Geografische Big Data Visualisierungen - Neue Design- und Interaktionskonzepte im Usability Vergleich

Research Paper

FH-Prof. Mag. DI Peter Hofer
FH Oberösterreich, Institut für Controlling, Rechnungswesen und Finanzmanagement, A-4400 Steyr, E-Mail: peter.hofer@fh-steyr.at

Stephan Karrer, B.A.
FH Oberösterreich, Institut für Controlling, Rechnungswesen und Finanzmanagement, A-4400 Steyr, E-Mail: stephan.karrer@fh-steyr.at

Katharina Ebner, M.A.
FH Oberösterreich, Institut für Controlling, Rechnungswesen und Finanzmanagement, A-4400 Steyr, E-Mail: k_ebner@gmx.at

Univ.-Prof. Dr. rer. nat. Hans-Christian Jetter M.Sc.
FH Oberösterreich, Institut für Human-Centered Computing, A-4232 Hagenberg
Universität Lübeck, Institut für Multimediale und Interaktive Systeme, D-23562 Lübeck, E-Mail: jetter@imis.uni-luebeck.de

Abstract
Dieser Artikel fokussiert auf die im Self-Service-Reporting von Big Data maßgeblichen geografischen Visualisierungen und deren Usability, im speziellen auf die häufig verwendeten Choropleth Maps und deren Kombinationen mit klassischen Geschäftsdiagrammen (Combined Visualizations). Die Ergebnisse unserer Online-Studie zeigen, dass Designvarianten in Form einer Kombination von klassischen Choropleth Maps mit Balken- bzw. integrierten Kreisdiagrammen eine signifikante Verbesserung der Effektivität, Effizienz und Zufriedenheit nach sich ziehen. Die Einbindung von Zoomfunktionen in die bestehenden Interaktionskonzepte, speziell bei einer Landmap mit integrierten Kreisdiagrammen, erhöht zusätzlich die Effektivität und Effizienz der Designvarianten und führt somit zu signifikanten Vorteilen in der Benutzerfreundlichkeit dieser geografischen Visualisierungen im Vergleich zum klassischen Standard in d3-Bibliotheken und Softwaretools im Reporting.

1 Einführung und Zielsetzung

Die Versorgung von Entscheidungsträgern mit relevanten Informationen zur Sicherstellung eines raschen und ergebnisorientierten Entscheidungsprozesses ist eine der Hauptaufgaben des Controllings. Dabei reicht es nicht aus, nur die richtigen Informationen zu liefern, sondern diese müssen auch dementsprechend aufbereitet werden. Durch die Verwendung von geeigneten Visualisierungen und die Darstellung von mehreren Grafiken in einem Dashboard können diese Anforderungen erfüllt werden (Eisl et al., 2018).

Dabei gewinnen vor allem moderne Big-Data-Visualisierungen im Zeitalter von Digitalisierung und Globalisierung für die Unternehmenssteuerung an Bedeutung. Die Eignung dieser Visualisierungstypen hängt dabei wesentlich von der Benutzerfreundlichkeit der Visualisierungen ab, welche in die Komponenten Effektivität, Effizienz und Zufriedenheit unterteilt werden kann. Die Reporting-Megatrends Echtzeit-Datenverarbeitung sowie Self-Service-BI verstärken die Wichtigkeit der Optimierung der Usability von interaktiven Big-Data-Visualisierungen, wie von den Autoren bereits in mehreren Publikationen verdeutlicht wurde (Hofer et al., 2018; Perkhofer, Walchshofer & Hofer, 2020).

Trotz des Nachteils hinsichtlich der Effektivität und der Effizienz sind aber immer noch klassische Geschäftsgrafiken, wie etwa Balken- oder Kreisdiagramme in den Unternehmen am häufigsten in Verwendung. Diese können durch eine interaktive Aufbereitung und eine Kombination mit anderen Diagrammvarianten aufgrund der Vertrautheit im Umgang mit diesen Visualisierungstypen, der folglich intuitiveren Interpretation und der simpleren Aufbereitung immer noch einen Mehrwert im Hinblick auf eine gute Entscheidungsbasis liefern. Eine 2016 durchgeführte Studie der Studiengänge Controlling, Rechnungswesen und Finanzmanagement (CRF) und Human-Centered-Computing (HCC) der FH OÖ zeigt, dass eine Kombination von Visualisierungen relativ häufig (30,3%) eingesetzt wird. Die Kombination von klassischen Geschäftsgrafiken und geografischen Visualisierungen ist dabei am verbreitetsten. (Hofer, Mayr & Walchshofer, 2018). Die gegenwärtige Relevanz der geografischen Visualisierungen zeigt sich ebenfalls in der Berichterstattung rund um die Coronakrise (Shneiderman, 2020).

Diese geografischen Visualisierungen ermöglichen den Entscheidungsträgern durch die präzise Darstellung von Informationen auf Landmaps, schnell und intuitiv Regionen zu analysieren, die auffällige oder abweichende Werte aufweisen. Durch die Kombination von Geschäftsgrafiken und geografischen Visualisierungen können die Vorteile beider Visualisierungen synergetisch genutzt werden. Währenddessen kann der Nachteil der Aggregation von Informationen, welcher bei Geschäftsgrafiken auftritt, durch die dynamische Kombination mit geografischen Visualisierungen ausgeglichen werden (Elmqvist & Fekete, 2010; MacEachren & Kraak, 2001). Eine richtig aufbereitete geografische Visualisierung kann somit einen wichtigen Mehrwert in der Berichterstattung für Entscheidungsträger liefern.

An dieser Stelle setzt dieses Research Paper an, da zwar die häufige Verwendung, nicht jedoch die geeignete Aufbereitung von geografischen Visualisierungen im Hinblick auf eine optimale Benutzerfreundlichkeit fundiert erforscht wurde. Für die Erforschung und Darstellung der Usability von geografischen Visualisierungen werden im Rahmen einer Online-Studie die signifikanten Unterschiede hinsichtlich der Effektivität, der Effizienz und der Zufriedenheit verschiedener Landmap-Varianten (Designvarianten) und verschiedener Zoomvarianten (Interaktionskonzepte) untersucht. Ziel ist es, die im Hinblick auf die Usability bestmögliche Variante hinsichtlich des Designs, der Interaktion sowie der Kombination aus Design und Interaktion für die Gestaltung einer geografischen Visualisierung zu finden. Das vorgestellte Paper fokussiert daher auf

- die optimale Gestaltung der unterschiedlichen geografischen Visualisierungen im Hinblick auf Design und Interaktionskonzept bzw. deren Kombination sowie
- einen anschließenden Vergleich dieser geografischen Visualisierungen im Hinblick auf deren optimalen Eignung, gemessen mittels Usability, in Abhängigkeit von bestimmten Aufgabenkategorien.

Hierfür werden mittels einer Literaturrecherche die wesentlichen geografischen Visualisierungen im Kapitel 2 vorgestellt. Darauf aufbauend werden im folgenden Kapitel die in der Studie verwendeten Designvarianten und Interaktionskonzepte erläutert und die darauf aufbauenden Hypothesen abgeleitet. Kapitel 4 behandelt die statistische Bewertung der Eignung dieser unterschiedlichen Design- und Interaktionskonzepte für differenzierte Big-Data-Aufgabenstellungen (Identifikation, Vergleichen, Zusammenfassen). Es wird ein finaler Usability-Vergleich der Landmap-Kombinationen präsentiert, der in wesentlichen Implikationen für Reporting im Big-Data-Umfeld resultiert. Limitationen und zukünftige Forschungsschwerpunkte für geografische Visualisierungen schließen dieses Paper ab.

2 Taxonomie von geografischen Visualisierungen und Interaktionen

2.1 Geografische Visualisierungen im Überblick

Landmaps ermöglichen in Bezug auf die Managementberichterstattung die geografische Visualisierung von Informationen. Dabei werden die Daten häufig in zweidimensionalen Landmaps integriert, wodurch der Benutzer schnell Informationen über bestimmte Regionen erhalten kann. Außerdem verbessern geografische Visualisierungen die Fähigkeit, komplexe Beziehungen visuell zu erkennen, die Suchzeit dabei zu verkürzen und Beziehungen aufzudecken, die ansonsten möglicherweise verborgen blieben (Dodge, 2008).

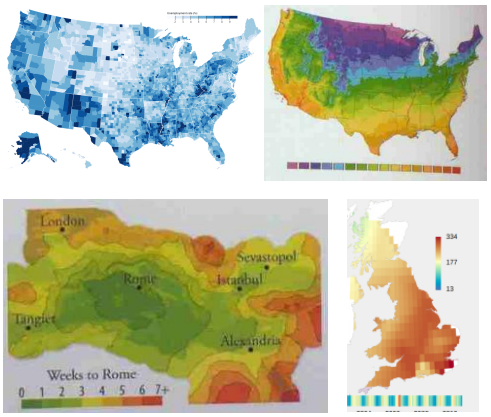
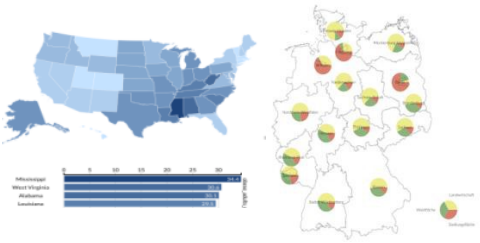
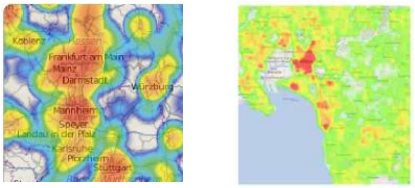
Kategorie	Beispiele	Literaturverweise
Regionen – Choropleth Maps – Isochrone Maps – Dasymetric Maps – Contour Maps		D3.js Bibliothek (Börner, 2015) (Few, 2009) (Mehler-Bicher & Böhm, 2012) (Nöllenburg, 2007) (Petrov, 2012) (Rahlf, 2018) (Schön, 2018) (Yau, 2013)
Combined Visualizations – Säulendiagramme – Kreisdiagramme – Tree Maps		D3.js Bibliothek (Few, 2009) (Rahlf, 2018) (Schön, 2018) (Schumann & Müller, 2000)
Heat Map – Isarithmic Maps – Hot Maps		D3.js Bibliothek (Börner, 2015) (Fisher, 2007) (Li et al., 2018) (Mehler-Bicher & Böhm, 2012)
Connections – Flow Maps – Spiral Tree		D3.js Bibliothek (Stephen & Jenny, 2017) (Verbeek, Buchin & Speckmann, 2011) (Yau, 2013)

Tabelle 1: Kategorisierung von geografischen Visualisierungen

Zu den wichtigsten geografischen Visualisierungen gehören Kartogramme, Choropleth-Maps, Reliefkarten und Proportionale-Symbolkarten (Börner, 2015). Hinzu kommen aber noch weitere Arten, die in der Tabelle 1 in mehrere Kategorien eingeteilt werden. Diese Kategorisierung hat nicht den Anspurch auf Vollständigkeit, sie bietet jedoch einen guten Überblick über die am häufigsten zitierten und daher offensichtlich wichtigsten Arten von geografischen Visualisierungen.

Da in der nachfolgenden Studie nur Choropleth Maps und Combined Visualizations verwendet werden, folgt nur zu diesen beiden Kategorien eine ausführlichere Beschreibung. Die sogenannten Heat Maps und Connections werden hingegen nur kurz erklärt. Während Heat Maps dem Betrachter durch starke Verdichtung und gut gewählte Farbkonzepte einen Überblick über mehrdimensionale, komplexe Datensätze geben, helfen Connections bei der Identifizierung von Material- oder Personenbewegungen zwischen unterschiedlichen Regionen (Perkhofer, Hofer & Walchshofer, 2019; Yau, 2013).

