Statistik – Projektaufgabe

Subashini Anbu
Data Analyst, Mathematikerin (M. Sc.)

The topics for Statistics Analysis

* ***1_Marketing.xlsx***
* ***2_Alternative.xslx***
* ***3_Vergleiche.xlsx***
* ***4_Optimierung.xlsx***

Beginn: Mittwoch ab 8:30 Uhr

Ende / Abgabe: Donnerstag bis 17:00 Uhr

Marketing

Projekt_Aufgabe_KW01

Agenda/Method

* Summary of Data (Marketing Kinder & Eltern)
* Bar & Kreis Diagram
* The X2-Square test & P-Value
* The Percentages of Kaufpräferenz
* Conclusion

Summary(X1_Marketing Kinder)

- Geschlecht Alter Präferenz
- Length:185 Length:185 Length:185
- Class :character Class :character Class :character
- Mode :character Mode :character Mode :character

The requirments of chi-square test

Sample Size: The overall sample size should be sufficiently large.

Independence of Observations: The variables should be independent.

Categorical Data: The data are classified into categories

The data are more than n>30. you need more than 50 data point for the whole table for every group and we need more than 5 data points. In addition the datas are independent variables. so

X2.Square test passt to this data.

Marketing Eltern

* **summary(Dataset)**
* Kind Kaufpräferenz
* Junge :50 Design :37
* Mädchen:40 Technik:53

```
counts:
Kaufpräferenz
 Design Technik
     37      53

percentages:
Kaufpräferenz
 Design Technik
  41.11   58.89

counts:
Kind
  Junge Mädchen
     50      40

percentages:
Kind
  Junge Mädchen
  55.56   44.44
>
```

Marketing Kinder

- **Alter**
- 10-13   14+   6-9
-    63   58   64
- **percentages:**
- Alter
- 10-13   14+   6-9
- 34.05 **31.35 34.59**

14+ is the **Min.**

6-9 is the **Max.**

It is age-specific preferences.

Geschlecht

Junge	Mädchen
92	93

percentages:

Geschlecht

Junge	Mädchen
49.73	50.27

counts:

Präferenz

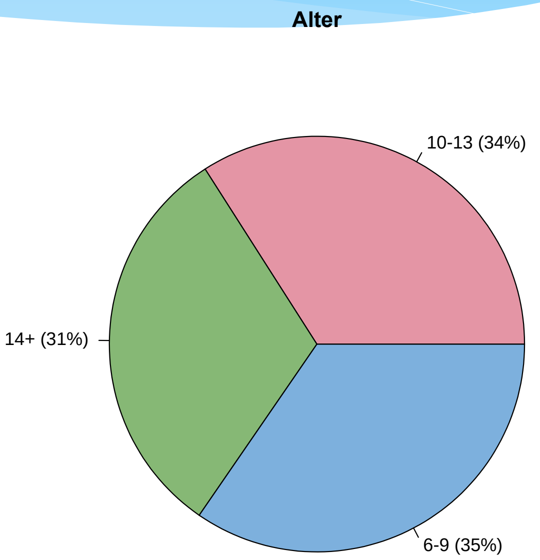
Design	Technik
108	77

percentages:

Präferenz

Design	Technik
58.38	41.62

The Percentages of



Alter and Gender (Marketing Kinder)

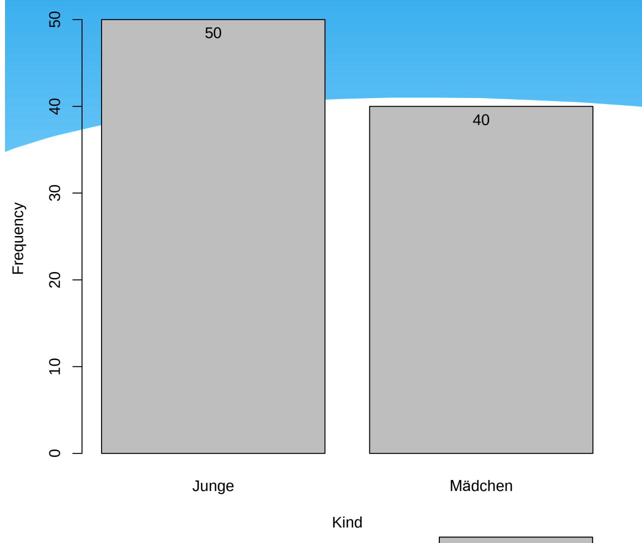
```
* Geschlecht 10-13 14+ 6-9
*    Junge        31   30   31
*    Mädchen      32   28   33

* Total percentages:
*            10-13  14+  6-9 Total
* Junge       16.8 16.2 16.8  49.7
* Mädchen     17.3 15.1 17.8  50.3
* Total       34.1 31.4 34.6 100.0

*       Pearson's Chi-squared test

* data:  .Table
* X-squared = 0.14194, df = 2, p-value = 0.9315 (H0)
```

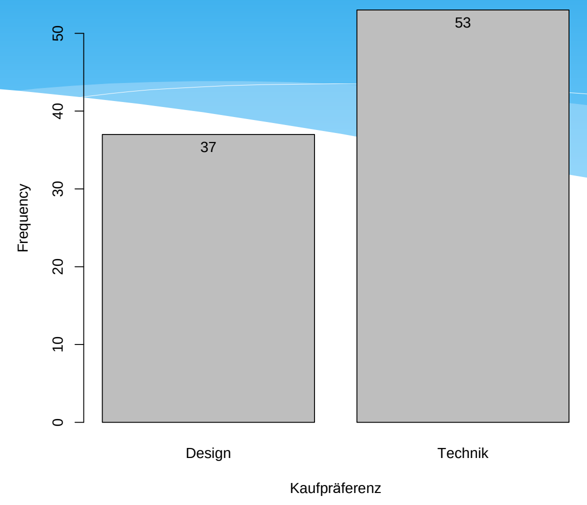
We stay in NullHypotheses and there is no difference between the ages, and it is not significant.

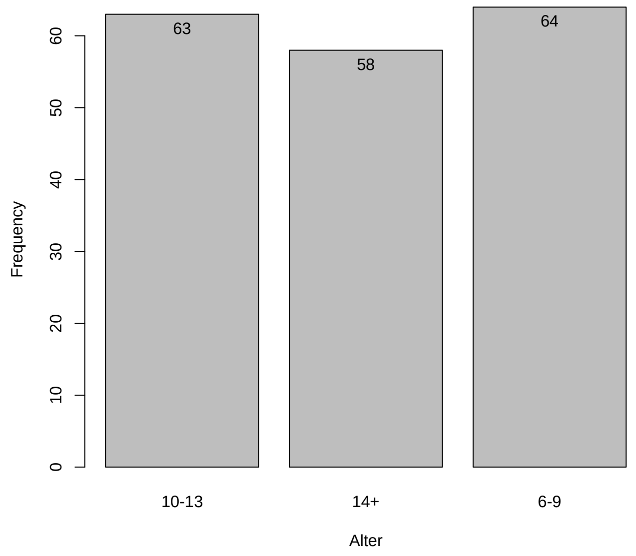
Alter and Preference

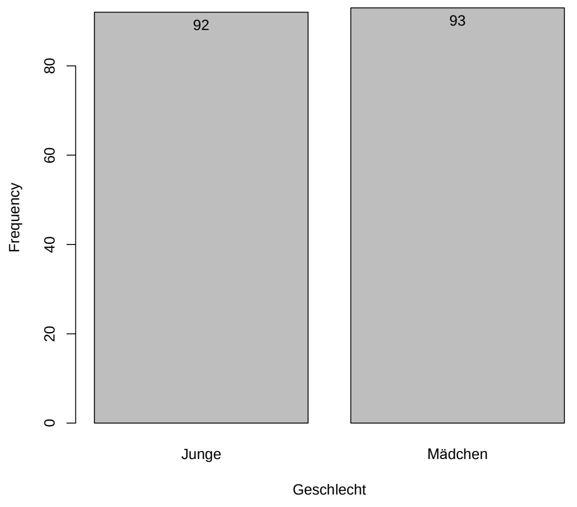
* Alter
* Präferenz 10-13 14+ 6-9
*   Design   36 43 29
*   Technik   27 15 35
* Total percentages:
*   10-13 14+ 6-9 Total
* Design   19.5 23.2 15.7 58.4
* Technik 14.6 8.1 18.9 41.6
* Total   34.1 31.4 34.6 100.0
*   Pearson's Chi-squared test
* data: .Table
* X-squared = 10.465, df = 2, p-value = 0.005341 (**it means it is significant and there is difference between the ages and the preference. So it goes to alternative hypotheses H1**)

