

**Gestaltung einer Informationsvisualisierung zur
Untersuchung des Stromverbrauchs privater
Haushalte in den jeweiligen Stadtteilen Kölns**

BACHELORARBEIT

ausgearbeitet von

Dorothea Buck

zur Erlangung des akademischen Grades

BACHELOR OF SCIENCE (B.Sc.)

vorgelegt an der

TECHNISCHE HOCHSCHULE KÖLN

CAMPUS GUMMERSBACH

FAKULTÄT FÜR INFORMATIK UND

INGENIEURWISSENSCHAFTEN

im Studiengang

MEDIENINFORMATIK BACHELOR

Erster Prüfer: Prof. Dr. Martin Eisemann
Technische Hochschule Köln

Zweiter Prüfer: Prof. Dr. Horst Stenzel
Technische Hochschule Köln

Gummersbach, 17. August 2016

Adressen: Dorothea Buck
Rösrather Str. 28
51107 Köln
info@dsbuck.de

Prof. Dr. Martin Eisemann
Technische Hochschule Köln
Institut für Informatik
Steinmüllerallee 1
51643 Gummersbach
martin.eisemann@th-koeln.de

Prof. Dr. Horst Stenzel
Technische Hochschule Köln
Institut für Informatik
Steinmüllerallee 1
51643 Gummersbach
horst.stenzel@th-koeln.de

Inhaltsverzeichnis

1	Einleitung	6
1.1	Forschungsfrage	6
1.2	Die Aufgabe und das Ziel	7
1.3	Begriffe	8
2	Grundlagen der Informationsvisualisierung	10
2.1	Definition Informationsvisualisierung / InfoVis	10
2.2	Definition Visual Analytics	11
2.3	Definition Geovisualisierung	11
2.4	Überblick und Vergleich von Visualisierungstechniken	13
2.4.1	Primitive Visualisierungstechniken	13
2.4.2	Multivariate Visualisierungstechniken	14
2.4.3	Fazit	14
2.5	Vergleich von Anwendungen	15
2.6	Referenzen zu angewandten Ansätzen	18
2.6.1	Fazit	21
3	Datensatzanalyse	22
3.1	Datenquellen	22
3.2	Werte (Kennzahlen)	24
3.3	Eigenschaften und Spezifikationen	25
3.4	Fragestellungen zu den Datenwerten	27
3.5	Erste Datensatzanalyse	29
3.5.1	Vorgehen und Testdatenauswahl	29
3.5.2	Naive Analyse mit primitiven Visualisierungstechniken	29
3.5.3	Begründung der Diagrammauswahl	30
3.6	Erkenntnisse aus der Analyse und Bedeutung	30
4	Konzept zur Entwicklung der InfoVis	34
4.1	D3 für RIS-Visualisierungen	34
4.2	Anforderungsanalyse	35
4.2.1	Benutzeranforderungen	35
4.2.2	Visualisierungsanforderungen	37
4.2.3	Anforderungen an den Datenraum	38
4.3	Layoutgestaltung	39
5	Prototypenumsetzung mit D3	42
5.1	Vorbereitung	42
5.2	Preprozess (Daten filtern)	43
5.3	Layout	43
5.4	Mapping (Codierung, Visualisierung)	46

5.5	Diagramme für die Detail-Panels	48
5.6	Nutzerinteraktion	50
5.7	Bugfixes	52
6	Formative Evaluation	55
6.1	Anforderungen	55
6.2	Evaluation mit Cognitive Walktrough	60
6.2.1	Vorbereitungsphase	61
6.2.2	Analysephase	64
6.3	Ergebnis	66
6.3.1	Verbesserungsvorschläge	68
7	Ergebnis und kritische Bewertung	70
8	Zusammenfassung	72
	Abbildungsverzeichnis	74
	Tabellenverzeichnis	75
	Literaturverzeichnis	78
	ANHANG	79
	Eidesstattliche Erklärung	98

Vorwort

Informationen sind zu einem wichtigen Mittel geworden, die Realität in einem Modell abzubilden, um mit diesen Entscheidungen fällen zu können. Dabei besteht die Problematik, auf welche Art und Weise die umfangreichen, heterogenen Daten aus den Statistiken informativ und unverfälscht an die Interessenten kommuniziert werden. In den Medien und im Journalismus wurde die Möglichkeit erkannt, Informationen anhand von einfachen Grafiken zu kommunizieren. Jedoch unterliegen diese Darstellungen, je nachdem wie die Information hervorgehoben wurde, häufig fehlerhaften und missverständlichen Visualisierungen. Eine andere Herangehensweise, die Informationen mit wissenschaftlichen Mitteln und speziellen Visualisierungstechniken zu vermitteln, übersteigt die vorhandene Lernbereitschaft eines Empfängers, für eine neue komplexe, nicht unbedingt schnell interpretierbare Visualisierung.

Der Bund und die Kommunen haben es sich zur Prämisse gemacht, Informationen der Öffentlichkeit zugänglich zu machen und durch mehr Transparenz, das Vertrauen der Bevölkerung zu steigern. Gleichzeitig hat der Bund gegenüber der Internationalen Energie Agentur (IEA), der EU und weiteren Vereinigungen die Pflicht, Daten zur Energiewende zu erheben und Bericht zu erstatten. Die Kommunen publizieren in den jährlichen Jahrbüchern die Entwicklung der Städte und visualisieren die Daten in verschiedenen Darstellungen. Meist tabellarisch dargestellt, um eine reine Übersicht der Werte zu geben oder in einfach zu interpretierenden, klassischen Diagrammen abgebildet. Für alle Daten und Statistiken dieser Arbeit, wurde Köln als Grundlage herangezogen. Im Kontext zu der Energiestatistik bietet Köln keine Zusammenhänge für private Haushalte. Bei der Recherche wurden statt dessen Energiestatistiken der Gebäudetechnik von öffentlichen Gebäuden, wie Schulen oder Bürogebäuden der Stadt Köln verwendet und durch entsprechende Daten- und Analyseergebnisse kommuniziert. Die Informationen zu dem Energieverbrauch privater Haushalte bleibt aus.

Eine weitere Methode der Recherche hat ergeben, dass in der Energiestatistik die privaten Haushalte gesondert betrachtet werden, neben den Sektoren Gewerbe, Handel, Dienstleistungen, Industrie und Verkehr. Allerdings lassen sich in den abgebildeten Energiestatistiken der Kommunen, keine Informationen zu dem Energieverbrauch der privaten Haushalte der Städte finden.

Abgeleitet auf Stadtteile und Demografie unter Berücksichtigung der Analyse der Darstellungen, entwickelte sich für diese wissenschaftliche Arbeit folgende Forschungsfrage: Wie wird eine Visualisierung von Informationen und Daten für Nicht-Experten entwickelt?

1 Einleitung

Thema dieser Arbeit ist eine Informationsvisualisierung zu gestalten, die es Nicht-Experten oder Novizen ermöglicht, einen Überblick des Stromverbrauchs von privaten Haushalten in Köln über die jeweiligen Stadtteile, Haushaltsgrößen und Energieträger, in Kilowattstunde (kWh) zu erhalten.

Die Forschungsfrage teilt sich in zwei Aspekte der Visualisierung auf. Erstens, in die Visualisierung, wie die Daten präsentiert werden und welche Interaktionen für die Analyse verwendet werden sollen und zweitens auf den Datensatz, welche Daten stehen zur Verfügung und welche Fragen lassen sich daraus bilden. Aus diesen Aspekten wird die Aufgabenstellung und die Zielsetzung abgeleitet.

1.1 Forschungsfrage

Fragestellung zur Visualisierung

Kann man und wenn ja, wie kann man eine vereinfachte, interaktive, Visualisierung für komplexe, mehrdimensionale (und relationale) Datenstrukturen über die Aussage: „Der Stromverbrauch der Einwohner der jeweiligen Stadtteile in Köln, nach Haushaltsgrößen und Energieträger in kWh“ gestalten?

Ist es dabei möglich, mit einem geringen Abstraktionsgrad, trotzdem den Überblick über den gesamten Beobachtungsraum zu behalten, ohne dabei einen hohen kognitiven Aufwand für den Nicht-Experten zu erzeugen?

Wie muss die Interaktion gestaltet werden, damit der Nutzer ohne Hürde, effizient die Daten in der Visualisierung, analysieren kann? Welche Hindernisse und wohlbekannte Problematiken in der Visualisierung von mehrdimensionalen Datensätzen können auftreten, die in der Gestaltung berücksichtigt und umgangen/vermieden werden müssen? Kann die Visualisierung als Werkzeug für Entscheidungen in den Domänen der Energie- und Bevölkerungsstatistik eingesetzt werden?

Fragestellung zum Datensatz

Können aus den gegebenen Datensätzen zu Stadtteil, Wohndichte, Haushaltsgröße und Stromverbrauch die Gründe für einen hohen Stromverbrauch der privaten Haushalte, in den jeweiligen Stadtteilen abgeleitet werden und wenn ja, eine entsprechende Lösung zur Minderung des Stromverbrauchs gefunden werden?

1.2 Die Aufgabe und das Ziel

Aufgabenstellung zur Visualisierung

Es soll nach Visualisierungstechniken aus den Wissenschaftsbereichen der Visual Analytics und Informationsvisualisierung ermittelt/gesucht werden, die für eine Gestaltung einer vereinfachten, interaktiven Visualisierung für komplexe mehrdimensionale (und relationale) Datenstrukturen geeignet sind. Mit den visuellen Darstellungen der Daten, aus der Energie- und Bevölkerungsstatistik, werden Aussagen über das Thema „Der Stromverbrauch der Einwohner der jeweiligen Stadtteile in Köln, nach Haushaltsgrößen und Energieträger in kWh“ getroffen.

Anhand der Dateneigenschaften, wird die Art der Visualisierung und dadurch die Interaktionsmethode bestimmt. Dabei werden zuerst die Daten spezifiziert, daraufhin die geeignete Visualisierungsmethoden untersucht und entschieden, mit welcher diese visualisiert werden. Abhängig von den Fragestellungen und Nutzeraufgaben wird durch die Visualisierung schon angedeutet, welche Interaktionsmethode passend implementiert werden sollte.

Aufgabenstellung zum Datensatz

Konzeption und Entwicklung einer interaktiven Informationsvisualisierung, die den Stromverbrauch der Stadt Köln nach Stadtteil und Haushaltsgröße expressiv und effizient für Nicht-Experten abbildet und untersucht werden kann. Die Zielgruppe der Visualisierung „Der Stromverbrauch der Einwohner der jeweiligen Stadtteile in Köln“, sollen hauptsächlich Nicht-Experten sein, somit hat die Visualisierung den Anspruch die komplexen Datenstrukturen in vereinfachte, aussagekräftige und klare Darstellungen zu projizieren. Es soll eine Visualisierung entwickelt werden, mit der die verschiedenen Fragestellungen zu Stadtteilgrößen, Haushaltsgrößen, sowie deren Beziehung zur Bevölkerungsdichte analysiert und beantwortet werden können. Parallel zu diesen Merkmalen sollen jeweils der Stromverbrauch, CO₂-Äquivalente und der Energiebedarf nach Energieträger abrufbar sein.

Zielsetzung

Der Nutzer soll ein Verständnis für die komplexen Datenstrukturen und deren Datenwerte entwickeln, sodass er in der Lage ist die Informationen zu interpretieren, Folgefragen zu stellen und neue Antworten zu finden. Dabei entwickelt der Nutzer ein mentales Modell über die Zusammenhänge der Bevölkerungsstatistik und des Stromverbrauchs im Sektor der privaten Haushalte in Köln. Die vordefinierten Fragestellungen werden beantwortet und die Hypothesen bestätigt oder widerlegt.

1.3 Begriffe

In den folgenden Kapiteln werden fachliche Begriffe verwendet, die vorher definiert werden sollten, um vollstes Verständnis in den Beschreibungen zu sichern und Missverständnisse zu vermeiden.

Novizen / Nicht-Experten Mit Nicht-Experten, werden Nutzer bezeichnet, die sich bisher mit den vorgelegten Informationen nicht beschäftigt haben. Damit sind sie unwissend über die Details der Daten und deren Bedeutung und können über Allgemeinwissen einen Bezug zu den Informationen setzen. Gleichzeitig wird hier definiert, dass die Nicht-Experten wenig Erfahrung mit Visualisierungstechniken haben und nur primitive, klassische Visualisierungsdiagramme, die mit wenigen Dimensionen und Variablen abgebildet werden, korrekt interpretieren können. Die Bezeichnung Novize, definiert einen Nutzer der vorher keine Visualisierungstechniken gesehen oder jemals verwendet hat. Dies ist auch auf den Datenbestand übertragbar.

Visualisierung [Keim u. a. (2006)] Visualisierung ist die Kommunikation von Daten, durch interaktive Interface und hat drei wichtige Ziele: präsentiert effizient und effektiv die Ergebnisse der Analyse, bestätigende Analyse um Hypothesen zu überprüfen, erforschende Datenanalyse als interaktive und normalerweise nicht gerichtete Suche nach Strukturen und Trends. Visualisierung wird verwendet um den Informationsraum zu untersuchen, wenn automatische Methoden scheitern. Außerdem können Hintergrundwissen, Intuition und Entscheidungen nicht automatisiert werden oder als Eingabe für zukünftige Entwicklungen für automatische Prozesse genutzt werden.

Geoinformationssysteme GIS / Rauminformationssysteme RIS [Bill (1999)] GIS / RIS werden für heterogene raumbezogene Daten verwendet und für Entscheidungen in der regionalen Planung oder Entwicklung in der Bevölkerungsstatistik, Wirtschaftsentwicklung, Siedlungsentwicklung, Infrastruktur oder andere regionale raumbezogene Entscheidungen zur Entwicklung oder Beobachtung herangezogen. Die Visualisierung erfordert hohe Interaktivität und wird hauptsächlich für Analysen verwendet.

Effektivität, Expressivität, Angemessenheit Die grundlegenden, allgemeinen Anforderungen sind nach Schumann u. Müller Expressivität, Effektivität und Angemessenheit. Expressiv ist eine Visualisierung nur, wenn sie die in einer Datenmenge enthaltenen Informationen darstellt. Die Datenmenge muss unverfälscht wiedergegeben werden. Als effektiv wird die Fähigkeit des Betrachters und des Ausgabegerätes unter Berücksichtigung der Zielsetzung und des Anwendungskontextes beschrieben. Je nach Eigenschaft einer Datenmenge, wird die beste Darstellungsform gewählt, in der die Informationen intuitiv von Betrachter wahrgenommen werden können. Angemessenheit beschreibt Nutzen und Aufwand, die zur Durchführung des Visualisierungsprozesses benötigt werden. Die Angemessenheit und Effektivität einer visuellen Darstellung werden mit dem Aufwand verknüpft. Kosten werden mit der zeitlichen, physischen und kognitiven Belastung des Betrachters ausgedrückt [Schumann u. Müller (2000), S.9ff].

Multidimensional / Multivariat Multidimensional und Multivariat müssen unterschieden werden. Dimensional beschreibt im Kontext der Visualisierung eine unabhängige Variable, die nicht veränderbar ist. Im Gegenzug beschreibt variat / variabel eine abhängige Variable. Ein multidimensionaler Datensatz ist ein Datensatz mit mindestens zwei skalaren Variablen (Dimensionalität eines Datensatzes = Variablen; $k = n + m$; die Gesamtanzahl der erfassten Dimensionen als Merkmalsraum) [Lehmann u. a. (2010)]. In dieser Arbeit werden in der Visualisierung sogenannte Modifikation verwendet, die die verschiedenen Dimensionen darstellen und die Interaktion bereitstellt, um dazwischen zu wechseln. Die Daten, die sich während des Modifikationswechsels verändern, sind die Multivariate.

Thematische Dimensionen, Layer- Folienprinzip oder Ebenenmodell [Lücking, Heinz-Josef, et. al. (2016)] In der GIS spricht man bei den oben genannten Modifikationswechseln als thematische Dimensionen, die in der Kartendarstellung als Ebenenmodell bezeichnet, übereinander gelegt werden. Durch Hinzunahme von Dimensionen werden die räumlichen Daten erweitert. In dieser Arbeit werden die Dimensionen nicht hinzugefügt, sondern gewechselt, um die Karte nicht zu überfüllen und den Kontext der Dimensionen bewusst nicht zu vermischen.

Quelldaten / Primärdaten und Sekundärdaten [Lücking, Heinz-Josef, et. al. (2016)] Mit den Begriffen Quell- und Primärdaten werden die erhobenen Rohdaten beschrieben. Meist sind diese Daten nicht verarbeitet und geben die Werte in roher und unverarbeiteter Form wieder. Sobald die Datenwerte bearbeitet und daraus weitere Datenwerte abgeleitet werden, nennt man diese Sekundärdaten. Die Quelldaten, die für diese Arbeit verwendet werden, sind Grundlage für weitere Daten, die aus der Energiestatistik und aus der Bevölkerungsstatistik abgeleitet werden.

2 Grundlagen der Informationsvisualisierung

In diesem Kapitel geht es darum, die Ansätze der Informationsvisualisierung und der Visual Analytics in Verbindung zur Geovisualisierung zu setzen. Auf dieser Grundlage kann eine Visualisierung entwickelt werden, mit welcher thematisches Wissen erlangt wird. Der Fokus liegt auf den raumbezogenen Daten aus der Bevölkerungsstatistik, der festgesetzten Zielgruppe, den Dateneigenschaften und den Fragestellungen aus der Einleitung in Kapitel 1. Mit diesen werden die Ansätze der Informationsvisualisierung eingegrenzt. Des weiteren werden zum Vergleich existierende GIS Anwendungen vorgestellt, welche zeigen weshalb in dieser Arbeit eine GIS angewendet wird, die nicht komplexe Visualisierungstechniken für Spezialisten verwendet. Auch die primitiven Visualisierungstechniken werden den Techniken gegenübergestellt, die für mehrdimensionale und multivariate Daten entwickelt wurden.

2.1 Definition Informationsvisualisierung / InfoVis

[Reiterer u. Jetter (2013)] In der InfoVis sind die Ausgangsdaten abstrakte, textuelle oder numerische Daten. Möglich sind auch Baum- und Netzwerkdatenstrukturen, die im Prozess der visuellen Darstellung zur gegenständlichen Betrachtung geführt werden können. Die abstrakten Daten werden in der Informationsvisualisierung (InfoVis) in visuelle Strukturen gebracht, die für das menschliche Auge erfassbar sind und infolgedessen auch für das Verständnis. Es werden folgende Strukturen der Visualisierung unterschieden:

- Eindimensionale Strukturen
- Zweidimensionale Strukturen
- Dreidimensionale Strukturen
- Multidimensionale Strukturen (kein Informationsverlust und Vergleiche mit vielen Dimensionen möglich)
- Netzwerke und Bäume (Darstellung von Beziehungen)

Die InfoVis unterstützt und verstärkt die menschliche Kognition. Im Besonderen die menschliche Fähigkeit zur Erkennung von Mustern, Anhäufungen, Ausreißern und Lücken in großen Datenmengen.

"Der Zweck aller interaktiven Visualisierung ist es den Benutzer bei der Durchführung von wissensintensiven Aufgabenstellungen zu unterstützen." [Reiterer u. Jetter (2013)]

2.2 Definition Visual Analytics

[Keim u. a. (2006)] Visual Analytics ist die Wissenschaft analytischer Argumentation, unterstützt durch interaktives visuelles Interface. Viele Probleme sind komplex und erfordern einen frühen Miteinbezug von Menschen im Datenanalyse-Prozess. Die visuellen Analysemethoden ermöglicht Entscheidern ihre Flexibilität, Kreativität und Hintergrundwissen mit großen Speicher- und Verarbeitungskapazitäten von Computern, komplexe Probleme einzusehen. Ziel der visuellen Analyse ist es, aus der Informationsüberladung eine Möglichkeit für Entscheider zu geben den Informationsfluss zu untersuchen, um effektiv und in Echtzeit auf die Situation zu reagieren.

Ursprünglich entwickelte sich Visual Analytics aus der Informationsvisualisierung und automatischen Datenanalyse. Die Visual Analytics beachtet interdisziplinäre Felder wie Mensch Computer Interaktion (MCI), Datenanalyse, Datenmanagement, geo-räumliche und temporäre Datenverarbeitung und Statistiken, mit welchen effektive Visualisierungen entwickelt werden können.

2.3 Definition Geovisualisierung

[MacEachren u. a. (2004)] Geovisualisierung ist eine interdisziplinäre Wissenschaft zu Geo- Informationssysteme (GIS), sie stammt aus dem Bereich der Kartografie und Geografie und verbindet heutige Wissenschaftsbereiche mit Informationsvisualisierung, explorative Datenanalyse (EDA) und Bildanalyse. Aus diesen Wissenschaftsbereichen adaptiert die Geovisualisation Methoden und Werkzeuge, um eine Visualisierung bereitstellen zu können, mit der Experten oder Novizen in der Lage sind die Präsentation visuell zu untersuchen, zu analysieren und neues Wissen zu erlangen. MacEachren u. a. beschreibt damit die vier Funktionen der Geovisualisierung und veranschaulicht dies in der Abbildung 2.1.

Mit den heutigen Ansätzen in der Informationsvisualisierung können interaktive Präsentationen entwickelt werden, die die Bereiche aus der MCI und Kognitions-Psychologie hinzunehmen, um

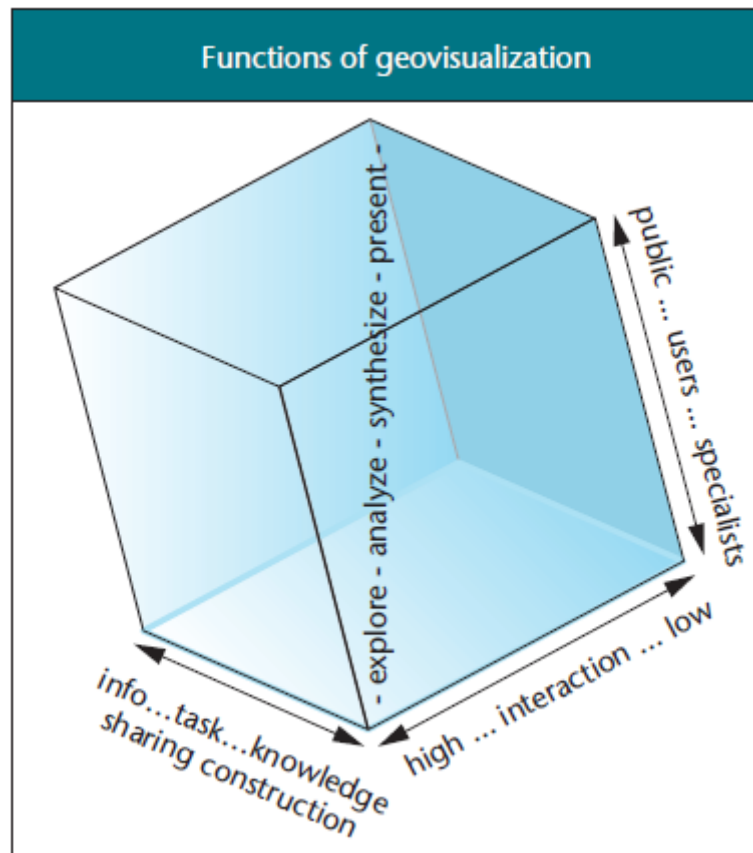


Abbildung 2.1: Die vier Funktionen der Geovisualisierung. Aus [MacEachren u. a. (2004)].

Aufgaben zur Untersuchung der Daten zu entwickeln, Wissen zu erlangen und Entscheidungen zu erleichtern.

Aufgrund der Interdisziplinarität in der Geovisualisierung besteht die Grundlage aller Informationsvisualisierungen aus Datensätzen unterschiedlichster Domänen, dies macht die Daten grundsätzlich heterogen, da sie in ihren Eigenschaften, Typen und Thematiken unterschiedlich sind. Dem entsprechend müssen Visualisierungen entwickelt werden, um den Anspruch der Korrektheit, Glaubwürdigkeit, Effektivität, Expressivität und Angemessenheit gerecht zu werden. In den letzten 50 Jahren wurden viele Visualisierungstechniken entwickelt, um multivariate und mehrdimensionale Datensätze angemessen zu präsentieren und besser analysieren zu können.

Zusammenfassend gliedert sich die Geovisualisierung hervorragend in die Disziplinen der Informationsvisualisierung und Visual Analytics ein. Sie kann somit von den Methoden und Werkzeugen, die entwickelt worden sind, profitieren.

2.4 Überblick und Vergleich von Visualisierungstechniken

Die Definition der primitiven Visualisierungstechniken wurden aus der vorangehenden Praxisprojekt Arbeit [Buck (2014)] übernommen. Dabei werden die drei gängigsten und einfachsten Diagramme beschrieben, die selbst Novizen kennen sollten und effizient interpretieren können.

2.4.1 Primitive Visualisierungstechniken

Kreisdiagramm

Ein Kreisdiagramm veranschaulicht Mengenverhältnisse in radialer Form. Durch die zusammenhängende Kreisform, werden die Daten als Ganzes assoziiert und lassen proportionale Unterschiede der Messwerte leicht erkennen. Die Darstellung wird meist mit unterschiedlichen Farben unterstützt, um die nominalen Daten voneinander abzugrenzen oder durch monochrome Farbwahl, um die Zusammengehörigkeit der Daten zu stärken. Es werden quantitative Daten mit nominalen, unabhängigen Variablen korreliert. In dieser Form ist die Darstellung von ordinalen Datentypen in geordneter oder indexierter Reihenfolge ineffizient. Die Ordnung nach Menge, von groß auf klein ist hingegen unterstützend, da Werte, die sich kaum voneinander unterscheiden, schwer abzuschätzen sind und nicht ohne Ergänzung ein Verhältnis erkennen lassen.

Säulen- und Balkendiagramm

Das Säulendiagramm bildet Mengenverteilungen von nominalen oder diskreten, unabhängigen Variablen (Abszisse) zum Vergleich ab. Für die Säulenform wird für jedes Merkmal der unabhängigen Variable eine Verbindung von der x-Achse bis zum Messpunkt der quantitativen y-Achse (Ordinate) gezogen. Der Abstand zwischen den Säulen verdeutlicht die Unabhängigkeit der Merkmale auf der Abszisse. Durch Korrelation mit weiteren unabhängigen Variablen vom gleichen Datentyp, können direkte Vergleiche mittels Gruppierung stattfinden. Balkendiagramme tauschen die Ordinate und die Abszisse im kartesischen Koordinatensystem. Bei der Darstellung ist die Anzahl der Merkmale auf der Ordinate geringer und die Messwerte bekommen eine höhere Relevanz. Im Säulendiagramm ist die Unterscheidung der Merkmale auf der unabhängigen Variablen wichtiger.

Linien- und Kurvendiagramm

Linien- und Kurvendiagramme eignen sich für die Repräsentation von Trends und lokalen Strukturen. Wie bei Säulendiagrammen ist die x-Achse die Abszisse, die eine ordinale, unabhängige Variable abbildet. Die Ordinate an der y-Achse bildet die quantitative Skala ab. Die Merkmalsausprägungen werden an den Schnittpunkten positioniert und die Nachbarn mit Linien oder Kurven verbunden bzw. interpoliert. Damit lassen sich die Punkte segmentweise gruppieren und Veränderungen der Menge entlang der Abszisse effektiv erkennen.

2.4.2 Multivariate Visualisierungstechniken

Scatterplot / Scatterplot Matrix [Chung Wong u. Bergeron]

Eine etablierte Methode ist die Scatterplot Methode (Streudiagramm). Die Visualisierung überträgt auf der x- und y-Achse die Variable und ist demnach für zweidimensionale Präsentationen geeignet. Eine Möglichkeit wäre auch, eine Dimension durch Farbe oder Größe der Punkte hinzuzufügen. Streudiagramme werden für statistische Korrelationsanalysen verwendet. Eine Erweiterung ist die Scatterplot Matrix (SPLOM), die multivariate Visualisierungen ermöglichen. Die Scatterplots werden pro Variablenpaar in allen Kombinationen horizontal oder vertikal adjazent in Beziehung gesetzt.

Parallel Coordinates [Chung Wong u. Bergeron]

Im parallelen Koordinatensystem werden die Achsen eines multidimensionalen Raums vertikal nebeneinander aufgestellt. Ein Skalar in einem kartesischen Koordinatensystem, entspricht einer sogenannten Polyline im parallelen Koordinatensystem. Durch die Auffächerung der Achsen in parallele vertikale Achsen, kann der Nutzer die mehrdimensionalen Daten besser visualisieren. Die Beziehung der Variablenpaare, zwischen zwei Achsen für alle Variable werden schneller erkannt und verstanden. Bis zu einer bestimmten Datenmenge ist die Visualisierung anwendbar. Wird die darstellbare Grenze überschritten, überdecken die Variablen Linien sich gegenseitig und sind nicht mehr einzeln zu erkennen.

2.4.3 Fazit

Es hat sich gezeigt, dass multidimensionale und -variate Visualisierungstechniken meist in wissenschaftlichen Bereichen angewendet werden und Nicht-Experten, diese Techniken meist unbekannt sind. Dazu sind diese Techniken alles andere als schnell zu interpretieren und fordern

einen hohen Kognitionsaufwand, bei dem der Novize in dem Fall, zuerst die Elemente, worauf die Daten abgebildet sind, zu verstehen versucht. Und selbst nachdem der Novize diese verstanden hat, kann er nicht sicher sein ob seine Interpretation korrekt ist. Die Arbeit von Lee u. a. (2016) bestätigt die Befürchtung, dass der Nutzer bei einer Verwendung von komplexen Visualisierungstechniken, nicht in der Lage sein wird, effizient Informationen aus der InfoVis abzuleiten. Im schlimmsten Fall führt es zu Fehlinterpretationen oder der Nutzer versteht nicht wie er die Informationen aus der Visualisierung ablesen soll. Aus diesem Grund wäre es für das Ziel dieser Arbeit gut, die primitiven Visualisierungstechniken zusammen mit den GIS zu nutzen, um damit für den Novizen eine expressive InfoVis zu gestalten.

2.5 Vergleich von Anwendungen

Neuere Visualisierungsmethoden werden mit Kombination von klassischen Techniken erweitert und unterstützen den Anwender bei der Ansicht und Untersuchung von komplexen Daten. Durch neuere Visualisierungstechniken wird eine effektivere Informationsvisualisierung, die nach den Methoden der MCI und analytischen Statistiken, erarbeitet werden. Damit werden die Datensätze aussagekräftiger und erzeugen weitere Ansichten, die untersucht werden können. Der Nutzer kann dadurch auch andere Ansätze der Analyse weiterverfolgen. Die folgenden GIS Anwendungen zeigen wie solche Visualisierungstechniken mit geografischen Karten kombiniert werden. Dabei wird auffallen, dass die Interaktionsmittel linking and brushing eine essentielle Bedeutung in der explorativen Untersuchung der thematischen Daten haben.

