Übung 2t – Test

Für die nachfolgenden Aufgaben sollen die Daten zuerst mit Werkzeugen der Deskriptiven Statistik untersucht werden.

Für alle Tests: Signifikanzniveau 1 – α = 95%

Die Ergebnisse dieser Übungen werden diesmal durch die Teilnehmer präsentiert, also müssen Sie alle Aktivitäten ausreichend dokumentieren! (Es gibt keine Musterlösung)

Übung 2t – Test

Abfüllung

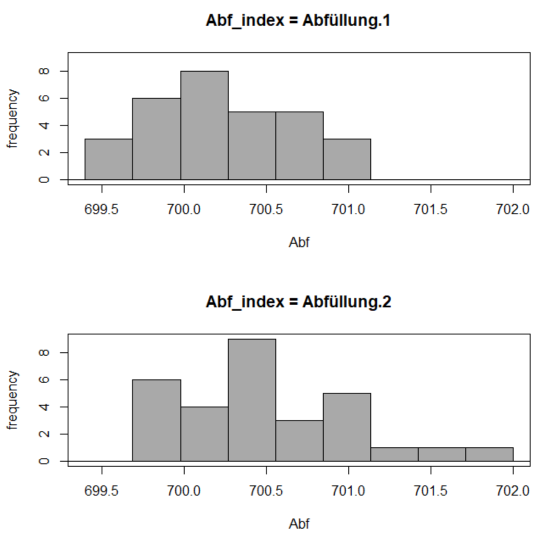
In einer Getränkeabfüllung sollen auf zwei Maschinen Flaschen mit 700 ml Inhalt abgefüllt werden.
Es besteht der Verdacht, dass die Maschinen unterschiedliche Füllmengen aufweisen.
Für beide Maschinen wurden mittels Zufallsstichprobe der Inhalt von jeweils 30 Flaschen vermessen.
Datensätze: Abfüllung1 und Abfüllung2 in *Übung_2t.xlsx*

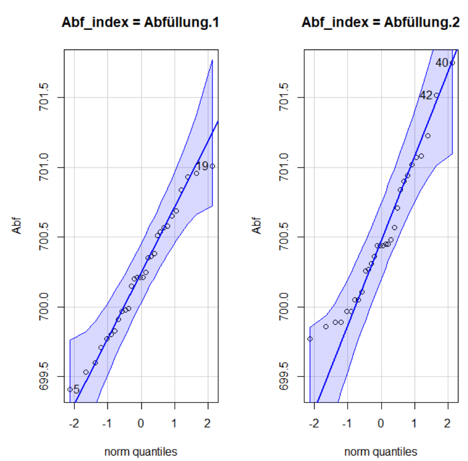
Wie gehen Sie vor? Überprüfen Sie den Verdacht.

Übung 2t – Test

Abfüllung – Grafik

- Das Histogramm für Abfüllung 1 ist unauffällig, Abfüllung 2 sieht ungewöhnlich aus=> Auf Normalität testen
- Die QQ-Diagramme deuten auf Normalverteilung in beiden Stichproben hin

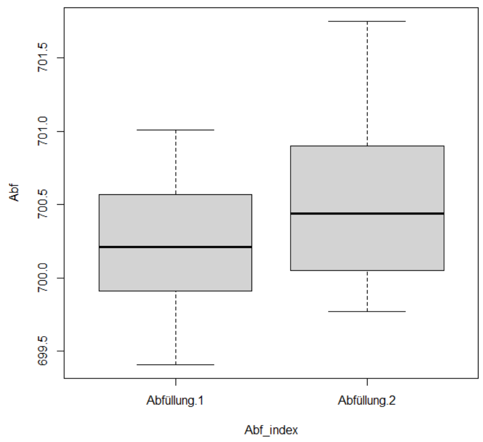




Übung 2t – Test

Abfüllung – Grafik

Aus dem Box Plot ist nicht direkt erkennbar, ob beide Stichproben gleiche Varianz aufweisen=> Varianzen überprüfen oder t-Test mit ungleichen Varianzen



Übung 2t – Test

Abfüllung – Normalverteilung

```
Anderson-Darling normality test

data:  Abf
A = 0.18713, p-value = 0.8957
 --------
Abf_index = Abfüllung.2

Anderson-Darling normality test

data:  Abf
A = 0.50882, p-value = 0.1833

 --------
 p-values adjusted by the Holm method:
            unadjusted adjusted
Abfüllung.1 0.8957     0.89570
Abfüllung.2 0.1833     0.36661
```

