

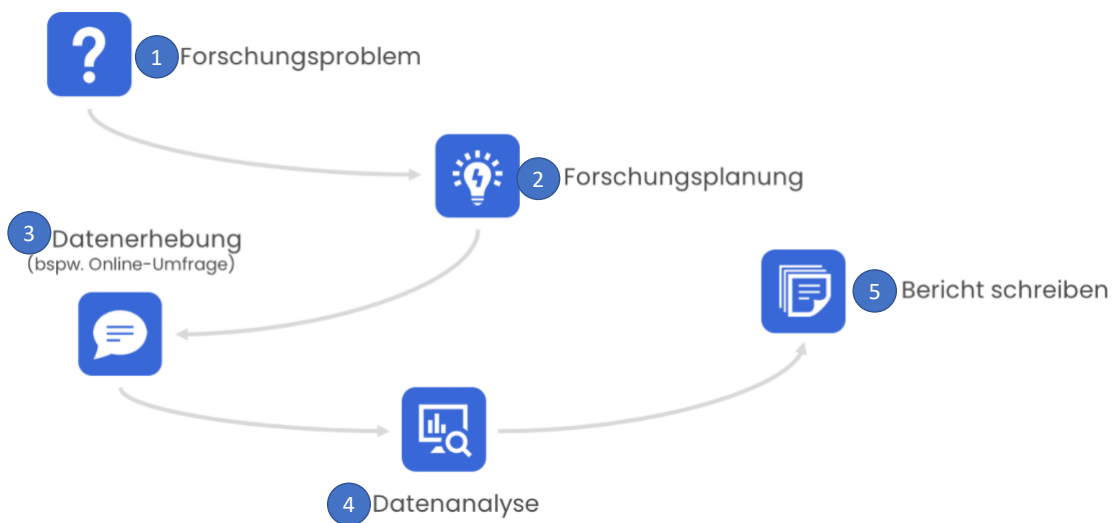
Empirische Sozialforschung

Skript 4

1

1

5 Hauptphasen einer empirischen Studie



2

2

5 Hauptphasen einer empirischen Studie:

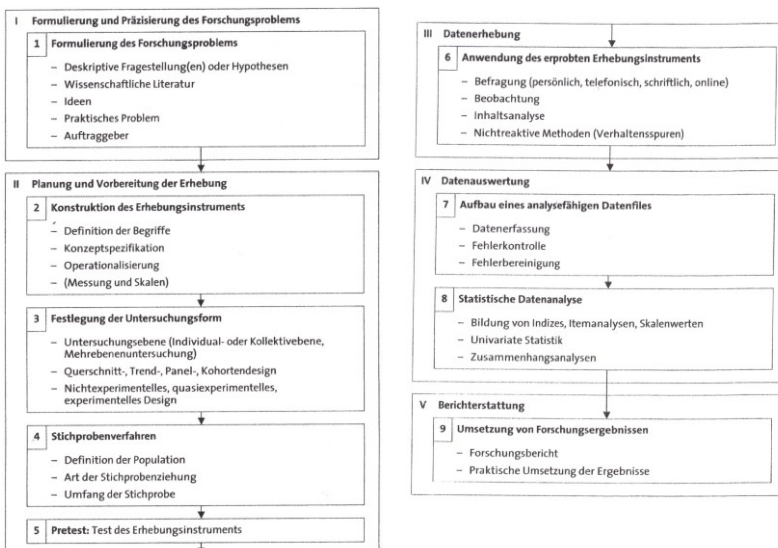
- I. Formulierung und Präzisierung des Forschungsproblems
- II. Planung und Vorbereitung der Erhebung
- III. Datenerhebung
- IV. Datenauswertung
- V. Berichterstattung

- Zu Beginn ist ein **Forschungsproblem** (*Entdeckungszusammenhang*) klar zu definieren. Die **Forschungshypothese** - abhängige und unabhängige Variablen - müssen präzise angegeben werden.
- Bei einer *deskriptiven Studie* sollte klar sein, bei welcher Population welche Merkmale geschätzt werden sollen. Am Anfang steht immer die Frage: **Was genau möchte ich wissen?**

3

3

5 Hauptphasen einer empirischen Studie:



Diekmann, A. (2014) Empirische Sozialforschung, Rowohlt Hamburg

4

4

I Formulierung und Präzisierung des Forschungsproblems

1. Formulierung des Forschungsproblems

Am Anfang deiner Forschung steht die Frage:

Was möchtest du untersuchen?

