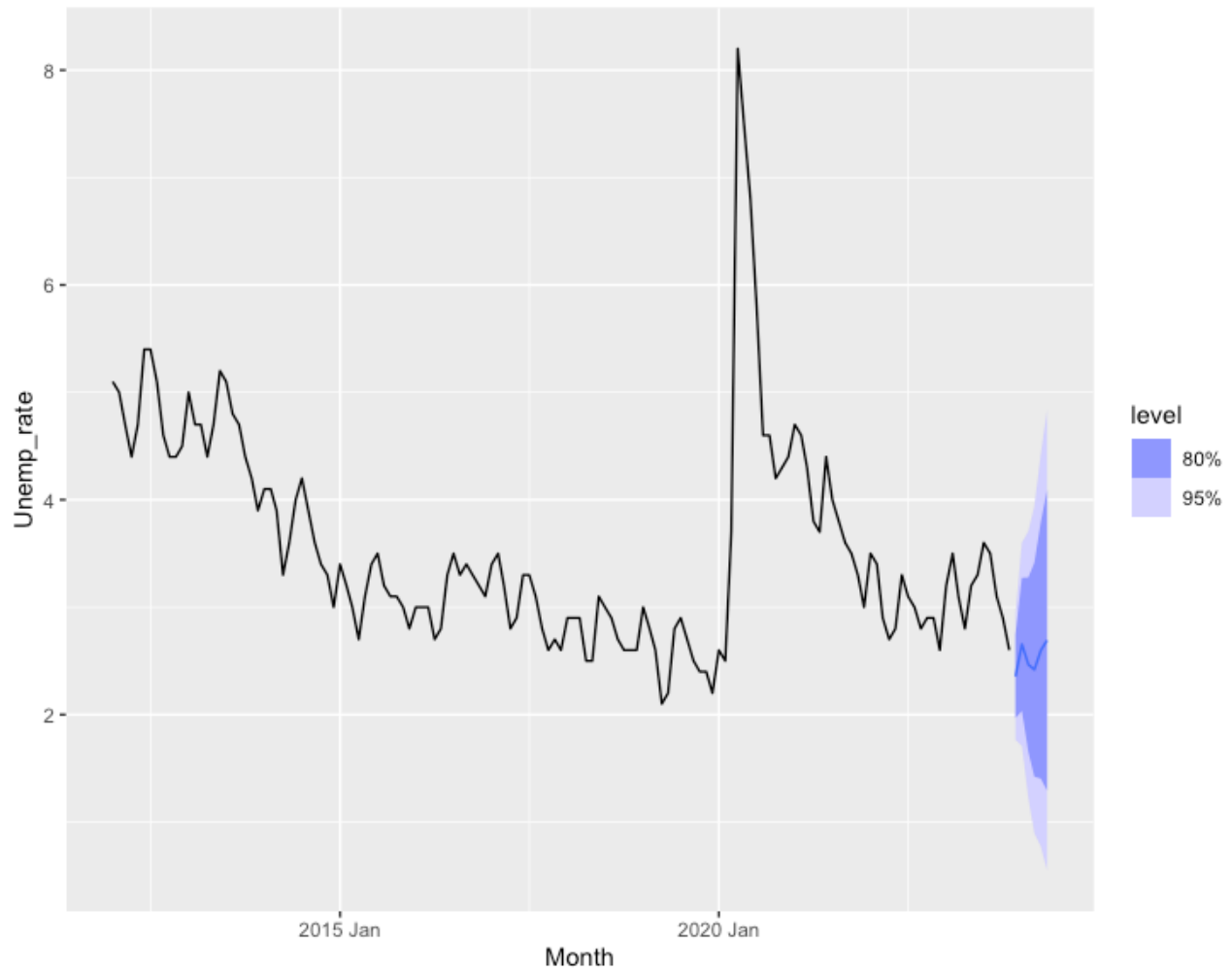


Section 2 – Exponential Smoothing (ETS) Models

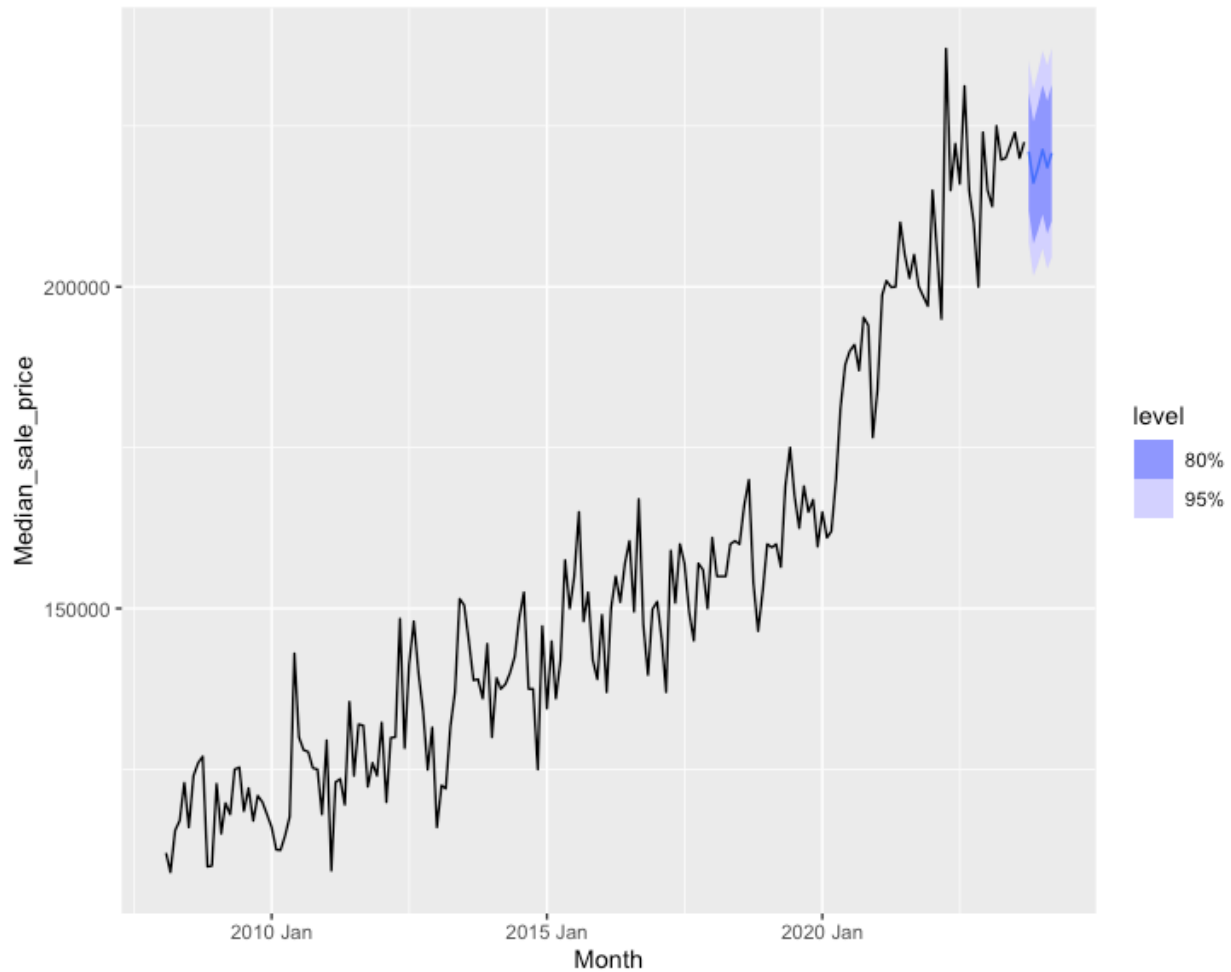
Amarillo Cattle ETS Out of Sample 6 Month Forecast