- Daten sind normalverteilt

Übung 2t – Test

Abfüllung – Varianz

```
F test to compare two variances

data:  variable by factor
F = 0.73172, num df = 29, denom df = 29, p-value = 0.4053
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.3482752 1.5373508
sample estimates:
ratio of variances
         0.7317248
```

- Es gilt die Nullhypothese, wir können von gleichen Varianzen ausgehen

Übung 2t – Test

Abfüllung – Hypothesen

H_1 **Die Mittelwerte der Abfüllungen sind nicht gleich**
(Die Differenz der Mittelwerte ist ungleich Null)

H_0 **Die Mittelwerte der Abfüllungen sind gleich**
(Die Differenz der Mittelwerte ist gleich Null)

Übung 2t – Test

Abfüllung – 2t-Test für unabhängige Stichproben

```
Two Sample t-test

data:  variable by factor
t = -2.1692, df = 58, p-value = 0.03418
alternative hypothesis: true difference in means between
group Abfüllung.1 and group Abfüllung.2 is not equal to 0
95 percent confidence interval:
 -0.51210335 -0.02056331
sample estimates:
mean in group Abfüllung.1 mean in group Abfüllung.2
                 700.2367                  700.5030
```

p=0,03418 = 3,42% < α Wechsel zur Alternativhypothese:
Die Mittelwerte der Stichproben sind nicht gleich (sie stammen aus verschiedenen Grundgesamtheiten) bzw. die Differenz der Mittelwerte ist ungleich Null.

Übung 2t – Test - Exkurs

Abfüllung – 2t-Test für unabhängige Stichproben

Beleg, dass Mittelwert Abfüllung 2 > Mittelwert Abfüllung 1
Differenz: Mittelwert 2 – Mittelwert 1

H_1 **Mittelwert 2 > Mittelwert 1 (Die Differenz der Mittelwerte ist größer Null)**

H_0 **Mittelwerte 2 <= Mittelwert 1 (Die Differenz der Mittelwerte ist gleich oder kleiner Null)**

Übung 2t – Test - Exkurs

Abfüllung – 2t-Test für unabhängige Stichproben

Beleg, dass Mittelwert Abfüllung 2 > Mittelwert Abfüllung 1
Differenz: Mittelwert 1 – Mittelwert 2

H_1 **Mittelwert 2 > Mittelwert 1 (Die Differenz der Mittelwerte ist kleiner Null)**

H_0 **Mittelwerte 2 <= Mittelwert 1 (Die Differenz der Mittelwerte ist gleich oder größer Null)**

Übung 2t – Test

Detergenzien

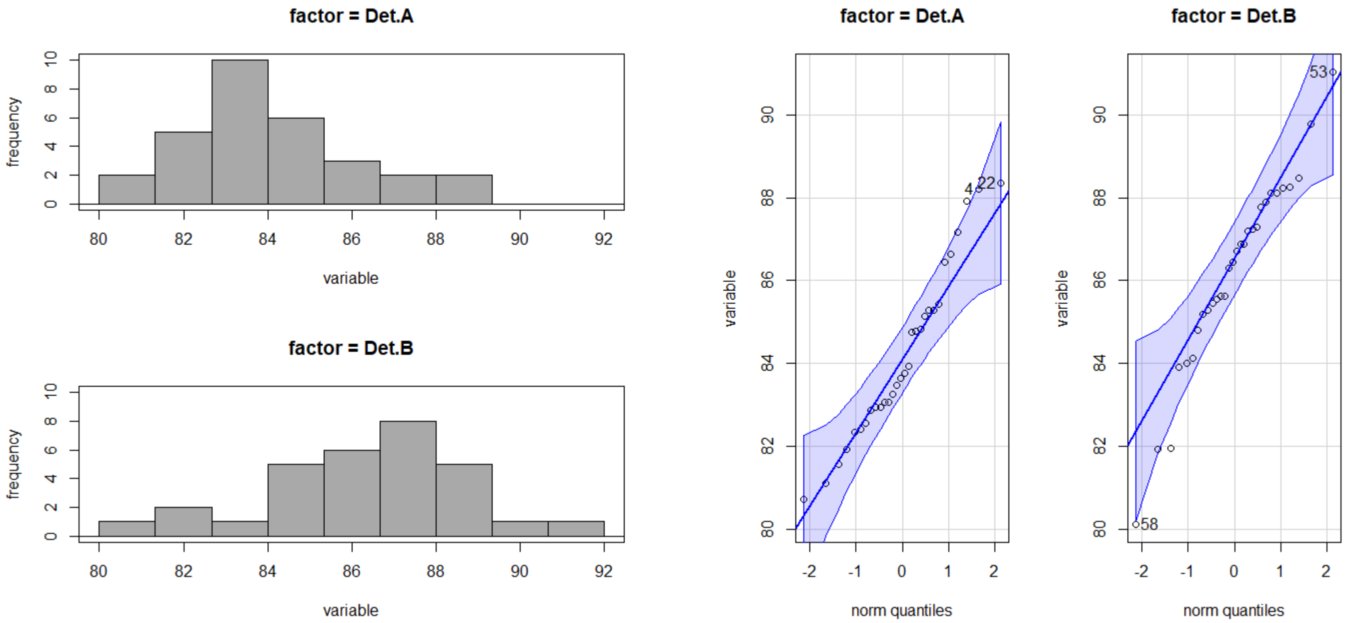
Ein chemischer Prozess soll durch die Zugabe unterschiedlicher Detergenzien verbessert werden. Es stehen zwei verschiedene Produkte (Datensätze: DetA und DetB in *Übung_2t.xlsx*) zur Verfügung, die geprüft werden sollen.

