



Methodencoach

Sie möchten ein Thema analysieren? Einen Bedarf für eine Dienstleistung ermitteln? Sie möchten ein Projekt bewerten oder einen Bericht schreiben? Häufig benötigen Sie dazu Statistiken oder Kennzahlen um eine These zu begründen. Der Methodencoach möchte Sie an die Themen „Statistik und empirische Sozialforschung“ heranführen und Ihnen praxistaugliche Leitlinien für die alltägliche Arbeit mit Daten und Statistiken an die Hand geben.

Die hier untenstehenden Punkte werden in diesem Dokument behandelt.

1. Was versteht man unter Statistik und Daten?	2
2. Was ist empirische Sozialforschung?	3
3. Einige Grundbegriffe der empirischen Sozialforschung	3
4. Wie komme ich an verlässliche Zahlen für die DG?	4
5. Prozessgenerierte Datensammlungen	5
6. Datenerhebung mittels Befragung	6
7. Online-Befragung	8
8. Weitere Datenerhebungstechniken	10
9. Wie führe ich eine Untersuchung sachgerecht durch?	11
10. Welche Untersuchungsformen gibt es?	14
11. Quantitative oder qualitative Forschungsansätze?	14
12. Wie baue ich eine einfache „Datenbank“ auf?	15
13. Statistik und Datenschutz?	16
14. Wie erstelle ich ein Formular oder einen Fragebogen? Einige Grundregeln.	17
15. Was ist bei der Formulierung von Fragen zu beachten?	18
16. Wie wird die Rücklaufquote bei postalischen Befragungen erhöht?	19
17. Wie werte ich die Ergebnisse einer Befragung aus?	20
18. Was ist Repräsentativität?	21
19. Vollerhebung oder Stichprobe?	22
20. Wie ziehe ich eine Stichprobe für eine Befragung? Die Auswahlverfahren.	23
21. Wie groß muss die Stichprobe sein?	24
22. Wo finde ich weitere praktische Informationen?	26



1. Was versteht man unter Statistik und Daten?

Statistik ist die Lehre von Methoden zur Analyse quantitativer Informationen (Zahlen und zählbare Informationen). Statistik ist also ein Oberbegriff für unterschiedliche Methoden der Datenanalyse (z.B. Mittelwertberechnungen, Kreuztabellierung oder Regressionsanalyse). In der Alltagssprache aber sind mit „Statistik“ oft aufbereitete Daten, also zusammenfassende Tabellen oder Diagramme, gemeint.

Als **Daten** versteht man logisch gruppierte Informationseinheiten. Sie sind das Ergebnis systematischer Erhebungen, etwa anhand von Befragungen, Beobachtungen, Interviews, Experimenten oder das Resultat von Verwaltungsvorgängen. Im letztgenannten Fall spricht man auch von „prozessgenerierten“ oder „prozessproduzierten Daten“. Diese Datenform ist in der Deutschsprachigen Gemeinschaft am häufigsten.

Daten für die Statistik sammeln und bearbeiten

Sie sollten darauf achten, dass Sie zunächst möglichst unbearbeitete Informationen sammeln. Diese „Daten“ werden in einem zweiten Schritt nach Bedarf aufbereitet.

Im Idealfall verfügen Sie für jeden Merkmalsträger - also für jede Person, jede Gebietseinheit, jeden Antrag oder jede Akte - über die relevanten Angaben in zunächst möglichst unbearbeiteter Form. Die Angaben sollten noch nicht zusammengefasst (aggregiert) sein.

Bevorzugen Sie zum Beispiel Datensätze, in denen zur Akte XYZ das Geburtsjahr und nicht etwa die Zugehörigkeit zu einer Altersklasse notiert wurde. So können Sie die Altersklassen später beliebig berechnen und Mittelwerte ausweisen. Auch haben Sie z.B. mehr Auswertungsoptionen, wenn Sie für eine Gemeinde die Flächengröße und Einwohnerzahl besitzen, als wenn Ihnen nur das „Produkt“ in Form der Einwohnerdichte vorliegt.

Daten speichern

Speichern Sie die „Datensätze“ z.B. in einem Tabellenblatt einer Excel-Datei. Benutzen Sie dabei für jeden Merkmalsträger (jede Akte, jede Person oder jeden Antrag,...), am besten eine Zeile. Teilen Sie jedem Merkmalsträger - und somit jeder Zeile - eine Identifikationsnummer zu.

Die Auswertungen dieser Datensätze, z.B. zusammenfassende Häufigkeitstabellen oder Diagramme, speichern Sie dann in anderen Tabellen- bzw. Diagrammblättern ab, ohne dabei die Rohwerte zu „zerstören“. Kommen später bei der Interpretation der Tabellen und Diagramme weitere Fragen auf, ermöglicht diese Vorgehensweise zusätzliche Auswertungen.

Tipp: Wenn Sie Ihre Daten in Excel bearbeiten, sollten Sie sich unbedingt das Arbeiten mit Pivot-Tabellen (auch Kreuztabellen genannt) aneignen. Mit Pivot-Tabellen erstellen Sie in Sekunden die unterschiedlichsten Häufigkeitstabellen. Auch die Berechnung von Prozenten und Mittelwerten geschieht automatisch. Einen kostenlosen Online-Kurs dazu finden Sie zum Beispiel auf <http://www.teialehrbuch.de/>



2. Was ist empirische Sozialforschung?

Empirische Sozialforschung ist die wissenschaftliche Untersuchung menschlichen Verhaltens und gesellschaftlicher Phänomene auf der Grundlage von verschiedenen Techniken und Methoden.

Sie dient vor allem der systematischen Prüfung von wissenschaftlichen Theorien oder Hypothesen. Sie findet Anwendung in allen Sozial- oder Gesellschaftswissenschaften (Humangeografie, Soziologie, Psychologie, Politologie, Wirtschaftswissenschaften, Demografie, Kommunikationswissenschaften, Ethnologie,...), aber auch in der Biologie und Medizin.

Die Instrumente und Ergebnisse der empirischen Sozialforschung sind jedoch nicht nur für die Wissenschaft von Nutzen. Tagtäglich begegnen wir den Ergebnissen empirischer Sozialforschung: Zahlen und Fakten, fundierte Schlussfolgerungen oder Trendanalysen zum Beispiel werden heute in allen gesellschaftlichen Bereichen verlangt.

Die Methoden der empirischen Sozialforschung werden z.B. für Folgendes genutzt: Zur Bewertung von Leistungen oder Maßnahmen, zu Strategie-, Budget- oder Marketingplanungen, für die Raum- und Regionalplanung, zur Bedarfsermittlung, zur Erfassung der Kundenzufriedenheit, zur Bewertung von Angebot und Nachfrage oder zur Ermittlung von Nutzerprofilen.

Tipp: Auch im Rahmen praxisorientierter Analysen ist die Auswahl der Erhebungs- und Auswertungsmethode von großer Bedeutung, denn je angreifbarer die Methode, desto anfechtbarer das Ergebnis. Verschaffen Sie sich daher einen Überblick über die Methoden der empirischen Sozialforschung!

3. Einige Grundbegriffe der empirischen Sozialforschung

Empirisch (griechisch für „Sinneserfahrung“): bedeutet durch Beobachtung, Befragung/Interview, Experiment oder durch die Sammlung sogenannter prozessgenerierter Daten und deren Auswertung gewonnenes Wissen. Empirische Sozialforschung bezeichnet die systematische Erhebung und Interpretation von Daten über soziale Tatsachen („erfahrbare“ Welt).

Merkmalsträger: Als Merkmalsträger wird in der Statistik eine Untersuchungseinheit bezeichnet, die bestimmte, für die Untersuchung relevante Ausprägungen von Merkmalen annehmen kann. Beispiel: Ein Individuum kann beispielsweise nach Alter, Berufsstatus, Einkommen usw. charakterisiert werden.

Grundgesamtheit: Menge aller statistischen Einheiten (auch Merkmalsträger), über die Aussagen getroffen werden sollen. Beispiel: alle Personen (statistische Einheiten), die am 1. Januar 2010 (zeitliche Identifikation) mit ihrem Hauptwohnsitz (sachliche Identifikation) in der DG (örtliche Identifikation) gemeldet sind.

Variable: Ein Merkmal, das verschiedene Ausprägungen annehmen kann. Häufig abgefragte soziodemographische Variablen sind z.B. Angaben wie Alter, Geschlecht, Familienstand, Wohngemeinde und Berufsstatus.



Ausprägung: Damit bezeichnet man die verschiedenen Werte, die eine Variable annehmen kann. Die Variable „Geschlecht“ kann z.B. die Ausprägungen „männlich“ oder „weiblich“ annehmen.

Theorie: Im strengen Sinne ist eine Theorie ein System logischer und widerspruchsfreier Aussagen (Hypothesen, Sätze) über einen Untersuchungsgegenstand.

Hypothesen: Aussagen, die einen Zusammenhang zwischen mindestens zwei Variablen postulieren. Es gibt „Wenn-Dann-Aussagen“ und „Je-Desto-Aussagen“. Beispielhypothese: je höher Variable 1 „Alter der Person“, desto höher Variable 2 „Anzahl grauer Haare“.

Typologie: Das Ziel einer Typologie oder Klassifikation besteht darin, die Untersuchungseinheiten nach theoretischen Gesichtspunkten Typen oder Klassen zuzuordnen. Beispiel: „Vollerwerbsbetrieb, Zuerwerbsbetrieb, Nebenerwerbsbetrieb“ als Typen landwirtschaftlicher Betriebe

Indikator: Ein Indikator ist ein beobachtbarer Sachverhalt, mit dem ein nicht beobachtbares (latentes) Konstrukt gemessen werden kann. Beispiel: Die Höhe des sozialen Engagements ist ein nicht beobachtbares Konstrukt. Es kann z.B. durch die ehrenamtlich geleisteten Stunden im Rahmen von Vereinsarbeit gemessen werden.

Index: Ein Index ist die Zusammenfassung von mehreren Einzelindikatoren zu einer neuen Variable. Beispiele: Entwicklungsindex einer Region, Verbraucherpreisindex, Alterungsindex,...

