

# Krankenhäuser, Unfälle und Statistik. Mini-Projekte zur Thematisierung von Zivilstatistik im Unterricht

SUSANNE PODWORNÝ UND DANIEL FRISCHEMEIER, PADERBORN

**Zusammenfassung:** In diesem Artikel stellen wir zwei Mini-Projekte über Zivilstatistik für den Unterricht der Sekundarstufe I vor. Dabei greifen wir auf einen Datensatz und ein Tool des Statistischen Bundesamts zurück, die frei verfügbar sind. Konkret stellen wir das Mini-Projekt „Unfallstatistik“ für die 5./6. Klasse vor und das Mini-Projekt „Krankenhäuser in Deutschland“ für den Unterricht in der 9./10. Klasse.

## 1 Einleitung

Ziviles Engagement gehört zu den konstitutiven Elementen einer Demokratie und lebt von mündigen Bürgerinnen und Bürgern. Um zum mündigen Bürger zu reifen (siehe u. a. Schiller 2017), bedarf es – gerade bei der gewaltigen Masse an Daten und Informationen in der heutigen Zeit – einer statistischen Grundbildung, der sogenannten Statistical Literacy (siehe z. B. Gal 2002). Mit Blick auf bestimmte Statistiken über gesellschaftlich relevante Phänomene (z. B. Migration, Arbeitsmarkt, Kriminalität, Umwelt, etc.), so genannten „Zivilstatistiken“, erlangen insbesondere der spezifische Kontext und das Kontextwissen eine enorme Bedeutung bei der Interpretation solcher Statistiken. Details zum Konzept der Zivilstatistik finden sich in Engel (dieses Heft) oder auch in Engel, Gal & Ridgway (2016). Die Thematisierung von Zivilstatistiken und die Allgegenwärtigkeit von großen Datenmengen – aber auch ihr Missbrauch – bilden neue Herausforderungen für die Stochastikausbildung (siehe z. B. Ridgway 2016). So liefern zivilstatistische Inhalte gute Anknüpfungspunkte für fächerübergreifenden Unterricht, haben aber oft deshalb keinen festen Platz im Curriculum (Engel, dieses Heft). Weiterhin müssen Lehrkräfte entsprechend zunächst erst einmal kompetent ausgebildet werden, um zivilstatistische Inhalte in den Unterricht implementieren zu können.

Die Autoren dieses Artikels haben eine universitäre Lehrveranstaltung konzipiert (siehe dieses Heft und Biehler, Frischemeier & Podworný 2018), um sowohl die fachlichen als auch die fachdidaktischen Kompetenzen von Lehramtsstudierenden im Bereich der Zivilstatistik auszubauen und zu fördern. Zwei Mini-Projekte, die in diesem Rahmen mit Lehramtsstudierenden durchgeführt worden sind, sind das Mini-Projekt „Unfallstatistik“ und das Mini-Projekt „Krankenhäuser in Deutschland“. Beide Mini-Pro-

jekte nutzen reale Daten, die vom Statistischen Bundesamt bereitgestellt werden sowie digitale Tools zur Analyse der Daten. Diese Mini-Projekte haben einen Umfang von ein bis drei Schulstunden und sind konzipiert worden, um im Unterricht der Sekundarstufe I zum einen die zivilstatistischen Kompetenzen von Schülerinnen und Schülern zu fördern und zum anderen um Brücken zum fächerübergreifenden Unterricht zwischen Mathematik und Sozialwissenschaften zu bauen. Hauptsächliches Ziel ist der Umgang mit echten Daten im Mathematikunterricht.

## 2 Das Mini-Projekt „Unfallstatistik“

Das Mini-Projekt „Unfallstatistik“ beruht auf dem Verkehrsunfallkalender des Statistischen Bundesamts, der unter <https://www.destatis.de/DE/Service/Verkehr/Verkehrsunfaelle.html> einsehbar ist. Dabei können verschiedene Straßenunfallgeschehen der Jahre 2007 bis 2017 dargestellt und miteinander verglichen werden.

Zur Implementation eines zivilstatistischen Mini-Projekts im Unterricht schlagen wir vor, zunächst eine kurze Begründung der Relevanz des zivilstatistischen Themas zu bringen ggf. in Kombination mit einem aktuellen Zeitungsartikel zur Motivation. Daraufhin sollten je nach Thema weitere Hintergrundinformationen (z. B. über die Erhebung der Daten) und die Daten selbst gegeben werden, wozu ggf. auch eine Variablenliste gehört. Hieran kann sich eine Analyse der Daten anschließen, die zum Teil entlang vorgegebener Fragen, aber auch entlang von selbstgewählten Fragen seitens der Schülerinnen und Schülern bestehen kann.

### Motivierung des Themas

Ein Aufhänger zur Motivation in das Mini-Projekt „Unfallstatistik“ kann eine Zeitungsmeldung oder die Meldung eines Online-Mediums (wie z. B. in <https://www.auto-motor-und-sport.de/verkehr/unfallstatistik-2017-zahl-der-verkehrstoten/> (aufgerufen am 23.10.2018)) sein.

