BAFT 2023

Forecasting Perishable Food Sales Quantity for Efficient Inventory Distribution to Large Retail Stores

Team 7

112078503 Audrey Shen

112078513 Pearl Lin

111078510 Tim Lin

111078517 Frank Tsai

Summary:

The forecasting initiative, designed for perishable goods distribution across various stores, seeks to enhance inventory management efficiency and minimize excess inventory by predicting weekly sales quantities. Our stakeholders are the inventory managers of suppliers, allocate inventory to stores, and surplus items result in handling costs. Given the perishable nature of the goods, failure to predict sales quantities can lead to excess inventory and additional costs.

We acquired data from Nuqleous, a company specializing in precision retail planning, following their recommendation to aggregate daily data into weekly data for a four-week forecasting window for perishable goods. Focusing on the SKU with the highest sales quantity, 1394919, we utilize a roll-forward forecasting approach which can promote accuracy by updating the recent data. External variables like average price are also incorporated to improve accuracy. Using Building Number 514 as an example, we showed the time plots and boxplots which identify the ETS method as the best model, considering outlier reduction, overfitting prevention, and minimizing residuals.

Continuous updates and real-time data refinement are crucial for model accuracy. The forecasting models, designed for a four-week horizon, offer a proactive strategy for suppliers to optimize inventory allocation, mitigate excess inventory costs, and reduce perishable item wastage.

There are still limitations involving the need for regular data updates to enhance predictive precision. Despite challenges, the forecasting models present a promising step towards efficient inventory management and sustainability in the supply chain.

Forecasting Perishable Food Sales Quantity for Efficient Inventory Distribution to Large Retail Stores

Business Problem and Goal:
Our stakeholders in this forecasting initiative are the **inventory managers of suppliers**. Their responsibility involves allocating inventory to various stores, and any surplus items incur handling costs borne by the inventory managers of suppliers. Additionally, since the products are perishable goods with a short shelf life, failure to predict sales quantities in advance can lead to excess inventory, resulting in additional costs for handling and disposing of expired items.

Our primary objective is to **employ predictive analytics for weekly sales forecasts, enhancing the efficiency of inventory allocation**. We strive to strike a balance between meeting demand and preventing unnecessary spoilage, aligning with our commitment to sustainability and operational efficiency. This proactive approach aims to **minimize excess inventory, reduce costs related to product spoilage, and optimize the overall performance of the supply chain**.

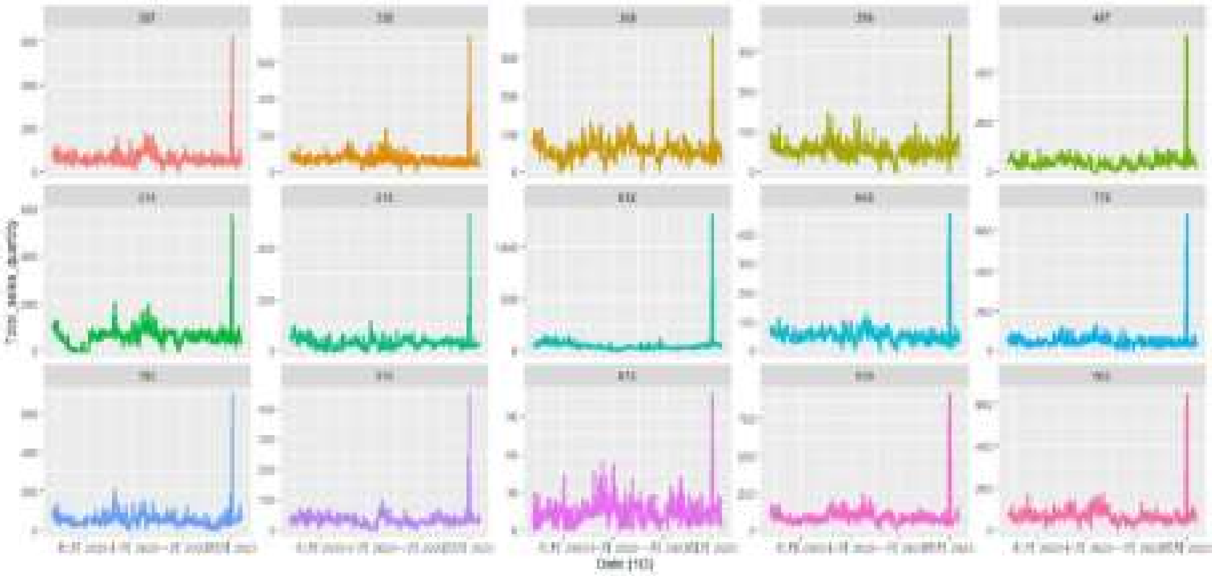
Forecasting Goal:
Our primary objective is to forecast the sales quantities of our **top-selling SKU product, 1394919** for each store. We define "top-selling" as identifying the product with **the highest sales quantity from 2021 to 2023**, across 15 stores. The decision to limit our focus to top-selling products, particularly the SKU 1394919, is intentional and serves as a minimum viable product (MVP) strategy. This approach allows us to establish the viability and effectiveness of our models in a controlled environment before considering their application to medium-selling items or expanding the scope of our analysis. According to the online meetings with Nuqleous meeting, we adopted their recommendation of setting warehouse stock levels for four weeks. Consequently, **our forecasting horizon will also be four weeks**.

We aim to **employ a roll-forward forecasting approach**, where part of the data is treated as newly acquired. The methodology for processing this data will be detailed in the data chapter. This approach allows us to carry out ongoing predictions and continuously refine our forecasts based on the latest information.

Data:
The raw data was provided by Nuqleous, who assists in precision retail planning through innovative solutions. The time period of the data spans from 2022/05/07 to 2023/04/20 with a total of 128,731 data rows. Our dataset includes a range of data columns, namely 'SKU Nbr', 'Building_Nbr', 'Purchase Channel', 'Date', 'Day of Week', 'Week and Year', 'This Year Sales ($)', 'Last Year Sales ($)', 'This Year Sales Qty', and 'Last Year Sales Qty'. Below are ten examples of data.

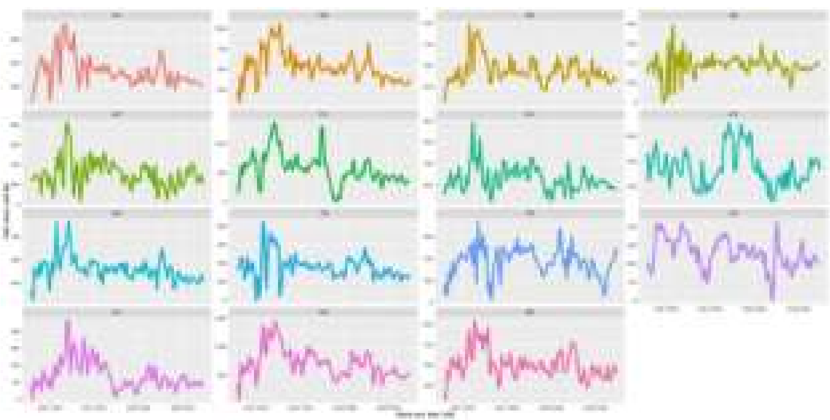
	SKU Nbr	Building Nbr	Purchase Channel	Date	Day of Week	Week and Year	This Year Sales ($)	Last Year Sales ($)	This Year Sales Qty	Last Year Sales Qty
1	1394919	207	Store Pickup	[illegible]	Saturday	202215	[illegible]	[illegible]	[illegible]	[illegible]
2	1394919	207	Store Delivery	[illegible]	Saturday	202215	[illegible]	[illegible]	[illegible]	[illegible]
3	1394919	207	In-Store	[illegible]	Saturday	202215	[illegible]	[illegible]	[illegible]	[illegible]
4	1394919	207	Store Pickup	[illegible]	Sunday	202215	[illegible]	[illegible]	[illegible]	[illegible]
5	1394919	207	Store Delivery	[illegible]	Sunday	202215	[illegible]	[illegible]	[illegible]	[illegible]
6	1394919	207	In-Store	[illegible]	Sunday	202215	[illegible]	[illegible]	[illegible]	[illegible]
7	1394919	207	Store Pickup	[illegible]	Monday	202215	[illegible]	[illegible]	[illegible]	[illegible]
8	1394919	207	Store Delivery	[illegible]	Monday	202215	[illegible]	[illegible]	[illegible]	[illegible]
9	1394919	207	In-Store	[illegible]	Monday	202215	[illegible]	[illegible]	[illegible]	[illegible]
10	1394919	207	Store Pickup	[illegible]	Tuesday	202215	[illegible]	[illegible]	[illegible]	[illegible]

Our business and forecasting goals require us to focus on columns which are Building Nbr, Week and Year, This Year Sales Qty, and This Year Sales ($). Additionally, ***SKU Nbr 1394919 has the highest sales*** **quantity** during the period from 2022 week 15 to 2023 week 12, we selected this SKU Nbr to conduct our forecasting. Below is the sample time chart for each Building Nbr series of SKU Nbr 1394919.



In order to meet our forecasting goals, we did the following preprocessing. First, we utilized the "Last year sales" column to **extend the data's time span,** resulting in a timeframe from 2021 week 15 to 2023 week 12. This doubles the data for forecasting purposes. Let us have more data for training. Second, To align with the supplier's management needs, we **aggregated the daily data into weekly data,** as Nuqleous suggested that a 4-week forecasting window would be sufficient. Finally, we **imputed missing values and removed extreme values**, please refer to Appendix 1 and 2 to see details. The following is an example of a time chart of each Building Nbr series after data processing.

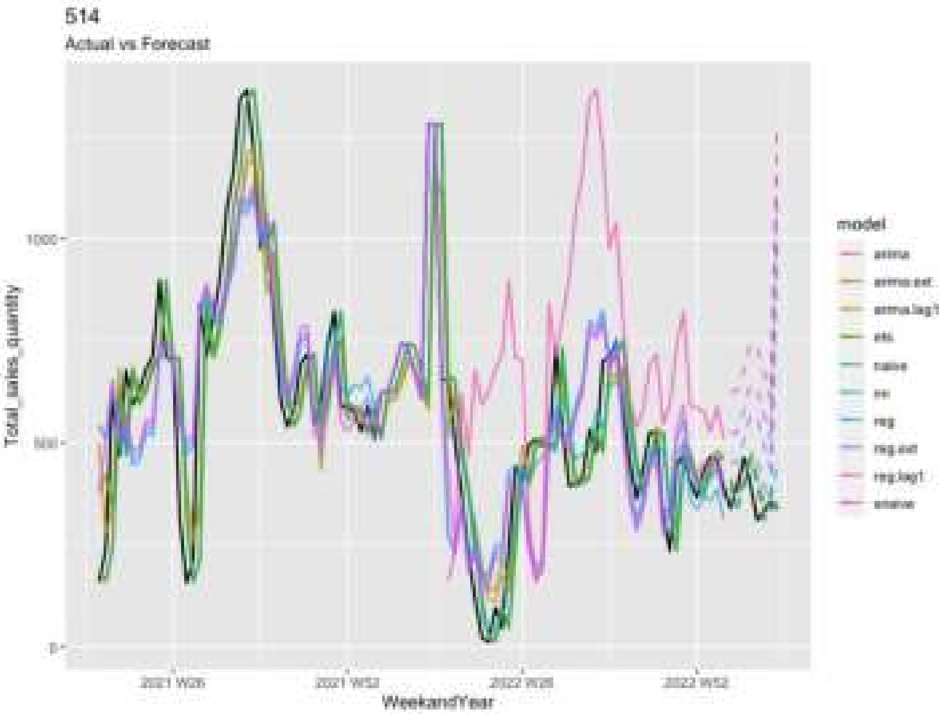
	Week and Year	Building_Nbr	Total_sales_quantity
1	2021 W15	207	100.0
2	2021 W16	207	148.0
3	2021 W17	207	339.0
4	2021 W18	207	332.0
5	2021 W19	207	553.0
6	2021 W20	207	572.0
7	2021 W21	207	675.0
8	2021 W22	207	650.0
9	2021 W23	207	476.0
10	2021 W24	207	507.0



Methods:

We will use the data from the **last eight weeks (2023 W5 to 2023 W12) as the validation period**, with the **training period spanning from week 15 of 2021 to week 4 of 2023**. We employed various forecasting methods, including NAIVE, Regression (TSLM), ARIMA, ETS, Neural Network, etc. Considering that price may influence sales quantity, we also include average price as an external variable in the model. Subsequently, we adopted a **roll-forward approach**, gradually incorporating data from 2023 W5 to 2023 W12 into the training period. To evaluate the performance, we also calculate the RMSE and residuals of each model.

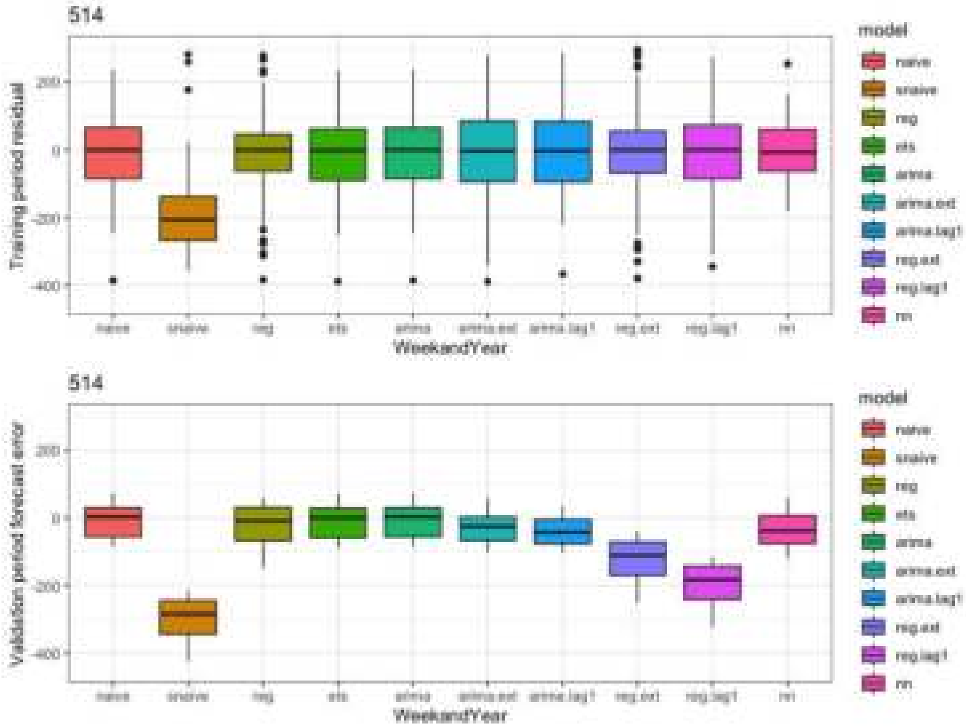
We used time plots to preliminarily observe the prediction results of each method. We take Building Nbr 514 as a representative for an explanation. (The other 14 buildings will be put in appendix 3.)



Observing this chart, we found that the forecast results of the SNAIVE method deviate significantly from the actual values. Therefore, for Building Nbr 514, our first step would be to eliminate this forecasting method.

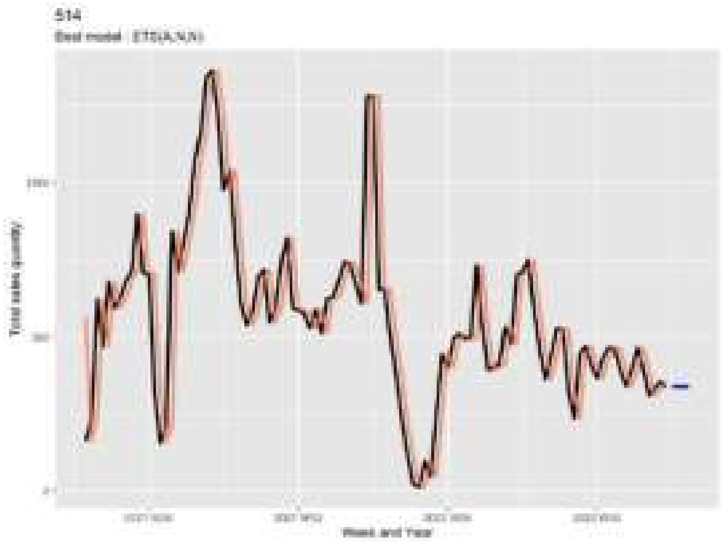
Next, we use boxplots to examine the residuals of each forecasting method. We follow the following steps to filter and choose the method. 1. Initially, we evaluate the performance on the validation set and **eliminate methods with a median significantly deviating from 0, excessive outliers, and large residuals**. 2. We compare the performance with the training period and **discard methods that exhibit overfitting** (perform well during the training period but poorly during the validation period). 3. We then **observe the total residuals** and choose the method with a smaller sum. Prioritizing the goal of minimizing inventory waste, we consider under-forecasting to be preferable over over-forecasting. 4. Additionally, since we aim to minimize error values, the RMSE metric imposes a substantial penalty for errors. Therefore, we use the RMSE metric as a supplementary criterion and **select the method**

with a smaller RMSE. 5. After four rounds of filtering if multiple methods remain, we opt for the **simpler one** (for example, choosing naive over ARIMA).



Taking the boxplot of Building Nbr 514 as an example: Initially, we retain the NAIVE, ETS, and ARIMA methods because they exhibit fewer outliers, and their medians do not significantly deviate from 0. After comparing these three methods with the training period, we observe no overfitting issues. Further comparison of the residuals for these three methods reveals that ETS has the smallest total sum (0.423). Therefore, we select ETS as the best model.