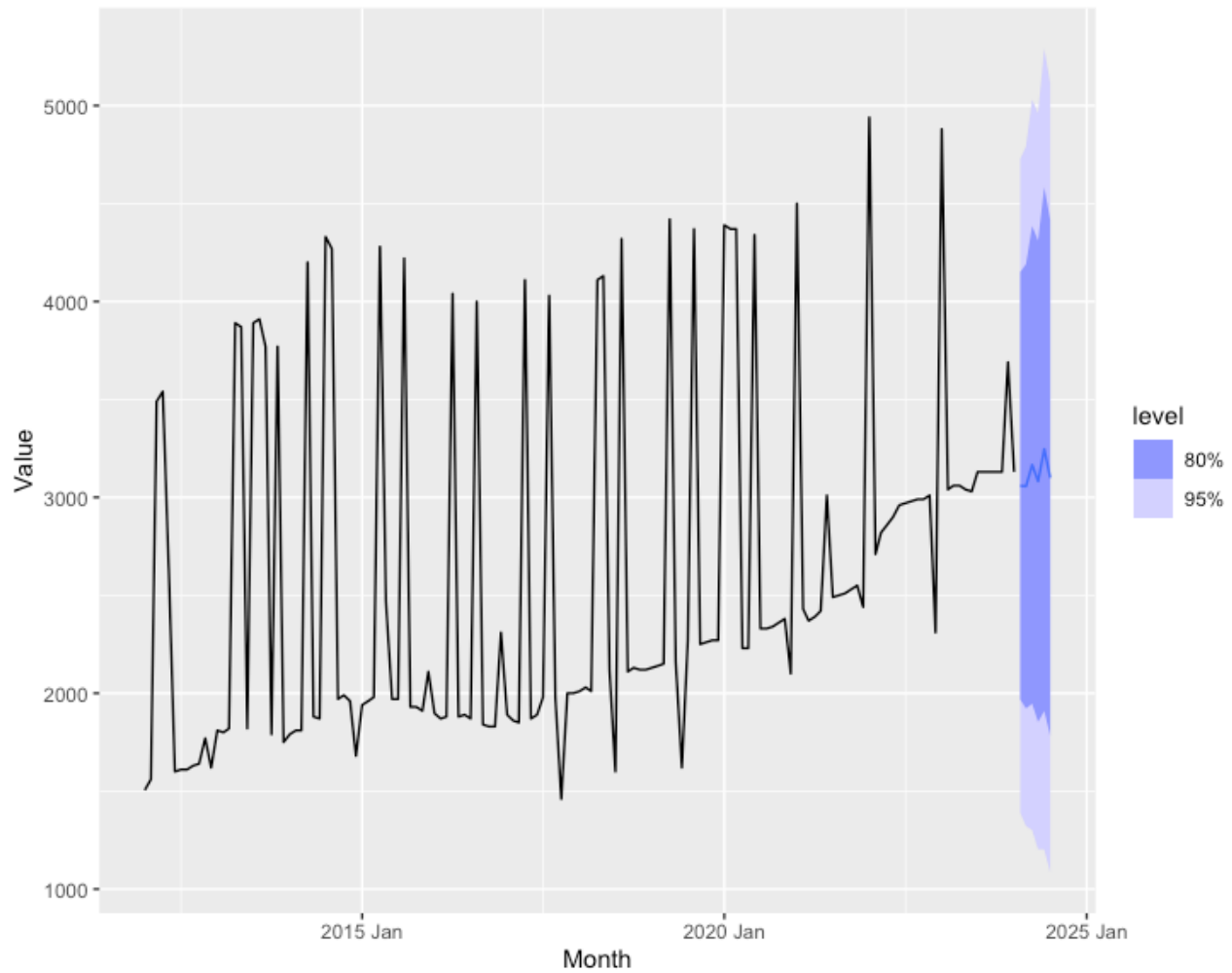
Amarillo Unemployment ETS Out of Sample 6 Month Forecast



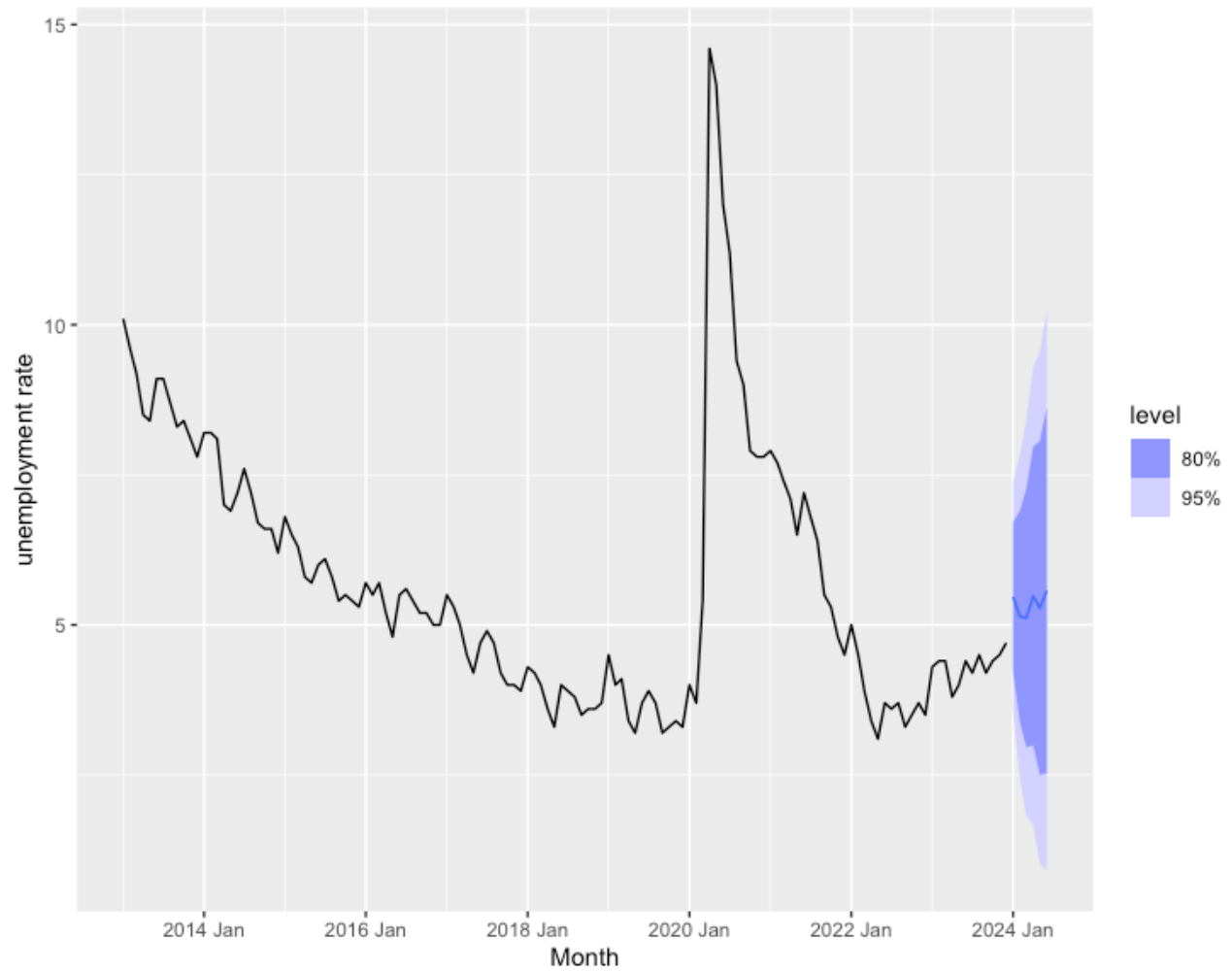
Amarillo Housing Market ETS Out of Sample 6 Month Forecast



