

# Hierarchical\_2

Janice Hsu

2023-09-15

## Contents

<b>1</b>	<b>Preliminary Analysis</b>	<b>2</b>
1.1	Data Introduction . . . . .	2
<b>2</b>	<b>Exploratory Data Analysis</b>	<b>3</b>
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
<b>3</b>	<b>Forecast</b>	<b>6</b>
<b>4</b>	<b>Forecast with ARIMA</b>	<b>10</b>
4.1	Plotting using ARIMA . . . . .	10
<b>5</b>	<b>Forecasting with ETS</b>	<b>17</b>
5.1	Plotting with ets . . . . .	17
<b>6</b>	<b>Accuracy assessment for ARIMA and ETS</b>	<b>22</b>

# 1 Preliminary Analysis

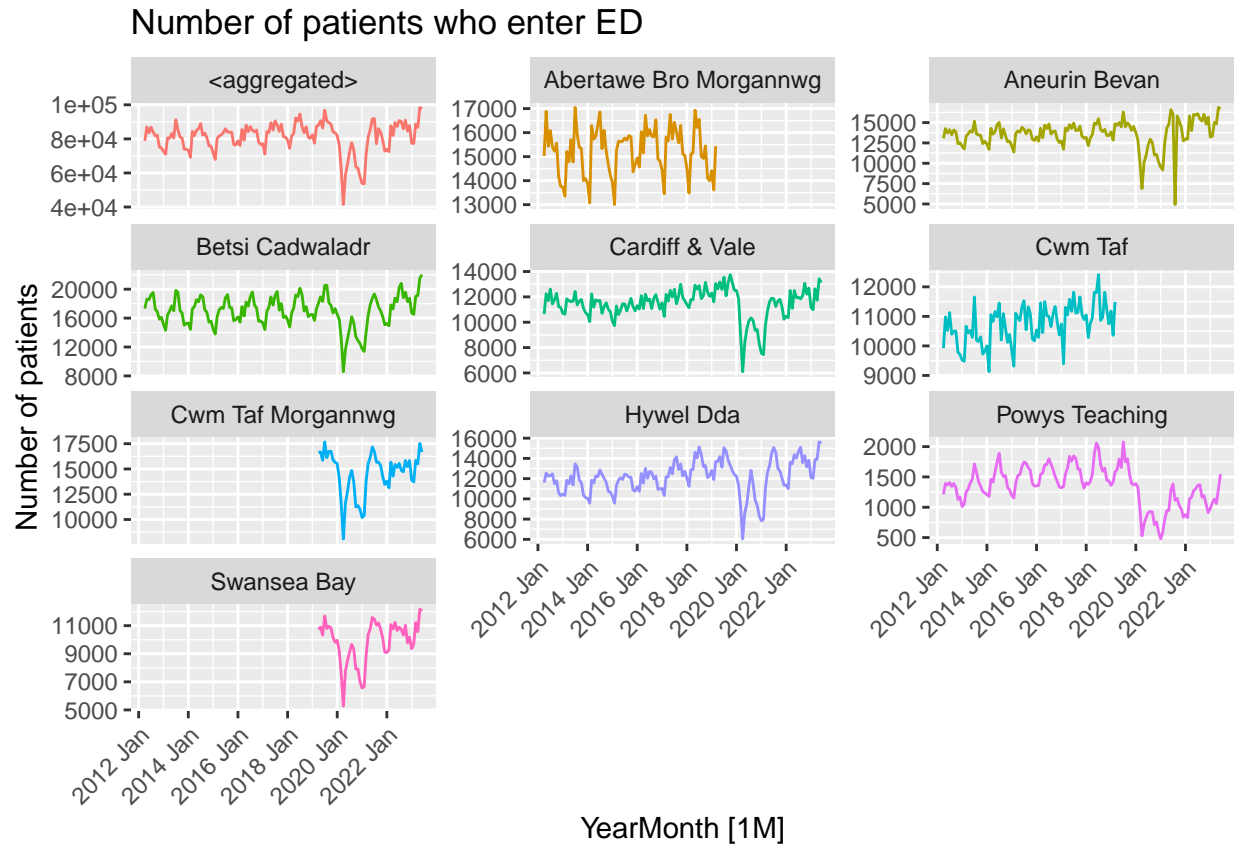
## 1.1 Data Introduction

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



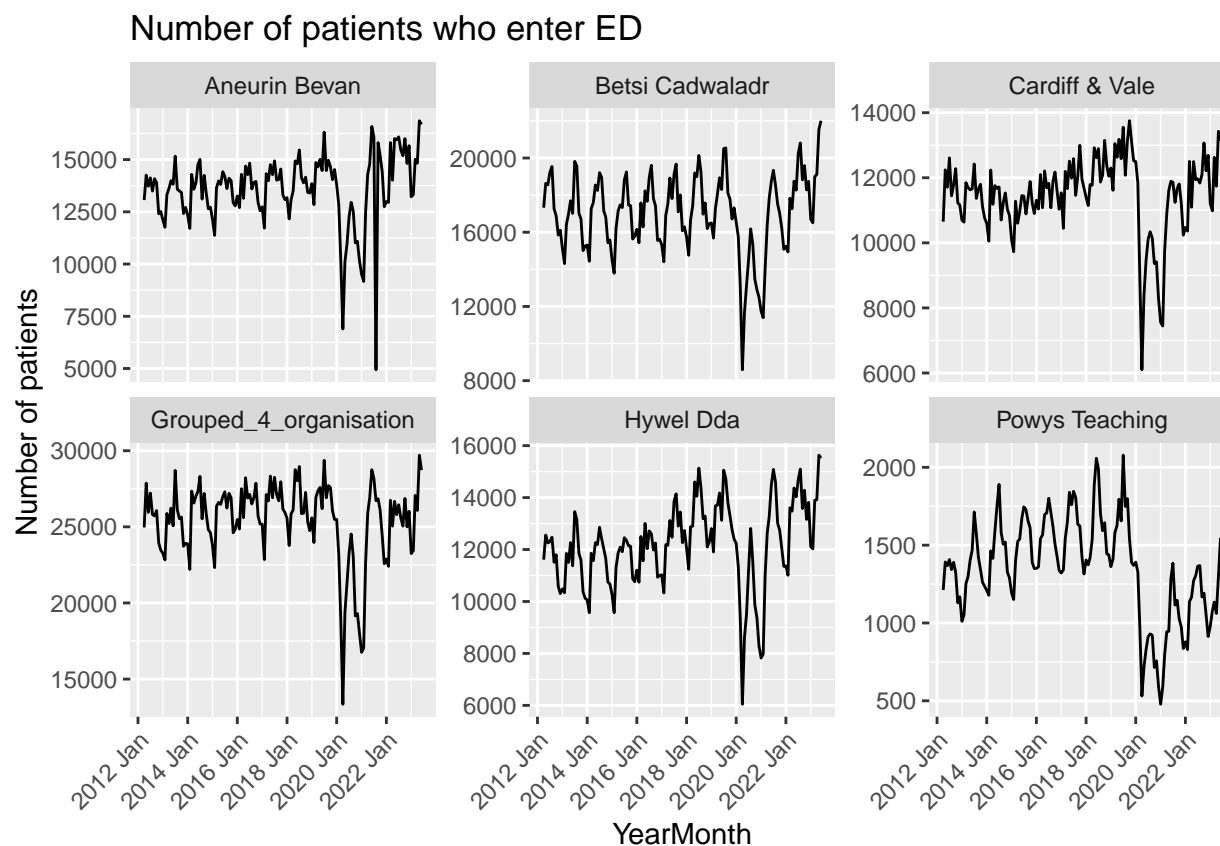
- A couple of Local Health Boards (LHBs) were redefined from the 1st of April 2019 onwards: Cwm Taf (27) → Cwm Taf Morgannwg (30) // Abertawe Bro Morgannwg (26) → Swansea Bay (31). Therefore, if you decide to forecast at LHB resolution, you might want to consider these 4 as a unique one.
- A the Princess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

