Time Series Analysis
Spring 2026
Assignment 1

Students:	Linus San Tommaso Fabro Christian Alexander Haborg Magnus Elgaard Bennetts
Department:	Applied Mathematics and Computer Science
Supervisor at DTU:	Christian Ankerstjerne Thilker
Submission Date:	March 9, 2026

Table of Contents

1 Plot the Data

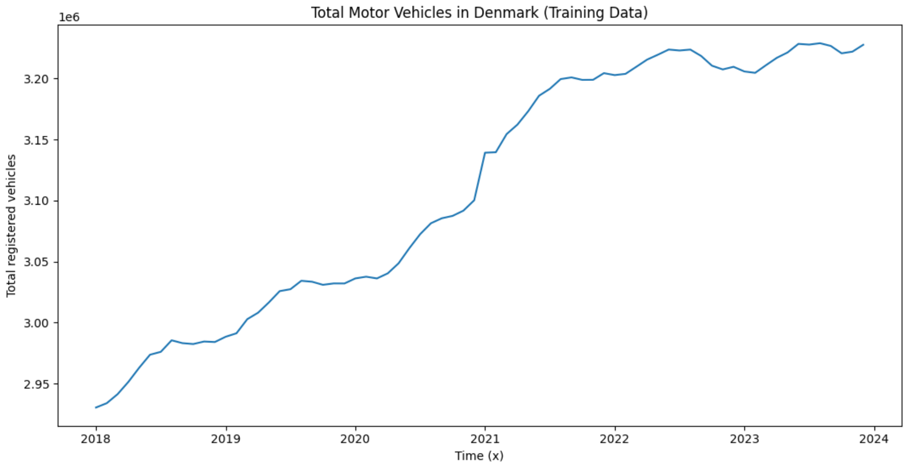

Figure 1: Total number of registered motor vehicles in Denmark (Training data: 2018–2023).

1.1 Description

The time series shows a clear upward trend from 2018 to 2023. The total number of registered motor vehicles increases steadily over the entire period, rising from approximately 2.93 million in early 2018 to around 3.23 million by the end of 2023.

The growth is not perfectly linear. There are small short-term fluctuations around the trend, indicating minor cyclical or seasonal movements. Around 2020–2021 there appears to be a slightly stronger increase compared to earlier years. After 2022, the growth continues but at a somewhat slower and more stable pace.

In general, the dominant feature of the series is a persistent positive trend.

2 Linear Trend Model

We consider the linear trend model

$$Y_t = \theta_1 + \theta_2 x_t + \varepsilon_t, \tag{2.1}$$

where $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ are assumed i.i.d. for $t = 1, \ldots, N$.

Matrix form for the first three time points

For $t = 1, 2, 3$ the model can be written in matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon},$$

where

$$\mathbf{y} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix}, \quad \boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}.$$

Inserting the time values

The time variable is defined as

$$x_t = \text{year}_t + \frac{\text{month}_t - 1}{12}.$$

For the first three observations we obtain

$$x_1 = 2018, \quad x_2 = 2018.083, \quad x_3 = 2018.167.$$

Hence, the design matrix becomes

$$\mathbf{X} = \begin{pmatrix} 1 & 2018 \\ 1 & 2018.083 \\ 1 & 2018.167 \end{pmatrix}.$$

Inserting the first three observations

Using the first three observed values (rounded to three significant digits),

$$\mathbf{y} = \begin{pmatrix} 2.93 \times 10^6 \\ 2.94 \times 10^6 \\ 2.95 \times 10^6 \end{pmatrix},$$

the linear trend model for the first three time points can be written as

$$\begin{pmatrix} 2.93 \times 10^6 \\ 2.94 \times 10^6 \\ 2.95 \times 10^6 \end{pmatrix} = \begin{pmatrix} 1 & 2018 \\ 1 & 2018.083 \\ 1 & 2018.167 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}.$$

3 OLS - Global Linear Trend Model

3.1 Estimation of the Parameters

We consider the global linear trend model

$$Y_t = \theta_1 + \theta_2 x_t + \varepsilon_t, \qquad \varepsilon_t \sim \text{i.i.d. } (0, \sigma^2).$$

The parameters θ_1 (intercept) and θ_2 (slope) are estimated using Ordinary Least Squares (OLS). The OLS estimator is given by

$$\hat{\boldsymbol{\theta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y},$$

where X is the design matrix containing a column of ones (for the intercept) and the time variable x_t, and Y is the vector of observed values.
The estimated parameter values are

$$\hat{\theta}_1 = -110{,}355{,}428.10 \qquad \hat{\theta}_2 = 56{,}144.56$$

Interpretation: The slope $\hat{\theta}_2$ indicates that the number of registered vehicles increases on average by approximately 56,145 vehicles per year. The large negative intercept $\hat{\theta}_1$ is due to the scaling of the time variable and has no direct economic meaning.