- Hier gehst du von einer groben Richtung in eine präzisere Forschungsfrage.
- Nachdem du dann den aktuellen Stand der Forschung untersucht hast (mit Hilfe von Literatur), bestimmst du das Forschungsdefizit. Das heißt: **Was wurde bisher noch nicht bzw. nicht ausreichend untersucht?**
- Daraus kannst du die Hypothesen deiner Arbeit formulieren.



5

5

II Forschungsplanung und Vorbereitung der Erhebung

2 Konstruktion des Erhebungsinstruments

- Basierend auf der Forschungsfrage wird nun die eigene Forschung geplant: Was ist die Untersuchungseinheit und wie soll die Erhebung durchgeführt werden?
- Die Auswahl der **Erhebungsmethode** ist dabei klar vom Ziel der Forschung abhängig: Sollen neue Aspekte erforscht oder sollen bestehende Theorien überprüft werden?
- **4 Methoden: Befragung** (persönlich, telefonisch, schriftlich), **Beobachtung, Inhaltsanalyse, Analyse von Verhaltensspuren**
- Wichtig ist, die *Problemformulierung zu operationalisieren*, d. h. einer Messung zugänglich zu machen. Bei komplexen mehrdimensionalen Begriffen (*wie Umweltbewusstsein, Betriebsklima etc.*) ist eine Konzeptspezifikation erforderlich.
- Festlegung des **Erhebungsinstruments** (Mess- und Skalierungsmethoden) Bei einer Befragung ist das der Fragebogen.

6

6

II Forschungsplanung und Vorbereitung der Erhebung

3 Festlegung der Untersuchungsform

- Das **Forschungsdesign** ist festzulegen. Hier stellt sich die Frage nach der **Untersuchungsebene**. Bei der **Kollektivhypothese** sind die Untersuchungseinheiten Kollektive. Bei der **Kontexthypothese** Individuen und Kollektive. Bei **Individualhypothesen** sind die Untersuchungseinheiten Individuen.
- Die zweite Entscheidung bezieht sich auf den **zeitlichen Aspekt** der Datenerhebung. Die wichtigsten Optionen sind **Querschnitt- und Längsschnitterhebungen**.
- Zu überlegen ist auch, ob eine **Vergleichs- oder Kontrollgruppe** berücksichtigt werden soll.

7

7

II Forschungsplanung und Vorbereitung der Erhebung

3 Festlegung der Untersuchungsform

Querschnittsstudie oder Längsschnittstudie?

- Die **Ergebnisse von Querschnittstudien bilden eine Momentaufnahme, während Längsschnittstudien soziale oder individuelle Wandlungsprozesse ausdrücken**. Auch Kombinationen von Längsschnitt- und Querschnittsdesigns sind möglich.
- Querschnittserhebungen ermöglichen es den Forschern, verschiedene Merkmale, wie die demografische Struktur der Verbraucher, ihre Interessen und Einstellungen, auf einmal zu untersuchen. Sie zielen darauf ab, Informationen über die Bevölkerung zum aktuellen Zeitpunkt zu liefern.
- Eine Längsschnittstudie (Verlaufsstudie, Longitudinalstudie, Längsschnittuntersuchung, Längsschnitterhebung) ermöglicht die Forschung zur Untersuchung sozialer und individueller Wandlungsprozesse in einer festgelegten Zeitspanne.