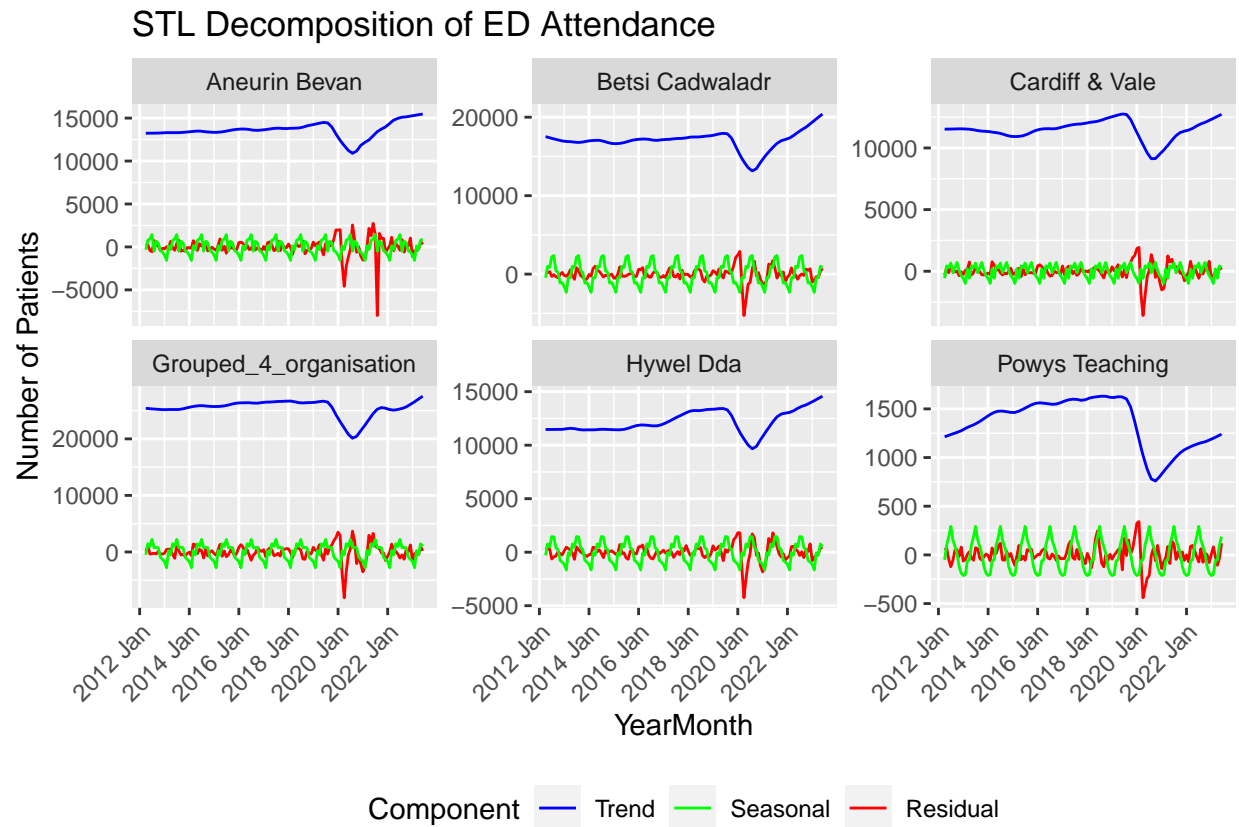
- There is a big decline during the Covid-19 period, and after the Covid-19, except the Powys Teaching, other local health boards have increased the number of attendance back to its previous years.
- There seems to be seasonality in the data for each health board, and I would like to investigate more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



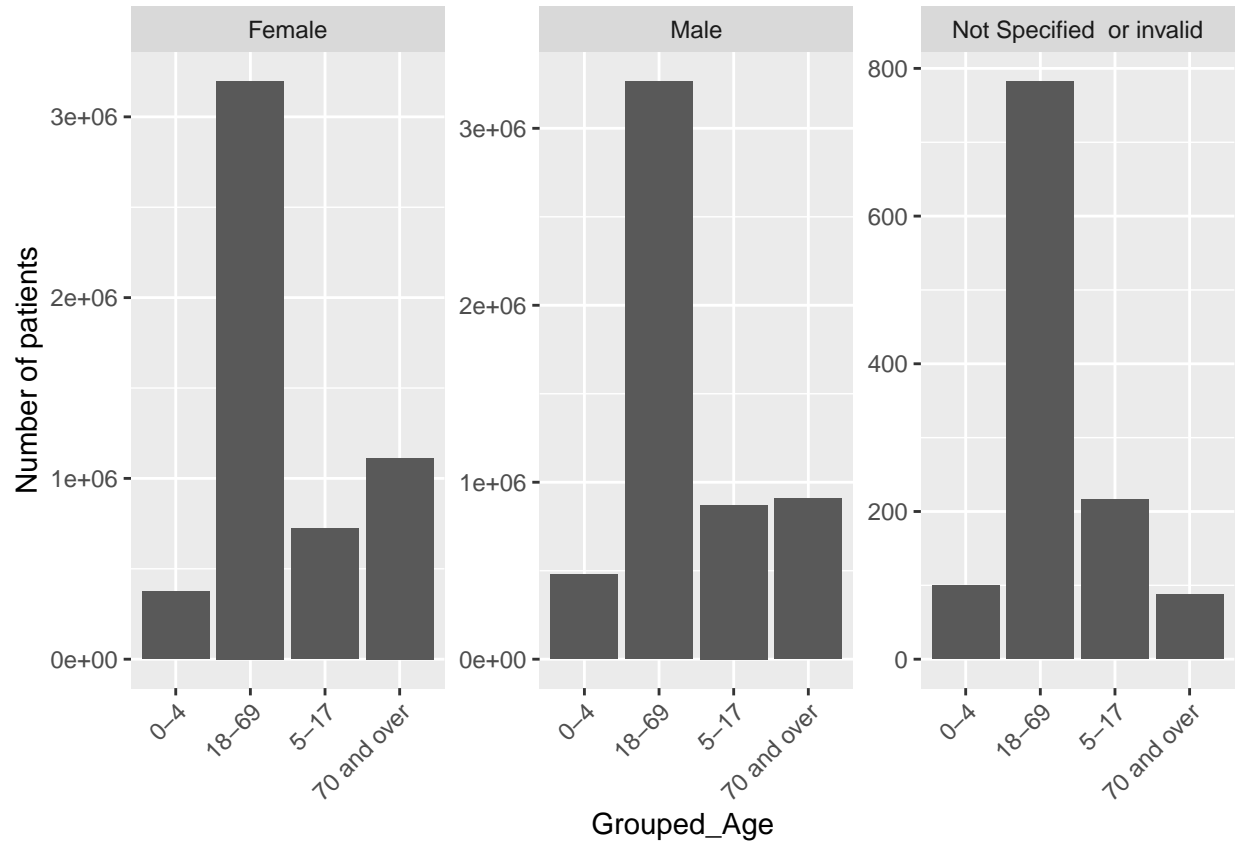
Based on the STL decomposition, it is evident that each health board follows a similar trend. Moreover, it is obvious that there is a significant downturn during the Covid-19 era. As for seasonality, there is a pronounced surge in the number of patient's attendance in the middle of the year (approximately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

```
## $'Aneurin Bevan'
##           Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## Jul 2023      15327.46  13401.32  17253.59  12381.69  18273.22
## Aug 2023      14758.62  12683.77  16833.47  11585.41  17931.83
## Sep 2023      14331.89  12174.36  16489.43  11032.23  17631.56
```

```

## Oct 2023      14077.53 11891.77 16263.28 10734.70 17420.35
## Nov 2023      13913.04 11715.52 16110.55 10552.23 17273.84
## Dec 2023      13809.83 11607.72 16011.95 10441.99 17177.68
##
## $'Betsi Cadwaladr'
##      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
## Jul 2023      22520.72 21208.70 23832.74 20514.16 24527.28
## Aug 2023      22195.49 20282.25 24108.74 19269.44 25121.55
## Sep 2023      20359.95 18117.18 22602.73 16929.92 23789.98
## Oct 2023      19284.65 16835.70 21733.59 15539.31 23029.99
## Nov 2023      17985.58 15400.91 20570.24 14032.67 21938.49
## Dec 2023      17199.82 14523.44 19876.20 13106.64 21293.00
##
## $'Cardiff & Vale'
##      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
## Jul 2023      12894.49 11932.17 13856.81 11422.743 14366.23
## Aug 2023      12453.99 11210.08 13697.89 10551.600 14356.38
## Sep 2023      12491.75 11090.31 13893.20 10348.425 14635.08
## Oct 2023      12649.40 11151.53 14147.27 10358.603 14940.20
## Nov 2023      12125.90 10566.68 13685.12  9741.274 14510.52
## Dec 2023      11831.67 10232.62 13430.73  9386.128 14277.22
##
## $Grouped_4_organisation
##      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
## Jul 2023      28201.58 26101.77 30301.39 24990.20 31412.96
## Aug 2023      27165.96 24527.91 29804.02 23131.41 31200.52
## Sep 2023      26742.12 23837.94 29646.29 22300.56 31183.67
## Oct 2023      26782.76 23735.25 29830.27 22122.00 31443.52
## Nov 2023      25645.29 22517.89 28772.70 20862.34 30428.25
## Dec 2023      25286.03 22113.33 28458.73 20433.81 30138.26
##
## $'Hywel Dda'
##      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
## Jul 2023      15984.72 14923.747 17045.70 14362.099 17607.35
## Aug 2023      16088.04 14582.199 17593.88 13785.055 18391.02
## Sep 2023      14280.20 12518.452 16041.94 11585.841 16974.55
## Oct 2023      13551.66 11633.763 15469.55 10618.491 16484.82
## Nov 2023      12548.97 10532.122 14565.81  9464.468 15633.47
## Dec 2023      11964.73  9883.705 14045.76  8782.076 15147.39
##
## $'Powys Teaching'
##      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
## Jul 2023      1593.479 1421.3226 1765.635 1330.1887 1856.768
## Aug 2023      1476.309 1242.9241 1709.693 1119.3776 1833.240
## Sep 2023      1391.293 1116.9363 1665.651  971.7004 1810.887
## Oct 2023      1260.690  956.2214 1565.159  795.0455 1726.335
## Nov 2023      1188.828  861.2557 1516.400  687.8494 1689.806
## Dec 2023      1020.005  674.2630 1365.747  491.2381 1548.772
##
## $Total
##      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
## Jul 2023      96354.13 89558.55 103149.71 85961.19 106747.1
## Aug 2023      89505.44 80770.23  98240.66 76146.08 102864.8
## Sep 2023      90104.99 80309.35  99900.63 75123.86 105086.1

