

Digitale *tools* zum Einsatz von Statistiken im Politikunterricht

Exploration, Analyse und Visualisierung von sozialwissenschaftlichen Daten

Florian Weber-Stein und Joachim Engel

1. Gesellschaftliche Datafizierung und Politische Bildung

Bereits im Vorwort seines Evaluationsberichts (Bundesgesundheitsministerium 2022) betonte der „Corona-Expertenrat“, dass „eine ausreichende und stringente begleitende Datenerhebung [fehlte], die notwendig gewesen wäre, um die Evaluierung einzelner Maßnahmen oder Maßnahmenpakete zu ermöglichen“ (ebd., S. 11). Exemplarisch wurde hier noch einmal deutlich, was mehr als zwei Jahre Pandemieerfahrung uns eindrücklich gelehrt haben: Das Abwägen von Risiken und Wahrscheinlichkeiten auf der Basis statistischer Daten hat für die Organisation des gesellschaftlichen Zusammenlebens und die Legitimierung politischer Entscheidungen eine fundamentale Bedeutung erlangt (Deutsche Arbeitsgemeinschaft Statistik 2021). Die Bereitschaft zur Einhaltung beschlossener Maßnahmen zur Eindämmung des Virus verlangt in demokratischen Gesellschaften von Bürger:innen ein Verständnis von Fakten, Zusammenhängen und darauf abgestimmten Regeln. Damit Maßnahmen wirksam sind, müssen die Entscheidungsträger:innen transparente und überzeugende Erklärungen für ihre Entscheidungen abgeben. Dies stellt hohe Anforderungen an Fähigkeiten zur Kommunikation statistischer Informationen seitens der Gesundheitsbehörden und Medien sowie an die Bürger:innen, diese Botschaften nachzuvollziehen und kritisch zu reflektieren.

Dieser Zusammenhang entspricht einem generellen Trend der *Datafizierung* fast aller Bereiche der Gesellschaft. Aus der gestiegenen Bedeutung von Daten für politische Entscheidungen folgt eine Verantwortung des Bildungssystems, die Bürger:innen mit den Fähigkeiten auszustatten, Daten und Statistiken kompetent zu ‚lesen‘, kritisch zu hinterfragen und ggf. eigene Datenrecherchen und Erhebungen zu initiieren. Die



Prof. Dr. Florian Weber-Stein

Professor für Politikwissenschaft und Politikdidaktik, Institut für Sozialwissenschaften, Pädagogische Hochschule Ludwigsburg



Prof. i.R. Dr. Joachim Engel

Professor für Mathematik und Mathematikdidaktik, Institut für Mathematik II, Pädagogische Hochschule Ludwigsburg

vom Stifterverband im Januar 2021 initiierte und von zahlreichen Institutionen und Personen unterstützte Data-Literacy-Charta (<https://www.stifterverband.org/charta-data-literacy>) fordert eine Vermittlung von Datenkompetenzen in allen Bildungsbereichen.

Die Autoren vertreten die Auffassung, dass zu diesem Zweck eine Integration von statistischer und sozialwissenschaftlicher Bildung von Bedeutung ist. Im Folgenden werden zunächst Konturen der durch Datafizierung gekennzeichneten Gegenwartsgesellschaft nachgezeichnet (2.), bevor der Diskussionsstand zum Verhältnis von statistischer und politischer Bildung knapp rekonstruiert wird (3.). Der Diskussionsstand spiegelt weder die aktuelle Verfügbarkeit von Daten noch Möglichkeiten ihrer dynamischen und handlungsorientierten Nutzung durch digitale *tools* wider. Vor dem Hintergrund einiger grundsätzlicher Überlegungen (4.) stellen wir drei *tools* zum Einsatz von Statistiken im Politikunterricht vor (5.). Ein Plädoyer für den Einbezug von Datenexploration in den sozialwissenschaftlichen Unterricht beschließt unsere Ausführungen (6.).

2. Datafizierung – Konturen und Trends

Innerhalb des allgemeinen Trends der Datafizierung lassen sich drei Entwicklungslinien differenzieren (Ridgway 2016, S. 528-549): Einerseits ist die Gegenwart gekennzeichnet durch ein *hohes Maß an Datentransparenz und -zugänglichkeit*: Zahlreiche Daten, bspw. von Behörden, Ministerien und statistischen Ämtern, liegen heute in öffentlich zugänglicher Form vor (*open data*). Sie können auf den Websites der Organisationen eingesehen, oftmals durch Apps dynamisch und individuell visualisiert, durchsucht und in verschiedenen Formaten heruntergeladen werden. Etliche zivilgesellschaftliche Akteure betreiben professionelles Daten-Monitoring, um Entwicklungen in den Bereichen soziale Ungleichheit, Entwicklung oder Demokratiequalität zu dokumentieren und stellen ihre Datensätze öffentlich zur Verfügung. Digitale Medien haben den Zugang zu Daten erheblich vereinfacht, es gibt kaum formelle Hürden oder Barrieren des Zugangs und diese eröffnen, den Idealen des *open data-movements* entsprechend, völlig neue Wege für das Engagement von Staatsbürger:innen in der Zivilgesellschaft.

