

Thema: Regression – Zusammenhänge erkunden

Thomas Müller

☒ TI-Nspire™ CAS

Schlagworte: Statistik, Regression

Unterrichtsmaterial

Aufgabe:

Im Airberlin-Magazin (3/2013, p51) finden sich technische Daten zur Luftflotte der Firma.

Damit die Daten nicht abgeschrieben werden müssen, gibt es die Angabe in der Datei „Flugzeug-Regression-Angabe.tns“

Hier findet man zu einigen Flugzeugtypen Daten wie Anzahl der Sitzplätze - Länge (m) - Spannweite (m) - Abflugmasse (t) - Reichweite (km) - Geschwindigkeit (km/h) - Verbrauch (Liter pro 100 PKM (Passagierkilometer))



Erläuterung zu PKM:

Darunter versteht man das Produkt der transportierten Personen bzw. Passagiere und der dabei zurückgelegten Entfernung in Kilometern.

Angenommen, ein Flugzeug hat 90 Passagiere für 4000 km an Bord. Dann sind das $90 \times 4000 = 360\,000$ PKM.

Aufgabenstellung:

- Ergänze in der Tabelle eine Spalte, in der der Gesamtverbrauch an Treibstoff bei voller Auslastung der Sitzplätze auf einem Flug von Wien nach London (rund 1300 km) ersichtlich ist. Berechne jeweils den Pro-Kopfverbrauch. Führe einen Vergleich mit einer 1300 km langen Autofahrt durch.
- Stelle die Abhängigkeit der Länge von der Sitzplatzzahl in einem Streudiagramm dar und stelle eine Vermutung auf, durch welche Funktion dieser Zusammenhang modelliert werden kann. Zeichne zunächst eine Passgerade und führe anschließend eine Regressionsrechnung durch.
- Untersuche weitere Zusammenhänge, wie die Abhängigkeit der Reichweite vom Verbrauch, die Abhängigkeit der Spannweite von der Länge und einen selbstgewählten Zusammenhang.

Didaktischer Kommentar

Diese Aufgabe entstammt dem Inhaltsbereich *Wahrscheinlichkeit und Statistik*.

Dabei sollen u.a. Werte aus tabellarischen Darstellungen abgelesen und entsprechend interpretiert werden. Statistische Kennzahlen sollen im jeweiligen Kontext interpretiert werden und einfache statistische Grafiken erstellt werden können.

In der vorliegenden Aufgabe können die Schüler und Schülerinnen selbstständig (ev. mit Hilfe der gegebenen Anleitung) die notwendigen statistischen Kennzahlen ermitteln und einzelne Visualisierungen gleichsam zur Kontrolle im Nachhinein oder als erste Orientierung im Vorhinein erstellen.

Vorschlag zur Umsetzung

Zu a)

Zur Berechnung des Verbrauchs auf der Strecke Wien-London gibt man in der Zelle j1 („wien-london“) die Formel ein $[(b1*1300)/(100))*h1]$ und kopiert sie anschließend in alle Zeilen.

	A	B	C	D	E	F	G	H	I	J	K	L
		plaetze	laenge	spannw	masse	reichw	geschw	verbra...	gesam...	wienlo...	jeperson	
1	Airbus 319	150	33.84	34.1	75.5	5560	858	4.3	645.	8385.	55.9	
2	Airbus 320-200	174	37.57	33.9	77.	5500	858	3.4	591.6	7690.8	44.2	
3	Airbus 321-200	220	44.51	34.1	93.	5700	870	3.1	682.	8866.	40.3	
4	Dash	76	32.83	28.42	29.57	2500	666	5.4	410.4	5335.2	70.2	
5	Boing 737-700	144	33.6	35.8	65.5	6100	840	4.7	676.8	8798.4	61.1	
6	Boing 737-800	186	39.5	35.8	78.	5420	840	3.1	576.6	7495.8	40.3	
7	Embraer	112	36.25	28.72	50.3	3700	860	5.5	616.	8008.	71.5	

Interessant ist der Prokopfverbrauch an Treibstoff, vgl. Spalte k („jeperson“)

Hier bietet sich ein Vergleich mit einer 1300 km langen Autofahrt an:

Annahme: 3 Personen und damit ein Verbrauch von 9,5 l je 100 Kilometer. $13 \times 9,5 = 123,5$ l Benzin. Je Person also insgesamt rund 41 l.

Ganz zwanglos ergeben sich nun Fragen über mögliche Zusammenhänge, z.B. ob der PKM-Verbrauch mit der Masse eines Flugzeuges zusammenhängt. Je größer die Masse, desto höher wird vermutlich der Treibstoffverbrauch sein.