GeoDa [Anseli u. a. (2006)]

Eine auf GIS spezialisierte Software, mit Adaption von statistischen, automatischen Analysemethoden. GeoDa besteht aus einer interaktiven Umgebung, die Karten mit statistischen Grafiken, mit der Technologie von dynamisch verknüpften Fenster kombiniert. Die Funktionen von GeoDa sind auf die Analyse von diskreten geografischen Daten ausgerichtet und Objekte werden durch ihre Position im Raum entweder als Punkte (Punktkoordinaten) oder Polygone (Polygongrenzenkoordinaten) gekennzeichnet.

GeoDa ist ein einfach zu bedienendes, visuelles und interaktives Software-Paket für nicht GIS Anwender. Der GeoDa Anwender interagiert mit einem Point and Click Interface und benötigt keine Kenntnisse aus der Programmierung. Das Programm nutzt linking and brushing in dynamischen Grafiken für die explorative Datenanalyse (ESDA). Die Verwendung von dynamischen linking und brushing ist die zentrale Organisationstechnik für die Datenvisualisierung und hat eine starke Tradition in explorativen Datenanalyse (EDA). GeoDa adaptiert Histogramm, Box-Plot, Streudiagramm, Karte und Gitter (für die Tabellenauswahl und Berechnungen). Eine

Choroplethenkarte, für die Signifikanz der Daten und Cluster-Karten für die lokalen Indikatoren der räumlichen Autokorrelation in Abbildung 2.2 zu sehen.

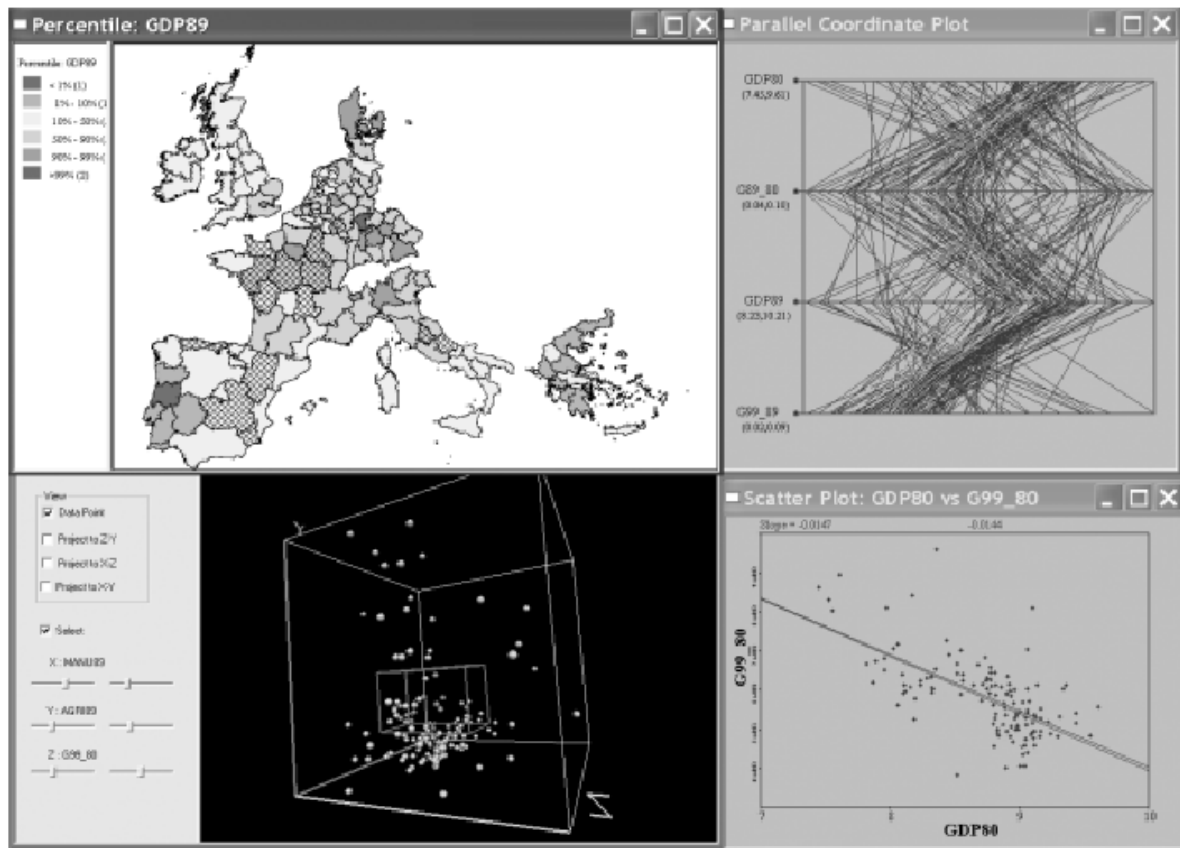


Abbildung 2.2: GeoDa adaptiert hier Streudiagramm (2D und 3D), Parallel Coordinates mit der Choroplethkarte. Aus [Anseli u. a. (2006)].

GeoVISTA Studio [MacEachren u. a. (2004)]

GeoVISTA Studio ist eine offene Software für geowissenschaftliche Analysen und unterstützt geografische Visualisierungen mit Komponenten basierten Gestaltung der Visualisierung. Damit können nach Domänen spezifische Datenkomponenten zur Visualisierung, Analyseaufgaben und zur Evaluierung eine Oberfläche zur Untersuchung von räumlichen Daten erstellt werden. Komplexe Funktionen lassen sich mit den Visualisierungen verknüpfen und verbessern die Analysefähigkeiten und erleichtert damit den Entscheidungsprozess. Damit wird ein Werkzeug bereitgestellt, welches es mit wenig Programmiererfahrung ermöglicht Geo-Informationssysteme zu gestalten. Zum Beispiel können Daten aus dem Anwendungsbereich des Gesundheitssystems mit geografischen und mehrdimensionalen Visualisierungstechniken kombiniert werden zeigt Abbildung 2.3.

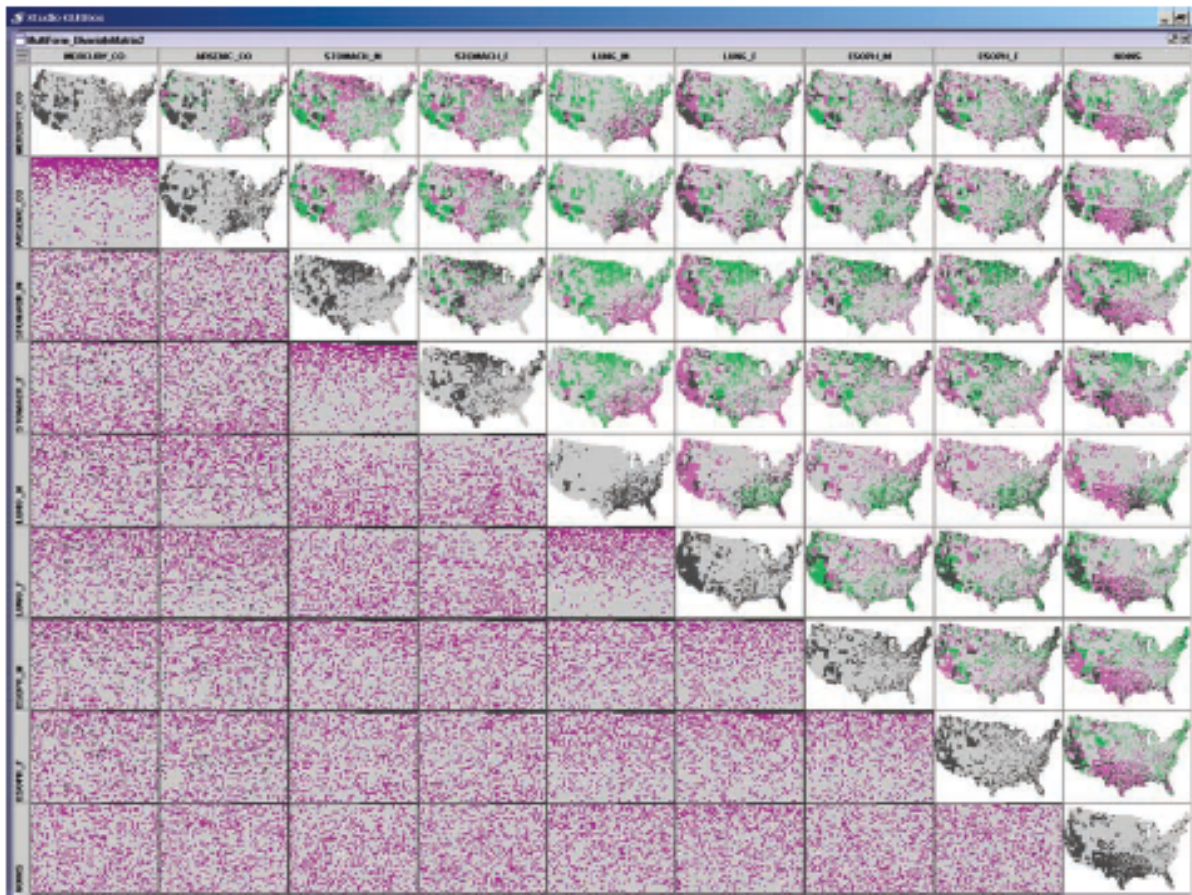


Abbildung 2.3: GeoVISTA kombiniert geografischen und mehrdimensionalen Visualisierungstechniken. Aus [MacEachren u. a. (2004)].

iVisDesigner [Ren u. a. (2014)]

iVisDesigner ist ein webbasiertes System, das es Benutzern ohne Programmierkenntnissen ermöglicht, Informationen für heterogene Datensätze interaktiv zu visualisieren. Der Nutzer erstellt Visualisierungen und Entwürfe für Endbenutzeroberflächen mit dem webbasierten Interface. Auch kann der Nutzer den Daten grafische Elemente zuordnen und das System ermöglicht zusätzlich, dass die Elemente gleichzeitig erstellt und bearbeitet werden können. Dies erleichtert den Aufwand Visualisierungen zu erstellen. Transformierte, aggregierte oder auf andere Weise erzeugte Daten können dem Datensatz hinzugefügt werden. Es stellt Funktionen wie Histogramme und Diagramm Layouts bereit. Grafische Elemente können mit ziehen (dragging) und brushing manipuliert werden, welches die zugrunde liegenden Datenattribute beeinflusst und die Gestaltung von interaktiven Visualisierungen ermöglicht. Bisher sind die Visualisierungen in iVisDesigner überwiegend statisch, bis die Visualisierungen neu erstellt werden müssen, wenn der Datensatz geändert wird.

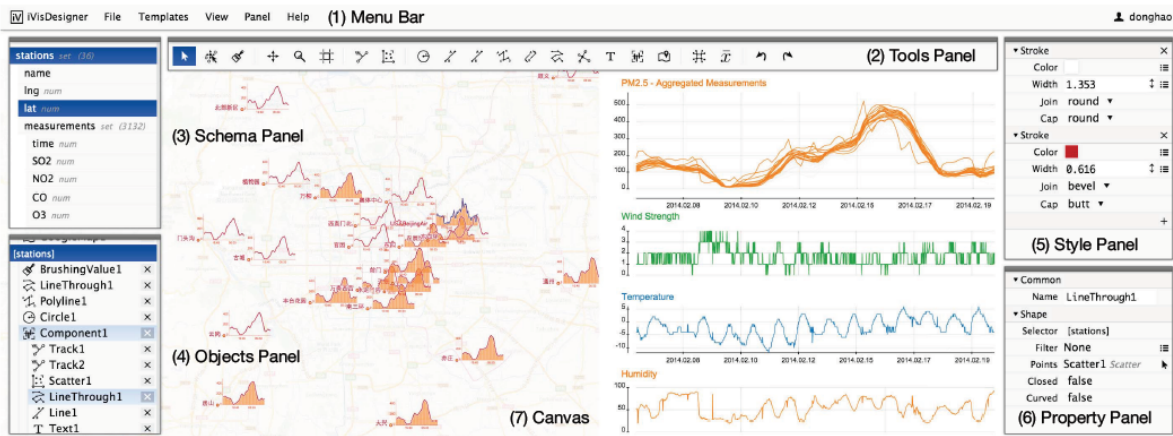


Fig. 2. The interface of iVisDesigner. (1) Menu Bar: Commands for loading or saving visualization designs, view settings, login and logout. (2) Tools Panel: Tools for moving objects around, creating new objects and changing the view. Grouped into different categories. (3) Schema Panel: Shows the structure of the dataset. Allows selection. (4) Objects Panel: Shows the objects currently in the visualization design. Allows selection. (5) Style Panel: Adjust graphical styles for currently selected objects. (6) Property Panel: Adjust properties of currently selected objects. (7) The canvas to draw the visualization. In this example, a visualization of Beijing Air Pollution data is presented. There are two linked views, the left view shows the timeline plots of PM2.5 indexes for each measurement station on top of a map, the right plots shows the trends of the PM2.5 indexes, wind strength, temperature and humidity. This visualization is designed solely through user interaction with iVisDesigner, without textual programming.

Abbildung 2.4: iVisDesigner Interface mit Beschreibung. Aus [Ren u. a. (2014)].

2.6 Referenzen zu angewandten Ansätzen

Visual Seeking Mantra [Shneiderman (1996)]

Aus der Beobachtung, dass sich bestimmte Interaktionen immer wiederholen, hat Shneiderman das Visual Information Seeking Mantra erfunden. Shneiderman beschreibt sieben Aufgaben mit hohen Abstraktionsgrad. Diese sind:

1. **Überblick:** Um einen Überblick über den Zusammenhang von Daten zu erhalten zoomt man heraus
2. **Zoom:** Man zoomt an ein Element heran, welches von Interesse ist und behält den Kontext bei
3. **Filter:** Uninteressante Elemente werden entfernt
4. **Details auf Anfrage:** Ein Element oder eine Gruppe von Elementen wird für mehr Details ausgewählt
5. **Beziehung:** Man beobachtet Beziehungen zwischen den Elementen
6. **Historie:** Um Operationen rückgängig zu machen, zu verfeinern oder nachzuvollziehen

7. Extrahieren: Entdeckte Elemente können festgehalten werden, durch speichern, ausdrucken, exportieren

Im Konzept (Kapitel 4) werden die Bearbeitungsziele für die Aufgaben abgebildet.

Shneiderman zählt ebenso sieben Datentypen auf, die gleichwohl wie bei Card u. Mackinlay, Chi und Schumann u. Müller [Buck (2014)] zur Klassifizierung für Visualisierungstechniken unterschieden werden und spricht diesen bestimmte datentypische Aufgaben zu, welche die Nutzer ausführen müssen, damit Datentyp bzw. spezifische Visualisierungsprobleme gelöst werden können. Die relevantesten werden hier beschrieben:

- 1. Eindimensional:** Die Daten werden in geordnete Listen abgebildet und sind mit einfachen Filter-, Überblick-, Auswahl- und Navigation-Aufgaben bearbeitbar. Die Aufgaben aus den eindimensionalen sind auch für den zweidimensionalen Datentyp geeignet.
- 2. Zweidimensional:** Abbildung von räumlichen Daten; Beinhaltet grundlegende Aufgaben wie Zählen, Filtern und Details auf Anfrage. Die Aufgaben dieses Datentypes können auch im ein-dimensionalen angewendet werden.
- 3. Zeitliche:** Zeitrahmen werden verwendet, um einen Datentyp zu erstellen, der von den eindimensionalen Daten getrennt ist. Sie haben eine Start- und Endzeit. Die Elemente können sich überlappen.
- 4. Multi-dimensional:** Bildet relationale und statistische Datenbanken in n-dimensionalen Datenräumen ab. Die Aufgaben umfassen die Ermittlung von Mustern, Cluster, Korrelationen, Lücken und Ausreißern und erfordern eine benutzergesteuerte Visualisierung.

What does the user want to see? What do the data want to be? [Pretorius u. Van Wijk (2009)]

Informationsvisualisierung Designer müssen ihre Nutzer, die Daten und ihre zu Grunde liegenden Werkzeuge in der Visualisierung überprüfen. Die erste Frage „Was möchte der Nutzer sehen?“ behandelt das Problem in Bezug auf die Anforderungen, die erfüllt werden müssen. Die zweite Frage „Was die Daten sein möchten?“ nähert sich dem Problem, indem man bedenkt, wie die Daten dem Benutzer präsentiert werden. Die folgende Methodik ist hilfreich, um interaktive Systeme zu entwerfen:

- 1. Identifizieren Sie Benutzeranforderungen.**
- 2. Entwickeln Sie alternative Entwürfe, die die Anforderungen erfüllen.**

3. Implementieren Sie die Entwürfe in einen interaktiven Prototypen.

4. Werten Sie die Prototypen aus.

In diesem Artikel wird argumentiert, dass es für die Gestaltung von InfoVis Techniken sinnvoll ist, nicht nur Nutzeranforderungen zu berücksichtigen, sondern auch die Daten sorgfältig zu prüfen, die visualisiert werden sollen. Dies bedeutet, auf die Daten einen genauen Blick zu haben, die kennzeichnenden Merkmale, die derzeit nicht ausreichen ans Licht gebracht wurden zu beachten und die Gestaltung von Visualisierungen für diese hervorzuheben.

Referenz Modell / Visualisation Pipeline [Haber u. McNabb (1990)]

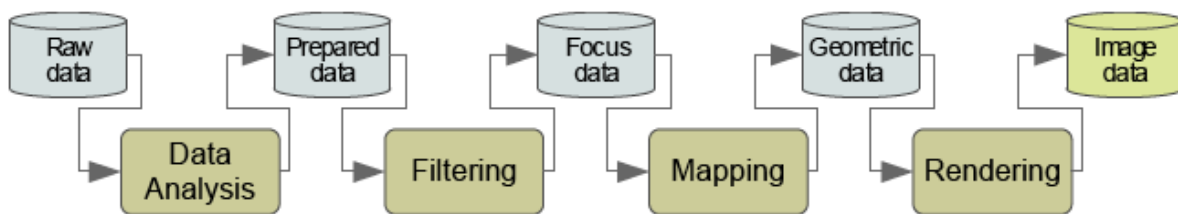


Abbildung 2.5: Die Visualisierungspipeline beschreibt (schrittweise) den Prozess der visuellen Darstellung von Daten. Aus [Hammerl, Michael, et. al. (2016)].

- 1. Datenanalyse:** Daten werden für die Visualisierung vorbereitet (zum Beispiel mit Glättungsfilter, Interpolation von fehlenden Werten, oder Fehlmessungen zu korrigieren) - in der Regel computerzentriert, ein wenig oder keine Benutzerinteraktion.
- 2. Filtering:** Die Auswahl der Datenteile werden sichtbar gemacht - in der Regel benutzerzentriert.
- 3. Mapping:** Fokusierte Daten werden auf geometrische Primitive abgebildet (zum Beispiel Punkte, Linien) und deren Attribute (zum Beispiel Farbe, Position, Größe); kritischster Schritt, um Expressivität und Effektivität zu erreichen.
- 4. Rendering:** geometrische Daten werden in Bilddaten umgewandelt.

Das Referenzmodell Schema stellt eine allgemeine Vorlage für die Strukturierung von Visualisierungsanwendungen dar und trennt die Datenmodelle, visuelle Modelle, Ansichten und interaktive Steuerelemente.

Daten und visuelle Modelle werden getrennt, um mehrere Visualisierungen von einer Datenquelle zu ermöglichen. Außerdem ermöglicht die Trennung der visuellen Modelle von Displays, mehrere

Ansichten einer Visualisierung und verwendeter modularer Steuerungen, um Benutzereingaben in einer flexiblen und wiederverwendbare Art und Weise zu handhaben.

2.6.1 Fazit

Pretorius und Van Wijk stimmen mit Haber und McNabb überein, dass die Daten zu allererst sorgfältig zu prüfen sind. Dies geschieht im Referenzmodell im ersten Schritt Datenanalyse, wobei die Daten nicht nur geprüft, sondern auch für die Visualisierung vorbereitet werden.

Es handelt sich bei dem Referenzmodell um ein allgemeines Vorgehensmodell in der Anwendungsentwicklung, weshalb das Referenzmodell sich in Schritt vier dem „prototyping“ einer InfoVis eingliedert. In diesem Projekt wird die Datenanalyse aus beiden Modellen und der Filtering Schritt ausgegliedert und als erster Schritt vor der Identifizierung der Benutzeranforderungen angeordnet. Mit der Synthese beider Schemata, entspricht dies den Vorschlägen. Siehe Abbildung 2.6. Für den Entwickler und Gestalter einer InfoVis wird es als sinnvoll erachtet, die Daten, ihre Bedeutung und die Merkmale genau zu kennen, um auf dieser Grundlage angemessene Benutzeranforderungen und Visualisierungsanforderungen abzuleiten.

Auf das Visual Seeking Mantra wird im Konzept Kapitel 4 näher eingegangen.

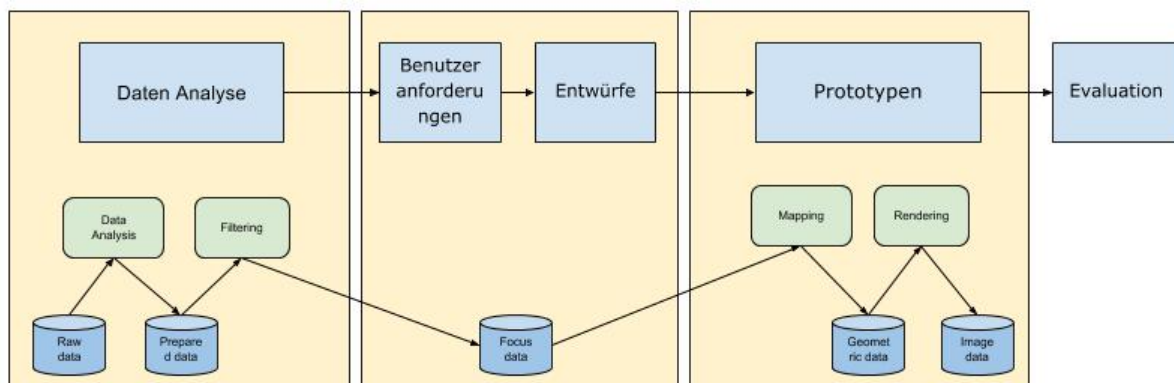


Abbildung 2.6: Synthese des Referenzmodell und Vorgehensweise zur Gestaltung einer InfoVis.
Erstellt von Dorothea Buck.

3 Datensatzanalyse

Die Grundlage aller Informationsvisualisierungen sind die Daten aus denen man die Informationen interpretieren kann. So ist es plausibel im ersten Schritt der Gestaltung die Daten genau zu spezifizieren. Nur so kann eine angemessene Visualisierung entwickelt werden, die von den Nutzer korrekt interpretiert werden kann. Nach Pretorius / van Wijk und Haber / McNabb sollte zuerst eine Datenanalyse vorgenommen werden, in der wie im Vorkapitel 2 Grundlagen der Informationsvisualisierung erklärt, die Daten überprüft und vorbereitet werden. In diesem Kapitel werden ausführlich die ersten Erkenntnisse dieses Schrittes beschrieben. Angefangen bei der Herkunft der Datensätze, welches die bedingten Entscheidung bei der Auswahl erklärt.

3.1 Datenquellen

Während der Suche nach den benötigten Datensätzen, wurden öffentliche Datenbanken abgefragt, die häufig die benötigten Daten zur Verfügung stellen konnten. Da dies jedoch nicht immer der Fall war, mussten weitere, unterschiedliche Quellen herangezogen werden. Zum Beispiel konnten fehlende Daten zum Stromverbrauch in Köln nicht in den statistischen Jahrbüchern gefunden werden. Sie wurden mit den Kennwerten der HEA - Fachgemeinschaft für effiziente Energieanwendung e.V. und BDEW Bundesverband der Energie- und Wasserwirtschaft ergänzt.

Statistische Bundesamt

Aus der GENESIS Datenbank des statistischen Bundesamt wurde die CSV Tabelle „Bevölkerung Deutschland, Fortschreibung des Bevölkerungsstandes“ (173-01-4) abgerufen, welche eine Liste des Bevölkerungsstand von 1990 bis 2013 aufführt <https://www.regionalstatistik.de> [Statistisches Bundesamt (2015)].

Offene Daten:Köln

Die Daten zu den statistischen Jahrbüchern der Stadt Köln werden auf Offene Daten:Köln <http://www.offenedaten-koeln.de/> in den Formaten CSV veröffentlicht und zum Download bereitgestellt. Dort sind die Daten „2012 Fläche Dichte Einwohner Stadtteile“ und „2012 Haushaltsgröße Stadtteil“, die zum 31.12.2012 erhoben wurden besonders interessant. Aus der „2012 Haushaltsgröße Stadtteil“ CSV sind alle Dimensionen relevant, in der anderen CSV jedoch nur die Dimension „Stadtfläche in Quadratkilometer“, „Einwohner je Quadratkilometer“ und „Einwohner insgesamt“ [Stadt Köln (2015)].

Arbeitsgemeinschaft Energiebilanzen e.V.

Auch die AGEB bietet einen Datenservice an, mit dem die Excel-Tabellen heruntergeladen werden können, dies gilt allerdings nur für die Erhebung der Energiebilanz und vierteljährliche Erhebungen des Primärenergieverbrauchs <http://www.ag-energiebilanzen.de/> [DIW u. EEFA (2014)]. Für die Tabelle „Endenergieverbrauch der Privaten Haushalte“ werden die Werte für das Jahr 2012 von „Energieverbrauch in PJoule“ in kWh umgerechnet.

BDEW Bundesverband der Energie- und Wasserwirtschaft e.V.

HEA - Fachgemeinschaft für effiziente Energieanwendung e.V.

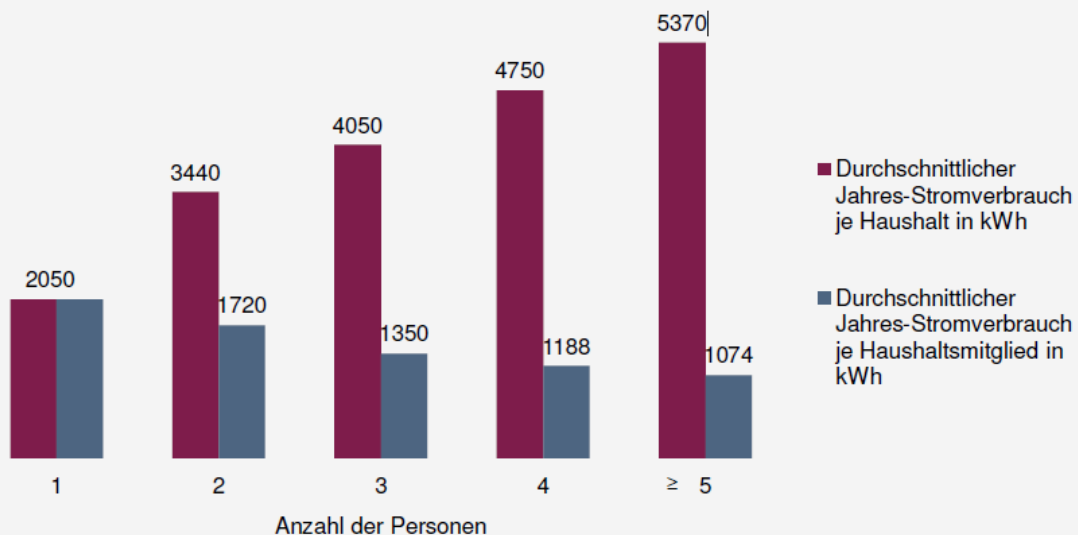
Nützliche Kennzahlen für den Sektor Haushalte erhält man aus der jährlichen Berichterstattung der BDEW und der HEA - Fachgemeinschaft für effiziente Energieanwendung e.V.. Allerdings werden keine Tabellen, sondern die Kennzahlen und Informationen in Berichtsform dargestellt. Aus den Diagrammen der Berichtshefte, werden nur die Durchschnittswerte des Stromverbrauchs in kWh pro Jahr je Haushaltsgrößen übernommen (Abbildung 3.1). Die Daten stammen aus der Erhebungsperiode 2007 - 2011 des NRW.STROMcheck 2012. <https://www.bdew.de/internet.nsf/id/daten-grafik-de> [Oberascher (2013)].

Bedingte Entscheidungen zur Datensatzauswahl

Laut der GENESIS Datenbank gibt es in Köln insgesamt 542261 Haushalte. Diese Daten wurden auf Grundlage des Zensus 2011 mit dem 09.05.2011 als Stichtag ermittelt. Im Vergleich dazu listet offene Daten Köln für Ende des Jahres 2011, insgesamt 515175 Haushalte auf. Es besteht eine Differenz von 27086. Auch für Ende 2010 lässt sich ein Unterschied von ca. 22000 erkennen.

Jahres-Stromverbrauch je Haushalt nach Haushaltsgröße und pro Haushaltsmitglied

(inklusive Warmwasserbereitung, ohne Heizstromverbrauch)



Quellen: forsa 2009, BDEW

Abbildung 3.1: Jahres Stromverbrauch je Haushalt nach HHG und HHM. Aus dem BDEW Beiblatt [Oberascher (2013)].

Aufgrund der quantitativen Unterschiede beider Quellen kann die Korrektheit der Werte an dieser Stelle nicht gewährleistet werden. Zudem führt die Tabelle der GENESIS keine Werte zu privaten Haushaltsgrößen und den von der Kommune gewählten Stadtteilen. Da das Interesse der Stadtteile einer Kommune in einer visuellen Analyse besteht und diese Tabellen in erster Linie die Quellen für den Zensus waren, bezieht sich diese Arbeit auf die Tabellen der Stadt Köln, der Quelle offene Daten.

3.2 Werte (Kennzahlen)

Stromverbrauch pro Haushaltsgrößen berechnen

In der Voranalyse werden einige Fragen aufgeworfen, die nur durch eine neue Wertermittlung mittels Berechnungen beantwortet werden kann, da eine Ermittlung durch Recherche hier nicht möglich ist. Diese Werte werden für die Korrelation bestimmter Datensätze benötigt, die den

Bezug des Energieverbrauchs der Stadtteile in Köln wiedergeben. Mit den Kennwerten “Stromverbrauch pro Jahr nach Haushaltsgrößen” der BDEW / HEA werden die Merkmale der Tabelle Haushaltsgröße erweitert. Dabei wird der Stromverbrauch von Ein- bis Fünfpersonenhaushalten für jeden Stadtteil Kölns berechnet. Die Tabelle Haushaltsgröße beinhaltet die Summen der Haushaltstypen, woraus die Anzahl der Einwohner herzuleiten ist. So wird ein Einpersonenhaushalt von einer Person bewohnt, ein Zweipersonenhaushalt von zwei Personen, usw. Für diese müssen die Werte des pro Kopf Energieverbrauchs herausgerechnet werden und in der Tabelle in eine eigene Spalte eingefügt werden. Wichtig hierbei ist zu erwähnen, dass die Berechnungen mit der Einwohnerzahl pro Haushaltsgröße durchgeführt wird. Mit derselben Berechnung bei Fünfpersonen+ Haushalte, wäre das Ergebnis nicht korrekt, da es hier möglich ist, dass mehr als fünf Personen diesen Haushalt bewohnen. Daher wird dieser ausgehend von der insgesamten Einwohnerzahl des Stadtteils $InsgesamteEinwohneranzahl_i$, von der restlichen Einwohnerzahl der ein bis vier Haushaltsgrößen abgezogen, wobei n für die Anzahl der Stadtteile steht. Die erfolgt nach folgender Berechnung:

$$InsgesamteEinwohneranzahl_i - \sum_{i=0}^n Einwohneranzahl(HHG_i)$$

Die Tabelle 3.1 zeigt am Beispiel des Kölner Stadtteils Neustadt-Süd die Operanden und die Ergebnisse der Berechnungen. Die Einwohneranzahl errechnet sich wie oben beschrieben und werden mit den kWh pro Kopf Werten multipliziert.