Dieses positive Bild wird jedoch durch einen zweiten Trend erheblich eingetrübt: Neben *open data* gibt es nicht öffentlich zugängliche, *intransparente und komplexe Datensätze aus neuen Datenquellen (big data)*: Im Zuge der Internetnutzung, zunehmend aber auch der digitalisierten Erfassung von Bewegungen und Handlungen durch in die Lebenswelt integrierte Geräte (v.a. Smartphones) und Sensoren (z.B. Fitnesstracker), fallen in enormer Menge Daten als ‚Nebenprodukte‘ des Verhaltens an. Diese Daten können in vielfältiger Weise gekoppelt werden (z.B. Inhaltsdaten und Geodaten) und weisen eine hohe ‚Granularität‘ (Feinkörnigkeit) auf, die es u.U. ermöglicht, anhand von eigentlich anonymen Daten einzelne Individuen zu identifizieren. *Big data* fallen in der Regel bei wirtschaftlichen Akteuren an, den großen Telekommunikationsanbietern, Internet-Plattformen sozialer Netzwerke und Suchmaschinenbetreibern, sie sind dementsprechend i.d.R. nicht öffentlich und ihre Analyse verlangt hochspezialisierte

Fähigkeiten und enorme Rechenkapazitäten. Nur unter bestimmten, gesetzlich spezifizierten, Bedingungen müssen Betreiber die Daten gegenüber staatlichen Behörden offenlegen wie bspw. im Falle eines konkreten Verbrechens-Verdachts auf richterliche Anordnung hin (z.B. Funkzellenabfrage).

Nicht nur die Anzahl und Art der Daten hat sich verändert, sondern auch die Wege ihrer Darstellung und Kommunikation. Es bestehen heute – dies ist der dritte Trend – *interaktive Möglichkeiten der Datenvisualisierung*. Die Kommunikation datenbasierter Informationen ist zu einem eigenen journalistischen Genre geworden. Häufig gestützt auf Ideale des *open data-movements* impliziert Datenjournalismus nicht nur die Informationsrecherche in Datenbanken, sondern auch die Sammlung, Aufbereitung, Analyse und adressatengerechte Darstellung öffentlich zugänglicher Daten. Die Pandemie hat zu einer enormen Verbreitung dieses Typus der Berichterstattung geführt. Kurz nach Ausbruch der Pandemie zirkulierten in den Medien unter dem Titel „Flatten the curve“ animierte visuelle Aufbereitungen, welche Auswirkungen auf die Infektionskurve verschiedene Grade der Reduktion der sozialen Kontakte haben könnten (vgl. Grenier 2020).

Datenvisualisierung ist aber kein Privileg professioneller Akteur:innen in Journalismus, politischer Kommunikation oder Marketing, auch User:innen haben heutzutage vielfältige Möglichkeiten, sich Daten über Apps gemäß individuellen Vorlieben und Bedürfnissen darstellen zu lassen.

Der Frage, welche Bedeutung bei der Aufnahme statistischer Information der Form ihrer Visualisierung zukommt, widmet sich mittlerweile die eigenständige wissenschaftliche Disziplin der *visualization studies*. Befunde deuten darauf hin, dass der graphischen Aufbereitung eine hohe Bedeutung für die Vermittlung der Inhalte zukommt und insbesondere mathematisch wenig vorgebildete, ‚numerophobe‘ Personen über geeignete Visualisierung angesprochen werden können (vgl. Segel/Heer 2010, Boy et al. 2017).

3. Statistik und Politische Bildung – eine Bestandsaufnahme

In der Politischen Bildung¹ ist der Einsatz von Statistiken ein Thema von randständigem Interesse. Monografien zum Thema sucht man vergebens, viele der zentralen Lexika und Handbücher widmen dem Thema „Daten“ und/oder „Statistik“ keine Erwähnung: In dem von Kerstin Pohl und Wolfgang Sander 2022 in fünfter, grundlegend überarbeiteter Auflage herausgegebenen Standardwerk „Handbuch Politische Bildung“ finden sich im Kap. V „Methoden und Medien“ zwar Artikel zum Lernen in „Gesprächen“, mit „Texten“, „Bildern“, „Filmen“ und „narrativen Medien“, jedoch nicht zu „Daten und Statistiken“ – nicht einmal im Sachwortregister findet sich ein Eintrag. Hier wie andernorts dominiert unumstritten der Fokus auf den Text als Leitmedium; sofern andere Medienformen thematisiert werden, dann bevorzugt Bildmedien (zur Schulung von *visual literacy*, bspw. durch eine Karikaturenanalyse, finden sich zahlreiche unterrichtspraktische Handreichungen).

Die wenigen systematischen Beiträge zum Einsatz von Statistiken lassen zwei didaktische Ansätze erkennen: Auf der einen Seite stehen Ansätze, die statistische *Les-*

fähigkeit schulen wollen (Lach/Massing 2017), worunter neben der Fähigkeit zur Aufnahme und Wiedergabe von Informationen auch deren kritische Hinterfragung sowie eine Reflexion hinsichtlich des Informationswertes verstanden wird. Diese Fähigkeit soll durch ein an Prüfkriterien orientiertes, systematisches Lesen von Statistiken sowie das Erstellen von Tabellen und einfachen grafischen Darstellungen anhand von Daten vermittelt werden. Ziel ist es, den „Lernenden die Scheu vor dem Umgang mit Statistiken zu nehmen und sie schrittweise mit der kritischen Analyse von Statistiken vertraut zu machen“ (ebd., S. 21).

Auf der anderen Seite stehen Ansätze, die den Fokus nicht auf das Verstehen, sondern die *Erhebung und Konstruktion von Daten* legen und dementsprechend einen stärker handlungsorientierten und wissenschaftspropädeutischen Zugang verfolgen (vgl. Schattschneider 2010). Gegenstand des Unterrichts soll die Erstellung von Fragebögen und Datenerhebung (hierzu: Porst/Holthof 2014a, 2014b), zumindest aber die eigenständige Analyse vorliegender Datenbanken durch niedrigschwellige und eigens für die politische Bildungsarbeit entwickelte Datenanalyse-Software sein (z.B. GrafStat).

