NTHU BAFT 2023

BUSINESS ANALYTICS USING FORECASTING
FINAL PROJECT REPORT

Forecasting Next-day Smartphone Usage for Self-Time Management

Team members:

Team 6

Lai Heng Sim (Madison)

Kuan-Chung Chen (Mario)

Yu-Pei Lin (Claire)

Yi-Jiuan Chung (Angel)

EXECUTIVE SUMMARY

Business problem

People often overuse their smartphones due to a lack of effective time management tools. Our solution involves creating a predictive model based on historical data to guide users in setting reasonable time limits, benefiting those interested in self-time management.

Data

The data source is from our four team members' daily phone usage records, and we are measuring the usage time every day (00:00 ~ 23:59). In addition, we also consider the external information such as holidays, events, class hours, working hours, previous-day social media and entertainment usage and phone pickup times that may affect our phone usage time.

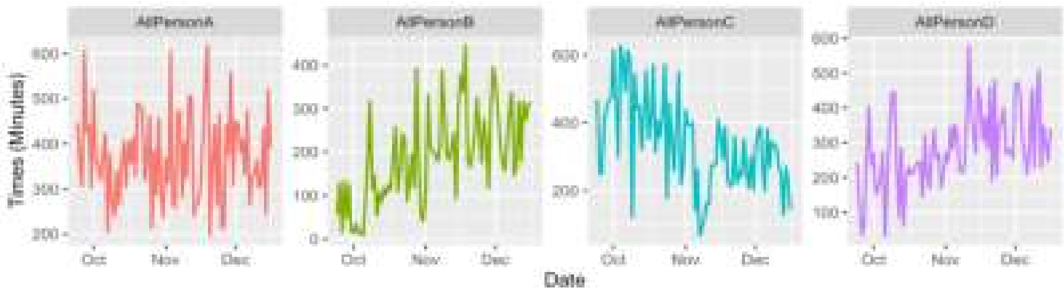


Figure.1 The actual phone's time usage charts of four group members.

Forecasting Solution

Users will get a predicted time usage report every 00:30, and make self-time management decisions according to the report, getting closer to their goals of proper time of phone usage.

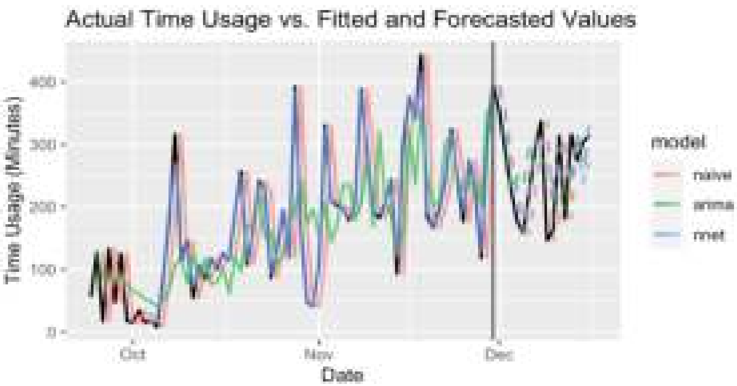


Figure.2 The sample of Person B's forecasting solution.

Recommendation

- Review and refresh the forecasting model every week to best fit the dynamic user behaviors.
- Integrate gamification and social features to incentivize user participation and engagement.
- Enhance forecasting models by incorporating user-selected reduction preferences and refining time units (e.g., 15-minute intervals) to improve accuracy.

1. BUSINESS GOAL

In our group, we have observed that we spend too much time using smartphones every day, and we want to establish time limits for the time usage of smartphones, cultivating better self-time management and enhancing our quality of life.

Regarding the existing time management function in smartphones, it only requires us to set the fixed usage time limit on the apps we would like to control. However, we do not know what is the proper time limit to set, and we are afraid that if we set the time limit too aggressively, it will lead to frustration and we may give up on time management eventually. To solve this problem, we would like to develop a forecasting model based on our historical behavior data to provide predicted time usage of phones for the following day, and we can set reasonable time limits according to it. We believe such methods can enhance our self-time management efficiency and, more importantly, boost our self-discipline and confidence. We also hope this can benefit those who are interested in self-time management in the future.

