

Hochschule Merseburg (FH)
University of Applied Sciences



Fachbereich Wirtschaftswissenschaften und Informationswissenschaften

Masterarbeit

**Entwicklung eines Konzepts zur datenbasierten Stadtentwicklung auf
kommunaler Ebene**

vorgelegt bei

Prof. Dr. Christian Schmeißer

Dipl. Biol. Stefan Rensch

Eingereicht von:

Name: Kevin Pocher

Abgabe: 02.09.2022

Sommersemester 2022

Danksagung

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich während der Anfertigung dieser Masterarbeit unterstützt und motiviert haben.

Zuerst gebührt mein Dank Herrn Rensch von der Stadt Mücheln, welcher mich bei der Themenfindung unterstützt hat und zudem auch meine Masterarbeit begleitet hat. Er war immer für Anfragen offen und konnte mir zahlreiche Ratschläge geben. Seine Anregungen und die immer konstruktive Kritik bei der Erstellung dieser Arbeit waren mir eine große Hilfe.

Ich bedanke mich auch bei meinem Team der GISA GmbH, welches mich unterstützt hat und mir immer einen guten Rat geben konnte.

Ebenso möchte ich meiner ehemaligen Tutorin während des Abiturs, Frau Beate Siewert, und meiner Familie danken, die mich während der ganzen Monate der Masterarbeit motiviert, begleitet und ebenfalls unterstützt haben.

Abschließend gebührt mein besonderer Dank Prof. Dr. Christian Schmeißer, welcher meine Masterarbeit betreut und begutachtet hat.

Kevin Pocher

Roßbach, 30.08.2021

Inhaltsverzeichnis

Abbildungsverzeichnis	I
Tabellenverzeichnis	II
Abkürzungsverzeichnis	III
1 Einleitung.....	1
1.1 Gegenstand der Forschung.....	1
1.2 Zielstellung der Arbeit	2
1.3 Methodik	2
1.4 Zu erwartender Nutzen	4
2 Theoretische Grundlagen.....	5
2.1 Digitale Transformation auf kommunaler Ebene	5
2.2 Open Data.....	7
2.3 Geodaten und Geoinformationssysteme	9
2.3.1 Digitale georeferenzierte Daten	9
2.3.2 Geoinformationssysteme	10
2.3.3 Verwendung von Geodaten in Kommunen.....	13
2.3.4 Level of Detail 2 – Gebäudemodelle	14
2.4 ETL-Prozess	16
3 Datengrundlage am Beispiel Mücheln (Geiseltal).....	20
3.1 Vorstellung der Stadt Mücheln (Geiseltal)	20
3.2 Untersuchungsobjekt	22
3.3 Bereitgestellte Daten.....	25
4 Analyse der Datenbasis auf kommunaler Ebene	27
4.1 Verwaltungsinterne Daten	27
4.2 Open Geodaten	32
5 Experimentelle Umsetzung	35
5.1 Vorbereitung der Umsetzung.....	35
5.1.1 Benötigte Software.....	35
5.1.2 Daten Mapping.....	38
5.2 Ermittlung von Gebäuden für Kindergärten.....	45
5.3 ETL-Prozess Stadt Mücheln	50
5.4 Weitere Funktionen.....	54
6 Diskussion der Ergebnisse	56

6.1	Auswertung der Ergebnisse.....	56
6.2	Weitere Möglichkeiten für die Stadt Mücheln.....	61
7	Fazit und Ausblick.....	63
8	Literatur.....	66
	Anhangsverzeichnis	71
	Anhang 1: LoD2-Daten Beispiele	71
	Anhang 2: Durchgeführte statistische Analysen.....	71
	Anhang 3: Beispiele zur Umsetzung und für Visualisierungen in QGIS	75
	Anhang 4: Ergebnisse der Untersuchung für Mücheln.....	79
	Eidesstattliche Erklärung	IV

Abbildungsverzeichnis

Abbildung 1: Framework für das Design Science Research nach Hevner.....	3
Abbildung 2: Datenstruktur von ALKIS am Beispiel Gebäude, nach niedersächsischer Vermessungs- und Katasterverwaltung.....	12
Abbildung 3: ETL-Prozess.....	17
Abbildung 4: Gemarkungen der Stadt Mücheln (Geiseltal).....	20
Abbildung 5: Kinder in Tageseinrichtungen: Deutschland, Stichtag, Altersgruppe.....	22
Abbildung 6: Geoobjekt mit demografischen Daten.....	24
Abbildung 7: Bevölkerungsanzahl nach Alter in Mücheln (Geiseltal) im Jahr 2022.....	28
Abbildung 8: Geburtenrate im Zeitraum 2000-2021 in der Stadt Mücheln (Geiseltal).....	29
Abbildung 9: Lineare Regression für die Zu- und Abwanderung der Bevölkerung für Mücheln (Gesamt) und Langeneichstädt.....	31
Abbildung 10: Aktualität des LoD2-Datensatzes am Beispiel Stöbnitzer Mitte.....	32
Abbildung 11: LoD2 Gebäudefunktion - „Gebäude für soziale Zwecke“ und OSM-Daten.....	33
Abbildung 12: Plugin für die Analysen im QGIS.....	36
Abbildung 13: Darstellung von Flurstücken der Stadt Mücheln im pgAdmin.....	37
Abbildung 14: Darstellung des LoD2-Datensatzes aus der PostGIS innerhalb von QGIS.....	38
Abbildung 15: Beispielhafte Gebäudegeometrie aus dem LoD2-Datensatz.....	39
Abbildung 16: Selektionsmaske des Programmes zum Daten Mapping.....	40
Abbildung 17: N-Gramm und Dice Koeffizient am Beispiel Straßename.....	42
Abbildung 18: Beispiel für eine Hashtabelle.....	43
Abbildung 19: Verkettung bei einer Kollision in einer Hashtabelle.....	44
Abbildung 20: Prozessablauf des Algorithmus zur Bestimmung der Gebäude für Kindergärten.....	46
Abbildung 21: QGIS Berechnung des Einzugsgebietes und Verschneidung der Gebäude.....	49
Abbildung 22: ELT-Prozess für die Zusammenführung und Analyse der Geodaten und demografischen Daten.....	52
Abbildung 23: Ausschnitt des Ergebnisdokument des Algorithmus für den OT Gröst.....	59
Abbildung 24: Ausschnitt des Ergebnisses des Algorithmus als Visualisierung im QGIS.....	60

Tabellenverzeichnis

Tabelle 1: Detailierungsstufen von LoD, nach Borrman, et al. 2015	15
Tabelle 2: Beispiele für Funktionen der Transformation	18
Tabelle 3: Datenstruktur der demografischen Daten aus der MESO DB.....	27
Tabelle 4: Berechnung des Bedarfs an Kindergartenfläche pro Ortsteil.....	50
Tabelle 5: Anzahl der Kinder je Ortsteil mit Berücksichtigung eines Einzugsgebiets	57
Tabelle 6: Ist- und Soll-Wert an Flächen für Kindertagesstätten in Müheln.....	58