```

## Oct 2023	89432.47	79003.05	99861.90	73482.04	105382.9
## Nov 2023	85475.78	74652.91	96298.65	68923.62	102027.9
## Dec 2023	83342.11	72270.11	94414.11	66408.95	100275.3



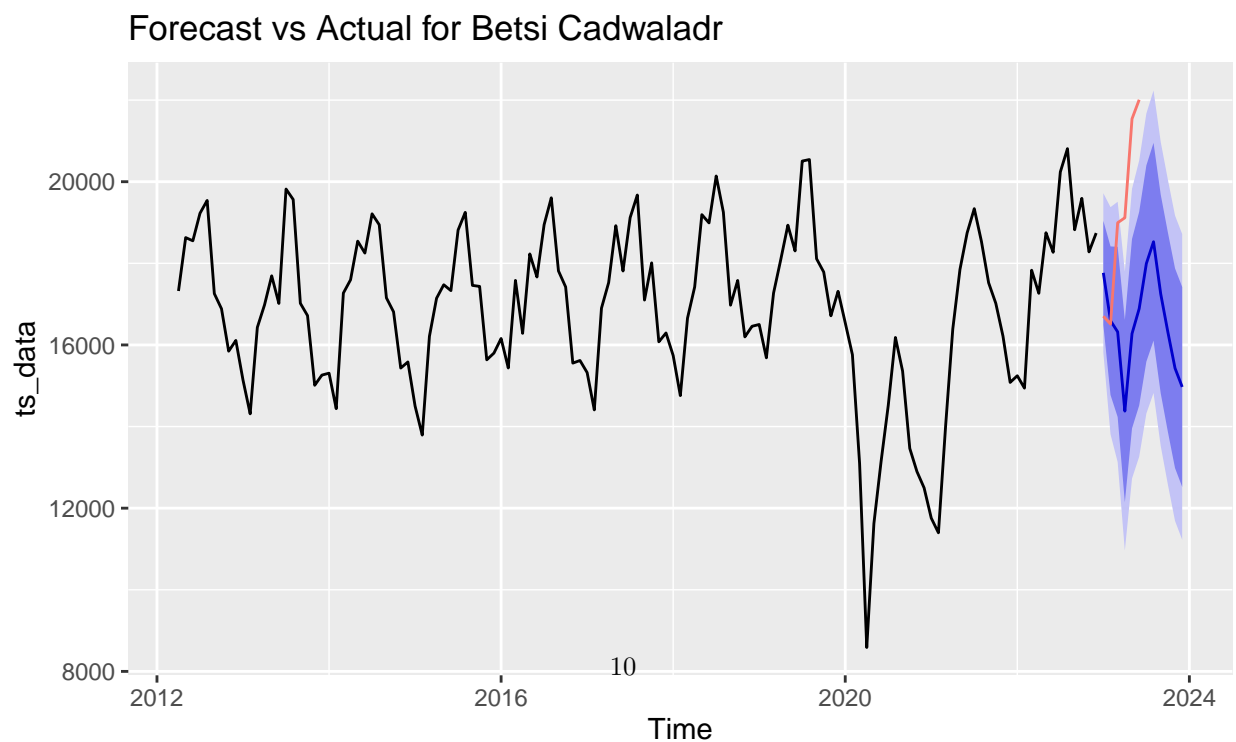
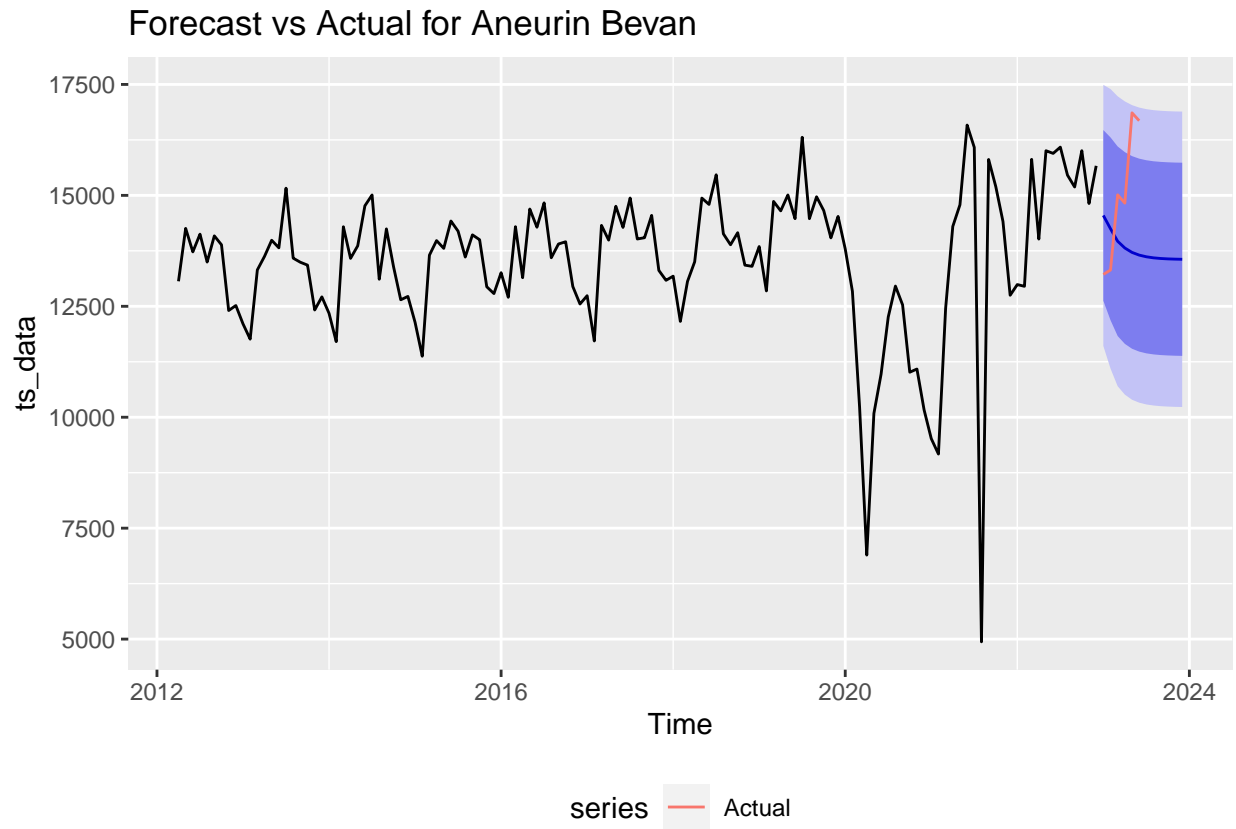