Tabelle 3.1: Ergebnisse der Berechnungen des Kölner Stadtteils Neustadt-Süd.

Stromverbrauch nach Haushaltsgröße				
Durchschnitt über 2007 - 2011				
	kWh	kWh pro Kopf	Einwohneranzahl	Neustadt-Süd kWh
1 Person HH	2050	2050.00	17078	35009900
2 Person HH	3440	1720.00	10230	17595600
3 Person HH	4050	1350.00	4983	6727050
4 Person HH	4750	1187.50	3672	4360500
5+ Person HH	5370	1074.00	2453	2634522
Durchschnitt	3932.00		7683.2	13265514.4

3.3 Eigenschaften und Spezifikationen

Die Spezifizierung der Ausprägungen werden hier festgelegt und damit die Konventionen zur Darstellung von Daten eingegrenzt. Die Konventionen sollen gewährleisten, dass Daten korrekt, effizient und nicht missinterpretiert werden.

Datensatz Spezifikation, von Merkmalsraum

Die Datensätze der Energiebilanz und der Bevölkerungsdaten sind Skalar (Tensor nullter Ordnung) [Lehmann u. a. (2010)]:

$$x, y \rightarrow f(x, y) : R^2 \rightarrow R$$

Skaleneigenschaften bzw. Ausprägungen

Die Daten weisen folgende Eigenschaften auf und werden entsprechenden Skaleneigenschaften zugewiesen. Zur Erinnerung werden die Eigenschaften wie folgt beschrieben: Ordinal (geordnete, feste Reihen), Nominal (Gruppen bildend), Quantitativ (Messgrößen)

Entsprechend der Definition der Eigenschaften Ordinal, Nominal und Quantitativ werden die Skaleneigenschaften den Merkmalen zugewiesen. Die Skaleneigenschaften der privaten Haushalte sollten Nominal sein, sie können zusätzlich aber auch Ordinal sein, wodurch die Expressivität der Visualisierung verbessert wird. Die Einwohneranzahl sollte Quantitativ sein. Die Bevölkerungsdaten können kategorisiert und vom Nutzer ausgewählt werden. Die Skaleneigenschaft der Energieträger müssen einzeln unterschieden werden, deshalb empfiehlt sich hier ebenso eine nominale Skala, die zusätzlich quantitative Eigenschaften besitzt, um damit Messwerte zu vergleichen, diese sind die Werte Endenergieverbrauch (EEV) aus der Tabelle „Endenergieverbrauch der Privaten Haushalte“. Alle Daten unterliegen der Intervall Dimension Jahre, als ordinale Achse. Die Tabelle 3.2 zeigt einen Überblick der Zuordnungen.

Die Variablen (Einwohneranzahl, Verbrauch, Wohndichte, THG-Emissionen) unter den Dimensionen für die regionale Visualisierung werden nominal zugeordnet. Was die Tabelle nicht deutlich macht. Normalerweise wäre dies ein Konventionsbruch, denn skalare Werte sollten nicht nominal dargestellt werden. Deshalb werden diese Daten zu der des Energieverbrauchs in privaten Haushalten als nominale Skala in Relation gesetzt. Dies geschieht bei der Visualisierungstechnik GIS und gilt daher nicht als Konventionsbruch, da die Daten in der Visualisierung Geo Spatial (räumlich) dargestellt werden. Deshalb ist diese Art der Darstellung eine adäquate Lösung und anderen Techniken vorzuziehen. Sie fördert das Verständnis und senkt den kognitiven Aufwand und entspricht damit den Kriterien der Effektivität und Expressivität die menschliche Wahrnehmung betreffend.

Tabelle 3.2: Zugewiesene Skaleneigenschaften der Dimensionen und Variablen.

Messniveau	Rechenoperation	Dimension	Variable
Qualitativ:			
Nominal	=,!=	Stadtteil, Energieträger	
Ordinal	<,>	Haushaltsgrößen	
Quantitativ:			
Intervall	+, -	Jahr	
Ratio / Verhältnis	*, /		
Absolut (Ordinal)	Maßeinheit		Einwohneranzahl, Verbrauch, Wohndichte, THG-Emissionen

3.4 Fragestellungen zu den Datenwerten

Für eine erste Analyse und Wissenserwerb über die Daten, wird versucht Antworten zu folgenden Fragen und Hypothesen zu erlangen. Erste Methoden werden hier visuell dargestellt und erleichtern die Interpretation und Kommunikation des erworbenen Wissens. Die relevanten und nützlichen Methoden, die während der ersten Analyse entstanden sind, können so in der Entwicklung adaptiert werden.

Datensatz Fragestellungen

1. Wieviele Einwohner hat Köln?
2. Wieviele Einwohner haben die einzelnen Stadtteile?
3. Wie sind die Anteile der Haushaltsgrößen in Köln und in den Stadtteilen verteilt?
4. Welcher Stadtteil ist der Größte, welcher der Kleinste?
5. Haben Stadtteile mit einer größeren Fläche eine niedrigere Wohndichte?
6. Haben diese Stadtteile, dann einen höheren Mehrpersonenhaushalt-Anteil?
7. Haben Stadtteile mit einer kleinen Fläche eine höhere Wohndichte?
8. Haben diese Stadtteile, dann einen höheren Einpersonenhaushalt-Anteil?

DETAIL LEVEL 1

9. Welche Stadtteile verbrauchen weniger und welche mehr Strom?

10. Haben Stadtteile mit einer kleinen Fläche einen höheren Energiebedarf?
11. Haben Stadtteile mit einer großen Fläche einen niedrigeren Energiebedarf?
12. Ist der Pro-Kopf-Verbrauch in Stadtteilen mit einer höheren Anzahl von Mehrpersonenhaushalten höher?
13. Welcher Stadtteil verbraucht im Durchschnitt weniger Energie und welcher mehr?
14. Wieviele Einpersonenhaushalte und Mehrpersonenhaushalte muss ein Stadtteil haben, damit alle Haushaltsgrößen in etwa denselben Verbrauch haben?
15. Ab welcher Haushaltsgröße von Mehrpersonenhaushalten, ist der Verbrauch größer als bei Einpersonenhaushalten?

ERWEITERT DETAIL LEVEL 2

16. Hängt die Flächengröße der Stadtteile, mit der Verteilung und dem Verbrauch der privaten Haushalte zusammen?
17. Im Vergleich der Ergebnisse zum Vorjahr, ist der Verbrauch größer?
18. Wieviel Energie wird nach Energieträger für Köln, je Stadtteil, je Haushalt und pro Kopf verbraucht?
19. Wie hoch sind die CO₂-Äquivalente für Köln, je Stadtteil, je Haushalt und pro Kopf?

Hypothesen

1. Mehrpersonenhaushalte verbrauchen weniger Energie pro Person, als private Haushalte mit ein oder zwei Personen.
2. Stadtteile mit einem höheren Anteil von Ein- bis Zweipersonenhaushalten, verbrauchen mehr Energie im Vergleich zu anderen Stadtteilen mit Mehrpersonenhaushalten.
3. Stadtteile mit einem, im Schnitt hohen Anteil von Ein- und Zweipersonenhaushalten, verbrauchen mehr Strom, als Stadtteile mit einem höheren Mehrpersonenhaushalte-Anteil.
4. Damit die Stadtteile in Köln, ihren Stromverbrauch einsparen, müssen die Einwohner in Mehrpersonenhaushalten wohnen.
5. Stadtteile mit weniger Einpersonenhaushalten, verbrauchen weniger Energie

6. Stadtteile mit weniger Einpersonenhaushalten, verbrauchen mehr Energie
7. Stadtteile mit mehr Einpersonenhaushalten, verbrauchen weniger Energie
8. Stadtteile mit mehr Einpersonenhaushalten, verbrauchen mehr Energie
9. Stadtteile mit weniger Mehrpersonenhaushalten, verbrauchen weniger Energie
10. Stadtteile mit weniger Mehrpersonenhaushalten, verbrauchen mehr Energie
11. Stadtteile mit mehr Mehrpersonenhaushalten, verbrauchen weniger Energie
12. Stadtteile mit mehr Mehrpersonenhaushalten, verbrauchen mehr Energie
13. Stadtteile mit einer hohen Bevölkerungsdichte pro qm, haben einen höheren Energieverbrauch, als Stadtteile mit einer geringeren Bevölkerungsdichte.

3.5 Erste Datensatzanalyse

3.5.1 Vorgehen und Testdatenauswahl

Die Tabelle der Haushaltsgrößen 2012 in Köln beinhaltet 86 Zeilen / Stadtteile und zählt in den Spalten die Anzahl der Haushalte insgesamt und die Größe der Personenhaushalte bis fünf. Für die ersten Analysen wurde eine Stichprobe mit 15 Stadtteilen ausgewählt und zwar die Stadtteile mit den signifikantesten Merkmalen. So ist Chorweiler stellvertretend für die Fünfpersonenhaushalte und die Altstadt-Nord stellvertretend für Einpersonenhaushalte. Außerdem wurden folgende Stadtteile ausgewählt: Neustadt-Süd, Mülheim, Neustadt-Nord, Zollstock, Altstadt-Nord, Deutz, Rodenkirchen, Chorweiler, Bayenthal, Rondorf, Raderberg, Marienburg, Raderthal, Hahnwald, Ehrenfeld.

Während die meisten Stadtteile in Köln eine hohe Anzahl von Einpersonen- und Zweipersonenhaushalte zählt, wurde bei der Auswahl versucht, eine ausgewogene Verteilung der Haushalte innerhalb der Stadtteile zu wählen, um die Verteilung der Haushaltsgrößen und den jeweiligen Energieverbrauch später vergleichen zu können.

3.5.2 Naive Analyse mit primitiven Visualisierungstechniken

Während der Analyse stellten sich die ersten Problematiken heraus. Um die Frage zu beantworten inwieweit der Stromverbrauch in Zusammenhang mit den Haushaltsgrößen korreliert,

werden in der Analyse die Haushaltsgrößen und deren Energieverbrauch gegenübergestellt. Mit traditionellen Balkendiagrammen lässt sich diese Frage nicht mit nur einer Visualisierung bearbeiten. Hierfür muss jedem Balkendiagramm ein Hilfsdiagramm mit den Verbrauchswerten zugeordnet werden. Bei primitiven Visualisierungstechniken sollte deshalb immer eine Analyse mit zwei parallelen Diagrammen durchgeführt werden. Für diese Problematik eignet sich die Streudiagramm-Visualisierung mehr, da diese Technik zwei Dimensionen abbilden kann.

Zusätzlich kommt in der Analyse und Detailabfrage die Problematik des Zusammenhangs, sowie die Unübersichtlichkeit hinzu. Besonders wenn pro Stadtteil die Frage gestellt wird, wie hoch der Verbrauch für die Energieträger pro Haushalt ist. Es ist eine Anforderung/Herausforderung, dass der Verbrauch im Zusammenhang mit den Haushaltsgrößen übersichtlich bleibt.

3.5.3 Begründung der Diagrammauswahl

Zur Visualisierung der Haushaltsgrößen-Anteile der Stadtteile wird ein Balkendiagramm genutzt (Abbildung 3.4), welches die Gesamtbreite der x-Achse nutzt. Die x-Achse bildet prozentual die Größe der Haushalte wieder, während die y-Achse die jeweiligen Stadtteile abbildet. So kann man visuell ein Gefühl dafür entwickeln, zu welchen Anteilen die jeweiligen Haushaltsgrößen zu den anderen Stadtteilen in Beziehung stehen, unabhängig von der tatsächlichen Stadtteilgröße, welche für die Hypothese nicht relevant ist. Auf diese Weise wird sich nur auf die Information konzentriert. Gleiches gilt für die Visualisierung, die parallel zur Haushaltsgröße erzeugt wird (Abbildung 3.3). Dieses Diagramm stellt den Verbrauch der jeweiligen Stadtteile dar und ist farblich mit den Haushaltsgrößen abgestimmt, so dass der direkte Vergleich gewogen werden kann.

Es werden der Durchschnittsverbrauch aller Stadtteile berechnet und die Gruppen verglichen. Ein Kreisdiagramm wie in Abbildung 3.4 ist eine effektivere Darstellung von eindimensionalen Datensätzen.

Das Balkendiagramm führt zur Fehleinschätzung des Durchschnittsverbrauch der zusammen genommenen Teststadtteile. Die Zweipersonenhaushalte und die Einpersonenhaushalte können im Vergleich überschätzt werden und so falsche Interpretationen hervorrufen.

3.6 Erkenntnisse aus der Analyse und Bedeutung

Die Haushaltdatensätze je Stadtteil, sind von der räumlichen Größe abhängig. So wird vermutet, dass der Energieverbrauch höher ist, je größer der Stadtteil, da es in größeren Stadtteilen mehr Personen gibt. In diesen Fällen ist die Bevölkerungsdichte zu berücksichtigen. Deswegen

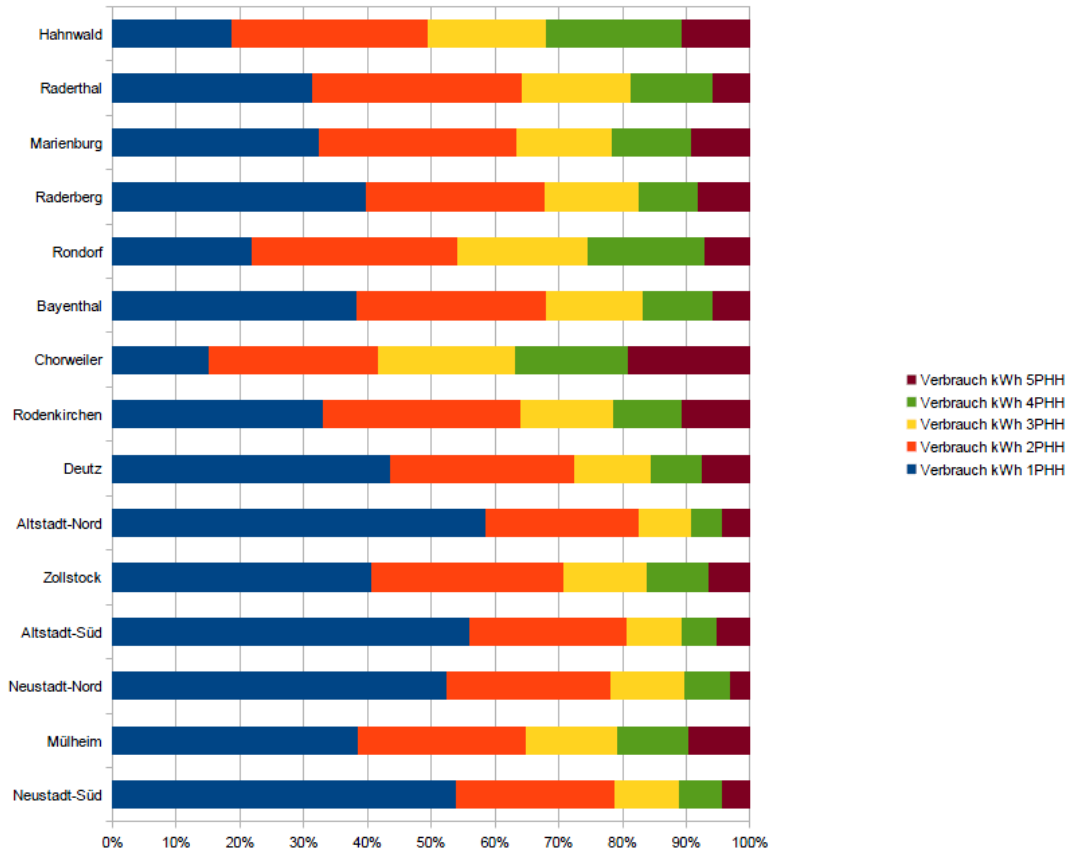


Abbildung 3.2: Verbrauch der jeweiligen Stadtteile. Erstellt von Dorothea Buck.

wird der Mittelwert der Personenhaushalte je Stadtteil berechnet und mit dem Mittelwert des Energieverbrauchs pro Kopf multipliziert. Daraus lässt sich herleiten, dass die Stadtteile mit den höchsten Anteilen von mehr als Dreipersonenhaushalten einen höheren Energieverbrauch aufweisen. Obwohl die Anzahl von Ein- und Zweipersonenhaushalte höher ist und sie pro Kopf mehr Energie verbrauchen als Fünfpersonenhaushalte. Ein weiterer wichtiger, zu berücksichtigender Aspekt ist, dass private Mehrpersonenhaushalte die Gemeinschaftskosten unter sich aufteilen und die gemeinsam genutzten Geräte wie TV, Kühlschrank und Licht dazu führen, dass der Pro-Kopf-Energieverbrauch sinkt, je größer der Haushalt ist [Oberascher (2013)].

Interessant ist daher das Beispiel Chorweiler, da dieser Stadtteil mehr Einwohner ab Dreipersonenhaushalte, als Ein- und Zweipersonenhaushalte hat und der Stromverbrauch höher ist, obwohl der Energieverbrauch pro Kopf geringer ist. Festgestellt wurde auch, dass alle weiteren Stadtteile, in denen die Einwohnerzahl von Ein- und Zweipersonenhaushalten höher ist, auch der Stromverbrauch signifikant höher ist.

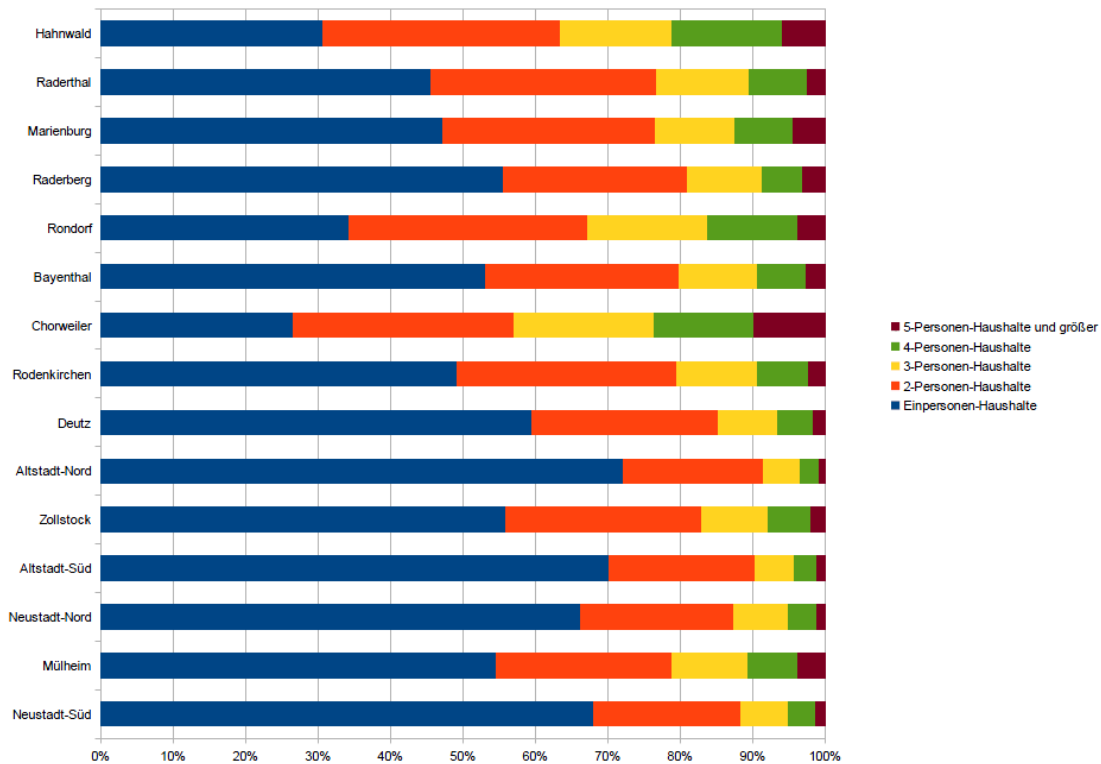


Abbildung 3.3: Haushaltsgrößen-Anteile der Stadtteile. Erstellt von Dorothea Buck.

Eindeutig und klar ist, dass der Verbrauch nach Stadtgröße korreliert. Je mehr Einwohner der Stadtteil hat, desto größer ist der Verbrauch. Dies bestätigt das Diagramm „Verbrauch nach Einwohner Gwh“. Im Diagramm „Verbrauch nach durchschnitts Hhgröße kWh“, haben die verhältnismäßig kleinen Stadtteile Hahnwald und Chorweiler im Vergleich zu den beiden größten Stadtteilen Neustadt-Süd und Mülheim, im Durchschnitt der jeweiligen Haushaltsgrößen einen höheren Verbrauch. Es ist zu beobachten, dass obwohl nach Einpersonenhaushalte und Mehrpersonen der Verbrauch pro Kopf absteigend ist, der Durchschnittsenergieverbrauch, aufgrund der Mehrpersonenhaushalte größer sein kann, als bei Stadtteilen mit mehr Einpersonenhaushalten. Daraus kann man folgende Hypothesen zum Datensatz stellen.

Hypothese: Stadtteile mit mehr Einpersonenhaushalte, verbrauchen gegenüber Stadtteilen mit höheren Mehrpersonenhaushalte Anteil, insgesamt weniger Energie.

Nullhypothese: Stadtteile mit einem höheren Anteil von Mehrpersonenhaushalten, verbrauchen weniger Energie, als Stadtteile mit einem höheren Einpersonenhaushalt-Anteil.

Hypothese: Der Verbrauch von Ein- und Zweipersonenhaushalte zusammen genommen, ist im Vergleich nicht viel höher, als der Gesamtverbrauch aller übrigen Haushaltsgrößen. (Obwohl der

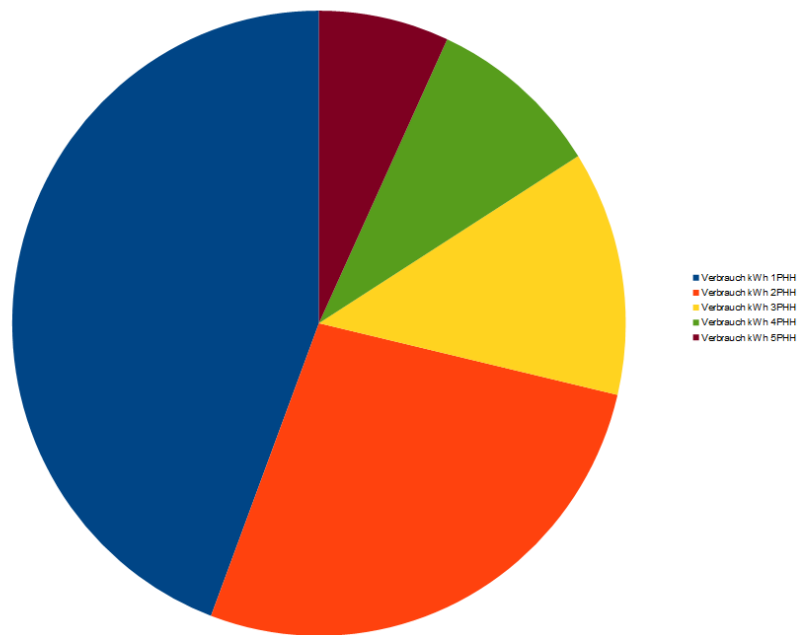


Abbildung 3.4: Kreisdiagramm für Durchschnittsverbrauch aller Stadtteile. Erstellt von Dorothea Buck.

Anteil der Ein- und Zweipersonenhaushalte gegenüber der ab Dreipersonenhaushalte deutlich über die 50% steigen.)

4 Konzept zur Entwicklung der InfoVis

Aus den Rückschlüssen der Datensatzanalyse und den Fragestellungen der Unterthemen Haushaltsgrößen, Wohndichte, Energieverbrauch, THG-Emissionen begrenzt auf die Region Köln, wird hier von einem thematischen Rauminformationssystem (RIS) gesprochen. Die thematische RIS ist mehrdimensional, da heterogene Primärdaten aus verschiedenen Domänen herangezogen werden. Die RIS-Visualisierung wird neben einer Choroplethkarte durch primitive Visualisierungstechniken unterstützt, dabei wird die InfoVis durch eine direkte numerische Wertausgabe erweitert. An dieser Stelle wird bewusst auf tabellarische Darstellungen verzichtet, da die Motivation dieser Arbeit im Verzicht dieser Art der Datenpräsentation besteht.

Nach dem Vorbild der GeoDa und GeoVISTA wird die thematische RIS mit getrennten Ansichten erweitert. So soll die Choroplethkarte nur die aktuell gewählte Thematik, im weiteren Modifikation genannt, visualisieren. Zusatzinformationen werden in eigenen Panels visualisiert. In den folgenden Abschnitten wird genauer auf die Anforderungen der InfoVis eingegangen und im letzten Abschnitt das Mockup/Layout skizziert.

4.1 D3 für RIS-Visualisierungen

Realisiert wird die Visualisierung, mit der Javascript Bibliothek D3 (Data-Driven Documents) von Mike Bostock [Bostock (2016)]. Damit ist die Visualisierung Betriebssystemunabhängig und kann überall angewendet werden. Die Technologie erfordert keine hohe Rechenleistung beim Rendern der Projektionen und ist schnell darin, eine Ausgabe des SVG-Datenabbildes zu erzeugen und ermöglicht damit eine performante Animation oder eine direkte Manipulation der Darstellung durch den Nutzer, gegebenenfalls auch mit hohen und mehrdimensionalen Datenmengen.

Die Bibliothek ist für die Entwicklung einer räumliche Informationsvisualisierung sehr gut geeignet. Sie bietet eine mit wenigen Schritten einfache Implementierung der geografischen Karte und ist mit D3 beliebig veränderbar.

Für einen verbesserten Nutzen und mehr Optionen bei den geografischen Daten, wird empfohlen die Werkzeuge Geospatial Data Abstraction Library (GDAL) zu installieren [Bostock (2016)].

Die Umwandlung geht auch ohne GDAL, wenn die verwendete SVG-Karte für jede Komponente, wie zum Beispiel Stadtteile einen Path / Polygon / Polyline definiert. Durch die Mächtigkeit des D3 DOM Quering mit der Methode `d3.select()`, lassen sich die Elemente ansprechen und Attribute hinzufügen. Fehlende Metadaten können mit der SVG-Karte verknüpft werden, so wie es in diesem Projekt gehandhabt wird.

4.2 Anforderungsanalyse

Die Anforderungsanalyse erfolgt nutzerzentriert und ist nach der Datenanalyse dieses erweiterten Referenzmodells (Abbildung 2.6) der zweite Schritt. Die Erarbeitung der Anforderungen, wird in drei Aspekte unterteilt, im ersten Aspekt geht es um die Benutzeranforderungen, sie werden in Nutzeraufgaben und Bearbeitungsziele aufgeteilt. Die Nutzeraufgaben definieren die Interaktionen, die der Nutzer benötigt um das Ziel, die Lösung der Aufgabe zu erreichen. Der zweite Aspekt sind die Anforderungen an das System / Visualisierung. Damit der Nutzer seine Nutzeraufgaben ausführen kann, wird ein Satz Werkzeuge in der Visualisierung implementiert, die ihm dabei helfen, die Bearbeitungsziele zu erreichen. Im dritten Aspekt, wird der im Vorfeld analysierte Datenraum, mit den Anforderungen der ersten beiden Aspekte verglichen. Hier wird ergänzend geklärt, welche Anforderungen an dem Datenraum gestellt werden, um die Benutzeranforderungen und die Visualisierungsanforderungen, auf dieser Grundlage umsetzen zu können.

4.2.1 Benutzeranforderungen

An dieser Stelle wird gefragt, welches Ziel und welche Erwartungen der Nutzer an die Analyse hat und wie er vorgeht, um sein Ziel zu erreichen. Das Ziel und die Erwartungen, die der Nutzer gegenüber den Daten hat, wurden schon im vorigen Kapitel 3 ausführlich in einer Liste von 3.4 Datensatz Fragestellungen formuliert. Daraus lässt sich die Zielsetzung ableiten und eine mögliche Erweiterung für die Visualisierung abwägen. Hier werden die Aufgabenstellungen, die der Nutzer an die Daten stellt und was der Nutzer während der Bearbeitung erlangen soll, beschrieben.

Bearbeitungsziele

Die Bearbeitungsziele wurden aus den Problemen der Datensatzanalyse hergeleitet. Die daraus Resultierenden Probleme erschweren eine korrekte Analyse/ Bearbeitung der Informationsvisualisierung für den Nutzer. Einige können den allgemeinen Problemklassen von [Schumann

u. Müller (2000)] zugeordnet werden: Identifikation, Unterscheidung, Lokalisierung, Einteilung, Bündelung, Anordnung, Vergleich, Assoziation und Korrelation.

Ergänzend zu den allgemeinen Problemklassen aus [Schumann u. Müller (2000)] wurden folgende allgemeine Ziele im Praxisprojekt [Buck (2014)] ermittelt: Überblick, Vollständigkeit, Zusammenhang, Überprüfbarkeit von Hypothesen, Analyse, Erlernbarkeit, keine Vorerfahrung nötig, Nachvollziehbarkeit, Verständlichkeit, Korrektheit und Vereinfachung. Dem Grad der Qualität entspricht, inwieweit die Anforderungen und Bearbeitungsziele gelöst wurden. Daneben ist auch die Korrektheit, die Expressivität, die Effektivität und die Angemessenheit der Visualisierung ein Qualitätsmerkmal.

Nutzeraufgaben

Da die Bearbeitungsziele oben festgelegt wurden, müssen dazu die entsprechenden Operatoren gefunden werden, womit die Ziele zufriedenstellend zu erreichen sind. Die Bearbeitungsziele können mit dem Visual Information Seeking Mantra [Shneiderman (1996)] verglichen und zugeordnet werden. Die Aufgaben überschneiden sich teilweise und sind nicht als feste Reihenfolge zu verstehen.

Das Mantra beschreibt sieben Aufgaben mit hohem Abstraktionsgrad, diese sind:

- 1. Überblick:** *Lokalisierung, Vollständigkeit, Unterscheidung*
- 2. Zoom:** *Identifikation, Vergleich*
- 3. Filter:** *Vereinfachung, Bewertung, Einteilung*
- 4. Details auf Anfrage:** *Analyse, Bündelung, Bewertung, Einteilung*
- 5. Beziehung:** *Korrelation, Zusammenhang, Anordnung, Bündelung*
- 6. Historie:** *Nachvollziehbarkeit, Korrektheit*
- 7. Extrahieren:** *Assoziation, Überprüfbarkeit von Hypothesen, Kommunikation*

Für die drei Bearbeitungsziele Korrektheit, Verständlichkeit und Erlernbarkeit sind keine Vorerfahrung nötig. Die Überprüfbarkeit von Hypothesen, ist ein allgemeines Ziel und wird daher während des ganzen Bearbeitungsprozesses angestrebt. Dies gilt auch für die Analyse, hier wird das Ziel als eine spezifischere Aufgabe verstanden und angewendet. Damit wird verdeutlicht, ab welchem Zeitpunkt die Daten eine höhere Relevanz für die Wissensgewinnung erhalten.