In dieser Meldung wird die Entwicklung der Zahl der (Unfall-)Verkehrstoten im Jahr 2017 thematisiert. Dieses kann als Motivation genutzt werden, sich allgemein mit der Entwicklung der Unfallzahlen in Deutschland auseinanderzusetzen.

## Hintergrundinformation

Eine wichtige Hintergrundinformation ist der Ablauf der Erhebung der Daten gemäß Auskunft des Statistischen Bundesamts:

Auskunftspflichtig sind – laut Gesetz – die Polizeidienststellen, deren Beamte den Unfall aufgenommen haben. Daraus folgt, dass die Statistik nur solche Unfälle erfasst, zu denen die Polizei herangezogen wurde; das sind vor allem solche mit schweren Folgen. Insbesondere Verkehrsunfälle mit nur Sachschaden oder mit nur geringfügigen Verletzungen werden zu einem relativ großen Teil der Polizei nicht angezeigt.

So muss bei der Interpretation der Daten stets berücksichtigt werden, dass es sich um einen Ausschnitt der tatsächlich geschehenen Unfälle handelt, nämlich nur solche, die polizeilich gemeldet wurden und bei denen Personenschäden auftraten.

Eine Variablenliste im eigentlichen Sinn existiert für dieses Tool nicht, allerdings kann die Gesamtanzahl der Unfälle mit Personenschaden verglichen werden mit Alkoholunfällen, Motorradunfällen oder Fahrradunfällen an den einzelnen Tagen eines Jahres.

## Analyse der Daten

In Abbildung 1 ist ein Ausschnitt des Unfallkalenders der Homepage des Statistischen Bundesamts zu sehen.

Für die einzelnen Tage der Jahre 2007 bis 2017 existiert für jedes Jahr ein Kalender, der die Verteilung der Anzahl der Verkehrsunfälle mit Personenschaden für die einzelnen Tage der Jahre zeigt. Links in Abbildung 1 ist dabei der Kalender zu sehen, in welchem Informationen über Personenschäden bei Verkehrsunfällen angezeigt werden. Beim rechten Ka-

lender in Abbildung 1 kann optional gewählt werden zwischen der Verteilung von Unfällen, die durch Alkoholeinfluss verursacht wurden und der Verteilung von Unfällen, die durch Motorradfahrer oder durch Fahrradfahrer verursacht wurden. Ein Vergleich der Häufigkeiten der verschiedenen (gemeldeten) Unfallarten ist somit möglich.

Im Wesentlichen geht es beim Unfallkalender des Statistischen Bundesamts um das Beschreiben und Interpretieren von Visualisierungen von Daten. Die bekannte Stufung in „reading the data“, „reading between the data“ und „reading beyond the data“ von Friel, Curcio, und Bright (2001) kann hier aus Metaperspektive mögliche Perspektiven beim Beschreiben und Interpretieren des Unfallkalenders aufzeigen. Kurz zur Erinnerung: Auf der unteren Stufe „reading the data“ geht es um das Ablesen von gegebenen Informationen (z. B. Was wird in dem Diagramm dargestellt? Welche Kategorien kommen vor? Welche Werte kommen vor?), eine Interpretation findet auf dieser Stufe noch nicht statt. Die mittlere Stufe „reading between the data“ erfordert bereits mathematische Kompetenzen wie beispielsweise das Vergleichen von Mengen und das Durchführen elementarer Rechenoperationen. Darüber hinaus wird auf dieser Stufe herausgearbeitet, welches der häufigste Wert/die häufigste Kategorie ist, welches der größte/kleinste Wert ist, wie viele Personen befragt worden sind und welchen Titel man dem Diagramm geben würde. Auf der nach Friel et al. (2001) höchsten Stufe „reading beyond the data“ werden Voraussagen und Schlussfolgerungen unter Einbeziehung von Hintergrundwissen vorgenommen. Um einen genaueren Einblick in die Anwendung dieser drei Stufen zu bekommen, schauen wir uns den Gesamtunfall-