Abkürzungsverzeichnis

ALKIS	Amtliche Liegenschaftskataster-Informationssystem
ATKIS	Amtliche Topographisch-Kartographische Informationssystem
BBSR	Bundesinstitut für Bau-, Stadt- und Raumforschung
ETL	Extract, Transform, Load
GDAL	Geospatial Data Abstraction Library
GIS	Geoinformationssystem
GML	Geography Markup Language
GRASS	Geographic Resources Analysis Support System
LVerGeo	Landesamt für Vermessung und Geoinformation Sachsen-Anhalt
LoD2	Level of Detail 2
MESO	Meldewesen Software
OGP	Open Government Data
OSM	Openstreetmap
OT	Ortsteil
OWSLIB	Open Web Services Library
SAGA	System for Automated Geoscientific Analyses
WFS	Web Feature Service
WMS	Web Map Service

1 Einleitung

1.1 Gegenstand der Forschung

Bedingt durch den starken Strukturwandel in Verbindung mit dem demographischen Wandel der letzten Jahre wurden die Herausforderungen für die Kommunen und Gemeinden in den ländlichen Räumen immer größer, die eigene Attraktivität zu bewahren bzw. zu verbessern. Die damit einhergehenden Veränderungen führen zu immer stärkerer Abwanderung der jüngeren Bevölkerung und zu einer Abschwächung der wirtschaftlichen Standorte.

Doch durch die immer weiter voranschreitende Digitalisierung entstehen neue Möglichkeiten, deren Nutzung nun nicht nur für größere Städte, sondern ebenso für Kleinstädten möglich werden kann. Dabei ist die Umwandlung von analogen zu digitalen Informationen bzw. Formaten einer der Effekte der Digitalisierung.¹ Zusätzlich werden die neuen Technologien immer komplexer und können viele Chancen bieten, welche die Lebensqualität im ländlichen Raum verbessern. Doch gerade in den ländlichen Räumen fehlt meist noch immer der Zugang sowie das technische Knowhow der neuen Technologien, um die neuen Möglichkeiten ergreifen zu können. Im Rahmen der Weiterentwicklung ist es von entscheidender Bedeutung, dass Bürger in diesen Räumen gleichwertigen Zugang zu digitalen Angeboten erhalten. Insofern sehen sowohl (Klein-)Städte als auch Gemeinden in ländlichen Räumen eine Chance in der Digitalisierung und den damit verbundenen neuen Technologien, um sowohl die eigenen Standorte wirtschaftlich aufzuwerten als auch als Wohnort attraktiver zu gestalten.² Darum ist es wichtig, dass die Städte bzw. Kommunen und Gemeinden in ländlichen Regionen den Ausbau der eigenen digitalen Infrastruktur eigenverantwortlich weiter vorantreiben und damit die Nutzung der sich bietenden digitalen Entwicklungsprojekte zur Verbesserung der örtlichen Struktur und Attraktivität möglich wird.

Die durch die Digitalisierung generierten Daten können für spezifische Themengebiete ausgewertet bzw. weiterverarbeitet werden, um einen Mehrwert für die Kommunen und die Bürger zu schaffen. Jedoch müssen die vorhandenen Daten zuvor in ein verarbeitbares Format gebracht werden. Beispielsweise stehen den Städten umfangreiche Geodatenätze zur Verfügung, welche für Auswertungen und Entscheidungsprozesse herangezogen werden können. Jedoch fehlen den meisten Städten die digitalen Lösungen, um mit den erhobenen Daten zielgerichtet zu arbeiten. So werden diese Daten derzeit meist nur für flächendeckende

¹ Vgl. Porsche; Steinführer; Sondermann 2019, S. 3.

² Vgl. Antonia Milbert 2021.

Register von Flurstücken und für wenige interne Anwendungsfälle der Kommunen verwendet. Wichtig wäre es also, hier neue Anwendungsfälle zu schaffen, welche den Kommunen intern neue digitale Prozesse aufzeigen, aber auch extern für Bürger einen Mehrwert darstellen.

1.2 Zielstellung der Arbeit

Im Rahmen dieser Forschungsarbeit wird untersucht, welche Daten und Indikatoren aus Geoinformationen und kommunalen Daten zur Analyse herangezogen werden können, um Algorithmen zur regionalen Stadtentwicklung zu entwickeln. Das Ziel dieser Masterarbeit soll deswegen sein, ein Konzept zu erarbeiten, welches den Kommunen neue Möglichkeiten im Umgang mit vorhandenen und speziell geobasierten Daten aufzeigt. Der Fokus soll dabei vor allem auf Kleinstädte und ländliche Gegenden gelegt werden. Das Konzept soll beispielhaft an der Stadt Mücheln (Geiseltal) aufgestellt und untersucht werden. Um dieses bewerkstelligen zu können, muss zuvor analysiert werden, welche Daten für Kommunen überhaupt zur Verfügung stehen und welche dieser Daten einen Mehrwert für die regionale Stadtentwicklung bedeuten können. Dabei sollen auch Open Data mit in die Analyse einbezogen werden. Open Data stellen dabei Daten dar, die für jegliche Anwendungsfälle kostenfrei verwendet, verarbeitet bzw. transformiert und weitergegeben werden können.³ Durch das Konzept sollen die vorhandenen und neuen Daten in eine neue Datenstruktur überführt werden, welche dann für Planungsalgorithmen und Analysen verwendet werden kann.

1.3 Methodik

Um die zuvor genannten Ziele erreichen zu können, wurde folgende Forschungsfrage aufgestellt.

- Welche Daten und Indikatoren können aus Geoinformationen in Verbindung mit kommunalen Daten zur Analyse herangezogen und abgeleitet werden, um darauffolgend Algorithmen zur Stadtentwicklung im ländlichen Raum zu entwerfen?

Um diese beantworten zu können, soll der Forschungsansatz des Design Science Research verwendet werden. Dieser Ansatz verfolgt das Ziel, neues Wissen aus Artefakten zu gewinnen. Dabei sollen Lösungen für aktuelle Probleme aus der Wirtschaft geschaffen werden. Innerhalb des Design Science Research gibt es verschiedene Anwendungsmöglichkeiten. In

³ Vgl. BMI 2021.

dieser Arbeit soll sich die Lösungsfindung auf das Framework nach Hevner et al. (2004) stützen. Die folgende Abbildung zeigt das Framework nach Hevner auf und verdeutlicht die verschiedenen Aspekte, welche beleuchtet werden müssen.

