

# Geoinformationen – die „Veredelung“ räumlich-statistischer Daten im digitalen Zeitalter

Moderne Informations- und Kommunikationstechnologien erfassen und nutzen oft Daten mit einem Raumbezug, wodurch relevante Geoinformationen für ökonomische, ökologische und gesellschaftliche Planungs- und Entscheidungsprozesse generiert werden. Geografische Informationssysteme (GIS) werden daher direkt oder indirekt in vielen Branchen und Disziplinen aus Wirtschaft, Verwaltung, Forschung und Lehre oder auch im privaten Bereich eingesetzt und sind fester Bestandteil des digitalen Zeitalters. Was verbirgt sich aber hinter GIS und Schlagworten wie „Geodaten“ oder „Georeferenz“ bzw. „Geokodierung“? Und wie werden aus Rohdaten maßgeschneiderte und nutzerorientierte Geoinformationen, die räumliche Muster und Beziehungen aufzeigen? In diesem Beitrag werden die wesentlichen Begriffe aus dem Umfeld der Geoinformationen erläutert. Anschließend werden Möglichkeiten skizziert, um verschiedene räumlich-statistische Datentypen aufzubereiten, zu analysieren und zu visualisieren. **Von Sarah Scholze**

## Digitale Geowelten und die Rolle des „Wo“

Mobile Endgeräte ausgestattet mit Global Positioning System (GPS)-Technologie, vernetzte Sensoren- und Fernerkundungstechnik oder das sogenannte GeoWeb (Google Maps, Open Street Map etc.) – all diese Informations- und Kommunikationstechnologien sind im digitalen Zeitalter stets präsent. Dabei werden permanent Daten erfasst und ausgetauscht, die oft über einen Raumbezug verfügen<sup>1)</sup> und somit auch Informationen zur Lage, Beschaffenheit und Nutzung eines Ortes bzw. Standortes liefern. Geodaten, also jene Daten mit Raumbezug, sind eine wesentliche Grundlage, um auftretende Phänomene, Prozesse und Entwicklungen zu beschreiben und nachzuvollziehen. Mit Hilfe von Geodaten können beispielsweise optimierte Routen in Bezug auf die Entfernung, Fahrtzeit oder Kosten in einem Straßennetz ermittelt werden. Unternehmen optimieren dadurch ihre Logistik und damit die Versorgung ihrer Kundschaft mit Produkten und Dienstleistungen. Im öffentlichen Sektor gibt es ebenfalls vielfältige Einsatzmöglichkeiten von Geodaten, wie beispielsweise beim geplanten Bau einer



Umgehungsstraße. Mit einer Kombination aus kleinräumigen Bevölkerungsdaten und Kartierungen von Naturschutzflächen kann identifiziert werden, wie viele Anwohnende am Ortsrand durch den Bau zusätzlich belastet werden und ob es naturschutzrechtliche Bedenken gibt.

Im Vordergrund steht stets das Ziel, anhand von Daten und Zahlen in zeitlicher und räumlicher Auflösung, die Welt zu verstehen, daraus Erkenntnisse zu gewinnen und entsprechende Entscheidungen für Handlungen und Maßnahmen zu treffen. Komplexe und vielschichtige Sachverhalte mittels statistischer Daten

1) In Publikationen wird oft unterstellt, dass 80 % aller Daten bzw. Informationen über einen Raumbezug verfügen. Einer Forschungsstudie von Hahmann und Burghardt (2012) zufolge kann der Anteil der raumbezogenen Daten nicht eindeutig quantifiziert werden.

zu beschreiben, räumlich aufzubereiten und dem Datenkonsumenten sowie einer breiten Öffentlichkeit verständlich zu kommunizieren, sind auch wichtige Aufgabenfelder der amtlichen Statistik und können mit geografischen Informationssystemen (GIS) effektiv umgesetzt werden. GIS ermöglichen es, räumliche Informationen („Wo ist etwas?“) mit erklärenden sachlichen Informationen („Was oder wie ist etwas?“ und ggf. „Wann ist etwas?“) digital zu verknüpfen, zu analysieren und in geeigneter Form zu visualisieren.

