

R-Einführung Teil 3

Inhalt:

1) Korrelationen

2) Wilcoxon-Rangsummen-Test

3) Chi-Quadrat-Test

1) Korrelationen

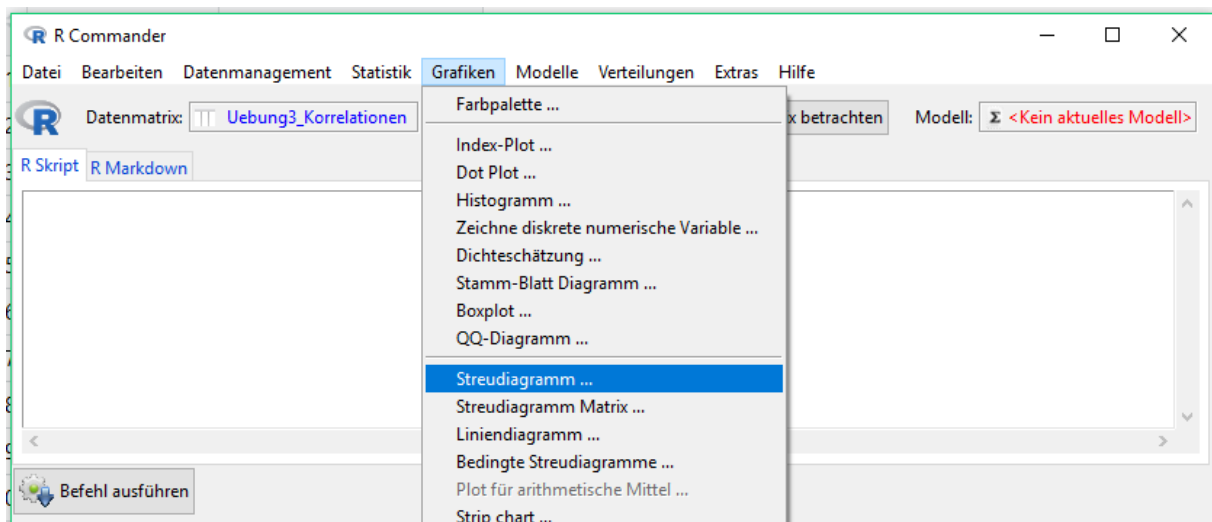
Datensatz: Uebung3_Korrelationen.csv

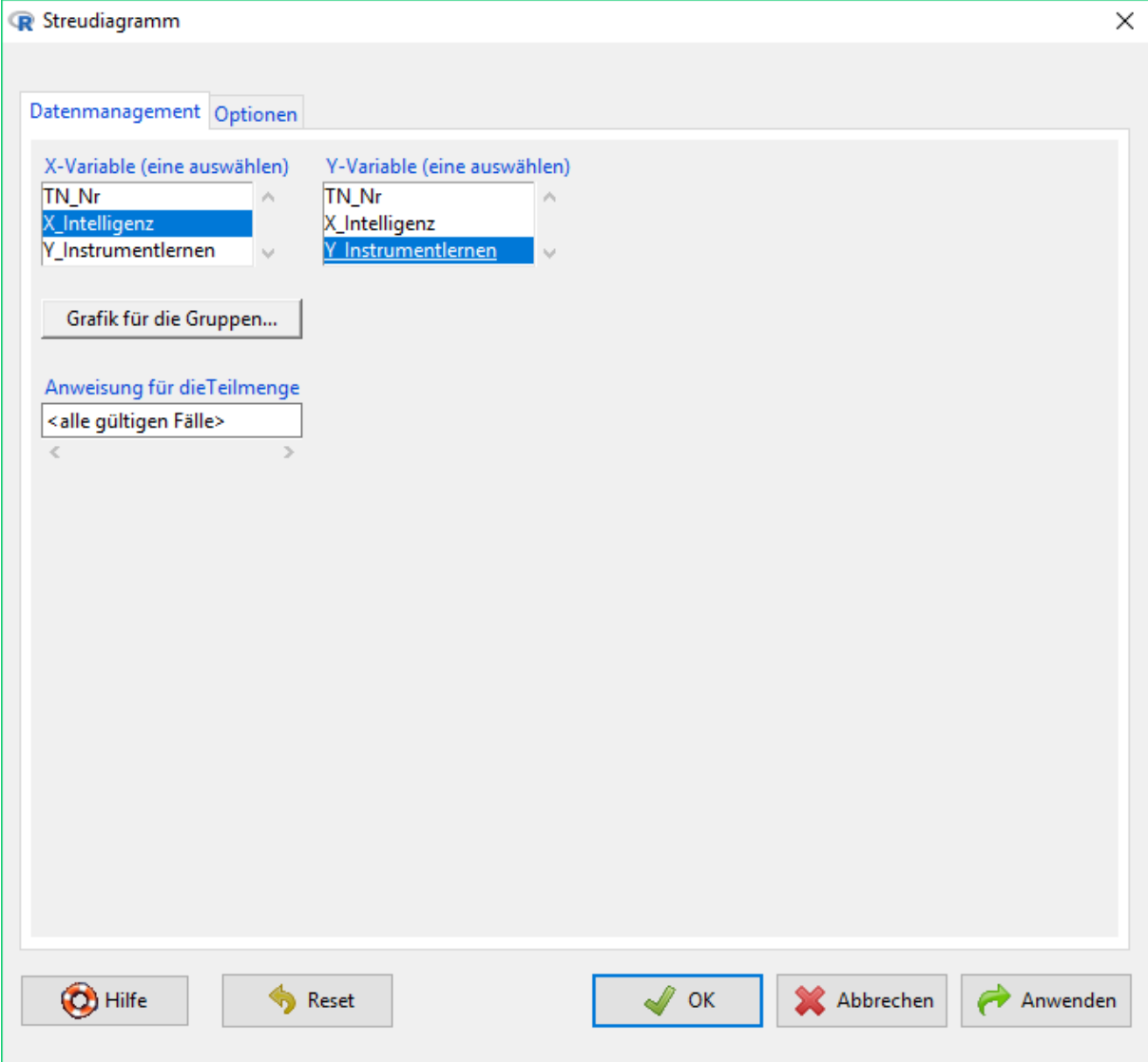
Inhalt: Zusammenhang Intelligenz