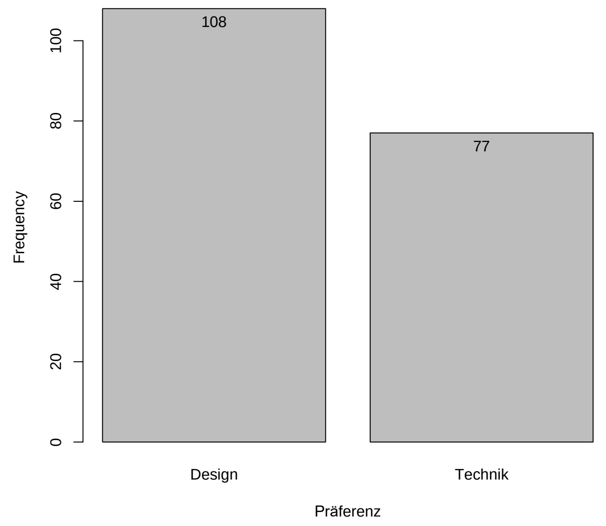
Bar Chart Marketing Kinder & Eltern



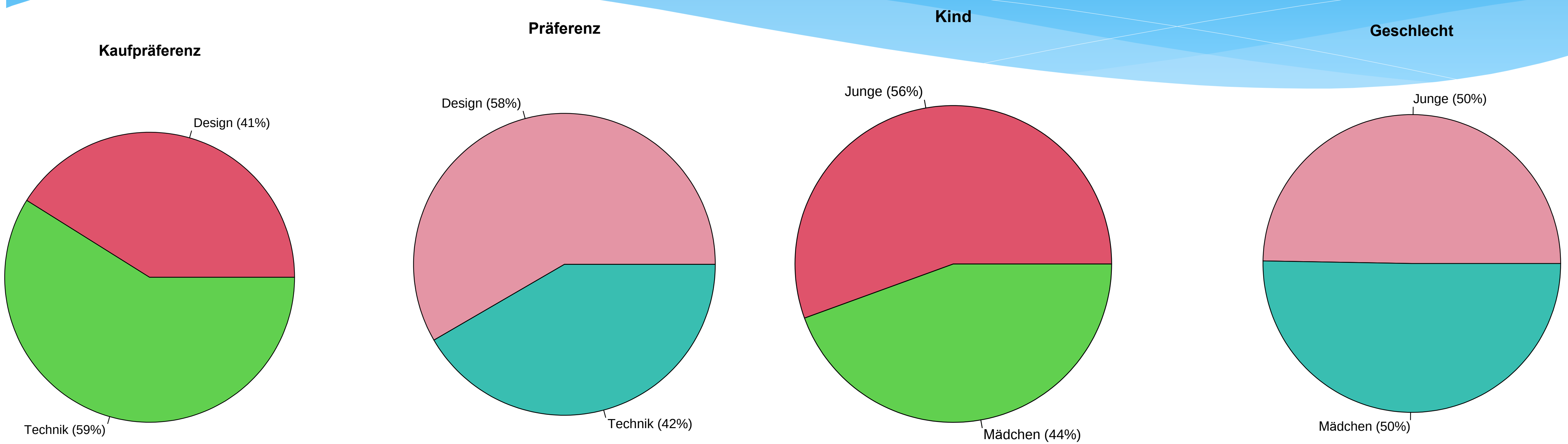








Kreis Diagram



According to the given data, the children are more interested in Design which is 58%, but they are less interested in technik which is 42%.
In another hand the parents are more interested in Technik which the pie chart shows 59%, and they are less interested in Design which is 41%.

Assume the Hypothesis

* The null hypothesis H0; there is no difference in buying preferences.
* Alternative Hypothesis H1: there is a difference by in Buying preferences.

The X2-Square test & P-Value

MarketingKinder

* Percentage table:
* Präferenz
* **Geschlecht** Design Technik Total Count
* Junge 66.3 33.7 100 92
* Mädchen 50.5 49.5 100 93
* data: .Table
* X-squared = 4.7316, df = 1, p-value = 0.02961 **(H1)**
* alternative hypothesis: two.sided
* 95 percent confidence interval:
* 0.01747414 0.29786013
* sample estimates:
* prop 1 prop 2
* 0.6630435 0.5053763

H0 is rejected, It goes to H1 but there is a difference between the preference of boys and girls.

Marketing Eltern

- Percentage table:
- Kind
- Kaufpräferenz Junge Mädchen Total Count
- Design 40.5 59.5 100 37
- Technik 66.0 34.0 100 53

2-sample test for equality of proportions without continuity correction

- data: .Table
- X-squared = 5.7369, df = 1, p-value = 0.01661 **(H1)**
- alternative hypothesis: two.sided
- 95 percent confidence interval:
- -0.45815317 -0.05179073
- sample estimates:
- prop 1 prop 2
- 0.4054054 0.6603774
- **H0 is rejected, It goes to H1, there is a difference in buyingpreference of the parents.**

The Percentage of Kaufpreferenz base on M.Eltern

Kaufpräferenz	Junge	Mädchen	Total	Count
Design	40.5	59.5	100	37
Technik	66.0	34.0	100	53

According to the given data the girls are more interested in Design which is 59.5% and they are less interested in technic which is 34%.

But the boys are more interested in technic which is 66% and they are less interested in Design which is 40.5.

As conclusion: the parents's kaufpreferenz depends on **gender.**

Conclusion

* According to the given data, the children are more interested in Design which is 58%, but they are less interested in technic which is 42%.
* In another hand the parents are more interested in Technic which the pie chart shows 59%, and they are less interested in Design which is 41%.
* As the data shows there are three different types of ages which **are taken age-specific preferences. And the gender are equal.**
* According to the given data the girls are more interested in Design which is 59.5% and they are less interested in technic which is 34%.
* But the boys are more interested in technic which is 66% and they are less interested in Design which is 40.5%.
* The gender is equal as in the Marketing Kinder has given. But in both Marketing are very significant.
* In both Marketing (Kinder and Eltern) the H0 is rejected and it goes to alternative Hyphotoses (H1). It means, that there is difference, between boys and girls's ages Buypreferences and the parents's Buypreferences depends on gender(boys and girls). And it is signifincant difference.

Alternative

Agenda/Method

* Summary of Data
* Active data set
* Shapiro-Wilk Normality test
* Histogram, QQ Diagram, Boxplot
* Wilcoxon.test (two.sided)
* Levene's Test
* Conclusion

Summary of Data

* The dataset X2_Alternative has 175 rows and 2 columns.
* Rcmdr> summary(X2_Alternative)

```
       Base           Alternative
 Min.   :1.963   Min.   :2.156
 1st Qu.:2.106   1st Qu.:2.259
 Median :2.200   Median :2.338
 Mean   :2.188   Mean   :2.348
 3rd Qu.:2.267   3rd Qu.:2.433
 Max.   :2.451   Max.   :2.672
```

Active Data Set

```
Rcmdr>  numSummary(X2_Alternative[,c("Alternative", "Base"), drop=FALSE],
Rcmdr+    statistics=c("mean", "sd", "se(mean)", "var", "IQR", "quantiles"),
Rcmdr+    quantiles=c(0,.25,.5,.75,1))
                mean        sd     se(mean)        var    IQR    0%    25%   50%
Alternative 2.348377 0.1130999 0.008549551 0.01279159 0.1740 2.156 2.2590 2.338
Base        2.188103 0.1016003 0.007680261 0.01032262 0.1615 1.963 2.1055 2.200
              75%  100%   n
Alternative 2.433 2.672 175
Base        2.267 2.451 175
```

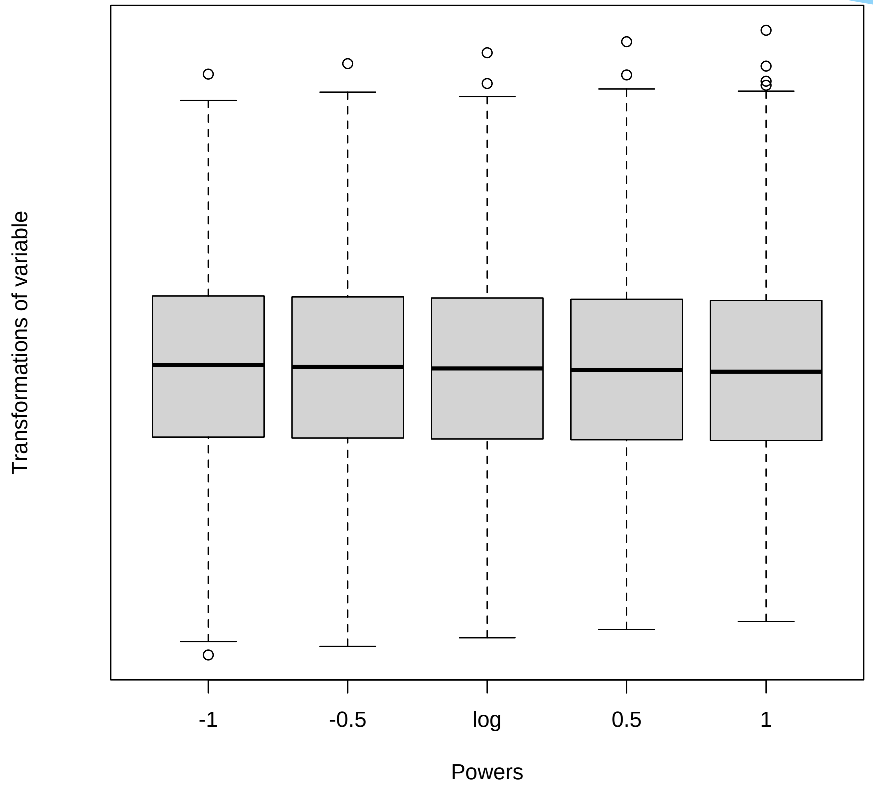
Shapiro-Wilk normality test

* data: Alternative
* W = 0.97119, p-value = 0.001079 (H1)
* data: Base
* W = 0.98925, p-value = 0.2073 (H0)
* p-values adjusted by the Holm method:

	unadjusted	adjusted
Alternative	0.0010786	0.0021573
Base	0.2073275	0.2073275

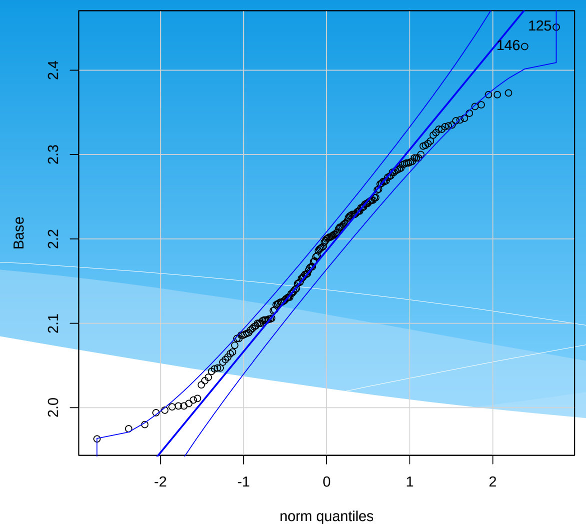
Note: the transformation (Alternative with Base) is not possible. And the Datas are not normal (Alternative) distributed so we choose a non-Parametric (Wilcoxon) test Method.

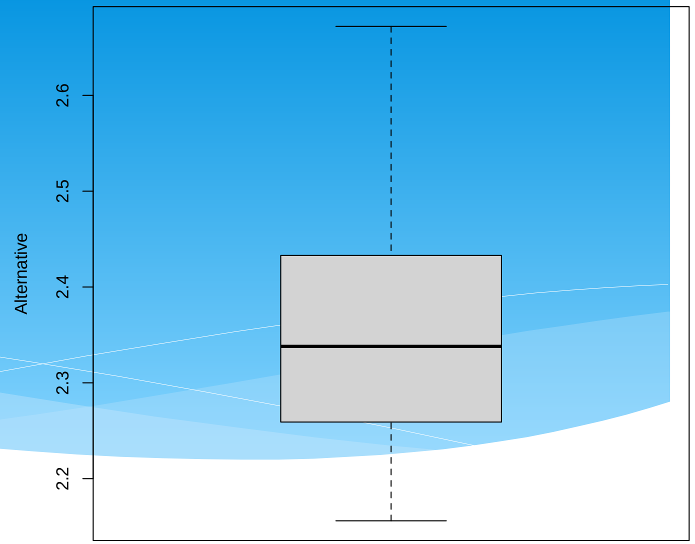
Symmetry Boxplot

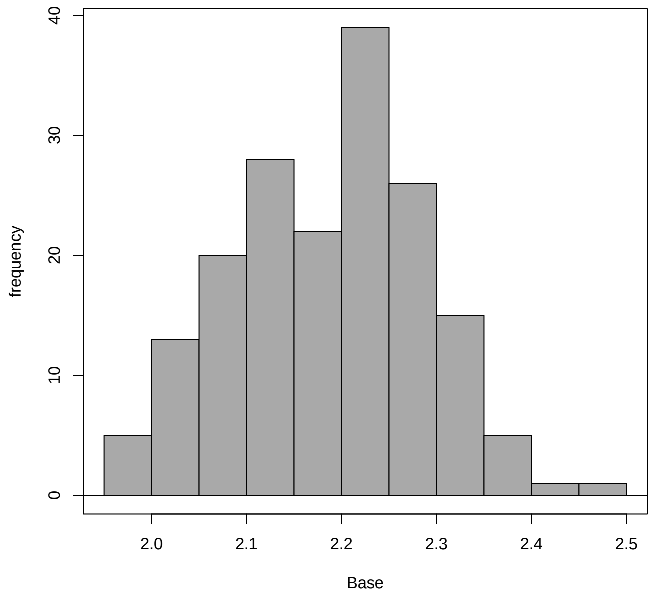


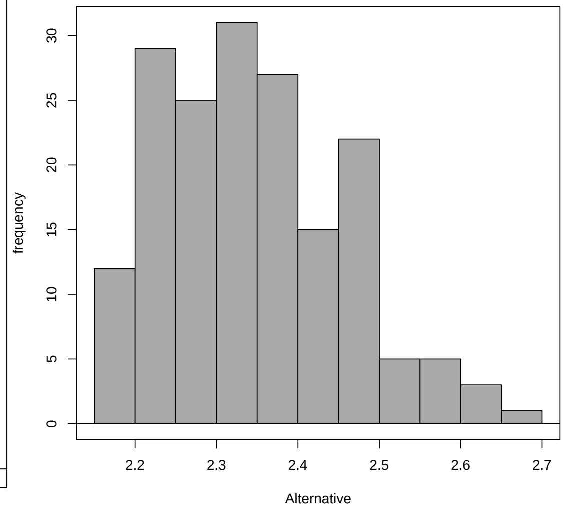
Due to of outliers in the transformation of variable, makes
The interpretation more difficult.

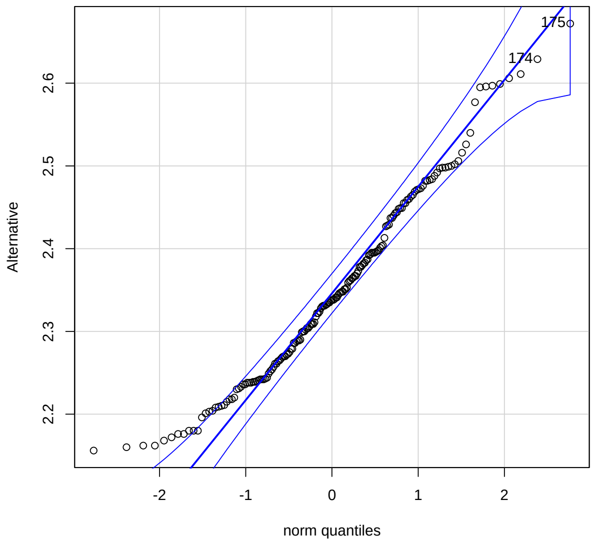
Histogram,QQ

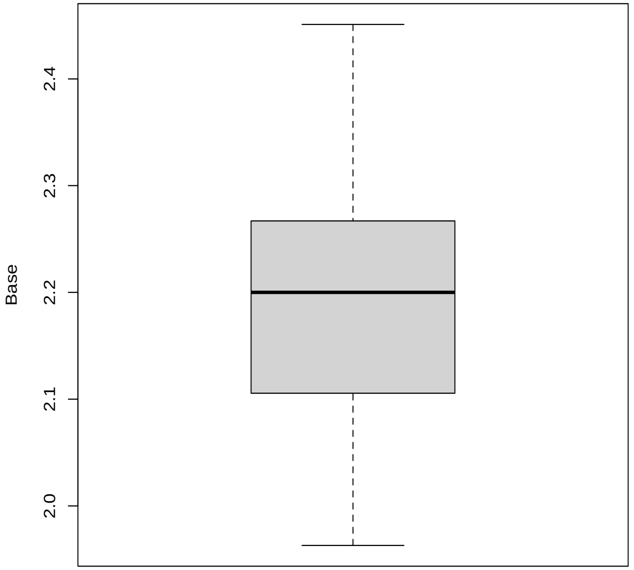












It should be check how the possibility to check transformation and outliers in the data.

Ausreißer Grubbs test for one outlier (Base & Alternative)

```
with(Dataset,grubbs.test(Alternative,type=10,opposite=FALSE,two.sided=FALSE))

         Grubbs test for one outlier

data:  Alternative
G = 2.86139, U = 0.95267, p-value = 0.335
alternative hypothesis: highest value 2.672 is an outlier

> with(Dataset,grubbs.test(Base,type=10,opposite=FALSE,two.sided=FALSE))

         Grubbs test for one outlier

data:  Base
G = 2.5876, U = 0.9613, p-value = 0.7938
alternative hypothesis: highest value 2.451 is an outlier
```

Both the Ausreiser are not significant, it means they don’t have any effect on change.

