

# Übung Lage- und Streuungsmaße

In der Datei *Übung\_Deskriptiv.xlsx* finden Sie verschiedene Datensätze.

Bestimmen Sie für diese Daten sinnvolle Lage- und Streuungsmaße.

Der Datensatz *Ausreisser* zeigt Ihnen die Wirkung von Ausreißern in Daten mit oder ohne Extremwert.

Die Datensätze **Box1** und *Box2* veranschaulichen die Wirkungsweise von Boxplots.

# Übung Lage- und Streuungsmaße

## *Wuerfel*

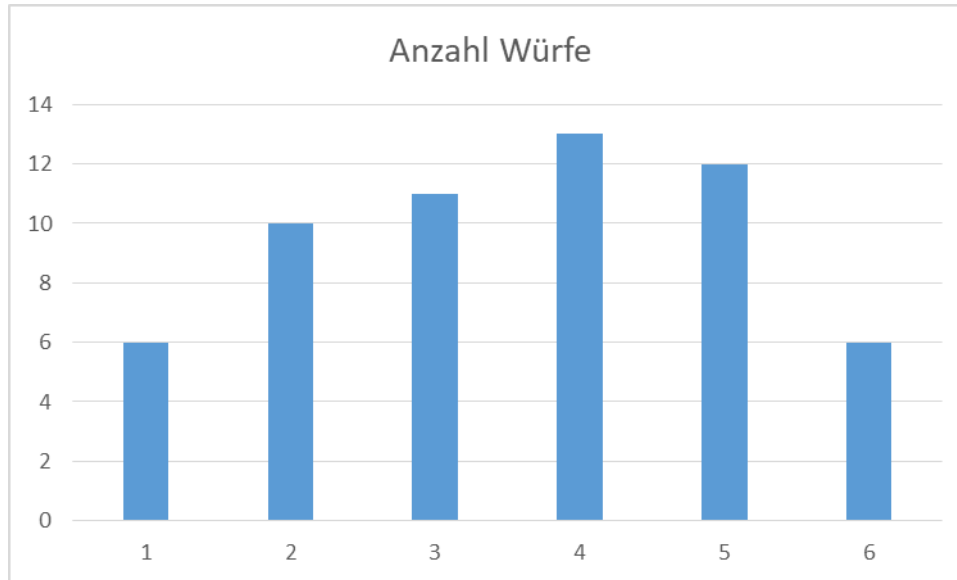
Wie der Datenname und die vorliegenden Werte zeigen handelt es sich um ein Würfelexperiment.

Hier bringen Modalwert und Balkendiagramm den höchsten Mehrwert.

Daten kumulieren, als Balkendiagramm darstellen und schauen, wo der Modalwert liegt.

