



REGRESSION UND KORRELATION

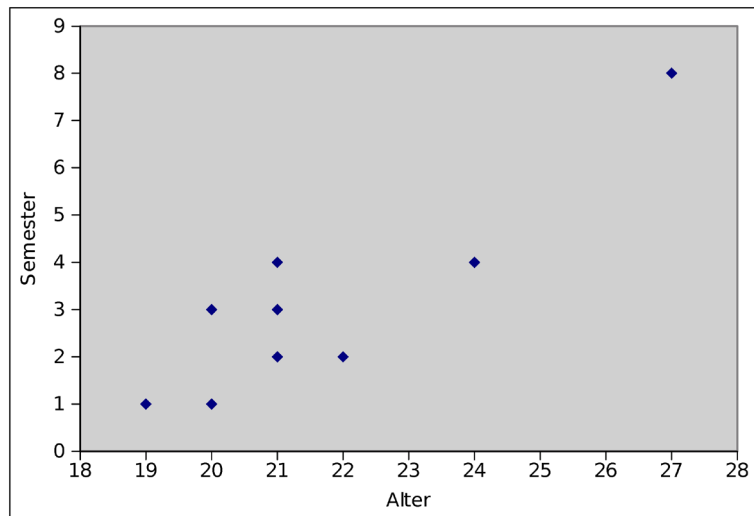
TORSTEN LINNEMANN

PUNKTWOLKEN

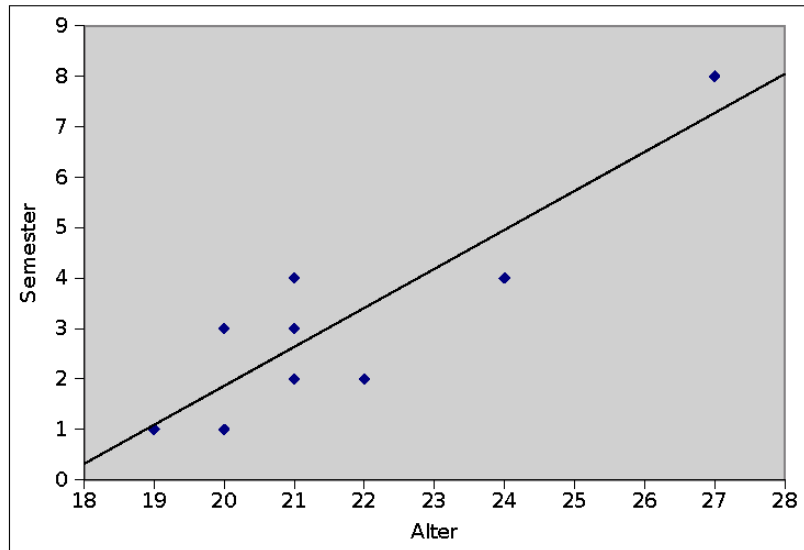
Beispiel 1.13

Wir beginnen mit einem kleinen Datensatz, in dem das Alter und die Semesterzahl einer Studentengruppe erhoben ist. Man vermutet: Je älter ein Student ist, desto höher ist sein Semester. Auch dass dieser Zusammenhang linear sein könnte, ist nicht unplausibel: Sieht man von Studienwechsel und Auslands-, Urlaubs- und Freisemestern ab, so erhöhen sich Alter und Semesterzahl proportional zueinander, sogar mit dem Proportionalitätsfaktor 1.

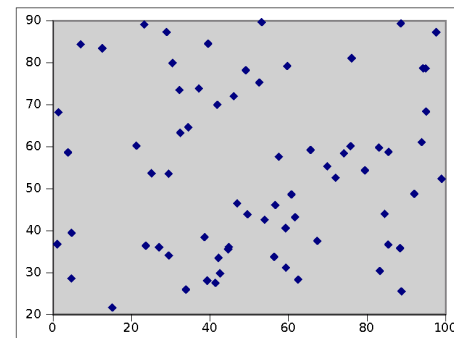
Alter	20	21	24	21	22	21	27	19	20	21
Semester	3	3	4	2	2	3	8	1	1	4



