

Title: "Supplier Quality Analysis"

author: "Siamak Goudarzi"

date: "2026-01-18"

output: pdf_document

Fragestellung

Dataset

Analyse

Schritt 1: Daten laden

Schritt 2: Erste Überprüfung und deskriptive Statistik

 Frequencies: Lieferant, Zuverlässigkeit

Interpretation der Ergebnisse

Schritt 3: Bivariate Deskriptivstatistik (Zusammenhang prüfen)

Gewählte Methode und deren Begründung

Hypothesen des Tests

 Two-Way Contingency Table: Lieferant, Zuverlässigkeit

Interpretation der Ergebnisse

Erstellung der Grafik

 Bar Plot: Lieferant

Interpretation der Grafik

Schritt 4: Statistischer Test (Chi-Quadrat-Test)

 Two-Way Contingency Table: Lieferant, Zuverlässigkeit

Interpretation der Ergebnisse

Schritt 5: Analyse des zweiten Teils

 Bar Plot: Produktgruppe

Interpretation der Ergebnisse

 Two-Way Contingency Table: Produktgruppe, Zuverlässigkeit

Interpretation der Ergebnisse

Abschließende Analyse:

Title: "Supplier Quality Analysis"

author: "Siamak Goudarzi"

date: "2026-01-18"

output: pdf_document

Fragestellung

Teil A:

Gibt es einen signifikanten Unterschied in der Produktqualität zwischen den drei Lieferanten?

Teil B:

Besteht ein Zusammenhang zwischen Produktgruppe und Qualität?

Dataset

Die Daten umfassen den Lieferantennamen (A, B, C), die Produktart und die Qualitätsstufe (hoch, mittel, niedrig).

Analyse

Da es sich bei den Daten um nominale/ordinale Daten handelt, wurde der Chi-Quadrat-Test bzw. die Kontingenzanalyse verwendet.

Schritt 1: Daten laden

```
> DS_Lieferanten <- readXL("/home/siamak/Downloads/Daten_KW03.xlsx",  
+ rownames=FALSE, header=TRUE, na="", sheet="1_Lieferantenbeurteilung",  
+ stringsAsFactors=TRUE)
```

Schritt 2: Erste Überprüfung und deskriptive Statistik

Frequencies: Lieferant, Zuverlässigkeit

```
> local({  
+ .Table <- with(DS_Lieferanten, table(Lieferant))  
+ cat("\ncounts:\n")  
+ print(.Table)  
+ cat("\npercentages:\n")  
+ print(round(100* .Table/sum(.Table), 2))  
+ })
```

```
counts:  
Lieferant  
Lieferant A Lieferant B Lieferant C  
160      200      199
```

```
percentages:
Lieferant
Lieferant A Lieferant B Lieferant C
      28.62      35.78      35.60
```

```
> local({
+   .Table <- with(DS_Lieferanten, table(Zuverlässigkeit))
+   cat("\ncounts:\n")
+   print(.Table)
+   cat("\npercentages:\n")
+   print(round(100*.Table/sum(.Table), 2))
+ })
```

```
counts:
Zuverlässigkeit
      Hoch Mittel Niedrig
      190     190     179
```

```
percentages:
Zuverlässigkeit
      Hoch Mittel Niedrig
      33.99  33.99  32.02
```

```
> library(abind, pos=17)
```

Interpretation der Ergebnisse

Lieferanten:

Insgesamt liegen 559 Datenpunkte vor.

Die Lieferanten B und C stellen mit jeweils ca. 36 % die größte Gruppe dar, während Lieferant A ca. 28,6 % der Daten beisteuert.

Zuverlässigkeit:

Die Qualitätsverteilung ist sehr gleichmäßig.

Die Kategorien „Hoch“ und „Durchschnittlich“ liegen mit jeweils 33,99 % exakt gleichauf, dicht gefolgt von „Niedrig“ mit 32 %.

>> Dies bedeutet, dass unsere Daten keine gravierenden Ungleichgewichte aufweisen.

Schritt 3: Bivariate Deskriptivstatistik

(Zusammenhang prüfen)

Gewählte Methode und deren Begründung

Da beide Variablen (Lieferant und Qualität) kategorial sind, wurde zur Beschreibung eine Kreuztabelle verwendet und zur Überprüfung der Hypothese der Chi-Quadrat-Test (χ^2 -Test) eingesetzt.

Hypothesen des Tests

Nullhypothese (H_0):

Die Produktqualität und der Lieferant sind voneinander unabhängig.

Alternativhypothese (H_1):

Die Produktqualität hängt vom Lieferanten ab.