2. FORECASTING GOAL

Our forecasting goal is to predict **daily time usage for smartphones** from a **forward-looking** perspective. The forecasting horizon is set on a **daily scale**, allowing users to plan and manage their app usage every day effectively. The primary purpose of this forecast is to provide the predicted time phone usage. Users can set the time usage limits according to the predicted result and their preferred daily reduction rate on time usage. We define the success of our forecasting goal into short-term and long-term success in the aspect of users. For short-term success, when users hit the everyday reduction rate target of phone time usage, users will gradually get confidence in self-management and be closer to their long-term goal. For long-term success, when the predicted values align with the user's ideal goals, then it represents a huge milestone in self-management. For the details of the user's interface, please see Appendix A.

	Forecasts of Time Usage in Dec. 17, 2023	Reduce 5% (Short-term goal)	Reduce 10% (Short-term goal)	Reduce 20% (Short-term goal)	Long-term Goal
Time Usage (mins)	200	190	180	160	120

Figure.3 The summary table example for forecasting goals in terms of users.

Users will get everyday forecasts at **12:30 am**, and then decide their preferred reduction rate to gradually achieve their long-term goal.

3. DATA

3.1 Data Description

- Source: Smartphones of each team member (both iOS and Android)
- What is measured: Daily smartphone usage (unit: minutes)
- Time period: Daily (00:00 ~ 23:59)
- Amount of data: from September 24th to December 16th (84 records)
- Sample of 10 rows for the entire data:

Date	PersonA	PersonB	PersonC	PersonD	AllPersonA	AllPersonB	AllPersonC	AllPersonD	IsHoliday	ClassHourA	ClassHourB	ClassHourC	ClassHourD	PickupsA	PickupsB	PickupsC	PickupsD	WorkingHourB	WorkingHourC	WorkingHourD	IsEventA	[illegible]	IsEventB	IsEventC	IsEventD
2023-09-24	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2023-09-25	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2023-09-26	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2023-09-27	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2023-09-28	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2023-09-29	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2023-09-30	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2023-10-01	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2023-10-02	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
2023-10-03	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

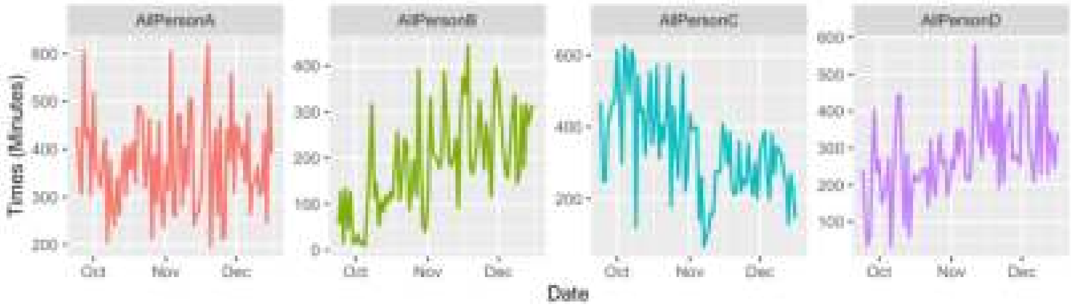


Figure.4 Time Plots of Each Relevant Series.

3.2 External Series

3.2.1 Dummy Variables (0=no, 1=yes)

- IsHoliday: Indicates whether that day is a public holiday.
- IsEvent: Indicates whether there is an event on that day that may affect phone usage, aside from holidays (e.g., reunion, travel, conference, etc.).

3.2.2 Continuous Variables

- Class hours (minutes)
- Working hours (minutes)
- Yesterday’s Time usage of social media and entertainment apps (minutes)
- Yesterday’s Pickups of phones

Social Media Usage Class Hours Pickup of Phones IsEvent

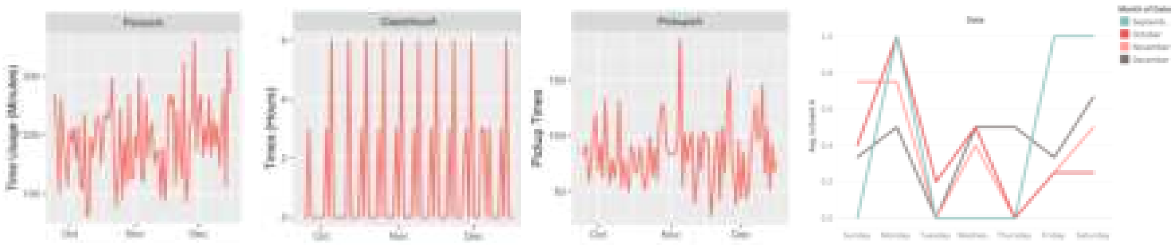


Figure.5 The Time Plots of External Series for Person A.