### GIS – die „Bühne“ der Geodaten

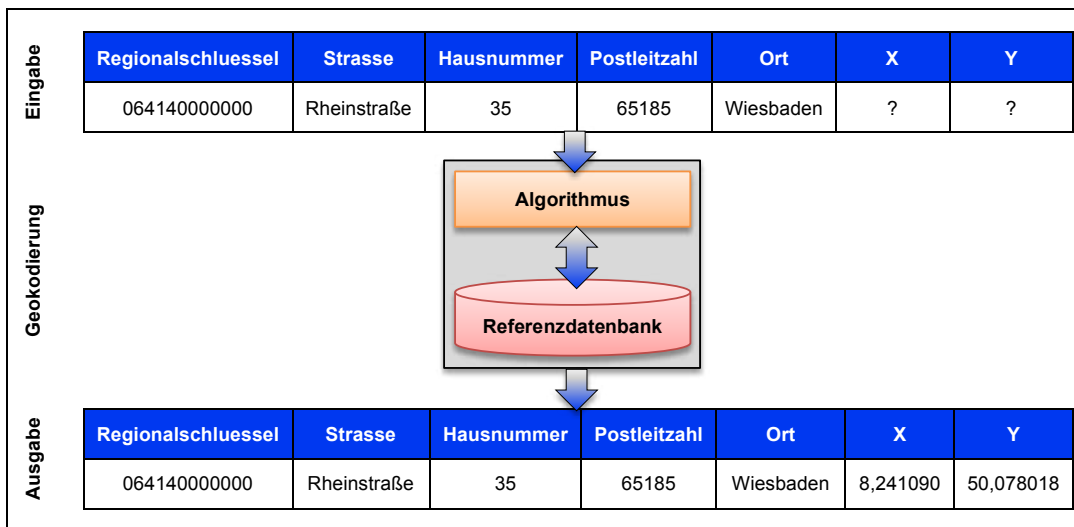
Eine große Anzahl von Daten der amtlichen Statistik basiert auf Einheiten mit einer räumlichen Komponente, wodurch sie mit einem Standort oder einem Gebiet verknüpft sind. Die statistischen Daten werden oft mittels einer Datenbank oder eines Statistikprogramms, wie SAS, verarbeitet und aufbereitet. Solche Systeme berücksichtigen jedoch nicht explizit die räumliche Dimension der Geodaten. GIS bieten dagegen Funktionalitäten, um diese Geodaten effizient zu erfassen, zu verarbeiten, zu analysieren und zu präsentieren (EVAP). Die grundlegenden Funktionen in einem GIS basieren auf dem Layer-Prinzip, das verschiedene Formate von Geodaten miteinander kombiniert, überlagert oder verrechnet. Dabei werden Geobasis- und Geofachdaten unterschieden. Geobasisdaten enthalten Informationen zur Lage auf der Erdoberfläche und repräsentieren Objekte aus der realen Welt, wie Straßen, Gebäude oder Grundstücke. Sie werden insbesondere durch das amtliche Vermessungswesen im Rahmen ihres gesetzlichen Auftrages erfasst und bereitgestellt. Zu den Produkten gehören beispielsweise das „Amtliche Topographisch-Kartographische Informationssystem (ATKIS)“ oder das „Amtliche Liegenschaftskatasterinformationssystem (ALKIS)“. Zudem existieren im GeoWeb auch nutzergenerierte Geobasisdaten, die unter dem Stichwort „Volunteered Geographic Information (VGI)“ zusammengefasst werden (Goodchild, 2007). Ein bekanntes VGI-Projekt ist die Open Street Map (OSM), bei der die Daten von Freiwilligen weltweit erhoben und für jedermann zur Verfügung gestellt werden.

Die Geobasisdaten liegen entweder als Raster- oder Vektordaten in digitaler Form vor. In einem Rasterdatenmodell werden die geografischen Objekte in Zellen oder Pixeln in einer gleichmäßigen Matrix angeordnet und z. B. in den Formaten GeoTiff oder JPEG gespeichert. Die Eigenschaften von geografischen Objekten (z. B. Hangneigung) sind in den Rasterzellen als Farb- oder Zahlenwerte kodiert. Klassische Beispiele für die Verwendung von Rasterdaten sind Luft- und Satellitenbilder. Vektordaten speichern die geografischen Objekte anhand der geometrischen Elemente Punkt, Linie und Fläche in GIS-spezifischen Formaten, wie Shapefile oder FileGeodatabase. Im Vektorformat werden beispielsweise die Verwaltungsgrenzen oder die geografischen Raster dargestellt. Mit Hilfe der Geobasisdaten können Informationen zu Entfernungen, Flächengrößen oder Nachbarschaftsbeziehungen ermittelt werden. Die thematischen Eigenschaften von geografischen Objekten werden anhand der Geofachdaten beschrieben, deren Inhalte beispielsweise aus den Bereichen der amtlichen Statistik, der Epidemiologie oder dem Marketing kommen können.

Interaktiv verbunden werden die Geobasis- und Geofachdaten anhand ihrer Georeferenz, bei der es sich entweder um einen direkten oder indirekten Raumbezug handelt. Der direkte Raumbezug beschreibt eine Position auf der Erdoberfläche und wird durch Geokoordinaten angegeben, die z. B. durch das GPS erfasst werden. Die meisten amtlichen Statistiken weisen jedoch einen indirekten Raumbezug auf, z. B. in Form eines Regionalschlüssels (06414000000) oder durch die Anschrift (Straße, Hausnummer, Postleitzahl und Ort). Der indirekte Raumbezug kann in einen direkten Raumbezug mit Hilfe der Geokodierung<sup>2)</sup> umgewandelt werden. Dadurch wird jede beschreibende Information zu einem räumlichen Punkt transformiert und kann auf einer Karte dargestellt werden (Goldberg, 2008: 5). In der einfachsten Form werden bei der Geokodierung 2 Datensätze, der zu geokodierende Datensatz und der Referenzdatensatz mit den Geokoordinaten, miteinander verglichen. Bei einem „Treffert“ werden die Geokoordinaten dem zu geokodierenden Datensatz zugewiesen (vgl. Abb. 1).

2) Die Begriffe Georeferenzierung und Geokodierung werden in der Literatur sowohl synonym als auch in unterschiedlicher Terminologie verwendet. Goldberg (2008: 3) bezeichnet die Georeferenzierung allgemein als die Zuweisung zu einem Raumbezug, während die Geokodierung als Methode der Georeferenzierung angesehen wird. Weitere Methoden der Georeferenzierung sind die Rektifizierung im Kontext von Luft- und Satellitenbildern sowie das Geotagging in Bezug auf das digitale Foto.

Abbildung 1: Schematischer Prozess der Geokodierung



Bei der Geokodierung sind korrekte Angaben der Adressen eine wesentliche Voraussetzung für ein gutes qualitatives Ergebnis. Fehlerhafte oder unterschiedliche Bezeichnungen müssen daher zuvor aufgedeckt und korrigiert werden. Die Geokodierung ist somit auch unter Qualitätsaspekten ein grundlegender Arbeitsschritt bei der Aufbereitung von adressenbasierten Statistiken. Anhand der georeferenzierten Daten können dann in einem GIS umfassende räumliche Analysen für viele Anwendungsbereiche durchgeführt werden.