8

8

II Forschungsplanung und Vorbereitung der Erhebung

4 Stichprobenverfahren

- Bestimmung und Größe der Stichprobe. Haupttypen sind: **Zufallsstichprobe**, **Quotensample** und **willkürliche Stichprobe**. Die Wahl hängt von den Forschungszielen und auch vom Forschungsetat ab.
 - Bei einer **Zufallsstichprobe**, also einer Zufallsauswahl, hat jedes Element der Grundgesamtheit die gleiche Wahrscheinlichkeit in die Stichprobe zu gelangen. *Ein Beispiel für eine einfache Zufallsstichprobe ist das Ziehen einer Kugel aus einer Urne, wie zum Beispiel bei der wöchentlichen Ziehung der Lottozahlen.*
 - Bei der **willkürlichen Stichprobe** werden Merkmalsträger unkontrolliert in die Stichprobe aufgenommen. Sie werden berücksichtigt, weil sie einfach verfügbar sind oder weil sie sich freiwillig melden.

9

9

II Forschungsplanung und Vorbereitung der Erhebung

4 Stichprobenverfahren

- **Quoten Sample** - Auswahlverfahren in der Umfrageforschung, bei dem die Auswahl der zu befragenden Personen dem Interviewer überlassen wird. Der Interviewer hat sich bei der Auswahl nach bestimmten Merkmalen der Befragten zu richten, die in einem bestimmten Verhältnis (Quote) im Sample auftreten sollen.
- Die Quoten werden z. Beispiel so gewählt, dass die Zusammensetzung des Samples in Bezug auf die ausgewählten Quotenmerkmale der Zusammensetzung der Grundgesamtheit (z.B. Bevölkerung der Bundesrepublik) entspricht.

10

10

II Forschungsplanung und Vorbereitung der Erhebung

4 Stichprobenverfahren

- Als **Quotenmerkmale** werden in der Regel Alter, Geschlecht, Beruf, Konfession, Schulbildung gewählt.
- Die Güte des Q.-S.s ist bis heute heftig umstritten. Für das Q.-S. werden von der Markt- und Meinungsforschung seine Wirtschaftlichkeit angeführt. Zudem sollen sich durch das Q.-S. Durchschnitte und Anteile recht gut erfassen lassen, ebenso Variablen, die mit den Quotenmerkmalen eng korreliert sind.
- Gegen das Q.-S. spricht, dass sich der Auswahlfehler nicht berechnen lässt, dass die Objektivität der Auswahl durch den Interviewer nicht kontrolliert werden kann, dass weitergehende Aufgliederungen über die Quotenmerkmale hinaus unzuverlässig sind.

11

11

II Forschungsplanung und Vorbereitung der Erhebung

5. Pretest

- In jedem Fall sollte ein neu erarbeitetes Erhebungsinstrument – auch Fragebogen oder Beobachtungsschema einem **Pretest** unterzogen werden, um zu prüfen, ob es mehrdeutige oder schlecht verständliche Fragen enthält.
- Als Pretests werden Verfahren bezeichnet, mit denen Datenerhebungen vor ihrer eigentlichen Durchführung schon optimiert werden.
- Beispiel: Bevor die Untersuchung anläuft, sollte der Fragebogen empirisch mit Hilfe eines sogenannten Pretests überprüft werden. Dazu wird der Fragebogen an Testpersonen erprobt.
- Grundsätzlich bestehen drei verschiedene Verfahren: Pretestverfahren im Untersuchungs-Feld, kognitive Verfahren sowie andere Verfahren wie beispielsweise eine Fokusgruppen oder Expertenbefragungen. Für Fragebogen lassen sich die beiden Hauptkategorien qualitative und quantitative Pretests unterscheiden.

12

12

III Datenerhebung

6 Anwendung des erprobten Erhebungsinstruments

- Die Erhebung wird umgesetzt. Es werden Teilnehmer identifiziert und auf diese zugegangen, damit die Fragen zur Forschung platziert werden können.
- Am Ende dieser Phase hat der Forschende also eine Reihe an auswertbaren Ergebnissen für die wissenschaftliche Arbeit.