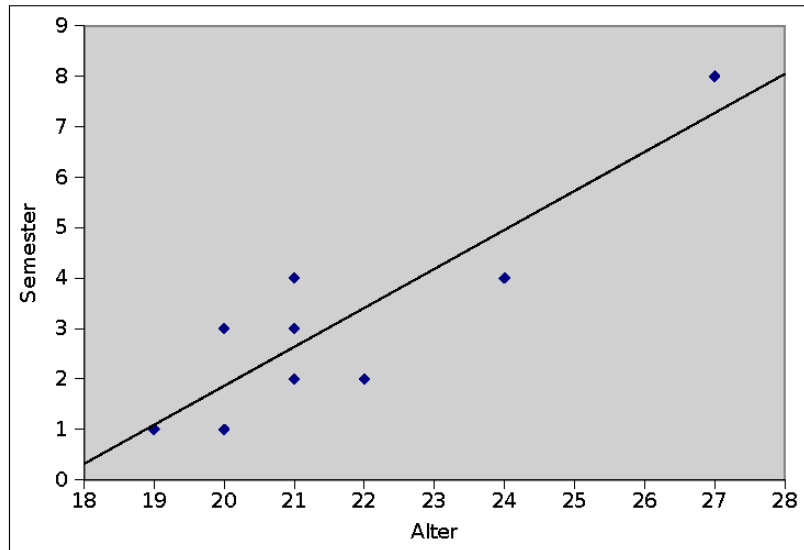
AUSGLEICHSGERADE MIT AUGENMASS



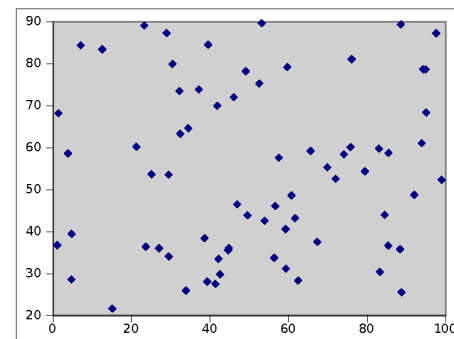
Keine sinnvolle
Ausgleichsgerade:



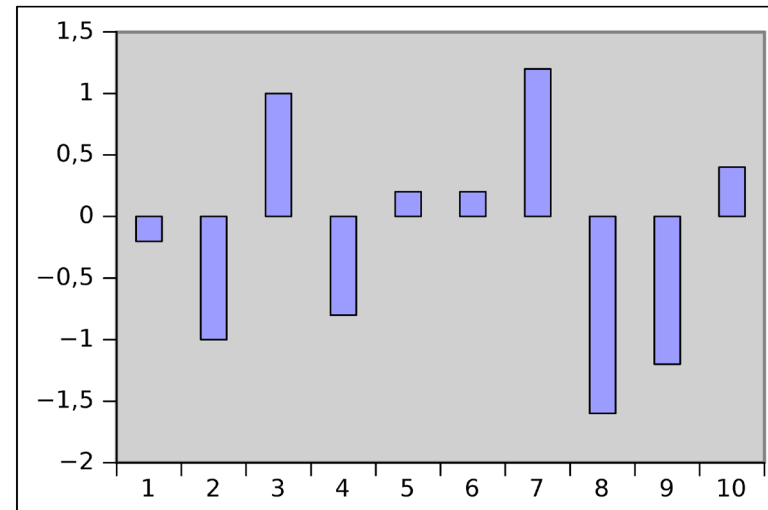
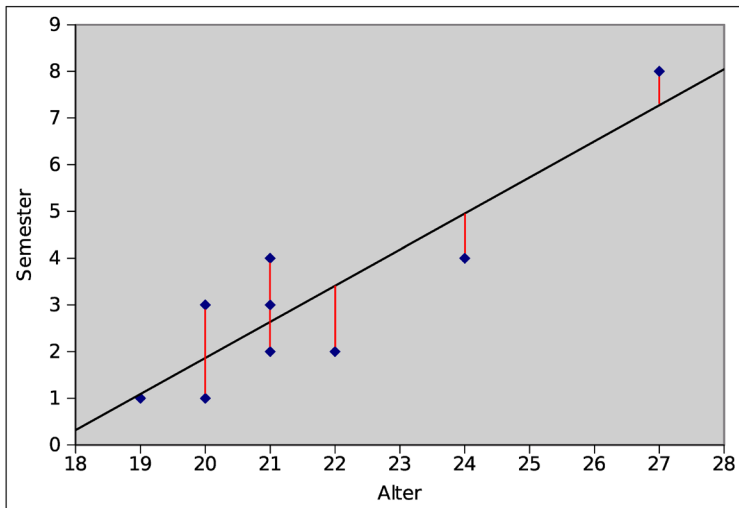
AUSGLEICHSGERADE MIT AUGENMASS