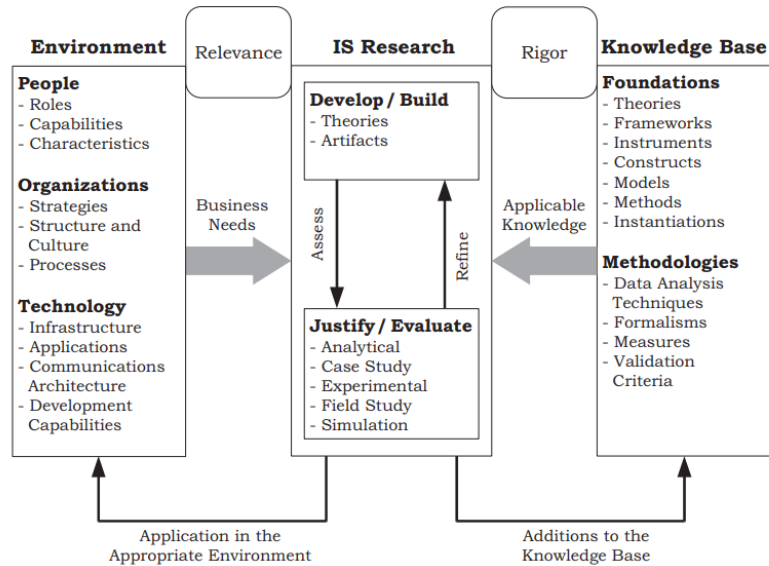


Abbildung 1: Framework für das Design Science Research nach Hevner⁴

Durch die Verwendung des Forschungsansatzes des Design Science Research soll zunächst die Problematik aufgezeigt werden, welche für Kleinstädte bei der Bereitstellung und Umsetzung von Anwendungsfällen für geodatenbasierte Datenanalysen im Umfeld der Stadtentwicklung auftritt. Grundsätzlich muss dafür jedoch noch aufgeklärt werden aus welchem Grund, gerade Kleinstädte bzw. Städte in ländlichen Räumen Schwierigkeiten mit der Stadtentwicklung haben. Somit wird sich der erste Teil dieser Arbeit mit dem Problemraum beschäftigen und die neben der bestehenden Problematik auf über die theoretischen Grundlagen aufklären.

Darauffolgend soll eine Ist-Analyse der Datenbasis auf kommunaler Ebene aufgestellt werden. Diese soll den Datenbestand von Kleinstädten aufzeigen. In Verbindung mit den Geobasisdaten werden dann die Zusammenhänge analysiert. Aus diesen sollen dann Logiken für die Umsetzungen eines Algorithmus abgeleitet werden.

Die dadurch aufgestellte Wissensbasis soll dann in einen Planungsalgorithmus überführt werden. Schließlich soll ein Experiment die Evaluation der Lösung ermöglichen. Dieses soll dabei als Feldexperiment aufgestellt und mit den vorhandenen und erarbeiteten Daten durchgeführt werden. Um das Vorhaben umsetzen zu können, werden einerseits die Geobasisdaten und andererseits die demographischen Daten der Stadt benötigt. Weiterhin sollen

⁴ Vgl. Alan R. Hevner, et al. 2004, S. 80.

verschiedene Tools, wie Open Source Software und verschiedene Datenbankmanagementsysteme eingesetzt und für Lösungsansätze getestet werden. Der letzte Teil der Arbeit wird somit die neue Wissensbasis analysieren und die Aussagekraft diskutieren.

1.4 Zu erwartender Nutzen

Als Ergebnis dieser Arbeit soll ein Konzept geliefert werden, welches vorrangig Kleinstädten und Städten in ländlichen Gegenden eine Möglichkeit für eine modernere Stadtentwicklung bietet. Dieses soll den Umgang mit georeferenzierten Daten in Verbindung mit Daten der Stadtverwaltung aufzeigen. Dabei sollen bereits bestehende Prozesse digitalisiert und neue Handlungsfelder aufgestellt werden. Zusätzlich soll ein Anwendungsfall exemplarisch umgesetzt werden.

Letztlich soll diese Arbeit den Städten eine Hilfestellung bzw. einen Ansatz liefern, um eine moderne Stadtentwicklung anzustreben, welche Abhilfe bei Themen wie Strukturwandel oder auch dem demografischen Wandel schafft.

2 Theoretische Grundlagen

2.1 Digitale Transformation auf kommunaler Ebene

Egal ob digitaler Wandel, digitale Transformation oder einfach nur Digitalisierung. Diese Begriffe sind allgegenwärtig und aus der heutigen Zeit nicht mehr wegzudenken. Die Digitalisierung beschreibt dabei die Umwandlung von analogen Daten in ein digitales Format, mit dem Zweck diese später kostengünstig zu reproduzieren, schnellstmöglich weltweit zu verteilen und maschinell auszuwerten.⁵ Ebenso ist der Begriff der digitalen Transformation in der stattfindenden Digitalisierung nicht mehr wegzudenken. Dabei steht die Transformation für die Verwendung von neuen digitalen Technologien, welche einen Mehrwert für den jeweiligen Anwendungsbereich bieten.⁶

Fast jeder Lebensbereich wird derzeit durch die Digitalisierung beeinflusst und birgt dabei verschiedenste Herausforderungen aber auch Chancen. Auch die Städte und Kommunen in Deutschland sind somit mit von der Digitalisierung betroffen. Die wichtigsten Themen, welche sich für die Städte aufzeigen, sind der Wandel in den Bereichen Arbeitsmarkt, Bildung, Forschung, Infrastruktur, Verwaltung und Wirtschaft.⁷ Innerhalb dieser Bereiche findet derzeit ein grundlegender Wandel der Strukturen statt. Seien es die neuen Modelle des mobilen Arbeitens oder auch die neuen Anforderungen der Bürger ihrer Stadt gegenüber. Dieser Wandel ist dabei ein relevantes Element, wenn es darum geht als Stadt bzw. auch als Wohnort attraktiv zu werden. Doch gerade durch die periphere Lage und die damit verbundenen Probleme, wie Breitbandausbau, Anbindung oder auch Daseinsvorsorge, stellt sich hier eine deutliche Benachteiligung der ländlichen Kleinstädte heraus. Wenn nun beachtet, dass Deutschland ein Land der Kleinstädte ist und diese ca. 58 Prozent der Städte und Gemeinden ausmachen und somit auch ca. ein Drittel der Bevölkerung diese besiedelt, wird klar, welche Relevanz sie bei der Stadtentwicklung haben. Gerade Kleinstädte, welche abseits von Großstädten liegen, weisen eine immer kleiner bzw. älter werdende Bevölkerung vor und durchleben einen stärkeren demografischen Wandel als andere Orte. Somit stehen diese Städte vor dem Problem die jüngere Bevölkerung an sich zu binden und die Abwanderung in andere Städte zu minimieren. Doch auch für die alternde Bevölkerung bedarf es ebenso einer konstanten Versorgung und einer dementsprechenden Infrastruktur. Somit müssen sich Kleinstädte in ländlichen Räumen Ihrer Lage bewusstwerden und entsprechend agieren. Die

⁵ Vgl. Harwardt 2022, S. 3.

⁶ Vgl. Harwardt 2022, S. 12.

⁷ Vgl. Heuermann; Tomenendal; Bressemer 2018, S. 13.

