1. Aufgabe: Auf einer Reifenrollentestanlage werden Versuche durchgeführt um die Langlaufeigenschaften zu untersuchen. Insbesondere wird die Untersuchung des Abriebs über simulierte Anfahr- Brems- und Kurvenmanöver auf verschiedenen Belägen und Lauflängen der Experimente ermöglicht. Das Ergebnis von jeweils 20 Versuchen zweier Reifentypen PL1 und PL2 auf dieser Anlage steht zum Vergleich an. Die Testreifen wurden jeweils zufällig und unabhängig voneinander aus der Monatsproduktion März des Reifenkonzerns Langstone ausgesucht. Seit Jahresbeginn war die Produktion der Reifentypen PL1 und PL2 nach gleicher Rezeptur der Gummimischungen gestartet worden.

Die Abriebergebnisse (gemessen an der Verringerung der Profiltiefe in mm) der gleichartigen Laufversuche auf der Anlage wurden in der Datei *Reifenversuch* dokumentiert. Die Versuche entsprechen einer Laufleistung von 10 000 km.

```
> read.table("D:\\Reifen.txt",header=T)
> Reifen

   PL1 PL2
1  2.8 2.9
2  3.0 3.6
3  2.9 3.8
4  2.6 3.3
5  2.9 3.4
6  3.3 3.0
7  2.8 3.3
8  2.7 2.9
9  3.4 3.2
10 3.2 3.5
11 3.1 3.2
12 3.2 2.8
13 3.3 3.6
14 3.2 3.6
15 3.1 3.5
16 3.3 3.7
17 3.0 3.4
18 3.1 3.5
19 3.2 3.7
20 2.9 3.2

> attach(Reifen)
```

a) Für welche Grundgesamtheit ist die Stichprobe repräsentativ?

b) Welches Skalenniveau haben die Variablen `PL1, PL2`?

c) Welche Stichprobensituation liegt vor?

d) Sie wollen testen, ob die Reifen von Typ 1 hinsichtlich der Abriebergebnisse, besser sind als die von Typ 2. Verwenden Sie bei allen Tests das Signifikanzniveau $\alpha = 0{,}05$.

   i. Welche Grafik würde man verwenden, um mögliche Unterschiede der Reifentypen hinsichtlich der Abriebergebnisse zu erkennen?

   ii. Welche Situationen sollte man vor der Verwendung eines Testes klären? Klären Sie die Situation durch Verwendung des jeweiligen für die Situation richtigen Testes. Formulieren Sie jeweils die Null- und die Alternativhypothesen und die Testentscheidungen.

   iii. Wählen Sie jetzt anhand der Klärung in ii. unter den folgende Tests den Richtigen aus. Formulieren Sie die Null- und die Alternativhypothese und die Testentscheidung.

   Die folgenden Tests ((a) bis (t)) wurden in R durchgeführt:

   (a) `> shapiro.test(PL1)`

```
        Shapiro-Wilk normality test

data:  PL1 W = 0.9568, p-value = 0.4827
```

   (b) `> shapiro.test(PL2)`

```
        Shapiro-Wilk normality test

data:  PL2 W = 0.9466, p-value = 0.3185
```

   (c) `> fligner.test(list(PL1,PL2))`

```
        Fligner-Killeen test of homogeneity of variances

data:  list(PL1, PL2)
 Fligner-Killeen:med chi-squared = 1.304, df = 1, p-value = 0.2535
```

   (d) `> var.test(PL1,PL2)`

```
               F test to compare two variances

data:  PL1 and PL2
 F = 0.5778, num df = 19, denom df = 19, p-value =0.2409
 alternative hypothesis: true ratio of variances is not equal to 1
 95 percent confidence interval:
 0.2287079 1.4598319
 sample estimates:
 ratio of variances
         0.5778192
```

(e) *> t.test(PL1,PL2,alternative="two.sided")*

```
        Welch Two Sample t-test

data:  PL1 and PL2 t = -3.7309, df = 35.461, p-value = 0.0006658
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.4708821 -0.1391179
sample estimates: mean of x mean of y
    3.050     3.355
```

(f) *> t.test(PL1,PL2,alternative="greater")*

```
        Welch Two Sample t-test

data:  PL1 and PL2 t = -3.7309, df = 35.461, p-value = 0.9997
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.443072       Inf
sample estimates: mean of x mean of y
    3.050     3.355
```

(g) *> t.test(PL1,PL2,alternative="less")*

```
        Welch Two Sample t-test

data:  PL1 and PL2 t = -3.7309, df = 35.461, p-value = 0.0003329
alternative hypothesis: true difference in means is less than 0 95
percent confidence interval:
      -Inf -0.166928
sample estimates: mean of x mean of y
    3.050     3.355
```

(h) *> t.test(PL1,PL2,var.equal=TRUE,alternative="two.sided")*

```
        Two Sample t-test

data:  PL1 and PL2 t = -3.7309, df = 38, p-value = 0.0006218
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.4704921 -0.1395079
sample estimates: mean of x mean of y
    3.050     3.355
```

(i) *> t.test(PL1,PL2,var.equal=TRUE,alternative="greater")*

```
        Two Sample t-test

data:  PL1 and PL2 t = -3.7309, df = 38, p-value = 0.9997
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.442825       Inf
sample estimates: mean of x mean of y
    3.050     3.355
```

(j) *> t.test(PL1,PL2,var.equal=TRUE,alternative="less")*

```
        Two Sample t-test

data:  PL1 and PL2 t = -3.7309, df = 38, p-value = 0.0003109
alternative hypothesis: true difference in means is less than 0 95
percent confidence interval:
      -Inf -0.167175
sample estimates: mean of x mean of y
    3.050     3.355
```