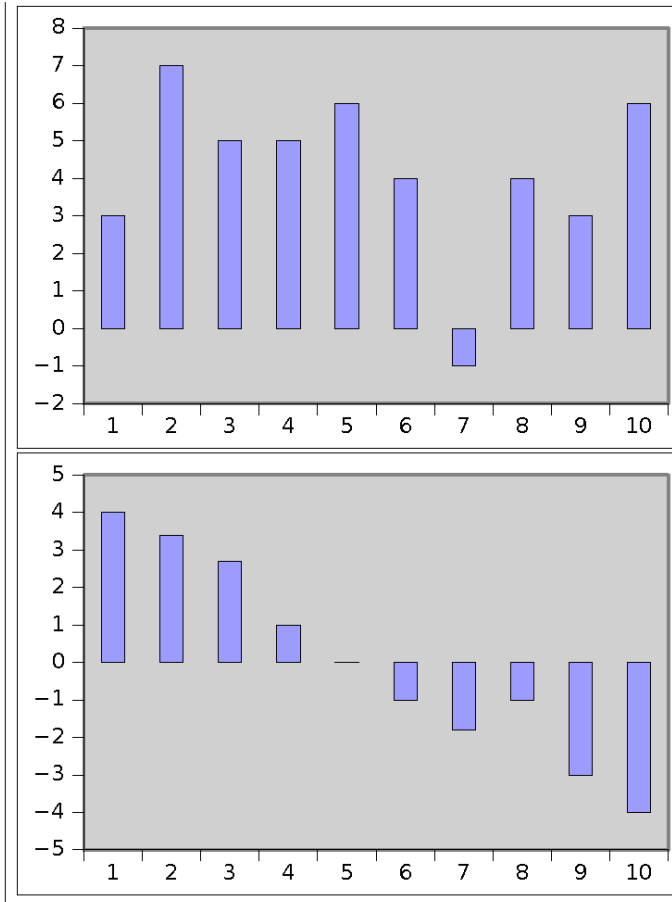
Keine sinnvolle
Ausgleichsgerade:



RESIDUEN – DIFFERENZEN VON DER AUSGLEICHSGERADEN



RESIDUEN – SYSTEMATISCHE FEHLER



REGRESSIONSGERADE

Als Regressionsgerade wird diejenige Gerade bezeichnet, bei der die Summe der Quadrate der Residuen besonders klein ist.

Die Berechnung dieser Regressionsgerade ist mit elementaren Mitteln möglich:

bei einem quadratischen Zusammenhang muss das Minimum (Scheitelpunkt) ermittelt werden.

Die genaue Formel verwenden wir hier nicht.

ARBEITSAUFTRAG

Öffnen Sie das Geogebra-Buch

«Regression April 2019»

<https://www.geogebra.org/m/mnap8zek>

Öffnen Sie die Datei «Trendlinie finden»

Passen Sie die Schieberegler so an, dass Ihre Ausgleichsgerade

$a(x)=mx+q$ möglichst gut ist.

KORRELATION

Einige Datensätze lassen sich recht gut mit Geraden annähern, beispielsweise die Alter \leftrightarrow Semesterzahl.

Andere weniger, wie wir bei der einen Punktwolke gesehen haben.

Um dies zu messen, gibt es den Korrelationskoeffizienten. Dieser berechnet die Abstände vom Mittelwert. Die Formel muss hier nicht verstanden werden.

Wichtig ist:

Korrelationskoeffizient 1: alle Punkte auf einer Geraden

Korrelationskoeffizient -1: alle Punkte auf einer Geraden mit negativer Steigung

Korrelationskoeffizient 0: kein Zusammenhang.

Korrelationen, die grösser als 0.7 sind, gelten als stark.

ARBEITSAUFTRAG

Bearbeiten Sie die Datei

Noten FMP

Wie ist der Zusammenhang zwischen den Mathematik- und den Französischnoten?

Wie ist der Zusammenhang zwischen den Mathematik- und den Physiknoten.

Verwenden Sie die Befehle

- Trendlinie
- Korrelationskoeffizient

ARBEITSAUFTRAG

Natürlich, manchmal sind die Zusammenhänge eher nicht linear.

Öffnen Sie die Datei korrelationen

- Ordnen Sie die Punkte näherungsweise auf einer Parabel an.
- Finden Sie eine Näherung mit einem Polynom der Ordnung 2 (eine quadratische Funktion)

- Ordnen Sie die Punkte auf einer Wellenlinie an. Verwenden Sie den Schieberegler Wellenlinie, um die bestpassende Sinusfunktion zu finden.