### Räumliche (Daten-) Analyse – das Herzstück eines GIS

GIS beinhalten unterschiedliche Werkzeuge zur Datenerfassung, kartografischen Bearbeitung sowie zur räumlich-statistischen Analyse. Die räumliche Analyse gilt dabei als das Herzstück eines GIS, da neue Informationen für entscheidungsunterstützende Abläufe und Erkenntnisse aus den vorliegenden Geodaten generiert werden können (Longley et al., 2015: 291). Ein bekanntes historisches Beispiel für eine gewinnbringende räumliche Analyse ist die des Cholera-Ausbruchs im Jahr 1854 in London. Der englische Arzt Dr. John Snow kartierte sowohl die Wohnorte von Cholera-Opfern als auch die dort vorhandene örtliche Wasserversorgung. Dadurch konnte er die Kontamination einer öffentlichen Pumpe in der Broad Street als Ursache für den Ausbruch der Krankheit aufzufindig machen und unter Einbeziehung des räumlichen Kontexts die Epi-

demie eindämmen. Dies veranschaulicht, dass durch die Verknüpfung unterschiedlicher Daten und deren Visualisierung komplexe räumliche Zusammenhänge aufgedeckt werden können, die sonst nicht ersichtlich gewesen wären.

Die Beantwortung der Fragen „Wo befindet sich etwas?“, „Wie steht etwas miteinander in Zusammenhang?“, „Welche Bedeutung hat dies und welcher mögliche Handlungsbedarf ergibt sich daraus?“ basiert auf einer Kombination unterschiedlicher Methoden und Techniken (z. B. lagebezogene Abfragen und Selektion, Verschneidung oder Interpolationen), die sich auf die räumliche Dimension beziehen (Longley et al., 2015: 291). Bei der Auswahl räumlicher Analysewerkzeuge spielen die verschiedenen räumlich-statistischen Datentypen eine grundlegende Rolle. Diese werden nach Cressie (1993) in kontinuierliche Daten („Continuous Data“), Punktmusterdaten („Point Pattern Data“) und aggregierte Regionaldaten („Areal Data“) klassifiziert.

Kontinuierliche Daten beschreiben Phänomene, die an jedem Punkt eines Gebietes existieren, wie der Niederschlag oder die Bevölkerungsdichte. In der Realität liegen für solche Ereignisse allerdings nur Stichprobenwerte für bestimmte Standorte auf der Erdoberfläche vor. Mittels geostatistischer Verfahren, wie Kriging, können Werte an den unbekannt Positionen aus den erhobenen Daten geschätzt werden, um anschließend die räumliche Variation in einem Untersuchungsgebiet zu ermitteln. Das Ergebnis einer solchen räumlichen Interpolation ist eine

kontinuierliche thematische Oberfläche, deren Darstellung beispielsweise in Wetterkarten vorkommt.

Punktmusterdaten und aggregierte Regionaldaten repräsentieren sogenannte diskrete Phänomene, die räumlich deutlich voneinander abgegrenzt werden können. In der amtlichen Statistik werden Ereignisse oft auf Basis administrativer Einheiten aggregiert, wodurch die tatsächliche räumliche Verteilung der Werte im Raum entweder nur annähernd oder gar nicht wiedergegeben wird. In der Realität konzentrieren sich z. B. Straßenverkehrsunfälle auf bestimmte Gebiete innerhalb einer administrativen Einheit. Die Kerndichteschätzung (KDE: Kernel Density Estimation) ist eine Möglichkeit, die kleinräumige Verteilung von punktbasierten Ereignissen zu ermitteln. Dabei wird anhand der zugrundeliegenden Punktmusterdaten und den jeweiligen Distanzen zueinander eine kontinuierlich eingefärbte Rasteroberfläche erzeugt, die die Dichteverteilungen in einem Gebiet darstellt. Kerndichtekarten können z. B. für räumliche Analysen von Straftaten, Krankheitsfällen oder auch Straßenverkehrsunfällen eingesetzt werden, wobei solche Ereignisse stets nur lokal an einem bestimmten Ort auftreten und somit nicht explizit kontinuierliche Phänomene repräsentieren. Kerndichtekarten können jedoch erste Hinweise z. B. auf lokale Unfallschwerpunkte (sogenannte Hot Spots) oder räumliche Zusammenhänge zwischen den Unfällen und physischen Strukturmerkmalen eines Gebietes liefern, um daraus Präventionsmaßnahmen abzuleiten.

