

Forecasting Monthly Housing Inventory Data

Data 690 Special Topics in Data Science : Time Series Forecasting
May 8th 2024

To :- Christopher McGraw

By :- Shreya, Sudini (LJ83576)

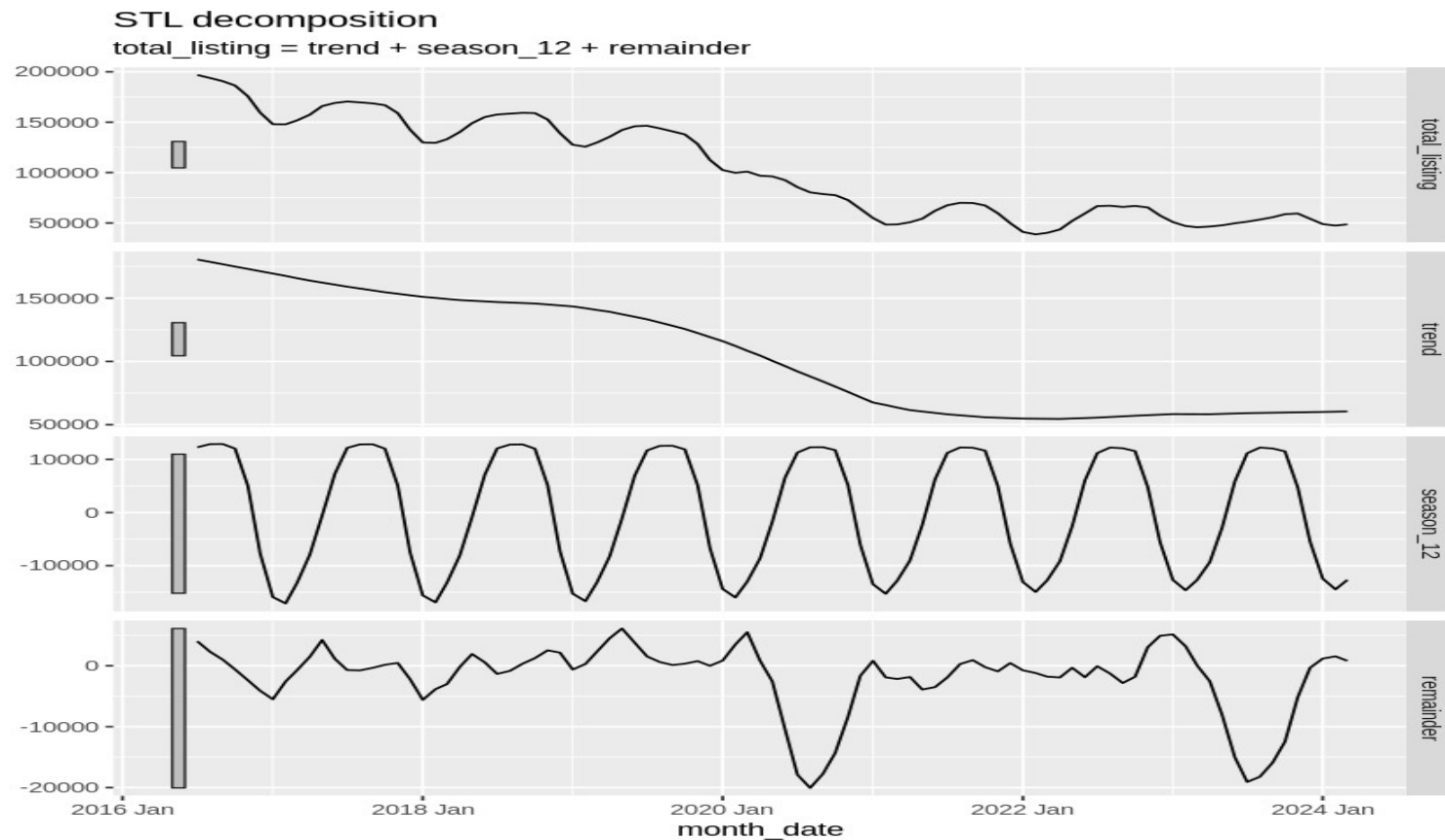
/ Problem Statement

Produce forecasts at the county level for MD, VA, PA, DE, and NJ. Use a variety of techniques to forecast one metric for the county, state, and five-state aggregate. You can only use the metrics that are counts for this option.

/ Data Source & Preparation

- Dataset :- montly_housing_inventory.csv
 - Metric:- active_listing_count
 - States:- MD, PA, VA, DE, NJ
-

STL Decomposition of the Metric



/ Outliers and Missing values

- There are no considerable outliers
 - 45 missing values in the active listing count
 - Filled with interpolate and ARIMA model
-

Models Used for Forecasting

- **ARIMA**
 - Flexibility
 - Integrated Approach
 - Interpretability
 - **ETS**
 - Robust to Noise
 - Ease of Use
 - **NAÏVE**
 - Simplicity
 - Baseline Model:
-

Forecast for Five State Aggregate Data



/ Modelling

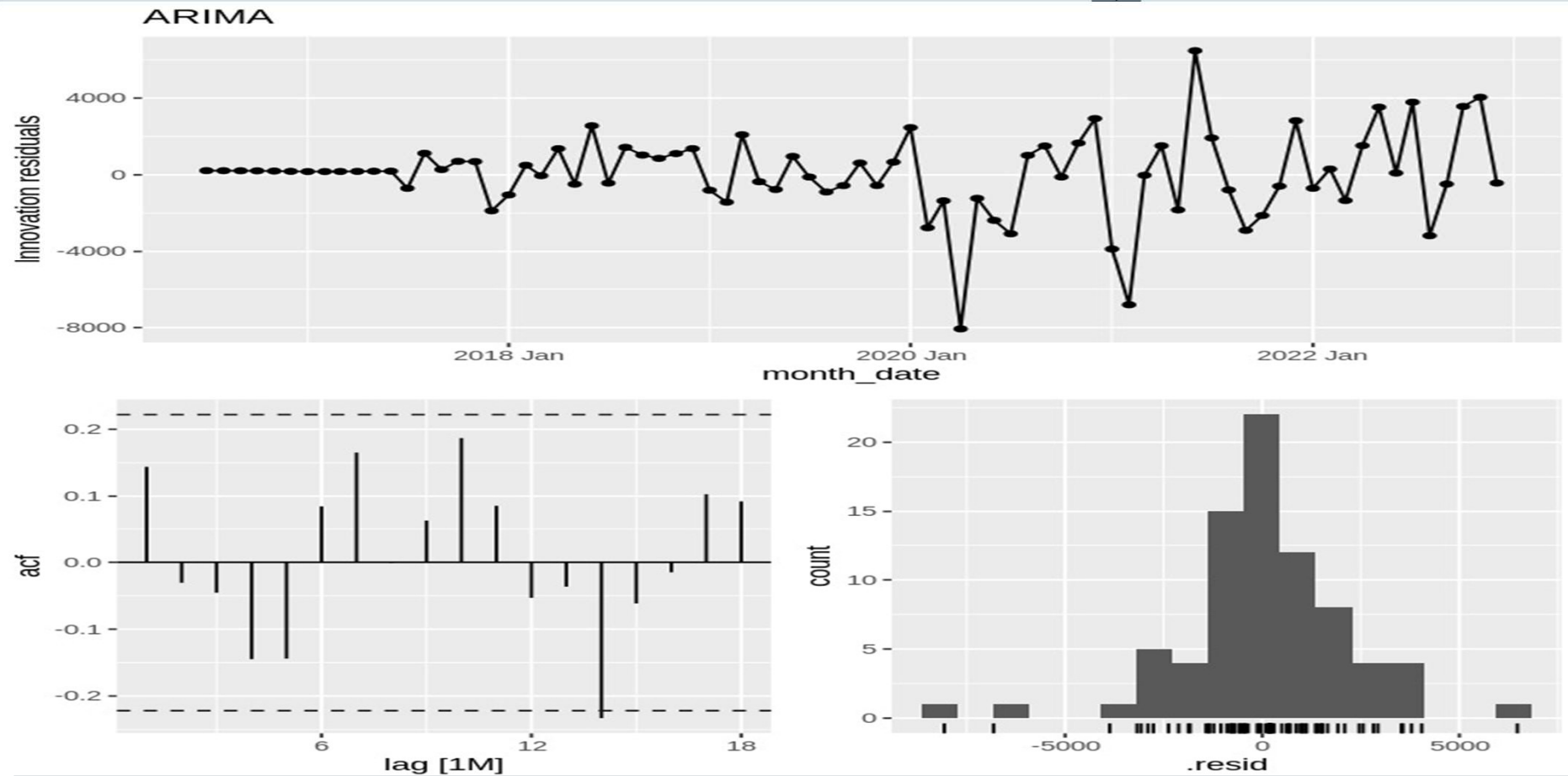
+

A tibble: 3 × 3

.model	RMSE	MAPE
<chr>	<dbl>	<dbl>
ARIMA	2125.340	1.903592
ETS	2670.962	2.476365
NAIVE	6511.471	5.480053