13

13

IV Datenauswertung

- Aus der Erhebung können jetzt die erzeugten Daten analysiert und aufbereitet werden.
- Während man *quantitative Ergebnisse* mit Hilfe von deskriptiver Statistik (Mittelwerte, Streuung, Korrelation, etc.) auswerten kann, geht es bei *qualitativen Ergebnissen* mehr darum, diese erst einmal zu strukturieren und dann über verschiedene Verfahren eine Analyse zu beginnen (bspw. Inhaltsanalyse).

14

14

IV Datenauswertung

7 Aufbau eines analysefähigen Datenfiles

Die Daten werden in eine analysefähige Form gebracht (PC , Tabelle etc.) und auf Fehler hin kontrolliert. Fehlerquellen werden bereinigt.

8 Statistische Datenanalyse

Der Umfang hängt von der Komplexität der Datenerhebung ab. Eine einfache deskriptive Datenerhebung erfordert relativ wenig Aufwand. Das Verfahren hängt vom gewählten Erhebungsinstrument (z. B. Befragung) ab.

15

15

V Berichterstattung

9 Umsetzung von Forschungsergebnissen

- **Abschlussarbeit schreiben und Ergebnisse präsentieren** - schlussendlich werden die Ergebnisse der Forschung dokumentiert und in einem „Aufsatz“ verfasst.
- Dazu werden die Forschungsergebnisse interpretiert und aus ihnen Schlussfolgerungen gezogen.
- Bei der Beschreibung gilt es gewisse Punkte zu beachten. So sind vor allem *Aufbau der Arbeit* und die *richtige Zitierweise* von Bedeutung.
- Letztendlich werden die Forschungsergebnisse in der Praxis umgesetzt und angewendet.

16

16

Auswertung der Ergebnisse der quantitativen und qualitativen Forschung

Ablauf der zusammenfassenden Inhaltsanalyse nach Mayring (eine Form der qualitativen Inhaltsanalyse):

1. Phase: Paraphrasierung: Einheitliche Sprachebene, Wiederholungen und Ausschmückungen löschen.
Bsp.: "Persönlicher Kontakt ist das Nonplusultra" und "Vertrauen kann man nur durch persönlichen Kontakt aufbauen"

2. Phase: Generalisierung: Die Paraphrasen müssen auf ein Abstraktionsniveau gebracht werden. Man generalisiert also die Paraphrasen so, dass sie allgemeiner werden.
Bsp.: "Persönlicher Kontakt ist unabdingbar" und "Vertrauen ist nur durch persönlichen Kontakt möglich"

3. Phase: Reduktion: Es werden nur Phrasen beibehalten, die wichtig sind. Doppelte oder nicht wichtige werden gelöscht. Und Paraphrasen, die gleich bzw. ähnlich sind, werden zusammengefasst.
Bsp.: "Persönlicher Kontakt ist wichtig, weil dann ein Vertrauen möglich ist"

Am Ende sind verschiedene Aussagen komprimiert als Kategoriensystem zusammengefasst. Jetzt wird überprüft, ob noch alle Aussagen aus der ersten Phase enthalten sind, ansonsten müsste der Prozess erneut starten. Die Zusammenfassung wird dann im weiteren Verlauf der Untersuchung genutzt, um diese zu interpretieren und Schlussfolgerungen für die Forschungsfrage schließen.

<https://www.empirio.de/empiriowissen/auswertungsmethoden-in-der-empirischen-forschung>

17

17

Auswertung der Ergebnisse der Forschung

Auswertungsmethoden quantitativer Forschung

Um quantitative Daten analysieren und auswerten zu können, gibt es zwei grundsätzliche Arten von Statistik: Deskriptive Statistik und Inferenzstatistik

In der **deskriptiven Statistik** werden Werte und Mengen der Ergebnisse "recht simpel" zusammengefasst. Es werden Häufigkeitsverteilungen in Diagrammen erstellt, Mittelwerte und Standardabweichungen berechnet sowie Korrelationen bestimmt.