4.2.2 Visualisierungsanforderungen

Mit den oben angesprochenen Nutzeraufgaben sind die Aufgaben gemeint, die nach dem Rendern auf der Präsentationsebene als Interaktion von Benutzer durchgeführt werden. Es ist schwierig Interaktionen und Aufgaben zu unterscheiden und bedarf genauer Differenzierung. An dieser Stelle wird die Frage nach dem wesentlichen Unterschied zwischen Interaktion und Visualisierungsaufgaben, die vom Nutzer direkt oder indirekt getriggert werden, gestellt. Die Interaktionen werden benötigt, um dem Nutzer bestimmte Aufgaben zu ermöglichen, die Datentyp spezifische Methoden zur Analyse und Navigation erfordert. Die Aufgaben wiederum, werden als Vorgehensweisen verstanden, die zu einem angestrebten Ergebnis führen. Um die nötigen Aufgaben entwickeln zu können, sollten die Nutzerziele definiert werden, die in Form der Fragenstellungen schon vordefiniert wurden.

Zu den Visualisierungsaufgaben gehören Interaktionen oder auch automatische Operationen, die die Projektion verändert. Mit diesen soll es möglich sein, in der Projektion navigieren und die Daten aus einer anderen Ansicht analysieren zu können. Auch können weitere Details entdeckt werden, die sonst verborgen blieben. Einige Interaktionstechniken haben sich in der Informationsvisualisierung als sehr nützlich erwiesen und werden deshalb in der Gestaltung berücksichtigt.

Linking and brushing Linking and brushing sind wie in den GIS Anwendungen eine der wichtigsten Interaktionstechniken, die die verschiedenen ausgelagerten Visualisierungstechniken mit der Karte verbindet (linking) und hervorhebt (brushing). Verbunden mit Details-on-demand, bei welchem die Auswahl der Stadtteile im Fokus stehen, werden Detailinformationen abgerufen.

Fokus + Context Die Grundidee mit Fokus-plus-Kontext-Visualisierungen ist es dem Nutzer zu ermöglichen, das Data-of-Interest (DOI) mit allen Details zu betrachten, während zur gleichen Zeit der Nutzer ein Überblick von allen umliegenden Informationen im Kontext zu sehen bekommt.

Rückverfolgung der Interaktionen / Interaktionen Rückgängig machen / Speicherung von Ergebnisse Obwohl laut Shneiderman eine Historie / Ermittlung der Nutzerinteraktionen schwierig sein kann, sollen ausgeführte Aufgaben, bis auf eine sinnvolle, machbare Anzahl verfolgbar und widerrufbar sein. Dies soll dem Nutzer ermöglichen bei Fehleinschätzungen wieder zurück zum Ausgangspunkt der Entscheidung zu gelangen oder eine Korrektur im begrenzten Rahmen der Visualisierung durchzuführen. Es sollte daher auch möglich sein Ergebnisse und Zwischenergebnisse festzuhalten / speichern, um diese bei einer späteren Analyse zu verwenden. Die Schritte die zu dem Ergebnis führen werden mit aufgezeichnet (Historie). Was zu einer Chance zur Evaluation der Ergebnisse führt. Ein solcher Umgang mit der Visualisierung führt dazu, das der Nutzer sich experimentierfreudig mit den Daten verhält. Durch die höhere

Datenmenge entsteht eine genauere Datenanalyse. Somit sind die Ergebnisse vertrauenswürdig und weniger anfällig für Fehlinterpretationen/Fehldeutungen [Shneiderman (1996)].

4.2.3 Anforderungen an den Datenraum

Der Nutzer als Nicht-Experte, soll in der Lage sein, mit einer interaktiven Informationsvisualisierung, ein mentales Modell zu bilden. Dabei erlangt er anwachsendes Wissen und Verständnis über die Datenstrukturen, sowie deren Zusammenhänge. An diesem Punkt entstehen durch Wissenslücken im mentalen Modell, neue Fragen die der Nutzer versucht zu beantworten. Ähnlich wie nach dem Visual-Seeking-Mantra führt der Nutzer in der Visualisierung wiederholt Operationen zum Filtern, Suchen, Ordnen / Sortieren, Vergleichen, Bewerten und Auswählen aus. Dies erfolgt iterativ, bis er alle relevanten Fragen und das Ziel seiner Ursprungsfrage beantwortet hat.

Ziel und Zielsetzung (Fokus)

Was ist mit Ziel und Zielsetzung gemeint? Unter Ziel versteht man das finale Ergebnis und unter Zielsetzung das leitende Ziel während der Analyse. So können die Visualisierungsaufgaben auch Zielen zugeordnet werden, die ein Nutzer für den Prozess der Datenuntersuchung braucht, um diese danach als Klassifizierungsrichtlinie zu verwenden. Damit werden die Aufgaben, passend für den Zeitraum der Nutzung, in die Gestaltung der Visualisierung eingebaut. Sind die Aufgaben zur Zielsetzung bekannt, kann die Darstellung der Daten in der InfoVis auch entsprechend der Aufgabe gestaltet werden.

Zur Zielsetzung mit den Fokus auf die Daten gehören Prozesse die den Überblick verschaffen, um die Bedeutung der Daten zu verstehen, indem Daten gelesen werden. Das ist das Stichwort für die Analyse, was der ganze Prozess mit dem Ziel ist, in Information / Assoziationen den Zusammenhang zu sehen und zu verstehen, um übergreifende Eigenschaften in den Ergebnissen zu erkennen.

Mindestanforderungen / Ziele

In diesem Abschnitt wird die Frage an den Datenraum und die Visualisierung gestellt: Welche Daten sollten mindestens dargestellt werden? Für die Implementierung wird daher eine Priorisierung angelegt, welche die relevantesten Datensätze auflistet. Diese Daten sind notwendig um die Kernfragen und Sekundärdaten zu errechnen. Die Priorisierung dient für die Entwicklung

der wichtigsten Funktionen nach der Priorisierung, sowie die minimalste Implementation der Visualisierung.

Prio 1:

Die Haushaltsgrößen mit der Anzahl der Bewohner, lässt sich auch einfach errechnen.

Die Kennzahlen des Stromverbrauchs pro Einwohner, in den jew. Haushaltsgrößen.

Prio 2:

Wohndichte für weiterführende Rückschlüsse.

Flächengröße der Stadtteile.

Prio 3:

CO₂-Äquivalente

Energieträger

Datensätze nach Jahren

4.3 Layoutgestaltung

In diesem Abschnitt werden die Anforderungen in konkreten Visualisierungsaufgaben und Interaktionen zusammengefasst. Dies entspricht in der Visualisierungsgestaltung dem dritten Schritt „Entwickeln Sie alternative Entwürfe, die die Anforderungen erfüllen“. Die Anforderungen werden für die InfoVis in Ansichten / Elemente, Datenvisualisierungen, Detailansichten, Interaktionen und Layout konzipiert. Die Punkte Features und Verworfen sind zusätzliche Ideen, die ausgewertet wurden und als nicht relevant erachtet werden. Hier werden die Beschreibungen zu Ansichten / Elemente, Detailansicht, und Layout aufgeführt. Alle anderen befinden sich im Anhang A.

Ansichten / Elemente:

1. Stadtkarte mit allen Stadtteilen, Bezirksnamen und Stadtteilnamen
2. Panel mit allgemeinen Kennzahlen von ganz Köln und aus der Auswahl errechnete Werte. Einwohneranzahl, Wohndichte, Fläche, Stromverbrauch und durchschnittlicher Stromverbrauch / Personen-Haushaltsgröße.

3. Detail-Panel für ausgewählte Stadtteile mit Diagramme über Anteile der Haushaltsgrößen, Stromverbrauch, CO₂-Äquivalente. Entsprechend den gewählten Modifikationen sollen für die Kategorie Haushalte, mit gruppiertem Balkendiagramm der direkte Vergleich je Haushaltsgröße möglich sein. Die Modifikation für den Stromverbrauch je Haushaltsgröße soll nicht mit der Skala angezeigt werden, da diese im Detail-Panel schon je Haushaltsgröße sortiert werden. Es sollen lediglich die Modifikationen für Einwohneranzahl, Wohndichte (km²), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamtstromverbrauch (kWh) in der Detailansicht zum Vergleich erzeugt werden. Der Stromverbrauch je Haushaltsgröße soll für die Modifikation, nicht den Haushaltsgrößen gegenübergestellt werden, sondern den ausgewählten Stadtteilen. In diesem Zusammenhang wäre auch interessant, wieviele Einwohner pro Stadtteil oder HHG vorhanden sind, da die Grafik nur den Energieverbrauch zeigt.
4. Checkboxes / Radiobuttons für Layer, auf der Stadtkarte wird für jeden Stadtteil ein bestimmter Farbcode eingeblendet, für den Stromverbrauch aller und je Haushaltsgrößen, Stromverbrauch pro Kopf, Einwohneranzahl, Wohndichte, oder eine Grafik für die Haushaltsgrößenverteilung. Es werden auf dem Layer noch keine Werte angezeigt, dieser dient nur zur Orientierung und dem Messgrößenüberblick.
5. Eine Legende zu den Farbcodes, der jeweils ausgewählten Layer-Daten.

Detailansicht

- In der Detailansicht sollen im oberen Bereich die Kennzahlen und Informationen der ausgewählten Stadtteile für die Zusammenfassung und Berechnung aufgeführt werden.
- Ein Säulendiagramm kann den direkten Vergleich pro Haushalt ein Stadtteil aufnehmen und ist damit in der Lage nach Auswahl der Karte, die Stadtteile nach Haushalt zu gruppieren, um nebenstehende Säulen in der Gruppe Haushalt direkt zu vergleichen. Das Diagramm erweitert sich nach rechts, je mehr Stadtteile ausgewählt werden.
- Das Diagramm nutzt die ganze Breite des rechten Fensterteils. Falls die Breite nicht ausreicht, werden Scrollbalken eingeblendet.
- Die Säulen sollten dynamisch transient sein und bei Größenunterschieden entsprechend skalieren. Bei Updates sollten die Säulen dynamisch skalieren.
- Die Informationen, die auf der Karte mit Farbwerten abgebildet werden, können hier nochmal neu für die Gruppenauswahl verwendet werden. Dabei ist die Detailansicht gesondert und der Fokus wird speziell auf die Auswahl gesetzt.

- Ein Problem besteht darin, wie alle Stadtteile in der Gruppierung unterschieden und identifiziert werden können. Durch Farben, Namen oder Zahlen, bei Interaktion z.B. Maus-hover?

Layout

Das Browserfenster wird in einen linken und rechten Bereich geteilt. Die Ansichten werden dort wie folgt positioniert. Die Skizze Abbildung 4.1 visualisiert das beschriebene Layout.

- Die Stadtkarte befindet sich im linken Teil des Fensters und sollte dieses ausfüllen.
- Checkboxes für Layer befinden sich über dem Panel mit den allgemeinen Kennzahlen, da bei der Auswahl der Modifikationen unterhalb der allgemeinen Kennzahlen, die Daten zu den Modifikationen angezeigt werden
- Unterhalb der Checkbox sollte sich die Legende zu den Modifikationen zugehörigen Farb-codes positionieren, damit bei der Modifikationsauswahl auf den ersten Blick auf den Interaktionsbereich, die Anpassung der Legende wahrzunehmen ist.
- Das Panel mit den allgemeinen Kennzahlen, befindet sich über der Detailansicht.
- Das Detail-Panel soll sich im rechten Browserfensterteil unter dem Panel mit den allgemeinen Kennzahlen befinden.

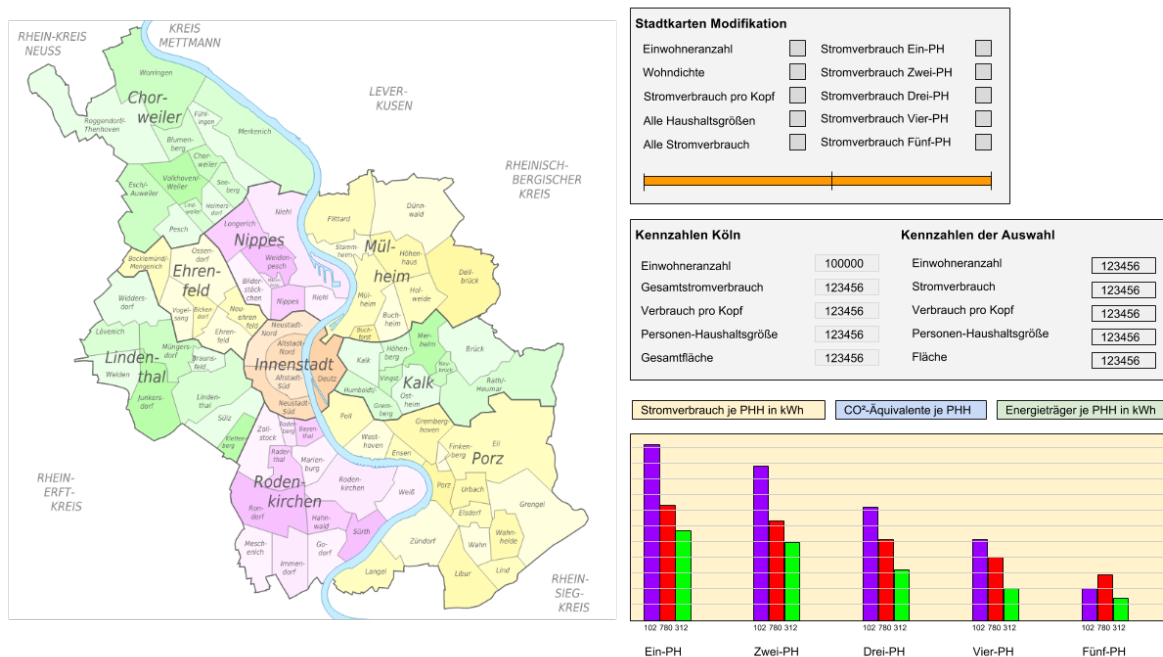


Abbildung 4.1: Skizze des Layout. Erstellt von Dorothea Buck.

5 Prototypenumsetzung mit D3

Schritt drei zur Gestaltung einer Informationsvisualisierung „Implementieren Sie die Entwürfe in einen interaktiven Prototypen.“ wird in diesem Kapitel dokumentiert. Im erweiterten Referenzmodell aus Kapitel 2 (Abbildung 2.6), werden die Schritte zur Strukturierung von Visualisierungsanwendungen, mit den fokussierten Daten aus der Filterung mit Mapping und anschließendem Rendering ausgeführt. Darunter fallen die Unteraufgaben Layout gestalten, Nutzerinteraktionen definieren und Bugfixes, die im Schema nicht beschrieben werden, jedoch im Konzept definiert wurden. Dieses Kapitel dokumentiert Vorüberlegungen zur Entwicklungsumgebung, beschreibt wesentliche Code-Snippets, die Struktur der Anwendung und Gestaltungsentscheidungen

5.1 Vorbereitung

Bevor mit der Programmierung gestartet werden kann, mussten einige Entscheidungen zur Entwicklungsumgebung getroffen werden. Die Frage, ob mit einem Build-System oder eine einfach implementierte HTML / Javascript Lösung eingesetzt werden sollte, musste beantwortet werden. Je nach Anzahl der Node-Module, mit hoher Komplexität und Deployment, wäre die Build Lösung nützlich. Somit wurde entschieden ein NodeJs-Projekt [Node.js Foundation (2016)] anzulegen, indem die Module D3 [Bostock (2016)], http-server [Robbins (2016)] und Bootstrap [Twitter, Inc. (2016)] zunächst über das Node-Package-Management (npm) [npm, Inc. (2016)] installiert wurden. Dabei sollte die Möglichkeit offen gehalten werden, bei Einsatz von mehreren Plugins ein Package zu erstellen, welches das Projekt transportabel machen. Dies setzt allerdings beim Deployment einen Server mit NodeJs Technologien voraus, bzw. eine Entwicklungsumgebung. Für den Fall, eine Implementierung als schnell und einfach anbieten zu können, wurde im Code später das Snippet zur D3.js `<scriptsrc = "https://d3js.org/d3.v3.min.js"charset = "utf-8"></script>` im HTML Head eingesetzt. Da alles auf Javascript und HTML/CSS basiert und dies eine einfache Browser-Lösung ist, sollte das Projekt auch ohne höhere Technologien, große Entwicklerkenntnisse sowie hohe Serveranforderungen ausführbar sein.

Schlussendlich wurde für das gesamte Projekt entschieden, kein Build-System zu integrieren, da es für diese Arbeit nicht nötig ist, das Projekt nicht gefährdet und keine weiteren Mitentwickler

beteiligt waren. Deswegen wurden von Bootstrap nur die vorkompilierten Stylesheets und Javascript Plugins verwendet. Entsprechend wurde die grundlegende HTML-Vorlage übernommen `index.html`, auf welcher die gesamte Visualisierung dargestellt wird.

Es ist gängig in der Entwicklung von Webanwendungen Git als Versions-Kontroll-System (VCS) zu nutzen. So wurde auf Github [GitHub, Inc. (2016)] ein Repository angelegt:

<https://github.com/dororo/Energy-InfoVis>, indem alle Schritte der Entwicklung nachvollziehbar sind und wenn nötig Korrekturen vorgenommen werden können. Ein weiterer Vorteil ist, dass der Code für andere Entwickler und Interessierte zugänglich ist. Der Dienstleister Github wurde ausgewählt, weil er als bekannter und zuverlässiger Repository-Datenbank-Inhaber große Beliebtheit genießt und durch viele weitere Dienste attraktiv für Entwickler ist. Auf diese Weise ist das Projekt, wie ursprünglich mit der Entscheidung ein Package zu erzeugen, mit einem Repository in einer anderen Form transportabel.

5.2 Preprozess (Daten filtern)

Für ein Data-Driven-Dokument (D3) benötigt man die Daten in Datentruckturen CSV, TSV oder JSON, da sonst die D3 Implementierung keine SVG Visualisierung erzeugen kann. In den vorherigen Kapiteln 3 Datensatzanalyse und 4 Konzept wurde festgelegt, welche Daten verwendet werden sollen. Mit der Vorbereitung und Priorisierung konnte eine CSV-Datei als Stammdatenquelle erstellt werden, welche anfangs als Testdaten eingesetzt worden sind und aufgrund der Vollständigkeit, als Hauptdatensatz übernommen wurde. Als letzte Änderung mussten die Spalten mit Stadtfläche, Einwohner pro km², Einwohneranzahl ergänzt werden. Ausgehend von dieser CSV-Datei wurden alle sonstigen Daten errechnet. Dies verhindert, dass für mehrere Datensatzkontexte einzelne CSV-Dateien importiert werden und die Visualisierung wird dynamisch. Damit diese wiederum kompatibel mit Datensätzen aus anderen Städten sind, solange dieselben Daten in der Quelldatei, in einer identischen Struktur zur Verfügung stehen und die analoge Stadtkarte ähnlich wie die aktuelle SVG der Stadt Köln ist. Der Vorteil ist, es entstehen keine redundanten Spalten, wie für die assoziative Identifizierung der Stadtteile mit Namen oder Nummern und sind dadurch eindeutig.

5.3 Layout

Das Layout wurde im vorigen Kapitel 4 festgelegt, jedoch wurden während der Entwicklung inhaltliche Entscheidungen das Detail-Panel betreffend geändert. Die Aufteilung des Layouts in einen linken und rechten Bereich wurde beibehalten, sowie der rechte Bereich des Panels für die Interaktion zu Datenkontextwechsel und Kennzahlen. Der rechte Bereich enthält auch das Panel,

in dem je nach Auswahl die Datenausgabe als Diagramme für nähere Untersuchungen geladen werden. Wie oben erwähnt, wird dafür als Style-Grundlage das Twitter-Framework-Bootstrap [Twitter, Inc. (2016)] verwendet. Es stellt standardisierte Vorlagen bereit, die übernommen werden können. Durch die große Verbreitung des Frameworks, sind versierte Internetnutzer mit der GUI vertraut und sollten wegen Ihrer Erfahrung, wenig Verständnisschwierigkeiten haben. Neben den vordefinierte Schriftstil, werden Farbschema und Erscheinung der Buttons, in diesem Fall Tabs und Formularelemente, die für Interaktionen verwendet werden, übernommen. Ein weiteres Argument für die Anwendung von Bootstrap ist das Grid-System, dies soll unter Berücksichtigung auf mobile Endgeräte Aufwand sparen. Während der Entwicklung fiel allerdings auf, dass einige spezielle Anpassungen der Anforderungen und Einschränkungen durchgeführt werden mussten, die den Aufwand wieder erhöhen, weil sie mit dem Grid-System nicht kompatibel sind.

Die Art der Einbindung der SVG-Stadtkarte ist nicht skalierbar und führt dazu, dass die Karte mit einer festen Breite definiert werden muss. Bei einer Auflösung von 1280 Pixel Breite schiebt sich der rechte Bereich leicht über die Grenze der Stadtkarte und daher werden bei responsiven Verhalten zusätzliche Breakpoints eingesetzt, welche allerdings zu weiteren Schwierigkeiten in der Darstellung führen. Diese Layoutproblematik wurde weitestgehend behoben, aber nicht komplett gelöst. Eine bessere Lösung wird zu einem späteren Zeitpunkt entwickelt, für den Moment reicht es manuell zu skalieren und lenkt damit nicht von der eigentlichen Priorität dieser Arbeit ab.

Aktuell werden die Breakpoints des Grid col-lg-6 mit ≥ 1201 Pixel bis 1523 Pixel neu festgelegt, in der das Panel mit den Radiobuttons, im bereits beschriebenen rechten Bereich angeordnet wird (Abbildung 5.2). Bis zu einer Breite von 1200 Pixel wird das Panel mit den Radiobuttons nach oben rechts, in anderer Boxabmessung, platziert. Wie die Abbildung 5.1 zeigt, ordnen sich die Panels für die Kennzahlen und die Diagramme unterhalb der Stadtkarte an.

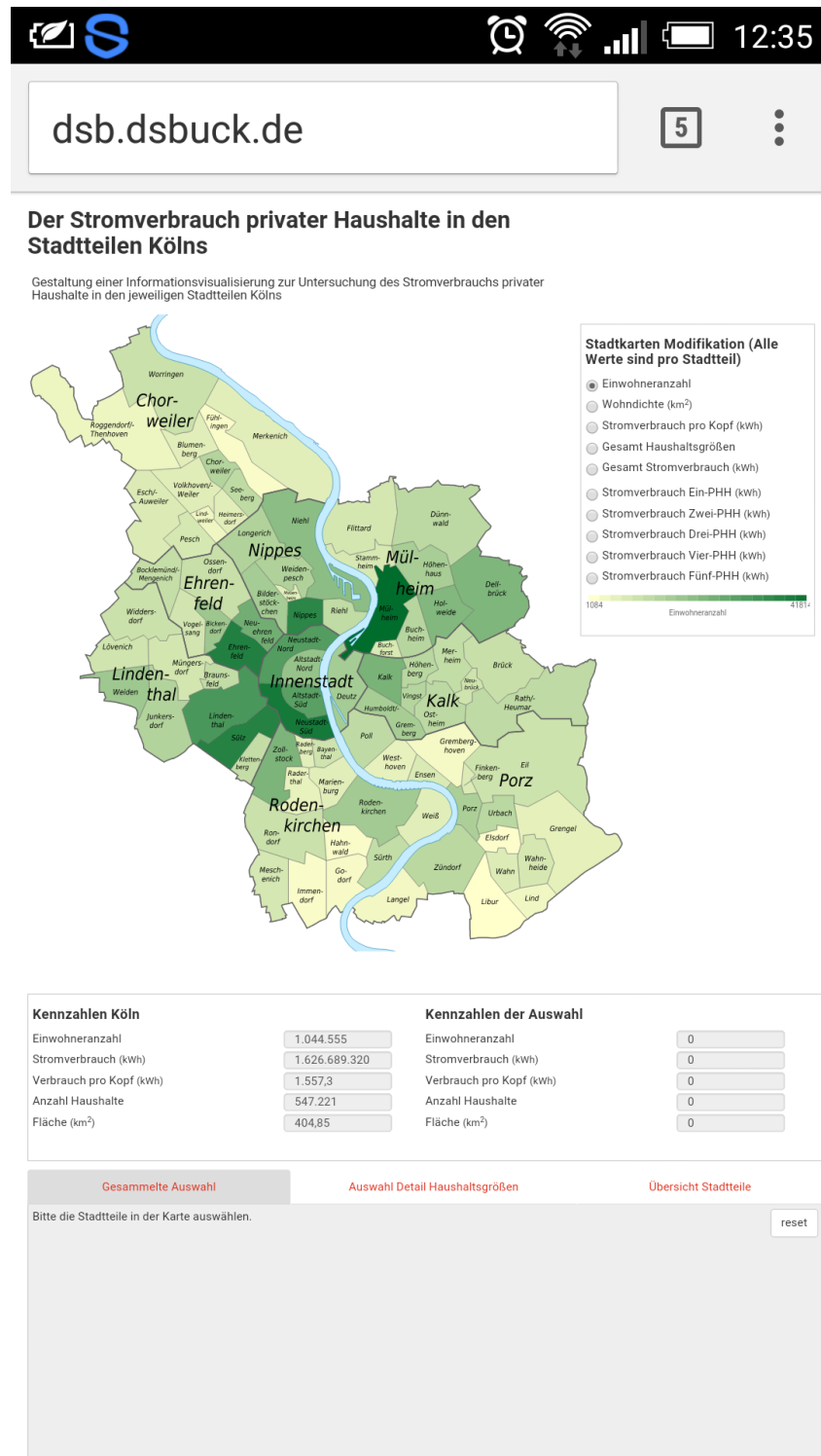


Abbildung 5.1: Das Layout bei mobiler Auflösung bis 1200 Pixel.

Der Stromverbrauch privater Haushalte in den Stadtteilen Kölns

Gestaltung einer Informationsvisualisierung zur Untersuchung des Stromverbrauchs privater Haushalte in den jeweiligen Stadtteilen Kölns

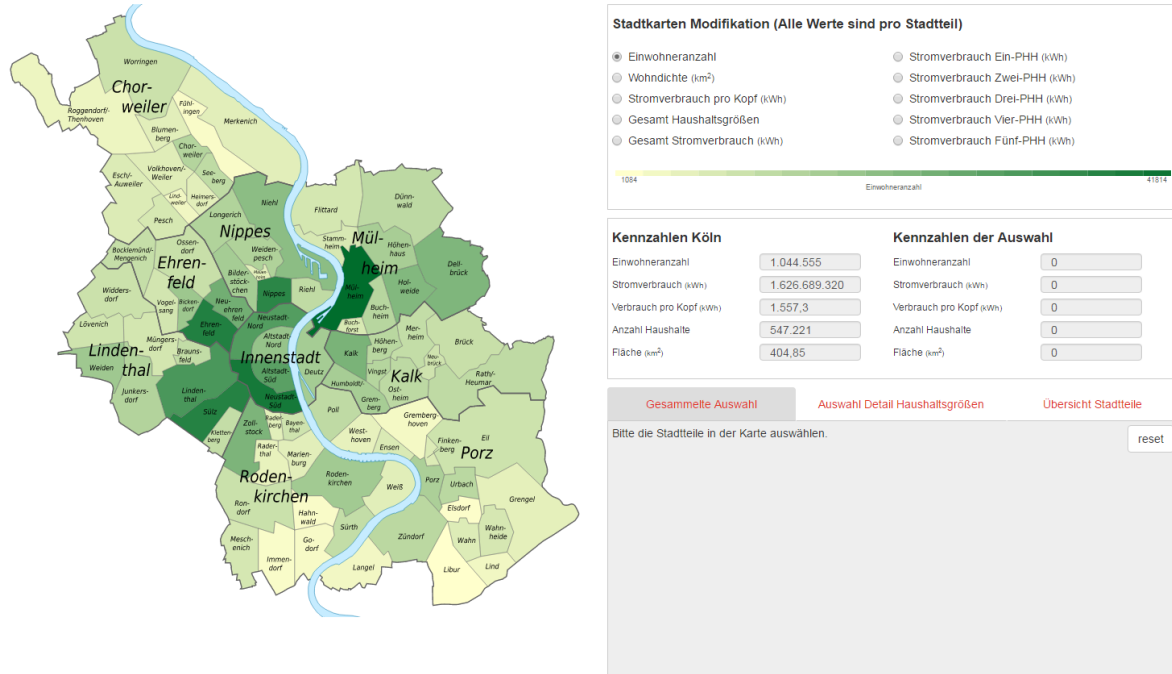


Abbildung 5.2: Fertiges Layout bei einer Desktop Auflösung ab 1201 Pixel.

5.4 Mapping (Codierung, Visualisierung)

Initialisieren bei Aufruf der Anwendung

Ausgehend von der Quelldatei, mit den Spalten „number, stadtteil, sumphh, einphh, zweiphh, dreiphh, vierphh, funfphh, prohh, stadtflaecheqkm, einwohnerqkm, einwohner“ genügen diese Variable für alle anderen Ausgaben, die daraus berechnet werden. Mit der `d3.csv()` Funktion werden die sogenannten Rohdaten aus der Quelldatei „2012-haushaltsgroesse-stadtteil.csv“ geladen und als Array an die Funktion `valuesDetail()` übergeben. In dieser Funktion werden die Kennwerte für das Detail-Panel berechnet, sowie einige Werte der globalen Variablen. Die statischen Kennzahlen werden in der Funktion direkt an die Ausgabefelder übergeben. Die Variablen stehen für die jeweiligen Modifikationen bereit und werden bei Aufruf mit den Radiobuttons oder bei der Auswahl der Stadtteile aufgerufen, dafür ist im Code die Funktion `switchLegends()` vorgesehen Listing 5.1.

```
1 function switchLegends (index) {
2   switch (index) {
3     case "0":
4       legende(resident, "linear", linearGreen, "Einwohneranzahl", 0);
```

```

5      initMap(data, resident, linearGreen, d3.max(resident), d3.min(resident)
      , "einwohner");
6      removeAllBarChart("#detail-panel3");
7      allCityBarChart(data, "#detail-panel3", "Einwohneranzahl");
8      d3.selectAll("#tab2").selectAll(".note").style("display", "block");
9      break;
10     [...]
```

Listing 5.1: switchLegends() Snippet mit Fall 1 für Einwohneranzahl

Das pickerArray() als Objektspeicher (Bucket) An dieser Stelle wurde getestet, ob die Daten in einem ObjektArray gesammelt werden können. Da das Array mit der Methode *addBar()* nicht erweiterbar ist und bei jeder Auswahl das vorhandene Objekt überschrieben hat, wurde darauf verzichtet. Das Problem bestand darin, dass kein Sammel-Array zur Stadtteil Auswahl generiert werden konnte und folglich bei der aufgehobenen Auswahl keine entfernt werden. Dies hätte statisch nach den Werten, die bei Auswahl übermittelt werden, durch die Identifizierung im D3 Data Binding durchgeführt werden müssen.