(k) *> t.test(PL1,PL2,alternative="two.sided",paired=T)*

```
        Paired t-test

data:  PL1 and PL2 t = -4.2366, df = 19, p-value = 0.0004465
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.4556817 -0.1543183
sample estimates: mean of the differences
     -0.305
```

(l) *> t.test(PL1,PL2,alternative="greater",paired=T)*

```
        Paired t-test

data:  PL1 and PL2 t = -4.2366, df = 19, p-value = 0.9998
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.4294843        Inf
sample estimates: mean of the differences
    -0.305
```

(m) *> t.test(PL1,PL2,alternative="less",paired=T)*

```
        Paired t-test

data:  PL1 and PL2 t = -4.2366, df = 19, p-value = 0.0002232
alternative hypothesis: true difference in means is less than 0 95
percent confidence interval:
      -Inf -0.1805157
sample estimates: mean of the differences
    -0.305
```

(o) `> wilcox.test(PL1,PL2,alternative="two.sided")`

```
        Wilcoxon rank sum test with continuity correction

data:  PL1 and PL2 W = 82, p-value = 0.001391
 alternative hypothesis: true location shift is not equal to 0

Warnmeldung:
 In wilcox.test.default(PL1, PL2, alternative = "two.sided") :
  kann bei Bindungen keinen exakten p-Wert Berechnen
```

(p) `> wilcox.test(PL1,PL2,alternative="greater")`

```
    Wilcoxon rank sum test with continuity correction

data:  PL1 and PL2 W = 82, p-value = 0.9994
 alternative hypothesis: true location shift is greater than 0

Warnmeldung:
 In wilcox.test.default(PL1, PL2, alternative = "greater") :
  kann bei Bindungen keinen exakten p-Wert Berechnen
```

(q) `> wilcox.test(PL1,PL2,alternative="less")`

```
    Wilcoxon rank sum test with continuity correction

data:  PL1 and PL2 W = 82, p-value = 0.0006953
 alternative hypothesis: true location shift is less than 0

 Warnmeldung:
 In wilcox.test.default(PL1, PL2, alternative = "less"):
  kann bei Bindungen keinen exakten p-Wert Berechnen
```

(r) `> wilcox.test(PL1,PL2,alternative="two.sided",paired=T)`

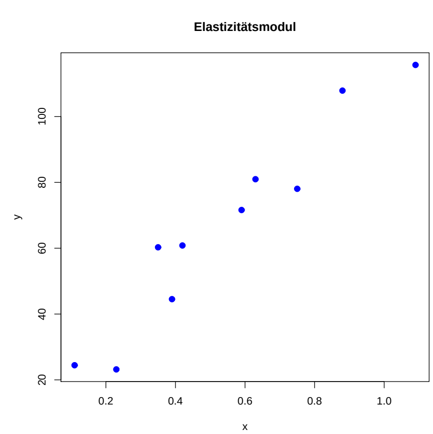
```
        Wilcoxon rank sum test with continuity correction

data:  PL1 and PL2 V = 22.5, p-value = 0.002167
 alternative hypothesis: true location shift is not equal to 0

Warnmeldung:
 In wilcox.test.default(PL1, PL2, alternative = "two.sided",paired=T) :
  kann bei Bindungen keinen exakten p-Wert Berechnen
```

(s)
```
> wilcox.test(PL1,PL2,alternative="greater",paired = T)

    Wilcoxon rank sum test with continuity correction

data:  PL1 and PL2 V = 22.5, p-value = 0.999
 alternative hypothesis: true location shift is greater than 0

Warnmeldung:
 In wilcox.test.default(PL1, PL2, alternative = "greater",paired = T):
  kann bei Bindungen keinen exakten p-Wert Berechnen
```

(t)
```
> wilcox.test(PL1,PL2,alternative="less",paired = T)

    Wilcoxon rank sum test with continuity correction

data:  PL1 and PL2 V = 22.5, p-value = 0.001083
 alternative hypothesis: true location shift is less than 0

 Warnmeldung:
 In wilcox.test.default(PL1, PL2, alternative = "less",paired = T) :
  kann bei Bindungen keinen exakten p-Wert Berechnen
```

2. **Aufgabe:** Das Elastizitätsmodul eines Polymers soll ermittelt werden. Dazu werden Zugversuche durchgeführt. Man spannt eine Probe in eine Prüfmaschine, wo sie mit konstanter Geschwindigkeit gezogen wird. Man notiert die benötigte Spannung y (in N/mm^2) und die erreichte Dehnung x (in %).

   a) Kann man x und y als Realisierungen von normalverteilten Zufallsvariablen auffassen oder gibt es signifikante Abweichungen von einer Normalverteilung ($\alpha = 0{,}05$). Verwenden Sie dazu die entsprechenden Tests aus den Tests (1)-(8) am Ende der Aufgabe. Geben Sie die Hypothesen und die Testentscheidungen an.
   b) Welche Grafik ist sehr gut geeignet, um mögliche Abweichungen von einer Normalverteilung zu erkennen?
   c) Welchen Zusammenhang zwischen x und y würde Sie aufgrund der folgenden Grafik vermuten?



d) Testen Sie zum Niveau $\alpha = 0{,}05$, ob es einen signifikanten gleichsinnigen Zusammenhang zwischen den beiden Größen gibt. Verwenden Sie dazu den entsprechenden Test aus den folgenden Tests (1)-(8). Geben Sie die Hypothesen und die Testentscheidung an.