### Straßenverkehrsunfälle: Die Kalender der Jahre 2017 und früher

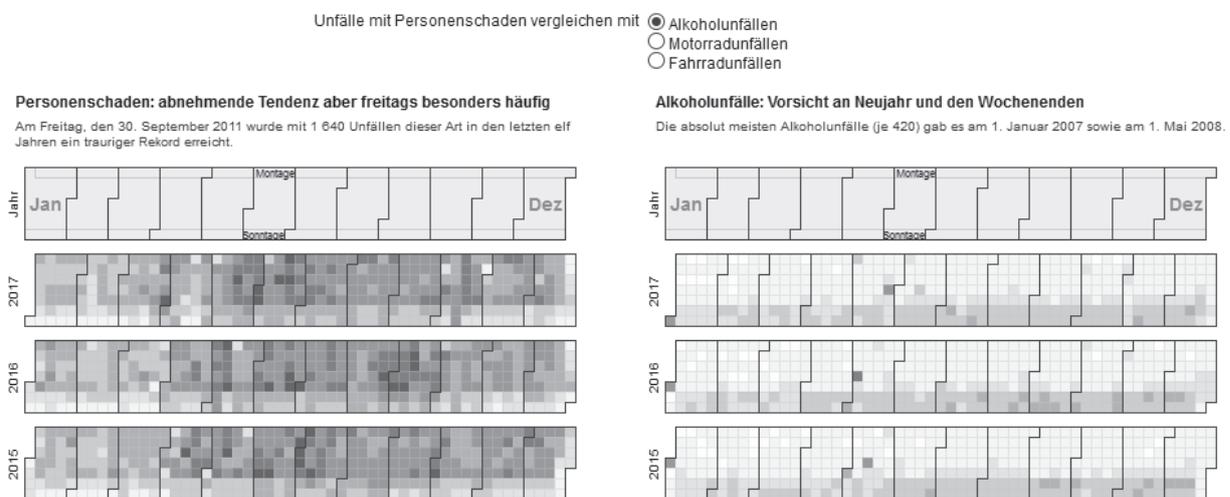


Abb. 1: Der Unfallkalender auf <https://www.destatis.de/DE/Service/Verkehr/Verkehrsunfaelle.html>

kalender des Jahres 2016 (siehe Abbildung 2) einmal genauer an.

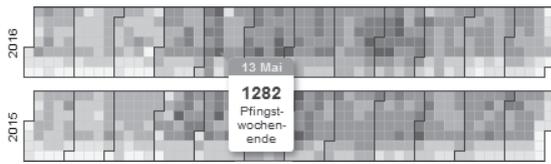


Abb. 2: Ausschnitt aus dem Unfallkalender (Jahre 2015 und 2016)

Auf der Stufe „reading the data“ könnte man feststellen, dass von links nach rechts gelesen die einzelnen Monate (Januar bis Dezember) eingezeichnet sind. Von oben nach unten betrachtet, lassen sich die Wochentage identifizieren (Montag bis Sonntag). Weiterhin könnten Lernende auf dieser Stufe feststellen, dass (für das Jahr 2016) die Anzahl der Verkehrsunfälle für jeden Tag des Jahres 2016 angegeben ist (die genaue Anzahl wird angegeben wie in Abbildung 2 zu sehen, wenn mit dem Mauszeiger über einen bestimmten Tag des Jahres gefahren wird, ansonsten gibt die Farbe einen Zahlenbereich an. Für die Zahlenbereiche existiert eine Legende, grundsätzlich gilt, je dunkler die Farbe ist, desto mehr Unfälle sind geschehen).

Auf der nächsten Stufe („reading between the data“) können dann erste Vergleiche zwischen den einzelnen Tagen angestrebt werden. Beispielsweise können einzelne Tage anhand der absoluten Häufigkeiten (Anzahl Verkehrsunfälle) verglichen werden. Dabei haben sich an den Tagen, die dunkler eingefärbt sind, vergleichsweise mehr Verkehrsunfälle ereignet. Mit dieser Farbkodierung können dann qualitativ nicht nur Tage, sondern auch Monate verglichen werden. So fällt bei qualitativer Betrachtung der Abbildung 2 auf, dass sich in den Sommermonaten anscheinend (absolut gesehen) mehr Unfälle ereignen als in den Wintermonaten. Auch an einzelnen Tagen scheint die absolute Anzahl an Unfällen hoch zu sein. Welche Erklärungen kann es dazu geben? Hier kommt die Stufe „reading beyond the data“ ins Spiel. Schauen wir erst einmal lokal auf die Unfallzahlen einzelner Tage. Als Tag mit vielen Unfällen sticht u. a. der 13. Mai 2016 heraus. Was ist an diesem Tag besonders? Wir können in Abb. 2 erkennen, dass der markierte Tag (13. Mai) ein Freitag ist, an dem sich 1282 (gemeldete) Unfälle ereignet haben. Mit Blick auf den Kalender erfahren wir, dass der 13. Mai 2016 ein besonderer Freitag ist (und das nicht aufgrund der Tatsache, dass es Freitag der 13. ist), nämlich der Freitag vor dem Pfingstwochenende. Mit diesem Hintergrundwissen können wir nun mutmaßen, dass viele Menschen das Pfingstwochenende nutzen, um beispielsweise in den

Kurzurlaub zu fahren. Das erhöhte Verkehrsaufkommen, das mit einer solchen Reisewelle einhergeht, kann dann absolut gesehen auch als Ursache für die vergleichsweise hohe Verkehrsunfallanzahl gesehen werden. Wie kann es jetzt – mit einem globaleren Blick – sein, dass es anscheinend in den Sommermonaten mehr Unfälle als in den Wintermonaten gibt? Auch hier lässt sich erst einmal nur auf Grundlage von Hintergrundwissen spekulieren: Es könnte daran liegen, dass im Winter aus Respekt vor möglicher Glätte mehr Personen auf den öffentlichen Nahverkehr umsteigen und nicht Auto fahren, wodurch die absolute Anzahl an Autofahrern (und damit vermutlich auch die absolute Anzahl an Unfällen) sinken würde. Ebenso kann es sein, dass Autofahrer im Winter vorsichtiger fahren als im Sommer und deshalb weniger Unfälle geschehen. Außerdem ist im Sommer Ferienzeit und aufgrund der langen Sommerferien immer noch die Urlaubszeit Nummer Eins, was entsprechende Bewegungen und Staus (und somit Unfallpotential) auf Bundesstraßen und Autobahnen nach sich zieht.