Unterscheiden sich die Stichproben hinsichtlich ihrer Mittelwerte?

Übung 2t – Test

Detergenzien – Grafik

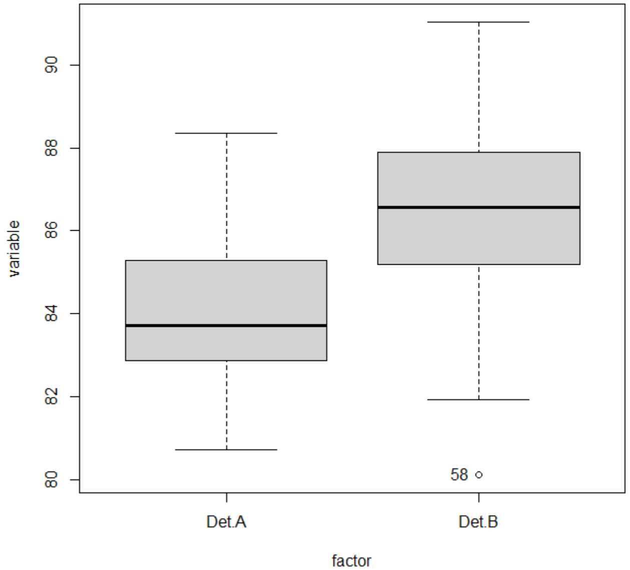
- Histogramm sind unauffällig
- QQ-Diagramme deuten auf Normalverteilung in beiden Stichproben hin

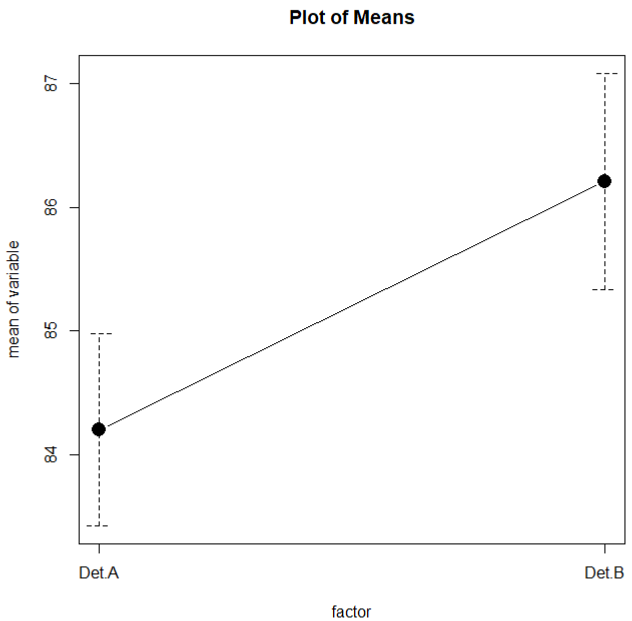


Übung 2t – Test

Detergenzien – Grafik

- Varianz ist auf Basis der Box Plots nicht zu beurteilen
- Plot of means deutet auf unterschiedliche Mittelwerte hin