(1) *> shapiro.test(x)*

```
Shapiro-Wilk normality test

data:  x W = 0.9757, p-value = 0.9383
```

(2) *> shapiro.test(y)*

```
      Shapiro-Wilk normality test

data:  y W = 0.9508, p-value = 0.678
```

(3) *> cor.test(x,y,alternative="two.sided")*

```
        Pearson's product-moment correlation

data:  x and y t = 10.29, df = 8, p-value = 6.854e-06
 alternative hypothesis: true correlation is not equal to 0
 95 percent confidence interval:
 0.8516640 0.9917581
sample estimates:
      cor
0.9642373
```

(4) *> cor.test(x,y,alternative="greater")*

```
        Pearson's product-moment correlation

data:  x and y t = 10.29, df = 8, p-value = 3.427e-06
 alternative hypothesis: true correlation is greater than 0
 95 percent confidence interval:
 0.8812376 1.0000000
sample estimates:
      cor
0.9642373
```

(5) *> cor.test(x,y,alternative="less")*

```
        Pearson's product-moment correlation

data:  x and y t = 10.29, df = 8, p-value = 1
 alternative hypothesis: true correlation is less than 0
 95 percent confidence interval:
 -1.000000  0.989553
sample estimates:
      cor
0.9642373
```

(6) *> cor.test(x,y,method="spearman",alternative="two.sided")*

```
        Spearman's rank correlation rho

data:  x and y S = 6, p-value < 2.2e-16
 alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.9636364
```

(7) *> cor.test(x,y,method="spearman",alternative="greater")*

```
        Spearman's rank correlation rho

data:  x and y S = 6, p-value < 2.2e-16
 alternative hypothesis: true rho is greater than 0
sample estimates:
      rho
0.9636364
```

(8) *> cor.test(x,y,method="spearman",alternative="less")*

```
        Spearman's rank correlation rho

data:  x and y S = 6, p-value = 1
 alternative hypothesis: true rho is less than 0
sample estimates:
      rho
0.9636364
```

3. Aufgabe: Für Kondensatoren gleicher Bauart wird eine Sollkapazität von 4,5μF gefordert. Aus der Tagesproduktion wurden zufällig und unabhängig voneinander 20 Kondensatoren entnommen und deren Kapazität (in μF) gemessen.

Die Messwerte wurden in R als Vektor Kapazitaet gespeichert.

Testen Sie, ob die erwartete Kapaziät sich signifkant von 4,5μF unterscheidet. Wählen Sie dazu aus (2)-(7) den richtigen Test aus. Formulieren Sie die Hypothesen und die Testentscheidung. Beachten Sie bei der Wahl des Tests den Test (1) und dessen Testergebnis.

Verwenden Sie bei allen Tests $\alpha = 0{,}05$.

(1)
```
> shapiro.test(Kapazitaet)


        Shapiro-Wilk normality test

data:  Kapazitaet W = 0.9404, p-value = 0.2437
```

(2)
```
> t.test(Kapazitaet, alternative="less",mu=4.5)


         One Sample t-test

data:  Kapazitaet t = -1.5171, df = 19, p-value = 0.07285
alternative hypothesis: true mean is less than 4.5
 95 percent confidence interval:
     -Inf 4.505241
sample estimates: mean of x
   4.4625
```

(3)
```
> t.test(Kapazitaet, alternative="greater",mu=4.5)


         One Sample t-test

data:  Kapazitaet t = -1.5171, df = 19, p-value = 0.9271 alternative
hypothesis: true mean is greater than 4.5
 95 percent confidence interval:
 4.419759      Inf
sample estimates: mean of x
   4.4625
```

(4) *> t.test(Kapazitaet, alternative="two.sided",mu=4.5)*

```
        One Sample t-test

data:  Kapazitaet t = -1.5171, df = 19, p-value = 0.1457
 alternative hypothesis: true mean is not equal to 4.5
 95 percent confidence interval:
   4.410764 4.514236
sample estimates: mean of x
   4.4625
```

(5) *> wilcox.test(Kapazitaet, alternative="less",mu=4.5)*

```
      Wilcoxon signed rank test with continuity correction

data:  Kapazitaet V = 56.5, p-value = 0.06265
 alternative hypothesis: true location is less than 4.5
```

(6) *> wilcox.test(Kapazitaet, alternative="greater",mu=4.5)*

```
        Wilcoxon signed rank test with continuity correction

data:  Kapazitaet V = 56.5, p-value = 0.9422
 alternative hypothesis: true location is greater than 4.5
```

(7) *> wilcox.test(Kapazitaet, alternative="two.sided",mu=4.5)*

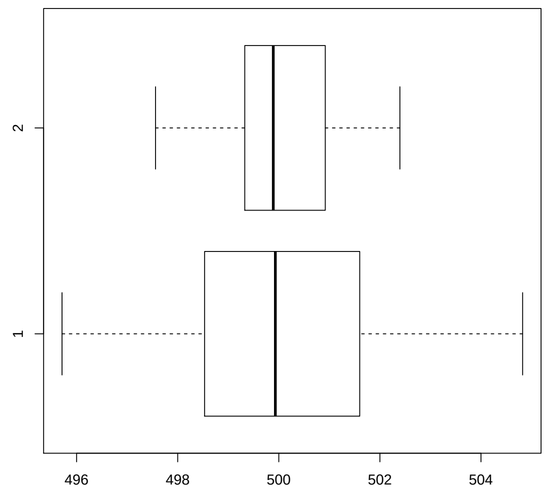
```
      Wilcoxon signed rank test with continuity correction

data:  Kapazitaet V = 56.5, p-value = 0.1253
 alternative hypothesis: true location is not equal to 4.5
```

