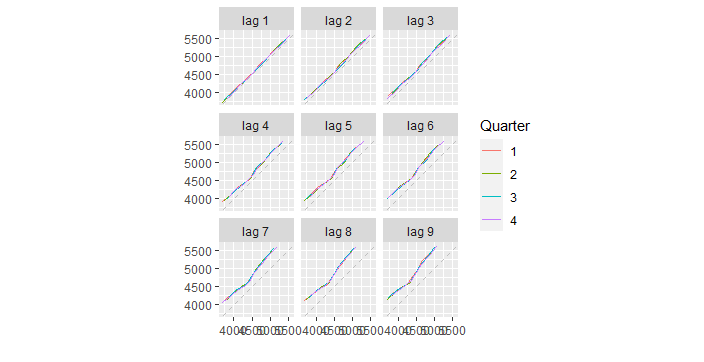
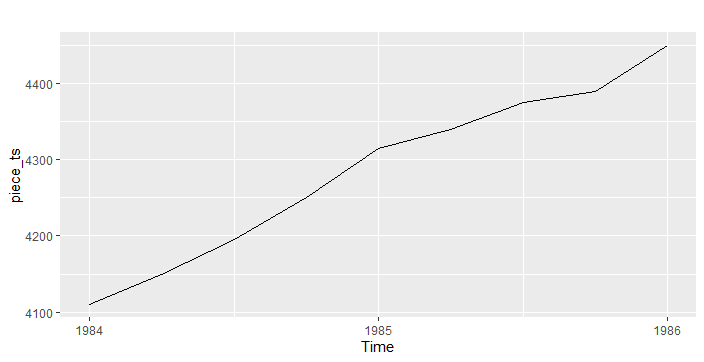
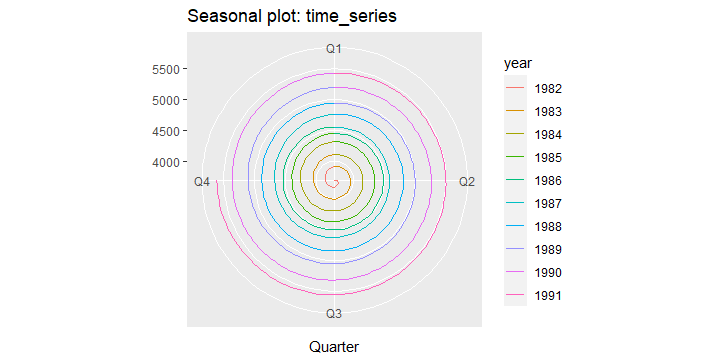
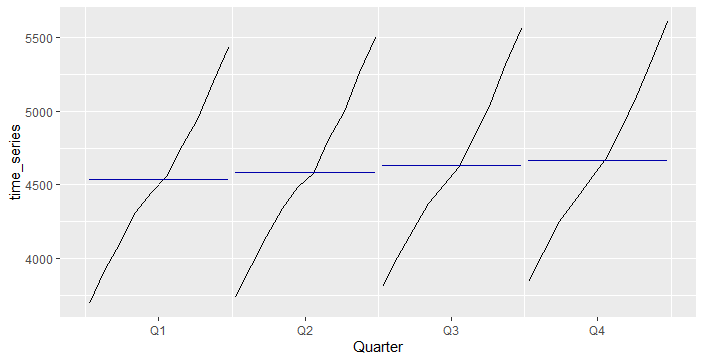
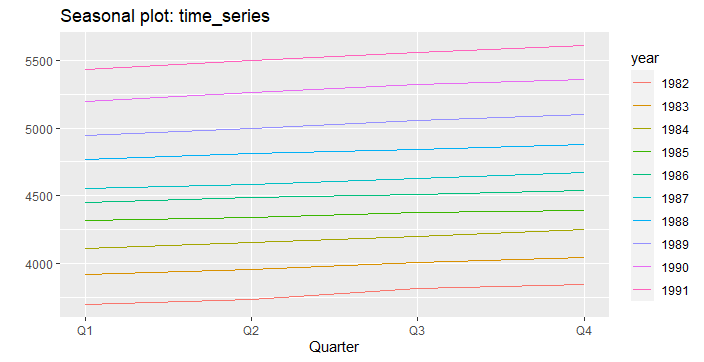
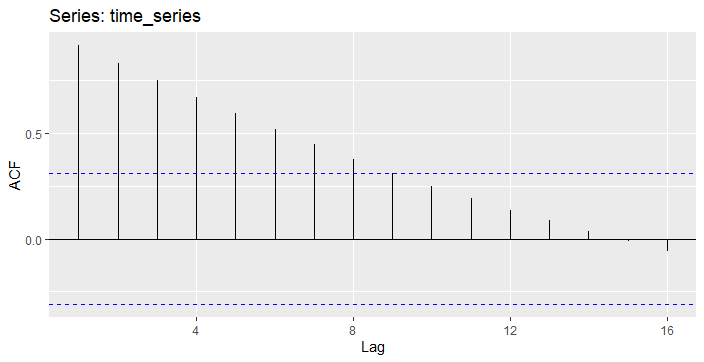
R Project-Time series

Investigating on our chosen time series. time\_series<-quarterly\_ts\_list[[703]]$x

It is a quarterly, almost trend with a small season affect and noise. Acf-plot indicated, slow decrease in the ACF as the lag increase is due to the trend, while “scalloped” shape is due to the small small seasonality or noise in the end.

