TÉCNICO LISBOA

DS 2025

Data Science Lab 5

Team nr: 2	**Student 1:** Åmund Grimstad	**IST nr:** 1116675
	Student 2: Arthur de Arruda Chau	**IST nr:** 1116090
	Student 3: Benjamin Raymond Kuhn	**IST nr:** 1115778
	Student 4: João Rafael Freitas Lourenço	**IST nr:** 425699

TIME SERIES ANALYSIS

Traffic Time Series

DATA PROFILING

Data Dimensionality and Granularity

Figure 1: Traffic Time Series at the most granular detail

Figure 2: Traffic Time Series at the second chosen granularity

Figure 3: Traffic Time Series at the third chosen granularity

Data Distribution

Figure 4: Boxplot(s) for Traffic Time Series

Figure 5: Histogram(s) for Traffic Time Series

Figure 6: Autocorrelation lag-plots for original Traffic Time Series

Figure 7: Autocorrelation correlogram for original Traffic Time Series

Data Stationarity

Figure 8: Components study for Traffic Time Series

Figure 9: Stationarity study for Traffic Time Series

DATA TRANSFORMATION

Aggregation

Figure 10: Forecasting plots after different aggregations on Traffic Time Series

Figure 11: Forecasting results after different aggregations on Traffic Time Series

Differentiation

Figure 12: Forecasting plots after first and second differentiation of Traffic Time Series

Figure 13: Forecasting plots after first and second differentiation of Traffic Time Series

Smoothing

Figure 14: Forecasting plots after different smoothing parameterisations on Traffic Time Series

Figure 15: Forecasting results after different smoothing parameterisations on Traffic Time Series

Scaling

Figure 16: Forecasting plots after different scaling parameterisations on Traffic Time Series

Figure 17: Forecasting results after different scaling parameterisations on Traffic Time Series

Inflation Rate Time Series

DATA PROFILING

Data Dimensionality and Granularity

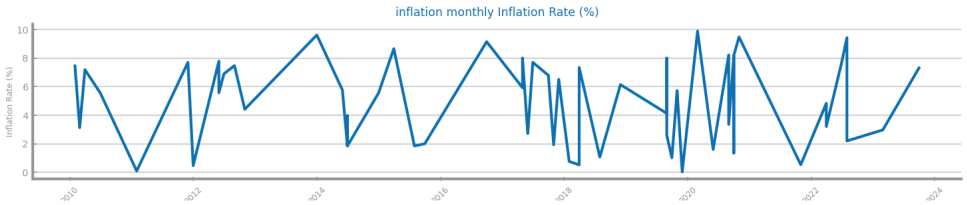

Figure 18: U.S. Inflation Rate dimensionality over time.