Insgesamt – abgesehen von dem Beschreiben und Interpretieren der Diagramme, wie wir es im obigen Abschnitt ausgeführt haben – bietet das Onlinetool des Statistischen Bundesamts nur wenig Explorationsmöglichkeiten, so z. B. das Filtern nach Unfällen, die durch Alkoholmissbrauch oder das Filtern nach Unfällen, die durch Motorradfahrer (siehe Abbildung 1, rechts) oder Fahrradfahrer verursacht wurden. Es werden keine statistischen Kennzahlen und auch keine relativen Häufigkeiten genutzt. Daher kann die Exploration des Unfallkalenders bereits relativ früh (z. B. in Jahrgangsstufe 5/6) erfolgen.

Die Daten selbst lassen sich als csv-Datei herunterladen und bieten somit eine Differenzierung für leistungsstarke Schülerinnen und Schüler an: Diese können die heruntergeladenen Daten mit einem Tabellenkalkulationsprogramm (z. B. Microsoft Excel) noch weiter und tiefgründiger explorieren. Um die Schülerinnen und Schüler bei der Exploration der Daten zu leiten, haben wir eine Leitfrage (*Wie entwickeln sich die Zahlen der Verkehrsunfälle im Laufe der Jahre?*) mit weiteren Unterleitfragen formuliert. Die Unterleitfragen können dabei zum Beispiel lauten:

- An welchen Tagen geschehen besonders viele Unfälle?
- Such Dir ein Jahr aus, auf das Du Dich in Deinen Untersuchungen konzentrierst. Wie ist der Zusammenhang zwischen den Wochentagen und der Anzahl der Unfälle mit Personenschaden in diesem Jahr?

- Welche Entwicklung kannst Du über den Zusammenhang zwischen Wochentagen und Anzahl der Alkoholunfälle über das Jahr hinweg erkennen?

Abschließend schlagen wir vor, offene Fragestellungen („Stell zwei eigene statistische Fragen zu Verkehrsunfällen in Deutschland und beantworte diese mit geeigneter statistischer Auswertung“) anzubieten, die Lernenden eine freie und offene Exploration der Daten ermöglichen.

### 3 Das Mini-Projekt „Krankenhäuser in Deutschland“

Ein weiteres Mini-Projekt, welches auf Daten des Statistischen Bundesamts beruht, ist das Mini-Projekt „Krankenhäuser in Deutschland“. Der entsprechende Datensatz ist als campus-file auf der Homepage des Statistischen Bundesamts unter <http://www.forschungsdatenzentrum.de/bestand/drg/cf/2010/index.asp> frei verfügbar (hierzu ist lediglich eine kostenlose Registrierung nötig). Der Datensatz bietet ca. 1,7 Millionen anonymisierte Patientendaten mit Variablen wie Region (Nord, Süd, Ost) und Gebiet des Krankenhauses, in das der Patient/die Patientin eingeliefert worden ist (Ballungsraum, Stadt, Land), Herkunft des Patienten, Geschlecht des Patienten, Alter des Patienten, Aufnahmeanlass des Patienten (Einweisung, Notfall, Verlegung, Geburt), Entlassungsgrund, Verweildauer, etc.<sup>1</sup>

Ein Aufhänger als Einstieg in das Mini-Projekt kann, wie auch beim Einstieg in das Unfallkalender-Projekt, ein Bericht aus den Medien sein.<sup>2</sup>

Hier finden sich beispielsweise Artikel, die sich kritisch damit auseinandersetzen, dass Krankenhäuser im ländlichen Raum wegen Unterfinanzierung schließen müssen und somit eine wohnortnahe medizinische Versorgung nicht mehr überall gewährleistet werden kann.

Aufgrund des sehr umfangreichen Datensatzes bietet es sich an, eine Zufallsstichprobe aus dem ursprünglichen Datensatz zu ziehen, um die Exploration zu erleichtern. Im Folgenden arbeiten wir mit einer Zufallsstichprobe im Umfang  $n = 10000$ . Auch bei einer bereits reduzierten Zufallsstichprobe – wie in diesem Fall –, bleibt der Einsatz einer Datenanalyse-Software für eine echte Datenexploration und für ein Drehen und Wenden der Daten unerlässlich. Einen Überblick über entsprechende Tools findet sich in Biehler, Ben-Zvi, Bakker, und Makar (2013) und eine Diskussion über den Einsatz verschiedener Werkzeuge bei der Exploration von Mikrodaten findet sich in Frische-meier, Biehler und Engel (2016).

Die Online-Datenanalyseplattform CODAP (<https://codap.concord.org>) bietet hier vielfältige Möglichkeiten der Datenexploration. Bei CODAP handelt es sich um eine kostenlose, browserbasierte Datenanalyse-Software, für die keine Anwendung installiert werden muss. Für den Unterricht stellt dies einen Vorteil dar, da weder Zeit für die Installation noch Kosten anfallen. Mit einem Klick auf den Link (hier z. B. <https://tinyurl.com/yavlcuc7>) lässt sich „unsere“ Stichprobe des Krankenhausdatensatzes mit CODAP im Browser öffnen<sup>3</sup>.