Finally, we forecast sales quantities for weeks 2023 W13 to 2023 W16 for each building using their respective best-fitted models.



Following this methodology, we chose the best model for each building and also used their respective best-fitted models to forecast sales quantities for weeks 2023 W13 to 2023 W16. The detailed results and methodology are included in the appendix for reference. (The regression equation is composed of trend and seasonality. The external variable is the average price of goods, and the lag1 is the lag of 1 week)

Building_Nbr	[illegible]	[illegible]	[illegible]	[illegible]
Model	NAIVE	NNAR(1,1,2)[52]	NNAR(2,1,2)[52]	reg_lag1 【TSLM(trend() + season() + avgprice_lag1)】
Building_Nbr	[illegible]	[illegible]	[illegible]	[illegible]
Model	reg 【TSLM(trend() + season())】	ETS(A,A,N)	NAIVE	NAIVE
Building_Nbr	[illegible]	[illegible]	[illegible]	[illegible]
Model	ARIMA(0,1,2)	reg.ext 【TSLM(trend() + season() + Average_price)】	NAIVE	ETS(A,N,N)
Building_Nbr	[illegible]	[illegible]	[illegible]	
Model	reg.ext 【TSLM(trend() + season() + Average_price)】	ARIMA(2,1,2)	reg 【TSLM(trend() + season())】	

Conclusions:

Advantages: In the future, suppliers can leverage the sales forecasting models identified for each building to predict sales volumes for the upcoming four weeks. **This forward-thinking approach aims to achieve better distribution of perishable goods, minimize food waste, and reduce associated costs related to inventory waste and spoilage.**

The sales forecasting model, implemented using a roll-forward method, enables suppliers to continuously project sales volumes through regular monthly updates. Additionally, our system has the capability to generate forecasts for the sales quantity of top-selling products over the next four weeks for each store. These forecasts are then shared through detailed reports sent to each store. By strategically distributing stock across various stores based on these forecasts, inventory managers can optimize their inventory allocation strategies.

Limitations: The accuracy of our forecasting models relies on timely data updates. However, following discussions with Nuqleous, it became evident that there is currently no established plan for regular data updates. Consequently, the absence of updated data poses a challenge in validating the accuracy of our predictions. While we have implemented a rigorous selection process for identifying the optimal forecasting models, the chosen models are still subject to ongoing refinement as additional real-time data becomes available. Continuous efforts are required to enhance the precision of our predictions.

As part of our strategic approach to adopting a Minimum Viable Product (MVP), we first chose to forecast inventory quantities for a single product. **Future research could delve into the exploration of whether applying a uniform decision logic to select models for all products yields more nuanced insights.** This consideration would allow for a more detailed examination of the applicability of our forecasting methodology across a broader product range.

Appendix

1. <u>**Missing Value**</u>

Addressing missing values is another step in our preprocessing. Given that the original data spans from 2022 week 15 to 2023 week 12, and we extend it to cover 2021 week 15 to 2023 week 12, there are gaps for 2022 week 13 and 2022 week 14. Additionally, values for the last two days (4/21, 4/22) of 2022 week 12 are missing. To rectify the missing values in week 12, we insert those by looking at the previous week (2022 week 11) to *calculate the percentage of weekly sales quantity* represented by the sales quantity on each of the last two days of the week. Then, the values for week 13 are replaced with the values of week 12, and week 14 are replaced with the values of week 15. (16 days of missing value are all being addressed)

2. <u>**Extreme value**</u>

After aggregating daily data into weekly data, there are extreme values in both 2022 week 10 and 2023 week 10. To address this, we replace the value of 2022 week 10 and 2023 week 10 by taking the *average of the values from their previous week and the following week.*

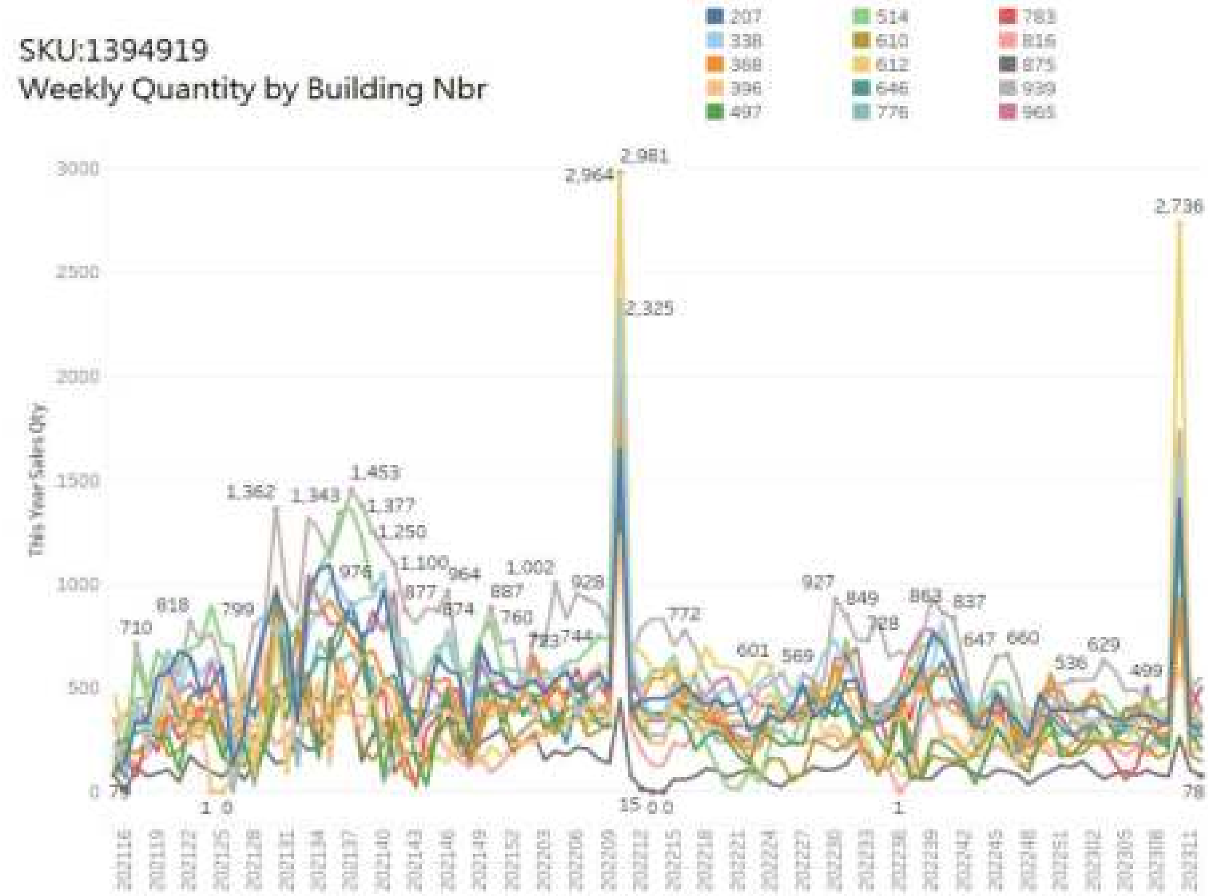
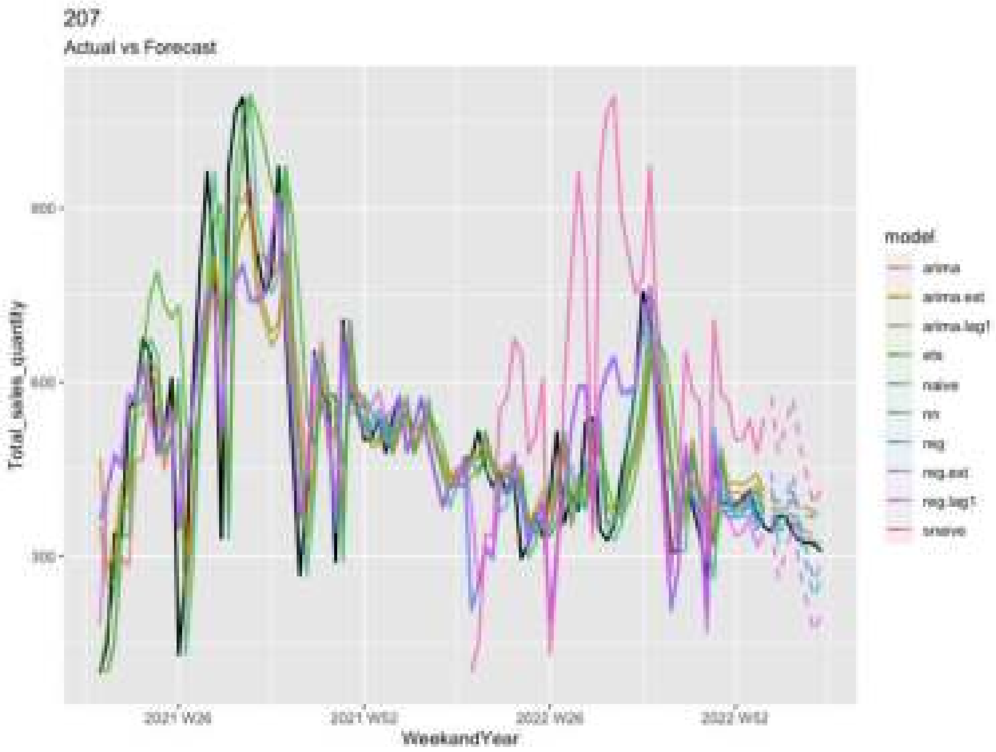


Fig1. Each building's weekly sales quantity

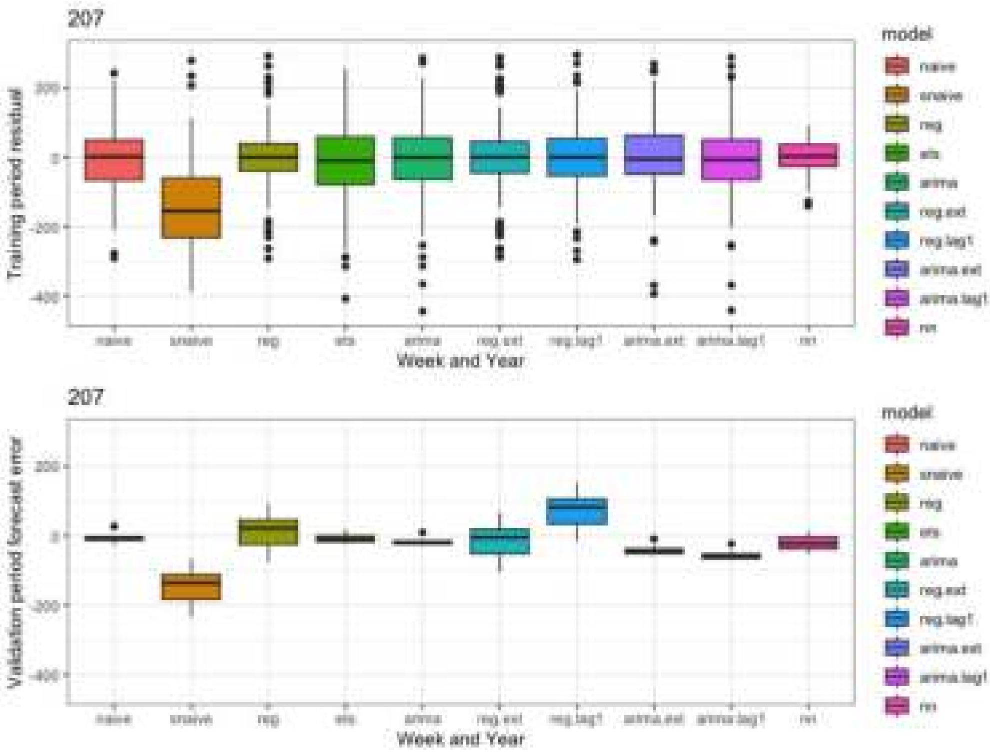
3. Performance evaluation:

1) Building_Nbr: 207

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method



- Performance evaluation

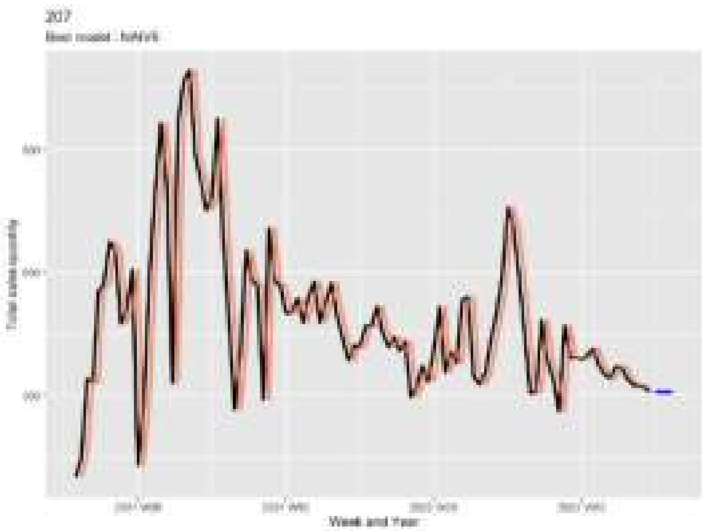
Type	Model	MAE	RMSE	MAPE
	arima	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	142	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
Training	naive	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	arima	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
Validation	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]

Model	total residual
arima	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
naive	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]

- Best model selection explanation

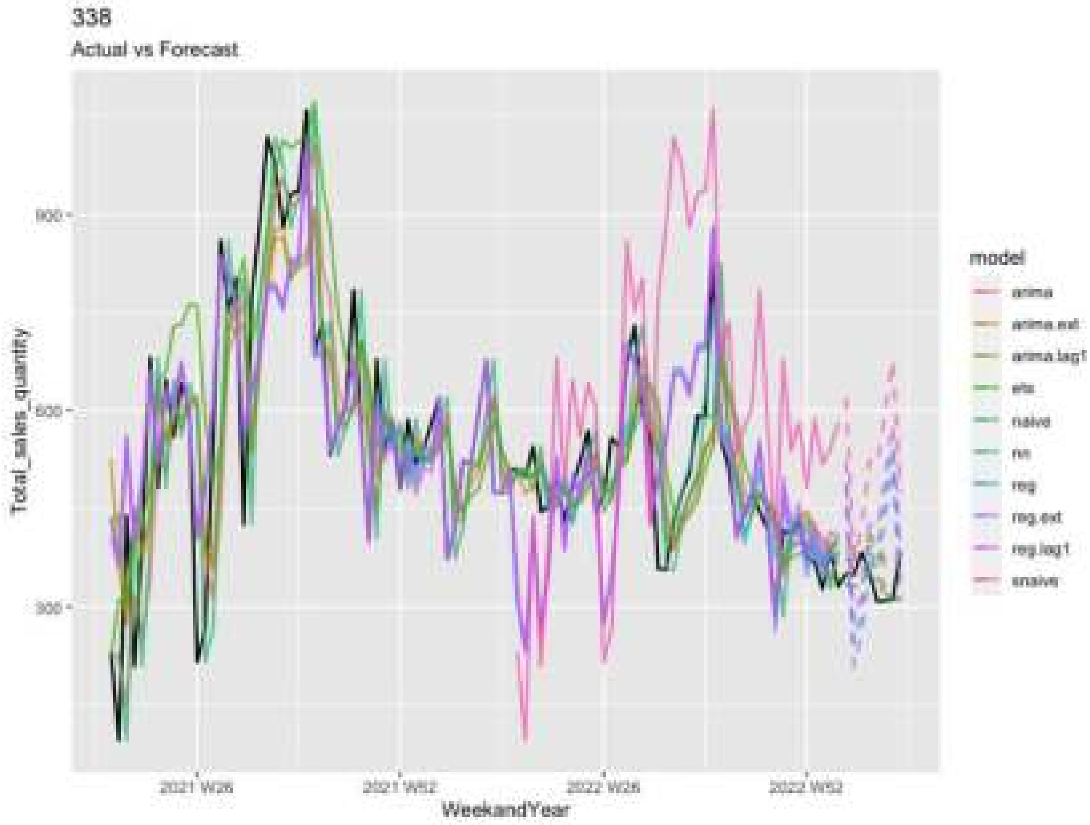
  From the training boxplot, we found that 'naive,' 'ets,' and 'neural network' perform the best with relatively fewer outliers. However, analyzing the validation boxplot, it is observed that 'neural network' performs better during training compared to validation, indicating potential overfitting (deviation from 0 during validation). Further examination of RMSE and residuals reveals that 'ets' has a residual of -3.21e+3, while 'naive' has 2.51e+2. The training RMSE for 'ets' is 164, and for 'naive' is 155. In the validation set, 'ets' has an RMSE of 15.6, while 'naive' has 15.5. Considering the residuals being closer to 0 and positive (indicating no over forecast) for 'naive,' along with its slightly better RMSE performance in both training and validation, 'naive' is chosen as it aligns better with the business goal.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