Beide Ansätze haben ihre Vor- und Nachteile. Für den ersten Ansatz spricht die Möglichkeit des unkomplizierten Einbezugs statistischer Quellen in klassische Unterrichtsszenarien und die Möglichkeit des systematischen (wiederholten) Übens statistischer Lesekompetenz. Gegen diesen Ansatz spricht allerdings die vorwiegend kognitive Inanspruchnahme und geringe Eigenständigkeit der Schüler:innen bei der vorstrukturierten Analyse. Für den zweiten Ansatz hingegen spricht der kritische Fokus auf die Frage, wie Daten erhoben werden, die Möglichkeit schülerorientierte Fragestellungen zu verfolgen und die Handlungsorientierung. Als Gegenargument wird jedoch nicht nur pragmatisch auf den hohen Zeitaufwand verwiesen, sondern auch systematisch die Befürchtung geäußert, dass die „großen methodischen Schwierigkeiten, die mit der Erstellung eines Fragebogens oder eines Interviewleitfadens sowie mit einer statistischen Auswertung [...] verbunden sind“ unterschätzt werden könnten: „Die Gefahr, dass ein unsachgemäßer Umgang zu Fehlkonzepten führt“ ist nach Einschätzung von Georg Weißeno größer als die anvisierte (aber unrealistische) „Zielvorstellung, in das forschende Denken und Handeln einzuführen“ (Weißeno 2010, S. 497f.).

Weißeno plädiert somit für eine Beschränkung auf die Vermittlung statistischer Lesefähigkeit in der Schule und will auch hier enge Grenzen ziehen. Er weist darauf hin, dass in der Schule ausschließlich beschreibende Statistik thematisiert werden könne (ebd., S. 498):

„Eine fundierte statistische Analyse erfordert umfangreiche Mathematikkenntnisse, die in der Schule auch in der Oberstufe nicht unbedingt vermittelt werden. Z.B. sind die probabilistische Testtheorie oder die multiple Regression (zu Recht?) keine Gegenstände des Mathematikunterrichts. [...] Korrelation, Signifikanztests, Reliabilitäten, Messmodelle etc. sind ohne statistische Kenntnisse verfrüht.“ (ebd., S. 496f.)

Der Nutzen beschreibender Statistik sei aber im Vergleich zu Texten eher gering. Die Mehrzahl der im Unterricht einsetzbaren Daten seien einfach und erlaubten „in den meisten Fällen eigentlich keine Rückschlüsse auf einen Ursache-Wirkungszusammen-

hang. Ein solcher steht aber meist im Zentrum politischer Debatten und Kontroversen“ (ebd., S. 496).

4. Statistik und Politische Bildung – Überlegungen zu einem zeitgemäßen Update

Sowohl die Unterscheidung zwischen kognitiven und handlungsorientierten didaktischen Ansätzen (a) als auch Weißenos generelle Bedenken gegenüber dem Einsatz von Statistiken im schulischen Politikunterricht (b) müssen im Licht der oben skizzenartig rekonstruierten Datenrevolution einer kritischen Neubewertung unterzogen werden.

Ad (a) Die Medialität von Statistiken hat sich grundlegend verändert. ‚Klassische‘ Ansätze statistischer Lesefähigkeit gingen vom Normalfall der abgedruckten Grafik oder Tabelle aus, die dem Interpreten oder der Interpretin als starre Form vorgegeben war und überdies stets in der Gefahr stand, bereits veraltete Daten zu präsentieren. In der Regel waren die Daten entweder tabellarisch angeordnet oder grafisch aufbereitet, Hinweise zur Kodierung von Variablen waren, sofern nicht mitgegeben, nur unter erheblichem Aufwand zu recherchieren, die Stichprobe sowie die Variablen waren nicht zu beeinflussen. – Demgegenüber sind Daten auf den Internetportalen zahlreicher Organisationen heutzutage nicht nur oftmals tagaktuell, sondern es kann auch zwischen der tabellarischen Ansicht und verschiedenen grafischen Visualisierungen gewechselt werden. Bei Interesse an einer umfassenderen Analyse können die Daten in verschiedenen Formaten exportiert werden. Hinweise zur Erhebung und Kodierung der Daten sind meistens unmittelbar verlinkt und u.U. können die zu untersuchende Stichprobe selbst festgelegt sowie die interessierenden Variablen individuell definiert werden. Anstelle von starren Grafiken können (Zeitreihen-)Daten im chronologischen Verlauf animiert werden und man kann sich von den Daten, etwas blumig formuliert, ‚Geschichten erzählen lassen‘ (Kosara/Mackinlay 2013). – Vor diesem Hintergrund wachsen der Schulung statistischer Lesefähigkeit ganz neue Dimensionen zu. Sie verliert das Starre und Formalistische und gewinnt Elemente des Spielerischen. Eine individualisierte und interessen geleitete Erkundung von Daten wird möglich, die kritische Befragung nach der Erhebung von Daten lässt sich meist ohne weiteres integrieren (notfalls per Email-Anfrage an die entsprechende Organisation). Kurzum: Der Gegensatz zwischen kognitiver Analyse und handlungsorientierter Erkundung schwächt sich ab.

Ad (b) Auch die von Weißeno vorgebrachten Bedenken gelten unter den neuen Bedingungen nur noch in abgeschwächtem Maße:

Zum einen hat sich durch *open data* die Anzahl und Verfügbarkeit großer und reichhaltiger (multivariater) Datensätze enorm vergrößert, aus denen mit Mitteln der deskriptiven Statistik politisch bedeutsame Informationen abgeleitet werden können. Die von Weißeno erwähnten, mathematisch anspruchsvollen, Hypothesen- und Signifikanztests, die Rückschlüsse einer kleinen Stichprobe auf eine Grundgesamtheit er-

lauben, haben angesichts der Verfügbarkeit großer Datensätze an Bedeutung verloren (Ridgway 2016, 546).

