

Gute Daten im Magistrat

Genderstatistik und Equality Daten

Leitfaden



Foto: Ingo Pertramer/PID

Stadträtin
Mag.ª Renate Brauner



Foto: Alexandra Kromus/PID

Stadträtin
Sandra Frauenberger

Vorwort

Mit dem Gleichstellungsmonitor und dem Wiener Integrations- und Diversitätsmonitor hat die Stadt Wien Grundlagen für die strategische Ausrichtung, Gestaltung und Evaluierung unserer Gleichstellungs- und Diversitätspolitik geschaffen. Doch wie kommt die Stadt Wien zu gender- und equalitygerechten Zahlen, und wie werden aus diesen Zahlen interpretierbare Daten?

Die Mitarbeiterinnen und Mitarbeiter des Magistrats erheben, analysieren, interpretieren und präsentieren täglich hunderte Daten – von der umfangreichen Studie bis hin zur einfachen Strichliste. Die Berücksichtigung von Themen wie Gender und Equality in der Datenerhebung und -auswertung ermöglicht es dabei, den alltäglichen Bedürfnissen und Lebensumständen der WienerInnen gerecht zu werden. Solche Daten dienen uns allen als Entscheidungsgrundlage und tragen dazu bei, Diskurse zu versachlichen. Der vorliegende Leitfaden soll die Mitarbeiterinnen und Mitarbeiter dabei unterstützen, Gender- und Equalityaspekte in ihrer Arbeit mit Daten zu berücksichtigen.

Die interdisziplinäre Zusammenarbeit zwischen den Fachdienststellen MD–OS, Dezernat Gender Mainstreaming, MA 23 – Wirtschaft, Arbeit & Statistik, MA 17 – Integration und Diversität sowie die partizipative Einbindung von mehr als 20 Fachdienststellen im Erstellungsprozess unterstreicht die Tatsache, dass Gender- und Equalitystatistik nicht als separate Statistiken zu verstehen sind sondern Teile der Statistik sind, wie sie tagtäglich in der Stadt Wien zur Anwendung kommt.

Statistik soll nicht als Selbstzweck verstanden werden. Nur durch ein genaueres Bild der BewohnerInnen, KundInnen und NutzerInnen unserer Dienstleistungen können wir die Leistungen und Angebote der Stadt Wien an den Bedürfnissen aller BürgerInnen ausrichten. So können Veränderungen in unserer Gesellschaft erreicht und Diskriminierungen verhindert werden.

Der vorliegende Leitfaden bietet Orientierung bei gender- und equalitygerechten Erhebungen und Auswertungen von Daten. Denn „gute Daten“ bilden eine gute Basis für zukünftige Entscheidungen.

Amtsführende Stadträtin für Finanzen,
Wirtschaft und Internationales

Amtsführende Stadträtin für Frauen,
Bildung, Integration, Jugend und
Personal

Inhalt

Einleitung und Grundlagen	6
Einleitung	7
Warum Daten wichtig sind	7
Genderstatistiken	8
Equality Daten	9
Genderstatistiken und Equality Daten – Möglichkeiten und Grenzen	10
„Gute“ Daten im Magistrat	10
Datenschutz	11
Die Fragen richtig stellen – Wissen schaffen	15
Wie komme ich zu Daten?	15
Daten vorhanden – was nun?	16
Daten erheben – was nun?	20
Studien vergeben – was nun?	24

Daten dokumentieren	26
Datendokumentation – warum?	26
Datendokumentation – was heißt das konkret?	26
Daten analysieren	29
Es liegen aufbereitete Daten vor	29
Es liegen unaufbereitete Rohdaten vor	29
Ordnen, Aggregieren und Erstellen von Tabellen	30
Bilder sagen mehr als 1000 Zahlen – Erste grafische Aufbereitungen	31
Erste rechnerische Auswertungen	32
Indizes und Quoten	38
Daten interpretieren	45
Dateninterpretation – warum?	45
Dateninterpretation – was heißt das konkret?	45
Daten präsentieren	49
Erstellung von Grafiken	49
Wahl des Diagramms auf Grund der Art der Daten	49
Erstellung von Tabellen	52
Literaturtipps	55

Einleitung und Grundlagen

Einleitung

Gendersensible Statistiken und Equality Daten sollen die Lebensrealitäten der Menschen möglichst wirklichkeitsgetreu abbilden und bestehende Ungleichheiten aufzeigen. Diese Statistiken dienen im Magistrat als Grundlage zur stärkeren Zielgruppenorientierung von Leistungen der Stadt Wien. Aufgrund ihrer zentralen Bedeutung handelt es sich bei der geschlechterbezogenen Datenerhebung und -analyse um einen der fünf Gender Mainstreaming-Grundsätze im Magistrat.

Der vorliegende Leitfaden soll die Mitarbeiterinnen und Mitarbeiter der Stadt Wien dabei unterstützen, „gute Daten“, die auch Gender- und Equality-Aspekte umfassen, zu finden und aufzubereiten. Dabei geht es sowohl darum, mit bestehenden Daten zu arbeiten, als auch selber Daten zu erheben – von der Strichliste bis zu größer angelegten Erhebungen.

Warum Daten wichtig sind

Im Arbeitsalltag des Magistrats spielen Daten eine große Rolle. Dabei muss es sich nicht immer um große Studien handeln: Auch in Jahres- oder Rechenschaftsberichten werden häufig Daten verwendet. BürgerInnenkontakte werden gezählt und interne Auswertungen vorgenommen, wie z.B. die Frage, wie viele KollegInnen der Dienststelle an verschiedenen Ausbildungen teilgenommen haben. Viele MitarbeiterInnen des Magistrats führen Register und wirken an der Erzeugung von Verwaltungsdaten mit, wenn etwa Anträge von BürgerInnen ausgewertet werden. Auch das Führen von Strichlisten, um NutzerInnen einer bestimmten Leistung zu zählen, gehört zur Arbeit mit Daten.

Beispiele, wo Daten regelmäßig in das Berichtswesen der Stadt Wien einfließen:

- Leistungsberichte der Magistratsdienststellen
- Kontrakte
- Wissenschaftsbericht
- Gender Budgeting-Berichte
- Gleichstellungsprogramm
- Statistisches Jahrbuch
- Gleichstellungsmonitor
- Integrations- und Diversitätsmonitor
- Sozialbericht
- interne Berichte
- (...)

Im Zuge der Verwaltungsmodernisierung und der Zunahme von Informations- und Kommunikationstechnologien gewinnt auch die Wichtigkeit von Daten weiter an Bedeutung. Daten können dazu beitragen, konkrete Bedürfnisse zu erfassen und die Leistungen der Verwaltung entsprechend zu gestalten. Aber erst, wenn diese Daten entsprechend aufbereitet werden, entstehen aus Daten Statistiken, die bei der Planung von konkreten Angeboten helfen können.

Die Qualität einer Leistung steigt mit ihrer Treffsicherheit. So wird durch gute Planung und Steuerung die Effizienz einer Maßnahme oder eines Angebotes erhöht. Denn eine Leistung anzubieten, die niemand braucht, ist wahrlich ineffizient. Die Bedürfnisse der BürgerInnen können etwa durch Zufriedenheitsbefragungen erhoben werden oder durch Rückmeldungen, Anfragen, Informationen aus Beratungen und dergleichen zusammengetragen werden.

Auch die Anforderungen der Gesellschaft an die Verwaltung verändern sich und werden komplexer. Dem Wunsch nach mehr Transparenz sowie nach Mitwirkungsmöglichkeiten für die BürgerInnen wird durch die Open-Government-Data (OGD)-Strategie entgegen gekommen, mit der Verwaltungsdaten zur (kosten-) freien Nutzung zur Verfügung gestellt werden.

<https://open.wien.gv.at/site>

Zudem sind Daten und Statistiken für die politische Ebene wichtig, um gesellschaftliche Entwicklungen erfassen zu können, die eventuell neue politische Maßnahmen und gesetzliche Grundlagen notwendig machen. Statistiken können

strukturelle Ungleichheiten in der Bevölkerung aufzeigen und über Zeitreihen oder Monitoringsysteme positive wie negative Veränderungen für bestimmte Bevölkerungsgruppen verdeutlichen. Um die Vielfalt der Menschen in der Bevölkerung möglichst gut abbilden zu können, ist es wichtig, Daten nach verschiedenen Merkmalen, wie bspw. Geschlecht, Herkunft oder Behinderung, aufzuschlüsseln und Statistiken entsprechend zu erstellen. Ein Schwerpunkt dieses Leitfadens liegt daher auf Genderstatistiken und Equality Daten.

Lesetipp: Zur Rolle von Statistik in der Verwaltung: Stadt Wien: „Statistik Journal Wien 1/2013. 150 Jahre Statistik Wien. Amtliche Statistik im Wandel“

<https://www.wien.gv.at/statistik/pdf/150-jahre-statistik-festschrift.pdf>

Bei der Arbeit mit Daten und Statistiken besteht die Aufgabe darin, in der Fülle und Dichte vorliegender Daten die benötigten und seriösen Informationen zu generieren und entscheidungsrelevante Informationen herauszufiltern.

Daten und Statistiken

Für die Begriffe „Daten“ und „Statistiken“ gibt es keine einheitlichen Definitionen. Daten werden meist als Zahlenwerte verstanden, die etwa durch zählen, messen, wiegen oder beobachten gewonnen werden (z.B. Einkommen, Kinderzahl). Mithilfe von verschiedenen Berechnungsarten werden Daten aufbereitet und möglichst übersichtlich in Statistiken dargestellt. Dies geschieht meistens in Form von Tabellen oder Grafiken. Es können mehrere Merkmale erfasst und dargestellt werden (z.B. Einkommen oder Kinderzahl mehrerer Personen nach Alter und Geschlecht aufgeschlüsselt).

Genderstatistiken

Genderstatistiken stellen keine gesonderten Statistiken dar. Vielmehr geht es darum, dass Statistiken gendersensibel sein sollen, um etwaige Unterschiede zwischen den Geschlechtern aufzeigen zu können. Unter Genderstatistik wird daher in einem ersten Schritt die Aufschlüsselung von personenbezogenen Daten nach dem biologischen Geschlecht (Frauen und Männer) verstanden. Dies betrifft sowohl die Erhebung als auch die Analyse der Daten. Genderstatistik meint aber mehr: Erkenntnisse der Frauen- und Gleichstellungsforschung müssen in die Fragestellung von Statistiken einfließen. Dies kann die Fragestellung ergänzen und auch die Statistik verändern oder erweitern.

Beispiele: Im Bereich der Arbeitsmarktbeteiligung ist die Erwerbstätigenquote nicht nur nach dem Geschlecht getrennt zu erheben, sondern auch das zeitliche Ausmaß der Arbeit (z.B. Teilzeit) darzustellen. Auch Kinderbetreuungsstatistiken sind ein gutes Beispiel für Genderstatistiken. Hier geht es nicht nur darum, wer betreut wird, sondern auch um die Öffnungszeiten und Schließtage der Betreuungseinrichtungen. Mit diesen Informationen kann eine bessere Annäherung an die Frage, ob eine Vereinbarkeit mit einem Vollzeitjob für beide Eltern möglich ist, gelingen – auch wenn hier noch weitere Sachverhalte, wie z.B. Arbeitszeitmodelle, eine Rolle spielen. Genderstatistik bedeutet, konkrete Gleichstellungsindikatoren zu entwickeln, wie eben beispielsweise den sogenannten VIF-Indikator (Vereinbarkeitsindikator für Familie und Beruf). Hiermit wird der „Anteil der betreuten Unter-6-jährigen Kinder in Kindertagesheimen, deren Öffnungszeiten mit der Vollzeitbeschäftigung der Eltern vereinbar sind, gemessen an den Unter-6-jährigen Kindern in Kindertagesheimen insgesamt“ dargestellt (Quelle: MA 23).

Im Bereich der Geschlechtergleichstellung ist der Zusammenhang von der Verteilung von bezahlter und unbezahlter Arbeit, der Erwerbsarbeitszeit, Betreuungsmöglichkeiten und dem geschlechtsspezifischen Einkommensunterschied von großer Bedeutung. Dies hat Auswirkungen auf die Verfügbarkeit von Zeit und die finanziellen Möglichkeiten von Frauen und Männern. NutzerInnenstatistiken von diversen (öffentlichen) Angeboten müssen etwa vor diesem Hintergrund betrachtet werden.

Auch für die 4-R-Methode, mit der im Magistrat Gender- und Equality-Analysen durchgeführt werden, werden Daten benötigt. Im Rahmen einer 4-R-Analyse wird folgende zentrale Frage gestellt: Wer (Repräsentation) bekommt Was (Ressourcen), Warum (Realität und rechtliche Situation) und warum nicht? Für jeden Bereich gibt es eine Reihe von Unterfragen. Für den Bereich der Repräsentation und der Ressourcenverteilung sind geschlechtersensible Statistiken unabdingbar, vgl. dazu: Handbuch „Gender Mainstreaming leicht gemacht“.

<http://www.wien.gv.at/menschen/gendermainstreaming/grundlagen/handbuch.html>

Wichtig ist der Ansatz, dass in allen Themenfeldern und Politikbereichen zumindest die personenbezogenen Daten nach Geschlecht getrennt erhoben werden müssen und nicht nur in Bereichen, in denen Geschlechterunterschiede bekannt sind. Es ist nicht immer im Vorhinein offensichtlich, wo die geschlechtsspezifischen Unterschiede liegen. Zudem werden diese Daten von allen Dienststellen jährlich für verschiedene (verpflichtende) Berichte benötigt (siehe Infobox oben).

Equality Daten

Als Equality Daten werden Daten und Kennzahlen bezeichnet, die benötigt werden, um Vielfalt (Diversität), Gleichstellung und Formen von Diskriminierung feststellen und messen zu können. Personenbezogene Daten, die Merkmale wie Herkunft, Religion und Weltanschauung, sexuelle Orientierung und Behinderung erfassen, werden – wie viele weitere Daten – als sensible oder besonders schutzwürdige Daten behandelt. Deshalb ist die Qualität bei der Verarbeitung von Equality Daten, von der Datensammlung bis zur Dateninterpretation besonders wichtig, um den Persönlichkeitsschutz zu gewährleisten. Konkret bedeutet das z.B. Datenschutzaspekte zu berücksichtigen, etwaige Anonymisierungen vorzunehmen und die Rückführbarkeit von Informationen über einzelne Personen zu unterbinden. Bei der Datenerstellung sollen Methoden angewendet werden, die (später) eine korrekte und gleichstellungsorientierte Analyse ermöglichen, wie z.B. die Verwendung von Vertrauensbereichen (= Konfidenzintervalle) zur Ausweisung der Stichprobengröße und -verhältnisse sowie der Aussagekraft von Ergebnissen (siehe Grafik Seite 47). Bei der Beschreibung und Interpretation der Daten sollte nur das beschrieben werden, was auch tatsächlich erfasst und abgebildet werden konnte.

Equality Daten sollen

- Kontext- und Hintergrundinformationen beinhalten (z.B. Geschlechterstruktur, Größe und Verteilung einer Gruppe oder bestimmte Merkmale einer Gruppe, gesetzliche Grundlagen etc.)
- zur Erstellung von Indikatoren geeignet sein,
- eine weitere und vertiefende Analyse und Forschung ermöglichen
- und möglichst repräsentativ sein.

Beispiel: Sie möchten zur Verbesserung Ihrer Angebote eine NutzerInnenanalyse Ihrer Kundinnen und Kunden durchführen. Dafür benötigen Sie möglichst viele Informationen über die KundInnenstruktur, um mögliche Unterschiede in der Verteilung bei den KundInnengruppen und/oder in der Nutzung Ihrer Angebote festzustellen. Neben Informationen zum Geschlecht und Alter können je nach Thema der Bildungshintergrund, der soziale Status, die Staatsangehörigkeit, die Herkunftssprache u.v.m. von Bedeutung sein. Sie können dabei alle KundInnen erfassen (= Vollerhebungsdatensatz) oder eine repräsentative Stichprobe (= Sample) aus ihrem KundInnenstock ziehen. Achtung, allein die Größe ihres Samples bedeutet nicht per se repräsentativ zu sein, auch innerhalb des Samples gilt es, repräsentative Größen festzulegen (z.B. Geschlechter- und Altersverteilung). Selbstverständlich können Sie sich auch für eine rein qualitative Analyse entscheiden, wobei hier die Wahl der Befragungstechnik der interviewten KundInnen von Relevanz ist, vgl.dazu auch das Kapitel 3.3 „Studien vergeben - was nun?“.

Zentrale Merkmale um Verschiedenheiten und Ungleichheiten in der Gesellschaft festmachen zu können sind:

Geschlecht
Staatsangehörigkeit und Geburtsland
Sprache
Religion und Weltanschauung
Behinderung
sexuelle Orientierung
Geschlechtsidentität
Alter
Elternschaft
Einkommen
Arbeitszeit
Bildungsstand
Stellung im Erwerbsleben
Haushaltsformen
Wohnform
(...)

Diese Merkmale können – wenn es die Datengrundlage zulässt – auch in kombinierter Form, z.B. Geschlecht und Behinderung, analysiert werden.

Genderstatistiken und Equality Daten – Möglichkeiten und Grenzen

Statistiken, bei deren Erstellung obige Aspekte der Gleichstellung berücksichtigt wurden, können strukturelle Ungleichheiten sichtbar machen.

Beispiel: Das mittlere Einkommen der Frauen in Wien ist um 18,4% geringer als das entsprechende Einkommen der Männer. Hierzu ist eine Gruppenbildung (die Frauen und die Männer) notwendig. Das heißt aber natürlich nicht, dass jede Frau automatisch weniger verdient als ein Mann. Individuelle Benachteiligungen oder Diskriminierungen sind aus den aggregierten Daten nicht erkennbar. Statistiken können nicht Beratungen und direkten Kontakt mit den BürgerInnen und das Anbringen ihrer konkreten persönlichen Probleme und Wünsche ersetzen.

Daten und Statistiken können einen Beitrag zur Versachlichung von Debatten liefern, indem sie entweder Vermutungen stützen oder auch widerlegen können.

„Gute“ Daten im Magistrat

Eine erste Annäherung, was „gute“ Daten sind, liefern die folgenden Tipps zur Datenqualität, die an den Verhaltenskodex für Europäische Statistik „European Statistics Code of Practice“¹ angelehnt sind:

- **Relevanz:** Relevante Daten werden den NutzerInnenanforderungen gerecht, sie produzieren also nicht am Bedarf vorbei.
- **Genauigkeit und Zuverlässigkeit:** Genaue Daten sind nah an der Wahrheit, die Schätzung sollte dicht am wirklichen Wert liegen. Der statistische Fehler (beispielweise aufgrund einer zu kleinen Stichprobe oder infolge von Antwortausfällen) sollte gering sein und bei der Publikation von Daten angegeben werden.
- **Regelmäßigkeit:** Die Daten erscheinen in regelmäßigen Abständen, so dass die NutzerInnen immer wieder auf diese Daten, der gleichen Quelle zurückgreifen können. Das ist insbesondere für Zeitreihen wichtig und stärkt die Aussagekraft.
- **Aktualität und Pünktlichkeit:** Aktuell sind Daten, wenn zwischen Erhebung und Veröffentlichung möglichst wenig Zeit vergeht. Die Pünktlichkeit bezieht sich auf die Einhaltung von angekündigten Veröffentlichungsterminen bzw. gesetzlichen Vorgaben.
- **Verfügbarkeit und Klarheit:** NutzerInnen sollen einen einfachen, umfangreichen und kostengünstigen oder kostenlosen Zugang erhalten. Klarheit bedeutet, dass Daten auch verstanden werden. Das wird durch zusätzliche Informationen über Konzepte und Methoden erreicht (Metadaten).
- **Vergleichbarkeit:** Daten sollten nicht nur in ihrem unmittelbaren Erhebungsbereich gelten, sondern darüber hinaus auch zeitlich und räumlich vergleichbar sein. Die Datenproduktion sollte daher möglichst einheitlich gestaltet sein.
- **Kohärenz:** Unterschiedliche Datenquellen sollen für gleiche Fragen gleiche Ergebnisse aufweisen und miteinander verknüpfbar sein. Dadurch wird es der Nutzerin bzw. dem Nutzer ermöglicht, diese in Beziehung zu setzen und tiefere Einsichten in die Realität zu erhalten.

Quelle und weitere Informationen: Stadt Wien (Hg.): Gendersensible Statistik: Lebensrealitäten sichtbar machen, Statistik Journal Wien 2/2014

<https://www.wien.gv.at/statistik/publikationen/genderstatistik.html>

Das Erheben und Sammeln von Daten soll kein Selbstzweck sein. Es geht darum, durch Daten und Statistiken Realitäten aufzuzeigen, die für Politik und Verwaltung Handlungsbedarf sichtbar machen. Daten und Statistiken können Grundlagen für die Ausgestaltung konkreter Maßnahmen bilden, sie können diese aber nicht ersetzen. Auch geht es nicht darum, möglichst viele Daten zu sammeln (Stichwort: Datenfriedhof), die niemand mehr durchschaut. Im Gegenteil, es geht um eine Reduktion von Komplexität und die Fähigkeit, entscheidungsrelevante Daten zu generieren.

1 <http://ec.europa.eu/eurostat/de/web/products-manuals-and-guidelines/-/KS-32-11-955>

Datenschutz

In Zusammenhang mit der Erstellung von Statistiken ist beim Thema Datenschutz primär sicherzustellen, dass im Zuge der statistischen Aufbereitung eines Datensatzes, sofern sich dieser auf Personen bezieht, nicht auf die Identität der Personen rückgeschlossen werden kann.

Personenbezogene Daten sind Angaben über Personen, durch die deren Identität bestimmbar ist. Das sind u.a. der Name, das Geburtsdatum oder die Sozialversicherungsnummer.

Sensible Daten sind Informationen über die ethnische Herkunft, die politische Meinung, religiöse Überzeugung, Gewerkschaftszugehörigkeit, Gesundheit oder die sexuelle Orientierung u.a. Diese Informationen sind besonders schutzwürdig und sollen im Rahmen der statistischen Tätigkeit keinesfalls in Verbindung mit einer bestimmten Person gebracht werden können.

Für die Erhebung und Weitergabe personenbezogener Daten bedarf es einer gesetzlichen Grundlage. Wenn Datensätze also für statistische Zwecke verwendet werden sollen, müssen diese Datensätze anonymisiert werden.

Bei **anonymisierten Daten** ist die Identität der Betroffenen für niemanden mehr feststellbar. Es liegt daher keine Datenschutzrelevanz vor. Datensätze werden anonymisiert, indem Personenbezüge entfernt, geschwärzt oder gelöscht werden und damit die Identität von Personen unkenntlich gemacht wird. Das Arbeiten mit anonymen Strichlisten und Befragungen oder zusammengefassten Daten ist daher datenschutzrechtlich völlig unbedenklich.

Hinweis

Zum Thema „Analyse bei kleinen Gruppen, Datenschutz, Mindestanzahl“:

Eine generelle Regel, ab wann die Veröffentlichung von Merkmalen aus Gründen des Datenschutzes bedenklich wird, gibt es leider nicht. Es muss in jedem Fall einzeln abgewogen werden, ob durch eine Veröffentlichung sehr kleiner Fallzahlen auf einzelne Individuen rückgeschlossen werden kann.