Wenn der Link geöffnet wird, wird direkt der Datensatz in CODAP geladen und die Daten sind, wie in Abbildung 3 zu erkennen, in einer Tabelle gespeichert. Jeder Fall wird als Zeile der Tabelle abgebildet.

Kollektion 1_Krankenhaus									
Cases (10000 cases)									
index	Fall_Nr	Region_K	Gebiet_K	Region_P	Gebiet_P	Geschlecht	Alter_gru	Aufnahm_eanlass	Aufnahm_ewegwicht
1	594778	Nord	Ballung	Nord	Ballung	m	7	E	NA
2	18997	Nord	Ballung	Nord	Ballung	w	8	E	NA
3	657488	Nord	Ballung	Nord	Ballung	w	8	N	NA
4	266578	Nord	Ballung	Nord	VerstΦ	m	9	E	NA
5	447858	Nord	Ballung	Nord	Ballung	w	9	E	NA
6	303812	Nord	Ballung	Nord	Ballung	w	5	N	NA
7	64938	Nord	Ballung	Nord	Ballung	w	9	N	NA
8	407754	Nord	Ballung	Nord	Ballung	m	4	N	NA
9	60855	Nord	Ballung	Nord	VerstΦ	w	7	E	NA
10	1420342	Ost	Ballung	Ost	Ballung	m	10	E	NA
11	376083	Nord	Ballung	Nord	Ballung	w	5	E	NA
12	1338859	Süd	Ballung	Süd	Ballung	m	10	V	NA
13	621922	Nord	Ballung	Nord	Ballung	w	8	N	NA

Abb. 3: Tabelle mit den Krankenhausdaten in CODAP

Um die Daten nun explorieren zu können, lassen sich die Daten ähnlich wie bei Fathom (Biehler, Hofmann, Maxara, & Prömmel, 2011) oder TinkerPlots (Konold & Miller, 2011) auf Mikroebene als Datenpunkte darstellen.

Das Mini-Projekt steht unter der offenen Fragestellung „Wie ist die Lage der Krankenhäuser in Deutschland?“. Um diese offene Fragestellung ein wenig zu leiten, haben wir folgende Unterfragestellungen ausgewählt:

- 1) Welches ist der häufigste Aufnahmeanlass?
- 2) Was lässt sich im Hinblick auf Krankenhausregion, Aufnahmeanlass und Patientenherkunft herausfinden?

Weiterhin geben wir den Schülerinnen und Schülern auch Gelegenheit zur Exploration der Daten nach selbstgewählten Fragestellungen („Stell zwei eigene statistische Fragen zur Lage der Krankenhäuser in Deutschland und beantworte diese mit geeigneter statistischer Auswertung“). Exemplarisch beantworten wir die Fragen 1) und 2) durch Exploration der Daten in CODAP. Hierbei sollte stets ein Bezug zum Kontext hergestellt werden, der durch die anfängliche Thematisierung der medizinischen Unterversorgung in ländlichen Gegenden motiviert ist.

Die erste Frage dient dazu, sich mit dem Datensatz vertraut zu machen. Um die Frage nach dem häufigsten Aufnahmeanlass zu beantworten, bietet es sich an, ein gestapeltes Punktdiagramm zur Verteilung des Merkmals „Aufnahmeanlass“ in CODAP zu erstellen, wie in Abbildung 5 zu sehen. Per Drag & Drop kann das gewünschte Merkmal in der Tabelle (siehe Abbildung 3) ausgewählt und auf die X- oder Y-Achse des Graphen gelegt werden. Zusätzlich können über Zusatzfunktionen (siehe rechte Leiste neben dem Graph, Abbildung 4) auch die jeweiligen absoluten und relativen Häufigkeiten angezeigt werden. Dazu wählt man die Option „Lineal“ (rechte Leiste in Abbildung 4) und dann „Prozent“.

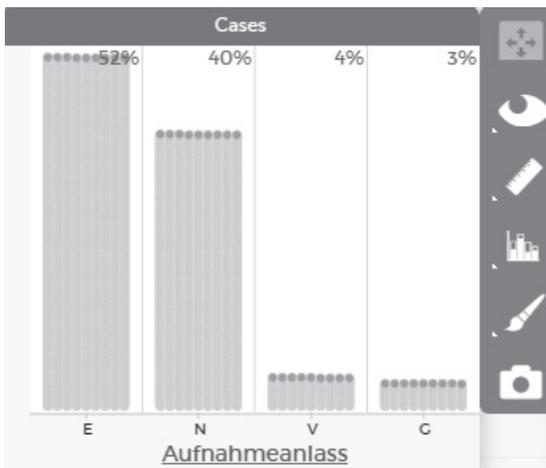


Abb. 4: Verteilung des Merkmals Aufnahmeanlass in CODAP

Wie in Abbildung 4 zu sehen ist, ist der mit in etwas mehr als der Hälfte aller Fälle (52 %) häufigste Aufnahmeanlass die Einweisung. 40 % der Aufnahmeanlässe sind Notfälle. Der Rest verteilt sich auf Verlegungen (4 %) und Geburten (3 %). Dieses Er-