3.3 Re-processing

Since the data record unit provided by the mobile phone is in hours (eg. 1 hour 30 minutes), we **converted the time unit into minutes** (eg. 90 minutes) for standardization.

3.4 Data Preparation

We compiled smartphone usage data from four team members, including daily records of overall usage time, social media app usage duration, and the frequency of phone pickups, all meticulously recorded in an Excel spreadsheet. Furthermore, variables like class hours, work hours, IsHoliday, and IsEvent were logged individually by each user based on their unique schedules and circumstances.

4. METHODS

4.1 Data Partitioning

We partitioned our data into training periods (before November 30th) and validation periods (after and including December 1st). We used the data in the training period to model and do the roll forward one-step-ahead forecast.

4.2 Forecasting Methods

We modeled the four series individually. For each series, we proposed using Naive, Seasonal Naive, and Sample Average as our modeling benchmarks. Moreover, we employed the following advanced modeling to forecast the future values of smartphone usage (coding formulas are listed in Appendix B):

1. Exponential Smoothing Method [`ets`]
2. Regression models both without and with external information [`tslm/tslm.ext`]
3. Two-layer models with regression (both without and with external information) [`twolayer / twolayer.ext`] (Note: ARIMA with trend and seasonality index)
4. ARIMA models (both without and with external information) [`arima/arima.ext`]
5. Neural Nets models (both without and with external information) [`nnet/nnet.ext`]

The hyperparameters and settings of the models above will be automatically selected by R using some information criteria.

4.3 Performance Measure

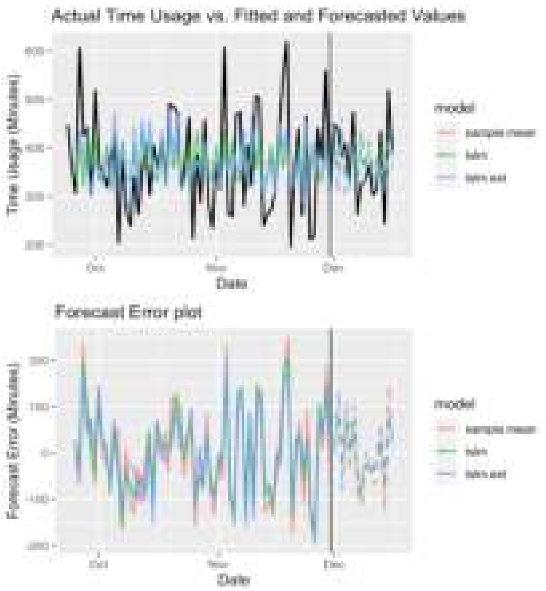
For each series, we fit a total of 3 benchmarking models and 9 alternative models. The final forecasting method for each series will be determined based on performance charts in both periods and performance metrics (RMSE and MAPE) for reference.

5. PERFORMANCE EVALUATION

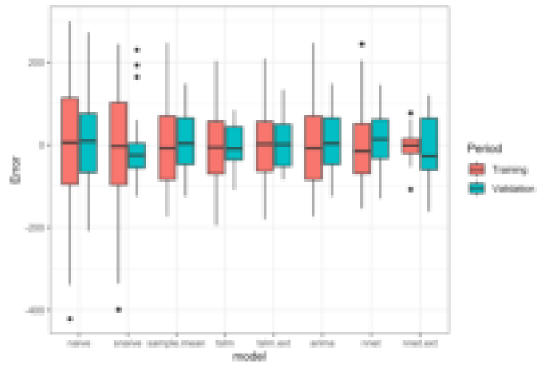
We performed the **top-performing benchmark** among three benchmark options and the **two best** models among the nine alternatives in the plot. Our model evaluations were guided by three main criteria: ensuring no outliers in the boxplot of forecast errors, preferring over-forecasting over under-forecasting, and aiming for minimal forecast errors to indicate superior model performance. For simplicity, we only present the results for Person A. Appendix D provides detailed performance and results for the other series.