# Übung Lage- und Streuungsmaße

## Wurfel



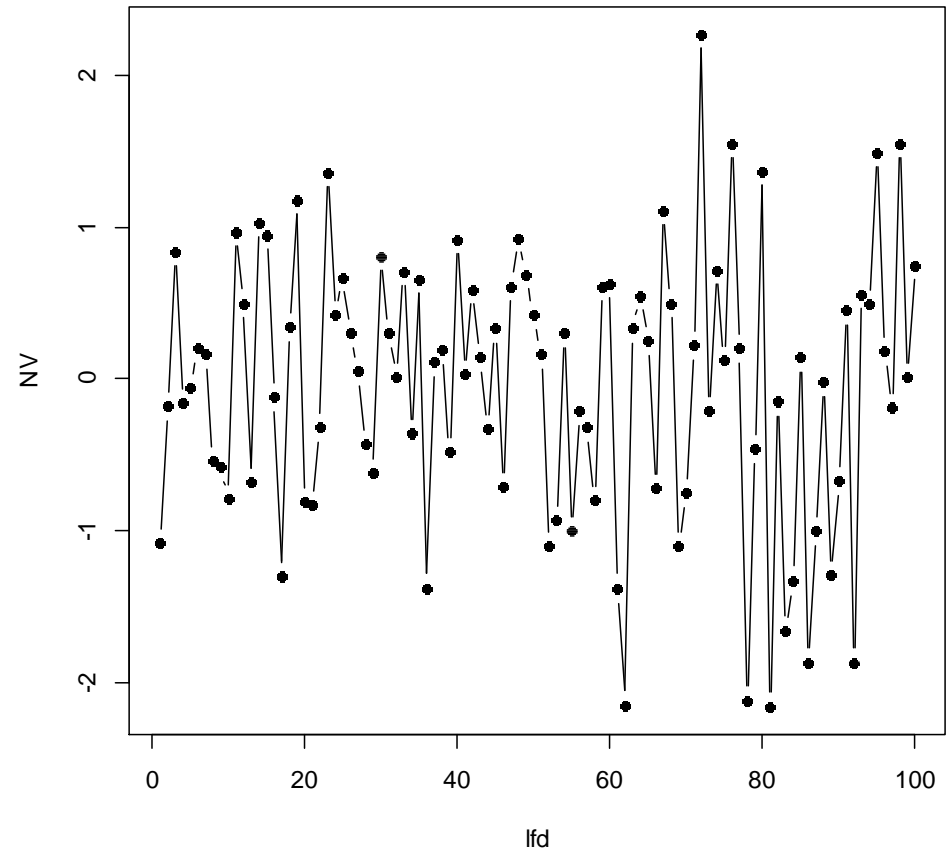
$$x_{mod} = 4$$

Im wahren Leben liegt die Häufigkeit der 4 schon sehr nahe bei den Werten für 3 und 5