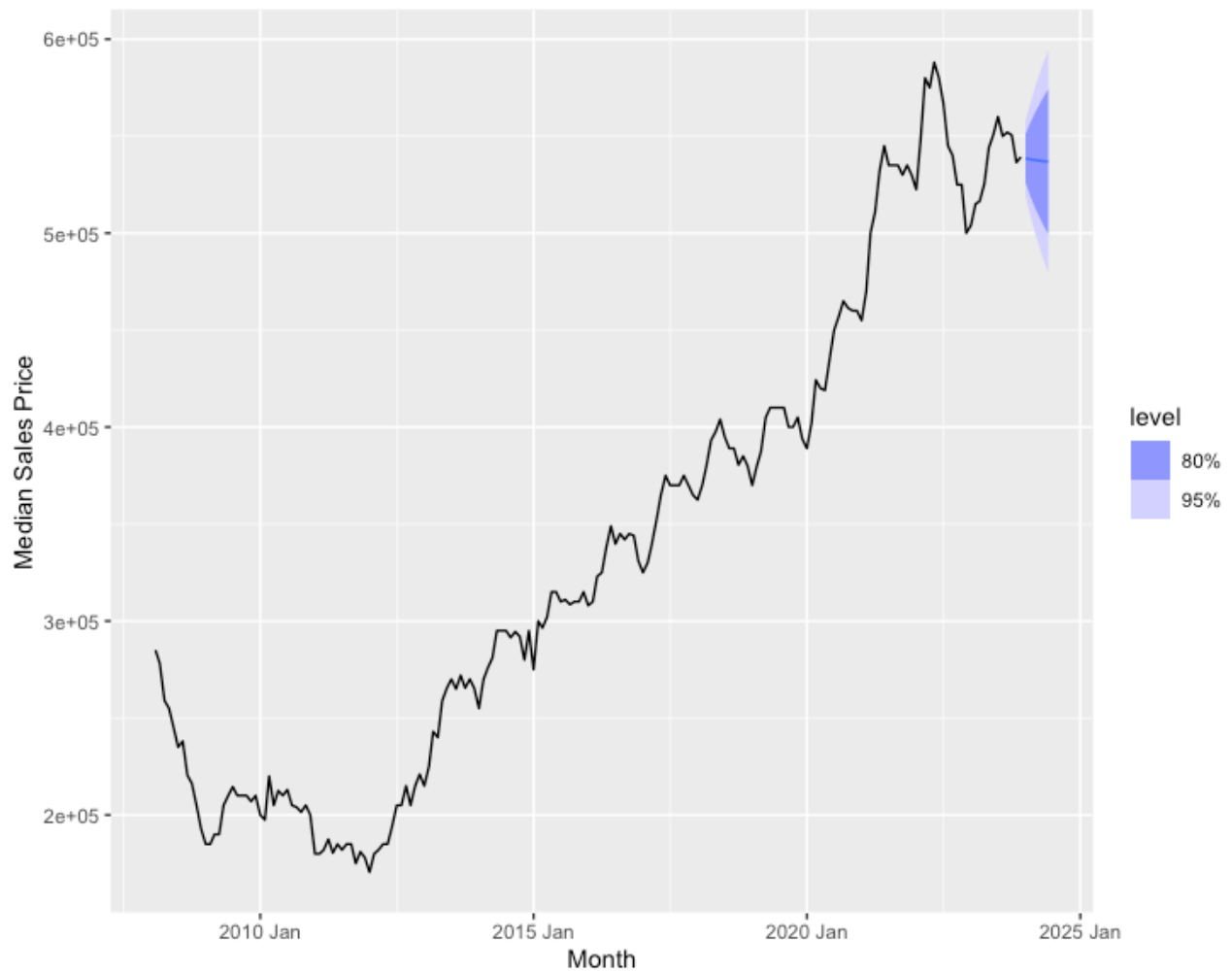
Sacramento Grapes ETS Out of Sample 6 Month Forecast