Tipp: Mit Indikatoren und Indexen lassen sich Zusammenhänge, Entwicklungen und Vergleiche besser darstellen und begreifen. Recherchieren Sie auf internationalen Portalen nach den **Referenzindikatoren** für das jeweilige Themengebiet! Diese Indikatoren werden Ihrer Analyse mehr Gewicht verleihen. Scheuen Sie jedoch auch nicht davor zurück, Ihre eigenen Indikatoren und Indexe zu entwickeln.

4. Wie komme ich an verlässliche Zahlen für die DG?

Grundsätzlich kann man zwischen „primären“ und „sekundären“ Daten unterscheiden. Primäre Daten sind solche, die selbst erhoben werden müssen. Häufig kann jedoch auf sekundäre Daten, d.h. bereits erfasste Daten, zurückgegriffen werden. Hierbei handelt es sich oftmals um sogenannte „prozessgenerierte“ Daten, die bei einem Verwaltungsvorgang als Nebenprodukt entstehen. Nicht selten werden beide Quellen genutzt.

Sekundäre Daten für die DG

Sekundäranalysen sind vergleichsweise kostengünstig und zeitsparend. Nachteil: Die Daten sind nicht immer optimal auf die Fragestellung angepasst und häufig nicht auf dem aktuellsten Stand. Zudem kann die Erhebungssituation nicht kontrolliert werden.

Wenn Sie sekundäre Daten recherchiert haben, können Sie diese neu auswerten. Rechnen Sie z.B. Prozente und Entwicklungen aus und setzen Sie verschiedene



Zahlenreihen zueinander ins Verhältnis (X pro Person / X anteilig an der Bevölkerung, X anteilig an der Gruppe Y).

Wo finden Sie die Daten? Zum Beispiel auf DGstat oder über die Statistik-Links dieser Website. Die veröffentlichten Angaben sind meist nur die Spitze des Eisberges. Fragen Sie einfach bei potenziellen „Datenproduzenten“ nach! Insbesondere föderale Ämter erfassen eine Vielzahl von Angaben, etwa über Verwaltungsprozesse in Zusammenhang mit Leistungen der sozialen Sicherheit.

Keine DG-Daten? Suchen Sie nach Statistiken zu vergleichbaren Regionen!

Die Deutschsprachige Gemeinschaft ist keine „von der Außenwelt abgeschottete Insel“. Um allgemeine Entwicklungen und Trends aufzuspüren, ist oft ein Blick auf Studien und Statistiken zu anderen Gebieten hilfreich.

Die Angaben aus anderen Regionen können Sie auch nutzen, um eine Größenordnung abzuschätzen. Ein Beispiel: Sie wissen, wie viele Personen im Bezirk Verviers im Jahr 2008 die Krankheit X gehabt haben? Schätzen Sie die Größenordnung für die DG ein, indem Sie zum Beispiel 28% des Wertes von Verviers nehmen; 28% ist der Bevölkerungsanteil der DG im Bezirk Verviers im Jahre 2008. Hierdurch erhalten Sie zumindest Näherungswerte.

Tipp: Verwenden Sie bevorzugt Angaben zu Gebieten mit ähnlichen Strukturmerkmalen. Das Ergebnis einer Arbeitsmarktanalyse über den Bezirk Arlon zum Beispiel – einem ähnlich wie die DG ländlich geprägten Gebiet mit relativ niedriger Arbeitslosenrate und vielen Grenzpendlern – gibt eher Hinweise über mögliche Tendenzen in der DG als eine Studie zum völlig anders strukturierten Großraum Brüssel.

Primäre Daten zur DG

Sie können sich nicht auf Sekundärdaten berufen? Sie möchten ganz spezifische Angaben erheben? Dann bietet sich die eigene Datenerhebung an. Auch hier stehen Ihnen viele Wege offen:

- **Der Aufbau einer eigenen prozessgenerierten Datenbank**
- **Die Befragung**
- **Besondere Erhebungswege: Beobachtung, Inhaltsanalyse,...**

Lesen Sie in der Folge mehr darüber.

5. Prozessgenerierte Datensammlungen

Prozessgenerierte Daten sind Daten, die als Ergebnis eines Verwaltungsvorganges im weitesten Sinne entstehen (z.B. Anmeldung des Wohnsitzes, Beantragung von Zuschüssen oder Zulassungen, eintragungspflichtige Ankäufe von Immobilien, Vorgänge in Krankenkassen und Familienkassen,...).

Die überwiegende Zahl der Angaben zur Deutschsprachigen Gemeinschaft ist auf prozessgenerierte Datensammlungen zurückzuführen. Die meisten Daten stammen dabei von föderalen Einrichtungen. Einiges wird auch in regionalen oder gemeinschaftlichen



Einrichtungen erhoben. Prozessgenerierte Daten können aber auch über die Angaben von Teilnahmeformularen oder die Erfassung eines Eintrittskartenverkaufs entstehen.

Vor- und Nachteile von prozessgenerierten Daten

Prozessgenerierte Daten werden selten eigens erhoben, um sozioökonomische Fragestellungen zu beantworten. Daher aufgepasst: Wenn Sie Sekundärdaten erhalten, sollten Sie sich immer die Mühe machen, die jeweiligen technischen Beschreibungen und Erläuterungen zu studieren. Denn nicht selten verbirgt sich hinter einem scheinbar klaren Begriff ein unerwarteter oder sogar widersprüchlicher Inhalt, der nur aus verwaltungstechnischer Sicht Sinn macht. Bleiben Fragen offen, zögern Sie nicht, sich direkt beim „Produzenten“ der Daten zu erkundigen.

Ein weiterer Nachteil prozessgenerierter Daten: So wie sich die Aufgaben und Zuständigkeiten des „Produzenten“ verändern, so verändern sich auch die Definitionen zu den Merkmalen. Infolgedessen steigen oder sinken Werte – scheinbar ohne Grund. Man spricht dann von „Brüchen in der Zeitreihe“.

Trotz alledem: Prozessgenerierte Statistiken bieten wertvolle Informationen. Sie sind meistens günstig, gut dokumentiert und transparent. Zudem sind sie häufig die einzige verfügbare Quelle, vor allem, wenn ein längerer Zeitraum beobachtet werden soll.

Bauen Sie Ihre eigene prozessgenerierte Datensammlung auf!

Diese Option steht Ihnen natürlich nur offen, wenn Sie selbst in irgendeiner Weise Zugang zu den Merkmalsträgern (Personen, Organisationen, Akten, Teilnehmer, Kunden,...) haben.

Dann aber ist der Aufbau einer eigenen Datensammlung ein guter und meist kostengünstiger Weg, um an Daten heranzukommen. Auch wenn eine gewisse Anfangsinvestition nötig ist: Wenn Sie sich einmal ein System aufgebaut haben, können Sie davon auf lange Sicht profitieren. Setzen Sie sich bei der Erstellung von Formularen oder bei der Planung eines Verwaltungsvorgangs einfach mal die Brille des „Sozialwissenschaftlers“ auf: Welche Angaben könnten später von Interesse sein? Wie sollten diese Angaben abgefragt werden und in welcher Häufigkeit? Wie können Sie diese Angaben so zentralisieren, dass sie hinterher einfach zur statistischen Analyse benutzt werden können?

Im Grunde ist auch die Abfrage von Daten im Zuge eines Verwaltungsvorgangs eine Befragung. Eine konkrete Vorgehensweise wird in den folgenden Punkten beschrieben.

6. Datenerhebung mittels Befragung

Befragungen in der DG sind vergleichsweise selten und finden meistens punktuell statt. Besonders internationale Einrichtungen wie die EU oder die OECD greifen auf Befragungen zur Datengewinnung zurück. Der Grund: Prozessgenerierte Daten unterschiedlicher Länder sind selten vergleichbar. Bei diesen internationalen Befragungen handelt es sich meistens um Stichprobenerhebungen. Diese Stichproben sind in der Regel so ausgerechnet, dass sie für ein ganzes Land oder eine größere Region ein



aussagekräftiges Ergebnis liefern. Daher ergeben sie meist kein detailliertes oder aussagekräftiges Ergebnis für eine kleine regionale Einheit wie die Deutschsprachige Gemeinschaft.

Auf der Website der Generaldirektion Statistik des Föderalministeriums für Wirtschaft können Sie mehr über die wichtigsten Befragungen in Belgien erfahren.

Eine Befragung vorbereiten

Bevor Sie mit der konkreten Planung der Befragung beginnen, sollten Sie sich genau überlegen, was Sie erreichen möchten: Was ist das Anliegen der Untersuchung? Welche Annahmen oder Hypothesen sind aus Ihrer Sicht „überprüfenswert“? Wo stehen qualitative, wo quantitative Aspekte der Fragestellung im Vordergrund? In welcher Form benötigen Sie die Antwort, um sie auswerten oder vergleichen zu können?

Sie wissen exakt, was Sie in Erfahrung bringen möchten? Dann sollten Sie sich Gedanken darüber machen, wer Ihnen diese Angaben liefern kann. Grenzen Sie die Gruppe, über die Sie etwas in Erfahrung bringen möchten, möglichst genau ein. Überlegen Sie, wie Sie ganz konkret an die Informationen heran kommen können.

Des Weiteren bleibt zu entscheiden, welche Untersuchungsform Sie nutzen (Kohorten, Längsschnittanalyse,...), ob Sie Einzelbefragungen oder Gruppenbefragungen durchführen, ob und wie Sie eine Stichprobe ziehen, ob Sie eher offene und wann geschlossene Fragen nutzen, wer die Auswertung übernimmt, usw.

Dabei wird Ihre Vorgehensweise in der Regel maßgeblich von den vorhandenen Rahmenbedingungen bestimmt. Wenn Sie zum Beispiel wenig Zeit und Geld haben, werden Sie eher die Gruppen befragen, zu denen Sie leichten Zugang haben (Antragsteller, Mitglieder, Kunden,...).

Unterschiedliche Befragungsformen

Wenn Sie sich für eine Befragung entscheiden, müssen Sie eine Befragungsform auswählen. Die häufigsten Befragungsformen sind:

- persönliche Befragung (Face to Face)
- Telefoninterview
- Postalische Befragung
- Online-Befragung

Für welche Befragungsform Sie sich entscheiden, hängt zum einen von der Erreichbarkeit der Grundgesamtheit – also der zu befragenden Personen/Einrichtungen – ab. Zudem kennzeichnen sich die verschiedenen Befragungsformen durch unterschiedliche Vor- und Nachteile.