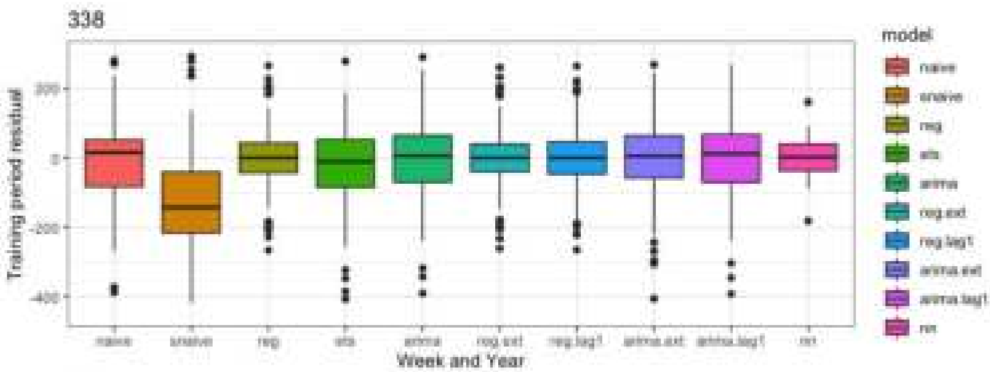


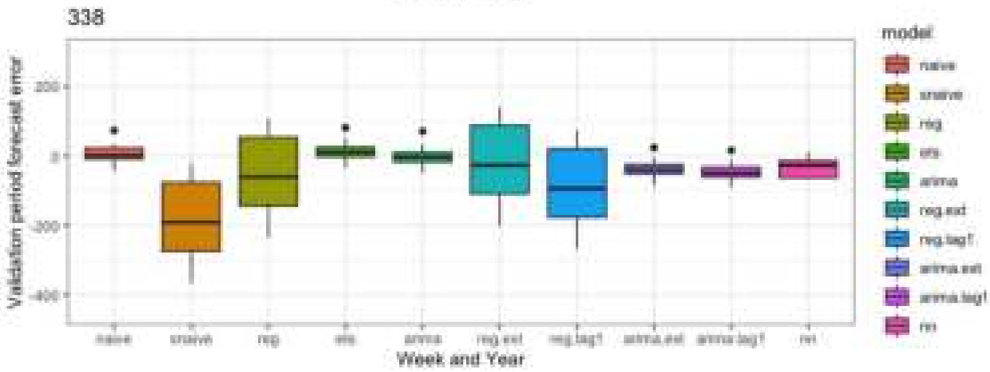
2) Building_Nbr: 338

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method





- Performance evaluation

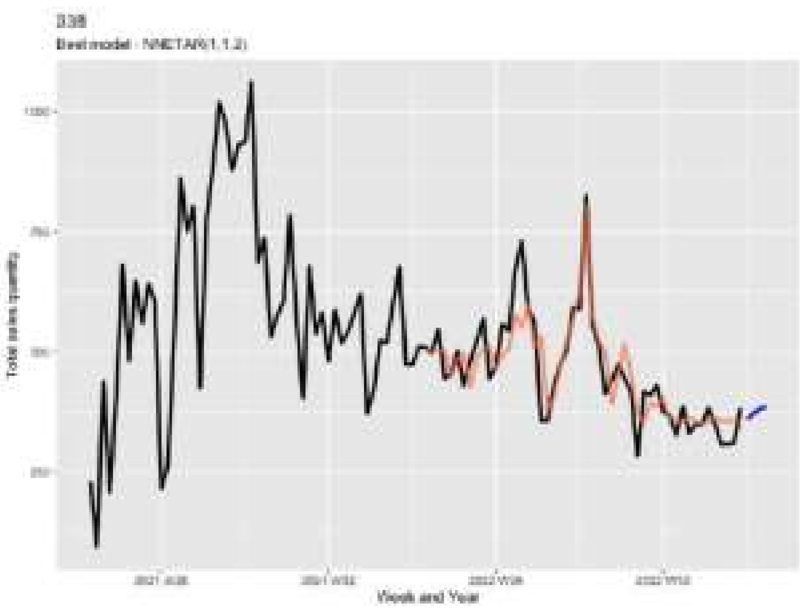
Type	Model	MAE	RMSE	MAPE
Training	arima	100	135	21.3
	arima.ext	95.3	129	22.6
	arima lag1	97.1	130	23.3
	ets	105	145	24.5
	naive	106	144	22.8
	nn	43.7	57.3	8.52
	reg	70.5	103	17.6
	reg.ext	70.2	102	17.4
	reg lag1	71.9	107	17.9
	snaive	214	253	47.1
Validation	arima	25.1	34.2	7.23
	arima.ext	41.2	48.9	12.5
	arima lag1	47.5	53.5	14.5
	ets	27.1	37.1	7.55
	naive	25.6	34.9	7.21
	nn	38.5	43.2	11.3
	reg	112	127	33.7
	reg.ext	103	117	30.8
	reg lag1	120	144	38.7
	snaive	102	218	55.7

Model	total residual
arima	2.10e+2
arima.ext	1.49e+2
arima lag1	3.22e+2
ets	2.05e+3
naive	1e+2
nn	2.14e+0
reg	6.82E-13
reg.ext	5.80E-14
reg lag1	1.08E-12
snaive	5.40e+3

- Best selection explanation

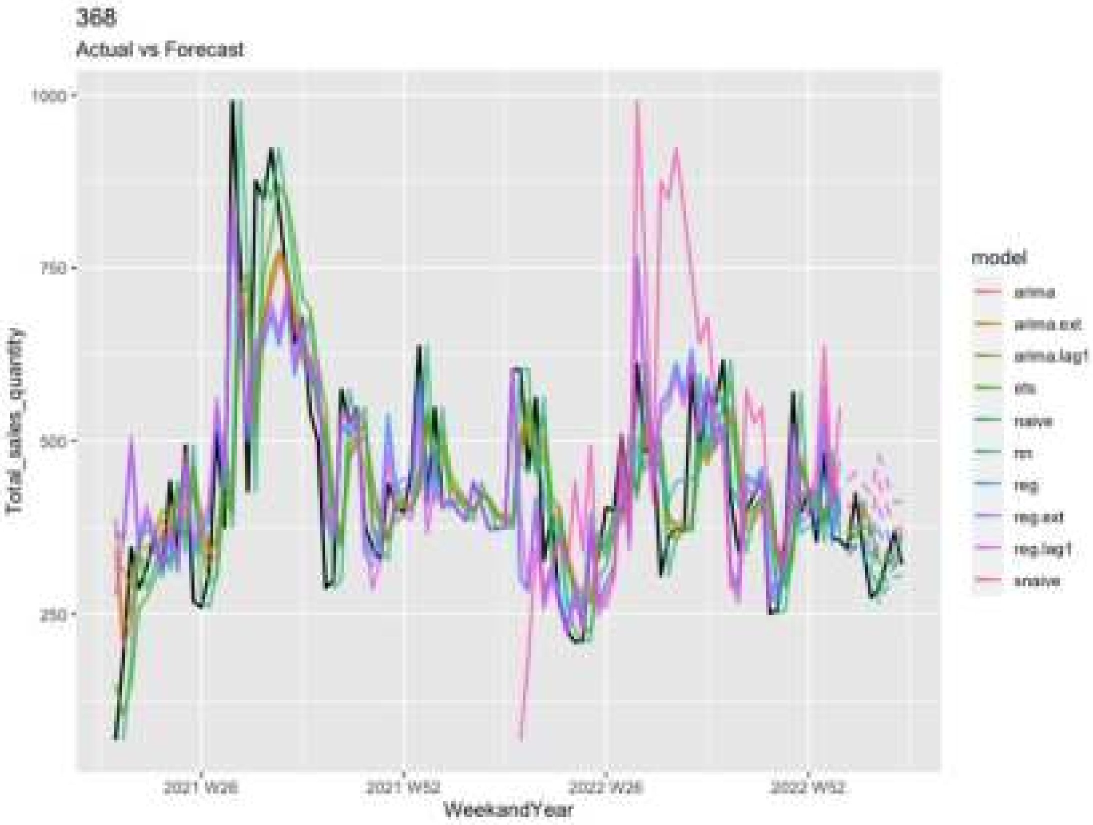
Based on the boxplot, 'naive' and 'neural network (nn)' are retained as they exhibit fewer outliers. Analyzing residuals, 'neural network' has a value of -2.14e+0, while 'naive' has 1e+2. In terms of training RMSE, 'neural network' is 57.3, and 'naive' is 144. For the validation set, 'neural network' has an RMSE of 43.2, and 'naive' has 34.9. Choosing 'neural network' is preferred due to its residual being closer to 0, indicating better performance in capturing the underlying patterns in the data.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

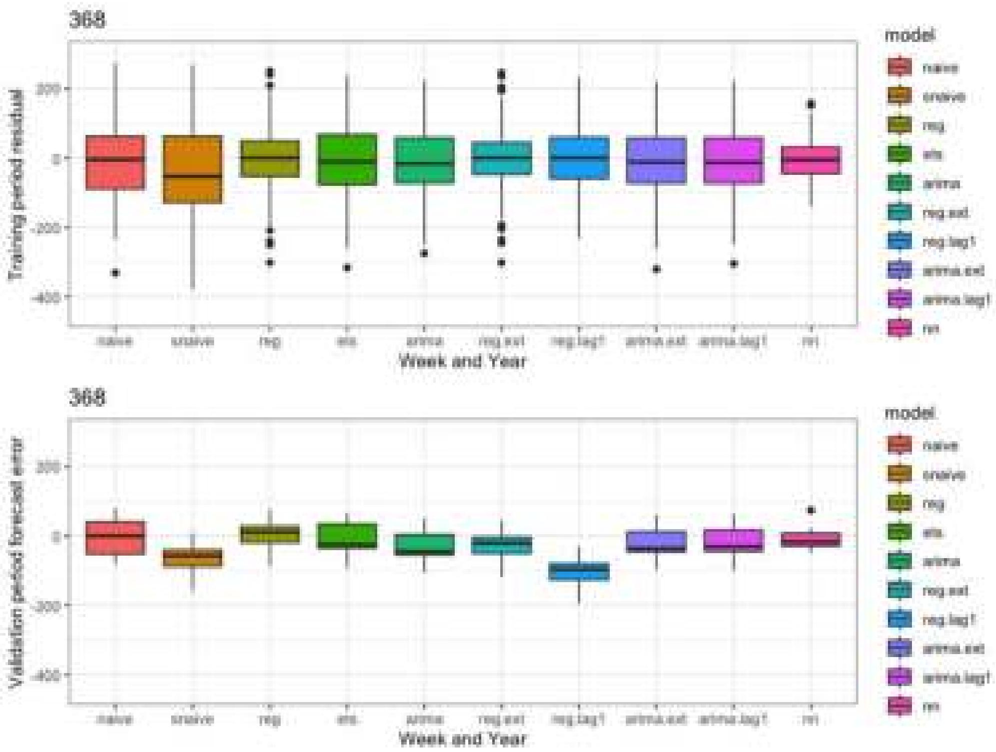


3) Building_Nbr: 368

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method



- Performance evaluation

Type	Model	[illegible]	RMSE	MAPE
Training	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	naive	[illegible]	143	[illegible]
	nn	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
Validation	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	nn	[illegible]	37	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]
	[illegible]	[illegible]	[illegible]	[illegible]

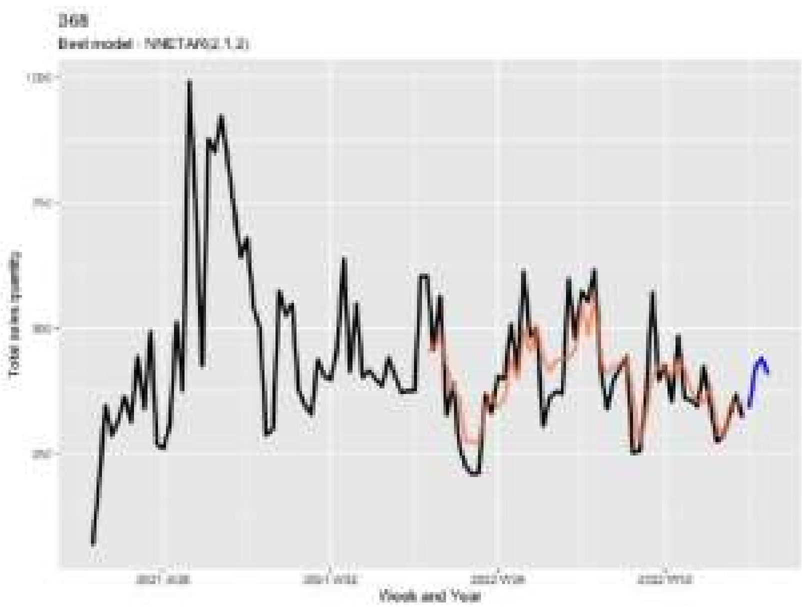
Model	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]

- Best selection explanation

Based on the boxplot analysis, 'naive,' 'neural network (nn),' and 'reg.lag1' are retained as they exhibit fewer outliers and do not deviate significantly from 0. Analyzing residuals, 'naive' has a value of 2.89e+2, 'neural network' has 3.74e+0, and 'reg.lag1' has 2.84E-13. In terms of training RMSE, 'naive' is 143, 'neural network' is 66.4, and 'reg.lag1' is 96.1. For the validation set, 'naive' has an RMSE of 54.2, 'neural network' has 37, and 'reg.lag1' has 114.

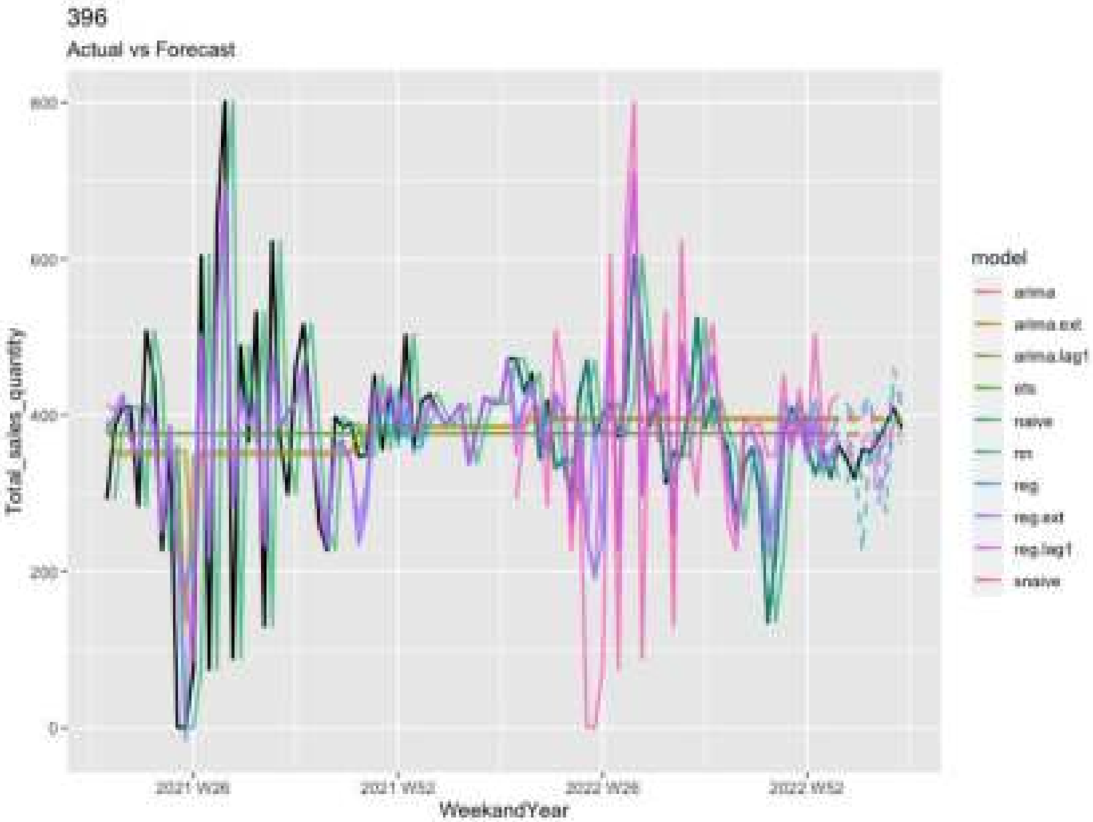
Choosing 'neural network' is preferred due to its residual being closer to 0, and it exhibits less over forecasting compared to the other models.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