2.2 Choropleth-Maps

Um regionale Daten zu visualisieren, werden in der Praxis am häufigsten Choropleth-Maps (Abbildung 1) verwendet. Dabei werden die Regionen entsprechend den zugrundeliegenden Daten unterschiedlich eingefärbt, um nicht überlappende Gebiete wie einzelne Staaten, Bundesländer oder Bezirke tatsächlich differenziert darstellen zu können. Höhere Werte werden mit einer höheren Farbsättigung als niedrigere Werte dargestellt (Yau, 2013). Sollen die Visualisierungen außerdem ungeordnete Daten zeigen, so ist besonders darauf zu achten, dass gut differenzierbare Farben verwendet werden, währenddessen bei geordneten Daten unterschiedliche Helligkeitsstufen oder Farbtöne ausreichen (Nöllenburg, 2007).

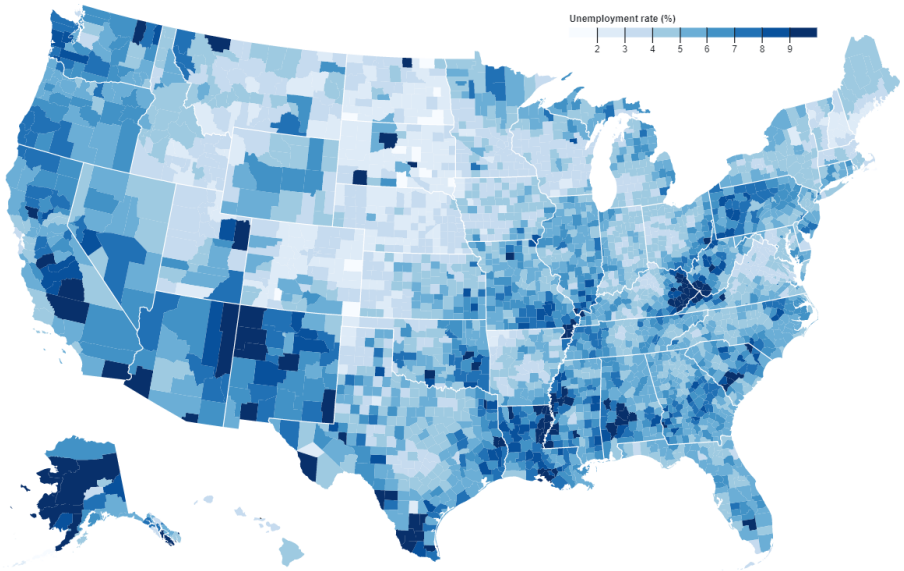


Abbildung 1: Choropleth-Map (Bostock, 2017)

Zwar eignen sich Choropleth-Maps sehr gut dazu, einen raschen Überblick über bestimmte Sachverhalte zu bekommen, jedoch können diese auch eine verzerrende Wirkung auf den Betrachter in der Interpretation haben. So stechen große Farbbereiche im Vergleich zu kleineren mehr hervor, wodurch der Betrachter eher größere als kleinere Gebiete betrachtet, obwohl beide Gebiete beispielsweise den gleichen Wert aufweisen (Few, 2009; Ritchie, Wigdor & Chevalier, 2019). Dieses Problem kann mit zusätzlichen Visualisierungen oder der Interaktionstechnik Perceptual Glimpses gelöst werden. Bei dieser Technik werden zwei Ansichten gleichzeitig dargestellt, wobei die erste die unverzerrte Ansicht und die zweite nur einen speziellen Teil der Daten zeigt. Die erste Ansicht visualisiert beispielsweise eine Choropleth-Map ohne Filterung von Daten, wodurch sie aufgrund der unterschiedlichen Größen der Regionen irreführend sein kann. Die zweite Ansicht zeigt ein Kartogramm, welches Quadrate in den richtigen Größen entsprechend den Werten jeder Region darstellt und die tatsächlichen Aussagen der Landmap dadurch verdeutlicht (Ritchie, Wigdor & Chevalier, 2019).

Bei Choropleth-Maps empfiehlt es sich außerdem, maximal sieben Datenklassen für ein gutes Design zu verwenden (Slocum, 2009). Da somit einem konkreten Wert keine bestimmte Farbe, sondern nur ein bestimmter Wertebereich zugewiesen werden kann, eignen sich Choropleth-Maps nur zur Darstellung von allgemeinen Trends oder zum groben Vergleich unterschiedlicher Regionen (Few, 2009; Nöllenburg, 2007).

2.3 Combined Visualizations

Bei Combined Visualizations werden geografische Visualisierungen um weitere Visualisierungen wie etwa Balken- oder Kreisdiagramme ergänzt, wodurch unterschiedliche Attribute und Größen in Bezug auf einen Standort visualisiert werden können (Schumann & Müller, 2000). So ermöglichen beispielsweise Kreisdiagramme in Kombination mit Landmaps die Darstellung von Häufigkeitsverteilungen beziehungsweise von bestimmten Zusammensetzungen (Mittag, 2016). Der interpretierbare Informationsgehalt ist somit bei dieser Art der Visualisierung größer als bei den Choropleth-Maps.

Bei der Anreicherung der Landmaps um zusätzliche Visualisierungen können jedoch auch Probleme auftreten. Beispielsweise können in die Regionen integrierte Kreis-, Säulen- oder Balkendiagramme schwer miteinander verglichen werden, falls die dargestellten Regionen sehr klein und die Größen der integrierten Geschäftsgrafiken sehr groß sind. An dieser Stelle kann es zu Überlappungen kommen, die eine Interpretation über verschiedene Regionen hinweg erschweren (Schumann & Müller, 2000). Diese zusätzlichen Visualisierungen können jedoch außerhalb der Landmap mit entsprechender Farbkodierung angebracht werden, um einen Mehrwert zu liefern (Mittag, 2016).

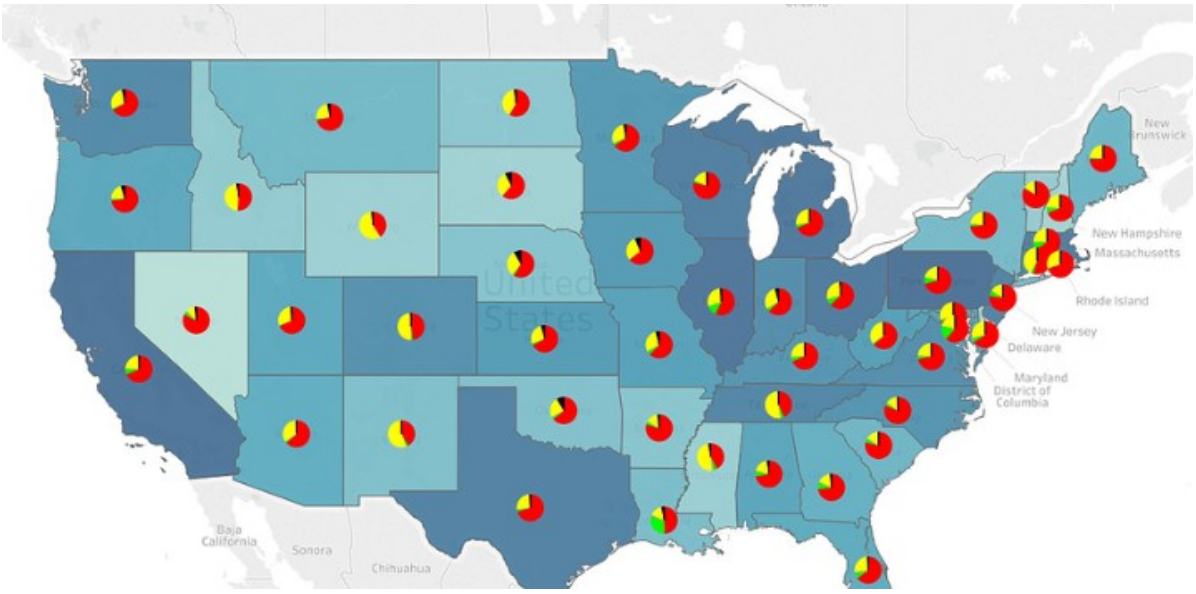


Abbildung 2: Choropleth in Kombination mit integrierten Kreisdiagrammen (van Panhuis, Cross & Burke, 2018)

2.4 Interaktionsmöglichkeiten für geografische Visualisierungen

In der Literatur finden sich viele verschiedene Interaktionsmöglichkeiten, die sich insbesondere für die Anwendung bei geografischen Visualisierungen eignen. Crampton (2002) fasst unterschiedliche Arten von Interaktivität, die häufig in der Geovisualisierung verwendet werden, zusammen und kategorisiert sie in vier Gruppen mit steigendem Interaktionsgrad. Dabei können mehrere Arten von Interaktivität miteinander kombiniert werden.

Interaktionskategorie	Interaktionsmöglichkeit
Interaktion mit der Datendarstellung durch Änderung:	- des Blickwinkels - der Datenausrichtung - des Zooms - der Skalierung - der Zuordnung von Symbolen
Interaktion mit der zeitlichen Dimension durch:	- Navigieren - Durchfliegen - Fixieren - Sortieren
Interaktion mit Daten durch:	- Datenbankabfragen - Filtern - Brushing - Hervorheben
Kontextualisierende Interaktionstechnik durch:	- Multiple Views - Linking von Datenebenen - Gleichzeitiges Anordnen von Daten

Tabelle 2: Interaktionsmöglichkeiten für geografische Visualisierungen (in Anlehnung an Crampton 2002)

Ein Großteil der Interaktionsmöglichkeiten in Tabelle 2 wurde ebenfalls in der Taxonomie von Shneiderman (1996) und Brehmer & Munzner (2013) dargestellt, sowie von Perkhofer, Hofer & Walchshofer (2019) zusammengefasst. Beispielsweise fallen laut dieser Taxonomie Hervorheben, Brushing und Linking allesamt unter den Begriff Selektieren.

Interaktionstechniken, die jedoch nur in der dargestellten Taxonomie von Crampton explizit erklärt wurden, sind:

- Durchfliegen: ermöglicht das Hin- und Herspringen zwischen unterschiedlichen Zeiträumen. Diese Interaktion hebt Änderungen während der Zeit hervor und wird häufig durch eine Dropdown-Liste möglich. Durchfliegen wird somit verwendet, wenn die Änderungen zwischen Zeiträumen im Fokus stehen.
- Sortieren: reiht Daten nach verschiedenen Kriterien. Diese Methode kann Trends aufdecken, die sonst verborgen bleiben würden (Crampton, 2002).

3 Studiendesign

Abgeleitet aus der Literatur und der Open-Source-Bibliothek D3.js wurde als Basiselement der Studie ein Grundtyp einer Choropleth-Map geschaffen, der die wichtigsten Empfehlungen im Hinblick auf Design und Interaktion berücksichtigt. Ausgehend von dieser Darstellung wurden durch Modifikation der geografischen Visualisierung zwei weitere Designs erstellt (kombinierte Visualisierungstypen, Farbnotation) und mit weiteren Interaktionen (Zoomvarianten) angereichert. Nach Skizzierung der Landmap-Visualisierungen wurden vier Hypothesen definiert, die die weitere Vorgehensweise grundlegend bestimmten. Das Umfragedesign wurde mithilfe der Software LimeSurvey erstellt und im Anschluss auf dem Online-Market-Place Amazon Mechanical Turk (MTurk) veröffentlicht. Dabei wurde auch die gewünschte Zielgruppe der Studie definiert. Jeder Teilnehmer musste mindestens einen US-

amerikanischen Bachelorabschluss vorweisen können. Die tatsächliche Auswahl der Teilnehmer passierte aber zufällig. Bei der Erstellung der Umfrage wurde die Aufgabentaxonomie von Brehmer und Munzner (2013) als Grundlage herangezogen, wodurch alle Designvarianten und Interaktionsmöglichkeiten hinsichtlich der Aufgabentypen Identifizierung, Vergleich und Zusammenfassung getestet wurden. Als Basis für die im Rahmen der Studie erstellten Landmaps dient ein von den Autoren generierter Datenbestand, der schon in anderen Studien zu verschiedensten Big Data Visualisierungen verwendet wurde und einen fiktiven Weinhandel darstellt. Dieser besteht aus 9.961 Datensätzen, wobei jeder Datensatz eine einzelne Kundenbestellung darstellt. Das Datenmuster besteht aus 5 Dimensionen und 14 Attributen.