Vor- und Nachteile verschiedener Befragungsformen (ergänzt nach: Scharf/Schubert 1995, S. 369)				
	schriftlich	mündlich	telefonisch	online
Antwortquote	-	+	+	0
Einheitlicher Erhebungsstichtag	-	+	+	+
Ausschluss unüberlegter Antworten	0	+	-	0
Einfluss von Dritten	-	+	0	0
Umfang des Fragebogens	-	+	-	-
Gefahr von Missverständnissen	-	+	0	-
Interviewereinfluss	+	-	0	+
Erreichbarkeit	+	-	+	+
Räumliche Repräsentation	+	-	+	+
Kosten	+	-	+	+

Legende: + : Vorteil; - : Nachteil; 0 : indifferent

Standardisierung der Befragung

Des Weiteren sollten Sie sich über das Ausmaß der Standardisierung Ihrer Befragung im Klaren sein. Wie standardisiert oder wie offen sollte die Befragungssituation sein? Wenn Sie ein Thema erst einmal „auskundschaften“ und möglichst viele Aspekte in Erfahrung bringen möchten, dann eignen sich weniger standardisierte Methoden. Steht die Frage nach dem „wie viel“ aber im Mittelpunkt oder müssen die Antworten vergleichbar sein, dann sind standardisierte Befragungen zu bevorzugen.

Befragung	Mündlich	Schriftlich
Standardisiert	(Telefon-)Interview mit Fragebogen	Postalische Befragung, Online-Befragung, Gruppen-Befragung mit Fragebogen
Teilstandardisiert	Leitfadengespräche (Gruppe oder Einzel)	Experten- oder Zielgruppenbefragung
Nichtstandardisiert	Experteninterview, Gruppendiskussion, narratives Interview	Informelle Umfrage bei Experten oder Zielgruppen

7. Online-Befragung

Eine Online-Befragung ist eine Befragungsmethode, bei der im Gegensatz zu einem gedruckten Fragebogen ein Online-Fragebogen im Internet ausgefüllt wird. Online-Untersuchungen erfreuen sich in den letzten Jahren einer stetig wachsenden Beliebtheit, da immer mehr Personen per E-Mail erreichbar sind und so zu Befragungen eingeladen werden können. Zudem bietet eine Online-Untersuchung eine Reihe von Vorteilen:



- Der zeitliche Gewinn bei der Erhebung, Auswertung und Präsentation der Daten.
- Die Kosten für den Druck, den Versand von Fragebögen sowie der Interviewer entfallen.
- Die Möglichkeit der Einbindung von interaktiven und audiovisuellen Elementen in die Befragung.
- Der Einfluss des Interviewers auf das Antwortverhalten des Befragten entfällt.
- Spezielle Interessengruppen oder auch geografisch verstreute Zielgruppen sind leichter zu erreichen.
- Eine zumeist hohe Akzeptanz aufgrund von Freiwilligkeit, Flexibilität und Anonymität, und damit einhergehend eine hohe Datenqualität.

Diesen Vorteilen stehen eine Reihe von Nachteilen gegenüber:

- Für die Gesamtbevölkerung ist keine Repräsentativität erreichbar.
- Eine Mehrfachteilnahme von Befragten ist technisch nur bedingt kontrollierbar.
- Die Programmierung der Online-Untersuchung braucht einen gewissen zeitlichen Vorlauf und technische Vorkenntnisse.
- Nicht alle Zielgruppen sind online und einzelne Befragte können technische Schwierigkeiten haben, wenn diese veraltete Software und ältere Anzeigegeräte nutzen.

Worauf müssen Sie bei einer Online-Befragung achten?

Grundsätzlich sollte man die Anweisungen in einer Online-Befragung kurz und maximal verständlich halten, gleiches gilt auch für die **Startseite**. Hier müssen die Teilnehmer über die zentralen Aspekte der Befragung informiert werden:

- Um welches Thema geht es? Welches Ziel verfolgt die Befragung?
- Wer ist verantwortlich für die Befragung?
- Wie lange wird die Befragung dauern?

Zudem sollte auf der Startseite ein kurzer Hinweis zur Anonymität und Datenschutz sowie ein Link zu den Ansprechpartnern der Studie gegeben werden. Falls vorhanden kann man an dieser Stelle auch auf eine Vergütung der Teilnahme oder eine Verlosung hinweisen.

Zu Beginn der **Befragung** werden die demografischen Variablen erhoben. Zum einen ermöglicht Ihnen ein solches Vorgehen die Abbrecher mit den Nicht-Abbrechern statistisch zu vergleichen und zu analysieren und zum anderen wird so die Motivation eines Teilnehmers überprüft. Zudem sollten die ersten Befragungsseiten so gestaltet sein, dass sich auf ihnen deutlich mehr Texte oder Informationen zu finden sind als auf den späteren Befragungsseiten.

Offene Fragen sollten Sie nur in einem begrenzten Umfang verwenden oder zumindest eher ans Ende der Befragung stellen. Dieser Fragetyp senkt erfahrungsgemäß die Teilnahmemotivation und führt zu höheren Abbruchquoten.

Generell sollten Sie in einer Online-Befragung versuchen eine hohe Benutzbarkeit zu erreichen. Neben verständlichen Items und Instruktionen sollte die Befragung leicht zu bedienen sein und möglichst auf jedem Computer funktionieren, ohne die Installation von zusätzlicher Software. Außerdem sollten Sie darauf achten, dass die Befragung selbst in



etwa eine Bildschirmbreite von 800 Pixeln hat, so dass möglichst wenig gescrollt werden muss.

Weitere nützliche Tools bei einer Online-Befragung sind ein Fortschrittsbalken, mit dem die Befragten über den jeweiligen Bearbeitungsstand informiert sind, sowie eine Filterführung, bei der die Teilnehmer Fragen, die nicht für sie relevant sind, erst gar nicht angezeigt bekommen.

Zum **Ende der Umfrage** sollten Sie den Teilnehmern die Möglichkeit bieten Anmerkungen zu machen, ihnen für ihre Teilnahme danken und noch einmal den Ansprechpartner nennen, falls noch weitere Rückfragen auftauchen oder andere Kontaktwünsche herrschen.

Wie ziehen Sie Ihre Stichprobe bei einer Online-Befragung?

Bei der Online-Befragung wird zwischen der aktiven und passiven Rekrutierung unterschieden.

Durch entsprechend aussagekräftig formulierte Einladungen in E-Mails wird bei einer **aktiven Stichprobenziehung** versucht, Personen zur Teilnahme zu bewegen. Entweder steht Ihnen hierfür eine vollständige *Adressliste* der Zielgruppe zur Verfügung (Mitarbeiterbefragung) oder Sie beziehen sich auf ein so genanntes *Online-Panel*, welches Ihnen durch ein Feldinstitut zur Verfügung gestellt wird. Hierbei handelt es sich um eine große Sammlung von E-Mail-Adressen von Personen, die prinzipiell bereit sind, an Untersuchungen teilzunehmen. Eine dritte Möglichkeit der aktiven Rekrutierung ist die *Schneeballtechnik*, bei der Sie Einladungs-mails an die Personen der Zielgruppe versenden, deren E-Mail-Adressen Ihnen bekannt sind, mit der Bitte diese Einladung an andere Personen weiterzuleiten.

Bei einer **passiven Auswahl** der Befragten wird entweder ein Hinweis auf eine Website gesetzt, ein Hinweis in anderen Online-Befragungen oder Foren gemacht oder aber eine Offline-Ansprache der Zielpersonen durch Aushänge oder Verteilen von Handzetteln und Flyern gewählt. Die passive Rekrutierung hat eines gemeinsam: Die Entscheidung zur Teilnahme liegt komplett beim Empfänger der jeweiligen Nachricht und die Nicht-Teilnehmer können nicht an die Befragung erinnert werden. Ein weiterer Unterschied zwischen aktiver und passiver Auswahl besteht darin, dass Sie die Stichprobe bei einer aktiven Ansprache quotieren um so eine repräsentative Schichtung der Befragten zu erreichen.

8. Weitere Datenerhebungstechniken

Neben den verschiedenen Befragungsformen gibt es weitere Datenerhebungstechniken:

Die **„dokumentierte Beobachtung“** (im Gegensatz zur „normalen“, alltäglichen Beobachtung) ist eine systematische Methode mit dem Ziel, zumeist quantitative Daten zu liefern. Das Handeln von Personen oder bestimmten Zuständen wird erfasst. Es kann dabei einerseits zwischen „direkter“ und „indirekter“ Beobachtung und andererseits zwischen „teilnehmender“ und „nicht teilnehmender“ Beobachtung unterschieden werden. Feldstudien und Kartierungen gehören ebenso wie Zählungen in diesen Zusammenhang.



Für die **Inhaltsanalyse** werden Dokumente, Akten, Printmedien, Urkunden, Karten usw. zur Datengewinnung benutzt.

Das **Experiment** versucht, den Einfluss eines Stimulus auf mindestens zwei Zielgruppen (Versuchs- und Kontrollgruppe) und innerhalb kontrollierter Bedingungen (z.B. im Labor) zu prüfen. Dabei muss die Aufteilung der Personen auf die zwei Zielgruppen zufällig sein. Der Einfluss des Stimulus wird vorher und nachher in beiden Gruppen gemessen. Experimente sind zum Beispiel in der Psychologie gängig.

Beim **Quasi-Experiment** wird wie beim Experiment der Einfluss eines Stimulus auf eine Versuchs- und Kontrollgruppe geprüft. Die Zielgruppe wird aber nicht zufällig ausgewählt, sondern ergibt sich aus der Situation heraus (Feldexperiment).

Ein Beispiel: In der Schule X wird die Schuluniform eingeführt, in der Schule Y nicht. Hypothese: Die Schuluniform hat einen Einfluss auf die Zufriedenheit der Schüler. Zur Überprüfung der Hypothese wird in beiden Schulen die Zufriedenheit einen Monat vor und ein Jahr nach der Einführung gemessen. Anschließend werden beide Schulen verglichen.

