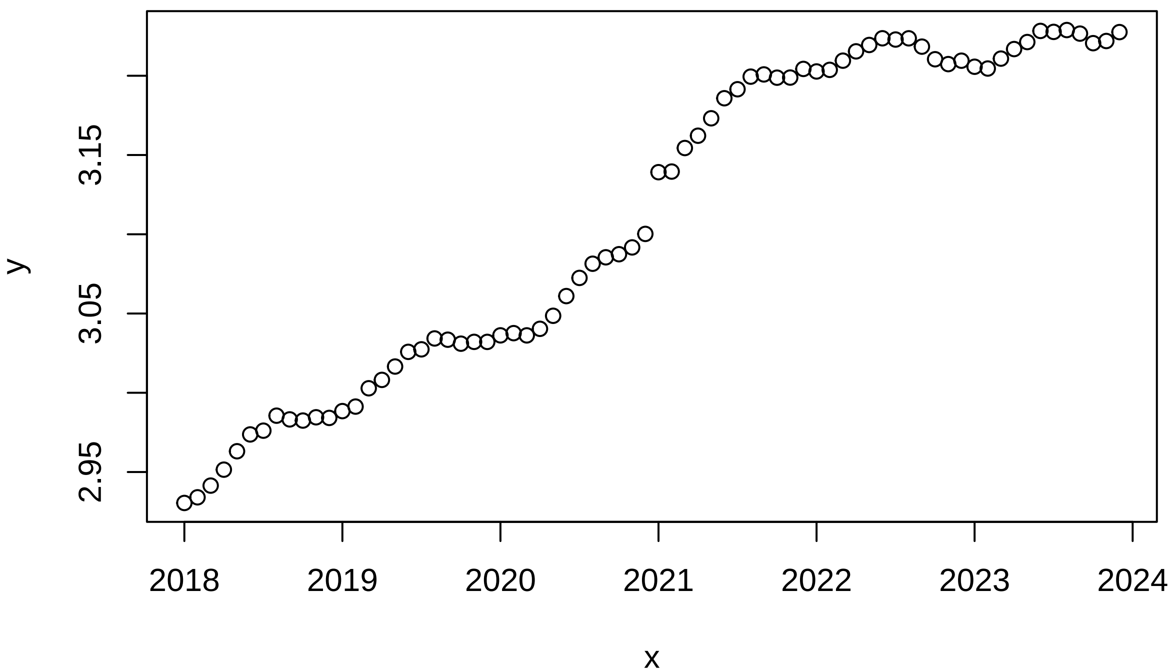
Assignment 1

2025-02-21

Group members: Keerthi (s243933), Katarina (s243906), Hubert (s243896), German (s243660)

1 Plot data

1.1.



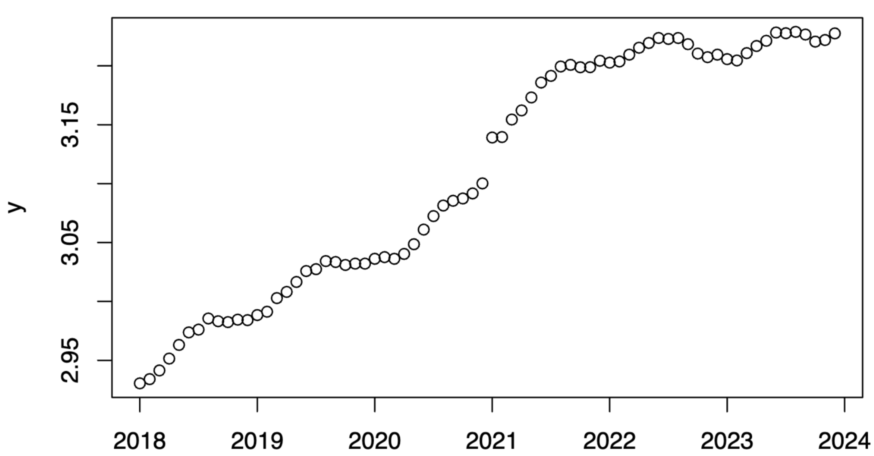


Figure 1: The plot of observations.

1.2.

There seems to be a general positive trend of number of vehicles in Denmark over the years. However, there is also a seasonal pattern within each year where it increases for approximately the first half of a year and then decreases. We also noticed that there is a jump at the start of 2021, and also after that the trend seems to have gotten flatter but the seasonal pattern still exists.

2 Linear trend model

2.1

$$y = \begin{bmatrix} 2.930 \\ 2.934 \\ 2.941 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 2018.000 \\ 1 & 2018.083 \\ 1 & 2018.167 \end{bmatrix}$$

2.2

```
# Fit a linear model
fit <- lm(total ~ year, data = Dtrain)
summary(fit)
```

```
##
## Call:
## lm(formula = total ~ year, data = Dtrain)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.049876 -0.019062 -0.006889  0.023099  0.053979
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.104e+02  3.594e+00  -30.71   <2e-16 ***
## year         5.615e-02  1.778e-03   31.57   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02613 on 70 degrees of freedom
## Multiple R-squared:  0.9344, Adjusted R-squared:  0.9335
## F-statistic:   997 on 1 and 70 DF,  p-value: < 2.2e-16
```

```
# abline(fit, col = "red")

# Extract the parameter estimates and their standard errors
theta_hat <- coef(fit)
theta_hat
```