Anpassung an die neuen Herausforderungen und der Einsatz von neuen Möglichkeiten ist dabei essenziell, um Potenziale und Strategien für eine moderne Stadtentwicklung zu nutzen.⁸

Derzeit werden eine Vielzahl an Projekten und Forschungsaufträgen unter dem Förderprogramm „Experimenteller Wohnungs- und Städtebau“ (ExWoSt) umgesetzt, welche durch modellhafte Vorhaben versuchen, das Potential von Kleinstädten mit Bezug auf die eigene Lage der Infrastruktur zu untersuchen. So konnte bereits nachgewiesen werden, dass eine konzeptionelle Neuausrichtung des Stadtumbaus nötig ist, um eine Transformation der städtebaulichen Entwicklung auf kommunaler Ebene erreichen zu können. Damit diese Transformation funktionieren kann, bedarf es zusätzlich einer Orientierung an den globalen Herausforderungen und der damit verbundenen Anpassung der kommunalen Stadtplanung. Andererseits müssen jedoch auch neue Technologien und deren Chancen angenommen und durch die Kommunen gefördert werden. Um dies erfüllen zu können, müssen die Kommunen auch das dementsprechende Personal und die benötigten Ressourcen bereitstellen, damit ein neues Stadtumbauprogramm erfolgreich umgesetzt werden kann.⁹ Eine der neuen Technologien, welche bereits durch das ExWoSt-Projekt untersucht wurde, sind webbasierte Medien und deren Wirkung auf die Bevölkerung vor Ort und darüber hinaus. Durch eine Vielzahl an untersuchten Referenzbeispielen und Fallstudien konnte hier festgestellt werden, dass durch den Einsatz dieser Medien neue Akteure, auch auf kommunaler Ebene, angezogen werden können. Diese tragen wiederum dazu bei, dass es zu einer Weiterentwicklung vor Ort kommt. Außerdem initiiert eine solche Form der Mediennutzung eine neue Art der Kommunikation, Partizipationsprozesse zwischen den Kommunen und der Bevölkerung. Dies trägt zu mehr Transparenz bei, was auch das Vertrauen bzw. die Akzeptanz von Entwicklungsprozessen steigert.

Die neuen Technologien können dabei nicht nur einen Mehrwert für die Kommunikation und die Transparenz darstellen, sondern auch neues Wissen zu generieren. Zusätzlich wurde versucht, mit verschiedensten Verfahren Wissen zu gewinnen, welches der kommunalen Selbstverwaltung von Nutzen sein kann. Die Verfahren, die untersucht wurden, waren bspw. Open Data, Sensordaten, Crowdsourcing oder auch Gamification. Hierbei konnte festgestellt werden, dass häufig ein Umdenken der aktuellen Struktur der Verwaltung notwendig ist, damit neue Anwendungen bzw. Verfahren integriert werden können.¹⁰

⁸ Vgl. BBSR 2016, S. 4.

⁹ Vgl. R. Riechel, Dr. H. Scheller, J. H. Trapp, Dr. J. Libbe, B. Walker 2022.

¹⁰ Vgl. BBSR 2018.

Allerdings musste auch festgestellt werden, dass es neue bzw. zusätzliche Ressourcen benötigt: neben den technischen Voraussetzungen bedarf es auch finanzieller Aufwendungen und eines spezifischen Know-How. Hier müssen unter Umständen auch externe Beteiligte herangezogen werden.¹¹ Dies wird vor dem Hintergrund deutlich, dass der Umgang mit digitalen Medien innerhalb der Städte und in ihren Verwaltungen problematisch ist.

Neben den verschiedenen Anwendungsfällen ist jedoch auch die Datenvielfalt ein wichtiges Thema. Die Datenlage der Kleinstädte bildet die Voraussetzung für die Entwicklung von Anwendungsfällen. Dabei lässt sich sagen, dass der Umfang der Datenbasis von Kleinstädten um ein Vielfaches geringer ausfällt als bei Großstädten und somit die Möglichkeiten begrenzter sind. Wichtig ist somit, für die Kleinstädte eine Konzeption eines Datenmodells aufzustellen, welches diesen Begrenzungen entgegenwirken kann.¹² Diese Thematik soll später noch einmal aufgegriffen werden, wenn es um die beispielhafte Datenanalyse der Kleinstadt Mücheln (Geiseltal) geht. Doch zunächst muss noch geklärt werden, wo die Geodaten beschafft werden können und wie sich geobasierte Daten zusammensetzen.

2.2 Open Data

Mit der voranschreitenden Digitalisierung werden gleichzeitig immer mehr Daten aus allen Lebensbereichen generiert. Seien es demographische, wirtschaftliche, Verwaltungs- oder andere Daten in allen Bereichen werden täglich große Mengen an Daten gesammelt.¹³ Diese bilden dabei die Grundlage einer modernen Gesellschaft bzw. auch eines modernen Staates. Somit können die Daten nicht nur einen Mehrwert für die Wirtschaft und Forschung bieten, sondern auch den Bürgern. Damit jedoch alle Beteiligten neben der Behörde die Möglichkeit erhalten auf diesen umfangreichen Datenbestand zuzugreifen, müssen die Daten innerhalb des gesetzten Rechtsrahmens entgeltfrei für eine freie Nutzung bereitgestellt werden.¹⁴ Hier greift der Ansatz der Open Data. Die Richtlinie (EU) 2019/1024 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors, definiert den Begriff Open Daten oder offene Daten als „Daten in einem offenen Format, die von allen zu jedem Zweck frei verwendet, weiterverwendet und weitergegeben werden können“. Zu diesen Daten zählen beispielweise Geodaten, Statistiken, Verkehrsinformationen, Forschungsergebnisse und

¹¹ Vgl. Prof. Dr. Claus-C. Wiegandt; Michael Lobeck 2017.

¹² Vgl. Antonia Milbert 2021, S. 45.

¹³ Vgl. Benjamin Seibel 2016, S. 5.

¹⁴ Vgl. BMI 2021, S. 5.

vieles mehr. Somit ist klar, dass es sich bei Open Data nicht nur um Datenbestände der öffentlichen Verwaltung, sondern auch um privatwirtschaftliche Beiträge handelt.¹⁵

Wenn nun der Ansatz der Open Data auf öffentliche Verwaltungs- und Forschungsdaten bezogen wird, dann wird dies als Open Government Data (OGD) bezeichnet.¹⁶ Dabei steht Open Government für eine Verwaltung, welche sich nach Außen öffnet und somit Transparenz, Partizipation und Kooperation vor allem mit der Bevölkerung ermöglicht.¹⁷ Dabei sollen durch diese Dimensionen die Einbeziehung der Bürger in Entscheidungsprozesse gefördert werden.¹⁸ Die einzelnen Dimensionen des Open Governments lassen sich wie folgt beschreiben:

- **Transparenz** als Ziel, das Handeln von Regierung und Verwaltung in eine nachvollziehbare Weise zu entwickeln. Dadurch sollen Informationen auch für Bürger bereitgestellt werden, sodass diese eine bessere Datengrundlage haben und Entscheidungen treffen können. Beispiele aus der Praxis stellen hier Open-Data-Portale der Städte dar.
- **Partizipation** als Ziel, Bürger für die Mitgestaltung der Stadt zu begeistern und diese in Entscheidungsprozesse einzubeziehen. Hier gibt es bereits verschiedene Formen z.B. Befragungen oder die Übergabe der Entscheidungsverantwortung an die Bürger. Beispiele für die Partizipation sind hier Ansätze der Bürgerbeteiligung in verschiedensten Bereichen, wie die Verwendung von öffentlichen Geldern (buergerhaushalt.org) oder auch eine Anlaufstelle für Bürgeranliegen (www.mangelmelder.de).
- **Kooperation** als Ziel soll die Zusammenarbeit der Bürger mit der Verwaltung ermöglichen. Beispiele für Plattformen, welche für die Kooperation mit den Bürgern genutzt werden können, sind Social Media, Wiki's, Shared Service Center, usw.¹⁹

Open Government Data müssen dabei folgende Kriterien erfüllen, damit die erhobenen und gespeicherten Daten der öffentlichen Verwaltung als solche zählen können. Somit müssen OGD folgende Punkte erfüllen:

- die Daten wurden von der Behörde erhoben oder Dritte dazu aufgefordert
- sind frei für alle zugänglich

¹⁵ Vgl. Jörn von Lucke und Christian Geiger 2010, S. 7.