- Training Set : 2016 – 2022 December

Residuals



/ Modelling

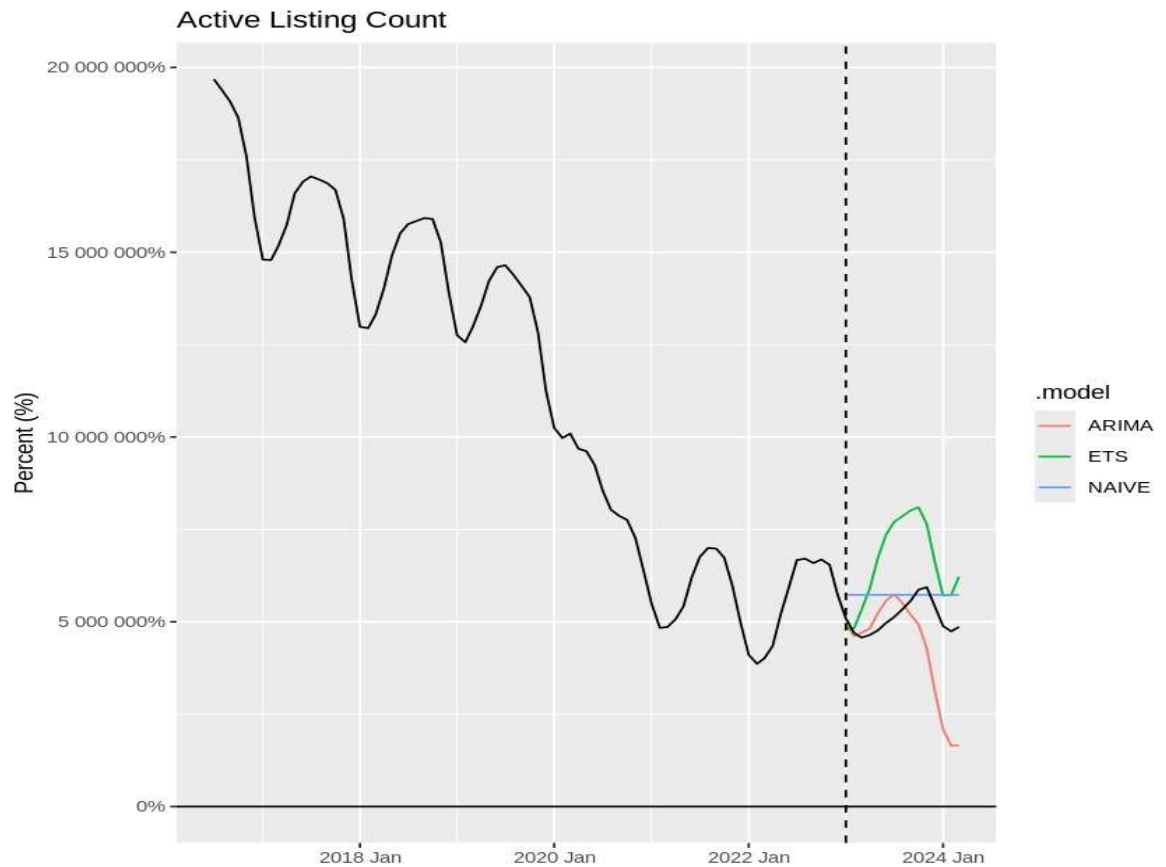
- Testing Set :- 2023 Jan – 2024 March

A tibble: 3 × 3

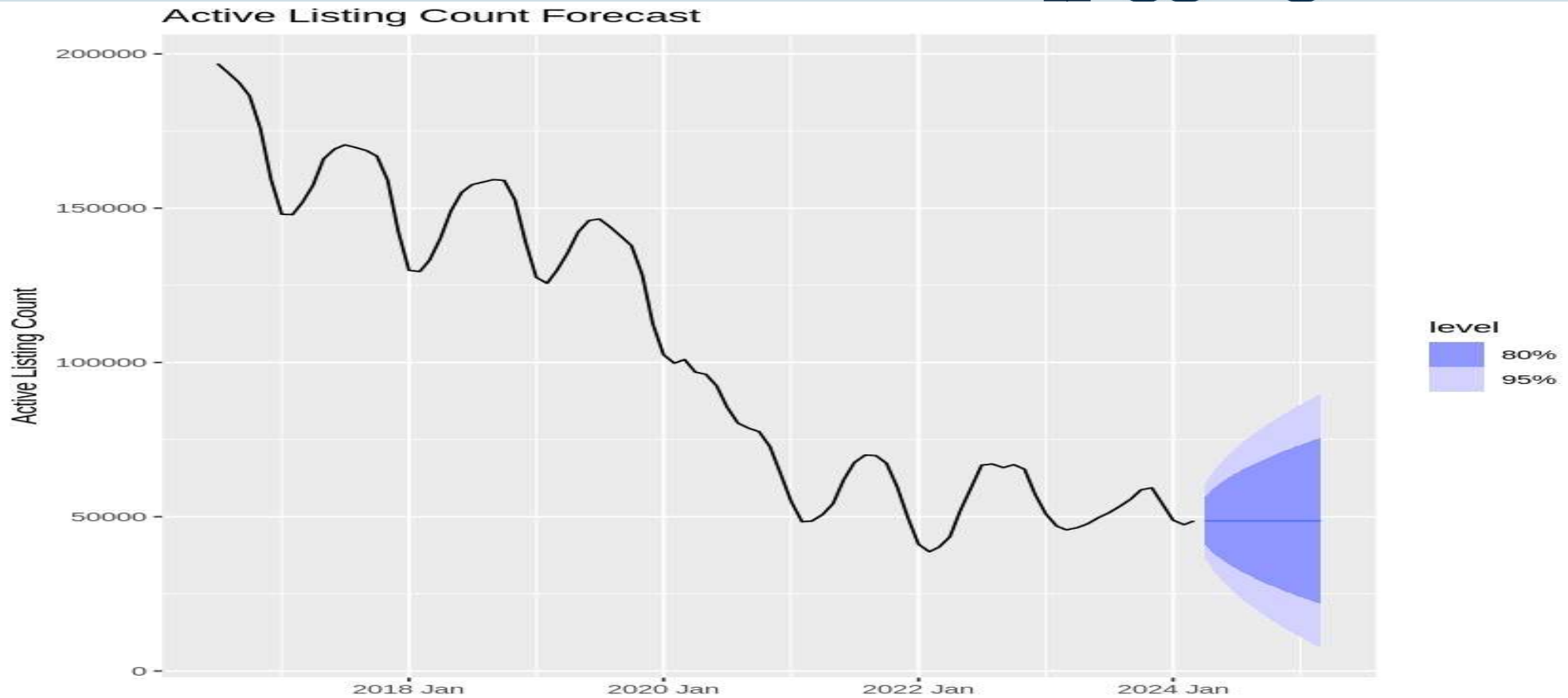
.model	RMSE	MAPE
<chr>	<dbl>	<dbl>
ARIMA	15859.099	21.99779
ETS	17039.648	29.12416
NAIVE	7600.066	13.90825

Forecasts

- On Test Set



Final Forecasts for 5 State Aggregate



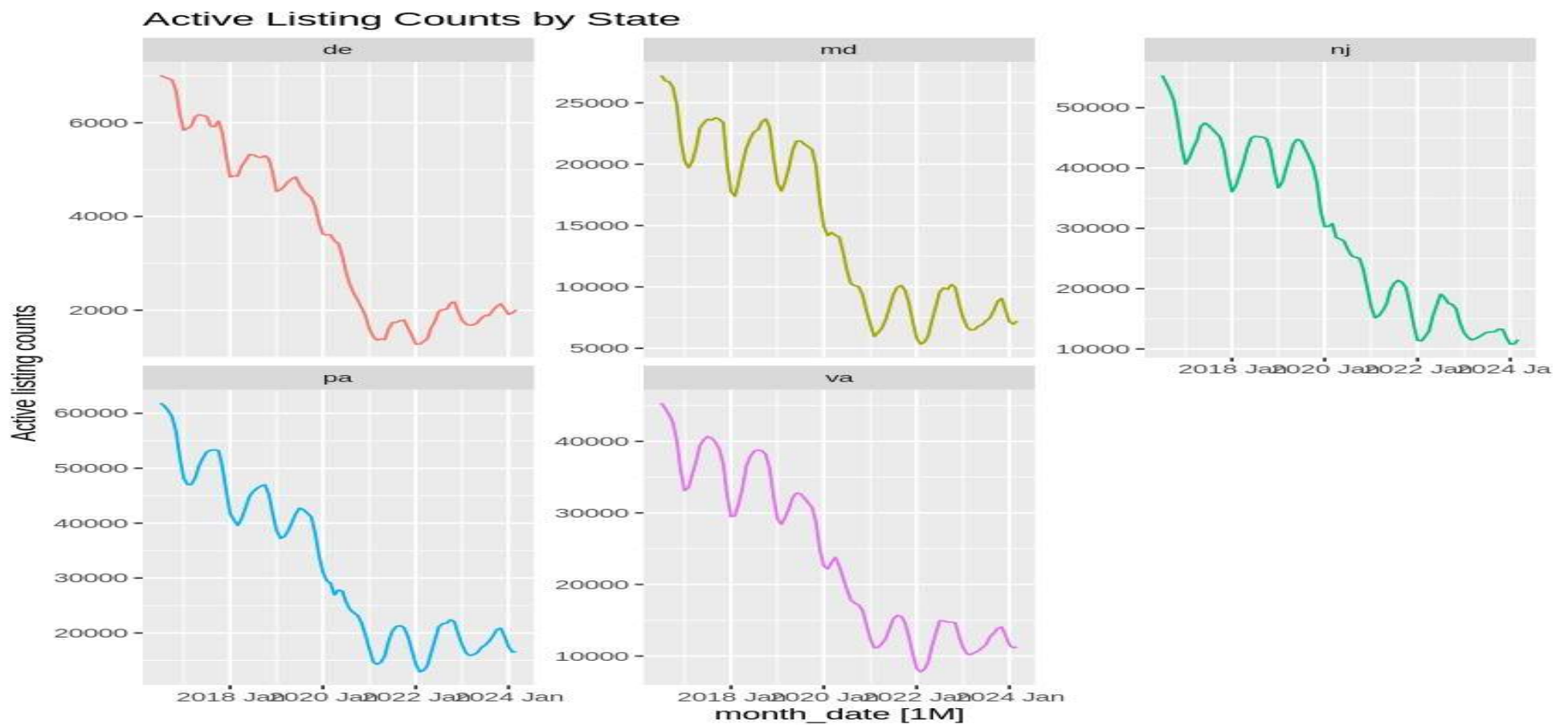
State Forecast

A tbl_ts: 6 × 3

month_date	state	total_listing
<mth>	<chr*>	<dbl>
2016 Jul	<aggregated>	196832
2016 Aug	<aggregated>	193834
2016 Sep	<aggregated>	190681
2016 Oct	<aggregated>	186352
2016 Nov	<aggregated>	175776
2016 Dec	<aggregated>	159476

