

# Semester project for ECON4170

## Table of contents

introduction . . . . .	1
Data . . . . .	1
Running Code . . . . .	2

### introduction

this is a semester project where we will be making a model to see if we can predict the price of electricity in Norway, one day (24 hours) ahead of time. we want to create a model that beats the naive benchmark.

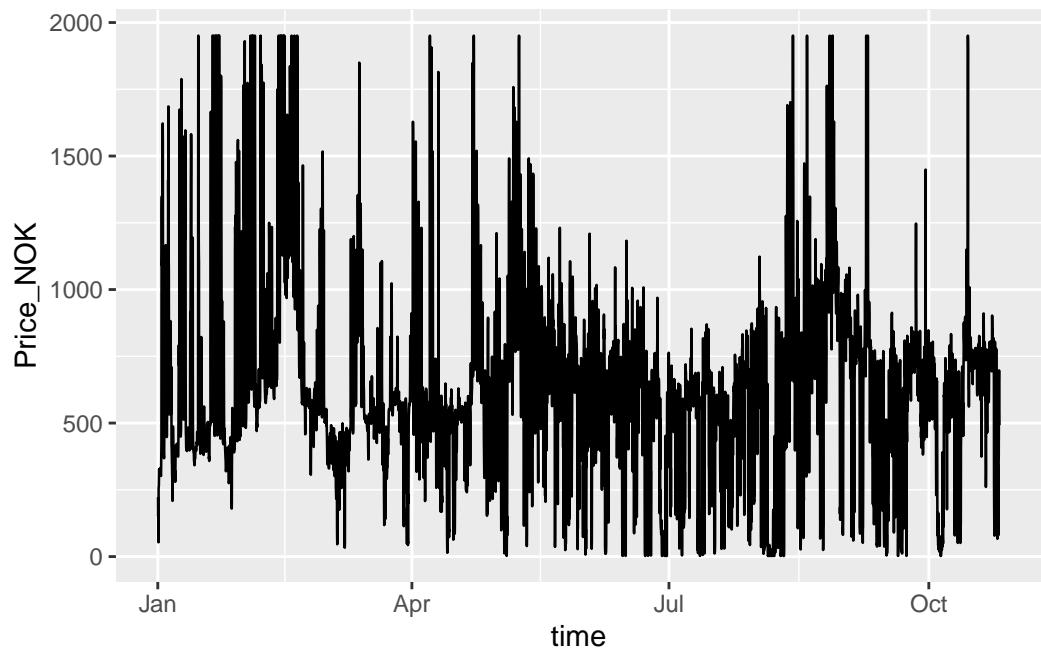
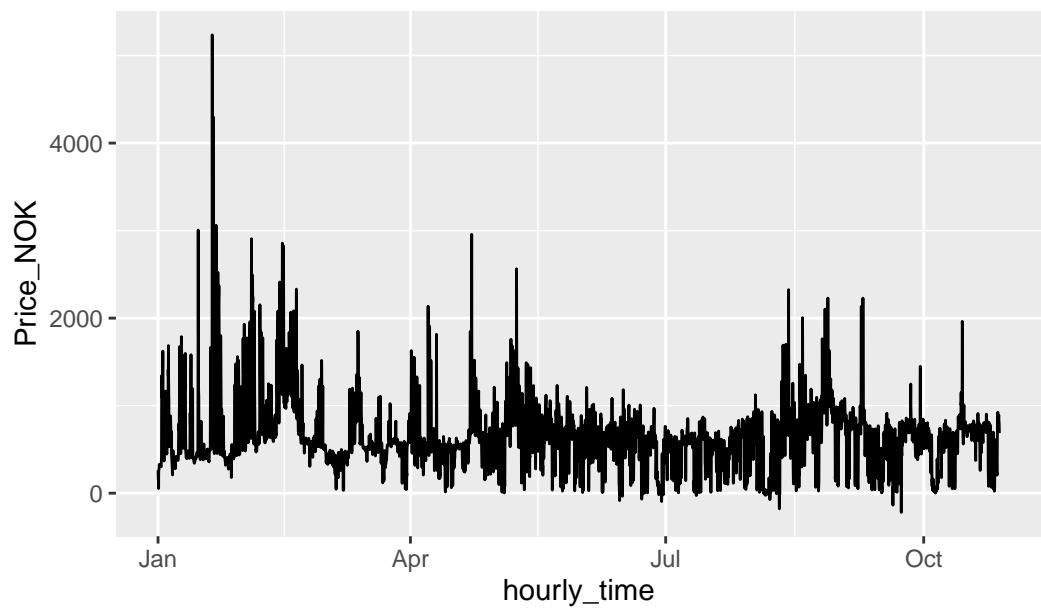
### Data

