Table of Contents

[Prison 3](#_Toc112940550)

[ARIMA 3](#_Toc112940551)

[MSE 3](#_Toc112940552)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 3](#_Toc112940553)

[MAE 3](#_Toc112940554)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 4](#_Toc112940555)

[ETS 4](#_Toc112940556)

[MSE 4](#_Toc112940557)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 4](#_Toc112940558)

[MAE 4](#_Toc112940559)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 5](#_Toc112940560)

[Labour 6](#_Toc112940561)

[ARIMA 6](#_Toc112940562)

[MSE 6](#_Toc112940563)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 6](#_Toc112940564)

[MAE 6](#_Toc112940565)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 7](#_Toc112940566)

[ETS 7](#_Toc112940567)

[MSE 7](#_Toc112940568)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 7](#_Toc112940569)

[MAE 8](#_Toc112940570)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 8](#_Toc112940571)

[Tourism 9](#_Toc112940572)

[ARIMA 9](#_Toc112940573)

[MSE 9](#_Toc112940574)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 9](#_Toc112940575)

[MAE 9](#_Toc112940576)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 9](#_Toc112940577)

[ETS 10](#_Toc112940578)

[MSE 10](#_Toc112940579)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 10](#_Toc112940580)

[MAE 10](#_Toc112940581)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 10](#_Toc112940582)

[Wikipedia 12](#_Toc112940583)

[ARIMA 12](#_Toc112940584)

[MSE 12](#_Toc112940585)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 12](#_Toc112940586)

[MAE 12](#_Toc112940587)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 12](#_Toc112940588)

[ETS 12](#_Toc112940589)

[MSE 12](#_Toc112940590)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling windows) 12](#_Toc112940591)

[MAE 12](#_Toc112940592)

[One-window only (errors are only calculated for the immediate 1-step ahead rolling window only) 12](#_Toc112940593)

[Running Time for ML method 13](#_Toc112940594)

Rolling window experiments

New Code files:

*run\_arima\_ets\_rolling\_origin.py* -> runs the base arima and ets rolling window forecasts

*run\_ml\_method\_rolling\_origin.py* -> runs the ML reconciliation for the rolling windows

*run\_hts\_benchmarks\_rolling\_origin.py* -> runs the reconciliation benchmarks for the rolling window

*calculate\_rolling\_origin\_errors.py* -> this code will read all the base and reconciled forecasts and output a csv file with the percentage relative improvement of each reconciliation method [the order of the columns shows order of the method performance]

The forecasts are only for one-step ahead

* Each rolling window has a set of fitted values, actual values corresponding to the fitted data, forecasts 1-step and the actual values for the 1-step ahead forecasts
* For all reconciliation approaches the reconciliation approach is trained/ parameters are found after every 10th rolling window (we discussed this approach to reduce the time complexity for re-estimating parameters for each rolling-window)
  + For example, if we have 20 rolling windows for 1-10 rolling windows the parameters are found for the reconciliation method using the 1st windows fitted values. And the forecasts are reconciled using these parameters for all 1-10 rolling windows. Then the parameters are re-estimated from the 11th rolling window and these parameters used to reconcile the forecasts across 11-20th rolling windows.

For ML methods:

Case 1 -> Only the bottom level error as the validation loss

Case 2 -> Whole hierarchy is considered in the validation loss

Number of rolling-windows per dataset

Prison – 24

Tourism – 120

Labour – 60

Wikipedia – 70

If we calculate errors only for the immediate 1-step ahead rolling window the number of samples data is as follows:

Prison – 3

Tourism – 12

Labour – 6

Wikipedia – 7

Base forecasts, reconciled forecasts and all error percentages are uploaded [here](https://drive.google.com/drive/folders/1NQqmZtG8DG3ZMERAbzUJnntFF6906xDk?usp=sharing).

# Prison

## ARIMA

### ARIMA - MSE

Training at every 10th iteration – mean across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintshrink** | **wls** | **bottomup** | **erm** | **case2\_lambda\_[0.1, 0.9]** |
| **Australia** | 5.82 | -21.17 | -24.63 | -129.82 | -750.49 | -22496.46 |
| **State** | 2.73 | -2.5 | -2.83 | -46.74 | -622.88 | -6070.32 |
| **Gender** | 4.63 | 3.12 | 4.26 | -35.98 | -614.71 | -5518.42 |
| **Legal** | 13.45 | 11.18 | 11.63 | -4.37 | -561.59 | -2404.44 |
| **Indigenous** | 11.29 | 13.12 | 13.43 | 0.0 | -545.22 | -1986.19 |
| **Overall** | 7.6 | -0.86 | -1.49 | -49.66 | -628.59 | -8793.85 |

