

hierarchical_forecasting_report

Janice Hsu

2023-09-30

Contents

1	Preliminary Analysis	2
1.1	Data Introduction	2
2	Exploratory Data Analysis	3
2.1	Number of patients entering ED under different hospital hierarchy	3
2.2	Group the changed Local Health Board together	4
2.3	Number of patients who enter ED under 6 different local health boards	4
2.4	Seasonality of number of attendances	5
2.5	Change the Age_Code structure into different groups	6
2.6	Plot Number of Patients in different age groups	6
3	Forecast with ARIMA	8
3.1	Plotting using ARIMA	8
4	Forecasting with ETS	12
4.1	Plotting with ets	12
5	Accuracy assessment for ARIMA and ETS	15
5.1	Forecasting using ARIMA and ETS	15
5.2	Calculating Accuracy Metrics for ARIMA and ETS	15
6	Reconciliation	19
6.1	Aggregate data	19
6.2	Time series cross validation	19
6.3	Reconcile	21
6.4	Include Age Group	22

1 Preliminary Analysis

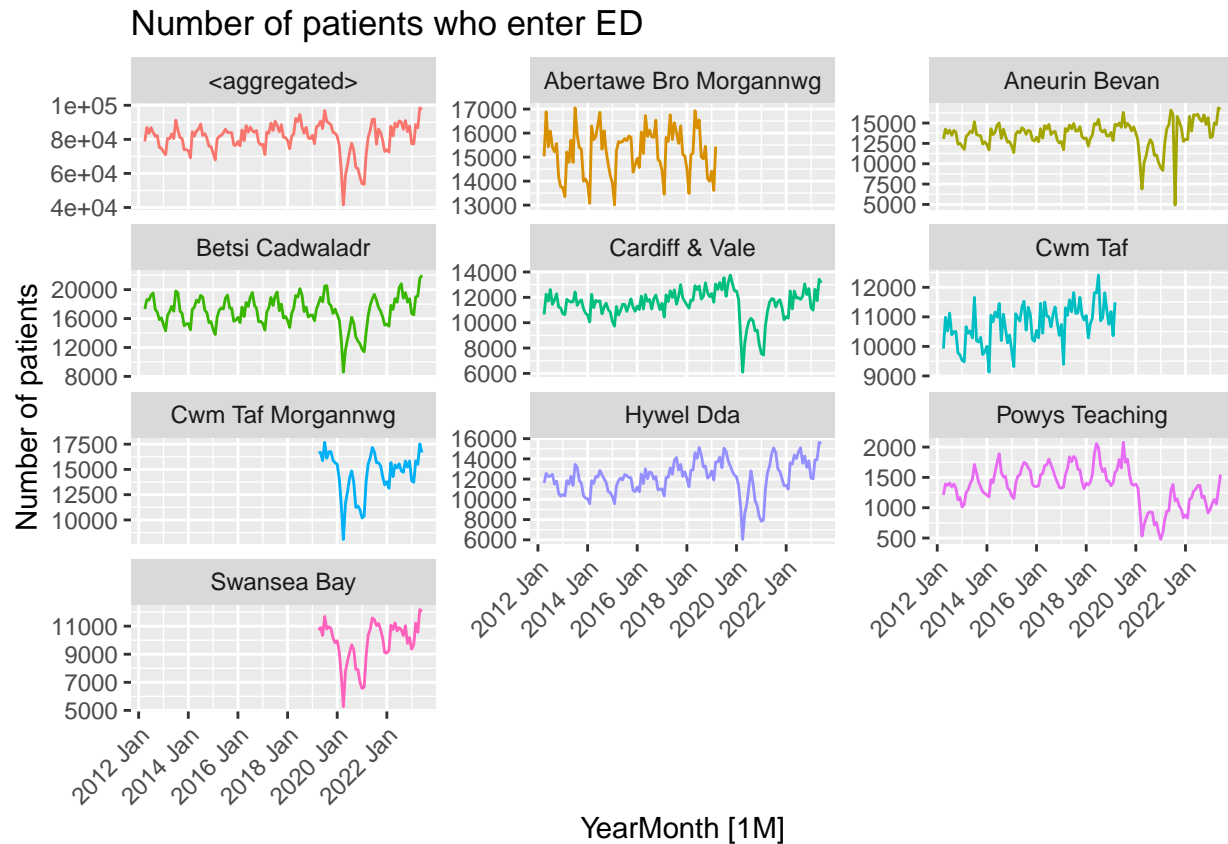
1.1 Data Introduction

The dataset contains 10 variables related to the hospitals and information of patients in Wales, UK. Here are the brief summary of the dataset:

- **Data:** This column represents the number of attendance in each emergency department.
- **YearMonth:** This column represents dates in the year-month format. Additionally, this dataset contains data from 2012 April to 2023 May.
- **Age_Code:** This column provides the age group that the patient is in. There are 17 different age groups. They are “0 to 4”, “5 to 17”, “18 to 24”, “25 to 29”, “30 to 34”, “35 to 39”, “40 to 44”, “45 to 49”, “50 to 54”, “55 to 59”, “60 to 64”, “65 to 69”, “70 to 74”, “75 to 79”, “80 to 84”, “85” and “Unknown”.
- **Sex_ItemName_ENG:** This column provides the information of patient’s gender.
- **Hospital_Code:** This column represents 42 different hospitals in Wales.
- **Hospital_ItemName_ENG:** This columns refers to the name of the 42 different hospitals in Wales.
- **Hospital_Hierarchy:** This column represents the code for the health board that the hospital belongs to.
- **Hospital_AltCode1:** This column provides an alternate code for the hospital.
- **Organisation:** This column represents the health board.
- **Organisation_Code:** A code for the organisation as well as the health board.
- There are three hierarchies in this dataset. On the top level, there is all the hospitals in Wales, while on the second hierarchy, there are 6 different health boards which also shown as the organisations. At the bottom level, there are 42 hospitals in total.

2 Exploratory Data Analysis

2.1 Number of patients entering ED under different hospital hierarchy



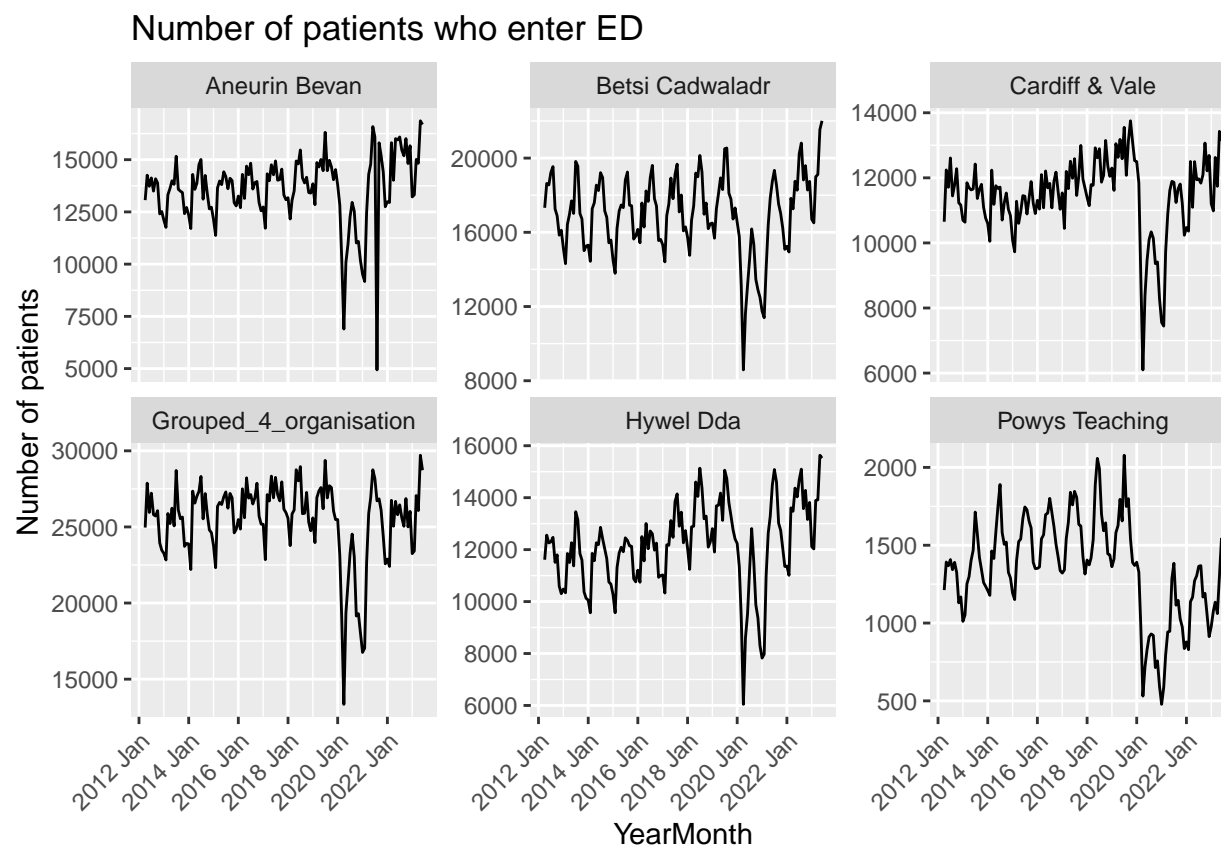
- A couple of Local Health Boards (LHBs) were redefined from the 1st of April 2019 onwards: Cwm Taf (27) → Cwm Taf Morgannwg (30) // Abertawe Bro Morgannwg (26) → Swansea Bay (31). Therefore, if you decide to forecast at LHB resolution, you might want to consider these 4 as a unique one.
- A the Princess of Wales Hospital changed its Local Health Boards
- So we analyse these 4 as one organisation

