Temperature

Milica Čudina

First, you read the data from our csv file into a data.frame called `temperature`:

```r
temperature<-read.csv("temperatures.csv")
```

If you want to see what your data.frame looks like, you can click on it in the **Global environment** in the upper right pane. The data.frame will get displayed in the upper left pane.

If you interested in the types and names of the variables in your data.frame, you can also run:

```r
ls.str(temperature)
```

```
## BodyTemp :  num [1:50] 97.6 99.4 99 98.8 98 98.9 99 97.8 96.8 99 ...
```

How many measurements are there?

```r
temp<-temperature$BodyTemp
n=length(temp)
print(n)
```

```
## [1] 50
```

Seeing as the data are numerical, we could do a summary:

```r
summary(temp)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   96.40   97.80   98.20   98.26   98.80  100.80
```
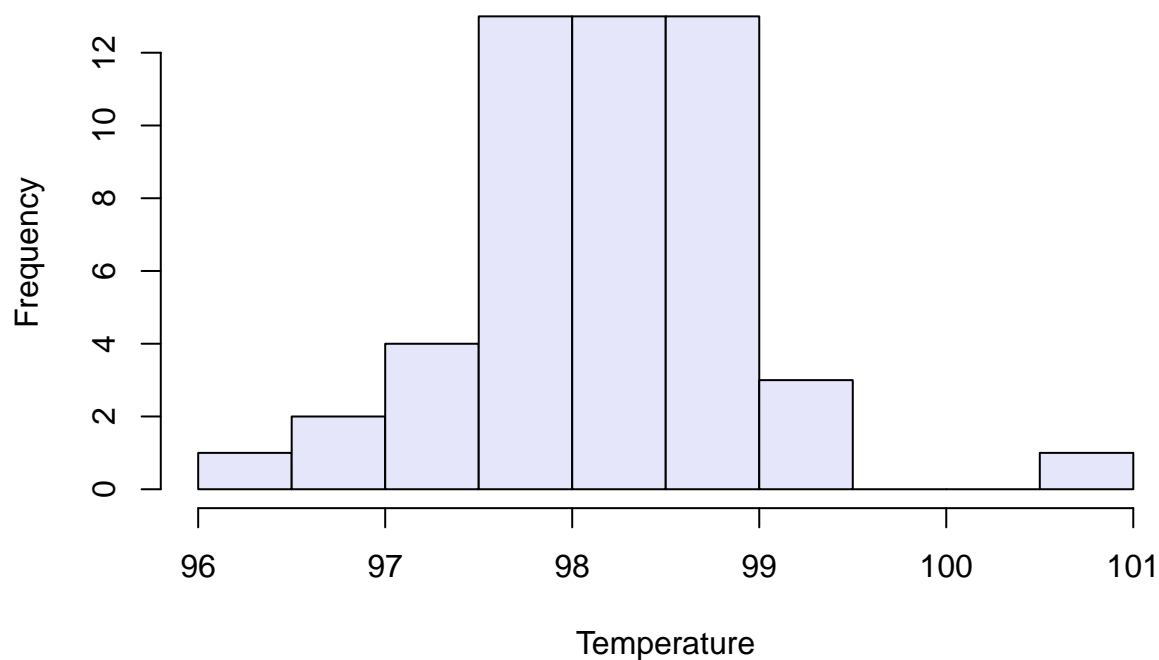
```r
sd(temp)
```

```
## [1] 0.7653197
```

The acutal first bit of exploratory data analysis would be to check out the histogram:

```r
hist(temp,
     main="Human temperatures",
     xlab="Temperature",
     col="lavender")
```