Training at each iteration – mean across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintshrink** | **wls** | **case2\_lambda\_[0.01, 5]** | **bottomup** | **erm** |
| **Australia** | 5.82 | -17.9 | -25.59 | -17.85 | -129.82 | -107.86 |
| **State** | 2.73 | -3.21 | -4.08 | -12.87 | -46.74 | -106.68 |
| **Gender** | 4.63 | 2.76 | 3.28 | -6.88 | -35.98 | -105.44 |
| **Legal** | 13.45 | 10.99 | 10.7 | 1.61 | -4.37 | -92.17 |
| **Indigenous** | 11.29 | 13.33 | 12.84 | 3.92 | 0.0 | -107.48 |
| **Overall** | 7.6 | -0.2 | -2.43 | -7.22 | -49.66 | -103.81 |

Training at each iteration – median across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **case1\_lambda\_1** | **mintshrink** | **wls** | **ols** | **bottomup** | **erm** |
| **Australia** | 7.44 | 55.27 | 31.32 | 10.13 | -54.26 | -75.98 |
| **State** | -22.9 | -17.84 | -16.52 | -11.09 | -35.91 | -141.73 |
| **Gender** | 2.13 | 4.24 | 7.31 | 2.89 | 0.1 | -108.98 |
| **Legal** | 2.58 | 9.88 | 4.15 | 13.96 | 16.7 | -47.39 |
| **Indigenous** | -3.85 | 5.24 | 1.67 | 1.9 | 0.0 | -83.84 |
| **Overall** | 18.17 | 13.21 | 12.97 | 7.15 | 6.49 | -63.16 |

NEW EXPERIMENTS – Input transformed fitted values instead of fitted values to the ML method

Training every 10th iteration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintshrink** | **wls** | **bottomup** | **erm** | **case2\_lambda\_[0.1, 0.9]** |
| **Australia** | 5.82 | -21.17 | -24.63 | -129.82 | -750.49 | -21353.95 |
| **State** | 2.73 | -2.5 | -2.83 | -46.74 | -622.88 | -5802.45 |
| **Gender** | 4.63 | 3.12 | 4.26 | -35.98 | -614.71 | -5296.7 |
| **Legal** | 13.45 | 11.18 | 11.63 | -4.37 | -561.59 | -2319.34 |
| **Indigenous** | 11.29 | 13.12 | 13.43 | 0.0 | -545.22 | -1927.79 |
| **Overall** | 7.6 | -0.86 | -1.49 | -49.66 | -628.59 | -8379.77 |

Training at each iteration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintshrink** | **wls** | **case2\_lambda\_[1, 4]** | **bottomup** | **erm** |
| **Australia** | 5.82 | -17.9 | -25.59 | -6.16 | -129.82 | -107.86 |
| **State** | 2.73 | -3.21 | -4.08 | -9.26 | -46.74 | -106.68 |
| **Gender** | 4.63 | 2.76 | 3.28 | -5.88 | -35.98 | -105.44 |
| **Legal** | 13.45 | 10.99 | 10.7 | 0.44 | -4.37 | -92.17 |
| **Indigenous** | 11.29 | 13.33 | 12.84 | 2.6 | 0.0 | -107.48 |
| **Overall** | 7.6 | -0.2 | -2.43 | -3.8 | -49.66 | -103.81 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **case1\_lambda\_[0.01, 5]** | **mintshrink** | **ols** | **wls** | **erm** | **bottomup** |
| **Australia** | 84.5 | 36.17 | 15.19 | 10.8 | -1.57 | -144.23 |
| **State** | 7.01 | -6.12 | -5.04 | -4.35 | -101.63 | -50.58 |
| **Gender** | 11.48 | -0.9 | 2.21 | 4.57 | -121.21 | -31.7 |
| **Legal** | -2.93 | 3.05 | 7.09 | 3.64 | -42.33 | -14.59 |
| **Indigenous** | 6.66 | 15.29 | 12.44 | 15.86 | -15.86 | 0.0 |
| **Overall** | 24.42 | 11.94 | 7.71 | 6.95 | -46.82 | -51.61 |

### ARIMA - MAE

Training at every 10th iteration – mean across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **ols** | **bottomup** | **erm** | **case2\_lambda\_[0.1, 0.9]** |
| **Australia** | -2.43 | -5.84 | 5.31 | -48.84 | -176.59 | -1230.18 |
| **State** | -0.35 | -1.2 | 0.45 | -18.28 | -179.9 | -590.09 |
| **Gender** | 3.16 | 2.72 | -0.49 | -11.54 | -164.07 | -509.28 |
| **Legal** | 4.67 | 4.85 | 3.44 | -1.71 | -178.96 | -321.27 |
| **Indigenous** | 4.55 | 4.61 | 1.75 | 0.0 | -176.6 | -269.01 |
| **Overall** | 3.01 | 2.6 | 1.88 | -9.24 | -175.51 | -452.15 |