5.1 Forecasting Solution for Person A

(1) Actual vs. Roll-forward Forecast Plot and Forecast Error Plot



(2) Boxplot of Forecast Error



(3) Performance Metrics

		Training		1-step-ahead roll forward	
		RMSE	MAPE	RMSE	MAPE
Selected model	TSLM	95.2	23.3	57.4	13.6
	TSLM+ext.	90.5	21.2	64.1	13.7
Benchmark	NAÏVE	153	37.2	123	27.0
	SNAÏVE	148	36.5	105	21.3
	Average	104	25.5	74.9	18.0

Figure.6 The performance charts of the forecasting solution for Person A.

For Person A, the benchmark was the **sample mean**. While Seasonal Naive performed the best if we focused on the errors, it exhibited numerous outliers. Consequently, we selected the sample mean as the benchmark. The top two models were **TSLM** and **TSLM including external information**. They achieved the lowest RMSE and MAPE among all the models, displaying the least errors overall in the box plot. The report of the model is displayed in Appendix C.

During the validation period, except for TSLM, none of the models surpassed the performance of the sample mean. Furthermore, according to the boxplot, NNET with external information performed well (overfitting) in the training period but did not perform as strongly in the validation.

5.2 Future Forecasts for Person A

We have chosen **TSLM** as our final model for Person A. The forecast for Dec 17 is 406.66 mins, and the 95% confidence interval is [212, 600].

6. CONCLUSIONS

6.1 Advantages and Limitations

Our project aims to identify optimal smartphone usage limits to help users address excessive phone usage. We use predictive models to forecast usage for the next day. Users then choose the degree of reduction in their ideal usage time. Through a gradual approach, we strive to assist users in establishing healthier smartphone habits by respecting their pace and reducing dependence systematically.

However, from an implementation perspective, our project relies heavily on user participation. It requires informing users about activities that may impact their smartphone usage on a given day and asking users to choose a preferred reduction in smartphone usage. Therefore, the success and availability of the system depend on the level of engagement demonstrated by the user.

Additionally, due to time constraints, the amount of data we've collected may be insufficient to observe trends and seasonality. Therefore, we believe that gathering more data would help improve the accuracy of our predictions.

6.2 Implementation and Recommendations

Our project operates in real-time, requiring continuous data collection and analysis. At the same time, reviewing and refreshing the forecasting model every week ensures best fits the dynamic user behaviors.

Moreover, we suggest further enhancing the forecasting models. This could involve incorporating the user's selected degree of reduction in smartphone usage as external information (which cannot be added at this stage as it requires daily collection) and adjusting the time units (considering, for instance, 15-minute intervals) to enhance prediction accuracy. Implementing these enhancements should help create a better and more accurate forecasting system.

Looking ahead, we recommend streamlining the process and minimizing disruptions. Additionally, exploring gamification elements or incorporating social features could serve as effective strategies to incentivize user participation.

To boost user confidence, we prefer to slightly over-forecast our forecasting results. This approach provides users with an additional safety margin and better addresses uncertainties in the data.

APPENDIX

Appendix A: The user interface of our time management solution

Figure.7 The user's interface of our time management solution.

Appendix B: The R code formula for each series (Person A as an example)

```
set.seed(123)
fit.personA <- data.person.train |> model(
naive = NAIVE(Value),
snaive = SNAIVE(Value),
sample.mean = ARIMA(Value ~ pdq(0,0,0) + PDQ(0,0,0)),
ets = ETS(Value),
arima = ARIMA(Value, stepwise = F),
tslm = TSLM(Value ~ trend() + season()),
twolayer = ARIMA(Value ~ trend() + season(), stepwise = F),
tslm.ext = TSLM(Value ~ trend() + season() + IsHoliday + ClassHourA + IsEventA +
                    PickupsA.Lag1 + PersonA.Lag1),
arima.ext = ARIMA(Value ~ IsHoliday + ClassHourA + IsEventA + PickupsA.Lag1 +
                    PersonA.Lag1, stepwise = F),
twolayer.ext = ARIMA(Value ~ trend() + season() + IsHoliday + ClassHourA +
                           IsEventA + PickupsA.Lag1 + PersonA.Lag1 + PDQ(0,0,0),
                           stepwise = F),
nnet = NNETAR(Value),
nnet.ext = NNETAR(Value ~ IsHoliday + ClassHourA + IsEventA + PickupsA.Lag1 +
                    PersonA.Lag1)
)
```