Zu b)

Doch zunächst soll der vermutete Zusammenhang zwischen Flugzeuglänge und Zahl der Sitzplätze untersucht werden: Dazu öffnet man eine neue Seite und wählt *Data & Statistics*, dann fügt man durch Anklicken der Variablenschalters „Click to add variable“ unten und links (linke Maustaste!) die gewünschten zu vergleichenden Eingabedaten an.

Der Eindruck des Zusammenhangs hängt natürlich auch von den Zeomeinstellungen auf den Achsen ab.

Vergleiche dazu die Abbildungen rechts: Beide stellen den Zusammenhang zwischen der Anzahl der Sitzplätze und der Längen der Flugzeuge dar. Lediglich die Maßstäbe auf den Achsen sind verschieden.

Empfehlenswert ist folgende Einstellung auf den Achsen „plaetze“ [0, 280], „laenge“ [0, 60]

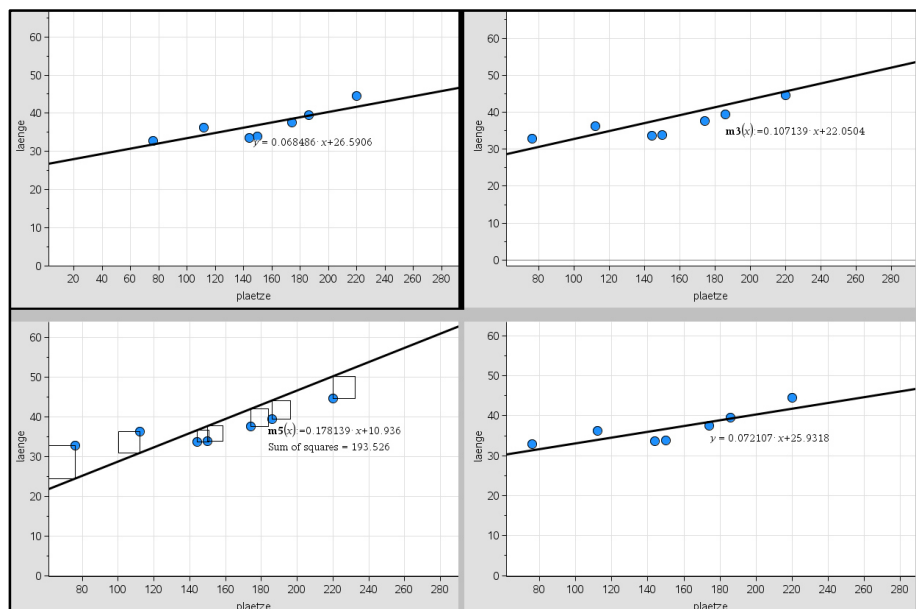
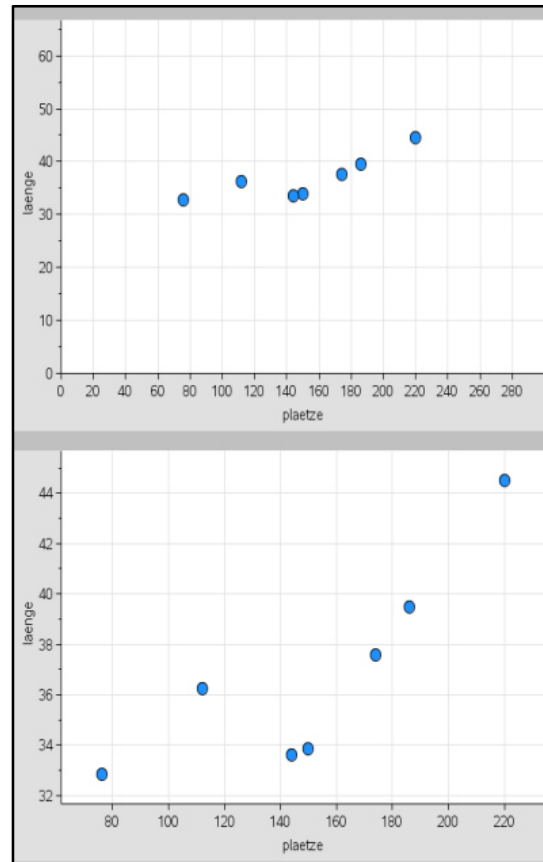
Die Punkte scheinen grob gesehen, entlang einer Geraden zu liegen.

Um dies tatsächlich nachprüfen zu können, bietet TI-Nspire im *Data & Statistics*-Menü unter *Analyze* das Einfügen einer beweglichen Geraden an: *Add Moveable Line*.