Folgende Berechnungsmethoden sind entscheidend in diesem Bereich der Statistik:

- **Mittelwerte:** Arithmetischer Mittelwert, Median und Modalwert
- **Streuung:** Spannweite (Differenz von höchstem und niedrigstem Wert), Standardabweichung und Varianz, Quartilsabstand
- **Korrelation:** Der Korrelationskoeffizient gibt den Zusammenhang von zwei Merkmalen an und befindet sich zwischen -1 (negativer Zusammenhang) und +1 (positiver Zusammenhang). 0 würde hierbei bedeuten, dass kein statistischer Zusammenhang ermittelt werden konnte. Das heißt die untersuchten Merkmale haben keine nachgewiesene Abhängigkeit voneinander.

<https://www.empirio.de/empiriowissen/auswertungsmethoden-in-der-empirischen-forschung>

18

18

Auswertung der Ergebnisse der Forschung

Auswertungsmethoden quantitativer Forschung

In der **Interferenzstatistik** geht es darum, ob die Ergebnisse nur Zufall waren oder ob wissenschaftliche Gesetzmäßigkeiten vorliegen. Das wichtigste Instrument in der Interferenzstatistik ist der Hypothesentest:

- **Hypothesentest:** Der Hypothesentest wird dazu genutzt, um zu sagen, ob von der vorliegenden Stichprobe auf die Gesamtheit geschlossen werden kann. Man stellt dafür der aktuell geltenden Hypothese (Nullhypothese) eine Alternativhypothese (H1) zu Seite, um dann abschließend sagen zu können, ob diese zu bestätigen oder zu widerlegen ist. Am Ende der Untersuchung kann ein Forscher also sagen, dass mit einer bestimmten Irrtumswahrscheinlichkeit von X Prozent die Hypothese bestätigt oder widerlegt werden kann.

<https://www.empirio.de/empiriowissen/auswertungsmethoden-in-der-empirischen-forschung>

19

19

Auswertung der Ergebnisse der Forschung

Auswertungsmethoden qualitativer Forschungsmethoden

Die Herausforderung in der qualitativen Forschung ist es, die erhobene Daten in irgendeiner Form analysieren zu können, da diese bspw. durch Interviews erst noch dokumentiert und zusammengetragen werden müssen.

Daher ist folgender Ablauf bei qualitativen Forschungen zu durchlaufen:

1. Aufzeichnung der Daten:

Mit Hilfe von Mitschnitten oder Aufzeichnungen können die Interviews festgehalten werden

2. Aufbereitung (Transkription) der Daten:

Man schreibt je Interview in einem Text-Dokument Wort für Wort mit und dokumentiert so sehr präzise, was in den Aufzeichnungen zu hören ist. Hier geht es allerdings rein um die Inhalte des Befragten, die zum Thema wichtig sind. Dialekte, Sprechpausen und Füllwörter (Bspw. „ehm“) können weggelassen werden.

<https://www.empirio.de/empiriowissen/auswertungsmethoden-in-der-empirischen-forschung>

20

20

Auswertung der Ergebnisse der Forschung

Der **Modus** (auch **Modalwert** genannt) ist in einer Datenreihe das Merkmal bzw. der Wert mit der größten Häufigkeit. Es kann auch mehrere **Modi** geben, wenn zwei oder mehrere verschiedene Merkmale gleich oft am häufigsten vorkommen. Bsp.: Anna und Jonas machen eine Umfrage zur Gewalt an ihrer Schule. Sie befragen 85 Schüler der Schule nach der Gewaltart, die sie an ihrer Schule am häufigsten wahrnehmen.