und Instrumentlernen, siehe letztes Semester, Sitzung zu Korrelationen

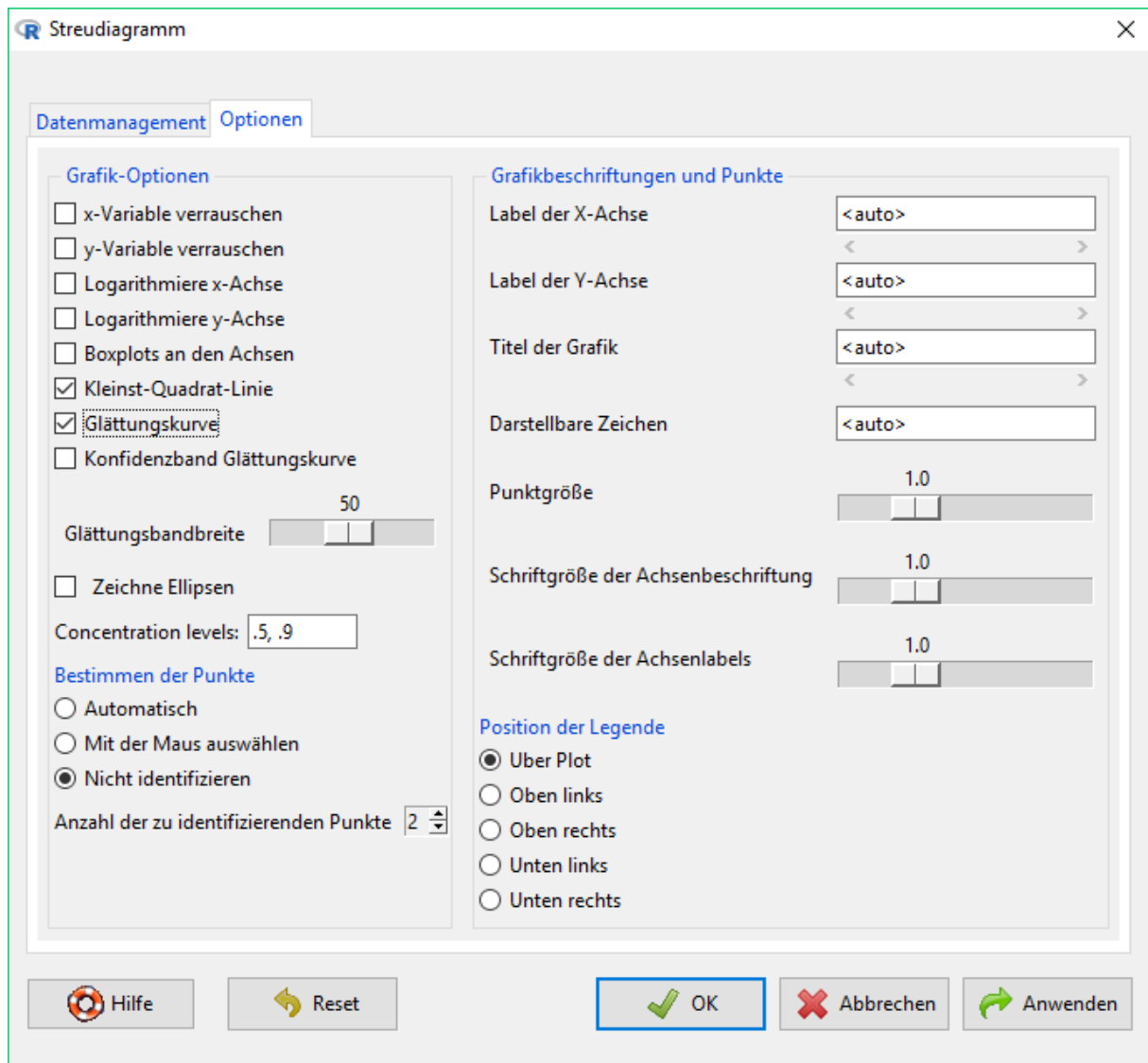
Ausgangspunkt:

