

```

data<-read.table("http://www.stat.uni-
muenchen.de/service/datenarchiv/miete/miete03.asc",header=T)
attach(data)
names(data)
[1] "nm"      "nmqm"    "wfl"     "rooms"   "bj"      "bez"     "wohngut" "wohnbest" "ww0"
"zh0"     "badkach0" "badextra" "kueche"

```

```

#####
## Marginale Korrelation
#####

```

```

cor(nm,wfl)
[1] 0.7074627
cor(rooms,wfl)
[1] 0.8406454
cor(nm,rooms)
[1] 0.5442473

```

```

#Die marginale Korrelation zwischen nm und wfl und rooms und wfl ist deutlich
#positiv zwischen nm und rooms eher mittel positiv

```

```

#####
## Bedingte Korrelation
#####

```

```

#Datensatz wird bzgl. der Faktorvariable rooms aufgesplittet (1 bis 6 Zimmer):
#man erhält eine Liste
dat<-split(as.data.frame(cbind(nm,wfl)),as.factor(rooms))

```

```

> head(dat)
$`1`
      nm wfl
9    446.33 52
11   337.26 35
14   264.93 23
28   354.91 38
...

$`2`
      nm wfl
1    741.39 68
2    715.82 65
7    204.85 55
12   756.73 68
...

$`3`
      nm wfl
3    528.25 63
4    553.99 65
8    426.93 79
10   381.45 77
...

$`4`
      nm wfl
5    698.21 100
6    935.65 81
16   757.74 82
24   759.57 89
...

$`5`
      nm wfl
80   526.70 131
129  619.94 115
203 1113.78 140
260  594.40 105
...

```

```

$`6`
      nm wfl
18    796.07 178
62   1012.86 116
128   639.13 166
357  1344.72 125
...

#über diese Liste werden mit lapply nun die Korrelationen zw. nm und wfl
#geschätzt: bedingte Korrelationsn
lapply(dat,cor)

$`1`
      nm      wfl
nm 1.0000000 0.6732726
wfl 0.6732726 1.0000000

$`2`
      nm      wfl
nm 1.0000000 0.5520034
wfl 0.5520034 1.0000000

$`3`
      nm      wfl
nm 1.0000000 0.5716983
wfl 0.5716983 1.0000000

$`4`
      nm      wfl
nm 1.0000000 0.5645286
wfl 0.5645286 1.0000000

$`5`
      nm      wfl
nm 1.0000000 0.4742612
wfl 0.4742612 1.0000000

$`6`
      nm      wfl
nm 1.0000000 0.1682122
wfl 0.1682122 1.0000000

#Die bedingte Korrelation nimmt mit der Anzahl der Zimmer tendenziell ab

#####
## Partielle Korrelation bezüglich Zimmer
#####

#Regression von nm bzw. wfl auf rooms (letztere als kategorial Variable; vgl.
#vektorielle Version), um Fehlerterme zu schätzen

errnm <- lm(nm~as.factor(rooms))$residuals
errwfl <- lm(wfl~as.factor(rooms))$residuals

cor(errnm,errwfl)
[1] 0.5507335

#Dasselbe mit rooms als stetige Variable aufgenommen
errnm<-lm(nm~rooms)$residuals
errwfl<-lm(wfl~rooms)$residuals

cor(errnm,errwfl)
[1] 0.5501128

#Der Unterschied zwischen rooms als kategoriale oder metrische Variable ist
#zu vernachlässigen. Nach herauspartialisieren des linearen Anteils von rooms
#nimmt die Korrelation von wfl und nm (ggü. der marginalen Korrelation)
#ab (0.7 auf 0.55).

detach(data)

```