Die statistische Auswertung der Studie wurde anschließend mit der ANOVA-Methode in Kombination mit dem Post-hoc-Student-Newman-Keuls-Test ausgewertet, um signifikante Unterschiede hinsichtlich der Effizienz und der Zufriedenheit der Designvarianten in Kombination mit Interaktionsvarianten identifizieren zu können. Für die Analyse der dichotomen Variable Effektivität kamen der Kruskal Wallis Test und der Mann-Whitney-U Test zur Anwendung.

3.1 Designvarianten

In diesem Kapitel werden die verschiedenen Designvarianten geografischer Visualisierungen, welche für die quantitative Online-Untersuchung dieser Arbeit kodiert wurden, dargestellt und die für deren Usability-Vergleich relevanten Hypothesen vorgestellt.

Insgesamt wurden für die Durchführung der Studie drei verschiedene Designvarianten erstellt. Die erste Variante, eine Choropleth Map, wurde auf Basis der D3.js-Bibliothek erstellt, um Interaktionen erweitert und z.B. im Hinblick auf Farbnotation an die Studienzwecke angepasst (Bostock, Ogievetsky & Heer, 2011). Die Choropleth-Map als Standardvisualisierung wird für den Vergleich mit zwei anderen, aus der Literatur abgeleiteten Varianten verwendet. Die zweite Designvariante stellt eine Choropleth-Map mit einem zusätzlichen Balkendiagramm außerhalb der Landmap dar. Das dritte Layout besteht aus einer Landmap in Kombination mit einem integrierten Kreisdiagramm in jeder Region. Abbildung 3 zeigt die angefertigten Skizzen für die drei verschiedenen Designvarianten.

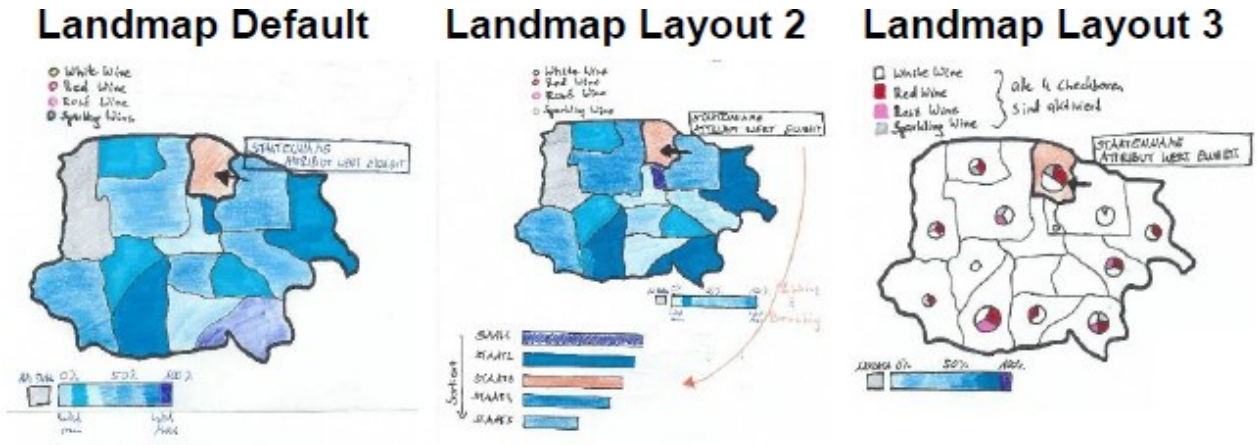


Abbildung 3: Skizzen der Landmap-Designvarianten

3.1.1 Designvariante 1 - Standard Choropleth-Map

Die Choropleth-Map ist die häufigste Art der geografischen Visualisierung zur Darstellung von Regionen und wird daher als erste Art der interaktiven geografischen Visualisierung für die folgende quantitative Online-Umfrage verwendet (Yau, 2013). Dazu wurde, basierend auf einer Standard-d3.js Visualisierung die Choropleth-Map kodiert, erweitert und an die Studienzwecke angepasst. Die größten Anpassungen gab es hier im Bereich der Interaktionstechniken, damit sowohl in dieser, als auch in den beiden folgenden Designvarianten das Selektieren von Regionen und damit verbunden ein Tooltip zur Datenanzeige genutzt werden können (Bostock, Ogievetsky & Heer, 2011). Die als Standardvisualisierung für die Studie erstellte Choropleth-Map, die als Ausgangsbasis für weitere Vergleiche mit den beiden anderen Landmaps dient, enthält somit schon wesentliche Elemente der Interaktion.

Zusätzlich zum Interaktionskonzept stellt eine passende Farbauswahl eine wesentliche Komponente für interaktive geografische Visualisierungen dar. Die Literatur bezieht sich dabei insbesondere auf den Farbton und die Helligkeit (Brewer, 1994). Es wurden bereits verschiedene Studien zur Färbung von Choropleth-Maps durchgeführt. Laut Few (2009) sollen prismatische Farben nur verwendet werden, um verschiedene Kategorien wie verschiedene politische Parteien nach den Wahlen darzustellen. Für Mengenvergleiche sollen unterschiedliche Farbintensitäten verwendet werden. Die Studie von Brewer (1994) zeigt dabei, dass es keine signifikanten Unterschiede zwischen den Farben gibt, wobei Graustufen weniger effektiv sind als andere Schattierungen.

Basierend auf diesen Erkenntnissen werden für diese Studie blaue Schattierungen für die Visualisierungen verwendet. Die verschiedenen Blautöne sind nicht einer definierten Anzahl von Klassen, sondern prozentual den verschiedenen Ausprägungen der Attribute der Datenbasis zugeordnet. Die kleinste Ausprägung wird mit der hellsten, und die größte Ausprägung wird mit der dunkelsten Schattierung angezeigt. Die Legende befindet sich unter der Visualisierung und ist mit den Maximal-, Minimal- und Mittelwerten jeder Farbkategorie gekennzeichnet. Vor der Legende der blauen Schattierungen wird ein graues Feld für Regionen mit fehlender Ausprägung angebracht.

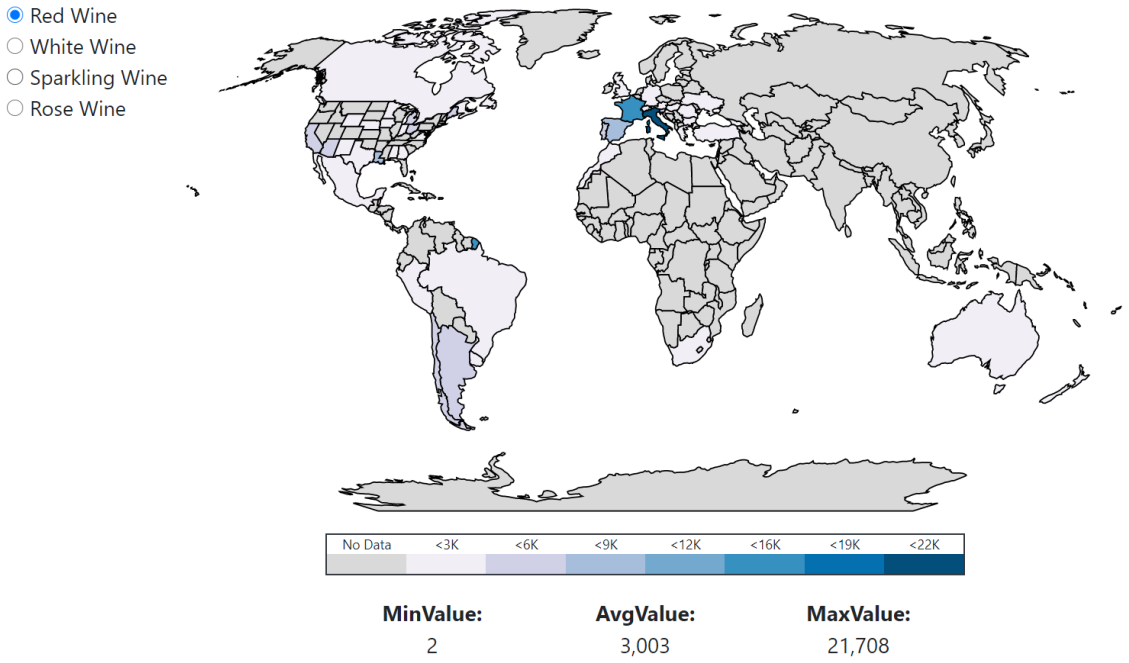


Abbildung 4: Designvariante 1 - Choropleth Map

Um unterschiedliche Dimensionsausprägungen (z.B. Weinsorte) abwechselnd anzuzeigen, werden in dieser Visualisierungsform Optionsfelder verwendet. Diese Optionsfelder ermöglichen ein Filtern der Daten, wobei mit Klicken auf eine Schaltfläche immer nur eine Selektionsausprägung ausgewählt werden kann.

Die Interaktionstechnik Selektieren wird für alle drei Designvarianten zur Auswahl der Daten je Region verwendet (Crampton, 2002; Dodge, 2008; Weaver, 2007). Wenn der Benutzer mit dem Cursor über eine Region fährt, wird die ausgewählte Region mit einem fetten schwarzen Rand hervorgehoben. Gleichzeitig mit dem schwarzen Rand zeigt ein Tooltip den Namen und die Attributsausprägung der ausgewählten Region an. Abbildung 4 zeigt die kodierte Choropleth Map, die für die anschließende Studie verwendet wurde.

3.1.2 Designvariante 2 – Choropleth-Map in Kombination mit Balkendiagramm

Das zweite Layout zeigt ebenfalls eine Choropleth-Map, welche aber zusätzlich um ein daneben angeordnetes Balkendiagramm ergänzt wird. Wie bei der Grundvisualisierung werden auch in dieser Variante alle Design- und Interaktionselemente, sprich dieselbe Farbgestaltung, dieselben Optionsfelder und dieselben Interaktionsmöglichkeiten, verwendet. Der einzige Unterschied besteht im zusätzlichen Balkendiagramm. Dieses Balkendiagramm zeigt die einzelnen Regionen und die dazugehörigen Attributsausprägungen in absteigender Reihenfolge. Diese Ergänzung sollte den Benutzer dieser Visualisierung dabei unterstützen, die erforderlichen Informationen zu erhalten. Nachdem mit dem Cursor eine Region markiert wird, wird nicht nur die Region mit einem fetten schwarzen Rand, sondern auch der Balken im Balkendiagramm entsprechend hervorgehoben. Verschiedene Teile der Visualisierung sind in dieser Variante somit verknüpft und Änderungen an einem Datenpunkt werden an allen verknüpften Stellen widergespiegelt (Isenberg & Fisher, 2009). Ein wesentlicher Nachteil, der speziell bei dieser Art der Visualisierung auftreten kann, ist die schlechtere Lesbarkeit aufgrund des zusätzlichen Platzbedarfs durch das zusätzliche Geschäftsdiagramm (Li, 2020). Mittels der zusätzlichen Zoom-Interaktionsmöglichkeiten, die im Folgekapitel näher erläutert werden, wird diesem Problem jedoch entgegengewirkt.