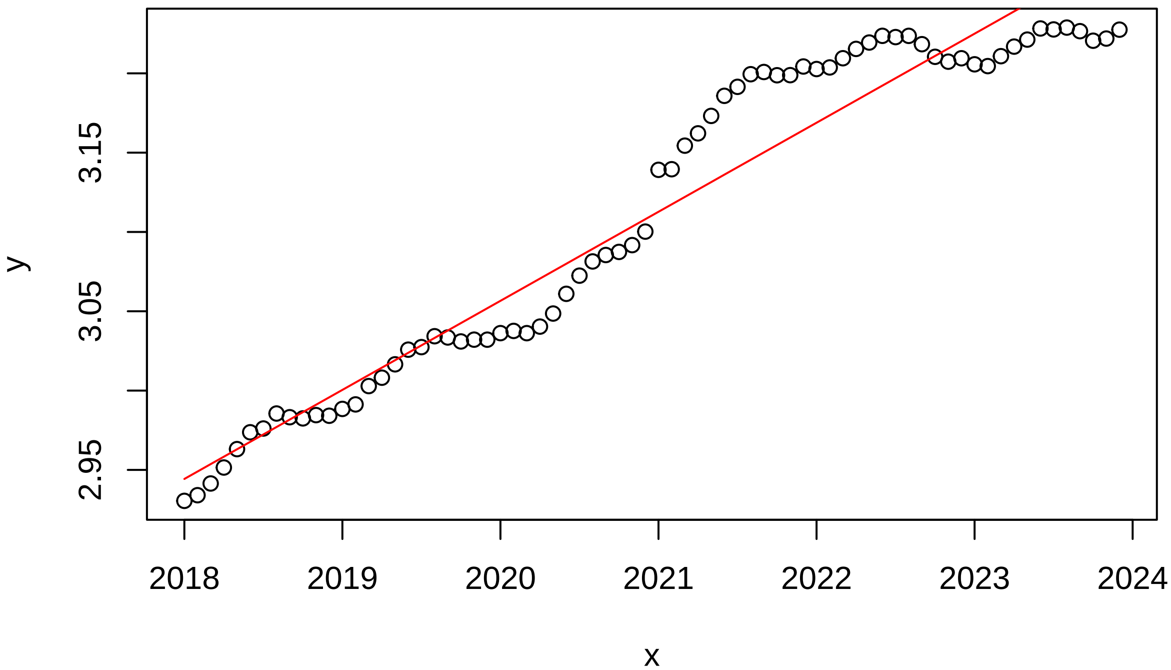
```
##   (Intercept)          year
## -110.35542813    0.05614456
```

```
theta_hat_se <- summary(fit)$coef[, 2]
theta_hat_se
```

```
## (Intercept)        year
## 3.593581122 0.001778156
```

$$\hat{\theta} = \begin{bmatrix} -110.4 \\ 0.05615 \end{bmatrix}, \quad \hat{\sigma} = \begin{bmatrix} 3.594 \\ 0.001778 \end{bmatrix}$$

```r
plot(x, y)
lines(Dtrain$year, predict(fit), col = "red")
```



3 WLS - local linear trend model

3.1

The variance-covariance matrix for the local model consists of the inverse of observation weights in the diagonal and zeros otherwise. N is equal to the number of observation, which is equal to 72.

$$\Sigma_{WLS} = \begin{pmatrix} \frac{1}{\lambda^{N-1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda^{N-2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda^0} \end{pmatrix}$$

Conversely, the variance-covariance matrix for the global model contains 1 in the diagonal and zeros otherwise.

$$\Sigma_{OLS} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

3.2

The highest weight is for the latest time point and it is equal to $\lambda^0 = 1$. The further in the past an observation point is, the more does its weight decrease. The weight distribution is visualized in Figure 1.

3.3

The sum of weights for the local model is $\sum_{n=1}^{N} \lambda^{n-1} = 9.994925$. The sum of weights for the global model is equal to the number of observations, 72.

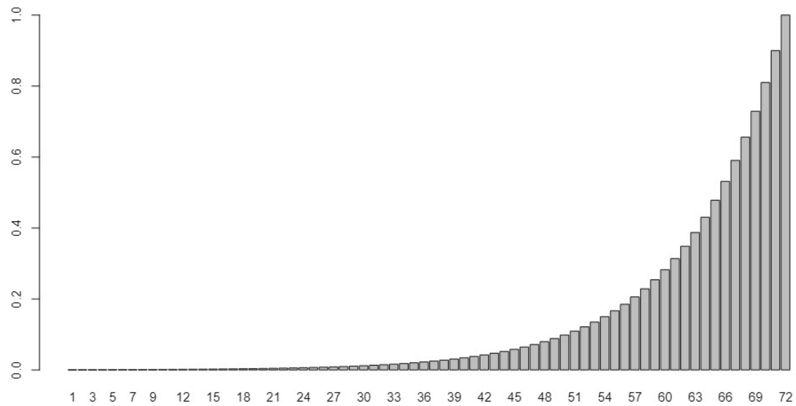


Figure 2: The weights of each observation in the train dataset.

3.4

According to the local model with $\lambda = 0.9$, the parameters are equal to $\theta_1 = -52.4828617$ and $\theta_2 = 0.0275299$.

3.5

The WLS model prioritizes recent data points, creating a more gradual slope that adapts to current trends. While OLS treats all data equally and provides a generalized fit, WLS gives higher weight to recent observations. We would choose WLS when new predictions are crucial, and OLS when analyzing from a long-term perspective.

Figure 3: The observations as well as the predictions for the OLS & WLS models, including prediction intervals.

4 Recursive estimation and optimization of λ

4.1