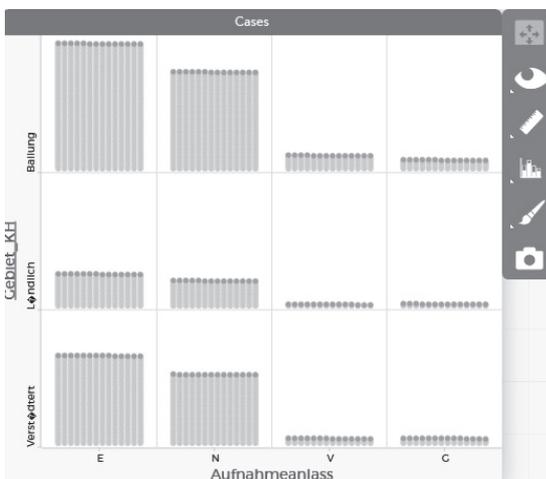


Abb. 5: Verteilung des Merkmals „Aufnahmeanlass“ getrennt nach Region des Krankenhauses in CODAP

gebnis sollte nicht weiter überraschend sein, denn Geburten und Verlegung kommen erwartungsgemäß (insbesondere verglichen mit Einweisung und Notfall) selten vor.

Beschäftigt man sich nun jedoch weiter mit dem Merkmal Aufnahmeanlass gemäß der zweiten Frage und nutzt dieses, um Unterschiede zwischen den verschiedenen Krankenhausregionen herauszufinden, so kann sich eine Graphik wie in Abbildung 5 ergeben.

An dieser Stelle kann im Unterricht thematisiert werden, dass ein Vergleich anhand von absoluten Häufigkeiten in diesem Fall nicht tragfähig ist, weil in den einzelnen Gruppen (Ballung, Ländlich, Verstädtert) unterschiedlich viele Fälle vorliegen. Es müssen in diesem Fall Zeilenprozent berechnet werden (Abb. 6).

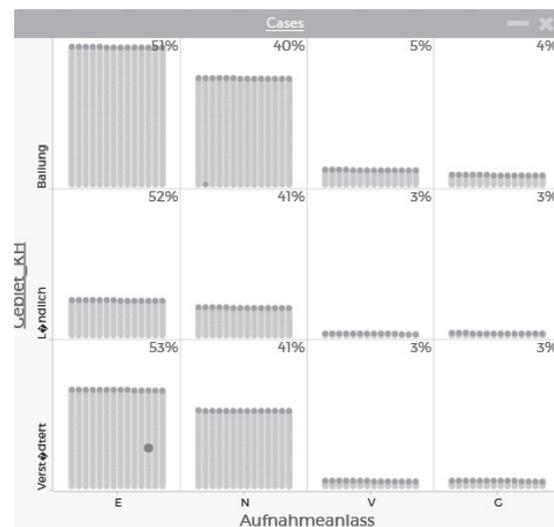


Abb. 6: Verteilung des Merkmals Aufnahmeanlass getrennt nach Region des Krankenhauses in CODAP mit Zeilenprozent

Anhand von Abbildung 6 ist erkennbar, dass sich der Aufnahmeanlass nicht wesentlich innerhalb der einzelnen Regionen unterscheidet. Auf der weiteren Suche nach Unterschieden zwischen den Regionen kann die Herkunft der Patienten im Hinblick auf die Region des aufgesuchten Krankenhauses verglichen werden. Die Erwartungshaltung ist hier zunächst, dass Patienten aus einer bestimmten Region ein Krankenhaus in dieser Region aufsuchen, z. B. ein Patient aus einer verstädterten Region in ein Krankenhaus dieser Region geht usw. Dies ist in Abbildung 7 zu sehen.

Die zuvor genannte Erwartungshaltung würde Einträge auf der Diagonalen bedeuten. In Abbildung 7 ist nun zu erkennen, dass es jedoch auch Bewegungen von Patienten in andere Regionen gibt, beispielsweise befinden sich nur 83 % der Patienten aus einer

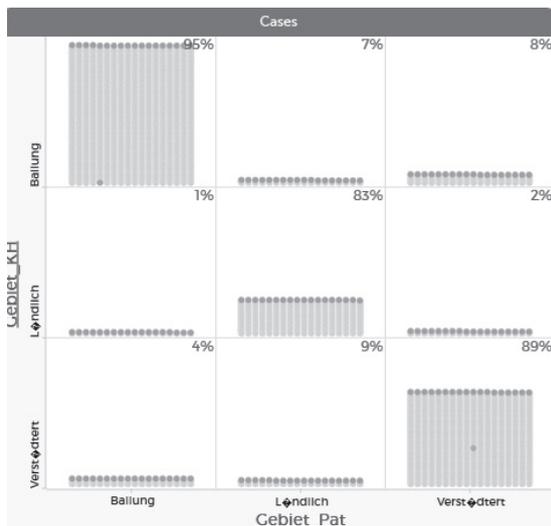


Abb. 7: Verteilung des Merkmals „Gebiet\_Pat“ getrennt nach Region des Krankenhauses in CODAP mit Spaltenprozenten

ländlichen Region in einem solchen Krankenhaus, aber 95 % der Patienten aus einer Ballungsregion befinden sich auch in einem Krankenhaus dieser Region.