Deshalb wurde ein neuer Ansatz gesucht und eine Lösung mit der neuen Funktion *pickerArray()* gefunden. Diese ersetzt das ursprüngliche Sammel-Array. Die Funktion *pickerArray()* beinhaltet das bigBucket Objekt-Holder, um daraus die Daten mit Objekten an die Funktion *d3.data()* anzubinden.

Die Stadtteile besitzen eindeutige Stadtteilnummern. Für das Bucket sind eigene Identifikations-Variablen für die Objekte und deren verschachtelte Werte definiert worden. Die Modifikationen werden als einzelne Objekte gespeichert und übernehmen die Reihenfolge aus der Quelldatei. Die Objekte können auf diese Weise, trotz relationalem Zusammenhang nicht zugeordnet werden. Die Lösung besteht darin, eine eigene Identifizierung hinzuzufügen, in der die Reihenfolge der Objekte inkrementiert und diese stellvertretend für die unterschiedliche Identifizierung der Objekte, relativ zu den Positionen der Indizes zugeordnet werden (Zeile 7 in Listing 5.2). So kann z.B. für die Auswahl des Stadtteils Altstadt-Süd, 101, stellvertretend der Index 75, nach der gewählten Modifikation referenziert werden.

groupedValues() Funktion für die Addition der Kennzahlen Die Ausgabefunktion *groupedValues()* wird bei jeder Auswahl eines Stadtteils aufgerufen und summiert die Werte für die Kennzahlengabe. Die Werte für die Ausgabe der Kennzahlen nach Auswahl, werden aus den vorigen Berechnungen der Variablen der Quelldatei errechnet. Die Kennwerte für die Summe Stromverbrauch (kWh) und Verbrauch pro Kopf (kWh) werden mittels *picker()* Funktion aus dem BigBucket ObjektArray entnommen. Da die Kennwerte dort vordefiniert, für jeden Stadtteil einzeln vorhanden und ideal zur Verrechnung sind. Die Interaktion „auswählen“ und „Auswahl aufheben“ auf der Stadtteilkarte, führt die entsprechende Operation auf die Daten aus. Hier

wird zum ersten mal der Stellvertreter Index zur Identifizierung der Position, welche relativ zur Auswahl des Stadtteils ist, verwendet.

Initialisieren der Karte mit `initMap()` Mit `initMap()` wird der Zugang zu den Polygonen der Stadtteilkarte mit D3 Funktionen hergestellt. Die Polygone erhalten Attribute die für die Implementierung der Interaktion benötigt werden, wie z. B. das Einfärben der Polygone für die Choropleth Darstellung, wie der Code in Listing 5.2 zeigt. In der Funktion `colorId()` wird die Farbstärke entsprechend den Werten zugeordnet. Es wird mit `d3.scale.linear()` und mit `d3.domain()` die Mindest- und Maximalgröße entsprechend auf `d3.range` nach Farbstärke je Stadtteil codiert und nach Stadtteil-ID neu eingefärbt.

Farbauswahl für Choroplethkarte Die Farben für die Choroplethkarte werden aus dem Online-Tool <http://colorbrewer2.org/> von Cynthia Brewer, Mark Harrower verwendet [Harrower u. Brewer (2003)]. Für jede Modifikation eignet sich ein sequentielles Schema, wobei die Farbwahl nach Multi-hue oder Single-hue Schema für die Modifikationen unterschiedlich sind und werden auf die Diagramme im Detail-Panel übertragen. Die Modifikationen Stromverbrauch nach Haushalt jedoch, teilen sich einen Single-hue Farbton. Die Farben helfen den Nutzer dabei zusätzlich, die Datensätze nach Modifikationen zu unterscheiden und in der Detail-Ansicht die Daten zuzuordnen.

Wechsel der Modifikationen/Dimensionen mit `switchLegends()` Die Legende im Radio-button-Panel erhält eine eigene Funktion, da diese mit Hilfe des D3 Frameworks `d3.colorLegend` von John Goodall [Goodall (2016)] erzeugt wird. Gesteuert werden die Funktionen für den jeweiligen Datensatzkontexte Einwohneranzahl, Wohndichte, Stromverbrauch pro Kopf, Gesamthaushaltsgrößen, Gesamtstromverbrauch und Stromverbrauch für die jeweiligen Haushaltsgrößen mit den Radiobuttons, welche durch die Funktion `switchLegends()` gesteuert wird. Ein Code-Ausschnitt zeigt Listing 5.1.

5.5 Diagramme für die Detail-Panels

Für die Detail-Panels werden gruppierte Säulendiagramme entwickelt. Dabei wurden mehrere Funktionen entwickelt, die verschiedene Arten von interaktiven Diagrammen implementieren.

Die Anforderung besteht darin, die jeweiligen Werte nach Auswahl der Stadtteile zu mappen. Im Säulendiagramm werden die gewählten Stadtteile für den direkten Vergleich gegenübergestellt (Tab1). Für die Eigenschaften der Werte im Bucket, ist dies grundsätzlich möglich, da alle Werte auch in der Karte dargestellt werden. Mit dem Säulendiagramm ist lediglich eine andere Darstellung gegeben, die andere Nutzeraufgaben, wie zum Beispiel den direkten Vergleich in einem Säulendiagramm ermöglicht. In Tab2 soll die Auswahl nach Haushaltsgrößen gespalten

werden, dazu mussten erneut konzeptionelle Überlegungen durchgeführt werden, welche Daten für die jeweiligen Modifikationen geeignet sind und ob neue Arrays dafür errechnet werden müssen.

Folglich wurden für die jeweiligen Tabs Säulendiagramme entwickelt, welche hier genauer beschrieben werden:

Tab1: Erlaubt eine gemischte Ansicht der ausgewählten Stadtteile und ermöglicht im Panel den direkten Vergleich über die Dimensionen hinweg. Jedoch werden nach Reihenfolge der Auswahl, neue eindimensionale Säulendiagramme gerendert. Die Skalen zeigen je nach Modifikation die maximalen Werte und lassen die Relation der Werte durchblicken. Nachteil ist die gemischte Reihenfolge, da besonders auf die Skalen geachtet werden muss und dies wiederum einen hohen kognitiven Aufwand für den Nutzer bedeutet.

Tab2: In der parallelen Detailansicht, sollen die Werte aus den jeweiligen Modifikationen nach Haushaltgrößen (HHG) zugeordnet werden. Nicht alle Daten der Modifikationen können nach der Dimension der HHG sinnvoll dargestellt werden, wie z.B die Wohndichte. Die Werte nach HHG zu sortieren hat hierbei keine Bedeutung, da die Stadtteilfläche nicht nach Haushaltsgrößen verteilt ist. In diesem Fall, wären die Diagramme entweder redundant, wie beim `headkwh` oder wurden in Tab1 schon einmal gerendert, wie bei den Modifikationen Stromverbrauch per HHG der Fall. Da Tab1 mit diesen Modifikationen eine ungeordnete Ansicht bereitstellt, wird die Lösung durch Tab2 erbracht, in der jede Auswahl der Stadtteile eigene Säulendiagramme erhalten und die Position der Säulen nicht effizient für den Vergleich sind.

Tab3: Löst die Sortierung selbst, allerdings wird die Skala hauptsächlich nach Stadtteilen sortiert. Benötigt wird aber eine nach HHG priorisierte Vergleichsmöglichkeit. Hierbei werden zur Verdeutlichung die Werte für alle Stadtteile als Übersicht dargestellt.

Die Funktionen für das Detail-Panel im Einzelnen

Insgesamt werden vier Methoden implementiert, die vier verschiedene, verhaltende Säulendiagramme bei der Interaktion mit der Karte erzeugen: `allCityBarChart()`, `groupedBarChart()`, `createBar()` und `initkwhPhhBar()`.

Tab1 sammelt die gewählten Werte der Kartenauswahl und instanziert für jede ein Säulendiagramm, mit der maximalen Skala des ausgewählten Datensatzes. Dies wird durch die Funktion `createBar()` erzeugt.

Tab2 wurde für zwei Diagramme entwickelt, die einzelne Werte, oder gruppierte Werte bei der Auswahl hinzufügen und wieder entfernen. Dies sind zum einen *initkwhPhhBar()* und zum anderen *groupedBarChart()*.

Tab3 greift wie die Funktion für Tab1 auf die Daten zu, die für die Choroplethkarte verwendet werden und gibt diese in *allCityBarChart()* mit einem Säulendiagramm wieder aus.

Die Funktionen für Tab2 wurden in drei Methoden aufgeteilt, die eigene Funktionen für die unterschiedlichen Arten der Interaktionen benötigen. Dabei müssen die Diagramme initialisiert werden, dies übernimmt die Methode *groupedBarChart()* oder *initkwhPhhBar()*, diese wird in die Methode *drawGroupBar()* oder *drawKwhBar()* implementiert und bei jeder Interaktion mit der Karte aufgerufen, wenn ein Stadtteil gewählt wird. An dieser Stelle wird das Datenobjekt aktualisiert und rendert neu. Bei der Interaktion „Stadtteilauswahl aus der Karte entfernen“ wird die Methode *exitGroupBar()* oder *exitKwhBar()* getriggert, welche ebenfalls die initialen Methoden implementiert. Das geänderte Datenobjekt wird hier auch aktualisiert und neu gerendert.

Alle vier Funktionen werden in der Funktion *polygonInteraction()* ausgeführt. Diese Funktion beinhaltet die Maus-Listener „on“ und „click“ und implementiert die Interaktionslogik der Visualisierung. Sie verbindet die Karte mit den Detail-Panel in *initMap()*, in der es implementiert ist und wird bei jedem Modifikationswechsel, zusammen mit der Karte neu aufgerufen.

5.6 Nutzerinteraktion

Es existieren zwei grundsätzliche Interaktionsfelder, das Panel mit den Radiobuttons, wodurch bei einem Klick auf einen der Buttons die Karte neu geladen wird und die Choropleth-Funktion aufruft, mit der die Polygon-Flächen der Stadtteile nach Datenwert eingefärbt werden. Dies erfolgt bei Aufruf durch die *initMap()* Funktion und implementiert die D3 Methode für das Browser eigene Tooltip, welche bei Maus-hover über die Stadtteile die Werte den Namen des Stadtteils und Wert ausgibt. In dieser Funktion werden die SVG-Elemente mit Klassen erweitert, die zur Identifizierung der Stadtteilauswahl an die *polygonInteraction()* übergeben werden.

```

1 for (var i = 0; i < data.length; ++i) {
2   colorId(innerSVG, String(value[i]), data[i]["number"], color, max, min);
3   //Tooltip and define polygons pointer
4   innerSVG.select("[id='"+data[i]["number"]+"']")
5     .attr("class", "pointer")
6     .attr("value", String(value[i]))
7     .attr("index", i)
8     .attr("name", data[i]["stadtteil"])
9     .append("title").text(function() {

```

```

10     return data[i]["stadtteil"]+": "+String(new Intl.NumberFormat().format(
11         value[i]));
12 }
13 polygonInteraction(innerSVG, ".pointer", color, max, min, modi, data);

```

Listing 5.2: Die For-Schleife durchläuft jeden Datensatz und setzt für die Polygone in der Karte die Attribute

In der *initMap()* Funktion wird die *polygonInteraction()* Funktion aufgerufen (Zeile 13 in Listing 5.2), die die Interaktion zur Auswahl der Stadtteile für das Detail-Panel implementiert. Ebenso ist dort der Hover-Effekt für die Karte implementiert. Bei Maus-hover werden die Polygone heller und ein Pointer erscheint, welcher durch CSS geändert wurde und auf die Interaktion verweist, welche dem Nutzer den Hinweis gibt, dass diese Polygone klickbar sind. Das Farbfeedback, gibt dem Nutzer zugleich eine visuelle Rückmeldung, welchen Stadtteil er gerade im Fokus hat. Die Polygone werden bei klick zur Auswahl eingefärbt und die ID als Index, welche in *initMap()* den Polygonen hinzugefügt wurde, an die Diagramme Funktionen z.B. an *groupedBarCharts()* übergeben (Listing 5.4). *GroupedBarCharts* ruft dann die Daten zur Id aus dem Bucket ab und fügt sie in ein neues Objekt ein, welches mit D3 in ein SVG-Element übertragen und gerendert wird.

```

1 function polygonInteraction (selection, pointer, color, max, min, modi, data) {
2     [...]
3     selection.select("#Viertel_Flaeche").selectAll(pointer)
4         .on("mouseover", function () {
5             d3.select(this).attr("fill-opacity", .7);
6         })
7         .on("mouseout", function () {
8             d3.select(this)
9                 .attr("fill-opacity", 1);
10        })
11        .on("click", function () {
12            value = d3.select(this).attr("value");
13            objectId = d3.select(this).attr("id");
14            index = d3.select(this).attr("index");
15            var kwhhead = picker("kwhhead", index);
16            var kwhallphh = picker("kwhallphh", index);
17            array = pickerArray();
18        [...]}

```

Listing 5.3: Ein Ausschnitt des *polygonInteraction()*. Zeile 3 bis 17 definiert das Verhalten und deklariert Variable bei Interaktionen

```

1     [...]
2     if (d3.select(this).attr("checked") == null) {
3         d3.select(this).attr("checked", true);
4         d3.select(this).attr("fill", selectedColor)

```

```

5     .attr("fill-opacity", 1);
6
7     createBar(modi, "#detail-panel", array, index, color[1]);
8 [...]

```

Listing 5.4: Bei Auswahl wird die Farbe geändert und in Zeile 7 wird im detail-panel ein Diagramm erzeugt.

```

1 [...]
2 else if (d3.select(this).attr("checked")){
3     colorId(selection, value, d3.select(this).attr("id"), color, max, min);
4     d3.select(this).attr("checked", null);
5
6     removeBarChart("#detail-panel", index, modi);
7     removeBarChart("#detail-panel2", index, modi);
8 [...]
9 }

```

Listing 5.5: Bei Auswahl aufheben wird die Farbe zurückgesetzt und die Diagramme entfernt.

5.7 Bugfixes

Während der Entwicklung sind einige Bugs entstanden und manche Features wurden nicht implementiert, da sie nicht zu den priorisierten Hauptfunktionen gehören oder sind als „nice to have“ Ideen eingefallen. Nachdem die Hauptfunktionen implementiert worden sind und der Prototyp der Informationsvisualisierung weitestgehend funktionsfähig ist. Wurden die notierten Bugs und ToDo's in einer Liste gesammelt und priorisiert. Anhand dieser Liste wurde der Code verbessert und überarbeitet. Die Liste ist im Folgenden aufgeführt. Es wurden als Notiz einige Anmerkungen gemacht, wie die Aufgaben gelöst wurden und welche Probleme aufgetaucht sind. Die Prioritäten 3 und 4 beinhalten Features, die allerdings nicht mehr berücksichtigt wurden. Im späteren Verlauf, nach der Evaluierung werden weitere Bugfixes und Verbesserungen in die ToDo-Liste aufgenommen. Die Liste ist damit nicht final und wird sich bis zur Beendigung des Projekts weiter verändern.

Priorisierte ToDo-Liste

P1

1. Detail-Panel Tab2 Funktion aus Tab3 umbauen, indem entweder die Objektfelder oder die Sichtbarkeit manipuliert wird. Der Umbau macht Tab3 unnötig. Einsatz in den Modifikationen Einwohner, allphh, kwhallphh. **NOTIZ:** Tab3 plottet nun die Modies als Ge-

samtübersicht der Stadtteile. Probleme gehabt die Werte zu gruppieren und im Diagramm hinzuzufügen. Die Gruppen werden nun extra in einem neuen Diagramm pro Stadtteil geplottet. Bei jedem Modifikationswechsel wird der Inhalt von Tab2 gelöscht. In Tab1 bleibt die Auswahl gespeichert. Kleines Feature eingebaut, bei Hover werden die Balken eingefärbt und der Pointer wird aktiv, so kann der Nutzer genauer, auch bei kleinen Werten die Werte für das Tooltip fokussieren. Workaround für Skalenwerte normalisieren. **TODO:** Bugfix mit rangeBound OuterPadding.

2. Recherche über Farbwerte und ordinale Farbskalen und lineare Farbskalen neu vergeben. Farbwahl für Choropleth divergent prüfen und für welche Daten ist es geeignet. (<http://colorbrewer2.org/>) **NOTIZ:** Es bestand das Problem, dass die Werte und die ausgewählten Farbwerte nicht den Eigenschaften der Datenwerte entsprach und somit der Mittelwert der Daten verschoben war. So sind z.B. einige Werte mehr in den unteren Bereichen angesiedelt, sodass die Ausreißerwerte die Choropleth Farbwerte zwar richtig, aber für die Interpretation verfälscht dargestellt wurden. Z.B. ließen sich die niedrigeren Wertbereiche, indem sich die meisten Werte befanden nicht mehr mit Genauigkeit unterscheiden. Deshalb war eine einfache lineare Farbwahl nicht angemessen. So musste die Farbwahl nochmals angepasst werden und jeweils nach Dateneigenschaft für eine Ordinale, Quantile oder Lineare mit drei Werten oder einer linearen mit einer Tonveränderung des min und max Tonwertes.
3. Legende verfeinern und normalisieren, die unteren Skalenwerte vergrößern, damit verhältnismäßig kleinere Werte abgelesen werden können. Besser durch Nutzerinteraktion, wenn es die Zeit zulässt. Da sonst bei Verzerrung der Skalen der Nutzer zu falschen voreiligen Rückschlüssen der Relationen kommt. **NOTIZ:** Durch die Farbwahl Multi-hue: ist die Differenz zwischen min und max Werten signifikanter. Die Natur der Daten ist sequentiell, so sollten die Farben auch linearer Natur sein. **TODO:** Den Median der Daten ermitteln, um Dateneigenschaften in Verbindung mit Farbskalen zu erläutern.

P2

1. Responsive Anpassungen (max 1024 und min 1440px Auflösung optimiert)
2. Quellcode clean up. Formeln auflisten und Variablen
3. Bezeichnungen beschreibender benennen und Messeinheiten
4. Designanpassung y-Grid hinzugefügt
5. Bugfixes

P3 Features

1. Tab1 Verbesserung der Darstellung sortiert. Idealerweise sollten pro Modi ein Diagramm geplottet werden, worin die Balken auch für den besseren direkten Vergleich gegenübergestellt werden
2. Die Auswahl sollte bei Modiwechsel bestehen bleiben, damit der Nutzer bei jeden Wechsel seine Auswahl neu markieren muss.

P4 Features

1. Transition für smooth einblenden der Choroplethkarte
2. Transition bei Detail-Panel-Diagramme
3. Methoden in Klassen aufteilen

6 Formative Evaluation

6.1 Anforderungen

In diesem Abschnitt soll überprüft werden, ob alle Anforderungen die zuvor beschrieben und skizziert wurden eingehalten worden sind. Die Tabellen listen alle Punkte, welche die Visualisierung beschreiben, welche Datensätze visualisiert werden sollen, welche neu errechnet werden, was in der Informationsvisualisierung umgesetzt wurde und was nicht mit einer kurzen Begründung auf. Der Schwerpunkt der Visualisierung wurde auf die Funktionen und Datendarstellung gesetzt, weshalb die meisten Änderungen in der Detailansicht und Datenvisualisierung gemacht worden sind. Obwohl das Design für eine Informationsvisualisierung für den Nutzer zur korrekten Ausführung und Interpretation auch wichtig ist, wird diese dennoch geringer priorisiert als die Funktionalitäten, die während einer Interaktion zu Datenausgaben führen. Diese sind hier nicht aufgeführt und können im Anhang B dieser Arbeit betrachtet werden.

Tabelle 6.1: Anforderungen Datenvisualisierung

Beschreibung	Datensätze	Umgesetzt	Nicht umgesetzt / Bugs
In den jeweiligen Panel (Detail, allgemeine Kennzahlen über Köln) die Kennzahlen und Zusatzinformationen anzeigen.	Einwohneranzahl, Wohndichte (km ²), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamtstromverbrauch (kWh) und Stromverbrauch je Haushaltsgröße.	HTML Elemente Label und Input für benannte Ausgabefelder, der berechneten Werte.	

Die Höhe der Bevölkerungsanzahl auf der Stadtkarte nach Intensivität der Farbsättigung codieren.	Selben Daten wie oben.	Choroplethenkarte mit linearer Farbe Mindest- und Maximalwerte codiert.	Es wird nicht Quantitativ oder Ordinal codiert.
In der Kartenansicht sollten noch keine Datenwerte pro Stadtteil angezeigt werden. Keine Vermischung von Überblick und Detailkontext.	Stadtteilnamen und Werte der jeweiligen abzubildenden Daten.	Standard Browser Tooltip in Choroplethenkarte und alle Säulendiagramme. Bei Hover wird der exakte Wert und bei der Karte zusätzlich der Stadtteilname eingeblendet.	
Überblick Layer blendet für jede Stadtteilfläche ein Kreisdiagramm.	Haushaltgröße, oder Stromverbrauch und Einwohneranzahl.	Statt dessen Tooltip.	Nicht umgesetzt.
Im Detail-Panel werden die Balken, Kreise und andere Datentyp angemessene Diagramme zur Informationsdarstellung erzeugt.	Einwohneranzahl, Wohndichte (km ²), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamtstromverbrauch (kWh).	Die Visualisierungen im Detail werden in Säulendiagrammen präsentiert, da hier eine bessere Gruppierung der Werte je Stadtteil möglich ist.	Keine Kreisdiagramme oder andere Diagramme.

Tabelle 6.2: Anforderungen Detailansicht

Beschreibung	Datensätze	Umgesetzt	Nicht umgesetzt / Bugs
--------------	------------	-----------	------------------------

In der Detailansicht sollen eventuell im oberen Bereich die Kennzahlen und Informationen der ausgewählten Stadtteile aufgeführt und zusammengefasst bzw. berechnet werden.	Einwohneranzahl, Wohndichte (km ²), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamtstromverbrauch (kWh).	Zeigt allgemeine Kennzahlen für ganz Köln und nach Auswahl zusammenaddiert.	Die Kennzahlen werden im Kennzahlen-Panel ausgegeben.
Ein Säulendiagramm kann für direkten Vergleich pro Haushalt ein Stadtteil aufnehmen und ist damit in der Lage nach Auswahl der Karte, die Stadtteile nach Haushalt zu gruppieren und nebenstehende Balken in der Gruppe Haushalt direkt zu vergleichen.	Daten nach Haushaltsgrößen aufgeteilt, in Einpersonen-, Zweipersonen-, Dreipersonen-, Vierpersonen-, Fünf- und Mehrpersonen-Haushaltsgrößen.	Im Detail-Panel unter dem Reiter Vergleich der Auswahl nach Haushaltsgrößen visualisiert in den Modi Einwohneranzahl, Summe aller Haushaltsgrößen, Gesamthaushaltsgrößen.	Die Säulengruppe wird nicht getrennt. Unterscheidung ist nur durch farblichen Rythmus zu erkennen.
Das Diagramm nutzt die ganze Breite des rechten Fensterteils. Falls es über die Breite geht, werden Scrollbalken eingeblendet.	Stammdaten CSV je nach Kontext gefiltert.	Nur im Reiter für Übersicht Stadtteile, da alle 86 Stadtteile visualisiert werden, braucht man eine gewisse Breite.	Werden nicht für Reiter Direktvergleich und Vergleich nach Auswahl eingesetzt.

Die Säulen sollten dynamisch transient sein und bei Größenunterschied sich entsprechend skalieren. Bei Updates sollten die Säulen dynamisch skalieren.	Alle abzubildenden Daten, ist Visualisierungsabhängig.	Skaliert nach Auswahl immer von null bis höchsten Wert. Nur unter Vergleich nach der Auswahl Reiter implementiert.	Keine Animation eingebaut, da die Art wie es implementiert wird nicht vollständig ermöglicht.
Die Informationen, die auf der Karte mit Farbwerten abgebildet werden, können hier nochmal anders für Gruppenauswahl verwendet werden. Es erlaubt die Detailansichten als gesondert und der Fokus wird speziell nur auf die Auswahl gesetzt.	Alle abzubildenden Daten, ist Visualisierungsabhängig.	Wird nur unter Direktvergleich verwendet. Hier wird entsprechend der Maximalwert übernommen, unabhängig von Wertgröße.	Alle anderen Säulendiagramm Darstellungen folgen einer bestimmten Farbreihenfolge.

Problem ist, wie die ganzen Stadtteile in der Gruppierung unterschieden und identifiziert werden. Durch Farben oder Namen, Zahlen, bei Interaktion. (hover?)	Alle abzubildenden Daten, ist Visualisierungsabhängig.	Farbliche Unterscheidung der einzelnen Haushalte wurde festgelegt und in einer Legende definiert. Zusätzlich wurde ein Hover-Effekt hinzugefügt der ein Hinweis auf interaktion mit dem Tooltip geben soll. Namen der Stadtteile werden auf der x-Achse angezeigt.	Die Gruppen sollten ein outpadding erhalten, da die Gruppierung nicht eindeutig zu erkennen ist.
--	--	--	--

Tabelle 6.3: Anforderungen Interaktion

Beschreibung	Datensätze	Umgesetzt	Nicht umgesetzt / Bugs
Bei Klick auf eine Fläche, wird diese farblich oder mit Stärke der Umrandung hervorgehoben. Im Detail-Panel werden die Diagramme für Haushaltsgrößen, Verteilung, Stromverbrauch hinzugefügt. Ein weiterer Klick entfernt es.	Einwohneranzahl, Wohndichte (km ²), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamtstromverbrauch (kWh).	Klickt man auf eine Fläche in der Choroplethkarte, wird diese farblich ausgefüllt in Violett. Damit aktiviert man die Fläche und erzeugt ein Plot im Detail-Panel, sowie die Funktion zur Berechnung der Kennzahlen die addiert werden bei mehrfach Auswahl. Entsprechende Deaktivierung entfernt das Objekt aus dem Diagramm.	Keine Umrandung, daher soll die Auswahlfarbe sich deutlich von allen anderen Choroplethfarben und Diagrammfarben unterscheiden.

Ein Dropdown oder Radiobutton für Datensatzauswahl, die entsprechende Diagramme, Kennzahlen anzeigen. Mit dieser Interaktion wird zwischen den Datensätzen gewechselt.	Selben Daten wie oben.	Der kategorische Datensatzwechsel oder sogenannte Modifikation wird im Panel mit Radiobuttons gewählt und die Karte neu geladen.	Ein Dropdown, zeigt nicht alle Modifikationsmöglichkeiten an und versteckt so Datenkontexte, die noch entdeckt werden müssten.
Eine andere Interaktion könnte parallele Ansichten erlauben, um Überblick und Kontext zu behalten.	Selben Daten wie oben.	Die Detailansichten werden in drei Reiter untergebracht (Direktvergleich nach Auswahl, Vergleich der Auswahl nach Haushaltsgrößen und Übersicht Stadtteile). Nur im mittleren Reiter wird die Auswahl nach Modifikationswechsel gelöscht.	Parallele Ansichten sind wegen den Monitorauflösungen beschränkt.

6.2 Evaluation mit Cognitive Walkthrough

Die Evaluation wird mit der Usability Methode Cognitive Walkthrough (CWT) durchgeführt, da es sich bei der Anwendung um einen Prototypen handelt, der sich noch in der Entwicklung befindet, ist eine formative Usability Methode angemessen [Sarodnick u. Brau (2011)]. Die zweite Begründung ist, dass die Funktionen und Gestaltungsentscheidungen nicht nach den Heuristiken konzipiert wurden und die Visualisierung deshalb nicht den Anforderungen der Gebrauchstauglichkeit entsprechen. Wie im Konzept in Kapitel 4 nachzulesen ist, wurde nach dem Nutzen anhand der Fragestellungen und Aufgaben der Visualisierung entwickelt. Die Evaluation sollte hauptsächlich den Fokus auf die Aufgabenangemessenheit und die Lösung der

Aufgaben durch Interaktion und geringem kognitiven Aufwand bei der Benutzung und Wahrnehmung bzw. Interpretation der Visualisierungen legen. Die vordefinierten Fragen werden für eine erneute Prüfung der Anforderungen herangezogen. Hauptsächlich soll evaluiert werden, ob die Visualisierung nützlich zur Beantwortung der Fragen sind. Letztendlich soll die Performanz getestet werden, die ein Nicht-Experte gegenüber einem Experten leisten kann.

Cognitive Walkthrough [Sarodnick u. Brau (2011)] Die Usability-Evaluations-Methode wird in zwei Phasen unterteilt, in eine Vorbereitungsphase und in eine Analysephase. In der Vorbereitungsphase wird für jede Aufgabe eine konkrete Handlungsabfolge entwickelt, die von Usability-Experten in der Analysephase bewertet werden. Die Handlungsabfolgen werden immer anhand von vier Fragestellungen analysiert. Diese werden aus der Sicht eines Nutzers überprüft, dessen vermutete Kenntnisse und Fähigkeiten berücksichtigt werden. Sollte bei einer Handlungsabfolge Schwierigkeiten auftreten, soll eine alternative Lösung vorgeschlagen werden. Der CWT fokussiert auf die Grundsätze der Dialoggestaltung der DIN ISO 9241-110 [Deutsches Institut für Normung (2006)], Punkt neun Lernförderlichkeit, welche besagt:

"Ein Dialog ist lernförderlich, wenn er den Benutzer beim Erlernen der Nutzung des interaktiven Systems unterstützt und anleitet."

Der Wissenserwerb zur Nutzung des Systems, durch exploratives Lernen, ist auch in der InfoVis eine Methode Wissen zu erlangen, mit der ein mentales Modell über die Domäne aus der die Daten stammen, entwickelt wird.

6.2.1 Vorbereitungsphase

Zielgruppe

Die Zielgruppe unterteilt sich in zwei Hauptgruppen. Zum einen die Experten, welche geübt im Umgang mit Datenpräsentationen sind und bestenfalls Kenntnisse über die Domäne besitzen. Zum anderen die Nicht-Experten oder Novizen, diese beziehen sich auf Nutzer, die selten bis gar keinen Kontakt mit rudimentären Datenvisualisierungen hatten und müssen daher erst eine Übersicht und Verständnis über die Domäne erlangen. In beiden Fällen erlernen die Nutzer, welche Ansichten und welche Daten die Informationsvisualisierung wiedergibt. Von beiden Gruppen wird erwartet, dass sie mit Browseranwendungen Erfahrungen haben, im Besonderen mit Interaktionselementen wie Radiobuttons, Pointer-Cursor und Schaltflächen. Da das Design mit dem bekannten und beliebten Twitter-Framework-Bootstrap umgesetzt wurde, sollten einige Schaltflächen, Designs und Tab-Design den Nutzern bekannt sein. Trotzdem sollte berücksichtigt werden, dass es auch Nutzer geben wird, denen das Design und die einhergehen-

de Bedeutung von Interaktionselementen unbekannt oder in anderer Darstellungsform geläufig sind, wie z.B. das Interface-Design von Software unter den jeweiligen Betriebssystemen.

Nutzerprofile

Das Profil eines Experten:

Andreas W.

Alter zwischen 50 und 60 Jahren, männlich.