2.2 Group the changed Local Health Board together

2.2.1 There are 6 Local Health Boards

```
## [1] "Betsi Cadwaladr"      "Hywel Dda"           "Grouped_4_organisation"  
## [4] "Cardiff & Vale"       "Aneurin Bevan"       "Powys Teaching"
```

2.3 Number of patients who enter ED under 6 different local health boards



Findings:

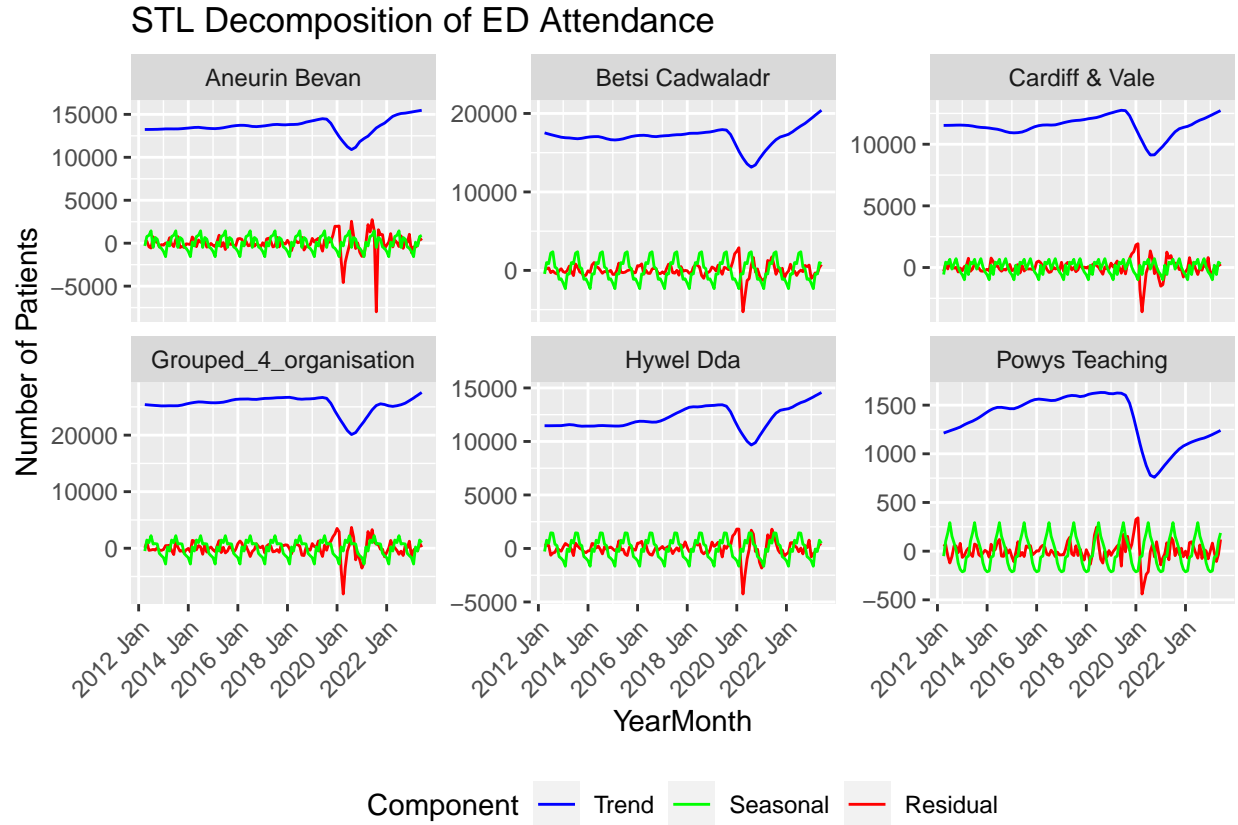
- There is a big decline during the Covid-19 period, and after the Covid-19, except the Powys Teaching, other local health boards have increased the number of attendance back to its previous years.
- There seems to be seasonality in the data for each health board, and I would like to investigate more on top of this.

2.4 Seasonality of number of attendances

To investigate deeper into the potential seasonality in the data, I would like to decompose the time series according to each health board. It allows us to discover the trend, seasonality and the residual components.

2.4.1 Decompose Time Series

2.4.2 Plotting



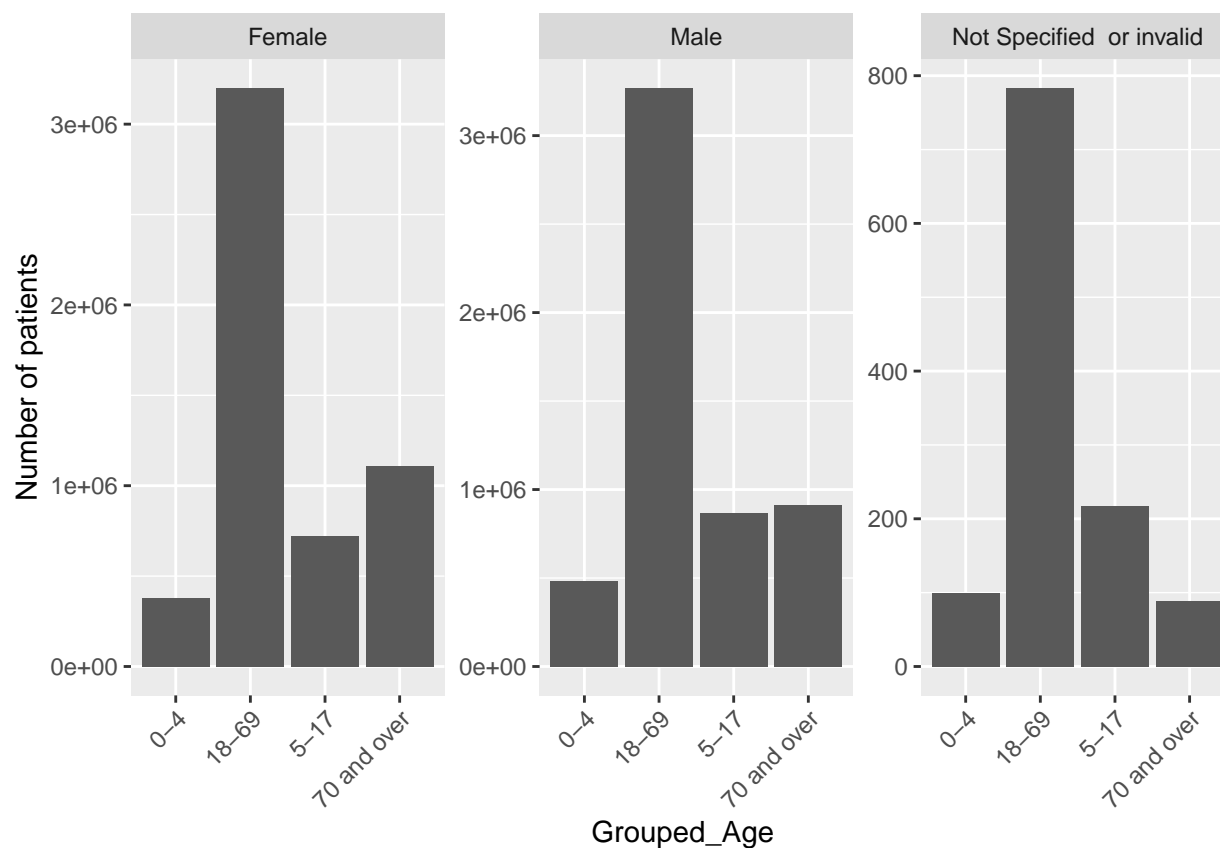
Based on the STL decomposition, it is evident that each health board follows a similar trend. Moreover, it is obvious that there is a significant downturn during the Covid-19 era. As for seasonality, there is a pronounced surge in the number of patient's attendance in the middle of the year (approximately in June or July). This seasonal pattern underlines the recurrent nature of patient admissions.

2.5 Change the Age_Code structure into different groups

```
## [1] "0 to 4" "18 to 24" "25 to 29" "30 to 34" "35 to 39" "40 to 44"  
## [7] "45 to 49" "5 to 17" "50 to 54" "55 to 59" "60 to 64" "65 to 69"  
## [13] "70 to 74" "75 to 79" "80 to 84" "85" "Unknown"
```

2.5.1 Age group: “0-4”, “5-17”, “18-69”, “70+”

2.6 Plot Number of Patients in different age groups



Findings:

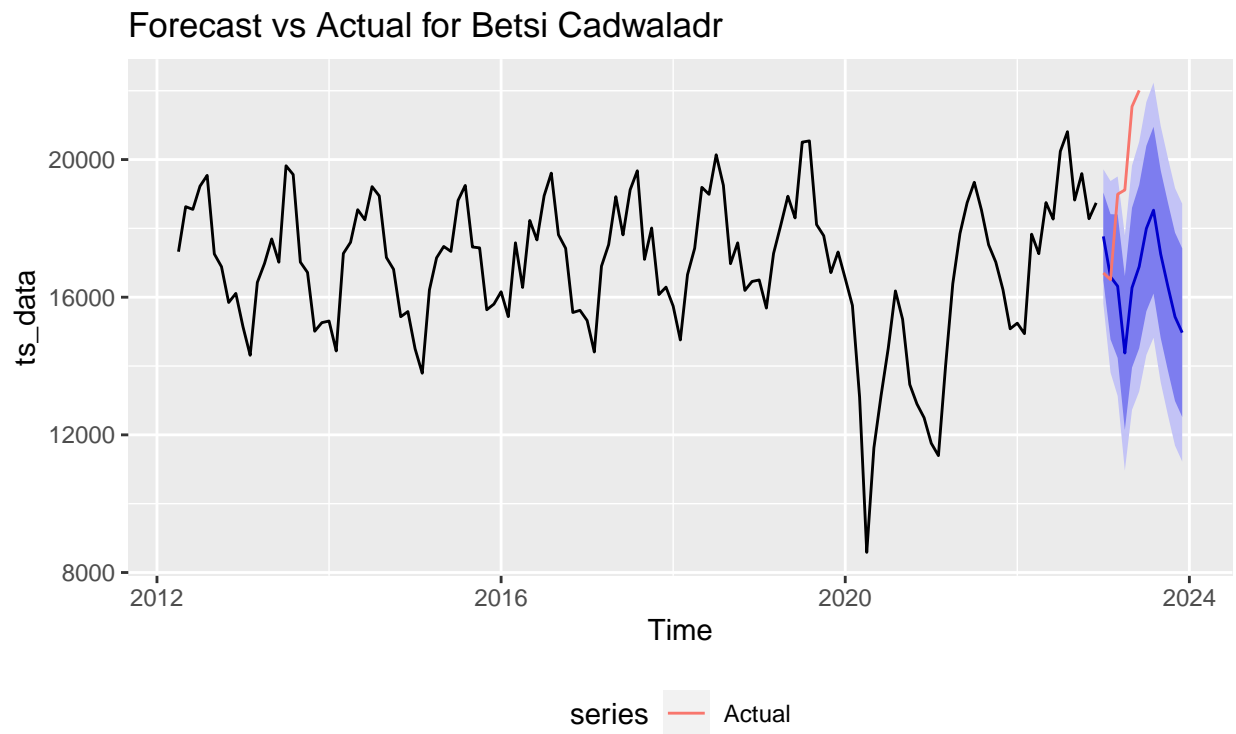
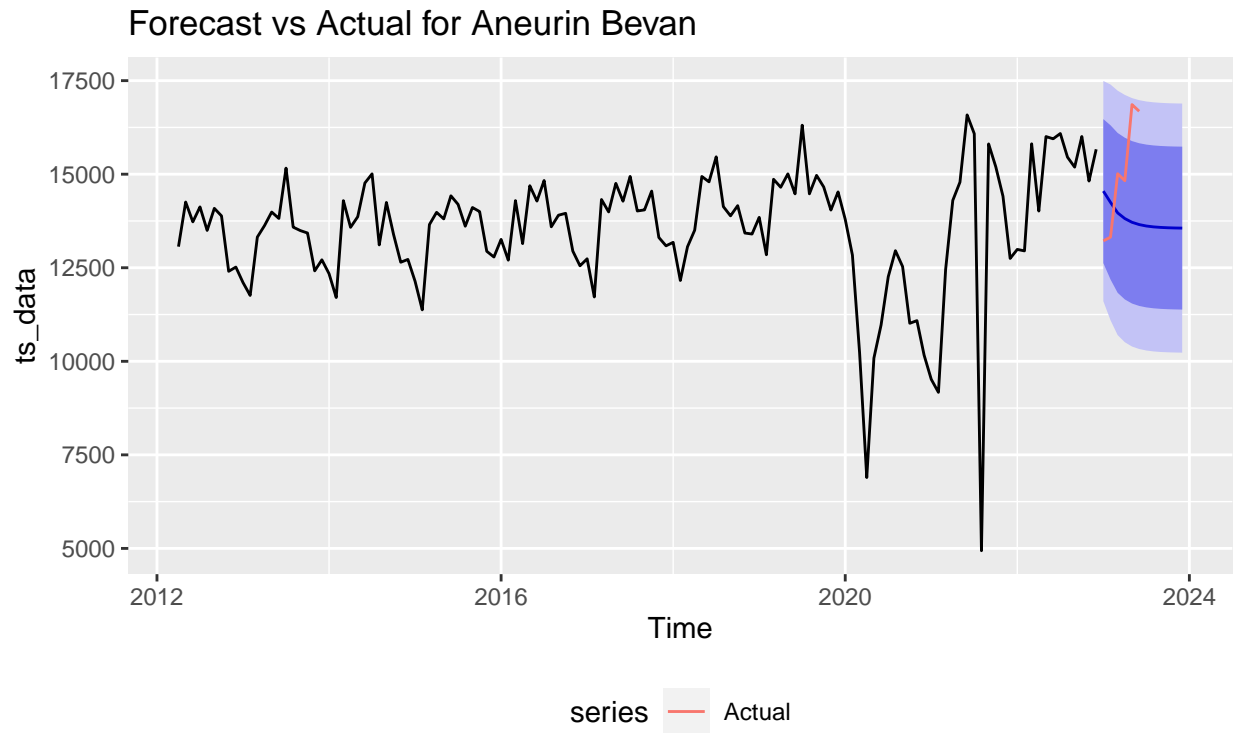
- The observation that the age group 18-69 has the most amount of patient attendance is expected, as it is the biggest group among all. However, it is noteworthy that the second biggest group are from the oldest age bracket, aligning with the general understanding of the health care need for the elders.

3 Forecast with ARIMA

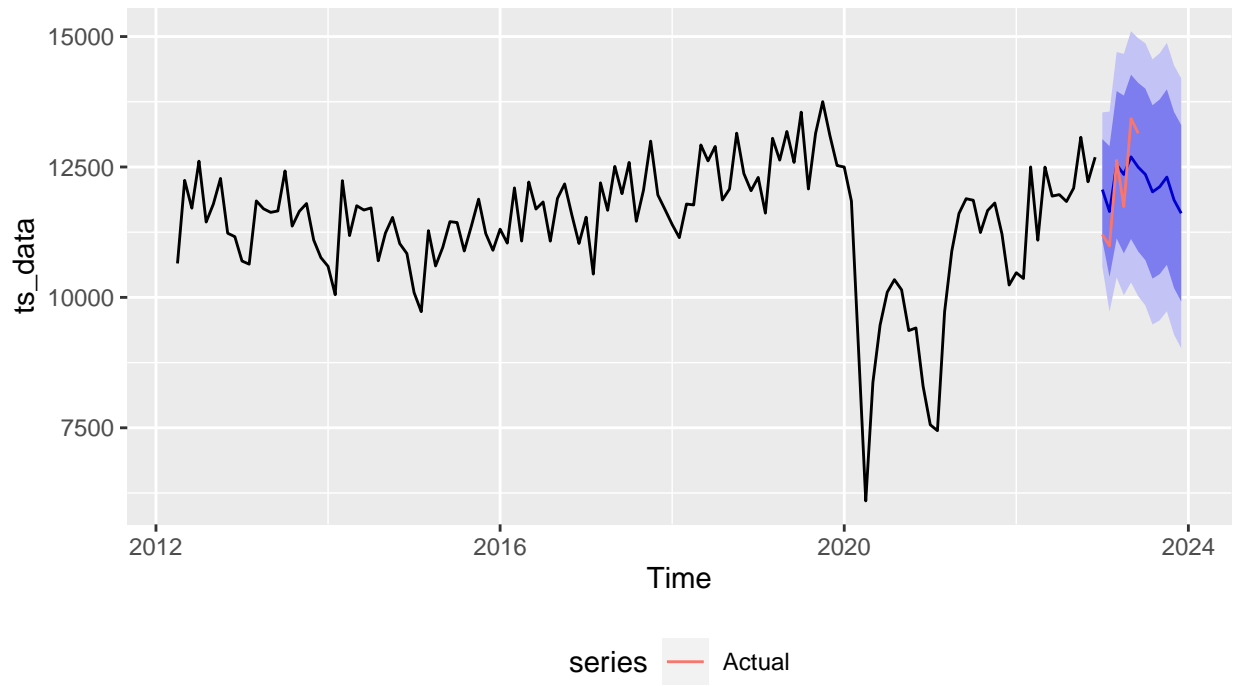
3.0.1 Define the forecast horizon and validation period

