Diagram

Description automatically generated

This process may involve loading in data,

identifying missing values,

filtering the time series,

and other pre-processing tasks.

The functionality provided by tsibble and other packages in the tidyverse substantially simplifies this step.

forecasts table -> fable

Sometimes one of these simple methods will be the best forecasting method available; but in many cases, these methods will serve as benchmarks rather than the method of choice. That is, any forecasting methods we develop will be compared to these simple methods to ensure that the new method is better than these simple alternatives. If not, the new method is not worth considering.

\*mean, \*drift \*naïve

Residuals are useful in checking whether a model has adequately captured the information in the data. For this purpose, we use innovation residuals.

If patterns are observable in the innovation residuals, the model can probably be improved. We will look at some tools for exploring patterns in residuals in the next section.

**A good forecasting method will yield innovation residuals with the following properties:**

1. **The innovation residuals are uncorrelated. If there are correlations between innovation residuals, then there is information left in the residuals which should be used in computing forecasts.**
2. **The innovation residuals have zero mean. If they have a mean other than zero, then the forecasts are biased.**

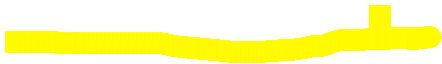
Adjusting for bias is easy: if the residuals have mean m, then simply subtract m from all forecasts and the bias problem is solved. Fixing the correlation problem is harder,

**In addition to these essential properties, it is useful (but not necessary) for the residuals to also have the following two properties.**

1. **The innovation residuals have constant variance. This is known as “homoscedasticity”.**
2. **The innovation residuals are normally distributed.**

Text

Description automatically generated



**Portmanteau tests for autocorrelation**

When we look at the ACF plot to see whether each spike is within the required limits, we are implicitly carrying out multiple hypothesis tests, each one with a small probability of giving a false positive.

When enough of these tests are done, it is likely that at least one will give a false positive, and so we may conclude that the residuals have some remaining autocorrelation, when in fact they do not



In order to overcome this problem, we test whether the first ℓ autocorrelations are significantly different from what would be expected from a white noise process. A test for a group of autocorrelations is called a **portmanteau test**, from a French word describing a suitcase or coat rack carrying several items of clothing.

**(DID NOT UNDERSTAND)**

**Text, letter

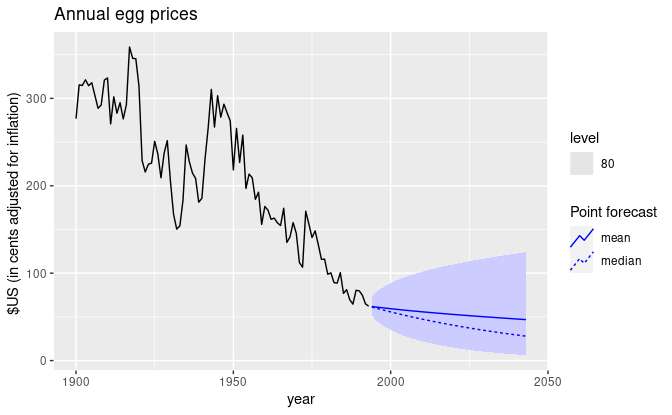
Description automatically generated**



**Graphical user interface, text

Description automatically generated**

**Difference between mean and median approach fro obtaining back-transformations of (Box Cox) for prediction intervals on original scale (inf transformation was used) – non-bias adjusted and the bias-adjusted backtransformations –**



## Forecasting with decomposition

To forecast a decomposed time series, we forecast the seasonal component,  S^t, and the seasonally adjusted component A^t, separately.

It is usually assumed that the seasonal component is unchanging, or changing extremely slowly, so it is forecast by simply taking the last year of the estimated component

 In other words, a seasonal naïve method is used for the seasonal component.

To forecast the seasonally adjusted component, any non-seasonal forecasting method may be used. For example, the drift method, or Holt’s method (discussed in Chapter [8](https://otexts.com/fpp3/expsmooth.html#expsmooth)), or a non-seasonal ARIMA model (discussed in Chapter [9](https://otexts.com/fpp3/arima.html#arima)), may be used.

**ERROR CALCULATIONS (apart from RMSE, MAPE and sMAPE)**

**1. Scaled Error**

A scaled error is less than one if it arises from a better forecast than the average one-step naïve forecast computed on the training data. Conversely, it is greater than one if the forecast is worse than the average one-step naïve forecast computed on the training data.