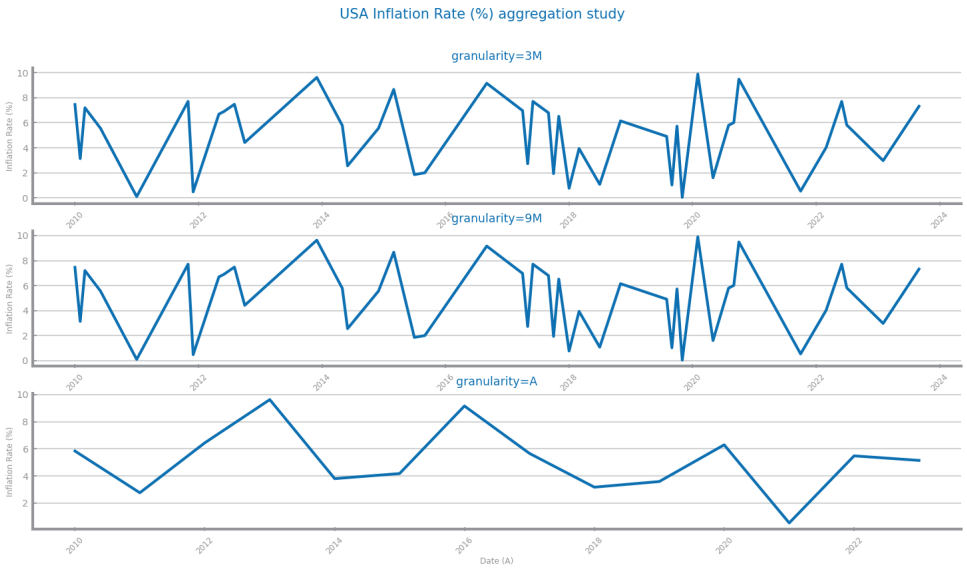

Figure 19: Inflation Rate Time Series at three different granularities

Data Distribution

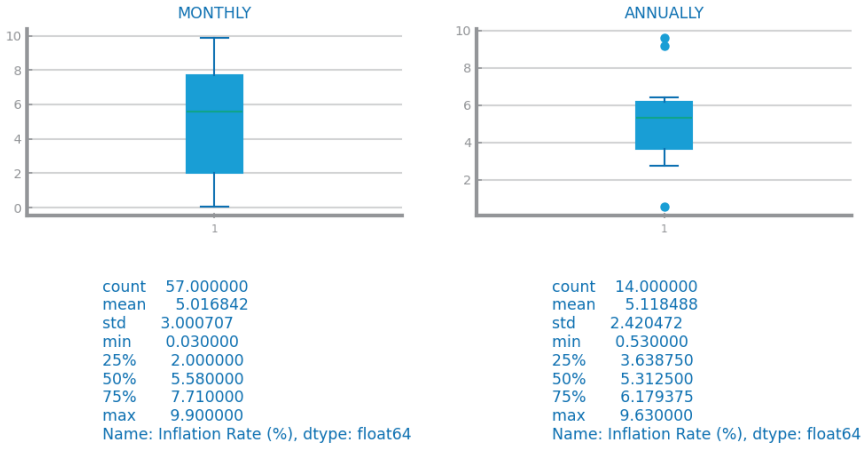

Figure 20: Boxplot(s) for Inflation Rate Time Series

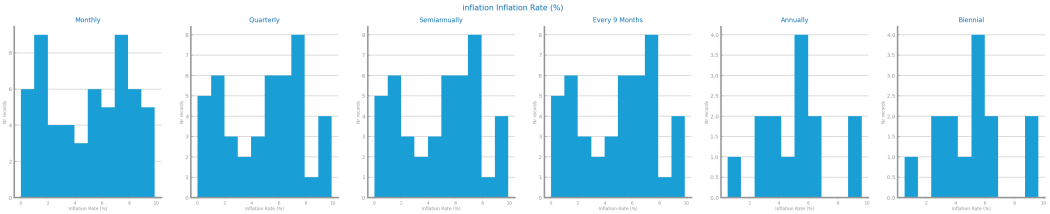

Figure 21: Histogram(s) for Inflation Rate Time Series

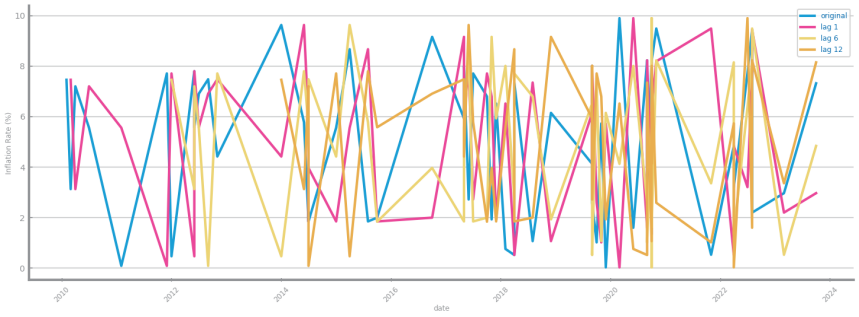