Vom Beruf Redakteur und Textchef im Institut der deutschen Wirtschaft IW Medien GmbH, in der Redaktion des Informationsdienst iwd.

Kenntnisse in InDesign und diversen Textverarbeitungsprogrammen. Redakteur für die Typo3 Webseite iwd.de. Die Infografiken stellt er nicht selber her, sondern erhält sie aus dem Wissenschaftsbereich und von den Grafikern der Medienproduktion des Instituts.

Aufgabe des iwd: „Bereits seit mehr als 40 Jahren berichtet der Informationsdienst iwd über die Forschungsergebnisse und Positionen des Instituts der deutschen Wirtschaft Köln zu aktuellen wirtschafts-, sozial-, gesellschafts- und bildungspolitischen Themen. Anders als die fachwissenschaftlichen Veröffentlichungen des IW Köln wendet sich der iwd an die breite Öffentlichkeit – mit journalistisch geschriebenen, allgemein verständlichen Texten, vielen Fakten und anschaulichen Infografiken. Zu den Nutzern zählen Meinungsbildner, Politiker, Mitarbeiter in Ministerien und von Verbänden, Hochschullehrer, Lehrer, Studenten sowie alle an Wirtschaftsthemen Interessierte. Zum 1. Juni 2016 wurde das Konzept für den iwd umgestellt, um die Informationen einem größeren Publikum zugänglich zu machen: Auf einer eigenen Website werden täglich neue Artikel veröffentlicht, die Printausgabe erscheint nun alle zwei Wochen – mit 12 statt nur 8 Seiten.“ [<https://www.iwd.de/ueber-den-iwd/>; Zugriff: 17.07.2016]

Profil eines Nicht-Experten:

Irene W.

Alter 18 Jahre, weiblich.

Auszubildende Mediengestalterin bei der Pixelpark AG und eingeschrieben im Berufskolleg Kartäuserwall in Köln.

Im ersten Jahr der Ausbildung erlangt sie Softwarekenntnisse zu diversen Adobe Produkten zur Layoutgestaltung und Grafikverarbeitung. In der Agentur erhält sie Aufgaben zur Erstellung von Grafiken und wird in die Konzeption miteinbezogen. Ebenso wird sie in den Bereich für Innovation integriert. Pixelpark AG ist eine Agentur die mit Projekten zur Informationsvisualisierung beauftragt wird, genau genommen die Darstellung von Daten des Bundeshaushalts [<https://www.bundeshaushalt-info.de/>; Zugriff: 17.07.2016], sowie Finanzdaten der Arbeitagentur, welche mit Raphaël 2.1.0 entwickelt worden sind. Da Irene im ersten Lehrjahr noch

nicht vertraut ist mit Informationsvisualisierungen und die Daten der Bevölkerungsstatistik Kölns und Energiewirtschaft, wird die Informationsbeschaffung, Einarbeitung und Wiedergabe bzw. Gestaltung eine kleine Herausforderung, jedoch können ihre grafischen Vorstellungsfähigkeiten und Affinität zu Internettechnologien von Vorteil sein.

Aufgaben und Handlungsschritte

Es muss festgelegt werden, welche Ziele mit den Informationen der Visualisierung erreicht werden sollen. Die Aufgaben werden aus den vordefinierten Fragestellungen übernommen, die während der Datenanalyse und zur Konzeption formuliert worden sind. Die Primäraufgabe löst nicht die Fragestellungen im Einzelnen, sondern prüft die Nützlichkeit der Visualisierung und damit, ob durch die Visualisierung das Ziel erreicht werden kann, neues Wissen mit den Informationen aus der InfoVis zu generieren. Es wurden Aufgabenstellungen zur Visualisierung und zum Datensatz gestellt, die hier besonders zur Beschreibung der Aufgaben Evaluierung unterschieden werden müssen. Die Fragestellungen, deren Ergebnisse es zu ermitteln gilt, lassen sich gut zu Aufgaben umschreiben. Auch die Hypothesen, die bestätigt oder widerlegt werden können, gehören zu den Aufgabenstellungen.

Aufgabe 1 Wähle einen Stadtteil nach Interesse aus und bewerte die Datenausgaben der jeweiligen Ansichten.

Aufgabe 2 Wähle mehrere Stadtteile nach Interesse aus und vergleiche die Datenwerte der Gruppen nach der Auswahl. Bewerte die Ausgabe.

Aufgabe 3 Leere die Ansichten und die Auswahl.

Aufgabe 4 Wähle eine der Modifikatoren bzw. Informationskontexte für die Karte. Wähle die Stadtteile nach Interesse aus und bewerte die Ausgaben der jeweiligen Ansichten.

Aufgabe 5 Wähle einzelne Stadtteile aus, ganz unabhängig in welchem Kontext die Karte aktiviert ist. Entferne einzelne Stadtteil aus der Auswahl.

Aufgabe 6 Wähle und sammle die Werte der Stadtteile in verschiedenen Kontexten. Bewerte die Datenwerte.

Aufgabe 7 Wähle und sammle die Werte der Stadtteile in verschiedenen Kontexten und bleib in der Ansicht „Direktvergleich nach Auswahl“. Entferne einzelne Werte der Auswahl beliebig aus der Ansicht.

Aufgabe 8 Wähle und sammle die Werte der Stadtteile in verschiedenen Kontexten und bleib in der Ansicht „Direkt Vergleich nach Auswahl“. Entferne die Auswahl komplett aus der Ansicht.

Aufgabe 9 Nachdem die Funktionsweisen der Visualisierung vertraut sind, versuche die Fragen 1-16 zu beantworten.

Handlungsabfolge

Bis zu Aufgabe 8 sind in diesem wenig komplexen System der Informationsvisualisierung alle Aufgaben einfach aufgebaut. Die eigentliche Herausforderung besteht in der kognitiven Leistung, die Fragen zu beantworten. Die ersten Fragen lassen sich einfach beantworten. Der Schwierigkeitsgrad steigt mit der Komplexität und der Anzahl der zu vergleichenden Werte und Kombination der Ansichten. Damit wird die Visualisierung auf die Probe gestellt, inwieweit sie dem Nutzer die Datenanalyse vereinfacht. Zumal dies Fragen sind, in denen die Datenwerte korreliert werden müssen. Ein Streupunktdiagramm wäre für diese Fragen angebracht, jedoch wurde in der Konzipierung davon abgesehen einen weiteren Diagrammtyp zu implementieren. Der Nutzer muss also versuchen mit der aktuell implementierte Informationsvisualisierung eine Antwort auf seine Fragen zu finden. Die Fragen 12, 14, 15 und 16 verlangen einen Kontextvergleich und sind daher schwerer zu beantworten. Hat der Nutzer die Bedeutung der Ansichten erkannt, ist er in der Lage diese Fragen zu beantworten. Da die einfachsten Fragestellungen schon aus der Oberflächenansicht der Kennzahlen abzulesen sind, werden sie nicht detailliert spezifiziert.

An dieser Stelle sollen die Fragen 12, 14, 15 und 16 auf ihre Machbarkeit und Performanz gegenüber der Aufgabenlösung geprüft werden. Die einfacheren Aufgaben sind Gegenwerte, bei der eine Minstdauer zu ermitteln ist, die eine Beurteilung mit Startwert von der Minimal- zur Maximaldauer einer Aufgabe quantifizierbar machen soll. Während die Antwort zu Frage 12 explorativ zu ermitteln ist, sind die anderen Fragen nur zu beantworten, wenn man sich genauer mit den Daten befasst hat und diese kennt. Sonst besteht das Problem, dass die Zusammenhänge nicht zu erkennen sind und nicht klar ist wo gesucht werden soll.

Die Lösungen zu den Fragen 12, 14, 15 und 16 werden im Anhang D beschrieben.

6.2.2 Analysephase

Die Analysephase wird von zwei Usability-Experten aus dem Fachbereich Medieninformatik durchgeführt und bewertet. Es wird abgeschätzt, ob die Nutzer die Handlungsschritte entsprechend zur Aufgabenlösung erkennen und durchführen. Dazu werden zu jeder Aufgabe die 4 Leitfragen geprüft und festgestellt, ob die Handlungsschritte eingehalten werden. Parallel dazu soll auch die Dauer, die ein Nutzer benötigt, dokumentiert werden. Damit soll die Effizienz der Visualisierung und das Ziel der Arbeit geprüft werden und die Frage, ob Nicht-Experten genauso effizient wie Experten Wissen erlangen können beantwortet werden. Dokumentiert und begrün-

det werden die Aufgaben / Handlungsschritte mit Erfolgs- und Misserfolgsstory. Bei Misserfolg werden Verbesserungen vorgeschlagen

Die 4 Leitfragen

1. Erkennt der Nutzer die Interaktionselemente und ihre Bedeutung, die zu bestimmten Ansichten führen oder Informationen ausgeben?
2. Wird der Nutzer die Elemente effizient oder explorativ dafür nutzen, Informationen aus der Visualisierung zu finden?
3. Versteht der Nutzer die Visualisierungen der Daten, sodass wenig Fehlinterpretationen entstehen können?
4. Ist die Visualisierung dafür geeignet, Wissen effizient zu erlangen und weiteres Wissen davon abzuleiten?

Erfolgs - Misserfolgsstory

Ein detaillierter Einblick in die Ergebnisse befindet sich im Anhang C. Die Tabellen sind in drei Spalten unterteilt, die einmal die Aufgabe, den Erfolg- oder die Misserfolgsstory beschreiben und nach den 4 Leitfragen evaluiert werden. Bei Misserfolg wird die letzte Spalte ausgefüllt und in diese die Verbesserungsvorschläge eingetragen. Die Fragen werden in eine Fragespalte und Antwortspalte aufgeteilt. In der Handlungsschritte- und Antwortspalte werden die Vorgehensweisen der Nutzer und ihre Antworten zeitmäßig erfasst.

Den Experten wurden zusätzliche Fragen betreffend der Wahrnehmungsreihenfolge der InfoVis gestellt. Es sollte zur Validierung verwendet werden, wie die Elemente der wahrgenommen werden und mit welchen Interaktionen er daraufhin beginnt die Informationsvisualisierung zu verwenden.

Würde der Nutzer die InfoVis zuerst obereflächlich betrachten, bevor er interagiert?

Nein, er bevorzugt, die InfoVis explorativ mit Interaktionen kennen zu lernen. Die Maus-Effekte verleiten den Nutzer dazu spielerisch zu interagieren und Funktionen kennen zu lernen.

Welche Elemente fallen ihm dabei chronologisch auf? Der Eye-catcher ist die Choroplethenkarte, als zweites fällt ihm der Inhalt in Tab1 auf, da er durch den Hover-Effekt auf der Karte unterbewusst verleitet wird eine Auswahl durchzuführen und als ein Feedback, das Laden der Diagramme in Tab1 zusätzlich auffällt. Als drittes fällt ihm relativ schnell das Panel mit den Modifikatoren auf, da dieses oben rechts prominent platziert wurde. Die Buttons laden

ebenso zur Interaktion ein. Beim Farbwechsel der Karte, stellt sich der Nutzer die Frage, welche Bedeutung sich dahinter verbirgt.

Ein kritischer Punkt während der Interaktion ist das Problem der responsiven Gestaltung bei bestimmten Auflösungen. Die untersten Panels, könnten bei einer mobilen Auflösung auf dem Desktop Browser übersehen werden, da diese im Browser nur mit scrollen sichtbar werden.

6.3 Ergebnis

Während des formativen Evaluation Cognitive Walkthrough mit den Nutzerprofilen eines Experten und einer Nicht-Expertin konnte überprüft werden, ob die gestaltete Informationsvisualisierung die Anforderungen erfüllen, indem die Gebrauchstauglichkeit getestet wurde und der Grad des kognitiven Aufwands der Nutzer gemessen wurde. Die geografische Darstellung der Daten in einer Choroplethkarte, wirkte auf den Nutzer ansprechend und forderte ihn durch den Hover-Effekt bei der ersten Maus-Interaktion auf, die Stadtteile anzuklicken. Damit wird eine explorative Untersuchung des Nutzers eingeleitet. Hierbei nimmt der Nutzer Schritt für Schritt alle interaktiven Elemente wahr und lernt auf diese Weise die Funktionen und Informationsausgaben der Visualisierung kennen. Gelenkt wird der Nutzer durch den Usability-Experten als Moderator, der ihm die Aufgaben stellt und bei der Exploration begleitet. Es hat sich herausgestellt, dass die Bedeutung der Interaktionselemente für den Nutzer bekannt sind und es deswegen keine Verzögerung gab. Vor allem der Nutzen des Resetbutton, wurde oft und effizient verwendet. Die Auswahl der Stadtteile über die Choroplethkarte ist die Hauptfunktion der Visualisierung und wurde in jeder Aufgabe gestellt. Durch die wiederholte Ausführung der Aufgaben und Unteraufgaben, gewöhnte sich der Nutzer an die Informationsvisualisierung, er konnte die Ausgaben immer besser verstehen und war damit in der Lage effizienter die Aufgaben zu lösen. Dies bemerkte man besonders bei der Bearbeitung der Fragen. Der Nutzer hatte wie erwartet, ein mentales Modell der Daten erzeugt und lernte während der Evaluation immer mehr dazu. Der Proband kommentiert: „Ich habe viel über Köln gelernt.“ Das zeigt, dass die Visualisierung nützlich ist, Informationen zu Bevölkerung und Energieverbrauch einer Stadt und ihrer Stadtteile, dem Nutzer näher zu bringen, ohne das eine fundierte Erfahrung nötig ist.

Dennoch konnten einige Schwächen durch die Evaluation ermittelt werden und mit Verbesserungsvorschlägen Lösungen entwickelt werden. Die Schwächen zeigten sich in Form einer fehlenden Farblegende als Übersichtshilfe der verschiedenen Modifikationen und Probleme in der Darstellung bei mobilen Auflösung. Auch wird vom Nutzer vermisst, dass die Auswahl in der Karte nach dem Wechsel in einen anderen Datensatzkontext und Features nicht beibehalten werden.

Interessanter und wichtiger ist allerdings die inhaltliche Bedeutung der Visualisierung, die mit den Fragen evaluiert wird. Dabei konnten einige Unklarheiten die z. B. in der Datenausgabe innerhalb der jeweiligen Tabs nach der benutzerdefinierten Auswahl vorkamen gefunden werden. Die Metriken und Bezeichnungen sind nicht sprechend und störten den Nutzer bei der Bearbeitung der Fragen. Oder die Datenbedeutung waren schwer nachvollziehbar, da die Ausgaben, wie in Tab3, schwer lesbar und manche Informationen schwer auffindbar waren und den Nutzer verwirrten, da die Ausgaben nicht den Erwartungen entsprachen. Teilweise wurden manche Datensätze nicht korrekt zur Lösung verwendet, wie z.B. die Modifikationen zum Stromverbrauch je Haushaltsgrößen. Diese wurden als Bewohnergrößen und nicht als Stromverbrauch in kWh fehlinterpretiert.

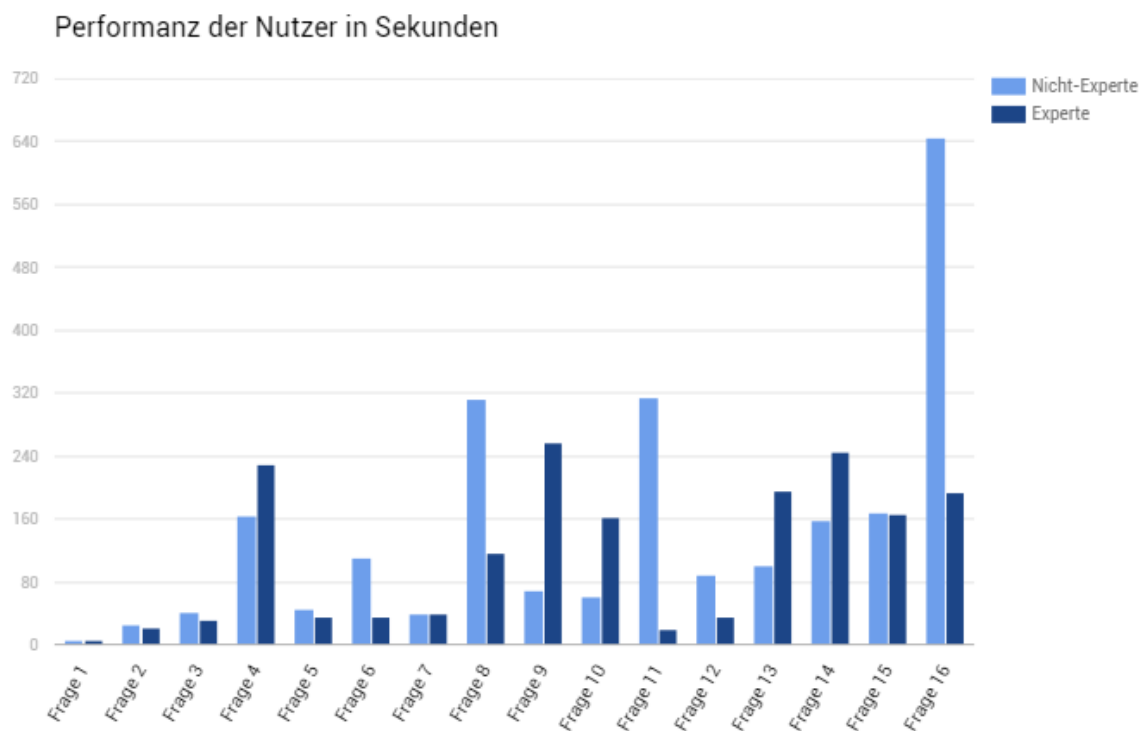


Abbildung 6.1: Die Dauer die der Experte und Nicht-Experte zur Lösung der Fragen benötigt hat, in Sekunden gemessen. Erstellt von Dorothea Buck.

Um die eigentliche Frage zu beantworten, ob der Nicht-Experte genauso souverän wie der Experte die Visualisierung zum Wissen schafft und wie viel besser oder schlechter er damit umgeht, wurde die Dauer gemessen, die beide zur Bearbeitung der jeweiligen Fragen benötigten. Folgend visualisiert das Diagramm (Abbildung 6.1) wieviele Sekunden die Nutzer gebraucht haben. Trotz der Tatsache, dass die Fragen 1 bis 11 und 13 leichter zu beantworten sind, wurde mehr Zeit beansprucht, als erwartet. Verwunderlich ist daher bei der einfacheren Frage 4, dass hier beide Nutzer länger gebraucht haben, obwohl die Frage durch einfaches ablesen der Choroplethkarte beantwortet werden kann. Die Ursache liegt darin, dass die Karte nicht einen Stadtteil als

geringsten oder mächtigsten darstellt, sondern mehrere Stadtteile mit ähnlichen Farben, d.h. Werten. Hier ist nicht sofort erkennbar, welche Stadtteile die Lösung sind, was die Nutzer zu anderen Lösungswegen zwingt. Bestätigt wurde, dass der Nicht-Experte bei manchen Fragen wenige Sekunden länger gebraucht hat, aber bei den komplizierten dagegen schneller als der Experte war, außer der letzten Frage. Diese letzte Frage diente dazu zu testen, ob der Nutzer die Visualisierung im Ganzen verstanden hat und in der Lage ist, sich darin zurecht zu finden. Der Experte brauchte in anderen Fragen länger, da er akkurat nach Gegenbeispielen gesucht hat, die die Lösung bestätigen und die optimalen Lösungswege zu finden. Dies galt vor allem für Fragen, bei denen die Antworten widersprüchlich waren.

6.3.1 Verbesserungsvorschläge

In diesem Abschnitt wird eine Liste der Verbesserungsvorschläge und Lösungen zu Problemen, die bei der Evaluation entdeckt wurden, aufgestellt. Pragmatischerweise soll diese Liste als eine ToDo-Liste verwendet werden. Sie soll genauso wie die bisherigen ToDo-Listen priorisiert werden. Dazu werden die Vorschläge nochmals hinterfragt und andere, angemessenere Verbesserungen entwickelt. Die inhaltlichen Verbesserungen werden in die erste Priorität sortiert, da die Bedeutung der Informationen das Herz der Visualisierung ist und dem Nutzer Korrektheit und Glaubwürdigkeit durch eine akkurate Präsentation der Informationen garantiert werden muss. Ebenso in die erste Priorität kommen die schwerwiegenden Vorschläge, die zur visuellen Optimierung der Visualisierung beitragen. Priorität zwei und drei sind Verbesserungen für Elemente, die schon vorhanden sind, jedoch anders gelöst oder erweitert werden und nach Wichtigkeit sortiert werden sollen.

Prio 1:

- Formatierung der Metriken sollte mit deutschen Bezeichnungen versehen werden.
- Titel „Kennzahlen der Auswahl“ muss umbenannt werden in „Summe der Auswahl“.
- Die Auswahl in der Karte sollte beibehalten werden, wenn die Modifikation gewechselt wird
- Der Fehler ist, dass eine Auswahl aus einer vorherigen Modifikation neu in das ObjektArray hinzugefügt werden muss, um sie komplett aus der Modifikation zu entfernen. Dieser Bug steht im Zusammenhang mit der Auswahl der Modifikationswechsel speichern und wieder anzeigen, wenn zurück gewechselt wird.
- Die Übersicht Tab3 erschwert die gezielte Suche nach Stadtteilen. Die Stadtteile sollten sortierbar sein und nach Hover/Auswahl Interaktion hervorgehoben werden. Eine alterna-

tive Lösung ist die existierende Sortierung nach Bezirken zu übernehmen und eine weitere x-Achse mit Bezirksnamen und Grid-Lines zur Eingrenzung einzufügen.

Prio2:

- Es sollte eine Legende oder eine farbliche Differenzierung der bestehenden Farben zu den jeweiligen Modifikatoren in der GUI ersichtlich sein. Man könnte die Schrift, eine Legende oder sonstige farbliche Abhebungen einbauen.
- Alternative für Kennzahlen nach Auswahl wäre ein Tooltip bei hover in der Tab1- oder Tab2-Ansicht
- Die Diagramme sollten in dem Fall ein Tooltip erhalten, welche beim Anklicken die einzelnen Kennzahlen, wie in der Box darüber ausgeben.

Prio3:

- Die Legende könnte prominenter platziert werden. Z.B. über der Karte
- Verdeckte Legende bei HD Auflösung, Legende sichtbar nach links oder nach oben setzen.
- Die Tab2- und Tab3-Ansichten werden nicht konsequent in den Vergleichen angewendet. Die unterschiedlichen Inhalte und Sichtbarkeit der Inhalte stellen ein Hindernis dar. Es sollte die Möglichkeit geben die Panels in eigene Fenster loszulösen. Das setzt allerdings eine passende Auflösung voraus.

7 Ergebnis und kritische Bewertung

Die Forschungsfrage wurde in der Einleitung detailliert formuliert und handelt von der Gestaltung einer Informationsvisualisierung für Nicht-Experten oder Novizen zu dem räumlichen Thema der Bevölkerungsstatistik Kölns und des allgemeinen Energieverbrauchs in privaten Haushalten der Stadt. Damit wurde zugleich auch die Aufgabe definiert. Hauptsächlich geht es in der Arbeit darum, 2D Visualisierungstechniken zu finden, die einfach und schnell interpretierbar sind, wenig Raum für Fehlinterpretationen lassen, komplexe mehrdimensionale und multivariate Datenstrukturen abbilden und dennoch einen geringen Abstraktionsgrad beibehalten können, sowie für den Überblick des Beobachtungsraumes interaktiv sind.

Bei der Recherche wurden mehrere Visualisierungstechniken nach Kategorisierung der Datentypen und Eigenschaften verglichen. Nach diesen spezifizierten und definierten Eigenschaften wurden angemessenen Visualisierungstechniken nach der Konvention der InfoVis ausgewählt. Ebenso konnte die Arbeit von Lee [Lee u. a. (2016)] bei der Abwägung hilfreich sein, da hier anhand der Studie sichtbar wird, wie Novizen mit multidimensionalen und -variate umgehen und welchen Hürden sie sich stellen müssen.

Mit den GIS-Anwendungen werden Beispiele von vorangegangenen Anwendungen gezeigt, die verdeutlichen, dass bei räumlich, thematischen Informationssystemen, die Karten mit weiteren Visualisierungstechniken erweitert werden und anzeigen, dass die essentielle Interaktion *linking and brushing* in jeder Anwendung grundlegend für die explorative Untersuchung einer InfoVis ist.

Im Kapitel 6 der Evaluierung wird verfolgt, inwieweit die Informationsvisualisierung auf den Nutzer wirkt. Es wird geprüft ob der Prototyp die Anforderungen und Benutzerziele erreicht hat. Bei der Evaluation wurde erkannt, dass der Nutzer sich sehr schnell mit den Interaktionselementen auseinandergesetzt hat, unterstützt durch die Wahl der Elemente, die im Allgemeinen für Browseranwendungen standardisiert sind. Bei jeder Interaktion, erhält der Nutzer ein Feedback, das zeigt, welche Datenpräsentationen und Bereiche im Browserfenster sich verändern. Dabei versucht der Nutzer aus den Präsentationen Informationen zu erhalten und hat so Schritt für Schritt die Hauptfunktionen der Informationsvisualisierung erlernt. Ein weiterer Erfolg konnte bei der Bearbeitung der Fragen verfolgt werden, auch wenn es noch einige Schwächen bei der Benennung der Datenausgaben gab, die zu einer fehlerhaften Nutzung führten.

In der Evaluierung zeigte sich, dass die Auswahl der primitiven Visualisierungstechniken in Kombination mit RIS, für Nicht-Experten und Novizen die bessere Wahl ist, wenn es darum geht Effizient, Expressiv und eine für den Nutzer angemessene InfoVis-Anwendung zu gestalten. Die meisten Bearbeitungsziele aus dem Konzept konnten damit erfüllt werden.

Identifikation, Unterscheidung, Lokalisierung, Einteilung, Bündelung, Anordnung, Vergleich, Assoziation und Korrelation.

Überblick, Vollständigkeit, Zusammenhang, Überprüfbarkeit von Hypothesen, Analyse, Erlernbarkeit, keine Vorerfahrung nötig, Nachvollziehbarkeit, Verständlichkeit, Korrektheit und Vereinfachung.

Außer der Anordnung, Einteilung und Bündelung, die nur bedingt anwendbar sind und eine gesonderte Interaktionsmöglichkeit erfordern, mit dem das Clustern verbessert werden kann. Und die Korrelation, welche zu den schweren Aufgaben in der InfoVis gehörte und auch nur sehr schwer und ungenau auszuwerten war. Zwar wurde anfangs beachtet, dass die Datensätze zu mehreren Dimensionen zusammengestellt werden können und damit Datentypen erzeugt wurden, die für wissenschaftliche Visualisierungstechniken prädestiniert wären und der Analyse für Korrelationen angemessen sind. Oder die statistische Korrelation mit dem Regressionsverfahren. Insgesamt wurden viele wissenschaftliche Methoden, die zu einer genaueren, zuverlässigeren und glaubwürdigeren Visualisierung hätte führen können, nicht implementiert. Andererseits stellte sich die Frage, welche analytischen Fähigkeiten der Nicht-Experte besitzt, und inwiefern diese notwendig sind. Sowie die realistische Machbarkeit der Gestaltung der Informationsvisualisierung vom Konzept bis zu einem nutzbaren Prototypen. Die Abwägungen zu Gunsten der Einfachheit für die Zielgruppe haben überwogen. Der Anspruch des Nicht-Experten für eine komplexe Fragestellung, die eine Korrelation mit mehrdimensionalen Darstellungen beinhaltet, sollte in der Weiterentwicklung geprüft werden.

8 Zusammenfassung

Die Hauptaufgabe in dieser Arbeit ist es, eine Informationsvisualisierung zu gestalten für Nutzer, die hier als Nicht-Experten und Novizen definiert werden. Das Recherchethema Energiewende wurde aus der Vorarbeit im Praxisprojekt [Buck (2014)] übernommen und spezialisiert sich auf das lokale Interessensgebiet Köln. Mit der Fragestellung „Wie der Stromverbrauch in den privaten Haushalten sich verteilt?“ werden die einzelnen Stadtteile und Haushaltsgrößen betrachtet. Nach der Festlegung des Visualisierungsthema, werden aus den Wissenschaftsbereichen InfoVis und Visual Analytics, Vorgehensweisen und grundlegende Erkenntnisse aus den letzten 50 Jahren recherchiert, die einen Überblick zu entworfenen Techniken, Problemfeldern, Datentypen, Anwendungen für diese Datentypen und Kategorien, sowie neuste Entwicklungen zeigen. Aus dieser Recherche konnte der Ausgangspunkt festgelegt werden, welcher immer die Daten sind. Mit einem etablierten Vorgehensmodell aus der InfoVis, dem Referenzmodell, für die Gestaltung und Visualisierungsentwicklung werden weitere Schritte abgeleitet. Die abgewogenen Vorgehensmodelle wurden für diese Arbeit kombiniert und hier dokumentiert.

Aufgrund der Datenspezifizierung konnte eine angemessene Visualisierungstechnik, GIS zur Realisierung gewählt werden. Der Fokus der Zielgruppe und Anwendungsaufgaben mussten dabei jederzeit berücksichtigt werden. Weswegen sich hier für primitive Visualisierungstechniken entschieden wurde, da diese für Novizen weniger kognitiven Aufwand erfordern. Der erste Schritt für die Datenanalyse konnte Ergebnisse zum Inhalt der Informationen und dem Umgang für eine korrekte Interpretation erarbeiten. Im zweiten Schritt werden die Anforderungen analysiert, die an der Visualisierung, den Nutzer und die Daten gestellt werden. Dabei wurde ein Konzept geschrieben, welches die Ziele der Informationsvisualisierung definiert und erklärt wie diese zu erreichen sind. Anhand der Anforderungsanalyse und dem Konzept, wurde das Layout, die Präsentation der Daten und Interaktionen zur Bearbeitung der Aufgaben beschrieben und skizziert. Schritt drei ist die Umsetzung eines Prototypen. Während der Implementierung mussten am Konzept programmierspezifische Änderungen vorgenommen werden. In einem iterativen Bugfix-Prozess, konnten die minimalen Anforderungen implementiert werden und sogar noch erweitert werden. Nicht alle Funktionen wurden übernommen und für die Weiterentwicklung in einer priorisierten ToDo-Liste dokumentiert. Der Stand des Prototypen konnte soweit umgesetzt werden, dass dieser für eine Evaluation bereit war. Der vierte Schritt ist die Evaluation des Prototypen und wurde mit dem Cognitive Walkthrough durchgeführt. Hier zeigen sich die Fähigkeit der InfoVis, Daten mit einfachen Visualisierungstechniken zu präsentieren und inwieweit

die Aufgabe in dieser Arbeit gelöst wurde. Die Evaluierung offenbart auch einige Schwächen, die in Form von Verbesserungsvorschlägen dokumentiert wurden.