- Die Daten liegen aufbereitet und im *Wide*-Format vor
- die CSV-Datei wurde in R-Studio geladen
- R-Commander wurde geladen
- die entsprechende Datenmatrix wurde in RCommander ausgewählt.

1.1 Streudiagramm erstellen für ersten Überblick (Commander)





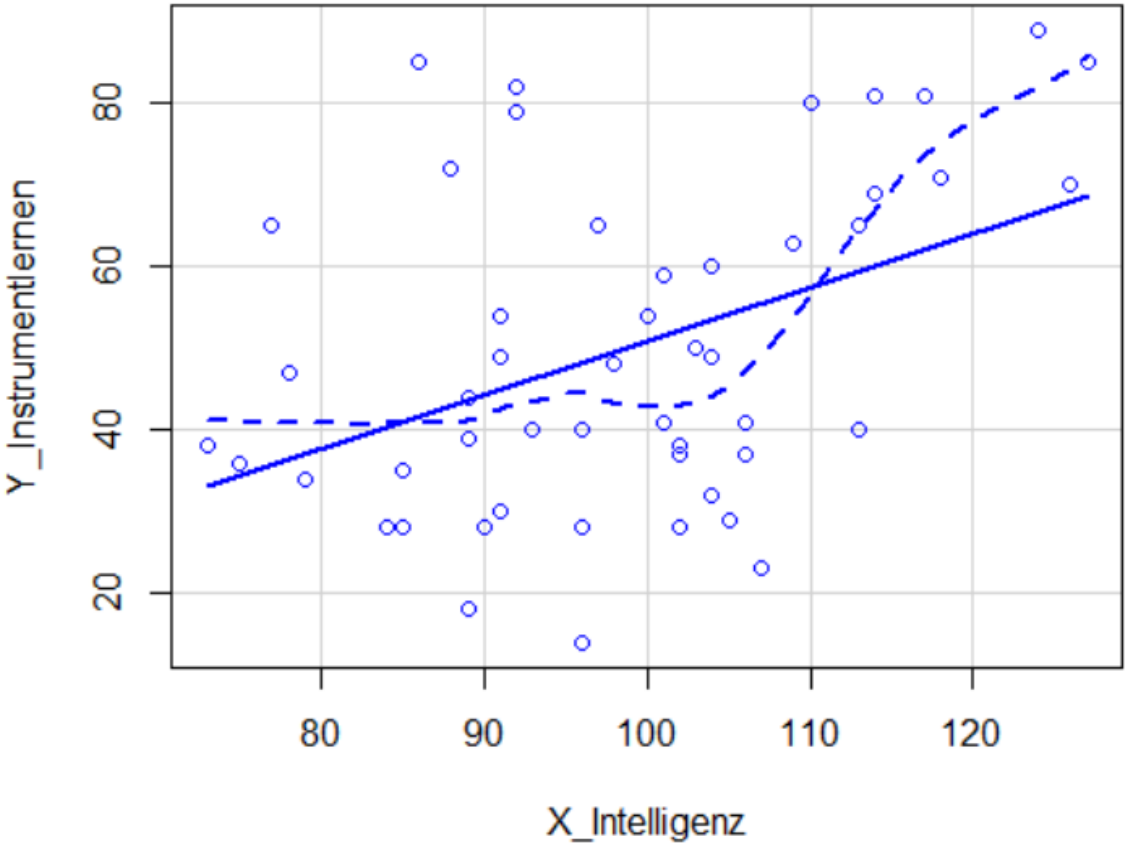


- Kleinst-Quadrat-Linie: Die Linie, von der die Datenpunkte am wenigsten abweichen
