**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

**Across Samples**