Zum anderen können durch moderne *Möglichkeiten der Datenvisualisierung* statistische Zusammenhänge, die vormals nur durch numerische Analyse aufgedeckt werden konnten, in Grundzügen visuell erschlossen werden.² Bspw. kann ein Zusammenhang (Korrelation) zwischen zwei Variablen in einem Streudiagramm erkannt und u.U. (je nach *software-tool*) eine Trendgerade eingezeichnet werden, ohne dass die mathematische Formel für die Berechnung bekannt oder präsent sein muss. Ebenso können Hintergrundvariablen ‚kontrolliert‘ werden, wenn aus einer größeren Population eine Stichprobe von Elementen mit vergleichbaren Eigenschaften bezüglich eines dritten Parameters ausgewählt und diese mit dem Rest der Population verglichen werden.

Schließlich können komplexe, untereinander zusammenhängende (multivariate) Daten dank innovativer Computergraphik visualisiert werden. Das Handling solcher Daten galt lange Zeit als anspruchsvoll, die Probleme im Umgang sind zu einem gewissen Grad eine Form ‚erlernter Hilflosigkeit‘. Nur wer auf die (mathematisch wesentlich einfacher handzuhabenden) bivariaten und linearen Zusammenhänge gepolt ist, schreckt vor der Analyse multivariater Daten zurück. Innerhalb der statistischen Lehr-Lernforschung gibt es erste Evidenz, dass statistisch „naive“ Nutzer:innen multivariate Daten mittels interaktiver Visualisierung adäquat erfassen und interpretieren können (vgl. Ridgway et al. 2007).

5. Drei *software-tools*: Hintergründe und Nutzungsmöglichkeiten

Wir verzichten aus Platzgründen an dieser Stelle auf eine lerntheoretische Begründung und die Skizze konkreter didaktischer Arrangements des Einsatzes von Statistik im sozialwissenschaftlichen Unterricht. Verviesen sei auf die Projektpublikationen (Nicholson et al. 2018, Ridgway 2022) und die Plattform *CivicStatMap* (<http://iase-web.org/islp/pcs>), die im Rahmen des Erasmus+ Projekts *ProCivicStat (Promoting civic engagement via explorations of evidence)* unter Leitung von Joachim Engel/PH Ludwigsburg (mit Partnern in Durham, Haifa, Paderborn, Porto und Szeged) entstanden sind. Auf der Plattform finden sich zahlreiche praxiserprobte Unterrichtsmaterialien, die sich für den Einsatz im sozialwissenschaftlichen Unterricht an Universitäten und Hochschulen sowie der gymnasialen Oberstufe sowie ggf. (bei entsprechender didaktischer Reduktion) der Sekundarstufe I eignen (Wassner/Prömmel 2022). Über eine Suchmaske können die Materialien nach Sprache, statistischen Inhalten, Themen, verwendeten *tools* und Schwierigkeitsgrad sortiert werden.

Im Folgenden stellen wir drei öffentliche zugängliche Web-Apps vor, die sich für einen niedrighschweligen und interaktiven Einbezug von sozialwissenschaftlichen Statistiken in den Politikunterricht eignen und (nach einer Einführung in statistische Grundkonzepte und die Funktionalitäten der Apps) eigenständige Datenexplorationen durch Schüler:innen ermöglichen. Wir stimmen Sören Torrau zu, dass unter „digitalen Vorzeichen“ (Torrau 2020, S. 243) Internet-Recherchen von Schüler:innen für die Unterrichtsgestaltung immer wichtiger werden und im Sinne der Schülerorientierung be-

deutsam sind. Neben der Fähigkeit, die „Gatekeeping-Funktion von Algorithmen durch breit aufgestellte Suchstrategien und kritisches Überprüfen ab[zuf]edern“ (ebd., S. 237) ist dazu auch eine basale statistische Alphabetisierung vonnöten.

5.1 Our World in Data

Hintergrund und Umfang: Das Datenportal *Our World in Data* (<https://ourworldindata.org>) wurde 2011 durch den Oxford-Ökonomen Max Roser begründet. Die Website ist über private Spenden und Stiftungen finanziert und verschreibt sich dem Ziel, Daten über die Gegenwartsgesellschaft und ihre zentralen Entwicklungen zu Aufklärungszwecken allgemein zugänglich zu machen. Die Website ist nach Themen-/Problemfeldern gegliedert („Articles by topic“). Jeder Artikel widmet sich einem zentralen globalen Problem – bspw. der demographischen Entwicklung, öffentlichen Gesundheit, Demokratie und Menschenrechten, Gewalt und Krieg sowie Bildung und Umweltthemen (und weiteren mehr).

Web-Oberfläche und Nutzung: Von Interesse für den Einbezug im Politikunterricht sind insbes. die *interaktiven Grafiken* („interactive charts“), die Daten in verschiedenen Diagrammformen aufbereiten und in den meisten Fällen auch historische Daten umfassen, so dass die Entwicklung von Trends in der Zeit beobachtbar ist. Über einen „Play“-Button kann die Entwicklung in der Zeit seit Beginn der Datenaufzeichnung animiert werden, ein Schieberegler erlaubt es, die Daten zu zurückliegenden Zeitpunkten spezifisch zu erfassen. Jedes Diagrammschaubild verfügt über fünf Reiter: Mittels der ersten drei Reiter kann zwischen den Darstellungsweisen Liniendiagramm („chart“), Abbildung der Daten auf einer Weltkarte („map“) oder einer tabellarischen Aufreihung der Messwerte („table“) gewählt werden; die beiden verbleibenden Reiter nennen und verlinken die Quellen der Datenerhebung und ermöglichen den Download der Grafiken als Bilddateien bzw. des Datensatzes zur weiteren Verwendung (siehe Abbildung 1). Stärken des Portals sind die große Anzahl verfügbarer Datensätze (in ihrer historischen Dimension), die einfache, interaktiv beeinflussbare Handhabung der Grafiken, die Hinweise zur Erfassung der Daten und Codierung der Variablen nebst Verlinkung der Quellen (inkl. Warnungen zur Datenqualität bzw. Hinweisen auf alternative Datensätze).