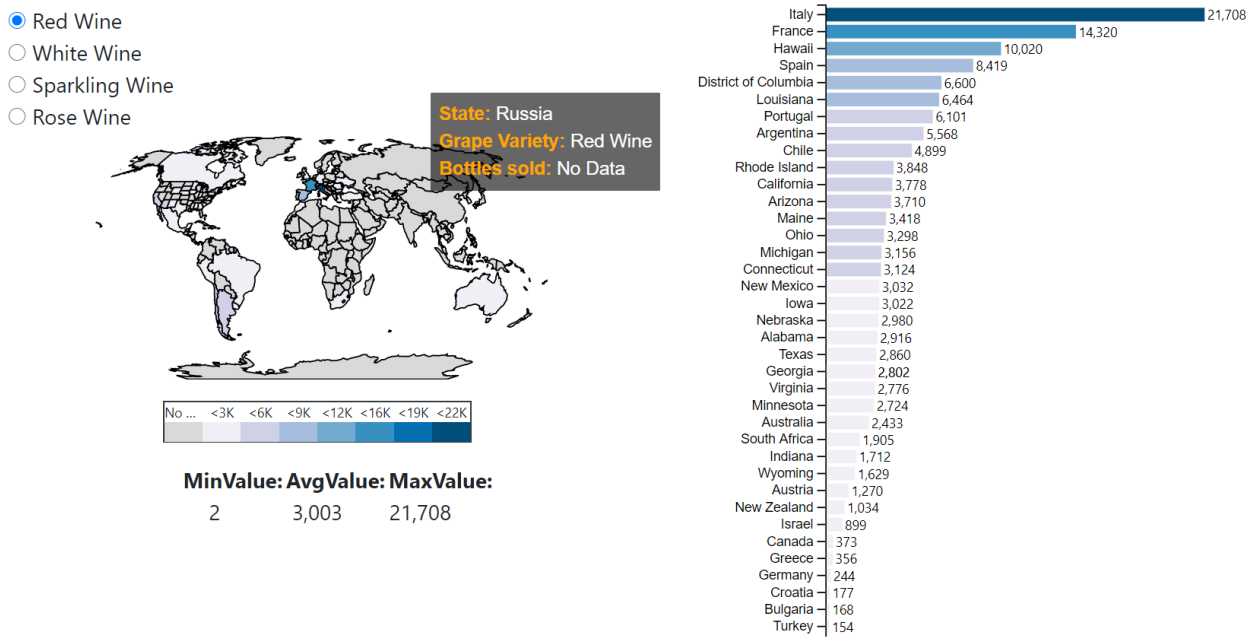


Abbildung 5: Designvariante 2 – Choropleth-Map mit Balkendiagramm

Im Hinblick auf die Untersuchung der Usability der beiden bisher vorgestellten Designvarianten lässt sich folgende Hypothese ableiten:

Hypothese H1: Eine Choropleth-Map in Kombination mit einem Balkendiagramm ist signifikant besser als eine Choropleth-Map ohne Balkendiagramm im Hinblick auf

- **a:** Effektivität,
- **b:** Effizienz und
- **c:** Zufriedenheit.

3.1.3 Designvariante 3 - Landmap in Kombination mit integrierten Kreisdiagrammen

Anders als bei den vorhergehenden Designvarianten zeigt Layout 3 keine Choropleth-Map, sondern eine Kombination aus einer Landmap und einem integrierten Kreisdiagramm je Region (Schumann & Müller, 2000). Im Gegensatz zu den beiden vorigen Varianten, in denen nur eine Dimensionsausprägung auswählbar ist, ist bei dieser Designvariante in den dargestellten Kreisdiagrammen mittels Mehrfachauswahl in den Kontrollkästchen die Darstellung von mehreren Dimensionsausprägungen möglich. Zusätzlich zur Darstellung des Kreisdiagramms können durch die Größe und die Farbe der Segmente des Kreisdiagramms weitere Informationen übermittelt werden (Abbildung 6). Analog zur zweiten Designvariante wird auch bei dieser Landmap die Notwendigkeit weiterer Interaktionsmöglichkeiten, vor allem einer Zoom-Interaktion, schnell ersichtlich. Der Grund dafür ist die Überlappung der Kreisdiagramme, die eine einfache Identifizierung von Details erschwert und die Usability daher beeinträchtigt (Li, 2020). Eine mögliche Variante des Zooms ist in Abbildung 7 ersichtlich.

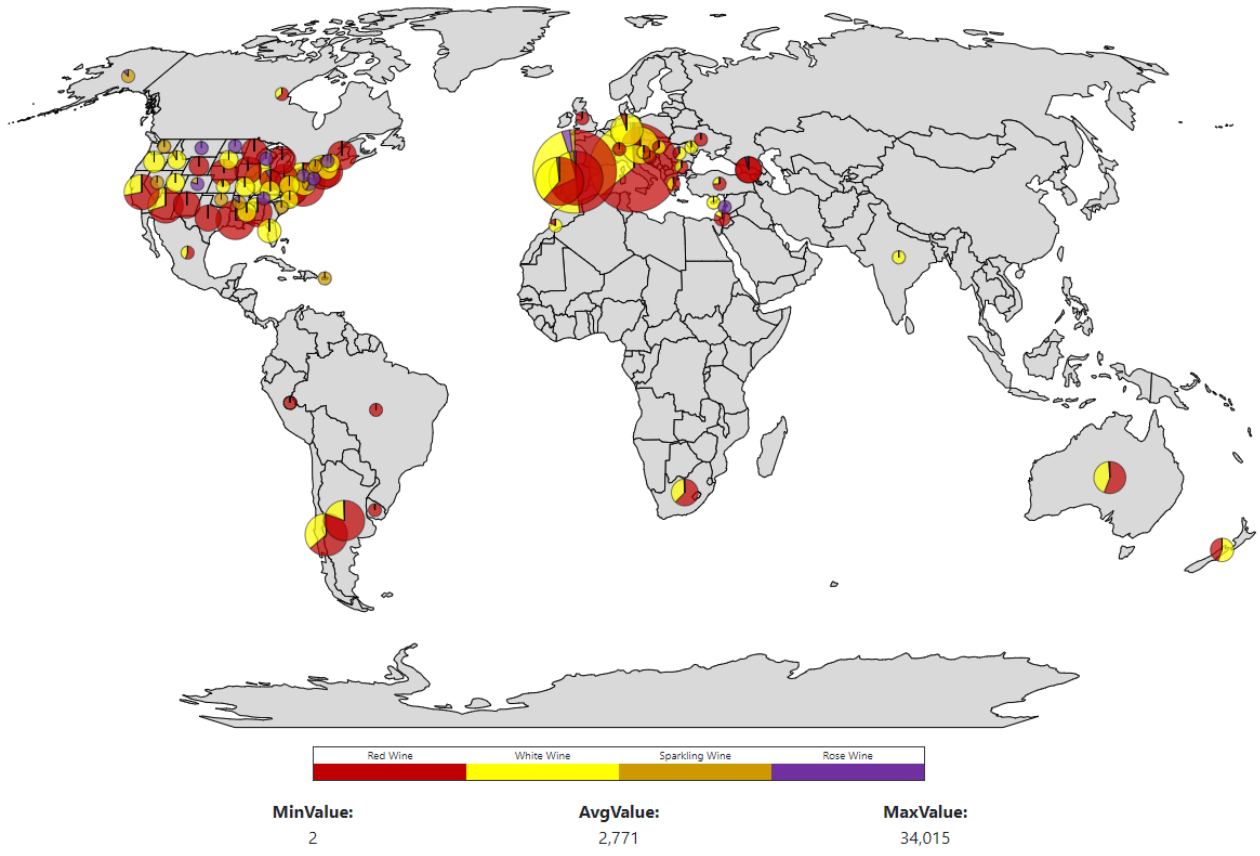


Abbildung 6: Designvariante 3 - Landmap mit integrierten Kreisdiagrammen

Um die Usability der Landmap in Kombination mit integrierten Kreisdiagrammen mit der Standard-Choropleth-Map vergleichen zu können, wird folgende Hypothese abgeleitet:

Hypothese H2: Eine Landmap in Kombination mit integrierten Kreisdiagrammen ist signifikant besser als eine Choropleth-Map im Hinblick auf

- **a:** Effektivität,
- **b:** Effizienz und
- **c:** Zufriedenheit.

3.2 Interaktionskonzepte

Um Detailinformationen aus einer Visualisierung ablesen zu können, bedarf es speziell bei integrierten Kombinationen aus Landmaps und Geschäftsgrafiken der Interaktionsmöglichkeit des Zooms (Shneiderman, 1996). Dabei gibt es unterschiedliche Möglichkeiten zur Umsetzung. So zeigt auch die zur Kodierung der Designvarianten herangezogene D3.js-Bibliothek verschiedene Arten des Zooms. Aus diesem Grund wurden für die Studie drei verschiedene Zoomvarianten kodiert. Die erste Zoomvariante ist ein stufenloses Scrollen mittels Mausrad, die zweite Variante entspricht einem Zoom mittels +/- Tasten und die dritte Methode ermöglicht das Zoomen durch einen Doppelklick. Bei den Varianten 2 und 3 ist jedoch kein stufenloses, sondern ein Zoomen auf drei Zoomebenen vorgesehen. Mittels eines Reset-Knopfs kann wieder in die ursprüngliche Ansicht gewechselt werden. Diese drei Interaktionskonzepte werden mit den drei Designvarianten kombiniert, wodurch neun verschiedene Visualisierungsformen getestet werden.

Designvariante	Interaktionsvariante
1 - Standard Choropleth-Map	1 - Stufenloses Scrollen mittels Mausrad
2 - Choropleth-Map + Balkendiagramm	2 - Dreistufiger Zoom mittels +/- Taste
3 - Landmap + integrierte Kreisdiagramme	3 - Dreistufiger Zoom mittels Doppelklick

Tabelle 3: Verwendete Merkmalsausprägungen der Visualisierungsformen

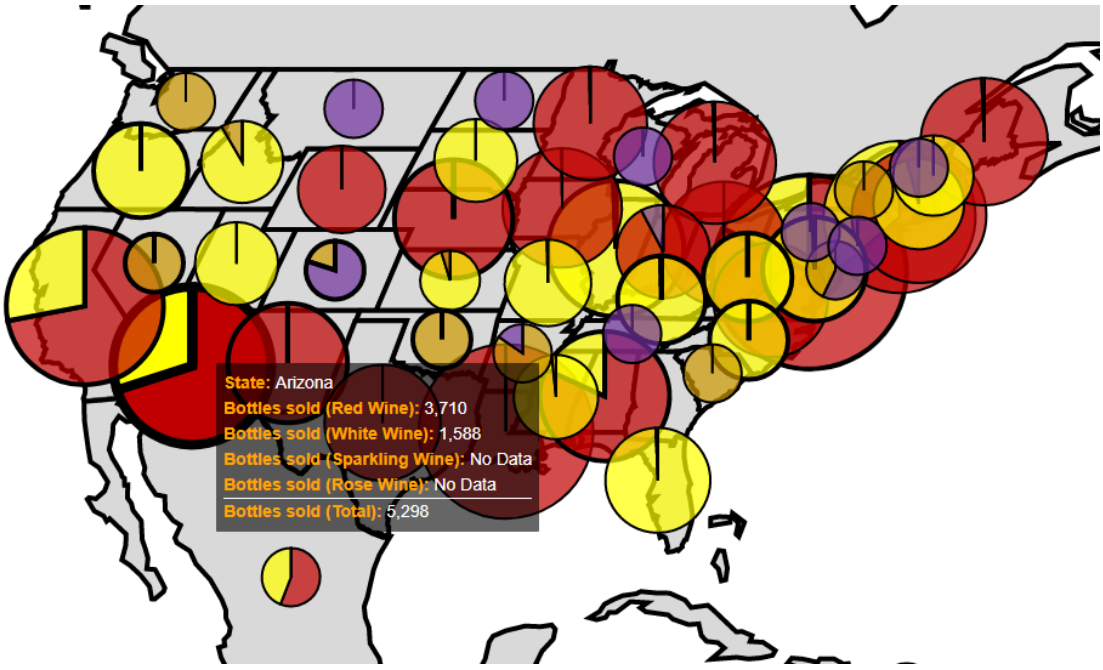


Abbildung 7: Interaktion Zoom inkl. Tooltip

Aufgrund der Wichtigkeit, nach einem ersten Überblick Details aus den Landmaps ableiten zu können, ist ein Usability-Vergleich der unterschiedlichen Zoomvarianten in Kombination mit den vorgestellten Designvarianten essentiell. Dieser Usability-Vergleich zwischen den drei Varianten wird mithilfe der folgenden beiden Hypothesen formuliert:

Hypothese H3: Eine Landmap mit der Zoomvariante Stufenloses Scrollen mittels Mausrad ist signifikant besser als eine Landmap mit dreistufiger Zoomvariante durch +/- Tasten im Hinblick auf

- **a:** Effektivität,
- **b:** Effizienz und
- **c:** Zufriedenheit.