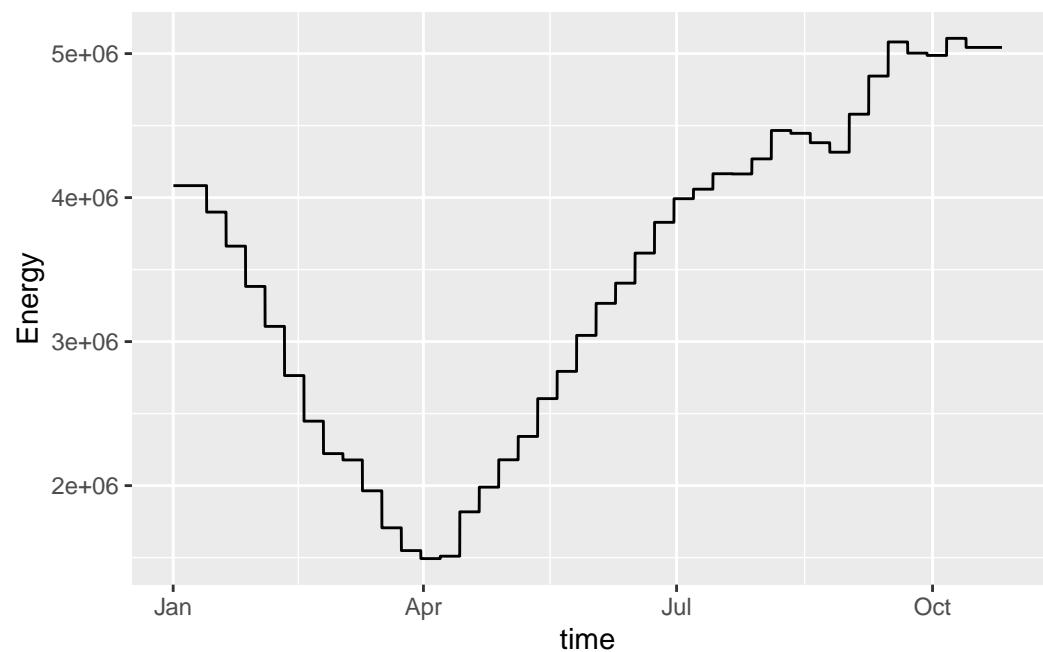
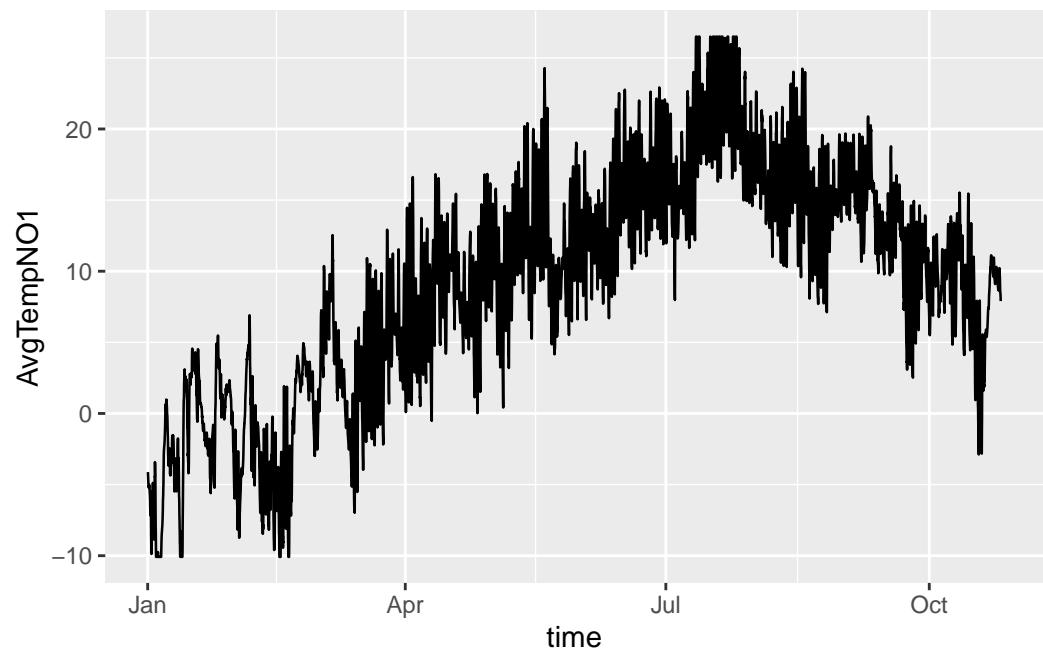
the data that we will be using comes from ENTSO-E. we have gotten a yearly report for this year (up to the 27th of October) with the time between measurements being every 15 minutes, across an entire 24-hour period. its an excel file (CSV). the original currency that is displayed in the excel file is in EUR.