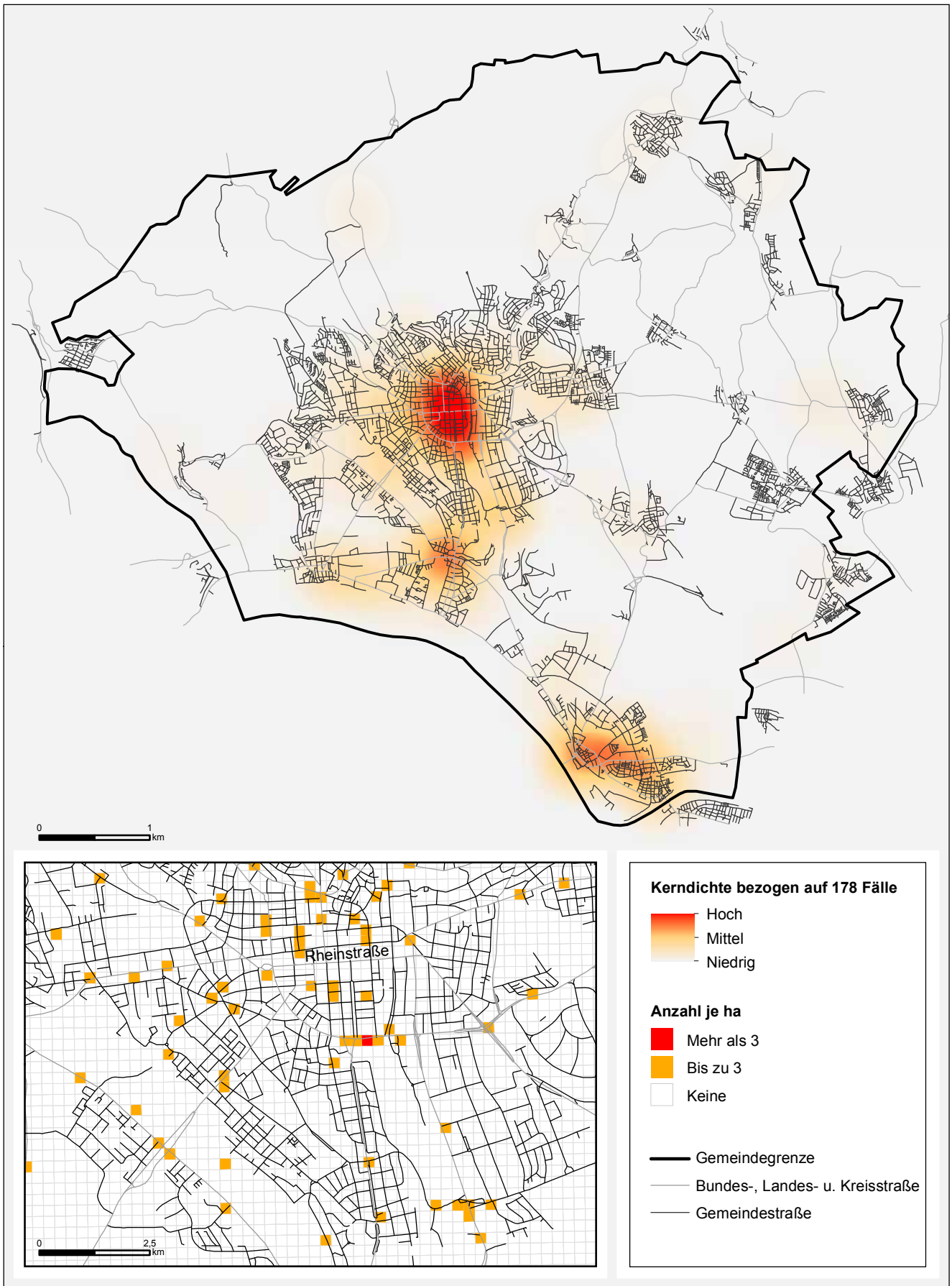
Abbildung 2 stellt die Straßenverkehrsunfälle mit Beteiligung von Fahrrad Fahrenden in der Stadt Wiesbaden auf Basis einer solchen Kerndichteberechnung dar. Der rote Bereich in der Innenstadt von Wiesbaden, im Einzugsgebiet der Biebricher Straße und im Kern von Mainz-Kastel zeigt eine hohe Konzentration von Unfällen, während der orangefarbene und graue Bereich eine geringe Dichte in den jeweiligen Gebieten aufweist. Im zugehörigen Kartenausschnitt wird der Unfallschwerpunkt in der Wiesbadener Innenstadt auf Basis des geografischen Rasters mit einer Seitenlänge von 100 m dargestellt, wodurch detaillierte Aussagen über das Auftreten von Unfällen vor Ort möglich sind.

Weitere Vorzüge von geografischen Rastern in der amtlichen Statistik werden bei Scholze (2015) diskutiert.

Räumliche Cluster auf Basis von aggregierten Regionaldaten können anhand von Hot-Spot-Analysen identifiziert werden. Dabei wird mittels räumlich-statistischer Verfahren (z. B. Gi\*-Statistik) überprüft, inwieweit benachbarte Gebiete in Bezug auf die zu untersuchenden statistischen Ereignisse miteinander korrelieren (Anselin, 1995; Getis und Ord, 1992). Abbildung 3 visualisiert den Anteil der Einpersonenhaushalte je km<sup>2</sup> in % im Regierungsbezirk Darmstadt. Die Darstellung im kleinen Kartenausschnitt vermittelt einen ersten Eindruck zur räumlichen Verteilung, welche jedoch abhängig von den definierten Klassengrenzen ist. Die resultierenden Ergebnisse der Hot-Spot-Analyse basieren auf der Berechnung der Gi\*-Statistiken und geben Auskunft darüber, ob sich die lokalen Werte benachbarter geografischer Rasterzellen statistisch signifikant vom Durchschnitt jener im Regierungsbezirk unterscheiden. Die roten Bereiche, wie etwa in den Städten und Umgebungen von Frankfurt am Main, Offenbach am Main oder Wiesbaden, kennzeichnen dabei die lokalen Hot Spots. Diese Gebiete weisen eine hohe Anzahl von Einpersonenhaushalten auf und sind gleichzeitig auch von Gebieten mit einer ähnlich hohen Anzahl von Einpersonenhaushalten umgeben. Umgekehrt handelt es sich bei den blauen Bereichen um lokale Gebiete mit niedriger Anzahl von Einpersonenhaushalten, die gleichzeitig auch von Gebieten mit einer ähnlich niedrigen Anzahl von Einpersonenhaushalten umgeben sind (sogenannte Cold Spots), wie etwa die Gemeinden Brensbach, Höchst im Odenwald, Fischbachtal oder Fränkisch-Crumbach. Mit einer Wahrscheinlichkeit von 90 % oder mehr sind diese auftretenden räumlichen Zusammenhänge zwischen dem Gebiet und dem Merkmal „Einpersonenhaushalte“ nicht zufälliger Natur. Gebiete mit hohen oder niedrigen Werten sind dabei nicht zwingend statistisch signifikante Hot oder Cold Spots, sondern können auch einer zufälligen räumlichen Verteilung unterliegen (siehe gelber Bereich in der Karte).