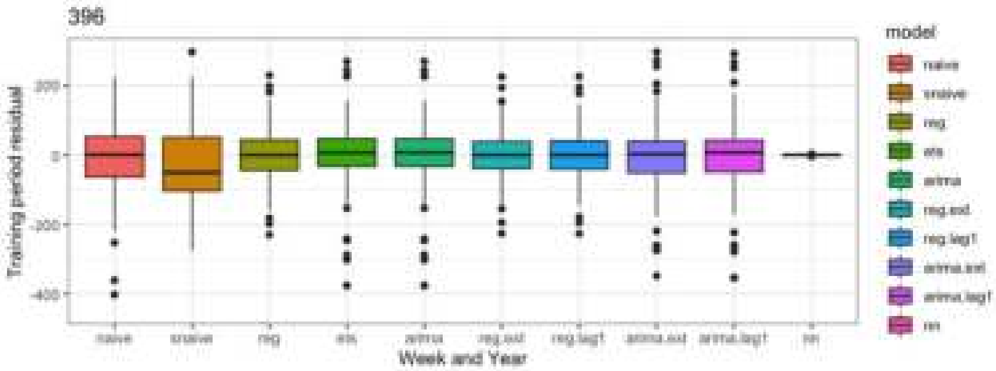


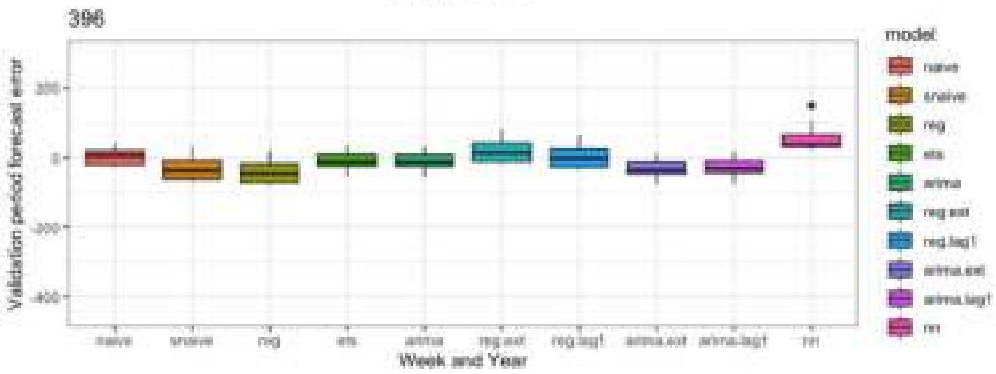
4) Building_Nbr: 396

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method





- Performance evaluation

Type	Model	MAE	RMSE	MAPE
Training	arima	82.5	124	Inf
	arima.ext	82.7	120	Inf
	arima.lag1	82.6	121	Inf
	ets	82.5	124	Inf
	naive	109	171	Inf
	nn	2.33	3.36	0.642
	reg	62.7	84.5	Inf
	reg.ext	56.3	78.1	Inf
	reg.lag1	50.4	80.9	Inf
	snaive	139	179	38.4
Validation	arima	24.1	28.8	6.85
	arima.ext	36.5	42.1	10.2
	arima.lag1	3.28E+01	39.1	8.30
	ets	23.9	28.5	6.78E+00
	naive	2.08E+01	22.7	5.75
	nn	5.05E+01	68.6	1.37E+01
	reg	4.50E+01	52.4	1.29E+01
	reg.ext	3.08E+01	39.3	8.2
	reg.lag1	2.04E+01	33.8	7.08
	snaive	38.9	45.3	11.4

Model	total_residual
arima	6.82E-13
arima.ext	4.78E-11
arima.lag1	4.92E-11
ets	2.91e+0
naive	7e+1
nn	-3.28e-1
reg	-2.84E-14
reg.ext	1.53E-13
reg.lag1	2.90E-13
snaive	4.14e+2

- Best selection explanation

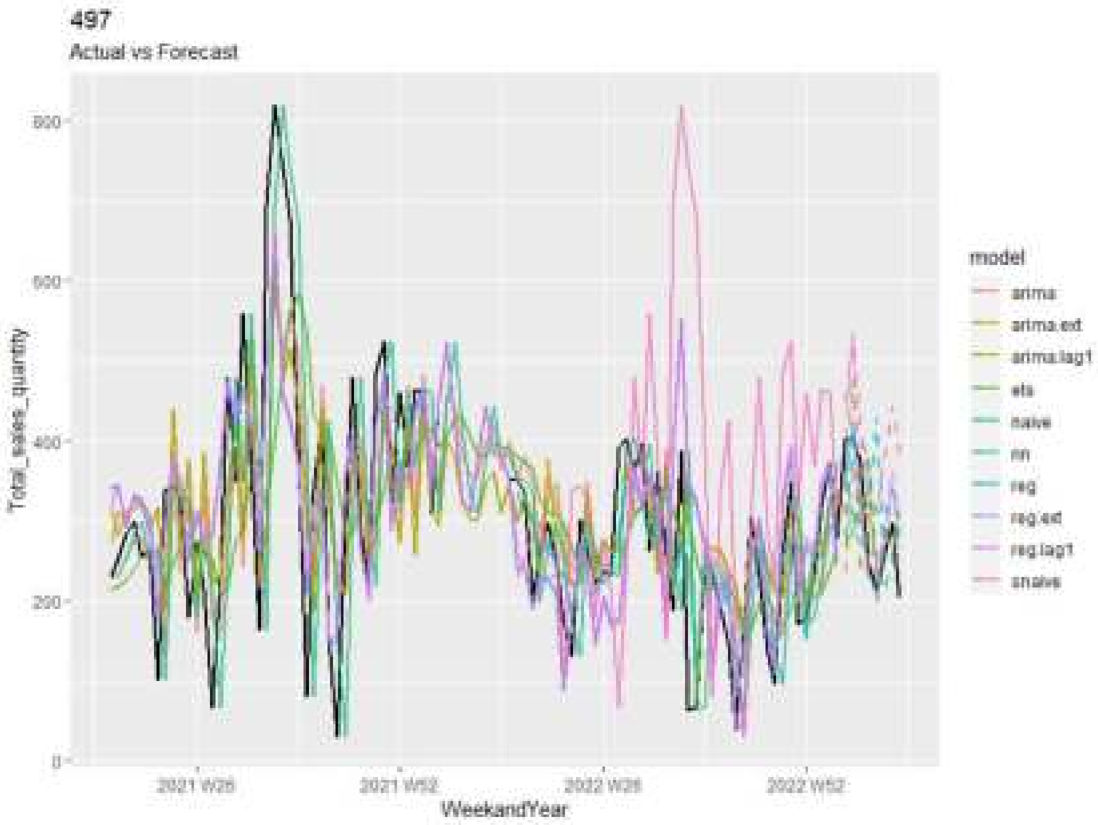
Based on the boxplot analysis for 'regression,' 'reg.ext,' and 'reg.lag1,' 'reg.lag1' is selected. The residuals for 'regression' are -2.84E-14, 'reg.ext' is 4.78E-11, and 'reg.lag1' is 4.92E-11. For the validation RMSE, 'regression' is 52.4, 'reg.ext' is 39.3, and 'reg.lag1' is 33.8. Choosing 'reg.lag1' is preferred as 'regression' exhibits deviations from 0 during validation, and there is a more severe over forecasting issue. Therefore, 'reg.lag1' is chosen for its residuals being closer to 0.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

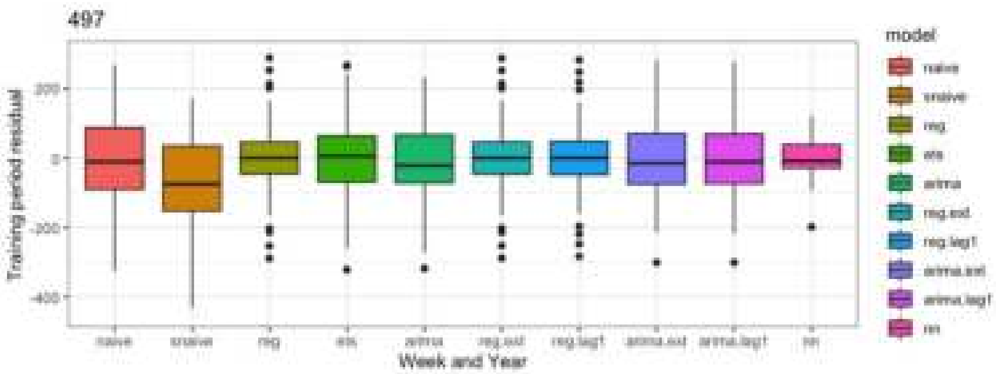
Although in the validation period, our forecasts were based on models that were estimated/trained only on data that was available at the "time of prediction". But when it comes to forecasting the future after we selected the best model, “Average_price” is not available at the time of prediction (we don’t have the future data of “Average_price”), therefore “Total sales quantity” of weeks 2023 W13 to 2023 W16 couldn’t be predicted.

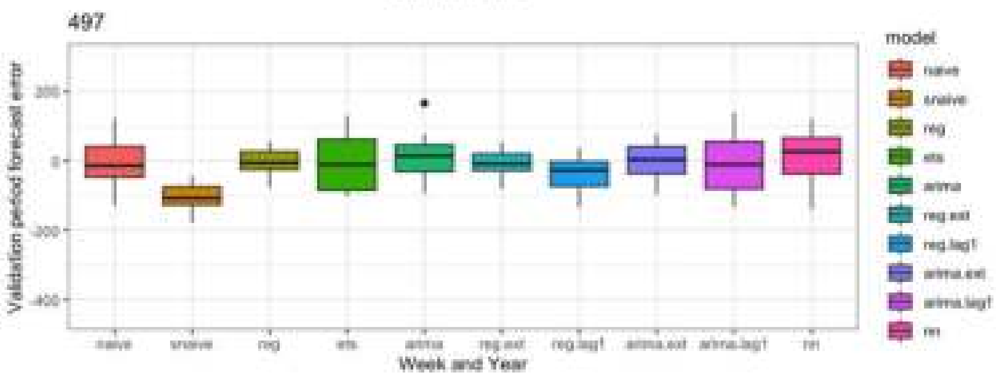
5) Building_Nbr: 497

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method





- Performance evaluation

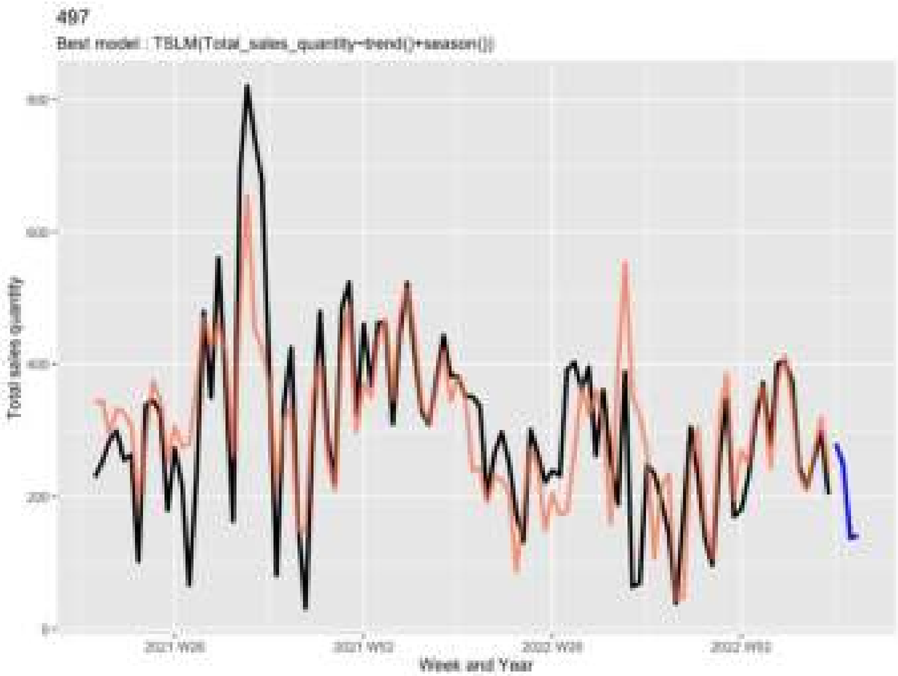
Type	Model	MAE	RMSE	MAPE
Training	arima	89.9	110	48.2
	arima.ext	90.4	110	48.7
	arima.lag1	90.2	118	49.7
	ets	90.4	135	58.3
	naive	108	142	53.4
	nn	42	55.7	28.1
	reg	66.8	94.7	37.8
	reg.ext	66.9	94.7	37.9
	reg.lag1	67	94.4	37.9
	snaive	164	225	112
Validation	arima	64.3	81.1	21.6
	arima.ext	49.2	56.6	17.8
	arima.lag1	82	91.8	28.6
	ets	76.2	84.6	27
	naive	61.9	75.9	23
	nn	70.5	96.9	27.6
	reg	35.9	43.9	13
	reg.ext	35.1	42.3	12.6
	reg.lag1	49.4	66.4	17.3
	snaive	101	113	39.6

Model	total_residual
arima	6.57e+1
arima.ext	1.30e+1
arima.lag1	2.02e-1
ets	2.40e+2
naive	4.9[illegible]e-1
nn	2.08e-1
reg	9.24E-14
reg.ext	2.42E-13
reg.lag1	-3.69E-13
snaive	4.27e+3

- Best selection explanation

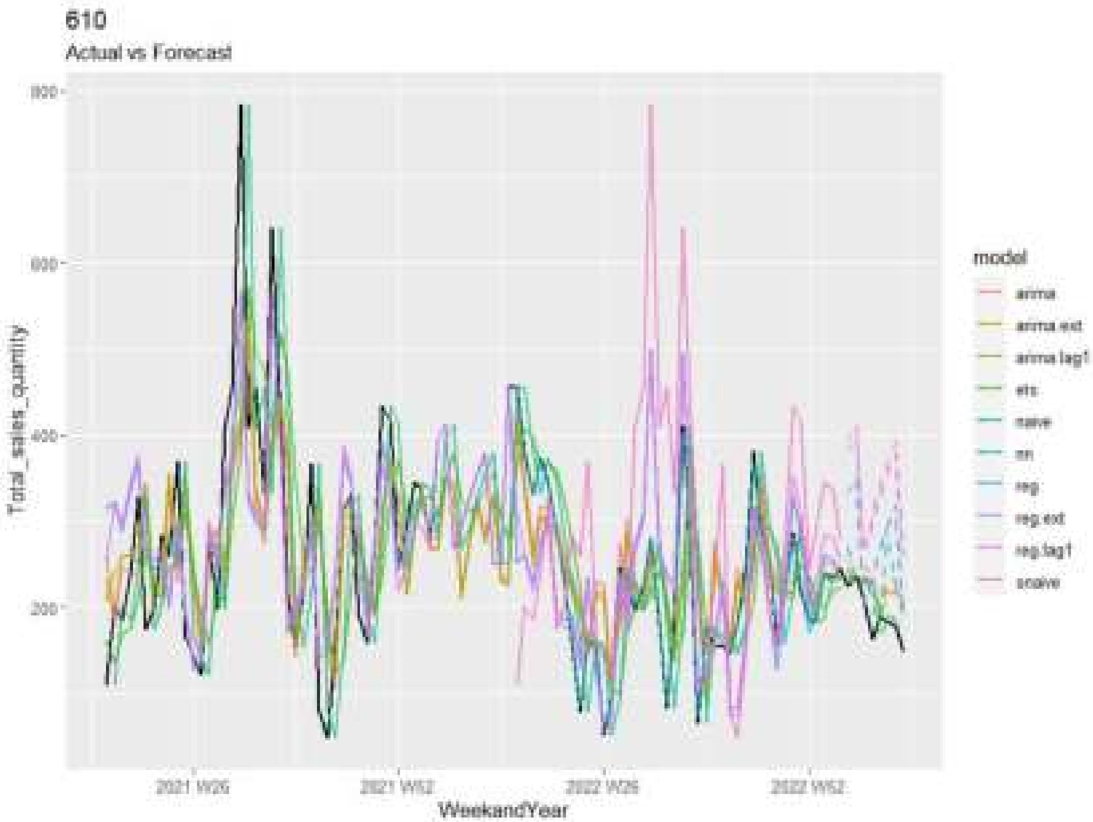
Based on the boxplot, 'reg,' 'reg.ext,' 'reg.lag1,' and 'nn,' 'reg' and 'reg.lag1' are selected. The residuals for 'reg' are 9.24E-14, 'reg.ext' is -2.42E-13, 'reg.lag1' is -3.69E-13, and 'nn' is 2.06e-1. Choosing 'reg' and 'reg.lag1' is preferred as 'nn' exhibits poor performance in validation, deviating from 0 or having a wide range of errors. Moreover, compared to 'reg.ext,' 'reg' is simpler.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