Nach je mehr Merkmalen man aufschlüsseln möchte, umso kritischer wird es. Deswegen sollte man zuerst nach jenen Merkmalen auswerten, die einem für die Analyse am wichtigsten sind (z.B. als ersten Schritt nach Geschlecht auswerten, dann nach Bildungsstand, dann nach Zählgebiet).

Es ist stets zu bedenken, dass eine selbst erstellte Statistik – auch wenn sie keine Rückschlüsse auf einzelne Personen zulässt – durch Kombination mit anderen Statistiken dazu führen kann, dass Informationen über einzelne Individuen zum Vorschein kommen.

Rechtmäßig veröffentlichte Daten (z.B. im Telefonbuch, auf offiziellen Seiten im Internet) sind nicht mehr schutzwürdig und können weiterverwendet werden.

Tipp: Wussten Sie, dass es in jeder Dienststelle Datenverantwortliche gibt? Falls Sie einmal Fragen zum Thema Datenschutz im Zusammenhang mit Ihrer Arbeit haben sollten, wenden Sie sich am besten an die Datenverantwortliche bzw. den Datenverantwortlichen Ihrer Dienststelle. Diese sind kompetente AnsprechpartnerInnen für alle MitarbeiterInnen in der Dienststelle (z.B. bei Unsicherheiten und Fragen zur Erstellung von Formularen, zur Anonymisierung, und zur Weitergabe von Daten) und gleichzeitig wichtige Schnittstelle zwischen ihren Dienststellen und der MA 26 – Datenschutz, Informationsrecht und Personenstand. Es ist sinnvoll, die Datenverantwortlichen bei Durchführung von großen Projekten von Anfang an einzubinden. Für weitere Informationen zum Thema Datenschutz steht die MA 26 gerne zur Verfügung.

Beispielhafte Situationen

Situation 1: Referat XXX bearbeitet Anträge zur Mindestsicherung. Die Dienststellenleiterin wünscht eine monatliche Auswertung darüber, wie viele Personen einen Antrag gestellt haben.

Nicht OK: Eine Referentin oder ein Referent leitet Daten über Anträge (oder die gesamten Anträge) des Monats in nicht-anonymisierter Form an ein anderes Referat (auch derselben Abteilung) zwecks statistischer Auswertung weiter.

OK: Die Referentin bzw. der Referent leitet folgende Info an das Statistik-Referat derselben Abteilung weiter: Bei mir haben diesen Monat drei Frauen und sechs Männern Anträge auf Mindestsicherung gestellt.

OK: Die Referentin bzw. der Referent löscht Name, Geburtsdatum und Straßennamen aus den Antragsformularen und leitet anonymisierte Fälle an das Statistik-Referat zur Auswertung weiter.

Hinweis: Das Prinzip des Magistrats als Einheit gilt also nicht im Zusammenhang mit dem Datenschutz.

Tipp: Was ist zu tun, wenn Sie Daten in nicht-anonymisierter Form zugeschickt bekommen, ohne dafür eine Rechtsgrundlage zu haben? Die Daten sollen nicht abgespeichert, ausgedruckt, protokolliert oder abgeheftet werden. Vernichten Sie die Daten und bitten Sie um anonymisierte Aufbereitung (u.U. mithilfe der MA 26 und/oder der MA 23). Lassen Sie eventuell das Antwortmail protokollieren und behalten Sie dieses als Selbstschutz.

Hinweis: Informationen über Anträge dürfen nur auf Datenträgern abgespeichert werden, die explizit dafür vorgesehen sind. Fragen Sie Ihre Vorgesetzte bzw. Ihren Vorgesetzten.

Situation 2: Um die Öffentlichkeitsarbeit in Bezug auf das Infoservice für Adoptiv- und Pflegekinder besser planen zu können, wird von der PR-Abteilung eine Statistik über die Inanspruchnahme des Infotelefon angefragt.

Nicht OK: Die Referentin bzw. der Referent leitet folgende Information an die PR-Abteilung weiter: Im Mai haben 15 Personen das Infotelefon angerufen. Darunter Frau Müller (verpartnert), Herr Meier (verheiratet).

OK: Die Referentin bzw. der Referent leitet folgende Information an die PR-Abteilung weiter: 15 Personen haben im Mai das Infoservice für Adoptiv- und Pflegekinder angerufen. 9 Frauen und 6 Männer. 4 der 15 InteressentInnen waren verpartnert.

Situation 3: In einer Abteilung mit nur einem einzigen männlichen Mitarbeiter soll die Zufriedenheit der MitarbeiterInnen mit einem Angebot ausgewertet werden.

Nicht OK: Aussagen mit Nennung des Geschlechts, da dadurch indirekt auf die Identität des männlichen Mitarbeiters geschlossen werden kann, z.B. von den vier befragten Personen waren drei Frauen mit dem Angebot zufrieden und ein Mann unzufrieden.

OK: Aussagen ohne Nennung des Geschlechts: Drei befragte Personen waren mit dem Angebot zufrieden, eine befragte Person war mit dem Angebot nicht zufrieden.

Situation 4: Es soll die BenutzerInnenfreundlichkeit eines Online-Formulars anhand der hinterlegten NutzerInnen-daten ausgewertet werden.

Nicht OK: In den Datensätzen werden die IP-Adressen der NutzerInnen, aufgrund derer Rückschlüsse auf einzelne Personen möglich sein könnten, gespeichert.

OK: IP-Adresse aus den Datensätzen löschen oder am besten von vornherein auf die Speicherung von IP-Adressen verzichten.

Beispiel für eine gelungene Umsetzung

Dies ist die Erhebung der Herkunft der Bediensteten der Stadt Wien im Rahmen des Wiener Diversitätsmonitorings. Mit etwa 65.000 Mitarbeiterinnen und Mitarbeitern ist die Stadt Wien (inkl. Krankenanstaltenverbund (KAV), Wiener Wohnen, Wien Kanal etc.) die größte Arbeitgeberin Wiens. Fragen bzgl. des gleichberechtigten Zugangs und der Teilhabe von Menschen mit Migrationshintergrund (Personen, die eine fremde Staatsangehörigkeit besitzen und/oder nicht in Österreich geboren wurden) spielen daher nicht nur bei der Leistungserbringung der Verwaltung eine Rolle, sondern insbesondere auch hinsichtlich der Funktion Wiens als Arbeitgeberin. Deshalb fand 2013 erstmals eine Auswertung der Personaldaten in Bezug auf die Herkunft der Bediensteten statt. Da die MA 2 – Personalservice nur die Staatsangehörigkeit der Bediensteten dokumentiert, musste eine Lösung gefunden werden noch weitere migrationsrelevante Merkmale für die Erhebung zu gewinnen. Mit einem statistischen Matching-Ver-

fahren durch die MA 23 – Wirtschaft, Arbeit und Statistik wurde die anonymisierten Personaldaten der MA 2 (ohne Zu- und Vornamen) mit dem Wiener Bevölkerungsregister zusammengeführt. So wurde es möglich, nicht nur das Merkmal „Staatsangehörigkeit“ aus den Personaldaten, sondern auch das Merkmal „Geburtsland“ der Bediensteten auszuwerten. Dies war nur bei jenen Bediensteten möglich, die ihren Hauptwohnsitz in Wien hatten, da das Wiener Bevölkerungsregister die Grundlage bildete. Aus Datenschutzgründen wurden im Rahmen der anonymisierten und aggregierten Erhebung folgende Restriktionen eingehalten:

- Der MA 17 wurden für das Monitoring von der MA 23 Daten in aggregierter Form und keine personenbezogenen Einzeldatensätze übermittelt.
- Fallzahlen < 6 wurden nicht übermittelt und nicht ausgewiesen.
- Alle Zahlen ≥ 6 wurden durch einen Zufallsmechanismus innerhalb der Grenzen ± 1 verzerrt, sodass anschließend nicht aus den Zeilen-/Spaltensummen auf einzelne Fallzahlen <6 zurückgeschlossen werden konnte, obwohl diese nicht ausgewiesen wurden.
- Dienststellen mit weniger als 50 Mitarbeiterinnen und Mitarbeitern wurden in einem Cluster „Sonstige“ zusammengefasst.
- Ausführliche Informationen zu dieser Erhebung finden Sie unter:
<https://www.wien.gv.at/menschen/integration/pdf/monitor-2014.pdf>

Weiterführende Informationen zum Thema Datenschutz

Intranetseite der MA 26: <http://www.intern.magwien.gv.at/ma26/index.html>

Datenschutzhandbuch der MA 26: <http://www.intern.magwien.gv.at/ma26/ds-handbuch.html>

Rechtsgrundlage für Datenschutz ist das Datenschutzgesetz 2000 (DSG 2000) zu nennen (insbes. § 1-14):
<https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=bundesnormen&Gesetzesnummer=10001597>

Hauptteil

Die Fragen richtig stellen – Wissen schaffen

Daten und Statistiken können erste Hinweise für grundlegende Einschätzungen liefern, etwa: Wie viele WienerInnen sind oder wären von einer Maßnahme betroffen? Sie können Aussagen unterstützen oder zur Versachlichung einer Debatte beitragen.

Bevor die Suche nach entsprechenden Daten gestartet wird, ist es wichtig, sich klar zu machen, was tatsächlich gefragt ist. Die Fragen „Was will ich wirklich wissen? Und was brauche ich dazu?“ sind dabei zentral.

Schon beim Formulieren der Frage „Was will ich wissen?“ sind Gender- und Equality-Aspekte zu berücksichtigen. Dies sollte die grundlegende Frage umfassen, ob es unterschiedliche Auswirkungen auf Frauen und Männer gibt. Und auch weitere Equality-Aspekte (Herkunft, Alter etc.) müssen bereits hier mitgedacht werden. Es kann in diesem Zusammenhang auch notwendig sein, die Fragestellung um Erkenntnisse aus der Gleichstellungsforschung zu erweitern. So wird etwa bei Kinderbetreuungsstatistiken nicht nur nach dem Geschlecht der betreuten Kinder gefragt, sondern die Fragestellung um die Öffnungszeiten und Schließtage der Betreuungseinrichtungen erweitert. Denn nur so kann abgeleitet werden, ob eine Vereinbarkeit mit einem Vollzeitjob für beide Eltern möglich ist.

Beispiel: Arbeitsmarkt: Was wollen Sie wissen? Geht es lediglich um die Frage, wie viele WienerInnen erwerbstätig sind? Dann reicht ein Blick auf die absolute Anzahl der Erwerbstätigen. Oder benötigen Sie einen Vergleichswert? Wie viel Prozent der WienerInnen im erwerbsfähigen Alter sind auch erwerbstätig? Hier braucht es die Erwerbstätigenquoten, wofür wiederum die Bevölkerungsdaten notwendig sind. Um strukturelle Unterschiede zwischen den Geschlechtern immer mitzudenken, müssen sowohl die Anzahl der Erwerbstätigen als auch die Erwerbstätigenquoten für beide Geschlechter betrachtet werden. Wollten Sie eigentlich wissen, wie viele WienerInnen einen Job/ ein Einkommen haben, von dem sie auch leben können? Hierzu muss zunächst definiert werden, was das heißt, etwa ein Einkommen über dem Existenzminimum. Dann können sich viele Fragen anschließen, z.B.: Kann die Arbeitszeit eine Rolle spielen, insbesondere der hohe Anteil von Frauen in Teilzeit? Wie hat sich der Niedriglohnbereich entwickelt? Gibt es vielleicht unterschiedliche Ergebnisse für ÖsterreicherInnen sowie MigrantInnen? Sind Jugendliche und/oder ältere Menschen stärker von Arbeitslosigkeit betroffen als solche im mittleren Erwerbsalter? Wie viele Menschen mit Behinderung erzielen kein Erwerbseinkommen über dem Existenzminimum?

Beispiel: Migrationshintergrund: Manchmal wird die Herkunft der MitarbeiterInnen erhoben, während in Wirklichkeit Interesse an einer Nacherhebung der Sprachkenntnisse der MitarbeiterInnen besteht. Von einem vorhandenen oder nicht vorhandenen Migrationshintergrund der MitarbeiterInnen ist jedoch nicht einfach auf Sprachkenntnisse zu schließen. Die Frage nach dem Migrationshintergrund kann in der anonymisierten Auswertung dem Entdecken von Barrieren und gläsernen Decken dienen, während die Frage nach Sprachkenntnissen auf die Feststellung von Kompetenzen des Personals abzielt. Beides sind unterschiedliche Fragestellungen.

Tipp: Es sollte immer überprüft werden, ob Daten tatsächlich die gestellte Frage beantworten. Etwas mehr Zeit auf diese – vielleicht einfach klingende – Frage zu verwenden, kann helfen insgesamt Zeit zu sparen und keine Fehlaussagen zu treffen.

Wie komme ich zu Daten?

Wenn die Forschungsfrage klar formuliert ist, stellt sich als nächstes die Frage „Wie komme ich zu den Daten?“. Hier gibt es drei Möglichkeiten:

- Daten vorhanden – was nun?
- Daten erheben – was nun?
- Studien vergeben – was nun?

Daten vorhanden – was nun?

In einem ersten Schritt sollte immer geprüft werden, ob es bereits Daten gibt, die helfen oder bestenfalls ausreichend sind, um die gestellte Frage zu beantworten.

Wo können Sie Daten finden?

MA 23: Wenn in der eigenen Dienststelle keine thematischen Daten vorhanden sind, ist die erste Anlaufstelle im Magistrat die MA 23 – Wirtschaft, Arbeit und Statistik. Überblicksdaten über die Stadt Wien sind auf der Homepage und im Statistischen Jahrbuch zu finden. Lassen sich die gewünschten Daten nicht sofort finden, helfen die KollegInnen gerne weiter.

<http://www.wien.gv.at/statistik/index.html>

Fachdienststellen: Auch die MitarbeiterInnen in den Fachdienststellen können häufig zu thematischen Fragestellungen mit Daten weiterhelfen. Im Bereich der Genderstatistik lohnt etwa ein Blick in den Wiener Gleichstellungsmonitor, abrufbar über die Webseite der Frauenabteilung.

<http://www.wien.gv.at/menschen/frauen/stichwort/gleichstellung/gleichstellungsmonitor>

Im Bereich der Equalitydaten liefert der Integrations- und Diversitätsmonitor der MA 17 viele aufbereitete Grundlagen auf der Homepage der MA 17 - Integration und Diversität.

<http://www.wien.gv.at/menschen/integration/grundlagen/monitoring/index.html>

Weitere Datenquellen in der Stadt Wien

Gleichstellungsrelevante Daten über die Bediensteten des Magistrats lassen sich bei der Dienststelle der Gleichbehandlungsbeauftragten finden.

<http://www.intern.magwien.gv.at/gleichbehandlung/index.html>

Zum Thema Menschen mit Behinderung finden sich Informationen auf der Homepage der Monitoringstelle für die Rechte von Menschen mit Behinderungen.

<http://www.wien.gv.at/verwaltung/antidiskriminierung/monitoring.html>

Auf der Webseite der MA 24 ist die Studie „Menschen mit Behinderung in Wien“ abrufbar, die Analysen zu elf verschiedenen Datenquellen umfasst.

<http://www.wien.gv.at/gesundheit/einrichtungen/planung/soziales/behinderung.html>

Auf der Webseite der MA 18 lassen sich verschiedene „thematische Karten“ zur Stadtforschung und Raumanalyse finden, wo Erhebungsdaten kleinräumig ausgewertet werden.

<http://www.wien.gv.at/stadtentwicklung/grundlagen/stadtforschung/gis/karten/index.html>

Auch „Open Government Data“, also Verwaltungsdaten, die der Öffentlichkeit frei zugänglich gemacht werden, können für die eigene Arbeit interessant sein.

<https://open.wien.gv.at/site/open-data>

Informationen zu Geo-Daten (ViennaGIS)

<http://www.wien.gv.at/viennagis>

Open Government Wien

<https://open.wien.gv.at/site/google-public-data-explorer>

Weiters finden sich diverse Grundlagen und Berichte mit Daten auf den Seiten vieler Magistratsabteilungen, wie beispielsweise auf der Homepage der Stadtforschung der MA 18.

<http://www.wien.gv.at/stadtentwicklung/grundlagen>

Publikationen und Studien der Energieplanung der MA 20

<http://www.wien.gv.at/stadtentwicklung/energieplanung/publikationen/index.html>

Sozial- und Gesundheitsberichte auf der Seite der MA 24

<http://www.wien.gv.at/gesundheit/einrichtungen/planung/soziales/sozialberichterstattung.html> und
<http://www.wien.gv.at/gesundheit/einrichtungen/planung/soziales/gesundheitsberichterstattung.html>

Webseite der Wohnbauforschung der MA 50

<http://www.wien.gv.at/wohnen/wohnbauforderung/wohnbauforschung.html>

Wiener Lebensqualitätsstudien

Im Rahmen der Wiener Lebensqualitätsstudien wird die Lebensqualität der Wiener Gesamtbevölkerung auf umfassende Art und Weise abgebildet. Die Stichprobe ist derart angelegt, dass sich auch sozialräumliche Unterschiede in der Lebensqualität untersuchen lassen. Außerdem ist durch die Größe der Stichprobe eine Betrachtung spezifischer Subgruppen der Bevölkerung möglich. Weiters lassen sich durch einen Vergleich mit den drei Vorgängerstudien auch langfristige Trends in der Lebensqualität abbilden. Ein derartig umfangreiches Monitoring-Projekt zur langfristigen Entwicklung der Lebensqualität in einer Stadt ist auch auf internationaler Ebene einzigartig.

Die beiden Studien zur Lebensqualität in Wien wurden von der MA 18 in Auftrag gegeben.

- „Lebensqualität in Wien 1995-2013. Sozialwissenschaftliche Grundlagenforschung II“ aus dem Jahr 2015
- „Studie zur Lebensqualität in Wien 2008. Sozialwissenschaftliche Grundlagenforschung I“ aus dem Jahr 2009

Datenquellen in anderen Organisationen

Statistik Austria

Tipp: Themensuche über den alphabetischen Index und direkt abrufbaren Daten. Zudem kann selbstständig in Datenbanken, wie STATcube, gesucht werden. Fragen Sie vor Anfragen bei Statistik Austria in der MA 23 nach, ob die Daten nicht ohnehin hier vorliegen.

<http://www.statistik.at>

OGD Österreich

<https://www.data.gv.at>

Eurostat

Hierbei handelt es sich um das Statistische Amt der Europäischen Union. Eurostat stellt neben Informationen über Staaten (und Beitrittskandidaten) der EU auch Regionaldaten und Städtestatistiken zur Verfügung.

<http://ec.europa.eu/eurostat/home>

EIGE (European Institute for Gender Equality)

Das Europäische Institut für Gleichstellung verfügt über viele Daten und stellt diese nach Themen geordnet sowie in einem Ressourcen-Center zur Verfügung:

<http://eige.europa.eu>

UNECE (United Nations Economic Commission for Europe)

Hier lassen sich grundlegende Informationen sowie weiterführende Tools und Handbücher finden, und zwar zu den Themen

Genderstatistiken <http://www.unece.org/stats/gender>

Migration <http://www.unece.org/stats/migration>

Worldbank

Hier gibt es eine große Anzahl an Kennzahlen von rund 210 Volkswirtschaften; unter anderem auch eine eigene Kategorie „Gender“ mit zahlreichen Indikatoren:

<http://data.worldbank.org/topic/gender>

FRA (Agentur der Europäischen Union für Grundrechte)

Auch hier lassen sich viele Daten aus den Bereichen Gender und Equality finden:

<http://fra.europa.eu/de/publications-and-resources/data-and-maps>

Universitäre und außeruniversitäre Forschungsinstitute: Forschungsinstitute stellen seltener einzelne Daten oder Statistiken zur Verfügung, sondern eher Berichte und Studien, aber auch diese können häufig weiter helfen.

Weitere Recherchemöglichkeiten je nach Thema sind internationale Organisationen (OECD, ILO), die Bundesministerien, Kammern und Verbände, der Hauptverband der Österreichischen Sozialversicherungsträger, u.v.m.

Daten zum Wiener Arbeitsmarkt finden sich zum Beispiel beim waff:

http://www.waff.at/html/index.aspx?page_url=Arbeitsmarktdaten&mid=356

Auch das Sozialministerium bietet Daten auf Bundesländerebene über das Online-Datenbankabfragesystem Bali:

<http://www.arbeitsmarktpolitik.at/bali/>

Relevante vorhandene Datenerhebungen

Volkzählungen versorgen die Verwaltung mit grundlegenden Informationen über die Bevölkerung, wie etwa demografische Merkmale, Informationen über Bildungsstand und Erwerbsstatus, oder zu Haushalts- und Familientypen. Alle Daten auf Personenebene sind nach dem Geschlecht und verschiedenen soziodemografischen Merkmalen (z.B. Alter und Staatsangehörigkeit) verfügbar. Die in Österreich erstmals als **Registerzählung** durchgeführte „Volkzählung“ erfolgte am Stichtag 31. 10. 2011. Die Statistik Austria stellt dabei den Landesstatistischen Diensten (in Wien in der MA 23 angesiedelt) für das jeweilige Bundesland anonymisierte Individualdaten aus der Registerzählung zur Verfügung. Auswertungen nach verschiedenen Merkmalen, wie Geschlecht, Alter, Herkunft etc. sind den Statistiken entnehmbar.

Verwaltungsdaten werden von öffentlichen Behörden in erster Linie nicht für statistische Zwecke erhoben, sondern dienen als Grundlage für die Erfüllung ihrer Aufgaben. So sind etwa die Einkommensdaten des Hauptverbandes österreichischer Sozialversicherungsträger ein Nebenprodukt der administrativen Vorgänge rund um die Sozialversicherung.

Auch die Daten von Fachdienststellen der Stadt Wien, die für die Aufgabenerfüllung erhoben werden (z.B. in den Jugendämtern, bei der Mindestsicherung oder den Kindergärten...) sind Verwaltungsdaten. Zur Sicherung der Datenqualität und der Verbesserung des Wissens über die Nutzerinnen und Nutzer von verschiedenen Leistungen sind Gender- und Equality-Aspekte hier unbedingt zu berücksichtigen.

Mittels **Haushaltsbefragungen** werden Informationen, die nicht oder nur beschränkt durch Verwaltungsdaten produziert, oder in Register eingetragen werden, erhoben. Ausgehend von einer Bevölkerungsstichprobe werden die aus der Befragung gewonnenen Ergebnisse mittels statistischer Methoden auf die gesamte Wohnbevölkerung hochgerechnet. Die Daten können nach unterschiedlichen Merkmalen erhoben werden. Deshalb sind die Ergebnisse von Haushaltsbefragungen aus der Gleichstellungsperspektive besonders interessant, da sich mehrere Informationen/Merkmale miteinander verknüpfen lassen und so genauere Informationen über die Haushalte und ihre Zusammensetzung erfasst werden können. Im Bereich der Genderstatistik sind bspw. Zeitverwendungserhebungen sehr interessant. Hier werden Personen gefragt, wie viel Zeit sie für verschiedene Tätigkeiten verwenden. Die Ergebnisse ermöglichen etwa Erkenntnisse über die Verteilung von bezahlter und unbezahlter Arbeit zwischen Frauen und Männern. Ergebnisse für Wien liegen bei der MA 23 vor und sind aufbereitet im Wiener Gleichstellungsmonitor und im Statistik Journal „Gendersensible Statistik“ nachlesbar.