Two-Way Contingency Table: Lieferant, Zuverlässigkeit

```
> local({
+   .Table <- xtabs(~Lieferant+Zuverlässigkeit, data=DS_Lieferanten)
+   cat("\nFrequency table:\n")
+   print(.Table)
+   cat("\nRow percentages:\n")
+   print(rowPercents(.Table))
+ })
```

Frequency table:

| | Zuverlässigkeit | | |
|-------------|-----------------|--------|---------|
| Lieferant | Hoch | Mittel | Niedrig |
| Lieferant A | 80 | 50 | 30 |
| Lieferant B | 70 | 60 | 70 |
| Lieferant C | 40 | 80 | 79 |

Row percentages:

| | Zuverlässigkeit | | | |
|-------------|-----------------|--------|---------|-------------|
| Lieferant | Hoch | Mittel | Niedrig | Total Count |
| Lieferant A | 50.0 | 31.2 | 18.8 | 100 |

| | | | | | |
|-------------|------|------|------|-----|-----|
| Lieferant B | 35.0 | 30.0 | 35.0 | 100 | 200 |
| Lieferant C | 20.1 | 40.2 | 39.7 | 100 | 199 |

Interpretation der Ergebnisse

Tabelle der Zeilenprozentsätze:

Lieferant A:

zeigt die beste Leistung. 50 % seiner Produkte weisen eine hohe Qualität („Hoch“) auf, und nur etwa 18,8 % fallen in die niedrige Qualitätsstufe.

Lieferant B:

hat eine durchschnittliche Leistung. Die Qualitätsverteilung ist nahezu gleichmäßig (35 % hoch und 35 % niedrig).

Lieferant C:

weist die schwächste Leistung auf. Nur 20 % seiner Produkte haben eine hohe Qualität, und rund 80 % liegen im mittleren oder niedrigen Qualitätsbereich.

>> Vorläufiges Fazit

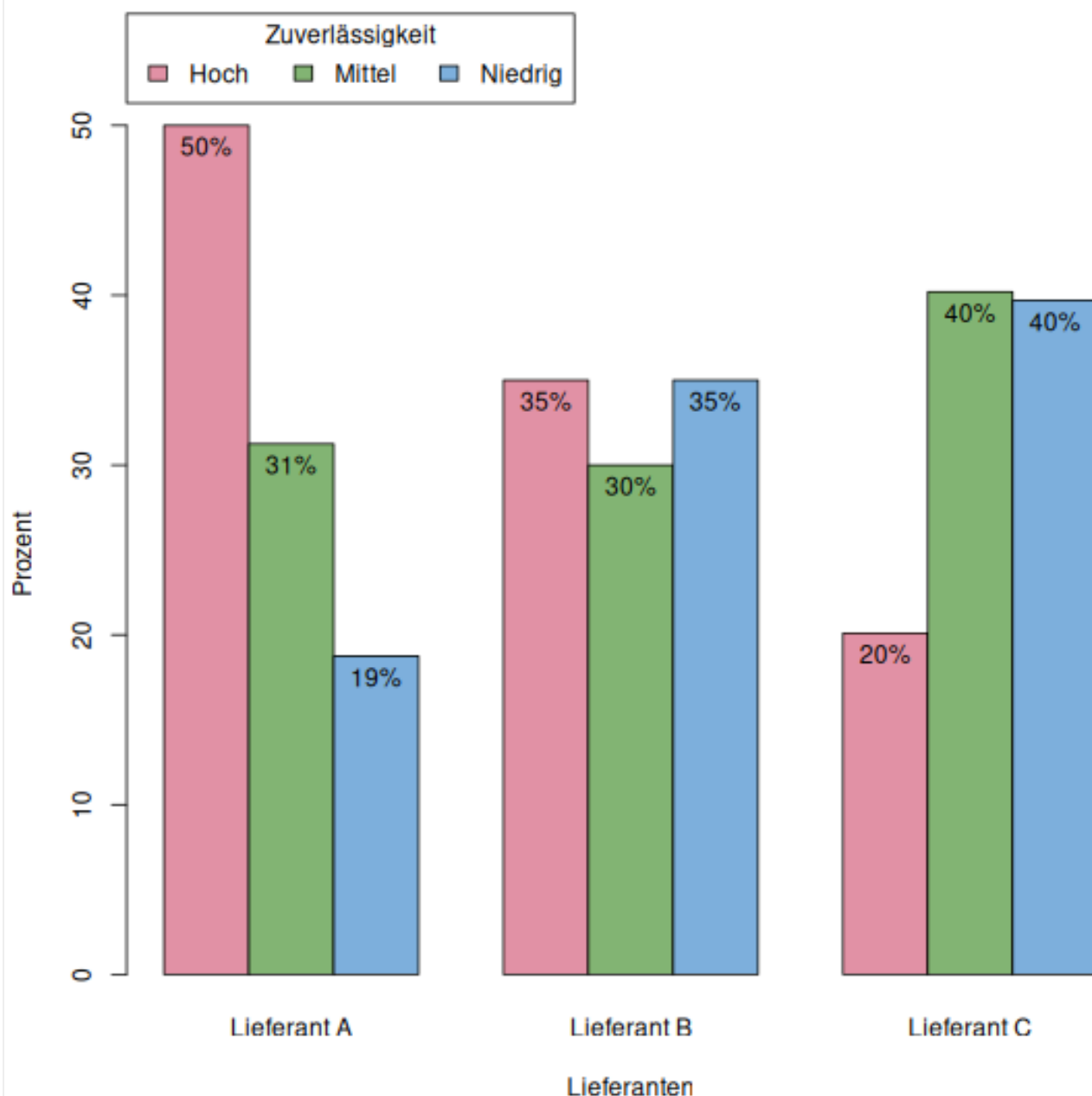
Aus der deskriptiven und visuellen Betrachtung wird deutlich, dass die Qualität vom Lieferanten abhängt. Nun muss mit einem statistischen Test überprüft werden, ob dieser Unterschied nicht zufällig ist.