Sacramento Unemployment ETS Out of Sample 6 Month Forecast

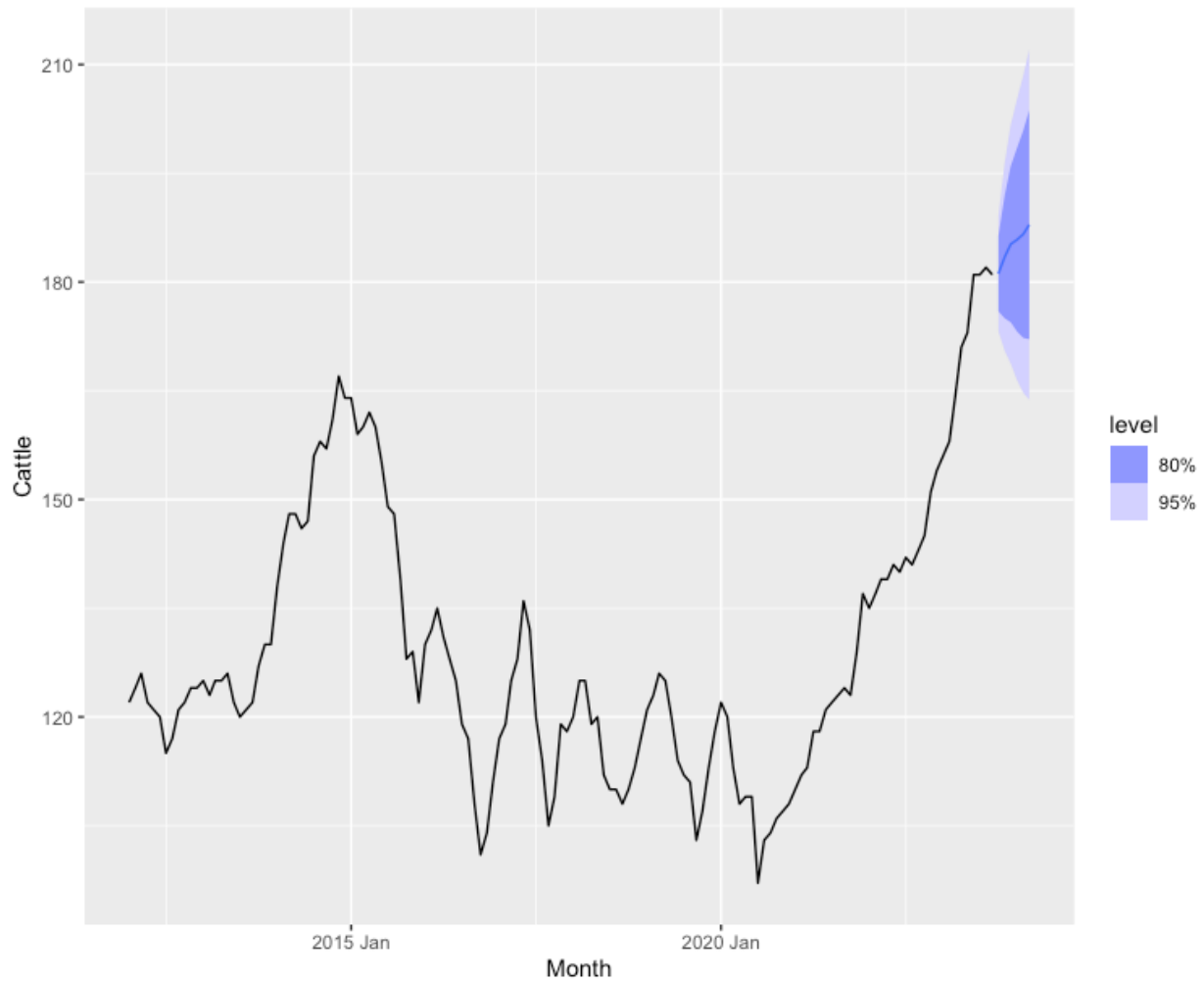


Sacramento Housing Market ETS Out of Sample 6 Month Forecast

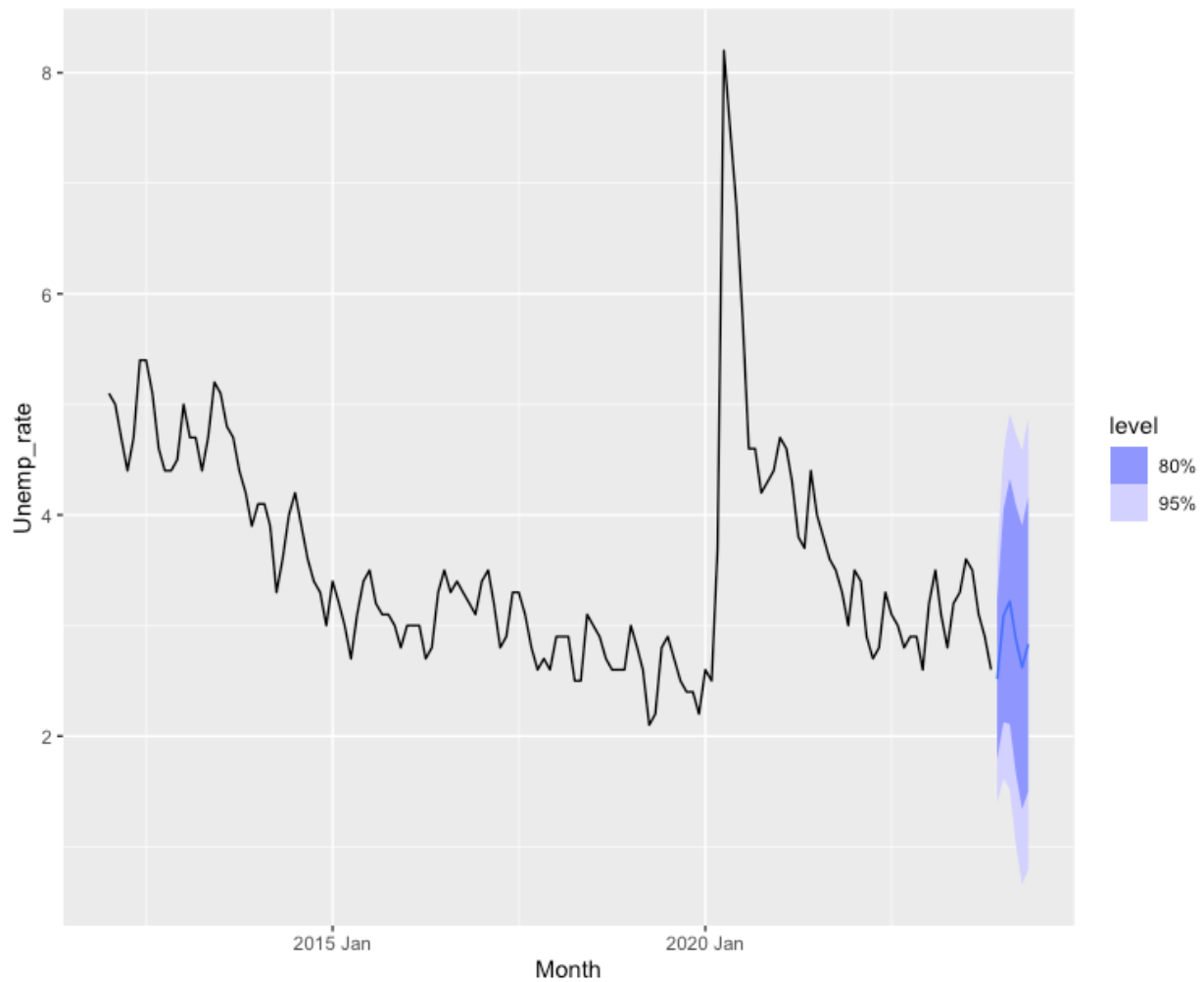


Section 3 – ARIMA Models

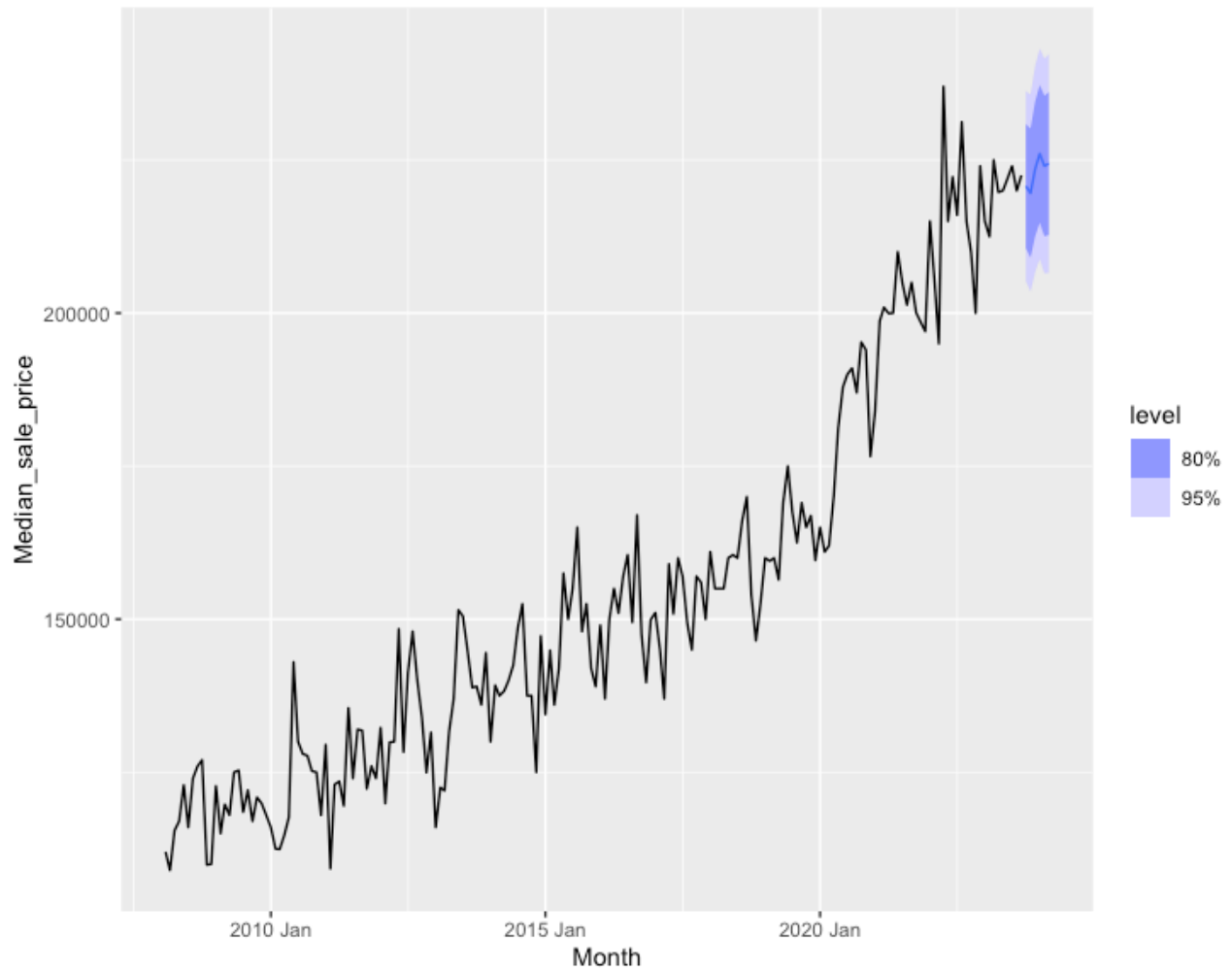
Amarillo Cattle - ARIMA Out of Sample 6 Month Forecast