Box Cox Transformation

```
summary(powerTransform(Alternative ~ 1, data=Dataset, family="bcPower"))
bcPower Transformation to Normality
   Est Power Rounded Pwr Wald Lwr Bnd Wald Upr Bnd
Y1   -3.5375          -1       -6.544       -0.531

Likelihood ratio test that transformation parameter is equal to 0
 (log transformation)
                           LRT df     pval
LR test, lambda = (0) 5.463342  1 0.019419

Likelihood ratio test that no transformation is needed
                           LRT df      pval
LR test, lambda = (1) 9.051774  1 0.0026244
```

Because the Box Cox transformation is negative, the parametric method is not possible.

Requirements for Wilcoxon test

Non-Parametric: The data should be non-parametric.

Paired Data: The data are paired in this case.

Independency: the paired observations should be independent of each other.

Continued data: The data should be measured on at least an ordinal scale.

No extreme outliers:

Assume Hypothesis

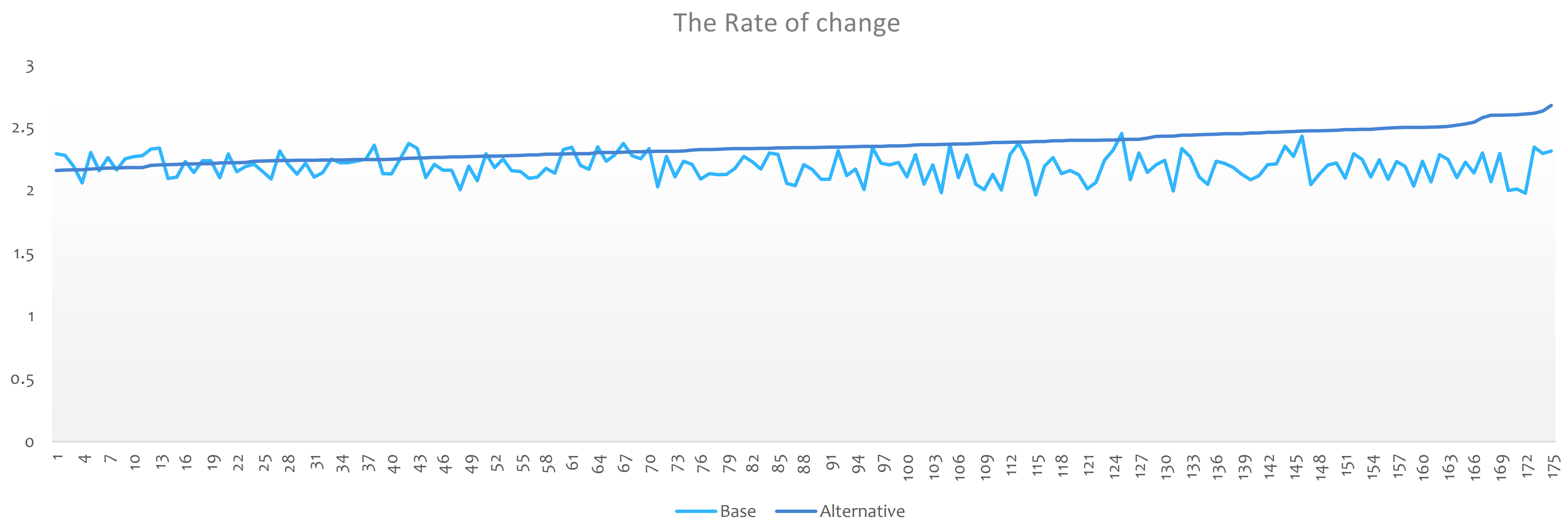
Null Hypothesis (H_0): There is no difference between paired observations.

Alternative Hypothesis (H_1): there is a significant difference between the paired observations.

wilcoxon.test (one side Test)

```
with(StackedData, wilcox.test(variable,
alternative='two.sided', mu=0.0))

	Wilcoxon signed rank test with continuity correction

data:  variable
V = 61425, p-value < 2.2e-16
alternative hypothesis: true location is not equal to 0
```

Rate of Change between Base and Alternative



As the chart shows , there is a posive change rate of change in Alternative and at the end the difference is increased.

Conclusion

The Data is not normally distributed, and both variables (Base and Alternative) didn’t pass to transformation, at this case the non-parametric Wilcoxon_Test was used. This means the Data as Non_Parametric Method tested.

As the Box-cox transformation has shown us that the transformation is not possible to be shift to the parametric method. In this case, we inform the production head that it is analyzed with Non-parametric methods.

Also the rate of change is **positive** analysed.

Vergleiche

Agenda/Method

* Summary of Dataset
* Shapiro-Wilk normality test
* Boxplot, QQ Diagram
* Anova_Test
* Multiple Comparisons of Means: Tukey Contrasts
* 95% family-wise confidence level
* Conclusion

Summary of Dataset

```
      Base              KB              KL               P        
 Min.   :2.100   Min.   :4.020   Min.   :2.620   Min.   :3.490  
 1st Qu.:2.195   1st Qu.:4.180   1st Qu.:2.902   1st Qu.:3.743  
 Median :2.260   Median :4.275   Median :3.035   Median :3.880  
 Mean   :2.262   Mean   :4.290   Mean   :3.018   Mean   :3.894  
 3rd Qu.:2.320   3rd Qu.:4.400   3rd Qu.:3.125   3rd Qu.:4.060  
 Max.   :2.440   Max.   :4.610   Max.   :3.350   Max.   :4.280  
       FL       
 Min.   :4.910  
 1st Qu.:5.000  
 Median :5.060  
 Mean   :5.070  
 3rd Qu.:5.115  
 Max.   :5.280  
```

Summary of Dataset

```
* Rcmdr+    .75,1))
*          mean          sd         var    IQR   0%    25%   50%   75% 100%  n
* Base 2.2622 0.08291353 0.006874653 0.1250 2.10 2.1950 2.260 2.320 2.44 50
* FL   5.0704 0.08573714 0.007350857 0.1150 4.91 5.0000 5.060 5.115 5.28 50
* KB   4.2896 0.14617518 0.021367184 0.2200 4.02 4.1800 4.275 4.400 4.61 50
* KL   3.0180 0.16293231 0.026546939 0.2225 2.62 2.9025 3.035 3.125 3.35 50
* P    3.8944 0.19991998 0.039968000 0.3175 3.49 3.7425 3.880 4.060 4.28 50
```

Shapiro-Wilk normality test

p-values adjusted by the Holm method:

	unadjusted	adjusted
Base	0.60325	1.0000
FL	0.06086	0.3043
KB	0.34109	1.0000
KL	0.87460	1.0000
P	0.52949	1.0000

- data: Base
- W = 0.98119, p-value = 0.6032 (H0)
- data: FL
- W = 0.95607, p-value = 0.06086 (H0)
- data: KB
- W = 0.97424, p-value = 0.3411 (H0)
- data: KL
- W = 0.98759, p-value = 0.8746 (H0)
- data: P
- W = 0.97946, p-value = 0.5295 (H0)

All the data are more than 0.05% which stays by Nullhypothese, it means they are not significant and no difference between them and don't go to alternative H1.