Erstellung der Grafik

Um den Zusammenhang visuell darzustellen

Bar Plot: Lieferant

```
> with(DS_Lieferanten, Barplot(Lieferant, by=Zuverlässigkeit,  
+ style="parallel", legend.pos="above", xlab="Lieferanten", ylab="Prozent",  
+ scale="percent", label.bars=TRUE))
```



plot of chunk unnamed-chunk-6

Interpretation der Grafik

In diesem Diagramm, das als paralleles Balkendiagramm dargestellt ist, entspricht jede Balkengruppe einem Lieferanten, und die Farben zeigen die Qualitätsstufen (Hoch, Mittel, Niedrig) an.

Lieferant A:

Der höchste Balken in dieser Gruppe gehört zur Kategorie Hoch. Dies zeigt, dass Lieferant A mit deutlichem Abstand den größten Anteil an hochwertigen Produkten in seinem Sortiment hat (gemäß unseren vorherigen Daten 50 %).

Lieferant B:

Die Balken dieses Lieferanten sind nahezu gleich hoch. Das bedeutet, dass es für diesen Lieferanten kaum Unterschiede gibt; er produziert in etwa gleich viele Produkte mit hoher, mittlerer und niedriger Qualität. Eine klare Qualitätsstabilität ist nicht erkennbar.

Lieferant C:

Hier sind die Balken für Mittel und Niedrig deutlich höher als der Balken für Hoch. Dies weist darauf hin, dass Lieferant C stark dazu tendiert, Produkte mit niedriger oder höchstens mittlerer Qualität anzubieten.

>> Fazit:

Das Balkendiagramm zeigt deutliche Unterschiede in der Qualitätsstruktur der Lieferanten.

Während Lieferant A einen klaren Schwerpunkt auf hohe Qualität legt, bietet Lieferant C überwiegend Produkte mit niedriger und mittlerer Qualität an.

Schritt 4: Statistischer Test (Chi-Quadrat-Test)

Nachdem wir das Problem nun visuell dargestellt haben, müssen wir unsere Ergebnisse präzise beweisen.

Two-Way Contingency Table: Lieferant, Zuverlässigkeit

```
> local({
+   .Table <- xtabs(~Lieferant+Zuverlässigkeit, data=DS_Lieferanten)
+   cat("\nFrequency table:\n")
+   print(.Table)
+   cat("\nRow percentages:\n")
+   print(rowPercents(.Table))
+   .Test <- chisq.test(.Table, correct=FALSE)
+   print(.Test)
+ })
```

Frequency table:

| | Zuverlässigkeit | | |
|-------------|-----------------|--------|---------|
| Lieferant | Hoch | Mittel | Niedrig |
| Lieferant A | 80 | 50 | 30 |
| Lieferant B | 70 | 60 | 70 |
| Lieferant C | 40 | 80 | 79 |

Row percentages:

```

                Zuverlässigkeit
Lieferant      Hoch Mittel Niedrig Total Count
Lieferant A  50.0   31.2   18.8   100   160
Lieferant B  35.0   30.0   35.0   100   200
Lieferant C  20.1   40.2   39.7   100   199

```

Pearson's Chi-squared test

```

data: .Table
X-squared = 39.987, df = 4, p-value = 0.00000004355

```

Interpretation der Ergebnisse

$X^2 = 39,987$:

Dieser Wert gibt die Differenz zwischen den beobachteten und den erwarteten Werten an. Je größer er ist, desto höher ist die Wahrscheinlichkeit einer Abhängigkeit.

p-Wert = 0,00000004355:

Dieser Wert ist sehr kleiner als 0,05.