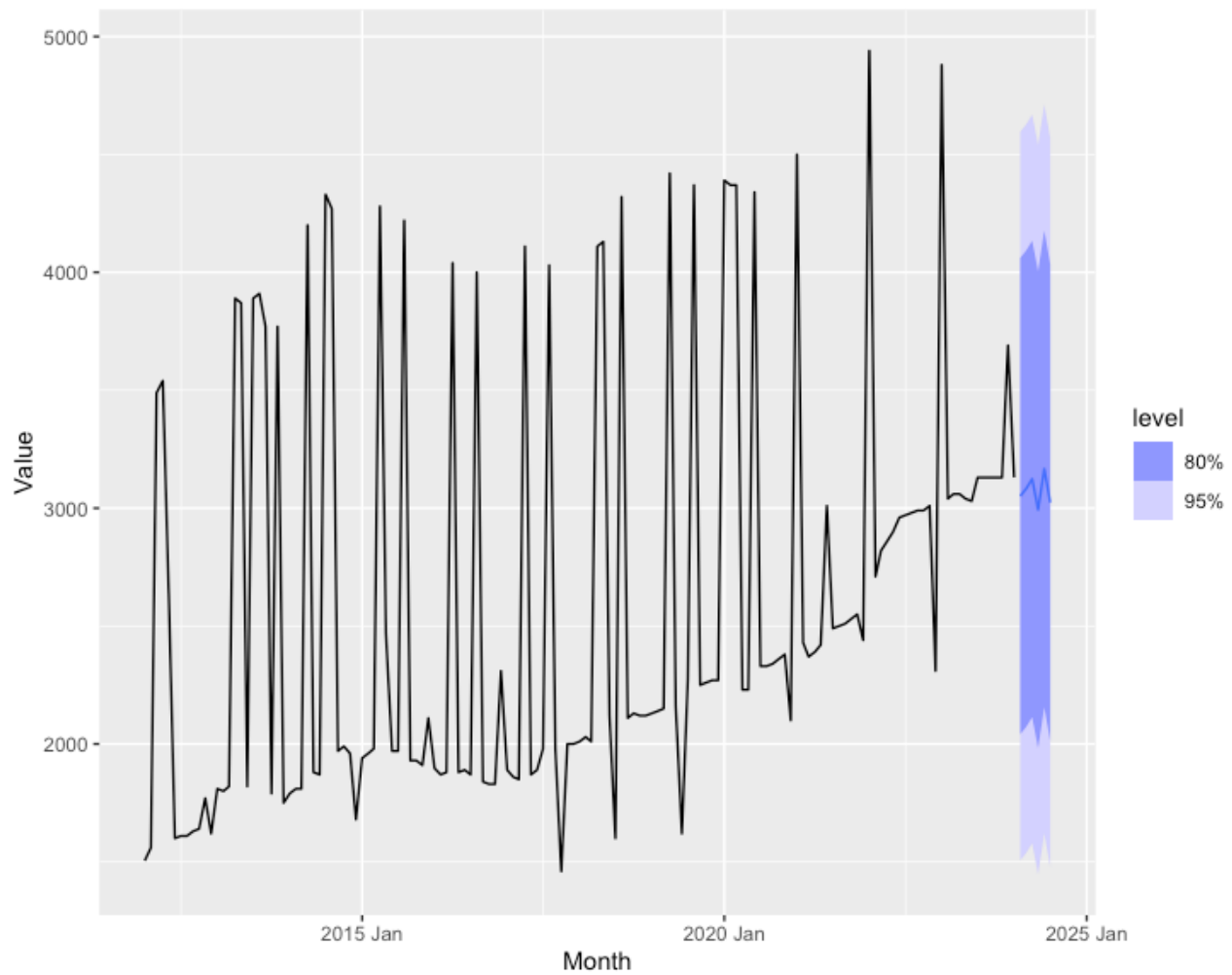
Amarillo Unemployment ARIMA Out of Sample 6 Month Forecast



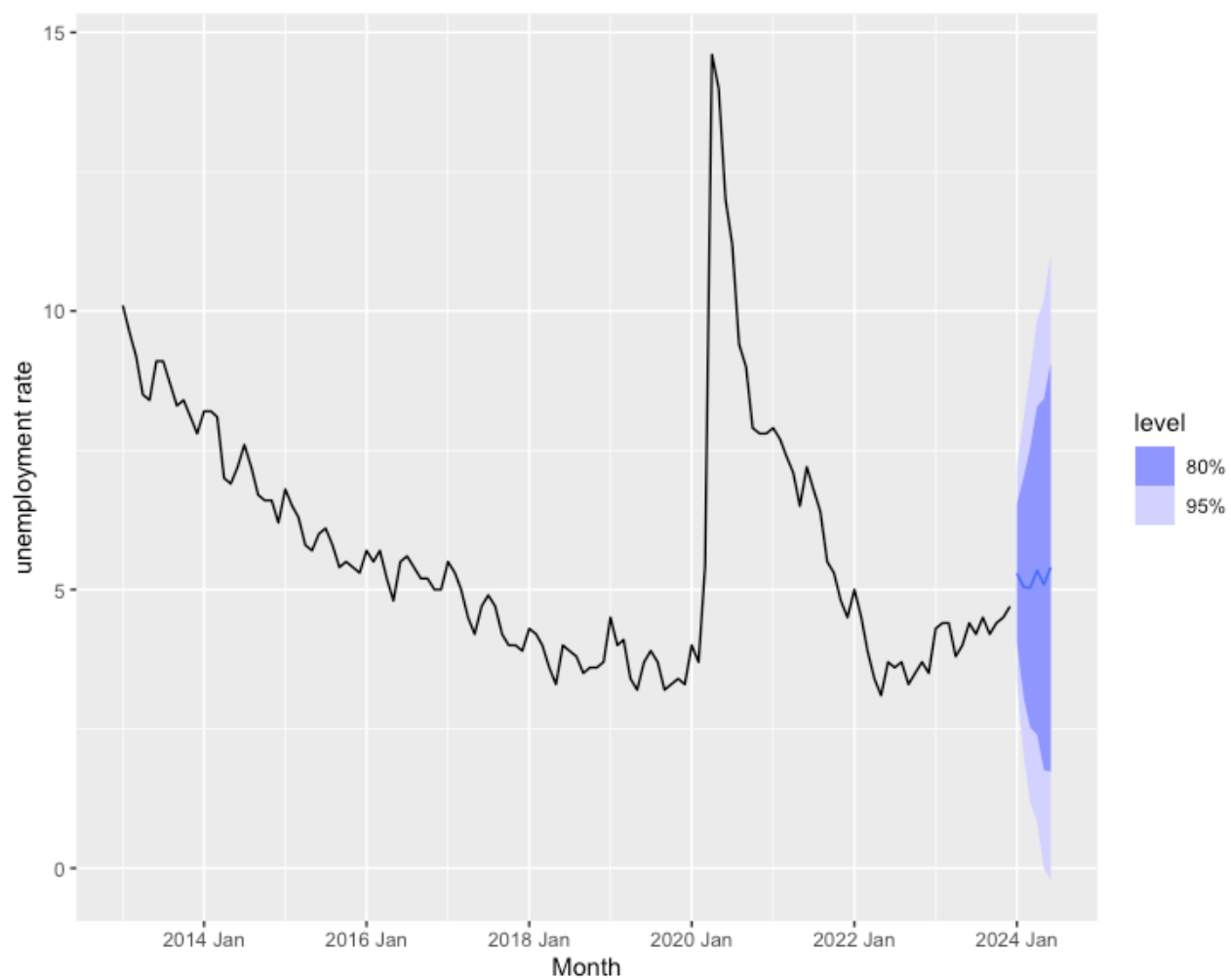
Amarillo Housing Market ARIMA Out of Sample 6 Month Forecast