Quelle und weitere Informationen in: Stadt Wien (Hg.): Gendersensible Statistik: Lebensrealitäten sichtbar machen, Statistik Journal Wien 2/2014

<https://www.wien.gv.at/statistik/publikationen/genderstatistik.html>

Unterschiedliche Methoden für unterschiedliche Ziele (quantitative oder qualitative Zugänge)

Die Anwendung quantitativer Methoden bei der Datenerfassung führt zu quantifizierbaren Ergebnissen, die gezählt werden können, wie zum Beispiel die Anzahl von Frauen und Männern im Gemeinderat oder die Löhne von Frauen und Männern. Quantitative Daten zeigen Veränderungen im Zeitverlauf auf, wie etwa die Entwicklung der Bildungsbeteiligung von Frauen.

Qualitative Methoden erfassen die Erfahrungen von Personen, ihr Verhalten, ihre Meinungen, Einstellungen und Gefühle. Beispiele sind Erfahrungen von Frauen im Bereich der unbezahlten Arbeit oder die Ansichten von Frauen und Männern über die Gründe von häuslicher Gewalt. Um Daten für qualitative Indikatoren zu erfassen, werden häufig partizipative Methoden wie Gruppendiskussionen oder auch Meinungsumfragen genutzt. Ein Beispiel ist ein Indikator über die Haltung junger Männer zur Gleichstellung.

Unterschiedliche Fragestellungen erfordern meist unterschiedliche Methoden und Daten. Dabei sollten quantitative und qualitative Ansätze nicht als Gegenteil gesehen werden, sondern – je nach Frage oder Zielvorstellung – im Forschungsdesign zusammen verwendet werden.

Der Text entstammt weitestgehend: OECD, Gendernet Practice Note: Gender Equality Indicators: What, why and how, unter: <http://www.oecd.org/social/gender-development/44952761.pdf>

Prüfen der Datenquelle

Bei der Verwendung aller Daten sollten Sie sich immer die Frage stellen, wie seriös die Datenquelle ist. Einige Hilfsfragen können sein:

Was wissen Sie über die Quelle? Handelt es sich um eine bekannte Institution, die auf der Grundlage von Gesetzen Daten erhebt oder handelt es sich um einen nicht näher bekannten Treffer einer Internetsuchmaschine?

Sind Quellenangaben vorhanden? Gibt es Datendokumentationen und Definitionen? Erscheinen diese plausibel?

Und auch bei als seriös eingeschätzten Quellen sollten einige Punkte bezüglich der Qualität der Aussagen beachtet werden:

- Es sollte nicht automatisch das erstbeste Ergebnis, das einem sinnvoll erscheint bzw. das die eigenen Vermutungen bestätigt, gewählt werden. Es muss weiter gesucht werden, ob es eventuell andere Studien und/oder Quellen zur gleichen Thematik gibt, die zu anderen Ergebnissen kommen. Eine vertiefte Suche und mehrere Quellen helfen auch, die eigene These abzusichern und eventuell noch weitere Erkenntnisse zu gewinnen.
- Wie aktuell sind die Daten? Aussagen aus dem Wien der 1980er Jahre werden eher weniger sinnvolle Rückschlüsse auf die heutige Situation zulassen. Auch wenn die Daten noch nicht ganz so alt sind, muss geprüft werden, ob es bspw. zwischenzeitlich Gesetzesänderungen gab. Grundsätzlich gilt: es müssen immer die aktuellsten Daten verwendet werden.
- Sind die Daten nach Geschlecht und weiteren Equality-Merkmalen aufgeschlüsselt? Bei seriösen Quellen lassen sich meist schnell Kontaktdaten für Nachfragen zu den Daten und Statistiken ausfinden. Wenn die Daten nicht aufgeschlüsselt öffentlich zugänglich sind, lohnt sich meist die Nachfrage, bei den Kontaktpersonen, ob die aufgeschlüsselten Daten vorliegen und unter welchen Bedingungen (eventuelle Kosten sind zu erfragen) diese zur Verfügung gestellt werden können.
- Handelt es sich um eine Stichproben- oder um eine Vollerhebung? Was heißt das jeweils für die Aussagen? (siehe Kapitel „Daten erheben - was nun?“)
- Bei einer Stichprobenerhebung: War die Stichprobe groß genug? Sind alle für die Fragestellung relevanten Bevölkerungsgruppen in der Befragung inkludiert?

Kritisches Prüfen des Fragebogens: Schauen Sie, ob wirklich gefragt wurde, was als Ergebnis behauptet wird. Hatten die Fragen vielleicht einen Bias, durch den möglicherweise Verzerrungen im Antwortverhalten entstanden sein können? (siehe Kapitel „Daten erheben - was nun?“, Bias – Verzerrungseffekte aus der Gender- und Equality-Perspektive, Seite 22)

Achtung:

Daten und Statistiken ohne Quellenangabe sollten nicht verwendet werden.

Vorsicht ist zudem bei Hochrechnungen mit ungesichertem Datenmaterial geboten, etwa bei „Dunkelziffern“, wie z.B. Schätzungen zum Frauenhandel. Sie können zwar Grundlagen für Einschätzungen bilden, die Aussagen sind jedoch immer nur eingeschränkt nutzbar. Das methodische Vorgehen, sowie die getroffenen Annahmen sollten genau überdacht und in die Überlegungen zur Verwertbarkeit der Aussagen einbezogen werden.

Daten erheben – was nun?

Wenn keine Daten vorhanden sind, die helfen können, bestehende Fragen zu beantworten, gibt es zwei Möglichkeiten: zu komplexeren Fragestellungen können Studien vergeben werden (siehe nächstes Kapitel) oder die Daten können eigenständig erhoben werden. In diesem Fall sind unterschiedliche Ebenen denkbar: von der einfachen Strichliste, etwa wenn NutzerInnen einer Maßnahme gezählt werden sollen, über KundInnenbefragungen bis zur umfassenden statistischen Erfassung aller Bearbeitungsfälle einer Dienststelle. Bevor Daten eigenständig erhoben werden, gilt es auch hier die Leitfrage zu stellen „Was möchte ich wirklich wissen?“. Anschließende Fragen können sein „Was ist überhaupt möglich?“ oder „Was ist im Rahmen meiner Tätigkeit möglich/ Wie viel Aufwand ist angemessen?“.

Vollerhebung oder Teilerhebung?

In einem nächsten Schritt sollte geklärt werden, ob es um Vollerhebungen oder um Teilerhebungen bzw. Stichproben geht. Bei einer Vollerhebung fließen alle Personen oder Daten einer Grundgesamtheit in die Erhebung ein. Grundgesamtheiten können zum Beispiel alle MitarbeiterInnen des Magistrats oder einer Dienststelle, alle EinwohnerInnen von Wien oder auch alle KundInnen der Dienststelle sein. Bei Teilerhebungen oder Stichproben, werden kleinere Einheiten aus der Grundgesamtheit ausgewählt; von den Ergebnissen werden Rückschlüsse auf die Grundgesamtheit gezogen.

Hinweise:

Rückschlüsse von Stichprobenerhebungen auf die Grundgesamtheit sind mit Vorsicht zu treffen. In der Statistik ist hier der so genannte Stichprobenfehler ein wichtiger Indikator, der angibt, wie nah die hochgerechneten Werte der Stichprobe am tatsächlichen Wert der Grundgesamtheit liegen. Die Größe des Stichprobenfehlers hängt von der Größe der Stichprobe sowie von der Streuung der Werte in der Grundgesamtheit ab. Die Qualität der Aussage steigt mit sinkendem Stichprobenfehler. Hierauf ist bei der Vergabe von Studien und Interpretation von Daten zu achten. Zudem müssen alle für die jeweilige Fragestellung relevanten Bevölkerungsgruppen in der Stichprobe inkludiert sein.

Im Arbeitsalltag ergibt sich diese Vorsicht bei Verallgemeinerungen auch intuitiv: So kann aus den Strichlisten einiger KollegInnen über NutzerInnen eines Angebotes natürlich nicht automatisch auf das Nutzungsverhalten aller WienerInnen geschlossen werden. Auch schließt bspw. die Auswertung aller Anfragen per Internetformular Fragen und Anliegen von Personen ohne Internetzugang und -kenntnisse aus. Solche Einschränkungen in den Aussagen müssen immer mitbedacht werden. Im Sinne von Gender- und Equality-Analysen nach der 4-R-Methode² gilt es immer, die Frage mit zu bedenken, wer warum etwas bekommt/nutzt und wer warum etwas nicht bekommt/nicht nutzt. Um eine Stichprobenauswahl der Grundgesamtheit anzunähern, kann mit Gewichtungen gearbeitet werden. Wenn in einer Stichprobe etwa nur 30% Männer vertreten sind, aber bekannt ist, dass der Männeranteil an der Wiener Wohnbevölkerung 48% beträgt, kann diese Verzerrung über eine Gewichtung ausgeglichen werden. Die Männer in der Stichprobe werden mit einem Gewichtungsfaktor multipliziert, damit ihr Anteil möglichst dem in der Grundgesamtheit entspricht. Achtung: Nicht jede Stichprobe lässt sich im Nachhinein mit einer Gewichtung „reparieren“ (z.B. wenn eine für die Umfrage relevante Personengruppe, bspw. PensionistInnen, gar nicht befragt wurde). Auch bei Stichproben ist auf die Vorschriften des Datenschutzes zu achten. Dies gilt insbesondere im Bereich der Gender- und Equality-Daten, wenn kleine Personengruppen nach mehreren Merkmalen betrachtet werden: etwa Frauen/Männer mit Migrationshintergrund und nach Altersklassen innerhalb einer (kleinen) Dienststelle. Hier werden sehr schnell Rückschlüsse auf einzelne/bestimmte Personen möglich sein.

Vergleichbarkeit der Daten

Werden die Daten eigenständig erhoben, ist unbedingt auf ihre Vergleichbarkeit bei verschiedenen Erhebungseinheiten zu achten. So können die Ergebnisse von Zufriedenheitsumfragen oder zur Stressbelastung variieren, wenn die gleiche Umfrage einmal in Sommermonaten und dann wieder im Jahresendstress durchgeführt wird, ohne dass sich an den inhaltlichen Punkten etwas verändert hat. Auch Nutzungsstatistiken können sehr vom Wetter abhängen, wie wenn bspw. die Nutzung eines Parks an verschiedenen Tagen erhoben oder verglichen wird und einmal die Sonne scheint und es ein anderes Mal schneit.

Tipp: Solche Fakten sollten in der Dokumentation der Statistik vermerkt werden, sodass eventuelle Unterschiede auch zu einem späteren Zeitpunkt oder von einer dritten Person nachvollzogen werden können.

Stichtag oder Zeitraum

Wenn Daten selbst erhoben werden, muss auch geklärt werden, ob ein Stichtag oder Zeitraum betrachtet wird. Stichtagsbetrachtungen sind weniger aufwendig, mitunter kostengünstiger und einfacher durchführbar. Typische Stichtagsdaten sind die Zahlen der EinwohnerInnen zu einem fixen Datum, die etwa für WählerInnenverzeichnisse oder für die Berechnung des Finanzausgleichs notwendig sind. Allerdings ist die Aussagekraft meistens nicht so hoch, wie bei der Betrachtung eines Zeitraumes. Hier fallen Ausreißer nicht so ins Gewicht bzw. können sie gesondert betrachtet werden (etwa erhöhter Beratungsbedarf an einem bestimmten Wochentag/Monatsanfang), kürzer andauernde Phänomene können besser erfasst werden. Die reine Erfassung von Vätern in Elternkarenz zum Stichtag 31.12, erfasst zum Beispiel all diejenigen Väter nicht, die zwar eine zweimonatige Karenz während des Jahres gemacht haben, jedoch nicht über den 31.12. Die Aussage kann jedoch trotzdem (ausreichende) Hinweise liefern, wenn die Ergebnisse über viele Jahre betrachtet werden und es lediglich um die Frage geht, ob ein Anstieg zu verzeichnen ist.

Vorteile der Datenerhebung

Daten selbst zu erheben bietet die gute Gelegenheit, diejenigen Daten zu erhalten, die gewünscht sind und die gesucht werden (bei Rückgriff auf bereits vorhandene Statistiken – so genannte sekundärstatistische Erhebungen – müssen eventuell Abstriche gemacht werden). So können Fragen in Umfragen sehr spezifisch formuliert werden oder Erhebungstabellen so definiert werden, wie sie für die Auswertung gebraucht werden (siehe Beispiele unten). Im Sinne der Gender Mainstreaming- und Diversitätsstrategie im Magistrat sollen Gender- und Equality-Aspekte routinemäßig miterhoben werden. Dies betrifft alle personenbezogene Daten, wie beispielsweise bei Nutzungsstatistiken, Zufriedenheitsbefragungen oder Kontakten von BürgerInnen mit Dienststellen. So handelt es sich bei der geschlechterbezogenen Datenerhebung und -analyse um einen der fünf Gender Mainstreaming-Grundsätze im Magistrat. Zudem ist es aus der Gleichstellungsperspektive immer wichtig, die (spätere) Verknüpfbarkeit von Daten und verschiedenen Merkmalen bei der Datenerhebung mitzudenken.

Tipps:

- Wenn Fragestellungen neu formuliert werden oder Umfragen neu designt werden, lohnt es sich, im Vorfeld mehr Zeit auf die Frage zu verwenden, ob wirklich alles erfasst wird. Punkte, die erst nach der Befragung auffallen, sind im Nachhinein kaum noch oder nicht mehr erfassbar.
- Das reine Ansammeln von großen Datenmengen sollte vermieden werden („Datenfriedhof“), dennoch können eventuell Synergieeffekte genutzt werden. So kann bei der Fragestellung überlegt werden, ob z.B. auch Daten miterhoben werden, die später einmal interessant sein könnten oder die etwa für eine andere Fragestellung der Dienststelle hilfreich sein könnten.

Bias – Verzerrungseffekte aus der Gender- und Equality-Perspektive

Ein wichtiger Punkt beim Erheben von Daten ist es, Faktoren, die einen Bias produzieren können, mitzudenken und zu vermeiden. Unter einem Bias wird ein Verzerrungseffekt verstanden, durch den Daten und Forschungsergebnisse (nachteilig) beeinflusst werden können.

Häufige Formen sind:

Androzentrismus: Probleme werden aus einer männlichen Perspektive untersucht, obwohl es sich um Fragen handelt, die Frauen gleichermaßen betreffen. Im Falle der Datenerhebung kann das heißen, dass Frauen in der Erhebung ausgelassen werden oder unterrepräsentiert sind, die Ergebnisse aber auf Frauen und Männer geschlossen werden (etwa ein Problem, wenn Medikamente nur an Männern getestet werden, aber gleichermaßen auch von Frauen eingenommen werden).

Euro-/Ethnozentrismus: Die Bewertungs- und Entscheidungsgrundlage basiert ausschließlich und unreflektiert auf den eigenen Werten, Normen und Vorurteilen. Dabei kann es v.a. bei Interpretationen auch zu rassistischen Formulierungen kommen. Beispiel: Die Gruppe xy besucht seltener ein Theater, weil das in ihrem Kulturkreis nicht üblich ist.

Insensibilität: Das Geschlecht oder andere Equality-Merkmale werden als Variable ignoriert. Dies kann z.B. der Fall sein, wenn Haushalte als kleinste Einheit untersucht und dadurch die Geschlechterbeziehungen innerhalb des Haushaltes ausgeblendet werden. Auch bei wienweiten Befragungen wird oft das WählerInnenverzeichnis zur Versendung der Fragebögen (etwa zur Zufriedenheit mit dem Wohnen) herangezogen und somit die Gruppe der nicht-stimmberechtigten unter 16-Jährigen und aller ausländischen StaatsbürgerInnen aus der Befragung ausgeblendet.

Doppelte Bewertungsmaßstäbe (beziehen sich auf die Interpretation der Daten): Bei Frauen und Männern werden identische Eigenschaften oder Verhaltensweisen unterschiedlich beurteilt. Ein Beispiel wäre, wenn Kinderbetreuungspflichten bei Frauen als normal und positiv, bei Männern aber als abweichendes Verhalten dargestellt würden.

Eine Verzerrung kann auch aufgrund dieser genannten Effekte schon bei der Auswahl der Stichprobe oder der Formulierung eines Fragebogens entstehen. Einfach, aber effektiv ist jeweils die Reflexion z.B. in Form einer Checkliste. Sollten sie bei bereits vorhandenen Daten einen Bias entdecken, können z.B. auch nachträglich Gewichtungen vorgenommen werden, um hier einen Ausgleich zu schaffen.

Quelle bzw. Lesetipp: Genderkompetenzzentrum: „Geschlechtsbezogener Verzerrungseffekt (Gender Bias)“, unter: http://www.genderkompetenz.info/w/files/gkompzpdf/gender_bias.pdf

Beispiele

Individuelles Gestalten von Tabellen: KundInnen, die ein Beratungsangebot nutzen (Strichlisten)
Einfache Erhebung: das Geschlecht ist routinemäßig zu erfassen.

Beispiel	Eingangskanal	Worum geht es?	Geschlecht bekannt?	Wenn ja: Frau/Mann
Individuelles Gestalten von Tabellen: KundInnen, die ein Beratungsangebot nutzen (Strichlisten)				
Einfache Erhebung: das Geschlecht ist routinemäßig zu erfassen.				

Erhebung weiterer Merkmale

Datum	Zeit	Eingangskanal	Art	Worum geht es?	an wen?	Geschlecht bekannt?	Wenn ja: Frau/Mann	Migrationshintergrund?	Alter	Bezirk

Unter „Eingangskanal“ können die Möglichkeiten „schriftlich, telefonisch oder persönlich“ gewählt werden. Unter „Worum geht es?“ können Stichworte zum Inhalt erfasst werden. Die Tageszeit kann hilfreich sein, um einen Überblick zu erhalten, wann Beratungsspitzen sind und wann welche KundInnen die Angebote nutzen. Unter „Art“ kann – wenn es sinnvoll erscheint – die Leistung weiter aufgliedert werden, etwa in „Rechtsberatung, Beschwerde, Frage etc.“. „An wen?“ wäre für größere Teams möglich, um zu schauen, wer welche Beratungen durchführt. Auch das Alter kann von Interesse sein. Hier können bspw. auch gröbere Kategorien gewählt werden, wie <20, 20-49, 50-69, >69 (oder auch ganz anders). Ist der Bezirk unter Umständen bei dieser Fragestellung auch interessant?

Achtung: Bei der Dokumentation von telefonischen Anfragen oder Umfragen muss das Geschlecht immer aktiv erfragt werden. Alleine durch die Stimmlage oder -höhe sind keine seriösen Rückschlüsse möglich.

Tipps:

- Wichtig ist, die Tabelle so zu gestalten, dass sie für die eigene Arbeit passt.
- Um später eine Auswertung der Tabelle zu erleichtern, empfiehlt es sich, hier Antwort-Kategorien vorzudefinieren. Immer an die spätere Auswertung denken.
- Detailinfos können teilweise sinnvoll sein, aber auch Auswertungen erschweren.
- In Teams mit mehreren Personen, ist es wichtig, dass alle die Kategorien gleich verstehen und entsprechend befüllen. Falls Adressen erhoben werden, ist es wichtig, diese gleich in der korrekten Schreibweise (Adressregister/Räumliches Bezugssystem Wien) zu erheben. Denn nur korrekt geschriebene Adressen können auch korrekt verortet werden. Das Adressservice Wien bietet eine Adresssuche mit Geocodierung, die auch in HTML-Formularen verwendet werden kann. Bei der Adresssuche mit Geocodierung wird auf Basis der Eingabe einer Adresse oder eines Gebietsnamens diese bzw. dieser gegen die Adress-Datenbank der Stadt Wien validiert und bei Gültigkeit werden Adressattribute wie beispielsweise Bezirk, Adresse, CountryCode, Straßename, Straßencode sowie die Koordinate (im gewünschten Zielkoordinatensystem) ausgegeben.

<https://open.wien.at/site/datensatz/?id=c223b93a-2634-4f06-ac73-8709b9e16888>

Gestaltung von Fragebögen

Die Gestaltung von Fragebögen (etwa für Zufriedenheitsumfragen) ist ein komplexes Themenfeld. Je besser die Fragen durchdacht sind, umso höher ist die Wahrscheinlichkeit, qualitativ hochwertige Antworten zu erhalten. Vor der Neugestaltung von Fragebögen sollte Unterstützung von ExpertInnen eingeholt werden. Grundsätzliche Punkte, die jedenfalls beachtet werden sollten sind:

- Die Fragen sollten möglichst kurz und verständlich formuliert sein.
- Testen Sie bei jeder Frage, ob Sie verwertbare Antworten bekommen.
- Die Fragen sollen nicht so formuliert werden, dass Antworten suggeriert werden (Bsp.: „Sind Sie auch der Meinung, dass ...“.)
- In einer Fragestellung sollten nicht zwei oder mehr Fragen gestellt werden.
- Begriffe müssen klar und verständlich verwendet werden. Dies betrifft insbesondere auch Gender- und Equality-Aspekte. Beispiele: Wenn nach „Arbeit“ gefragt wird, ist dann die (bezahlte) Erwerbstätigkeit oder auch unbezahlte Arbeit gemeint.
- Hinterfragen Sie zentrale Begrifflichkeiten gründlich. Stimmen die gewählten Begriffe mit der tatsächlichen Fragestellung überein? Beispiel: wenn Sie Informationen über Menschen mit Migrationshintergrund in Wien erhalten möchten, fragen Sie nicht nur nach der Staatsangehörigkeit. Menschen, die im Ausland geboren und nach Wien gekommen sind und die österreichische Staatsangehörigkeit angenommen haben, würden so nicht erfasst werden.
- Fragen sollten möglichst unmissverständlich gestellt werden. Es gilt nach Merkmalen zu fragen, die allgemein bekannt sind, weitgehend einheitlich definiert werden und damit möglichst eindeutig zu beantworten sind. Im obigen Beispiel: bei der Frage nach Geburtsland und Staatsangehörigkeit (ggf. mehrerer Generationen), ist allen Befragten klar, worum es geht. Die Antwort auf die eigentliche Frage nach dem Migrationshintergrund erhalten Sie erst durch die Auswertung der Antworten.

Achtung: Es ist nicht ausreichend, das Geschlecht in Fragebögen nur soweit abzufragen, dass nach der Auswertung gesagt werden kann, wie viele Frauen und Männer an der Umfrage teilgenommen haben. Fragebogen und Auswertungstool müssen so designt sein, dass eine Auswertung aller Antworten nach dem Geschlecht vorgenommen werden kann. Nur so werden spezifische Interessen und Bedürfnisse sichtbar und Maßnahmen können entsprechend gestaltet werden.

Geschlecht, Geschlechtsidentitäten und Geschlechterrollen sind mit sozialen und ökonomischen Rechten und Möglichkeiten, Kompetenzen und Barrieren verbunden. Um Gleichstellung, Vielfalt, aber auch Unterschiede in der Teilhabe und in den Formen von Diskriminierung messen zu können, bedarf es einer durchgehenden Auswertung von Frauen und Männern, und je nach Fragestellung oder Inhalt dazugehöriger weiterer Variablen, die einen Einfluss auf Geschlechterrollen haben können.