Einsatz im Unterricht: Im Unterricht eignet sich das Portal vor allem in Recherchephasen, in denen Schüler:innen eigenständig Daten zu bestimmten strukturellen Entwicklungen (z.B. Entwicklung der Weltbevölkerung, demografische Altersstruktur in verschiedenen Gesellschaften), politischen Herausforderungen (z.B. Entwicklung der Anzahl und Letalität von Terroranschlägen, Proliferation von Nuklearwaffen) oder aktuellen Herausforderungen (z.B. Corona-Impfquote verschiedener Länder, Übersterblichkeit seit Ausbruch der Pandemie usw.) ermitteln. Wichtig ist es, dass die Schüler:innen die Daten – auch aus qualitativ hochwertigen Datenbanken und verlässlichen Quellen – niemals unhinterfragt und als vermeintlich objektive Fakten einfach übernehmen: Alle Daten lassen sich kritisch bezüglich ihrer Erhebung, der Quantifizierung qualitativer Merkmale usw. befragen! (Z.B. Wie ist „Terroranschlag“ definiert?

Welche Informationen zur datenerhebende Institution, „Global Terrorism Database“, sind aufzutreiben? Handelt es sich um eine regierungsunabhängige Organisation? usw.).

Abbildung 1: Todesfälle durch terroristische Anschläge 2000-2019



Quelle: <https://ourworldindata.org/terrorism>.

Statistische Datenanalyse: Die Mittel der statistischen Analyse sind stark begrenzt. Es lassen sich weder Parameter zur Bestimmung der Lage (z.B. Durchschnittswert, Mittelwert) noch zur Streuung der Daten bestimmen; außerdem ist es nicht direkt möglich, zwei Variablen zueinander ins Verhältnis zu setzen, um deren Zusammenhang zu analysieren. Vergleiche sind nur hinsichtlich verschiedener Messzeitpunkte (diachron) sowie zwischen verschiedenen Elementen der Stichprobe (synchron) möglich.

5.2 Gapminder

Hintergrund und Umfang: Die öffentliche zugängliche Web-App *Gapminder* (<https://www.gapminder.org/tools>) wird von einer weltanschaulich unabhängigen schwedischen Stiftung, der Gapminder Foundation, betrieben. Diese verfolgt das Ziel, Fehlinformationen und Missverständnisse über globale Entwicklungen zu bekämpfen, indem sie *open data* aus zahlreichen Bereichen des öffentlichen Lebens kostenlos als Unterrichtsressourcen zur Verfügung stellt. Es sind über 500 Trends von Indikatoren zu Wirtschaft, Politik, Gesundheit und Gesellschaft abrufbar, die Daten reichen (wie bei *Our World in Data*) zum Teil bis zum Jahr 1800 zurück.

Web-Oberfläche und Nutzung: Die Anwendung „Tools“ bietet vielfältige Möglichkeiten der Visualisierung. In der Standardeinstellung werden Blasendiagramme („bubble charts“) gezeigt, die die Verteilung von Daten in einem Koordinatensystem aus zwei Variablen (Streudiagramm) anzeigen (Abb. 2). Zusätzliche Informationen können über Größe und Farbe der Blasen abgebildet werden. Über einen „Play“-Button kann die Entwicklung der Daten im Zeitverlauf dargestellt werden. Es können somit insgesamt fünf Variablen zugleich visualisiert werden.³ Neben der Grundeinstellung können weitere Darstellungsformen wie geographische Zuordnung von Datenblasen auf

einer Weltkarte („Maps“), Liniencharts im Zeitverlauf („Trends“, siehe Abb. 3) oder Rangfolgecharts („Ranks“) gewählt werden.

Abbildung 2: Streudiagramm Pro-Kopf-Einkommen vs. Lebenserwartung



Quelle: Grafik auf:
<https://www.gapminder.org/tools>

Abbildung 3: Liniendiagramm: Veränderung des Pro-Kopf-Einkommens seit 1800



Quelle: Grafik erstellt auf:
<https://www.gapminder.org/tools>

Einsatz im Unterricht: Zentraler Vorteil gegenüber *Our World in Data* ist, dass *Gapminder* die Möglichkeit bietet, zwei Variablen in einem Streudiagramm zu „plotten“, d.h. miteinander in Beziehung zu setzen. Dadurch können statistische Zusammenhänge erkundet werden. Bspw. ist in Abb. 2 erkennbar, dass zwischen den Variablen Lebenserwartung in Jahren („Life expectancy“) und dem inflationsbereinigten Pro-Kopf-Einkommen im Jahr 2022 ein deutlicher Zusammenhang besteht: Je höher das Pro-Kopf-Einkommen, desto höher ist tendenziell auch die Lebenserwartung in einem Land.