Übung 2t – Test

Detergenzien – Normalverteilung

```
Shapiro-Wilk normality test
 --------
 p-values adjusted by the Holm method:
      unadjusted adjusted
Det.A 0.22931    0.45862
Det.B 0.41896    0.45862
```

- p-Werte > α: Nullhypothese, die Daten sind normalverteilt

Übung 2t – Test

Detergenzien – Varianz

```
F test to compare two variances

data:  variable by factor
F = 0.78497, num df = 29, denom df = 29, p-value = 0.5186
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.3736179 1.6492177
sample estimates:
ratio of variances
         0.7849695
```

- $p>\alpha$: Es gilt die Nullhypothese, wir können von gleichen Varianzen ausgehen

Übung 2t – Test

Detergenzien – Hypothesen

H_1 **Die Mittelwerte der Stichproben der beiden Detergenzien sind nicht gleich (Die Differenz der Mittelwerte ist ungleich Null)**

H_0 **Die Mittelwerte der Stichproben der beiden Detergenzien sind gleich (Die Differenz der Mittelwerte ist gleich Null)**

Übung 2t – Test

Detergenzien – 2t-Test für unabhängige Stichproben

```
       Two Sample t-test

data:  variable by factor
t = -3.5106, df = 58, p-value = 0.0008724
alternative hypothesis: true difference in means between group Det.A
and group Det.B is not equal to 0
95 percent confidence interval:
 -3.1492926 -0.8620408
sample estimates:
mean in group Det.A mean in group Det.B
           84.20033            86.20600
```

p=0,0008724 = 0,09% < α Wechsel zur Alternativhypothese:
Die Mittelwerte der Stichproben sind nicht gleich (sie stammen aus verschiedenen Grundgesamtheiten) bzw. die Differenz der Mittelwerte ist ungleich Null.

Übung 2t – Test

Lack

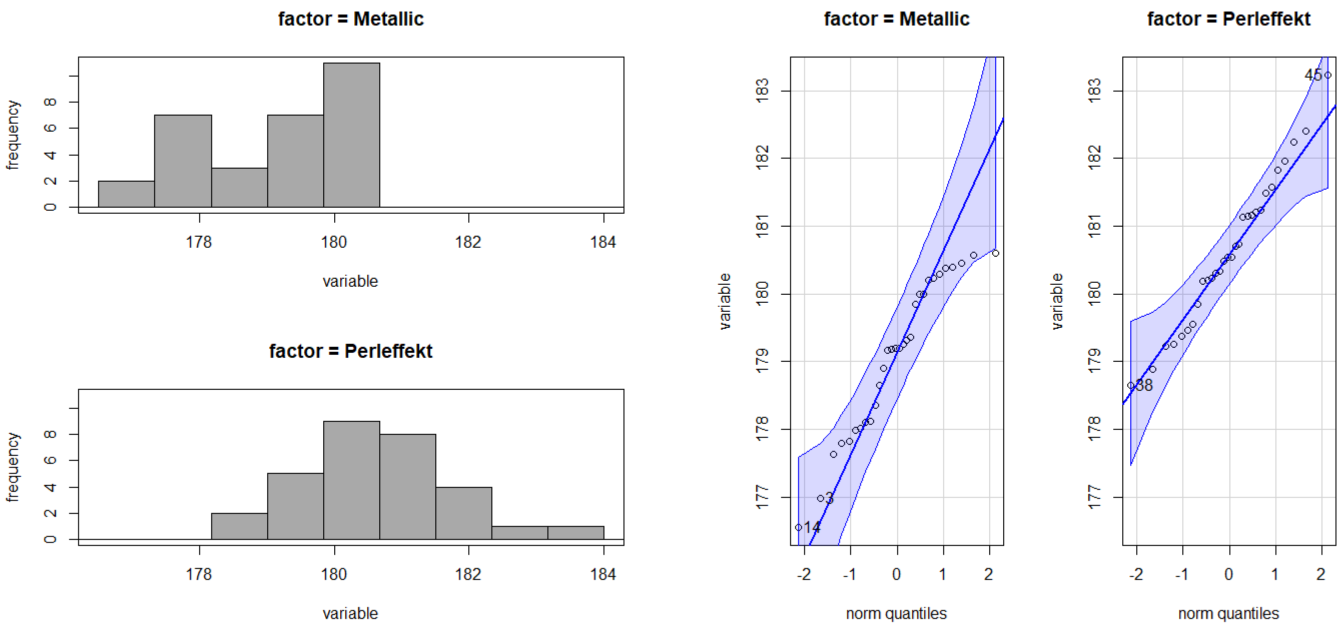
In einer Lackiererei sollen für zwei verschiedene Lackierungsarten die Trocknungszeiten überprüft werden. Es wurden für beide Effekte Stichproben (n=30) genommen (Datensätze: Metallic, Perleffekt in *Übung_2t.xlsx*)