9. Wie führe ich eine Untersuchung sachgerecht durch?

Sie möchten eine Fragestellung untersuchen, wissen aber nicht, wie Sie vorgehen sollen? Hier stellen wir Ihnen eine idealtypische Vorgehensweise vor, die Ihnen helfen kann, Ihre Untersuchung zu strukturieren. Beachten Sie dabei immer, welche Ressourcen (Personal, Zeit, Finanzen) Ihnen zur Verfügung stehen und passen Sie den „Forschungsprozess“ an.

Schritt 1: Klärung des Entdeckungs- und Verwertungszusammenhangs

Klären Sie, welches Problem untersucht werden soll. Für wen und mit welchem Ziel? Wie müssen die Informationen aussehen, um den jeweiligen Zweck zu erfüllen?

Schritt 2: Präzisierung der Fragestellung und Identifikation von relevanten Dimensionen

Was genau soll untersucht werden? Eine Fragestellung sollte relevant, begrenzt, präzise formuliert und vor allem erforschbar sein. Zudem müssen Sie sich fragen, welche Dimension wichtig ist. Mal liegt der Schwerpunkt auf der Beschreibung, mal gilt es, eine Hypothese zu testen und mal steht der Bezug zu einem Projekt im Vordergrund.

Suchen Sie präzise Informationen zu einer Situation zum Zeitpunkt X, zu einer Meinung oder zu einer Entwicklung (Diagnose)? Möchten Sie zunächst das Themenfeld erkunden (Exploration)? Möchten Sie einen kausalen Zusammenhang oder eine Hypothese prüfen? Zielt Ihre Fragestellung darauf ab, eine Prognose zu erstellen? Oder möchten Sie eine konkrete Maßnahme bewerten (Evaluation)?

Schritt 3: Zuordnung von geeigneten Begriffen zu den relevanten Dimensionen

Existieren schon geeignete Begriffe und Definitionen, um eine klare Fragestellung zu formulieren? Oder müssen diese noch eingeführt werden? Definieren Sie die zentralen Begriffe Ihrer Fragestellung.



Schritt 4: Einordnung der Problemstellung in vorhandene Kenntnisse, ggf. Hypothesenbildung und Entscheidung über das Forschungsdesign

Welche theoretischen Kenntnisse sind vorhanden? Sind genügend Vorkenntnisse vorhanden oder müssen Sie das Themenfeld vorab weiter kennenlernen, um die Problemstellung mit der erforderlichen Genauigkeit formulieren zu können?

Falls Sie eine Hypothese testen: Welche Hypothesen wurden bereits getestet? Welche Aspekte sind noch offen? Welche Untersuchungsform ist angemessen und welche Methode der Informationsgewinnung ist sinnvoll und möglich?

Schritt 5: Auswahl von Indikatoren (falls erforderlich)

Nicht immer sind die Dinge, die untersucht werden sollen, direkt zu beobachten (z.B. Innovationskraft von Unternehmen). Daher ist es notwendig, Indikatoren auszuwählen, mit denen dann wiederum Indexe oder Typologien gebildet werden können (z.B. Anzahl Patentanmeldungen, Ausgaben für Forschung und Entwicklung). Die Indikatorentwicklung ist ein mitunter schwieriges Unterfangen.

Schritt 6: Festlegung des erforderlichen Differenzierungsgrades der Informationen sowie Angabe der Messinstrumente

Wie kann die Ausprägung der Variablen erfasst werden? Welche Erhebungsinstrumente und Skalen eignen sich? Sind die verwendeten Instrumente zuverlässig? Mit welchen Messverfahren kann gearbeitet werden, und wie muss die Situation sein, in der die Messung geschieht?

Schritt 7: Auswahl der „Merkmalsträger“, der Grundgesamtheit und ggf. Entscheidung über Art und Umfang einer Stichprobe

Wer sind die Merkmalsträger (Personen, Organisationen, Zeitschriften,...)? Sollen alle untersucht werden oder nur einige? Wie kann eine Auswahl getroffen werden?

Schritt 8: Erhebung und Aufbereitung der Daten

Zunächst stellt sich die Frage, ob eine Primärerhebung nötig ist oder ob bestehende Informationen ausreichen (Sekundärauswertung). Ist eine Primärerhebung nötig, so sind unter Berücksichtigung der vorhandenen Ressourcen die entsprechenden Datenerhebungsmethoden auszuwählen.

Häufig unterschätzt wird der Arbeitsschritt der Datenaufbereitung. Bevor die Daten für die Auswertung bereit stehen, müssen sie bereinigt, selektiert, umkodiert, neu berechnet oder übersetzt werden.

Schritt 9: Auswertung und Verdichtung der Information

Im Unterschied zur Interpretation geht es bei der (statistischen) Auswertung des Materials zunächst darum, die Daten lesbar, also verständlich zu machen. Hierdurch erst werden aus Daten „Statistiken“. Dies geschieht zum Beispiel anhand von zusammenfassenden Tabellen und Häufigkeitsdiagrammen, durch das Berechnen von Prozentsätzen oder Indexen oder über die Anwendung von statistischen Modellen.

Im Wesentlichen ist bei der statistischen Auswertung zwischen der deskriptiven (beschreibenden) Statistik und der (schließenden) „Inferenzstatistik“ zu unterscheiden.

Letztere zielt darauf ab, mit Hilfe von Wahrscheinlichkeitsrechnungen eine Aussage zu einer begrenzten Zahl von Merkmalsträgern auf der Grundlage einer Stichprobe auf eine größere Gesamtheit zu verallgemeinern. Wichtigste Bedingung: Die Stichprobe muss nach dem Zufallsprinzip ausgewählt werden und jedes Element der Grundgesamtheit muss die gleiche Chance haben, ausgewählt zu werden. Nur wenn diese Bedingung erfüllt ist, ist die sogenannte Inferenzstatistik möglich. In bestimmten Fällen kann jedoch auch eine bewusste Auswahl der Stichprobe zielführend sein.

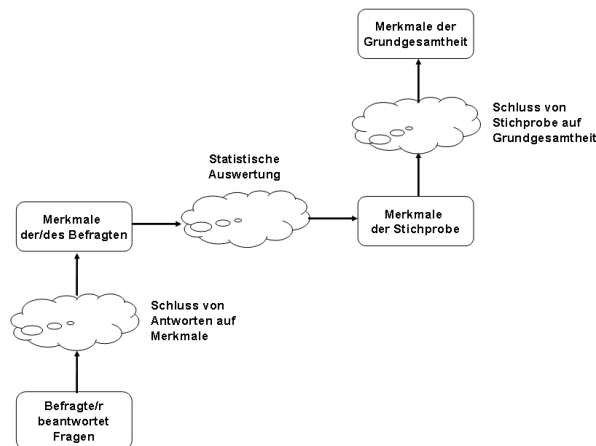


Abbildung: Forschungsablauf; nach Groves et al. 2004, S. 39.

Schritt 10: Interpretation der Ergebnisse

Bei der Interpretation der Daten wird das Material wieder in den inhaltlichen Kontext gebracht. Der Bezug zur Fragestellung steht im Vordergrund. Was sagen die statistischen Ergebnisse aus und wie sind sie zu erklären? Und waren die angewendeten Methoden und Verfahrensweisen angemessen, um die Fragestellung zu erörtern und möglichst zu beantworten? Welche Konsequenzen bzw. Handlungsempfehlungen ergeben sich daraus?

Schritt 11: Dokumentation des Forschungsprozesses und der Ergebnisse

Für eine seriöse Untersuchung muss nicht nur das Ergebnis sondern auch der Weg dahin dokumentiert werden. Bei umfangreichen Untersuchungen sollte ein entsprechendes Kapitel Hinweise zur Methode und zu den benutzten Instrumenten geben. Bei kürzeren Analysen sind methodische Anmerkungen und Fußnoten ein Zeichen für die Qualität der Untersuchung. Die Angabe der Quelle ist für die Zuverlässigkeit bzw. Nachvollziehbarkeit einer Information wesentlich. Aussagen aus der Literatur sollten Sie nicht unkommentiert bzw. unreflektiert darstellen. Den Stellenwert der Aussage sollten Sie zum einen mit dem Autor belegen und zum anderen mit einer Bewertung der Aussage, z.B. „MUSTERMANN (2008) führt als Argument an, dass ...“. Hierin ist zunächst die Quelle ersichtlich und außerdem die Angabe, dass es sich um die Meinung des zitierten Autors handelt.



10. Welche Untersuchungsformen gibt es?

Um eine Fragestellung zu untersuchen, gibt es unzählige Vorgehensweisen. Neben der Auswahl einer Datenerhebungstechnik ist die Suche nach der geeigneten Untersuchungsform entscheidend. Die Untersuchungsform legt fest, wann, wo und wie oft die Daten zu erheben sind. Nicht selten werden dazu verschiedene Untersuchungsformen kombiniert. Hier die meistgenutzten Formen:

Querschnittuntersuchungen: Dabei wird die empirische Untersuchung (z.B. Befragung, Inhaltsanalyse) **nur einmal** durchgeführt. So entstehen gesellschaftliche „Momentaufnahmen“ von derzeit gültigen Fakten, Meinungen oder Verhaltensweisen. Beim Querschnittsdesign verwendet man Zufallsstichproben, um die Repräsentativität zu gewährleisten.

Längsschnittuntersuchungen: Längsschnittuntersuchungen sind Untersuchungen, bei denen zu einer Grundgesamtheit zu **mindestens zwei Zeitpunkten** Daten erhoben werden. Im Vergleich zum Experiment oder zum Quasi-Experiment fehlt die Kontrollgruppe. Um dieses Manko auszugleichen, muss darauf geachtet werden, genügend Fälle zu untersuchen. Es gibt **verschiedene Formen von Längsschnittuntersuchungen:**

- Beim **Panel** wird jede Erhebung an denselben Untersuchungsobjekten zu verschiedenen Zeitpunkten mit demselben Erhebungsinstrument durchgeführt („**Personenidentische Mehrfachbefragung**“).
- Die **Kohortenstudie** ist eine spezielle Form der Paneluntersuchung, bei der alle Personen einer Stichprobe derselben Kohorte angehören. Eine Kohorte ist eine Gruppe von Personen, in deren Lebensläufen ein bestimmtes biographisches Ereignis annähernd zum selben Zeitpunkt aufgetreten ist. Beispiele: Geburtskohorten, Einschulungskohorten, Scheidungskohorten und viele andere mehr.
- Eine **Trendstudie** ist eine Wiederholungsbefragung, gelegentlich auch **replikativer Survey** genannt. Anders als bei Panelstudien werden hier nicht immer die gleichen Personen befragt, sondern nur wiederholt Stichproben aus der gleichen Grundgesamtheit gezogen. Durch Wiederholung der Befragung kann jede Querschnittbefragung in eine Trendstudie umgewandelt werden. Trendstudien sind also genaue „Nachbauten“ vormals durchgeführter Studien. Der große Vorteil von Trendstudien ist ökonomischer Art: Sie sind wesentlich günstiger als Panelstudien.