Training at each iteration – mean across windows

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **ols** | **case1\_lambda\_1** | **case2\_lambda\_[0.01, 5]** | **bottomup** | **erm** |
| **Australia** | 1.32 | -4.43 | 5.31 | -9.98 | -11.29 | -48.84 | -47.88 |
| **State** | 0.01 | -1.35 | 0.45 | -10.34 | -9.46 | -18.28 | -50.13 |
| **Gender** | 2.63 | 2.51 | -0.49 | -4.46 | -6.37 | -11.54 | -46.11 |
| **Legal** | 4.27 | 4.5 | 3.44 | -2.89 | -4.25 | -1.71 | -47.44 |
| **Indigenous** | 4.63 | 4.62 | 1.75 | -1.99 | -2.08 | 0.0 | -53.24 |
| **Overall** | 3.19 | 2.56 | 1.88 | -4.62 | -5.32 | -9.24 | -49.5 |

Training at each iteration – median across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintshrink** | **case1\_lambda\_1** | **bottomup** | **wls** | **erm** |
| **Australia** | 4.82 | 32.96 | 3.44 | -24.57 | 16.79 | -33.02 |
| **State** | -5.86 | -9.0 | -12.85 | -23.13 | -9.64 | -71.14 |
| **Gender** | 4.87 | 8.41 | 3.52 | -5.79 | 8.66 | -40.7 |
| **Legal** | 4.34 | 3.08 | 1.59 | 1.01 | 1.43 | -33.08 |
| **Indigenous** | -2.32 | 2.04 | -8.43 | 0.0 | 2.28 | -44.62 |
| **Overall** | 5.9 | 5.81 | 4.33 | 3.92 | 3.16 | -36.72 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **case1\_lambda\_[0.01, 5]** | **mintshrink** | **wls** | **ols** | **bottomup** | **erm** |
| **Australia** | 65.95 | 22.05 | 4.51 | 6.0 | -53.67 | -2.02 |
| **State** | -13.37 | -8.93 | -6.03 | -9.19 | -32.13 | -49.65 |
| **Gender** | 2.66 | 5.16 | 6.76 | 2.39 | -11.24 | -38.24 |
| **Legal** | 2.01 | -0.17 | 1.07 | 1.31 | -2.02 | -26.03 |
| **Indigenous** | 1.59 | 4.01 | 3.28 | 2.44 | 0.0 | -20.46 |
| **Overall** | 5.33 | 2.83 | 2.08 | 0.84 | -11.36 | -27.4 |

## ETS

### ETS - MSE

Training at every 10th iteration – mean across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **ols** | **wls** | **bottomup** | **erm** | **case1\_lambda\_[0.1, 0.9]** |
| **Australia** | 4.21 | 4.67 | -5.3 | -105.84 | -2884.11 | -18110.55 |
| **State** | 8.2 | 4.46 | 4.52 | -40.55 | -2018.79 | -5729.73 |
| **Gender** | 13.42 | 8.12 | 10.87 | -30.36 | -1893.14 | -5309.36 |
| **Legal** | 11.68 | 9.0 | 9.88 | -6.36 | -1538.18 | -2578.87 |
| **Indigenous** | 12.27 | 9.78 | 9.84 | 0.0 | -1412.38 | -2149.98 |
| **Overall** | 9.09 | 6.78 | 4.34 | -46.57 | -2084.76 | -8362.88 |

Training at each iteration – mean across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **ols** | **wls** | **case2\_lambda\_[0.1, 0.9]** | **bottomup** | **erm** |
| **Australia** | 6.5 | 4.67 | -4.72 | 15.8 | -105.84 | -154.59 |
| **State** | 8.16 | 4.46 | 2.95 | -14.85 | -40.55 | -122.43 |
| **Gender** | 13.27 | 8.12 | 9.29 | -5.07 | -30.36 | -133.07 |
| **Legal** | 13.38 | 9.0 | 9.27 | -4.23 | -6.36 | -140.22 |
| **Indigenous** | 13.74 | 9.78 | 9.64 | 1.41 | 0.0 | -133.49 |
| **Overall** | 10.31 | 6.78 | 3.82 | 0.68 | -46.57 | -139.15 |