- Data

Visualizations



Modelling for Maryland State

Training Set

A tibble: 3 × 3

<code>.model</code>	<code>RMSE</code>	<code>MAPE</code>
<code><chr></code>	<code><dbl></code>	<code><dbl></code>
ARIMA	883.5645	2.785318
ETS	1037.3064	3.463531
NAIVE	2244.5827	6.624438

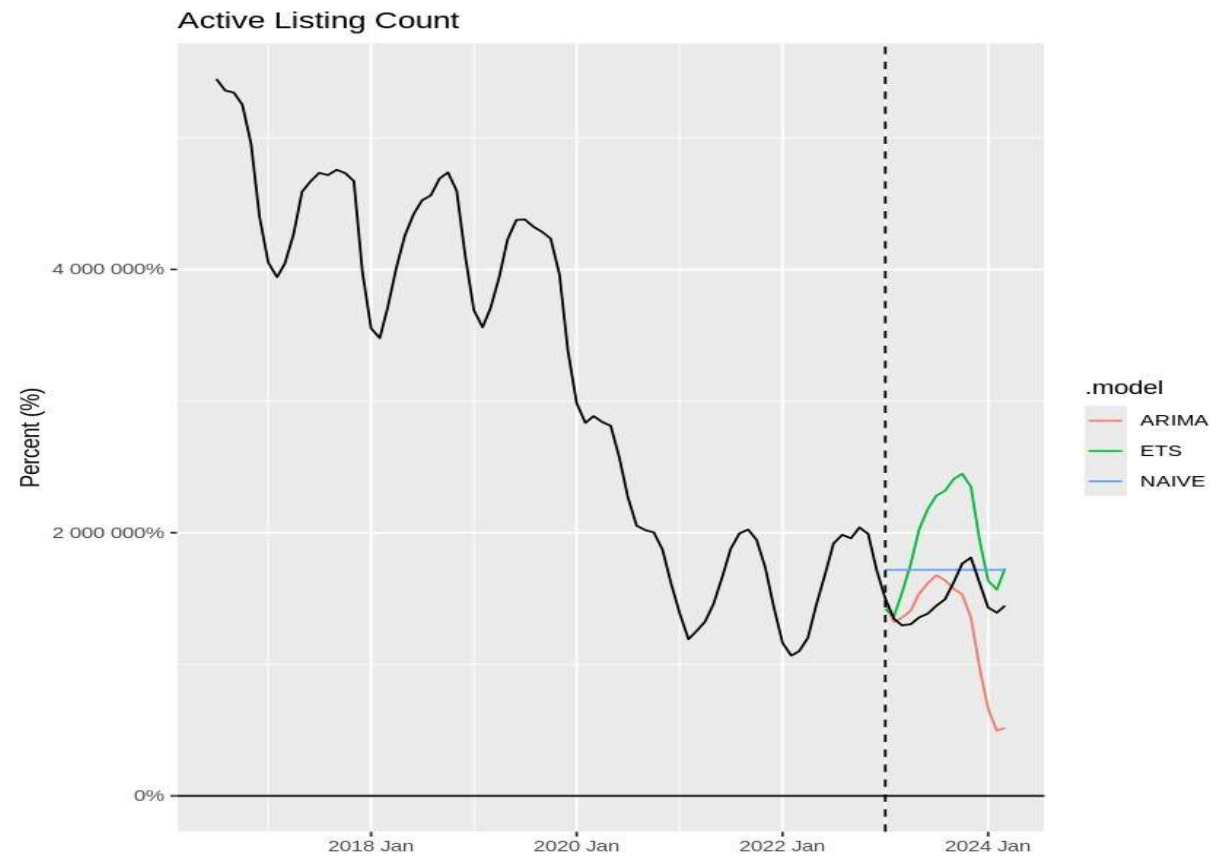
Testing Set

A tibble: 3 × 3

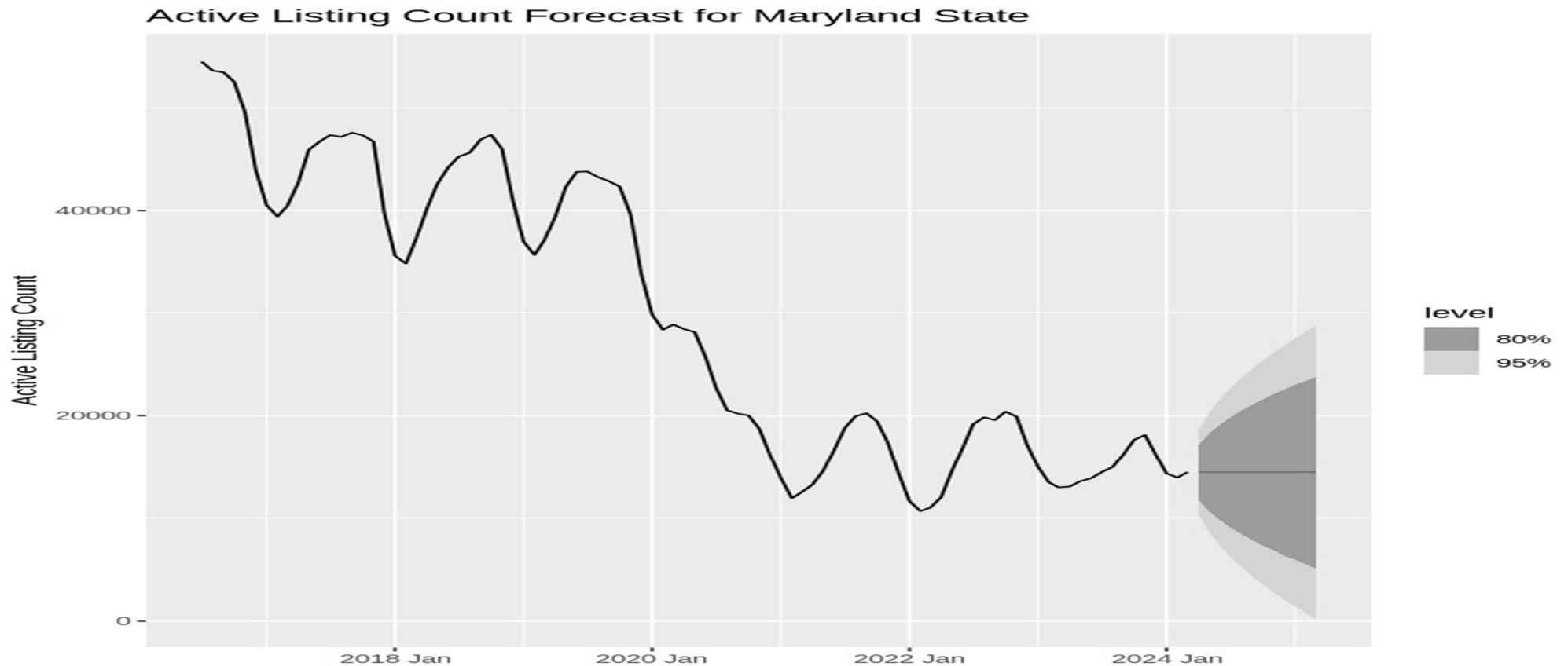
<code>.model</code>	<code>RMSE</code>	<code>MAPE</code>
<code><chr></code>	<code><dbl></code>	<code><dbl></code>
ARIMA	4552.876	22.40922
ETS	5394.558	30.93497
NAIVE	2817.013	18.17258

Forecast

- On Test Set



Out of Sample Forecast for MD State



Modelling for Virginia State

Training Set

A tibble: 3 × 3

.model	RMSE	MAPE
<chr>	<dbl>	<dbl>
ARIMA	1036.102	2.117894
ETS	1598.267	3.254986
NAIVE	3352.477	6.165440