# Übung Lage- und Streuungsmaße

NV Erstmal ein Liniendiagramm

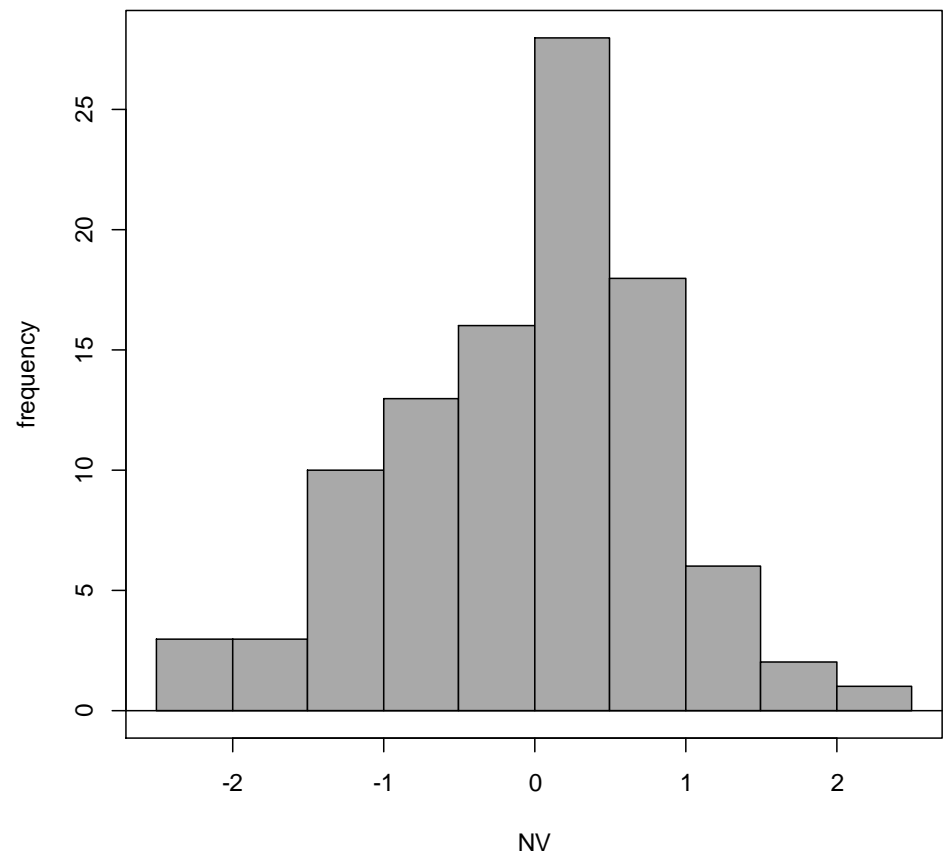
Nichts Aufsehenerregendes



# Übung Lage- und Streuungsmaße

*NV dann ein Histogramm*

Sieht aus wie eine  
Normalverteilung  
(dazu später mehr)



# Übung Lage- und Streuungsmaße

NV

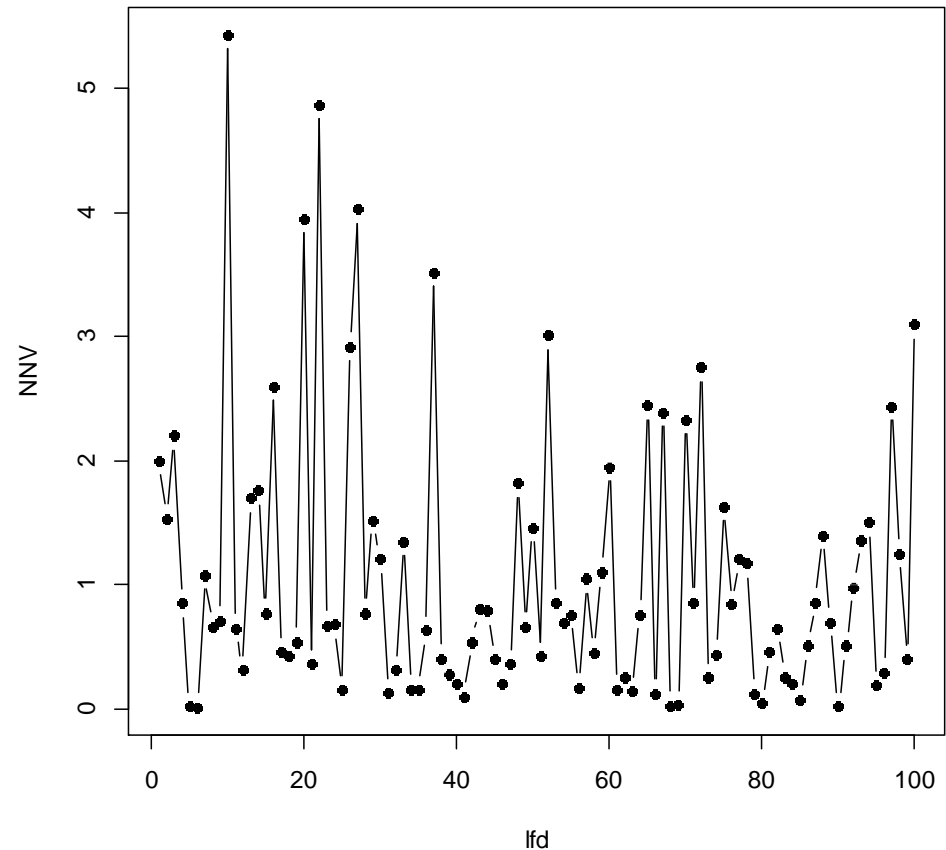
lfd		NV	
Min.	: 1.00	Min.	:-2.1600
1st Qu.	: 25.75	1st Qu.	:-0.6725
Median	: 50.50	Median	: 0.1150
Mean	: 50.50	Mean	:-0.0429
3rd Qu.	: 75.25	3rd Qu.	: 0.5675
Max.	:100.00	Max.	: 2.2700