Einsatzmöglichkeiten für Hot-Spot-Analysen gibt es in vielen unterschiedlichen Bereichen, wie

Abbildung 2: Straßenverkehrsunfälle mit Beteiligung von Fahrrad Fahrenden<sup>1)</sup> in Wiesbaden 2015

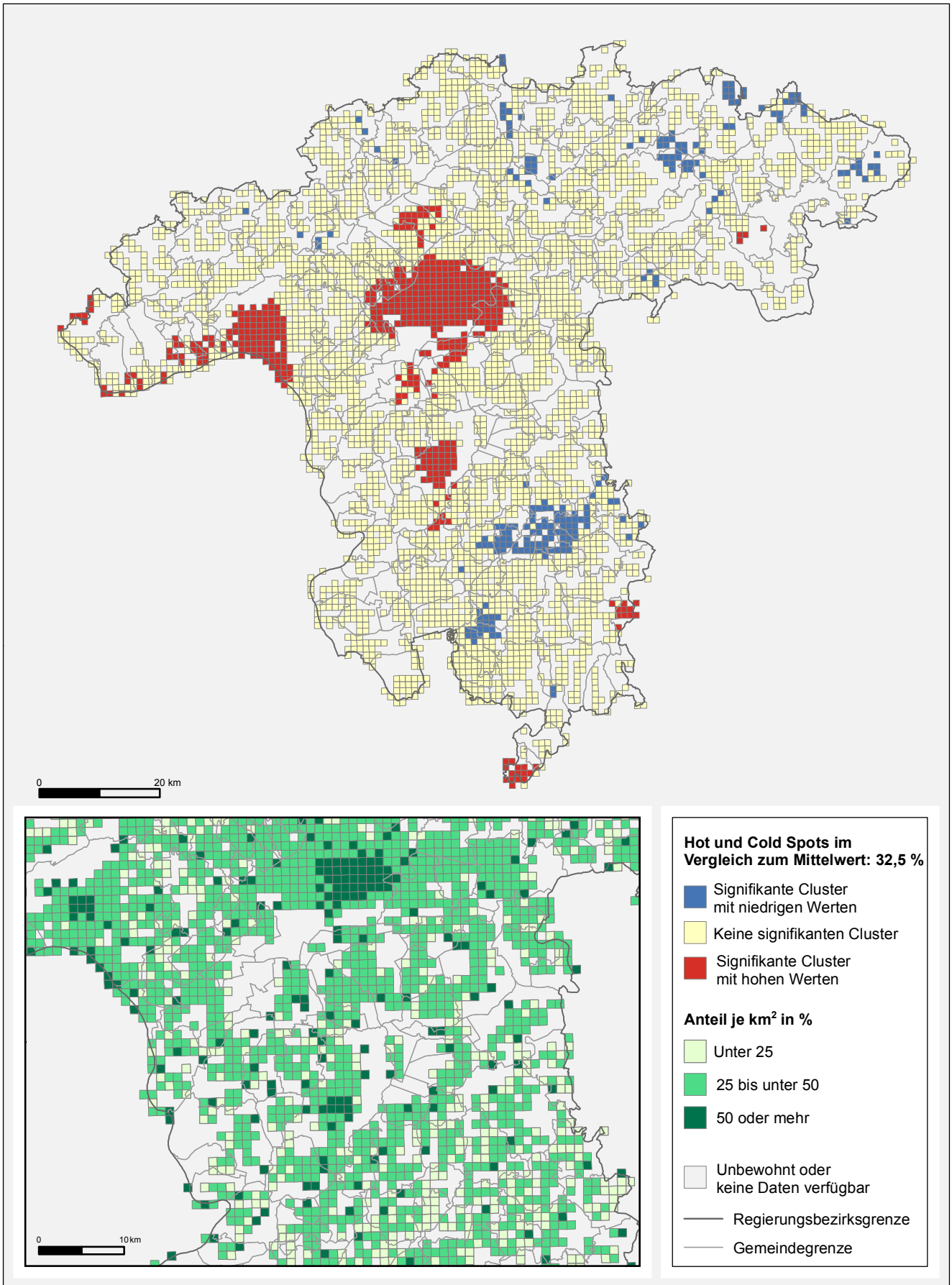


1) Fahrrad und Pedelec ohne Versicherungskennzeichen

© Hessische Verwaltung für Bodenmanagement und Geoinformation, 2016  
© Hessisches Statistisches Landesamt, Wiesbaden, 2017  
Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.



Abbildung 3: Einpersonenhaushalte im Regierungsbezirk Darmstadt zum 9.5.2011 auf Basis 1-km-Rasterzellen<sup>1)</sup>



1) Die statistische Signifikanz basiert auf den Konfidenzintervallen 90 %, 95 % und 99 %. Diese sind in der kartografischen Darstellung zusammengefasst.

beispielsweise der Kriminalitätsanalyse, Epidemiologie, Wirtschaft, Demografie oder auch bei Wahlen. Die räumliche Analyse soll Aufschluss darüber geben, wie sich statistische Ereignisse räumlich verteilen und ob es eventuell Cluster, d. h. Gebiete mit einer relativen Konzentration von Merkmalen, gibt. Eine entscheidende Rolle spielt dabei auch die Visualisierung, um die räumlichen Ergebnisse zu präsentieren und darüber hinaus mögliche Einflussfaktoren gezielt vor Ort zu identifizieren, also die Frage nach dem „Warum“ zu klären.