## 4 Forecast with ARIMA

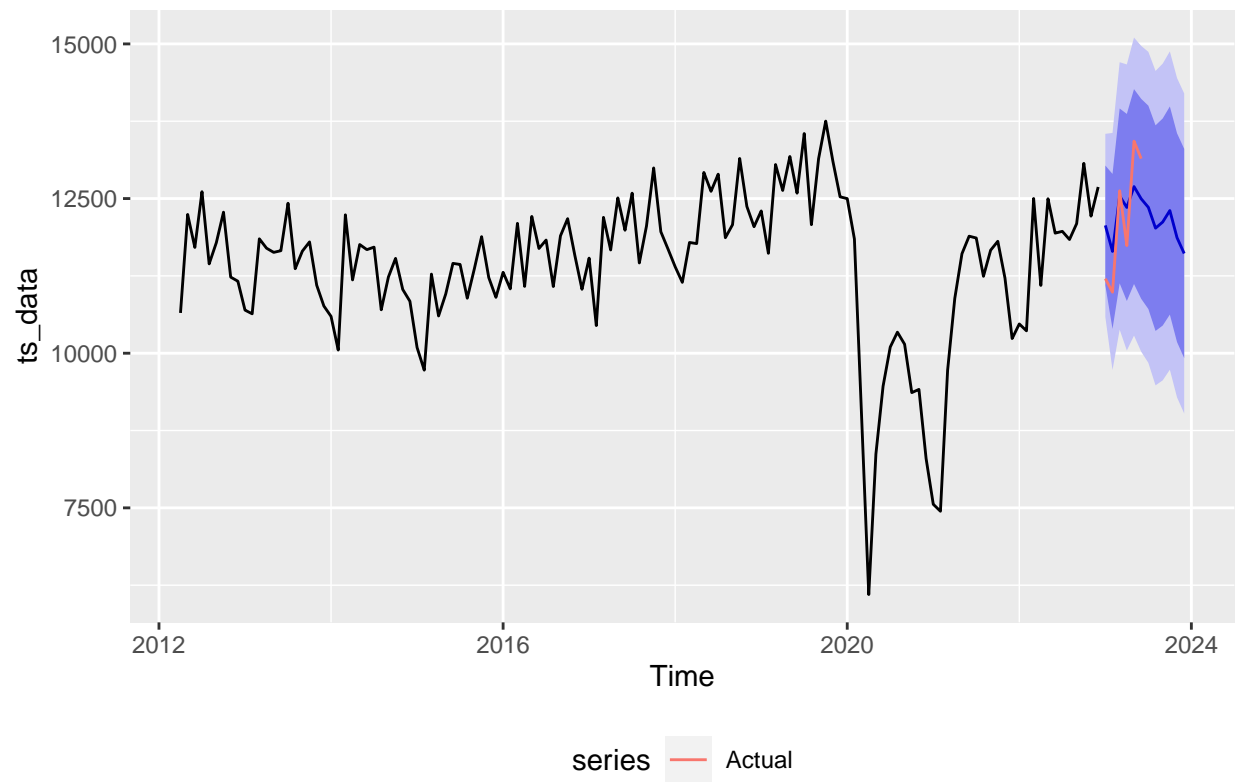
### 4.0.1 Define the forecast horizon and validation period