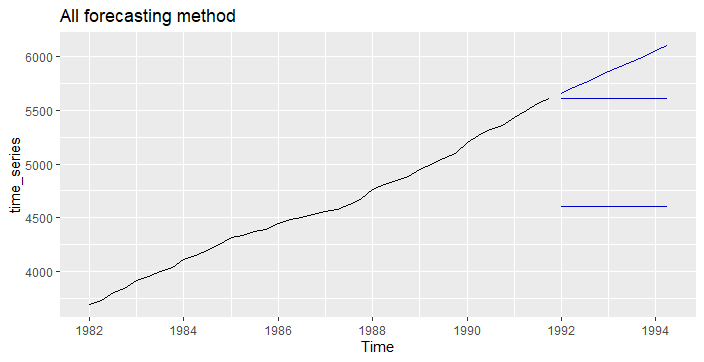
 





We have two kind of trend function linear or exponential. By Inserting data in excel and drawing plot, and trying to add a trend line including a forecast to a line chart. then a simple trend line matched. So, we can assume the series as a trend series with some noisy affect mostly in the end. Then the time series should check for stationary. By looking at ACF plot, easily find it is not stationary

Put three forecast modeling 1-mean 2-naïve 3-drift on the training data the result is represented:

1-drift(random walk)

2-naïve

3-mean

Residual: Difference between observation value and corresponding fitted value. for each prediction the residual is indicated in the following:

A good prediction should have:

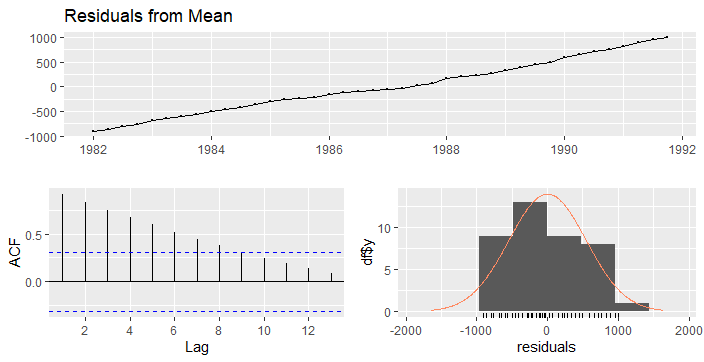
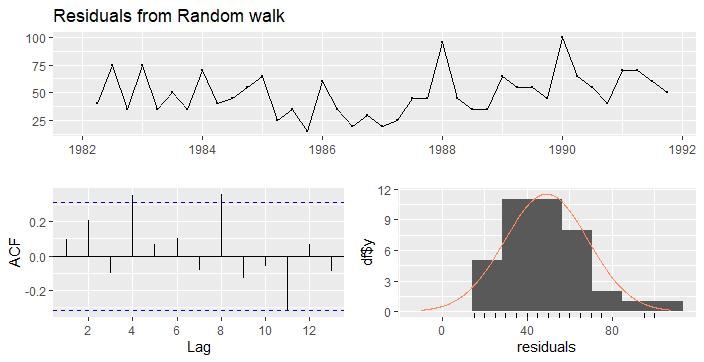
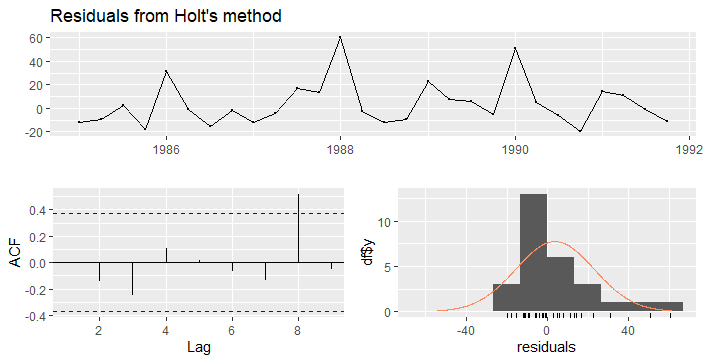
1-The residuals are uncorrelated. If there are correlations between residuals,

then there is information left in the residuals which should be used in

computing forecasts.