Bei den Tests (4)-(7) erscheinen immer die folgenden beiden Warnmeldungen:

```
Warnmeldungen:

 1:  kann bei Bindungen keinen exakten p-Wert Berechnen
 2:  kann den exakten p-Wert bei Nullen nicht berechnen
```

4. Aufgabe: Eine Brauerei erweitert ihre Produktion. Dazu wird neben der alten Abfüllanlage auch eine neue Abfüllanlage in Betrieb genommen. Bei beiden Anlagen wurden aus einer Tagesproduktion jeweils 50 Flaschen zufällig und unabhängig voneinander entnommen. Die abgefüllte Menge wurde gemessen. Die 50 Messwerte der alten Anlage sind in der Variable alt und die der neuen Abfüllanlage in der Variable neu gespeichert.

a) Was können Sie aus folgender Graphik ablesen?
1 - alte Abfüllanlage
2 - neue Abfüllanlage



b) Welche Stichprobensituation liegt vor?

c) Welche Hypothesen werden durch die folgenden Tests getestet und wie lauteten die Testentscheidungen bei $\alpha = 0{,}05$?

```
> shapiro.test(alt)

                Shapiro-Wilk normality test

data:  alt W = 0.9826, p-value = 0.6658
```

```
> shapiro.test(neu)

                Shapiro-Wilk normality test

data:  neu W = 0.9774, p-value = 0.4478
```

d) Begründen Sie, dass um zu testen, ob die Varianzen für beide Abfüllanlagen gleich sind oder nicht, der Test von Fligner hier nicht benötigt wird.

e) Es wird vermutet, dass die Varianz der Füllmenge bei der alten Anlage signifikant größer ist als bei der neuen Abfüllanlage. Wählen Sie aus den folgende 3 Tests den richtigen aus. Welche Hypothesen werden getestet und wie lautet die Testentscheidung?

(1) > *var.test(alt,neu)*

```
            F test to compare two variances

data:  alt and neu
 F = 2.9784, num df = 49, denom df = 49, p-value = 2.502e-06
 alternative hypothesis: true ratio of variances is not equal to 1
 95 percent confidence interval:
 2.309864 7.172841
sample estimates: ratio of variances
          4.070416
```

(2) > *var.test(alt,neu,alternative="less")*

```
         F test to compare two variances

data:  alt and neu
 F = 2.9784, num df = 49, denom df = 49, p-value = 1
 alternative hypothesis: true ratio of variances is less than 1
 95 percent confidence interval:
 0.000000 6.542337
sample estimates: ratio of variances
          4.070416
```

(3) > *var.test(alt,neu,alternative="greater")*

```
        F test to compare two variances

data:  alt and neu
 F = 2.9784, num df = 49, denom df = 49,  p-value = 1.251e-06
 alternative hypothesis: true ratio of variances is greater than 1
 95 percent confidence interval:
 2.532472      Inf
sample estimates: ratio of variances
          4.070416
```

f) Es wird behauptet, dass die erwartete Abfüllmenge in beiden Anlagen unterschiedlich ist. Testen Sie diese Behauptung. Wählen Sie dazu aus den folgenden 9 Tests unter Beachtung der Ergebnisse aus c) und des Testergebnisses eines der drei Tests in e) den richtigen Test aus. Formulieren Sie die Hypothesen und treffen Sie die Testentscheidung bei $\alpha = 0{,}05$.

(1) `> t.test(alt,neu,alternative="less",var.equal = FALSE)`

```
        Welch Two Sample t-test

data:  alt and neu t = -0.2686, df = 71.706, p-value = 0.3945
alternative hypothesis: true difference in means is less than 0
 95 percent confidence interval:
      -Inf 0.4756359
sample estimates: mean of x mean of y
 500.0224  500.1138
```

(2) `> t.test(alt,neu,alternative="greater",var.equal = FALSE)`

```
        Welch Two Sample t-test

data:  alt and neu t = -0.2686, df = 71.706, p-value = 0.6055
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.6584094       Inf
sample estimates: mean of x mean of y
 500.0224  500.1138
```

(3) `> t.test(x1,x2,alternative="two.sided",var.equal = FALSE)`

```
          Welch Two Sample t-test

data:  alt and neu t = -0.2686, df = 71.706, p-value = 0.789
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.7697533  0.5869797
sample estimates: mean of x mean of y
 500.0224  500.1138
```

(4)
```
> t.test(alt,neu,alternative="less",var.equal = TRUE)

        Two Sample t-test

data:  alt and neu t = -0.2686, df = 98, p-value = 0.3944
alternative hypothesis: true difference in means is less than 0 95
percent confidence interval:
      -Inf 0.4736516
sample estimates: mean of x mean of y
 500.0224  500.1138
```

(5)
```
> t.test(alt,neu,alternative="greater",var.equal = TRUE)

         Two Sample t-test

data:  alt and neu t = -0.2686, df = 98, p-value = 0.6056
 alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.6564251        Inf
sample estimates: mean of x mean of y
 500.0224  500.1138
```

(6)
```
> t.test(alt,neu,alternative="two.sided",var.equal = TRUE)

        Two Sample t-test

data:  alt and neu
 t = -0.2686, df = 98, p-value = 0.7888
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.7666446  0.5838711
sample estimates: mean of x mean of y
 500.0224  500.1138
```