>> Fazit:

Da der p-Wert deutlich unter dem 5%-Fehlerniveau (0,05) liegt, wird die Nullhypothese (H_0) der Unabhängigkeit verworfen. Somit ist statistisch mit hoher Wahrscheinlichkeit bewiesen, dass die Produktqualität stark vom Lieferanten abhängt.

Schritt 5: Analyse des zweiten Teils

Der Einfluss der Produktgruppe auf die Qualität

Ausgewählte Methode:

Ähnlich wie im vorherigen Teil wurde der Chi-Quadrat-Unabhängigkeitstest verwendet, um zu untersuchen, ob die Produktart (Elektronik, Mechanik usw.) ein Faktor für den Qualitätsverlust ist.

Hypothesenprüfung:

Hypothese 0 (H_0):

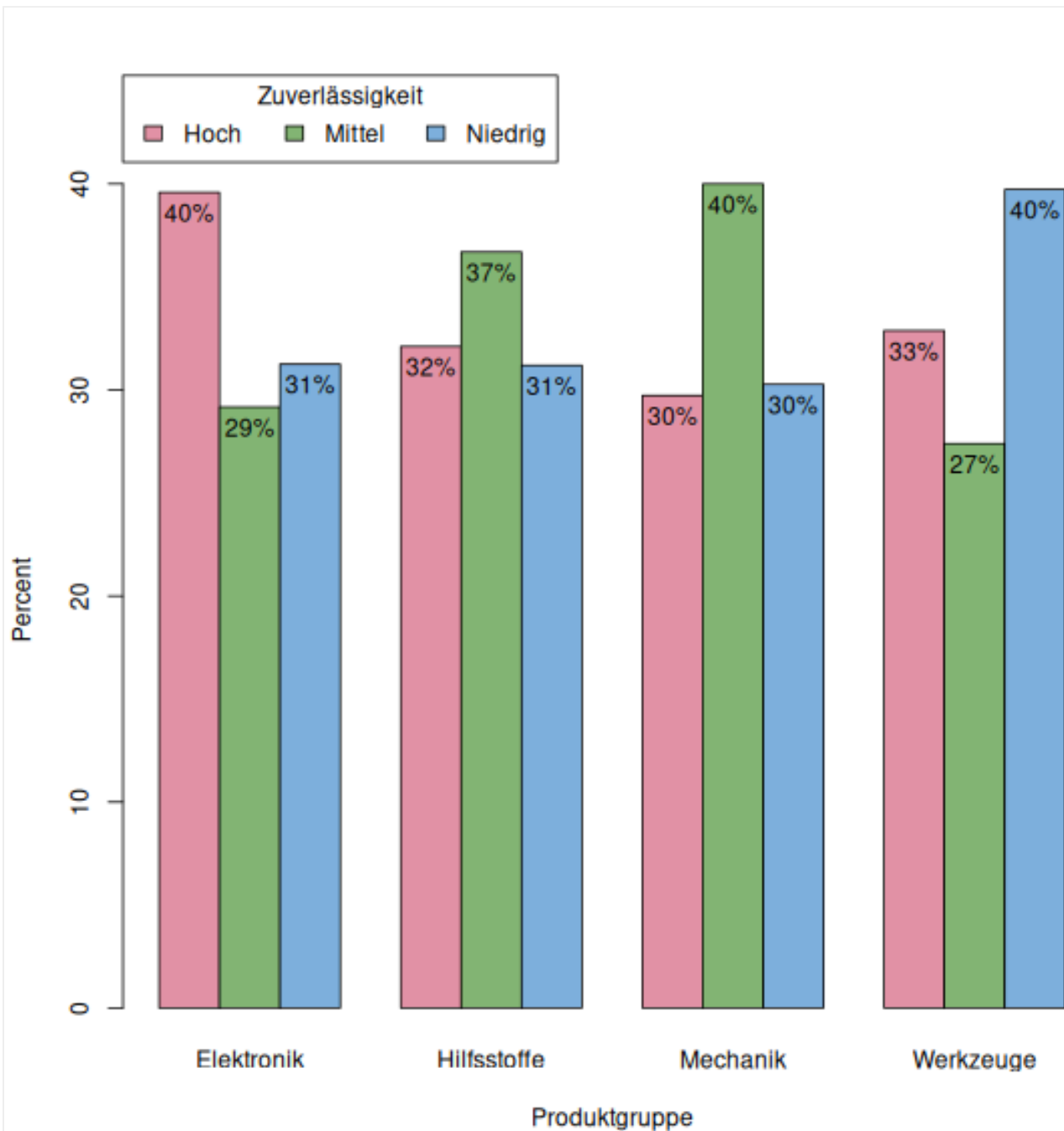
Die Produktqualität und die Produktgruppe sind unabhängig.