Im Folgenden dazu ein Beispiel aus der Lebensqualitätsstudie der Stadt Wien bei Fragen zu Pflege-, Betreuungs- und Hausarbeitsleistungen: Hier war es notwendig, neben der Kategorie Geschlecht, der Frage nach betreuungspflichtigen Kindern, zu pflegenden Angehörigen, zur Erwerbstätigkeit und Hausarbeit auch die Variable „Zeitbudget“ einzubauen. Ergebnis: Frauen verfügen aufgrund der (unbezahlt) geleisteten Pflege, Hausarbeit und Betreuungen durchschnittlich über weniger (Frei-)Zeit als Männer. Diese Erkenntnis ermöglichte eine klarere Auswertung von Unterschieden in der Partizipation von Frauen und Männern am kulturellen, sozialen und politischen Leben. **Tipp:** Neben der Frage „Wieviel Zeit für diesen Leistungen benötigen Sie wöchentlich?“ (= Zeitbudget) empfiehlt sich aus Gleichstellungssicht auch die Frage, an welchen Wochentagen und zu welchen Uhrzeiten diese Zeiten eingesetzt werden müssen, um herausarbeiten zu können, wieviel Zeit etwa für diese unbezahlten Arbeiten an Wochenenden und in der Nacht anfällt.

Hinweis: Es ist nicht abzustreiten, dass Genderstatistiken in ihrer jetzigen Form von einer Zweigeschlechtlichkeit (Frauen und Männer) ausgehen und diese Vorstellung mitunter sogar noch verfestigen. Alle Menschen, die sich z.B. als weiblich **und** männlich oder als **weder** weiblich **noch** männlich definieren, können sich hier nicht zuordnen und werden in der Folge mit ihren spezifischen Anliegen und Bedürfnissen nicht sichtbar. Es ist daher zu überlegen, bei Fragebögen neben der Kategorie „Frau/Mann“ als weitere Ankreuzmöglichkeit die Kategorie „trans*“ anzubieten.

Tipp: Wenden Sie sich bei Fragen an die Wiener Antidiskriminierungsstelle für gleichgeschlechtliche und transgener Lebensweisen (WAS^t):
<http://www.wien.gv.at/kontakte/wast/>

Studien vergeben – was nun?

Wenn Sie Studien vergeben, die auch Daten bzw. empirische Inhalte beinhalten, dann sollten neben den Grundlagen des Vergaberechts die Grundprinzipien der Wiener Gleichstellungs- und Diversitätspolitik (z.B. Definitionen, Antidiskriminierung) bei der Vergabe, Durchführung und der Abnahme einer Studie berücksichtigt werden.

Studien sind oft nicht „reine Zahlenberichte“, sondern umfassen neben quantitativer Forschung auch qualitative Inhalte und Methoden. Zur Erfassung und Beobachtung von Gleichstellung und Nicht-Diskriminierung in Studien sollten für die Daten folgende Grundsätze gelten:

- Auch externe AuftragnehmerInnen der Stadt Wien müssen grundsätzlich in Studien Daten nach dem Geschlecht und weiteren Equality-Merkmalen erheben und analysieren. Bei Aufträgen und Vergaben sollte dies als Standard in der Leistungsbeschreibung festgehalten sein.
- Die zentralen Variablen zur Feststellung von (Un-)Gleichheiten müssen in der Studie enthalten und z.B. bereits im Fragebogen erfasst sein. Aus einer Gender- und Equality-Perspektive kann es notwendig sein, die Fragestellungen zu erweitern oder zu differenzieren um überhaupt Unterschiede und Gemeinsamkeiten feststellen zu können.
- **Beispiel:** Das Beschäftigungsausmaß soll nach Möglichkeit nicht nur in Teilzeit und Vollzeit, sondern auch als Stundenausmaß ausgewiesen werden. In den bestimmten offiziellen Statistiken zur Erwerbstätigkeit gelten etwa Personen mit bereits einer Wochenstunde Beschäftigung als erwerbstätig.
- Die verwendeten Methoden müssen beschrieben werden.
- Bei kleineren, nicht repräsentativen Stichproben sind Differenzierungsgrade nur soweit zulässig, als dass es die Stichprobengröße seriös erlaubt.
- Ebenso sollten dabei Aussagen und Interpretationen ausschließlich auf die erfasste Gruppe bezogen werden.
- Ein Bezug zu einer Referenzgruppe sollte jeweils hergestellt werden können.
- Bei der Auswahl, Darstellung und Interpretation der Daten sind Stereotypisierung und Kulturalisierung zu vermeiden. Beispielsweise gibt es über viele Herkunfts- und Religionsgruppen wenige empirische Informationsgrundlagen, weshalb die Datenauswahl und die Methode oft auf Basis von Annahmen vorgenommen werden, im schlimmsten Fall sogar nach äußeren Merkmalen.

Beispiel: Es existieren in Österreich keine vollständigen Daten über Menschen mit muslimischem Glaubensbekenntnis. Letztmalig wurde flächendeckend das Religionsbekenntnis bei der Volkszählung 2001 erfasst. Heute ist es oft üblich, die Staatsangehörigkeit aus Ländern, in denen Musliminnen und Muslime die Mehrheit in der Bevölkerung bilden, als Merkmal für ein muslimisches Religionsbekenntnis heranzuziehen. Vor dem Hintergrund, dass mit der fremden Staatsangehörigkeit nur ein Teil der Bevölkerung abgebildet wird und aufgrund der Tatsache, dass eine Staatsangehörigkeit wenig bis gar keine Auskunft über den Glauben einer Person gibt, geschweige denn, das Glaubensbekenntnis über dessen tatsächliche Ausübung, sollte bei diesem Beispiel entweder die Methode angepasst oder von einer Interpretation Abstand genommen werden.

- Bei Interviews ist darauf zu achten, dass die Zielgruppe auch tatsächlich die Möglichkeit hat, das Interview vollständig zu beantworten. Das heißt, manchmal gibt es sprachliche, zeitliche und finanzielle Barrieren. Um diesen entgegenzuwirken, können z.B. muttersprachliche InterviewerInnen eingesetzt werden, Interviews auch außerhalb von Arbeitszeiten geführt oder kleine Aufwandsentschädigungen geleistet werden. Auch sollte bei den InterviewerInnen ein Grundwissen über gender und -kultursensible Befragungstechniken vorliegen. So kann es bei Befragungen zu sehr persönlichen Dingen einen Unterschied machen, ob die Person alleine interviewt wird oder ob die Partnerin bzw. der Partner beim Interview anwesend ist. Auch das Geschlecht des/der InterviewerIn kann bei sensiblen Fragestellungen, wie bspw. sexuelle Gewalterfahrung, eine Rolle spielen.
- Die geschlechtergerechte Sprache – und damit eine direkte Ansprache von Frauen und Männern – ist sowohl in Fragebögen, bei den Interviews als auch im Text der Studie zu verwenden.

Beispiel: Sie wollen eine Studie zu Ausbildungs- und Berufsplänen von Jugendlichen mit und ohne Migrationshintergrund beauftragen. Die vorhandenen Datenquellen lassen nicht ausreichende Rückschlüsse zu, weshalb Jugendliche befragt werden sollen. Bereits hier gilt es, die geeigneten Methoden und Vorgehensweisen zu finden. Sollen es Einzelinterviews, persönlich oder per Telefon geführte Interviews sein, mit offenen oder standardisierten Fragen? Oder Fokusgruppenbefragungen? Welche Interviewsprachen sind notwendig? Wen zähle ich zu der Zielgruppe „Jugendliche mit Migrationshintergrund“? Welche Stichprobengröße lassen sowohl die finanziellen also auch die zeitlichen Ressourcen zu?

In der Praxis weisen derartige Studien folgende Mängel und Fehler auf: Die Altersgruppe der Jugendlichen ist mit 15 bis 24 definiert, in der Stichprobe finden sich jedoch mehrheitlich nur Interviewte an den beiden Polen. Die Interviews werden auf Deutsch durchgeführt, weshalb nur langansässige bzw. in Österreich geborene migrantische bzw. aus Deutschland zugewanderte Jugendliche erfasst werden. Die Stichprobe ist klein, sodass bei weiteren Differenzierungen nach Geschlecht und Herkunft keine aussagekräftigen Ergebnisse mehr zu erhalten sind. Um derartige Fehler zu vermeiden, empfiehlt es sich von Beginn an und gemeinsam mit dem bzw. der AuftragnehmerIn grundlegende Eckpunkte festzulegen.

Tipps:

- Definitionsgrundlagen sollten Sie direkt am Beginn festlegen. Welche Definition für Migrationshintergrund wird verwendet? In weiterer Folge ist darauf zu achten, dass die Daten dieser Definition entsprechen.
- Sie sollten bei der Auftragsvergabe vereinbaren, dass Sie z.B. den Fragebogen oder andere wesentliche Teilschritte nochmals überprüfen, bevor er in die Feldphase geht. Hierbei ist darauf zu schauen, ob Gender- und Equality-Aspekte hinreichend berücksichtigt sind (ggf. Mustertabellen zukommen lassen).
- Mit den AuftragnehmerInnen bereits im Vertrag festlegen, in welcher Form die Daten und der Bericht zur Verfügung gestellt werden. Achtung! Stellen Sie als AuftraggeberIn vertraglich sicher, dass Sie ein Anrecht auf den originalen Datensatz haben, da Sie diesen für spätere Analysen und Auswertungen brauchen könnten.
- Bias vermeiden! Bewusstmachen und Reflektieren von Stereotypen, Vorurteilen und Ungleichheiten sowohl auf Seiten der AuftragnehmerIn als auch der AuftraggeberIn (siehe unter Bias – Seite 21)
- Die Stadt Wien hat für die unterschiedlichsten Bereiche viele gute Tipps und Leitfäden erstellt, die auch für AuftragnehmerInnen hilfreich sein können. Z.B.:
http://www.intern.magwien.gv.at/gendermainstreaming/tools_downloads.html
<http://www.intern.magwien.gv.at/index/wien-spricht-anders/gerecht.html>

Auch auf europäischer Ebene gibt es ein Handbuch zur Erstellung von Gleichstellungsdaten:

<http://ec.europa.eu/social/BlobServlet?docId=1427&langId=de>

Hinweis: Der Statistikerlass (MA 5-336/2012) legt die Zusammenarbeit der städtischen Dienststellen mit der MA 23 im Bereich der Statistik und der Registerzählung fest. Die Kolleginnen und Kollegen unterstützen Sie bei der Studienvergabe sowohl hinsichtlich der Festlegung von statistischen Methoden, Standards und Verfahren, als auch bei Vereinbarungen über die Lieferung von statistik- und registerbasierten Verwaltungsdaten an externe Institutionen. Bitte wenden Sie sich bereits vor der Beauftragung an die MA 23, sie helfen damit Doppelgleisigkeiten im Magistrat zu vermeiden. Bitte vergessen Sie nicht, der MA 23 relevante statistische Übersichten, Berichte, und Analysen, die Daten in aggregierter Form enthalten, in elektronischer Form zu übermitteln.

Daten dokumentieren

Datendokumentation – warum?

Eine gute Datendokumentation hilft dabei, Daten verständlich und interpretierbar zu machen. Denn nur wenn bekannt ist, um welche Daten es sich handelt, wie diese entstanden sind und was bei der Verwendung berücksichtigt werden sollte, können diese optimal genutzt werden. Die Dokumentation sollte so gestaltet sein, dass die Daten auch zu einem späteren Zeitpunkt oder für andere NutzerInnen gut verständlich und auffindbar sind.

Datendokumentation – was heißt das konkret?

Welche Informationen bzw. wie ausführlich diese Informationen dargestellt werden sollten, hängt von der Komplexität und dem Umfang des Datensatzes ab. Eine aussagekräftige Dokumentation sollte folgende Informationen (Metadaten) enthalten:

Metadaten sind Daten, die eine Beschreibung der Merkmale und Informationen über die eigentlichen Daten enthalten, aber nicht direkt die Daten selbst sind.

Allgemeine Informationen

Aus welcher Datenquelle stammen die Daten? Was wurde erhoben und aus welchem Grund? Gibt es eine rechtliche Grundlage dafür? Wofür und von wem werden die Daten genutzt?

Hinweis: Häufig werden für bestimmte Analysen bzw. zur Beantwortung verschiedener Fragestellungen keine eigenen Daten erhoben, sondern auf bereits vorhandene Daten zurückgegriffen (Sekundärstatistik).

Werden Daten eigens für einen bestimmten Untersuchungszweck erhoben, handelt es sich um eine **primärstatistische Erhebung**. Der Begriff **Sekundärstatistik** beschreibt hingegen die statistische Auswertung von bereits vorhandenem Datenmaterial, das ursprünglich für andere Zwecke oder Fragestellungen erhoben wurde.

Datenquelle und Erhebungszweck geben Hinweise auf die Zuverlässigkeit der Informationen und für welche Fragestellungen die Informationen Aussagekraft haben. So wird beispielsweise eine Vielzahl an gesetzlich geregelten Erhebungen von der Verwaltung durchgeführt, die mitunter über die Zuerkennung einer Leistung (Mindestsicherung, Wohnkostenzuschuss, Forschungsförderung etc.) entscheiden oder die Höhe einer Abgabe bzw. Steuer festlegen. Dementsprechend kann auch von einer sehr guten Datenqualität der dargestellten Merkmale ausgegangen werden. Daten, für die keine Datenquelle bekannt ist, sollten nicht verwendet werden!

Erhebungsmethode

Wie wurden die Daten erhoben? Handelt es sich um eine Voll- oder Teilerhebung? Was sind die Erhebungsmerkmale und wie sind diese definiert? Welche Einheiten werden in den Daten erfasst?

Hinweis: In der Statistik ist es wichtig zu wissen, was die Grundgesamtheit darstellt, also was die „100%“ ausmacht. Das können beispielsweise „die gesamte Bevölkerung Wien“, „alle Erwerbstätigen“ oder „alle befragten Personen“ sein. Um zu wissen, welche Personen konkret darin erfasst werden, sollten Begriffe genau definiert werden (Beispiel Erwerbstätige: Zählen „inaktive“ Erwerbstätige, die Kindergeld beziehen oder Präsenzdienst leisten dazu oder nicht?). Häufig wird bei Befragungen aus pragmatischen Gründen nur eine Stichprobe untersucht, die exemplarisch für das Ergebnis der Grundgesamtheit stehen soll. In so einem Fall sollte unbedingt die Methode der Stichprobenziehung (z.B. Zufallsauswahl), die Grundlage (z.B. Zentrales Melderegister) und die Erhebungstechnik (z.B. Face-to-Face Interview) angeführt werden, da auf Basis dieser Information besser eingeschätzt werden kann, ob die Stichprobenbasis der jeweiligen Grundgesamtheit entspricht, oder ob bestimmte Personengruppen darin über- oder unterrepräsentiert sein könnten (z.B. sind erwerbstätige Personen in Singlehaushalten schwerer anzutreffen als Elternteile mit Kindern). Auch der Stichprobenumfang sollte angeführt werden, da dieser maßgeblich für die Schwankungsbreite des Ergebnisses der Hochrechnung ist.

Wichtig: Insbesondere bei der späteren Verwendung der Daten ist es für die Interpretation hilfreich, auch außergewöhnliche Bedingungen oder Ereignisse zu dokumentieren wie z.B. besonders gutes oder schlechtes Wetter bei einer PassantInnenzählung. Um zeitliche Vergleiche anzustellen ist es notwendig Zeitreihenbrüche zu dokumentieren, die unterschiedliche Ursachen haben können. Es könnte eine statistische Methodenumstellung durch-

geführt worden sein (neue Fragestellung, andere Befragungsart, geänderte Stichprobenzusammensetzung etc.) oder auch gesetzliche Änderungen stattgefunden haben (z.B. Zeitreihenbruch bei den KarenzgeldbezieherInnen durch die Einführung des Kinderbetreuungsgeldes im Jahr 2002). Insbesondere bei Länderzuordnungen können sich Änderungen aufgrund von Beitritten zu internationalen Organisationen oder Bündnissen ergeben (wie z.B. die EU-Beitritte von Bulgarien und Rumänien im Jahr 2007 oder Kroatien im Jahr 2013). Solche Änderungen können dazu führen, dass Ergebnisse nicht oder nur bedingt mit Ergebnissen aus Vorperioden verglichen werden können.

Datenqualität

Wie zuverlässig sind die Daten? Zu welchem Zeitpunkt und in welchen Abständen stehen die Daten zur Verfügung? Sind die Daten zeitlich und räumlich vergleichbar?

Hinweis: Bei zuverlässigen Datenquellen (z.B. Statistik Austria, Eurostat, etc.) steht in der Regel eine Datendokumentation mit Angaben zur Qualität der Daten, der Periodizität und zur Vergleichbarkeit in zeitlicher Hinsicht bzw. mit anderen Quellen zur Verfügung.

Achtung: Es gibt auch Fälle, bei denen bekannt ist, dass Personengruppen untererfasst sind und die Ursache nicht mit der statistischen Darstellung zusammenhängt. Ein Beispiel ist die Unterschätzung weiblicher Obdachloser, da diese seltener als männliche Obdachlose das Angebot an Schlafstellen nutzen und daher seltener in den Nutzungsstatistiken aufscheinen. Solche Hinweise können nicht direkt aus den Statistiken herausgelesen werden und beruhen auf Erfahrungswerten oder der Expertise der jeweiligen Fachdienststellen. Diese Informationen sind für die weitere Interpretation der Daten bzw. zur Beurteilung der Zuverlässigkeit der Informationen und der Nutzbarkeit für bestimmte Fragestellungen besonders wertvoll und sollten daher bereits in der Datendokumentation festgehalten werden.

Hinweis: Bei Teilerhebungen sollte – insbesondere bei Umfrageergebnissen - auch auf die Ausschöpf- bzw. Rücklaufquoten der Stichprobe geachtet werden, also wie viele Befragungen tatsächlich realisiert werden konnten. Generell liegen die Rücklaufquoten bei schriftlichen Befragungen deutlich niedriger als bei mündlichen Befragungen. Eine hohe Ausschöpfungsquote ist auch deshalb wichtig, weil es im Falle von „systematischen Ausfällen“ zu einer Verzerrung der Ergebnisse, einem so genannten „selection bias“, kommt (z.B. wenn bei einer Befragung zu Lebensbedingungen besonders viele Personen mit hohem Einkommen die Befragung verweigern oder abbrechen). Die Ergebnisse gelten damit nicht für die gesamte Grundgesamtheit (z.B. die gesamte Wohnbevölkerung), sondern es könnten nur mehr Aussagen über einen Teil der Bevölkerung gemacht werden. Genauso wie bei der Stichprobengröße ist es nicht möglich, eine konkrete Zahl für eine ausreichend hohe (und damit repräsentative) Ausschöpfungsquote zu nennen, da hierfür auch andere Kriterien wichtig sind, wie z.B. die Stichprobenauswahl und die Art der Fragestellung. Wichtig für die Beurteilung der Datenqualität ist daher vor allem die Transparenz hinsichtlich des Auswahlverfahrens der Stichprobe, um die Ergebnisse nachvollziehbar zu machen.

Bei **systematischen Stichprobenausfällen** handelt es sich um nicht zufällig verteilte Ausfälle, die sich strukturell hinsichtlich der Befragungsinhalte von den Befragten unterscheiden.

Tipps:

- Originaldaten unverändert archivieren und gegebenenfalls Schreibschutz verwenden!
- Dateinamen überlegt wählen! Um Dokumente wiederzufinden ist neben einer sauberen Ordnerstruktur auch eine eindeutige Benennung der Dokumente nötig. Die Namen von Ordnern und Dokumenten sollten bereits einen Hinweis auf deren Inhalt geben und nicht kryptisch benannt oder fachspezifisch abgekürzt werden (z.B. „Useranforderung“ anstelle von „UA“ oder „Personalstamm“ anstatt „PS“). Eine sinnvolle Namensgebung könnte zum Beispiel folgendermaßen aussehen: **Name_des_Dokuments_JJJJMMTT_Version**
Auf die Verwendung von Sonderzeichen (Umlaute, scharfes „ß“, Dollarzeichen, u.ä.) sollte verzichtet werden und anstelle von Leerzeichen besser Unterstriche „_“ gesetzt werden, da es sonst bei weiteren Handlungen (Verlinkungen, Berechtigungsvergaben etc.) zu Problemen kommen kann.
- Wenn bekannt, notieren Sie am besten gleich auch die Kontaktdaten der Ansprechperson(en) zu den Daten, die bei Nachfragen kontaktiert werden können.
- Erhaltenswerte Daten im ELAK markieren! Wichtig ist auch der Vermerk, ob die Daten freigegeben wurden und uneingeschränkt veröffentlicht werden können. Sollte es dazu Einschränkungen geben, sind diesbezügliche Vereinbarungen möglichst konkret festzuhalten. Wenn nötig, ist es auch möglich eingeschränkte Leseberechtigungen für die jeweilige Aktenzahl im ELAK zu vergeben. Definieren Sie geeignete Suchbegriffe und -kriterien, um die Auffindbarkeit zu erleichtern. Beachten Sie dabei aber auch die Erfordernisse des Datenschutzes (z.B. Zugriffs- und Vertretungsrechte).
- Es hat sich für einige Abteilungen als praktikabel erwiesen, ein einheitliches Formular für die Datendokumentation als Tabellenblatt bzw. „Dokumentationsblatt“ zu verwenden.
- Veraltete Speichermedien rechtzeitig umspeichern!
- Sollten sie „wiki“ in Ihrer Abteilung im Rahmen des Wissensmanagement verwenden: Nützen Sie es auch für Ihre Datendokumentationen!
- Bitte nicht vergessen: Alle für den Bundesländer- und Städtevergleich relevanten statistische Daten, Berichte und Analyse sind an die MA 23 – Wirtschaft, Arbeit und Statistik zu übermitteln (siehe Erlass MA 5-336/2012). Wenn die statistische Unterlage aufgrund der Vertraulichkeit des Inhalts nicht veröffentlicht werden soll, ist diese mit einem Vermerk „vertraulich“ zu versehen.

Daten analysieren

Um aus Daten Erkenntnisse gewinnen zu können oder um die Informationen erst einmal zu verdichten, ist es sinnvoll die vorliegenden Daten mithilfe unterschiedlicher Techniken aufzubereiten, bevor sie inhaltlich interpretiert werden. Die Daten werden strukturiert, geordnet, charakterisiert und durch Kennzahlen, Tabellen, Grafiken, etc. „beschrieben“.

Hinweis: Die MA 23 unterstützt Sie gerne bei der projektbezogenen Analysen und Auswertung von statistik- und registerbasierten Verwaltungsdaten und steht auch als interne Dienstleisterin für diese Analysen und Auswertungen zur Verfügung. (siehe STATISTIKERLASS MA 5-336/2012)

Es liegen aufbereitete Daten vor

Heutzutage können Sie eine Fülle an bereits aufbereiteten Daten und Statistiken (z.B. in Form von Tabellen und/oder Diagrammen) in sehr kurzer Zeit finden (siehe Kapitel „Einleitung“). Bevor Sie jedoch beginnen, mit einer Tabelle loszuarbeiten, sollten Sie sich die Frage stellen, ob die Statistik überhaupt Antworten auf die Fragestellung (Kapitel „Die Fragen richtig stellen“) geben kann. (siehe Kapitel „Was will ich wissen?“)

Liegen die Daten schon aufbereitet vor (z.B. eine heruntergeladene Tabelle), können einzelne Schritte der Datenaufbereitung entfallen. Auf jeden Fall ist es sinnvoll, einen Blick auf die Metadaten zu werfen (siehe Kapitel „Daten dokumentieren“).