¹⁶ Vgl. Kubicek; Jarke 2020, S. 2.

¹⁷ Vgl. Jörn von Lucke 2017, S. 153.

¹⁸ Vgl. Bundesministerium des Innern 2012, S. 26.

¹⁹ Vgl. Bundesministerium des Innern 2012, S. 26f.

- maschinenlesbar
- aktuell
- enthalten keinen Personenbezug
- frei verwendbar
- enthalten keine sicherheitsrelevanten Informationen.^{20,21,22}

Es lassen sich verschiedene Ziele für die Anwendung von Open Data in der öffentlichen Verwaltung ableiten. Zum einen soll das Vertrauen der Bürger gegenüber der Politik verstärkt werden, dies geschieht durch die schon genannte Transparenz der Daten. Zum anderen soll ein datenangetriebenes Wirtschaftswachstum durch die Nutzung von OGD ermöglicht werden, indem durch offene Daten neue Anwendungsmöglichkeiten auch für kleine und mittlere Unternehmen zur Verfügung stehen. Zu diesen neuen Anwendungsmöglichkeiten zählen auch heute schon mobile Anwendungen, welche verschiedenste Dienste mit georeferenzierten Daten anbieten.²³ Ein weiteres Ziel für OGD stellt die Effizienzsteigerung in der öffentlichen Verwaltung dar. Diese soll vor allem durch weniger Suchaufwand und die Reduzierung von Redundanzen ermöglicht werden.²⁴ Ein Beispiel für die umfangreiche Nutzung von OGD liefert die Stadt Leipzig, welche schon jetzt eine Vielzahl an Datensätzen zur Verfügung stellt.²⁵

2.3 Geodaten und Geoinformationssysteme

In diesem Kapitel soll nun aufgezeigt werden, wie sich georeferenzierte Daten zusammensetzen und welchen Wert diese bereits für Städte und Kommunen besitzen. Des Weiteren soll auch aufgezeigt werden, welche Lösungen es im Bereich der Geoinformationssysteme bereits gibt und wie diese eine moderne Stadtentwicklung ermöglichen können.

2.3.1 Digitale georeferenzierte Daten

Unter dem Begriff der Geodaten werden digitale Informationen verstanden, welche ein Geoobjekt mit dessen räumlichem und zeitlichem Bezug beschreiben, sodass dessen Koordinaten räumlich zugeordnet werden können. Dabei stellen Geoobjekte reale Objekte oder

²⁰ Vgl. BMI 2021, S. 6.

²¹ Vgl. Kubicek; Jarke 2020, S. 3.

²² Vgl. Jörn von Lucke und Christian Geiger 2010, S. 8f.

²³ Vgl. Kubicek; Jarke 2020, S. 5.

²⁴ Vgl. BMI 2021, 10f.

²⁵ Vgl. Stadt Leipzig.

auch Ereignisse aus der realen Welt dar, welche in einen räumlichen Bezug überführt werden. Ein Geoobjekt besitzt jedoch auch immer eine bestimmte Thematik, welche durch mehrere Merkmale, Sachdaten oder Attribute aufgezeigt wird. Somit können dem Objekt verschiedene themenbasierte Informationen zugeordnet werden, Einfache Beispiele für ein Geoobjekt sind bspw. Gebäude. Diese haben Attribute, wie Hausnummer, Dachhöhe, Grundfläche, Baujahr, Bewohner und viele mehr. Eine andere Art des Geoobjektes ist ein Ereignis oder auch Phänomene, wie bspw. der demografischen Wandel oder die Veränderung der Wetterbedingungen an einem Standort.^{26, 27, 28, 29} Damit die georeferenzierten Daten für verschiedenste Anwendungen genutzt werden können, müssen diese einen eindeutigen Bezug zur Lage des Objektes vorweisen. Dies ist zwingend notwendig, damit die Daten später auch flächenübergreifend in einen Bezug gebracht und analysiert bzw. dargestellt werden können. Dafür gibt es amtlich festgelegte Bezugssysteme für die Lage, Höhe und Schwere, die eine Einheitlichkeit ermöglichen.³⁰ Dabei wird bei Geoobjekten des Weiteren unter Vektoren und Rastermodellen unterschieden. Dabei handelt es sich bei dem Vektormodell, um Vektoren innerhalb eines Koordinatenbezugssystems, welche durch einen Start und Endpunkt ermittelt werden können. Die verschiedenen Vektoren können sich als viele Punkte vorstellen, welche die Eckpunkte einer Geometrie darstellen. Hingegen stellt das Rastermodell eine Fläche fester Formen und Größen dar, die als ein Raster von Quadraten aufgestellt wird. In der Praxis spielt das Rastermodell eine wichtige Rolle bei der Bilderverarbeitung. Ein Beispiel wären hier die Verarbeitung von Satellitenaufnahmen in kleinen quadratischen Pixeln.³¹

2.3.2 Geoinformationssysteme

Um zu verstehen was ein Geoinformationssystem für eine Bedeutung hat und welche Funktionsweisen mit diesem einhergehen, wird nun zuerst der Begriff Informationssystem im Allgemeinen definiert. Ein Informationssystem ist ein computergestütztes System, welches Informationen verarbeiten kann, diese also ermittelt, transportiert, transformiert, speichert und auch bereitstellen kann. Dabei lassen sich Informationssysteme im heutigen Alltag in fast allen Bereichen der Wirtschaft aber auch Verwaltung finden.³² Die fundamentale

²⁶ Vgl. Klenk; Nullmeier; Wewer 2020, S. 102.

²⁷ Vgl. Lange 2020, S. 133.

²⁸ Vgl. GeoZG 2009, § 3 Abs. 1.

²⁹ Vgl. Ralf Bill 2018, S. 790.

³⁰ Vgl. LVErmGeo 2021, S. 8.

³¹ Vgl. Lange 2020, S. 129f.

³² Vgl. Ferstl; Sinz 2013, S. 3.