2. The residuals have zero mean. If the residuals have a mean other than zero,

then the forecasts are biased.

MEAN Accuracy: ME RMSE MAE MPE MAPE MASE

Training set 0.000 539.000 452.950 -1.37982 9.962504 2.324476

Test set 1277.375 1283.997 1277.375 21.67875 21.678746 6.555310

ACF1 Theil's U

Training set 0.9162347 NA

Test set 0.6711852 22.90204

NIVE Accuracy:

ME RMSE MAE MPE MAPE MASE

Training set 49.10256 52.77553 49.10256 1.064104 1.064104 0.2519875

Test set 271.87500 301.45999 271.87500 4.575473 4.575473 1.3952245

ACF1 Theil's U

Training set 0.09605304 NA

Test set 0.67118520 5.552135

DRIFT Accuracy:

ME RMSE MAE MPE MAPE

Training set -2.332044e-13 19.34411 15.70020 -0.01066136 0.3445598

Test set 5.091346e+01 56.12294 50.91346 0.85900172 0.8590017

MASE ACF1 Theil's U

Training set 0.08057122 0.09605304 NA

Test set 0.26128077 0.56456475 1.00684

Trying to model time series with multiple linear regression got error:

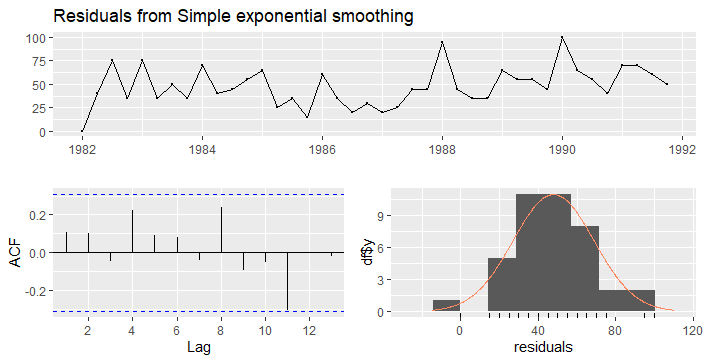
autoplot(time\_series, series="Data") +autolayer(fitted(time\_series), series="Fitted") +xlab("Year") + ylab("") +ggtitle("regression model")+guides(colour=guide\_legend(title=" "))

Error:

Error: $ operator is invalid for atomic vectors

The chosen time series is trend+ noisy so we can holt exponential forecast method but first being sure that there is no season affect on it because holt method working on time series with no season affection.

Error plot for simple exponential indicated, by consideration the error plot we can find simple exponential is not a suitable method for our series, simple exponential working perfect on the series who are not seasonal and trend. For using this method first, we should take out seasonal and trend affection then using method and after putting again the seasonal and trend affect on it.

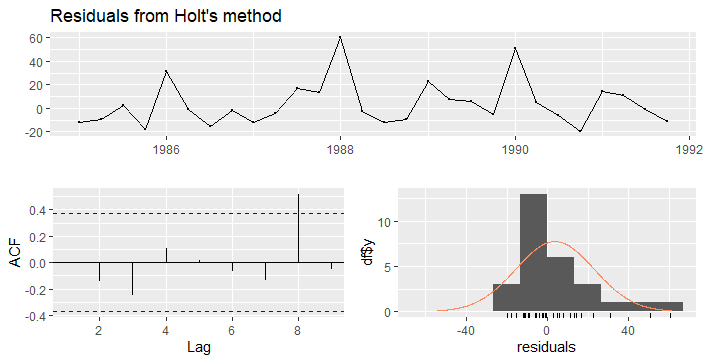


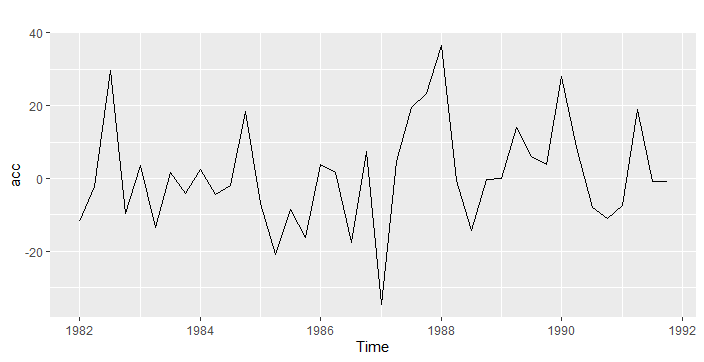
Holt method:

The chosen data is trendy+ noisy, the holt method should work perfect but it does not the validation should be checked. Why? Error mean=1.17??