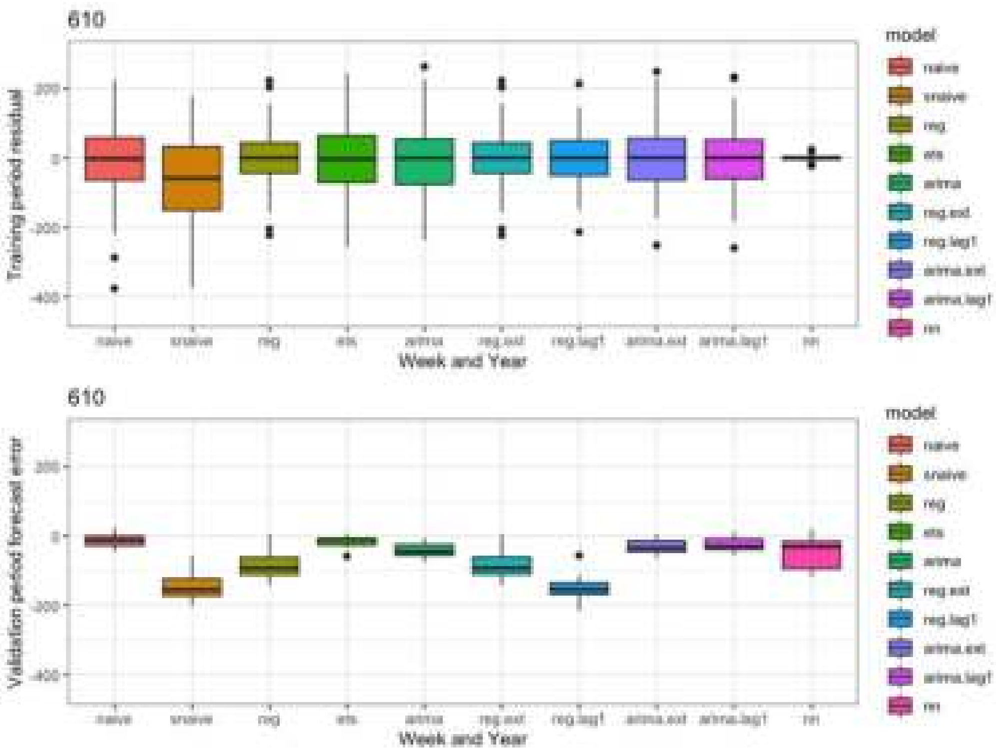


6) Building_Nbr: 610

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method



- Performance evaluation

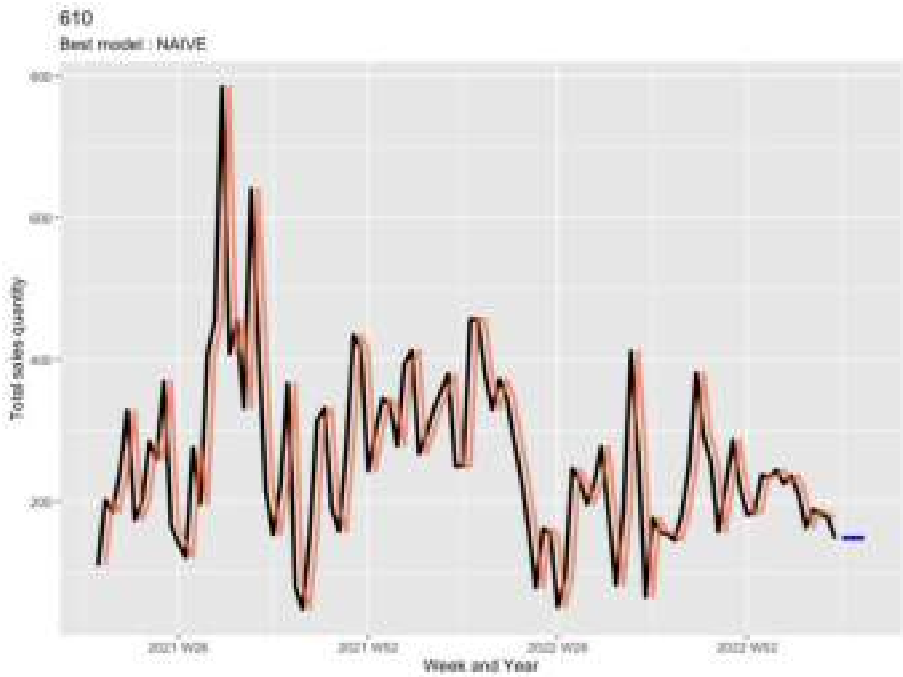
Type	Model	MAE	RMSE	MAPE
Training	arima	74.304091	90.016581	38.511525
	arima.ext	73.691973	97.323042	39.380601
	arima lag1	73.246295	96.183157	38.854993
	ets	83.304346	110.076677	42.353966
	naive	87.817204	116.956855	41.711197
	nn	6.144375	7.636152	3.819631
	reg	54.117071	73.737681	27.057006
	reg.ext	54.106405	73.237559	27.054204
	reg lag1	56.025078	71.510848	27.529772
	snaive	130.547619	166.707805	74.018951
Validation	arima	41	46.7	23.3
	arima.ext	31.4	37.6	19.1
	arima lag1	28.5	33.4	18.3
	ets	19.3	28.1	11.2
	naive	21	24.6	11.7
	nn	42.3	52.4	23.6
	reg	83.4	93.3	43.8
	reg.ext	82.9	92.7	43.6
	reg lag1	148	155	78.3
	snaive	144	150	78.3

Model	total residual
arima	8.37e+1
arima.ext	9.87e+1
arima lag1	6.37e+1
ets	1.50e+2
naive	1.38e+2
nn	1.51e+0
reg	6.79e-13
reg.ext	2.04E-13
reg lag1	2.70e-13
snaive	2.50e+3

- Best selection explanation

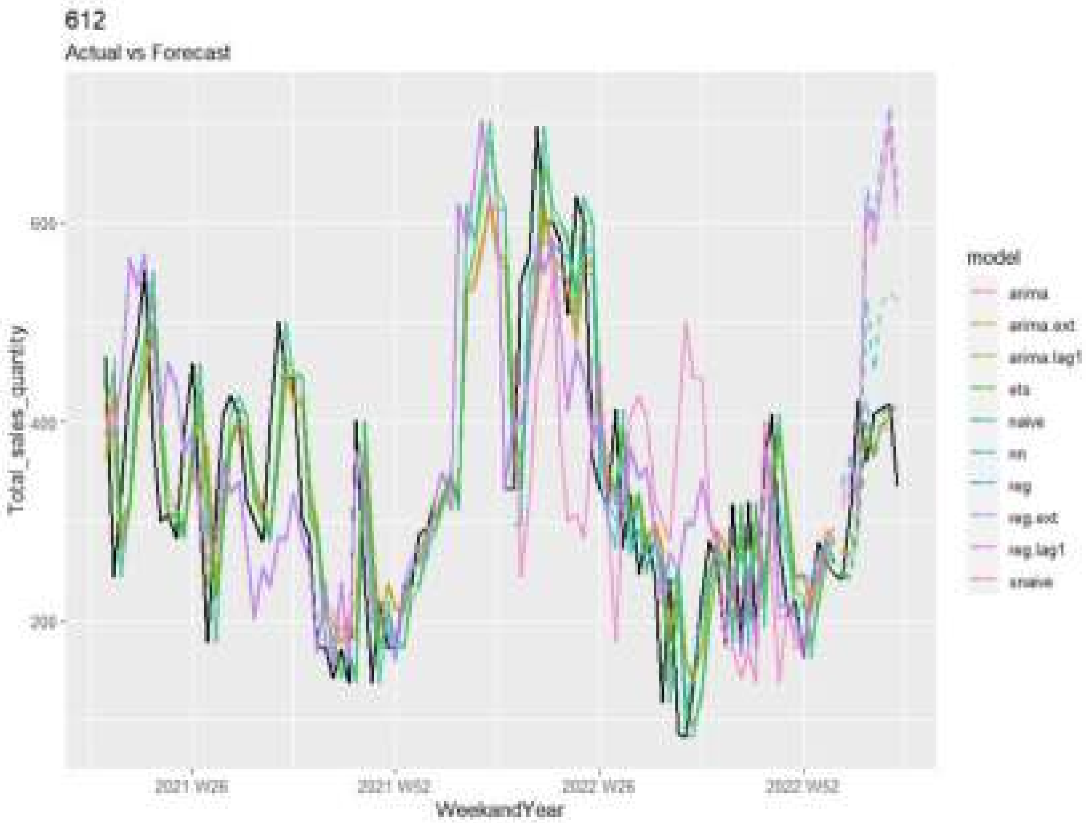
We choose naive as the best model. It seems that the 'ets' boxplot appears favorable initially; however, it has a significant issue with over forecasting and includes outliers.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

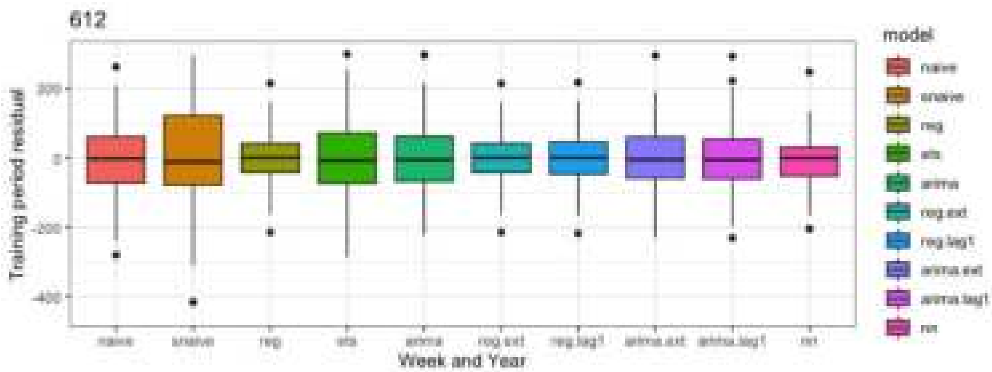


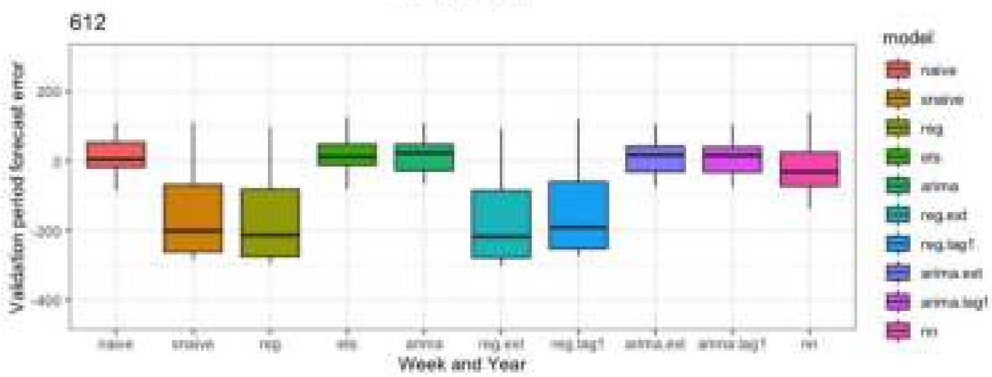
7) Building_Nbr: 612

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method





- Performance evaluation

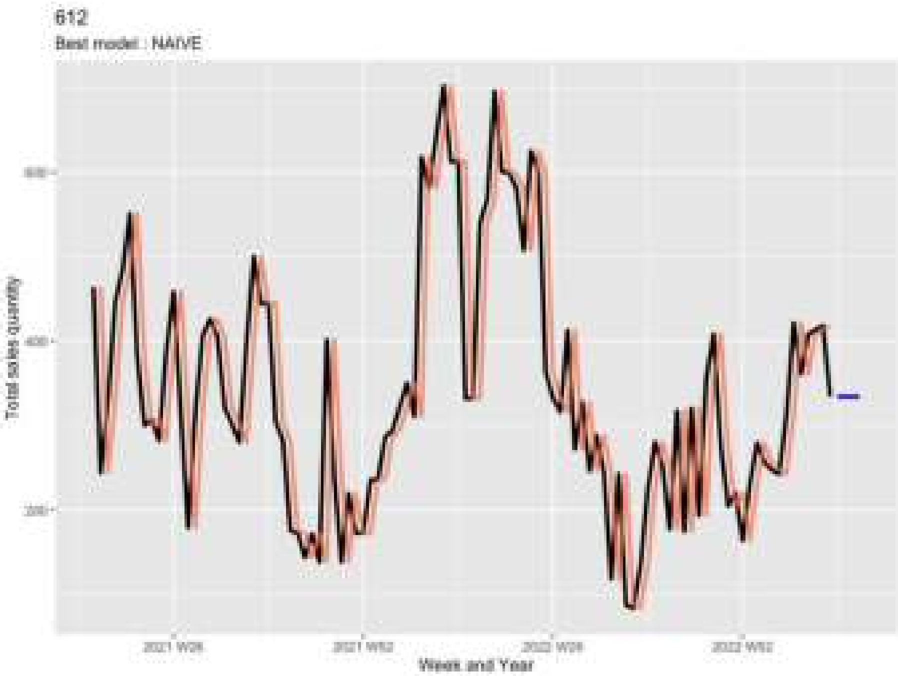
Type	Model	MAE	RMSE	MAPE
Training	arima	78.01856	95.97127	20.27119
	arima.ext	75.73368	95.51918	27.90720
	arima.lag1	74.24162	93.73112	27.25819
	ets	79.50863	99.79117	20.10913
	naive	78.88889	102.78723	28.11164
	nn	81.35834	92.62775	25.0880
	reg	58.88849	80.11861	22.71284
	reg.ext	50.82249	80.10979	22.6931
	reg.lag1	58.33794	78.73447	22.56288
	snaive	130.30952	189.99573	56.51
Validation	arima	46.0	54.2	12.9
	arima.ext	45.6	53.5	12.7
	arima.lag1	45.1	53.4	12.7
	ets	47	61.3	12.8
	naive	48.1	60.9	13.3
	nn	101	115	27.3
	reg	181	212	51.9
	reg.ext	184	215	52.7
	reg.lag1	174	195	46.8
	snaive	181	202	49

Model	total residual
arima	1.40e+2
arima.ext	1.30e+2
arima.lag1	-2.44e+2
ets	2.23e+2
naive	-2.17e+2
nn	8.60e-1
reg	4.83E-13
reg.ext	0.53E-14
reg.lag1	2.84E-14
snaive	5.41e+2

- Best selection explanation

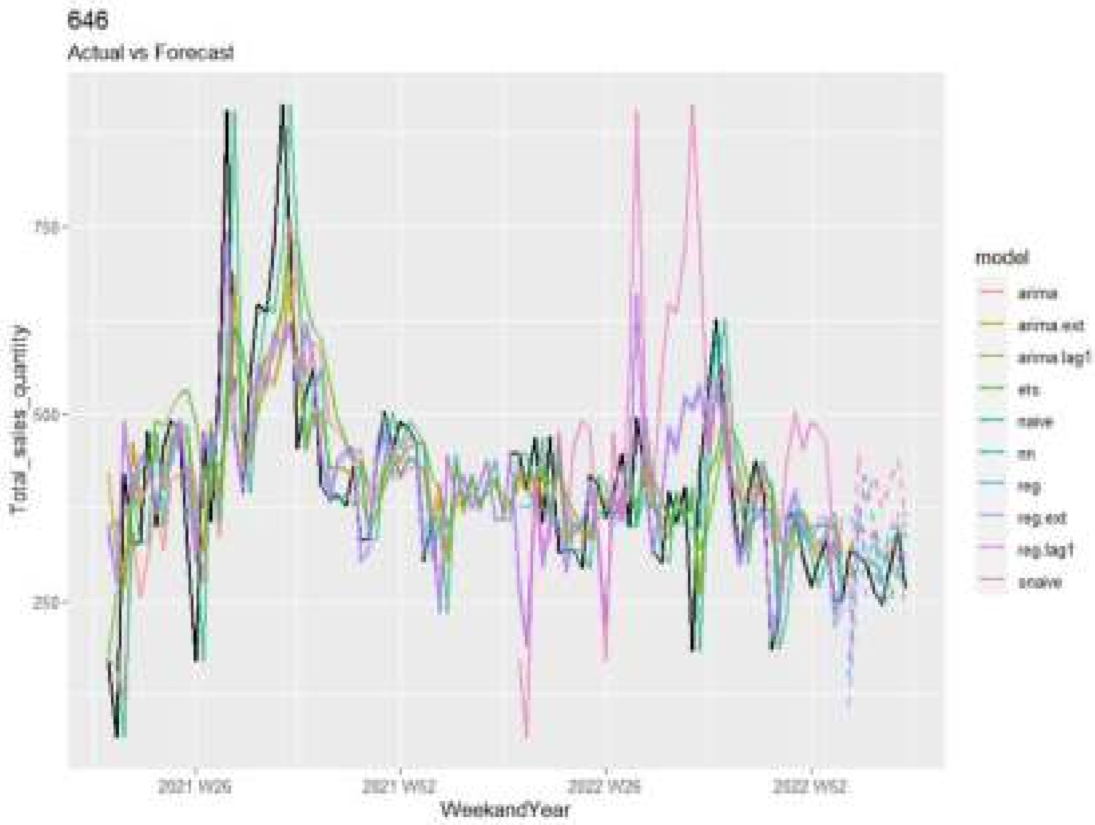
Based on the boxplot, we selected 'naive' and 'ets.' Upon further observation of residuals, 'naive' has a residual of -2.17e+2, while 'ets' has -2.23e+2. Choosing 'naive' is preferred in this case.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