## Digitale Karten als „Gala“ der Geoinformationen

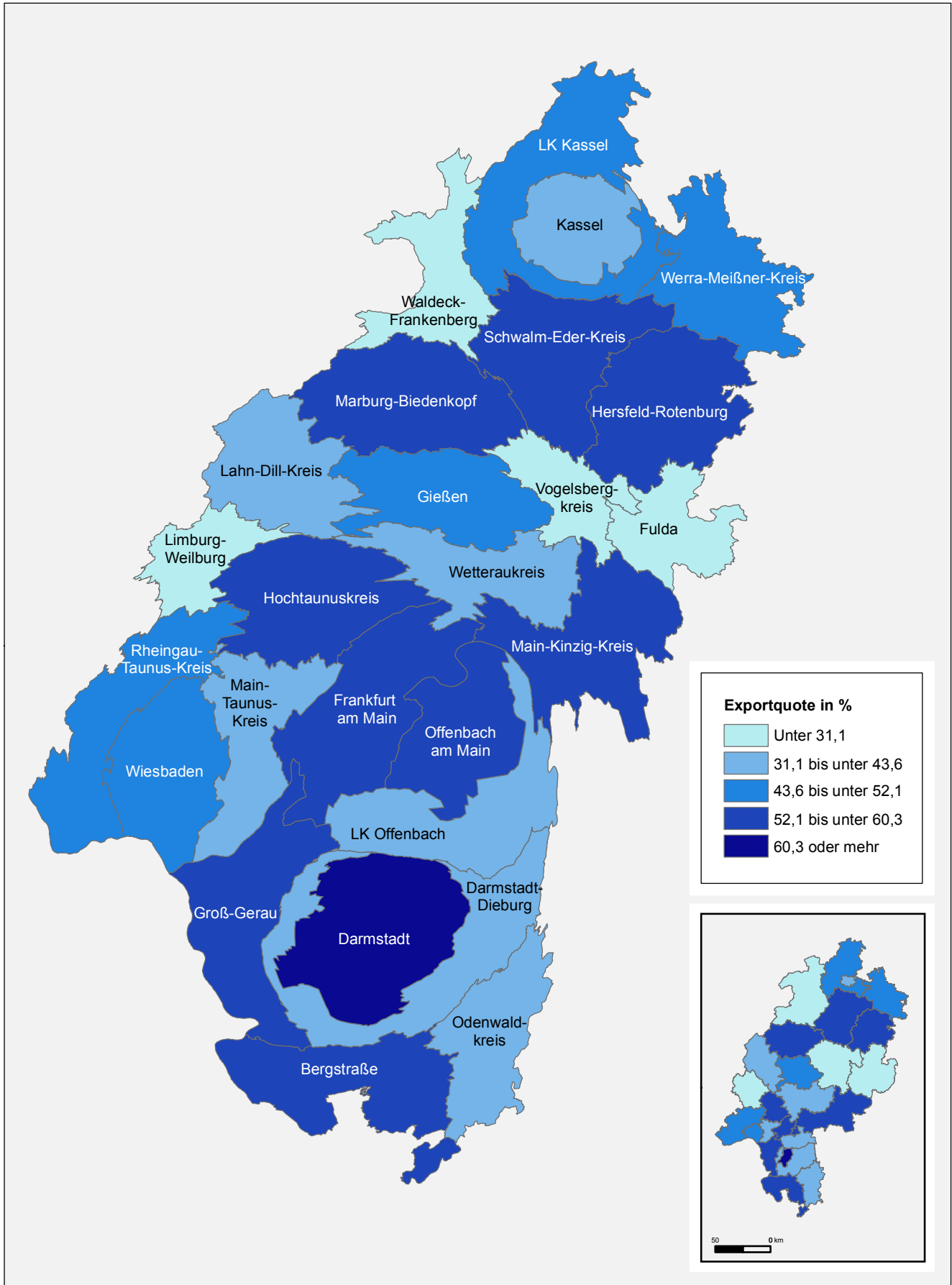
Geodaten weisen eine komplexe Datenstruktur auf und beinhalten abstrakte Sachverhalte, weshalb räumliche Muster und Trends in Form von regional aufgeschlüsselten Tabellen nicht bzw. nur schwer „sichtbar“ werden. Je größer und unstrukturierter dabei die Datenmengen sind, desto komplizierter ist die Vermittlung der relevanten Informationen. Die Aufbereitung des Datenmaterials muss daher so erfolgen, dass die räumlichen Informationen schnell und intuitiv interpretiert werden können. Eine effektive Möglichkeit, Geoinformationen zu vermitteln, ist die Visualisierung in Form von thematischen Karten. Dabei werden die statistischen Werte (z. B. Exportquote) mit den Lageinformationen (z. B. Landkreis) verlinkt, wodurch beispielsweise Vergleiche mit benachbarten Kreisen durchgeführt werden können. Am häufigsten wird die Choroplethenkarte eingesetzt, bei der die statistischen Werte klassifiziert und mit farblichen Abstufungen dargestellt werden. Die Klassifikation der Werte beeinflusst das Erscheinungsbild der resultierenden Karte und somit auch die zu vermittelnde Aussage, da unterschiedlich festgelegte Intervalle auch verschiedene räumliche Verteilungsmuster und Zusammenhänge suggerieren (Monmonier, 2005: 218 f.). Die Auswahl einer Klassifikationsmethode ist somit einerseits abhängig von der zugrundeliegenden Datenstruktur, andererseits vom Zweck sowie der zugehörigen Botschaft der Karte. Die Klassifikation mittels „Jenks Natürlicher Unterbrechungen“ ist beispielsweise eine gebräuchliche Methode für ungleichmäßig verteilte Datenwerte und geeig-