Figure 22: Autocorrelation lag-plots for original Inflation Rate Time Series

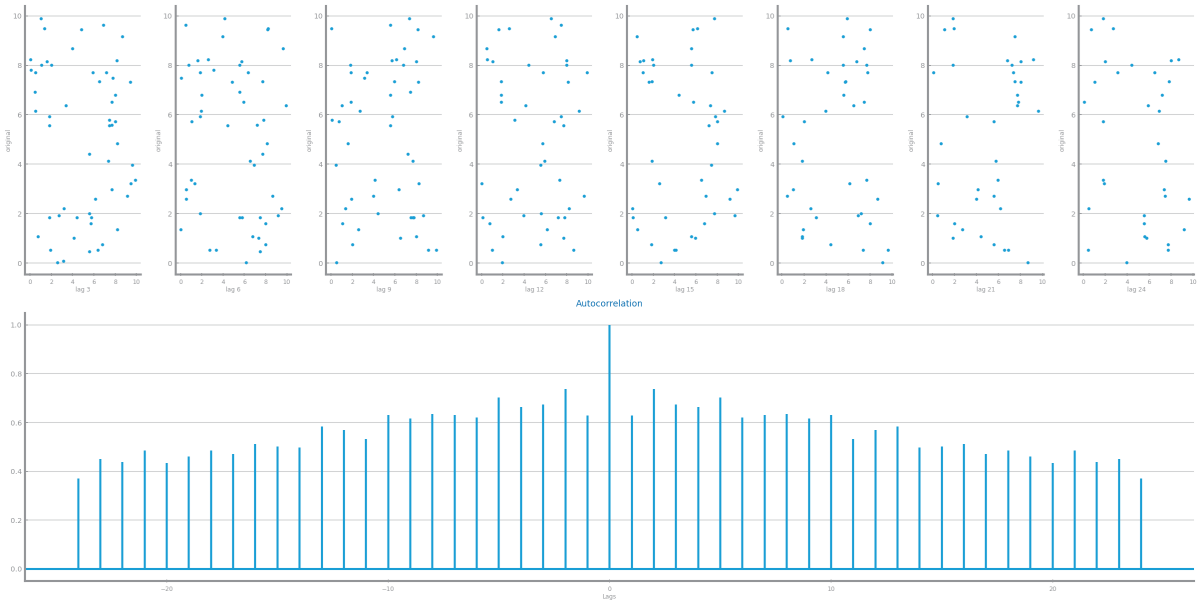

Figure 23: Autocorrelation correlogram for original Inflation Rate Time Series

Data Stationarity

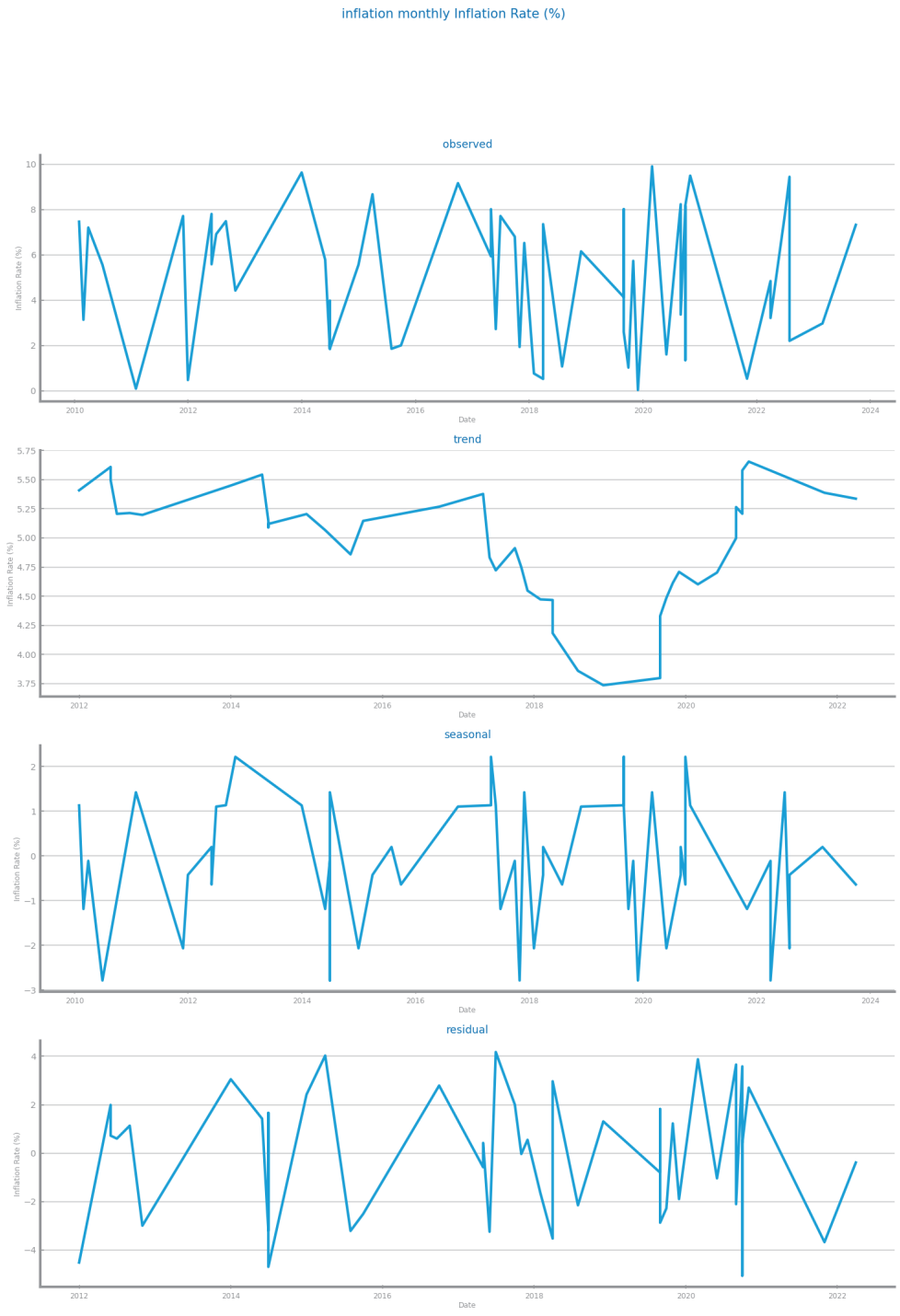

Figure 24: Components study for Inflation Rate Time Series