Training at each iteration – median across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **ols** | **case1\_lambda\_[0.01, 5]** | **bottomup** | **erm** |
| **Australia** | 61.11 | 34.72 | 29.99 | 6.68 | -85.07 | -53.34 |
| **State** | 1.7 | 12.45 | 3.34 | 3.3 | -16.93 | -54.61 |
| **Gender** | 5.15 | 16.15 | -3.0 | 1.34 | -13.71 | -73.12 |
| **Legal** | 6.59 | 7.0 | 2.14 | 8.65 | 3.27 | -56.51 |
| **Indigenous** | 7.16 | 8.2 | 4.48 | -3.74 | 0.0 | -72.32 |
| **Overall** | 22.76 | 13.81 | 11.81 | -3.89 | -8.87 | -69.48 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **case1\_lambda\_[0.01, 5]** | **mintshrink** | **wls** | **ols** | **bottomup** | **erm** |
| **Australia** | 78.85 | 44.46 | 27.49 | 11.46 | -39.47 | 46.27 |
| **State** | 21.35 | 2.12 | -0.94 | -6.94 | -25.33 | -193.36 |
| **Gender** | 19.98 | 1.81 | 1.34 | -3.1 | -17.07 | -258.45 |
| **Legal** | 19.78 | 17.63 | 9.19 | 2.95 | 5.07 | -79.56 |
| **Indigenous** | 9.61 | 13.0 | 3.87 | -1.42 | 0.0 | -68.78 |
| **Overall** | 37.81 | 21.73 | 12.01 | 2.94 | -17.34 | -72.38 |

### ETS - MAE

Training at every 10th iteration – mean across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **ols** | **bottomup** | **erm** | **case1\_lambda\_[0.1, 0.9]** |
| **Australia** | 12.88 | 6.56 | 4.82 | -43.41 | -433.72 | -1089.23 |
| **State** | 0.77 | -1.1 | 0.65 | -16.76 | -352.53 | -585.11 |
| **Gender** | 2.47 | 1.04 | -1.08 | -14.21 | -323.75 | -520.72 |
| **Legal** | 3.68 | 3.77 | 2.7 | -2.57 | -292.72 | -331.13 |
| **Indigenous** | 5.98 | 5.32 | 3.09 | 0.0 | -290.57 | -274.32 |
| **Overall** | 4.54 | 3.25 | 2.01 | -9.71 | -319.57 | -454.6 |

Training at each iteration – mean across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **ols** | **case2\_lambda\_[0.1, 0.9]** | **bottomup** | **erm** |
| **Australia** | 15.21 | 8.1 | 4.82 | 5.88 | -43.41 | -40.73 |
| **State** | 2.67 | -1.2 | 0.65 | -4.95 | -16.76 | -52.12 |
| **Gender** | 4.37 | 1.08 | -1.08 | -4.05 | -14.21 | -54.53 |
| **Legal** | 5.89 | 4.11 | 2.7 | -4.03 | -2.57 | -53.72 |
| **Indigenous** | 7.34 | 5.51 | 3.09 | -2.36 | 0.0 | -52.67 |
| **Overall** | 6.39 | 3.52 | 2.01 | -2.78 | -9.71 | -52.13 |

Training at each iteration – median across windows

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **ols** | **case1\_lambda\_[0.01, 5]** | **bottomup** | **erm** |
| **Australia** | 37.64 | 18.98 | 16.08 | 3.32 | -36.41 | -24.21 |
| **State** | 3.43 | 7.76 | -0.03 | -0.95 | -1.69 | -26.13 |
| **Gender** | 4.02 | 7.78 | -1.54 | 1.6 | -6.91 | -27.28 |
| **Legal** | 4.61 | 3.98 | 3.09 | -1.06 | 2.63 | -27.83 |
| **Indigenous** | 1.93 | 2.0 | -1.58 | -6.07 | 0.0 | -41.51 |
| **Overall** | 6.43 | 4.88 | -0.54 | -0.65 | -1.21 | -27.41 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **case1\_lambda\_1** | **mintshrink** | **wls** | **ols** | **bottomup** | **erm** |
| **Australia** | 61.09 | 25.92 | 14.31 | 5.23 | -19.75 | 30.25 |
| **State** | 12.22 | 2.16 | 1.22 | -3.16 | -4.95 | -39.3 |
| **Gender** | 7.48 | -1.01 | -2.13 | -8.87 | -6.78 | -49.01 |
| **Legal** | 9.31 | 7.01 | 1.56 | -1.98 | -2.23 | -34.94 |
| **Indigenous** | 3.75 | 5.82 | 1.72 | -0.78 | 0.0 | -30.39 |
| **Overall** | 12.93 | 6.57 | 2.29 | -2.1 | -4.41 | -29.68 |