SD = 0,887

# Übung Lage- und Streuungsmaße

NVV Erstmal ein Liniendiagramm

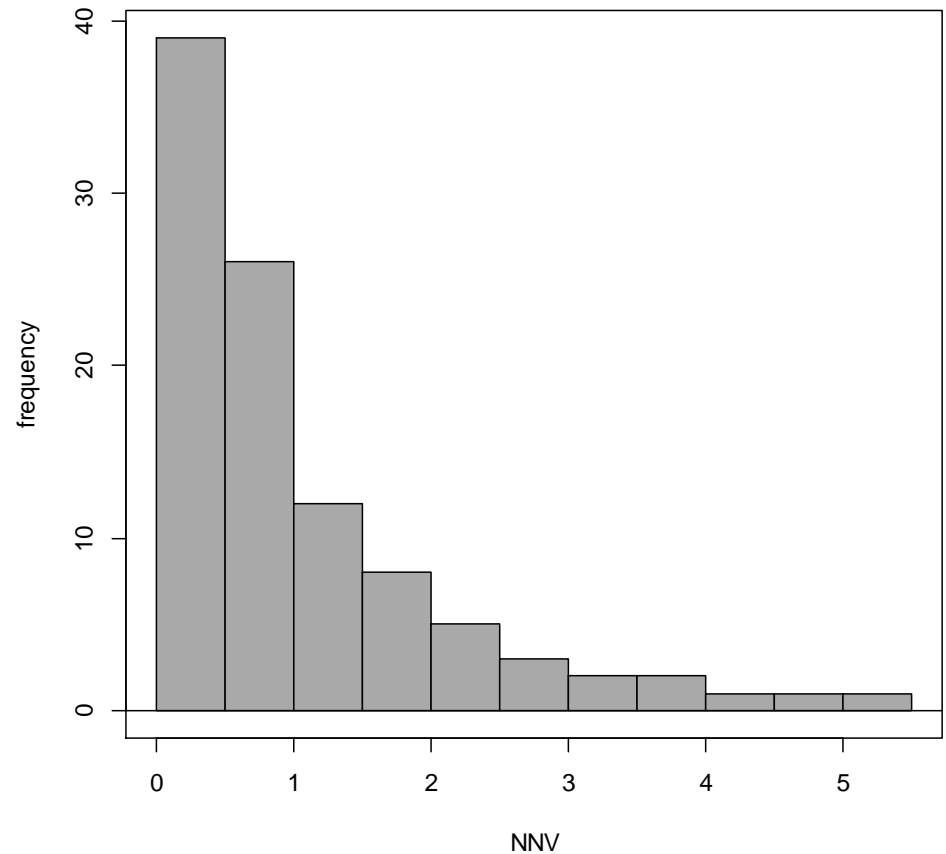
Verschiedene Niveaus?



# Übung Lage- und Streuungsmaße

*NVV dann ein Histogramm*

Eindeutig keine  
Normalverteilung





# Übung Lage- und Streuungsmaße

NVV

lfd		NNV	
Min.	: 1.00	Min.	:0.0100
1st Qu.	: 25.75	1st Qu.	:0.2875
Median	: 50.50	Median	:0.6850
Mean	: 50.50	Mean	:1.0469
3rd Qu.	: 75.25	3rd Qu.	:1.4150
Max.	:100.00	Max.	:5.4300