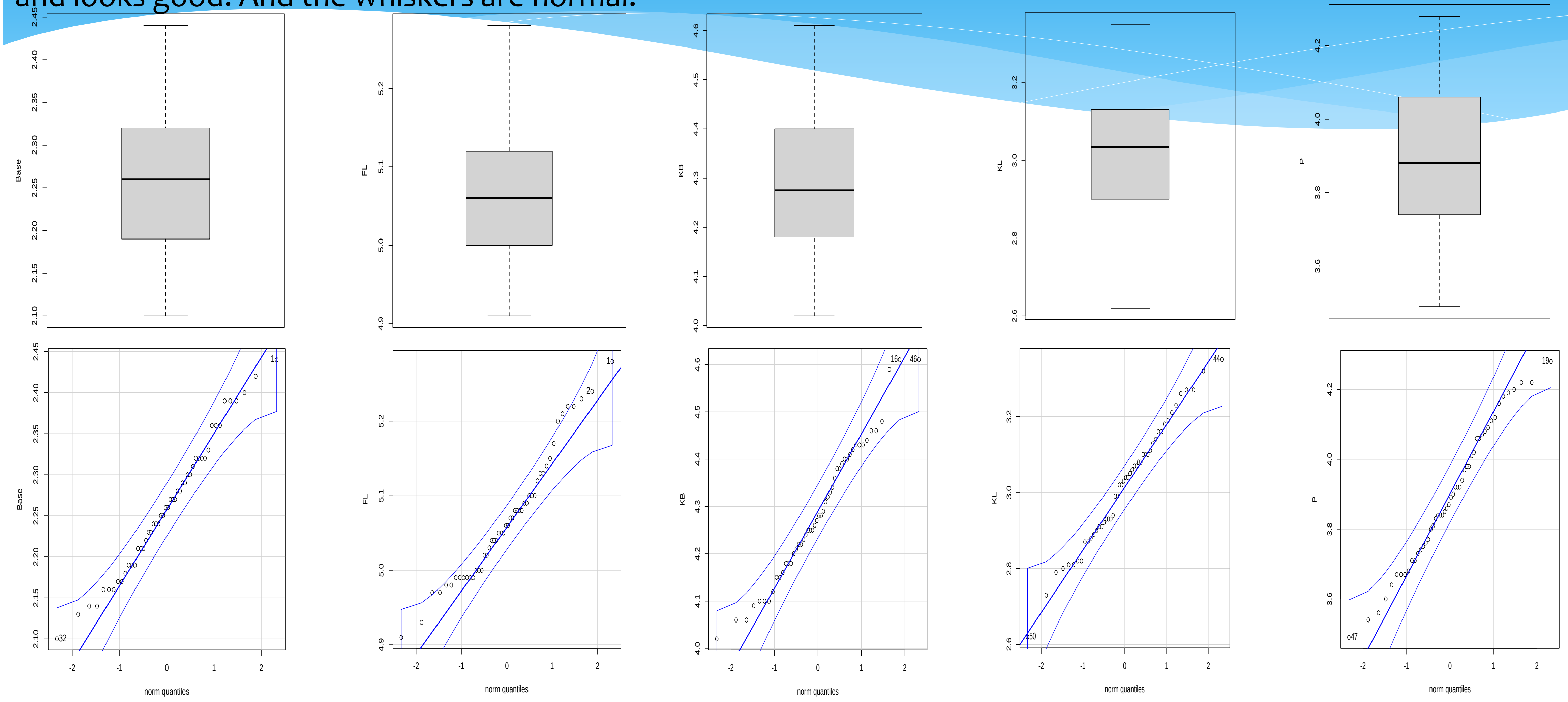
Pearson's product-moment correlation

```
with(Dataset, cor.test(Base, FL, alternative="two.sided",
method="pearson"))

       Pearson's product-moment correlation

data:  Base and FL
t = 15.573, df = 48, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8519903 0.9503274
sample estimates:
      cor
0.9136632
```

There is high positive correlation between them, and we can say that they are dependent variables

Boxplot & QQ Diagram

Only there are a few outliers in the FL QQ diagram, otherwise, the rest are normal. and the Boxplots are Asymmetric and looks good. And the whiskers are normal.



Paired t. Test

* Dependency of the samples
* The difference pairs to be formed are required •
* At least interval scaled
* Normal distribution (or $n \geq 30$)
* As positive a correlation as possible (otherwise the power Suffer)

Assume Hypothesis

- Null Hypothesis (H_0): the mean difference is equal to zero (0)
- Alternative Hypothesis (H_1): The mean difference is not equal to Zero (0)

Paired t-test (Base & FL)

```
with(Dataset, (t.test(Base, FL, alternative = "two.sided", conf.level =
.95,
Rcmdr+     paired=TRUE)))
       Paired t-test
data:  Base and FL
t = -564.93, df = 49, p-value < 2.2e-16 alternative hypothesis: true mean
difference is not equal to 0
95 percent confidence interval:
 -2.818189 -2.798211
sample estimates:
mean difference
        -2.8082
```

In order the Base Model and FL model are dependent are test with Paired t-test. `(H1)They are not equal.`

Requirements of one factor Anova Test

* At least interval scaled dependent variable
* Characteristic expressions must be independent of each other be (if not: ANOVA with repeated measures…)
* Normal distribution of the dependent variable in all groups
* Equal variance in all groups

Bartlett test of homogeneity of variances

```
Tapply(variable ~ factor, var, na.action=na.omit, data=StackedData)
Rcmdr+    # variances by group
        KB         KL          P
0.02136718 0.02654694 0.03996800

Rcmdr>  bartlett.test(variable ~ factor, data=StackedData)

      Bartlett test of homogeneity of variances

data:  variable by factor
Bartlett's K-squared = 5.0163, df = 2, p-value = 0.08142
```

There is no difference between the variances. And they have homogeneuos variance.

Assume Hypothesis

* Null Hypothesis (H0): there is no significant difference among the means of the groups.
* Alternative Hypothesis (H1): there is a significant difference in at least

One pair of group means.

AnovaModel.3 <- aov(variable ~ factor, data=StackedData)

```
AnovaModel.2 <- aov(variable ~ factor, data = StackedData)

Rcmdr>  summary(AnovaModel.2)
             Df Sum Sq Mean Sq F value Pr(>F)    
factor        2  42.35  21.177   722.9 <2e-16 *** (Significant, there is difference. H1)
Residuals   147   4.31   0.029                   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Rcmdr>  with(StackedData, numSummary(variable, groups = factor, statistics=c('mean', 
Rcmdr+     'sd')))
     mean        sd data:n
KB 4.2896 0.1461752     50
KL 3.0180 0.1629323     50
P  3.8944 0.1999200     50
```

Multiple Comparisons of Means: Tukey Contrasts

P-KL has positive effect and is significant

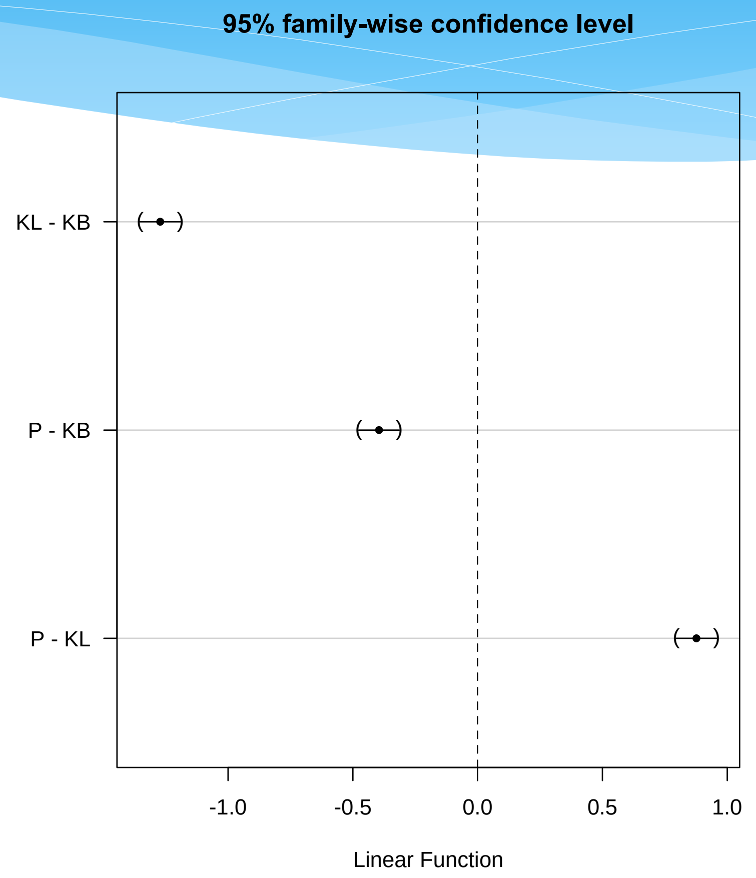
```
 Simultaneous Tests for General Linear Hypotheses
Fit: aov(formula = variable ~ factor, data = StackedData)
Linear Hypotheses:
             Estimate Std. Error t value Pr(>|t|)
KL - KB == 0 -1.27160    0.03423  -37.15   <2e-16 ***
P - KB == 0  -0.39520    0.03423  -11.54   <2e-16 ***
P - KL == 0   0.87640    0.03423   25.60   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)
```

Multiple Comparisons of Means: Tukey Contrasts

```
Simultaneous Confidence Intervals
Fit: aov(formula = variable ~ factor, data = StackedData)
Quantile = 2.3679
95% family-wise confidence level
Linear Hypotheses:
             Estimate lwr     upr
KL - KB == 0 -1.2716  -1.3527 -1.1905
P - KB == 0  -0.3952  -0.4763 -0.3141
P - KL == 0   0.8764   0.7953  0.9575

 KB  KL   P
"a" "b" "c"
```

95% family-wise confidence level

According to the family-wise confidence level, the combination of P-Kl Models is positive but the others Models combinations are negative.



Conclusion

According to the Anova test in the case of independent Models the P-KL models have positive combination s but the rest combinations have negative combination. And in the case of dependent variables models they are more significant to each other.

So there is are a difference in the effectiveness of the change.

We can conclude that the Base Model with characteristics P and KL are the best .