# Labour

## ARIMA

### MSE

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **ols** | **wls** | **bottomup** | **mintsample** | **erm** | **case1\_lambda\_[1, 4]** |
| **Total Employees** | -0.19 | 3.53 | -10.01 | -36.14 | -54.93 | -1198.93 | -73639.06 |
| **Main Occupation** | 4.79 | 4.57 | 3.23 | -3.31 | -46.91 | -717.9 | -9851.65 |
| **Employment Status** | 6.89 | 4.55 | 4.95 | 0.67 | -51.51 | -721.96 | -5501.62 |
| **Gender** | 3.9 | 2.27 | 2.64 | 0.0 | -53.97 | -686.01 | -3578.87 |
| **Overall** | 4.35 | 3.8 | 1.36 | -6.8 | -51.42 | -791.84 | -17681.76 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **bottomup** | **case2\_lambda\_[0.1, 0.9]** | **ols** | **mintsample** | **erm** |
| **Total Employees** | 45.98 | 43.9 | 58.97 | 41.71 | 11.42 | -11.4 | 27.47 |
| **Main Occupation** | -1.18 | -3.63 | -10.64 | -6.23 | -2.89 | -17.83 | -112.52 |
| **Employment Status** | 5.39 | 5.49 | 1.72 | 3.79 | 4.45 | -46.36 | -167.97 |
| **Gender** | 2.67 | 2.52 | 0.0 | 0.98 | 1.59 | -49.06 | -164.46 |
| **Overall** | 10.56 | 9.52 | 8.74 | 7.53 | 3.1 | -32.99 | -115.92 |

### MAE

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **ols** | **bottomup** | **mintsample** | **erm** | **case1\_lambda\_[1, 4]** |
| **Total Employees** | 6.91 | 3.92 | 3.01 | -4.46 | -23.52 | -210.6 | -2548.04 |
| **Main Occupation** | 2.79 | 2.04 | 2.53 | -0.65 | -21.76 | -166.55 | -694.84 |
| **Employment Status** | 3.4 | 2.75 | 2.38 | 1.01 | -26.51 | -169.08 | -494.27 |
| **Gender** | 2.22 | 1.67 | 1.35 | 0.0 | -24.4 | -165.92 | -382.14 |
| **Overall** | 3.03 | 2.24 | 2.05 | -0.13 | -24.41 | -170.05 | -632.95 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **wls** | **mintshrink** | **ols** | **bottomup** | **case1\_lambda\_[0.1, 0.9]** | **mintsample** | **erm** |
| **Total Employees** | 29.36 | 30.09 | 7.53 | 44.28 | 32.83 | 9.06 | 13.23 |
| **Main Occupation** | -3.86 | -3.34 | -3.29 | -7.97 | -8.43 | -17.21 | -58.22 |
| **Employment Status** | 2.81 | 1.63 | 3.88 | -0.9 | -1.51 | -26.62 | -62.09 |
| **Gender** | 0.99 | 0.04 | 1.57 | 0.0 | -0.26 | -27.88 | -67.67 |
| **Overall** | 2.78 | 2.17 | 1.77 | 1.51 | 0.22 | -22.39 | -57.64 |

## ETS

### MSE

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **ols** | **wls** | **bottomup** | **mintsample** | **erm** | **case1\_lambda\_[0.1, 0.9]** |
| **Total Employees** | 7.46 | 3.91 | 2.88 | -14.45 | -103.28 | -9904.66 | -75173.41 |
| **Main Occupation** | 1.54 | 1.61 | 1.03 | -3.47 | -115.96 | -15244.29 | -10316.04 |
| **Employment Status** | 4.9 | 3.2 | 3.71 | 2.5 | -99.16 | -12379.33 | -5778.02 |
| **Gender** | 1.07 | 0.42 | 0.03 | 0.0 | -116.41 | -11541.61 | -3781.49 |
| **Overall** | 3.45 | 2.18 | 1.9 | -2.76 | -108.89 | -12542.81 | -18673.49 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **case1\_lambda\_[0.01, 0.09]** | **wls** | **mintshrink** | **ols** | **bottomup** | **mintsample** | **erm** |
| **Total Employees** | 39.69 | 48.36 | 42.8 | 15.85 | 26.54 | 38.58 | -116.19 |
| **Main Occupation** | -1.45 | -6.63 | -6.89 | -5.35 | -14.66 | -79.96 | -155.6 |
| **Employment Status** | 12.59 | 4.99 | 5.45 | 4.19 | 4.73 | -43.55 | -108.06 |
| **Gender** | 4.14 | 1.26 | 2.12 | -0.34 | 0.0 | -51.33 | -139.12 |
| **Overall** | 9.58 | 5.89 | 5.52 | 1.64 | 0.5 | -46.21 | -131.01 |