Hypothese H4: Eine Landmap mit der Zoomvariante Stufenloses Scrollen mittels Mausrad ist signifikant besser als eine Landmap mit dreistufiger Zoomvariante durch Doppelklick im Hinblick auf

- **a:** Effektivität,
- **b:** Effizienz und
- **c:** Zufriedenheit.

3.3 Studiendurchführung

Nach der Kodierung der Kombinationen der drei Landmap-Designvarianten und der drei verschiedenen Zoomvarianten wurde mittels der Open-Source-Software LimeSurvey ein Fragebogen erstellt. Die verwendeten Fragen wurden dabei gemäß Brehmer und Munzner (2013) in die drei Aufgabentypen, Identifizierung, Vergleich und Zusammenfassung, eingeteilt. Für jede Aufgabe und jede Visualisierungskombination gab es zur Validierung letztendlich zwei Beispiele.

Insgesamt ergaben sich aus den drei Designvarianten, den drei Interaktionsmöglichkeiten, den drei Aufgabentypen und zwei Befragungen je Aufgabentyp 54 Kombinationen, um die Usability aller Kombinationen testen und vergleichen zu können.

Die einzelnen Aufgabenstellungen wurden, da es sich um geografische Abbildungen handelt, so formuliert, dass bestimmte Staaten oder Länder als Antwortmöglichkeit in Frage kommen. So wurden für den Aufgabentypen Identifikation zunächst große Staaten und bei der zweiten Wiederholung dieser Aufgabe- und Visualisierungskombination kleine Staaten als mögliche Antworten festgelegt. Beim Vergleich wurden in beiden Durchgängen Kontinente abgefragt, während bei der Zusammenfassung erst Kontinente und anschließend Staaten als richtige Ergebnisse definiert wurden. Anders als bei den Aufgabentypen Identifikation und Vergleich wurde die Auswahl der Kontinente und Staaten bei der Zusammenfassung bewusst getroffen. Um einer möglichen Stichprobenverzerrung vorzubeugen, wurden die Aufgabenstellungen den Teilnehmern aber in zufälliger Reihenfolge präsentiert. Dabei hatten die Teilnehmer stets die drei Antwortmöglichkeiten „die Aussage ist richtig“, „die Aussage ist falsch“ und „die Aussage kann nicht beantwortet werden“.

Die quantitative Online-Studie wurde anschließend im August 2019 mit dem auf LimeSurvey erstellten Fragebogen auf dem Online-Marktplatz MTurk gestartet. Die Studie wurde dabei in drei Teile gegliedert, um die Konzentrationsfähigkeit der Probanden durchgehend gewährleisten zu können. In jedem Teil wurde immer nur eine Zoomvariante abgefragt. Am Beginn der Umfrageteilnahme wurden die Studienteilnehmer noch mit einer Erläuterung zum Datensatz und zum Ablauf der Studie sowie mit verschiedensten Beschreibungen und einem kurzen Video zu Visualisierungen und Interaktionstechniken unterstützt.

Die Ergebnisse der Studie wurden anschließend in Microsoft Excel deskriptiv ausgewertet. In SPSS (Version 24.0.0.0) wurden im Anschluss die statistischen Berechnungen für signifikante Unterschiede hinsichtlich der abhängigen Variablen Effizienz und Zufriedenheit mithilfe der ANOVA-Methode in Kombination mit dem Post-hoc-Student-Newman-Keuls-Test (SNK) durchgeführt. Für die dichotome Variable Effektivität wurde der Kruskal-Wallis-H-Test in Kombination mit dem Mann-Whitney-U-Test zur Berechnung herangezogen. Die detaillierten Ergebnisse dieser statistischen Methoden werden im Kapitel der statistischen Auswertung und der Ergebnisse näher erläutert.

3.4 Datenbasis und Teilnehmer

Der zur Studie herangezogene Datenbestand wurde von den Autoren generiert und stellt einen fiktiven Weinhandel dar. Dieser Datensatz wurde bereits für andere Studien zu verschiedenen Visualisierungstypen im Rahmen des Forschungsprojekts USIVIS verwendet. Der ursprüngliche Datensatz dieser Studien wurde jedoch um zusätzliche geografische Informationen erweitert, um für diese Studie verwendet werden zu können. Der Datenbestand besteht insgesamt aus 9.961 Datensätzen, wobei jeder Datensatz eine einzelne Kundenbestellung darstellt. Das Datenmuster besteht aus 5 Dimensionen und 14 Attributen. Die Dimensionen sind die Rebsorte, die Produktgruppe, das Herkunftsland, der Kontinent, der Händler. Zu den Attributen zählen der Listenpreis, die Anzahl der verkauften Flaschen, der Bruttoumsatz, der Nettoumsatz, die Marge in % und absolut, der Kundenrabatt in % und absolut, die Umsatzkosten, der Bruttogewinn, die Versandkosten, der Deckungsbeitrag in % und absolut und die Qualität des Weins in Punkten. Die Visualisierungen der Studie zeigen jedoch nur die Absatzmengen der Weinhändler in jeder Region.

Insgesamt nahmen an dieser internationalen Online-Studie 143 Probanden teil. Von diesen wurden im Rahmen der Durchführung die demografischen Informationen Geschlecht, Alter, Bildung, Berufserfahrung, Position im Beruf, Tätigkeitsbereich, Branche und Herkunftsland abgefragt. In Bezug auf das Geschlecht war die Testgruppe ungefähr gleichmäßig verteilt (männlich = 50%, weiblich = 47%, keine Angabe = 3%). Die Altersspanne der Gruppe betrug 24 bis 73 Jahre mit einem Durchschnittsalter von 42 Jahren. Am häufigsten vertreten war die Altersgruppe von 31-40 Jahre (Abbildung 8).

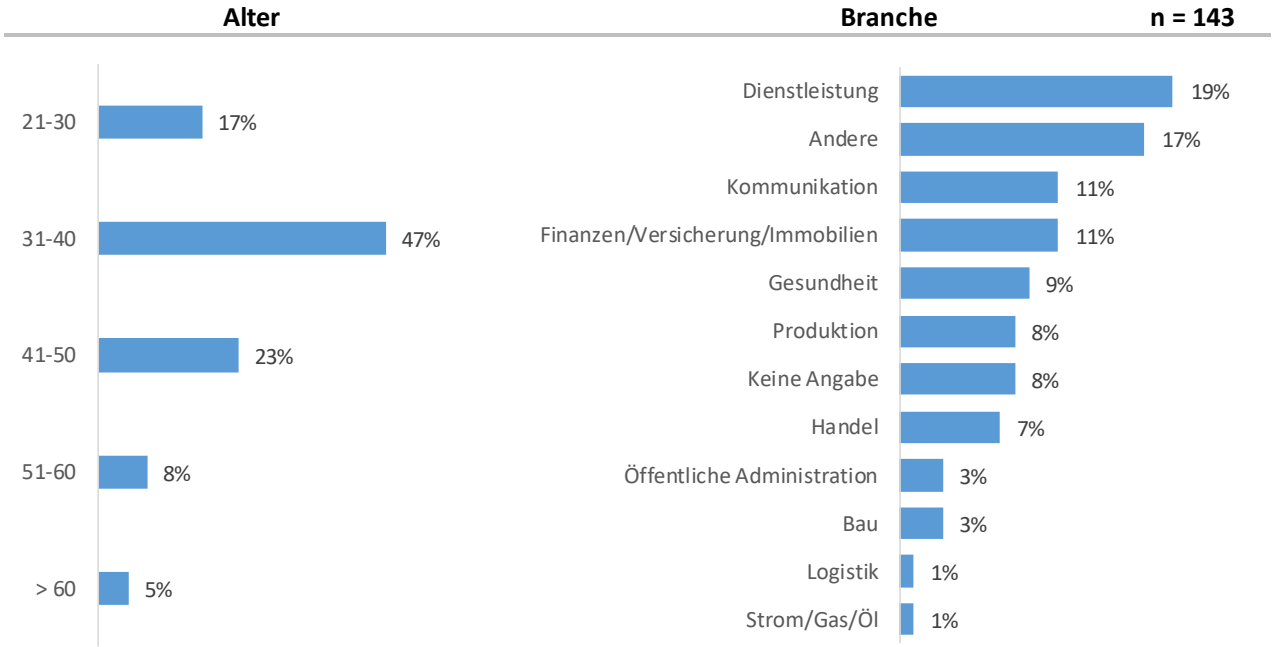


Abbildung 88: Alters- und Branchenverteilung der Probanden

71% der Studienteilnehmer wiesen den geforderten Bachelorabschluss auf, der Rest konnte sogar einen höheren Bildungsabschluss vorweisen. Die Mehrheit der Teilnehmer konnte außerdem auf Berufserfahrung (93%) zurückgreifen, wobei es keine großen Unterschiede hinsichtlich der Branchenverteilung gab (Abbildung 8). Fast die Hälfte der Teilnehmer waren Angestellte (48%). Der Rest gliedert sich in Selbständige (17%), mittleres Management (12%), unteres Management (10%) und oberes Management (5%). Die meisten Befragten stammen aus den USA (92%), während der Rest Indien (5%) und das Vereinigte Königreich (2%) als Herkunftsland angab.

3.5 Statistische Auswertung

Zur Auswertung der Umfragedaten wurde deskriptive Statistik, sowohl für die demografischen Daten als auch für die Usability-Analysen jeder Visualisierung und Aufgabe angewandt. In Vorbereitung auf die Datenanalysen wurden die Ausreißer durch Weglassen der niedrigsten 5% und der höchsten 5% im Hinblick auf die benötigte Zeit pro Visualisierung eliminiert.

Zur Messung der Effektivität wird die Antwortrichtigkeit als dichotome Variable mit 0 = „falsch“ und 1 = „richtig“ definiert und bestimmt. Die Antwortzeit wird vom Umfragetool aufgezeichnet und kann als stabiler Grad für die Effizienz verwendet werden. Die Zufriedenheit wird durch Befragung nach jeder Visualisierung als Ordnungsgrad mit einer 5-Punkte-Likert-Skala ermittelt. Diese Skala reicht von 0 = „sehr unzufrieden“ bis 4 = „sehr zufrieden“.

Die abhängigen Variablen sind Effektivität, Effizienz und Zufriedenheit. Die unabhängigen Variablen sind die verschiedenen Designvarianten, die verschiedenen Interaktionsvarianten und die verschiedenen Kombinationen von Design- und Interaktionsvarianten. Darauf basierend wurden für die Analysen von Effizienz und Zufriedenheit ANOVA und Post hoc SNK herangezogen, während für die Analyse der Effektivität der Kruskal Wallis Test und der Mann-Whitney-U Test zur Anwendung kamen. Als Signifikanzniveau wurde für die ANOVA 0,05 festgelegt.

Die Effektgröße Eta beschreibt zusätzlich zur Signifikanz des jeweiligen Usability-Unterschieds die Stärke dieses Effekts zwischen den verschiedenen Designvarianten, Aufgaben und Interaktionsmöglichkeiten und reicht von 0 bis 1. Als Grenzen für die Effektgröße für die Effizienz und die Zufriedenheit wird, gemäß der Skala von Cohen (1977), 0,01 für einen kleinen Effekt, 0,06 für einen mittleren Effekt und 0,14 für einen großen Effekt definiert. Für die Effektgröße der Effektivität (Mann-Whitney-U-Test) wird die Skala von Cohen (1992) mit den Grenzwerten 0,10 für einen kleinen Effekt, 0,30 für einen mittleren Effekt und 0,50 für einen großen Effekt angewandt.