- Glättungskurve: Kurve, von der die Datenpunkte am wenigsten abweichen

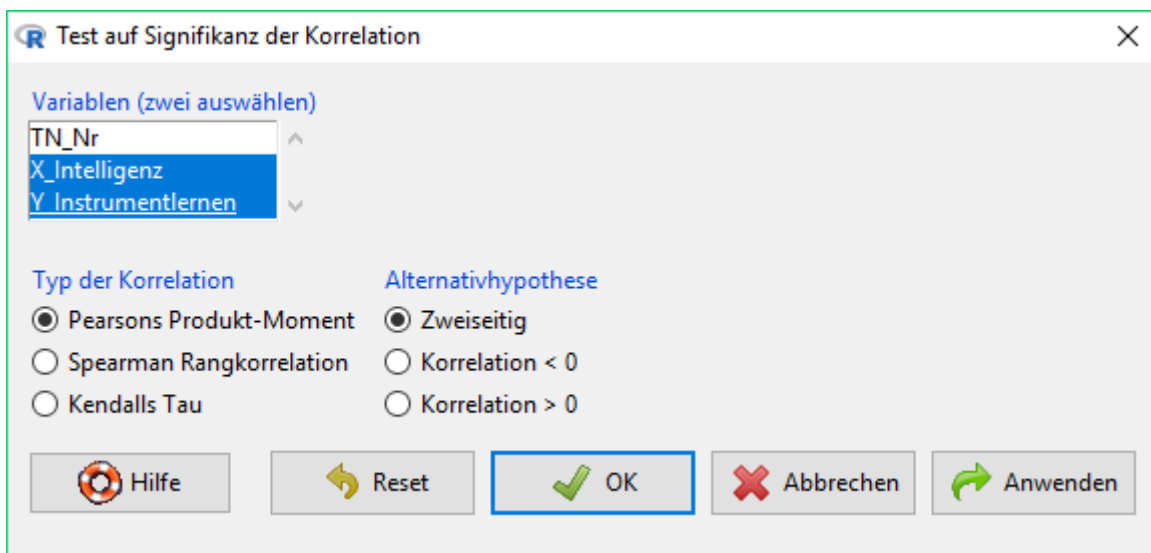
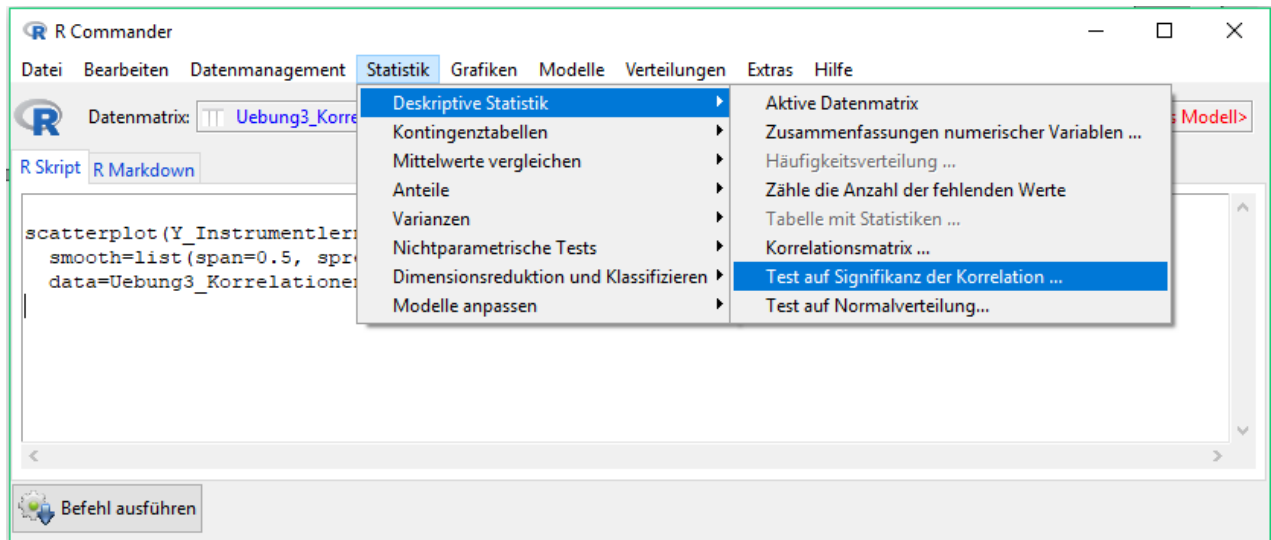
Resultierender Code:

```
scatterplot(Y_Instrumentlernen~X_Intelligenz, regLine=TRUE,
smooth=list(span=0.5, spread=FALSE), boxplots=FALSE,
data=Uebung3_Korrelationen)
```