3.2 Parameter Estimates and Standard Errors

The residuals of the model are defined as

$$\varepsilon = Y - \hat{Y},$$

where $\hat{Y} = X\hat{\boldsymbol{\theta}}$ are the fitted values. The estimated error variance is then

$$\hat{\sigma}^2 = \frac{\varepsilon^\top \varepsilon}{n - p},$$

where n is the number of observations and $p = 2$ is the number of estimated parameters (intercept and slope).
The variance-covariance matrix of the OLS estimator is

$$\widehat{\text{Var}}(\hat{\boldsymbol{\theta}}) = \hat{\sigma}^2 (X^\top X)^{-1},$$

and the standard errors of the individual parameter estimates are obtained as the square roots of the diagonal elements:

$$\widehat{\text{se}}(\hat{\theta}_1) = \sqrt{[\widehat{\text{Var}}(\hat{\boldsymbol{\theta}})]_{11}} = 3{,}593{,}581.12,$$

$$\widehat{\text{se}}(\hat{\theta}_2) = \sqrt{[\widehat{\text{Var}}(\hat{\boldsymbol{\theta}})]_{22}} = 1{,}778.16.$$

Interpretation: Both parameters are statistically significant. The slope $\hat{\theta}_2$ confirms a strong upward trend in vehicle registrations, while the large negative intercept $\hat{\theta}_1$ arises from the scaling of the time variable.

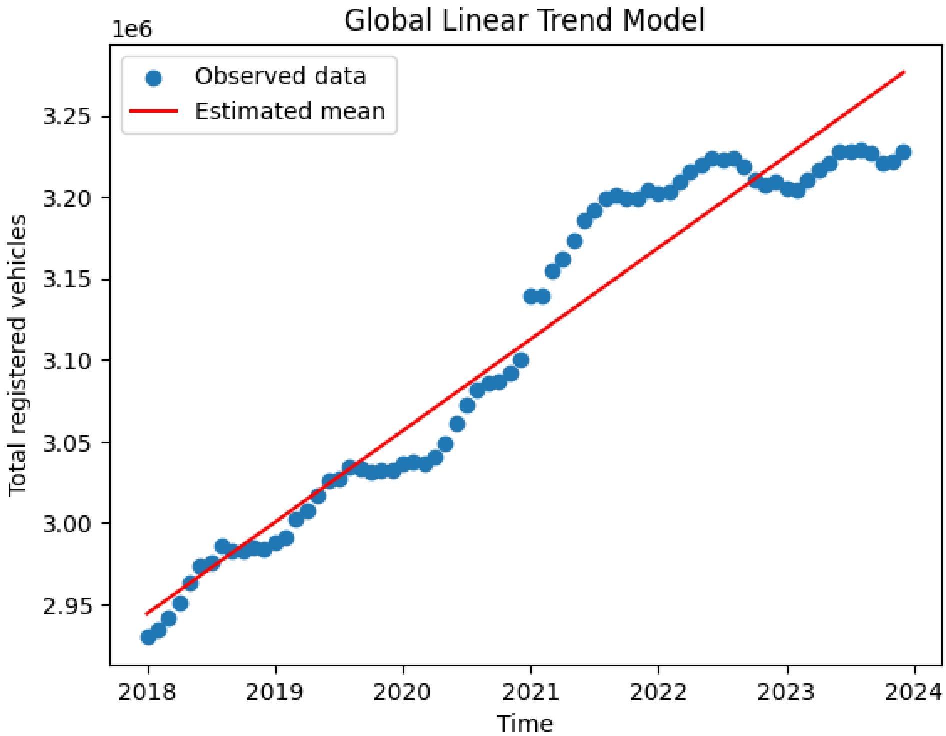

Figure 1: Observed number of registered vehicles (points) and estimated global linear trend (red line).