**Scaled error (qj) also has MASE, and RMSSE (Mean Absolute Scaled Error) and (Root Mean Squared Scaled Error)**

The accuracy() function will automatically extract the relevant periods from the data to match the forecasts when computing the various accuracy measures.

Code for Forecast and Accuracy check for non-seasoned data

**google\_fit <- google\_2015 %>%**

**model(**

**Mean = MEAN(Close),**

**`Naïve` = NAIVE(Close),**

**#** *IF SEASONAL LINE -*

`Seasonal naïve` = SNAIVE(Beer),

**Drift = RW(Close ~ drift())**

**)**

**google\_fc <- google\_fit %>%**

**forecast(google\_jan\_2016)**

**google\_fc %>%**

**autoplot(bind\_rows(google\_2015, google\_jan\_2016),**

**level = NULL) +**

**labs(y = "$US",**

**title = "Google closing stock prices from Jan 2015") +**

**guides(colour = guide\_legend(title = "Forecast"))**

**accuracy(google\_fc, google\_stock)**

**Comparing similar benchmark methods –**

**Text, letter

Description automatically generated**

**Graphical user interface, text, application, email

Description automatically generated**

**The skill\_score() function, will always compute the CRPS for the appropriate benchmark forecasts, even if these are not included in the fable object**.

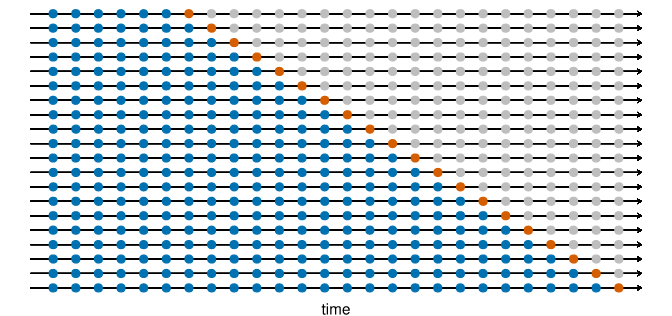
**When the data are seasonal, the benchmark used is the seasonal naïve method rather than the naïve method. To ensure that the same training data are used for the benchmark forecasts, it is important that the data provided to the accuracy() function starts at the same time as the training data.**

**The skill\_score() function can be used with any accuracy measure. For example, skill\_score(MSE) provides a way of comparing MSE values across diverse series.**

However, it is important that the test set is large enough to allow reliable calculation of the error measure, especially in the denominator. For that reason,

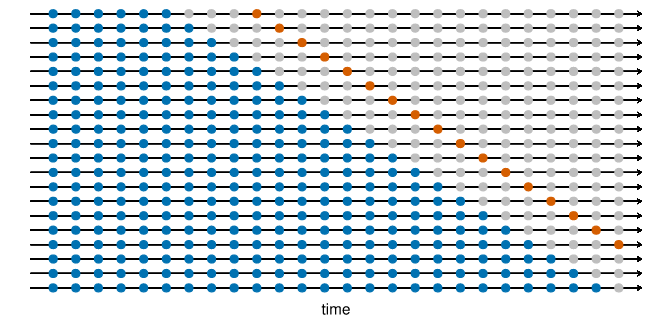
MASE or RMSSE are often preferable scale-free measures for point forecast accuracy.

## Time series cross-validation

****

The forecast accuracy is computed by averaging over the test sets. This procedure is sometimes known as **“evaluation on a rolling forecasting origin” because the “origin” at which the forecast is based rolls forward in time.**

With time series forecasting, one-step forecasts may not be as relevant as multi-step forecasts. In this case, the cross-validation procedure based on a rolling forecasting origin can be modified to allow multi-step errors to be used. Suppose that we are interested in models that produce good 44-step-ahead forecasts. Then the corresponding diagram is shown below

****

**In the following example, we compare the accuracy obtained via time series cross-validation with the residual accuracy. The stretch\_tsibble() function is used to create many training sets. In this example, we start with a training set of length .init=3, and increasing the size of successive training sets by .step=1.**



# Time series cross-validation accuracy

google\_2015\_tr <- google\_2015 %>%

stretch\_tsibble(.init = 3, .step = 1) %>%



relocate(Date, Symbol, .id)