3.0.2 Splitting the data into training and testing sets

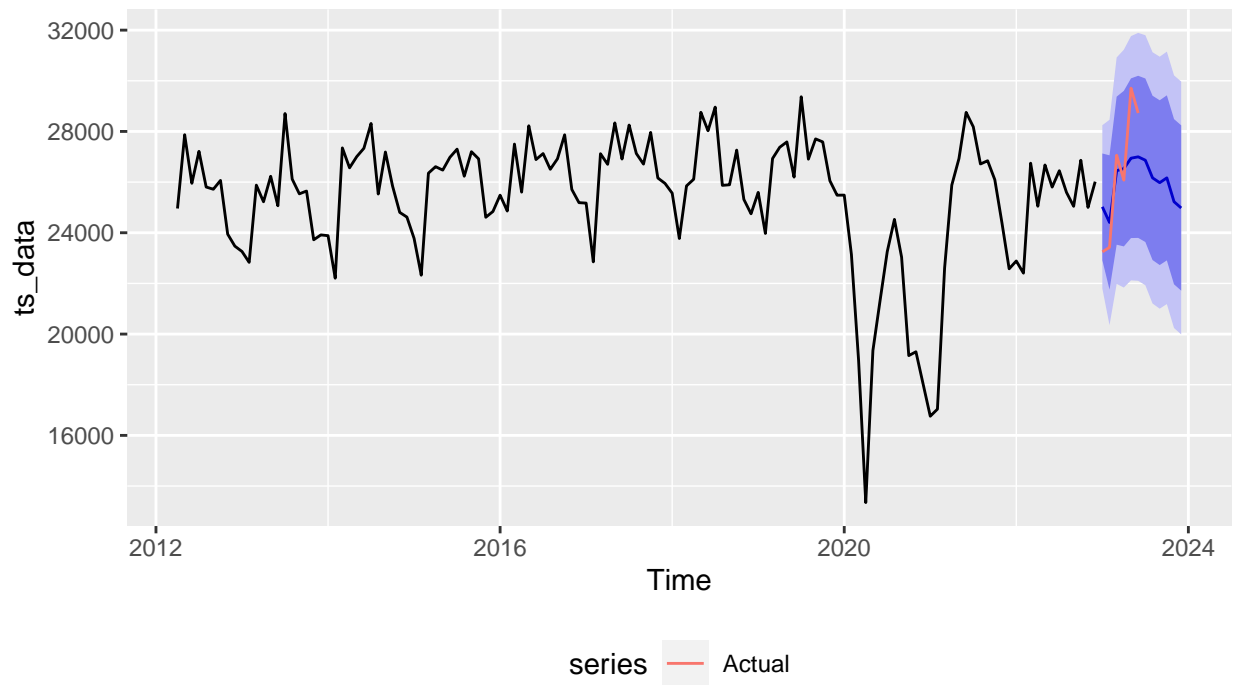
3.1 Plotting using ARIMA



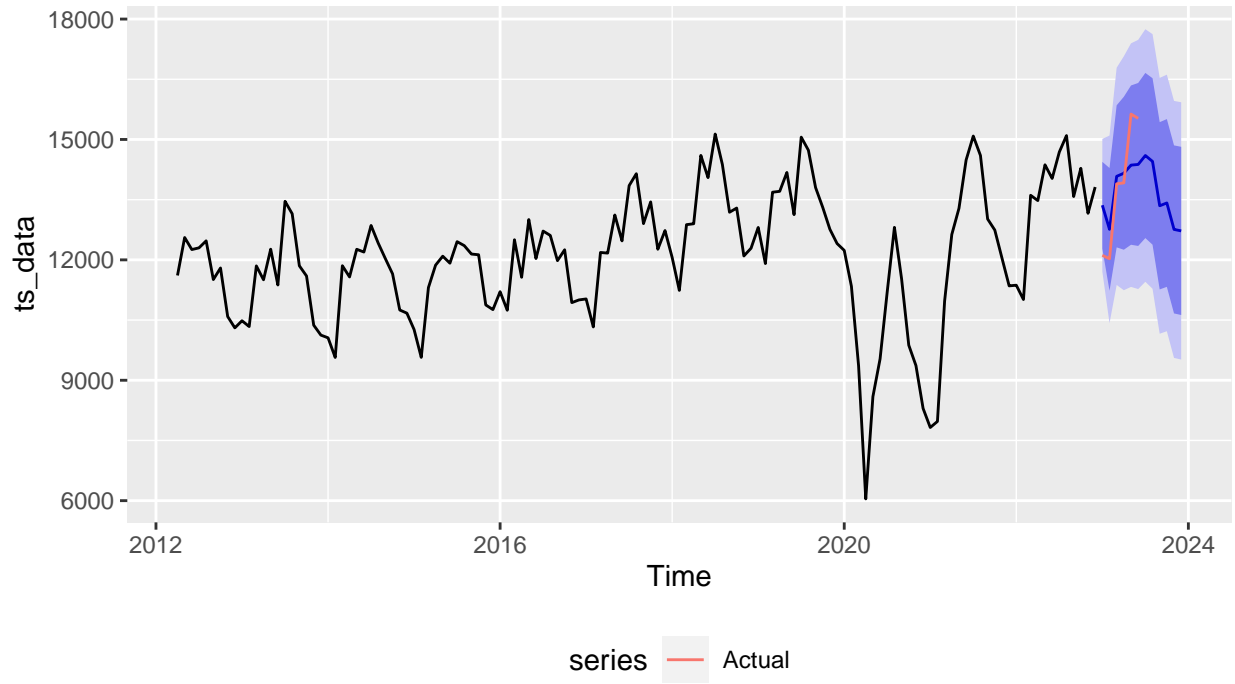
Forecast vs Actual for Cardiff & Vale



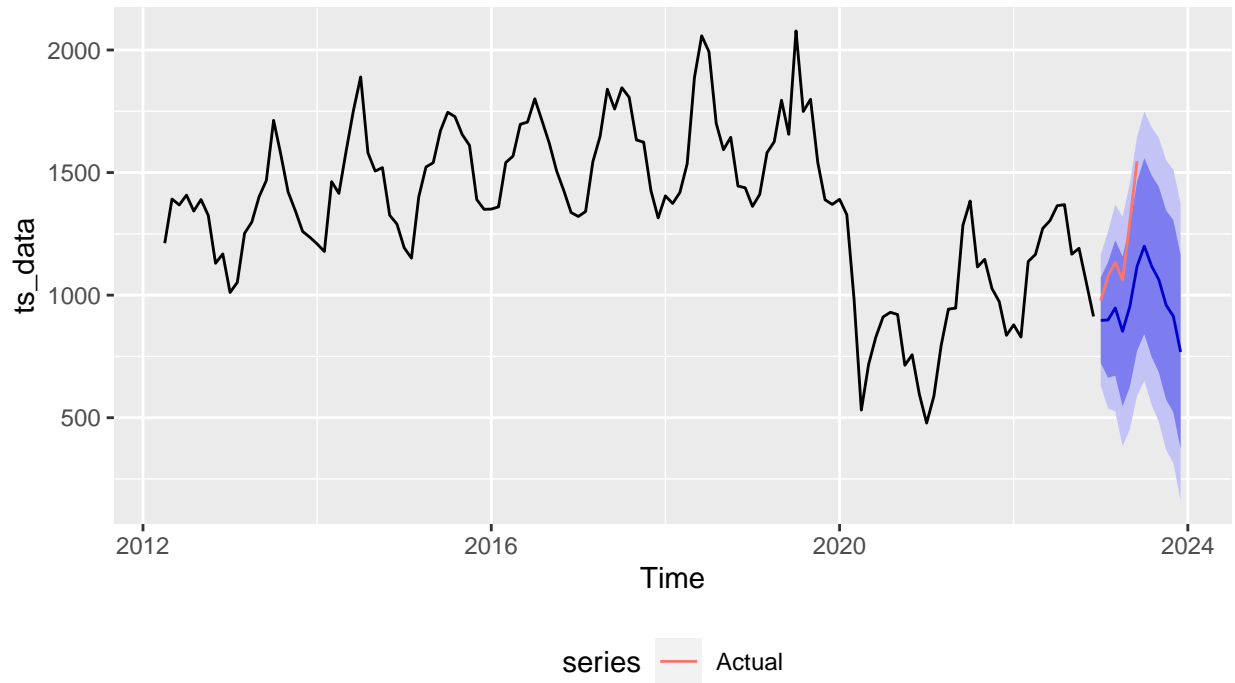
Forecast vs Actual for Grouped_4_organisation