Hier sind Median und Spannweite sinnvoll

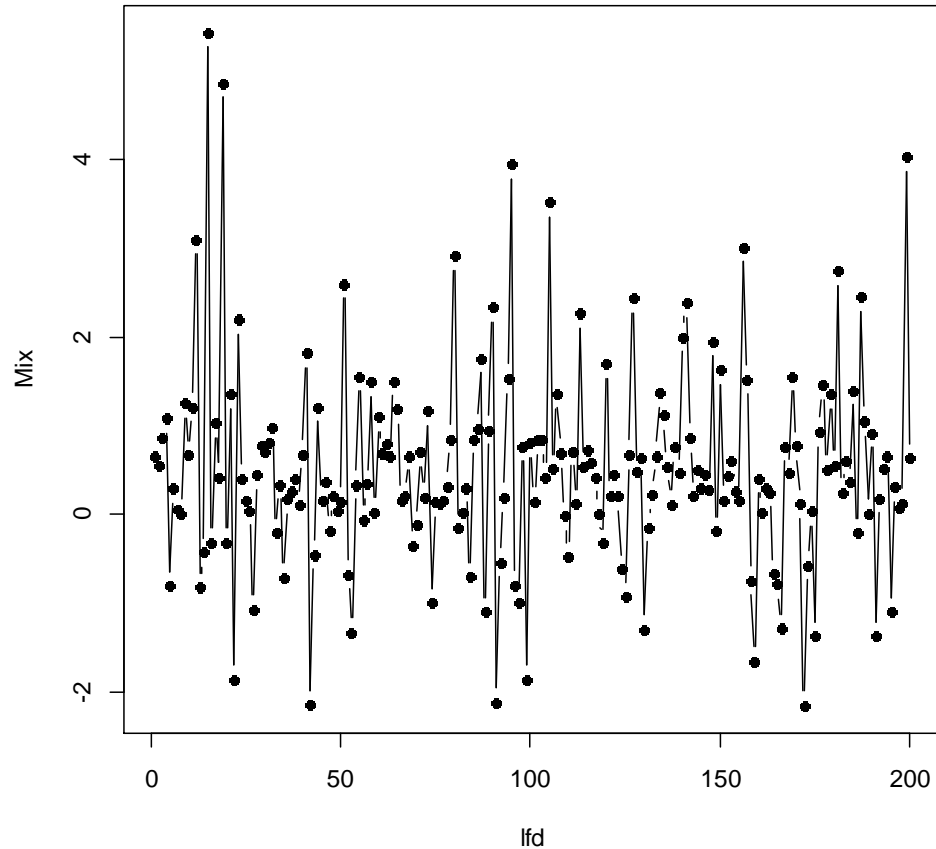
$$\tilde{x} = 0,685$$

$$R = \text{Max.} - \text{Min.} = 5,52$$

# Übung Lage- und Streuungsmaße

*Mix* Erstmal ein Liniendiagramm

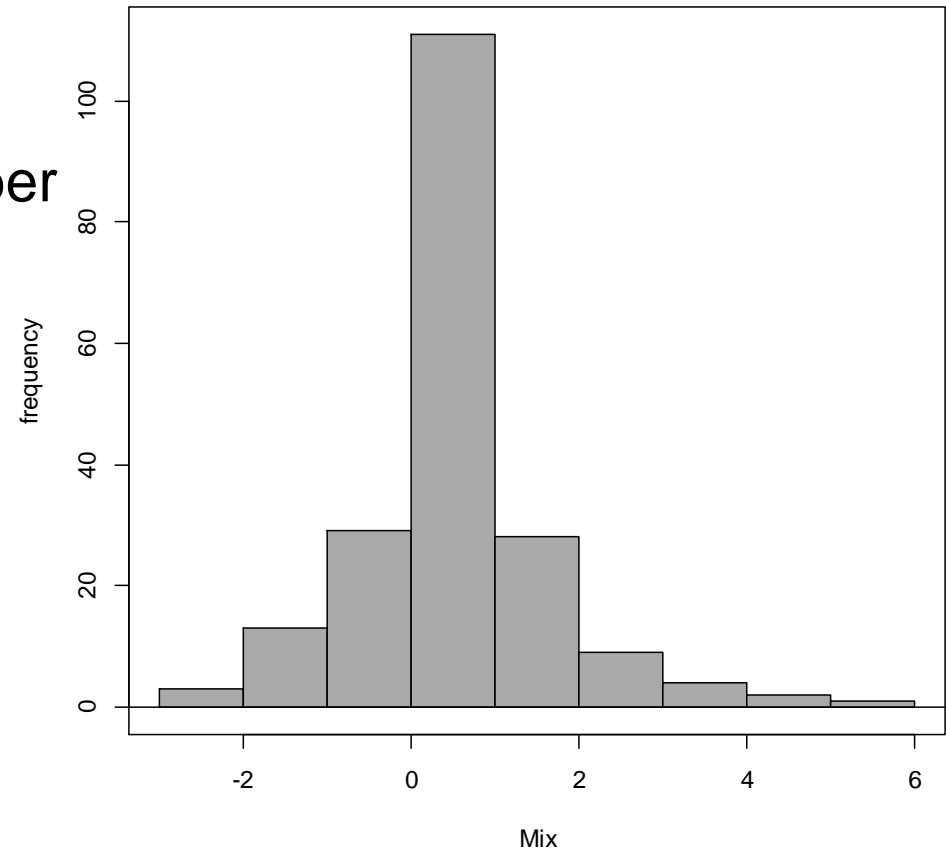
Keine Auffälligkeiten



# Übung Lage- und Streuungsmaße

*Mix dann ein Histogramm*

Eindeutig keine  
Normalverteilung, aus  
Symmetriegründen können aber  
Mittelwert und Standard-  
Abweichung Sinn machen



# Übung Lage- und Streuungsmaße

*Mix*

<i>lfd</i>		<i>Mix</i>	
<i>Min.</i>	: 1.00	<i>Min.</i>	:-2.160
<i>1st Qu.</i>	: 50.75	<i>1st Qu.</i>	: 0.020
<i>Median</i>	:100.50	<i>Median</i>	: 0.410
<i>Mean</i>	:100.50	<i>Mean</i>	: 0.502
<i>3rd Qu.</i>	:150.25	<i>3rd Qu.</i>	: 0.860
<i>Max.</i>	:200.00	<i>Max.</i>	: 5.430

Hier sind Mittelwert und Standardabweichung sinnvoll, alternativ gehen auch Median und Spannweite