Gewaltart	Bedrohung	Erpressung	Körperverletzung	Raub	Verspottung
Anzahl	13	17	28	19	8



<https://de.bettermarks.com/>

21

21

Auswertung der Ergebnisse der Forschung

Der **Median** (auch **Zentralwert** genannt) ist der Wert in der Mitte einer der Größe nach geordneten Datenreihe.

- Das heißt, mindestens 50% der Daten sind kleiner als der Median oder gleich dem Median und mindestens 50% der Daten sind größer als der Median oder gleich dem Median. Der Median ist somit unempfindlich gegenüber Extremwerten.
- Der Median wird verwendet für Daten, die in eine „natürliche“ Reihenfolge gebracht und mit Zahlenwerten versehen werden können.

Beispiele: Schulnoten (sehr gut, gut, ..., ungenügend), Antworten auf Befragungen (stark zustimmend, zustimmend, ..., stark ablehnend)

Bei einer **ungeraden Anzahl** an Datenwerten ist der Median der Wert in der Mitte.

Bei einer **geraden Anzahl** an Datenwerten entspricht der Median dem Durchschnitt der beiden mittleren Werte.

<https://de.bettermarks.com/>

22

22

Auswertung der Ergebnisse der Forschung

Der **Median** wird genutzt, um einen einzelnen Wert der Datenreihe qualitativ einzuordnen. Sophie möchte wissen, ob sie vergleichsweise viel oder weniger viel für die Mathematikarbeit gelernt hat. Sie fragt ihre Klassenkameraden, wie viele Stunden sie jeweils gelernt haben.

Name	Sophie	Laura	Tom	Anne	Marie	Julian	Alex
Stunden	7	5	16	6	7	4	6

Um den Median zu bestimmen, sortierst du die Werte der Größe nach.

Name	Julian	Laura	Anne	Alex	Marie	Sophie	Tom
Stunden	4	5	6	6	7	7	16

Da hier eine ungerade Anzahl an Daten vorliegt, kannst du den Median direkt ablesen. Es ist der Wert in der Mitte (6).

Sophie hat mit 7 Stunden also eher viel für die Mathematikarbeit gelernt.

<https://de.bettermarks.com/>

23

23

Auswertung der Ergebnisse der Forschung

Median: Paul Ulrich möchte wissen, ob seine Familie vergleichsweise viele oder wenige Kinder hat. Er zählt, wie viele Kinder die Familien in seiner Nachbarschaft haben.

Familie	Meyer	Kalesch	Eckard	Ulrich	Baar	Weber
Anzahl der Kinder	0	5	1	2	1	2

Um den Median zu bestimmen, sortierst du die Werte wieder der Größe nach

Familie	Meyer	Eckard	Baar	Weber	Ulrich	Kalesch
Anzahl der Kinder	0	1	1	2	2	5

Da hier eine gerade Anzahl an Daten vorliegt, kannst du den Median nicht direkt ablesen. Du berechnest den Durchschnitt der beiden Werte in der Mitte:

$$\frac{1+2}{2} = \frac{3}{2} = 1,5$$

Pauls Familie hat mit zwei Kindern also vergleichsweise viele Kinder.

<https://de.bettermarks.com/>

24

24

Auswertung der Ergebnisse der Forschung

Mittelwert:

Der Mittelwert eignet sich für quantitative Daten wie z.B. Körpergröße, Einkommen, Preise. Der Mittelwert ist empfindlich gegenüber Extremwerten.

Ausreißer innerhalb einer Datenreihe, "ziehen" den Mittelwert nach oben oder unten.

Den **Mittelwert** (auch **arithmetisches Mittel** genannt) berechnest du, indem du alle Beobachtungswerte x_1, x_2, \dots, x_n addierst und die Summe durch die Anzahl n der Beobachtungen dividierst:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Amelie macht in den Schulferien fünf Tage Urlaub bei ihrer Freundin.

Sie notiert sich jeden Tag, wie viel Taschengeld sie ausgegeben hat.