Überprüfen Sie mit einem geeigneten statistischen Testverfahren, ob sich die Trockenzeiten in Abhängigkeit des Lackeffektes signifikant voneinander unterscheiden. Wie lauten die Nullhypothese und die Alternativhypothese für Ihr Testverfahren?

Übung 2t – Test

Lack – Grafik

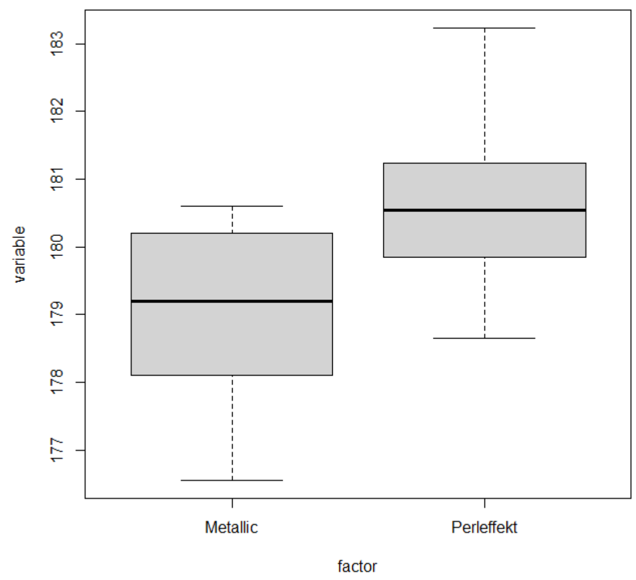
- **Histogramm Metallic ist auffällig**
- **QQ-Diagramme deuten auf Normalverteilung**

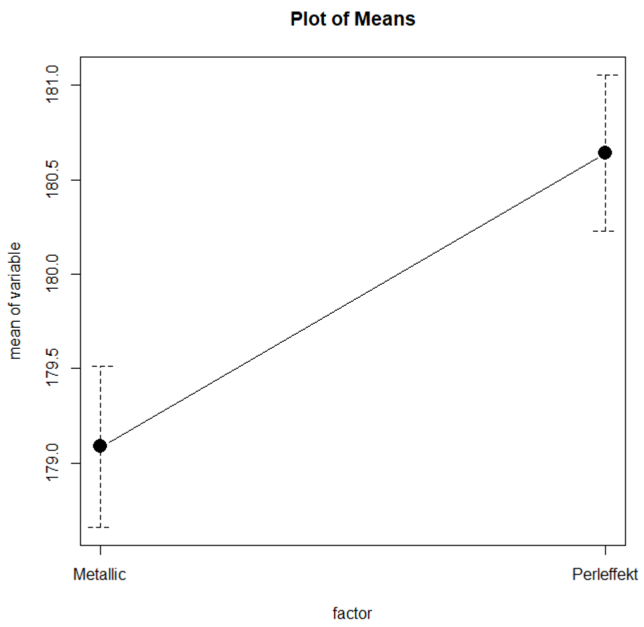


Übung 2t – Test

Lack – Grafik

- Box Plots deuten auf ähnliche Streuung hin
- Plot of Means deutet auf unterschiedliche Mittelwerte hin





Übung 2t – Test

Lack – Normalverteilung

```
 factor = Metallic

         Anderson-Darling normality test

data:  variable
A = 0.61109, p-value = 0.1019
 --------
 factor = Perleffekt

Anderson-Darling normality test

data:  variable
A = 0.17166, p-value = 0.9225
 --------
 p-values adjusted by the Holm method:
           unadjusted adjusted
Metallic   0.10191    0.20382
Perleffekt 0.92248    0.92248
```