Wichtig bei der Untersuchung von Zusammenhängen zwischen zwei Größen ist es, die Schüler:innen darauf hinzuweisen, dass ein statistischer Zusammenhang (Korrelation) ein Indikator für eine Ursache-Wirkungs-Beziehung (Kausalität) sein kann, dies aber keinesfalls immer der Fall sein muss. Eine wichtige Rolle spielt dabei die Möglichkeit von explanativen Drittvariablen (*confounders*), die einen beobachteten Zusammenhang zwischen zwei Variablen erklären können. Hier hilft es, mit *Gapminder* weitere Variablen über die Größe und Farbgebung der Blasen zu kodieren oder gezielt Untergruppen (z.B. OECD Länder) auszuwählen. Ein erster Schritt, um eine kausale Wirkung begründet zu vermuten, besteht darin, eine Hypothese zu einem kausalen Mechanismus aufzustellen, die den Zusammenhang plausibilisiert. Mit Blick auf den in Abb. 2 dargestellten Zusammenhang ließe sich bspw. vermuten, dass mit steigendem Einkommen auch die Möglichkeiten gesunder Ernährung sowie die Chance, in einem gut ausgestatteten Gesundheitssystem zu leben, ansteigen. – Generell bietet es sich an, Schüler:innen hinsichtlich der zu untersuchenden Zusammenhänge zunächst immer erst Hypothesen bilden zu lassen (so auch: Porst 2014, S. 81), die dann ‚getes-

tet‘ werden, um ein zielloses und willkürliches Aufeinanderbeziehen von Variablen zu vermeiden (vgl. Weber-Stein/Engel 2021).⁴

Statistische Datenanalyse: Auch *Gapminder* bietet bezüglich der Datenanalyse nur begrenzte Möglichkeiten, die Berechnung statistischer Kennwerte ist nicht möglich. Allerdings sind die Möglichkeiten der visuellen Exploration nicht zu unterschätzen, Korrelationen zwischen zwei Variablen können recht gut visualisiert werden.

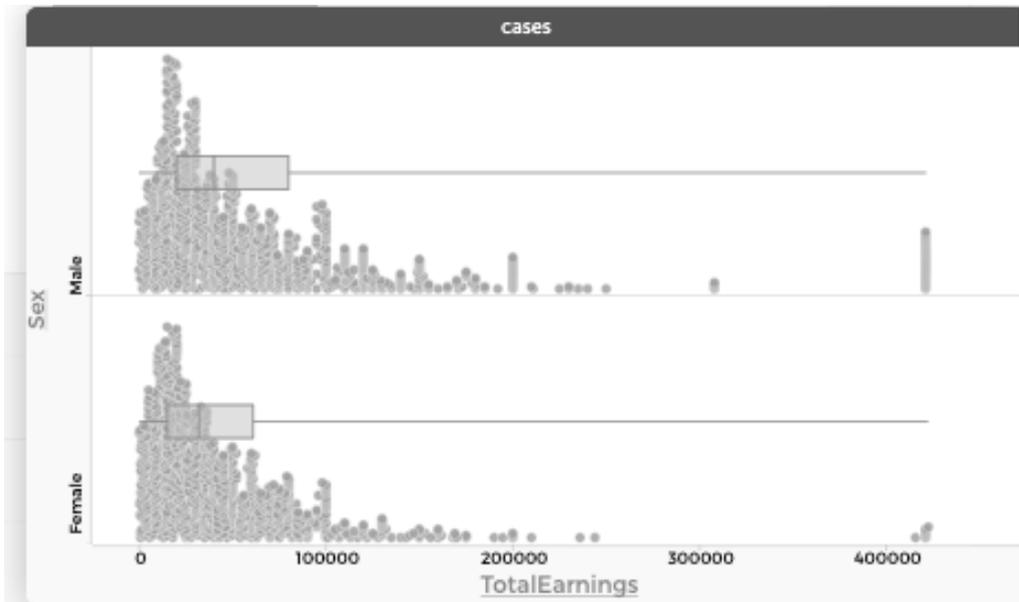
5.3 CODAP

Hintergrund: CODAP (<https://codap.concord.org/>) steht für *Common Online Data Analysis Platform* und ist eine kostenlose, browser-basierte Web-App, die mittels Unterstützung der US-amerikanischen National Science Foundation als didaktisches Werkzeug zur Datenexploration entwickelt wurde und beständig weiterentwickelt wird. Anders als bei den beiden bisherigen Plattformen, müssen die meisten Daten in CODAP erst (per „drag and drop“) importiert werden. Ein Datenimport ist über das (MS-Excel-kompatible) .csv-Format möglich (viele Datenbanken, die im Internet heruntergeladen werden können, besitzen dieses Format). In begrenztem Maße können .html-basierte Daten (Tabellenform) aus dem Internet eingelesen werden („webscraping“). Über Java-Script-Plugins können die Grundfunktionalitäten der App erweitert werden. Eine Warnung muss ausgesprochen werden: Nicht alle im Internet verfügbaren Datenbanken können direkt sinnvoll mit CODAP bearbeitet werden. Oftmals ist zuvor eine „Kuratierung“ durch die Lehrkraft notwendig, z.B. Begrenzung der Variablen und der Fälle auf ein für die Schüler:innen handhabbares Maß.

Web-Oberfläche und Nutzung: Importierte Datenbanken werden im *default*-Modus als Tabelle angezeigt. Es können per Mausklick Variablen den beiden Achsen eines Streudiagramms zugeordnet werden. Grafische Darstellungen und Tabellen sind gekoppelt: Werden beispielweise Ausreißer-Werte in einer Grafik per Maus markiert, sind diese anschließend in allen anderen Grafiken und der Tabelle hervorgehoben (und können ausgeblendet, gelöscht usw. werden). Die grafischen Darstellungen können mittels verschiedener Funktionen individuell bearbeitet werden (Durchführung numerischer Analysen, Ausblendung eines Teils der Stichprobe usw.).