Testing Set

A tibble: 3 × 3

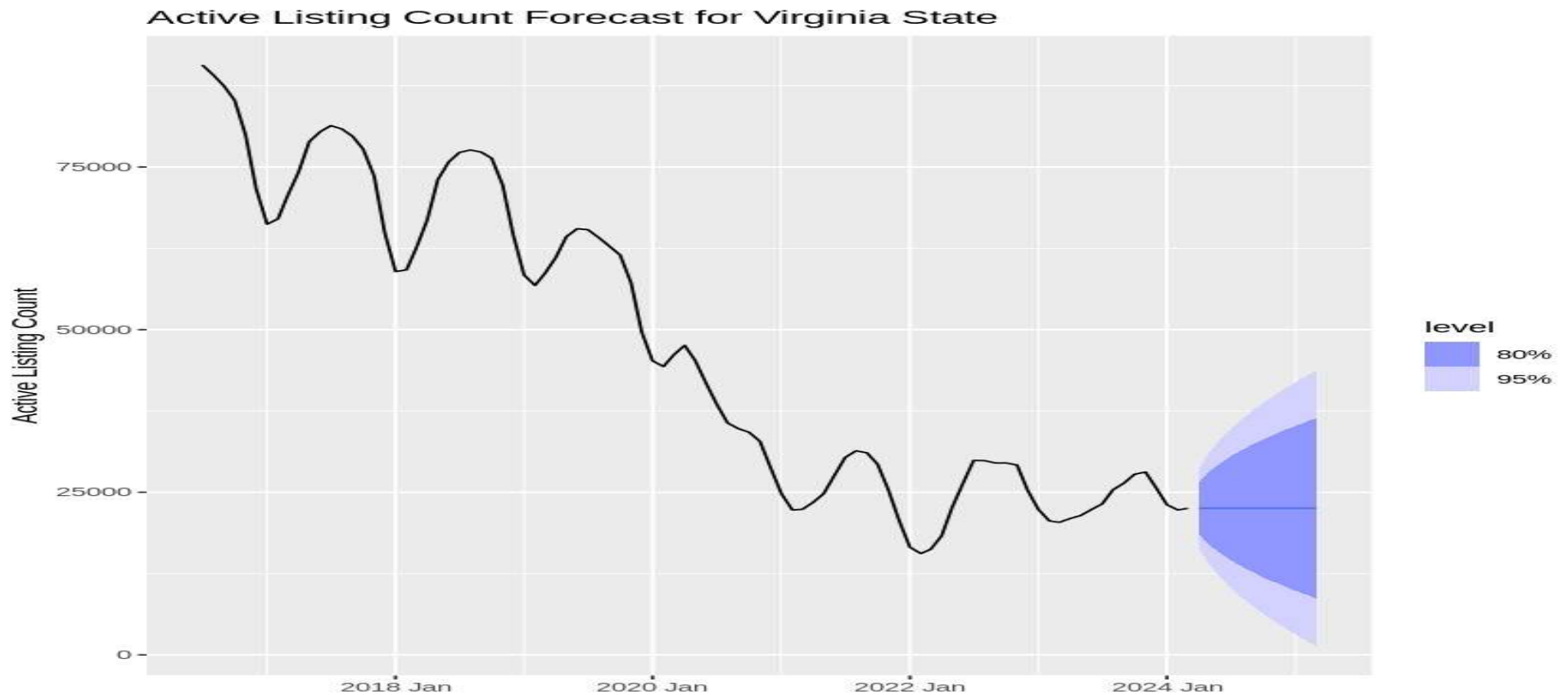
.model	RMSE	MAPE
<chr>	<dbl>	<dbl>
ARIMA	8819.108	28.47173
ETS	8872.994	33.16598
NAIVE	3069.868	12.18663

/ Forecast

- On Test Set



Out of Sample Forecast for VA State



Modelling for Pennsylvania State

Training Set

A tibble: 3 × 3

<code>.model</code>	<code>RMSE</code>	<code>MAPE</code>
<code><chr></code>	<code><dbl></code>	<code><dbl></code>
ARIMA	1442.814	2.028322
ETS	1452.443	1.957435
NAIVE	3767.456	5.252145

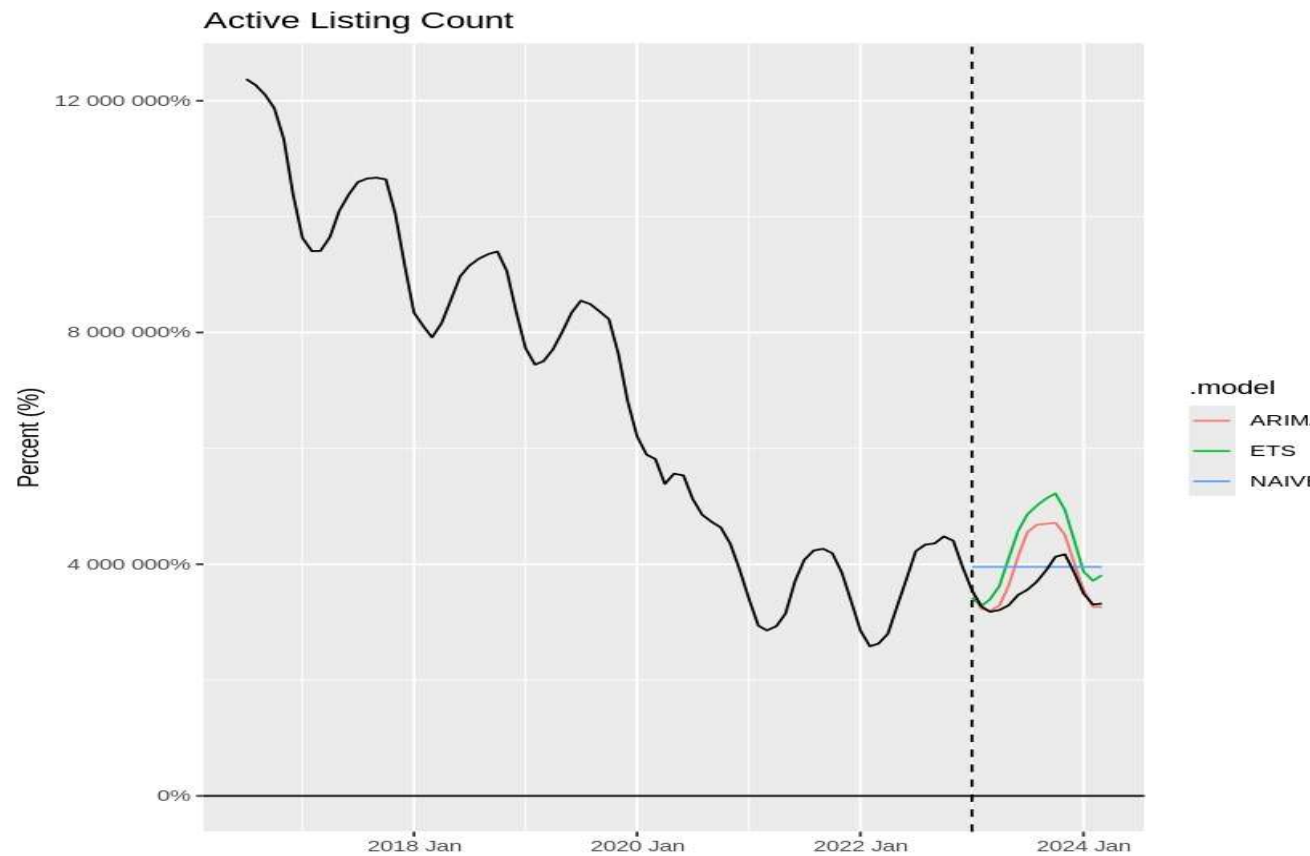
Testing Set

A tibble: 3 × 3

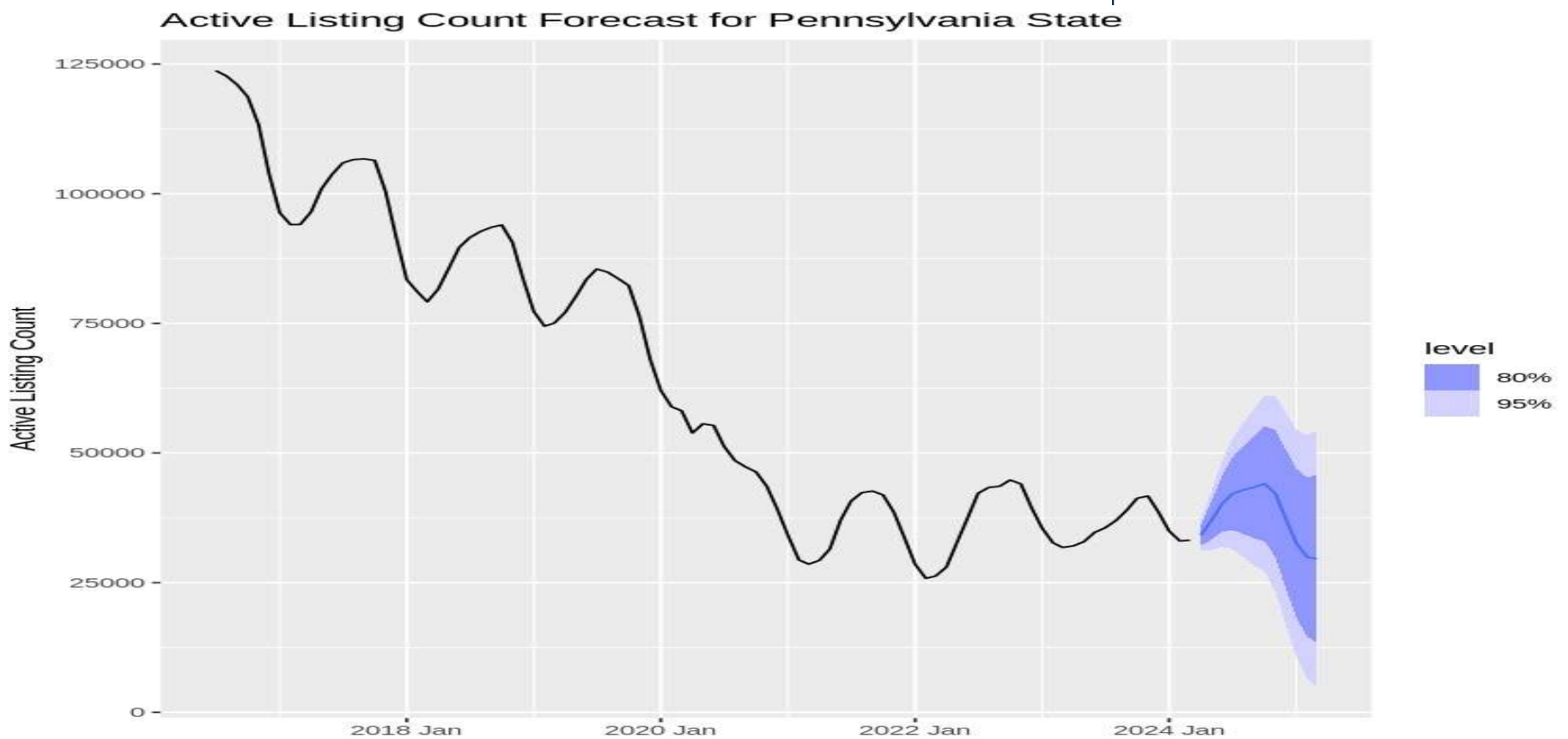
<code>.model</code>	<code>RMSE</code>	<code>MAPE</code>
<code><chr></code>	<code><dbl></code>	<code><dbl></code>
ARIMA	4936.519	9.38575
ETS	8046.889	18.76447
NAIVE	5040.283	13.15297