11. Quantitative oder qualitative Forschungsansätze?

Quantitative Forschungsansätze möchten Sachverhalte in Form von Modellen, Zusammenhängen und zahlenmäßigen Ausprägungen möglichst genau beschreiben und vorhersagbar machen. Dazu werden mittels festgelegter Datenerhebungsinstrumente Informationen vorrangig in Form von numerischen („harten“) Daten gewonnen. Die Datenerhebung wird dazu standardisiert (z.B. Antwortvorgaben bei Befragungen). Die quantitativen Forschungsansätze sind immer dann geeignet, wenn die Frage nach einem statistischen Zusammenhang oder nach dem „wie viel“ im Mittelpunkt steht. Die



Auswertung geschieht mit normierten, mathematisch-statistischen Verfahren. Eine Repräsentativität ist durch Zufallsstichproben und große Stichproben (samples) möglich.

Die **qualitativen Forschungsansätze** hingegen versuchen, eine möglichst breite Informationssammlung aus möglichst vielfältigen Perspektiven zusammenzutragen. Es geht darum, Zusammenhänge zu beschreiben, zu interpretieren und zu verstehen, Klassifikationen oder Typologien aufzustellen und Hypothesen zu entwickeln. Dabei erfolgt die Datenerhebung nicht oder kaum standardisiert (z.B. ausführliche Antwortmöglichkeiten bei Interviews). Die Frage nach dem „wie viel“ steht dabei nicht im Mittelpunkt. Die Auswertung der „weichen“ Daten geschieht durch sogenannte interpretativ-verstehende Verfahren. Es ist keine Repräsentativität im statistischen Sinne zu erreichen.

In der Praxis ist die strikte Trennung zwischen qualitativem und quantitativem Arbeiten in der Regel nicht aufrecht zu erhalten. Beide Ansätze ergänzen sich. Qualitative Forschung steht zeitlich meist vor der standardisierten Forschung, um ein Thema zu erkunden und den genauen Informationsbedarf zu ermitteln, also als Vorstudie, zur Exploration als Basis für eine groß angelegte quantitative Untersuchung. Zudem ist qualitative Forschung immer dann sinnvoll, wenn die Zahl der zu beobachtenden Merkmalsträger gering, die Zielgruppe sehr heterogen und die Komplexität des Themas besonders hoch ist.

Achtung:

Auch „standardisierte Forschungsansätze“ beinhalten „qualitative“ Informationen (Handlungsmotive, Meinungen zu Sachverhalt X). Die Datenerhebung ist jedoch festgelegter, „objektiver“ und vergleichbarer.

Andererseits benutzen die qualitativen Forschungsansätze im Gegenzug vielfach quantifizierende Methoden (z.B. Wortzählungen bei Inhaltsanalysen). Dabei ist die Vorgehensweise jedoch offener und statistisch vergleichbare Ergebnisse werden nicht produziert.

12. Wie baue ich eine einfache „Datenbank“ auf?

Wenn Sie über Daten verfügen, wollen Sie diese in der Regel sinnvoll ablegen und nutzbar machen. Wenn die Daten zunächst nur auf Papier (Fragebögen, Anträge,...) oder in einer Vielzahl von unterschiedlichen Dateien vorliegen, müssen Sie die „Datensätze“ in einer Datenbank zentralisieren. Denken Sie schon bei der Erstellung des Formulars, des Fragebogens oder des Antrags an diesen Schritt! Mit speziellen Softwareprogrammen zur Gestaltung von Formularen oder Umfragen lässt sich dieser Schritt sogar automatisieren (Beispiel: „Adobe PDF Formular-Maker“,...).

Wenn Sie nicht über eine Datenbank-Software verfügen, und wenn die Datensammlung nicht zu komplex und zu groß ist, dann bietet sich die Bearbeitung mit einer Tabellenkalkulation (z.B. Excel) als Datenbankersatz an. Wenn Sie beispielsweise jedes Jahr 1.500 Datensätze (Einschreibungen, Personen, Anträge, Fragebögen,...) und zu jedem Datensatz 20-30 Variablen (Alter, Geschlecht, Name, Aktennummer,...) speichern, ist dies mit einer Tabellenkalkulation durchaus zu meistern. Excel erlaubt maximal 65.500 Zeilen und 256 Spalten.



Wann verwenden Sie die Tabellenkalkulation (z.B. Excel,...) als Datenbank?

- Wenn Sie mit relativ kleinen Datenmengen arbeiten (bis maximal 65.500).
- Wenn nur wenige Anwender auf diese Datenbank zugreifen müssen, und dies nicht gleichzeitig.
- Wenn Sie „einfach“ strukturierte Daten besitzen (etwa 20-30 Variablen).
- Wenn Sie Daten aus anderen Datenbanken (beispielsweise Access, dBase) einlesen möchten und mit den Excel-Funktionen analysieren möchten.
- Wenn sie überwiegend deskriptive Statistik erstellen (Häufigkeiten, Prozente, Entwicklungen,...)

Wie verwenden Sie die Tabellenkalkulation als Datenbank? Einige Grundprinzipien:

- In die erste Zeile schreiben Sie die Titel der Merkmale (auch Variablen genannt). Typische Merkmale sind zum Beispiel die Identifikationsnummer, Kurzbeschreibung, Geschlecht, PLZ, Straße, Geburtsdatum, Land, Eintragsdatum, usw.
- Für jeden Datensatz benutzen Sie eine neue Zeile.
- Geben Sie jedem Datensatz eine Identifikationsnummer, die nur einmal vergeben wird.
- Achten sie darauf, in jeder Spalte nur ein Merkmal zu speichern. Speichern Sie zum Beispiel nicht die ganze Adresse in einer Zelle, sondern Postleitzahl, Ort, Land und Straße jeweils in getrennten Spalten.

Tipp: Weitere Erläuterungen dazu finden Sie unter anderem auf der Website <http://www.teialehrbuch.de>. Der kostenlose und frei verfügbare Kurs „Excel Professionals 2003“ beschreibt in Kapitel 3, wie Excel als Datenbank genutzt werden kann.

13. Statistik und Datenschutz?

Ob und wie sehr Sie Ihre Daten schützen müssen oder auch möchten, hängt selbstverständlich von Ihren Daten ab und ist mit ihrem Datenschutzbeauftragten abzuklären.

Im Allgemeinen gilt jedoch, dass Sie personenbezogene Daten unzugänglich aufbewahren sollten. Mit einer Datenschutz-Erklärung signalisieren Sie den Datenlieferanten zusätzlich, dass Sie personenbezogene Daten nicht an Dritte weiterleiten und vertraulich behandeln.

Möchten Sie die Daten zur Erstellung von Statistiken nutzen? Das ist meistens gar kein Problem, wenn Sie die Daten vorher anonymisieren. Sie erstellen zu diesem Zweck eine neue Tabelle, in der Sie alle personenbezogenen Angaben (Namen, Adressen, Firmenbezeichnungen, usw.) entfernen.

Andere Merkmale (Berufsbezeichnungen, Tätigkeitsbeschreibungen,...), die möglicherweise auf eine bestimmte Person oder Organisation schließen lassen, sind durch allgemeine Kodierungen zu ersetzen oder in Kategorien zusammenzufassen. So wird zum Beispiel aus dem Blumenkübel-Schweißer eine „Fachkraft im Metallsektor“, oder aus dem Weinfachhandel ein Einzelhändler im Lebensmittelsektor.



Wie sehr Sie verallgemeinern, hängt von der Struktur Ihrer Daten ab. Die Grundregel dabei lautet, dass Sie die Daten so zusammenfassen und kategorisieren, dass nur mit „unverhältnismäßigem Aufwand“ der Bezug zu einer bestimmten Person hergestellt werden kann.

Um Statistiken innerbelgisch aber auch über die Grenzen hinaus vergleichen zu können, sollten Sie vorhandene Nomenklaturen nutzen. International anerkannte Kodierungen und Klassifizierungen finden Sie auf dem Klassifikationsserver „RAMON“ von Eurostat: <http://ec.europa.eu/eurostat/ramon> Darüber hinaus sind auf der Website des Föderalen Öffentlichen Dienstes Wirtschaft - Generaldirektion Statistik, spezifisch belgische Nomenklaturen verfügbar.

14. Wie erstelle ich ein Formular oder einen Fragebogen? Einige Grundregeln.

Bevor Sie mit der Erstellung des Fragebogens beginnen, beachten Sie bitte einige Grundregeln:

- Je präziser Sie ein Thema untersuchen möchten, desto enger muss die Fragestellung abgesteckt sein. Ein Fragebogen kann nicht alles leisten!
- Planen Sie genügend Zeit zur Erstellung des Fragebogens ein! Unklare Formulierungen führen zu nichtssagenden Ergebnissen. Testen Sie daher den Fragebogen mehrfach.
- Faustregel: Je weniger persönlichen Nutzen sich die Befragten von der Befragung versprechen, desto weniger Zeit werden sie bereit sein zu opfern. Im Rahmen eines Zuschussantragsformulars können Sie daher mehr erfragen, als zum Beispiel im Rahmen einer Konsumentenzufriedenheitsbefragung.
- Aber: Nicht die Länge des Fragebogens, sondern die Zeit, die zur Beantwortung benötigt wird, ist entscheidend. Einfache Fragestellungen und Ankreuz-Kästchen sind schneller ausgefüllt als offene Fragen.
- Achtung: Statt fertige Fragen zu übernehmen, die „halbwegs“ passen, sollten Sie besser die Fragen selbst „maßschneidern“. Gute Fragen erhöhen die Rücklaufquote! Als Vorbereitung macht es dennoch Sinn, sich die Fragebögen von angesehenen Einrichtungen anzusehen.
- Die optische Gestaltung des Fragebogens ist ganz entscheidend! Gestalten Sie ein freundliches Titelblatt, informieren Sie darauf über die durchführende Institution mit Adresse und Kontaktmöglichkeiten; heben Sie klar hervor, wer was wo eintragen muss. Auch Überschriften, Nummerierungen, Seitenzahlen, Pfeile oder Hinweiskästchen vereinfachen - in Maßen verwendet - das Lesen.
- Machen Sie sich von Beginn an Gedanken über die Auswertung des Fragebogens. Wie werden Sie die Daten speichern? Mit welchen Hilfsmitteln werden Sie Ihre Daten auswerten? Bedenken Sie auch, dass offene Fragen und Mehrfach-Nennungen schwieriger auszuwerten sind. Bitte sparsam nutzen!
- Der Fragebogen sollte logisch aufgebaut sein (zum Beispiel über Themenblöcke). Die Struktur sollte den Befragten vertraut sein und so das Ausfüllen erleichtern.
- Eine „Aufwärmphase“ mit leichten Fragen („Eisbrecher-Fragen“) zu Beginn des Fragebogens ist sinnvoll. Auch der Wechsel zwischen unterschiedlichen Frageformen ist sinnvoll, weil so „Ermüdungserscheinungen“ vorgebeugt wird.
- Am Ende des Fragebogens sollten Sie mindestens eine halbe Seite frei lassen, die Befragungspersonen nutzen können, Ihnen einmal so richtig „die Meinung zu sagen“.



