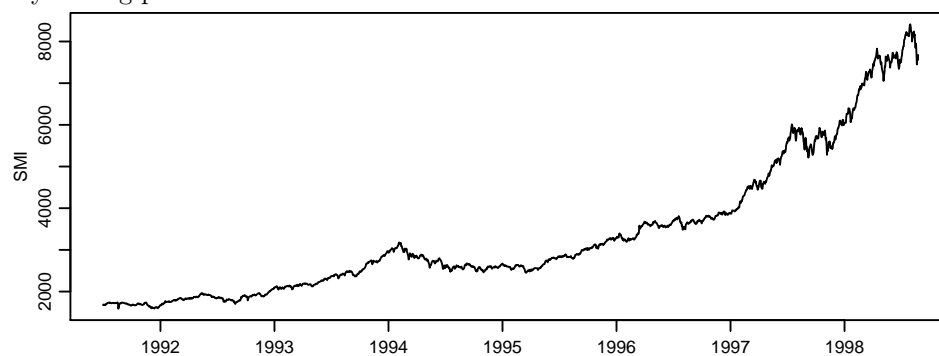


Series 1

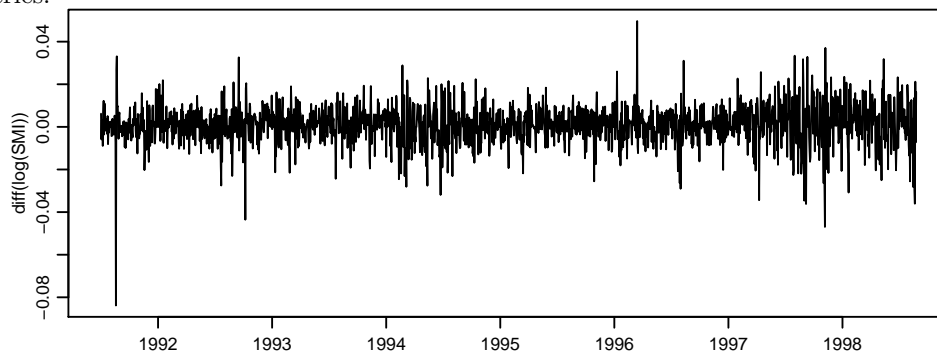
1. The Dataset `rain.txt` contains quarterly rainfall data (in mm) from the year 1965 until 1977. The dataset can be found here: <http://stat.ethz.ch/Teaching/Datasets/WBL/>
 - a) Read in the data and define it correctly as a time series of class `ts`.
R-Hint: `ts.d <- ts(data= ..., ...)`
 - b) Plot the time series.

2. Here are some examples of time series. Decide on the frequency and *delat*.
 - a) Sunshine duration per month in Basel from 1990 to 2000.
 - b) Number of newborn babies in the city of Zurich per year from 2000 to 2011.
 - c) Time Series with the number of reservations in a restaurant for every night during 4 weeks.
 - d) Time Series with the water runoff of a river. The data has been collected every day for four years.
 - e) Time Series with the number of reservations in a restaurant for every night during 4 years.

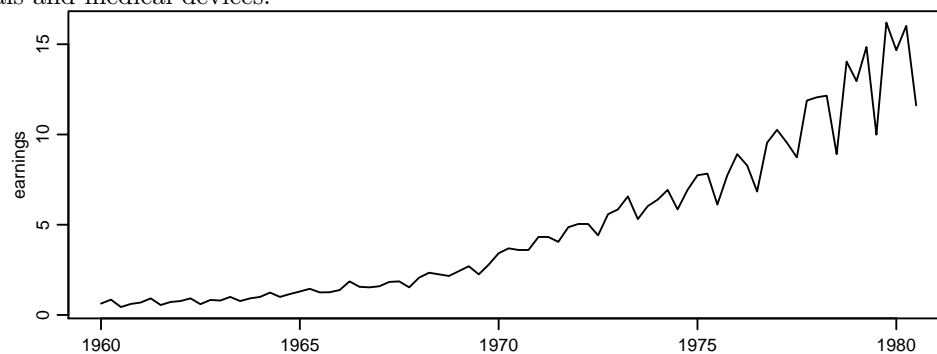
3. Have a look at the following set of time series. For every series, shortly answer the following questions (motivate your answers):
 - Is the time series stationary?
 - Is there a trend?
 - Can one find some seasonal effect? If yes, what is the period?
 - Which transformation should be applied (if required)?
 - a) Daily closing prices of the Swiss Market Index:



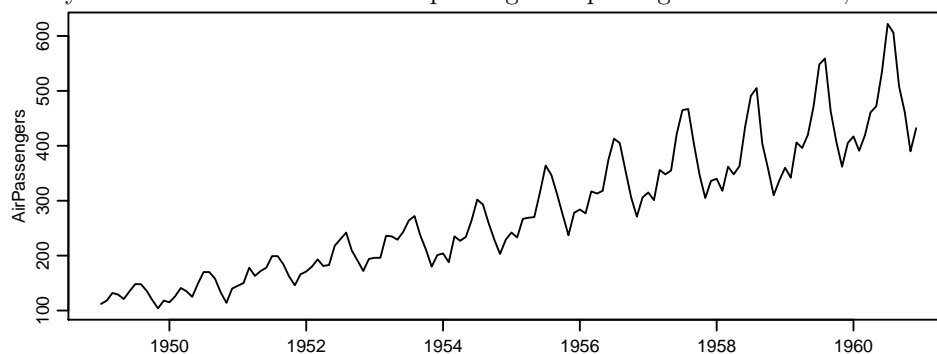
- b) The underlying time series is the same as in Part a), the daily closing prices of the Swiss Market Index. Here, the data is plotted after taking differences to lag 1 of the logarithm of the time series.



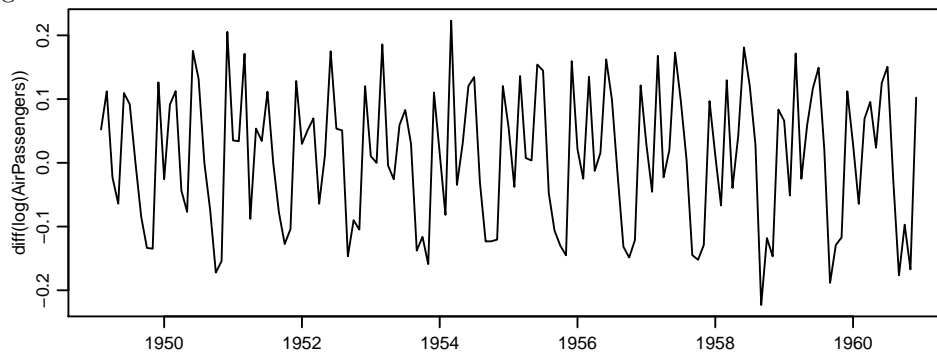
- c) Quarterly earnings per share of Johnson & Johnson, an American manufacturer of pharmaceuticals and medical devices:



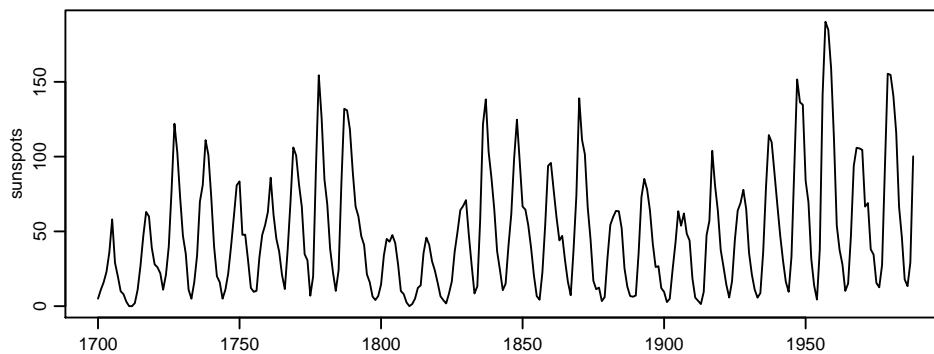
- d) Monthly totals of international airline passengers departing from the USA, 1949 to 1960:



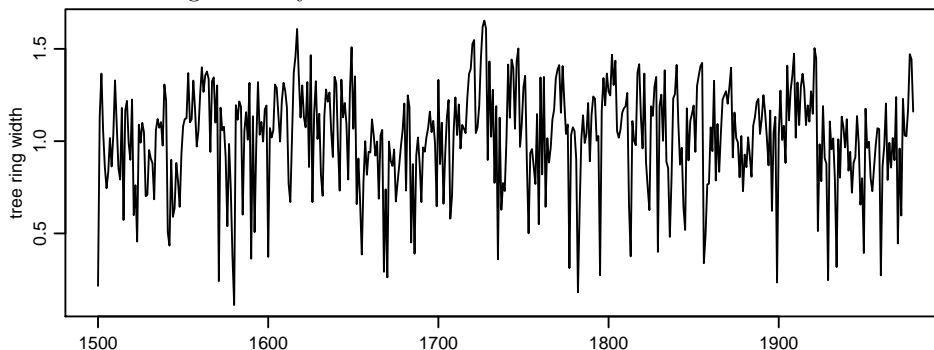
- e) Again the “air passengers” time series of Part d), plotted after taking differences to lag 1 of the logarithm of the time series.



- f) Yearly numbers of sunspots:



g) Widths of tree rings in the years from 1500 to 1979:



4. We would like to illustrate various methods for descriptive decomposition and elimination of trends using the dataset `hstart`.

This dataset contains monthly data on the start of residential construction in the USA within the time frame of January 1966 to January 1974. The data have undergone some transformation unknown to us (perhaps an index over some baseline value has been calculated, or perhaps the data are to be read as $x \cdot 10^7$ construction permits). In our opinion, this makes these data a good didactic choice for illustrating the theory.

(Source: U. S. Bureau of the Census, *Construction Reports*.)

Importing the data (without `header=T!`) and preparing them:

```
hstart <- read.table("http://stat.ethz.ch/Teaching/Datasets/WBL /hstart.dat")
hstart <- ts(hstart[,1], start=1966, frequency=12)
```

- a) Make a time series plot. Is this a stationary time series? If not, what kind of non-stationarity is evident? Into which components might this time series be decomposed sensibly?

b) Differences

Try to remove the trend and seasonal effects by computing differences. After removing seasonal effects, choose some linear trend elimination method from the course notes and plot the outcome.

R hint:

```
H.y <- diff(hstart, lag=12) # time series differences  $Y_t = X_t - X_{t-k}$ ,  $lag = k$ 
H.z <- diff(H.y, lag=...)
par(mfrow=c(2,1))
plot(H.y)
plot(H.z)
```

c) STL decomposition

Decompose the time series in trend, seasonal component and remainder using the non-parametric STL method, and add data from this decomposition to the plot from part a).

R hint:

```
The decomposition is made using H.stl <- stl(hstart, s.window="periodic")
```

Note: The smoothing parameter for the seasonal effect is chosen by means of `s.window`. If `s.window="periodic"`, the seasonal effect is estimated by averaging. It is also possible to specify a value for the smoothing parameter (an odd number). Try e.g. `H.stl.var <- stl(hstart, s.window = 15)`, and compare the result of this to `H.stl`. Incidentally, `summary()` can be used for displaying the values of `window`.

The trend estimation parameter can be set using `t.window`. Unlike `s.window`, this argument does have a default value (cf. the help file). Perhaps you could try to vary this parameter as well. The documentation for R and the help files give more details.

Trend, seasonal component and remainder of the STL are stored in

```
H.st$time.series["trend"]
```

```
H.st$time.series["seasonal"]
```

```
H.st$time.series["remainder"]
```

Have a look at the output of `str(H.stl)` for more details.

Plot the remainder from the STL-Decomposition vs. time.