Human temperatures

I do not see the histogram to be unimodal or symmetric, however the data are plentiful enough at 50 for me to think that it's okay to use $t-$procedures. With the built-in `t.test`, I can figure out the confidence interval, e.g.
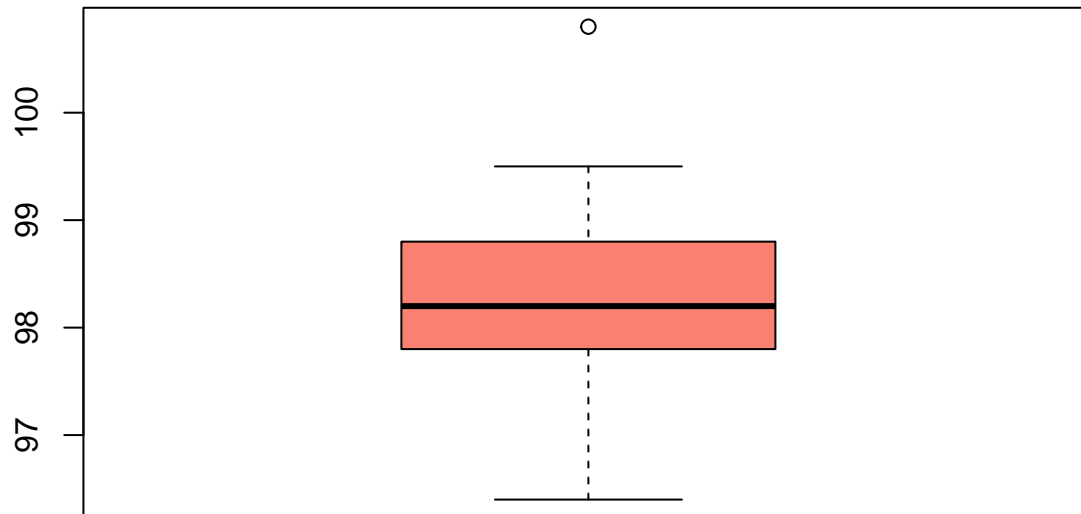
```
t.test(temp)
```

```
##
##  One Sample t-test
##
## data:  temp
## t = 907.86, df = 49, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  98.0425 98.4775
## sample estimates:
## mean of x
##    98.26
```

Note the estimated mean $\hat{\mu} = 98.26$. The 95%-confidence interval is $(98.0425, 98.4775)$.
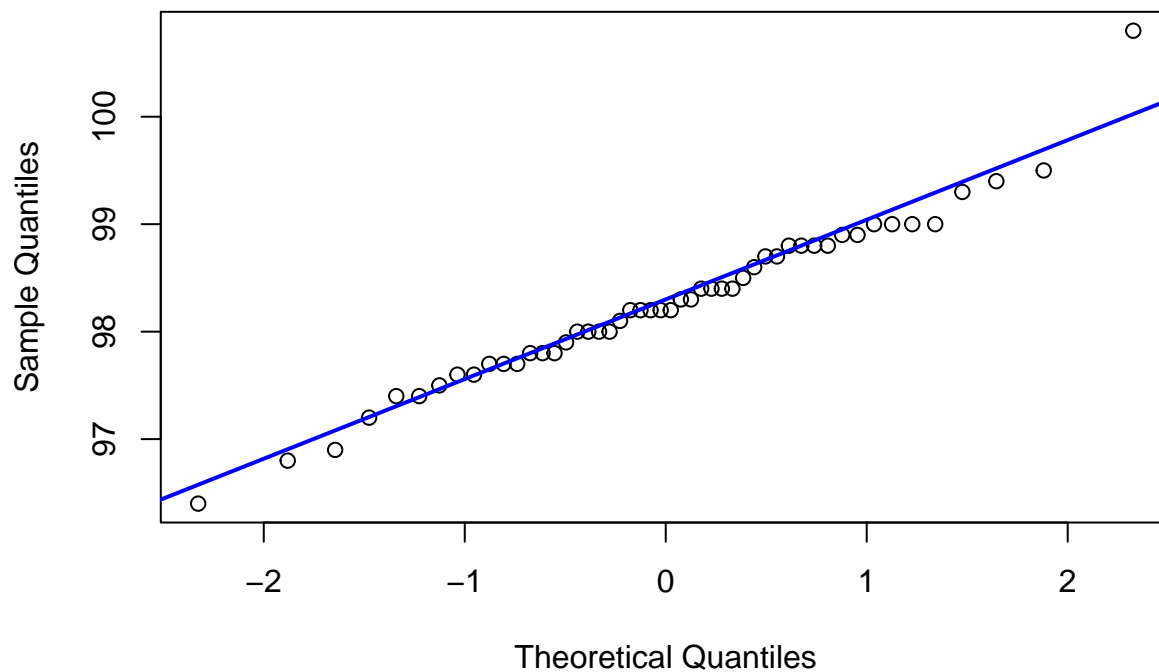
Now, we might be interested in normality, nonetheless. We can check it out further visually:

```
boxplot(temp,
        col="salmon")
```

```
qqnorm(temp)
qqline(temp, col="blue", lwd=2)
```