15. Was ist bei der Formulierung von Fragen zu beachten?

Die Qualität eines Fragebogens oder Formulars steht und fällt mit der Qualität der Fragen. Im Wesentlichen unterscheidet man zwischen drei Fragetypen:

Geschlossene Fragen – alle möglichen Antworten sind vorgegeben

Der Befragte kann zwischen verschiedenen vorgelegten Antwortalternativen wählen. Ergänzen Sie gegebenenfalls mit den Kategorien „weiß nicht“ oder/und „nicht zutreffend“.

Offene Fragen - Fragen ohne Antwortvorgabe

Hier wird keine Antwort vorgegeben. Nutzen Sie offene Fragestellungen immer dann, wenn die Antwortmöglichkeiten nicht vollständig bekannt sind oder spontane Antworten ermittelt werden sollen. Beachten Sie aber: Der Auswertungsaufwand für offene Fragen ist deutlich höher als für geschlossene Fragen – daher in Maßen verwenden! In aller Regel werden Sie die vielfältigen Antworten im Anschluss wieder zu Kategorien zusammenfassen müssen.

Hybridfragen (halboffene Fragen)

Sie enthalten Antwortvorgaben sowie die Möglichkeit der eigenen Ergänzung. Sie können nicht alle Antwortmöglichkeiten aufzählen? Geben Sie die häufigsten Antworten vor und ergänzen Sie zum Beispiel mit „andere:...“.

Achtung: Bei der Auswertung sollten Sie die halboffene Frage zunächst wie zwei Fragen auswerten. Die erste Auswertung behandelt den geschlossenen Teil der Frage (A,B,C oder „andere“), die zweite Auswertung befasst sich mit den Textbeiträgen in „andere“.

Formulierungen:

- Vermeiden Sie Fachtermini / Fremdwörter.
- Formulieren Sie Ihre Fragen klar und unmissverständlich!
- Die Fragen sollten kurz und prägnant sein!
- Vermeiden Sie Abkürzungen!
- Beziehen Sie Ihre Fragen immer nur auf einen einzigen Sachverhalt!
- Vermeiden Sie Suggestivfragen („Sind Sie nicht auch der Meinung, dass...“).
- Vermeiden Sie doppelte Verneinungen in der Frage.
- Vermeiden Sie bei postalischen Befragungen wenn möglich Filter („Bitte weiter mit Frage xx!“). Sie sind große Fehlerquellen!
- Sollen Meinungen, Einstellungen oder Befindlichkeiten erhoben werden, können sogenannte „Skalen“ verwendet werden, die passend zum Gegenstand entwickelt werden müssen.
- Dabei werden eine Reihe von „Ist-Aussagen“ über den Untersuchungsgegenstand formuliert, auf die der Befragte zum Beispiel mittels fünfstufiger Skala (z.B. „trifft genau zu / trifft eher zu / unentschieden / trifft eher nicht zu / trifft überhaupt nicht zu“) antwortet.
- Einfache „ja / nein“-Aussagen sind ungünstiger. Sie enthalten weniger Information und erregen oft Unwillen bei der Beantwortung.

Die Aufnahme einer Mittelkategorie wie „unentschieden“, „weiß nicht“, „weder noch“ ist eine Glaubensfrage. Einerseits begünstigt eine solche Kategorie das Ausweichen vor einer



klaren Stellungnahme, andererseits kann ihr Fehlen zu Ärger bei den Beantwortern führen.

16. Wie wird die Rücklaufquote bei postalischen Befragungen erhöht?

Ein gewisses Problem stellen die unterschiedlichen Rücklaufquoten bei den einzelnen Befragungsformen dar: Die höchste Rücklaufquote erhält man normalerweise bei mündlichen Befragungen. Im Idealfall sind hier 100% Rücklauf möglich; in der Realität sind es freilich zumeist weniger. Geringer ist die Antwortrate bei Telefonbefragungen. Dies ist abhängig vom Thema sowie dem Zeitpunkt und den Umständen der Befragung. Verglichen mit persönlich-mündlichen oder telefonischen Befragungen werden über postalischen Umfragen insgesamt eher niedrige Ausschöpfungsquoten erreicht: Bei Online-Befragungen liegen sie im Allgemeinen etwas höher als bei schriftlichen Befragungen. Dort können sie teilweise bei nur 5% liegen. In beiden Fällen antworten nur die Personen, die an dem Thema interessiert sind und somit eine bestimmte Motivation für eine Antwort besitzen. Das muss aber nicht so sein. Hier einige Tipps zur Erhöhung der Rücklaufquote:

- Adressieren Sie möglichst persönlich (also z.B. „Frau Helene Mustermann“ statt „Frau Mustermann, Helene“)
- Betonen Sie den „Nutzen“, den die Teilnehmer erwarten können. Neben der Aussicht auf materielle Gewinne stellt auch die Aussicht auf Verbesserung einer Situation einen Nutzen dar.
- Das Befragungsmaterial besteht mindestens aus dem Fragebogen, einem Kuvert, einem Anschreiben, einem Rückantwortkuvert und dem Datenschutzblatt.
- Das Versandkuvert sollte einen seriösen Eindruck vermitteln und sich von der Vielzahl der Werbebriefe auf den ersten Blick unterscheiden.
- Eine Briefmarke ist besser als ein Postwertzeichenstempel. Noch besser: Sonderbriefmarken; sie wirken persönlicher.
- Das Anschreiben muss kurz sein (maximal 1 DIN A4-Seite) und der Zielperson glaubhaft vermitteln, dass die Teilnahme ihr selbst oder wenigstens ihr vergleichbaren Menschen konkrete Vorteile bringt. Das Anschreiben beinhaltet einen seriösen Briefkopf, eine in den Text integrierte Anschrift und eine persönliche Anrede, eine Zusicherung der Anonymität, Name und Telefonnummer eines Ansprechpartners für Rückfragen und dessen Bürozeiten, eine „handgeschriebene“ oder eingescannte Unterschrift. Ein Begleitschreiben enthält idealerweise die folgenden Angaben:

1. Wer ist verantwortlich für die Befragung (Kontaktadresse)?
2. Anrede des Befragten
3. Grund der Untersuchung
4. Antwortappell
5. Rücklauftermin
6. ggf. Anleitung zum Ausfüllen des Fragebogens
7. Dauer des Ausfüllens
8. Zusicherung der Anonymität
9. Dank für die Mitarbeit



- Mit einem Datenschutzblatt demonstrieren Sie den Zielpersonen Ihre Seriosität. Sie teilen darin mit, wie die Daten verarbeitet werden (Trennung von Adresse und Inhalt, elektronische Auswertung in anonymer Form und zu Gruppen zusammengefasst), welche Sicherheitsvorkehrungen getroffen werden (Anonymität, keine Weitergabe an Dritte, Datenschutzbeauftragte).
- Das beigefügte Rückantwortkuvert sollte mit dem Aufdruck „Gebühr zahlt Empfänger“ versehen sein.
- Vorankündigung: Nutzen Sie dazu die Presse, das Radio, E-Mail, Fachzeitschriften etc. Bei vorhandenen Telefoneinträgen kann man die Befragungen gegebenenfalls auch direkt vorher ankündigen.
- Interessantes Thema: Das Thema der Befragung sollte das Zielpublikum interessieren. Dieser Faktor hat einen erheblichen Einfluss auf die Rücklaufquote!
- Schaffen Sie eine Referenz zu „anerkannten Autoritäten“. Die Kooperation mit Universitäten oder lokalen Fachinstituten, mit Verbänden oder Ärzten sollte hervorgehoben werden.
- Schaffen Sie materielle oder immaterielle Anreize zur Teilnahme, aber wahren Sie trotzdem ein seriöses Bild. Die Anreize sollten an die Zielgruppe und das Thema angepasst sein.
- Nachfassen: Keine Antwort? Dann senden Sie nach einer Woche ein erstes Erinnerungsschreiben (Postkarte) oder führen Sie ein Erinnerungstelefonat. Nach drei Wochen folgt ein weiteres Erinnerungsschreiben mit Ersatzfragebogen. Auch ein drittes Erinnerungsschreiben ist durchaus üblich. Planen Sie das Nachfassen mit ein!
- Damit Sie wissen, wer den Fragebogen zurückgesendet hat, können Sie jeden Fragebogen mit einer Identifikationsnummer kennzeichnen. Wenn Sie ihr Erinnerungsschreiben gezielt an die Personen richten möchten, die sich noch nicht beteiligt haben, müssen Sie Ihre Fragebögen kennzeichnen! Wird diese Nummer lediglich zur Kontrolle des Rücklaufs benutzt und später vernichtet, steht sie nicht im Widerspruch zum vertraulichen Umgang mit den Daten und ihrer anonymisierten Verarbeitung.
- Berücksichtigt werden sollte im Übrigen der Versandzeitpunkt. Für Privatpersonen sollte der Fragebogen am Wochenende (Freitag) vorliegen. Ferienzeiten sind dabei zu beachten. Für Betriebe sollte der Fragebogen hingegen am Wochenanfang eingehen.