Holt winter method:

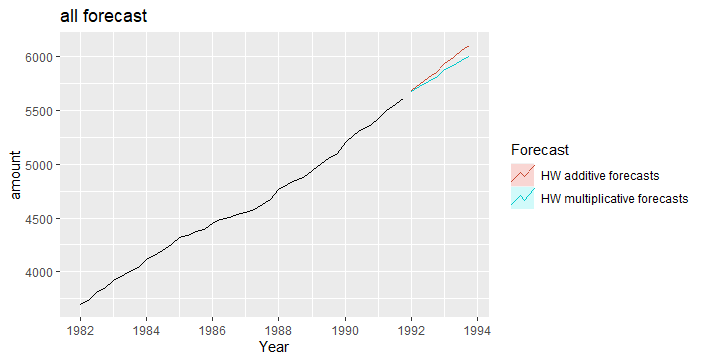
Mean error=0.87





We will use time series cross validation to choose best method for forecast:

* Holt with out damped->mean cross-validation^2=1158.11, mean(abs(cross-validation)=19.56)
* Holt with-damped->mean cross-validation^2=1143.77, mean(abs(cross-validation)=19.89)
* Holt with- damped, Seasonal= "additive"->mean cross-validation^2=1143.77, mean(abs(cross-validation)=19.89)
* Holt with out damped, Seasonal= "additive"->mean cross-validation^2=1158.11, mean(abs(cross-validation)=19.56)
* Holt with out damped, Seasonal= "multiplicative"->mean cross-validation^2=1537.32, mean(abs(cross-validation)=24.17)
* Holt with- damped, Seasonal= "multiplicative "->mean cross-validation^2=1256.40, mean(abs(cross-validation)=19.11)

In the picture below and by checking mean(abs(cross-validation)=24.17), we can see Multiplicative Holt can be a perfect method for forecasting.

And for finding hyperparameters I was using ETS: ets(time\_series, model="AAA", damped=NULL, alpha=NULL, beta=NULL,

gamma=NULL, phi=NULL, lambda=NULL, biasadj=FALSE,

additive.only=FALSE, restrict=TRUE,

allow.multiplicative.trend=FALSE)

and the hyperparameters: ETS(A,A,A)

Call:

ets(y = time\_series, model = "AAA", damped = NULL, alpha = NULL,

Call:

beta = NULL, gamma = NULL, phi = NULL, additive.only = FALSE,

Call:

lambda = NULL, biasadj = FALSE, restrict = TRUE, allow.multiplicative.trend = FALSE)

Smoothing parameters:

alpha = 0.9999

beta = 0.3011

gamma = 1e-04

Initial states:

l = 3647.7926

b = 51.3435

s = -10.667 0.6207 2.2618 7.7845

sigma: 16.0735

AIC AICc BIC

378.8032 384.8032 394.0032

Arima Model:

Checking residual is mandatory, p-value????how we can relay??

is a forecasting model is working with stationary time series so for our chosen time series we should first make it stationary by differencing method.

Summary:Forecast method: ARIMA(1,1,0)(2,0,0)[4] with drift combination of models

Model Information:

Series: time\_series

ARIMA(1,1,0)(2,0,0)[4] with drift

Coefficients:

ar1 sar1 sar2 drift

0.365 0.2405 0.5429 54.1809

s.e. 0.160 0.1218 0.1465 11.6569

sigma^2 = 249.8: log likelihood = -162.97

AIC=335.94 AICc=337.76 BIC=344.26

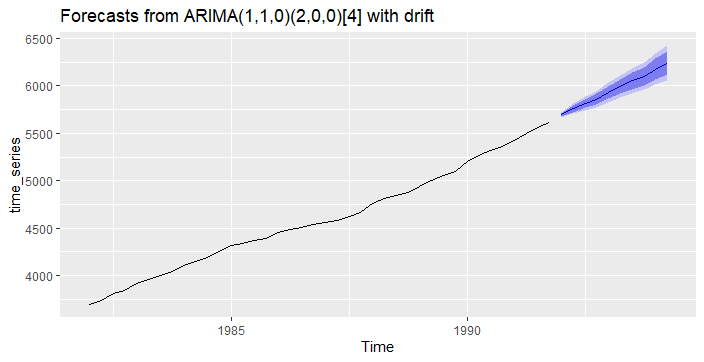
Error measures:

ME RMSE MAE MPE MAPE

Training set 0.3591122 14.78319 11.59596 0.0002349683 0.2600844

MASE ACF1

Training set 0.05950887 -0.0488566



ME RMSE MAE MPE MAPE

Training set 0.3591122 14.78319 11.59596 0.0002349683 0.2600844

Test set -19.7160711 24.84769 21.11690 -0.3337997482 0.3583757

MASE ACF1 Theil's U

Training set 0.05950887 -0.048856599 NA

Test set 0.10836899 0.004548282 0.4642263