Appendix C: The report of two selected effective models for Person A

TSLM

```
Series: Value
Model: TSLM

Residuals:
     Min       1Q   Median       3Q      Max
-195.071  -69.564   -5.925   56.942  203.780

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   321.75574   39.34056   8.179 2.73e-11 ***
trend()        -0.02129    0.64164  -0.033   0.9736
season()week2 113.39122   46.63217   2.432   0.0181 *
season()week3  35.96807   46.61451   0.772   0.4434
season()week4  92.65602   46.60568   1.988   0.0514 .
season()week5  13.63613   45.40239   0.300   0.7650
season()week6  11.45742   45.37971   0.252   0.8015
season()week7  84.77871   45.36610   1.869   0.0666 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 101.4 on 59 degrees of freedom
Multiple R-squared: 0.1644, Adjusted R-squared: 0.06525
F-statistic: 1.658 on 7 and 59 DF, p-value: 0.13719
```

TSLM.ext

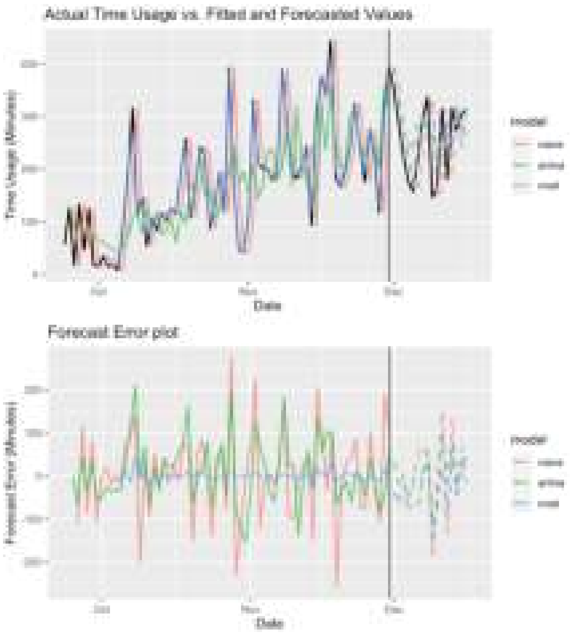
```
Series: Value
Model: TSLM

Residuals:
     Min       1Q   Median       3Q      Max
-180.085  -63.885    1.699   56.263  208.802

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   347.4819   129.5114   2.683  0.00966 **
trend()        -0.3268     0.7174  -0.455  0.65058
season()week2  45.5483   106.0844   0.429  0.66937
season()week3 -27.7633   103.2830  -0.269  0.78910
season()week4  46.1073   104.4598   0.441  0.66069
season()week5 -30.0233   103.0583  -0.291  0.77192
season()week6 -23.5688    68.6518  -0.343  0.73270
season()week7  22.2507   105.8442   0.210  0.83429
IsHoliday     -39.6640    42.2474  -0.939  0.35199
ClassHourA    -15.5802    17.3626  -0.897  0.37352
IsEventA      -56.3395    31.7802  -1.773  0.08190 .
PickupsA.Lag1   0.5235     0.5238   0.999  0.32203
PersonA.Lag1    0.1798     0.2410   0.746  0.45879
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 100.8 on 54 degrees of freedom
Multiple R-squared: 0.2446, Adjusted R-squared: 0.07671
F-statistic: 1.457 on 12 and 54 DF, p-value: 0.16991
```

Appendix D: Performance Evaluation (Cont.)

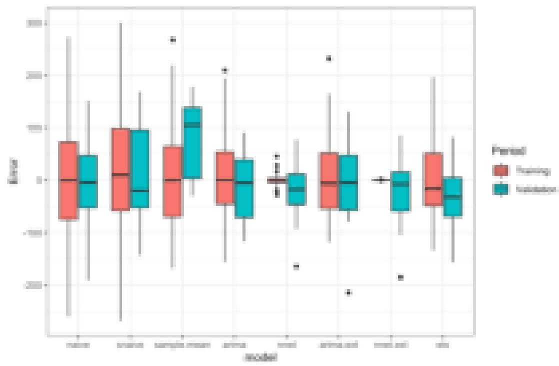
We perform the results of Person B, Person C, and Person D here.