Ausblick

Wie im Kapitel 5 Umsetzung und Kapitel 6 Evaluation zum Ende angedeutet, soll die InfoVis-Anwendung weiterentwickelt werden. Da es sich um einen Prototypen handelt, müssen die fehlenden Funktionen und Verbesserungsvorschläge der Evaluation implementiert werden. Des Weiteren sollte die Anwendung auf Hinblick der Visualisierung für andere Städte getestet werden. Dafür wird überprüft, in welcher Form andere Kommunen Regionalkarten und Bevölkerungsdaten zur Verfügung stellen. Auf diese Voraussetzungen sollte die InfoVis angepasst werden. Es ist vorgesehen eine offene Schnittstelle für die Dateneingabe von Bevölkerungsdaten, mit den Merkmalen Haushaltsgrößen, Bevölkerungsdichte und Einwohneranzahl der jeweiligen Stadtteile einer Stadt in die Visualisierung zu implementieren. Auf diese Weise kann die Visualisierung als Werkzeug verwendet werden, welches Ansichten der Ergebnisse zum Stromverbrauch der Städte neu erzeugt und exportiert und in der Lage ist zwischen den Städten Vergleiche aufzustellen.

Mit dieser Erweiterung hat die Informationsvisualisierung „Energy-InfoVis“, wie sie in GitHub benannt wurde, das Potenzial in der Entwicklung der Energiewende in Deutschland, den Energiebedarf verschiedener Städte zu vergleichen. Dies könnte zum Beispiel für Kommunalpolitiker von Nutzen sein, um die Daten für Projekte zur Energieeinsparung zu verwenden. Dabei werden die Verbraucher einer Stadt vom Energieversorger beim SmartMetering-Rollout mit Smart Metern ausgestattet mit denen zusätzlich die Städte verglichen werden können, die nicht am Rollout teilnehmen. Dieser Vergleich kann aussagen, ob ein hoher oder niedriger Unterschied des Energiebedarfs der privaten Haushalte besteht. Der Anwender kann mit der Kommunikation für die Energieverbrauchsergebnisse verschiedenste Zielgruppen ansprechen. Er erreicht damit eine Sensibilisierung der Nutzer und bringt ihnen durch Energy-InfoVis einen sorgsam Umgang mit den Ressourcen optisch verständlich nahe. Langfristig kann mit Energy-InfoVis die Entwicklung der stadtbezogenen Energiebilanzen beobachtet werden.

Abbildungsverzeichnis

2.1	Funktionen der Geovisualisation	12
2.2	GeoDa	16
2.3	GeoVISTA	17
2.4	iVisDesigner	18
2.5	Visualisation Pipeline	20
2.6	Erweitertes Referenzmodell	21
3.1	Jahres Stromverbrauch je Haushalt nach HHG und HHM BDEW Beiblatt	24
3.2	Verbrauch nach PHH in Prozent	31
3.3	Anteile PHH nach Prozent	32
3.4	Durchschnittsverbrauch der PHH in kWh	33
4.1	Layout Mockup	41
5.1	Layout auf mobilen Gerät	45
5.2	Layout auf Desktop Auflösung	46
6.1	Diagramm zur Nutzerperformanz	67

Tabellenverzeichnis

3.1	Ergebnisse der Berechnungen des Kölner Stadtteils Neustadt-Süd.	25
3.2	Zugewiesene Skaleneigenschaften der Dimensionen und Variablen.	27
6.1	Anforderungen Datenvisualisierung	55
6.2	Anforderungen Detailansicht	56
6.3	Anforderungen Interaktion	59

Literaturverzeichnis

- [Anseli u. a. 2006] ANSELI, Luc ; SYABRI, Ibnu ; KHO, Youngihn: GeoDa: An Introduction to Spatial Data Analysis. In: *Geographical Analysis* (2006)
- [Bill 1999] BILL, Ralf: *Bill, Ralf. Grundlagen der Geoinformationssysteme. Bd. 2: Grundlagen der Geo -Informationssysteme. Band 2. Analysen, Anwendungen und neue Entwicklungen.* 2., völlig neubea. Heidelberg : Wichmann, 1999. – ISBN 9783879073269
- [Bostock 2016] BOSTOCK, Mike: *Data-Driven Documents (D3).* <https://d3js.org/>. Version: 2016. – [Zugriff: 16. August 2016]
- [Buck 2014] BUCK, Dorothea: *Visuelle Datenanalyse mit Generative Gestaltung für eine ästhetische und abstrakte Informationsvisualisierung: Auswahl, Bewertung und Korrelation der interessierten Daten aus der Energiestatistik und Strukturdaten aus Köln.* Köln, Fachhochschule Köln Campus Gummersbach, Diss., 2014
- [Card u. Mackinlay 1997] CARD, S. K. ; MACKINLAY, J.: The structure of the information visualization design space. In: *VIZ '97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium*, 1997, S. 92–99
- [Chi 2000] CHI, Ed H.: A taxonomy of visualization techniques using the data state reference model. In: *Information Visualization, 2000. InfoVis 2000. IEEE Symposium on*, 2000. – ISBN 0-7695-0804-9, 69–75
- [Chung Wong u. Bergeron] CHUNG WONG, Pak ; BERGERON, R. D.: 30 Years of Multidimensional Multivariate Visualization. In: *Scientific Visualization*, S. 3–33
- [Deutsches Institut für Normung 2006] Norm DIN EN ISO 9241-110 2006. *Ergonomie der Mensch-System-Interaktion - Teil 110: Grundsätze der Dialoggestaltung (ISO 9241-110:2006); Deutsche Fassung EN ISO 9241-110:2006*
- [DIW u. EEFA 2014] DIW, Berlin ; EEFA: *Auswertungstabellen zur Energiebilanz für die Bundesrepublik Deutschland 1990 bis 2013.* Arbeitsgemeinschaft Energiebilanzen e.V. <http://www.ag-energiebilanzen.de>. Version: 2014. – [Zugriff: 25. Februar 2015]
- [GitHub, Inc. 2016] GITHUB, INC.: *GitHub.* <https://github.com/>. Version: 2016. – [Zugriff: 31. Juli 2016]
- [Goodall 2016] GOODALL, John: *d3.colorLegend.* <https://github.com/jgoodall/d3-colorlegend>. Version: 2016. – [Zugriff: 16. August 2016]
- [Haber u. McNabb 1990] HABER, Robert B. ; MCNABB, David A.: *Visualization Idioms: A Conceptual Model for Scientific Visualization Systems.* http://www2.cs.uni-paderborn.de/cs/ag-domik-static/visualisierung/vis-report/hypervis/concepts/idio_sh.htm. Version: 1990 (Visualization in scientific computing)

-
- [Hammerl, Michael, et. al. 2016] HAMMERL, MICHAEL, ET. AL. ; INSTITUTE OF SOFTWARE TECHNOLOGY & INTERACTIVE SYSTEM (Hrsg.): *InfoVis:Wiki*. Vienna University of Technology, 2016. – <http://www.infovis-wiki.net/> [Zugriff: 15. August 2016]
- [Harrower u. Brewer 2003] HARROWER, Mark ; BREWER, Cynthia A.: ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. In: *The Cartographic Journal* 40 (2003), Nr. 1, S. 27–37. <http://dx.doi.org/10.1179/000870403235002042>. – DOI 10.1179/000870403235002042
- [Keim u. a. 2006] KEIM, D. A. ; MANSMANN, F. ; SCHNEIDEWIND, J. ; ZIEGLER, H.: Challenges in Visual Data Analysis. In: *Tenth International Conference on Information Visualisation (IV'06)*, 2006, S. 9–16
- [Lee u. a. 2016] LEE, Sukwon ; KIM, Sung-Hee ; HUNG, Ya-Hsin ; LAM, Heidi ; KANG, Youn-ah ; YI, Ji S.: How do People Make Sense of Unfamiliar Visualizations?: A Grounded Model of Novice's Information Visualization Sensemaking. In: *IEEE transactions on visualization and computer graphics* 22 (2016), Nr. 1, S. 499–508. <http://dx.doi.org/10.1109/TVCG.2015.2467195>. – DOI 10.1109/TVCG.2015.2467195. – ISSN 1077–2626
- [Lehmann u. a. 2010] LEHMANN, Dirk J. ; ALBUQUERQUE, Georgia ; EISEMANN, Martin ; TATU, Andrada ; KEIM, Daniel ; SCHUMANN, Heidrun ; MAGNOR, Marcus ; THEISEL, Holger: Visualisierung und Analyse multidimensionaler Datensätze. In: *Informatik-Spektrum* 33 (2010), Nr. 6, S. 589–600. <http://dx.doi.org/10.1007/s00287-010-0481-z>. – DOI 10.1007/s00287-010-0481-z. – ISSN 0170–6012
- [Lücking, Heinz-Josef, et. al. 2016] LÜCKING, HEINZ-JOSEF, ET. AL.: *GISWiki*. 2016. – <http://giswiki.org/> [Zugriff am: 15. August 2016]
- [MacEachren u. a. 2004] MACEACHREN, Alan M. ; GAHEGAN, Mark ; AL. et.: Geovisualization for knowledge construction and decision support. In: *Computer Graphics and Applications, IEEE* (2004)
- [Node.js Foundation 2016] NODE.JS FOUNDATION: *Node.js*. <https://nodejs.org/>. Version: 2016. – [Zugriff: 6. Juni 2016]
- [npm, Inc. 2016] NPM, INC.: *Node Package Management (npm)*. <https://www.npmjs.com>. Version: 2016. – [Zugriff: 6. Juni 2016]
- [Oberascher 2013] OBERASCHER, Claudia: *Stromverbrauch und Stromverwendung der privaten Haushalte in Deutschland: Ergebnisse einer Studie im Auftrag von HEA, BDEW und EnergieAgentur.NRW, 2012*. <https://www.bdew.de/internet.nsf/id/daten-grafik-de>. Version: 2013. – [Zugriff: 21. März 2015]
- [Pretorius u. Van Wijk 2009] PRETORIUS, A. J. ; VAN WIJK, Jarke J.: What does the user want to see? What do the data want to be? Version: 2009. <http://dx.doi.org/10.1057/ivs.2009.13>. In: *Information Visualization 2009*. 2009. – DOI 10.1057/ivs.2009.13, S. 14
- [Reiterer u. Jetter 2013] REITERER, Harald ; JETTER, Hans-Christian: *Informationsvisualisierung: Anspruch und Charakteristika der Forschungsdisziplin*. Konstanz : Bibliothek der Universität Konstanz, 2013

- [Ren u. a. 2014] REN, Donghao ; HOLLERER, Tobias ; YUAN, Xiaoru: iVisDesigner: Expressive Interactive Design of Information Visualizations. In: *IEEE transactions on visualization and computer graphics* 20 (2014), Nr. 12, S. 2092–2101. <http://dx.doi.org/10.1109/TVCG.2014.2346291>. – DOI 10.1109/TVCG.2014.2346291. – ISSN 1077–2626
- [Robbins 2016] ROBBINS, Charlie: *http-server*. <https://www.npmjs.com/package/http-server>. Version: 2016. – [Zugriff: 6. Juni 2016]
- [Sarodnick u. Brau 2011] SARODNICK, Florian ; BRAU, Henning: *Methoden der Usability Evaluation*. 2. Aufl. s.l. : Verlag Hans Huber, 2011 http://sub-hh.ciando.com/book/?bok_id=240821. – ISBN 978–3–456–84883–9
- [Schumann u. Müller 2000] SCHUMANN, Heidrun ; MÜLLER, Wolfgang: *Visualisierung: Grundlagen und allgemeine Methoden*. Berlin and Heidelberg and New York and Barcelona and Hongkong and London and Mailand and Paris and Singapur and Tokio : Springer, 2000. – ISBN 9783540649441
- [Shneiderman 1996] SHNEIDERMAN, B.: The eyes have it: a task by data type taxonomy for information visualizations. In: *1996 IEEE Symposium on Visual Languages*, 1996, S. 336–343
- [Stadt Köln 2015] STADT KÖLN: *Bevölkerung Köln*. Offene Daten Köln. <http://www.offenedaten-koeln.de/>. Version: 2015. – [Zugriff: 19. März 2015]
- [Statistisches Bundesamt 2015] STATISTISCHES BUNDESAMT: *Bevölkerung: Deutschland, Fortschreibung des Bevölkerungsstandes*. <https://www.regionalstatistik.de>. Version: 2015. – [Zugriff: 24. März 2015]
- [Twitter, Inc. 2016] TWITTER, INC.: *d3.colorLegend*. <http://getbootstrap.com/>. Version: 2016. – [Zugriff: 6. Juni 2016]

Anhang

Anhang A

Daten Visualisierung

- In den jeweiligen Panel (Detail, Allgemeine Kennzahlen über Köln) die Kennzahlen und Zusatz Informationen anzeigen (Jahr, Verbrauchsverhalten und weitere Metadaten und Werte.)
- Die Höhe der Bevölkerungsanzahl auf der Stadtkarte nach Intensität der Farbsättigung codieren
- In der Kartenansicht sollten noch keine Datenwerte pro Stadtteil angezeigt werden. Keine Vermischung von Überblick und Detailkontext.
- Überblick Layer blendet für jede Stadtteil Fläche ein Kreisdiagramm für jede Haushaltsgröße ein, oder Stromverbrauch und Einwohneranzahl mit Farbcode. Jede Einzeldarstellung einer Haushaltsgröße ist eine alternative Darstellung zu dem Kreisdiagramm, wenn hier auch mit der Flächenfarbe codiert wird. Die Kenngröße der Wohndichte kann auch als Farbcode visualisiert werden. (Vielleicht mit den genannten Informationen eine neue Darstellung finden, die mit Wohndichte kombiniert wird. Zusammenhang zwischen Wohndichte und Stromverbrauch / Einwohneranzahl mit Farbcode lösen)
- Im Detail-Panel werden die Balken, Kreis und andere Datentyp angemessene Diagramme zur Information Darstellung erzeugt. Hier können konkrete Werte angezeigt werden. Welche Daten zeigt Punkt 3 in Ansichten / Elemente.

Interaktion

- Bei klick auf eine Fläche, wird diese mit farblich oder Stärke der Umrandung hervorgehoben. Im Detail Panel werden die Diagramme für Haushaltsgrößen, Verteilung, Stromverbrauch hinzugefügt. Ein weiterer klick entfernt es.
- Ein Dropdown oder Radiobutton für Datensatz Auswahl, die entsprechende Diagramme, Kennzahlen anzeigen. Mit dieser Interaktion wird zwischen den Datensatz gewechselt. Eine andere Interaktion könnte parallele Ansichten erlauben, um Überblick und Kontext zu behalten. (Überlegung die layer Option Auswahl mit Datensatz Dropdown zusammen führen.)

Features

- Bei Hover erscheint ein Tooltip, mit den Kennzahlen oder vergrößert die Grafik, die unter Darstellung Option ausgewählt wurde.
- Bei Implementierung von Jahren, können links und rechts Pfeile implementiert werden, die Ansichten zwischen Jahren wechseln. Nachteil, erlaubt keine Darstellung von ganzheitlichen Verlauf über die Jahre, diese muss extra als Zusatz Visualisierung hinzugefügt werden.

- Optional, ein Panel bzw. Toolbar für Historie, Anfragen Pfad, Design Option, Eingabe / Ausgabe der CSV, speichern, drucken

Verworfen

- Im Detail-Panel können weitere Bearbeitungsaufgaben implementiert werden. Mit dem hinzufügen und arrangieren mit drag and drop von Diagramme oder für Notizen
- Fokussierung bedeutet, dass bei klick oder eine andere Interaktion auf eine Fläche in der Karte, zu der entsprechenden Detailansicht eines Stadtteil im Detail-Panel bewegt wird. Wenn man ein extra Fenster für Detail erzeugt, funktioniert diese Interaktion nicht mehr? -> Gar nicht mehr nötig, weil im Detail-Panel das Säulendiagramm für Vergleiche verbessert wird.
- Sollte vielleicht doch ein extra Fenster für Detail geöffnet werden? Und bei der Detail Untersuchung braucht man die Karte? Hier nur die Daten und nichts anderes anzeigen. Mit extra Fenster kann das danach realisiert werden. Und der Nutzer hat die Möglichkeit selbst zu entscheiden, ob die Karte permanent sichtbar ist. Dabei kann er das Fenster relativ zu Monitorgröße anpassen und einen größeren Ausschnitt der Detailansichten erhalten. Wie er die Anzeigen positioniert, hängt von dem Bearbeitungsziel des Nutzers ab, wie er für sich ein effizientes Setting zu Bearbeitung und Untersuchung gestaltet. Man könnte gleich wie bei Browser integrierten Entwicklertool-Panel, eine Option für Loslösung des Panel aus dem Browserfenster anbieten.
- Weiterführend, kann man Fenster auch für Gruppen Ansichten für den jeweiligen Untersuchungsfokus verwenden, dafür müssen Instanzen der Fensteransichten möglich sein
- Scrollbalken sollten im Detail-Panel implementiert werden, weil die Stadtkarte als Ausgangspunkt und Orientierung und eventuell für Navigation, Fokussierung der einzelnen Stadtteile verwendet wird.
- Schließt man eine Detailansicht eines Stadtteil im Detail-Panel mit x wird die Hervorhebung in der Karte zurück genommen.
- Man könnte nach data of Interest (DOI) ein Dropdown bereit stellen, mit dem der Inhalt des Diagramms nach allgemeinen Stromverbrauch, oder Stromverbrauch nach Energieträger, Co²-Äquivalente, Haushaltsgrößen Anzahl anzeigen. Diese zusätzlichen Diagramme können untereinander eingeblendet werden.

Anhang B

Ansichten und Elemente

Beschreibung	Datensätze	Umgesetzt	Nicht Umgesetzt / Bugs
SVG Stadtkarte Köln, mit Bezirke, Stadtteile und Grenzen	Stadtteil Nummer, Bezirksnamen und Stadtteilnamen	Externe SVG Datei referenziert. Die Karte enthält von sich aus die Bezirksnamen und Stadtteilnamen als path. Im Stammdatensatz werden die Nummer, Stadtteilnamen assoziativ gespeichert.	
Panel mit allgemeinen Kennzahlen von ganz Köln und aus der Auswahl addierte Kennzahlen	Einwohneranzahl, Wohndichte, Fläche, Gesamt Stromverbrauch und Durchschnitt von Stromverbrauch / Personen-Haushaltsgröße	Die Kennzahlen werden aus den Stammdaten errechnet. Folgende 5 Werte werden addiert: Einwohneranzahl, Summe Stromverbrauch, Durchschnittsverbrauch Pro Kopf, Gesamtanzahl der Haushalte und Gesamt Fläche.	
Detail panel für ausgewählte Stadtteile mit Diagramme	Einwohneranzahl, Wohndichte (km ²), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamt Stromverbrauch (kWh) und Stromverbrauch je Haushaltsgröße	Die Ausgabe ist Abhängig von der gewählten Kartenmodifikation. Es wurden Balkendiagramme implementiert, die nach Auswahl in der Karte hinzugefügt oder entfernt werden.	Keine Kreisdiagramme implementiert, weil keine angemessene Darstellung der Daten
Checkboxes / Radiobuttons für Layer, auf der Stadtkarte wird für jedes Stadtteil ein bestimmter farbcodes eingeblendet	Stromverbrauch aller und je Haushaltsgrößen, Stromverbrauch pro Kopf, Einwohneranzahl, Wohndichte	Radiobuttons realisiert. Aus den Stammdaten wurden die relevanten Werte je Stadtteil errechnet und in Arrays gespeichert. Hinzu kommen die Werte für Summe der Haushalte.	
Eine Legende zu den farbcodes, der jeweils ausgewählten Layer Daten	Analog zu den Panel mit den Modifikationen werden dieselben Daten angebunden wie für die Choropleth Darstellung der Karte	Die Legende wurde unterhalb der Radiobuttons platziert und geben dieselben Daten wieder, wie die Choropleth Darstellung auf die Karte abgebildet worden ist.	

Daten Visualisierung

Beschreibung	Datensätze	Umgesetzt	Nicht Umgesetzt / Bugs
In den jeweiligen Panel (Detail, Allgemeine Kennzahlen über Köln) die Kennzahlen und Zusatz Informationen anzeigen	Einwohneranzahl, Wohndichte (km2), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamt Stromverbrauch (kWh) und Stromverbrauch je Haushaltsgröße	HTML Elemente Label und Input für benannte Ausgabefelder, der berechneten Werte.	
Die Höhe der Bevölkerungsanzahl auf der Stadtkarte nach Intensivität der Farbsättigung codieren	Selben Daten wie oben	Choroplethenkarte mit linearer Farbe mindest und maximal Werte codiert	Es wird nicht Quantitativ oder Ordinal codiert
In der Karten Ansicht sollten noch keine Datenwerte pro Stadtteil angezeigt werden. Keine Vermischung von Überblick und Detail Kontext	Stadtteilnamen und Werte der jeweiligen abzubildenden Daten	Standard Browser Tooltip in Choroplethenkarte und alle Balkendiagramme. Bei Hover wird der Exakte Wert und bei Karte zusätzlich der Stadtteilname eingeblendet	
Überblick Layer blendet für jede Stadtteil Fläche ein Kreisdiagramm	Haushaltsgröße, oder Stromverbrauch und Einwohneranzahl	statt dessen Tooltip	nicht umgesetzt
Im Detail panel werden die Balken, Kreis und andere Datentyp angemessene Diagramme zur Information Darstellung erzeugt.	Einwohneranzahl, Wohndichte (km2), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamt Stromverbrauch (kWh)	Die Visualisierung im Detail werden in Balkendiagramme präsentiert, da hier bessere Gruppierung der Werte je Stadtteil möglich ist.	Keine Kreidiagramme oder andere Diagramme

Detail Ansicht

Beschreibung	Datensätze	Umgesetzt	Nicht Umgesetzt
In der Detail Ansicht sollen eventuell im oberen Bereich die Kennzahlen und Informationen aufgeführt werden, von den ausgewählten Stadtteilen und zusammengefasst bzw. gerechnet werden.	Einwohneranzahl, Wohndichte (km ²), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamt Stromverbrauch (kWh)	Zeigt allgemeine Kennzahlen für ganz Köln und nach Auswahl zusammenaddiert.	Die Kennzahlen werden im Kennzahlen Panel ausgegeben
Ein Balkendiagramm kann für direkten Vergleich pro Haushalt ein Stadtteil aufnehmen und ist damit in der Lage nach Auswahl der Karte, die Stadtteile nach Haushalt zu gruppieren und nebenstehende Balken in der Gruppe Haushalt direkt zu vergleichen.	Daten nach Haushaltsgrößen aufgeteilt, in einperson, zweipersonen, dreipersonen, vierperson und fünf und mehrperson Haushalte	Im Detail Panel unter dem Reiter Vergleich der Auswahl nach Haushaltsgrößen visualisiert in den Modi Einwohneranzahl, Summe aller haushaltsgrößen, Gesamthaushaltsgrößen.	Die Balkengruppe wird nicht getrennt. Unterscheidung ist nur durch farblichen Rythmus zu erkennen
Das Diagramm nutzt die ganze Breite des rechten Fensterteils. Falls es über die Breite geht, werden Scrollbalken eingeblendet.	Stammdaten CSV je nach Kontext gefiltert	Nur im Reiter für Übersicht Stadtteile, da alle 86 Stadtteile visualisiert werden, braucht man eine gewisse breite.	werden nicht für Reiter Direktvergleich und Vergleich nach Auswahl eingesetzt
Die Balken sollten dynamisch transient sein und bei größenunterschied sich entsprechend skalieren. Bei Updates sollten die Balken dynamisch Skalieren.	Alle abzubildenden Daten, ist Visualisierungsabhängig	Skaliert nach Auswahl immer von null bis höchsten Wert. Nur unter Vergleich nach der Auswahl Reiter implementiert.	Keine Animation eingebaut, da die Art wie es implementiert wird nicht vollständig ermöglicht.
Die Informationen, die auf der Karte mit farbwerten abgebildet werden, können hier nochmal anders für Gruppenauswahl verwendet werden. Es erlaubt die Detail Ansichten als gesondert und der Fokus wird speziell nur auf die Auswahl gesetzt.	Alle abzubildenden Daten, ist Visualisierungsabhängig	Wird nur unter Direktvergleich verwendet. Hier wird entsprechend der maximal wert übernommen, unabhängig von Wertgröße.	Alle anderen Balkendiagramm darstellungen folgen einer bestimmten farbreihenfolge
Problem ist, wie die ganzen Stadtteile in der Gruppierung unterschieden und identifiziert werden. Durch Farben oder Namen, Zahlen, bei Interaktion (hover)?	Alle abzubildenden Daten, ist Visualisierungsabhängig	Farbliche unterscheidung der einzelnen Haushalte wurde festgelegt und in einer Legende definiert. Zusätzlich wurde ein hover effekt hinzugefügt der ein Hinweis auf interktion mit dem Tooltip geben soll. Namen der Stadteile werden auf der x-Achse angezeigt	Die Gruppen sollten ein outerpadding erhalten, da die Gruppierung nicht eindeutig zu erkennen ist.

Interaktion

Beschreibung	Datensätze	Umgesetzt	Nicht Umgesetzt
Bei Klick auf eine Fläche, wird diese mit farblich oder Stärke der Umrandung hervorgehoben. Im Detail Panel werden die Diagramme für Haushaltsgrößen, Verteilung, Stromverbrauch hinzugefügt. Ein weiterer Klick entfernt es	Einwohneranzahl, Wohndichte (km ²), Stromverbrauch pro Kopf (kWh), Summe aller Haushaltsgrößen und Gesamt Stromverbrauch (kWh)	Klickt man auf eine Fläche in der Choroplethkarte, wird diese farblich ausgefüllt in Violet. Damit aktiviert man die Fläche und erzeugt ein Plot im Detail Panel, sowie die Funktion zur Berechnung der Kennzahlen die addiert werden bei mehrfach Auswahl. Entsprechende deaktivierung entfernt den Das Objekt aus dem Diagramm.	Keine Umrandung, daher soll die Auswahl Farbe sich deutlich von allen anderen Choropleth Farben und diagramm Farben unterscheiden
Ein Dropdown oder Radiobutton für Datensatz Auswahl, die entsprechende Diagramme, Kennzahlen anzeigen. Mit dieser Interaktion wird zwischen den Datensatz gewechselt.		Der Kategorische Datensatz Wechsel oder sogenannte Modifikation wird im Panel mit Radiobuttons gewählt und die Karte neu geladen.	Ein Dropdown, zeigt nicht alle Modifikationsmöglichkeiten an und versteckt so Daten Kontexte, die noch entdeckt werden müssten
Eine andere Interaktion könnte parallele Ansichten erlauben, um Überblick und Kontext zu behalten.		Die Detail Ansichten werden in drei Reiter untergebracht (Direkt Vergleich nach Auswahl, Vergleich der Auswahl nach Haushaltsgrößen und Übersicht Stadtteile). Nur im mittleren Reiter wird die Auswahl nach Modifikationswechsel gelöscht.	Parallele Ansichten sind wegen den Monitorauflösungen beschränkt.

Layout

Beschreibung	Umgesetzt	Nicht Umgesetzt / Bug
Die Stadtkarte befindet sich im linken Teil des Fensters und sollte diese soweit ausfüllen.	ist umgesetzt, die Karte nimmt die Hälfte des Browserfensters ein bei genügend großer Auflösung, ansonsten die ganze breite im Mobilen Browser	hier gibt es noch probleme mit der Skalierung im Responsiven
Checkboxes für Layer befinden sich über dem Panel mit den allgemeinen Kennzahlen, da bei Auswahl der Modis unterhalb der allgemeinen Kennzahlen, die Daten zu den Modi angezeigt werden	Das panel mit den Modifikationen befindet sich oberhalb der Kennzahlen Panel. Zuerst von oben der Input, dann die Ausgabe Ansichten. Im responsiven wird das Modi Panel rechts oben platziert und die Ausgabe Ansichten unterhalb der Karte	die Modifikationen werden mit Radioboxes ausgewählt / geladen. Radioboxes, da diese für einfache Auswahl verwendet werden. Checkboxes würde mehrfachauswahl bedeuten.
Unterhalb der Checkbox sollte sich die Legende zu den Modi zugehörigen Farbcodes positioniert werden um bei der Modi Auswahl, mit den ersten Blick auf den Interaktionsbereich die Anpassung der Legende wahrzunehmen.	Die Legende befindet sich in der Modifikation Panel unterhalb der Radiobutton Auswahl. Mit dem Rahmen lässt sich eine Zugehörigkeit und der Kontext zu den Buttons und der Interaktion direkter in Verbindung setzen.	
Das Panel mit den allgemeinen Kennzahlen befindet sich beginnend in Leserichtung im rechten Bereich des Browserfenster, an der rechten Kante der Stadtkarte über der Detail Ansicht.	Die reinen Datenwerte der Kennzahlen und nach Auswahl errechneten Kennwerte befinden sich in der Box für Kennzahlen, darin werden diese links und rechts aufgeteilt. Die Box befindet sich wie beschrieben zwischen den Modifikations Panel und Detail Panel	
Das Detail Panel soll sich im rechten Browser Fensterteil befinden unter dem Panel mit den allgemeinen Kennzahlen.	Genauso wie die anderen Boxen befindet sich das Panel für die Detail Ausgabe in der rechten Hälfte des Browserfensters, unterhalb der Kennzahlen Box..	

Features

Beschreibung	Umgesetzt	Nicht Umgesetzt
Bei Hover erscheint ein Tooltip, mit den Kennzahlen oder vergrößert die Grafik, die unter Darstellung Option ausgewählt wurde.	Tooltip bei Hover über die Balkendiagramme und auf der Karte werden angezeigt.	Die Grafik wird nicht vergrößert bei Hover
Bei Implementierung von Jahren, können Links und rechts Pfeile implementiert werden, die Ansichten zwischen Jahren wechseln. Nachteil, erlaubt keine Darstellung von ganzheitlichen Verlauf über die Jahre, diese muss extra als Zusatz Visualisierung hinzugefügt werden		Es wurden nur die Daten für das Jahr 2012 gezeigt. Daten über Jahre hinweg wurde nicht implementiert, da dies eine prototypische Entwicklung ist und nicht in der Priorität steht.
Optional, ein panel bzw. toolbar für Historie, Anfragen Pfad, Design Option, Eingabe / Ausgabe der CSV, speichern, drucken	Eine Art Historie im weitesten Sinne kann die Detail Ansicht Direkt Vergleich nach Auswahl geben, da hier über die Modies hinweg jedes angewählte Stadtteil behält, bis man es bewusst löscht durch doppeltes klicken aus der Modifikation heraus oder durch reseten.	Anfragen Pfad, Design Option, Eingabe / Ausgabe der CSV, speichern, drucken sind nicht implementiert.