$$\bar{x} = 0,502$$

$$\tilde{x} = 0,410$$

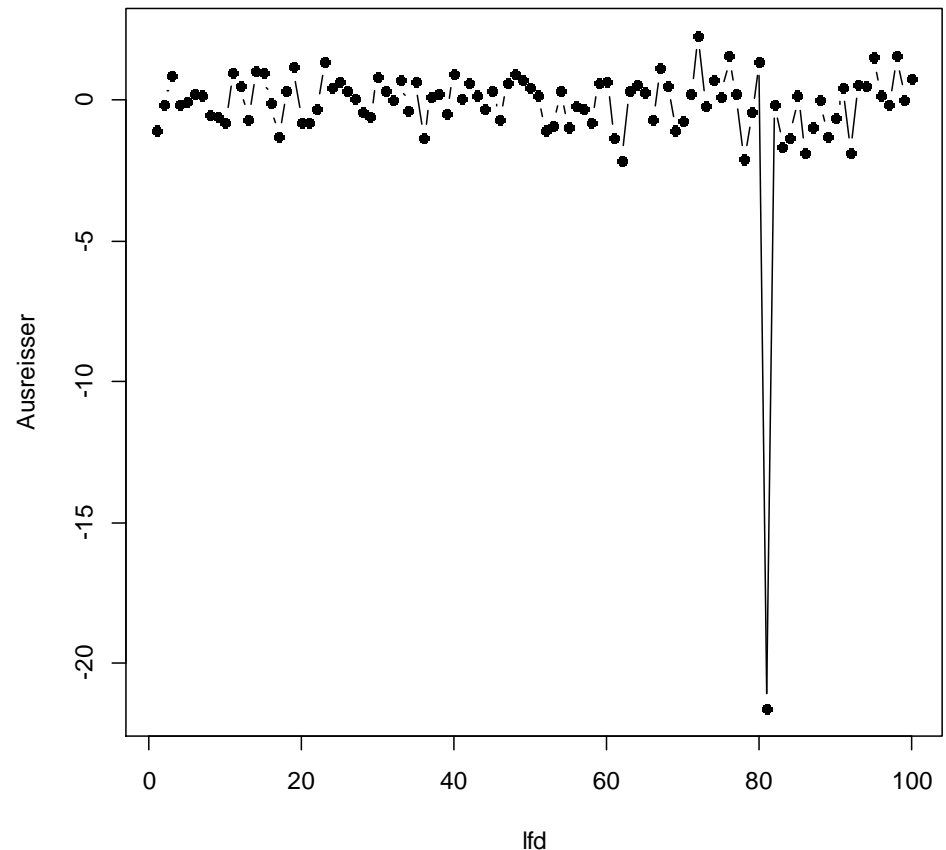
$$sd = 1,1321$$

$$R = Max. - Min. = 7,59$$

# Übung Lage- und Streuungsmaße

*Ausreisser* Erstmal ein Liniendiagramm

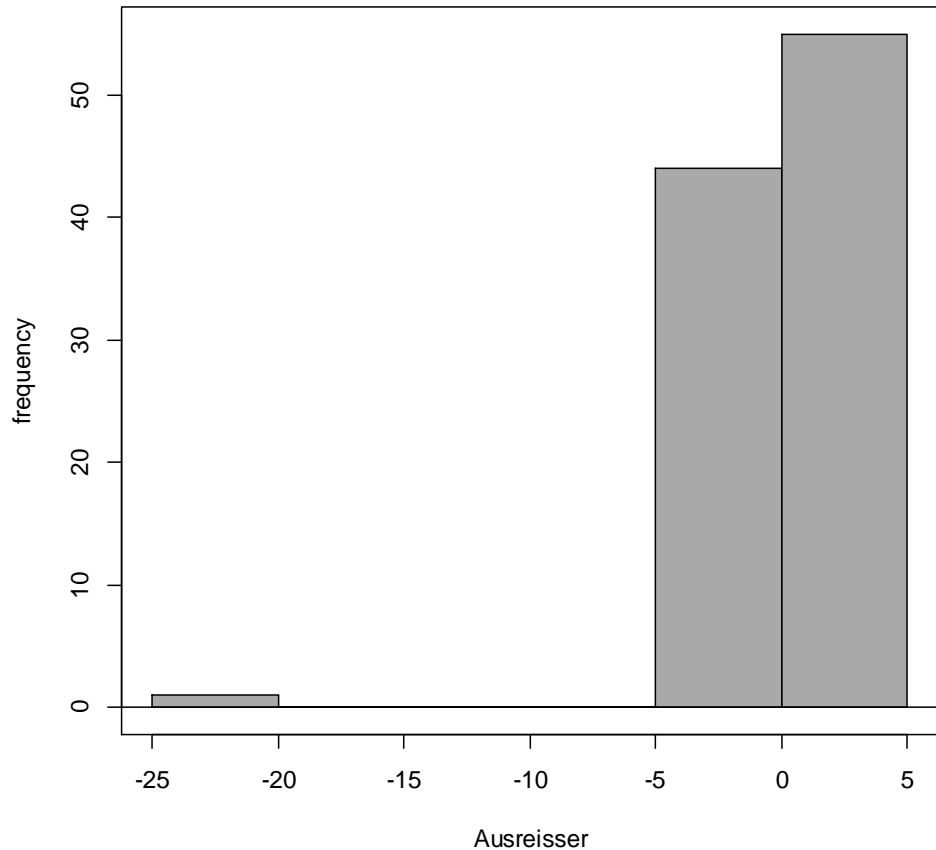
Ups



# Übung Lage- und Streuungsmaße

*Ausreisser dann ein Histogramm*

Eindeutig ein Problem...