3.3 Forecast for 2024 with Prediction Intervals

The predicted number of registered vehicles for January to December 2024, along with the 95% prediction intervals, is shown in Table 3.3. The prediction intervals were calculated using:

$$\hat{Y}_{t+l} \pm t_{n-2,0.975}\sqrt{\hat{\sigma}^2\left[1 + x_{t+l}^\top (X^\top X)^{-1} x_{t+l}\right]}.$$

Month	Forecast	Lower PI	Upper PI
2024-01	3,281,154	3,227,579	3,334,728
2024-02	3,285,832	3,232,198	3,339,467
2024-03	3,290,511	3,236,815	3,344,208
2024-04	3,295,190	3,241,430	3,348,950
2024-05	3,299,869	3,246,044	3,353,693
2024-06	3,304,547	3,250,656	3,358,439
2024-07	3,309,226	3,255,267	3,363,185
2024-08	3,313,905	3,259,876	3,367,934
2024-09	3,318,583	3,264,483	3,372,683
2024-10	3,323,262	3,269,090	3,377,435
2024-11	3,327,941	3,273,694	3,382,188
2024-12	3,332,620	3,278,297	3,386,942

Table 3.1: Forecasted total registered vehicles for 2024 with 95% prediction intervals

3.4 Fitted Model and Forecasted Values

Figure 2 shows the OLS global linear trend fitted to the training data, along with the forecasted values for January to December 2024. The 95% prediction intervals are represented as the shaded area around the forecasted trend.

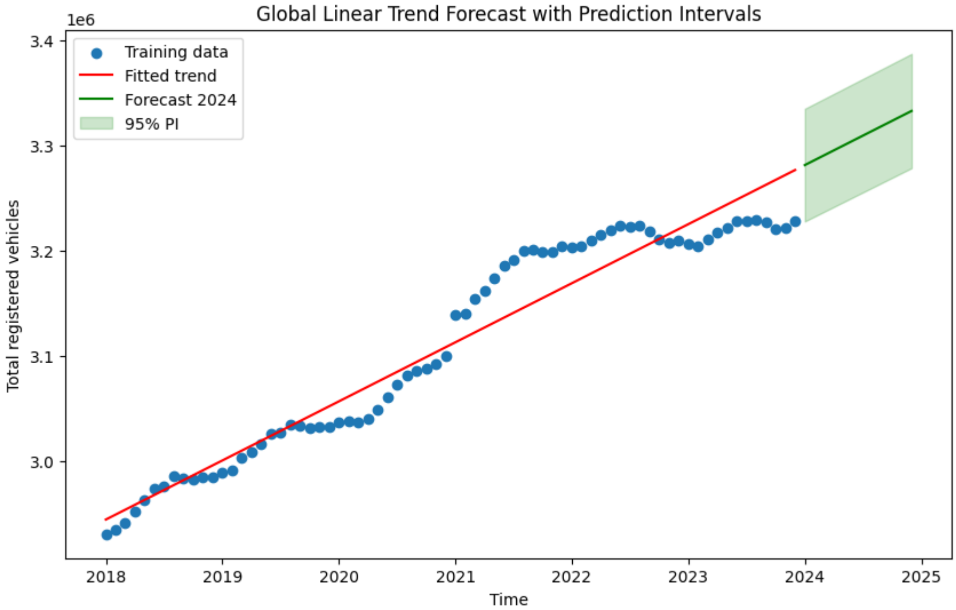

Figure 2: OLS fitted trend (red line), observed training data (points), forecast for 2024 (green line), and 95% prediction intervals (shaded area).

3.5 Forecast Evaluation

The forecast for 2024 based on the global linear trend model predicts a continued increase in the total number of registered vehicles, as shown in Table 3.3 and Figure 2.
While the model captures the overall upward trend, there are several limitations to consider:

- **Seasonality ignored:** Monthly fluctuations and seasonal effects are not accounted for, which may lead to systematic under- or overestimation in specific months.
- **Linear assumption:** The model assumes a constant growth rate. Any acceleration or deceleration in registrations is not captured.
- **Residual autocorrelation:** The model residuals may exhibit autocorrelation, violating the independence assumption of OLS and potentially underestimating uncertainty.

Conclusion:
The forecast provides a reasonable estimate of the long-term trend but may be less accurate for short-term monthly predictions. The 95% prediction intervals help account for this uncertainty.

3.6 Residual Diagnostics

The residuals of the OLS global linear trend model were examined to assess whether the model assumptions are fulfilled. Residuals are defined as

$$\varepsilon_t = Y_t - \hat{Y}_t,$$

where $\hat{Y}_t$ are the fitted values.