### 4.0.2 Splitting the data into training and testing sets

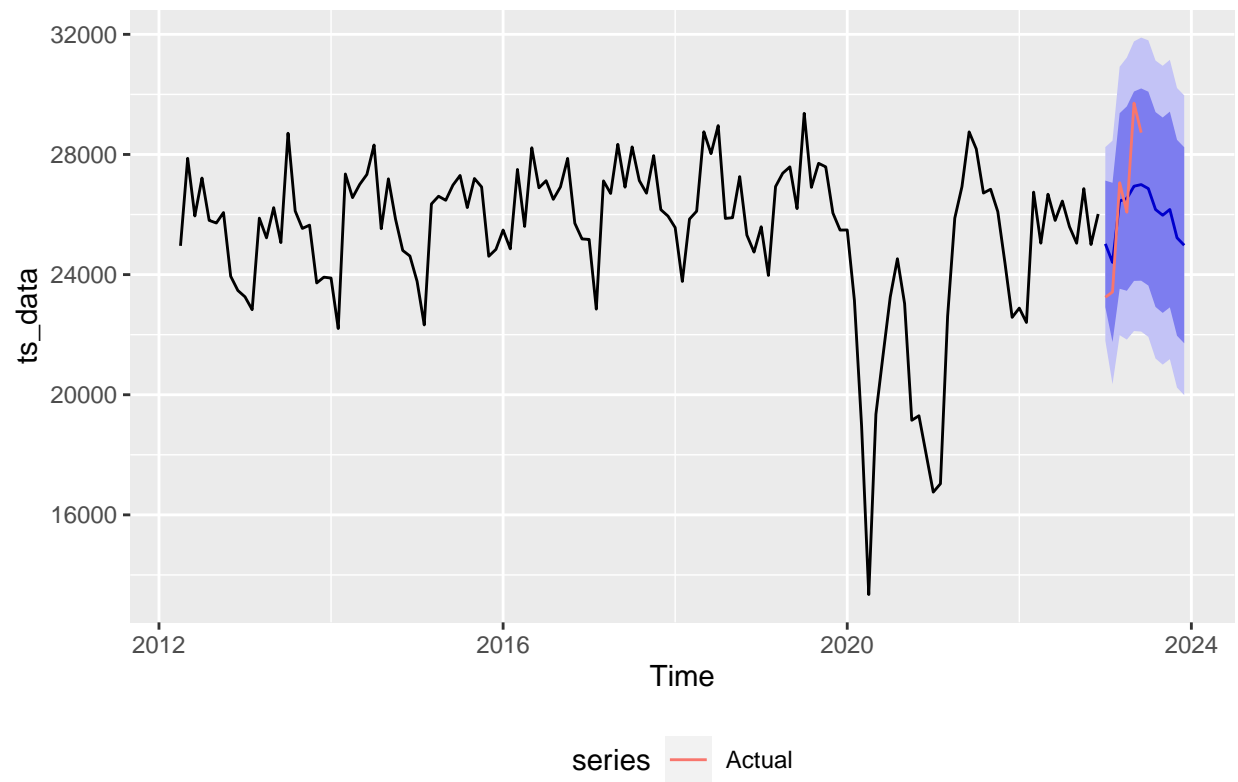
### 4.1 Plotting using ARIMA



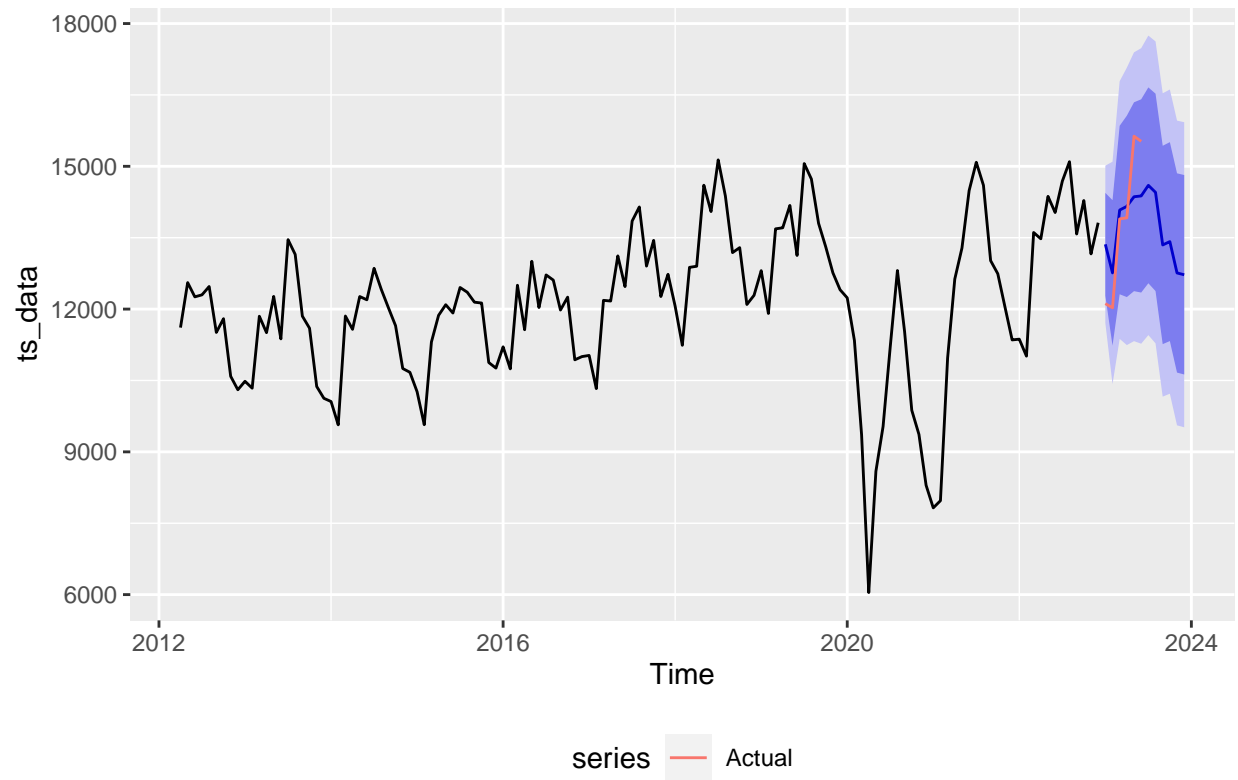
Forecast vs Actual for Cardiff & Vale



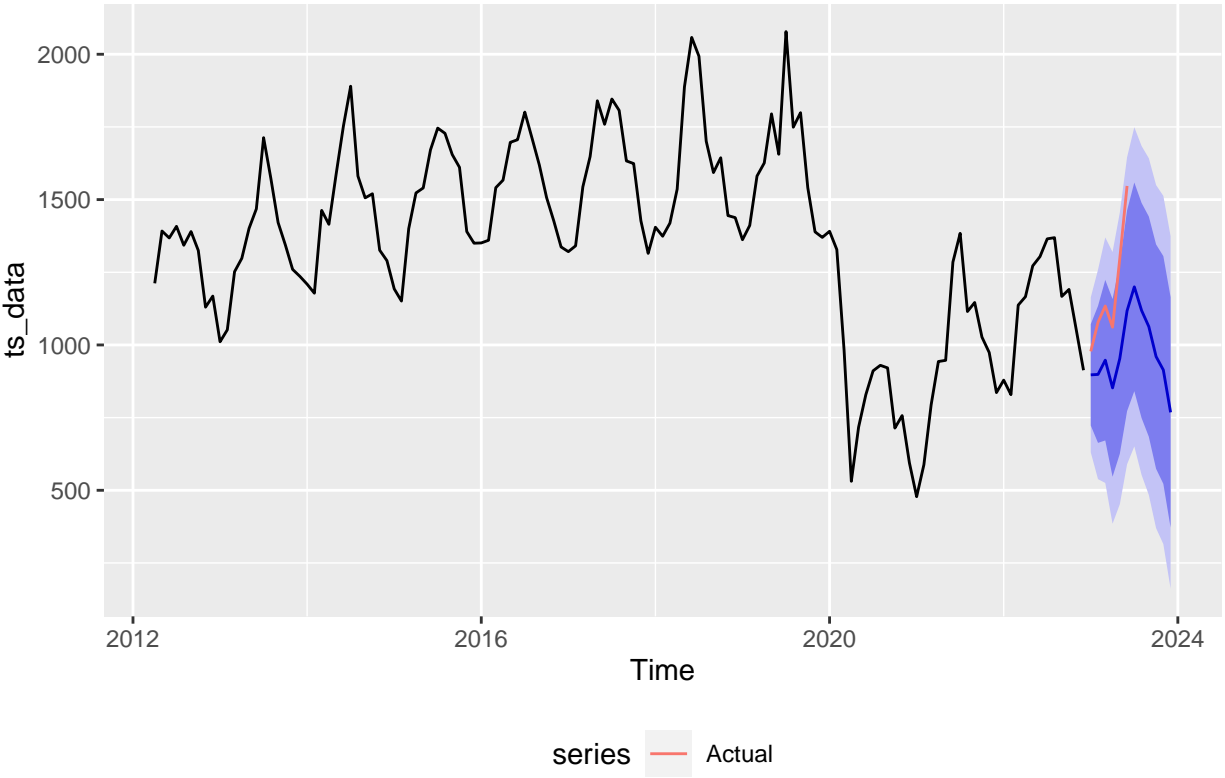
Forecast vs Actual for Grouped\_4\_organisation