Es empfiehlt sich, stichprobenartig ein paar Überprüfungen an der Tabelle vorzunehmen, z.B.:

- Ergibt die Summe aus den Spalten „Frauen“ und „Männer“ jeweils die Zahl in der Spalte „Insgesamt“?
- Stimmen die Größenordnungen?
- Anteile müssen z.B. immer kleiner 1 bzw. kleiner 100 im Fall von Prozentzahlen sein.

Bei Zeitreihen ist es sinnvoll, die Daten mit denselben Daten der Vorjahre zu vergleichen. Nicht plausible Entwicklungen (z.B. überdurchschnittliche Schwankungen) können auf folgende Ursachen hindeuten (Kapitel „Daten dokumentieren“):

- Zeitreihenbruch aufgrund von:
 - anderer Klassifizierung (z.B. seit dem EU-Beitritt eines Landes wird dieses zur EU und nicht mehr zu den Drittstaaten gezählt)
 - anderer Erhebungsmethode (z.B.: Fragestellung bei einer Umfrage hat sich geändert, Umstellung von händischer auf elektronische Zählweise)
 - anderer Berechnungsmethode (z.B.: für das BIP gelten seit November 2014 neue Berechnungsregeln)
- Besondere Vorkommnisse im Zusammenhang mit dem Erhebungstag und/oder der Erhebungsmethode (z.B.: besonders schlechtes Wetter am Tag einer PassantInnenzählung)
- Fehler in den Daten (Kommaverschiebungen, Formelfehler,...)

Es liegen unaufbereitete Rohdaten vor

Daten können in absoluter Rohform, das heißt als einzelne unsystematisch gereichte Zahlenwerte vorliegen. In diesem Fall ist es notwendig, die Daten als ersten Schritt aufzubereiten, also

- zu ordnen
- zu aggregieren (zusammenzufassen)
- eine Tabelle zu erstellen
- ein Diagramm zu erstellen

Ordnen, Aggregieren und Erstellen von Tabellen

Beispiel: Alters- und Geschlechtsverteilung

Sie wollen eine Gruppe mit 38 Personen nach dem Merkmal „Alter“ auswerten. Besonders interessieren Sie Kindergartenkinder und Personen im erwerbsfähigen Alter. Da es auch interessant wäre, das Geschlecht der Personen zu kennen, erfassen Sie dieses gleich mit (siehe Kapitel „Einleitung“).

Folgende Altersangaben und Geschlechtszugehörigkeiten (f/m) mehrerer Personen wurden erhoben:

56f (56 Jahre/Frau), 48m (48 Jahre/Mann), 0f, 23m, 1m, 3f, 4m, 76m, 38m, 5m, 19f, 61f, 8m, 15m, 36m, 18m, 3m, 18m, 31f, 34m, 34f, 37f, 7m, 2f, 40m, 41m, 43f, 45m, 20m, 45f, 50f, 1m, 39f, 53f, 63f, 5f, 74f, 17f.

Diese Daten können geordnet werden, indem sie z.B. aufsteigend nach dem Alter geordnet und anschließend in Altersklassen zusammengefasst werden. Um die oben gestellte Frage beantworten zu können, müssen die Altersklassen richtig gewählt werden. In diesem Fall sollte es nachher möglich sein, „unter 6-Jährige“ (Kindergartenkinder), sowie „15-64-Jährige“ (Personen im erwerbsfähigen Alter) auszuwerten.

Die Tabelle könnte beispielsweise so aussehen:

Altersklassen	Frauen	Männer	Insgesamt
< 3 Jahren	2	2	4
3 bis unter 6	2	3	5
6 bis 10	0	2	2
11 bis 14	0	0	0
15 bis 19	2	3	5
20 bis 29	0	2	2
30 bis 39	4	3	7
40 bis 49	2	4	6
50 bis 59	3	0	3
60 bis 64	2	0	2
> 64 Jahre	1	1	2
Summen	18	20	38

Quelle: erfundenes Beispiel

Sie müssen sich bewusst sein, dass durch die Aggregation der Daten immer Informationen über die konkreten Merkmalsausprägungen verloren gehen (hier wäre es die Information des genauen Alters).

Genauso wie bei bereits aufbereiteten Daten empfiehlt es sich, stichprobenartige Überprüfungen zu machen. Um die Richtigkeit der Tabelle zu überprüfen können Sie Spaltensummen (18, 20) bilden und diese mit der Tabellensumme (38) abgleichen ($18+20=38$).

Die Klassen bzw. Kategorien sollten sinnvoll gewählt werden. Werden die Klassen zu groß gewählt (0-29, 30-49, 50-79), gehen Informationen verloren. Sie sehen dann z.B. nicht mehr, dass es keine einzige männliche Person im Alter von 60-64 Jahren gibt.

Tipp: Wenn Sie beim Arbeiten mit einer Tabelle in Excel feststellen, dass es besser wäre, Zeilen und Spalten zu tauschen, können Sie sich der Funktion „Transponieren“ behelfen: Tabelle markieren und kopieren – Rechtsklick dort, wo die Tabelle eingefügt werden soll – Einfügeoptionen: Transponieren (T).

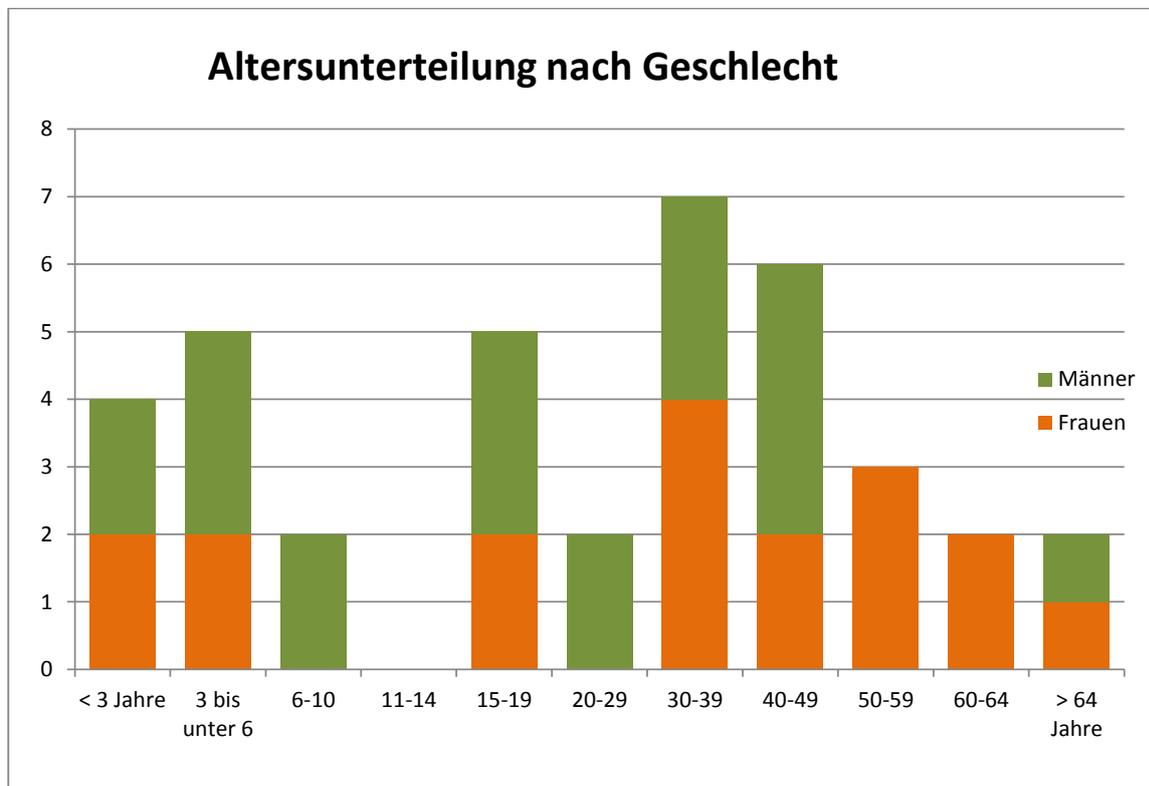
Bilder sagen mehr als 1000 Zahlen – Erste grafische Aufbereitungen

Für die Darstellung von diskreten Merkmalen (d.h. bei abzählbar vielen Elementen wie z.B. dem Geschlecht) eignet sich das Säulendiagramm, um einen Überblick über die Verteilung der Daten zu bekommen. Bei stetigen Merkmalen (d.h. bei unendlich vielen bzw. bei nicht abzählbar vielen Elementen wie z.B. der Körpergröße) ist ein Histogramm dazu geeignet, die Verteilung der Daten darzustellen. Hierbei entspricht nicht die Höhe, sondern die Fläche des Balkens der Häufigkeitsausprägung. Ein typisches Beispiel für die Verwendung von Histogrammen sind Einkommensverteilungen.

Histogramme lassen sich nur mit Statistik-Software erstellen, weshalb in weiterer Folge mit einem einfachen Säulendiagramm gearbeitet wird.

Um die Daten unseres obigen Beispiels zu visualisieren werden die Daten zuerst in Klassen eingeteilt. Eine sinnvolle Klassenbreite erhalten Sie am besten durch Ausprobieren! In jedem Fall sollten die Gruppen überschneidungsfrei sein (0-9, 10-19, 20-29)!

Die obige Tabelle enthält eine solche Einteilung. Anschließend wird ein Säulendiagramm über diese Klassen erstellt, woran Sie schnell erkennen, dass es viele Personen im Alter um die 40 Jahre bzw. im Kleinkindalter gibt.



Grafik: erfundenes Beispiel

Erste rechnerische Auswertungen

Die Daten sind jetzt strukturiert und aufbereitet. Nun können als nächste Schritte erste Berechnungen und Auswertungen folgen.

Absolute und relative Häufigkeiten

Die Anzahl der Merkmalsausprägungen in den einzelnen Klassen (Anzahl der Personen in den Altersgruppen) werden auch als **absolute Häufigkeiten** (sieben Personen sind zwischen 30 und 39 Jahren alt) bezeichnet. Um einen ersten Eindruck über die Datenverteilung zu erhalten, ist die Angabe von **relativen Häufigkeiten** hilfreich. Sie werden berechnet, indem die absolute Häufigkeit eines Merkmals durch die Anzahl der zugrundeliegenden Gesamtmenge geteilt wird. Üblicherweise wird noch mit 100 multipliziert, um eine Prozentzahl zu erhalten. Im obigen Beispiel: 18,4% (=rel. HK) aller Personen sind zwischen 30 und 39 Jahren alt (Berechnung: $7/38 \cdot 100 = 18,4\%$). 22% der Frauen ($4/18 \cdot 100$) und 15% der Männer ($3/20 \cdot 100$) sind zwischen 30 und 39 Jahren alt.

Verhältnisse bilden

Das Bilden von Verhältnissen ist in der Datenanalyse ein sehr häufig verwendeter Vorgang. Durch das In-Beziehung-setzen zweier Zahlen zu einander werden unterschiedliche Aspekte zahlenmäßig abgebildet.

Beispiel: Höhe der Bildungsabschlüsse nach Geschlecht:

Bildungsabschlüsse	Insgesamt	Frauen	Männer
Insgesamt	85	44	41
ohne Tertiärabschluss	43	21	22
mit Tertiärabschluss	42	23	19

Quelle: erfundenes Beispiel

Insgesamt gibt es in der betrachteten Gruppe 85 Personen (Tabellensumme). Diese Grundgesamtheit wird in der Tabelle nach zwei verschiedenen Merkmalen (Geschlecht und Bildungsabschluss) aufgeschlüsselt. Die Spaltensummen geben Auskunft über die Anzahl der Frauen und Männer während die Zeilensummen Auskunft über die Anzahl der Personen ohne bzw. mit Tertiärabschluss geben. Die Summe der Zeilensummen sowie die Summe der Spaltensumme muss wieder die Grundgesamtheit (Tabellensumme) ergeben.

Eine Teilgröße zur Grundgesamtheit in Beziehung setzen:

Mehr als die Hälfte aller Personen ist weiblich (Anteil der Frauen an allen Personen: $44/85 = 0,52$)

49% aller Personen haben einen Tertiärabschluss (Anteil der Personen mit Tertiärabschluss an allen Personen: $42/85 = 0,49$)

Zwei Teilgrößen zueinander in Beziehung setzen:

Die Anzahl der Männer mit Tertiärabschluss in Relation gesetzt zur Anzahl der Frauen mit Tertiärabschluss: $19/23 = 0,83$

Verschiedenartige Größen zueinander in Beziehung setzen (unabhängig vom obigen Beispiel)

Beispiel: Zahl der Geburten / Zahl der Frauen im Alter von 15-44 Jahren

Beispiel: Anzahl des konsumierten Kaffees / Preis eines Kaffees

Beispiel: Zahl der Lehrstellensuchenden pro offene Lehrstelle

Die Verwendung von Beziehungszahlen ist nur aussagekräftig, wenn die beiden Größen tatsächlich in einer Beziehung zueinander stehen. Im oberen Beispiel eher nicht sinnvoll wäre es, die Zahl der Geburten an der Wohnbevölkerung insgesamt anzugeben. Ist die Wohnbevölkerung z.B. an sich groß, besteht aber aus vielen Kindern und/

oder vielen SeniorInnen, muss die Anzahl der Geburten nicht höher sein, als in einer kleineren Wohnbevölkerung, die jedoch dieselbe Menge an Frauen im gebärfähigen Alter aufweist.

Wie Sie anhand der gebildeten Verhältniszahlen sehen, kommt es immer darauf an, was wozu in Relation gesetzt wird. Es ist wichtig, in ergänzenden Texten genau zu beschreiben, wie das Verhältnis gebildet wurde.

Mittelwerte

In der Folge werden die gängigsten Mittelwerte und ihre Unterschiede dargestellt. Wir verwenden dabei die Bezeichnung „Mittelwert“ für Werte, die genau genommen die „zentrale Tendenz“ eines Datensatzes angeben. Für die Interpretation bzw. die Beurteilung der Aussagekraft all dieser Mittelwerte ist es jedoch jedenfalls wichtig zu wissen, wie weit die Werte des jeweiligen Datensatzes vom berechneten Mittelwert entfernt sind (Streuung).

Arithmetisches Mittel (Durchschnitt)

Das arithmetische Mittel (Durchschnitt) ist der gebräuchlichste Mittelwert. Es wird errechnet, indem die beobachteten Werte addiert und anschließend durch die Anzahl der beobachteten Werte dividiert werden.

Beispiel: Wie hoch ist das durchschnittliche Nettomonatseinkommen?

Nettomonatseinkommen in Euro, geordnet nach Höhe

E_1	E_2	E_3	E_4	E_5	E_6	E_7
1.200	1.300	1.600	1.700	1.800	2.000	2.300

Quelle: Zahlen frei gewählt

$$\begin{aligned} \text{arithmetisches Mittel (Durchschnitt)} &= \frac{\text{Einkommen}_1 + \dots + \text{Einkommen}_n}{n} \\ &= \frac{E_1 + E_2 + E_3 + E_4 + E_5 + E_6 + E_7}{7} \end{aligned}$$

$$\frac{1200 + 1300 + 1600 + 1700 + 1800 + 2000 + 2300}{7} = 1.700 \text{ Euro}$$

Wichtig: Bei der Angabe von Mittelwerten nicht auf die Einheit vergessen.

Das arithmetische Mittel ist eine beliebte Maßzahl. Es sollte jedoch nicht überinterpretiert werden.

Beispiel: Anzahl Kinder pro Frau

Frauen in Österreich bekommen durchschnittlich 1,44 Kinder. Dass dies ein rein rechnerischer Wert ist, liegt auf der Hand.

Ausreißer in der Datenverteilung verzerren das arithmetische Mittel stark, wie das folgende Beispiel zeigt.

Beispiel: Nettomonatseinkommen in Euro, geordnet nach Höhe

E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8
1.200	1.300	1.600	1.700	1.800	2.000	2.300	4.900

Quelle: Zahlen frei gewählt

Der Ausreißer ist in diesem Beispiel E_8 mit 4.900 Euro (sonst entsprechen die Werte dem vorherigen Beispiel). Das arithmetische Mittel (Durchschnitt) entspricht jetzt 2.100 Euro, liegt also um 400 Euro höher als im obigen Beispiel und das aufgrund eines einzigen Wertes! Es beziehen aber nur zwei (E_7 und E_8) dieser acht Personen ein Einkommen, das über diesem Mittelwert liegt (die anderen sechs liegen alle darunter).

Hinweis: Ein Vorteil des arithmetischen Mittels ist, dass es gut für mathematische Berechnungen verwendbar ist. Angenommen es gibt zwei Verteilungen: E_1 bis E_4 und E_5 bis E_8 . Die beiden arithmetischen Mittel sind 1.450 und 2.750 Euro. Soll nun der Mittelwert dieser beiden Verteilungen berechnet werden, muss nur das arithmetische Mittel aus den beiden Durchschnitten berechnet werden. So gelangen Sie zum selben Ergebnis wie wenn Sie von vorne herein den Durchschnitt der gesamten Verteilung berechnet hätten.

$$\begin{aligned} \text{arith. Mittel (aM) d. Gesamtverteilung} &= \frac{\text{aM Verteilung 1} + \text{aM Verteilung 2}}{2} \\ &= \frac{1.450 + 2.750}{2} = 2.100 \text{ Euro} \end{aligned}$$

Berechnungen wie diese sind allerdings nur bei ähnlich großen und ähnlich gelagerten Verteilungen sinnvoll.

Der Median (Zentralwert)

Bei Datenverteilungen mit einigen wenigen Ausreißern ist der Median als Mittelwert besser geeignet als das arithmetische Mittel. Der Median liegt in der Mitte der Daten, wenn diese der Größe nach geordnet sind. Die Hälfte der Beobachtungswerte ist kleiner, die andere Hälfte größer als der Median. Der Median wird z.B. oft in Zusammenhang mit Einkommensverteilungen verwendet. Dabei wird auch vom mittleren Einkommen oder dem Medianeinkommen gesprochen. Es gibt zwei leicht unterschiedliche Berechnungsformeln, je nachdem ob es sich um eine ungerade oder eine gerade Anzahl von Daten handelt.

Bei einer **ungeraden Anzahl** von Daten (W_1 bis W_n) wäre der Median jener Wert, der genau in der Mitte liegt. Wenn der Wert in der Mitte aufgrund der großen Anzahl von Werten nicht einfach durch Abzählen eruiert werden kann, wird am besten nach folgender Formel vorgegangen (Werte aus dem Beispiel von oben): Zur Anzahl der Werte (n) wird eins hinzu gezählt und anschließend wird diese, jetzt gerade Anzahl, halbiert.

$$\text{Median} = E_{\frac{n+1}{2}} = E_{\frac{7+1}{2}} = E_4 = 1.700 \text{ €}$$

Bei einer **geraden Anzahl** von Daten (E_1 bis E_8) ist der Median definiert als der Durchschnitt jener zwei Werte, die in der Mitte liegen:

Mit 1.750 Euro liegt er nur 50 Euro über dem arithmetischen Mittel oder dem Median, welche sich aus den Werten E_1 bis E_7 ergeben.

$$\text{Median (Medianeinkommen)} = \frac{1}{2} (E_{\frac{n}{2}} + E_{\frac{n}{2}+1}) = \frac{1}{2} (E_{\frac{8}{2}} + E_{\frac{8}{2}+1}) = \frac{1}{2} (E_4 + E_5) = 1.750 \text{ €}$$

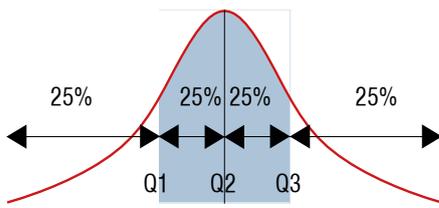
Hinweis: Im Gegensatz zum arithmetischen Mittel ist der Median nicht so gut für mathematische Berechnungen geeignet.

Beispiel: Wenn bei der Analyse des geschlechtsspezifischen Einkommensunterschieds nur der Median der Einkommen der Frauen und der Median der Einkommen der Männer bekannt sind, kann aus diesen beiden Werten nicht der Gesamtmedian errechnet werden.

Streuungsmaße – Quantile – Quartile

Mittelwerte allein können einen Sachverhalt oft nicht vollständig beschreiben. Es ist auch wichtig zu wissen, wie die Werte um den Mittelwert herum verteilt sind. Manchmal sind auch gerade Randverteilungen von Interesse, da es sich dabei etwa um Personen handelt, die besonders von bestimmten Problemen betroffenen sind (z.B. die Personen mit den niedrigsten Einkommen). Nur Mittelwerte anzuschauen reicht in diesen Fällen nicht aus. Hier ist es notwendig auch statistische Maße anzugeben, die die Streuung und Verteilung der Daten in weiterer Folge besser abbilden.

Quantile sind bestimmte Schwellenwerte, die die Verteilung untergliedern. Sie werden oft im Zusammenhang mit Einkommens- oder Vermögensverteilungen verwendet.



Quartile sind die am häufigsten verwendeten Quantile. Sie teilen die Beobachtungswerte in vier gleich große Gruppen ein. Als Quartile werden die Werte Q_1 , Q_2 und Q_3 bezeichnet. 25% aller Werte liegen unter Q_1 , 25% zwischen Q_1 und Q_2 , wieder 25% zwischen Q_2 und Q_3 und 25% über Q_3 . Der Median ist nichts anderes als das mittlere Quartil, das die Verteilung in zwei gleich große Hälften teilt. Der Quartilabstand ($Q_3 - Q_1$) wird auch als Streuungsmaß verwendet.

Grafik: erfundenes Beispiel

Beispiel: Von der Statistik Austria wird das 1. Quartil der Nettojahreseinkommen der weiblichen Pensionisten im Jahr 2013 mit 9.844 Euro angegeben, jenes der Männer lag bei 15.041 Euro. Demnach lagen jeweils 25% aller Nettojahreseinkommen unterhalb dieser beiden Werte.

http://www.statistik.at/web_de/statistiken/soziales/personen-einkommen/jaehrliche_personen_einkommen/020056.html

Gewichtetes (gewogenes) arithmetisches Mittel

Das gewichtete (gewogene) arithmetische Mittel wird angewandt, wenn Sie bei der Berechnung eines arithmetischen Mittels wichtige Daten betonen und den Einfluss weniger wichtiger Daten reduzieren möchten.

$$\text{Gewogenes arithmetisches Mittel} = g_1x_1 + g_2x_2 + \dots + g_nx_n$$

Die Summe der Gewichte g muss eins oder 100% ergeben.

$$g_1 + g_2 + \dots + g_n = 1$$

Die Gewichte werden so gewählt, dass wichtige Daten ein höheres Gewicht vorangestellt bekommen als unwichtigere und in der Folge das Ergebnis aussagekräftiger wird.

Das normale arithmetische Mittel ist genau genommen nur ein Spezialfall des gewichteten arithmetischen Mittels bei dem alle Gewichte gleich sind ($g = \frac{1}{n}$, also z.B. $\frac{1}{3}$ bei 3 Werten).