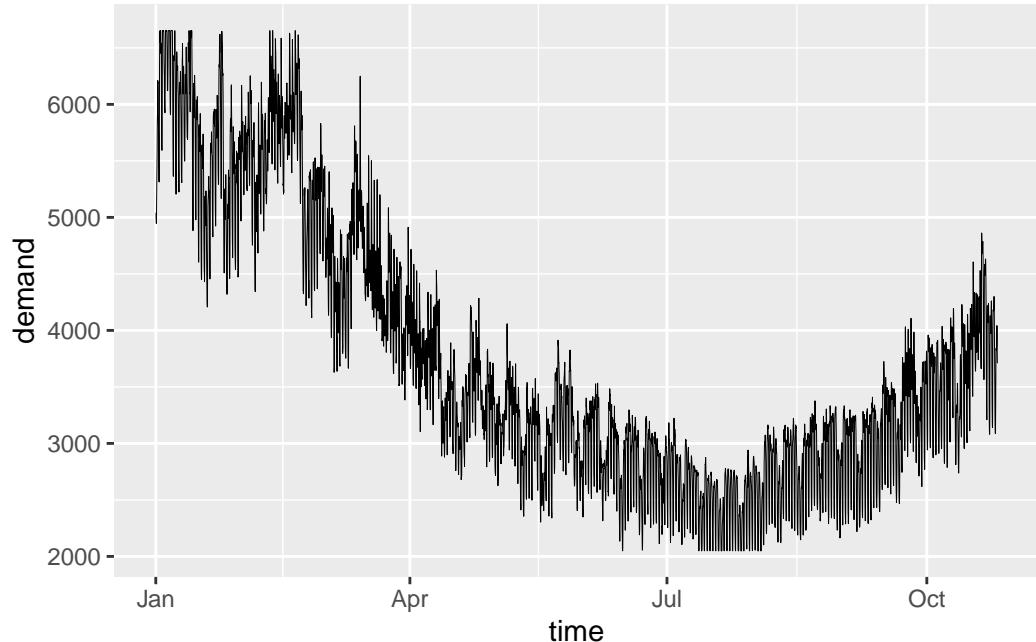
we will transform this data to fit our needs as we go. to start, we will be adding a new column that converts the currency and price, to something that displays NOK. To do this, we will use todays current exchange rate. (11,68 NOK per EUR, as of the 26th of October 2025)