### MAE

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **ols** | **bottomup** | **mintsample** | **erm** | **case1\_lambda\_[0.1, 0.9]** |
| **Total Employees** | 4.95 | 2.59 | 2.64 | -5.29 | -41.53 | -606.89 | -2647.04 |
| **Main Occupation** | 1.62 | 1.46 | 1.74 | -0.24 | -41.46 | -705.46 | -724.82 |
| **Employment Status** | 2.5 | 2.05 | 1.91 | 1.57 | -35.61 | -612.52 | -510.83 |
| **Gender** | 1.08 | 0.66 | 0.52 | 0.0 | -38.0 | -633.79 | -397.41 |
| **Overall** | 1.91 | 1.4 | 1.37 | 0.08 | -38.27 | -641.56 | -657.36 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **case1\_lambda\_[0.01, 0.09]** | **wls** | **mintshrink** | **bottomup** | **ols** | **mintsample** | **erm** |
| **Total Employees** | 35.52 | 30.52 | 25.2 | 25.27 | 8.35 | 33.17 | -31.24 |
| **Main Occupation** | 2.76 | -2.11 | -2.62 | -3.5 | -3.91 | -30.87 | -60.8 |
| **Employment Status** | 6.96 | 2.94 | 2.94 | 2.19 | 3.08 | -17.57 | -49.14 |
| **Gender** | 2.26 | 1.25 | 1.22 | 0.0 | 0.5 | -19.97 | -58.49 |
| **Overall** | 5.94 | 2.86 | 2.4 | 1.49 | 0.85 | -18.32 | -54.26 |

# Tourism

## ARIMA

### MSE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintsample** | **mintshrink** | **wls** | **bottomup** | **case1\_lambda\_[1, 4]** |
| **Australia** | -2.49 | -9.7 | -30.11 | -54.42 | -122.15 | -183.5 |
| **States** | 9.11 | 10.65 | 5.12 | -3.14 | -31.01 | -46.77 |
| **Regions** | 5.1 | 4.16 | 7.49 | 6.36 | 0.0 | -9.36 |
| **Overall** | 2.5 | -0.86 | -12.34 | -27.47 | -71.71 | -109.47 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **case2\_lambda\_[0.01, 0.09]** | **mintsample** | **ols** | **mintshrink** | **wls** | **bottomup** | **erm** |
| **Australia** | 12.96 | 6.31 | -2.3 | -11.48 | -25.31 | -51.96 | -75.45 |
| **States** | 18.3 | 20.17 | 11.27 | 12.69 | 5.47 | -12.66 | -22.86 |
| **Regions** | 7.47 | 10.18 | 3.87 | 8.15 | 5.23 | 0.0 | -38.98 |
| **Overall** | 13.63 | 10.47 | 2.09 | -2.5 | -13.13 | -34.67 | -56.74 |

### MAE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintsample** | **mintshrink** | **wls** | **bottomup** | **case1\_lambda\_[1, 4]** |
| **Australia** | -0.52 | -4.99 | -11.74 | -20.05 | -44.94 | -60.81 |
| **States** | 5.25 | 6.06 | 3.18 | -0.45 | -11.56 | -18.13 |
| **Regions** | 0.8 | 0.82 | 2.85 | 2.14 | 0.0 | -3.62 |
| **Overall** | 1.78 | 1.34 | 0.77 | -1.84 | -9.73 | -15.95 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **case1\_lambda\_[0.01, 0.09]** | **mintshrink** | **mintsample** | **wls** | **ols** | **bottomup** | **erm** |
| **Australia** | 9.21 | 5.94 | -2.82 | 4.73 | 0.75 | -7.77 | -31.61 |
| **States** | 12.01 | 10.3 | 9.94 | 8.14 | 6.67 | 3.45 | -11.82 |
| **Regions** | 1.86 | 2.94 | 3.85 | 1.39 | -0.09 | 0.0 | -20.42 |
| **Overall** | 5.97 | 5.49 | 4.27 | 3.83 | 1.9 | -0.5 | -20.16 |

## ETS

### MSE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintsample** | **mintshrink** | **wls** | **bottomup** | **case1\_lambda\_[1, 4]** |
| **Australia** | -1.23 | -1.35 | -23.67 | -29.27 | -59.15 | -131.81 |
| **States** | 3.37 | -0.14 | -4.28 | -6.43 | -20.4 | -49.64 |
| **Regions** | 3.33 | -1.26 | 3.73 | 3.37 | 0.0 | -15.98 |
| **Overall** | 0.8 | -0.99 | -13.78 | -17.56 | -38.68 | -89.94 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **case2\_lambda\_[0.1, 0.9]** | **mintsample** | **ols** | **mintshrink** | **wls** | **bottomup** | **erm** |
| **Australia** | 14.16 | 11.91 | 0.41 | -4.87 | -7.3 | -20.72 | -113.74 |
| **States** | 5.51 | 3.55 | -0.75 | -4.54 | -6.1 | -15.75 | -124.26 |
| **Regions** | 3.45 | 2.03 | 2.78 | 2.57 | 2.33 | 0.0 | -105.37 |
| **Overall** | 11.06 | 8.96 | 0.4 | -3.99 | -5.99 | -17.36 | -115.2 |