Beispiel:

Das gewichtete arithmetische Mittel wird in der Praxis oft verwendet, z.B. bei der Berechnung des Verbraucherpreisindex. Dabei werden die Warengruppen nach dem prozentuellen Anteil an den gesamten Konsumausgaben gewichtet. Für Lebensmittel wird ein höherer Anteil aufgewendet als für Telefonapparate, wodurch sie höher gewichtet werden.

Nahrungsmittel: $g = 10,43183 \%$

Telefonapparate und Telefaxgeräte: $g = 0,15861 \%$

Der Modus

Eine weitere Kennzahl ist der Modus. Das ist jener Wert, der am häufigsten vorkommt. Er ist jedoch nur bei Häufigkeitsverteilungen mit einem einzigen Hoch-/ oder Tiefpunkt (eingipfelige Verteilung, z.B. Normalverteilung) eindeutig.

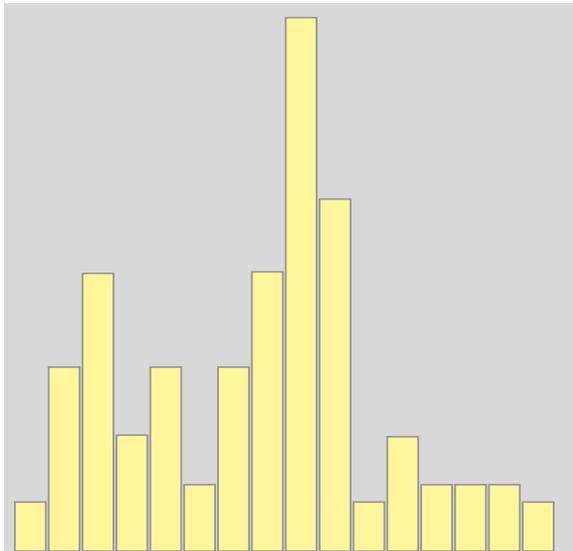


Abbildung: Zweigipflige Verteilung (erfundenes Beispiel)

Beispiel: Zufriedenheitsbefragung

Die meisten Befragten geben an, mit dem öffentlichen Verkehrsmittelangebot sehr zufrieden zu sein.

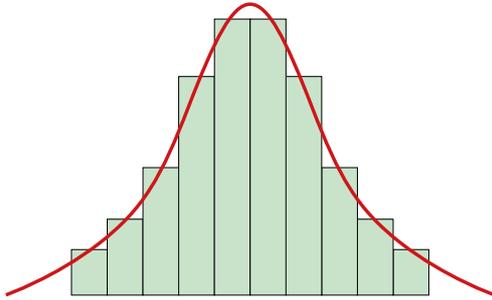
So wie der Median ist auch der Modus nicht gut für mathematische Berechnungen geeignet. Bei extrem flachen Verteilungen kann es vorkommen, dass es überhaupt keinen Modus gibt (wenn jeder Wert nur einmal vorkommt). Manchmal ist es sinnvoll, auch den 2.- oder 3.-häufigsten Wert anzugeben. Liegen diese weit entfernt vom häufigsten Wert, liegt eine mehrgipflige Verteilung vor.



Abbildung: Flache Verteilung mit mehreren Hochpunkten (erfundenes Beispiel)

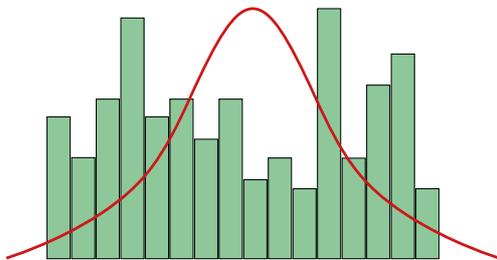
Welcher Mittelwert ist denn nun der „Richtige“?

Das kann leider nicht so einfach beantwortet werden und hängt von der jeweiligen Fragestellung ab. Die Verteilung der Daten hat entscheidende Auswirkungen auf die Wahl des Mittelwertes. Schon aus diesem Grund ist es sinnvoll, ein Histogramm zu erstellen.



Eine Normalverteilung liegt vor, wenn sich die Verteilung der Ausprägungen einer Gauß'schen Glockenkurve annähert. In diesem Fall ist das arithmetische Mittel gut geeignet, da es jenen Wert wiedergibt, bei dem die meisten Ausprägungen vorhanden sind (Hochpunkt der Gauß'schen Glockenkurve).

Grafik: erfundenes Beispiel



Liegt eine Datenverteilung mit hohen Randverteilungen (viele hohe und/oder niedrige Werte) vor, erweist sich das arithmetische Mittel als wenig aussagefähig, da es entweder in eine Richtung verzerrt wird, oder auf einen Wert fällt, wo es nur wenige Ausprägungen gibt.

Grafik: erfundenes Beispiel

In welchem Fall ist welcher Mittelwert der Richtige?

Ausgangslage	geeigneter Mittelwert
Die Daten liegen in einer Verteilung nahe einer Normalverteilung vor.	Arithmetisches Mittel, Median und Modus sind gut geeignet.
In der Verteilung gibt es extreme Ausreißer.	Median besser geeignet als arithmetisches Mittel
In der Verteilung gibt es mehrere Hoch-/Tiefpunkte (mehrgipflige Verteilung).	Modus (+ eventuell zweit- und dritthäufigste Werte) angeben. Andere Mittelwerte mitunter wenig aussagekräftig.
Der Mittelwert soll ein wirklicher Wert der Verteilung sein.	Andere Mittelwerte als das arithmetische Mittel besser geeignet
Man möchte mathematische Berechnungen mit dem Mittelwert durchführen.	Arithmetisches Mittel gut geeignet Modus und Median nicht so gut geeignet

Tipp: Oft kann es auch hilfreich sein, sich mehrere verschiedene Mittelwerte anzuschauen, um ein Gefühl für die Daten bzw. die Sachlage zu bekommen.

Indizes und Quoten

Der Begriff „Index“ hat zwei unterschiedliche Bedeutungen.

Einerseits sind Indizes aus mehreren Statistiken und Indikatoren zusammengesetzte Maßzahlen, die dazu verwendet werden, auch Aussagen über komplexe Fragestellungen treffen zu können. Mit ihrer Hilfe werden z.B. Vergleiche von Personengruppen oder Regionen, aber auch wiederholte Vergleiche über einen Zeitraum hinweg durchgeführt.

Beispiel:

Der sogenannte „Gleichstellungsindex Arbeitsmarkt“ des AMS wurde entwickelt, um die Gleichstellung von Frauen und Männern am Arbeitsmarkt rasch überprüfen und beobachten zu können. Es fließen insgesamt 30 Indikatoren in den Gesamtindex ein, wobei neben den Themenfeldern Arbeit und Einkommen auch die Themenfelder Bildung und Familie als Schnittstellen zum Arbeitsmarkt berücksichtigt werden:

http://www.forschungsnetzwerk.at/downloadpub/AMS_Gleichstellungsindex_AM_Endbericht.pdf

Andererseits ist ein Index eine Möglichkeit, die Veränderung einer Kennzahl anzugeben. Dafür wird der Wert eines frei wählbaren Jahres meist auf den Wert 100 gesetzt (Basisjahr) und anschließend die Veränderung in Bezug auf dieses Basisjahr angegeben. Für die Analyse ist es sinnvoll, sowohl prozentuelle Veränderungen (Index) als auch die zugrundeliegenden Absolutwerte zu betrachten.

Eine Quote gibt den Anteil einer Teilgesamtheit an einer Grundgesamtheit an (z.B.: Arbeitslosenquote, Stimmenanteil bei Wahlen, Teilzeitquote, Frauenquote, Steuerquote). Je nachdem was als Nenner gewählt wird, drückt die Quote etwas anderes aus. Nicht alles, was als Quote bezeichnet wird ist strenggenommen auch eine Quote.

Zum Beispiel wird zur Berechnung der Erwerbstätigenquote in Wien die Erwerbstätigen in Wien durch die Einwohner Wiens dividiert. Ein großer Teil der Erwerbstätigen in Wien wohnt jedoch außerhalb Wiens wodurch der Zähler des Bruches nicht Teil des Nenners ist und damit keine Quote im eigentlichen Sinn vorliegt.

Umgangssprachlich wird anstelle des Begriffes „Quote“ oft der Begriff „Rate“ verwendet (Arbeitslosenrate), der jedoch im Gegensatz zur Quote, die sich immer auf einen bestimmten Zeitpunkt bezieht, eigentlich eine Veränderung über einen bestimmten Zeitraum beschreibt (z.B: Wachstumsrate, Inflationsrate).

Die Veränderung einer Quote wird mithilfe von Prozentpunkten angegeben. Meistens werden bei der Berechnung der Erwerbstätigenquote (auch Beschäftigungsquote) alle Beschäftigten gleich gezählt, egal ob sie 10 Stunden oder 40 Stunden pro Woche arbeiten. Für die Berechnung von Gender- und Equalitydaten empfiehlt es sich, die Beschäftigungsverhältnisse auch in Vollzeitäquivalente umzurechnen.

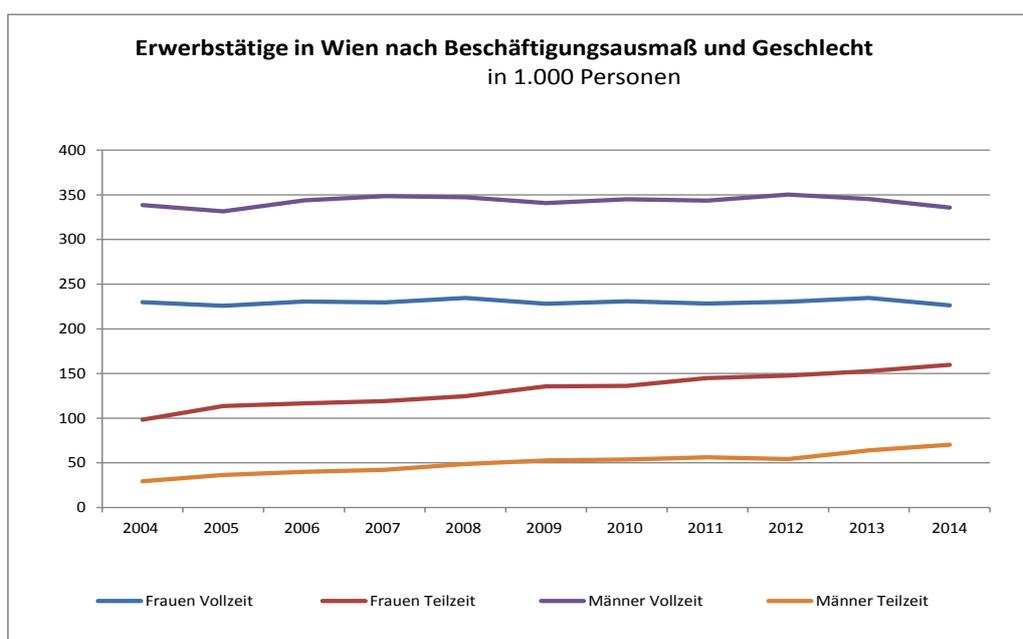
Beispiel:

Erwerbstätige in Wien nach Beschäftigungsausmaß und Geschlecht / in 1.000 Personen

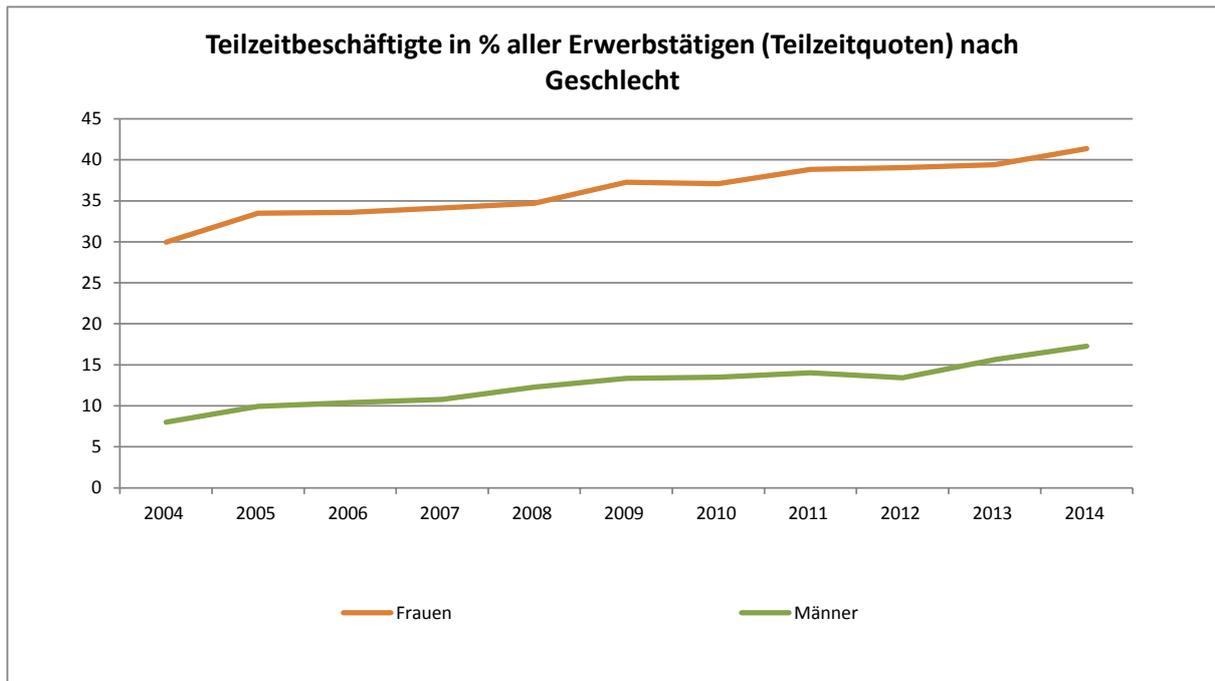
Jahr	Frauen				Männer			
	Vollzeit	Teilzeit	Teilzeit- quote in %	Veränderung in Prozentpunkten ggü. Vorjahr	Vollzeit	Teilzeit	Teilzeit- quote in %	Veränderung in Prozentpunkten ggü. Vorjahr
2004	229,7	98,2	29,9		338,5	29,4	8,0	
2005	225,8	113,7	33,5	3,5	331,5	36,5	9,9	1,9
2006	230,5	116,6	33,6	0,1	343,9	39,9	10,4	0,5
2007	229,6	119	34,1	0,5	348,6	42,2	10,8	0,4
2008	234,7	124,7	34,7	0,6	347,3	48,7	12,3	1,5
2009	228,1	135,6	37,3	2,6	340,9	52,6	13,4	1,1
2010	230,9	136,2	37,1	-0,2	345	53,8	13,5	0,1
2011	228,2	145	38,9	1,8	343,5	56,1	14,0	0,5
2012	230,4	147,7	39,1	0,2	350,4	54,3	13,4	-0,6
2013	234,5	152,6	39,4	0,4	345,2	63,9	15,6	2,2
2014	226,4	159,8	41,4	2,0	335,8	70,2	17,3	1,7

Quelle: Statistik Austria - Mikrozensus Arbeitskräfteerhebung

In den beiden folgenden Abbildungen werden die obigen Daten grafisch veranschaulicht. Die erste Grafik bezieht sich auf die Absolutzahlen der Erwerbstätigen, während die zweite die Entwicklung der Teilzeitquoten wieder gibt.



Quelle: Statistik Austria – Mikrozensus Arbeitskräfteerhebung



Quelle: Statistik Austria – Mikrozensus Arbeitskräfteerhebung

Folgende Schlüsse lassen sich daraus ziehen:

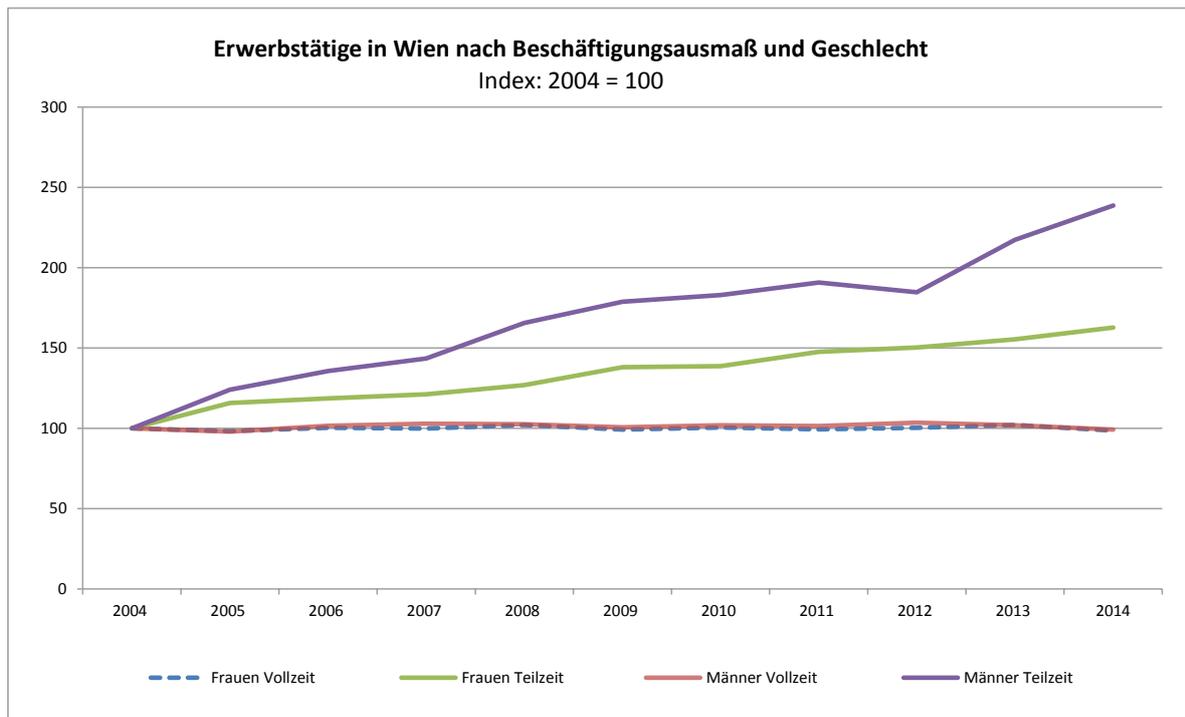
- Die Zahl der Teilzeit arbeitenden Männer ist zwischen 2004 und 2014 um 40.800 gestiegen. Die Teilzeitquote der Männer hat sich seit dem Jahr 2004 um 9,3 Prozentpunkte erhöht und sich dadurch mehr als verdoppelt.
- Die Zahl der Frauen, welche Teilzeit beschäftigt sind, ist zwischen 2004 und 2014 um 61.600 auf 159.800 gestiegen. Die Teilzeitquote der Frauen ist damit um 11,4 Prozentpunkte auf 41,4% gestiegen.

Als Indizes dargestellt, sehen dieselben Daten folgendermaßen aus. Als Basisjahr wurde das Jahr 2004 gewählt. Quoten werden üblicherweise nicht als Index dargestellt.

Erwerbstätige in Wien nach Beschäftigungsmaß und Geschlecht. Index: 2004 = 100

Jahr	Frauen		Männer	
	Vollzeit	Teilzeit	Vollzeit	Teilzeit
2004	100,0	100,0	100,0	100,0
2005	98,3	115,8	97,9	124,1
2006	100,3	118,7	101,6	135,7
2007	100,0	121,2	103,0	143,5
2008	102,2	127,0	102,6	165,6
2009	99,3	138,1	100,7	178,9
2010	100,5	138,7	101,9	183,0
2011	99,3	147,7	101,5	190,8
2012	100,3	150,4	103,5	184,7
2013	102,1	155,4	102,0	217,3
2014	98,6	162,7	99,2	238,8

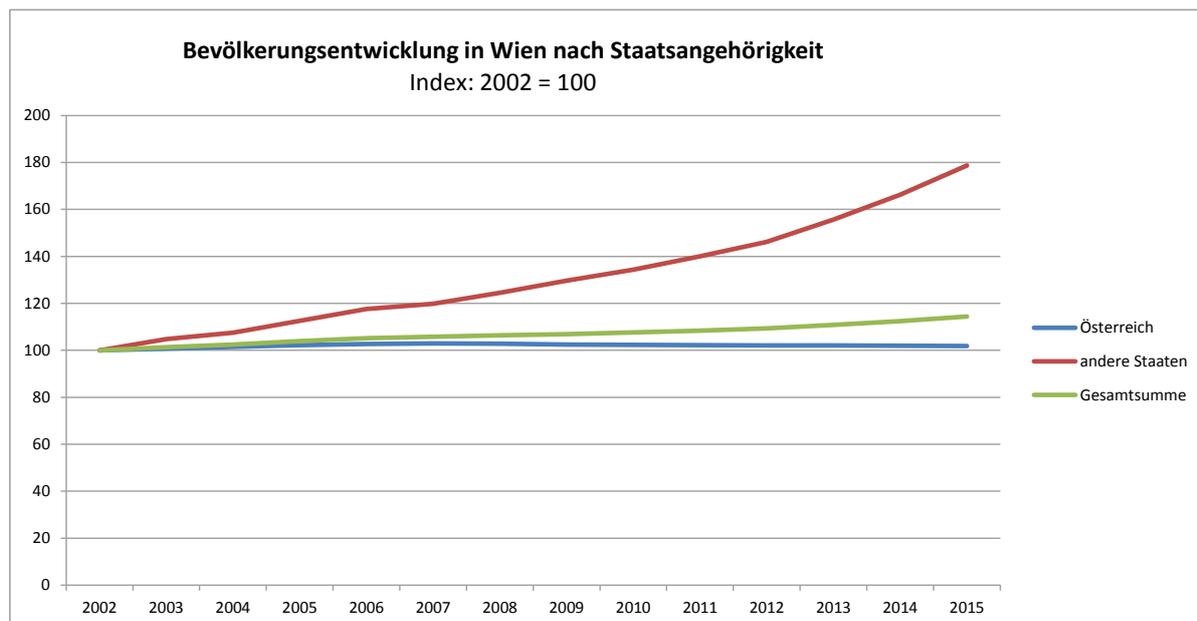
Quelle: Statistik Austria - Mikrozensus Arbeitskräfteerhebung



Quelle: Statistik Austria – Mikrozensus Arbeitskräfteerhebung

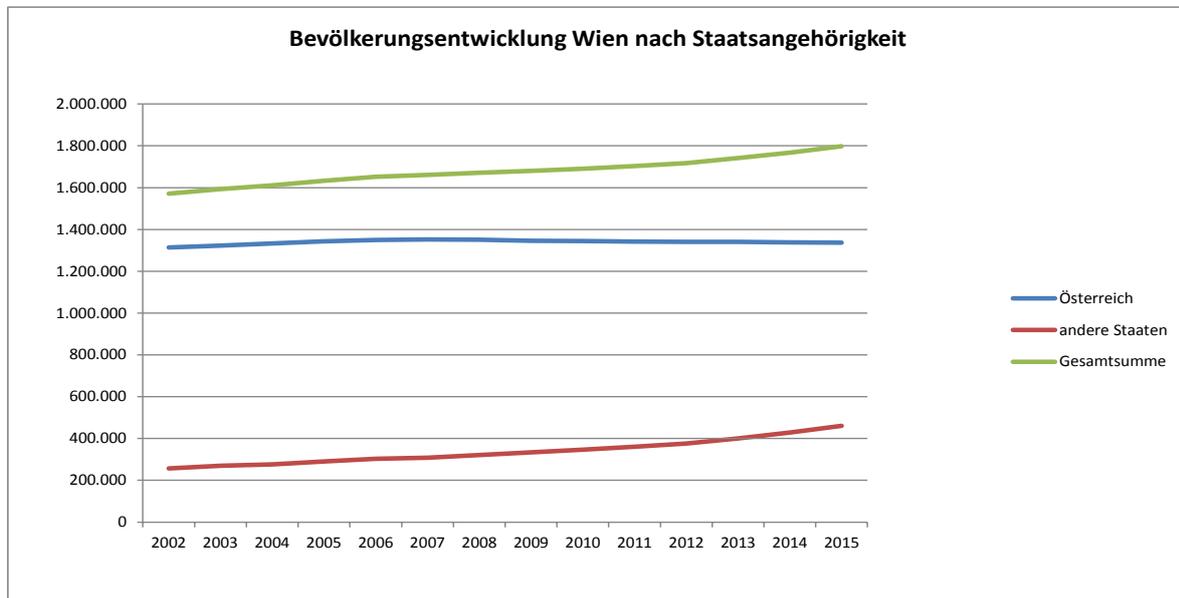
- Hier ist sehr schön zu erkennen, dass bei beiden Geschlechtern die Anzahl der Teilzeit arbeitenden Personen zwischen 2004 und 2013 stark zugenommen hat. Bei den Männern war diese Entwicklung noch deutlich ausgeprägter, was allerdings auf das niedrigere Ausgangsniveau zurückzuführen ist:
 - Zwischen 2004 und 2014 ist die Anzahl der Frauen in Teilzeitbeschäftigung um über 60% gestiegen.
 - Die Zahl der teilzeitbeschäftigten Männer ist im selben Zeitraum um annähernd 140% Prozent gestiegen.
- Der Einbruch der Teilzeitbeschäftigung bei den Männern im Jahr 2012 ist in der Darstellung als Index deutlicher erkennbar.