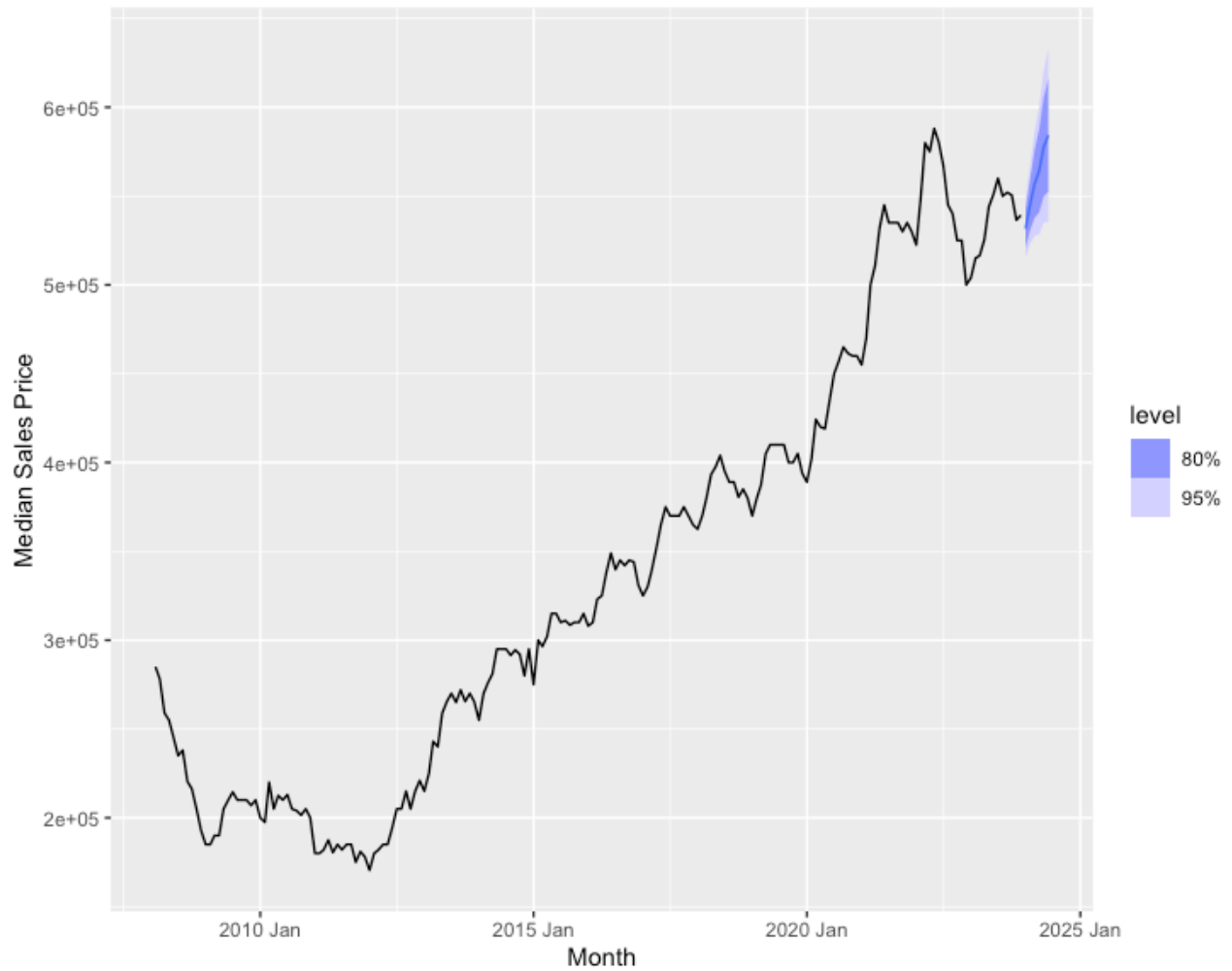
Sacramento Grapes ARIMA Out of Sample 6 Month Forecast



Sacramento Unemployment ARIMA Out of Sample 6 Month Forecast

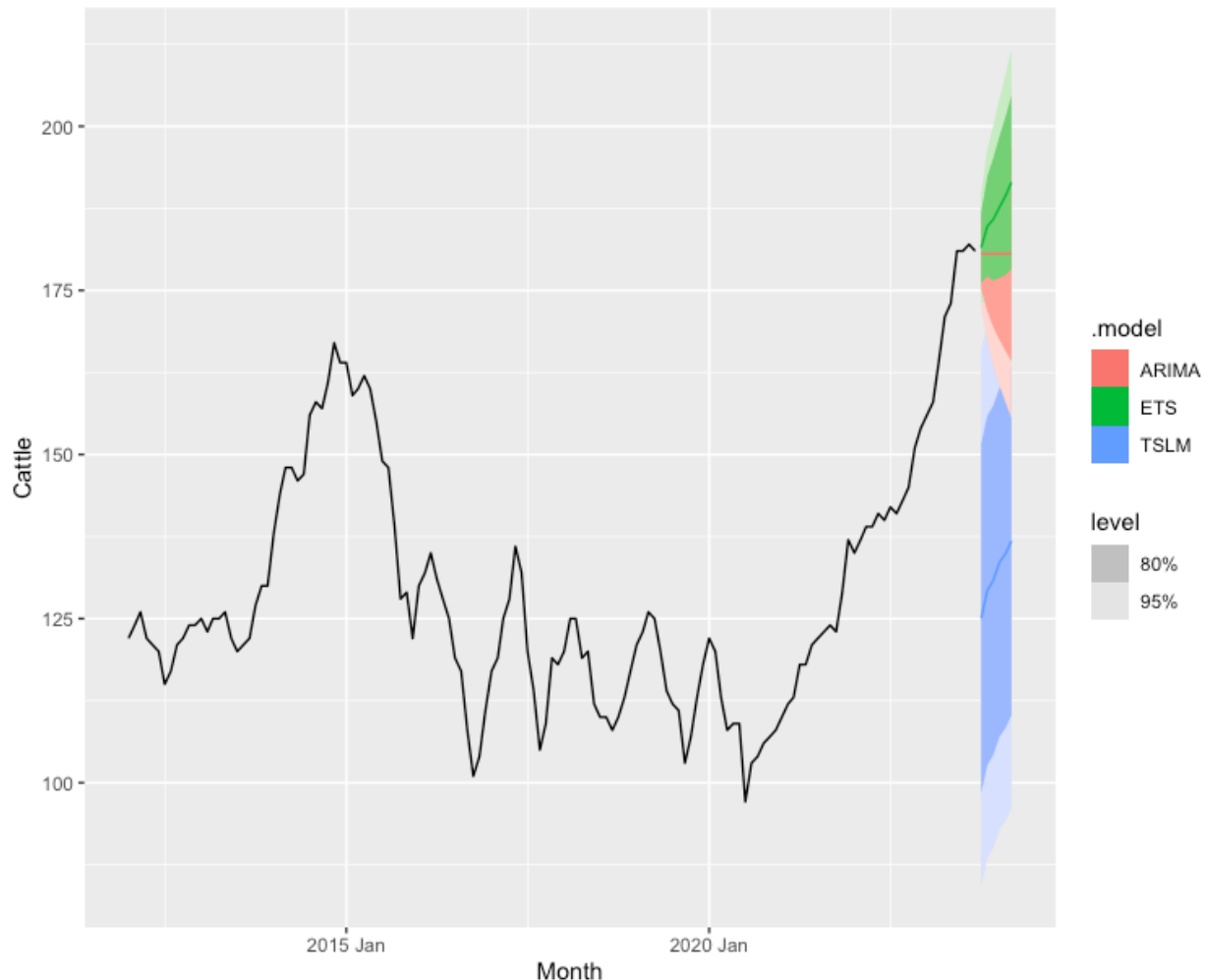


Sacramento Housing ARIMA Out of Sample 6 Month Forecast



Section 4- Compare your three “winning” models and combine forecasts in one graph. Which model would you choose among the three to forecast each time series?