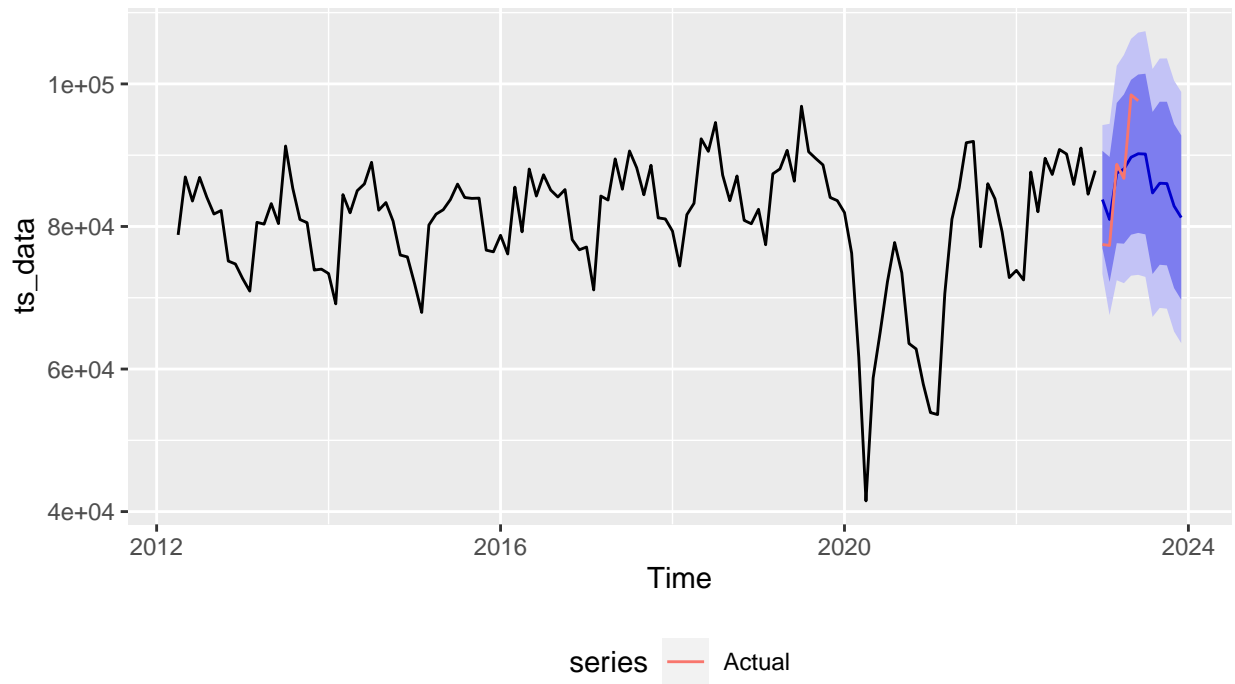
Forecast vs Actual for Hywel Dda



Forecast vs Actual for Powys Teaching



Forecast vs Actual for Total

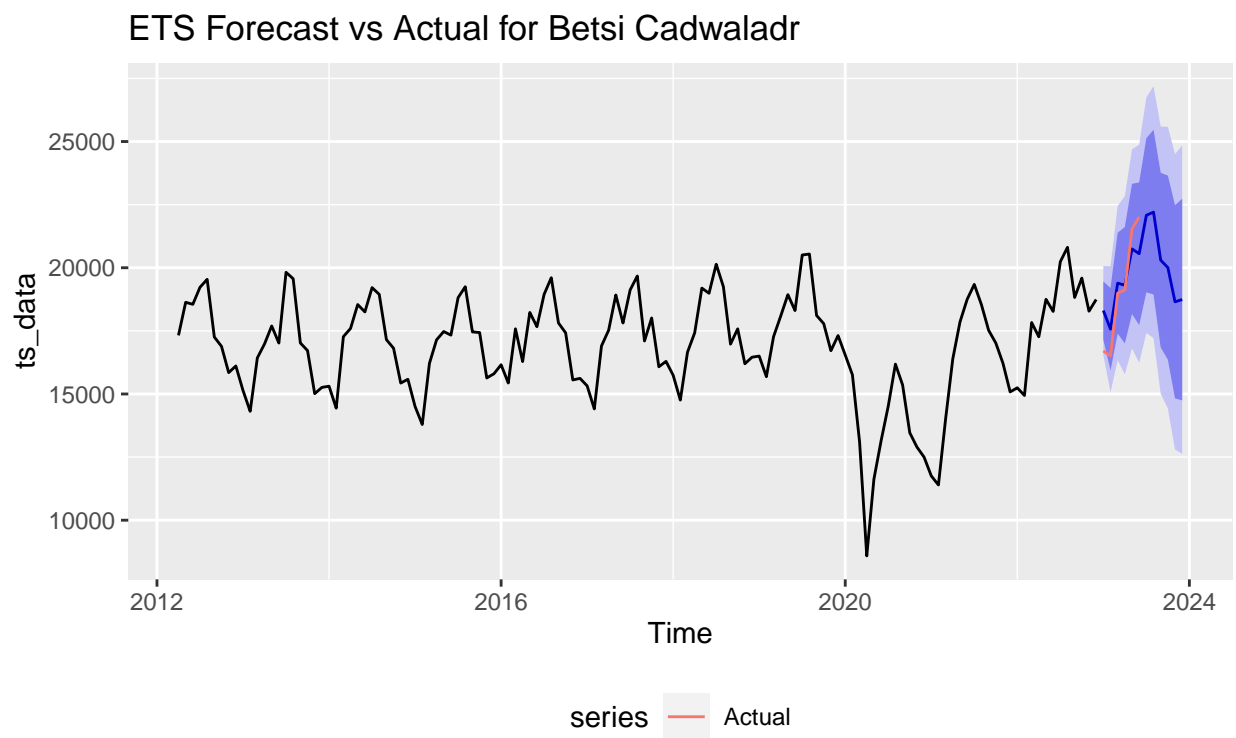
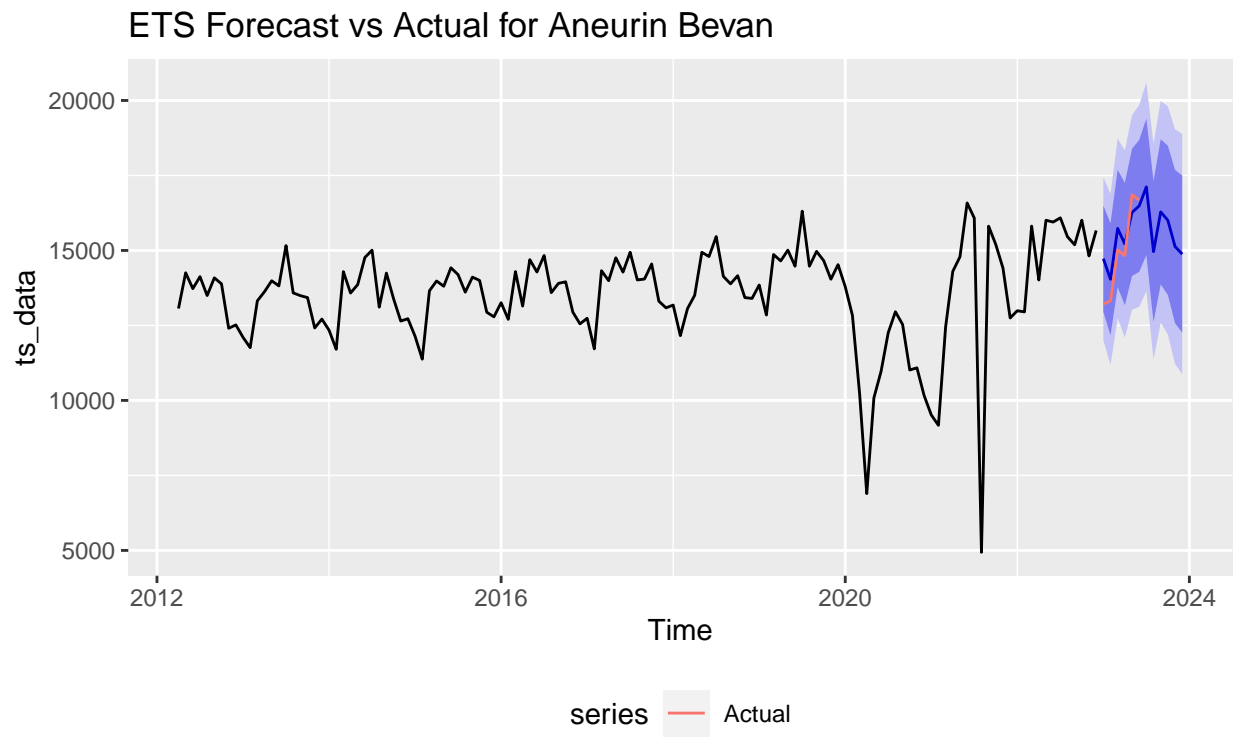


Due to the impact of Covid-19, the trend of forecast in the future appears to be lower than expected. However, the ARIMA model seems to suggest the consistent direction as the actual data for Cardiff & Vale, Grouped_4_Organisation, Hywel Dda, Powys Teaching and Total (All-Wales). Noticeably, they all have a downturn in the latter half of 2023.

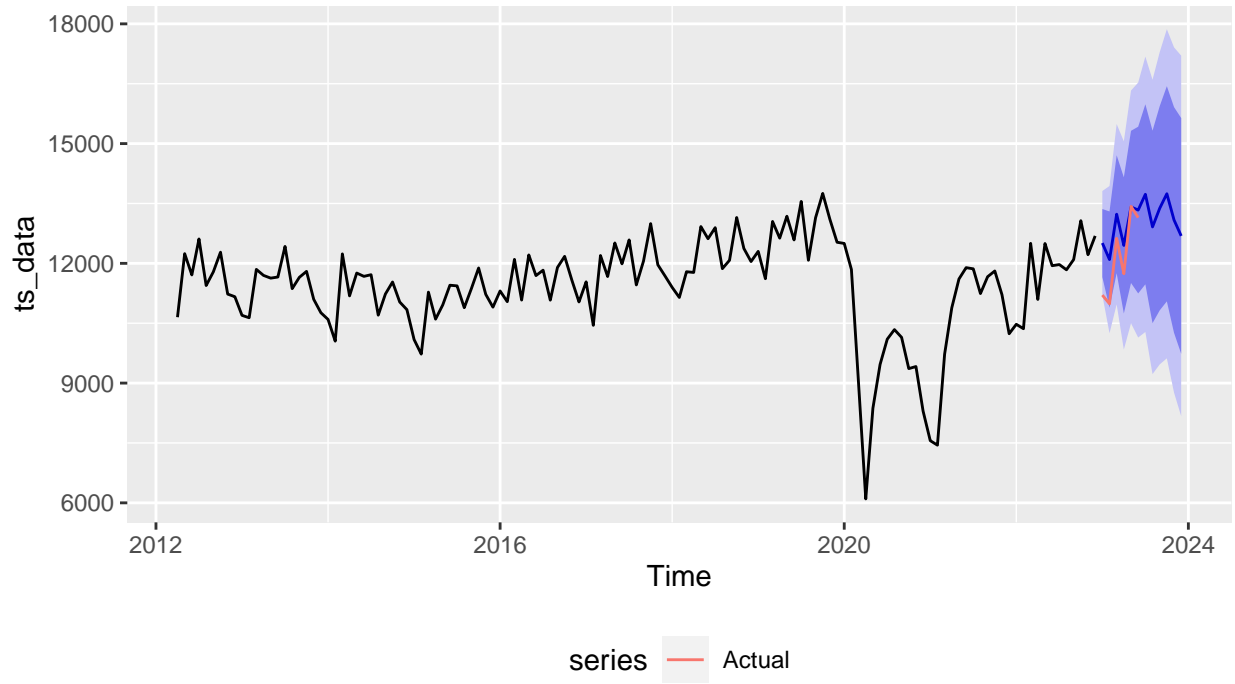
However, the forecast for Aneurin Bevan and Betsi Cadwaladr are not that accurate according to the divergence between the predicted value and the actual data.