Output:



1.2 Berechnen einer Korrelation und Test auf Signifikanz (Commander)



Resultierender Code:

```
with(Uebung3_Korrelationen, cor.test(X_Intelligenz, Y_Instrumentlernen,
alternative="two.sided", method="pearson"))
```

Output:

Pearson's product-moment correlation

data: X_Intelligenz and Y_Instrumentlernen

t = 3.2706, df = 48, p-value = 0.001991

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.1685717 0.6303378

sample estimates:

cor

0.4268891

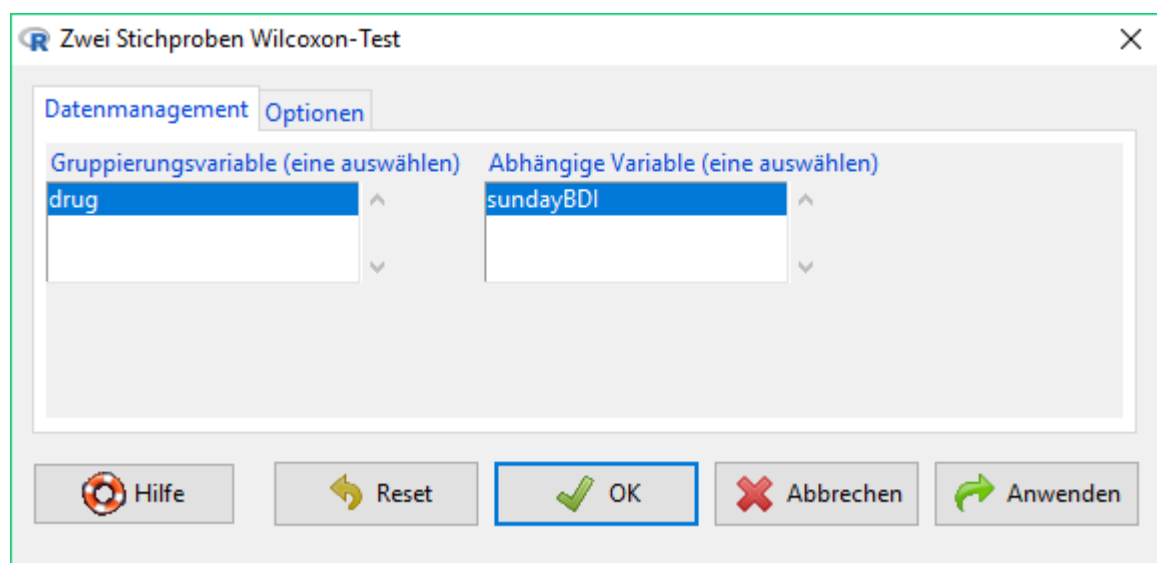
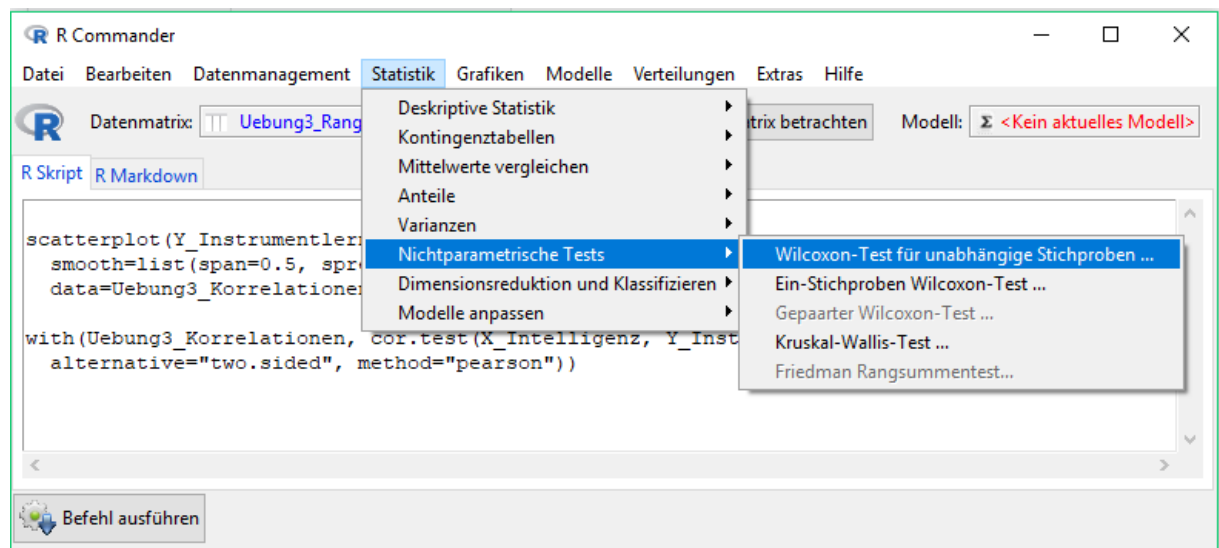
2) Wilcoxon-Rangsummentest