### MAE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **ols** | **mintsample** | **mintshrink** | **wls** | **bottomup** | **case1\_lambda\_[1, 4]** |
| **Australia** | -0.49 | -0.93 | -9.62 | -11.81 | -25.39 | -49.03 |
| **States** | 2.14 | 0.69 | -1.67 | -2.75 | -9.43 | -19.85 |
| **Regions** | 0.6 | -1.92 | 1.25 | 1.1 | 0.0 | -5.72 |
| **Overall** | 0.82 | -1.09 | -1.23 | -1.94 | -6.46 | -16.24 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **case2\_lambda\_[0.01, 5]** | **mintsample** | **mintshrink** | **wls** | **ols** | **bottomup** | **erm** |
| **Australia** | 12.14 | 6.61 | 4.71 | 4.6 | 0.59 | 3.03 | -38.68 |
| **States** | 2.28 | 2.56 | -1.33 | -1.73 | -0.9 | -4.51 | -47.43 |
| **Regions** | 1.44 | 0.56 | 0.75 | 0.54 | 0.81 | 0.0 | -48.17 |
| **Overall** | 3.91 | 2.36 | 1.02 | 0.78 | 0.3 | -0.57 | -45.98 |

# Wikipedia

## ARIMA

**Fitted Values as Input**

### MSE – training at every 10th iteration

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **bottomup** | **ols** | **case1\_lambda\_[0.01, 0.09]** | **erm** |
| **Total** | 14.96 | 11.85 | 7.61 | 4.37 | 12.1 | -923.18 |
| **Access** | 18.98 | 14.15 | 8.59 | 6.29 | 12.31 | -600.83 |
| **Agent** | 8.9 | 3.62 | -2.67 | -2.54 | 3.64 | -620.11 |
| **Language** | 14.41 | 8.75 | 3.9 | 1.55 | 4.62 | -643.66 |
| **Purpose** | 11.02 | 7.05 | 6.26 | 0.14 | -42.55 | -921.73 |
| **Article** | 4.35 | 1.34 | 0.0 | 0.02 | -69.99 | -849.05 |
| **Overall** | 13.1 | 8.71 | 4.39 | 2.08 | -4.97 | -744.63 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **case2\_lambda\_[0.01, 0.09]** | **mintshrink** | **bottomup** | **wls** | **ols** | **erm** |
| **Total** | 45.04 | 33.16 | 16.7 | 20.36 | 5.37 | -182.29 |
| **Access** | 34.74 | 25.89 | 28.1 | 24.74 | 11.47 | -165.56 |
| **Agent** | 8.04 | 0.45 | 0.88 | -2.14 | -9.79 | -244.83 |
| **Language** | 33.79 | 27.53 | 23.26 | 19.6 | 5.35 | -195.86 |
| **Purpose** | 11.82 | 13.35 | 12.85 | 7.93 | -0.39 | -301.0 |
| **Article** | 0.89 | 4.29 | 0.0 | -2.84 | -1.36 | -355.78 |
| **Overall** | 26.43 | 19.75 | 15.98 | 13.9 | 2.8 | -218.81 |

### MAE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **bottomup** | **ols** | **case2\_lambda\_[0.01, 0.09]** | **erm** |
| **Total** | 6.2 | 4.64 | 3.7 | 1.33 | -14.1 | -207.32 |
| **Access** | 8.8 | 6.15 | 1.03 | -0.55 | -2.34 | -164.29 |
| **Agent** | 4.37 | 1.74 | -3.49 | -6.93 | -8.64 | -175.93 |
| **Language** | 5.35 | 2.56 | 0.2 | -5.58 | -8.36 | -176.39 |
| **Purpose** | 3.7 | 1.72 | 0.24 | -14.71 | -15.88 | -213.48 |
| **Article** | 1.54 | 0.53 | 0.0 | -12.38 | -21.41 | -236.03 |
| **Overall** | 4.09 | 2.2 | 0.09 | -8.69 | -14.04 | -204.83 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **bottomup** | **mintshrink** | **case2\_lambda\_[0.01, 0.09]** | **wls** | **ols** | **erm** |
| **Total** | 18.18 | 19.02 | 31.54 | 14.35 | 5.49 | -57.75 |
| **Access** | 23.09 | 16.94 | 23.09 | 14.69 | -0.41 | -57.79 |
| **Agent** | 9.86 | 3.2 | 6.39 | 0.94 | -14.78 | -78.67 |
| **Language** | 9.28 | 9.91 | 8.96 | 4.08 | -7.87 | -81.97 |
| **Purpose** | 5.02 | 2.57 | -3.58 | 1.07 | -15.15 | -113.08 |
| **Article** | 0.0 | 0.38 | -5.57 | -0.98 | -13.99 | -117.06 |
| **Overall** | 7.99 | 6.24 | 5.18 | 3.68 | -10.01 | -94.01 |