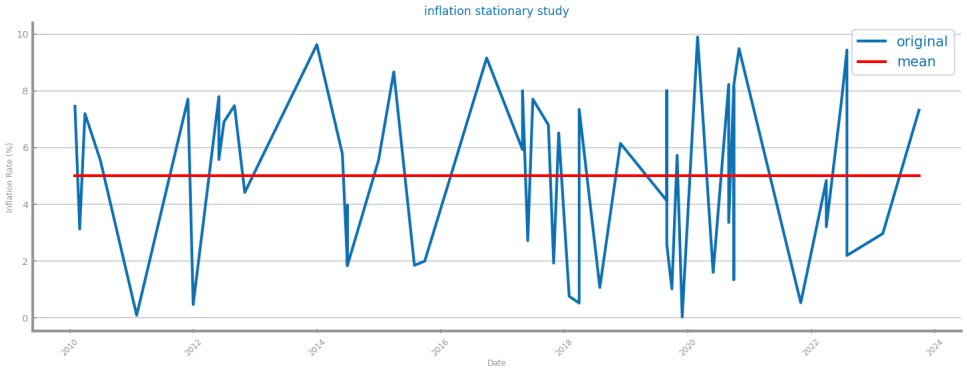

Figure 25: Stationarity study for Inflation Rate Time Series

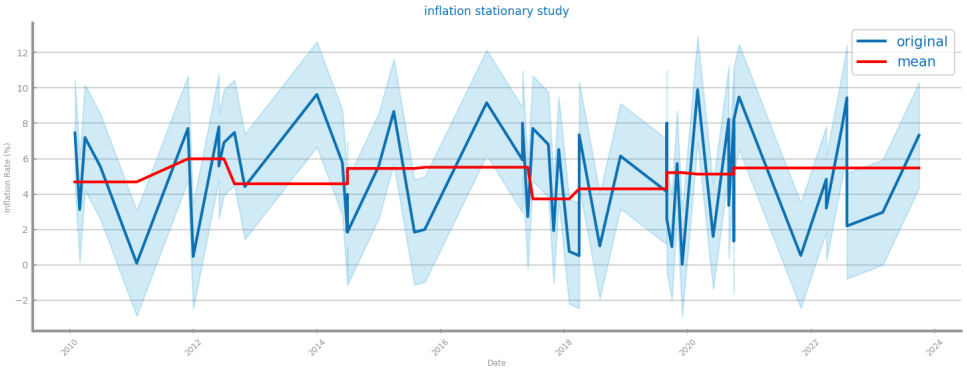

Figure 26: Stationarity study with bins for Inflation Rate Time Series

```
ADF Statistic: -10.410
p-value: 0.000
Critical Values:
        1%: -3.553
        5%: -2.915
        10%: -2.595
The series is stationary
```