17. Wie werte ich die Ergebnisse einer Befragung aus?

Verschaffen Sie sich zunächst einen Überblick über Ihren Datensatz. Wie viele Erhebungsbögen liegen vor? Sind stets alle Fragen beantwortet worden? Geben Sie die Fragen und Antworten in eine Datenverarbeitungs-Software ein (z.B. Excel, SPSS). Stellen Sie fest, wie viele gültige Antworten Sie pro Frage erhalten haben.

Fassen Sie die erhobenen soziodemographischen Merkmale der Befragten (z.B. Herkunft, Alter, Familienstand, usw.) zu einer „Beschreibung der Stichprobe“ zusammen. So können Sie ggf. Aussagen darüber treffen, inwieweit die Struktur Ihrer Stichprobe derjenigen der (bekannten) Grundgesamtheit entspricht.

Im Rahmen der Auswertung sollten Sie die Fragen und Antworten zunächst einzeln betrachten. Hierfür bietet die Statistik die folgenden Möglichkeiten:



Deskriptive Statistik (beschreibende Statistik)

- Bestimmung von Häufigkeiten durch „Auszählen“ der Antworten: absolute Werte oder relative Werte (%).
- Lagemaße: Mittelwerte der Verteilung einer Variablen, arithmetisches Mittel, Modalwert, Median.
- Kreuztabellierung (tabellarische Darstellung der gemeinsamen Häufigkeitsverteilung zweier Variablen); z.B.: „gewählte Partei nach Geschlecht“.
- Streuungsmaße: Breite der Verteilung einer Variable, Minimum / Maximum, Spannweite, Standardabweichung, Varianz.
- Klassifizierung, Clusteranalyse: Oberbegriff für verschiedene mathematische Verfahren zur Klassifizierung von Daten, d.h. Klassenbildung anhand bestimmter Merkmale; z.B.: „EU-Regionen mit besonderen Arbeitsmarktproblemen anhand verschiedener Indikatorwerte zum Arbeitsmarkt (Arbeitslosenquote der über 50-Jährigen usw.)“

Inferenzstatistik (schließende Statistik)

- Korrelation (Darstellung der Intensität des statistischen Zusammenhangs zwischen zwei Variablen); z.B.: „Je weiter das Grundstück vom Stadtkern entfernt ist, desto niedriger sind die Grundstückspreise“.
- Regression (Darstellung des Einflusses einer Variablen auf eine andere mittels einer mathematischen Gleichung); z.B.: „Der Grundstückspreis nimmt pro Kilometer Entfernung vom Stadtkern um 2 € pro m² ab“.

Grundsätzlich empfiehlt es sich, die Darstellung und die Interpretation der Ergebnisse als zwei getrennte Arbeitsschritte vorzunehmen. So bleibt ihre Vorgehensweise nachvollziehbar.

18. Was ist Repräsentativität?

Von Repräsentativität wird gesprochen, wenn sich aus einer Stichprobe zutreffende Rückschlüsse auf eine Grundgesamtheit ziehen lassen. Im engeren Sinne ist eine Stichprobe dann repräsentativ, wenn alle Merkmalsträger der Grundgesamtheit die gleiche Chance besessen haben, Teil dieser Stichprobe zu werden.

Generell gilt das „Gesetz der großen Zahl“. Dies besagt, dass sich die Eigenschaften der Stichprobe mit wachsendem Stichprobenumfang den Eigenschaften der Grundgesamtheit annähern. Je stärker die Werte der untersuchten Variablen streuen, desto größer sollte der Stichprobenumfang sein. Wichtig für die Größe der Stichprobe ist vor allem, wie genau die Gegebenheiten der Grundgesamtheit durch die Stichprobe abgebildet werden können. Der minimal erforderliche Stichprobenumfang kann mathematisch berechnet werden.

Die Auswahl einer Teilgesamtheit sollte so vorgenommen werden, dass aus dem Ergebnis der Teilerhebung möglichst exakt und sicher auf die Verhältnisse der Gesamtmasse geschlossen werden kann. Gemäß einer abgeschwächten Begriffsdefinition ist Repräsentativität bereits dann gewährleistet, wenn eine Stichprobe ein verkleinertes



Abbild einer Grundgesamtheit darstellt. Damit können neben allen Wahrscheinlichkeitsauswahlverfahren auch die Quotaverfahren, die zu den nicht-zufallsgesteuerten Auswahlverfahren zählen, zur Bildung einer repräsentativen Stichprobe führen (siehe Frage 19). Allerdings weisen diese Stichproben lediglich eine Repräsentativität hinsichtlich der Quotenmerkmale auf.

Sie möchten wissen, ob ein Ergebnis vertrauenswürdig oder aussagekräftig ist?

Ebenso wichtig wie die Frage nach der „Repräsentativität“ der Stichprobe ist die Frage nach der Qualität der Ergebnisse. Dies sollten Sie anhand von unterschiedlichsten Gütekriterien prüfen. Fragen Sie zum Beispiel, ob die Methode wissenschaftlichen Vorgaben entspricht und ob sie transparent bzw. nachvollziehbar ist. Fragen Sie nach der Quelle und den Hintergründen der Befragung. Nur wenn Sie sich näher mit der Methode auseinandersetzen, werden Sie erfahren, ob und wie vertrauenswürdig oder aussagekräftig die Ergebnisse sind. Dabei müssen die folgenden Aspekte bedacht werden:

- Entspricht die angewandte Methode wissenschaftlichen Maßstäben?
- Sind die benutzten Quellen verlässlich und vertrauenswürdig?
- Können Sie aus der Teilgesamtheit transparente und nachvollziehbare Schlüsse auf die Gesamtheit ziehen?
- Objektivität: Empirische Messungen sollten möglichst objektiv sein, also bspw. unabhängig vom Durchführenden. Ideal wäre es, wenn z.B. Interviewer A, B und C zum selben Ergebnis kommen.
- Reliabilität: Die Ergebnisse sollten zuverlässig sein, das heißt z.B., dass wiederholte Messungen am gleichen Objekt zum selben Ergebnis kommen.
- Validität: Die Validität drückt aus, wie gültig die Ergebnisse einer Studie sind. Es muss gewährleistet sein, dass das gewählte Messinstrument „das misst, was es tatsächlich messen soll.“

19. Vollerhebung oder Stichprobe?

Häufig ist es nicht möglich, alle betroffenen Personen (-gruppen) bzw. Fälle oder Merkmalsträger in die Untersuchung einzubeziehen. Dann ist eine Stichprobe notwendig. Bevor Sie sich mit der Stichprobenziehung auseinandersetzen, müssen Sie Ihre Grundgesamtheit definieren. Wie groß ist die Gruppe der Objekte (Personen, Organisationen,...), für die die Aussagen der Untersuchung gelten sollten?

Wenn diese Grundgesamtheit klein und die interessierenden Merkmale sehr unterschiedlich sind (Meinungen oder Einstellungen zu unterschiedlichen Bereichen, allgemeine Gewohnheiten oder Verhaltensmuster, Verbesserungsvorschläge,...), dann sollten Sie über eine Vollerhebung nachdenken. Doch was bedeutet „klein“ konkret? Die Grundregel lautet: Die Vollerhebung sollte mit einem realistischen Aufwand durchgeführt werden können.

Die Frage, ob eine Vollerhebung oder eine Stichprobe vorzuziehen ist, richtet sich also nach dem jeweiligen Thema sowie nach dem erforderlichen Aufwand und der gewählten Methode. Daher ist keine grundsätzliche Empfehlung möglich, auch eine „Faustregel“ kann nicht genannt werden. Man kann höchstens davon ausgehen, dass ab einer



Grundgesamtheit von etwa 500 – 1000 aus praktischen Gründen eine Stichprobe sinnvoll sein kann.

Eine Vollerhebung vereinfacht die deskriptiven Datenauswertungen (Häufigkeiten, Prozentsätze,...), da alle statistischen Informationen vorliegen und nicht durch Wahrscheinlichkeitsrechnungen abgeschätzt werden müssen. Jedoch erreichen Sie auch bei einer Vollerhebung nicht unbedingt alle Personen, da immer mit Ausfällen gerechnet werden muss, z.B. weil die Personen auch bei wiederholtem Besuch nicht anzutreffen sind oder die Kontaktdaten nicht mehr stimmen. Zudem steigen die Kosten und der Zeitaufwand für eine Vollerhebung mit der Größe der Grundgesamtheit überproportional an.

Ein weiteres Problem: sind die Interviewer unzureichend geschult, kann es zu ungenauen Ergebnissen kommen. In diesem Fall kann eine professionell durchgeführte Teilerhebung unter Umständen die besseren Ergebnisse liefern.

20. Wie ziehe ich eine Stichprobe für eine Befragung? Die Auswahlverfahren.

Sie haben sich für eine Stichprobenziehung entschieden? Dann müssen Sie nun das Auswahlverfahren festlegen und sich über die Größe und Repräsentativität der Stichprobe Gedanken machen.

Die am häufigsten verwendeten Auswahlverfahren sind die einfache Zufallsauswahl, die Schichtung, Klumpenverfahren und Quotenverfahren. Werden mehrere Verfahren kombiniert, spricht man von einem mehrstufigen Auswahlverfahren.

Bei der **einfachen Zufallsauswahl** erfolgt die Auswahl rein zufällig. Jede Person bzw. jedes Untersuchungsobjekt hat die gleiche Auswahlwahrscheinlichkeit. Beispielsweise ist die zufällige Ziehung von Werkstücken aus einer Charge eine einfache Zufallsauswahl aus dieser Charge.

Bei der **geschichteten Zufallsstichprobe** (auch: stratifizierte Zufallsstichprobe) in der Statistik wird die entsprechende Grundgesamtheit in mehrere kleinere Gruppierungen - die sogenannten Schichten - unterteilt. Danach zieht man aus jeder Schicht eine einfache Zufallsstichprobe. Diese Stichproben werden dann beim Schluss auf die Grundgesamtheit entsprechend den Umfängen der einzelnen Schichten, die bekannt sein müssen, gewichtet.