Hypothese 1 (H_1):

ie Produktqualität hängt von der Produktgruppe ab.

Bar Plot: Produktgruppe

```
> with(DS_Lieferanten, Barplot(Produktgruppe, by=Zuverlässigkeit,  
+ style="parallel", legend.pos="above", xlab="Produktgruppe", ylab="Percent",  
+ scale="percent", label.bars=TRUE))
```



plot of chunk unnamed-chunk-8

Interpretation der Ergebnisse

Anders als im Lieferantendiagramm (wo Lieferant A sich deutlich von Lieferant C unterschied), weisen hier alle vier Produktgruppen (Elektronik, Hilfsstoffe, Mechanik, Werkzeuge) eine weitgehend ähnliche Struktur auf.

In allen Gruppen unterscheiden sich die Säulen für die Qualitätsstufen Hoch, Mittel und Niedrig in ihrer Höhe nicht signifikant von den anderen Gruppen.

>> Visuelles Ergebnis: Das Diagramm zeigt, dass die Wahrscheinlichkeit, ein gutes oder schlechtes Produkt zu erhalten, unabhängig vom bestellten Produkt in allen Gruppen nahezu gleich ist.

Two-Way Contingency Table: Produktgruppe, Zuverlässigkeit

```
> local({
+   .Table <- xtabs(~Produktgruppe+Zuverlässigkeit, data=DS_Lieferanten)
+   cat("\nFrequency table:\n")
+   print(.Table)
+   cat("\nRow percentages:\n")
+   print(rowPercents(.Table))
+   .Test <- chisq.test(.Table, correct=FALSE)
+   print(.Test)
+ })
```

Frequency table:

| | Zuverlässigkeit | | |
|---------------|-----------------|--------|---------|
| Produktgruppe | Hoch | Mittel | Niedrig |
| Elektronik | 76 | 56 | 60 |
| Hilfsstoffe | 35 | 40 | 34 |
| Mechanik | 55 | 74 | 56 |
| Werkzeuge | 24 | 20 | 29 |

Row percentages:

| | Zuverlässigkeit | | | Total | Count |
|---------------|-----------------|--------|---------|-------|-------|
| Produktgruppe | Hoch | Mittel | Niedrig | | |
| Elektronik | 39.6 | 29.2 | 31.2 | 100 | 192 |
| Hilfsstoffe | 32.1 | 36.7 | 31.2 | 100 | 109 |
| Mechanik | 29.7 | 40.0 | 30.3 | 100 | 185 |
| Werkzeuge | 32.9 | 27.4 | 39.7 | 100 | 73 |

Pearson's Chi-squared test

data: .Table

X-squared = 8.9333, df = 6, p-value = 0.1774

Interpretation der Ergebnisse

Im Gegensatz zum Lieferantenbereich sind die Ergebnisse hier deutlich anders:

Prozentsätze: In allen Produktgruppen (Elektronik, Werkzeuge, Mechanik usw.) ist die Qualitätsverteilung nahezu identisch.

Beispielsweise liegt der Anteil hochwertiger Produkte in der Kategorie „Elektronik“ bei etwa 39 % und in der Kategorie „Werkzeuge“ bei etwa 32 %. Die Unterschiede sind minimal.

Interpretation des p-Werts: Der Wert beträgt 0,1774.

Da dieser Wert größer als 0,05 ist, wird die Nullhypothese (H_0) nicht verworfen. Dies bedeutet, dass die in Abbildung A dargestellten Unterschiede statistisch nicht signifikant und rein zufällig sind.

>> Schlussfolgerung (Frage B): Im Gegensatz zu den Lieferanten hat die Produktgruppe keinen signifikanten Einfluss auf die Qualität ($p = 0,177$).

Das bedeutet, dass das Qualitätsproblem nicht mit der Produktart zusammenhängt, sondern direkt mit dem Qualitätsmanagement der Lieferantenunternehmen.

Abschließende Analyse:

Zur Qualitätsverbesserung sollte der Fokus auf einem Wechsel oder einer Anpassung des Lieferanten (insbesondere Lieferant C) liegen, nicht auf einer Änderung der Produktart.