(7)
```
> wilcox.test(alt,neu,alternative="less")

        Wilcoxon rank sum test with continuity correction

data:  alt and neu W = 1165, p-value = 0.2801
 alternative hypothesis: true location shift is less than 0
```

(8)
```
> wilcox.test(alt,neu,alternative="greater")

        Wilcoxon rank sum test with continuity correction

data:  alt and neu W = 1165, p-value = 0.7222
 alternative hypothesis: true location shift is greater than 0
```

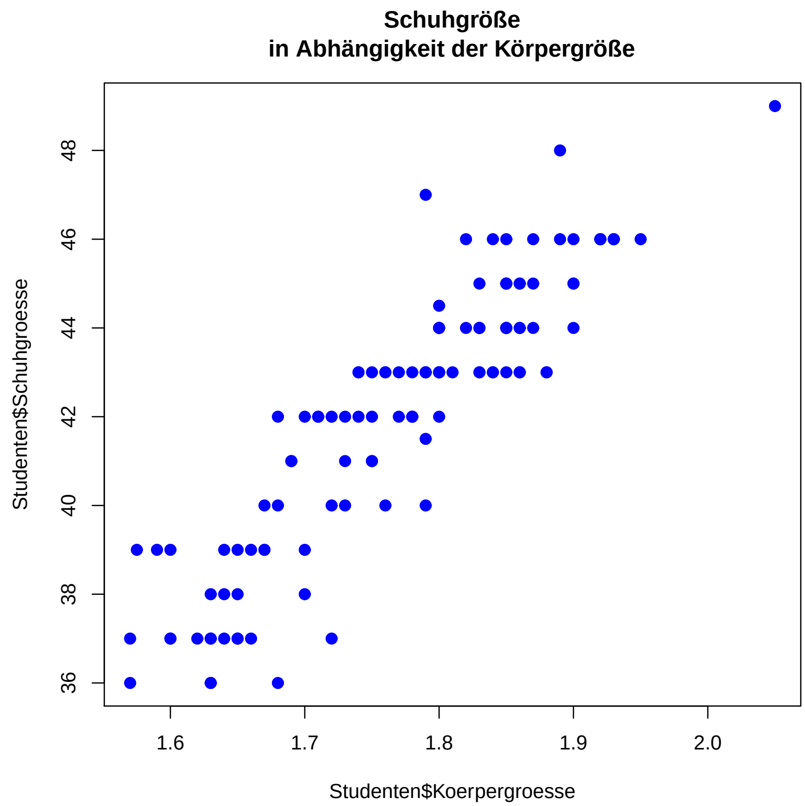
(9)
```
> wilcox.test(alt,neu,alternative="two.sided")

        Wilcoxon rank sum test with continuity correction

data:  alt and neu W = 1165, p-value = 0.5602
 alternative hypothesis: true location shift is not equal to 0
```

5. Aufgabe: Von den Studenten der Lehrveranstaltung Stochastik und Statistik für Ingenieure wurden unter anderem die Körper- und die Schuhgröße erfragt. Die Daten von 121 Studenten sind in StudentenIng.txt zu finden.

```
> Studenten<-read.table("D:/StudentenIng.txt",header=T)
> Studenten

      Koerpergroesse   Schuhgroesse Geschlecht
 1           1.600         37.0          w
 2           1.830         43.0          m
 3           1.800         43.0          m
 4           1.680         42.0          m
 5           1.650         37.0          w
 .             .             .           .
 .             .             .           .
 .             .             .           .
 118         1.600         39.0          w
 119         1.790         41.5          m
 120         1.690         41.0          m
 121         1.800         44.5          m
```

a) Was können Sie aus der folgenden Grafik ablesen?

```
> plot(Studenten$Schuhgroesse,Studenten$Koerpergroesse,col="blue",
lwd=6,cex=0.5,main="Schuhgrösse in Abhängigkeit der Körpergrösse")
```



b) Welche Hypothesen werden durch die folgenden Test getestet und wie lauteten die Testentscheidungen bei $\alpha = 0{,}01$?

```
> shapiro.test(Studenten$Koerpergroesse)

        Shapiro-Wilk normality test

data:  Studenten$Koerpergroesse
 W = 0.9736, p-value = 0.01778

> shapiro.test(Studenten$Schuhgroesse)

        Shapiro-Wilk normality test

data:  Studenten$Schuhgroesse
 W = 0.9439, p-value = 7.391e-05
```

c) Welchen der 2 folgenden Tests würden Sie verwenden, um zu testen, ob es eine signifikante Abhängigkeit zwischen der Schuhgröße und der Körpergröße gibt? Wie lauten die Hypothesen und wie die Tesentscheidung bei $\alpha = 0{,}05$?

(1) *> cor.test(Studenten$Koerpergroesse, Studenten$Schuhgroesse)*

```
        Pearson's product-moment correlation

data:  Studenten$Koerpergroesse and Studenten$Schuhgroesse
 t = 23.0919, df = 119, p-value < 2.2e-16
 alternative hypothesis: true correlation is not equal to 0
 95 percent confidence interval:
  0.8653509 0.9322260
sample estimates:
      cor
0.9041851
```

(2) *> cor.test(Studenten$Koerpergroesse, Studenten$Schuhgroesse, method="spearman")*