Tag	1	2	3	4	5
Ausgaben (in €)	24	15	20	28	12

Um die durchschnittlichen Ausgaben von Amelie zu bestimmen, berechnest du den Mittelwert:

$$\bar{x} = \frac{24 + 15 + 20 + 28 + 12}{5} = 19,80$$

Amelie hat also durchschnittlich **19.80€** am Tag ausgegeben.

<https://de.bettermarks.com/>

25

25

Auswertung der Ergebnisse der Forschung

Streuung:

Während Mittelwerte und Median Auskunft über die Mitte einer Verteilung geben, deuten die **Streuungsmaße** an, wie weit die Werte um die Mitte herum streuen.

- Je weiter die Werte sich vom Mittelwert entfernen, desto größer ist die Streuung und desto flacher die Verteilung.
- Je näher die Werte sich am Mittelwert befinden, desto kleiner ist die Streuung und desto schmalgipfliger die Verteilung.
- Das bekannteste Streuungsmaß ist die **Standardabweichung**

Zu ihrer Berechnung benötigt man zwei Werte: **Mittelwert** und **Varianz**.

Aus diesem Grund vollzieht sich die Berechnung in **drei Schritten**:

1. Berechnung des arithmetischen Mittels: $\bar{x}(A)$

2. Berechnung der Varianz: $s^2(A)$

3. Berechnung der Standardabweichung: $s(A)$

Die Standardabweichung sagt aus, wie weit (im Durchschnitt) alle gemessenen Ausprägungen eines Merkmals vom Durchschnitt entfernt sind. Die Standardabweichung lässt sich mit einer Statistiksoftware wie SPSS berechnen.

<https://de.bettermarks.com/>

26

26

Auswertung der Ergebnisse der Forschung

Die Varianz gibt an, wie sich die Beobachtungswerte um den Mittelwert aller Beobachtungen verteilen.

Sie gehört zu den Streuungsmaßen.

Beispiel

Wir haben 5 Personen gefragt, wie viele Stunden Sport sie pro Woche treiben.

Person	1	2	3	4	5
Stunden Sport/Woche	2	3	7	5	3

Mittelwert: $\frac{2+3+7+5+3}{5} = 4$

Varianz: $\frac{(2-4)^2+(3-4)^2+(7-4)^2+(5-4)^2+(3-4)^2}{5-1} = \frac{4+1+9+1+1}{4} = \frac{16}{4} = 4$

<https://www.scribbr.de/statistik/varianz/>

27

27

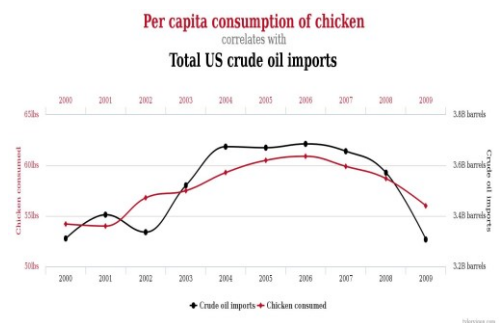
Auswertung der Ergebnisse der Forschung

Korrelation oder Kausalität

Kausalität bedeutet, dass eine Änderung einer Variable eine Änderung einer anderen Variable bewirkt. Das wird Ursache-Wirkungs-Prinzip genannt. **Korrelation** bedeutet, dass es einen statistischen Zusammenhang zwischen Variablen gibt. Kausalität würde bedeuten: *Der Pro-Kopf-Verzehr von Hühnchen beeinflusst die Rohölimporte!*

1. Je mehr Hühnchen man verzehrt,
2. desto mehr Gewicht nimmt man zu,
3. desto weniger möchte man sich sportlich betätigen,
4. desto mehr nutzt man das Auto oder den Bus,
5. desto höher wird die Benzinnachfrage und
6. desto mehr Rohöl muss importiert werden?

Aber stehen diese beiden Größen wirklich in einem kausalen Zusammenhang zueinander?



<https://www.uni-due.de/dataedu/korrelation-vs-kausalitat/>

28

28