Einsatz im Unterricht: Im Gegensatz zu den beiden anderen Apps ist die Datenexploration mittels CODAP nicht voraussetzungslos. Einerseits erfordert die technische Handhabung des Programms etwas Übung, in die Grundfunktionen der App sollte vorab eingeführt werden. Informatische Kenntnisse sind hingegen nicht erforderlich. Andererseits sollten basale statistische Grundkenntnisse (Lage- und Streuungsparameter, Boxplots, ggf. Korrelation) begleitend (oder in Kooperation mit Mathematiklehrer:innen) vermittelt werden, um die Funktionen angemessen nutzen zu können. – Sind diese beiden Voraussetzungen erfüllt, bietet CODAP deutlich anspruchsvollere und flexiblere Möglichkeiten der Datenexploration als die beiden anderen Apps. Erste Erfahrungen zum Einsatz im schulischen Unterricht liegen aus dem US-amerikanischen und deutschen Raum vor (vgl. Louie 2018).

Abbildung 4: Einkommensverteilung amerikanischer Haushalte differenziert nach Geschlecht (binär)



Grafik gestaltet mit <https://codap.concord.org/>

Statistische Datenanalyse: Ein großes Plus gegenüber den beiden anderen Apps besteht darin, dass mit den Daten aktiv gearbeitet werden kann: So können bspw. Mittelwerte (arithmetisches Mittel, Median) ermittelt und in ein Diagramm eingezeichnet werden, ebenso Boxplots zum Vergleich von Verteilungen (vgl. Abb. 4). Neue Variablen können definiert werden und die Datentabelle kann mit einfachen Handhabungen (*data moves*) umstrukturiert werden, was bei der Analyse hierarchisch strukturierter Datensätze (z.B. Länder können nach Regionen und diese nach Kontinenten sortiert werden) sehr hilfreich ist. Außerdem kann eine Regressionsgerade eingezeichnet und der Determinationskoeffizient (r^2) bestimmt werden.

Abbildung 5 fasst die Eigenschaften der drei vorgestellten *tools* noch einmal zusammen:

Abbildung 5: Eigenschaften und Einsatzmöglichkeiten der vorgestellten statistischen tools

	Our World in Data	Gapminder	CODAP
Sprache	Englisch	Englisch	Englisch, Deutsch und weitere
Zugang	Browserbasiert, open access und open source	Browserbasiert, open access und open source	Browserbasiert, open access und open source
Verfügbarkeit von Daten auf Website	Ja (über 3.000 Datenbanken zu fast 300 Themenfeldern)	Ja (über 500 Datenbanken zu zahlreichen Themenfeldern)	Bedingt (Zugang über Plug-Ins)
Vorausgesetzte technische Fertigkeiten	keine	keine	Handhabung bedarf der Übung, ist aber recht intuitiv
Statistische Voraussetzungen	<ul style="list-style-type: none"> – kritische Grundhaltung zur Erhebung von Daten – Liniendiagramme 	<ul style="list-style-type: none"> – Kritische Grundhaltung zur Erhebung von Daten – Zeitreihendiagramme – Streudiagramme – ggf. Korrelation 	<ul style="list-style-type: none"> – Kritische Grundhaltung zur Erhebung von Daten – Grundkenntnisse deskriptiver Statistik: Mittelwert, Median, Perzentile, Boxplot, Streudiagramm – ggf. Korrelation
Statistische Anwendungen	<ul style="list-style-type: none"> – Vergleich von Trends in der Zeit – evidenzbasierte Beschreibung von Entwicklungen und Dynamiken 	<ul style="list-style-type: none"> – Einfache Exploration statistischer Zusammenhänge – Beschreibung reichhaltiger Diagramme 	<ul style="list-style-type: none"> – Darstellen und Vergleichen von Verteilungen – Analyse von Untergruppen – Definition neuer Variablen – Geraden oder Kurven einfügen

Eigene Darstellung

6. Fazit

Wie effektiv waren Demokratien in der Eindämmung der Corona-Pandemie? Wie kann Effektivität in diesem Kontext operationalisiert werden? Welche Indikatoren für effektives Pandemiemanagement gibt es (und welche Vor- und Nachteile haben sie jeweils)? Haben Demokratien, die bei der Eindämmung besonders erfolgreich waren, demokratische Standards verletzt? (hierzu: Weber-Stein/Engel 2022) – Zahlreiche aktuelle und kontroverse Themen im sozialwissenschaftlichen Unterricht setzen die Analyse und Interpretation komplexer Daten voraus. Anstelle die Schüler:innen ausschließlich mit datenbasiertem Wissen aus zweiter Hand (bspw. aus Texten, die die Daten interpretieren) arbeiten zu lassen, ist es eine motivierende und stärkende Erfahrung, Daten selbst zu explorieren. Die dazu notwendige Verbindung von methodischen und inhaltsbezogenen Kompetenzen ist herausfordernd. Moderne digitale tools stellen aber einen niedrighschwelligigen Zugang zu aktuellen Datenbanken bereit und ermöglichen eine Exploration auf verschiedenen Niveaustufen.

Anmerkungen

- 1 Im Folgenden thematisieren wir ausschließlich die Debatte in der deutschen Politikdidaktik. – Für den internationalen Diskurs um *Data Literacy* als Aspekt von *Digital Citizenship* ließe sich analog zeigen, dass statistische Bildung allenfalls eine untergeordnete Rolle spielt.
- 2 „Data visualisations hold the promise of direct access to rich, authentic and contemporary content to support exploration and decision-making“ (Ridgway 2016, S. 540).
- 3 Für Möglichkeiten der Datenexploration mit Gapminder siehe den Erklärfilm „How does Income Relate to Life Expectancy?“, <https://www.gapminder.org/answers/how-does-income-relate-to-life-expectancy/> [11.7.22].
- 4 Zur Bearbeitung mit Gapminder haben die Autoren eine Lernaufgabe zum Thema „Demokratiequalität“ entwickelt. Das Arbeitsblatt kann bei den Autoren angefordert werden und wird zeitnah auf *CivicStatMap* (<http://iase-web.org/islp/pcs>) zum download bereitstehen.