Aufgabe des Informationssystems ist es dabei, auf einen vorliegenden Datenbestand zurückzugreifen und diesen dann im Sinne der Anwendung auszuwerten, um daraus Informationen zu generieren. Je nach Art der gespeicherten Informationen, wie beispielsweise numerische, textuelle, bildhafte Daten, lassen es sich verschiedene Arten von Informationssystemen unterscheiden. Beispiele wären hier Informationssysteme in Banken bei der Verwaltung von Kundendaten, in Reisebüros mit verschiedenen Informationen über Verkehr, Reiseziele oder auch Buchungen und viele andere Arten. Das Informationssysteme, welches in dieser Arbeit eine Rolle spielt, ist das Geoinformationssystem (GIS). Diese Art des Informationssystems setzt sich mit der Verarbeitung von raumbezogenen Daten bzw. Geoobjekten der realen Welt auseinander.³³

Bei einem Geoinformationssystem handelt es sich um ein computergestütztes Informationssystem, welches aus Daten, Software und Hardware besteht. Diese ermöglichen es georeferenzierte Daten persistent zu speichern, neu zu strukturieren, Analysen der Daten durchzuführen und diese auch für verschiedenste Anwendungsfälle zu visualisieren.³⁴ Durch die Analyse der unterschiedlichen Geodaten innerhalb eines GIS, können Lösungen für Probleme der realen Welt geschaffen werden. Dabei spielt es auch eine große Rolle, dass alle Geodaten eine gemeinsame Eigenschaft haben, und zwar einen räumlichen Bezug zu einem realen Objekt in der Welt. Durch die Verknüpfung von Geodaten und Fachdaten können somit neue Informationen bzw. auch neues Wissen geschaffen werden. Zu den Anwendungen, welche durch ein GIS ermöglicht werden können, zählen vielfältige Karten und Pläne:

- Karten für touristische Zwecke
- Wind- Wasser, Wetter- bzw. Klimakarten
- Pläne für Standortoptimierungen (Leerstandsmanagement, Renovierungen, usw.)
- Risikoanalysen für den Katastrophenschutz (Hochwassersimulationen)
- Abbildung der demographischen Begebenheiten
- Und viele weitere.³⁵

Die zuvor genannten Beispiele für die Anwendung eines GIS zeigen, dass das Potential sehr groß sein kann. So lassen sich gerade in der Stadtentwicklung viele der Funktionen eines GIS verwenden. Die Grundlage für ein GIS im Bereich der Stadtentwicklung stellen dabei die Geobasisdaten des Amtlichen Liegenschaftskataster-Informationssystem (ALKIS) und

³³ Vgl. Lange 2020, S. 373.

³⁴ Vgl. Hintze; Lakes 2009.

³⁵ Vgl. Gerold Noack 2019, S. 65f.

des Amtliche Topographisch-Kartographischen Informationssystem (ATKIS) dar.³⁶ Dabei beinhaltet das ALKIS alle Daten rund um die Liegenschaft der verschiedenen Regionen.³⁷ Somit enthält das ALKIS Daten über verschiedene Arten von Objekten. Die Daten zu diesem Objekt werden in die Objektartenbereiche Flurstücke, Eigentümer, Gebäude, tatsächliche Nutzung, Bauwerke und weiteres unterteilt. Diese Bereiche enthalten dann jeweils verschiedene Unterkategorien, welche dem Objekt verschiedene Attribute zuweisen.³⁸ Folgende Abbildung soll die Datenstruktur für ein Gebäude mit der Funktion „Gebäude für soziale Zwecke“ aufzeigen.

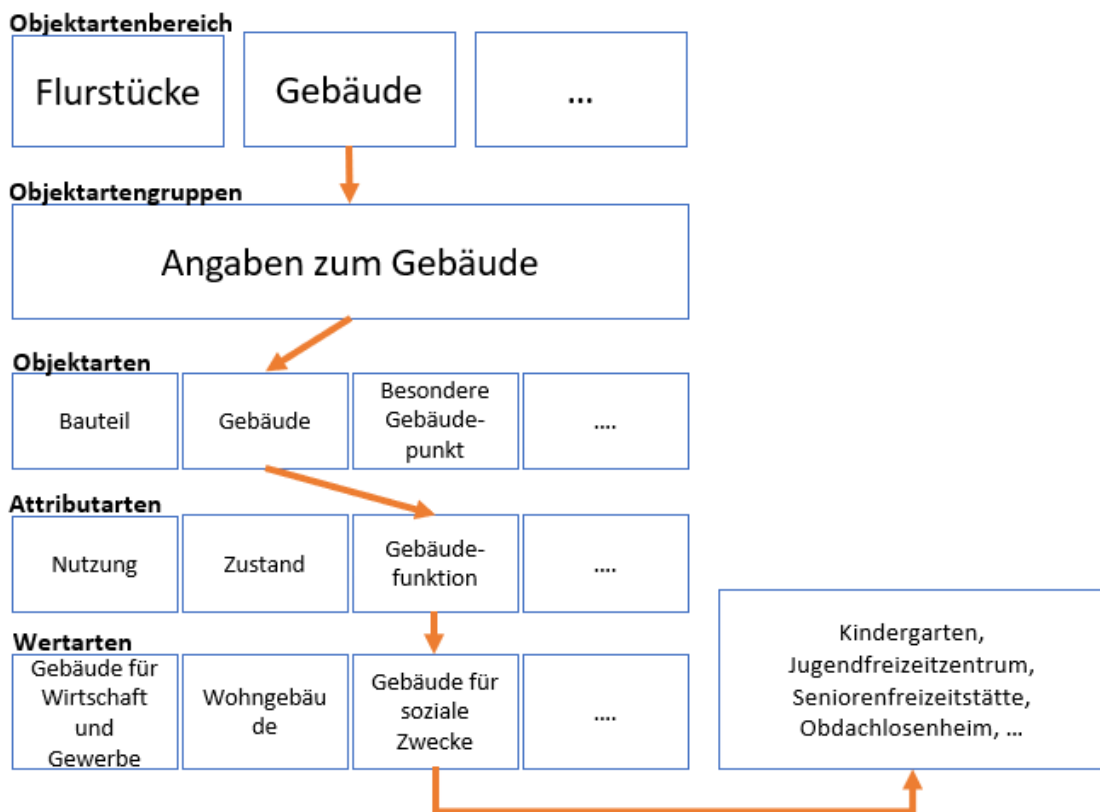


Abbildung 2: Datenstruktur von ALKIS am Beispiel Gebäude, nach niedersächsischer Vermessungs- und Katasterverwaltung³⁹

Anhand der Abbildung lässt sich die Aufgliederung der Datenstruktur in themenbasierte Abschnitte erkennen.

³⁶ Vgl. Ralf Bill 2018, S. 792.

³⁷ Vgl. Lange 2020, S. 220.

³⁸ Vgl. Lange 2020, S. 225f.

³⁹ Vgl. Niedersächsische Vermessungs- und Katasterverwaltung 2010, S. 61.