Figure 27: Augmented Dickey-Fuller test results for Inflation Rate Time Series

DATA TRANSFORMATION

Aggregation

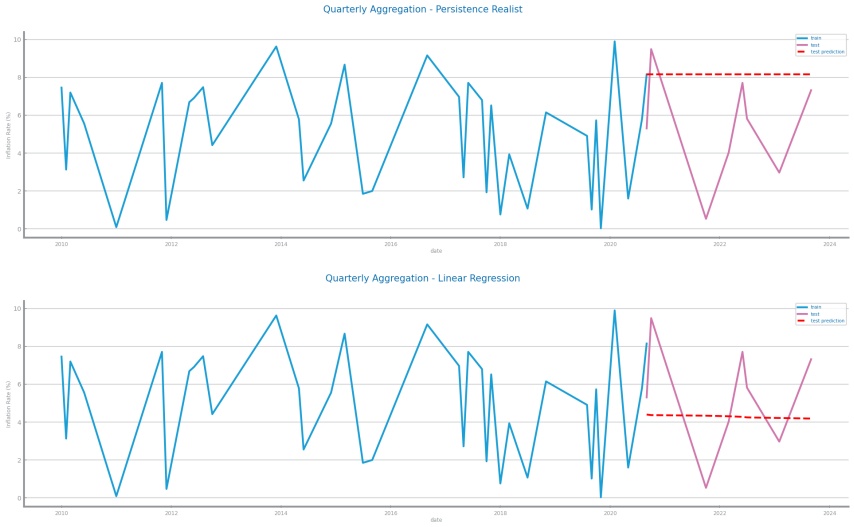

Figure 28: Forecasting predictions from Persistence Realist and Linear Regression on **Quarterly Inflation Aggregation**

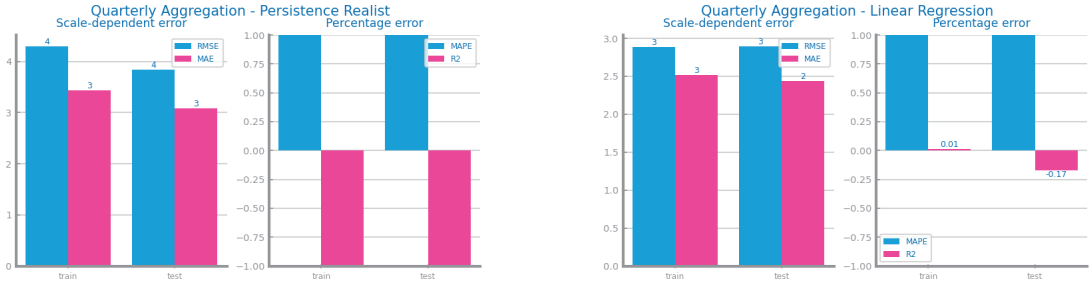

Figure 29: Performance of Persistence Realist and Linear Regression on **Quarterly Inflation Aggregation**

Figure 30: Forecasting predictions from Persistence Realist and Linear Regression on **Semi-Annual Inflation Aggregation**

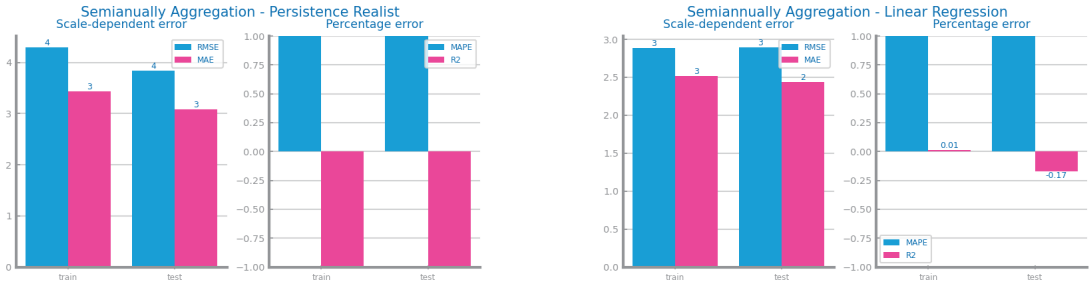

Figure 31: Performance of Persistence Realist and Linear Regression on **Semi-Annual Inflation Aggregation**

Figure 32: Forecasting predictions from Persistence Realist and Linear Regression on **Annual Inflation Aggregation**