Beispiel: Bevölkerungsentwicklung Wiens 2002-2014 nach Staatsangehörigkeit



Quelle: Statistik Austria - Mikrozensus Arbeitskräfteerhebung

Das obige Diagramm könnte den Eindruck erwecken, Personen mit ausländischer Staatsangehörigkeit stellten die Mehrheit aller in Wien lebenden Personen dar. Sieht man sich die dahinterliegenden Absolutzahlen an, kommt man zu einem ganz anderen Schluss. Hier ist klar ersichtlich, dass die meisten Personen in Wien die österreichische Staatsbürgerschaft besitzen.



Quelle: Statistik Austria - Mikrozensus

Berechnung einer durchschnittlichen Wachstumsrate – geometrisches Mittel

Wenn Sie für einen bestimmten Zeitraum wissen wollen, wie stark etwas im Durchschnitt gestiegen oder gesunken ist, benötigen Sie das geometrische Mittel.

Beispiel: Um wie viel Prozent sind die BesucherInnenzahlen durchschnittlich pro Jahr gestiegen?

Jahr	Anzahl der BesucherInnen gesamt	Geschlecht		jährliche Wachstumsrate im Vergleich zum Vorjahr (BesucherInnen gesamt)	Wachstumsfaktor (BesucherInnen gesamt)
		Frauen	Männer		
2010	20	10	10		
2011	28	15	13	0,4	1,4
2012	35	19	15	0,3	1,3
2013	30	21	11	-0,1	0,9
2014	32	26	9	0,1	1,1
2015	37	30	7	0,2	1,2

Quelle: erfundenes Beispiel.

Die Wachstumsrate beschreibt den Zuwachs einer Größe in einem Zeitraum und wird berechnet als aus der Differenz des Wertes des jeweiligen Jahres und des Vorjahres dividiert durch den Wert des Vorjahres. Die Wachstumsfaktoren erhalten Sie, indem Sie zur Wachstumsrate 1 addieren oder indem Sie einfach den Wert des Folgejahres durch den Wert des vorangegangenen Jahres dividieren.

Multiplizieren Sie den Ausgangswert mit den Wachstumsfaktoren aller Jahre, erhalten Sie den Endwert:

$$20 \cdot 1,4 \cdot 1,3 \cdot 0,9 \cdot 1,1 \cdot 1,2 = 37$$

Würden Sie den jährlichen Wachstumsfaktor mithilfe des arithmetischen Mittels berechnen, kämen Sie auf folgendes Ergebnis:

$$\frac{1,4 + 1,3 + 0,9 + 1,1 + 1,2}{5} = 1,15$$

Als Prozentzahl ausgedrückt also 15% Wachstum durchschnittlich pro Jahr. Wenn Sie daraufhin die Probe durchführen und den Anfangswert fünf Mal mit dem errechneten Wachstumsfaktor von 1,15 multiplizieren, erhalten Sie:

$$20 \cdot 1,15 \cdot 1,15 \cdot 1,15 \cdot 1,15 \cdot 1,15 = 40,2$$

Es sollte aber 37 herauskommen! Der durchschnittliche Wachstumsfaktor muss also kleiner als 1,15 sein und kann mit dem arithmetischen Mittel nicht berechnet werden!

Den richtigen Wert für den durchschnittlichen Wachstumsfaktor erhalten Sie mithilfe des geometrischen Mittels. Für die Berechnung des geometrischen Mittels ziehen Sie die n-te Wurzel aus dem Produkt der Wachstumsfaktoren, in unserem Beispiel also die 5-te Wurzel aus den fünf Wachstumsfaktoren.

$$\text{Geometrisches Mittel} = \sqrt[n]{WF_1 \cdot WF_2 \cdot \dots \cdot WF_n} = \sqrt[5]{1,4 \cdot 1,3 \cdot 0,9 \cdot 1,1 \cdot 1,2} = 1,13$$

$$\text{Probe: } 20 \cdot 1,13 \cdot 1,13 \cdot 1,13 \cdot 1,13 \cdot 1,13 = 37$$

Der durchschnittliche prozentuelle BesucherInnenzuwachs beträgt folglich

Das obige Beispiel verdeutlicht außerdem anschaulich warum es wichtig ist, sich nicht nur die Frage nach dem gesamten durchschnittlichen BesucherInnenzuwachs zu stellen, sondern die Daten auch nach dem Geschlecht aufzuschlüsseln. Die Anzahl der männlichen Besucher ist nämlich im selben Zeitraum gesunken! Das Wachstum ist ausschließlich auf den starken Anstieg bei den Frauen zurückzuführen.

Daten interpretieren

Dateninterpretation – warum?

Bei der Interpretation von Daten geht es darum, auf die ursprüngliche Fragenstellung zurückzukommen und auf Basis von Zahlen und Daten Antworten zu finden. Die Aufgabe besteht darin, Ergebnisse aus den dargestellten Zahlen und Indikatoren herauszulesen und sprachlich zu „übersetzen“. Mit dem nötigen Fachwissen können die Daten verständlich beschrieben werden sowie Vermutungen und Fragestellungen auf Basis der statistischen Informationen überprüft und beantwortet werden. Werden diese Ergebnisse in einem Text verarbeitet, können stimmige Schlussfolgerungen vermittelt werden und auch von fachfremden Personen richtig eingeschätzt und genutzt werden.

Dateninterpretation – was heißt das konkret?

Daten können anhand von aufeinander aufbauenden Schritten interpretiert werden, die zu einem kontinuierlichen Erkenntnisfortschritt führen. Zunächst ist es wichtig, die Daten korrekt und nachvollziehbar zu beschreiben. Die nächsten Schritte bestehen darin, Vergleiche anzustellen sowie mögliche Zusammenhänge und Ursachen zu erkennen.

Schritt 1: Sachverhalt erklären

Hier gilt es, den Ergebnissen einen Rahmen zu geben, indem nochmals auf die Datenquelle und die Erhebungsmethoden Bezug genommen werden kann und die dargestellten Fakten erklärt werden. Besonders interessante, eventuell auch überraschende Informationen sollten gesondert herausgestellt werden.

Konkret bedeutet das, gezielt auf dargestellte Informationen hinzuweisen:

- Welche Fragen sollen beantwortet werden, worum geht es bei der Statistik?
- Von wem wurden die Daten erfasst/beauftragt?
- Was wird dargestellt (Jahreszahlen oder Stichtage? Prozent-, Index- oder Absolutwerte?) und welche Einheiten (Jahre, Kilogramm, Stunden, Euro etc.) sind angegeben?
- Welche Merkmale wurden erfasst (Alter, Geschlecht, Geburtsland etc.) und welche Gender- und Equalitäsaspekte können aus den Daten herausgelesen werden?
- Was sagen die Zahlen aus? Welche konkreten Aussagen können gemacht werden?

Tipp: Nehmen Sie Missverständnisse vorweg und stellen Sie gleich zu Beginn klar, was mit den vorhandenen Daten NICHT erklärt werden kann! Beispiel: „Die Einkommen der Selbständigen sind in der Statistik nicht enthalten, es werden lediglich die Einkommen der unselbständig Beschäftigten abgebildet.“

Schritt 2: Entwicklungen beschreiben, Veränderungen erkennen, Vergleiche anstellen

Im nächsten Schritt werden Entwicklungen beschrieben und beobachtbare Tendenzen und Trends festgehalten. Zu diesem Zweck ist es hilfreich, weiteres Datenmaterial zu recherchieren (siehe Kapitel „Wie komme ich zu Daten?“), Entwicklungen im Zeitverlauf zu betrachten oder Vergleiche mit anderen Beobachtungseinheiten (Personengruppen, Regionen etc.) anzustellen. Dabei ist es wichtig zu klären, ob die Daten vergleichbar sind und keine statistischen oder inhaltlichen Gründe gegen eine Vergleichbarkeit sprechen (siehe Kapitel „Daten erheben - was nun?“). Sind Einschränkungen bei der Vergleichbarkeit vorhanden, sollte ebenfalls darauf hingewiesen werden.

Konkret bedeutet das, über einzelne Informationen hinausgehende Aussagen herauszuarbeiten:

- Welche Schlussfolgerungen können in Bezug auf die Fragestellung gemacht werden?
- Welche Entwicklung lässt sich aus den Zahlen herauslesen: zeigt sich eine eindeutige Entwicklung in eine bestimmte Richtung? Gibt es auffällige Unregelmäßigkeiten oder widersprüchliche Tendenzen?
- Ist die Entwicklung bei allen Personen gleichermaßen erfolgt, oder betrifft die Veränderung nur ein Geschlecht, eine Altersgruppe oder lediglich Personen mit/ohne Migrationshintergrund (etc.)?
- Sind die Informationen plausibel und entsprechen den Erwartungen bzw. Erfahrungswerten?
- Können die Informationen mit anderen Erhebungen verglichen werden und kommen diese zu ähnlichen Ergebnissen oder gibt es markante Unterschiede? Können Besonderheiten durch die Erhebungsmethode oder Informationen außerhalb der Statistik (z.B. Gesetzesänderungen, regionale Besonderheiten etc.) erklärt werden?

Hinweis: Nicht immer ist es einfach, Veränderungen richtig einschätzen zu können. So ist bei der Beurteilung von Abweichungen zwischen zwei Vergleichsgrößen oder bei Veränderungen im Zeitverlauf keine generelle Aussage möglich, ab welchem Wert es sich um eine geringfügige oder gravierende Abweichung handelt bzw. ab wann es sich um eine kleine oder große Veränderung handelt. Bei Quoten (oder anderen zusammengesetzten Kennzahlen) ist zudem vorab eine Aufschlüsselung erforderlich, welche Größe sich überhaupt verändert hat: Der Zähler? Der Nenner? Oder Beides?

Beispiel: Die Teilzeitquote wird als Anteil der Erwerbstätigen in Teilzeit an allen Erwerbstätigen berechnet. Eine Steigerung der Quote kann dadurch zustande kommen, dass Erwerbstätige von Vollzeitarbeit zu Teilzeitarbeit wechseln, ohne dass die Zahl der Erwerbstätigen sich ändert (Anzahl der teilzeitbeschäftigten Personen im Zähler steigt, Zahl der Erwerbstätigen im Nenner bleibt gleich). Es könnte aber auch sein, dass die Zahl der Teilzeitbeschäftigten gleich geblieben ist, aber die Zahl der Vollzeitbeschäftigten gesunken ist (Zähler bleibt gleich, aber die Zahl der erwerbstätigen Personen im Nenner sinkt). Dies hätte ebenfalls ein Steigen der Teilzeitquote zur Folge.

Es besteht dabei einerseits die Herausforderung, relevante Entwicklungen zu erkennen und andererseits nicht den Fehler zu begehen, aus „einer Mücke einen Elefanten zu machen“. Diese Einschätzung ist basierend auf Fachkenntnissen zu treffen bzw. mit Expertinnen und –experten der jeweiligen Fachabteilung sowie den Gender- und Equalityexpertinnen und -experten im Magistrat zu beraten.

Achtung: Insbesondere bei Stichprobenerhebungen ist zu beachten, dass die Hochrechnung auf die Grundgesamtheit immer mit einer gewissen Unschärfe behaftet ist und nur innerhalb bestimmter Fehlergrenzen möglich ist (umso kleiner die Stichprobe, umso höher ist der Fehlerbereich). Daher kann es sein, dass scheinbar stattgefundenen Änderungen innerhalb dieser Fehlergrenzen liegen und in der Realität gar keine Veränderung stattgefunden hat. Eine Möglichkeit, diese Unschärfe sichtbar zu machen, ist die Darstellung von hochgerechneten Stichprobenergebnissen nicht als einzelne Zahl, sondern als Vertrauensbereich („Konfidenzintervall“), der durch die statistischen Fehlergrenzen berechnet werden kann.

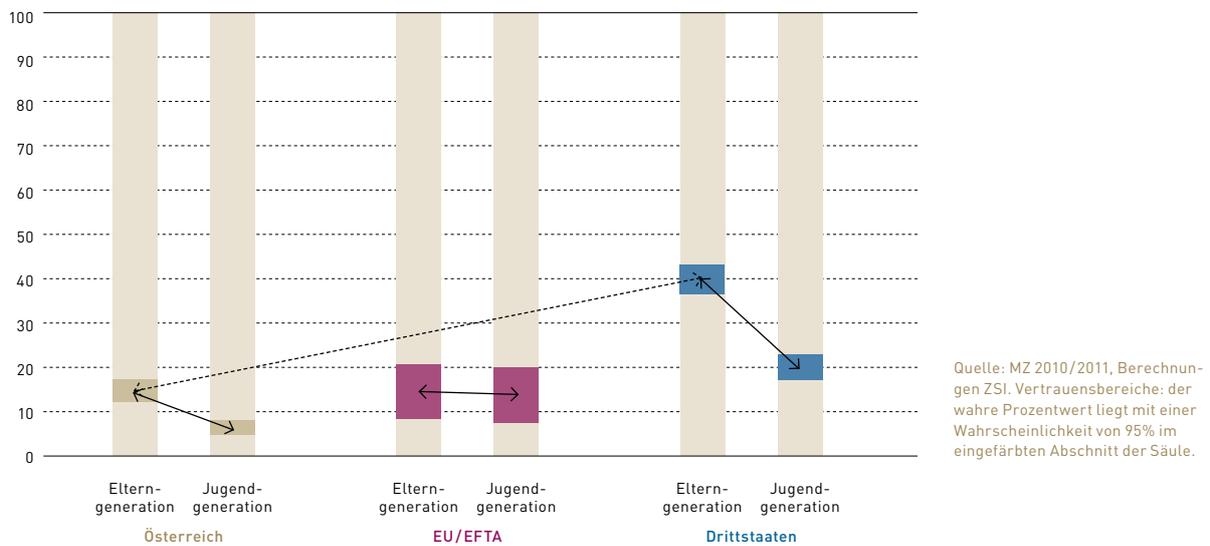
Konfidenzintervall: Darstellung von Ergebnissen aus Stichproben, um die Häufigkeit eines bestimmten Merkmals festzustellen. Das Konfidenzintervall (der Vertrauensbereich) gibt den Bereich an, in dem sich die untersuchte Häufigkeit mit einer bestimmten Wahrscheinlichkeit befindet.

Quelle: 3. Wiener Integrations- & Diversitätsmonitor, S. 220

<https://www.wien.gv.at/menschen/integration/pdf/monitor-2014.pdf>

Beispiel: Die Abbildung zeigt den aus einer Stichprobe hochgerechneten Anteil der Bevölkerung, der höchstens einen Hauptschulabschluss hat und unterscheidet dabei zwischen der Elterngeneration (das sind die heute 45- bis 59-Jährigen) und der Jugendgeneration (den 15- bis 29-Jährigen) nach Herkunft. Die farbigen Bereiche zeigen die Konfidenzintervalle der hochgerechneten Ergebnisse. Je breiter die Bereiche sind, umso größer sind die statistischen Unsicherheiten, die sich aus kleinen Fallzahlen bei der Stichprobe ergeben. Unterschiede zwischen den verschiedenen Gruppen werden deutlich, wenn sich die Vertrauensbereiche nicht überlappen. So zeigt sich sowohl bei Personen ohne Migrationshintergrund als auch bei Personen mit Herkunft aus Drittstaaten, dass der Pflichtschulanteil bei der Jugendgeneration deutlich niedriger ist als bei der Elterngeneration. Zwischen der älteren und der jüngeren Generation konnte demnach tatsächlich ein Bildungsfortschritt erzielt werden.

Anteil mit maximal Hauptschulabschluss nach Staatsangehörigkeit bzw. Geburtsland der Eltern



Quelle: Stadt Wien, Monitoring Integration und Diversität 2009-2011, S. 61.

Lesbeispiel: Der Anteil der Elterngeneration (das sind die heute 45- bis 59-Jährigen) ohne Migrationshintergrund (Österreich) und maximal Hauptschulabschluss liegt zwischen 12% und 16%.

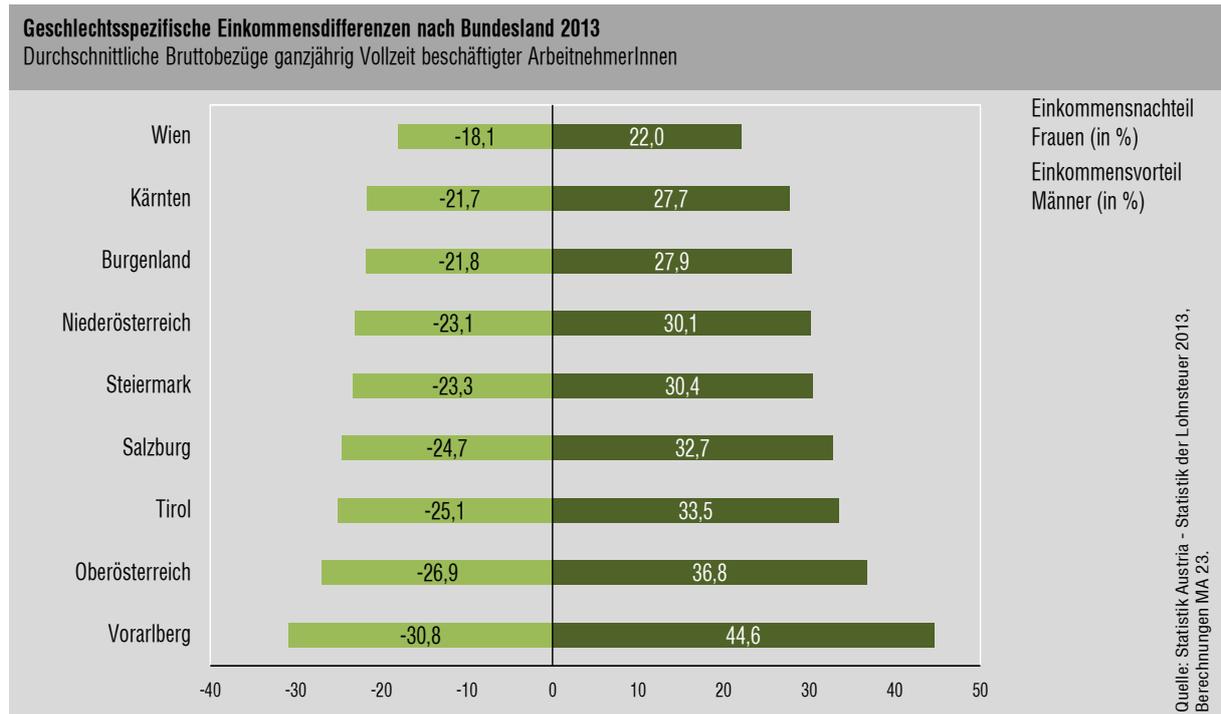
Tipp: Absprache mit den Statistikexpertinnen und -experten der MA 23.

Schritt 3: Zusammenhänge erkennen, Ursachen ermitteln

Zuletzt werden Schlüsse aus den Daten gezogen, Ursachen und Zusammenhänge beschrieben und auf die damit verbundenen Konsequenzen hingewiesen. Je differenzierter die Daten vorliegen, umso besser können bestimmte Umstände auf ihre Ursachen „abgeklopft“ werden. Ist es beispielsweise möglich, die Einkommensunterschiede zwischen Frauen und Männer nach verschiedenen Merkmalen darzustellen, ermöglicht dies Aussagen über den Einfluss einzelner Faktoren, wie z.B. Arbeitszeit, Branche, Bildung, etc.

Wichtig ist die Daten in einen größeren Kontext zu stellen und die Rahmenbedingungen abzuklären. Sinkt der Männeranteil an den Erwerbstätigen kann das bedeuten, dass Frauen sich allgemein stärker am Arbeitsmarkt beteiligen. Diese Entwicklung kann aber beispielsweise auch ein spezifisches Ergebnis eines Wandels der Wirtschaftsstruktur sein, in dessen Folge Branchen schrumpfen, in denen vorwiegend Männer beschäftigt waren.

Manchmal ist es auch sinnvoll, Daten in einen anderen Bezug zu stellen bzw. Verhältniszahlen auch einmal umzudrehen. Ein Beispiel ist die Berechnung des Gender Pay Gaps, das den prozentuellen Einkommensnachteil der Frauen, gemessen an den Einkommen der Männer, misst. Umgekehrt kann aber auch der Einkommensvorteil der Männer, gemessen am Einkommen der Frauen, berechnet werden. Die Aussage ist ebenfalls richtig, aber die Einkommensunterschiede werden in einem anderen Licht dargestellt, was wiederum auch Einfluss auf die daraus gezogenen Schlüsse haben kann. Die Einkommen der Männer könnten demgemäß als „Besserstellung“ oder aber die Einkommen der Frauen als „Benachteiligung“ interpretiert werden.



Quelle: Statistik Austria - Mikrozensus

Lesebeispiel: Die Einkommen der Frauen liegen in Wien um 18,1% unter den Einkommen der Männer. Umgekehrt betrachtet sind die Einkommen der Männer in Wien durchschnittlich um 22,0% höher als jene der Frauen.

Sollte die Statistik markante Mängel haben, widersprüchliche Informationen vorliegen oder Unklarheiten zu den Ergebnissen bestehen, sollten diese ebenfalls erläutert werden.