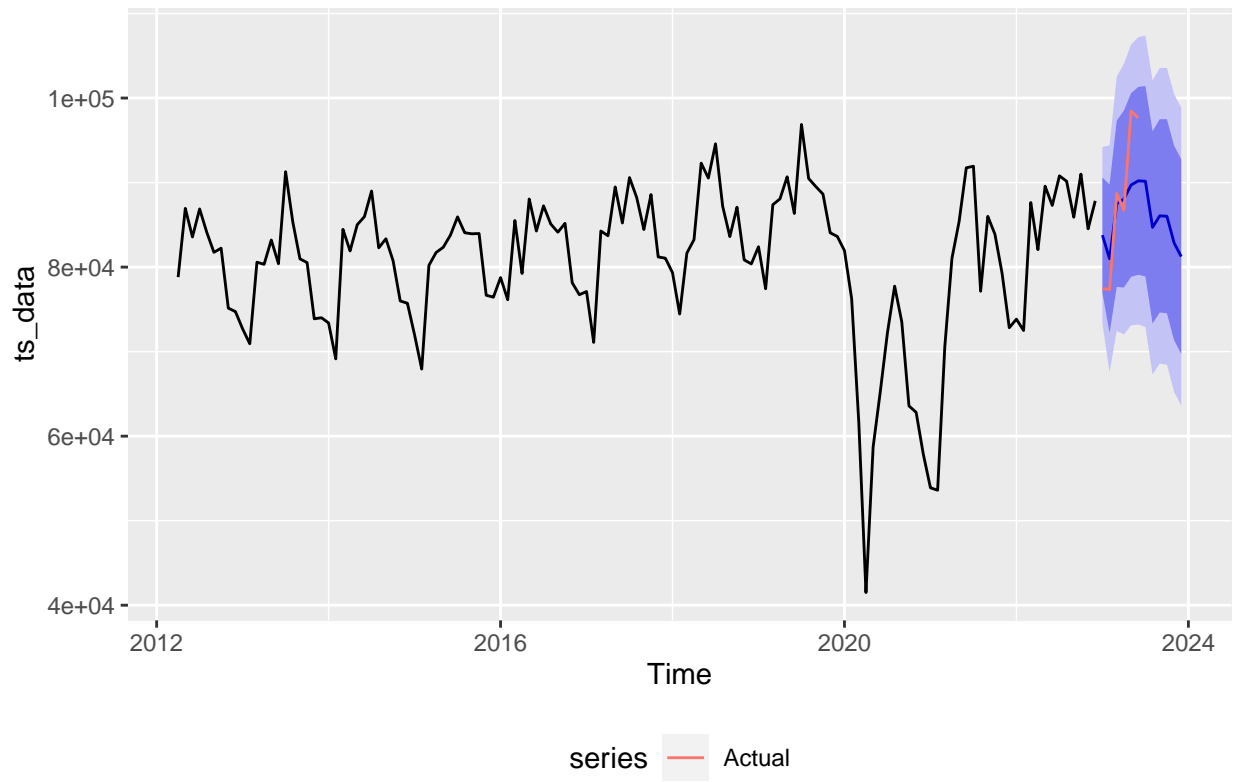
Forecast vs Actual for Hywel Dda



Forecast vs Actual for Powys Teaching



Forecast vs Actual for Total



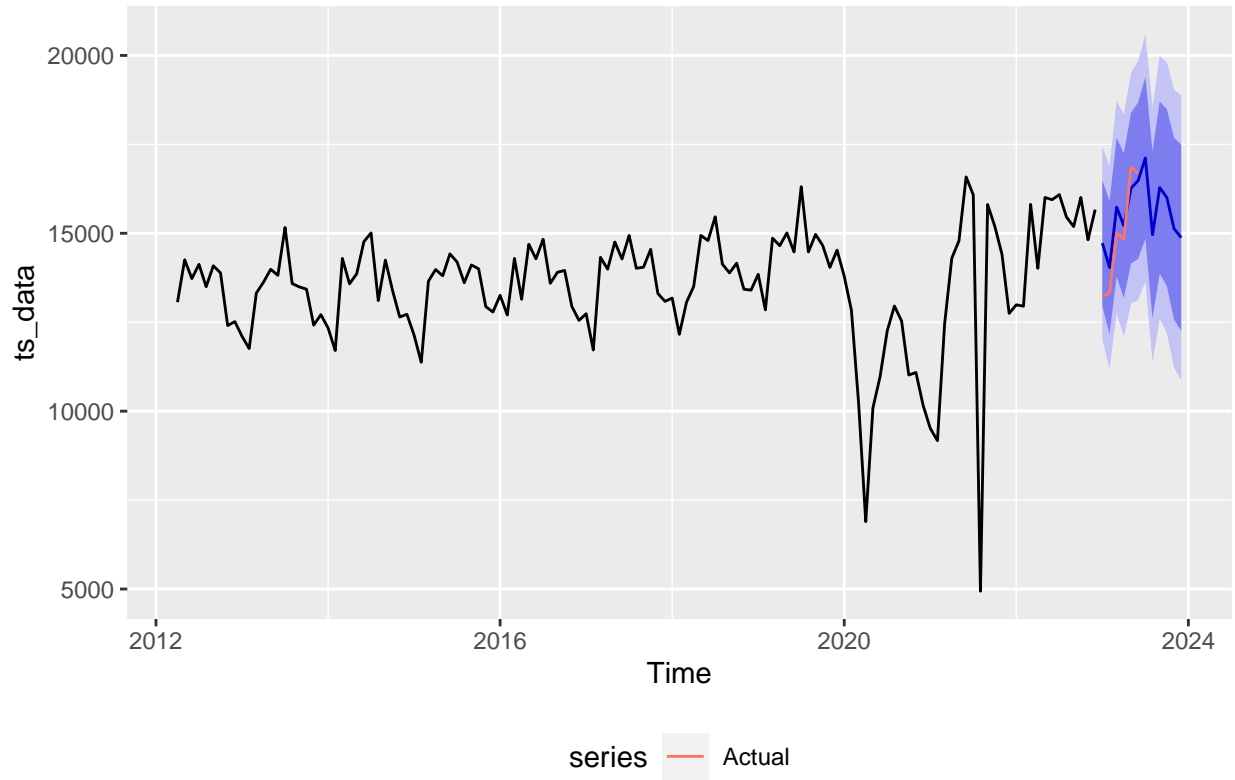




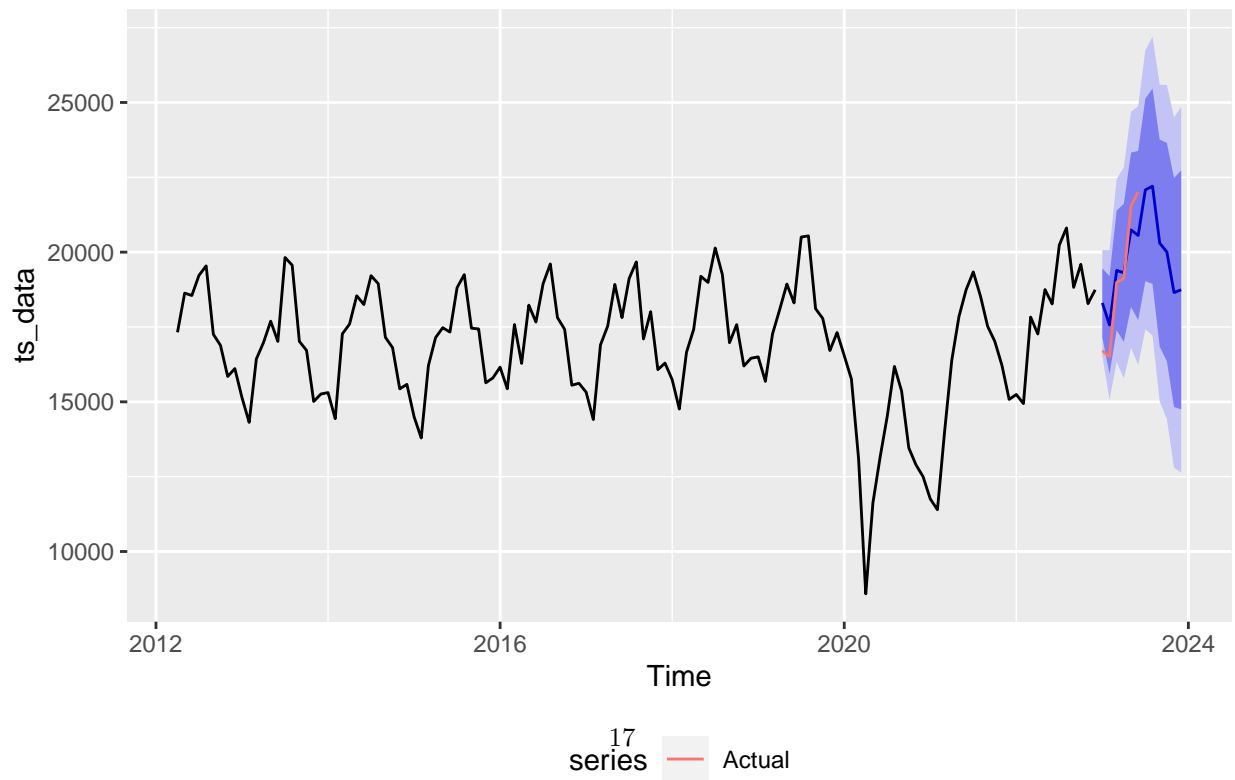
## 5 Forecasting with ETS

### 5.1 Plotting with ets

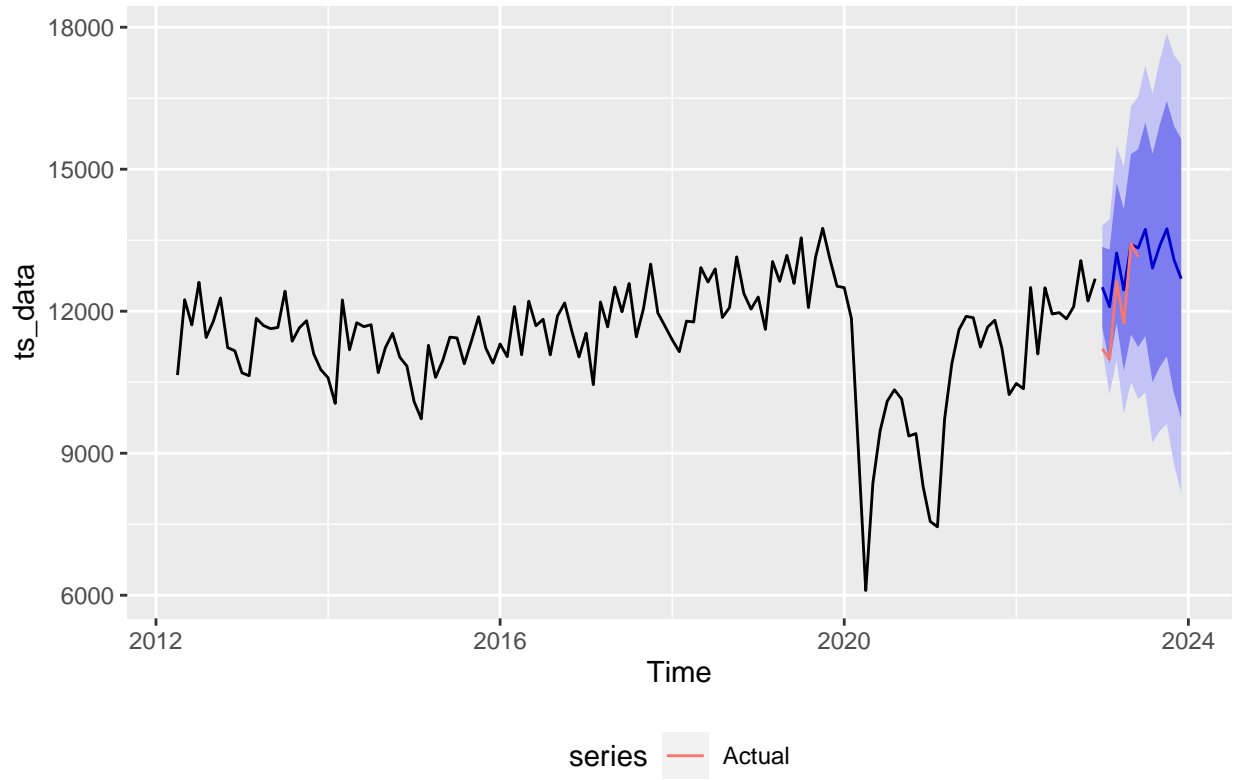
ETS Forecast vs Actual for Aneurin Bevan