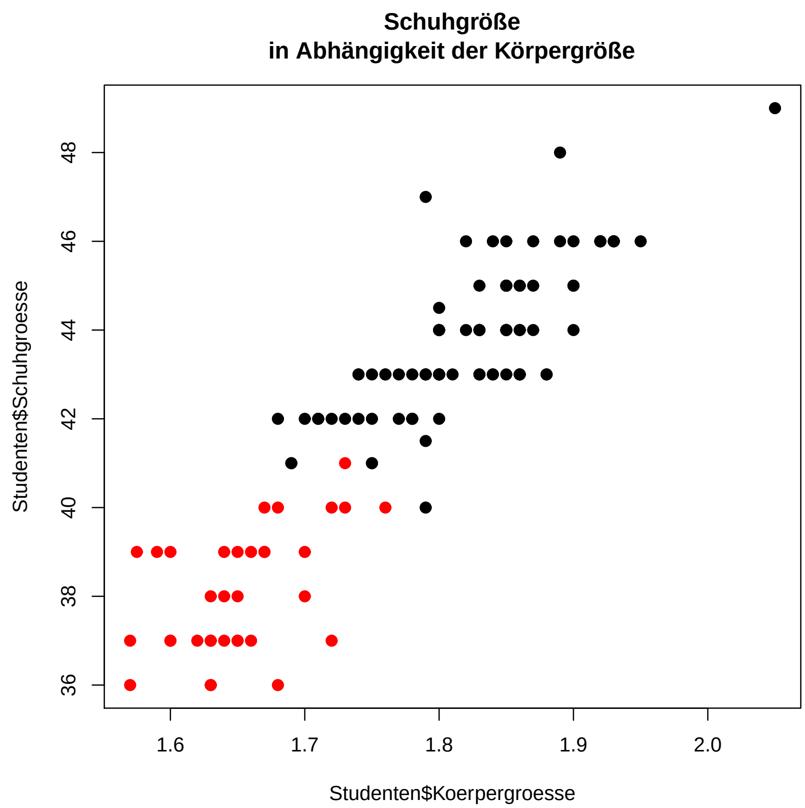
```
        Spearman's rank correlation rho

data:  Studenten$Koerpergroesse and Studenten$Schuhgroesse
 S = 28484.59, p-value < 2.2e-16
alternative hypothesis: true rho is not equal to 0 sample estimates:
      rho
0.9035206

Warnmeldung:
 In cor.test.default(Studenten$Koerpergroesse,Studenten$Schuhgroesse,:
  Kann exakten p-Wert bei Bindungen nicht berechnen
```

d) Welchen Test könnte man noch verwenden um eine mögliche signifikante Abhängigkeit zwischen Körper- und Schuhgröße festzustellen?

6. Aufgabe: Betrachten Sie erneut die Daten der Aufgabe 5.

a) Was können Sie aus der folgenden Grafik ablesen?

```
> plot(Studenten$Schuhgroesse,Studenten$Koerpergroesse,col=Geschlecht,
lwd=6,cex=0.5,main="Schuhgrösse in Abhängigkeit der Körpergrösse")
```



b) Unter den Studenten in der Stichprobe sind 34 Frauen und 87 Männer.

```
> sum(Studenten$Geschlecht=="w")
[1] 34
> sum(Studenten$Geschlecht=="m")
[1] 87
```

Bei den Frauen gibt es 28 mit Schuhgröße in der Kategorie klein (kleiner 40) und keine in der Kategorie groß (größer 44).

```
> sum(Studenten$Geschlecht[Studenten$Schuhgroesse<40]=="w")
[1] 28
> sum(Studenten$Geschlecht[Studenten$Schuhgroesse>44]=="w")
[1] 0
```

Insgesamt sieht es wie folgt aus:

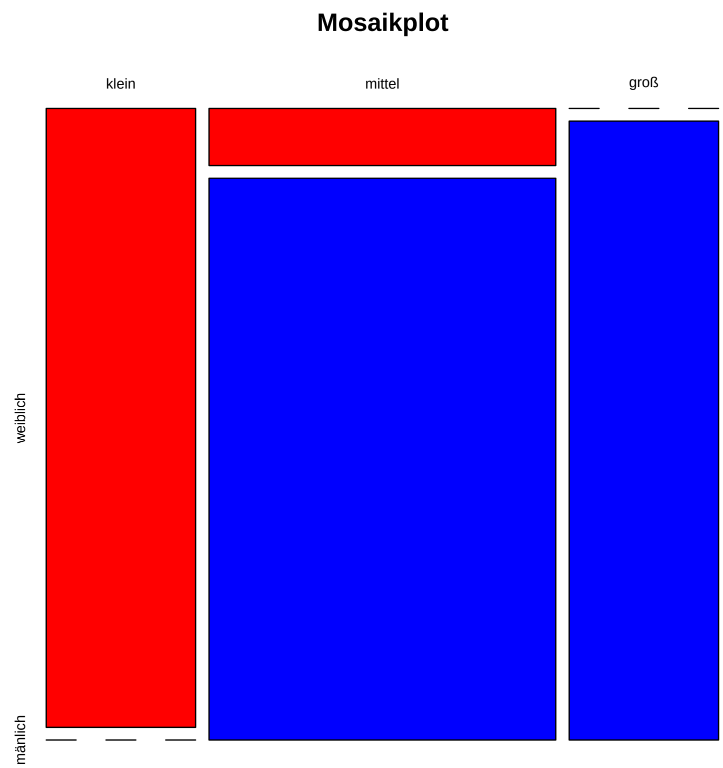
	Schuhgröße		
Geschlecht	klein	mittel	groß
weiblich	28	6	0
männlich	0	59	28

```
> Tabelle <- matrix(c(28,6,0,0,59,28), 2, 3, byrow=TRUE)
> rownames(Tabelle) <- c("weiblich","mänlich")
> colnames(Tabelle) <- c("klein","mittel","gross")
> Tabelle  # Counts
```

```
         klein mittel gross
weiblich    28      6     0 mänlich      0     59   28
```