Konkret geht es also darum, über die dargestellten Informationen hinausgehend Schlüsse zu ziehen, die für die jeweilige Fragestellung behilflich sein können:

- Kann ein Zusammenhang mit anderen Informationen zu diesem Themenbereich hergestellt werden? Durch welche anderen Daten können Schlussfolgerungen bestätigt oder in Frage gestellt werden?
- Sind Ursachen für die beschriebenen Entwicklungen bekannt?
- Welche Zusatzinformationen wären notwendig, um präzisere Aussagen zur dargestellten Entwicklung bzw. deren Ursachen treffen zu können? Welche Fragen bleiben offen?
- Kann eine „Gesamtaussage“ aus den Daten getroffen werden?

Tipp: Wenn Sie Ihre Ergebnisse möglichst verständlich vermitteln möchten und diese auch für fachfremde bzw. weniger zahlenaffine Personen nützlich sein sollen, ist es sinnvoll die Informationen zu relevanten Daten und Indikatoren auch sprachlich auszuformulieren und grafische Darstellungen durch ein Lesebeispiel zu ergänzen.

Daten präsentieren

Erstellung von Grafiken

Grafiken sind ein geeignetes Werkzeug, um statistische Informationen übersichtlich und schnell verständlich darzustellen. Im Optimalfall benötigt die oder der Betrachtende nur wenige Momente um die vermittelten Inhalte zu verstehen.

Achtung: Bei der Erstellung statistischer Grafiken muss immer die korrekte Informationsvermittlung im Vordergrund stehen. Diese darf nicht zugunsten von Effekten oder Dekorationen undeutlich gemacht oder gar verfälscht werden. Darstellungen können, ja sogar sollen, optisch ansprechend gestaltet werden - die Übersichtlichkeit und Klarheit der Informationsvermittlung muss aber immer oberste Priorität haben.

Kriterien, die eine gute Grafik erfüllen muss:

- ansprechend gestaltet
- leicht und schnell verständlich
- übersichtlich
- aussagekräftig

Zu erwähnen ist die Schriftgröße. Bei PowerPoint wird eine Mindestgröße von 16 Punkt empfohlen.

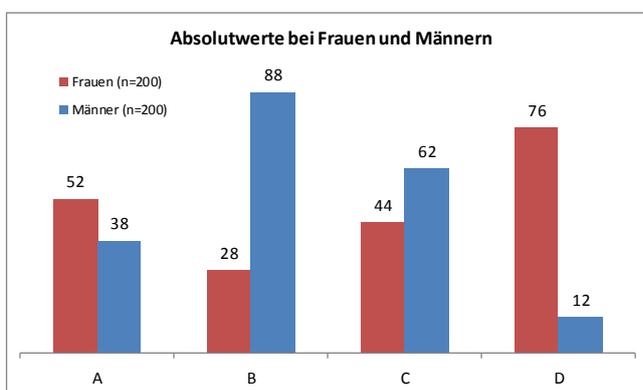
An dieser Stelle der Verweis auf den Styleguide für AutorInnen:

<http://www.wien.gv.at/styleguide/autor>

Wahl des Diagramms auf Grund der Art der Daten

Besonderes Augenmerk muss auf die Wahl des passenden Diagrammtyps gelegt werden. Nicht jeder Diagrammtyp ist für alle Darstellungen geeignet. Hier werden die gängigsten Diagrammtypen und ihre möglichen Anwendungen beschrieben.

Säulen- und Balkendiagramme



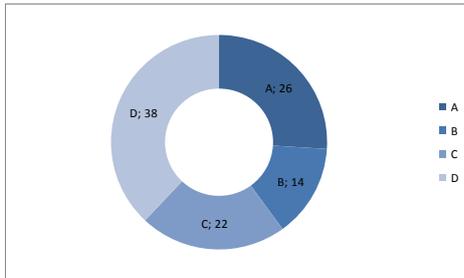
Frauen und Männer haben bei den Optionen A bis C unterschiedliche Ergebnisse bzw. Präferenzen. Bei D sind es 76 Frauen, aber nur 12 Männer. Umgekehrt sieht es bei B aus, wo es 88 Männer aber nur 28 Frauen sind. Es wird deutlich, dass eine getrennte Darstellung der Geschlechter sinnvoll ist um gezieltere Schlüsse ziehen oder Handlungen für Männer und Frauen setzen zu können. Quelle: erfundenes Beispiel.

In Säulen- und Balkendiagrammen werden die Häufigkeiten von verschiedenen Merkmalsausprägungen durch Säulen bzw. Balken dargestellt. In einem Säulendiagramm sind die länglichen Flächen nebeneinander angeordnet, in einem Balkendiagramm übereinander. Die Entscheidung, ob Sie ein Säulen- oder Balkendiagramm wählen, hängt z.B. davon ab, ob es sich um eine Zeitreihe (eher Säulen, die von links nach rechts gelesen werden), oder um eine ab- bzw. aufsteigende Reihung (eher Balken, die von oben nach unten bzw. umgekehrt gelesen werden) handelt. Aber auch Faktoren, wie etwa die Länge der Beschriftungen bei den Ausprägungen können zur Entscheidung beitragen.

Mögliche Anwendung:

- Vergleiche zweier oder mehrerer Werte
- Absteigende oder aufsteigende Reihungen
- Zeitreihen (jedoch sollte es sich aus Gründen der Übersichtlichkeit um maximal zwei Ausprägungen handeln)
- Vergleich von Prozentverteilungen durch Stapelung (besser geeignet als Tortendiagramme)

Torten- und Donutdiagramme



Quelle: erfundenes Beispiel

Torten- bzw. Donutdiagramme werden im Regelfall dafür verwendet, um die verschiedenen Anteile eines Ganzen (100 Prozent) darzustellen. Donutdiagramme bieten den Vorteil, dass in den leeren Platz in der Mitte die Fallzahl geschrieben werden kann. Torten- bzw. Donutdiagramme können als grafische Abwechslung zu Säulen- bzw. Balkendiagrammen verwendet werden, sie bieten aber keinen Mehrwert gegenüber den zuletzt genannten. Im Gegenteil: Größenverhältnisse können durch gerade Linien (Säulen oder Balken) schneller und besser eingeschätzt werden als durch die Winkel und Flächen von Kreissektoren. Speziell bei Zeitreihen oder sonstigen Vergleichen sind Säulen- bzw. Balkendiagramme

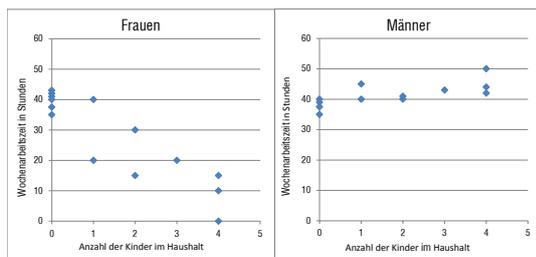
zu bevorzugen, da sie platzsparend sind und sich Veränderungen leicht ablesen lassen.

Mögliche Anwendung:

- Aufschlüsselung eines Ganzen in Prozent (aber: nicht ratsam bei Zeitreihen oder Vergleichen)

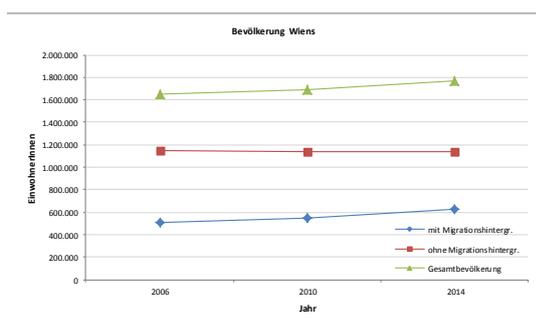
Punkt- und Liniendiagramme

Punkt- und Liniendiagramme kommen etwa dann zur Anwendung, wenn es zwei Variablen gibt und es für einen x-Wert verschiedene y-Werte (und/oder umgekehrt) geben kann. In solchen Fällen sind weder Balken- noch Liniendiagramme sinnvoll.



Quelle: erfundenes Beispiel

In Abbildung 3 sieht man jeweils ein Diagramm in dem auf der x-Achse die Anzahl der Kinder im Haushalt und auf der y-Achse die Anzahl der gearbeiteten Stunden (Erwerbstätigkeit) aufgetragen ist. Beispielsweise sieht man bei den Frauen, dass eine Frau mit zwei Kindern 30 Stunden arbeitet und eine andere Frau, die ebenfalls zwei Kinder hat, 15 Stunden. Es gibt für einen x-Wert (2 Kinder), zwei verschiedene y-Werte (40 Stunden, sowie 20 Stunden).



Quelle: Statistik Austria

Für die Darstellung von Zeitreihen sind Liniendiagramme eine gute Wahl, besonders wenn es mehr als zwei Ausprägungen gibt. Bei dem Beispiel in Abbildung 4 handelt es sich um die Veränderung dreier Ausprägungen im Zeitverlauf. Dieses Beispiel könnte auch mittels Balkendiagramm dargestellt werden. Bei einem solchen wäre die Nachvollziehbarkeit der Veränderungen im Zeitverlauf aber bereits unübersichtlich, da es sich hier um drei Ausprägungen handelt. Das Heranziehen eines Liniendiagramms macht die Veränderungen schnell ersichtlich. Um bei Liniendiagrammen klarzustellen, dass nicht jeder Punkt der Linie

eine Ausprägung darstellt (was selten der Fall ist), bietet es sich an die Werte auf der Linie durch Markierungsoptionen hervorzuheben.

Mögliche Anwendungen:

- Zeitreihen, speziell wenn mehrere Parameter im Zeitverlauf dargestellt werden sollen
- Darstellung von Durchschnitts (in Kombination mit einem Säulendiagramm)
- Darstellung von Quoten sowie Raten

Quote bezeichnet einen Prozentanteil einer Gesamtmenge oder -anzahl, wobei sich die Quote auf Bestandsgrößen beziehen kann. Im Gegensatz zur Rate, die sich immer auf eine Zeiteinheit bezieht, ist die Quote immer auf die Größe zu einem Zeitpunkt bezogen.

Kartographische Darstellungen



Durch Einfärbung abgegrenzter Bereiche (Gemeindebezirke oder kleinere Ebenen) in unterschiedlichen und/oder abgestuften Farbtönen, lässt sich leicht verständlich und übersichtlich darstellen wie stark ein bestimmter Parameter in einem Gebiet ausgeprägt ist.

Mögliche Anwendung:

- Bei geografisch zuordenbaren Ausprägungen, insbesondere bei kleinräumigen Analysen.

Diagrammerstellung und -formatierung

Bei Beschriftung und Layout müssen die notwendigen Informationen in der Grafik enthalten sein. Sie sollten aber auf unnötige Elemente verzichten, um die Aufmerksamkeit auf die Information zu fokussieren.

Nicht fehlen sollten:

- Überschrift
- Datenquelle
- Maßeinheit
- Legende (nach Möglichkeit direkt bei den Parametern)
- Datenbeschriftung (vorzugsweise direkt bei der Ausprägung. Falls dies aus Platzgründen nicht möglich ist, sollte ein Hilfsgitter verwendet werden, damit sich die Größenordnung der Ausprägung einschätzen lässt; siehe Abbildung 4)

Tipps:

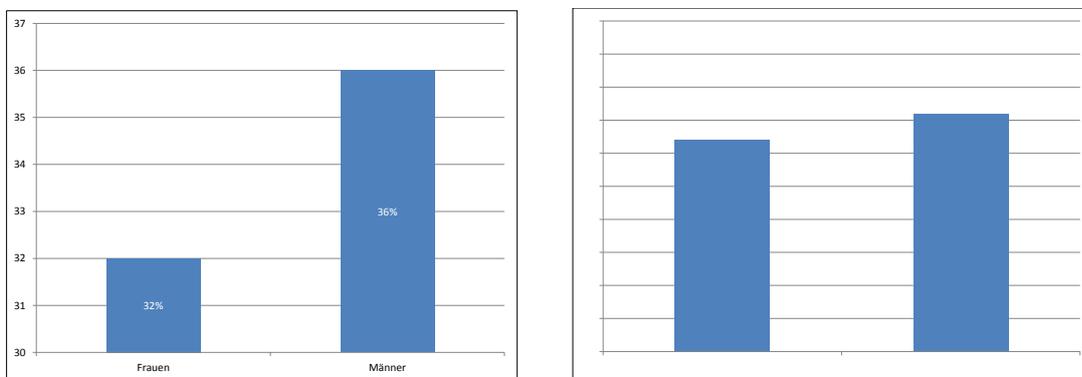
- Durch das Hinzufügen der Werte direkt bei den Ausprägungen in der Grafik, lässt sich, speziell bei Säulen- bzw. Balkendiagrammen, eine Achse inklusive Achsenbeschriftung einsparen.
- Durch Angabe des Absolutwertes, etwa in der Überschrift oder in der Legende, ersparen Sie sich eine grafische Darstellung desselben (siehe Abbildung 1).
- Die Legende können Sie, bei ausreichend Platz, direkt in die Grafik schreiben, oder platzsparend in einen freien Bereich des Diagramms verschieben.
- Achten Sie darauf, dass Beschriftungen einen Kontrast zum Hintergrund haben um gut lesbar zu sein.
- Zeigen Sie Diagramme, vor Verwendung, Personen ohne Vorwissen zu dem Thema, um die Verständlichkeit abzuklären.

- Bedenken Sie bei der Farbwahl eine mögliche rot/grün-Blindheit beim Zielpublikum
- Wenn anstatt eines Farbausdruck lediglich Graustufen gedruckt werden sollen, empfiehlt es sich zur besseren Erkennbarkeit von Liniendiagrammen unterschiedliche Linienarten (strichliert, punktiert etc.) zu verwenden.

Achtung:

- Vermeiden Sie Verzerrungen, Schatten- sowie 3-D-Effekte. Diese haben keinen Zusatznutzen und lenken von der eigentlichen Information ab oder führen zu Verzerrungen der selbigen.
- Stellen Sie Achsen immer ab dem Nullpunkt dar. So vermeiden Sie Verzerrungen durch abgeschnittene Achsen. Dies ist auch aus der Gender- und Equality-Sicht wichtig, denn je nach Wahl der Achsenabschnitte kann eine (ungewollte) Überbetonung von Unterschieden zwischen zwei Merkmalsausprägungen vermittelt werden.

Beispiel: Beiden Grafiken liegen dieselben Daten zugrunde. Die Unterschiede zwischen den Geschlechtern wirken in der linken Grafik (Achsenabschnitt 30 bis 38) aber deutlich größer als in der rechten Grafik (Achse beginnt mit dem Nullpunkt). Die rechte Grafik vermittelt ein realistischeres Bild.



Quelle: erfundenes Beispiel

Erstellung von Tabellen

Bei der Erstellung und Formatierung von Tabellen ist es wichtig die präsentierten Daten verständlich und übersichtlich aufzubereiten, sowie optisch ansprechend zu formatieren. Dadurch lässt sich das Interesse des angesprochenen Personenkreises maximieren.

Einwände man habe nicht genug Platz um gender- und diversitätsgerechte Tabellen darzustellen, lassen sich leicht abwehren: es genügt die wichtigsten Informationen in kleinen Tabellen darzustellen oder relevante Informationen herauszugreifen und komplexe, große Gesamttabellen mit allen Merkmalsausprägungen im Anhang anzuführen. In knappen Werbefoldern oder Presstexten kann darauf verwiesen werden wo detaillierte Informationen zu finden sind.

Beachten sollten Sie:

- Tabellenbezeichnung: der Name der Tabelle muss unbedingt mit den tatsächlich dargestellten Daten übereinstimmen. Im Zweifel sollte die Tabellenbezeichnung lieber etwas länger, dafür aber korrekt sein.
- Tabellenköpfe: kurze aber klare Bezeichnungen. Durch eine farbliche Hinterlegung und „fett“-Formatierung können Sie den Tabellenkopf vom Rest der Tabelle abheben.
- Einheiten im Tabellenkopf statt bei jedem Wert: Das Verschieben der Maßeinheiten in den Tabellenkopf, trägt zu einem sauberen Layout der Tabelle bei und verbessert den Lesefluss.
- Keine „darunter“-Rubriken: stattdessen beispielsweise: Insgesamt / Frauen / Männer. Die Führung einer „darunter“-Rubrik vermittelt den Eindruck als würde es sich um einen Sonderfall handeln.
- Da prinzipiell alle personenbezogenen Daten nach Geschlechtern getrennt erfasst und auch dargestellt werden, gehören die Merkmalsausprägungen (Frauen, Männer sowie der Gesamtwert) immer in Tabellenspalten und nicht in -zeilen. Denn nur so können weitere Merkmale (Migrationshintergrund, Alter,

- Zuzüge, Wegzüge...) nach Geschlechtern getrennt betrachtet werden (siehe Tabelle 1).
- Summenzeilen bzw. -spalten: Beschriftung und Werte „fett“ formatieren. Dadurch lässt sich schneller erkennen welche Bereiche Summen, Salden, etc. sind.
 - Aufschlüsselungen können Sie einrücken: Die Beschriftungen zu Aufschlüsselungen von Summen, Salden, etc. werden leicht eingerückt formatiert um den Bezug zu der darüberstehenden Kategorie klar hervorzuheben. Beispiel: In der unten stehenden Tabelle lässt sich schnell erkennen, dass die Wanderungssaldo die Differenz aus Zuzügen und Fortzügen ist.
 - Bei Aufschlüsselung nach Gruppen (etwa „wohnhafte Personen mit EU- Staatsangehörigkeit“) sollte auch der prozentuelle Anteil an der Grundgesamtheit (in diesem Fall die gesamte „Wohnbevölkerung“, also die Summe aller in Wien wohnhaften Personen aller Staatsangehörigkeiten) angegeben werden. Wird lediglich die Absolutzahl ausgewiesen, ist nicht auf den ersten Blick erkennbar, in welcher Relation diese zur restlichen Bevölkerung steht. Um ein vollständiges Bild zu erhalten, geben Sie den Anteil der EU-BürgerInnen an der ausländischen Wohnbevölkerung als auch der EU BürgerInnen an der Gesamtbevölkerung an.
 - Aber auch die Angabe des prozentuellen Anteils der fremden Staatsangehörigkeiten an der Untergruppe „AusländerInnen“ könnte bei den BetrachterInnen der Tabelle zu einem verzerrten Bild über die tatsächlichen prozentuellen Anteile von bestimmten Personengruppen an der Gesamtbevölkerung führen.
 - Tausenderpunkte dienen der schnelleren Lesbarkeit großer Zahlen.
 - Nachkommastellen sollten bei jenen Zahlen hinzugefügt werden, bei denen dadurch eine Präzisierung des Wertes erreicht werden kann. Nachkommastellen sollten jedoch weggelassen werden, wenn sie sinnlos sind, etwa wenn eine Ausprägung nur aus ganzen Zahlen bestehen kann.
 - Zahlen rechtsbündig, Texte linksbündig oder zentriert formatieren. Dies trägt zu einem ansprechenden Layout bei.

Beispiel einer formatierten Tabelle

Bevölkerung in Wien 2013			
* Bevölkerungsstand zum 1.1.2014			
Demographische Informationen	Insgesamt	Frauen	Männer
Bevölkerung *	1.766.746	916.150	850.596
Bevölkerung mit fremder Staatsangehörigkeit *	428.213	212.443	215.770
Bevölkerung mit Migrationshintergrund *	629.541	324.410	305.131
Geburtenbilanz	2.359	15	2.344
Lebendgeborene	18.691	8.826	9.865
Gestorbene	16.332	8.811	7.521
Wanderungsbilanz	22.711	10.798	11.913
Zuzüge	94.034	43.056	50.978
Fortzüge	71.323	32.258	39.065

Quelle: Statistik Austria, Berechnung MA 23

Beispiel einer Formatierungsoptimierung

Insgesamt	Frauen	Männer	Frauen	Männer
200	90	110	45,0 %	55,0 %
155	125	30	80,6 %	19,4 %
3578	2000	1578	55,9 %	44,1 %



Insgesamt	Frauen	Männer	Frauen	Männer
	absolut		in %	
200	90	110	45,0	55,0
155	125	30	80,6	19,4
3.578	2.000	1.578	55,9	44,1



Quelle: erfundenes Beispiel

Literaturtipps

European Institute for Gender Equality (2013): Gender Equality Index Report. Eigenverlag

Online unter: <http://eige.europa.eu/sites/default/files/documents/Gender-Equality-Index-Report.pdf>

Hametner, Kristina (2005): Kommentar zum Artikel „Gendersensible Statistik“. In: Statistische Mitteilungen der Stadt Wien, Heft 1,2/2005, Wien, 4-6

Online unter: <http://www.wien.gv.at/statistik/pdf/genderstatistik.pdf>

Makkonen, Timo (2007): Europäisches Handbuch zu Gleichstellungsdaten. Europäische Kommission, Eigenverlag

Online unter: <http://bookshop.europa.eu/de/europaeisches-handbuch-zu-gleichstellungsdaten-pbKE7606381/>

Stadt Wien (2011): Handbuch „Gender Mainstreaming – leicht gemacht“. Eigenverlag, Wien

Online unter: <https://www.wien.gv.at/menschen/gendermainstreaming/grundlagen/handbuch.html>

Stadt Wien (2011): Leitfaden für geschlechtergerechtes Formulieren und eine diskriminierungsfreie Bildsprache. Eigenverlag, Wien

Online unter: <https://www.wien.gv.at/medien/service/medienarbeit/richtlinien/leitfaden-gender/>

Stadt Wien (2014): Gendersensible Statistik: Lebensrealitäten sichtbar machen. Statistik Journal Wien 2/2014. Eigenverlag, Wien

Online unter: <https://www.wien.gv.at/statistik/pdf/genderstatistik-broschuere.pdf>

Stadt Wien (2014): Wiener Gleichstellungsmonitor. Eigenverlag, Wien

Online unter: <https://www.wien.gv.at/menschen/frauen/stichwort/gleichstellung/gleichstellungsmonitor/>

Stadt Wien (2014): 3. Wiener Integrations- und Diversitätsmonitor 2011-2013. Eigenverlag, Wien

Online unter: <https://www.wien.gv.at/menschen/integration/grundlagen/monitoring/>

Statistics Sweden (1996): Engendering Statistics – A Tool for Change. Eigenverlag, Schweden

Online unter: http://www.scb.se/statistik/_publikationer/LE0202_1996A01_BR_X93%C3%96P9601.pdf

UNECE (2010): Developing Gender Statistics: A Practical Tool. Eigenverlag, Genf

Online unter: https://www.unece.org/fileadmin/DAM/stats/publications/Developing_Gender_Statistics.pdf

Wroblewski, Angela/Leitner, Andrea/Steiner, Peter (2005): Gendersensible Statistik – Vom Sex-Counting zur Genderanalyse“. In: Statistische Mitteilungen der Stadt Wien, Heft 1,2/2005, Wien, S. 7-45

Online unter: <http://www.wien.gv.at/statistik/pdf/genderstatistik.pdf>

Impressum:

Eigentümerin und Herausgeberin:

MD-OS Dezernat Gender Mainstreaming, Rathaus, Stiege 8, Hochparterre, Zimmer 104, 1082 Wien

Telefon: (+43 1) 4000-75114, E-Mail: gm@md-os.wien.gv.at

Text- und Redaktionsteam: Christian Fendt (MA 23), Theodora Manolakos (MA 17), Jana Schultheiß (MA 24), Eva Seibold (MA 23), Edith Waltner (MA 23) unter Mitarbeit von Dominik Sandner (MD-OS, Dezernat Gender Mainstreaming)

Layout: Kurt Wurscher (MD-OS)

Wien, Dezember 2015