Das ATKIS hingegen liefert Datensätze zu digitalen Landschaftsmodellen, topografischen Karten und Geländemodellen. Beispiele für Datensätze aus dem ATKIS wären topografische Informationen zu Gewässern, bebauten Flächen, Vegetation oder auch Verkehrswegen.⁴⁰

2.3.3 Verwendung von Geodaten in Kommunen

Geodaten finden schon seit jeher einen wichtigen Nutzen in der öffentlichen Verwaltung, bis heute hat sich die Anwendergemeinschaft auf andere öffentliche und wirtschaftliche Einrichtungen erweitert. Durch die Verwendung von Geodaten in einem GIS wird es den Verwaltungen und anderen Nutzergruppen erlaubt, schnelle grafische Visualisierungen von Sachverhalten aufzuzeigen, aber auch die Analyse dieser Sachverhalte mit einem räumlichen Bezug durchzuführen. Durch diese Analysen wird in der öffentlichen Verwaltung eine Effizienzsteigerung erzielt.⁴¹ Im Fokus liegt dabei die Verwendung von Geodaten in Bereichen wie Bau und Planung, Naturschutz und Umwelt, Verkehr, Gewässer, amtliche Statistik und vieles mehr. Im Bereich der Stadtentwicklung für Bau-, Planung- und Dokumentationskarten stellt GIS schon seit langem eine wichtige Anwendung dar. Die Baupläne werden nicht nur über GIS konstruiert und verwaltet, sondern auch über Dienste nach außen bereitgestellt.⁴² Durch die internetbasierte Infrastruktur wird es nicht nur den öffentlichen Verwaltungen oder Behörden ermöglicht, die Geodaten zu verwenden, sondern auch den Bürgern der Einstieg in die gemeinsame Nutzung der Geodaten eröffnet. Auch heute gibt es in diesem Bereich schon viele Angebote für Bürger, seien es öffentliche Geoportale oder auch Bürgerinformationsportale. Diese stellen den Bürger verschiedene Anwendungen wie Hochwassersimulationen, Bauplandaten, Schadstoffbelastung, Verkehr, usw. zur Verfügung.⁴³ Eine weitere wichtige Anwendung für die Verwaltung stellt die Verwendung von Geodaten im Bereich Tourismus dar. Beispielsweise können die Städte thematische Tourismuskarten mit den wichtigsten touristischen Zielen bereitstellen. Aber auch Radtourenplaner, Wanderportale o.ä. sind Möglichkeiten für Städte mit Geodaten die touristische Attraktivität zu stärken.⁴⁴ Ein weiteres Thema wäre noch das Risikomanagement mit Bezug auf Geodaten. Hier finden Geodaten vor allem Anwendung bei der Simulation von Katastrophen. Ein Beispiel wäre die zuvor schon genannt Simulation von Hochwassersituationen. Mithilfe der Analyse der Simulationen kann die Gefahr für das zu betrachtende Gebiet durchdacht werden, um

⁴⁰ Vgl. Niedersächsische Vermessungs- und Katasterverwaltung 2010, S. 5.

⁴¹ Vgl. Kay Ruge 2014, S. 229.

⁴² Vgl. Ostrau; Dr. Stefan, S. 34.

⁴³ Vgl. Klenk; Nullmeier; Wewer 2020, S. 103f.

⁴⁴ Vgl. Ostrau; Dr. Stefan, S. 50f.

später den Handlungsbedarf einschätzen zu können. Somit können Schutzmaßnahmen rechtzeitig ergriffen werden und in kritischen Gebieten präventiv gehandelt werden.⁴⁵

Durch die Nutzung der verschiedenen Anwendungen, welche zum Teil schon heute existieren, können die Verwaltungen bereits heute erheblich profitieren. Beispiele für den Mehrwert sind u.a. :

- Steigerung der Verwaltungseffizienz, durch das Bündeln der verschiedenen Daten zu einer übersichtlichen Struktur
- Verbesserung der Zugänglichkeit durch die Bündelung
- Entscheidungsgrundlage für verschiedene Themen wie Bauplanung, Verkehrswege, Kindergarten, usw.
- Verbesserung regionaler Wettbewerbsfähigkeit
- Transparenz für den Bürger.^{46, 47}

2.3.4 Level of Detail 2 – Gebäudemodelle

Für die Umsetzung der geplanten Algorithmen dieser Forschungsarbeit spielen die Level of Detail (LoD)- Datensätze, zu Deutsch „Detaillierungsgrad“, eine wichtige Rolle. Aus diesem Grund wird im Folgenden die Herkunft und der Aufbau der LoD-Datensätze mit Bezug zur Forschungsarbeit aufgezeigt.

Je nach Anwendungsfall werden unterschiedlich detaillierte Modelle von einem Objekt benötigt. Dafür wurden fünf verschiedene Detaillierungsstufen, die sogenannten Level of Details geschaffen. Dabei beschreibt jede Detaillierungsstufe eine Komplexität der geometrischen und semantischen Zusammensetzung des Objektes. Somit steigt die qualitative und quantitative Granularität der Darstellung des Objektes, je höher die Detaillierungsstufe wird. Wie bereits erwähnt, wird nicht für jedes Vorhaben die tiefste Detaillierungsstufe benötigt und somit werden die geringeren Stufen, welche auch eine geringe Komplexität aufweisen, für andere Anwendungen benötigt als die jeweils höheren Detailstufen. Die fünf Stufen wurden dabei aufsteigend benannt in LoD0 bis LoD4. Folgende Tabelle soll die einzelnen Eigenschaften der jeweiligen Detaillierung aufzeigen.⁴⁸

⁴⁵ Vgl. Ostrau; Dr. Stefan, S. 55.

⁴⁶ Vgl. Kay Ruge 2014, 230f.

⁴⁷ Vgl. Volker Galle 2014, S. 5.

⁴⁸ Vgl. Borrmann, et al. 2015, S. 186.

Detailierungsstufe	Eigenschaften
LoD0	<ul style="list-style-type: none"> • Regionalmodelle als generalisiertes Modell • dargestellt als zweidimensionale Projektion der Grundfläche des Objektes
LoD1	<ul style="list-style-type: none"> • Klötzchen-, Block- oder Massenmodell als dreidimensionale Formen
LoD2	<ul style="list-style-type: none"> • Erweiterung des LoD1 um Dach- und Fassadenstrukturen bzw. Oberflächentexturen • Auch Oberflächenmodell mit nahezu realistischen 3D-Visualisierungen
LoD3	<ul style="list-style-type: none"> • Architekturmodell • Präzisiert die Oberflächenstruktur des LoD2, bspw. Türen oder Fenster
LoD4	<ul style="list-style-type: none"> • Innenraummodell • Innere Bauwerksstrukturen werden ergänzt und somit werden auch Räume und Stockwerke beschrieben

Tabelle 1: Detailierungsstufen von LoD, nach Borrmann, et al. 2015⁴⁹

Im Anhang 1 wird jeweils ein Beispiel für das gleiche Gebäude mit der Detaillierung LoD1 und LoD2 als 2D-Visualisierung bzw. bei LoD2 auch als 3D-Ansicht präsentiert. Der Unterschied dieser Datensätze ist anhand der verschiedenen Elemente der jeweiligen Detaillierung zu erkennen. So werden im LoD2-Gebäude, neben der Grundfläche, auch das Dach und dessen Form aufgezeigt. Zusätzlich kann durch LoD2 in der 3D-Ansicht zwischen verschiedenen Elementen, wie Mauern, Dächer oder Grund unterschieden werden.