Optimization

Agenda/Method

* Summary of Dataset
* Test of Normality
* Shapiro-Wilk Normality test
* Boxplot, QQ Diagram
* One-way Anova test
* Conclusion

Summary of Dataset

```
Rcmdr>  summary(X4_Optimierung)
 Flügellänge            Papier          Körperbreite             FZ       
 Length:240         Length:240         Length:240         Min.   :1.650  
 Class :character   Class :character   Class :character   1st Qu.:4.103  
 Mode  :character   Mode  :character   Mode  :character   Median :5.570  
                                                          Mean   :5.559  
                                                          3rd Qu.:6.952  
                                                          Max.   :9.550  
```

summary(Dataset)

```
Flügellänge    Papier     Körperbreite        FZ
 130 mm:120    80 g:120    20 mm:120     Min.   :1.650
 80 mm :120    90 g:120    35 mm:120     1st Qu.:4.103
                                         Median :5.570
                                         Mean   :5.559
                                         3rd Qu.:6.952
                                         Max.   :9.550
```

Summary of Dataset

```
numSummary(X4_Optimierung[,"FZ", drop=FALSE], statistics=c("mean", "sd",
Rcmdr+    "se(mean)", "var", "IQR", "quantiles"), quantiles=c(0,.25,.5,.75,1))
    mean       sd   se(mean)      var  IQR   0%    25%  50%    75% 100%   n
 5.55875 1.991605 0.1285576 3.966491 2.85 1.65 4.1025 5.57 6.9525 9.55 240
```

normalityTest(FZ ~ variable, test="shapiro.test", data=Dataset)

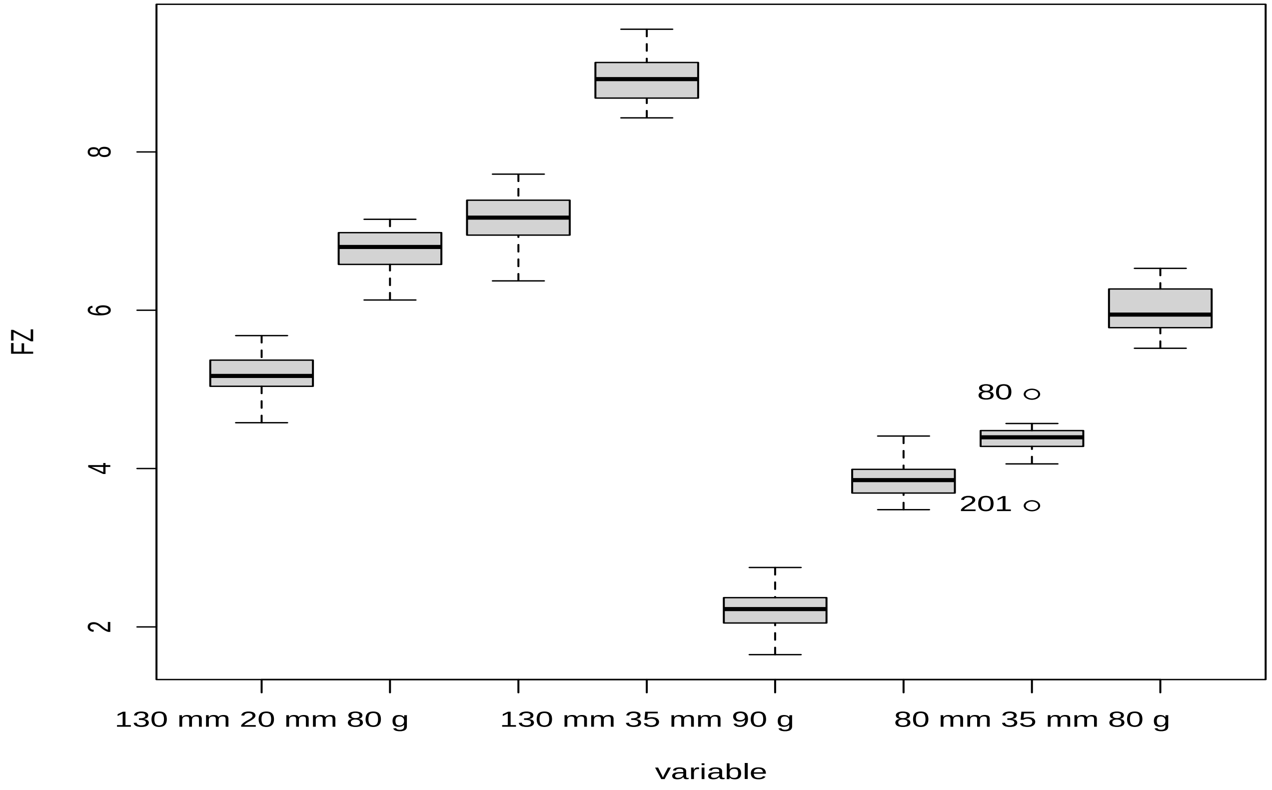
```
variable = 130 mm 20 mm 80 g
data:  FZ
W = 0.97672, p-value = 0.7332 (H0)
 variable = 130 mm 20 mm 90 g
data:  FZ
W = 0.96903, p-value = 0.5131 (H0)
 variable = 130 mm 35 mm 80 g
data:  FZ
W = 0.97962, p-value = 0.8155 (H0)
 variable = 130 mm 35 mm 90 g
data:  FZ
W = 0.96041, p-value = 0.3174 (H0)
variable = 80 mm 20 mm 80 g
data:  FZ
W = 0.98607, p-value = 0.954 (H0)
 variable = 80 mm 20 mm 90 g
data:  FZ
W = 0.9618, p-value = 0.3441 (H0)
```

normalityTest(FZ ~ variable, test="shapiro.test", data=Dataset)

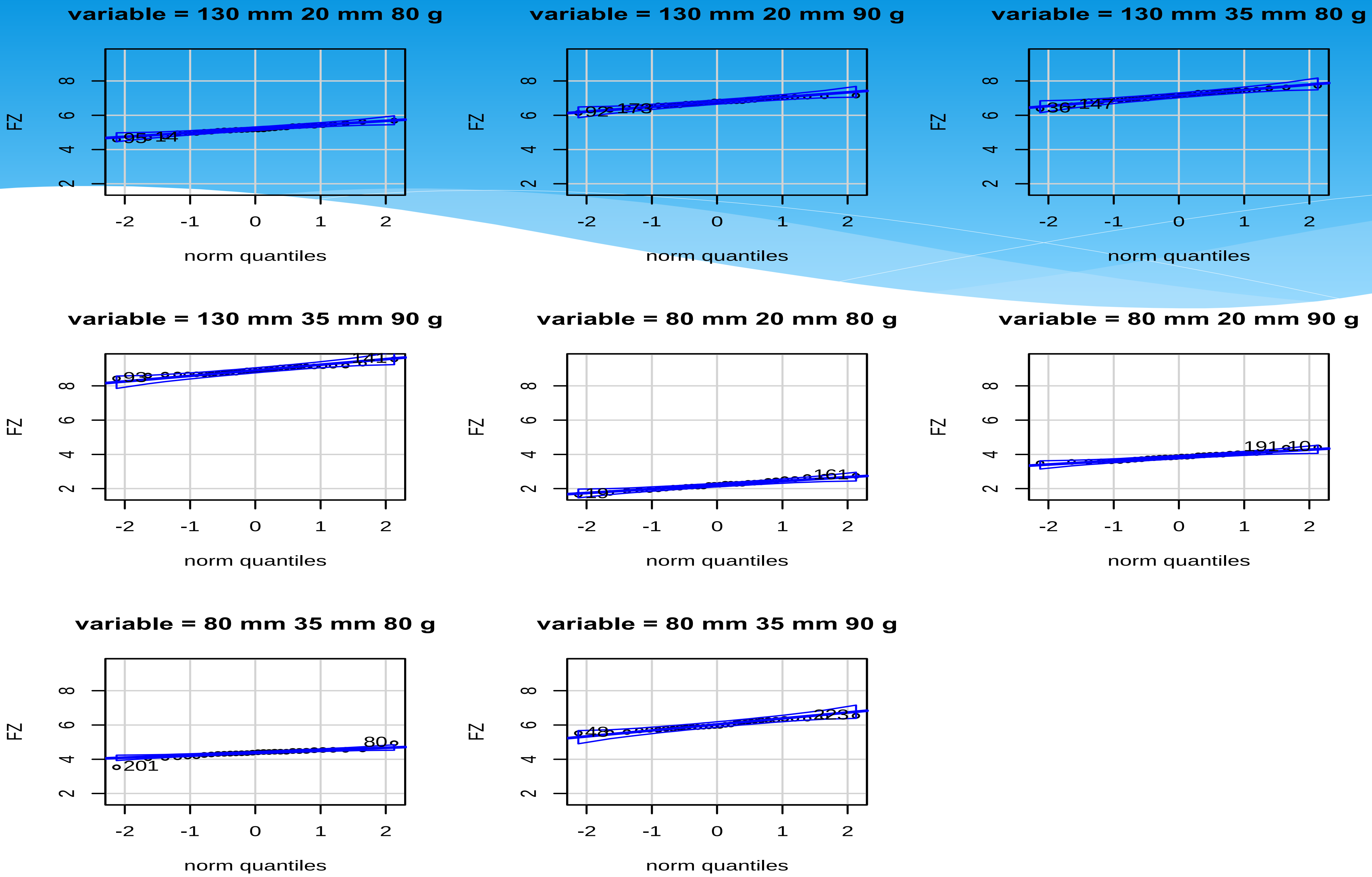
```
 variable = 80 mm 35 mm 80 g
data:  FZ
W = 0.87906, p-value = 0.002683 (H1) this variable is not normal and is significant.
 variable = 80 mm 35 mm 90 g
data:  FZ
W = 0.96219, p-value = 0.3519 (H0)
 p-values adjusted by the Holm method:
                  unadjusted adjusted
130 mm 20 mm 80 g 0.7331531  1.000000
130 mm 20 mm 90 g 0.5130966  1.000000
130 mm 35 mm 80 g 0.8154718  1.000000
130 mm 35 mm 90 g 0.3174075  1.000000
80 mm 20 mm 80 g  0.9539671  1.000000
80 mm 20 mm 90 g  0.3441299  1.000000
80 mm 35 mm 80 g  0.0026832  0.021466
80 mm 35 mm 90 g  0.3519448  1.000000
```

Except one variable which is significat (H1),but all other variables stay by H0 which means are not significant and are equal.