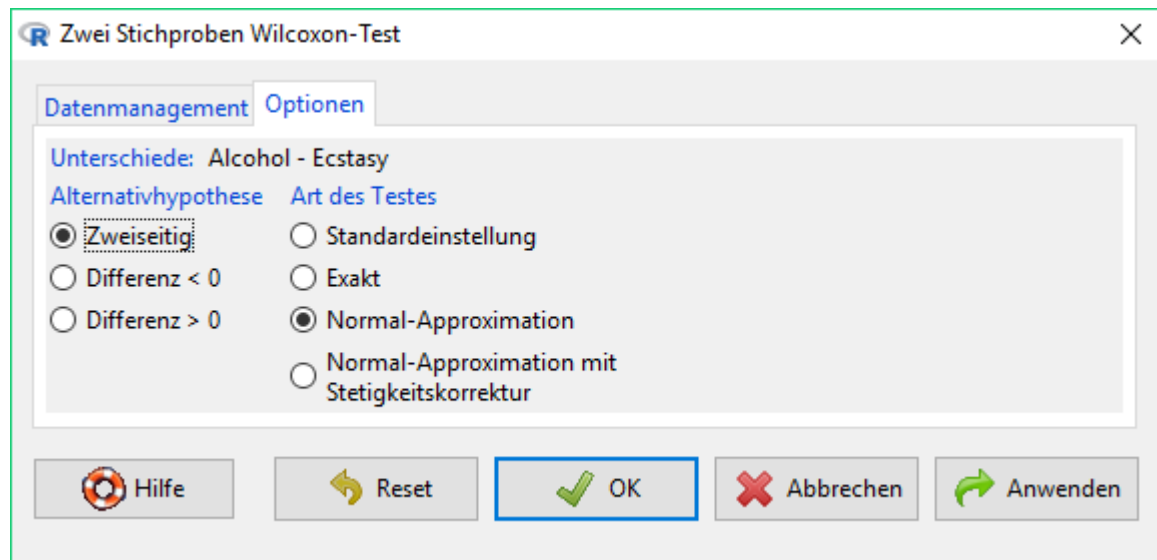
Datensatz: Uebung3_Rang.csv

Inhalt: Werte auf Beck Depression Inventory nach Konsum von Alkohol vs. Ecstasy (siehe Andy Field, Discovering statistics using R, Kapitel15)

Ausgangspunkt:

- Die Daten liegen aufbereitet und im *Long*-Format vor
- die CSV-Datei wurde in R-Studio geladen
- R-Commander wurde geladen
- die entsprechende Datenmatrix wurde in RCommander ausgewählt.





Art des Tests: Ein Blick in die Hilfe für die Funktion `wilcox.test` verrät uns, dass der exakte Test verwendet werden kann, sofern keine Rangbindungen vorliegen. Ansonsten sollte die Normal-Approximation verwendet werden (siehe Sitzung zu Rangtests im letzten Semester).

Files **Plots** **Packages** **Help** **Viewer**

Search: wilcox.test

R: Wilcoxon Rank Sum and Signed Rank Tests

wilcox.test {stats} R Documentation

Wilcoxon Rank Sum and Signed Rank Tests

Description

Performs one- and two-sample Wilcoxon tests on vectors of data; the latter is also known as 'Mann-Whitney' test.

Usage

```
wilcox.test(x, ...)
```

```
## Default S3 method:
wilcox.test(x, y = NULL,
            alternative = c("two.sided", "less", "greater"),
            mu = 0, paired = FALSE, exact = NULL, correct = TRUE,
            conf.int = FALSE, conf.level = 0.95, ...)
```

```
## S3 method for class 'formula'
wilcox.test(formula, data, subset, na.action, ...)
```

Arguments

By default (if `exact` is not specified), an exact p-value is computed if the samples contain less than 50 finite values and there are no ties. Otherwise, a normal approximation is used.

Generierter Code:

```
with(Uebung3_Rang, tapply(sundayBDI, drug, median, na.rm=TRUE))
wilcox.test(sundayBDI ~ drug, alternative='two.sided', exact=FALSE,
  correct=FALSE, data=Uebung3_Rang)
```

Output:

```
Rcmdr> with(Uebung3_Rang, tapply(sundayBDI, drug, median, na.rm=TRUE))
Alcohol Ecstasy
  16.0    17.5
```

```
Rcmdr> wilcox.test(sundayBDI ~ drug, alternative='two.sided', exact=FALSE,
Rcmdr+   correct=FALSE, data=Uebung3_Rang)
```

```
Wilcoxon rank sum test
```