net, um Cluster aufzuzeigen. Die Werte werden dabei den Klassen so zugeordnet, dass die Unterschiede innerhalb der Klassen minimiert und zwischen den Klassen maximiert werden. Weitere gängige Klassifikationsmethoden sind „Gleiches Intervall“, „Quantil“ oder „Standardabweichung“ (Mitchell, 1999: 48 ff.). Die definierten Klassenintervalle werden in Farbskalen übersetzt, die die Aussagen der Karte unterstützen sollen (Harrower und Brewer, 2003: 27). Sequenzielle Farbverläufe von „hell“ bis „dunkel“ symbolisieren „niedrige“ bzw. „hohe“ metrische Werte und assoziieren eine logische Reihenfolge. Eine Datenreihe mit positiven und negativen Werten oder einem Schwellenwert wird anhand divergierender kontrastierender Farbschemata dargestellt, um entgegengesetzte Richtungen zu verdeutlichen. Qualitative Farbschemata suggerieren hingegen aufgrund unterschiedlicher Farben, dass kein Zusammenhang zwischen den Kategorien besteht und sind daher für Themen, wie z. B. Religion oder Landnutzungsregionen, geeignet<sup>3)</sup>.

Choroplethenkarten basieren i. d. R. auf einer Kartengrundlage mit flächentreuer Projektion und vermitteln ein räumliches Bild der dargestellten Phänomene in den entsprechenden Bezugsflächeneinheiten, wie beispielsweise die Bevölkerung in Gemeinden je km<sup>2</sup>. Viele sozioökonomische Indikatoren, wie Wahlergebnisse in den einzelnen Wahlbezirken oder die Exportquote, beziehen sich jedoch nicht direkt auf die Fläche, sondern auf die thematischen Inhalte selbst. Burgdorf (2009) empfiehlt daher eine Ergänzung flächentreuer Choroplethenkarten mit amorphen Kartogrammen, deren Bezugsflächen proportional zu einer thematischen Variablen skaliert werden. Bei Kartogrammen werden der genaue Lagebezug und die topografische Orientierung vernachlässigt und die Flächen, Formen oder Entfernungen verzerrt. Die Verzerrung wird oft in Abhängigkeit eines Wertes der Objekte vorgenommen. Abbildung 4 zeigt die Exportquote in Hessen als amorphes Kartogramm und zum Vergleich im kleinen Kartenausschnitt als klassische Choroplethenkarte. Bei einem amorphen Kartogramm, das auf dem Algorithmus von Gastner und Newman (2004) basiert, werden jene Flächen, die einen hohen Wert aufweisen

3) Bei der Auswahl von geeigneten Farbskalen unterstützt die interaktive Anwendung ColorBrewer (siehe <http://colorbrewer2.org/>).

Abbildung 4: Exportquote im Verarbeitenden Gewerbe im 1. Halbjahr 2015



© GeoBasis-DE / BKG 2016 (Daten verändert)

© Hessisches Statistisches Landesamt, Wiesbaden, 2017  
 Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.



im Verhältnis zu anderen Flächen in der Karte größer dargestellt und umgekehrt. Die Größe der Wissenschaftsstadt Darmstadt spiegelt somit hohe Datenwerte der Variable „Exportquote“ wider. Das Merkmal Größe wird hier verwendet, um gezielt den Blick auf die thematischen Indikatoren zu lenken und dadurch Aufmerksamkeit zu erreichen. Das Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumordnung (BBR) veröffentlicht daher verschiedene Themen in Publikationen in Form von amorphen Kartogrammen, wie beispielsweise zur Entwicklung am Wohnungsmarkt (BBSR, 2016).

Die Visualisierung und Verbreitung von Geoinformationen erfolgt aufgrund zunehmender Vernetzung und Mobilität vermehrt in Web-Map-Applikationen, die von einfachen Kartendarstellungen mit Zoom- oder Pan-Funktionen bis hin zu komplexen interaktiven Karten reichen, bei denen der Nutzer direkt mit der Karte interagieren kann. Die Statistischen Ämter des Bundes und der Länder haben z. B. ausgewählte Ergebnisse des Zensus 2011 auf Basis von geografischen Rastern aufbereitet und bieten diese in einem interaktiven Atlas an<sup>4)</sup>. Dieser ermöglicht, neben einfachen WebGIS-Funktionen, auch die Bevölkerungszahl für nutzerdefinierte Gebiete anhand des implementierten Einwohnerrechners zu ermitteln. Eine weitere Kommunikationsform ist die sogenannte „Story Map“<sup>5)</sup>, bei der zu vermittelnde Themen und Botschaften in einer Kombination aus Karten, Grafiken, begleitendem Text und Multimedia-Elementen einer breiten Öffentlichkeit präsentiert werden. Dabei können auch komplexe und heterogene große Datenmengen klar und benutzerfreundlich aufbereitet werden, um räumliche Sachverhalte und relevante Informationen auf einen Blick zu erfassen.