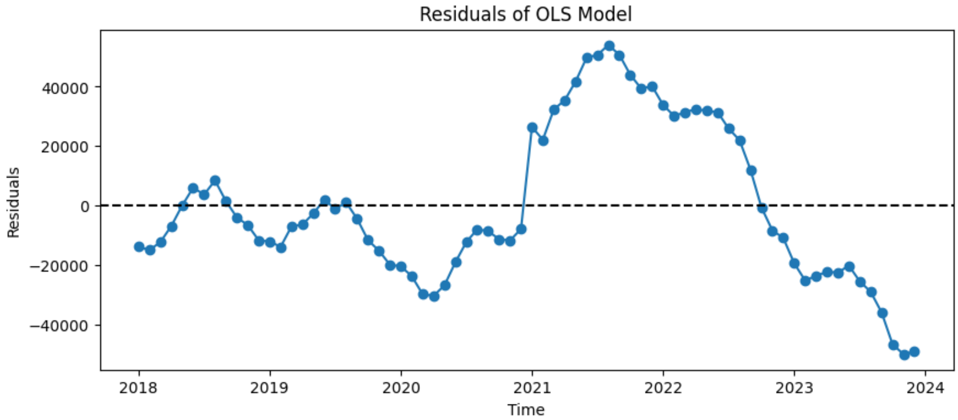

Figure 3: Residuals of the OLS global linear trend model over time.

Analysis

- **Zero mean:** The residuals are centered around zero, consistent with OLS properties.
- **Homoskedasticity:** The magnitude of the residuals increases between 2021 and 2022 and decreases afterward, indicating heteroskedasticity.
- **Independence:** The residuals display clear temporal patterns, violating the independence assumption. This suggests autocorrelation.

Conclusion

The OLS global linear trend model captures the overall upward trajectory of vehicle registrations, but some OLS assumptions are only approximately satisfied. In particular, the presence of heteroskedasticity and autocorrelation indicates that prediction intervals may underestimate uncertainty, especially in periods of higher residual variance. For more accurate short-term forecasts, a model incorporating seasonality and autocorrelation could be considered.

4 WLS – Local Linear Trend Model

We now estimate the linear trend model using Weighted Least Squares (WLS), where more recent observations receive higher weights. The weight structure is defined as

$$\lambda_0 = 1, \quad \lambda_1 = \lambda, \quad \lambda_2 = \lambda^2, \quad \ldots$$

with $\lambda = 0.9$.
Thus, the most recent observation has weight 1, and weights decrease exponentially as we move backward in time.

4.1 Variance-Covariance Matrix

In the local WLS model, the error variance-covariance matrix is

$$= \sigma^2 \begin{pmatrix} \lambda^{-(N-1)} & 0 & 0 & \cdots & 0 \\ 0 & \lambda^{-(N-2)} & 0 & \cdots & 0 \\ 0 & 0 & \lambda^{-(N-3)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

For the global OLS model:

$$= \sigma^2 I_N,$$

where all observations have equal variance and equal weight.
In contrast, the WLS model allows variances to differ across time and assigns more importance to recent observations.

4.2 λ-Weights

The weights are defined as

$$w_t = \lambda^{N-t}, \quad t = 1, \ldots, N.$$

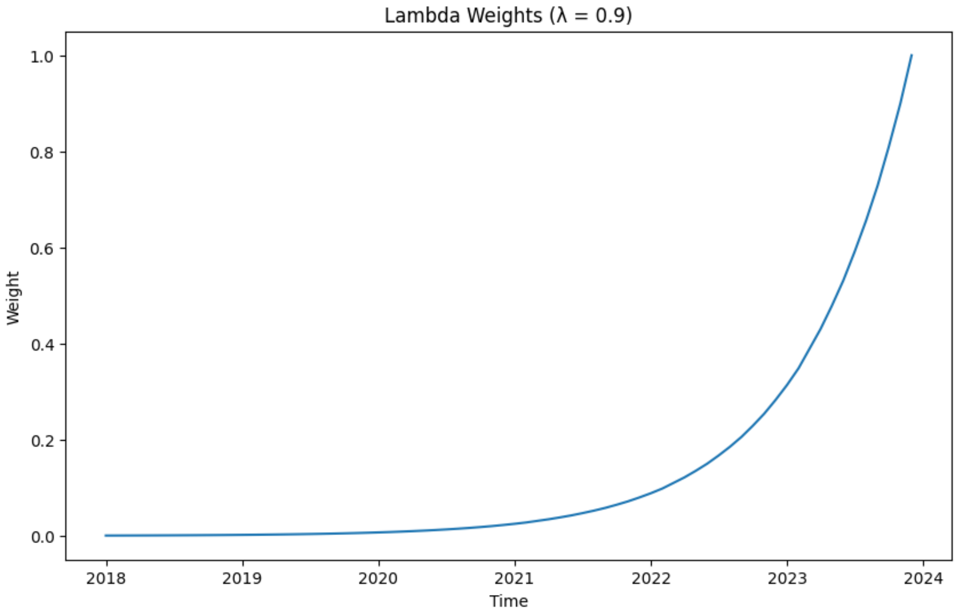

Figure 1: λ-weights over time for $\lambda = 0.9$.