Eine **Klumpen-Stichprobe** ist eine Form der Zufallsauswahl. In einer Klumpen-Stichprobe wird die Grundgesamtheit in viele kleine (oft geografisch abgegrenzte) Teilgesamtheiten zerlegt, die sogenannten Klumpen (beispielsweise Schulen für eine Befragung von Schülern). Dabei geht man im Gegensatz zur geschichteten Stichprobe so vor, dass nur ein Teil der Klumpen zufällig ausgewählt wird und in die Stichprobe gelangt. Die Elemente der Klumpen werden dann allerdings vollständig erfasst (alle Schüler der ausgewählten Schulen werden befragt). Gegenüber der einfachen Stichprobe erreicht man mit einer Klumpenstichprobe vielfach eine erhebliche Kostenreduktion je erhobener Einheit, für die allerdings ein erhöhter Stichprobenfehler, der auch als Klumpeneffekt bezeichnet wird, in Kauf genommen werden muss.



Während bei der Schichtung das Schichtungsmerkmal so festgelegt wurde, dass es möglichst große Unterschiede zwischen den Schichten gibt (erste Schicht: Altersklasse X, zweite Schicht: Altersklasse Y), sollten bei der Klumpenauswahl die Unterschiede zwischen den Klumpen möglichst gering und die Unterschiede innerhalb der Klumpen möglichst groß sein (alle Bewohner der Provinz A, alle Bewohner der Provinz B).

Ein weiteres Beispiel: Für eine Bürgerbefragung im ländlichen Gebiet werden verschiedene Dörfer zufällig ausgewählt. Jedes Dorf ist ein „Klumpen“, in dem alle Bürger befragt werden.

Bei **Quotenverfahren** werden Personen oder Objekte bewusst so ausgewählt, dass die Stichprobe bestimmte Merkmale aufweist, deren Werte in der Grundgesamtheit bekannt sind. Ein Beispiel: Die Stichprobe wurde so ausgewählt, dass 50 von 100 Personen männlich und 50 von 100 weiblich sind, weil auch in der Grundgesamtheit 50% männlich und 50% weiblich sind. Sobald die Quote von 50 weiblichen Personen erreicht ist, wird keine weitere weibliche Person mehr befragt.

Ein großer Nachteil der Quotenverfahren ist, dass ein bewusstes Auswahlverfahren keine Zufallsstichprobe darstellt und somit keine Angaben zur Genauigkeit von Schätzwerten berechnet werden können. Die Ergebnisse von Quotenverfahren sind für die sogenannte schließende Statistik (Inferenzstatistik) unbrauchbar.

Tipp: Zwei Methoden der Zufallsauswahl.

Manuell auslosen: Geben Sie jedem Merkmalsträger (Person, Fragebogen,...) eine Nummer. Schreiben Sie diese auf Zettel und losen Sie wie in der Lotterie aus.

Zufallszahlen in Excel: Geben Sie jedem Merkmalsträger eine Nummer und lassen Sie sich anschließend mittels Excel-Funktion die erforderlichen Zufallszahlen geben. Die Funktion lautet: „=ZUFALLSZAHN(b-a)+a“, wobei a die kleinste Zahl und b die größte Zahl ist. Kennzeichnen und filtern Sie dann alle ausgewählten Merkmalsträger, etwa über eine „Sverweis“-Formel oder Duplikatensuche.

21. Wie groß muss die Stichprobe sein?

In der Praxis bestimmen meistens sachliche, personelle, zeitliche und finanzielle Restriktionen den Stichprobenumfang. Trotz allem gilt, dass die Stichprobe als Teilmenge der Grundgesamtheit, die untersuchungsrelevanten Merkmale möglichst genau abbilden (repräsentieren) soll. Von der Verteilung der Merkmale einer Stichprobe wird also auf die Grundgesamtheit geschlossen. Das Ergebnis stellt Schätzwerte dar.

Zur Berechnung der Stichprobengröße gibt es keine allgemeingültige Regel, da der letztlich erforderliche Stichprobenumfang immer von der Streuung der Untersuchungsvariablen und der Anzahl von Variablen pro Hypothese abhängt. Je mehr Variablen eine Hypothese kombiniert und je größer die Zahl der Ausprägungen jeder einzelner Variable, desto höher sollte der Stichprobenumfang sein. Häufig kennt man allerdings eben jene Streuung vor der Untersuchung nicht. In diesem Fall ist eine Vorstudie erforderlich.

Genauso wie auf der Grundlage des Stichprobenumfangs auf die Exaktheit der Stichprobenergebnisse hinsichtlich ihrer Verteilung in der Grundgesamtheit geschlossen werden kann, ist es auch umgekehrt möglich, ausgehend von einer beabsichtigten



Genauigkeit der Stichprobenergebnisse in Bezug auf die Grundgesamtheit den dafür notwendigen Stichprobenumfang zu berechnen (Wessel 1996, S. 200).

Für beide Herangehensweisen existieren Berechnungsformeln, z.B. in:

Atteslander, P. (1971): Methoden der empirischen Sozialforschung. Berlin.
Bahrenberg, G., Giese, E., Nipper, J. (1990): Statistische Methoden in der Geographie. Stuttgart.

Friedrichs, J. (1990): Methoden empirischer Sozialforschung. Opladen.

Hantschel, R., Tharun, E. (1980): Anthropogeographische Arbeitsweisen. Braunschweig

Holm, K. (1982): Die Befragung. Bd. 1: Der Fragebogen – Die Stichprobe. München.

Rinne, H. (1995): Taschenbuch der Statistik für Wirtschafts- und Sozialwissenschaftler. Frankfurt a. M.

Wessel, K. (1996): Empirisches Arbeiten in der Wirtschafts- und Sozialgeographie. Paderborn.

Eine Regel aber gilt: Je größer die Grundgesamtheit, desto kleiner ist proportional gesehen die benötigte Stichprobe (bei gleichem „Konfidenzintervall“ und gleichem „Konfidenzlevel“).

Ein Beispiel:

Sie benötigen eine Stichprobe, mit der Sie sich zu 95% (Konfidenzlevel) sicher sein wollen, dass die Antworten von jeder theoretisch möglichen Stichprobenkombination maximal 3% nach oben oder nach unten von ihrer Antwort abweichen. Ob sie jetzt zufällig Personen X, Y oder Z in die Stichprobe auswählen, oder ob sie zufällig die Personen J, K und L in die Stichprobe auswählen – die Abweichung wird zu 95% nicht mehr als +/- 3% sein. Somit können Sie sich zu 95% sicher sein, dass auch die Antwort der Grundgesamtheit in diesem Bereich liegen würde.

Wenn also 70% der Grundgesamtheit Antwort A wählen würden, dann möchten Sie sich zu 95% sicher sein, dass die Antworten jeder Stichprobe zwischen 67% und 73% (Konfidenzintervall) liegen. Wie groß muss die Stichprobe sein?

Die benötigte Stichprobengröße entwickelt sich dann wie folgt:

Grundgesamtheit	Stichprobe	Anteil an der Grundgesamtheit
100	92	92%
1.000	516	51,6%
5.000	880	17,6%
10.000	964	9,64%
100.000	1056	1,56%
1.000.000	1066	0,1066%
10.000.000	1067	0,01067%
Bei Konfidenzintervall 3 (+/-3%) und Konfidenzlevel: 95%		

Bei Umfragen kann das Problem auftauchen, dass bei einem Teil der Befragten, die zu einer Stichprobe gehören, wegen Nichtanwesenheit oder Verweigerung keine Antwort erzielt wird oder einem (schriftlichen oder mündlichen) Frageprogramm nur unvollständig entsprochen wird. Dieses „Non-Response-Problem“ kann die Repräsentativität einer Stichprobe gravierend beeinträchtigen, weil eine Nichtbeantwortung meist mit den Befragungsgegenständen verknüpft ist. In diesem Falle kann man aus den ursprünglich Nichtantwortenden eine Stichprobe auswählen und ein anderes Befragungsmodell, z.B.



einer fernmündlichen Befragung anwenden. Bei unvollständiger Bedienung eines Frageprogrammes durch einzelne Befragte werden hierdurch „Surrogate“ für die fehlenden Daten ermittelt und in die Schätzung einbezogen.

Lassen Sie sich Ihre Stichprobengröße im Internet berechnen:

<http://www.gmi-mr.de/resources/sample-size-calculator.php#terms>

Sie kennen nun ihre theoretische Stichprobengröße? Bitte bedenken Sie zudem, dass Sie nicht alle Personen erreichen werden. Um die theoretische Stichprobengröße auch wirklich zu erreichen, müssen Sie deutlich mehr Personen befragen; die zu erwartende Ausfallquote beeinflusst den notwendigen Auswahlstichprobenumfang erheblich!

22. Wo finde ich weitere praktische Informationen?

Kostenloser Online-Kurs zu Excel und Pivot-Tabellen <http://www.teialehrbuch.de/>

Lexikon der Sozialforschung <http://www.lrz-muenchen.de/~wlm/ilmes.htm>

Methoden der Sozialforschung: <http://www.stangltaller.at/ARBEITSBLAETTER/FORSCHUNGSMETHODEN/>

Sozialwissenschaftliche Fachportal <http://www.sowiport.de/>

Statistik <http://web.neuestatistik.de>

Statistik <http://www.multivariate.de/>

Stichprobenrechner im Internet <http://www.gmi-mr.de>

Auswahlverfahren: http://web.neuestatistik.de/inhalte_web/content/start.html

Einige Literaturhinweise:

Backhaus, K., B. Erichson, W. Plinke & R. Weiber (2000): Multivariate Analysemethoden. Springer-Verlag, Berlin.

Gebhardt, H., R. Glaser, U. Radtke & P. Reuber (Hrsg.) (2007): Geographie – Physische Geographie und Humangeographie. Elsevier Spektrum Akademischer Verlag, Heidelberg.

Groves, R.M. et al. (2004). Survey Methodology. New York: Wiley.

Meier Kruker, V. & J. Rauh (2006): Arbeitsmethoden der Humangeographie. Wissenschaftliche Buchgesellschaft, Darmstadt.

Scharf, A., B. Schubert (1995): Marketing – Einführung in Theorie und Praxis. Schäffer & Poeschel, Stuttgart.

Wessel, K. (1996): Empirisches Arbeiten in der Wirtschafts- und Sozialgeographie. Schöningh-Verlag, Paderborn.