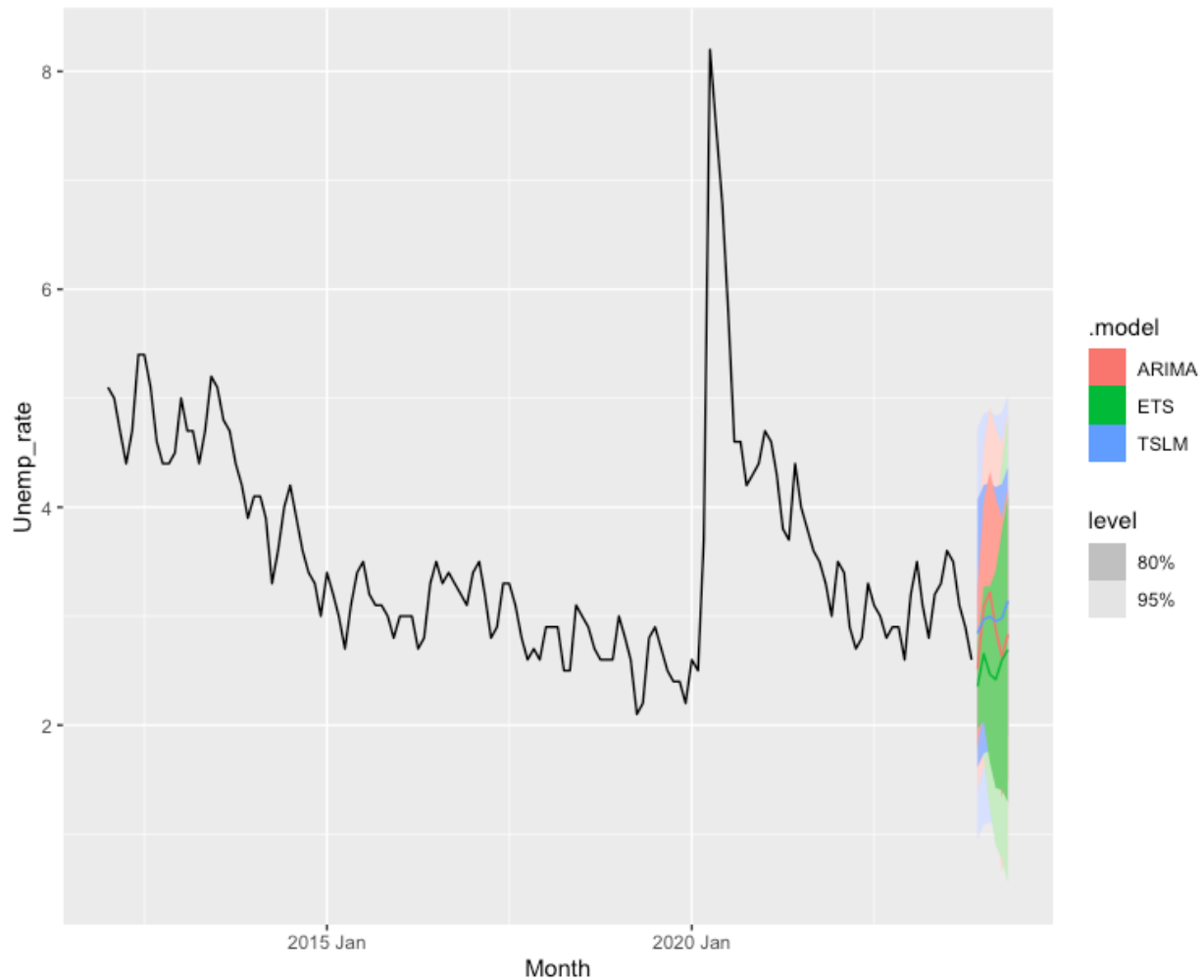
Amarillo Cattle Combined Forecasts



When choosing the “winning” model between the above forecasts, I would choose the ARIMA model. The ARIMA model for Amarillo cattle has the lowest AIC of all the other models. The choice between ARIMA, ETS, and trend models for time series forecasting is going to be determined by individual properties and forecasting requirements of the data. ARIMA provides the best fit to the data, maintains simplicity and is likely to lead to better predictive performance.



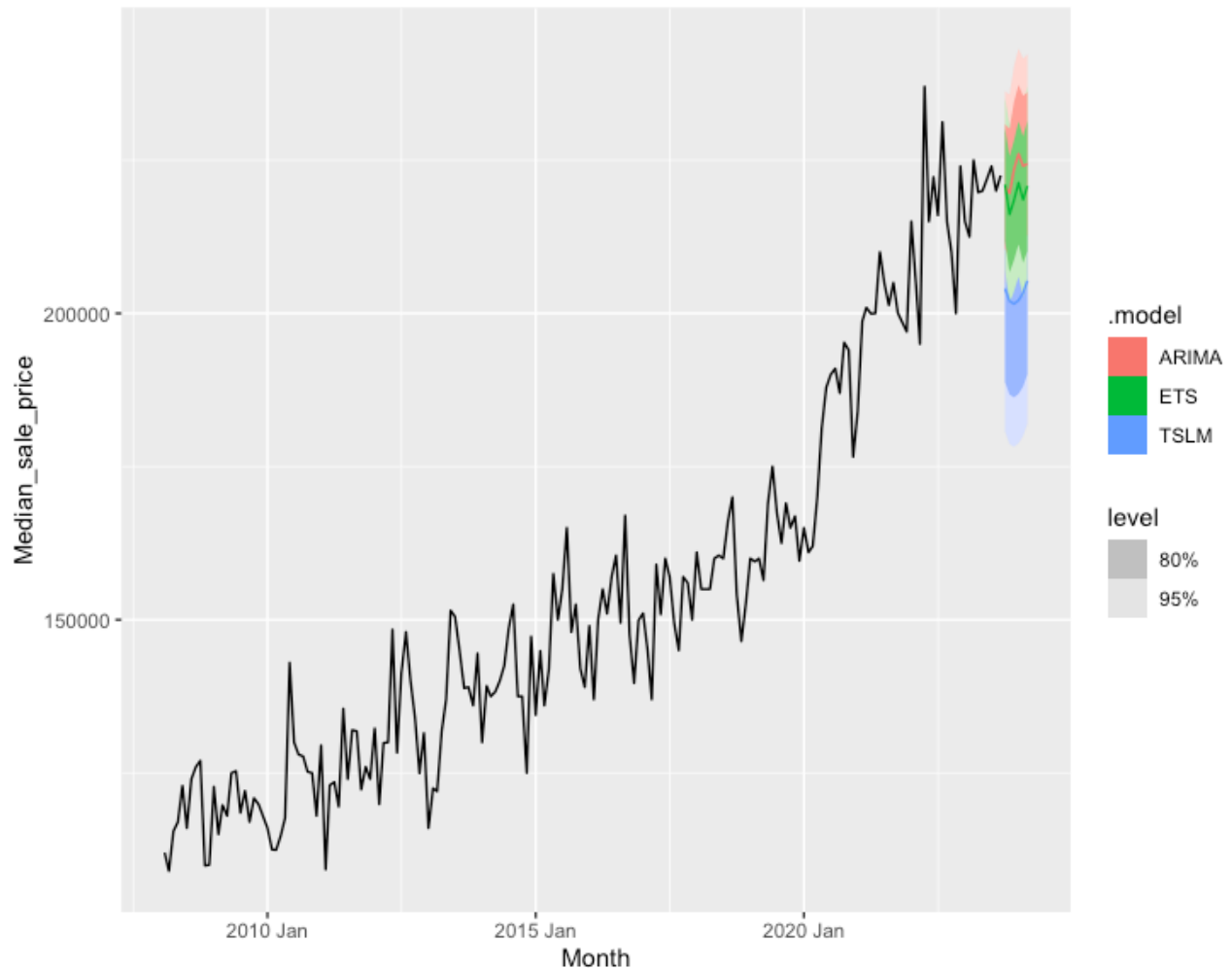
Amarillo Unemployment Combined Forecasts



When choosing the “winning” model between the above forecasts, I would choose the ARIMA model. The ARIMA model for Amarillo Unemployment has the lowest AIC, AICc and BIC of all the other models. The choice between ARIMA, ETS, and trend models for time series forecasting is going to be determined by individual properties and forecasting requirements of the data. ARIMA provides the best fit to the data, maintains simplicity and is likely to lead to better predictive performance.