Forecast

- On Test Set



Out of Sample Forecast for PA State



Modelling for Delaware State

Training Set

A tibble: 3 × 3

<code>.model</code>	<code>RMSE</code>	<code>MAPE</code>
<code><chr></code>	<code><dbl></code>	<code><dbl></code>
ARIMA	199.7413	2.768512
ETS	208.8094	2.875389
NAIVE	361.8855	4.502324

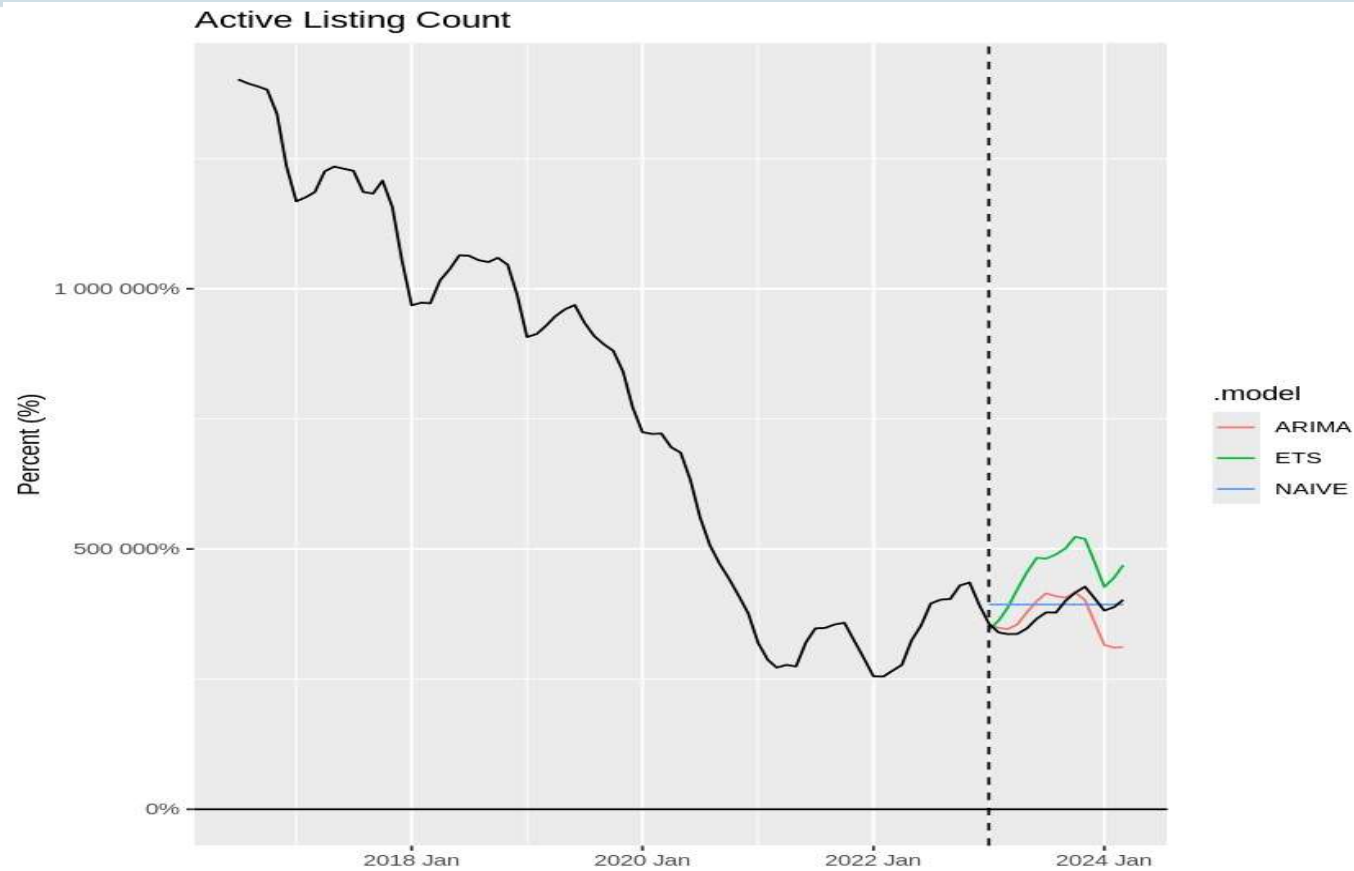
Testing Set

A tibble: 3 × 3

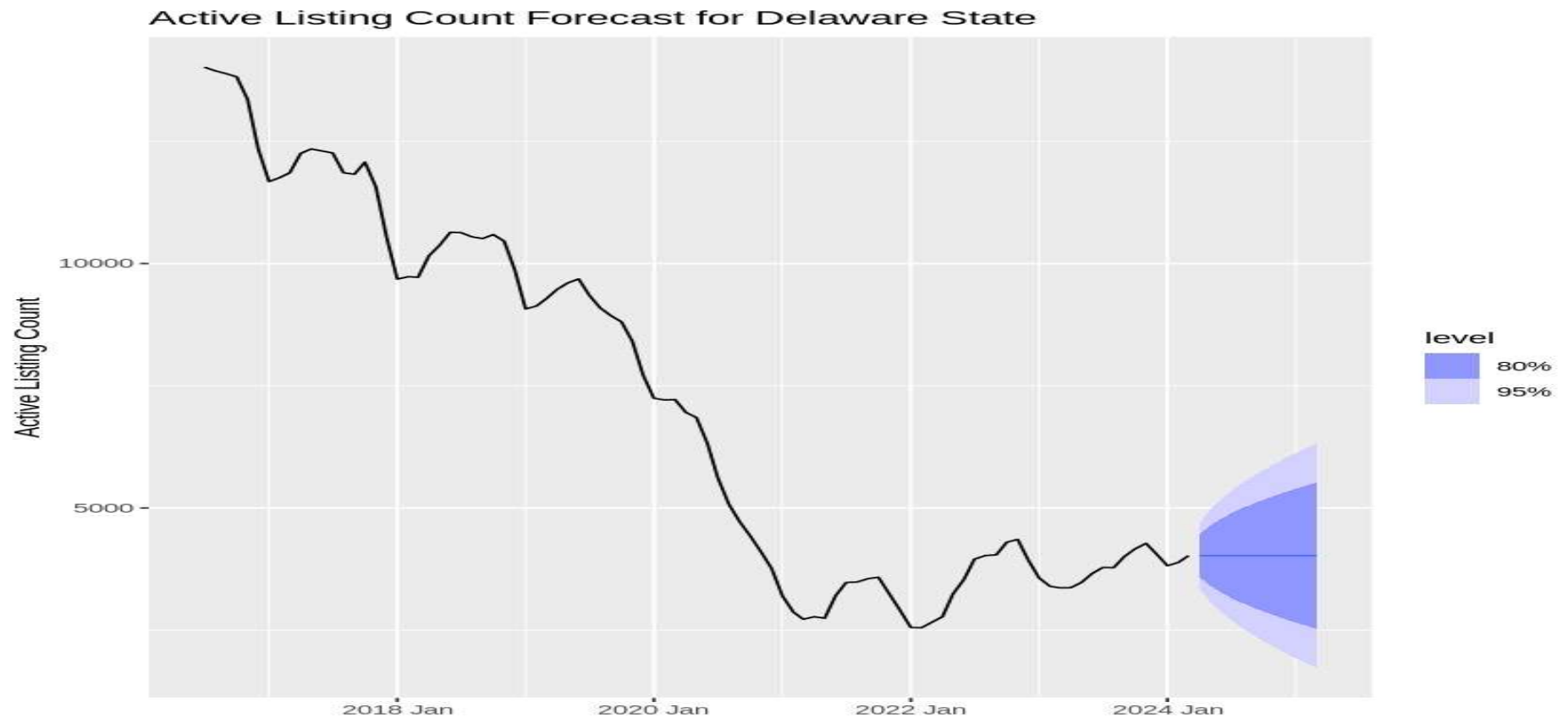
<code>.model</code>	<code>RMSE</code>	<code>MAPE</code>
<code><chr></code>	<code><dbl></code>	<code><dbl></code>
ARIMA	419.5162	8.440094
ETS	831.0740	20.189967
NAIVE	328.2710	7.577924

Forecast

- On Test Set



Out of Sample Forecast for DE State



Modelling for New Jersey State

Training Set

A tibble: 3 × 3

<code>.model</code>	<code>RMSE</code>	<code>MAPE</code>
<code><chr></code>	<code><dbl></code>	<code><dbl></code>
ARIMA	1362.770	1.980402
ETS	1639.513	2.484375
NAIVE	3782.200	5.699805