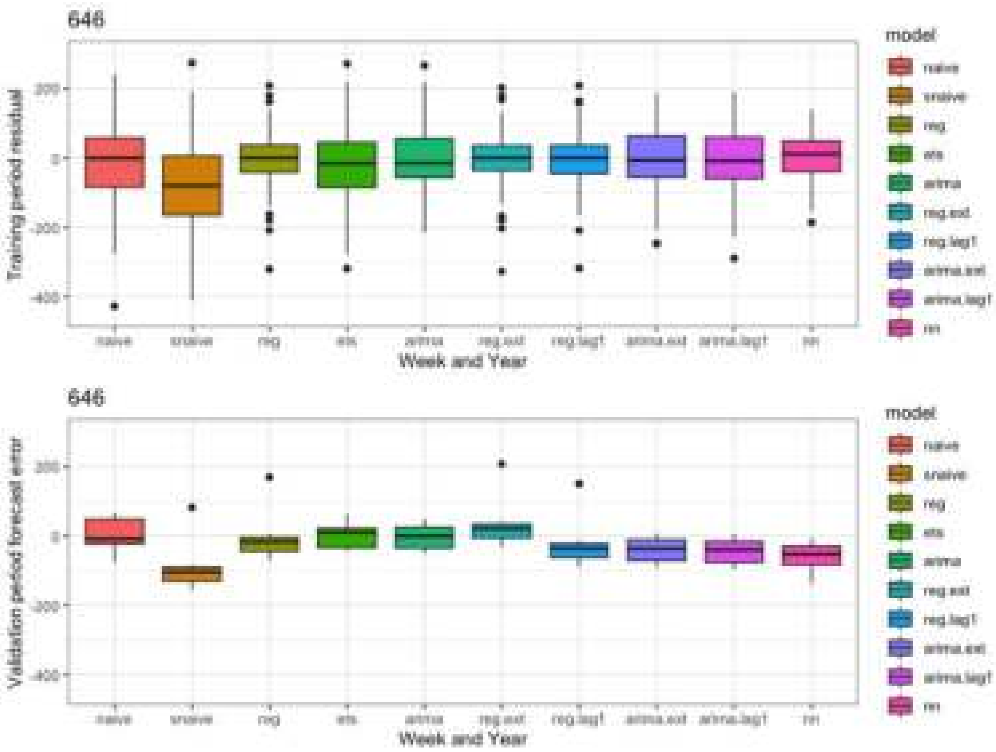


8) Building_Nbr: 646

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method



- Performance evaluation

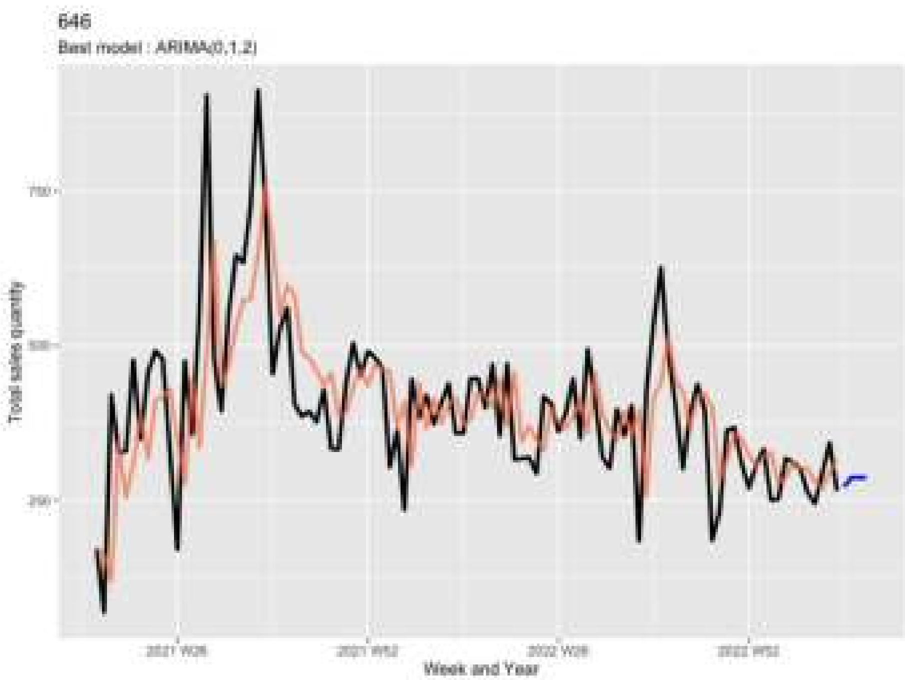
Type	Model	MAE	RMSE	MAPE
Training	arima	63.1	111	22.9
	arima.ext	62.7	109	25.6
	arima lag1	63.2	108	26.2
	ets	64.0	114	25.7
	naive	93.2	125	25.3
	lm	52.0	60.1	16.6
	reg	60	86.8	18.5
	reg.ext	56.6	96.2	17.0
	reg lag1	60.3	86.6	18.9
	snaive	151	200	40
Validation	arima	20.0	31	10.1
	arima.ext	41.9	52.0	15.5
	arima lag1	44.4	55.3	16.4
	ets	31.2	35.6	10.6
	naive	39.2	46.1	13.6
	lm	57.7	60.1	20.4
	reg	47	68.5	16.1
	reg.ext	46.1	77.6	15.1
	reg lag1	60.1	72.9	20.8
	snaive	111	114	30.9

Model	total residual
arima	2.68e+2
arima.ext	1.51e+2
arima lag1	9.99e+1
ets	1.67e+3
naive	7.9e+1
lm	1.17e+0
reg	4.29E-13
reg.ext	7.87E-13
reg lag1	3.41E-13
snaive	3.62e+3

- Best selection explanation

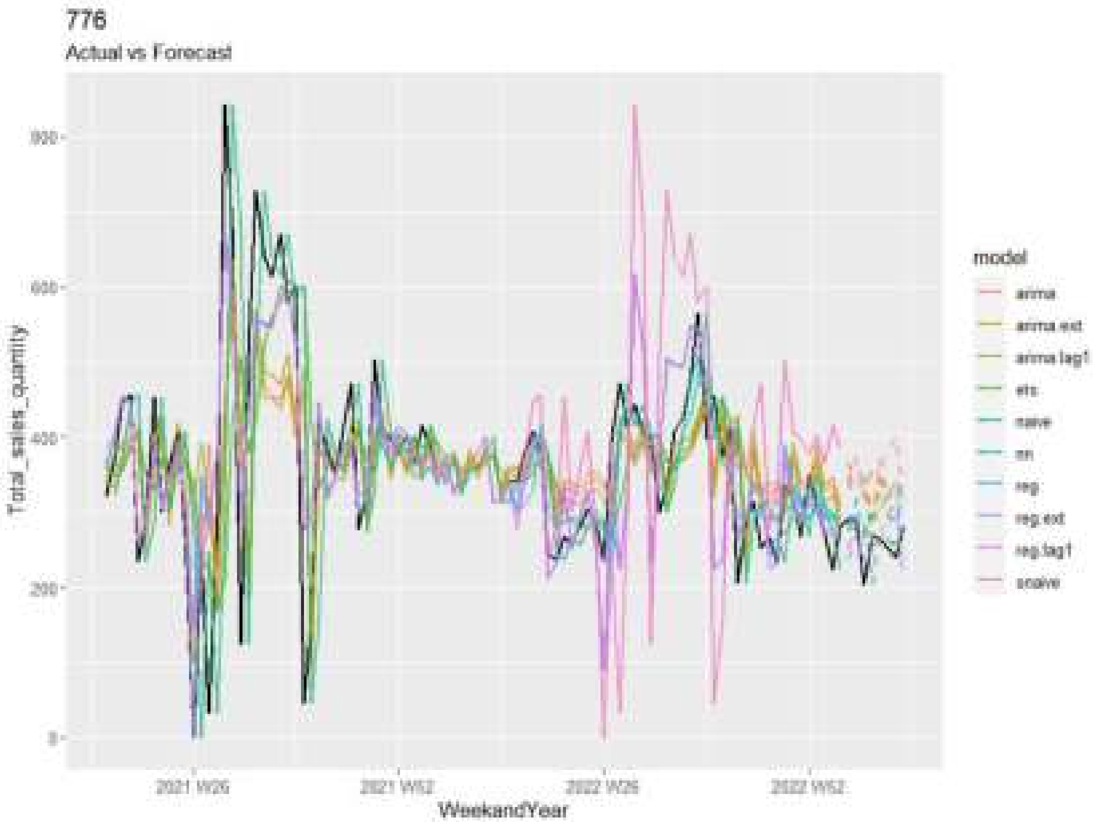
After observing the boxplot, we selected 'ets' and 'arima.' Upon further observation of residuals, 'ets' has a value of -1.67e+3, while 'arima' has 2.68e+2. Choosing 'arima' is preferred as its residual is positive, closer to zero, and the median in the boxplot is closer to 0.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

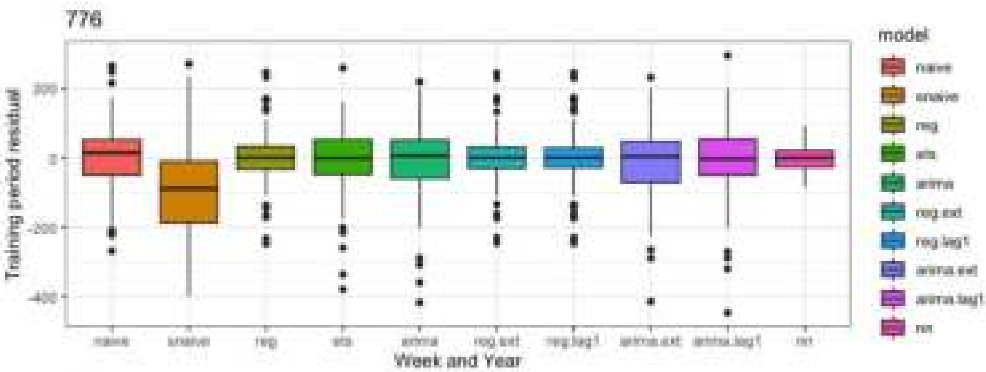


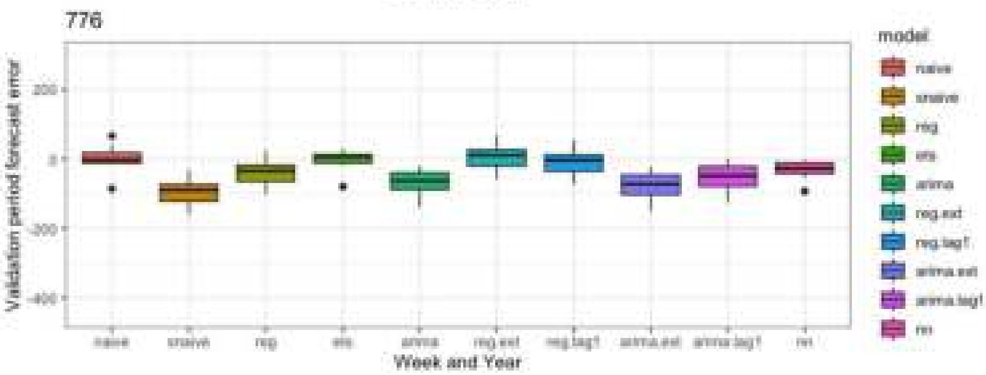
9) Building_Nbr: 776

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method





- Performance evaluation

Type	Model	MAE	RMSE	MAPE
Training	arima	91.57532	119.05767	Inf
	arima.ext	91.04365	115.07645	Inf
	arima lag1	78.55335	114.83048	Inf
	ets	90.87852	133.27101	Inf
	naive	92.35484	144.84016	Inf
	nn	34.12264	42.10087	10.87960
	reg	57.55319	83.88575	Inf
	reg.ext	52.471	60.45592	Inf
	reg lag1	53.92806	82.22312	Inf
	snaive	148.18048	186.16276	43.49093
Validation	arima	87.5	76.2	27.5
	arima.ext	77	96.2	31.3
	arima lag1	53	65.5	22
	ets	22.1	31.9	9.44
	naive	29	41.6	12.1
	nn	31.7	42	13.6
	reg	46	54.3	18.7
	reg.ext	33.1	39.9	12.6
	reg lag1	31.5	39.9	12.5
	snaive	83.9	102	37.6

Model	total residual
arima	1.64e+1
arima.ext	1.02e+1
arima lag1	5.58e+0
ets	2.59e+2
naive	-4.1e+1
nn	3.95e-2
reg	4.83E-13
reg.ext	-1.03E-13
reg lag1	2.84E-13
snaive	2.37e+3

- Best selection explanation

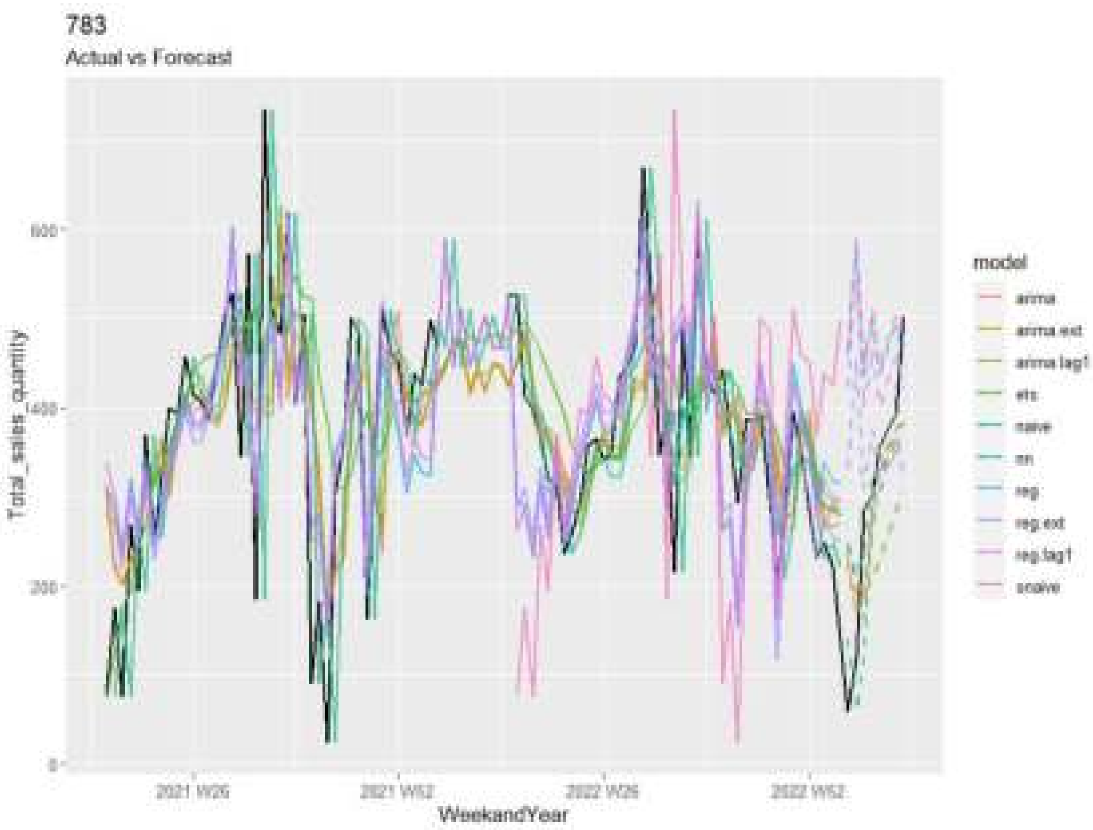
Despite the seemingly good performance of 'ets' in the boxplot, it is observed that 'ets' has too many outliers in the training set, indicating insufficient training. Upon further observation of residuals, 'naive' has a value of -4.1e+1, while 'reg ext' is -1.03e-13. Choosing 'reg ext' is preferred due to its smaller residual and lack of outliers in the training set.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

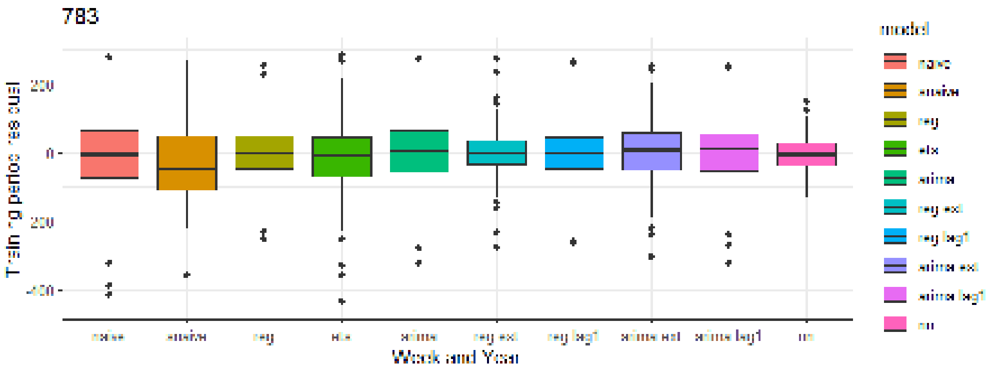
Although in the validation period, our forecasts were based on models that were estimated/trained only on data that was available at the "time of prediction". But when it comes to forecasting the future after we selected the best model, “Average_price” is not available at the time of prediction (we don’t have the future data of “Average_price”), therefore “Total sales quantity” of weeks 2023 W13 to 2023 W16 couldn’t be predicted.