- Beide Datensätze sind normalverteilt

Übung 2t – Test

Lack – Varianz

```
F test to compare two variances

data:  variable by factor
F = 1.0646, num df = 29, denom df = 29, p-value = 0.8673
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.5067089 2.2367061
sample estimates:
ratio of variances
          1.064593
```

- $p>\alpha$: Es gilt die Nullhypothese, wir können von gleichen Varianzen ausgehen

Übung 2t – Test

Lack – Hypothesen

H_1 Die Mittelwerte der Stichproben der beiden Lacke sind nicht gleich (Die Differenz der Mittelwerte ist ungleich Null)

H_0 Die Mittelwerte der Stichproben der beiden Lacke sind gleich (Die Differenz der Mittelwerte ist gleich Null)

Übung 2t – Test

Lack – 2t-Test für unabhängige Stichproben

```
        Two Sample t-test

data:  variable by factor
t = -5.3743, df = 58, p-value = 0.000001426
alternative hypothesis: true difference in means between group Metallic and
group Perleffekt is not equal to 0
95 percent confidence interval:
 -2.1318858 -0.9747808
sample estimates:
  mean in group Metallic mean in group Perleffekt
                179.0873                 180.6407
```

$p=0{,}000001426 = 0{,}00014\% < \alpha$

Wechsel zur Alternativhypothese:
Die Mittelwerte der Stichproben sind nicht gleich (sie stammen aus verschiedenen Grundgesamtheiten) bzw. die Differenz der Mittelwerte ist ungleich Null.

Übung 2t – Test

Lager

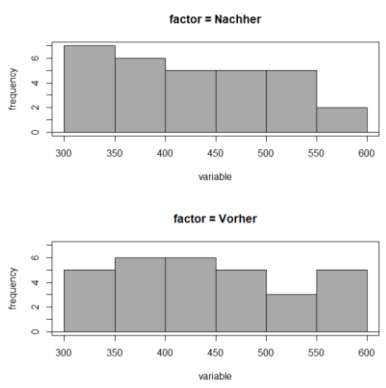
In einem automatischen Lager sollen die Lieferzeiten verbessert. Dafür wird eine Wegeoptimierung in der Lagersoftware eingeführt. Um die Verbesserung zu belegen untersucht das Unternehmen für 30 definierte Testbestellungen die Lieferzeiten vorher und nachher. (Datensätze: Vorher, Nachher in *Übung_2t.xlsx*)

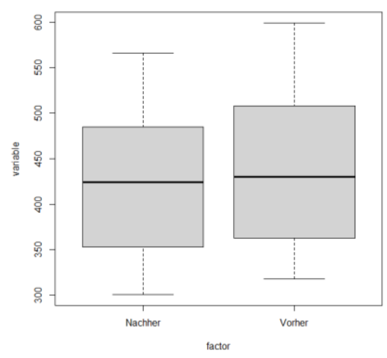
Können Sie die Verbesserung nachweisen?