Histogram,Boxplot & QQ Diagram



QQ Diagram



There are som a few outliers in the QQ Diagram but they don’t have effect on normality.

Ausreißer Grubbs test for one outlier

```
with(Dataset,grubbs.test(FZ,type=10,opposite=FALSE,two.sided=FALSE))

Grubbs test for one outlier

data:  FZ
G = 2.00404, U = 0.98313, p-value = 1
alternative hypothesis: highest value 9.55 is an outlier
```

The P-value is greater than alfa. And it is not significant.
And it does not have any effect.

summary(one factorial Anova)

```
Rcmdr>  .myAnova
             Df Sum Sq Mean Sq F value Pr(>F)
variable      7  932.1  133.16    1943 <2e-16 ***
Residuals   232   15.9    0.07
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Rcmdr>  AnovaModel.4 <- aov(FZ ~ variable, data=Dataset)
Rcmdr>  .myAnova <- summary(AnovaModel.4)
Rcmdr>  .myAnova
             Df Sum Sq Mean Sq F value Pr(>F)
variable      7  932.1  133.16    1943 <2e-16 ***
Residuals   232   15.9    0.07
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

summary(one factorial Anova)

```
colnames(.effectSize) <- c("part. eta sq.")
Rcmdr>  .effectSize
            part. eta sq.
variable           0.9832
Residuals          0.0168
Rcmdr>  # NUMERIC SUMMARY OF GROUPS
Rcmdr>  with(Dataset, numSummary(FZ, groups=variable, statistics=c("mean",
"sd")))
                     mean        sd data:n
130 mm 20 mm 80 g 5.183000 0.2571884     30
130 mm 20 mm 90 g 6.758667 0.2477865     30
130 mm 35 mm 80 g 7.157333 0.3052295     30
130 mm 35 mm 90 g 8.918667 0.2500860     30
80 mm 20 mm 80 g  2.222333 0.2754643     30
```

summary(one factorial Anova)

```
Multiple Comparisons of Means: Tukey Contrasts


Fit: aov.default(formula = FZ ~ variable, data = Dataset)

Linear Hypotheses:
                                         Estimate Std. Error t value
130 mm 20 mm 90 g - 130 mm 20 mm 80 g == 0  1.57567    0.06759  23.312
130 mm 35 mm 80 g - 130 mm 20 mm 80 g == 0  1.97433    0.06759  29.211
130 mm 35 mm 90 g - 130 mm 20 mm 80 g == 0  3.73567    0.06759  55.270
80 mm 20 mm 80 g - 130 mm 20 mm 80 g == 0  -2.96067    0.06759 -43.804
80 mm 20 mm 90 g - 130 mm 20 mm 80 g == 0  -1.31800    0.06759 -19.500
80 mm 35 mm 80 g - 130 mm 20 mm 80 g == 0  -0.82533    0.06759 -12.211
80 mm 35 mm 90 g - 130 mm 20 mm 80 g == 0   0.82433    0.06759  12.196
130 mm 35 mm 80 g - 130 mm 20 mm 90 g == 0  0.39867    0.06759   5.898
130 mm 35 mm 90 g - 130 mm 20 mm 90 g == 0  2.16000    0.06759  31.958
80 mm 20 mm 80 g - 130 mm 20 mm 90 g == 0  -4.53633    0.06759 -67.116
80 mm 20 mm 90 g - 130 mm 20 mm 90 g == 0  -2.89367    0.06759 -42.813
80 mm 35 mm 80 g - 130 mm 20 mm 90 g == 0  -2.40100    0.06759 -35.523
80 mm 35 mm 90 g - 130 mm 20 mm 90 g == 0  -0.75133    0.06759 -11.116
130 mm 35 mm 90 g - 130 mm 35 mm 80 g == 0  1.76133    0.06759  26.059
80 mm 20 mm 80 g - 130 mm 35 mm 80 g == 0  -4.93500    0.06759 -73.015
80 mm 20 mm 90 g - 130 mm 35 mm 80 g == 0  -3.29233    0.06759 -48.711
80 mm 35 mm 80 g - 130 mm 35 mm 80 g == 0  -2.79967    0.06759 -41.422
80 mm 35 mm 90 g - 130 mm 35 mm 80 g == 0  -1.15000    0.06759 -17.015
80 mm 20 mm 80 g - 130 mm 35 mm 90 g == 0  -6.69633    0.06759 -99.074
80 mm 20 mm 90 g - 130 mm 35 mm 90 g == 0  -5.05367    0.06759 -74.770
80 mm 35 mm 80 g - 130 mm 35 mm 90 g == 0  -4.56100    0.06759 -67.481
80 mm 35 mm 90 g - 130 mm 35 mm 90 g == 0  -2.91133    0.06759 -43.074
80 mm 20 mm 90 g - 80 mm 20 mm 80 g == 0    1.64267    0.06759  24.304
80 mm 35 mm 80 g - 80 mm 20 mm 80 g == 0    2.13533    0.06759  31.593
80 mm 35 mm 90 g - 80 mm 20 mm 80 g == 0    3.78500    0.06759  56.000
80 mm 35 mm 80 g - 80 mm 20 mm 90 g == 0    0.49267    0.06759   7.289
80 mm 35 mm 90 g - 80 mm 20 mm 90 g == 0    2.14233    0.06759  31.696
80 mm 35 mm 90 g - 80 mm 35 mm 80 g == 0    1.64967    0.06759  24.407
```

```
                                           Pr(>|t|)
130 mm 20 mm 90 g - 130 mm 20 mm 80 g == 0 <0.000001 ***
130 mm 35 mm 80 g - 130 mm 20 mm 80 g == 0 <0.000001 ***
130 mm 35 mm 90 g - 130 mm 20 mm 80 g == 0 <0.000001 ***
80 mm 20 mm 80 g - 130 mm 20 mm 80 g == 0  <0.000001 ***
80 mm 20 mm 90 g - 130 mm 20 mm 80 g == 0  <0.000001 ***
80 mm 35 mm 80 g - 130 mm 20 mm 80 g == 0  <0.000001 ***
80 mm 35 mm 90 g - 130 mm 20 mm 80 g == 0  <0.000001 ***
130 mm 35 mm 80 g - 130 mm 20 mm 90 g == 0 <0.000001 ***
130 mm 35 mm 90 g - 130 mm 20 mm 90 g == 0 <0.000001 ***
80 mm 20 mm 80 g - 130 mm 20 mm 90 g == 0  <0.000001 ***
80 mm 20 mm 90 g - 130 mm 20 mm 90 g == 0  <0.000001 ***
80 mm 35 mm 80 g - 130 mm 20 mm 90 g == 0  <0.000001 ***
80 mm 35 mm 90 g - 130 mm 20 mm 90 g == 0  <0.000001 ***
130 mm 35 mm 90 g - 130 mm 35 mm 80 g == 0 <0.000001 ***
80 mm 20 mm 80 g - 130 mm 35 mm 80 g == 0  <0.000001 ***
80 mm 20 mm 90 g - 130 mm 35 mm 80 g == 0  <0.000001 ***
80 mm 35 mm 80 g - 130 mm 35 mm 80 g == 0  <0.000001 ***
80 mm 35 mm 90 g - 130 mm 35 mm 80 g == 0  <0.000001 ***
80 mm 20 mm 80 g - 130 mm 35 mm 90 g == 0  <0.000001 ***
80 mm 20 mm 90 g - 130 mm 35 mm 90 g == 0  <0.000001 ***
80 mm 35 mm 80 g - 130 mm 35 mm 90 g == 0  <0.000001 ***
80 mm 35 mm 90 g - 130 mm 35 mm 90 g == 0  <0.000001 ***
80 mm 20 mm 90 g - 80 mm 20 mm 80 g == 0   <0.000001 ***
80 mm 35 mm 80 g - 80 mm 20 mm 80 g == 0   <0.000001 ***
80 mm 35 mm 90 g - 80 mm 20 mm 80 g == 0   <0.000001 ***
80 mm 35 mm 80 g - 80 mm 20 mm 90 g == 0   <0.000001 ***
80 mm 35 mm 90 g - 80 mm 20 mm 90 g == 0   <0.000001 ***
80 mm 35 mm 90 g - 80 mm 35 mm 80 g == 0   <0.000001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)


