We clean the memory and charge the libraries

```
rm(list=ls())
library(fpp2)
```

```
## Warning: package 'fpp2' was built under R version 4.0.5
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
## -- Attaching packages --------------------------------------------- fpp2 2.4 --
```

```
## v ggplot2   3.3.5     v fma       2.4
## v forecast  8.16      v expsmooth 2.3
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

```
## Warning: package 'forecast' was built under R version 4.0.5
```

```
## Warning: package 'fma' was built under R version 4.0.5
```

```
## Warning: package 'expsmooth' was built under R version 4.0.5
```

```
##
```

We assign a variable with the time series of the train data

```
tr_NO2_24h_00 <- read.csv("~/It Academy/Data Science/Aire/tr_NO2_24h_00.csv", row.names=1)
```

We execute auto.arima to find the optimal parameters of the model

```
mod.tr_NO2_24h_00 <- auto.arima(tr_NO2_24h_00, trace=TRUE, allowdrift=FALSE)
```

```
##
##  Fitting models using approximations to speed things up...
##
##  ARIMA(2,1,2)                    : 6491.072
##  ARIMA(0,1,0)                    : 6705.866
##  ARIMA(1,1,0)                    : 6672.009
##  ARIMA(0,1,1)                    : 6619.224
##  ARIMA(1,1,2)                    : 6498.167
##  ARIMA(2,1,1)                    : 6489.622
##  ARIMA(1,1,1)                    : 6505.887
##  ARIMA(2,1,0)                    : 6614.865
##  ARIMA(3,1,1)                    : 6497.743
##  ARIMA(3,1,0)                    : 6595.524
##  ARIMA(3,1,2)                    : 6494.934
##
##  Now re-fitting the best model(s) without approximations...
##
##  ARIMA(2,1,1)                    : 6495.288
##
##  Best model: ARIMA(2,1,1)
```

The best model is ARIMA (2,1,1). We create a model with this parameters and the train data

```
mod.tr_NO2_24h_00 <- Arima(tr_NO2_24h_00[,1], order=c(2,1,1))
```

We evaluate the accuracy of the prediction model, in this case the one step forecasting is executed on the train data

```
accuracy(mod.tr_NO2_24h_00)
```

```
##                      ME     RMSE      MAE       MPE     MAPE      MASE
## Training set -0.1270531 10.45401 8.327531 -8.482339 24.96462 0.8945777
##                      ACF1
## Training set 0.001653211
```

We assign a variable with the time series of the test data

```
te_NO2_24h_00 <- read.csv("~/It Academy/Data Science/Aire/te_NO2_24h_00.csv", row.names=1)
```

We perform a one step forecast in the data test using the model with the fitted parameters from the train data

```
mod.te_NO2_24h_00<- Arima(te_NO2_24h_00[,1], model=mod.tr_NO2_24h_00)
```

We evaluate the accuracy of the model on the test data,

```
accuracy(mod.te_NO2_24h_00)
```

```
##                      ME     RMSE      MAE       MPE     MAPE      MASE
## Training set 0.08205777 12.92776 10.48077 -8.869367 26.72146 0.9063232
##                     ACF1
## Training set 0.01880632
```

We write a csv with the values of the time series forecast of the data test

```
write.csv(mod.te_NO2_24h_00[["fitted"]],"C:\\Users\\hinoj\\OneDrive\\Documentos\\It Academy
\\Data Science\\Aire\\fr_NO2_24h_00.csv")
```