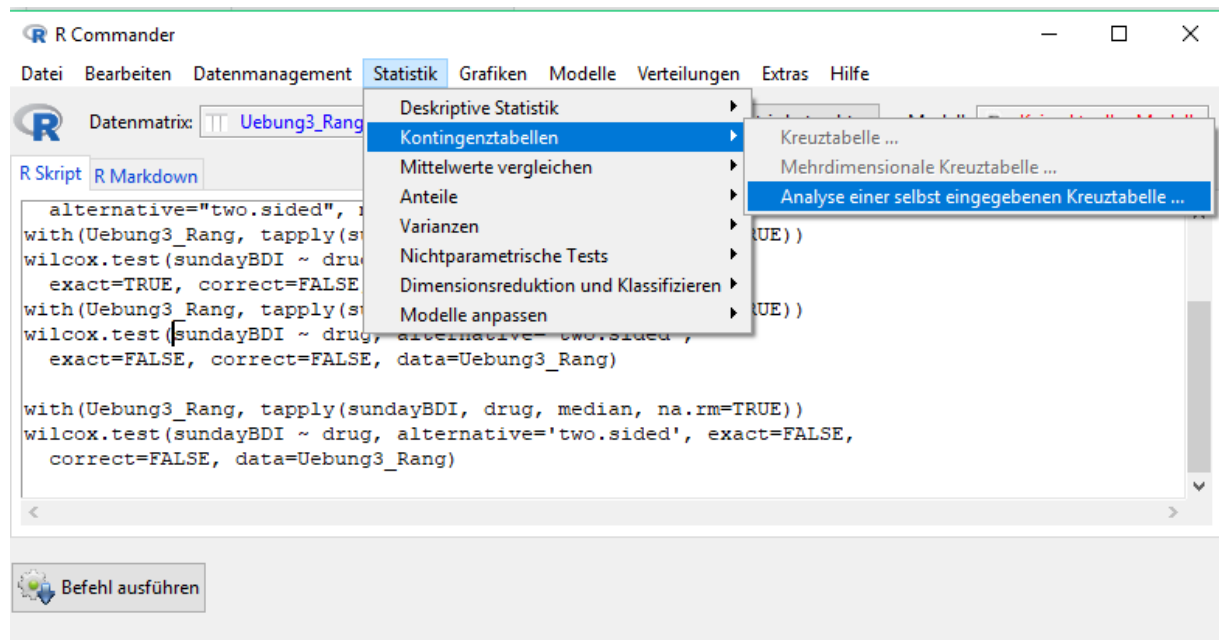
```
data:  sundayBDI by drug
W = 35.5, p-value = 0.2692
alternative hypothesis: true location shift is not equal to 0
```

3) Chi-Quadrat-Test über Häufigkeiten

Ausgangspunkt: Häufigkeitstabelle. Die Daten müssen nicht aufbereitet werden, sondern können direkt in den Commander eingegeben werden.

Beispiel: Häufigkeit von Therapieabbruch nach Ausbildung der Therapeut*innen (siehe letztes Semester, Sitzung zum Chi-Quadrat-Test):

Absolute Häufigkeiten	Psychologische Psychotherapeuten	Medizinische Psychotherapeuten	Psychiater
Abbruch	2	8	11
Kein Abbruch	38	32	33



Eingabe der Kreuztabelle

Tabelle Statistik

Name der Zeilenvariable (optional): Abbruch

Namen der Spaltenvariable (optional): Ausbildung

Anzahl der Zeilen: 2

Anzahl der Spalten: 3

Häufigkeiten eingeben:

	PPT	MPT	P
Ja	2	8	11
Nein	38	32	33

Hilfe Reset OK Abbrechen Anwenden

Eingabe der Kreuztabelle

Tabelle Statistik

Berechne Prozente

Zeilenprozente

Spaltenprozente

Prozentuierung auf die gesamte Tabelle

Keine Prozente

Hypothesentest

Chi-Quadrat-Test auf Unabhängigkeit

Komponenten der Chi-Quadrat Statistik

Zeige erwartete Häufigkeiten

Fishers exakter Test

Hilfe Reset OK Abbrechen Anwenden

(Anmerkung: Es kann auch Fishers exakter Test gewählt werden)

Generierter Code:

```
library(abind, pos=17)
.Table <- matrix(c(2,8,11,38,32,33), 2, 3, byrow=TRUE)
dimnames(.Table) <- list("Abbruch"=c("Ja", "Nein"), "Ausbildung"=c("PPT",
  "MPT", "P"))
.Table # Counts
.Test <- chisq.test(.Table, correct=FALSE)
.Test
remove(.Test)
remove(.Table)
```

Output:

```
Rcmdr> library(abind, pos=17)
Rcmdr> .Table <- matrix(c(2,8,11,38,32,33), 2, 3, byrow=TRUE)
Rcmdr> dimnames(.Table) <- list("Abbruch"=c("Ja", "Nein"), "Ausbildung"=c("PPT",
Rcmdr+   "MPT", "P"))
Rcmdr> .Table # Counts
      Ausbildung
Abbruch PPT MPT P
   Ja    2  8 11
   Nein  38 32 33
Rcmdr> .Test <- chisq.test(.Table, correct=FALSE)
Rcmdr> .Test
      Pearson's Chi-squared test

data:  .Table
X-squared = 6.3519, df = 2, p-value = 0.04175
```