Die Gerade soll nun so bewegt werden, dass die *Summe der Abstände* zu den Datenpunkten möglichst gering ist. Je geringer diese ist, desto besser nähert die Gerade die „Punktwolke“ an. (Abstand = Differenz oder Rest = Residual)

Da Punkte unterhalb und oberhalb der Geraden liegen, ist eine einfache Addition der orientierten Abstände nicht sinnvoll, da einander negative und positive Werte zum Teil aufheben. Eine Abhilfe schaffen die Quadrate dieser Abstände. Diese lassen sich durch *Analyze >> Residuals >> Show Residual Squares* einblenden. Bei Bewegungen der Geraden wird die Summe der Quadrate der Abstände eingeblendet. Ziel ist es nun, die Gerade so zu positionieren, dass sich eine möglichst kleine Summe ergibt.

Die Gerade mit der kleinsten Quadratsumme heißt Regressionsgerade. Man kann sie direkt über *Analyze >> Regression >> Show Linear* erzeugen ($mx + b$ entspricht $y = kx + d$ in der „österreichischen“ Notation einer linearen Funktion).



CAS – Projekt T³ Österreich



Im *Lists & Spreadsheet* - Fenster erhält man die Gleichung über *Statistics >> Stat Calculations >> Linear Regression (mx + b)*

L	M
	=LinRegMx
Title	Linear Re..
RegEqn	m*x+b
m	0.072107
b	25.9318
r ²	0.696302
r	0.834447
Resid	{-2.9078...

Die Regressionsgerade hat in diesem Fall die Gleichung $y = 0,072x + 25,93$

Dabei entspricht x der Platzanzahl und y der Länge eines Flugzeuges.

Will man theoretisch vorberechnen, wie lange ein Flugzeug etwa sein würde, das 400 (= x) Personen Platz bieten sollte, dann könnte es eine Länge von 55(= y) Metern haben.

Eine Regression beschreibt eine Ursache-Wirkungs-Beziehung (Platzanzahl → Länge).

Der *Korrelationskoeffizient* r ($=0,83$) gibt die Güte dieses Zusammenhangs an. Je näher bei +1 (-1 für fallende Gerade) desto besser kann eine (steigende) Gerade an die Datenpunkte angepasst werden.

Zur c)

Als Beispiel für einen weiteren Zusammenhang sei das Datenpaar (Masse/Verbrauch) angeführt:

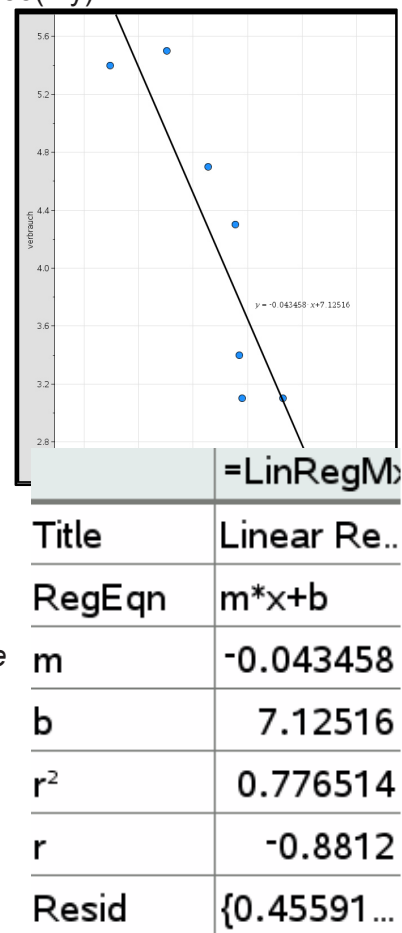
Hier ist der Korrelationskoeffizient beinahe -0,9. Es besteht also ein äußerst guter (negativer) linearer Zusammenhang: Je größer die Masse, desto geringer ist der PKM-Verbrauch, also: Große Flugzeuge fliegen wirtschaftlicher als kleine.

ACHTUNG: Besonders sollte hier thematisiert werden, dass dies trotz der Sprechweise „je größer ... desto geringer“ *keine* indirekte Proportionalität darstellt.

Dateinamen:

„Flugzeug-Regression-Angabe.tns“

„Flugzeug-Regression-Loesung.tns“



Technologiehilfe

Zoomeinstellungen auf den Achsen lassen sich entweder *direkt* mit gedrückter linker Maustaste oder mit Rechtsmausklick direkt auf die Achse verändern >> Zoom >> Windows Settings (Abbildung rechts).