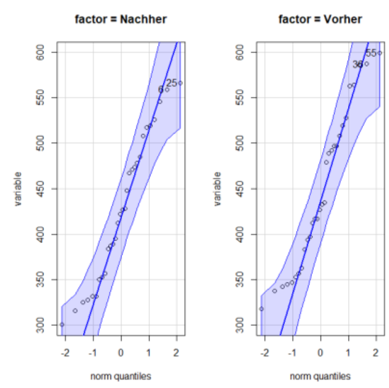
Übung 2t – Test

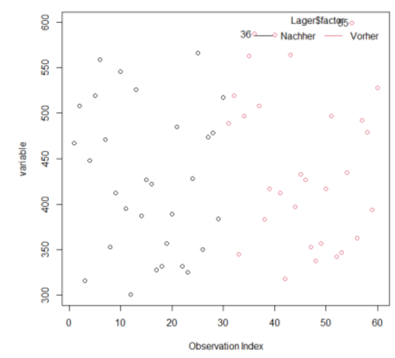
Lager – Grafik

- Keine Auffälligkeiten





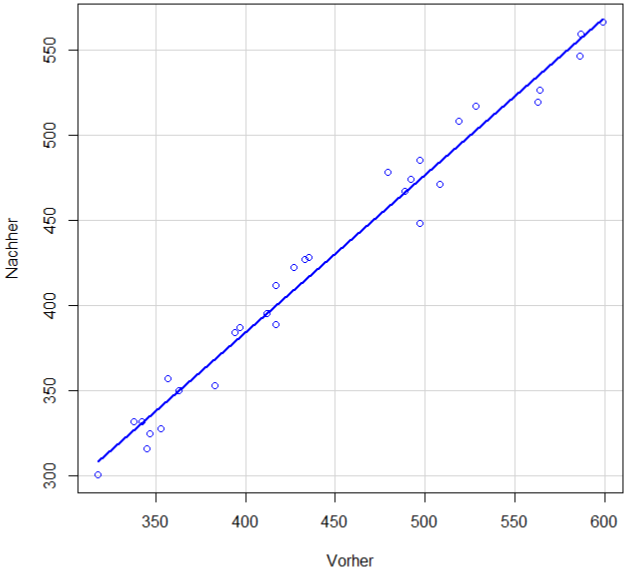




Übung 2t – Test

Lager – Grafik

- Starke positive Korrelation?



Übung 2t – Test

Lager – Korrelation

	Nachher	Vorher
Nachher	1.0000000	0.9888622
Vorher	0.9888622	1.0000000

$r > 0{,}7$ Starke positive Korrelation nach Pearson

Übung 2t – Test

Lager – Normalverteilung

```
Shapiro-Wilk normality test

data:  Diff
W = 0.94225, p-value = 0.1046
```

$p > \alpha$: Die Differenzen sind normalverteilt

Übung 2t – Test

Lager – Hypothesen

H_1 $\quad \bar{x}_{Nachher} < \bar{x}_{vorher}$

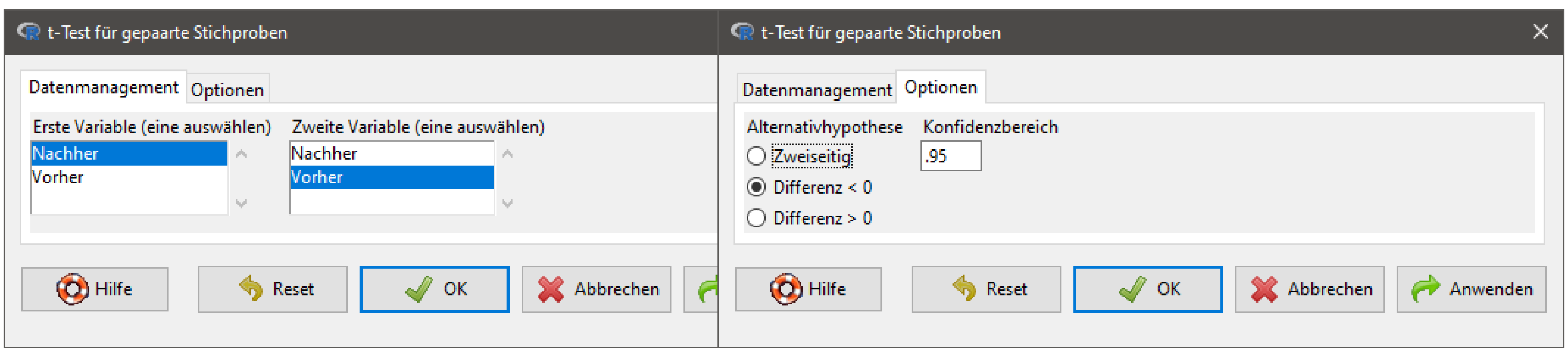
Lieferzeit nach Optimierung ist kleiner als vorher

$$\bar{x}_d = \bar{x}_{Nachher} - \bar{x}_{Vorher} < 0$$

H_0 $\quad \bar{x}_{Nachher} \geq \bar{x}_{Vorher}$

Lieferzeit nach Optimierung ist größer oder gleich der Lieferzeit vorher

$$\bar{x}_d = \bar{x}_{Nachher} - \bar{x}_{Vorher} \geq 0$$



Übung 2t – Test

Lager – t-Test

```
        Paired t-test

data:  Nachher and Vorher
t = -7.8859, df = 29, p-value = 0.000000005351
alternative hypothesis: true mean difference is less than 0
95 percent confidence interval:
      -Inf -15.27232
sample estimates:
mean difference
      -19.46667
```

$p < \alpha$: Wir wechseln zur Alternativhypothese
t(Nachher) – t(Vorher) < 0,
d.h. Mittelwert t(Vorher) > Mittelwert t(Nachher)
Die Verbesserung wirkt.