4 Forecasting with ETS

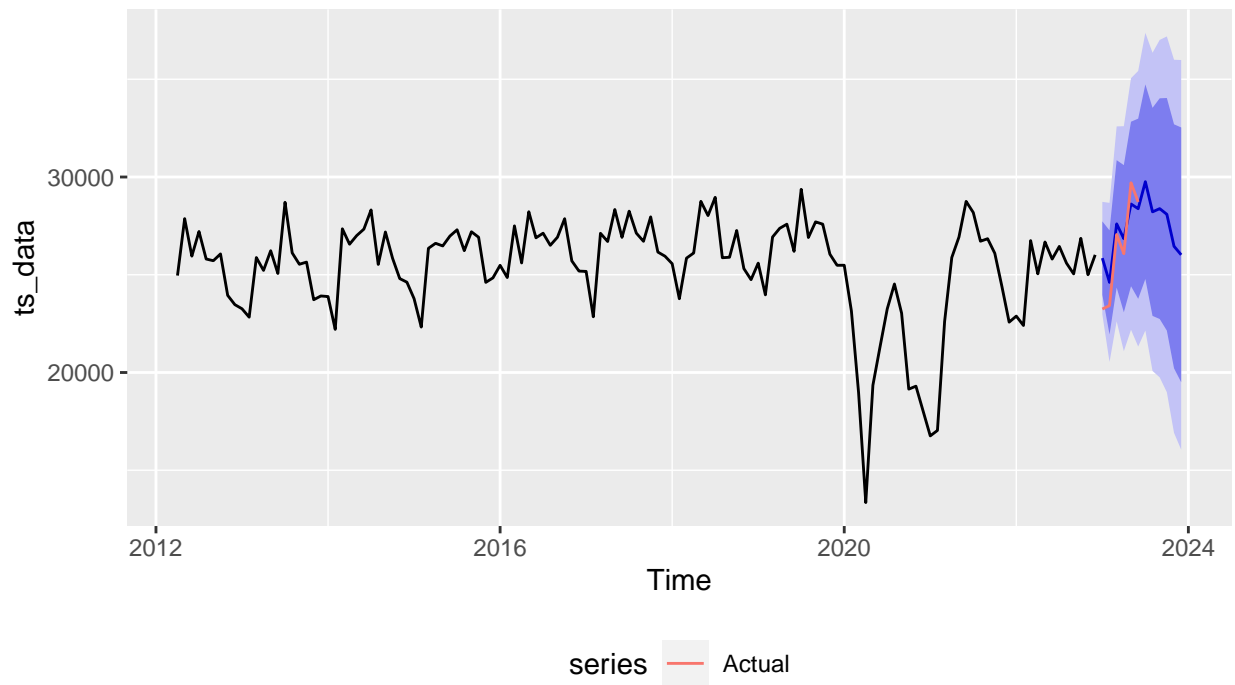
4.1 Plotting with ets



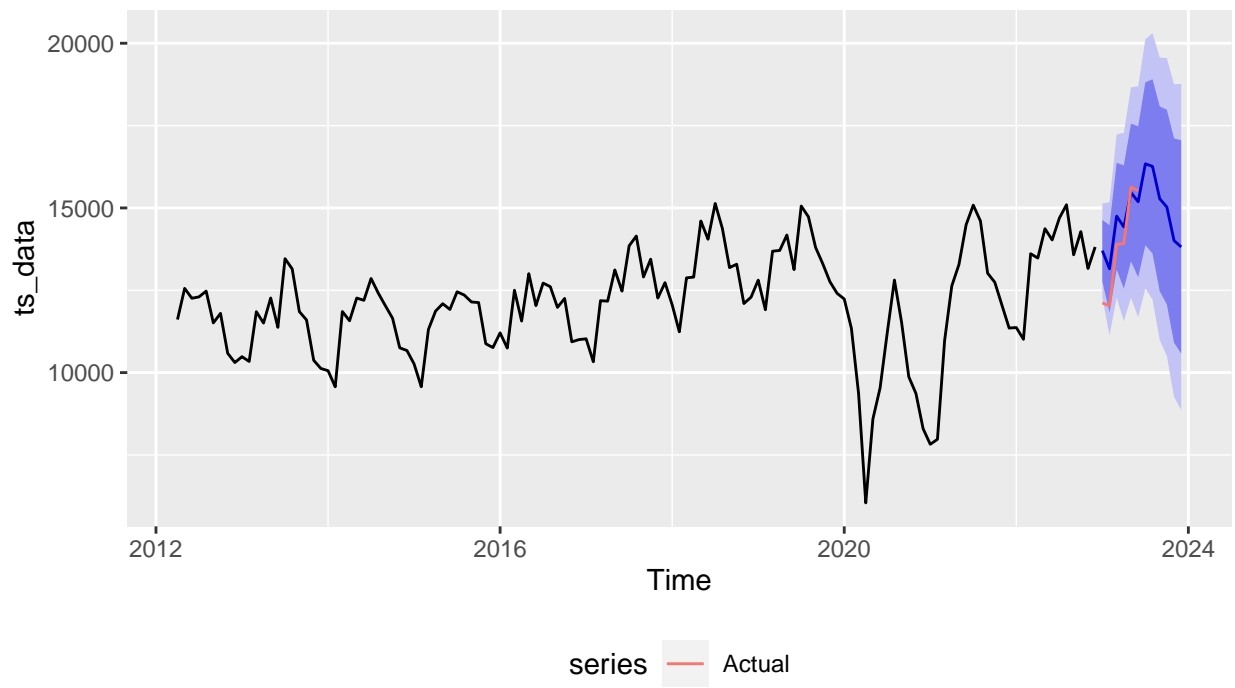
ETS Forecast vs Actual for Cardiff & Vale



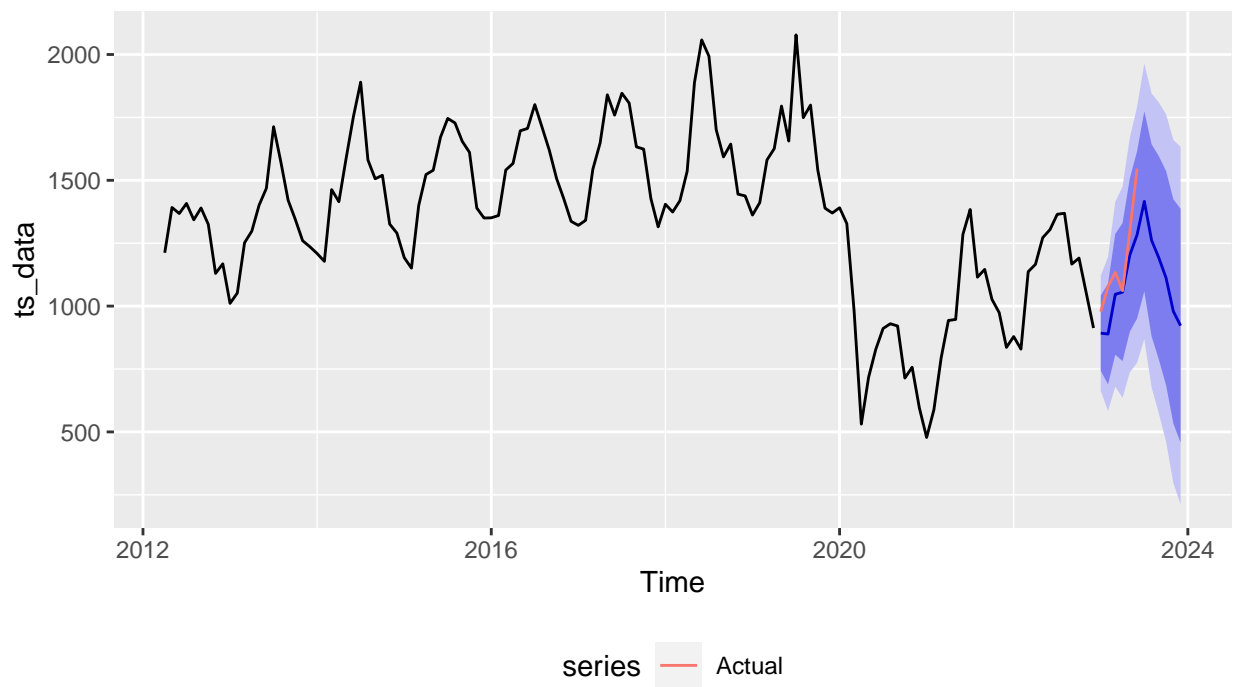
ETS Forecast vs Actual for Grouped_4_organisation

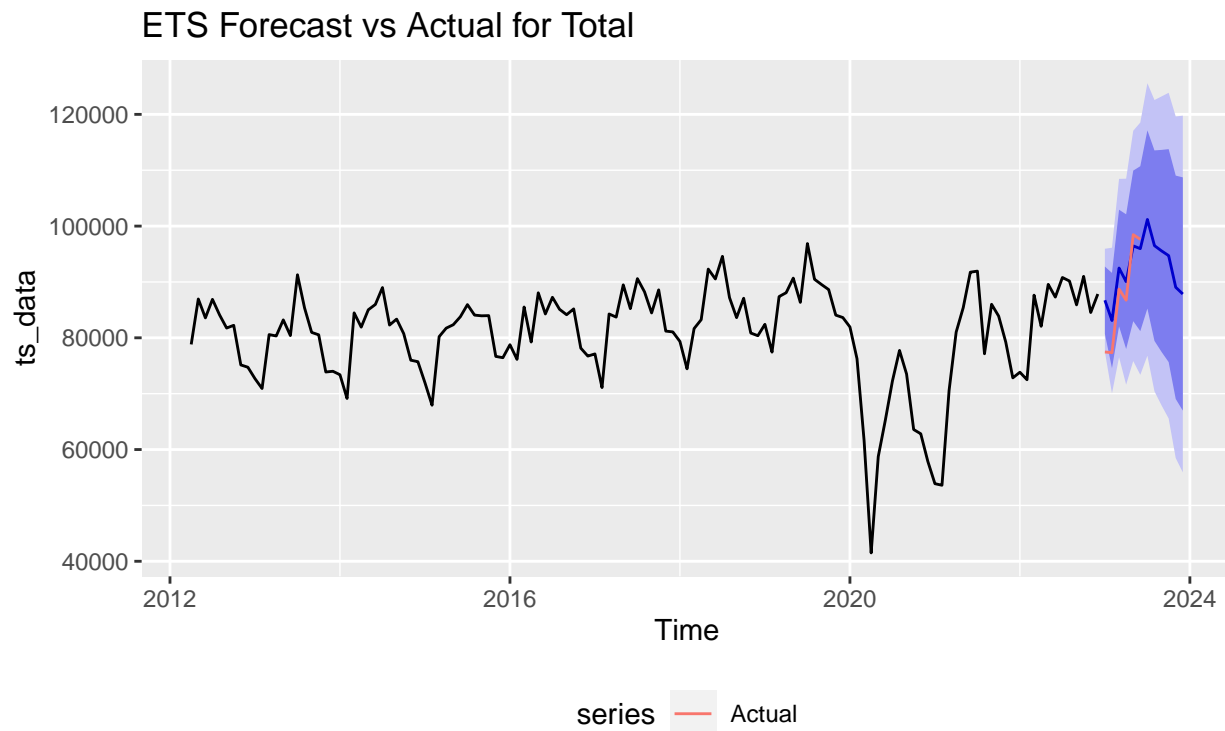


ETS Forecast vs Actual for Hywel Dda



ETS Forecast vs Actual for Powys Teaching





The prediction from the ETS model aligns closely with the actual data upon the visual inspection. Therefore, It might suggest that a higher accuracy for the ETS model compared to the ARIMA model.