# Übung Lage- und Streuungsmaße

## *Aussreisser*

<i>lfd</i>	<i>Aussreisser</i>
Min. : 1.00	Min. : -21.6000
1st Qu.: 25.75	1st Qu.: -0.6725
Median : 50.50	Median : 0.1150
Mean : 50.50	Mean : -0.2373
3rd Qu.: 75.25	3rd Qu.: 0.5675
Max. : 100.00	Max. : 2.2700

Wir wollen die Wirkung auf den Mittelwert sehen:

$$\bar{x} = -0,2373$$

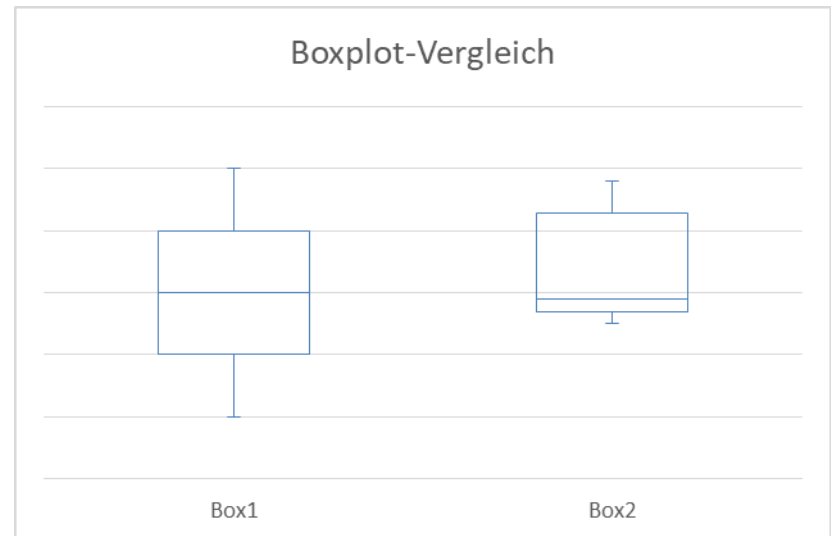
$$sd = 2,3234$$

Bei den Daten handelt es sich um den Datensatz NV mit einem Schreibfehler

# Übung Lage- und Streuungsmaße

*Box1, Box2*

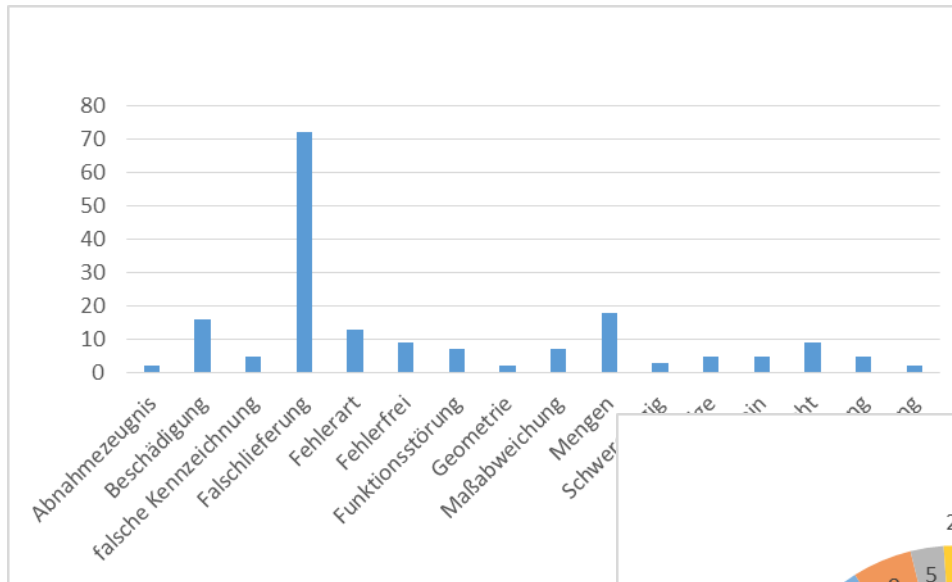
*Ähnliche Lage, aber deutliche  
Unterschiede bei der Symmetrie*





# Übung Lage- und Streuungsmaße

## Fehler



*Reale Skurrilitäten:  
Fehlerart und Fehlerfrei  
tauchen als Eintrag  
in einer Fehlerliste auf!  
(Ein Pareto wäre schön)*