Anhang C

Nicht-Experten Test

Aufgabe	Erfolg- oder Misserfolg	Verbesserungsvorschlag
1. Wähle ein Stadtteil nach Interesse aus und bewerte die Datenausgaben der jeweiligen Ansichten.	1. Bei Hover Effekt auf die Stadtteile der Karte, erwartet der Nutzer eine Interaktion bei einem Klick. Es fallen sofort die Tabs auf und das bei der Karten Interaktion der Inhalt des Tab1 geladen wird. 2. Explorativ werden die anderen Tabs aktiviert und die Inhalte verglichen. 3. Bei genauen betrachten der Diagramme in den jeweiligen Ansichten, versteht der Nutzer die Bedeutung. Kritisch sind die Metriken, die Formatierung k, m und M führen zu Fehlinterpretationen. Bei Kennzahlen nach Auswahl wird zuerst vermutet, dass je nach zuletzt geklickte Auswahl die Werte angezeigt werden. 4. Hauptsächlich wird die Auswahl Funktion dazu verwendet, die Stadtteil Diagramme in Tab1 zu sammeln, um sie zu vergleichen	<ul style="list-style-type: none"> • Formatierung der Metriken sollte noch angepasst werden auf deutsche Bezeichnungen. • Titel "Kennzahlen der Auswahl" muss umbenannt werden in "Summe der Auswahl" • Alternative für Kennzahlen nach Auswahl wäre ein Tooltip bei hover in der Tab1 oder Tab2 Ansicht
2. Wähle mehrere Stadtteile nach Interesse aus und vergleiche die Datenwerte der Gruppen nach der Auswahl. Bewerte die Ausgabe.	1. siehe Aufgabe 1 2. Durch explorative Interaktion mit der Karte, hat der Nutzer die Funktion erkannt und nutzt die Funktion Auswahl. 3. Jedoch erst nach dem Hinweis, dass es eine Legende gibt in Tab2, versteht der Nutzer die Bedeutung. Diese wurde bei einer HD Auflösung in der Ansicht verdeckt und es musste darauf hingewiesen werden. Die Überschrift in Tab2 ist nicht sprechend genug. 4. Während dem Vergleich der Visualisierungen, erlangt der Nutzer Wissen und entwickelt Fragen. Er erkennt Widersprüche in den Datensätzen, die es zu begründen gilt.	<ul style="list-style-type: none"> • Verdeckte Legende bei HD Auflösung, Legende sichtbar nach links oder nach oben setzen. • Die Tab2 und Tab3 Ansichten werden nicht konsequent in den vergleichen angewendet. Die unterschiedlichen Inhalte und Sichtbarkeit der Inhalte stellen ein Hinderniss dar.
3. Leere die Ansichten und die Auswahl	1. Sofort erkannt 2. Bei erster Gelegenheit in Aufgabe 1 verwendet. Um Übersicht wiederherzustellen bei explorativer Nutzung 3. -- 4. Als nützliches Tool wird es schon bei den ersten beiden Aufgaben oft verwendet.	
4. Wähle eine der Modifikatoren bzw. Informationskontexte für die Karte. Wähle die Stadtteile nach Interesse und bewerte die Ausgaben der jeweiligen Ansichten.	1. Interaktion mit Karten Modifikation Radiobuttons sofort als Interaktionselement erkannt und gewählt (Gesamt Stromverbrauch). Stadtteil Auswahl zuvor in Aufgabe 1 und 2 erlernt. 2. Tab1 für den Vergleich der Auswahl bewertet und nach höchsten Werten verglichen. 3./4. Durch die Auswahl nach Stichproben der Dauer und Anzahl der Proben, festigt sich das Verständnis der Visualisierung und den erzielten Ergebnisse.	
5. Wähle die Stadtteile, ganz unabhängig in welchen Kontext die Karte aktiviert ist. Entferne einzelne Stadtteile aus der Auswahl.	1. Die Interaktion Auswahl in der Karte aufheben wird durch Vorerfahrung aus anderen Visualisierungen sofort erkannt. In Verbindung mit der Ansicht in Tab 1 ist der vice versa Rückschluss vorhanden. Durch Auswahl hinzufügen und Auswahl aufheben entfernen der Diagramme. 2. Der Nutzer entfernt die Auswahl durch Intraktion in der Karte 3. -- 4. Durch die nicht-lineare Aufhebung der Auswahl, kann der Nutzer die energisch auswählen und hat keine Scheu Fehler zu machen, da eine Fehlauswahl sofort zurück	<ul style="list-style-type: none"> • Die Auswahl in der Karte sollte beibehalten werden, wenn die Modifikation gewechselt wird

	genommen werden kann. Es ermöglicht ihn zu filtern und zu sortieren.	
6. Wähle und sammle die Werte der Stadtteile in verschiedenen Kontexten und bleib in der Ansicht "Gesammelte Auswahl". Entferne die Auswahl beliebig einzeln aus der Ansicht	<p>1. Auswahl und Wechsel zwischen den Modifikationen stellt sich als kein Problem dar. Siehe vorherige Aufgaben. Ebenso Auswahl aufheben.</p> <p>2. --</p> <p>3. Der Nutzer hat erkannt, dass bei Modiwechsel, die Diagramme in Tab1 bestehen bleiben, aber die Kartenauswahl verloren geht. Bei dem Versuch die Auswahl nochmal durchzuführen, wurde beobachtet, dass die Auswahl hinzugefügt wird und bei aufheben, wird das Diagramm und die Kartenauswahl ganz entfernt. Erkennt, dass das Entfernen nur in den jeweiligen Modifikationen gelingt.</p> <p>4. Effizient ist der Zusammenhang der Interaktion, durch Feedback in der Karte und entsprechendes Verhalten in der Ansicht.</p>	Hier wird eine unvollständige Implementierung getestet, die aber korrigiert werden muss. Der Fehler ist, dass man eine Auswahl aus einer vorigen Modifikation neu auswählen (hinzufügen) muss, um sie komplett für die Modifikation zu entfernen. Dieser Bug steht in Zusammenhang mit der Auswahl bei Modifikationswechsel speichern und wieder Anzeigen, wenn zurück gewechselt wird.
7. Wähle und sammle die Werte der Stadtteile in verschiedenen Kontexten und bleib in der Ansicht "Direkt Vergleich nach Auswahl". Entferne die Auswahl komplett aus der Ansicht.	<p>1. Siehe Aufgabe 7 und Aufgabe 3</p> <p>2. Der Nutzer verwendet den Reset Button, um bei Fehler und im explorativen Vergleich, neu beginnen zu können.</p> <p>3. --</p> <p>4. Der Reset Button ist ein Werkzeug, was der Nutzer häufig verwendet und schnell zu einer neu Abfrage der Daten verleitet.</p>	
8. Nachdem man sich mit der Funktionsweise der Visualisierung vertraut gemacht hat, versuche die Fragen 1 bis 16 zu beantworten.	Bemerkungen des Nutzers: In den Balkendiagrammen ist nicht der gewollte Effekt, man muss darauf hinweisen, länger das Feld länger zu fokussieren für das Tooltip.	Die Diagramme sollten in dem Fall ein Tooltip erhalten, was bei Klick die einzelnen Kennzahlen wie in der Box darüber ausgibt.

Frage	Handlungsschritte und Antwort
1. Wieviele Einwohner hat Köln?	Sofort erkannt in der Kennzahlen Köln Spalte (5s)
2. Wieviele Einwohner haben die einzelnen Stadtteile?	Es wurde die Kennzahl Ausgabe für die jeweiligen Auswahl für die Antwort verwendet. (26s)
3. Wie sind die Anteile der Haushaltsgrößen in Köln und in den Stadtteilen verteilt?	In der Modifikation Stromverbrauch PHH durch Auswahl eines Stadtteils verglichen. Die Frage konnte teils beantwortet werden, es wurde die Modifikation Stromverbrauch PHH falsch interpretiert. (42s)
4. Welches Stadtteil ist das Größte, welches das Kleinste?	Der Nutzer wusste zunächst nicht, wie man die Frage beantwortet. Erste Versuch war die Antwort über die Karte zu schätzen. Die Auswahl und der Vergleich der geschätzten größten und kleinsten wurden gesammelt und in Tab1 verglichen. Am Ende wurde erkannt, dass die Werte bei der Einzel-Auswahl über die Kennzahlen abgelesen werden kann. (2,45min)
5. Haben Stadtteile mit einer größeren Fläche eine niedrigere Wohndichte?	Anhand der Karte wurde erkannt, dass große Flächen heller sind. Bestätigt wird dies durch den Vergleich der Stichproben und Auswahl Ansicht Tab1 (45s)
6. Haben diese Stadtteile, dann einen höheren Mehrpersonen-Haushalt Anteil?	Zusätzlicher Vergleich mit Stromverbrauch von Mehrpersonen-Haushalt. Es wurde erkannt, dass die Haushaltsgrößen nicht im Zusammenhang mit der Wohndichte stehen. Hier wurde wieder anstelle der Übersicht für Haushaltsgrößen, der Stromverbrauch je nach Haushaltsgröße heran gezogen. (45s)
7. Haben Stadtteile mit einer kleinen Fläche eine höhere Wohndichte?	Es wurden alle Stadtteile in der Karte Wohndichte gewählt mit kleiner Fläche. Die Antwort ist, es kommt auf die Stadtteile an. Überwiegend gilt, dass kleine Stadtteile eine höhere Wohndichte haben im Durchschnitt. (40s)

8. Haben diese Stadtteile, dann einen höheren Einpersonen-Haushalt Anteil?	Auswahl in der Gesamt Haushaltsgrößen Ansicht Tab2 und Vergleich auf der Karte. Neustadt-Süd und Mauenheim als repräsentanten ausgewählt. davon zusätzlich der ein-personen Stromverbrauch und fünf-personen Stromverbrauch hinzugefügt. Erkannt wurde das einpersonen Haushalte generell einen höheren Anteil haben. Im Vergleich der Choropleth Karte wird dies ersichtlich. (5,13s)
9. Welche Stadtteile verbrauchen weniger und welche mehr Strom?	Die Stromverbrauch pro Kopf Modifikation wird gewählt und anschließend der Gesamt Stromverbrauch und anhand der Karten vergleicht. Es wird erkannt, dass Mühlheim ein Ausreißer in der Korrelation ist. (1,09s)
10. Haben Stadtteile mit einer kleinen Fläche einen höheren Energiebedarf?	Vergleich über die Karte gezogen in den Modifikationen, Gesamt Stromverbrauch und Stromverbrauch je PHH und ebenso Stromverbrauch pro Kopf. Die Innenstadt mit den kleineren Flächen mit den höchsten Verbrauch erkannt. (1.02min)
11. Haben Stadtteile mit einer großen Fläche einen niedrigeren Energiebedarf?	Aus Gesamt Stromverbrauch Auswahl große Stadtteile im Vergleich zu pro Kopf Ansicht. Die Flächengröße auf der Karte vergleichen zum Gegenteil mit kleinen Stadtteilen. Es wurden die Kennzahlen der Auswahl gemerkt, sowie Größe der Stadtteile. Der Nutzer bemerkt, das die Antwort hier zum Widerspruch der aus Frage 10 ist. Der Stromverbrauch von großen Stadtteilen ist höher. (5,15min)
12. Welches Stadtteil verbraucht im Durchschnitt weniger Energie und welche mehr?	Unter Gesamt Stromverbrauch im Tab3 Übersicht der Stadtteile, wurde auf der Karte die Innenstadt verglichen. Es wird nach dem Stadtteil mit dem niedrigsten und den höchsten Stromverbrauch gesucht. (1,30min)
13. Ist der pro Kopf Verbrauch in Stadtteilen mit einer höheren Anzahl von Mehrpersonen-Haushalten höher?	Zum Vergleich in Tab1 werden nach Stichproben die Stadtteile nach Stromverbrauch pro Kopf und Stromverbrauch fünf-PHH gesammelt. Nach Vergleich in Tab1, lautet die Antwort, Nein, die Anzahl der Mehrpersonen Haushalte ist nicht höher. (1,40min)
14. Wieviele Einpersonenhaushalte und Mehrpersonenhaushalte muss ein Stadtteil haben, um etwa denselben Verbrauch zu haben?	Vergleich zwischen Stromverbrauch ein-PHH und fünf-PHH mit dem Stromverbrauch pro Kopf auf der Karte. Repräsentativ wurden Dünwald und Rodenkirchen gewählt. Nach dem Choropleth wird nach Stadtteilen mit der gleichen Farbstärke gewählt und in Tab1 und Tab2 verifiziert. Die Antwort ist Stadtteil abhängig. Es wurde erkannt, dass einpersonen Haushalte mehr Verbrauchen. (2,39min)
15. Ab welcher Mehrpersonen-Haushaltsgröße, ist der Verbrauch größer als bei Einpersonen-Haushalten?	Mit Referenz zu Frage 14 werden Rodenkirchen und Dünwald wieder verglichen. Mit Stromverbrauch nach ein-PHH und zwei-PHH, sowie pro Kopf verbrauch wird wieder verglichen in der Tab1 Ansicht. die Werte werden vergglichen und die Antwort geschätzt. Der Einpersonen-Haushalt Verbrauch ist ca. 4x größer als ein Mehrpersonenen-Haushalt. (2,28min)
16. Hängt die Flächengröße der Stadtteile, mit der Verteilung privater Haushalte und den Verbrauch zusammen?	Es wurden Stichproben aus Einwohneranzahl und Stromverbrauch aller PHH ausgewählt. Der Vergleich erfolgt in Tab1. Zusätzlich wird die Wohndichte hinzugefügt und verglichen. In der Übersicht alle Haushalte aller Stadtteile, wurde wieder erkannt, dass ein-PHH dominieren. Eil als Stichprobe zeigt allerdings nicht unbedingt dominanter. Zuletzt wird noch der Vergleich von pro Kopf Verbrauch, zu Wohndichte und Stromverbrauch der PHH Größen. Indirekt hängt die Flächengröße zusammen, aber je mehr Stromverbrauch, umso mehr Einpersonen-Haushalte sind vorhanden. Hat aber nicht mit den Flächen zu tun, es hängt von Stadtmitte und Außenstadtteile zusammen. (10,44min)

Bemerkungen: Sortierung von PHH für alle Stadtteile Ansicht, Karte und Balken Selektion der Stadtteile sollte für Orientierung verlinkt werden.

Experten Test

Aufgabe	Erfolg- oder Misserfolg	Verbesserungsvorschlag
1. Wähle ein Stadtteil nach Interesse aus und bewerte die Datenausgaben der jeweiligen Ansichten.	<ol style="list-style-type: none"> 1. Ja, die Elemente sind aus anderen Anwendungen bereits bekannt. 2. Der Nutzer versucht eine Abschätzung im Vergleich des Stromverbrauchs der ein-PHH und zwei-PHH eines Stadtteils mit dem Gesamtverbrauch und Zusammenhang mit der Einwohneranzahl zu überschlagen, und verifiziert die Daten mit den Verbrauch pro Kopf 3. Er betrachtet die Ausgaben, liest sie korrekt und interpretiert die Ausgaben korrekt. 4. Die Suche nach Informationen wird explorativ durchgeführt. 	
2. Wähle mehrere Stadtteile nach Interesse aus und vergleiche die Datenwerte der Gruppen nach der Auswahl. Bewerte die Ausgabe.	<ol style="list-style-type: none"> 1. Ja, der Nutzer wählt bewusst aus einem Kontext und aktiviert Tab2 2. Ja, gezielte Nutzung der Interaktionselemente 3. Er versteht die Informationen und kommt zum richtigen Schluss 4. Vergleicht zur Konsolidierung der Interpretation zwei Stadtteile 	
3. Leere die Ansichten und die Auswahl	<ol style="list-style-type: none"> 1. Es wird die Auswahl in der Karte aufgehoben 2. Bei einer größeren Auswahl, würde er den Reset Button nutzen, bei einer kleinen Karten Auswahl, entfernt er diese aus der Karte. 3. -- 4. -- 	
4. Wähle eine der Modifikatoren bzw. Informationskontexte für die Karte. Wähle die Stadtteile nach Interesse und bewerte die Ausgaben der jeweiligen Ansichten.	<ol style="list-style-type: none"> 1. / 2. Ja, wie zuvor werden die Elemente wie vorgesehen erkannt und genutzt. Gewählt werden Gesamthaushaltsgrößen 3. In Tab1 wird die Bedeutung erkannt, in Tab2. Es wird korrekt interpretiert, in Tab3 Gab es Probleme das entsprechend ausgewählte Stadtteil zu finden und erkennt Tab2 als geeigneter. 4. Jedoch wird der Nutzen einer Übersicht des Tab3 erkannt und zieht das Wissen aus Tab2 mit Tab3 in Vergleich und bekommt eine Schätzung der Größen im Bezug zur Gesamt Verteilung der Stadt. Durch ein höheres Verständnis wird die Effizienz gesteigert. 	<ul style="list-style-type: none"> Die Übersicht Tab3 erschwert die gezielte Suche nach Stadtteilen. Die Stadtteile sollten sortierbar sein und nach Hover/Auswahl Interaktion hervorgehoben werden
5. Wähle die Stadtteile, ganz unabhängig in welchen Kontext die Karte aktiviert ist. Entferne einzelne Stadtteile aus der Auswahl.	<ol style="list-style-type: none"> 1. Siehe Aufgabe 3 2. Der Nutzer sucht und wählt nach Stärke der Choroplethkarte aus. Die Legende wird generell nicht sehr beachtet. Entfernen der Auswahl kann effizienter genutzt werden durch den Reset Button. Siehe Aufgabe 3 3. Die Choropleth Darstellung ist schnell und einfach zu interpretieren. 4. Der Nutzer merkt an, dass er viel über Köln gelernt hat. 	<ul style="list-style-type: none"> Die Legende könnte prominenter platziert werden. Z.B. über der Karte
6. Wähle und sammel die Werte der Stadtteile in verschiedenen Kontexten und bleib in der Ansicht "Gesammelte Auswahl". Entferne die Auswahl beliebig einzeln aus der Ansicht	<ol style="list-style-type: none"> 1. Siehe Aufgabe 3. Da bei der Auswahl in Tab1 sofort das Ergebnis der Interaktion zu sehen ist und entsprechende Aufneung zu erkennen ist das die Auswahl aus Tab1 gelöscht wird. 2. Der Reset Button wird gerne verwendet, da es schneller ist als Rückläufiges wählen. 3. Der Nutzen als Historie und Überblick einzelner Werte über die Kontexte hinweg wird erkannt. und wird auch als Speicher der Auswahl verwendet 4. Kritisch: Die Modifikatoren die Farbcodes erst bei der Auswahl rendern 	<ul style="list-style-type: none"> Es sollte eine Legende oder eine farbliche Differenzierung der bestehenden Farben zu denjeweiligen Modifikatoren in der GUI ersichtlich sein. Man könnte die Schrift, oder eine Legende oder sonstige farbliche Abhebung einbauen.

	und bei Wechsel geändert werden. Hilft die Farbe aus der Tab1 der jew. Datensätze zuzuordnen. Der Nutzer kennt diese aber nicht auswendig und die Beschriftung ist schwer zu lesen.	
7. Wähle und sammle die Werte der Stadtteile in verschiedenen Kontexten und bleib in der Ansicht "Direkt Vergleich nach Auswahl". Entferne die Auswahl komplett aus der Ansicht.	1. Ja, wie in Aufgabe 6 und 3. Der Nutzer verwendet die Interaktion angemessen 2. Ja, es wird entschieden ab wann es angemessen ist 3. -- 4. Ja, es wird als geeignetes Werkzeug anerkannt	
8. Nachdem man sich mit der Funktionsweise der Visualisierung vertraut gemacht hat, versuche die Fragen 1 bis 16 zu beantworten.	Der Nutzer hat die Daten explorativ kennen gelernt und ein mentales Modell entwickelt, sowie die Interaktionsmöglichkeiten und Ausgabe Felder hat er verstanden und werden berücksichtigt.	

Frage	Handlungsschritte und Antwort
1. Wieviele Einwohner hat Köln?	Wie beim Nicht-Experten, sofort gewusst wo es abzulesen war (6 sek)
2. Wieviele Einwohner haben die einzelnen Stadtteile?	Bei der Auswahl über die Choroplethkarte bemerkt der Nutzer, dass ein Tooltip erscheint und erkennt, dass dies die Werte der aktuellen Modifikation Einwohnerzahl sind. (22 sek)
3. Wie sind die Anteile der Haushaltsgrößen in Köln und in den Stadtteilen verteilt?	Es wird erst nach der entsprechenden Modifikation gesucht und Stadtteile ausgewählt, durch ausprobieren wird versucht die Frage durch die Kennzahlen, mit Vergleich der Diagramme in Tab2 und Tab3 zu beantworten. Mit Gesamt Haushaltsgrößen und Tab3 kann die Verteilung aller Stadtteile verglichen werden. (31 sek)
4. Welches Stadtteil ist das Größte, welches das Kleinste?	Es wird die Wohndichte Modifikation gewählt und in das Tab3 gewechselt. Die Ansicht zeigt Einwohner pro qkm wird festgestellt, deswegen wird über die Karte nach der Antwort gesucht und nach geschätzter größter und geschätzter kleinster Stadtteil gewählt. Tab3 ist hier Kritisch, da nicht erwartet wird die Fläche anzulesen zu können. Mit der Tooltip Methode wird versucht zu merken welche Schätzungsweise die Frage beantworten könnte. Anschließend wird erkannt welche Funktion bei der Auswahl der Stadtteile die Kennzahlen Ausgabe hat und darüber die Frage beantwortet. (3,49 min)
5. Haben Stadtteile mit einer größeren Fläche eine niedrigere Wohndichte?	Mit der Wohndichte Modifikation und Abschätzung der Choroplethkarte wird die Frage beantwortet. (35 sek)
6. Haben diese Stadtteile, dann einen höheren Mehrpersonen-Haushalt Anteil?	Es wird in Gesamt Haushaltsgröße gewechselt und große Stadtteile werden zum Gegenvergleich gewählt. Der Nutzer ist unentschlossen ob Tab2 oder Tab3 die richtige Antwort gibt. (36 sek)
7. Haben Stadtteile mit einer kleinen Fläche eine höhere Wohndichte?	Es wird wieder in die Modifikation Wohndichte gewechselt und die Choroplethkarte verglichen und die geschätzten Stadtteile gewählt. In Tab2 wird verglichen. (40 sek)
8. Haben diese Stadtteile, dann einen höheren Einpersonen-Haushalt Anteil?	Aus voriger Auswahl gilt in Tab1 als Referenz zur Beantwortung der Frage. Es wird in Tab2 verglichen. Der Nutzer erkennt das generell Ein-PHH überwiegen und bestätigt dies in der Übersicht der Stadtteile. (1,56 min)
9. Welche Stadtteile verbrauchen weniger und welche mehr Strom?	Wechselt in Modifikation Gesamt Stromverbrauch. Betrachtet die Karte und sieht schnell die Antworten. Zur Verifizierung wird eine Auswahl getroffen in pro Kopf Darstellung und in Tab2 nach den mit wenigsten Verbrauch verglichen. Der Inhalt in Tab2 für Gesamtverbrauch ist verwirrend, aber es werden die Signifikanten Werte erkannt. (4,17 min)
10. Haben Stadtteile mit einer kleinen Fläche einen höheren Energiebedarf?	Es wird aus Gesamtverbrauch Modifikation die entsprechend kleinen Stadtteile gewählt und zum Gegenvergleich Große Stadtteile und in Tab1 verglichen. Es werden Ausreißer erkannt, jedoch wird

	geantwortet, dass im allgemeinen im 4 Millionen Bereich der Bedarf ist. Es kommt auf die Stadtteile an, aber generell nein. (2,42 min)
11. Haben Stadtteile mit einer großen Fläche einen niedrigeren Energiebedarf?	Der Vergleich erfolgt über dieselbe Auswahl aus Frage 10. Die Antwort ist gleich. (20 sek)
12. Welches Stadtteil verbraucht im Durchschnitt weniger Energie und welche mehr?	Es wird der Stromverbrauch in der Modifikation pro Kopf bewertet (36 sek)
13. Ist der pro Kopf Verbrauch in Stadtteilen mit einer höheren Anzahl von Mehrpersonen-Haushalten höher?	Aus der Modifikation Stromverbrauch ein-PHH und Fünf-PHH werden Stichproben gewählt, ebenso aus pro Kopf Verbrauch. In Tab2 hierzu ist wieder verwirrende Ausgabe, da alle Werte pro Stadt gleich sind. In Tab1 werden die Auswahl verglichen. Probleme die Stromverbrauch pro PHH zu unterscheiden da der Farbton der Gleiche ist, trotz Beschriftung. Antwort wird abgeschätzt (3,16 min)
14. Wieviele Einpersonenhaushalte und Mehrpersonenhaushalte muss ein Stadtteil haben, um etwa denselben Verbrauch zu haben?	Der Stromverbrauch für Ein-PHH und Fünf-PHH werden wieder gewählt für den Vergleich, dann in Tab2 begutachtet. Aus der Modifikation Gesamtverbrauch wird ein Stadtteil als Beispiel gewählt und aus Einwohneranzahl und Gesamt Haushaltsgrößen hinzugenommen. Dellbrück stellte sich als schlechtes Beispiel heraus Begründet sich daher, dass die Farbwerte nicht genau abgeschätzt werden können. Es wird ein anderes Stadtteil nach Stromverbrauch PHH gewählt und verglichen. Und geschätzt anhand Tab1, dass ca 7x mehr mehrpersonenhaushalte benötigt werden. (4,05 min)
15. Ab welcher Mehrpersonen-Haushaltsgröße, ist der Verbrauch größer als bei Einpersonenhaushalten?	Selbe Antwort wie zu Frage 14. Es wird Mülheim zum Vergleich hinzugenommen und ein effizienterer Weg zur Lösung gesucht. (2,47 min)
16. Hängt die Flächengröße der Stadtteile, mit der Verteilung privater Haushalte und den Verbrauch zusammen?	In der Modifikation Gesamt Haushaltsgrößen wird die Übersicht der Stadtteile betrachtet und bewertet. Anschließend der Stromverbrauch pro Kopf Karte und in Vergleich zur Übersicht Tab3. Ein Beispiel wird in der Karte gewählt aus Gesamt Haushaltsgröße und Stromverbrauch der Haushalte mit den selben Stadtteilen und im diese in Detail in Tab2 betrachtet, anschließend in Tab1 die Auswahl verglichen. Urteil ist, dass die Flächengröße nicht mit den Stromverbrauch zusammen hängt. (3,14 min)

Anhang D

Ideale Handlungsabfolge

Frage 12 Lösung:

Wähle die Modifikation Stromverbrauch pro Kopf, wechsel in die Ansicht Übersicht Stadtteile und vergleiche die Werte mit den höchsten Werten. Der Durchschnittswert sagt aus welche Stadtteile die meisten Mehrpersonenhaushalte hat. Achtung, der Stromverbrauch der einzelnen Haushaltsgrößen sind nicht repräsentativ für den pro-Kopf-Verbrauch je Stadtteil. Dann in die Ansicht Direktvergleich wechseln. Hier sieht man, dass der Durchschnittswert nicht die maximale Messhöhe erreicht. Also die Frage mit einem Nein zu beantworten ist.

Für den Gegenvergleich welche Stadtteile den höchsten pro Kopf Verbrauch aufweist, wählt man die Modifikation Gesamt Stromverbrauch und wählt anhand der Übersicht Stadtteile und der Karte die 5 mit den höchsten Werten, wechselt dann in die Modifikation Summe aller Haushaltsgrößen und erkennt auch dort dieselben höchstwerte zu den zuvor ausgewählten Stadtteilen. Anschliesend wieder in die Modifikation Stromverbrauch pro Kopf und wählt nochmals genau diese zuvor gewählten Ausreißer. Man erkennt dass, hier der größte Durchschnitt pro Kopf Verbrauch ist. Die Höhe des pro Kopf Verbrauchs ist also unabhängig von der Menge der Haushaltsgrößen. Auch bei durchklicken der Stromverbrauch Modi je Haushalt lässt sich die Menge des Stromverbrauchs bei Mehrpersonen Haushalte eher sinkend anzeigen. Sowie in der Kartenansicht bei Stromverbrauch pro Kopf die Stadtteile mit den höchsten Einpersonen Haushalten einen höheren Verbrauch zeigen.

Frage 14 Lösung:

Man wählt die Modifikation Gesamt Stromverbrauch und vergleicht die Werte im Reiter Übersicht. Suche nach repräsentativen Stadtteilen mit möglichst gleich hohen Energie Werten. Kalk Zollstock und Altstadt-Nord lassen sich gut vergleichen. Diese wählt man in der Karte aus für den direkt Vergleich in Tab1. Im mittleren Reiter Tab2 kann man sehen wie hoch der Verbrauch der einzelnen Haushaltsgrößen ist und gibt schon mal einen Hinweis der Haushaltsgrößen. Unter Summe der Haushalte im zweiten Reiter bekommt man den Vergleich der Haushaltsgrößen der entsprechenden Stadtteile. Eine eindeutige Antwort gibt es hier nicht. Die Antwort ist eine allgemeine Schätzung. Da für die Dimensionen der jeweiligen Stadtteile, die Größen der Haushalte variieren.

Frage 15 Lösung:

Diese Frage lässt sich auch mit der Lösung aus Frage 14 beantworten. Dort wird nach Verbrauchswerten gesucht die gleich sind und die Ansicht für die jeweiligen Stadtteile nach Haushalte gesplittet verglichen. Von geschätzten Ergebnis aus kann man grob sagen ab welcher Mehrpersonenhaushaltsgröße der Verbrauch ansteigt, wenn man die Auswahl im Detail direkt vergleicht. Lässt sich auch mathematisch lösen mit Dreisatz.

Frage 16 Lösung:

Stichprobenartig wählt man unter der Modifikation Wohndichte repräsentative Stadtteile, die sich von der Flächengröße und Wohndichte unterscheiden. In diesen Fall Neustadt-Süd, Mauenheim, Mülheim und Dellbrück. Nun wählt man unter Summe aller Haushalte die Vorauswahl der Stichprobe aus und wechselt in die Ansicht Vergleich der Auswahl Tab1, in der alle fünf Haushaltsgrößen visualisiert werden. Der Vergleich zwischen Mauenheim und Neustadt-Süd zeigt, dass die Fläche und die Anzahl der Haushalte bzw. der Einwohner die Wohndichte zusammenhängen. Ein Flächenmäßig großes Stadtteil hat entsprechend mehr Einwohner für die selbe Stärke der Wohndichte, wie ein kleines Stadtteil, mit ebenso entsprechend vielen Einwohner, um auf dieselbe Wohndichte zu kommen. Der Verbrauch korreliert allerdings nicht mit der Wohndichte, wenn man in die Visualisierung Gesamt Stromverbrauch nach Auswahl vergleicht. Der Verbrauch korreliert mehr mit der Einwohneranzahl bzw. Anzahl der Haushalte.

Mülheim und Dellbrück halten die Regel ein, je kleiner die Fläche, desto größer die Wohndichte, wobei die Anzahl der Haushalte auch kleiner wird, je größer das Stadtteil ist, was überraschenderweise nicht in die Hypothese passt und divergent zu sein scheint. Es ist also ein recht schwierig zu beantwortende Frage zu sein, da es mehrere Aspekte und Antworten zu geben scheint, die nicht eindeutig sind. Grundsätzlich bestehen Zusammenhänge, aber man kann diese nicht als Regel über alle Stadtteile Kölns festlegen.

Eidesstattliche Erklärung

Ich versichere, die von mir vorgelegte Arbeit selbständig verfasst zu haben.

Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer entnommen sind, habe ich als entnommen kenntlich gemacht. Sämtliche Quellen und Hilfsmittel, die ich für die Arbeit benutzt habe, sind angegeben.

Die Arbeit hat mit gleichem Inhalt bzw. in wesentlichen Teilen noch keiner anderen Prüfungsbehörde vorgelegen.

Gummersbach, 17. August 2016

Dorothea Buck