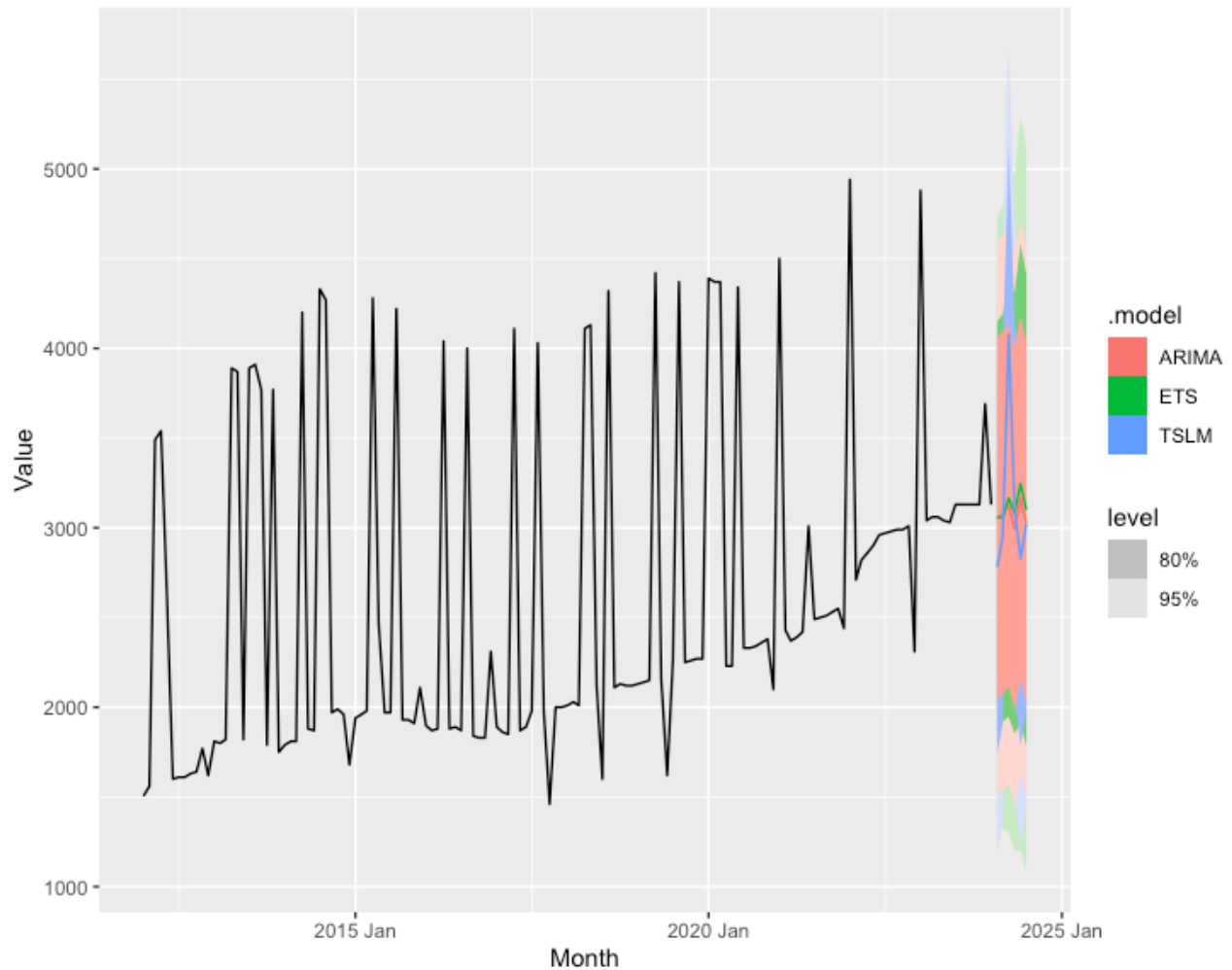
Amarillo Housing Market Combined Forecasts



When choosing the “winning” model between the above forecasts, I would choose the ARIMA model. The ARIMA model for the Amarillo Housing Market scores the best across all AICs of the other models, which typically indicates the best option. The choice between ARIMA, ETS, and trend models for time series forecasting is going to be determined by individual properties and forecasting requirements of the data. ARIMA provides the best fit to the data, maintains simplicity and is likely to lead to better predictive performance.



Sacramento Grapes Combined Forecasts



When choosing the “winning” model between the above forecasts, I would choose the ARIMA model. The ARIMA model for the Sacramento Grapes scores the best across all AICs of the other models, which typically indicates the best option. The choice between ARIMA, ETS, and trend models for time series forecasting is going to be determined by individual properties and forecasting requirements of the data. ARIMA provides the best fit to the data, maintains simplicity and is likely to lead to better predictive performance.



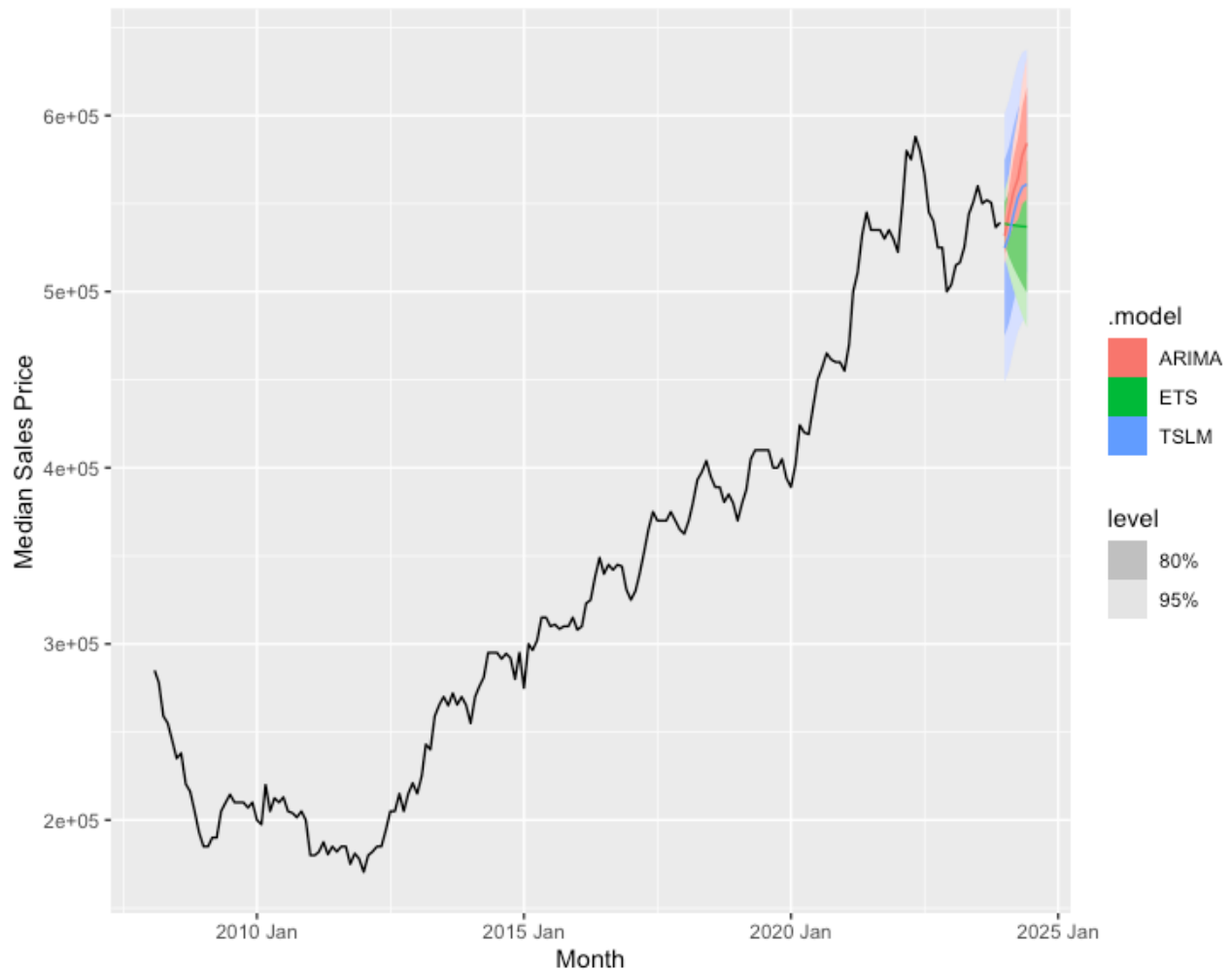
Sacramento Unemployment Rate Combined Forecasts



When choosing the “winning” model between the above forecasts, I would choose the ARIMA model. The ARIMA model for Sacramento Unemployment Rate has the lowest AIC of all the other models, which typically indicates the best option. The choice between ARIMA, ETS, and trend models for time series forecasting is going to be determined by individual properties and forecasting requirements of the data. ARIMA provides the best fit to the data, maintains simplicity and is likely to lead to better predictive performance.



Sacramento Housing Market Combined Forecasts



When choosing the “winning” model between the above forecasts, I would choose the ARIMA model. The ARIMA model for the Sacramento Housing Market has the lowest AIC of all the other models, which typically indicates the best option. The choice between ARIMA, ETS, and Fourier models for time series forecasting is going to be determined by individual properties and forecasting requirements of the data. ARIMA provides the best fit to the data, maintains simplicity and is likely to lead to better predictive performance.

