**Changes:**

1. Data Cleaning for all datasets
2. Wikipedia dataset – hierarchy structure changed based of the paper, Total X Access X Agent X Language X Purpose X Article

**Data Before and After Cleaning**

**Prison**

**Chart, line chart

Description automatically generatedChart, line chart

Description automatically generated**

**Tourism**

**Chart, histogram

Description automatically generatedChart, bar chart, histogram

Description automatically generated**

**Labour**

**Graphical user interface, histogram

Description automatically generatedGraphical user interface, histogram

Description automatically generated**

**Wikipedia**

**Chart, histogram

Description automatically generated Chart, line chart, histogram

Description automatically generated**

**Summary Data and Results**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dataset | Frequency | Horizon | Number of Samples | Minimum training sample length | Number of levels | Number of total time series |
| Prison | 4 (quarterly) | 8 | 3 | 24 | 5 | 121 |
| Tourism | 12 (monthly) | 12 | 10 | 144 | 3 | 85 |
| Wikipedia | 7 (weekly) | 7 | 10 | 324 | 6 | 1095 |
| Labour | 4 (quarterly) | 12 | 5 | 68 | 4 | 57 |

**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 - Latest**

Full Horizon

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dataset** | **ARIMA** | **ETS** | **DeepAR** | **WaveNet** |
| Prison | OLS  Best ML Rank – 4 (Case 2 Lambda 0.1-0.9) | Case 1 Lambda 0. 1-0. 9 No skip /  Case 1 Lambda 1 |  |  |
| Tourism | OLS  Best ML Rank – 3 Case 1 Lambda 0.01-5 | OLS  Best ML Rank- 2  Case 2 Lambda 1 |  |  |
| Labour | OLS  Best ML Rank – 4 (Case 2 Lambda 0.1-0.9) | Case 2 Lambda 0.01-0.09 |  |  |
| Wikipedia | Case 1 Lambda 0.01-0.09 / Case 2 Lambda 1-4 | Bottom-Up  Best ML Rank – 4 (Case 1 Lambda 0.01-5) |  |  |

Short Horizon

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dataset** | **ARIMA** | **ETS** | **DeepAR** | **WaveNet** |
| Prison | Case 1 Lambda 0.1-0.9 | Case 2 Lambda 0.01-0.09 |  |  |
| Tourism | OLS  Best ML Rank – 3 Case 2 Lambda 0.01-5 | Case 2 Lambda 1 |  |  |
| Labour | OLS  Best ML Rank – 3 (Case 2 Lambda 0.01-5) | Case 1 Lambda 0.01-5 |  |  |
| Wikipedia | Case 1 Lambda 0.1-0.9 | Case 2 Lambda 1-4 |  |  |

**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