Was können Sie aus folgender Grafik ablesen?

```
> mosaicplot(t(Tabelle),main="Mosaikplot",color=c(2,4))
```



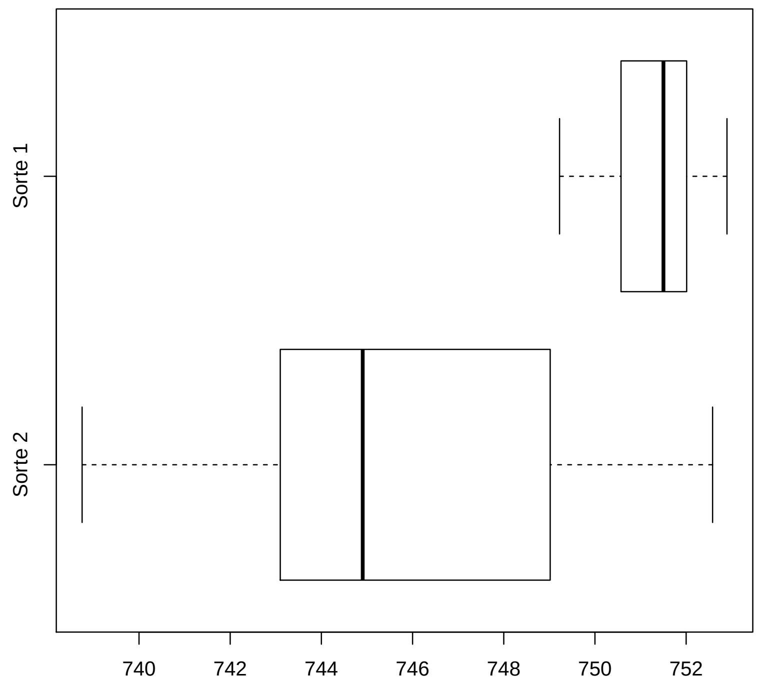
c) Testen Sie zum Niveau $\alpha = 0{,}01$, ob Schuhgröße und Geschlecht abhängig sind.

```
> chisq.test(Tabelle, correct=FALSE)

        Pearson's Chi-squared test

data:  Tabelle X-squared = 94.0436, df = 2, p-value < 2.2e-16
```

7. Aufgabe: Eine neue Sorte von Reagenzgläsern soll bezüglich ihrer Schmelztemperatur mit einer gebräuchlichen Sorte verglichen werden. Aus der Tagesproduktion wurden zufällig und unabhängig voneinander jeweils 10 Reagenzgläser entnommen und deren Schmelztemperatur bestimmt.

In der Variable Sorte1 stehen die Schmelztemperaturen der 10 neuen Reagenzgläser und in Sorte2 die der 10 herkömmlichen Reagenzgläser.

a) `> boxplot(Sorte2,Sorte1,horizontal=T,names=c("Sorte 2","Sorte 1"))`



Vergleichen Sie die Verteilungen der Schmelztemperaturen der beiden Sorten (Lage- und Streuungsvergleich).

b) Welcher Test wird in (b1) durchgeführt? Formulieren Sie die Hypothesen, die Testentscheidung und das Testergebnis für $\alpha = 0{,}05$.

(b1) `> shapiro.test(Sorte1)`

```
            Shapiro-Wilk normality test

data:  Sorte1 W = 0.9617, p-value = 0.8045
```

(b2) `> shapiro.test(Sorte2)`

```
            Shapiro-Wilk normality test

data:  Sorte2 W = 0.9653, p-value = 0.8446
```

c) Testen Sie, ob sich die Varianzen der Schmelztemperaturen der beiden Sorten signifikant unterscheiden. Verwenden Sie dazu einen der beiden folgenden Tests ((c1) oder (c2)). Begründen Sie Ihre Wahl kurz. Formulieren Sie die Hypothesen, die Testentscheidung und das Testergebnis bei einem Signifikanzniveau von $\alpha = 0{,}05$.

(c1)
```
> fligner.test(list(Sorte1,Sorte2))
        Fligner-Killeen test of homogeneity of variances

data:  list(Sorte1, Sorte2)
 Fligner-Killeen:med chi-squared = 7.457,df = 1, p-value = 0.006319
```

(c2)
```
> var.test(Sorte1,Sorte2)
              F test to compare two variances

data:  Sorte1 and Sorte2
 F = 0.0583, num df = 9, denom df = 9, p-value = 0.0002444
 alternative hypothesis: true ratio of variances is not equal to 1
 95 percent confidence interval:
 0.01449045 0.23487026
sample estimates: ratio of variances
        0.05833845
```

d) Es wird behauptet, dass die erwartete Schmelztemperatur der neuen Reagenzgläser (Sorte 1) signifikant größer ist als die der herkömmlichen (Sorte 2). Testen Sie die Behauptung zum Niveau $\alpha = 0{,}05$. Verwenden Sie unter Berücksichtigung der Tests in b) und c) den richtigen der nun folgenden 9 Tests (d1) bis (d9). Formulieren Sie die Hypothesen, die Testentscheidung und das Testergebnis.