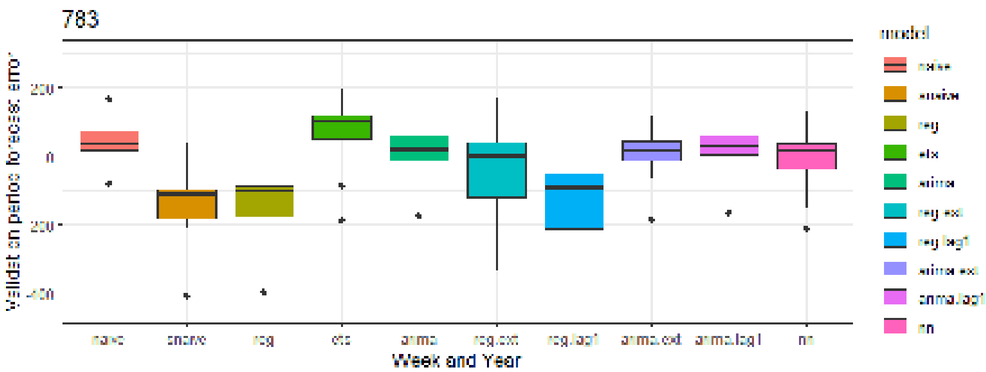
10) Building_Nbr: 783

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method





- Performance evaluation

Type	Model	MAE	RMSE	MAPE
Training	arima	79.5126	107.4004	35.01734
	arima.ext	81.15109	108.11047	36.38247
	arima lag1	81.2468	108.18283	36.82485
	ets	82.69722	116.0771	45.44606
	naive	97.60215	137.881	38.73193
	nn	41.26508	52.15033	13.54214
	reg	63.1468	87.74634	30.80748
	reg.ext	56.17569	80.98035	27.38532
	reg lag1	61.72401	87.08885	30.36887
	snaive	143.35714	185.07355	48.55024
Validation	arima	66.3	87	52.4
	arima.ext	66.5	89.5	55.0
	arima lag1	73.4	90.6	53
	ets	127	133	73.8
	naive	65.5	81.7	37.3
	nn	73.5	92.9	50.7
	reg	191	240	158
	reg.ext	123	160	104
	reg lag1	164	214	140
	snaive	190	247	182

Model	total residual
arima	3.15e+2
arima.ext	2.59e+2
arima lag1	2.77e+2
ets	1.73e+3
naive	6.5e+1
nn	3.77e+0
reg	7.11E-13
reg.ext	2.56E-13
reg lag1	1.31E-12
snaive	3.07e+2

- Best selection explanation

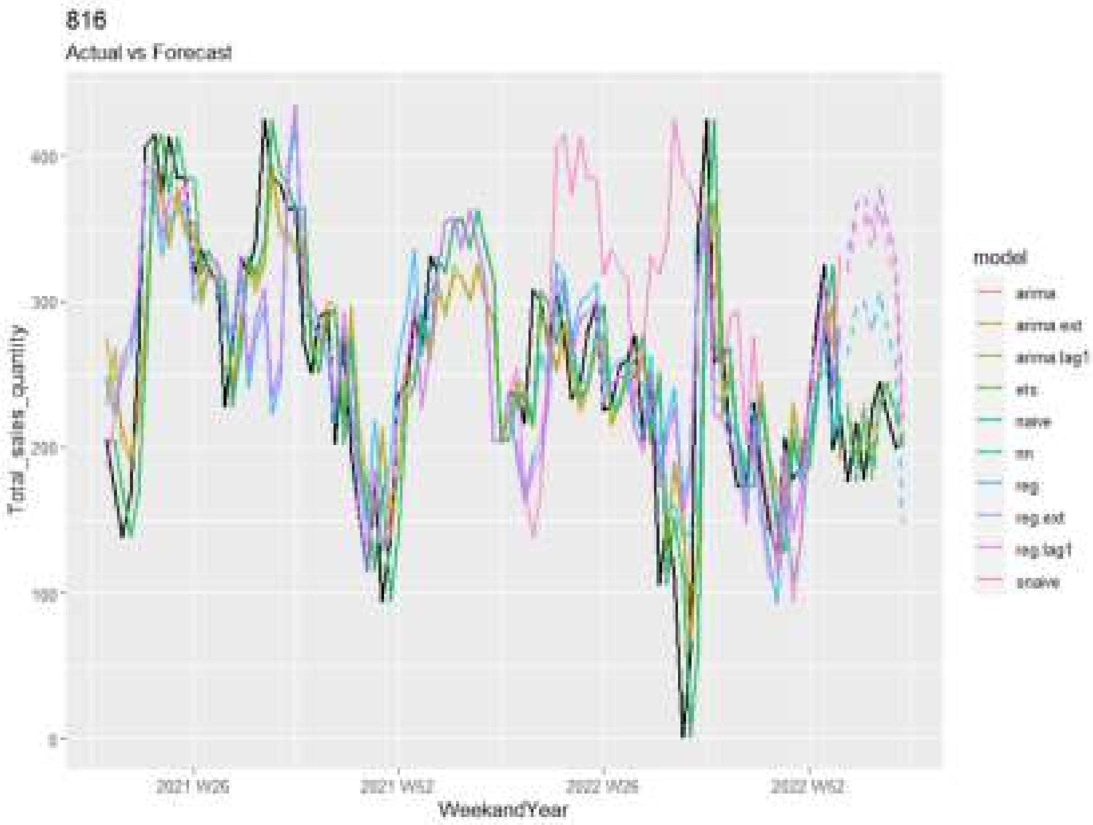
After observing the boxplot, 'arima' and 'arima ext' were selected. Further examination of residuals reveals that 'arima' has a value of 3.15e+2, while 'arima ext' is 2.59e+2. Choosing 'arima ext' is preferred due to its smaller residual.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

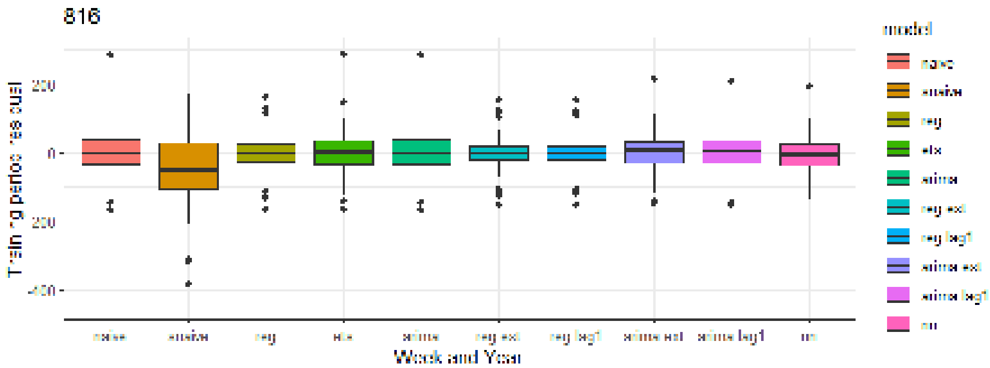
Although in the validation period, our forecasts were based on models that were estimated/trained only on data that was available at the "time of prediction". But when it comes to forecasting the future after we selected the best model, “Average_price” is not available at the time of prediction (we don’t have the future data of “Average_price”), therefore “Total sales quantity” of weeks 2023 W13 to 2023 W16 couldn’t be predicted.

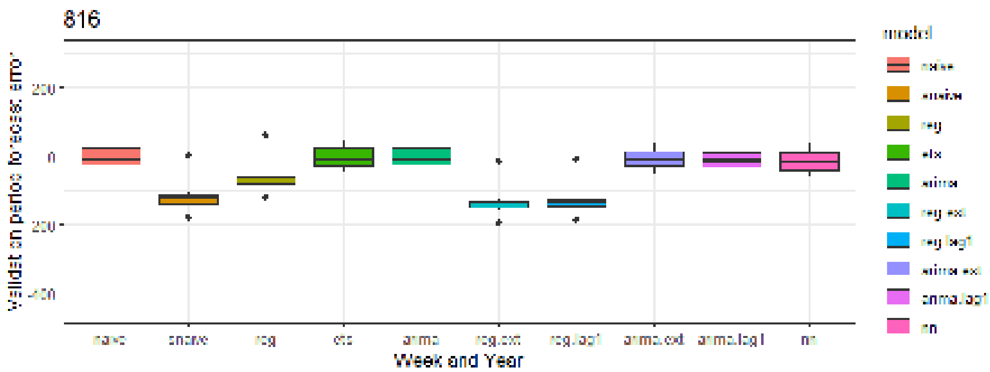
11) Building_Nbr: 816

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method





- Performance evaluation

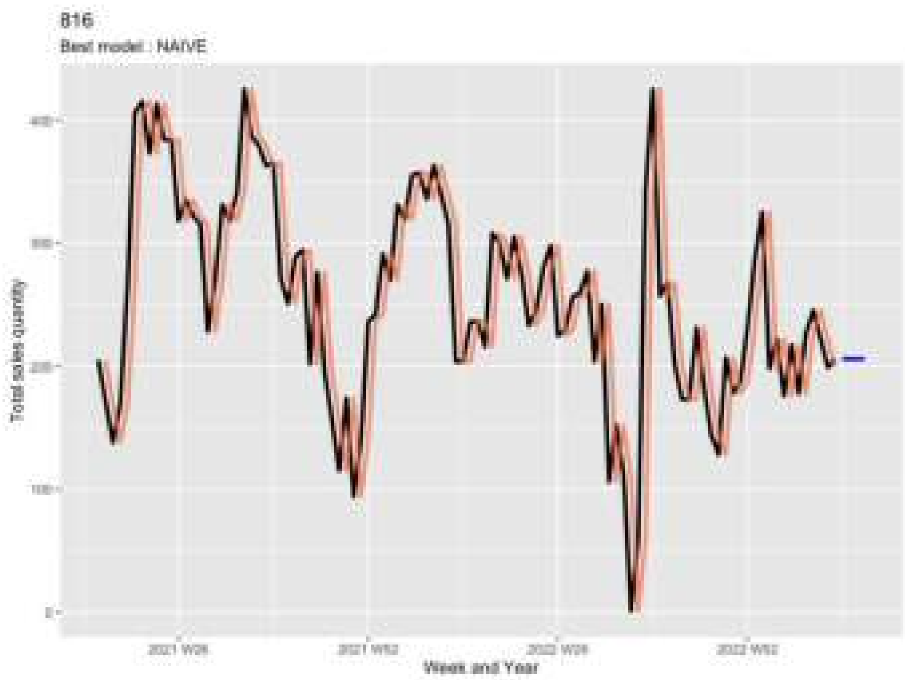
Type	Model	MAE	RMSE	MAPE
Training	arima	45.67238	67.90582	138.4801
	arima.ext	43.44527	57.11583	175.5502
	arima.lag1	43.05609	57.805	179.9311
	ets	45.43664	67.86448	140.1002
	naive	46.16129	63.21282	138.9476
	nn	42.02555	57.65348	348.0578
	reg	38.0156	53.65907	193.6767
	reg.ext	33.98471	50.43754	178.7287
	reg.lag1	34.3741	49.96182	170.673
	snaive	53.35714	126.9584	971.131
Validation	arima	30.8	33.8	15.1
	arima.ext	27.3	32.1	13.8
	arima.lag1	28.2	30.3	11.4
	ets	30.1	32.9	14.8
	naive	30.0	33.0	15.1
	nn	30.6	34.4	15.4
	reg	72.6	75.0	36
	reg.ext	130	139	63.8
	reg.lag1	125	134	61.3
	snaive	115	125	56.7

Model	total_residual
arima	1.72e+1
arima.ext	5.88e+1
arima.lag1	0.08e+0
ets	1.88e+1
naive	1.7e+1
nn	-4.00e-1
reg	5.68E-14
reg.ext	5.83e-13
reg.lag1	1.99E-13
snaive	-2.38e+3

- Best selection explanation

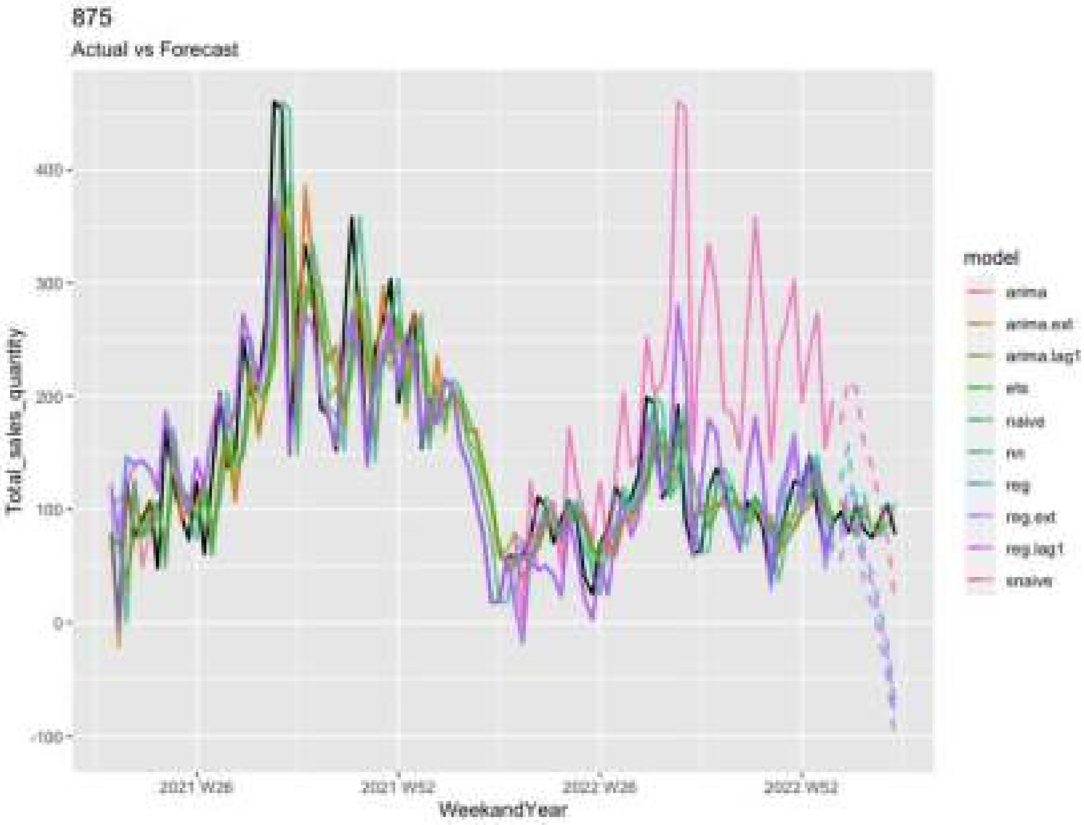
After observing the boxplot, we selected 'naive,' 'ets,' and 'arima.' Upon further examination of residuals, 'naive' has a value of 1.7e+1, 'ets' has 1.88e+1, and 'arima' has 1.72e+1. Choosing 'naive' is preferred based on its smaller residual.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

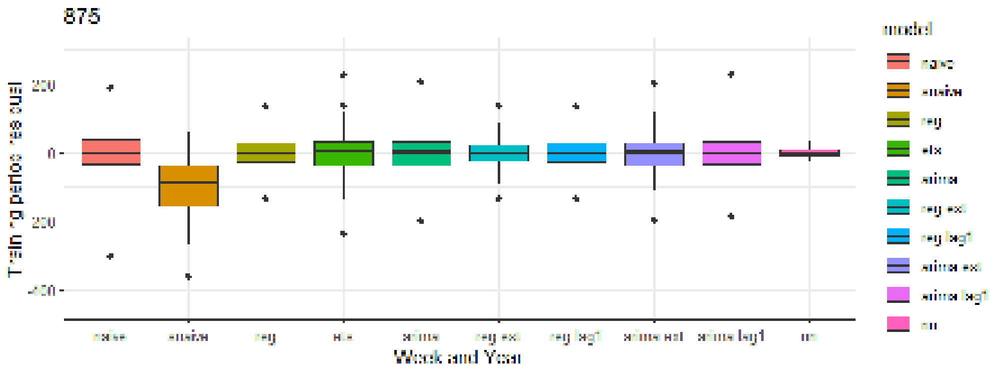


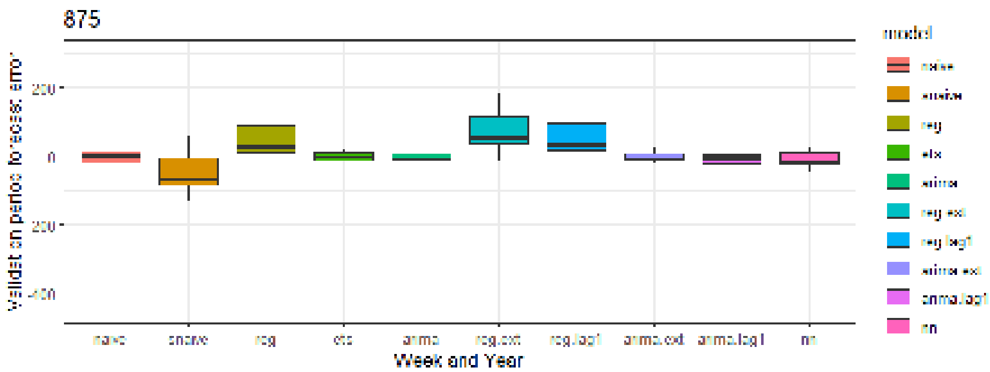
12) Building_Nbr: 875

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method





- Performance evaluation

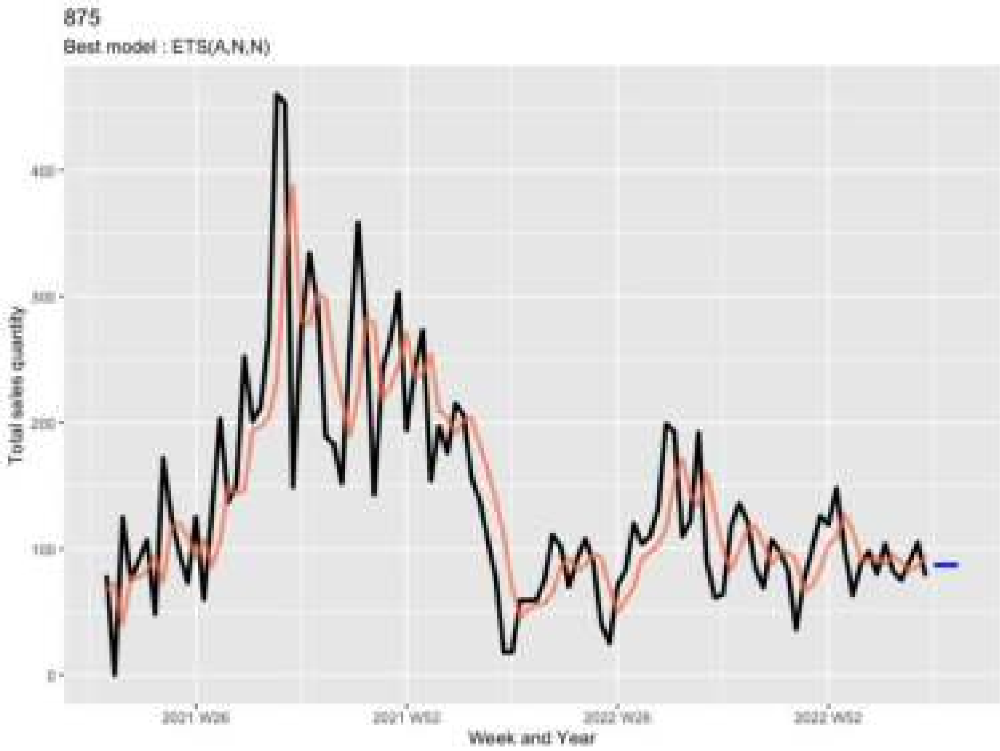
Type	Model	MAE	RMSE	MAPE
Training	arima	43	56.5	Inf
	arima.ext	41.7	54.9	Inf
	arima.lag1	41.0	56.6	Inf
	ets	44.3	61	Inf
	naive	47.7	65.0	Inf
	nn	11	14.2	14.2
	reg	31.1	42.1	Inf
	reg.ext	28.9	38.8	Inf
	reg.lag1	30.6	41.9	Inf
	snaive	103	130	117
Validation	arima	13.8	15.4	15.4
	arima.ext	14	15.8	15.6
	arima.lag1	14.1	16.9	16.0
	ets	13.5	14	15.4
	naive	17.2	19.4	19.7
	nn	26.1	34.7	30.7
	reg	59	77.0	67.9
	reg.ext	77.8	97.1	88.9
	reg.lag1	63.4	80.2	73.1
	snaive	69.8	78.7	80.5