## ETS

### MSE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **bottomup** | **ols** | **case2\_lambda\_[1, 4]** | **erm** |
| **Total** | 0.78 | 0.37 | -0.42 | -1.87 | -72.13 | -4629.76 |
| **Access** | 5.46 | 5.19 | 4.94 | 2.82 | -33.71 | -3481.03 |
| **Agent** | 5.78 | 5.49 | 4.55 | 2.34 | -32.79 | -3518.68 |
| **Language** | 1.9 | 1.55 | 0.75 | -0.22 | -44.75 | -3428.47 |
| **Purpose** | -1.7 | -1.93 | -3.31 | -1.07 | -57.88 | -4194.51 |
| **Article** | 1.8 | 1.3 | 0.0 | 0.54 | -67.11 | -4161.6 |
| **Overall** | 2.74 | 2.41 | 1.57 | 0.56 | -49.72 | -3881.1 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling windows)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **bottomup** | **wls** | **ols** | **case2\_lambda\_1** | **erm** |
| **Total** | 21.74 | 28.98 | 20.14 | 4.24 | 14.17 | -858.75 |
| **Access** | 13.41 | 11.65 | 10.79 | 5.63 | 9.29 | -436.19 |
| **Agent** | -8.68 | -14.63 | -11.66 | -8.34 | -18.34 | -616.72 |
| **Language** | 6.54 | 2.3 | 4.12 | 2.78 | -2.23 | -737.14 |
| **Purpose** | -2.18 | -6.45 | -4.34 | -2.05 | -11.76 | -726.46 |
| **Article** | 4.84 | 0.0 | 3.8 | 1.18 | -0.13 | -608.2 |
| **Overall** | 6.13 | 3.89 | 3.82 | 0.57 | -1.22 | -639.7 |

### MAE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **wls** | **mintshrink** | **bottomup** | **ols** | **case1\_lambda\_[0.1, 0.9]** | **erm** |
| **Total** | 1.26 | 1.4 | 0.96 | -0.76 | -32.95 | -558.68 |
| **Access** | 2.81 | 2.69 | 2.06 | 0.63 | -19.18 | -456.07 |
| **Agent** | 2.95 | 2.83 | 1.96 | -0.42 | -16.73 | -463.58 |
| **Language** | 0.91 | 0.8 | 0.6 | -1.92 | -16.25 | -481.69 |
| **Purpose** | 0.5 | 0.47 | -0.29 | -4.7 | -16.65 | -516.87 |
| **Article** | 0.08 | 0.1 | 0.0 | -4.55 | -19.33 | -521.76 |
| **Overall** | 1.02 | 0.99 | 0.57 | -2.82 | -19.12 | -503.64 |

#### One-window only (errors are only calculated for the immediate 1-step ahead rolling window only)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **mintshrink** | **wls** | **bottomup** | **ols** | **case2\_lambda\_1** | **erm** |
| **Total** | 12.2 | 10.86 | 15.34 | 1.45 | 8.68 | -194.36 |
| **Access** | 10.22 | 8.98 | 7.43 | 3.28 | 1.85 | -140.94 |
| **Agent** | -1.09 | -2.5 | -6.23 | -4.57 | -12.12 | -162.42 |
| **Language** | 0.68 | -0.4 | -3.73 | -1.08 | -6.85 | -180.84 |
| **Purpose** | -0.86 | -1.39 | -2.91 | -5.52 | -6.42 | -172.07 |
| **Article** | 1.37 | 1.13 | 0.0 | -4.24 | -0.9 | -153.76 |
| **Overall** | 2.32 | 1.57 | 0.07 | -2.82 | -3.18 | -163.52 |

# Running Time for ML method

This table shows the average running time per rolling window when training and executing a trained ML reconciliation method. The average training time also includes the hyper-parameter tuning time.

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset** | **Base Model** | **Average training time** | **Average execution time once trained** |
| Prison | ARIMA | 25.47 minutes | 0.024 minutes |
|  | ETS | 22.22 minutes | 0.024 minutes |
| Labour | ARIMA | 15.65 minutes | 0.023 minutes |
|  | ETS | 15.09 minutes | 0.023 minutes |
| Tourism | ARIMA | 11.46 minutes | 0.028 minutes |
|  | ETS | 11.79 minutes | 0.023 minutes |
| Wikipedia | ARIMA | 77.79 minutes | 0.033 minutes |
|  | ETS | 72.34 minutes | 0.035 minutes |