Immerhin 16 % der Patienten aus einer ländlichen Region gehen in ein Krankenhaus der Region Stadt oder Ballung, in umgekehrter Richtung gehen nur 1 %, bzw. 2 % der Patienten aus diesen Regionen in ein Krankenhaus der ländlichen Region. Dies wirft nun weiterführende Fragen auf, vor allem danach, was Gründe für diese Beobachtung sind. Möglicherweise spielen hier Drittvariablen eine Rolle oder es lassen sich Gründe im Kontext vermuten. Eine mögliche Hypothese könnte sein, dass Krankenhäuser in Ballungsregionen größer sind und deshalb mehr Fachabteilungen haben und Patienten aus ländlichen Regionen eher in Ballungsregionen verlegt werden. Dieser Hypothese kann man anhand der Daten nachgehen, indem man den Aufnahmeanlass hinzufügt und nur nach der Ausprägung Verlegung filtert, wie in Abbildung 8 geschehen.

In Abbildung 8 erkennt man, dass 31 % der Patienten aus einer ländlichen Region in eine Ballungsregion verlegt werden, umgekehrt betrifft dies nur 1 % der Patienten. Die aufgestellte Hypothese wird also durch diese Darstellung gestützt und kann im Folgenden weiter exploriert werden. Dabei kann auch der Frage nachgegangen werden, welche Rolle andere Merkmale spielen, wie beispielsweise, ob eine Operation durchgeführt wurde oder ob die Verweildauer ebensolche Unterschiede bezüglich der Region zeigt. Auch hierfür lassen sich mit Hilfe von Kontextwissen Vermutungen aufstellen und datenbasiert begründen.

Wir haben in diesem Abschnitt nur ansatzweise das Explorationspotential der Krankenhausdaten und das Werkzeugpotential von CODAP angedeutet. Im weiteren Unterrichtsverlauf können weiterführende Explorations mit CODAP vorgenommen werden, so zum Beispiel auch Vergleiche von Verteilungen oder die Exploration des Zusammenhangs zweier numerischer Variablen. Generell kann dieses Mini-Projekt sowie die Exploration dieser Daten in CODAP unseres Erachtens gut in der Jahrgangsstufe 9/10 thematisiert werden. Aufgrund der Reichhaltigkeit der Daten und des Explorationspotentials von CODAP können sowohl zu gegebenen Fragestellungen als auch zu von den Schülerinnen und Schülern selbstgenerierten Fragestellungen substantielle statistische Explorations vorgenommen werden. Diese können vom Erstellen einer Verteilung eines kategorialen Merkmals (z. B. Aufnahmeanlass, Abb. 5) und dem Beschreiben und Interpretieren dieses Diagramms reichen bis hin zu Untersuchungen von Zusammenhängen zwischen zwei Variablen (anhand verschiedener Prozente) wie z. B. in den Abbildungen 7 und 8.

#### 4 Zusammenfassung

In diesem Artikel haben wir exemplarisch zwei Mini-Projekte und mögliche Explorations, wie sie im Unterricht realisiert werden können, vorgestellt. Beide Projekte unterscheiden sich in ihrer Komplexität. Die Exploration des Unfallkalenders mit dem Tool des Statistischen Bundesamts ist stärker gelenkt und kann somit gut in den unteren Klassenstufen der Sekundarstufe I genutzt werden. Die Exploration der Krankenhausdaten mit der kostenlosen Anwendung

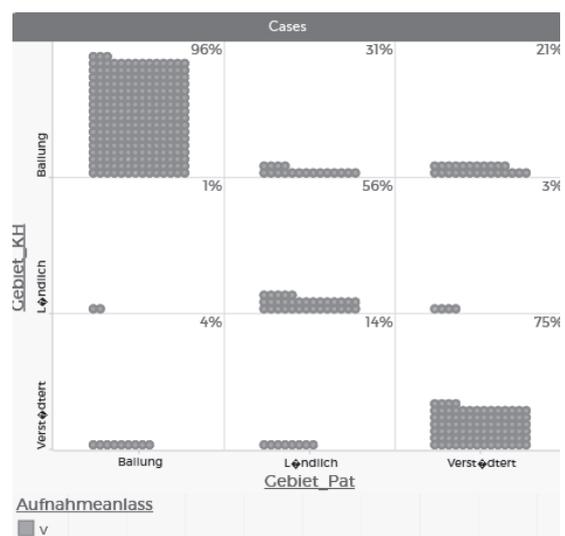


Abb. 8: Verteilung des Merkmals „Gebiet\_Pat“ getrennt nach Region des Krankenhauses in CODAP mit Spaltenprozenten gefiltert nach dem Aufnahmeanlass „Verlegung“