(d1)
```
> t.test(Sorte1,Sorte2,alternative="less",var.equal = FALSE)
        Welch Two Sample t-test

data:  Sorte1 and Sorte2 t = 4.0737, df = 10.047, p-value = 0.9989
alternative hypothesis: true difference in means is less than 0
 95 percent confidence interval:
   -Inf 8.6578
sample estimates: mean of x mean of y
 751.3083  745.3155
```

(d2) *> t.test(Sorte1,Sorte2,alternative="greater",var.equal = FALSE)*

```
        Welch Two Sample t-test

data:  Sorte1 and Sorte2 t = 4.0737, df = 10.047, p-value = 0.001108
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 3.32771     Inf
sample estimates: mean of x mean of y
 751.3083  745.3155
```

(d3) *> t.test(Sorte1,Sorte2,alternative="two.sided",var.equal = FALSE)*

```
        Welch Two Sample t-test

data:  Sorte1 and Sorte2 t = 4.0737, df = 10.047, p-value = 0.002216
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 2.717014 9.268497
sample estimates: mean of x mean of y
 751.3083  745.3155
```

(d4) *> t.test(Sorte1,Sorte2,alternative="less",var.equal = TRUE)*

```
        Two Sample t-test

data:  Sorte1 and Sorte2 t = 4.0737, df = 18, p-value = 0.9996
alternative hypothesis: true difference in means is less than 0
 95 percent confidence interval:
     -Inf 8.543723
sample estimates: mean of x mean of y
 751.3083  745.3155
```

(d5) *> t.test(Sorte1,Sorte2,alternative="greater",var.equal = TRUE)*

```
        Two Sample t-test

data:  Sorte1 and Sorte2 t = 4.0737, df = 18, p-value = 0.0003564
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 3.441787      Inf
sample estimates: mean of x mean of y
 751.3083  745.3155
```

(d6) `> t.test(Sorte1,Sorte2,alternative="two.sided",var.equal = TRUE)`

```
        Two Sample t-test

data:  Sorte1 and Sorte2 t = 4.0737, df = 18, p-value = 0.0007128
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 2.902105 9.083406
sample estimates: mean of x mean of y
 751.3083  745.3155
```

(d7) `> wilcox.test(Sorte1,Sorte2,alternative="less")`

```
        Wilcoxon rank sum test

data:  Sorte1 and Sorte2 W = 89, p-value = 0.9992
 alternative hypothesis: true location shift is less than 0
```

(d8) `> wilcox.test(Sorte1,Sorte2,alternative="greater")`

```
        Wilcoxon rank sum test

data:  Sorte1 and Sorte2 W = 89, p-value = 0.001045
 alternative hypothesis: true location shift is greater than 0
```

(d9) `> wilcox.test(Sorte1,Sorte2,alternative="two.sided")`

```
        Wilcoxon rank sum test

data:  Sorte1 and Sorte2 W = 89, p-value = 0.002089
 alternative hypothesis: true location shift is not equal to 0
```

8. Aufgabe: Zum Vergleich der Wirksamkeit zweier Rostschutzmittel R_1 und R_2 wurden Proben genommen und die Ergebnisse den Kategorien *geringe*, *mittlere* und *starke* Wirksamkeit zugeordnet.

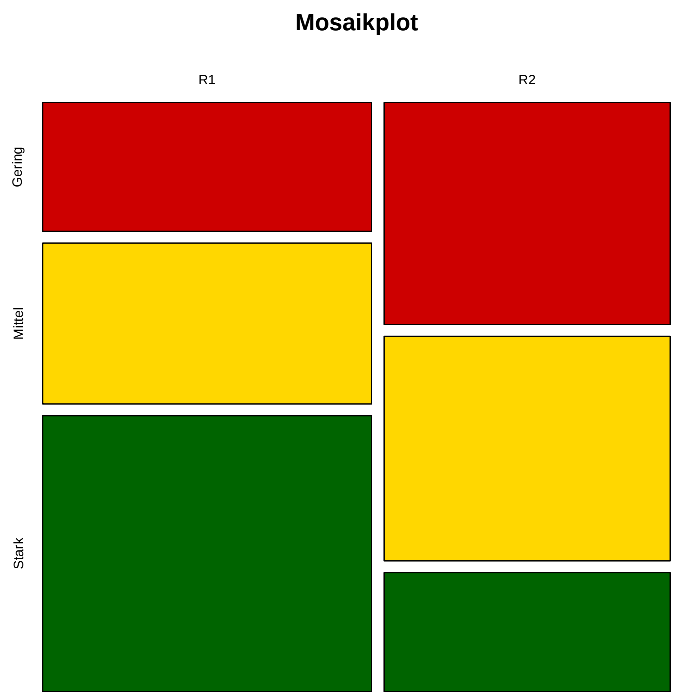
Insgesamt ergab sich die folgende Kontingenztabelle:

	Wirksamkeit		
Rostschutzmittel	Gering	Mittel	Stark
R_1	56	70	120
R_2	84	85	45

```
> Tabelle <- matrix(c(56,70,120,84,85,45), 2, 3, byrow=TRUE)
> rownames(Tabelle) <- c("R1","R2")
> colnames(Tabelle) <- c("Gering","Mittel","Stark")
> Tabelle

   Gering Mittel Stark
R1     56     70   120 R2     84     85    45
```

a) Welches Rostschutzmittel würden Sie aufgrund des folgenden Plots bevorzugen. Begründen Sie Ihre Wahl kurz.

```
> mosaicplot((Tabelle),main="Mosaikplot",color=c("red3","gold","darkgreen"))
```



b) Welche Hypothesen werden im Folgenden getestet? Wie lautet die Testentscheidung und das Testergebnis bei einem Signifikanzniveau von $\alpha = 0{,}01$?

```
> chisq.test(Tabelle, correct=FALSE)

        Pearson's Chi-squared test

data:  Tabelle X-squared = 39.1057, df = 2, p-value = 3.223e-09
```