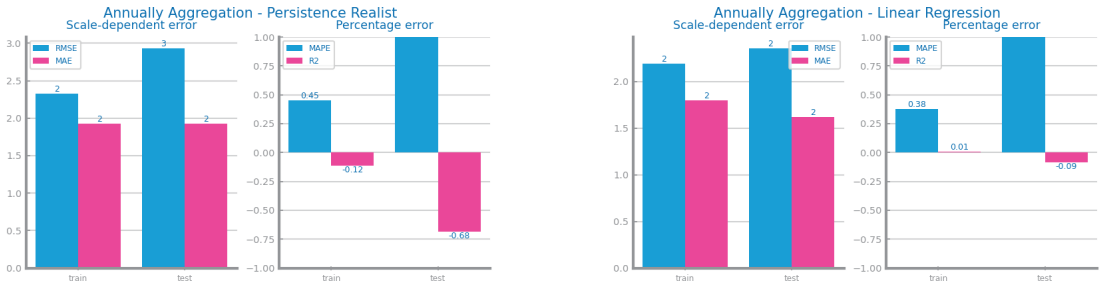

Figure 33: Performance of Persistence Realist and Linear Regression on **Annual Inflation Aggregation**

For all aggregation techniques, the Linear Regression model yields a lower RMSE than the Persisent Realist Model for the test set. Among the Linear Regression models, the Annual aggregation yields the lowest RSME for the test set (1.6).

Differentiation

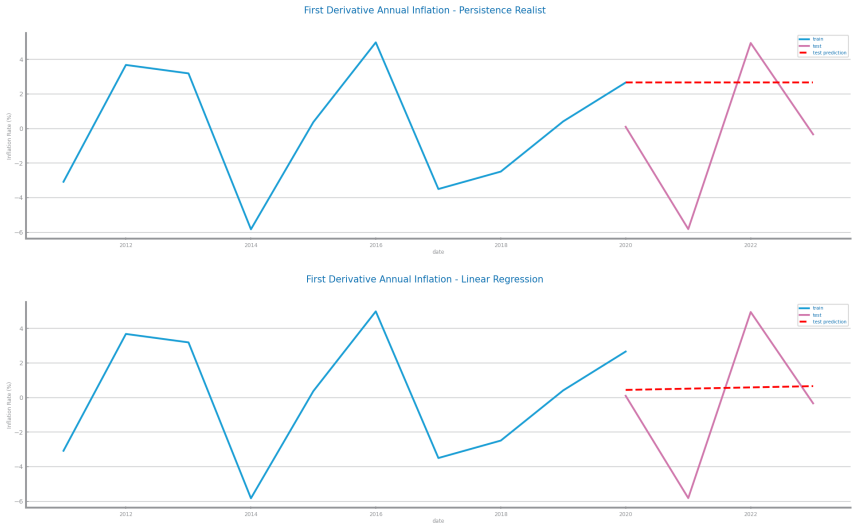

Figure 34: Forecasting predictions from Persistence Realist and Linear Regression on **Annual Inflation Differentiation (1st Order)**

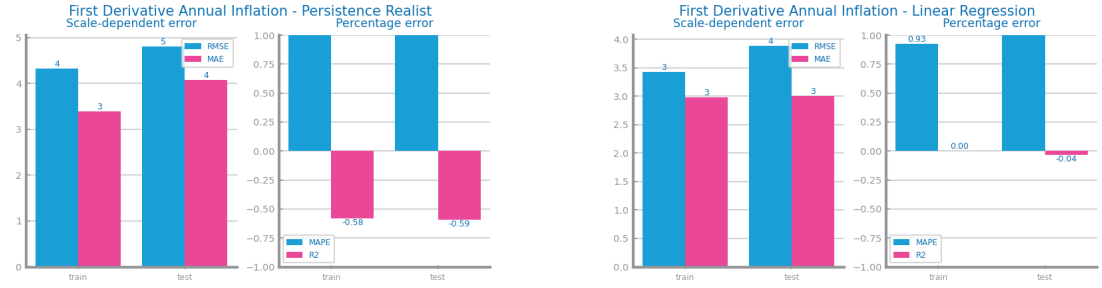

Figure 35: Performance of Persistence Realist and Linear Regression on **Annual Inflation Differentiation (1st Order)**