Model	total_residual
arima	3.30e+1
arima.ext	-3.51e+1
arima.lag1	1.57e+2
ets	5.07e+1
naive	9e+0
nn	-9.70e-1
reg	1.65E-13
reg.ext	-1.28E-13
reg.lag1	3.02E-13
snaive	-3.96e+3

- Best selection explanation

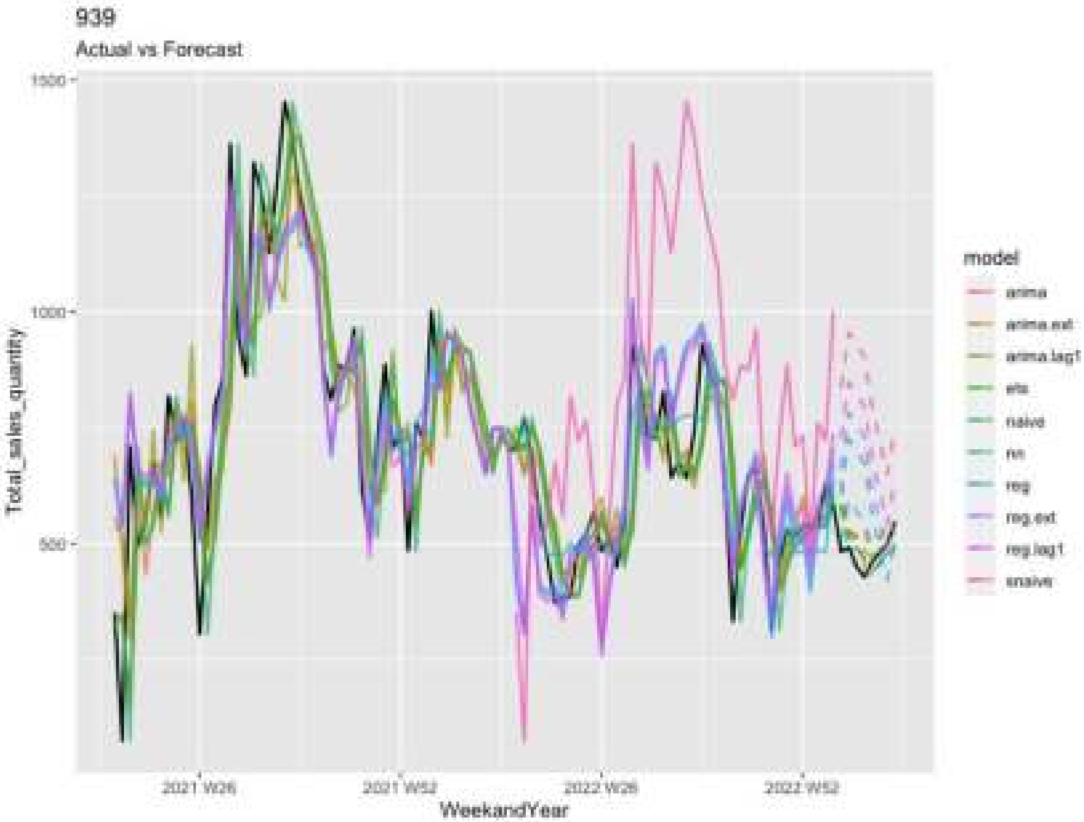
We choose ets as the best model through the boxplot.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

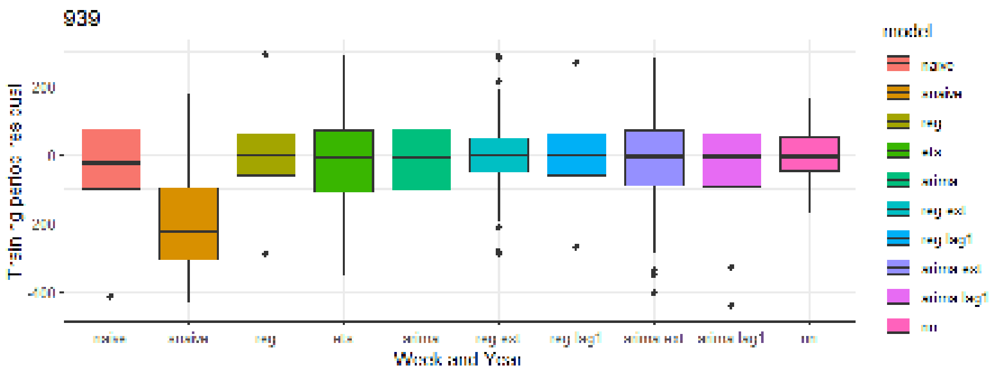


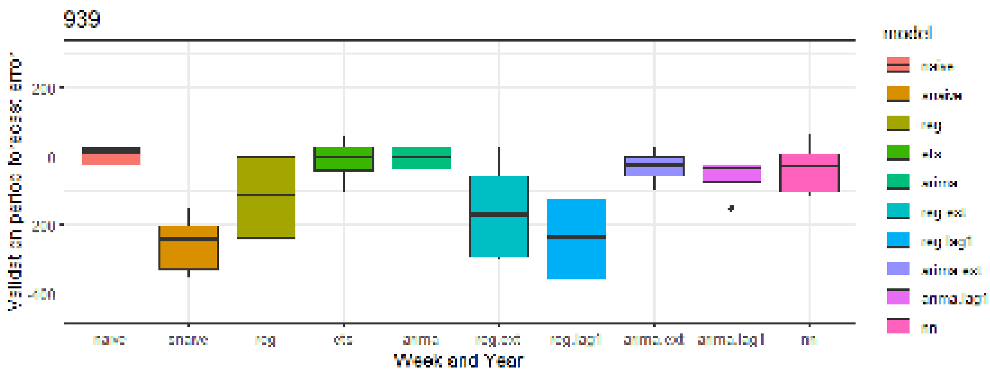
13) Building_Nbr: 939

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method





- Performance evaluation

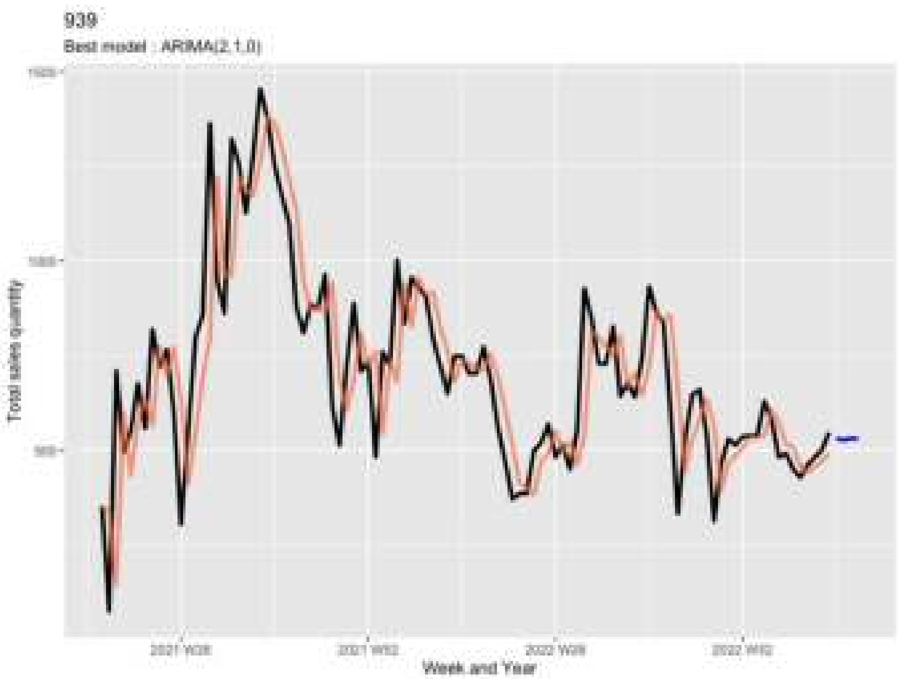
Type	Model	MAE	RMSE	MAPE
Training	arima	122	167	21.9
	arima.ext	116	160	23
	arima lag1	116	158	22.9
	ets	123	169	21.9
	naive	130	175	22.8
	nn	63.0	90.1	11.9
	reg	83.3	124	18.7
	reg.ext	82.2	123	18.7
	reg lag1	82.2	119	18.7
	snaive	293	340	50.6
Validation	arima	42.4	49.7	8.01
	arima.ext	37.9	49.0	8.08
	arima lag1	57.2	73.6	12.2
	ets	44	51.1	8.18
	naive	36.6	45.2	7.82
	nn	61.3	75.0	13.1
	reg	134	167	29
	reg.ext	170	204	36.0
	reg lag1	229	258	48.1
	snaive	337	350	71.7

Model	total residual
arima	3.49e+2
arima.ext	7.30e+2
arima lag1	3.61e+2
ets	3.99e+2
naive	2.3e+2
nn	7.46e-2
reg	5.68E-14
reg.ext	1.42E-12
reg lag1	-2.84E-13
snaive	9.81e+3

- Best selection explanation

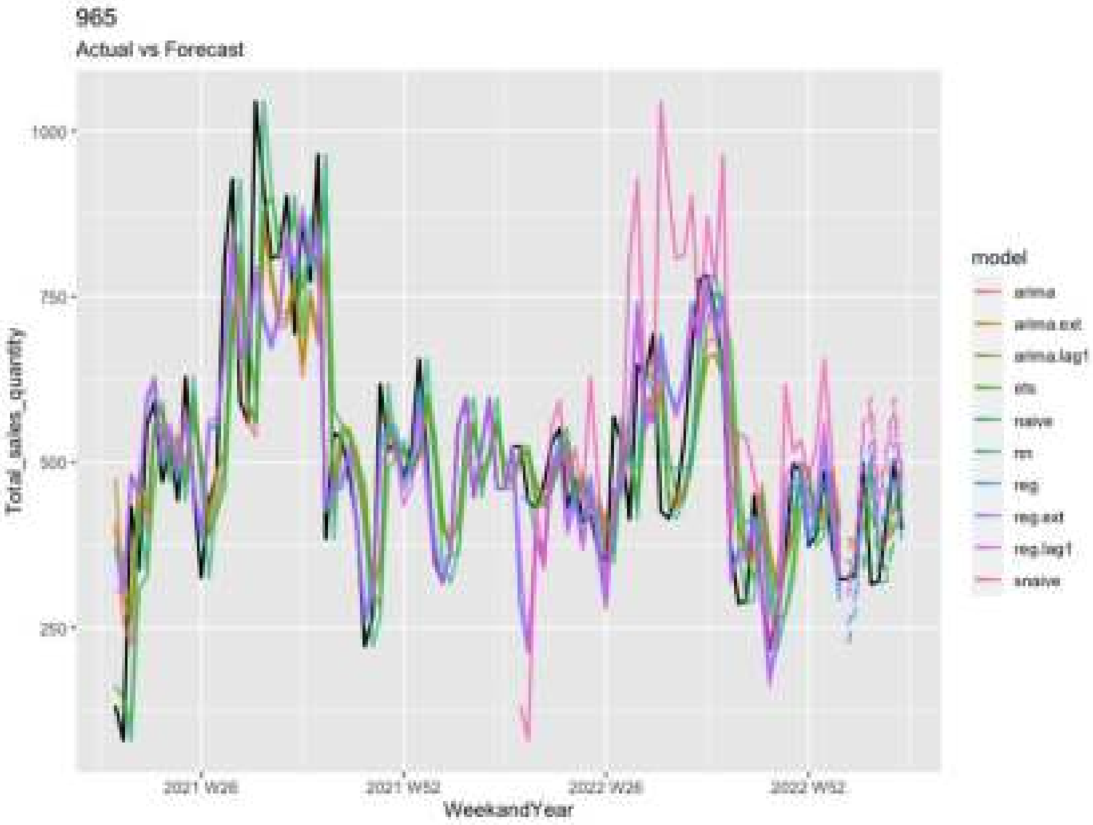
From the boxplot, it is evident that 'ets' and 'arima' perform the best. Upon further observation of residuals, 'ets' has a value of 3.99e+2, while 'arima' has a slightly lower value of 3.49e+2. Therefore, 'arima' is selected.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16

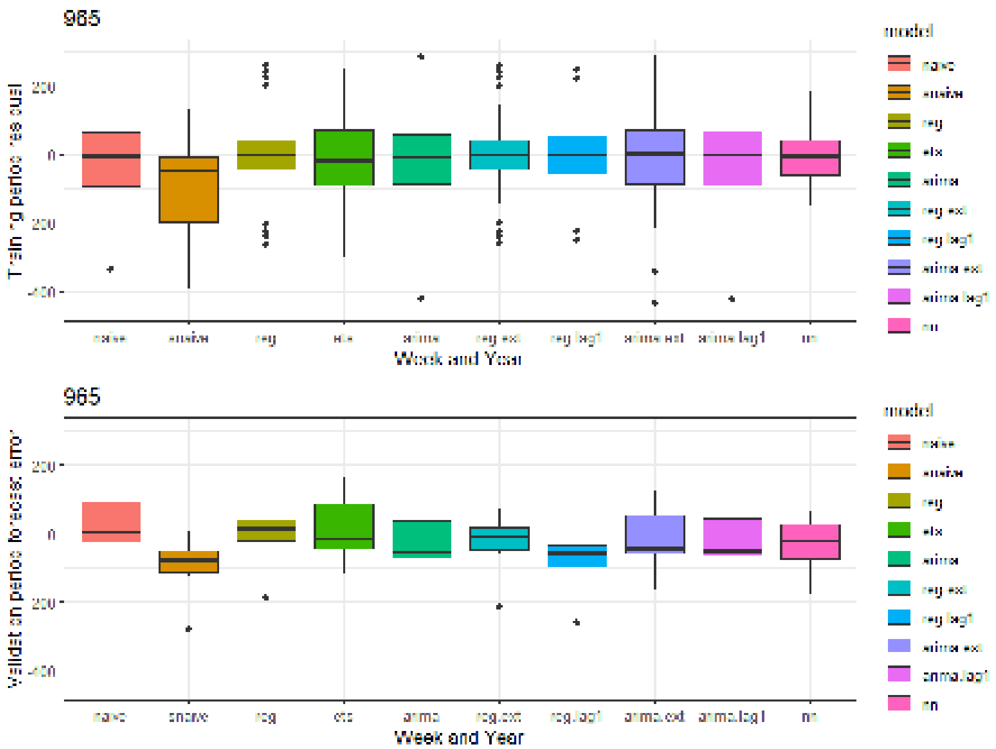


14) Building_Nbr: 965

- Time plots to observe the prediction results of each method



- Boxplots to examine the residuals of each forecasting method



- Performance evaluation

Type	Model	MAE	RMSE	MAPE
Training	arima	105	136	25.2
	arima.ext	103	137	25.4
	arima lag1	104	137	26
	ets	109	144	23.5
	naive	111	151	23.7
	nn	80.7	75.1	14.2
	reg	86.3	91.8	17
	reg.ext	85.0	91.6	17
	reg lag1	87.4	89.3	17
	snaive	155	215	35
Validation	arima	92.0	94.5	22.4
	arima.ext	78.1	69.4	20.0
	arima lag1	80.7	90.1	21.4
	ets	77.3	92.6	19
	naive	81.6	106	20.3
	nn	80.7	64.5	17
	reg	57.8	80.9	17.2
	reg.ext	57.0	65.5	17.2
	reg lag1	83.2	112	23.4
	snaive	98.1	121	26.6

Model	total_residual
arima	3.36e+2
arima.ext	3.81e+2
arima lag1	2.88e+2
ets	3.07e+2
naive	1.91e+2
nn	9.90e-2
reg	5.97E-13
reg.ext	-3.41E-13
reg lag1	-5.68E-13
snaive	3.05e+3

- Best selection explanation

Boxplot analysis reveals that 'regression' and 'regression external' perform the best. Observing the residuals, 'regression'' has a value of 5.97e-13, while 'regression external' has -3.41e-13. We select 'reg' due to its positive residual sum and simpler model structure.

- Total sales quantity prediction for weeks 2023 W13 to 2023 W16