**Summary Data and Results**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Frequency | Horizon | Number of Samples | Minimum training sample length |
| Prison | 4 (quarterly) | 8 | 3 | 24 |
| Tourism | 12 (monthly) | 12 | 10 | 144 |
| Wikipedia | 7 (weekly) | 7 | 10 | 324 |
| Labour | 4 (quarterly) | 12 | 5 | 68 |

**Error calculation**

* Step 1 – For each time series in the hierarchy calculate the error metric (MSE, MAPE)
* Step 2 – Calculate the mean error across the time series in each level (this gives the mean error for each level)
* Step 3 – Calculate the overall error by getting the mean error across all time series errors we have in Step 1
* Step 4 – Repeat step 1-3 for all samples
* Step 5 – Calculate the mean error for each level and overall, across the samples
* Step 6 – Calculate the percentage improvement for each level and overall

**Results Across Samples**

Full Horizon (*In the full horizon ML reconciliation performs best for most 3 out of datasets when the base model is ETS. However, when the base forecasts are from ARIMA the ML reconciliation performs best only for the Tourism dataset*)

|  |  |  |
| --- | --- | --- |
| **Dataset** | **ARIMA** | **ETS** |
| Prison | OLS  Best ML Rank – 4 (Case 1 Lambda 1-4) | Case 1 Lambda 0.01-0.09 |
| Tourism | Case 1 Lambda 0.01-5 | Case 2 Lambda 1 |
| Labour | OLS  Best ML Rank – 4 (Case 2 Lambda 1) | Case 1 Lambda 0.01-0.09 |
| Wikipedia | OLS  Best ML Rank – 5 (Case 2 Lambda 0.1-0.9) | OLS  Best ML Rank – 5 (Case 2 Lambda 1) |

Short Horizon (*In the short horizon ML reconciliation performs best for most 3 out of datasets when the base model is ETS – similar observation as the full horizon. When the base forecasts are from ARIMA the ML reconciliation performs best for Prison and Tourism datasets*)

|  |  |  |
| --- | --- | --- |
| **Dataset** | **ARIMA** | **ETS** |
| Prison | Case 1 Lambda 1-4 | Case 1 Lambda 0.01-0.09 |
| Tourism | Case 2 Lambda 0.01-5 | Case 1 Lambda 1-4 |
| Labour | OLS  Best ML Rank – 2 (Case 2 Lambda 1) | Case 1 Lambda 0.01-0.09 |
| Wikipedia | OLS  Best ML Rank – 5 (Case 2 Lambda 0.1-0.9) | OLS  Best ML Rank – 5 (Case 2 Lambda 0.01-5) |

**Lambda Ranges – Lambda value for each sample from the overall best ML method**

**Prison**

ARIMA

Chart, treemap chart

Description automatically generated

ETS

Chart, funnel chart

Description automatically generated

**Tourism**

ARIMA

Chart, bar chart

Description automatically generated

ETS

Chart, bar chart

Description automatically generated

Labour

ARIMA

Chart, waterfall chart

Description automatically generated

ETS

**Chart, waterfall chart

Description automatically generated**

**Wikipedia**

ARIMA

Chart, bar chart

Description automatically generated

ETS

Chart, bar chart

Description automatically generated

**Lambda Ranges – Lambda value and the best ML method per sample**

**Prison**

ARIMA

Chart, treemap chart

Description automatically generated

ETS

A picture containing table

Description automatically generated

**Tourism**

ARIMA

Chart, bar chart

Description automatically generated

ETS

Chart, waterfall chart

Description automatically generated

Labour

ARIMA

Chart, bar chart, waterfall chart

Description automatically generated

ETS

**Chart, bar chart, funnel chart

Description automatically generated**

**Wikipedia**

ARIMA

Chart, bar chart

Description automatically generated

ETS

Chart, bar chart, waterfall chart

Description automatically generated