Normal Q–Q Plot

We can also test for normality using Shapiro-Wilk:
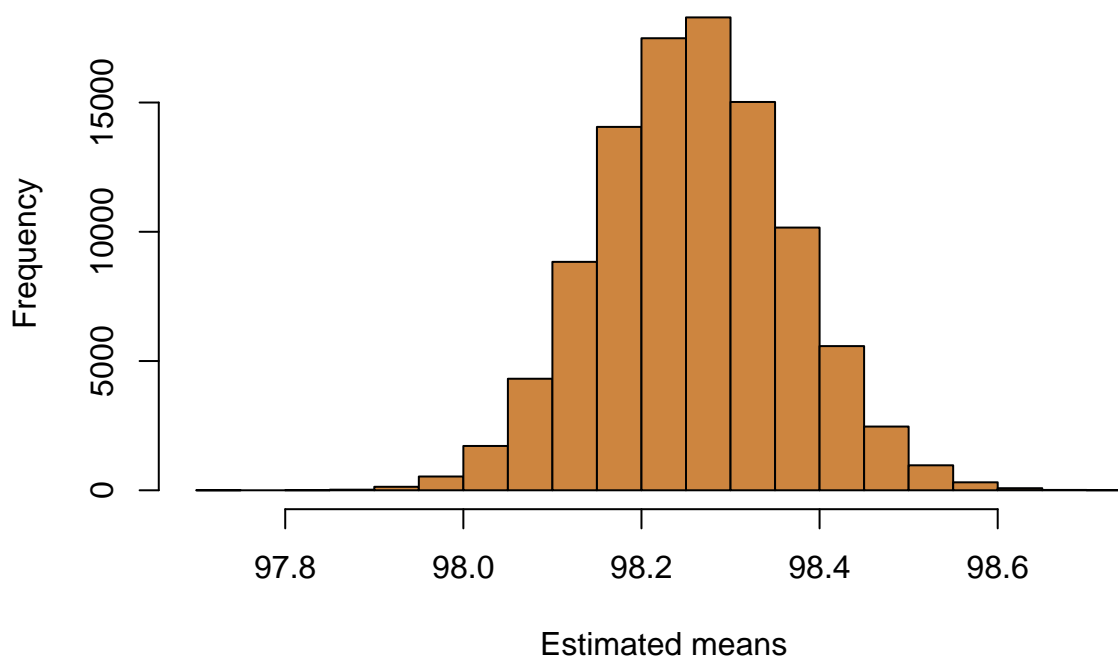
```
shapiro.test(temp)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  temp
## W = 0.97322, p-value = 0.3115
```

Since the $p-$value is sizeable, there is not evidence to reject the normality hypothesis.

So, let's check if *bootstrap* would yield a similar confidence interval. As a rule, it's good to see if a "new tool" conforms with the results of an "old tool" in a setting when they both apply.

```
n.boot=10^5
mu.hats=replicate(n.boot, mean(sample(temp, n, replace=TRUE)))
#(mu.hats)
hist(mu.hats,
     main="Histogram of estimates of the mean",
     xlab="Estimated means",
     col="peru")
```

Histogram of estimates of the mean


```
mu.hat=mean(mu.hats)
print(mu.hat)
```

```
## [1] 98.25951
```

Of course, the boot $\hat{\mu}_{BOOT}$ conforms to the $\hat{\mu}$ from above. What about the confidence intervals?

First, the quantile $95\%-$confidence interval:

```
quantile(mu.hats, c(0.025, 0.975))
```

```
##   2.5%  97.5%
## 98.052 98.474
```

Second, the $2SE$ confidence interval:

```
se=sd(mu.hats)
(mu.hat-se*qnorm(0.975))
```

```
## [1] 98.04944
```

```
(mu.hat+se*qnorm(0.975))
```

```
## [1] 98.46958
```