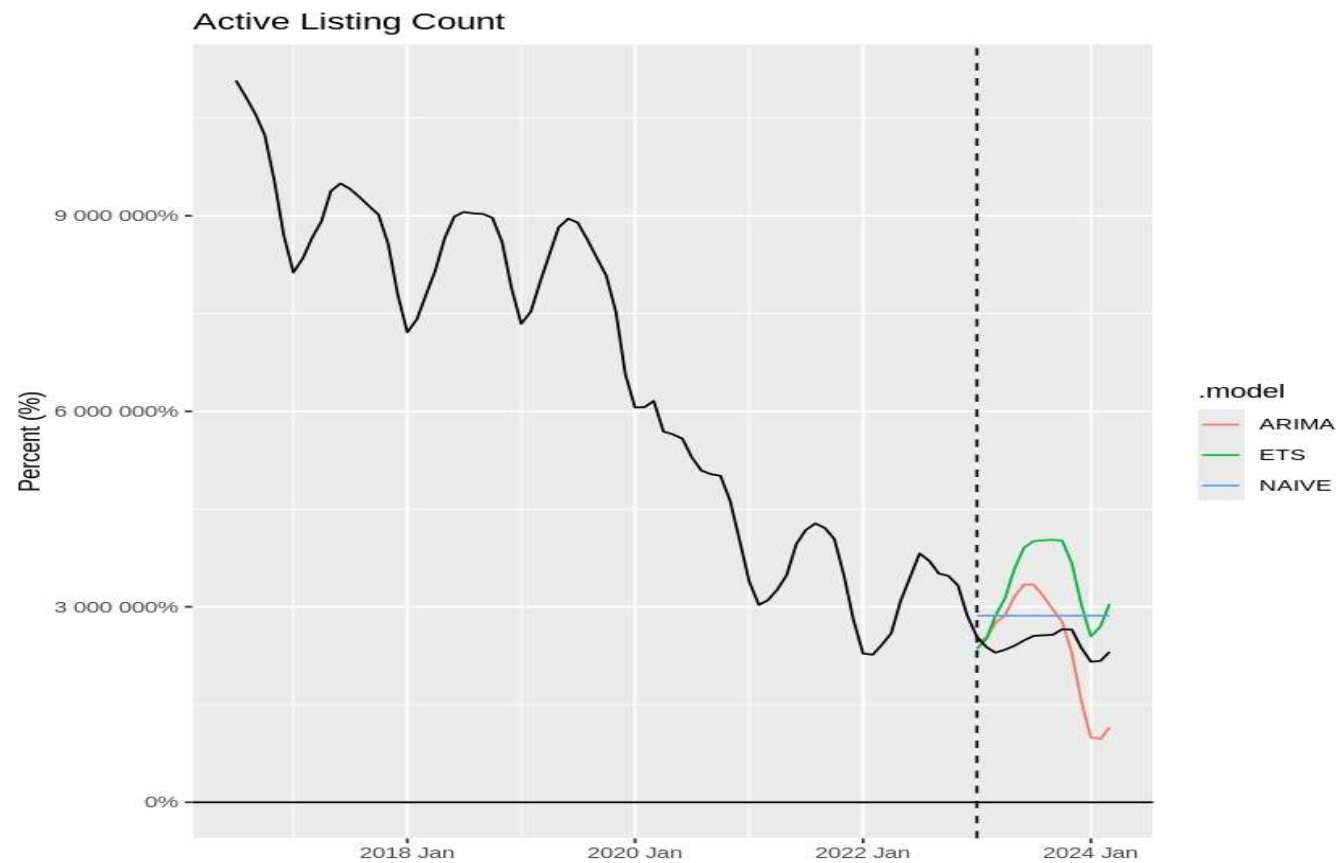
Testing Set

A tibble: 3 × 3

<code>.model</code>	<code>RMSE</code>	<code>MAPE</code>
<code><chr></code>	<code><dbl></code>	<code><dbl></code>
ARIMA	7245.257	26.56299
ETS	10029.780	36.01875
NAIVE	4604.539	18.34390