D.1 The forecasting results of Person B:

(1) Actual vs. Roll-forward Forecast Plot and Forecast Error Plot



(2) Boxplot of Forecast Error



(3) Performance Metrics

		Training		1-step-ahead roll forward	
		RMSE	MAPE	RMSE	MAPE
Selected model	ARIMA	76.0	67.3	60.2	23.9
	NNETAR	10.6	8.0	61.3	21.9
Benchmark	NAÏVE	104	73.4	87.5	31.5
	SNAÏVE	119	75.5	90.1	32.2
	Average	106	157.0	104	29.7

Figure.8 The performance charts of the forecasting solution for Person B.

For Person B, the benchmark was **the Naive forecast** due to larger errors in both Seasonal Naive and Sample Mean. Moreover, the median of the Seasonal Naive model indicates over-forecasting, while the Sample Mean exhibited a considerable under-forecasting tendency in the validation period.

In summary, **ARIMA and NNET** emerged as the most effective models with the lowest RMSE and MAPE among all models. ARIMA demonstrated no outliers, and its median was closer to "0" compared to ETS in the boxplot of the forecast error. While NNET had smaller errors, it exhibited signs of overfitting during the training period.

We have chosen **ARIMA** as our final model for Person B. The forecast for Dec. 17 is 267.98 mins, and 95% confidence interval is [118, 418].

The report of two selected effective models for Person B:

ARIMA

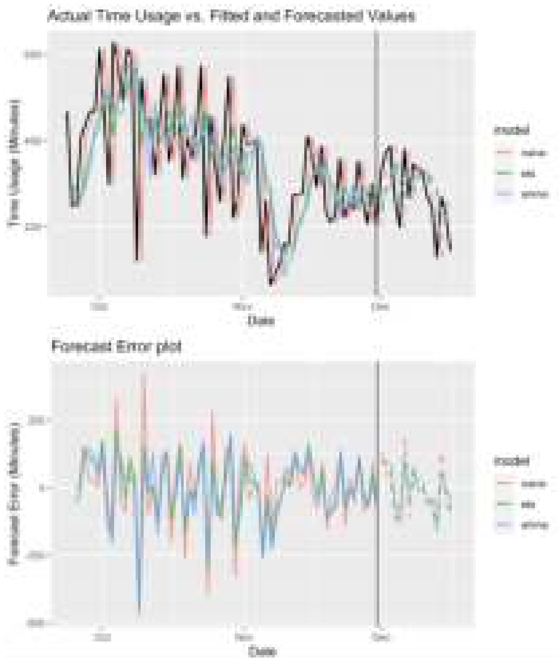
```
Series: Value
Model: ARIMA(0,1,1)(1,0,2)[7]

Coefficients:
          ma1    sar1     sma1    sma2
      -0.8747  0.6370  -0.7890  0.5850
s.e.   0.0600  0.1666   0.1719  0.1812

sigma^2 estimated as 6245:
log likelihood=-384.25
AIC=778.5   AICc=779.5   BIC=789.45
```

NNET

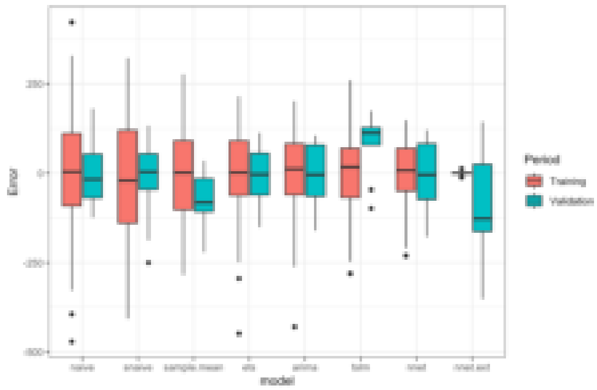
```
Series: Value
Model: NNAR(9,1,5)[7]

Average of 20 networks, each of which is
a 9-5-1 network with 56 weights
options were - linear output units

sigma^2 estimated as 112.8
```

D.2 The forecasting results of Person C:

(1) Actual vs. Roll-forward Forecast Plot and Forecast Error Plot



(2) Boxplot of Forecast Error



(3) Performance Metrics

		Training		1-step-ahead roll forward	
		RMSE	MAPE	RMSE	MAPE
Selected model	ETS	123	38.8	78.3	29.4
	ARIMA	118	38.0	83.4	29.8
Benchmark	NAÏVE	154	43.1	91.9	31.3
	SNAÏVE	172	61.5	98.4	34.5
	Average	140	48.9	104	41.4

Figure.9 The performance charts of the forecasting solution for Person C.