CODAP bietet aufgrund des Tools und des multivariaten Datensatzes größere Entfaltungsmöglichkeiten für Schülerinnen und Schüler und ist deshalb eher für die oberen Klassenstufen der Sekundarstufe I geeignet. Dennoch war es uns wichtig, Entfaltungsmöglichkeiten auch unabhängig vom Tool anzubieten. So sind wesentliche Konzepte dieser Mini-Projekte, dass sie zunächst mit einer offenen Leitfrage starten („Wie entwickeln sich die Zahlen der Verkehrsunfälle im Laufe der Jahre?“ bzw. „Wie ist die Lage der Krankenhäuser in Deutschland?“) und dann durch geführte Unterfragen begleitet werden (z. B. „Welches ist der häufigste Aufnahmeanlass?“). Schließlich wird den Schülerinnen und Schülern auch die Exploration der Daten nach selbstgewählten Fragestellungen ermöglicht („Stell zwei eigene statistische Fragen zur Lage der Krankenhäuser in Deutschland und beantworte diese mit geeigneter statistischer Auswertung“).

Beide Mini-Projekte wurden bereits in einem Seminar „Statistische Allgemeinbildung im Mathematikunterricht“ für Lehramtsstudierende Mathematik an Haupt-, Real- und Gesamtschulen erfolgreich erprobt (siehe Frischemeier, Podworny und Biehler in diesem Heft) und sollen mit Blick auf den Einsatz in der Schule zum einen die zivilstatistischen Kompetenzen von Schülerinnen und Schülern der Sekundarstufe I fördern und zum anderen Brücken zum fächerübergreifenden Unterricht zwischen Mathematik und Sozialwissenschaften bauen.

Die beiden Mini-Projekte „Unfallstatistik“ und „Krankenhäuser in Deutschland“ sind nur ein Teil von einer Fülle an Materialien, Mini-Projekten und Datensätzen für den Mathematikunterricht, die im Rahmen des EU-Projekts ProCivicStat erarbeitet und zusammengestellt wurden (<http://iase-web.org/islp/pcs>).

## Anmerkungen

- 1 Hintergrundinformationen zum Datenerhebungsverfahren, zur Variablenliste sowie zur Lage der Krankenhäuser in Deutschland finden sich u. a. hier: <https://tinyurl.com/y8zs76ss> (aufgerufen am 7.7.2018)
- 2 Zur Lage der Krankenhäuser im ländlichen Raum, siehe z. B. <https://tinyurl.com/y9bjue8w> (aufgerufen am 7.7.2018)
- 3 Über eine Einbindung der Google-Cloud können Datensätze für andere über einen geteilten Link zugänglich gemacht werden, so wie hier vorliegend geschehen.

## Literatur

- Biehler, R.; Ben-Zvi, D.; Bakker, A.; Makar, K. (2013): Technology for Enhancing Statistical Reasoning at the School Level. In: M. A. Clements, A. J. Bishop, C. Keitel-Kreidt, J.; Kilpatrick; Leung, F. K.-S. (Hrsg.): *Third International Handbook of Mathematics Education* (S. 643–689). New York: Springer Science + Business Media.
- Biehler, R.; Frischemeier, D.; Podworny, S. (2018): Civic Engagement in Higher Education: A university course in civic statistics for mathematics preservice teachers. In: *Zeitschrift für Hochschulentwicklung*, 13(2), S. 169–182.
- Biehler, R.; Hofmann, T.; Maxara, C.; Prömmel, A. (2011): Daten und Zufall mit Fathom: Unterrichtsideen für die SI mit Software-Einführung. Braunschweig: Schroedel.
- Engel, J.; Gal, I.; Ridgway, J. (2016): Mathematical Literacy and Citizen Engagement: The Role of Civic statistics. Paper presented at the 13th International Congress on Mathematical Education, Hamburg.
- Friel, S. N.; Curcio, F. R.; Bright, G. W. (2001): Making sense of graphs: Critical factors influencing comprehension and instructional implications. In: *Journal for Research in Mathematics Education*, 32(2), S. 124–158.
- Frischemeier, D.; Biehler, R.; Engel, J. (2016): Competencies and dispositions for exploring micro data with digital tools. Paper presented at the IASE Roundtable Conference „Promoting understanding of statistics about society“, Berlin.
- Gal, I. (2002): Adults’ Statistical Literacy: Meanings, Components, Responsibilities. In: *International Statistical Review*, 70(1), S. 1–51. doi:10.2307/1403713
- Konold, C.; Miller, C. (2011): *TinkerPlots 2.0*. Emeryville, CA: Key Curriculum Press.
- Ridgway, J. (2016): Implications of the Data Revolution for Statistics Education. In: *International Statistical Review*, 84(3), S. 528–549. doi:10.1111/insr.12110
- Schiller, A. (2017): The Importance of Statistical Literacy for Democracy – Civic-Education by Statistics. Paper presented at Challenges and Innovations in Statistics Education, Szeged.

## Anschrift der Verfasser

Susanne Podworny  
 Universität Paderborn  
 Institut für Mathematik  
 Warburger Straße 100  
 33098 Paderborn  
 podworny@math.upb.de

Daniel Frischemeier  
 Universität Paderborn  
 Institut für Mathematik  
 Warburger Straße 100  
 33098 Paderborn  
 dafr@math.upb.de