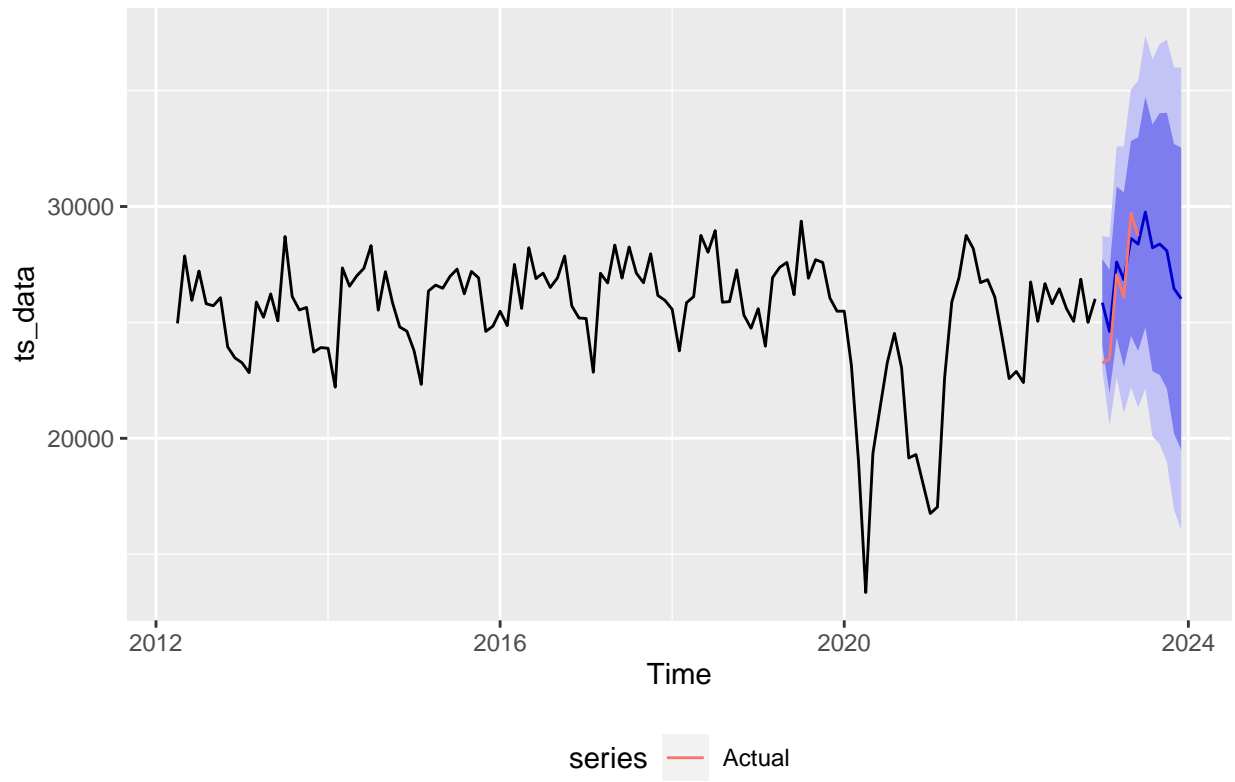
ETS Forecast vs Actual for Betsi Cadwaladr



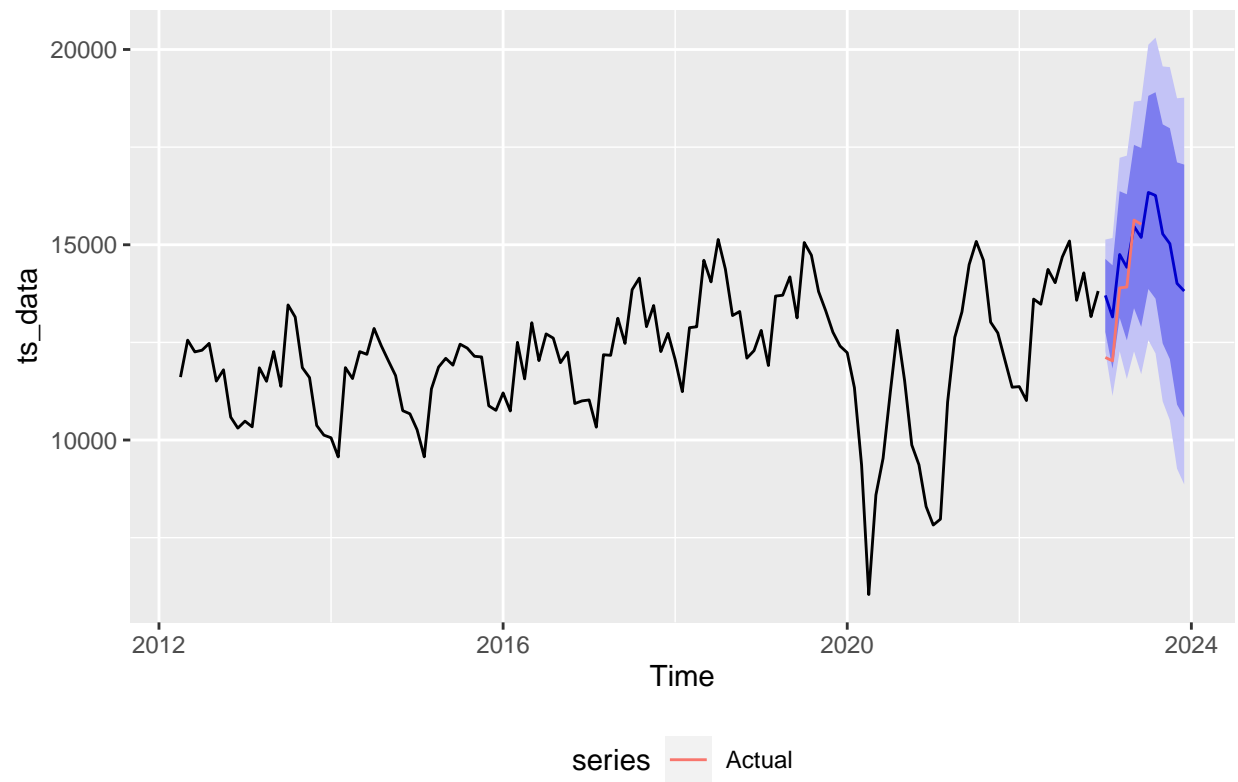
ETS Forecast vs Actual for Cardiff & Vale