## Running Code

Electricity prices (hourly)







```
[1] -0.2799558
```

```
[1] -0.8358425
```

```
[1] 0.2793221
```

```
[1] 0.8555199
```

```
# A tibble: 32 x 6
  price   HDH  load reservoir  RMSE complexity
  <int> <int> <int>      <int> <dbl>      <int>
1     4     2     2          1 113.         9
2     4     1     2          1 113.         8
3     4     2     2          2 113.        10
4     4     2     1          1 113.         8
5     4     1     2          2 114.         9
6     4     1     1          1 114.         7
7     4     2     1          2 114.         9
8     4     1     1          2 114.         8
9     2     2     2          1 117.         7
10    2     1     2          1 117.         6
# i 22 more rows
```

```
# A tibble: 1 x 6
  price   HDH  load reservoir  RMSE complexity
  <int> <int> <int>      <int> <dbl>      <int>
1     4     1     1          1 114.         7
```

```

$price_lags
[1] 1 2 24 168

$HDH_lags
[1] 0

$load_lags
[1] 0

$reservoir_lags
[1] 0

Call:
lm(formula = struktur, data = final_data)

Residuals:
    Min      1Q   Median      3Q      Max 
-1059.80 -137.15   -1.07  127.71 1259.61 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 3.318e+02 7.856e+01 4.224 2.43e-05 ***
price_lag1  4.924e-01 2.595e-02 18.974 < 2e-16 ***
price_lag2 -6.472e-02 2.544e-02 -2.544 0.010986 *  
price_lag24 5.221e-02 1.233e-02  4.233 2.34e-05 *** 
price_lag168 2.455e-02 1.045e-02  2.350 0.018808 *  
HDH         1.512e+01 1.551e+00  9.751 < 2e-16 *** 
demand      -2.374e-02 1.413e-02 -1.680 0.093027 .  
Energy       -2.199e-05 1.393e-05 -1.579 0.114353  
dummy_hour1 -3.816e+01 2.149e+01 -1.775 0.075874 .  
dummy_hour2 -4.437e+01 2.192e+01 -2.024 0.043005 *  
dummy_hour3 -7.823e+01 2.239e+01 -3.494 0.000479 *** 
dummy_hour4 -8.691e+01 2.277e+01 -3.818 0.000136 *** 
dummy_hour5 -8.915e+01 2.290e+01 -3.893 9.99e-05 *** 
dummy_hour6 -7.245e+01 2.271e+01 -3.191 0.001425 ** 
dummy_hour7 -7.758e+00 2.210e+01 -0.351 0.725540  
dummy_hour8  9.484e+01 2.170e+01  4.372 1.25e-05 *** 
dummy_hour9  1.373e+02 2.210e+01  6.213 5.52e-10 *** 
dummy_hour10 9.391e+01 2.226e+01  4.220 2.48e-05 *** 
dummy_hour11 4.611e+01 2.234e+01  2.064 0.039061 *  
dummy_hour12 2.871e+01 2.257e+01  1.272 0.203499  
dummy_hour13 3.174e-01 2.272e+01  0.014 0.988856  
dummy_hour14 -1.915e+01 2.279e+01 -0.840 0.400778  
dummy_hour15 -2.240e+01 2.279e+01 -0.983 0.325733  
dummy_hour16  7.288e+00 2.269e+01  0.321 0.748013  
dummy_hour17 5.284e+01 2.269e+01  2.329 0.019869 *  

```

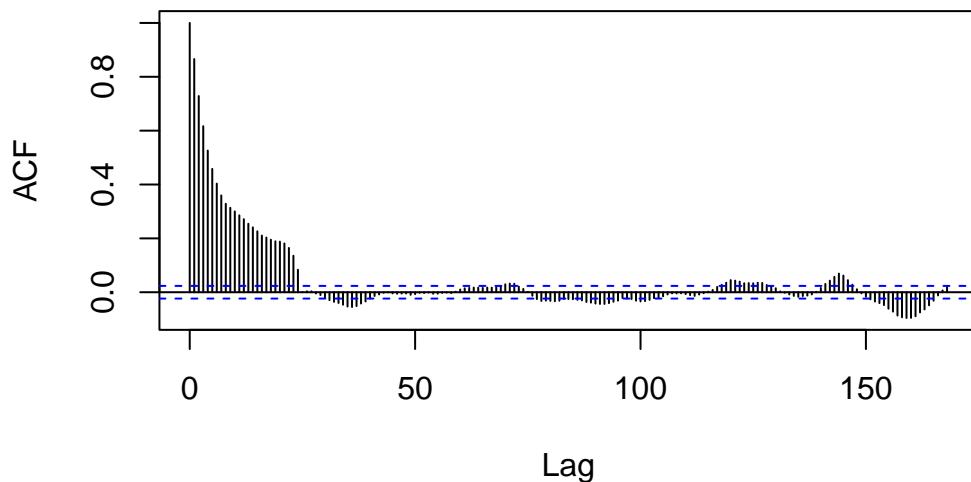
dummy_hour18	9.042e+01	2.272e+01	3.979	7.00e-05	***
dummy_hour19	1.173e+02	2.252e+01	5.210	1.95e-07	***
dummy_hour20	1.141e+02	2.224e+01	5.130	2.97e-07	***
dummy_hour21	9.999e+01	2.192e+01	4.562	5.16e-06	***
dummy_hour22	9.578e+01	2.166e+01	4.422	9.92e-06	***
dummy_hour23	4.006e+01	2.146e+01	1.867	0.061941	.
dummy_weekday2	3.777e+01	1.182e+01	3.195	0.001405	**
dummy_weekday3	-4.561e+00	1.170e+01	-0.390	0.696628	
dummy_weekday4	-3.473e+01	1.180e+01	-2.942	0.003270	**
dummy_weekday5	-1.881e+02	1.180e+01	-15.941	< 2e-16	***
dummy_weekday6	-1.159e+02	1.292e+01	-8.973	< 2e-16	***
dummy_weekday7	7.785e+01	1.277e+01	6.098	1.13e-09	***
dummy_month2	1.860e+01	2.058e+01	0.904	0.366017	
dummy_month3	-1.264e+02	3.105e+01	-4.072	4.71e-05	***
dummy_month4	-5.656e+01	3.578e+01	-1.581	0.113960	
dummy_month5	6.352e+01	3.153e+01	2.014	0.044013	*
dummy_month6	3.396e+01	3.047e+01	1.114	0.265147	
dummy_month7	6.009e+01	3.447e+01	1.743	0.081322	.
dummy_month8	1.623e+02	3.405e+01	4.767	1.91e-06	***
dummy_month9	6.347e+01	3.425e+01	1.853	0.063890	.
dummy_month10	1.028e+01	3.164e+01	0.325	0.745362	