4 Ergebnisse des Usability-Vergleichs zu Design, Interaktion und Kombinationen

Dieses Kapitel präsentiert und beschreibt die Ergebnisse der zuvor beschriebenen quantitativen Online-Studie. Die Ergebnisse sind in die Effektivität, die Effizienz und die Zufriedenheit unterteilt, wobei diese gesondert für die Designvarianten, die Interaktionsmöglichkeiten und die Kombination daraus dargestellt werden. Mittels der statistischen Auswertungen werden die je Bereich vorgestellten Hypothesen verifiziert bzw. falsifiziert. Hypothesen, die verifiziert werden, sind mit einem ✓ gekennzeichnet, diejenigen, die falsifiziert werden, sind mit einem ✗ gekennzeichnet.

4.1 Designvarianten

Sowohl der Kruskal-Wallis-Test bei der Effektivität ($p = 0{,}004$) als auch die ANOVA-Analysen bei der Effizienz ($p = 0{,}000$) und der Zufriedenheit ($p = 0{,}000$) weisen signifikante Unterschiede zwischen den Layouts auf. Designvariante 2 (Choropleth-Map + Balkendiagramm) und Variante 3 (Landmap + integrierte Kreisdiagramme) zeigen dabei höhere Effektivitätswerte als die Designvariante 1 (Standard Choropleth-Map). Der SNK-Test der Effizienz, aber auch jener der Zufriedenheit, ordnet jede Designvariante in einzelne, signifikant unterscheidbare Gruppen ein, wobei hier Variante 3 die geringste Zeit für die Beantwortung der Fragen beansprucht. Die höchste Zufriedenheit der Teilnehmer weist die Designvariante 2 auf, während Design 3 hier die niedrigste Bewertung erhält.

Die Effektgröße Eta zeigt bzgl. der Effektivität (0,06) keinen, bei der Effizienz einen mittleren (0,07) und bei der Zufriedenheit einen großen Effekt (0,27).

Designvariante	Mittelwert Effektivität	Mittelwert Effizienz	Mittelwert Zufriedenheit
1	88%	44,4	2,56
2	92%	33,7	3,47
3	92%	30,0	1,92
p	0,004	0,000	0,000
Eta	0,06	0,07	0,27
Gruppen (Mann-Whitney U)	1. Landmap 2 & 3 2. Landmap 1		
Gruppen (SNK)		1. Landmap 3 2. Landmap 2 3. Landmap 1	1. Landmap 2 2. Landmap 1 3. Landmap 3

Tabelle 4: Designvarianten - Mittelwerte, Signifikanzniveaus und Gruppenvergleiche im Überblick

Abgeleitet aus den statistischen Ergebnissen (Tabelle 4) lassen sich die Hypothesen zu den Designvarianten wie folgt verifizieren oder falsifizieren:

H1: Eine Choropleth-Map in Kombination mit einem Balkendiagramm ist signifikant besser als eine Choropleth-Map ohne Balkendiagramm hinsichtlich der

- **a:** Effektivität – VERIFIZIERT
- **b:** Effizienz – VERIFIZIERT
- **c:** Zufriedenheit – VERIFIZIERT

Effektivität	Effizienz	Zufriedenheit
✓	✓	✓

Tabelle 5: Ergebnisse zu Hypothese 1

H2: Eine Landmap in Kombination mit integrierten Kreisdiagrammen ist signifikant besser als eine Choropleth-Map hinsichtlich der

- **a:** Effektivität – VERIFIZIERT
- **b:** Effizienz – VERIFIZIERT
- **c:** Zufriedenheit – FALSIFIZIERT (signifikant schlechter)

Effektivität	Effizienz	Zufriedenheit
✓	✓	✘

Tabelle 6: Ergebnisse zu Hypothese 2

Zusammenfassend kann gesagt werden, dass die Designvarianten 2 und 3 effektiver und effizienter als Designvariante 1 sind. Variante 2 zeigt außerdem eine höhere Zufriedenheit der Teilnehmer als Variante 1. Die deskriptive Statistik verdeutlicht außerdem, dass die Choropleth-Map in Kombination mit dem Balkendiagramm (Designvariante 2) auch mehr Zufriedenheit als die Landmap mit integrierten Kreisdiagrammen (Designvariante 3) bei den Befragten auslöst. Hinsichtlich der Effektivität erweisen sich die Designvarianten 2 und 3 als gleich effektiv, wohingegen bei der Effizienz die Variante 3 leichte Vorteile aufweist.

4.2 Interaktionskonzepte

Die vorgestellten Zoomvarianten unterscheiden sich signifikant, sowohl hinsichtlich der Effektivität (Kruskal-Wallis-Test: p = 0,016) als auch der Effizienz (ANOVA: p = 0,000). Die höchste Effektivität weisen dabei die Zoomvarianten 2 (dreistufiger Zoom mittels +/- Taste) und 3 (dreistufiger Zoom mittels Doppelklick) auf. In Bezug auf die Effizienz werden die Zoomvarianten zwei Gruppen zugeordnet. Die Zoomvarianten 2 und 3 gehören laut SNK-Test zur Effizienz zur ersten Gruppe und sind daher effizienter als die Zoomvariante 1. In Bezug auf die Zufriedenheit (ANOVA: p = 0,270) gibt es keine signifikanten Unterschiede. Die Effektstärke Eta zeigt bei der abhängigen Variable Effektivität keinen Effekt (Mann-Whitney-U), bei der Effizienz hingegen weist der SNK-Test einen kleinen Effekt auf.

Zoom	Mittelwert Effektivität	Mittelwert Effizienz	Mittelwert Zufriedenheit
1	88%	38,7	2,63
2	92%	35,4	2,70
3	92%	34,0	2,61
p	0,016	0,000	0,270
Eta	0,06 - 0,08	0,01	
Gruppen (Mann-Whitney U)	1. Zoom 3 & 2 2. Zoom 1		
Gruppen (SNK)		1. Zoom 3 & 2 2. Zoom 1	

Tabelle 7: Interaktionsmöglichkeiten - Mittelwerte, Signifikanzniveaus und Gruppenvergleiche im Überblick

Abgeleitet aus den Ergebnissen der Tabelle 7 können die folgenden Hypothesen wie folgt falsifiziert werden:

H3: Eine Landmap mit der Zoomvariante Stufenloses Scrollen mittels Mausrad ist signifikant besser als eine Landmap mit dreistufiger Zoomvariante durch +/- Tasten hinsichtlich der

- **a:** Effektivität – FALSIFIZIERT (signifikant schlechter)
- **b:** Effizienz – FALSIFIZIERT (signifikant schlechter)
- **c:** Zufriedenheit – FALSIFIZIERT (nicht signifikant)

Effektivität	Effizienz	Zufriedenheit
✘	✘	✘

Tabelle 82: Ergebnisse zu Hypothese 3

H4: Eine Landmap mit der Zoomvariante Stufenloses Scrollen mittels Mausrad ist signifikant besser als eine Landmap mit dreistufiger Zoomvariante durch Doppelklick hinsichtlich der

- **a:** Effektivität – FALSIFIZIERT
- **b:** Effizienz – FALSIFIZIERT
- **c:** Zufriedenheit – FALSIFIZIERT (nicht signifikant)

Effektivität	Effizienz	Zufriedenheit
✘	✘	✘

Tabelle 9: Ergebnisse zu Hypothese 4

Dreistufige Zoomvarianten durch +/- Tasten und Doppelklick erweisen sich somit als effektiver und effizienter als das Zoomen durch Stufenloses Scrollen mittels Mausrad. Zusätzlich zeigt die deskriptive Statistik, dass diese beiden Zoomvarianten (2 & 3) gleich effektiv sind (92%). Bei der Effizienz zeigt die Interaktionsmöglichkeit mittels Doppelklick, bei der Zufriedenheit die Interaktionsmöglichkeit mittels +/- Tasten jedoch etwas bessere Werte.

4.3 Kombination Designvarianten & Interaktionskonzepte

Die effektivste interaktive Visualisierung ist die Designvariante 3 (Landmap mit integrierten Kreisdiagrammen) in Kombination mit der Zoomvariante 3 (Dreistufiger Zoom mittels Doppelklick: 94%). Der Kruskal-Wallis-Test bestätigt den signifikanten Unterschied zwischen den Layouts in Kombination mit den Zoomvarianten ($p = 0{,}007$).

Die effizienteste interaktive Visualisierung ist Variante 3 in Kombination mit Zoomvariante 2 (dreistufiger Zoom mittels +/- Taste: 29,3 Sek.). Die Ergebnisse der ANOVA-Auswertung zeigen signifikante Unterschiede hinsichtlich der Kombination von Layouts und Zoomvarianten ($p = 0{,}00$). Der SNK-Test für die Kombination von Designvariante und Zoomvariante bildet hierbei vier Gruppen. Die erste Gruppe mit den signifikant besten Effizienzwerten enthält alle drei Kombinationen mit Designvariante 3, zwei Kombinationen mit Layout 2 und keine Kombination mit Layout 1.

Die höchste Zufriedenheit wird von der Designvariante 2 (Choropleth mit Balkendiagramm) in Kombination mit der Zoomvariante 1 (Stufenloses Scrollen mittels Mausrad) erzielt, währenddessen Layout 3 in Kombination mit der Zoomvariante 2 die geringste Zufriedenheit zeigt. Das Ergebnis der ANOVA bestätigt die Signifikanz der Unterschiede hinsichtlich der Zufriedenheit bei der Kombination des Layouts und der Zoomvariante ($p = 0{,}00$). Im SNK-Test wird die Designvariante 2 in Kombination mit allen drei Zoomvarianten mit der höchsten Zufriedenheitsrate in die erste Gruppe eingeordnet. Alle Kombinationen mit der Designvariante 3 werden in die letzte Gruppe eingeordnet und sind somit am wenigsten zufriedenstellend.

Design + Zoom	Kombination	Mittelwert Effektivität	Mittelwert Effizienz	Mittelwert Zufriedenheit
1	Design 1 Zoom 1	84%	50,3	2,44
2	Design 1 Zoom 2	90%	41,2	2,80
3	Design 1 Zoom 3	90%	41,6	2,45
4	Design 2 Zoom 1	90%	34,4	3,54
5	Design 2 Zoom 2	92%	35,6	3,41
6	Design 2 Zoom 3	93%	31,1	3,45
7	Design 3 Zoom 1	89%	31,2	1,92
8	Design 3 Zoom 2	92%	29,3	1,89
9	Design 3 Zoom 3	94%	29,4	1,94
Design + Zoom	Kruskal Wallis	p = 0,007 Chi Quadrat = 21,018		
	ANOVA		p = 0,000 F = 25,776	p = 0,000 F = 123,555
	Gruppen (SNK)		1. Layout 8 & 9 & 6 & 7 & 4 2. Layout 6 & 7 & 4 & 5 3. Layout 2 & 3 4. Layout 1	1. Layout 5 & 6 & 4 2. Layout 2 3. Layout 1 & 3 4. Layout 8 & 7 & 9

Tabelle 10: Kombination Designvarianten & Interaktionsmöglichkeiten - Mittelwerte, Signifikanzniveaus und Gruppenvergleiche im Überblick

Durch die Beantwortung der Hypothesen H1-H4 können in Kombination mit der deskriptiven Statistik die Aussagen im Hinblick auf die Effektivität, Effizienz und Zufriedenheit für die Kombinationen von Designvarianten und Interaktionskonzepten wie folgt zusammengefasst werden:

Eine interaktive geografische Visualisierung, die für die Aufgaben Identifizieren, Vergleichen und Zusammenfassen von Big Data verwendet wird, soll durch klassische Geschäftsgrafiken (z.B. Balken- oder Kreisdiagramme) erweitert werden und die User zusätzlich mit mehrstufigen Zoomvarianten (Interaktion durch +/- Schaltflächen oder Doppelklick) unterstützen, um die besten Ergebnisse hinsichtlich der Effektivität und der Effizienz erreichen zu können. In Bezug auf die Zufriedenheit gilt es zu beachten, dass die Kombination aus einer Choropleth-Map und einem Balkendiagramm das beste Ergebnis liefert. Die Landmap mit integrierten Kreisdiagrammen schneidet in diesem Teilbereich der Usability-Messung, unabhängig von der Zoomvariante, deutlich schlechter ab. Die Choropleth-Map ohne zusätzliche Visualisierung liefert in allen Bereichen schlechtere Ergebnisse als eine Vergleichsvariante. Deren alleinige Verwendung sollte deshalb für die Managementberichterstattung vermieden werden.