For Person C, the benchmark was **the Naive forecast** due to larger errors in both Seasonal Naive and Sample Mean. Moreover, the median of the Seasonal Naive model and Sample Mean indicates over-forecasting tendency in the validation period.

ETS and ARIMA have displayed the most effective models, displaying the lowest RMSE and MAPE among all models. Both models exhibit similar patterns in the charts. However, the Interquartile Range (IQR) for the error of ETS in the validation period is smaller than that of ARIMA. Additionally, apart from ETS and ARIMA, none of the other models outperformed the benchmark – the NAIVE model.

We have chosen **ETS** as our final model for Person C. The forecast for Dec. 17 is 207.98 mins, and 95% confidence interval is [75, 341].

<u>The report of two selected effective models for Person C:</u>

ETS

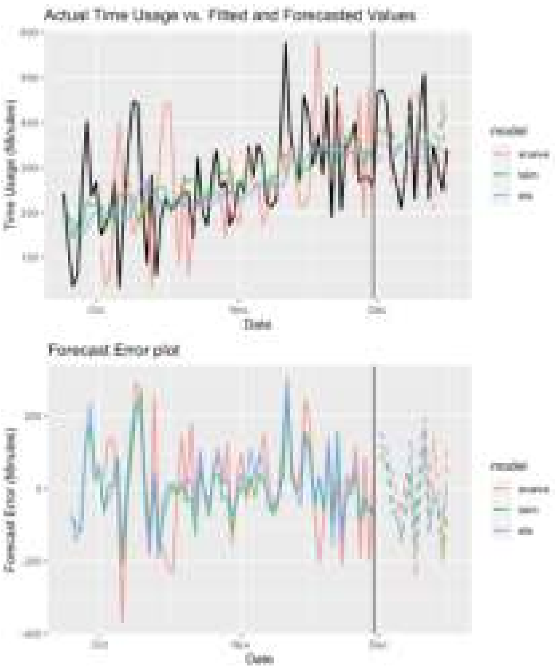
```
Series: Value
Model: ETS(M,N,N)
  Smoothing parameters:
    alpha = 0.3984098

  Initial states:
     l[0]
 300.3202

  sigma^2:  0.1173

     AIC     AICc      BIC
921.7557 922.1367 928.3698
```

ARIMA

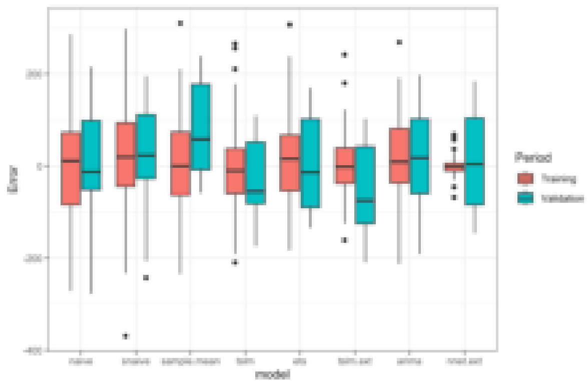
```
Series: Value
Model: ARIMA(2,1,0)

Coefficients:
          ar1      ar2
      -0.7030  -0.4648
s.e.   0.1082   0.1076

sigma^2 estimated as 14559:
log likelihood=-409.35
AIC=824.69   AICc=825.08   BIC=831.26
```

<u>D.3 The forecasting results of Person D:</u>

(1) Actual vs. Roll-forward Forecast Plot and Forecast Error Plot



(2) Boxplot of Forecast Error



(3) Performance Metrics

		Training		1-step-ahead roll forward	
		RMSE	MAPE	RMSE	MAPE
Selected model	TSLM	92.3	43.9	95.6	28.8
	ETS	101	46.6	105	28.4
Benchmark	NAÏVE	123	54.7	137	36.4
	SNAÏVE	139	60.1	115	28.2
	Average	108	57.7	124	24.8

Figure.10 The performance charts of the forecasting solution for Person D.

For Person D, the benchmark was **the Naive forecast**. Moreover, the median of the Seasonal Naive model and Sample Mean indicates under-forecasting tendency and Naive forecast indicates over-forecasting in the validation period.