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 256.5 on 6910 degrees of freedom  
 Multiple R-squared: 0.476, Adjusted R-squared: 0.4726  
 F-statistic: 139.5 on 45 and 6910 DF, p-value: < 2.2e-16

### Series residuals(model)



Series: y  
 Regression with ARIMA(1,0,1)(1,0,0)[24] errors

Coefficients:

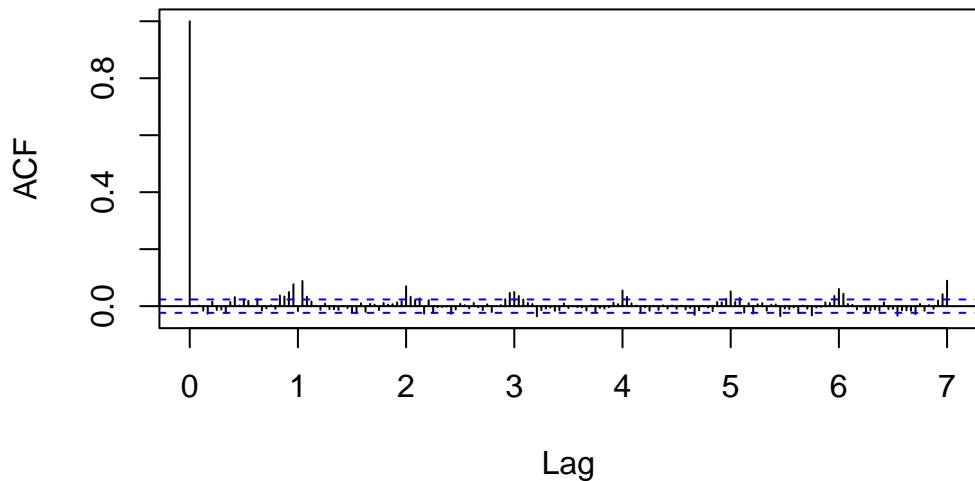
	ar1	ma1	sar1	HDH	demand	Energy	dummy_hour1	dummy_hour2
	0.8763	0.1580	0.2022	10.8585	0.1456	0	-42.8998	-49.9640
s.e.	0.0066	0.0131	0.0121	2.7898	0.0163	0	9.2341	14.1068
	dummy_hour3	dummy_hour4	dummy_hour5	dummy_hour6	dummy_hour7			
	-77.0449	-93.6553	-101.5176	-94.0449	-44.0434			
s.e.	17.3438	19.6703	21.1179	21.7754	21.7111			
	dummy_hour8	dummy_hour9	dummy_hour10	dummy_hour11	dummy_hour12			
	55.0187	123.5390	99.4008	23.1643	-37.1758			
s.e.	21.8025	22.8966	24.0340	24.9260	25.5196			
	dummy_hour13	dummy_hour14	dummy_hour15	dummy_hour16	dummy_hour17			
	-96.2569	-144.8887	-168.7837	-146.4707	-88.2298			
s.e.	25.8442	25.7827	25.4658	24.7747	23.9819			
	dummy_hour18	dummy_hour19	dummy_hour20	dummy_hour21	dummy_hour22			
	-20.2081	48.3186	83.0012	89.2535	94.6486			
s.e.	22.9291	21.2793	19.2498	16.8262	13.6879			
	dummy_hour23	dummy_weekday3	dummy_weekday4	dummy_weekday5				
	50.4089	-32.4839	-51.1664	-82.0030				
s.e.	9.1620	15.6174	20.5469	22.1721				
	dummy_weekday6	dummy_weekday7	dummy_month3	dummy_month4	dummy_month5			
	-61.3754	-25.0752	-91.3649	58.7388	320.4472			
s.e.	20.8816	15.6693	50.1698	50.1996	57.0708			
	dummy_month6	dummy_month7	dummy_month8	dummy_month9	dummy_month10			
	297.0823	339.3710	468.4285	317.9811	200.2162			
s.e.	74.1563	89.3955	94.5056	100.2327	101.3992			