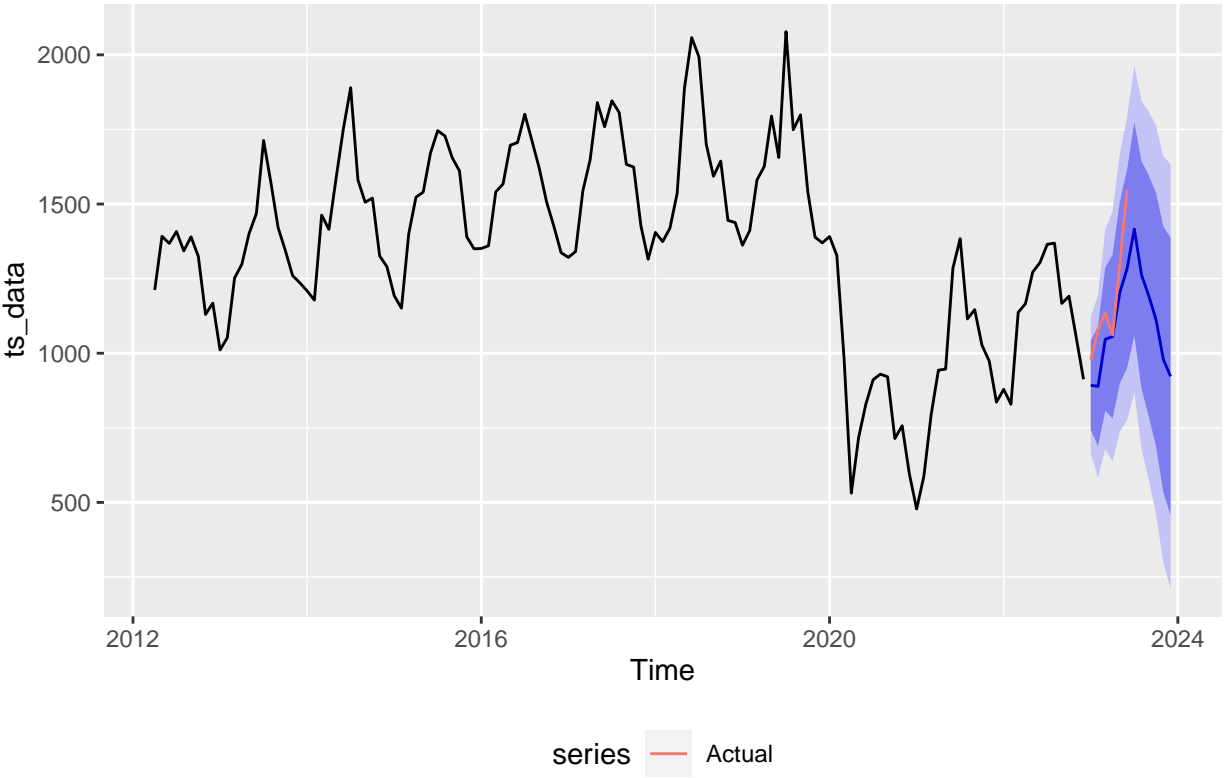
ETS Forecast vs Actual for Grouped\_4\_organisation

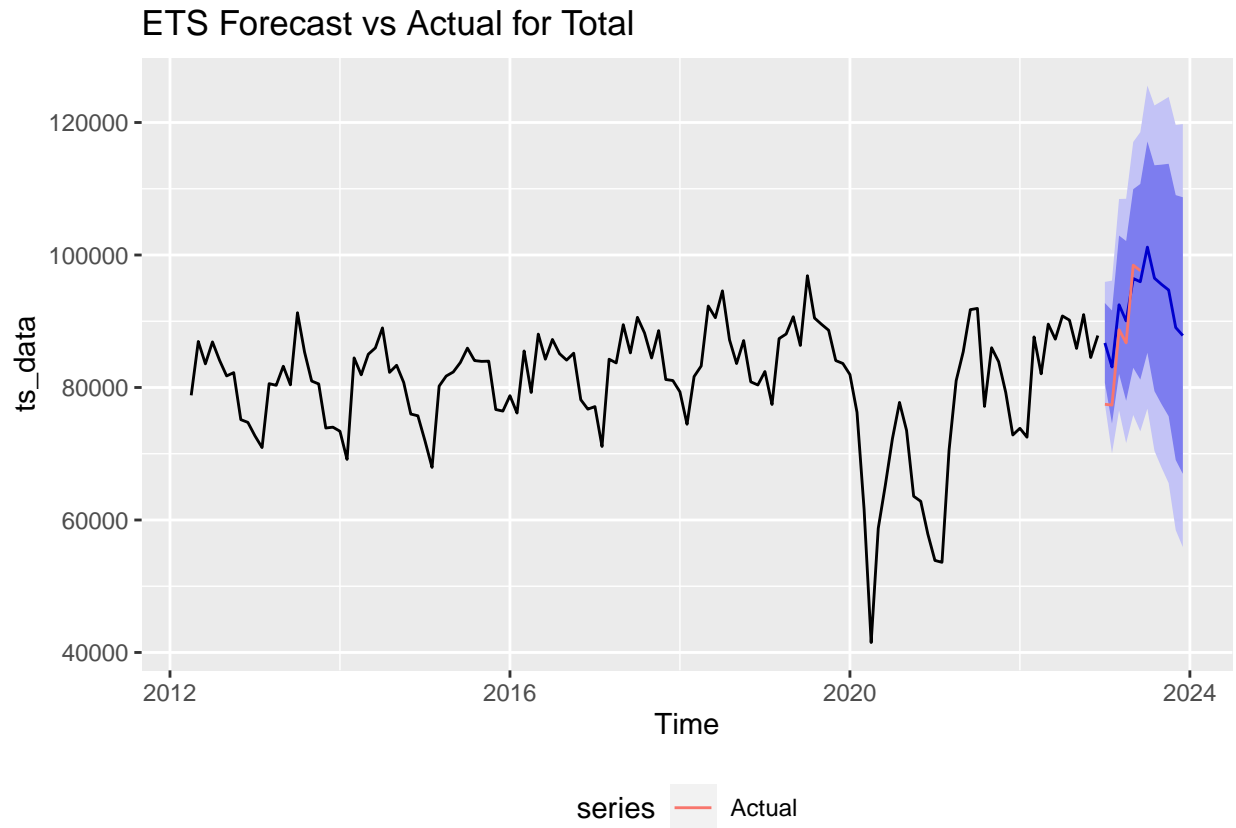


ETS Forecast vs Actual for Hywel Dda



ETS Forecast vs Actual for Powys Teaching





## 6 Accuracy assessment for ARIMA and ETS

```
##
## Accuracy metrics for: Aneurin Bevan
## -----
## ARIMA:
## $MAE
## [1] 688.4172
##
## $RMSE
## [1] 801.6437
##
## $MAPE
## [1] 0.04825847
##
## ETS:
## $MAE
## [1] NaN
##
## $RMSE
## [1] NaN
##
## $MAPE
## [1] NaN
```

```

##
##
##
## Accuracy metrics for: Betsi Cadwaladr
## -----
## ARIMA:
## $MAE
## [1] 915.7354
##
## $RMSE
## [1] 1050.122
##
## $MAPE
## [1] 0.04895248
##
## ETS:
## $MAE
## [1] NaN
##
## $RMSE
## [1] NaN
##
## $MAPE
## [1] NaN
##
##
##
## Accuracy metrics for: Cardiff & Vale
## -----
## ARIMA:
## $MAE
## [1] 654.8113
##
## $RMSE
## [1] 799.0302
##
## $MAPE
## [1] 0.05682325
##
## ETS:
## $MAE
## [1] NaN
##
## $RMSE
## [1] NaN
##
## $MAPE
## [1] NaN
##
##
##
## Accuracy metrics for: Grouped_4_organisation
## -----
## ARIMA:

```

```

## $MAE
## [1] 1089.509
##
## $RMSE
## [1] 1315.071
##
## $MAPE
## [1] 0.0434871
##
## ETS:
## $MAE
## [1] NaN
##
## $RMSE
## [1] NaN
##
## $MAPE
## [1] NaN
##
##
##
## Accuracy metrics for: Hywel Dda
## -----
## ARIMA:
## $MAE
## [1] 763.0849
##
## $RMSE
## [1] 905.503
##
## $MAPE
## [1] 0.05915982
##
## ETS:
## $MAE
## [1] NaN
##
## $RMSE
## [1] NaN
##
## $MAPE
## [1] NaN
##
##
##
## Accuracy metrics for: Powys Teaching
## -----
## ARIMA:
## $MAE
## [1] 120.9989
##
## $RMSE
## [1] 146.7962
##

```



```

## $MAPE
## [1] 0.09808414
##
## ETS:
## $MAE
## [1] NaN
##
## $RMSE
## [1] NaN
##
## $MAPE
## [1] NaN
##
##
##
## Accuracy metrics for: Total
## -----
## ARIMA:
## $MAE
## [1] 4297.46
##
## $RMSE
## [1] 5014.551
##
## $MAPE
## [1] 0.05203606
##
## ETS:
## $MAE
## [1] NaN
##
## $RMSE
## [1] NaN
##
## $MAPE
## [1] NaN

```