Highest weight: The latest observation (time point N) has the highest weight equal to 1.

4.3 Sum of λ-Weights

The sum of the weights is

$$\sum_{k=0}^{N-1} \lambda^k = \frac{1-\lambda^N}{1-\lambda}.$$

For $\lambda = 0.9$ and $N = 72$:

$$\sum w_t = \frac{1-0.9^{72}}{1-0.9} = \frac{1-0.000507}{0.1} = 9.995.$$

Thus, the effective total weight of the WLS model is approximately 10.
In contrast, in the OLS model all weights equal 1, so

$$\sum w_t^{\text{OLS}} = N = 72.$$

4.4 Weighted Least Squares (WLS) Estimator

The WLS estimator is:

$$\hat{\beta}_{\text{WLS}} = (\mathbf{X}^\top \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \Sigma^{-1} \mathbf{y}$$

where $\mathbf{X}$ is the design matrix, $\mathbf{y}$ is the response vector, and Σ is the error variance-covariance matrix.

$$\begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} = (\mathbf{X}^\top \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \Sigma^{-1} \mathbf{y} = \begin{bmatrix} -52{,}482{,}861.79 \\ 27{,}529.90 \end{bmatrix}$$

4.5 Forecast for the Next 12 Months

Using the WLS estimates, forecasts are computed as

$$\hat{Y}_{t+l} = \hat{\theta}_1^{\text{WLS}} + \hat{\theta}_2^{\text{WLS}}(t+l), \quad l = 1, \ldots, 12.$$

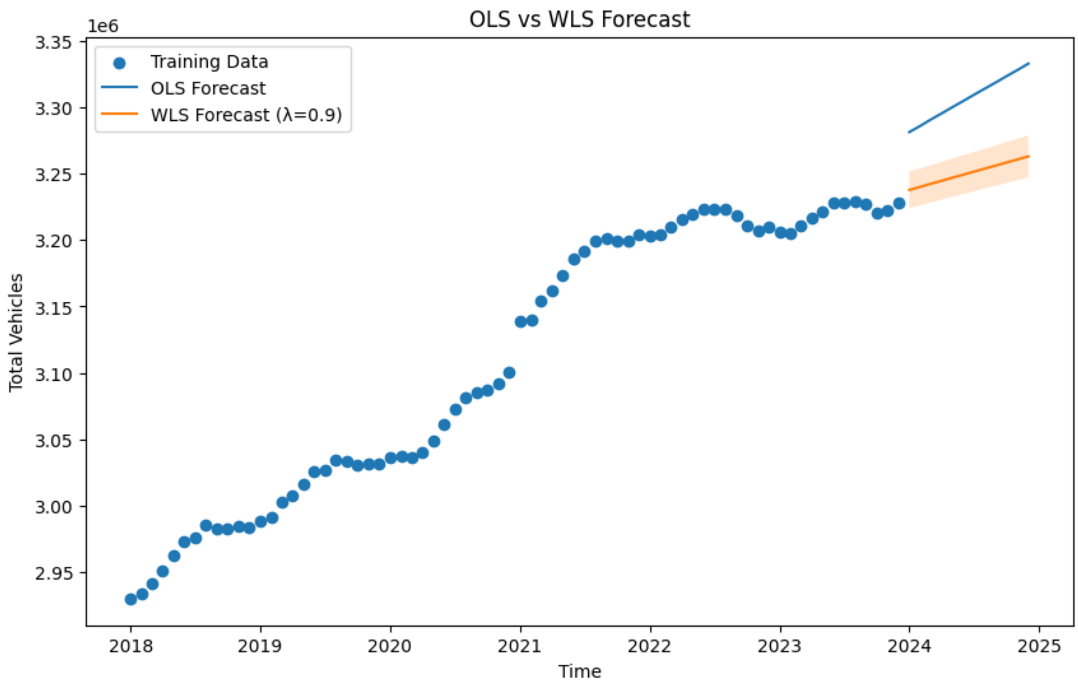

Figure 2: Training data with OLS and WLS forecasts ($\lambda = 0.9$).

Comment

The WLS forecast reacts more strongly to recent movements in the data due to the exponential weighting scheme. If structural changes or recent shifts are present, the WLS forecast may be preferable. However, if the long-term trend is stable, the OLS forecast may provide more stable predictions.
Given the observed variability in the most recent residuals, the WLS model appears more appropriate for short-term forecasting.