Solution to Series 1

1. `> dd <- read.table("http://stat.ethz.ch/Teaching/Datasets/WBL/rainDay.txt", header=T)`

a) Read in the data `rainDay.txt` and tell R that the column `DATE` is a date.

```
> dd$DATE <- as.Date(dd$DATE, format="%d.%m.%Y")
```

b) Define your data (without the `DATE`-column) correctly as a time series of class `ts`.
Since it is daily data, the frequency is 365.

```
> ts.dd <- ts(dd[,2], start=2000, freq=365)
> str(ts.dd)
 Time-Series [1:2922] from 2000 to 2008: 0 12.9 0 0.05 3.55 2.05 3.5 7.65 1.1 9.8 ...
```

c) Use the R-Functions `weekdays()`, `months()` and `quarters()` to create these factors. Combine them together with the rainfall data and the date into one dataframe.

```
> t.weekday <- factor(weekdays(dd$DATE, abbreviate=TRUE))##,levels=c("Mon","Tue","Wed","Thu","
> t.month <- factor(months(dd$DATE, abbreviate=TRUE))## levels=c("Jan","Feb","Mar","Apr","May"
> t.quarter <- quarters(dd$DATE, abbreviate=FALSE)
> dd.new  <- data.frame(date=dd$DATE, rain=ts.dd, weekday=t.weekday, month=t.month,
                        quarter=t.quarter)
```

2. **a)** Sunshine duration per month in Basel from 1990 to 2000.
This is monthly data, so the frequency is 12 and `deltat` $\frac{1}{12}$.

b) Number of newborn babies in the city of Zurich per year from 2000 to 2011.
The frequency is 1, `deltat` is 1 as well.

c) Number of reservations in a restaurant for every night during 4 weeks.
An obvious time unit would be one week, so frequency is 7 and `deltat` is $\frac{1}{7}$.

d) Water runoff of a river. The data has been collected every day for four years.
The time unit here would be one year. Frequency is 365 and `deltat` is $\frac{1}{365}$.

e) Number of reservations in a restaurant for every night during 4 years.
It depends on wether we have a seasonal effect. If there are no obvious differences throughout the year (Christmas, Summer Holidays), then we can say one week is one time unit. Otherwise it would be one year. The crucial question is: "How long does it take until I get similar data?"

3. **a)** The series is non-stationary. There is a nonlinear trend, mostly decreasing. There is no seasonal component.

b) The trend from before was removed by differencing. We achieved an approximately constant variance by taking the logarithm. Although the series thus might seem to be stationary, it is known, that the Log-Returns of financial time series are NOT stationary, since there are volatility clusters, e.g. $Var(X_t) \neq Var(X_t|X_{t-1}, X_{t-2}, ...)$, where X_t are the Log-Returns.

c) The series is non-stationary. There is an (exponentially) increasing trend and a (multiplicative) seasonal component with period 1 year.

d) Dito, as in Task c)

e) The series is non-stationary. There is no trend, but a seasonal effect with period 1 year.

f) The series is probably stationary. There are no trend and seasonal effects, but a (non-seasonal!) periodicity with a period of approximately 11 years. Maybe, the variance could be varying over time.

g) The series is probably stationary. There is no clear trend (maybe a slight one), and no seasonal effect visible.