5 Fazit und Ausblick

Im Zeitalter der Digitalisierung und der damit einhergehenden disruptiven Veränderung der Unternehmenssteuerung ist insbesondere das Controlling verantwortlich, die Informationsversorgung im Reporting kontinuierlich im Hinblick auf Benutzerfreundlichkeit und Agilität zu verbessern und somit die Entscheidungsqualität des Managements zu erhöhen. Megatrends im Reporting wie Big Data und Self Service bringen zwar neue Erkenntnisse durch zusätzliche Information und neue Analysemöglichkeiten, das Management ist aber durch das exponentiell ansteigende Datenvolumen und deren selbständiger Analyse gefordert, teilweise sogar überfordert.

Die Gestaltung von Berichten und die damit verbundene Präsentation von Information an das Management muss daher noch stärker in den Fokus des Controllings rücken, um durch effektive und effiziente Informationsverarbeitung den Entscheidungsfindungsprozess des Managements signifikant zu verbessern. Design Guidelines wie die IBCS (Hichert, 2011) unterstützen das Controlling in dieser Aufgabe und können mittels empirischer Eye-Tracking-Tests noch benutzerspezifisch validiert werden. Im Reporting von Big-Data ist neben diesen Gestaltungsempfehlungen eine benutzerzentrierte interaktive Datenaufbereitung, welche den Prozess der Datenexploration und somit den Self-Service-Gedanken unterstützt, unverzichtbar geworden. Neue interaktive Visualisierungstypen wie mehrdimensionale Visualisierungen oder die in diesem Beitrag behandelten Landmaps stellen, insbesondere als Kombination mit klassischen Geschäftsgrafiken in Dashboards, einen Schlüsselfaktor für das Controlling dar.

Eine zunehmende Nutzung dieser Visualisierungen wird nur durch gutes Design und einer damit verbundenen hohen Benutzerfreundlichkeit der neuen Big-Data-Visualisierungen erreicht. Ziel dieser Studie war es daher, die Usability von unterschiedlichen Layout- und Interaktionskonzepten von Landmaps zu vergleichen und zu optimieren, um die Datenvisualisierung mittels interaktiver Karten einem breiteren Publikum zugänglich zu machen.

Die häufig verwendeten Choropleth-Maps visualisieren die Datenausprägungen unterschiedlicher Regionen differenziert, für geordnete Daten mittels unterschiedlicher Helligkeitsstufen bzw. Farbtöne, für ungeordnete Daten mittels differenzierbarer Farben. Um Informationspathologien wie die falsche Einschätzung der Datenausprägung in Abhängigkeit von der Größe der Region (zusätzlich zur Einfärbung der Region) zu vermeiden, verwenden Combined Visualizations eine Verknüpfung aus interaktiver Choropleth-Map und klassischen Geschäftsgrafiken (wie z.B. Balkendiagramm). Zusätzliche Zoominteraktionskonzepte wie „Stufenloses Scrolling“, „+/- Zooming auf mehreren Ebenen“ oder „Zooming durch doppelten Mausklick auf mehreren Ebenen“ sind speziell bei Kombinationen von Landmaps mit integrierten, in die einzelnen Regionen eingebetteten Geschäftsdiagrammen (z.B. Kreisdiagramme) essentiell, um nach einem ersten Überblick wesentliche Detailinformationen für die Attributausprägungen der Regionen und deren Zusammensetzung zu erhalten.

Die drei Landmap-Layouts (Standard-Choropleth-Map, Kombination Choropleth mit einem Balkendiagramm, Kombination Landmap mit integrierten Kreisdiagrammen je Region) wurden mit den erwähnten drei Zoomingvarianten kombiniert und in einer Online-Studie (MTURK) einem Usability-Test unterzogen. Eine Landmap mit integrierten Kreisdiagrammen je Region und der dreistufigen Zoomvariante durch Doppelklick ergibt die effektivsten Ergebnisse (94%), wobei dieser Visualisierungstyp auch am effizientesten war (29,3 sec). Signifikante Unterschiede zur Kombination mit anderen Zoom-Konzepten bzw. zum Layout Choropleth-Map mit Balkendiagramm und Zoom-Interaktionen lassen sich aber im Hinblick auf Effektivität und Effizienz nicht feststellen. Bei der dritten Usability-Komponente, der Zufriedenheit, weist die Choropleth-Map mit Balkendiagramm die höchsten Zufriedenheitswerte auf, speziell in Kombination mit der Interaktion „Stufenloses Scrollen“. Bei der Analyse der Auswirkung der Zoom-Interaktionen auf die Zufriedenheit konnten aber keine signifikanten Unterschiede beobachtet werden.

Die Erkenntnisse der Studie verdeutlichen daher umso mehr die Notwendigkeit, Design und Interaktionskonzept von geografischen Visualisierungen zu optimieren. Bei unterschiedlich komplexen Aufgabentypen (Identifikation, Vergleichen, Zusammenfassen) und mehreren Attributausprägungen ist die Anreicherung des Layouts um klassische Visualisierungen wie Balken- und Kreisdiagramme für eine signifikante Steigerung der Usability essentiell. Durch die Reduktion der Datenmenge auf ausgewählte Regionen ermöglichen diese in Kombination mit Interakti-

onskonzepten ein sehr schnelles und effektives Arbeiten mit den Daten. Im Gegensatz zur Literatur ist ein stufenloses Zooming nicht zwingend erfordert, die kombinierten geografischen Visualisierungen performen mit wenigen Zoom-Stufen sogar besser als die empfohlene Interaktionsvariante.

Limitationen und Forschungsausblick

Das gewählte Forschungsdesign und die Datenstichprobe zeigen Limitationen, die nun diskutiert werden sollen. Die Erhebung der Daten mit einer standardisierten Online-Umfrage und dem Tool Amazon Mechanical TURK verhindert ein weiteres individuelles Nachfragen während der Beantwortung durch die Teilnhehmer. Die Antworten könnten daher durch Fehlinterpretationen beeinflusst werden. Um Missverständnisse zu vermeiden, wurde eine Vorstudie durchgeführt, um die korrekte Formulierung der Aussagen zu überprüfen. Zur Unterstützung erhielten die Teilnehmer Erklärungen zum Datenbestand, zum Forschungsziel und zur Aufgabenstellung. Zusätzlich wurden Videos zur Verfügung gestellt, um dem Benutzer auf Wunsch die jeweiligen interaktiven Visualisierungen zu zeigen und so Lerneffekte zu erzielen.

Die Datenstichprobe enthält zwar verschiedene Länder aus der ganzen Welt, die Mehrheit bildeten aber europäische Länder. Die Auswahl der Länder wurde nach dem Zufallsprinzip getroffen, allerdings repräsentiert die Datenbasis nicht alle Länder weltweit. Die (nicht steuerbare) Konzentration der Teilnehmer der Studie auf die Vereinigten Staaten kann eventuell auch zu einer Verzerrung der Selektion führen, die möglicherweise die Ergebnisse beeinflusst.

Die vorgestellte Studie analysiert die Usability von drei, auf Basis der Literaturrecherche und des Praxiseinsatzes, ausgewählten geographischen Visualisierungen und Interaktionskonzepten. Zusätzlich können in aufbauenden Studien weitere interaktive geographische Visualisierungstypen untersucht werden, wie z.B. die im Literaturteil beschriebene Kombination mit der Interaktionsmöglichkeit „Perceptual Glimpses“ oder die häufig in der Praxis verwendeten dreidimensionalen Landkarten.

Ein weiterer Ansatzpunkt für zukünftige Forschungsarbeit stellt die Layoutgestaltung der geografischen Visualisierungen für unterschiedlich große Displays bzw. Ausgabeformate (z.B. Tablets, Mobiltelefone, Surface Hubs) dar. Interaktive Landmaps sollten, analog dem Responsive Design Ansatz, je nach Ausgabeformat unterschiedlich dargestellt werden (z.B. Anzahl der Pixel, Leistung, Skalierbarkeit, Farbgebung) und dann auch auf ihre Benutzerfreundlichkeit getestet werden.

In diesem Zusammenhang gilt es ebenfalls, die Integration der im Rahmen der Studie auf data driven documents aufgebauten geografischen Visualisierungen in bestehende BI-Tools zu erforschen. Designentwürfe von geografischen Visualisierungen in Tableau und Microsoft Power BI können dem Appendix entnommen werden. Erste Erfahrungen zeigen, dass eine Landmap in Kombination mit einem zusätzlichen Balkendiagramm durch die intuitiv mögliche Dashboardgestaltung in beiden BI-Tools rasch umsetzbar ist. Die Darstellung von Kreisdiagrammen in einzelnen Regionen kann ebenfalls in beiden Tools effizient umgesetzt werden (Abbildung 12). Ein Wechsel von integrierten Kreis- auf Balkendiagramme ist nach unseren Erfahrungen nur in Tableau, dort allerdings eingeschränkt möglich (Abbildung 10). Die kombinierte Darstellung einer Choropleth Map mit integrierten Geschäftsgrafiken ist im Power BI nach aktuellem Stand nicht umsetzbar, in Tableau lässt sich diese Kombination aus spezieller geografischer Visualisierung und klassischer Geschäftsgrafik bereits realisieren (Abbildung 9). Basierend auf diesen Erkenntnissen sollte die rasche Integration in Softwaretools jener kombinierten Visualisierungen, die sich in unserem Paper als besonders benutzerfreundlich herauskristallisiert haben, für eine Weiterentwicklung des Reportings in naher Zukunft hohe Priorität seitens der Softwarefirmen erhalten.

Anhang

Visualisierungsmöglichkeiten in Tableau:

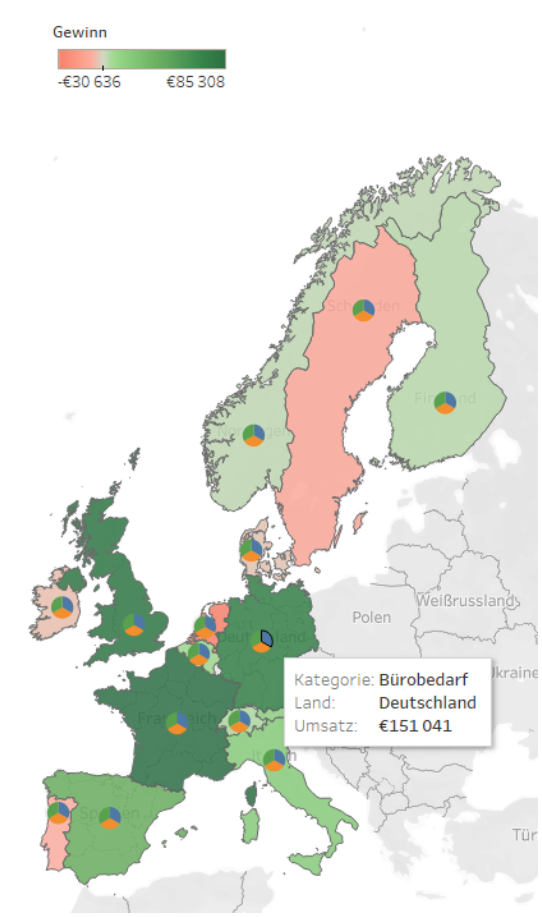


Abbildung 9: Choropleth Map + Kreisdiagramme

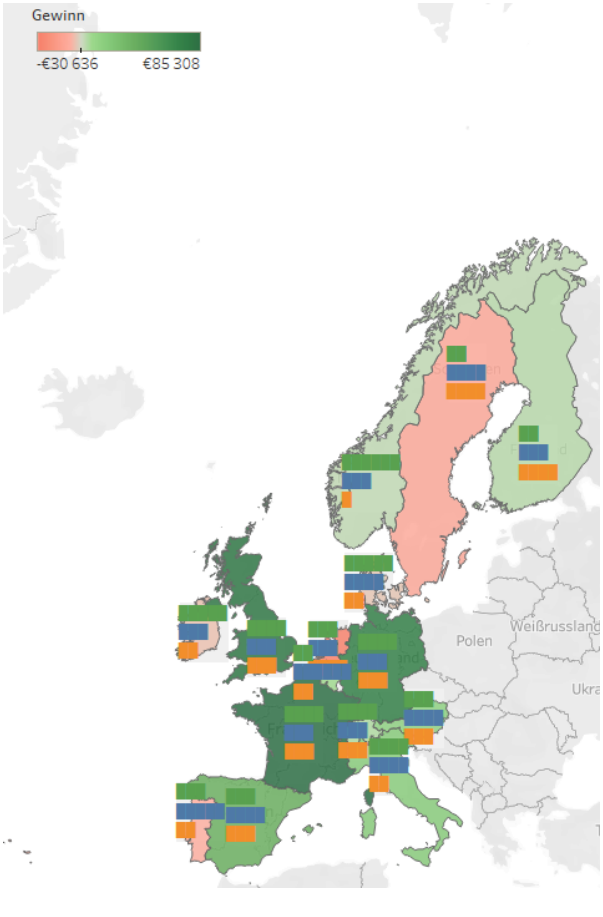


Abbildung 10: Choropleth Map + Balkendiagramme

Visualisierungsmöglichkeiten in Microsoft Power BI

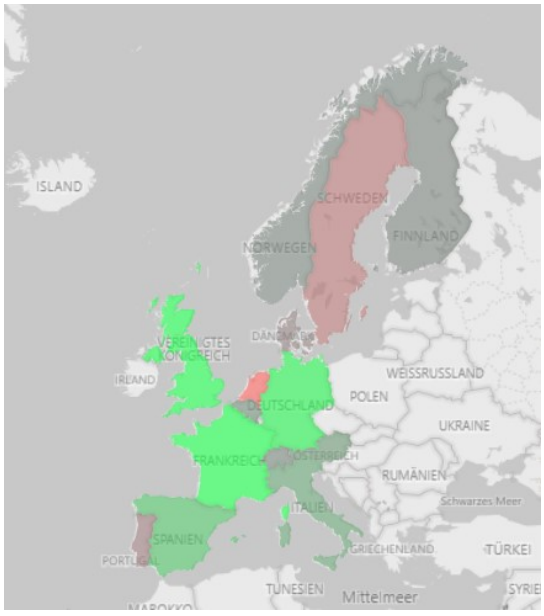


Abbildung 11: Choropleth Map

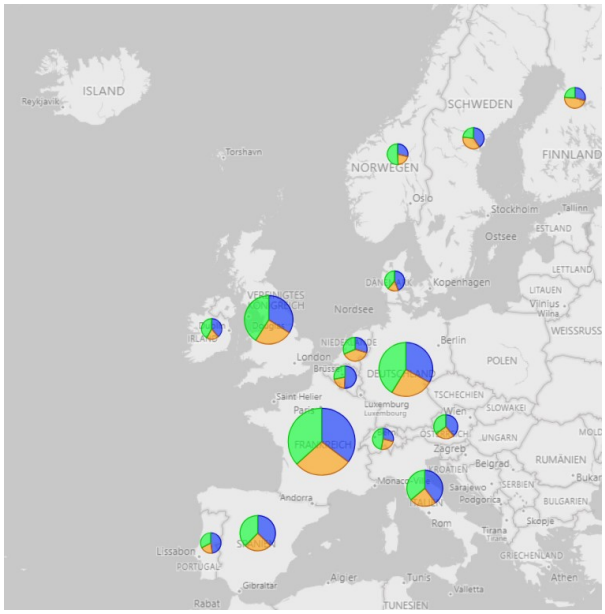


Abbildung 12: Landmap + Kreisdiagramme

Literaturverzeichnis

Börner, K. (2015). *Atlas of knowledge: Anyone can map*. The MIT Press.

Bostock, M., Ogievetsky, V. & Heer, J. (2011). D³: Data-Driven Documents. *IEEE transactions on visualization and computer graphics*, *17*(12), 2301–2309. https://doi.org/10.1109/TVCG.2011.185

Bostock, M. (2017). *Choropleth: Unemployment rate by county, August 2016*. https://observablehq.com/@d3/choropleth. Zugegriffen: 29.06.20.

Brehmer, M. & Munzner, T. (2013). A multi-level typology of abstract visualization tasks. *IEEE transactions on visualization and computer graphics*, *19*(12), 2376–2385. https://doi.org/10.1109/TVCG.2013.124

Brewer, C. A. (1994). Color Use Guidelines for Mapping and Visualization. In A. M. MacEachren & F. D. R. Taylor (Hg.), *Modern cartography: Vol. 2. Visualization in modern cartography* (1. Aufl., Bd. 2, S. 123–147). Pergamon. https://doi.org/10.1016/B978-0-08-042415-6.50014-4

Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences*. Elsevier. https://doi.org/10.1016/C2013-0-10517-X

Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, *112*, 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Crampton, J. W. (2002). Interactivity Types in Geographic Visualization. *Cartography and Geographic Information Science*, *29*(2), 85–98. https://doi.org/10.1559/152304002782053314

Dodge, M. (2008). *Geographic visualization: Concepts, tools and applications*. John Wiley & Sons, Ltd.

Eisl, C., Perkhofer, L., Hofer, P. & Losbichler, H. (2018). *Exzellenz im Reporting Design: Leitfaden für messbar bessere Berichte*. Haufe Group.

Elmqvist, N. & Fekete, J. (2010). Hierarchical aggregation for information visualization: overview, techniques, and design guidelines. *IEEE transactions on visualization and computer graphics*, *16*(3), 439–454.

Few, S. (2009). Introduction to Geographical Data Visualization. *Visual Business Intelligence Newsletter*. http://mail.perceptualedge.com/articles/visual_business_intelligence/geographical_data_visualization.pdf. Zugegriffen: 21.07.20.

Fisher, D. (2007). Hotmap: looking at geographic attention. *IEEE transactions on visualization and computer graphics*, *13*(6), 1184–1191. https://doi.org/10.1109/TVCG.2007.70561

Hichert, R. (2011). Regeln zur (einheitlichen) Gestaltung von Managementberichten und Geschäftspräsentationen. In J. Weber & U. Schäffer (Hg.), *Controlling* (S. 232–237). Schäffer/Pöschel. https://www.ibcs.com/wp-content/uploads/2011/09/Weber_Controlling_Hichert_2011.pdf. Zugegriffen: 06.07.20.

Hofer, P., Mayr, A. & Walchshofer, C. (2018). Interaktive Visualisierungen von Big Data: Ergebnisse einer empirischen Studie. *CFOaktuell*(2), 78–82.

Hofer, P., Walchshofer, C., Eisl, C., Mayr, A. & Perkhofer, L. (2018). Sankey, Sunburst & Co: Interaktive Big Data Visualisierungen im Usability Test. In L. Nadig & U. Egle (Hg.), *Proceedings of CARF 2018: Controlling, Accounting, Risk and Finance* (S. 92–112). Verlag IFZ.

Isenberg, P. & Fisher, D. (2009). Collaborative Brushing and Linking for Co-located Visual Analytics of Document Collections. *Computer Graphics Forum*, *28*(3), 1031–1038. https://doi.org/10.1111/j.1467-8659.2009.01444.x

Li, M., Choudhury, F., Bao, Z., Samet, H. & Sellis, T. (2018). ConcaveCubes: Supporting Cluster-based Geographical Visualization in Large Data Scale. *Computer Graphics Forum*, *37*(3), 217–228. https://doi.org/10.1111/cgf.13414

Li, Q. (2020). *EMBODYING DATA: Chinese aesthetics, interactive visualisation and gaming technologies*. Springer Singapore.

MacEachren, A. M. & Kraak, M.-J. (2001). Research Challenges in Geovisualization. *Cartography and Geographic Information Science*, *28*(1), 3–12. https://doi.org/10.1559/152304001782173970

Mehler-Bicher, A. & Böhm, K. B. (2012). Architektur eines Open Source BI Systems mit Geo-Erweiterung. *Multikonferenz Wirtschaftsinformatik 2012*. http://www.digibib.tu-bs.de/?docid=00047904.

Mittag, H.-J. (2016). Univariate Häufigkeitsverteilungen. In H.-J. Mittag (Hg.), *Springer-Lehrbuch. Statistik: Eine Einführung mit interaktiven Elementen* (4. Aufl., S. 41–62). Springer Spektrum. https://doi.org/10.1007/978-3-662-47132-6_4

Nöllenburg, M. (2007). Geographic Visualization. In A. Kerren (Hg.), *Lecture notes in computer science: 4417 : Tutorial. Human-centered visualization environments: GI-Dagstuhl Research Seminar, Dagstuhl Castle, Germany, March 5-8, 2006 ; revised lectures* (Bd. 4417, S. 257–294). Springer. https://doi.org/10.1007/978-3-540-71949-6_6

Perkhofer, L., Hofer, P. & Walchshofer, C. (2019). BIG Data Visualisierungen 2.0: Optimale Gestaltung und Einsatz neuartiger Visualisierungsmöglichkeiten. In L. Nadig (Hg.), *Proceedings of CARF 2019: Controlling, Accounting, Risk and Finance* (S. 76–104). Verlag IFZ.

Perkhofer, L., Walchshofer, C. & Hofer, P. (2020). Does design matter when visualizing Big Data? An empirical study to investigate the effect of visualization type and interaction use. *Journal of Management Control*, *31*(1), 55–95. https://doi.org/10.1007/s00187-020-00294-0

Petrov, A. (2012). One Hundred Years of Dasymetric Mapping: Back to the Origin. *The Cartographic Journal*, *49*(3), 256–264. https://doi.org/10.1179/1743277412Y.0000000001

Rahlf, T. (2018). *Datenvisualisierung mit R*. Springer Berlin Heidelberg.

Ritchie, J., Wigdor, D. & Chevalier, F. (2019). A Lie Reveals the Truth. In S. Brewster, G. Fitzpatrick, A. Cox & V. Kostakos (Hg.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19* (S. 1–13). ACM Press. https://doi.org/10.1145/3290605.3300423

Schön, D. (2018). *Planung und Reporting im BI-gestützten Controlling*. Springer Fachmedien Wiesbaden.

Schumann, H. & Müller, W. (2000). *Visualisierung - Grundlagen und allgemeine Methoden*. Springer Verlag.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualization. In *IEEE 1996: Proceedings, August 14-16, 1996, Blue Mountain Lake, New York* (S. 336–343). IEEE Computer Society Press.

Shneiderman, B. (2020). *Data Visualization's Breakthrough Moment in the COVID-19 Crisis*. https://medium.com/nightingale/data-visualizations-breakthrough-moment-in-the-covid-19-crisis-ce46627c7db5. Zugegriffen: 13.07.20.

Slocum, T. A. (2009). *Thematic cartography and geovisualization* (3. Aufl.). *Prentice Hall series in geographic information science.*

Stephen, D. M. & Jenny, B. (2017). Automated layout of origin–destination flow maps: U.S. county-to-county migration 2009–2013. *Journal of Maps*, *13*(1), 46–55. https://doi.org/10.1080/17445647.2017.1313788

van Panhuis, W., Cross, A. & Burke, D. (2018). *Counts of Measles reported in UNITED STATES OF AMERICA: 1888-2002*. https://visguides.org/t/visualising-disease-distribution-across-the-us/455. Zugegriffen: 13.07.20.

Verbeek, K., Buchin, K. & Speckmann, B. (2011). Flow map layout via spiral trees. *IEEE transactions on visualization and computer graphics*, *17*(12), 2536–2544. https://doi.org/10.1109/TVCG.2011.202

Weaver, C. (2007). Patterns of coordination in Improvise visualizations. In R. F. Erbacher, J. C. Roberts, M. T. Gröhn & K. Börner (Hg.), *SPIE Proceedings, Visualization and Data Analysis 2007* (64950K). SPIE. https://doi.org/10.1117/12.704570

Yau, N. (2013). *Data points: Visualization that means something*. Wiley.