Notes for Hyndman and Athanasopoulos – Chapter 2

* Frequency of a time-series is the number of observations before the seasonal pattern repeats.
  + Yearly data: Frequency = 1
  + Quarterly data: Frequency = 4
  + Monthly data: Frequency = 12
  + Weekly data: Frequency = 52
  + Daily data might have two types of seasonality:
    - Weekly seasonality (frequency = 7)
    - Annual seasonality (frequency = 365)
* Time series patterns
  + *Trend:* A trend exists when there is a long-term increase or decrease in the data. It does not have to be linear.
  + *Seasonal:* A seasonal pattern occurs when a time-series is affected by seasonal factors such as the time of the year or the day of the week. Seasonality is always fixed and of known frequency.
  + *Cyclic:* A cycle occurs when the data exhibit rises and falls that are not of a fixed frequency
* If the fluctuations are not of a FIXED FREQUENCY, then they are cyclic; if the frequency is unchanging and associated with some aspect of the calendar then the pattern is seasonal.
* An accurate forecasting method must take into account the time patterns in the data and be able to capture the patterns properly.
* A seasonal plot is similar to a time plot except that the data are plotted against individual seasons in which the data was observed.
* Scatterplots are useful for visualizing relationship between time-series. It is common in these cases to display the correlation coefficient to measure the strength of the relationship between these two variables.
* Autocorrelation measures the linear relationship between lagged values of a time-series.
* The autocorrelation coefficients are plotted to show the autocorrelation function (AFC). The plot generated is known as the *correlogram*.
* When data has a trend, the autocorrelation for small lags tend to be large and positive because observations nearby in time are also nearby in size (due to changing mean of the series as time goes by).
* When data is seasonal, the autocorrelations will be larger for the seasonal lags (at multiples of the seasonal frequency) than for other lags.
* Time series that show no autocorrelation are called **white noise**.

Notes for Hyndman and Athanasopoulos – Chapter 3

**Simple forecasting methods**

* Average method for forecasting:
  + The forecasts of all future values are equal to the average of the historical data.
* Naïve method of forecasting:
  + Naïve forecast consists in setting all forecasts to be the value of the last observation. This is an optimal technique when data follows a random walk.
* Seasonal naïve method:
  + The forecast is equal to the last observed period from the same season of the past year.
* Drift method:
  + A variation of the naïve method is to allow for forecasts to increase or decrease over time, where the amount of change over time (**drift**) is set to be the average change seen in historical data. Equivalent to drawing a line between the first and last observations and extrapolating into the future.
* These simple forecasting methods usually serve as benchmark to analyze more complicated models.

**Transformation and adjustment**

* Adjusting the historical data can often lead to a simpler forecasting task. The purpose of these adjustments and transformations is to simply the patterns in historical data by removing known sources of variation.
* Calendar adjustments:
  + A common source of variation derives from the fact that some months have more days than others.
* Population adjustments:
  + Any data that is affected by population changes can be adjusted to give per-capita data. This removes the effects of total population changes.
* Inflation adjustments:
  + Financial time-series are usually adjusted so that all values are stated in dollar values from a particular year.

**Mathematical transformations**

* Logarithmic transformations.
* Power transformations.
* Box-cox transformation: a family of transformations that combines natural logarithms and power transformations. For different values of lambda, the series will change in shape. A good value of lambda makes the seasonal variation about the same across the whole series.
* One issue with mathematical transformations is that the back-transformed forecast will not be the mean of the forecast distribution. Therefore, the result might need to be adjusted to ensure that the mean of the distribution will be recovered. This process is called bias-adjusting.