### „Small is Smart“ – kleinräumige Informationen im „Rampenlicht“

Geografische Informationssysteme bieten mittels räumlicher Analysemethoden und Visualisierungstechniken effektive Möglichkeiten, sta-

tistische Daten zu veredeln und daraus „smarte“ Geoinformationen zu generieren. Dadurch können räumliche Variationen, Muster und Trends identifiziert, beschrieben und erläutert werden, um letztendlich erfolgsbringende Entscheidungen zu treffen. Voraussetzung dafür sind qualitativ hochwertige und relevante georeferenzierte Daten, die maßgeschneidert auf bestimmte Fragestellungen oder ein Projekt bereitgestellt werden. Dabei steht, neben dem Umfang, der Vielfalt und Aktualität der Daten, die verwendete räumliche Auflösung im Rampenlicht. In der amtlichen Statistik werden für ausgewählte Themenbereiche kleinräumige Daten auf Basis geografischer Raster ab einer Mindestgröße von 1 ha angeboten. Dazu regelt das E-Government-Gesetz (§ 14 E-GovG) die Georeferenzierung von statistischen Daten und Registern, wonach die Koordinaten ergänzt werden sollen. Dies ermöglicht einerseits detaillierte und präzise räumliche Analysen, andererseits spielt gerade bei räumlich hochaufgelösten statistischen Daten die Frage der Gewährleistung der Geheimhaltung eine zentrale Rolle. In diesem Kontext besteht die Möglichkeit, die kleinräumigen Daten mittels räumlich-statistischer Verfahren aufzubereiten und dem Datenkonsumenten ein auf dem Analyseergebnis basierendes kartografisches Endprodukt bereitzustellen. So veröffentlicht Statistics Canada in Kooperation mit der Canadian Community Health Survey (CCHS) beispielsweise kleinräumige Indikatoren zum Gesundheitswesen in thematischen Karten als Ergebnis einer Hot-Spot-Analyse<sup>6)</sup>, wodurch Rückschlüsse auf einzelne Personen nicht möglich sind. Eine Verschneidung von kleinräumigen Daten mit anderen Informationsinhalten aus öffentlichen oder privaten Quellen ist, im Hinblick auf maßgeschneiderte und „intelligente“ statistische Produkte für den Endnutzer, ebenfalls denkbar (Wonka, 2008: 116). Dies ermöglicht eine Erweiterung des Nutzerkreises für amtlich statistische Daten und erhöht somit auch den Marktwert von Geoinformationen.

**Sarah Scholze**; Tel: 0611 3802-282;  
E-Mail: sarah.scholze@statistik.hessen.de

4) Der interaktive Zensusatlas (inkl. Einwohnerrechner) ist online verfügbar unter <https://atlas.zensus2011.de>.

5) Beispiele für Story Maps können online unter <https://storymaps.arcgis.com/de/gallery/> eingesehen werden.

6) Thematische Karten mit Indikatoren aus dem Gesundheitswesen sind online verfügbar unter <http://www.statcan.gc.ca/pub/82-583-x/2011001/themat-eng.htm>.

## Literaturverzeichnis

- Anselin, L. (1995): Local Indicators of Spatial Association – LISA. In: *Geographical Analysis* 27(2), 93–115.
- BBSR – Bundesinstitut für Bau-, Stadt- und Raumforschung (Hrsg., 2016): *Neue Ansichten auf die Wohnungsmieten*, BBSR-Analysen KOMPAKT 08/2016, Bonn. Online unter: <http://www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/AnalysenKompakt/2016/ak-08-2016.html>, [Stand: 13.02.2017].
- Burgdorf, M. (2009): Kartogramme: aus der Form geraten oder auf den Punkt gebracht? In: *Informationen zur Raumentwicklung* (10/11), 689–699. Online unter: [http://www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/R/2009/10\\_11/Inhalt/DL\\_Burgdorf.pdf?\\_\\_blob=publicationFile&v=2](http://www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/R/2009/10_11/Inhalt/DL_Burgdorf.pdf?__blob=publicationFile&v=2), [Stand: 13.02.2017].
- Cressie, N. A. C. (1993): *Statistics for spatial data*, New York.
- Gastner, M. T. und Newman, M. E. J. (2004): Diffusion-based method for producing density-equalizing maps. In: *Proceeding of Academy for Science of the United States of America (PNAS)*, 101, 7499–7504. Online unter: <http://www.pnas.org/content/101/20/7499.full.pdf>, [Stand 03.02.2017].
- Getis, A. und Ord, J. K. (1992): The Analysis of Spatial Association by Use of Distance Statistics. In: *Geographical Analysis* 24(3), 189–206.
- Goldberg, D. (2008): *A Geocoding Best Practices Guide*, North American Association of Central Cancer Registries, Springfield.
- Goodchild, M. F. (2007): Citizens as Sensors: The World of Volunteered Geography. In: *GeoJournal* 69(4), 211–221.
- Hahmann, S. und Burghardt, D. (2012): Forschungsergebnisse zur Frage: Haben 80 % aller Daten einen Raumbezug? In: *gis.Science* (3), 101–108.
- Harrower, M. A. und Brewer, C. A. (2003): ColorBrewer.org: An Online Tool for Selecting Color Schemes for Maps. In: *The Cartographic Journal* 40(1), 27–37.
- Longley, P. et al. (2015): *Geographic Information Science and Systems*, New York.
- Mitchell, A. (1999): *The ESRI Guide to GIS Analysis, Volume 1: Geographic Patterns and Relationships*, Redlands.
- Monmonier, M. (2005): Lying with Maps. In: *Statistical Science* 20(3), 215–222.
- Scholze, S. (2015): Statistische Ergebnisse „im Quadrat“ – Geografische Raster am Beispiel des Zensus 2011. In: *Staat und Wirtschaft in Hessen* (5), 211–218.
- Wonka, E. (2008): Von der Amtsstatistik zu Geoinformationen. In: Dutter, R. (Hrsg.): *Festschrift – 50 Jahre Österreichische Statistische Gesellschaft*, Wien, 115–131.