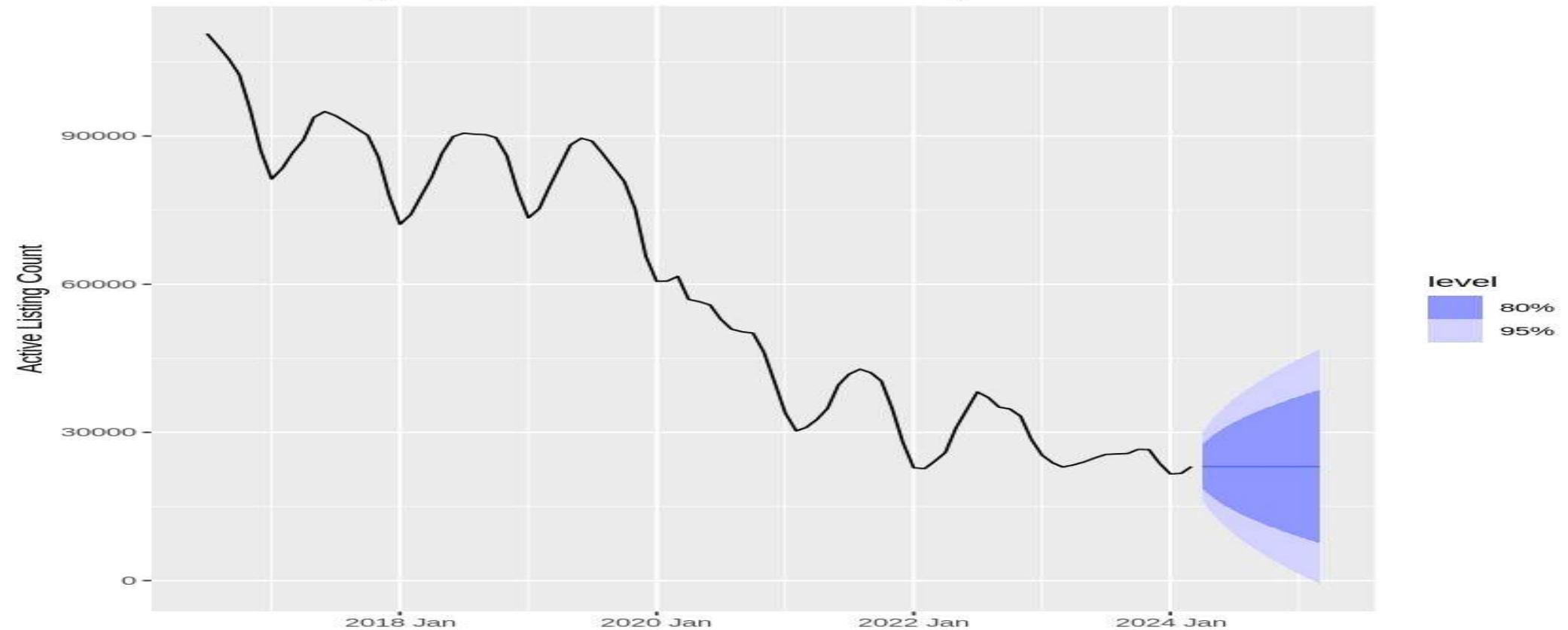
Forecast

- On Test Set



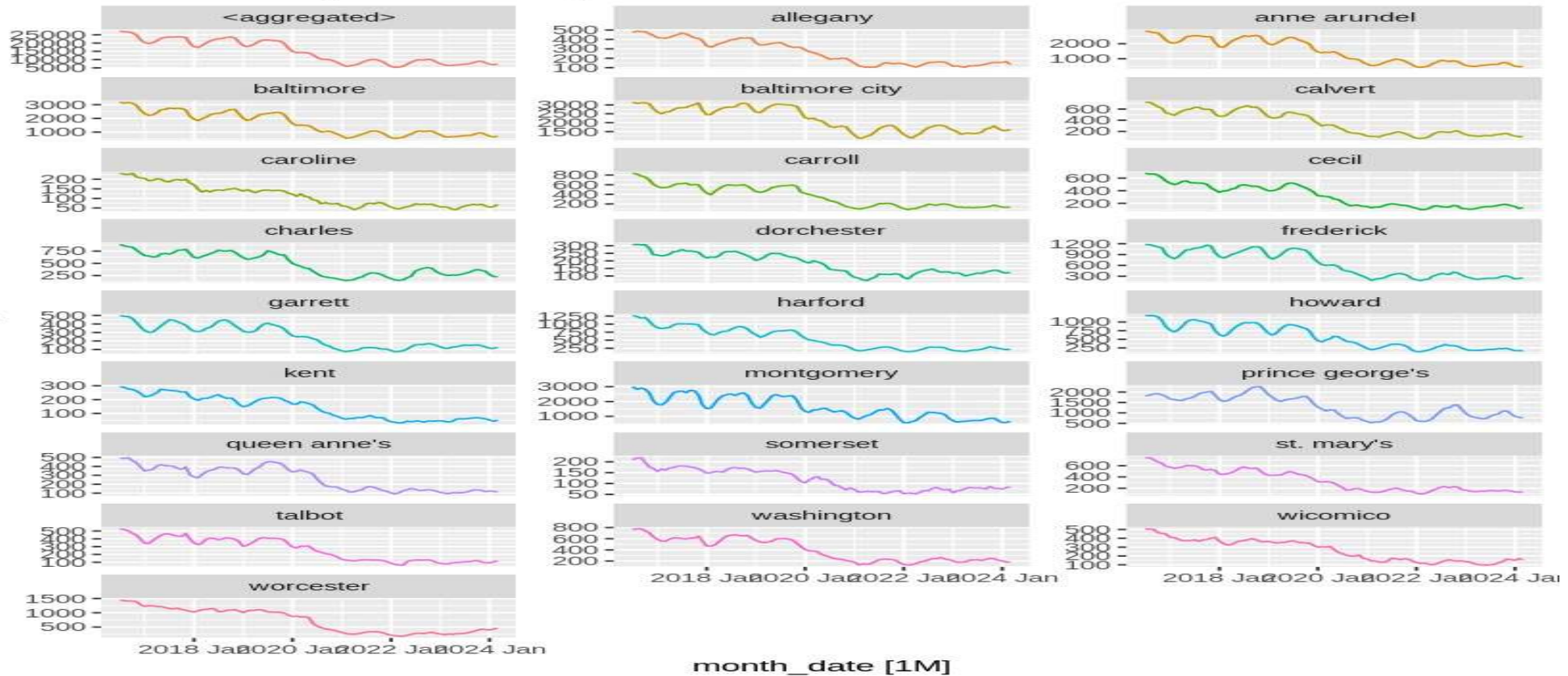
Out of Sample Forecast for NJ State

Active Listing Count Forecast for New Jersey State

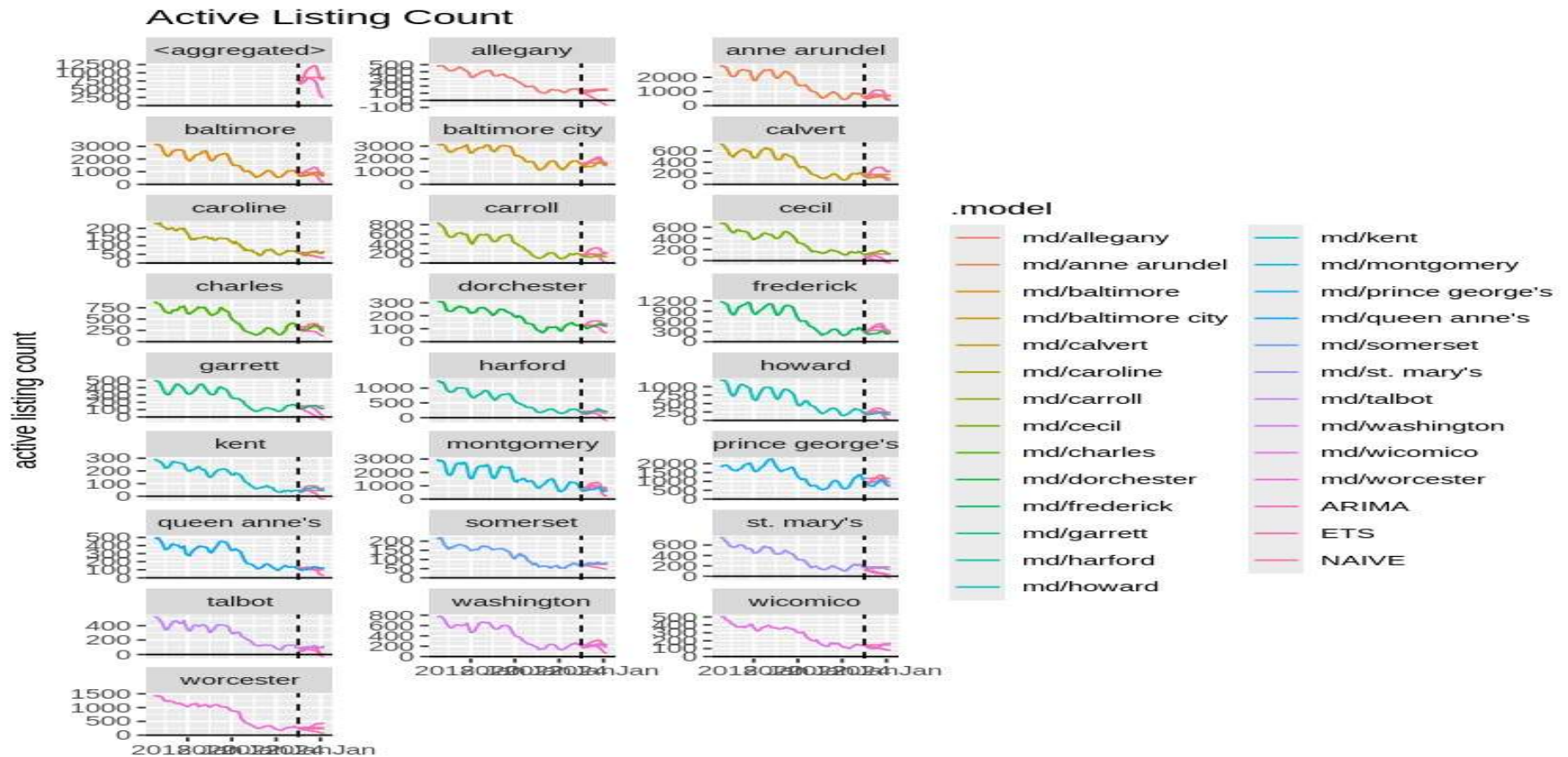


Counties in Maryland

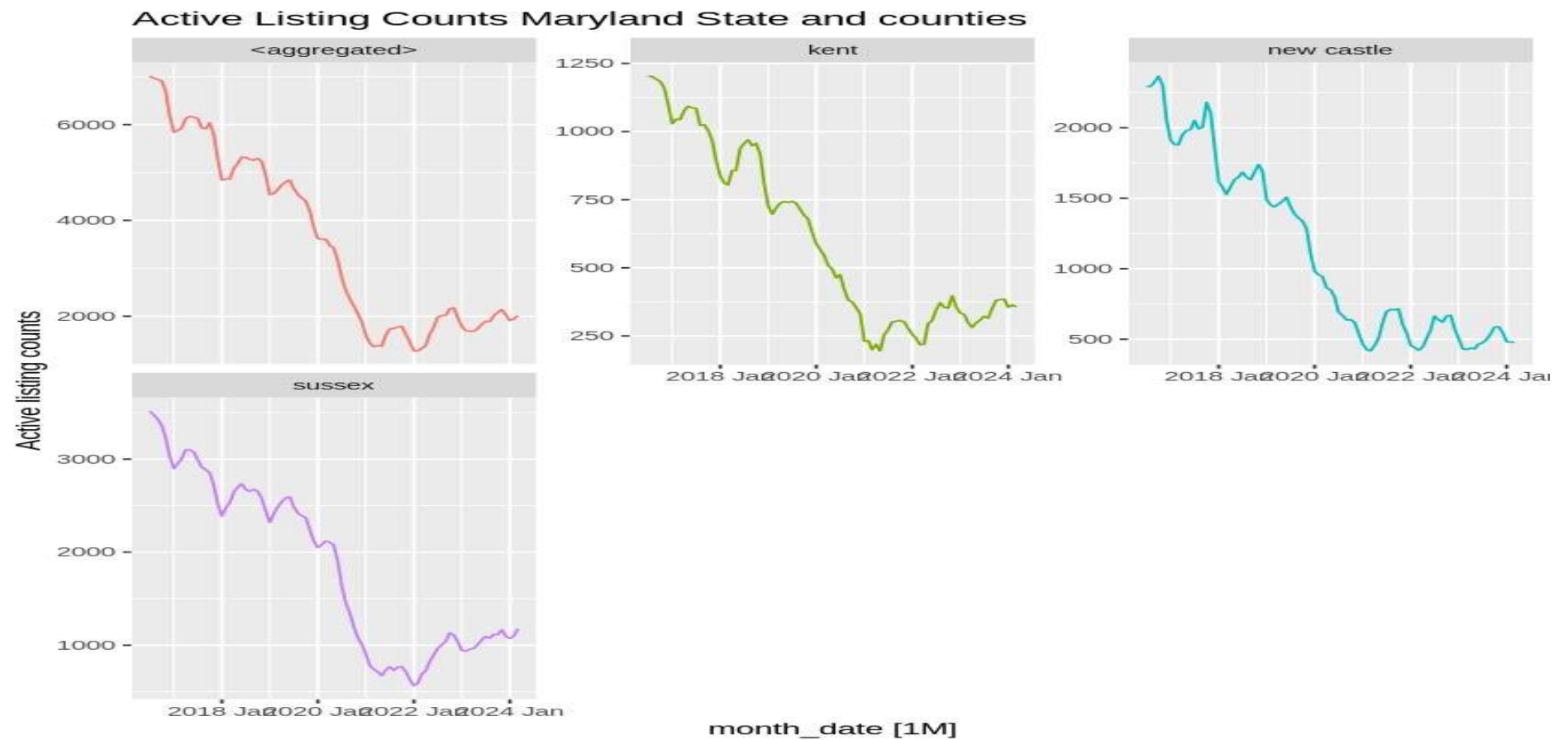
Active Listing Counts Maryland State and counties