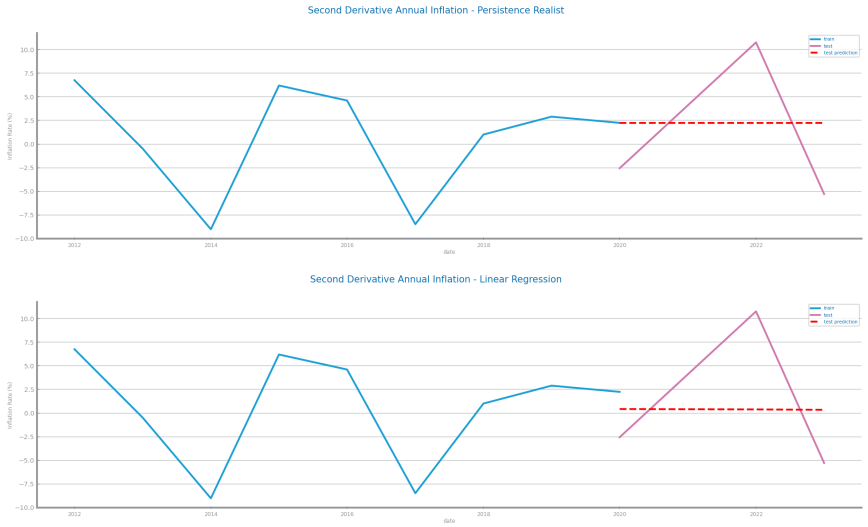

Figure 36: Forecasting predictions from Persistence Realist and Linear Regression on **Annual Inflation Differentiation (2nd Order)**

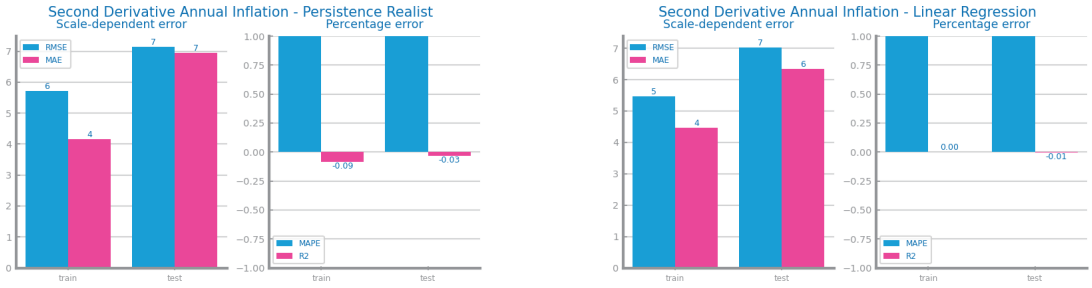

Figure 37: Performance of Persistence Realist and Linear Regression on **Annual Inflation Differentiation (2nd Order)**

Figure 38: Forecasting predictions from Persistence Realist and Linear Regression on **Annual Inflation Differentiation with Lag 2**

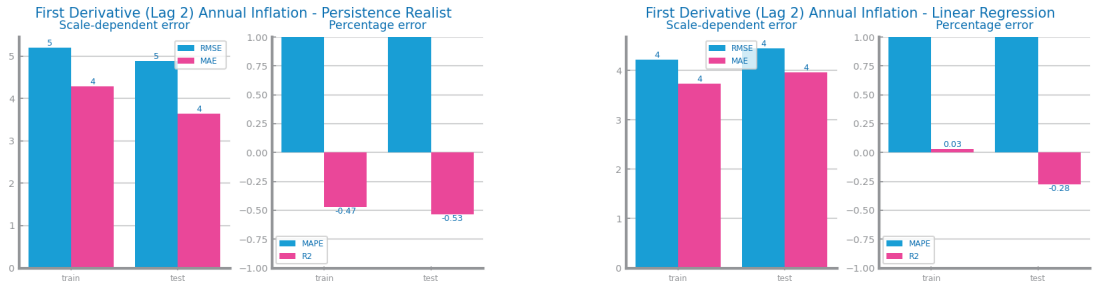

Figure 39: Performance of Persistence Realist and Linear Regression on **Annual Inflation Differentiation with Lag 2**

For diff_1 and diff_2, the Linear Regression Model yields a lower RMSE for the test set than the Persistent Realist models. For the diff_lag2 treatment, the Persistent Realist has a lower RMSE than the Linear Regression Model. But overall, the best model, a Linear Regression Model trained on the diff_1 treatment, is worse than the best model without differential treatment. So, we do not differentiate.

Smoothing

Figure 40: Forecasting plots after different smoothing parameterisations on Inflation Rate Time Series

Figure 41: Forecasting results after different smoothing parameterisations on Inflation Rate Time Series

Scaling

Figure 42: Forecasting plots after different scaling parameterisations on Inflation Rate Time Series

Figure 43: Forecasting results after different scaling parameterisations on Inflation Rate Time Series