Additionally, all the predictions have a downturn of trend in the latter half of 2023. According to the historical data, this might be due to that the peak appears in the middle of the year followed by a trough at the end and the start of the year.

5 Accuracy assessment for ARIMA and ETS

5.0.1 Lists to store forecasts and error metrics

5.1 Forecasting using ARIMA and ETS

5.2 Calculating Accuracy Metrics for ARIMA and ETS

5.2.1 Displaying Accuracy Metrics for each Column

```
##
## Accuracy metrics for: Aneurin Bevan
## -----
## ARIMA:
## $MAE
## [1] 1748.634
##
## $RMSE
## [1] 1992.543
##
## $MAPE
```

```

## [1] 0.1127431
##
## ETS:
## $MAE
## [1] 688.4172
##
## $RMSE
## [1] 801.6437
##
## $MAPE
## [1] 0.04825847
##
## -----
##
## Accuracy metrics for: Betsi Cadwaladr
## -----
## ARIMA:
## $MAE
## [1] 3157.257
##
## $RMSE
## [1] 3756.403
##
## $MAPE
## [1] 0.1557319
##
## ETS:
## $MAE
## [1] 915.7354
##
## $RMSE
## [1] 1050.122
##
## $MAPE
## [1] 0.04895248
##
## -----
##
## Accuracy metrics for: Cardiff & Vale
## -----
## ARIMA:
## $MAE
## [1] 600.4287
##
## $RMSE
## [1] 648.3758
##
## $MAPE
## [1] 0.04999195
##
## ETS:
## $MAE

```



```

## [1] 654.8113
##
## $RMSE
## [1] 799.0302
##
## $MAPE
## [1] 0.05682325
##
##
## -----
##
## Accuracy metrics for: Grouped_4_organisation
## -----
## ARIMA:
## $MAE
## [1] 1386.277
##
## $RMSE
## [1] 1598.559
##
## $MAPE
## [1] 0.0519391
##
## ETS:
## $MAE
## [1] 1089.509
##
## $RMSE
## [1] 1315.071
##
## $MAPE
## [1] 0.0434871
##
##
## -----
##
## Accuracy metrics for: Hywel Dda
## -----
## ARIMA:
## $MAE
## [1] 804.5973
##
## $RMSE
## [1] 923.8532
##
## $MAPE
## [1] 0.05832447
##
## ETS:
## $MAE
## [1] 763.0849
##
## $RMSE
## [1] 905.503

```

```

##
## $MAPE
## [1] 0.05915982
##
##
## -----
##
## Accuracy metrics for: Powys Teaching
## -----
## ARIMA:
## $MAE
## [1] 238.2648
##
## $RMSE
## [1] 264.6673
##
## $MAPE
## [1] 0.1922351
##
## ETS:
## $MAE
## [1] 120.9989
##
## $RMSE
## [1] 146.7962
##
## $MAPE
## [1] 0.09808414
##
##
## -----
##
## Accuracy metrics for: Total
## -----
## ARIMA:
## $MAE
## [1] 4781.817
##
## $RMSE
## [1] 5601.875
##
## $MAPE
## [1] 0.05378115
##
## ETS:
## $MAE
## [1] 4297.46
##
## $RMSE
## [1] 5014.551
##
## $MAPE
## [1] 0.05203606
##

```

```
##
## -----
```

- Most of the error metrics for ETS model are lower than the ARIMA model, except for Cardiff & Vale.
- The ARIMA model is based on the autoregressive and moving average components of the time series, in line with the differencing method in order to make the data stationary. On the other hand, the ETS model focuses on the error, trend and seasonality components of the time series.
- According to the interest based on the NHS team, they would like to understand what is the best error metric to assess the quality of the forecast for the non-zero count data. Therefore, I would like to provide some insights for the error metrics:

1. Root Mean Squared Error (RMSE)

- **Description:** Measures the square root of the average squared differences between forecast and actual values.
- **Pros:** It gives more weight to large errors.
- **Cons:** It can be influenced significantly by outliers. RMSE might be inflated, if the count data is prone to significant spikes or declines, which should not be the problem for this data except for the Covid-19 era.

2. Mean Absolute Percentage Error (MAPE)

- **Description:** Measures the average of the absolute percentage errors.
- **Pros:** It is a relative metric, and is scale-independent and easy to interpret.
- **Cons:** MAPE can overemphasize relative errors on small counts.

3. Mean Absolute Error (MAE)

- **Description:** Measures the average of the absolute differences between forecast and actual values.
- **Pros:** It's less sensitive to outliers than RMSE. It provides a straightforward average error size.
- **Cons:** It does not emphasize large errors.

Recommendation:

- If you want to penalize larger errors more severely, **RMSE** might be the preferred choice if the larger errors need to be considered, such as the drops in the Covid-19 era.
- For a more balanced view of the average error, **MAE** can be considered.
- **MAPE** is useful and insightful if a relative metric is needed to as it gives error as a percentage of the actual counts.

However, to get a comprehensive overview of the forecast accuracy, it is beneficial to examine on multiple error metrics. Additionally, the choice of the metrics also depend on the specific objectives.

6 Reconciliation

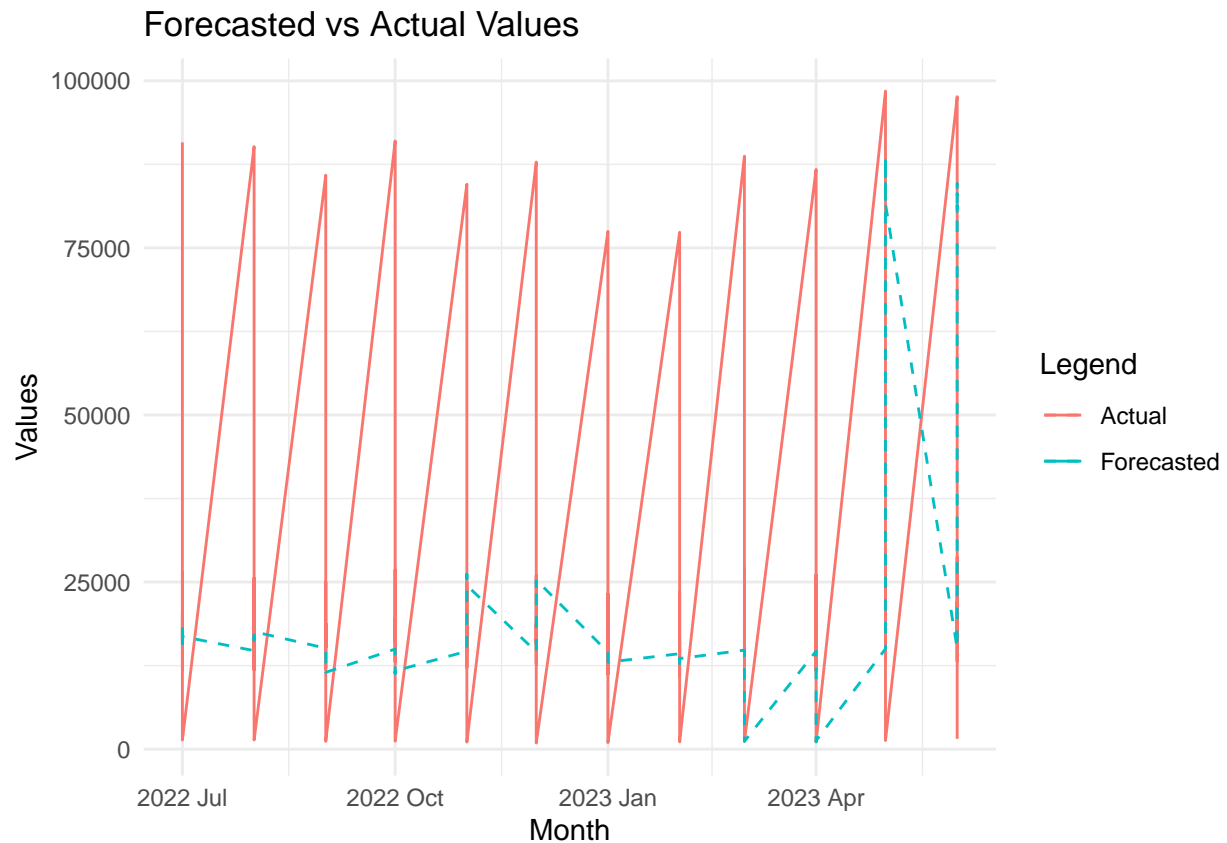
6.1 Aggregate data

6.2 Time series cross validation

```
## Length of forecast: 84
```

```
## Length of test data: 84
```

```
## [1] 0
```



```
##
## Welch Two Sample t-test
##
## data: comparison_data$Actual and comparison_data$Forecasted
## t = 0.31272, df = 165.49, p-value = 0.7549
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -6712.277 9238.730
## sample estimates:
## mean of x mean of y
## 25155.69 23892.46
```

The p-value is very high, suggesting that we do not reject the null hypothesis. It indicates the equal means between the actual and forecasted values, and they are not significantly different from each other.