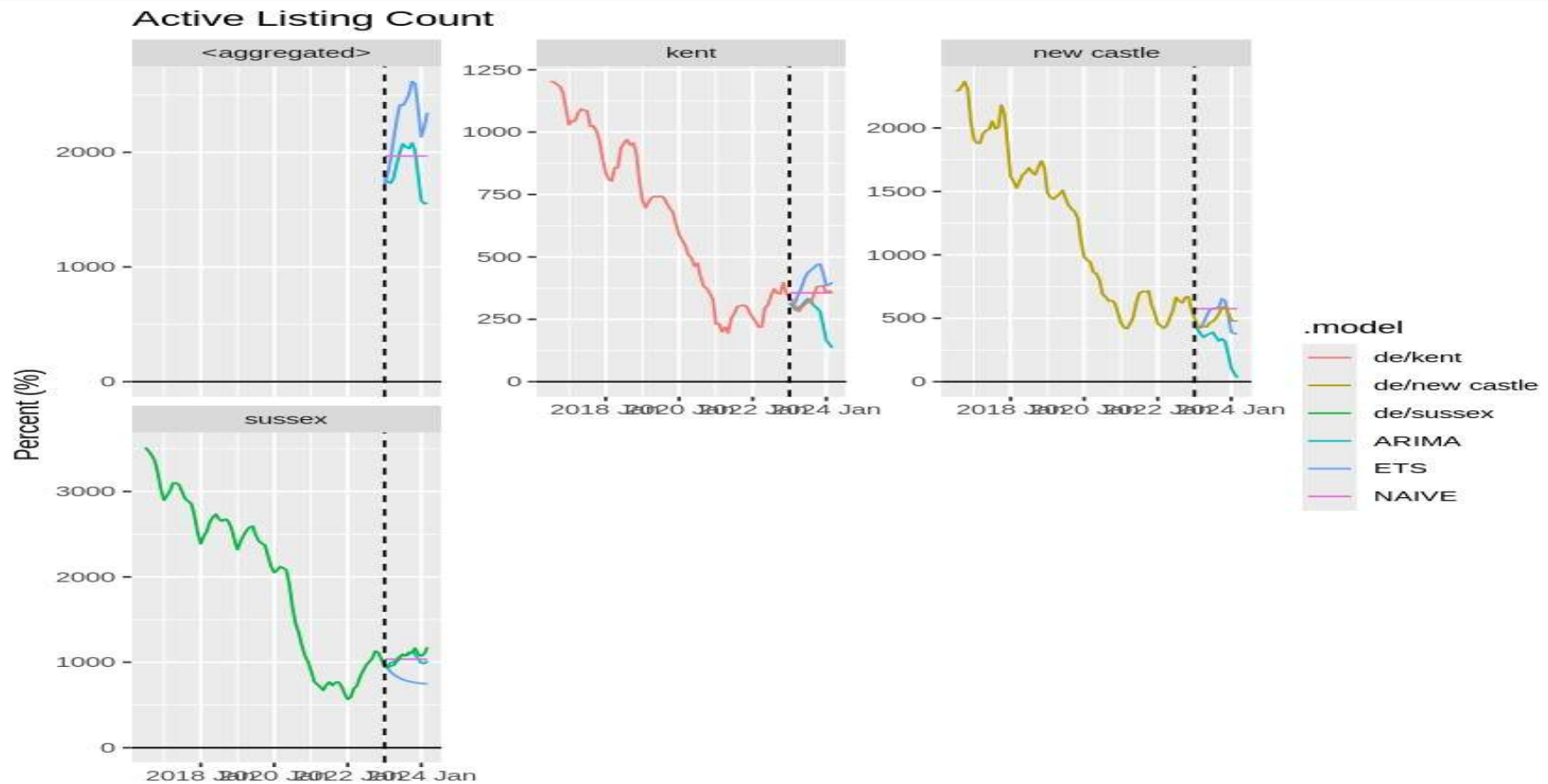
Forecast on Testing Data Set, MD



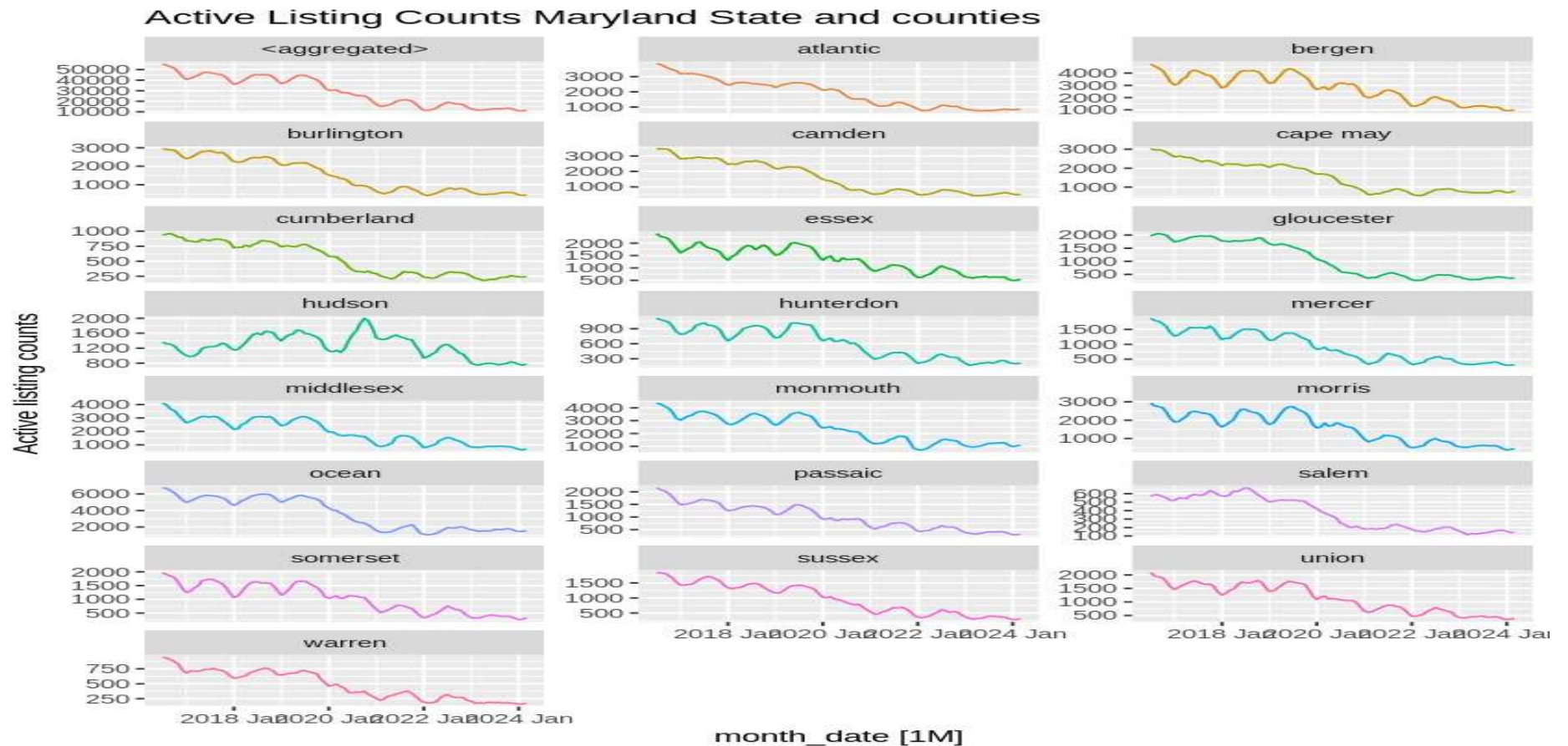
Counties in Delaware



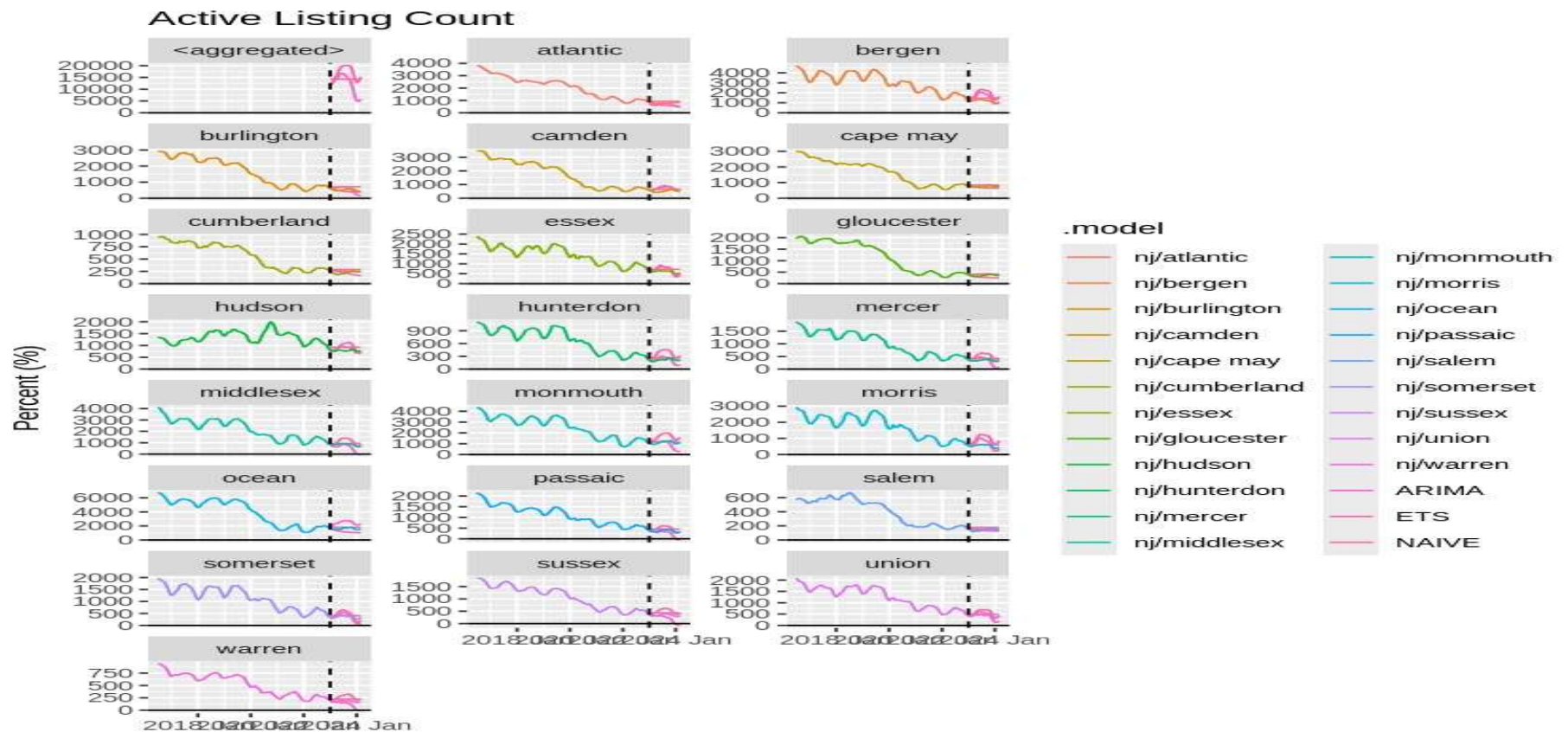
Forecast on Testing Data Set, DE



Counties in New Jersey



Forecast on Testing Data Set, NJ



/

Thank You!