         Simultaneous Confidence Intervals

Multiple Comparisons of Means: Tukey Contrasts


Fit: aov.default(formula = FZ ~ variable, data = Dataset)
```

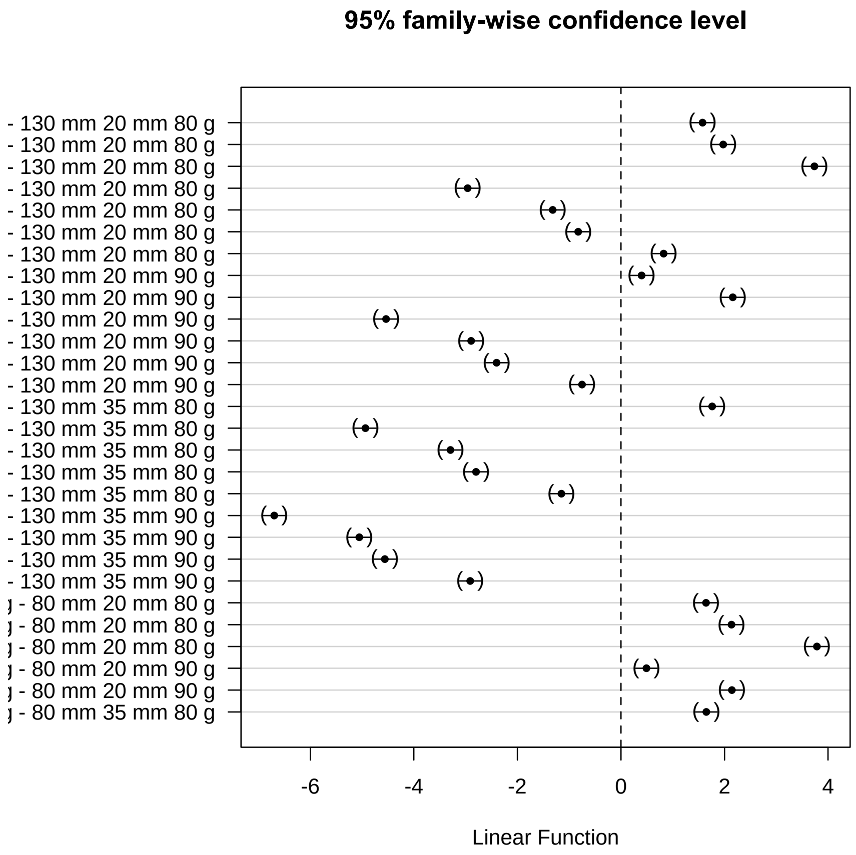
One way Factor Anova

```
Fit: aov.default(formula = FZ ~ variable, data = Dataset)

Quantile = 3.0565
95% family-wise confidence level


Linear Hypotheses:
                                          Estimate lwr     upr
130 mm 20 mm 90 g - 130 mm 20 mm 80 g == 0  1.5757   1.3691  1.7823
130 mm 35 mm 80 g - 130 mm 20 mm 80 g == 0  1.9743   1.7677  2.1809
130 mm 35 mm 90 g - 130 mm 20 mm 80 g == 0  3.7357   3.5291  3.9423
80 mm 20 mm 80 g - 130 mm 20 mm 80 g == 0  -2.9607  -3.1673 -2.7541
80 mm 20 mm 90 g - 130 mm 20 mm 80 g == 0  -1.3180  -1.5246 -1.1114
80 mm 35 mm 80 g - 130 mm 20 mm 80 g == 0  -0.8253  -1.0319 -0.6187
80 mm 35 mm 90 g - 130 mm 20 mm 80 g == 0   0.8243   0.6177  1.0309
130 mm 35 mm 80 g - 130 mm 20 mm 90 g == 0  0.3987   0.1921  0.6053
130 mm 35 mm 90 g - 130 mm 20 mm 90 g == 0  2.1600   1.9534  2.3666
80 mm 20 mm 80 g - 130 mm 20 mm 90 g == 0  -4.5363  -4.7429 -4.3297
80 mm 20 mm 90 g - 130 mm 20 mm 90 g == 0  -2.8937  -3.1003 -2.6871
80 mm 35 mm 80 g - 130 mm 20 mm 90 g == 0  -2.4010  -2.6076 -2.1944
80 mm 35 mm 90 g - 130 mm 20 mm 90 g == 0  -0.7513  -0.9579 -0.5447
130 mm 35 mm 90 g - 130 mm 35 mm 80 g == 0  1.7613   1.5547  1.9679
80 mm 20 mm 80 g - 130 mm 35 mm 80 g == 0  -4.9350  -5.1416 -4.7284
80 mm 20 mm 90 g - 130 mm 35 mm 80 g == 0  -3.2923  -3.4989 -3.0857
80 mm 35 mm 80 g - 130 mm 35 mm 80 g == 0  -2.7997  -3.0063 -2.5931
80 mm 35 mm 90 g - 130 mm 35 mm 80 g == 0  -1.1500  -1.3566 -0.9434
80 mm 20 mm 80 g - 130 mm 35 mm 90 g == 0  -6.6963  -6.9029 -6.4897
80 mm 20 mm 90 g - 130 mm 35 mm 90 g == 0  -5.0537  -5.2603 -4.8471
80 mm 35 mm 80 g - 130 mm 35 mm 90 g == 0  -4.5610  -4.7676 -4.3544
80 mm 35 mm 90 g - 130 mm 35 mm 90 g == 0  -2.9113  -3.1179 -2.7047
80 mm 20 mm 90 g - 80 mm 20 mm 80 g == 0    1.6427   1.4361  1.8493
80 mm 35 mm 80 g - 80 mm 20 mm 80 g == 0    2.1353   1.9287  2.3419
80 mm 35 mm 90 g - 80 mm 20 mm 80 g == 0    3.7850   3.5784  3.9916
80 mm 35 mm 80 g - 80 mm 20 mm 90 g == 0    0.4927   0.2861  0.6993
80 mm 35 mm 90 g - 80 mm 20 mm 90 g == 0    2.1423   1.9357  2.3489
80 mm 35 mm 90 g - 80 mm 35 mm 80 g == 0    1.6497   1.4431  1.8563

130 mm 20 mm 80 g 130 mm 20 mm 90 g 130 mm 35 mm 80 g 130 mm 35 mm 90 g
              "a"               "b"               "c"               "d"
 80 mm 20 mm 80 g  80 mm 20 mm 90 g  80 mm 35 mm 80 g  80 mm 35 mm 90 g
              "e"               "f"               "g"               "h"
```

80 mm 35 mm 90 g - 80 mm 20 mm 80 g == 0   3.7850   3.5781 3.9919

This factor has the maximum effect.

95% family-wise confidence level



Conclusion

As a conclusion, this 80 mm 35 mm 90 g - 80 mm 20 mm 80 g == 0 3.7850 3.5781 3.9919 factor is a big influence in the improvement.

We would like to say the factors has effect on the flight improvement.

Short Report to the Company

In the Marketing what we have found is how children and parents prefer to buy products. In both Marketing (Kinder and Eltern) the H0 is rejected and it goes to alternative Hyphotoses (H1). It means, that there is difference, between boys and girls's ages Buypreferences and the parents's Buypreferences depends on gender(boys and girls). And it is signifincant difference.

Alternative: according to our analysis, The Data is not normally distributed, and both variables (Base and Alternative) didn't pass to transformation, at this case the non-parametric Wilcoxon_Test was used. This means the Data as Non_Parametric Method tested.

As the Box-cox transformation has shown us that the transformation is not possible to be shift to the parametric method. In this case, we inform the production head that it is analyzed with Non-parametric methods.

Also, the rate of change is **clearly identified.**

Verlgeiche: According to the Anova test in the case of independent Models the P-KL models have positive combinations but the rest combinations have negative combinations. And in the case of dependent variables models they are more significant to each other.

So there is are a difference in the effectiveness of the change.We can conclude that the Base Model with characteristics P and KL are the best.

Optimization: Except one variable that is different and significant, all the other factors has equal effect on the flight time improvement.

Furthermore, each separate part hast its ellaborated report as conclusion.

At the end everything is based on statistical data from the company.

If there is any question related to the anaylsis please keep in touch with us.