$\sigma^2 = 13572$ : log likelihood = -42946.26  
 AIC=85978.53 AICc=85979.08 BIC=86272.96

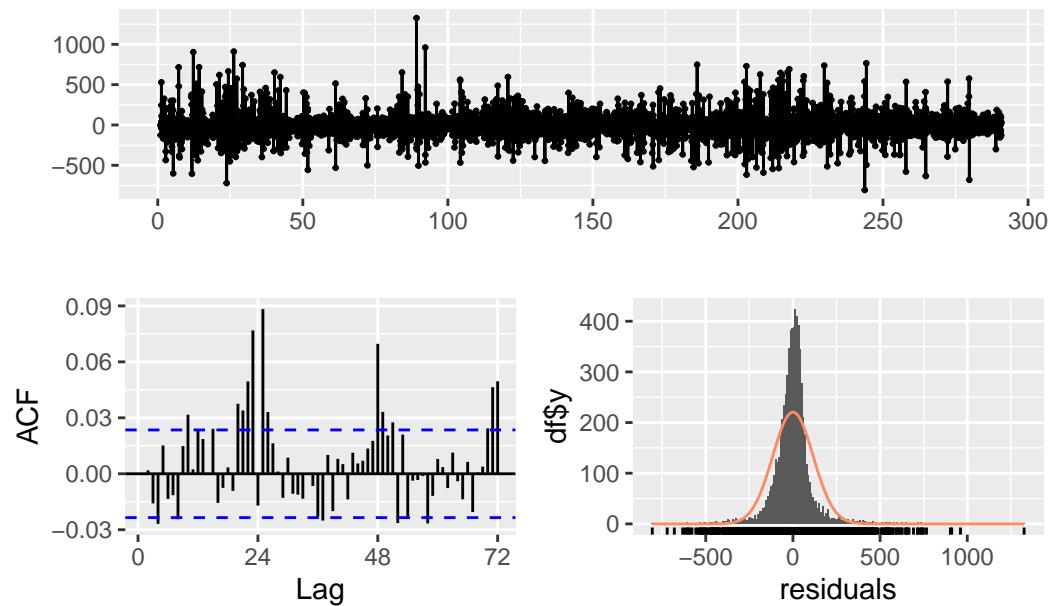
Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE
Training set	0.1429675	116.1464	71.84922	-37.79588	51.272	0.3618096
	ACF1					
Training set	0.0004310809					

### Series residuals(sarimax)



Residuals from Regression with ARIMA(1,0,1)(1,0,0)[24] errors



Ljung-Box test

```
data: Residuals from Regression with ARIMA(1,0,1)(1,0,0)[24] errors
Q* = 236.07, df = 45, p-value < 2.2e-16
```

Model df: 3. Total lags used: 48

[1] 39

```

[1] "HDH"           "demand"        "Energy"         "dummy_hour1"   "dummy_hour2"
[6] "dummy_hour3"

#CHhecks

#simple model

###trying with sarimax because strong autocorrelation in adl model###

final_data <- clean_data_dummies %>%
  mutate(y_next = lead(Price_NOK, 24)) %>%
  drop_na()

#build y and x for the sarimax model#
y<- ts(final_data$y_next, frequency = 24)

struktur_xreg <- y_next ~ HDH + demand + dummy_hour - 1
x <- model.matrix(struktur_xreg, data = final_data)
x <- x[, colnames(x) != "dummy_hour0", drop = FALSE]
sarimax_simpler <- auto.arima(y, xreg = x, seasonal = TRUE, stepwise = TRUE, approximation = T
                                max.p = 3, max.q = 3, max.P = 1, max.Q = 1)
summary(sarimax_simpler)

#acf with new model#
acf(residuals(sarimax), lag.max = 168)

checkresiduals(sarimax)

```