Literatur

- Boy, Jeremy et al. (2017): Showing People Behind Data: Does Anthropomorphizing Visualizations Elicit More Empathy for Human Rights Data?, in: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, S. 5462 – 5474. DOI: 10.1145/3025453.3025512
- Bundesgesundheitsministerium (2022): Evaluation der Rechtsgrundlagen und Maßnahmen der Pandemiepolitik. Bericht des Sachverständigenausschusses, in: <https://www.bundesgesundheitsministerium.de/service/begriffe-von-a-z/s/sachverstaendigenausschuss-infektionsschutzgesetz.html> [05.09.2022].
- Deutsche Arbeitsgemeinschaft Statistik (2021): *Stellungnahme der Deutschen Arbeitsgemeinschaft Statistik (DAGStat): Daten und Statistik als Grundlage für Entscheidungen: Eine Diskussion am Beispiel der Corona-Pandemie*, in: https://www.dagstat.de/fileadmin/dagstat/documents/DAGStat_Covid_Stellungnahme.pdf [05.09.2022].
- Grenier, Elizabeth (2020): Coronavirus: ‚Social Distancing‘ - Visualisierungen im Netz, DW, Digitalkultur, in: <https://p.dw.com/p/3ZWN4> [05.09.2022].
- Kosara, Robert/Mackinlay, Jock (2013): Storytelling: The Next Step for Visualization, in: Computer 5/2013, S. 44-50. DOI: 10.1109/MC.2013.36.
- Lach, Kurt/Massing, Peter (2017): Umgang mit Statistiken und Tabellen, in: Peter Massing/Siegfried Frech (Hg.), *Methodentraining II für den Politikunterricht*, 2. Aufl. Bonn: Bundeszentrale für politische Bildung, S. 21-30.
- Louie, Josephine (2018): Strengthening Data Literacy across the Curriculum, in: <https://www.edc.org/strengthening-data-literacy-across-curriculum> [11.07.22].
- Nicholson, James et al. (2018): Understanding Civic Statistics: A Conceptual Framework and its Educational Applications. A Product of the ProCivicStat Project, in: <http://IASE-web.org/islp/pcs> [11.07.2022].
- Porst, Rolf/Holthof, Ruth (2014a): Der sozialwissenschaftliche Forschungsprozess – und was davon wie im Sozialkundeunterricht machbar und möglich ist, in: GWP – Gesellschaft. Wirtschaft. Politik 1/2014, S. 79-91. DOI: 10.3224/gwp.v63i1.16064
- Dies. (2014b): Verfahren zur Erhebung sozialwissenschaftlicher Daten, in: GWP – Gesellschaft. Wirtschaft. Politik 3/2014, S. 405-416. DOI: 10.3224/gwp.v63i3.16713
- Ridgway, Jim et al. (2007): Reasoning with multivariate evidence, in: International Electronic Journal of Mathematics Education 3/2007, S. 245-269. DOI: 10.29333/iejme/212
- Ridgway, Jim (2016): Implications of the Data Revolution for Statistics Education, in: International Statistical Review 3/2016, S. 528-549. DOI: 10.1111/insr.12110

- Ridgway, Jim (2022): *Statistics for Empowerment and Social Engagement – Teaching Civic Statistics to Develop Informed Citizens*, Wiesbaden: Springer 2022, i.E.
- Schattschneider, Jessica (2010): Empirische Tools, in: Anja Besand/Wolfgang Sander (Hg.), *Handbuch Medien in der politischen Bildung*, Schwalbach/Ts.: Wochenschau, S. 186-192.
- Segel, Edward/Heer, Jeffrey (2010): Narrative Visualization: Telling Stories with Data, in: *IEEE Transactions on Visualization and Computer Graphics* 6/2010, S. 1139-1148.
DOI: 10.1109/TVCG.2010.179
- Torrau, Sören (2020): „Und dann google ich...“ Recherchestrategien von Schüler*innen im Internet, in: *GWP – Gesellschaft, Wirtschaft, Politik* 2/2020, S. 235-245. DOI: 10.3224/gwp.v69i2.12
- Wassner, Christian/Prömmel, Andreas (2022): Civic Statistics at School – Promoting Real Data Competence in the Classroom, in: Jim Ridgway (Hg.), *Statistics for Empowerment and Social Engagement – Teaching Civic Statistics to Develop Informed Citizens*, Wiesbaden: Springer, i.E.
- Weber-Stein, Florian/Engel, Joachim (2021): Civic Statistical Literacy und Politische Bildung im Informationszeitalter. Kooperative statistik- und politikdidaktische Erkundungen im Feld der Demokratiemessung, in: Carl Deichmann/Marc Partetzke (Hg.), *Demokratie im Stresstest? Reaktionen von Politikdidaktik und politischer Bildung*, Wiesbaden: Springer VS, S. 165-192.
DOI: 10.1007/978-3-658-33077-4_10.
- Weber-Stein, Florian/Engel, Joachim (2022): The Covid-19 Crisis as a Challenge for the Integration of Statistical and Citizenship Education, in: Sue Peters u.a. (Hg.): *Bridging the Gap: Empowering & Educating Today’s Learners in Statistics. Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11)*, Rosario, Argentina 2022, DOI: 10.52041/iase.icots11.T1E1.
- Weißeno, Georg (2010): Statistik/Schaubild, in: Anja Besand/Wolfgang Sander (Hg.), *Handbuch Medien in der politischen Bildung*, Schwalbach/Ts.: Wochenschau, S. 495-502.