The confidence interval for the difference between the means spans from -6712.277 to 9238.730, which includes 0, further indicating that the difference between the means of the two groups is not statistically significant.

Additionally, Mean of Actual Values and Mean of Forecasted Values shows that, they are very close to each other. And that probably gives the reason why the result of the accuracy test is 0.

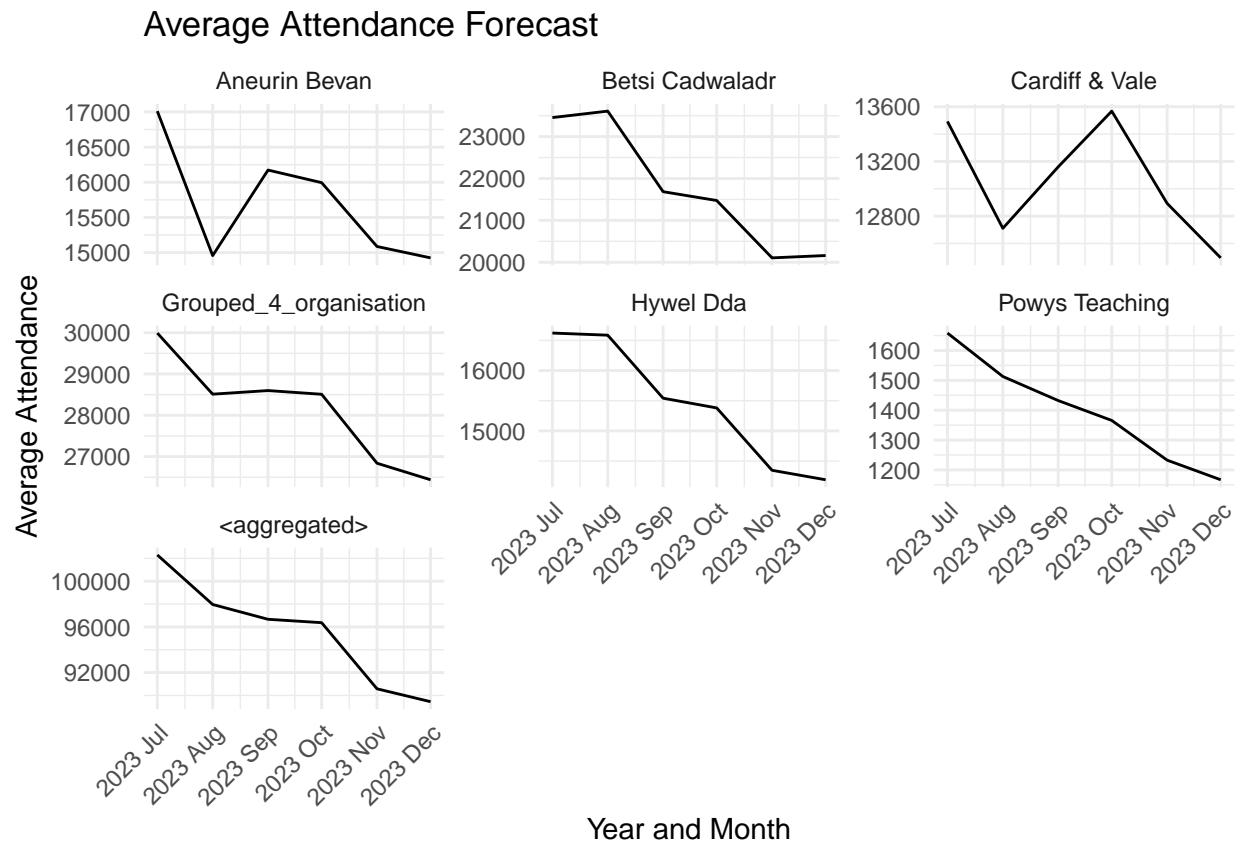
```
## Mean Absolute Percentage Error: 355.5203 %
```

Mean Squared Error: 1512259654

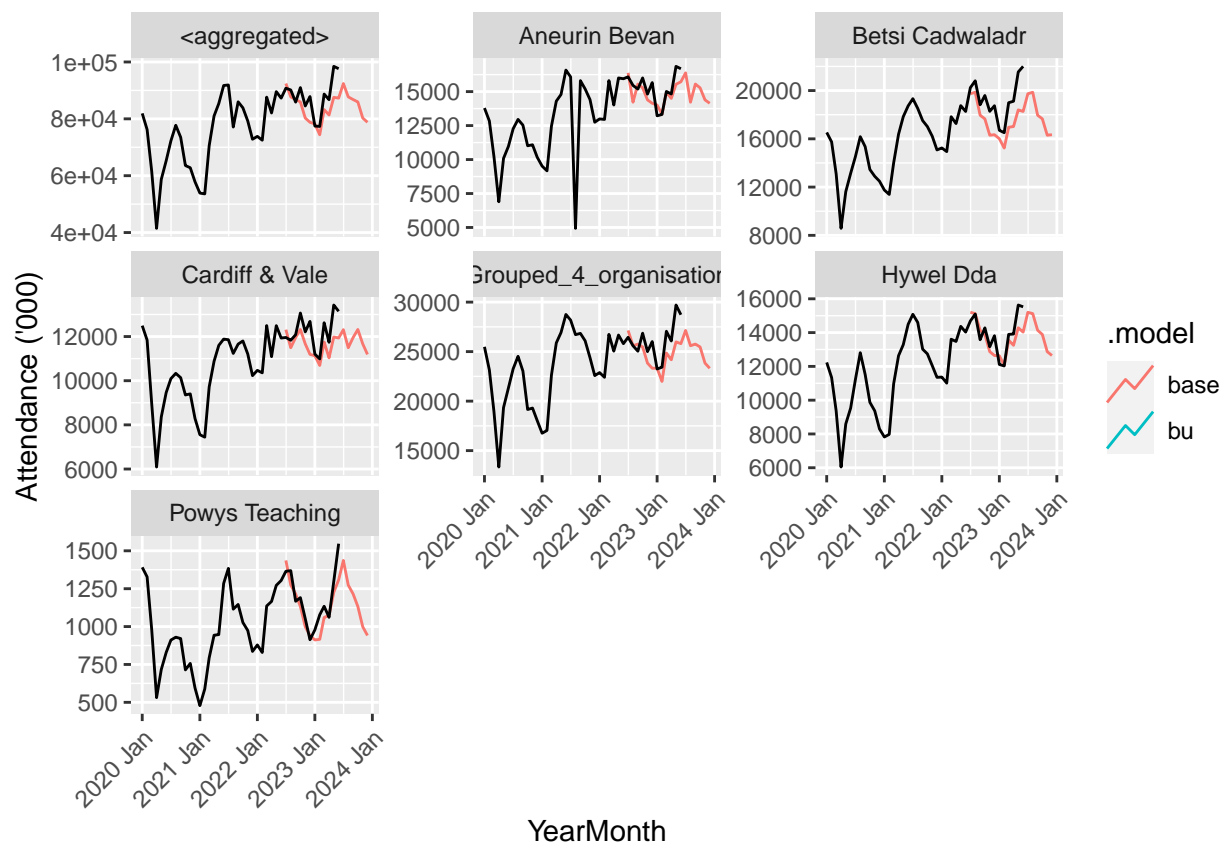
- Base forecasting: applied when forecasting simple time series.
- Reconciliation forecasting (Hierarchical forecasting): it includes forecasting on different levels of aggregation, and it should be applied when dealing with hierarchical or grouped time series.

6.3 Reconcile

6.3.1 1. Fitting models and Forecasting



6.4 Include Age Group



```
## # A fable: 8,640 x 7 [1M]
## # Key:      Aggregated_Organisation, Grouped_Age, Hospital_ItemName_ENG, .model
## #      [480]
##   Aggregated_Organisation Grouped_Age Hospital_ItemName_ENG .model YearMonth
##   <chr*>                  <chr*>      <chr*>                  <chr>    <mt>
## 1 Aneurin Bevan          0-4         Nevill Hall Hospital    base     2022 Jul
## 2 Aneurin Bevan          0-4         Nevill Hall Hospital    base     2022 Aug
## 3 Aneurin Bevan          0-4         Nevill Hall Hospital    base     2022 Sep
## 4 Aneurin Bevan          0-4         Nevill Hall Hospital    base     2022 Oct
## 5 Aneurin Bevan          0-4         Nevill Hall Hospital    base     2022 Nov
## 6 Aneurin Bevan          0-4         Nevill Hall Hospital    base     2022 Dec
## 7 Aneurin Bevan          0-4         Nevill Hall Hospital    base     2023 Jan
## 8 Aneurin Bevan          0-4         Nevill Hall Hospital    base     2023 Feb
## 9 Aneurin Bevan          0-4         Nevill Hall Hospital    base     2023 Mar
## 10 Aneurin Bevan         0-4         Nevill Hall Hospital    base     2023 Apr
## # i 8,630 more rows
## # i 2 more variables: attendance <dist>, .mean <dbl>

## [1] "<aggregated>" "0-4"          "18-69"        "5-17"         "70 and over"
```