TSLM and ETS have proven to be the most effective models, showcasing the lowest RMSE and MAPE among all considered models. The majority of models indicate over-forecasting during the validation period, except for the ARIMA and NNET.EXT models.

We have chosen **TSLM** as our final model for Person D. The forecast for Dec. 17 is 407.62 mins, and 95% confidence interval is [208, 607].

<u>The report of two selected effective models for Person D:</u>

TSLM

```
Series: Value
Model: TSLM

Residuals:
    Min      1Q  Median      3Q     Max
-209.47  -60.04  -10.19   37.85  265.87

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    157.275     38.134   4.124 0.000118 ***
trend()          2.575      0.622   4.141 0.000112 ***
season()week2   54.294     45.203   1.201 0.234502
season()week3   41.719     45.185   0.923 0.359624
season()week4   44.254     45.177   0.980 0.331293
season()week5   39.526     44.010   0.898 0.372779
season()week6  -21.349     43.989  -0.485 0.629234
season()week7   10.075     43.975   0.229 0.819573
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 98.32 on 59 degrees of freedom
Multiple R-squared: 0.2681, Adjusted R-squared: 0.1812
F-statistic: 3.087 on 7 and 59 DF, p-value: 0.0076973
```

ETS

```
Series: Value
Model: ETS(A,N,N)
  Smoothing parameters:
    alpha = 0.1289138

  Initial states:
     l[0]
 190.8189

  sigma^2:  10472.45

     AIC     AICc      BIC
905.8697 906.2506 912.4838
```

Note that **we do not consider the ensemble model** since the correlations of the residuals (forecasts error in the training period) of the above models are all positively correlated, as shown in the figure below. The figure displays the scatterplot and correlation of the residuals of Person A among different models in the training period, and it has a similar pattern for the other key series.

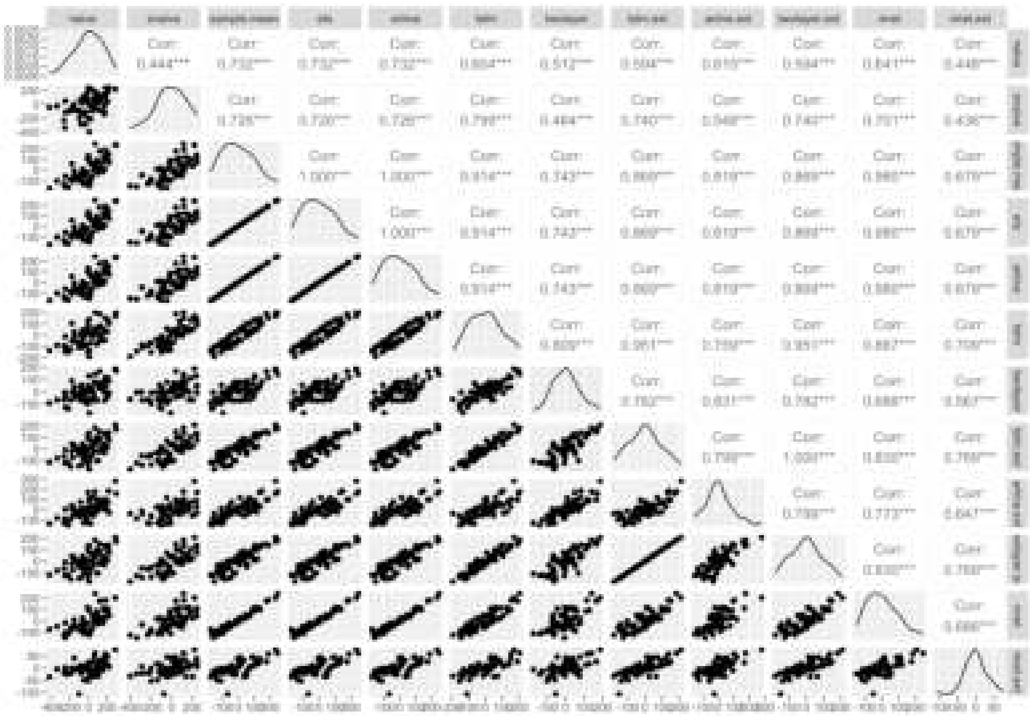


Figure.11 The scatterplot and correlation of the residuals of Person A among different models in the training period.