Die Standarddatensätze für LoD1 und LoD2 sind bereits bundesweit verfügbar und liegen somit als GML oder Shape-Datei bereit. Beispielsweise bietet Sachsen-Anhalt über sein Geoportal, das „Ivermgeo“, die Möglichkeit, die 3D-Gebäudemodelle als Testdaten herunterzuladen.⁵⁰

Für die Vorhaben dieser Forschungsarbeit wird der LoD2-Datensatz herangezogen, da dieser zum einen alle Gebäude des ALKIS umfasst und zum anderen vereinfachte Dachformen und Fassadenstrukturen beinhaltet. Diese sind für spätere Umsetzungen essenziell, um definierte Faktoren für Algorithmen, wie Höhe des Hauses, Dachform, Dachneigung, usw., auswerten zu können. Der LoD2-Datensatz wird dabei durch den Gebäudegrundriss definiert und durch

⁴⁹ Vgl. Borrmann, et al. 2015, S. 186f.

⁵⁰ https://www.lvermgeo.sachsen-anhalt.de/de/download_lod2.html.

Laserscanning-Daten ergänzt, um somit die Dacheigenschaften zu bestimmen. Wichtig ist hierbei, dass Gebäude bzw. Objekte, welche spezielle Dachformen besitzen oder unter 13 m² liegen immer ein Flachdach zugewiesen bekommen. Diese könnten später Berechnungen verfälschen oder müssen zuvor identifiziert und beseitigt werden. Die Gebäudemodelle entsprechen der Genauigkeit der Flurkarte, von welcher die Grundrisse des Gebäudes entnommen werden. Die Höhengenaugigkeit liegt bei ca. 20-30cm des realen Gebäudes, jedoch werden Abweichung bis zu 1 m nicht ausgeschlossen.⁵¹ Die Datensätze werden für die Aktualität jährlich erneuert und somit können auch neue Objekte identifiziert und einbezogen werden. Allein in Sachsen-Anhalt konnten 2021 somit ca. 1,77 Mio. Gebäude verzeichnet werden.⁵²

2.4 ETL-Prozess

Um eine Vielzahl an Daten aus verschiedenen Systemen bzw. Quellen verarbeiten zu können, müssen diese zuvor auf eine einheitliche und konsistente Datenbasis überführt werden. Dieser Prozess, von der Datenextraktion aus unterschiedlichsten Quellen über die Aufbereitung bzw. Formatierung dieser Daten bis hin zum Laden der Daten, wird als ETL-Prozess bezeichnet. Das Ziel eines ETL-Prozesses ist es somit, dass Probleme und Konflikte zwischen den Daten behandelt werden, damit die extrahierten Daten in strategische Informationen umgewandelt werden können, welche später dem Nutzer einen Mehrwert bieten.⁵³ Die folgende Abbildung soll einen vereinfachten ETL-Prozess zur Verdeutlichung aufzeigen. Dabei steht die Abkürzung ETL für:

- Extract (die regelmäßige Beschaffung der Daten aus verschiedenen Systemen),
- Transform (die Festlegung grundlegender Formatierungen und Bereinigung der Daten),
- Load (das Laden der Daten in ein Zielsystem).⁵⁴

⁵¹ Vgl. Bayerische Vermessungsverwaltung 2018, S. 2.

⁵² Vgl. Bundesamt für Kartographie und Geodäsie 2022, S. 4.

⁵³ Vgl. Ponniah 2010, S. 283.

⁵⁴ Vgl. Farkisch 2011, S. 59.

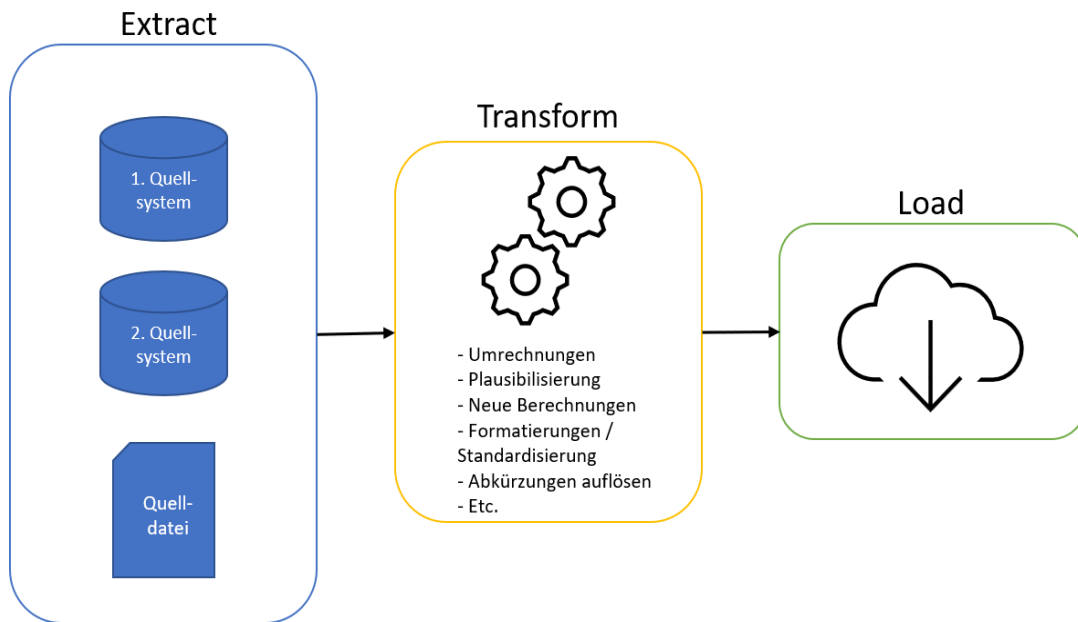


Abbildung 3: ETL-Prozess
Datenquelle: eigene Darstellung

Wie der Abbildung zu entnehmen ist, werden die Daten in der Extraktion aus verschiedenen Quellsystemen gesammelt. Dabei stellen die Quellsysteme üblicherweise eine Datenbank dar, von welcher Daten abgerufen werden. Doch ebenso ist es möglich auch nur bestimmte Dateiformate, wie Excel-Dateien o.ä. als Quelle der Daten zu verwenden. Bei der Extraktion werden regelmäßige Selektionen geplant, bei denen die Quelldaten gesammelt werden und somit das Zielsystem immer auf dem neuesten Stand ist. Bereits bei der Selektion sollte bekannt sein, welche Daten eine Relevanz für das spätere Zielsystem darstellen und demnach sollten auch nur diese Daten für die Selektionen eingeplant werden. Auch sollte zuvor geplant werden, in welchen Zyklen die Daten extrahiert werden. Dabei kommt es auf den Anwendungsfall an, ob die Daten periodisch, auf Abruf, ereignisgesteuert oder sofort bei der Änderung der Quelldaten abgerufen werden sollen.⁵⁵ Hier sollten die Beteiligten im Hinblick auf den Anwendungsfall, den sinnvollsten Zyklus wählen. Wenn die Quelldaten selektiert wurden, werden diese an den Arbeitsbereich der Transformation weitergeleitet.

Die obige Abbildung zeigt einige Beispiele, welche Prozesse bzw. Änderungen während der Transformation stattfinden können. Ziel der Transformation besteht dabei in der Anpassung der Quelldaten auf ein einheitliches Standardformat, welches dann in das Zielsystem überführt werden kann. Somit spielt es in der Transformation eine große Rolle wie die definierte Datenstruktur aufgebaut werden soll. Dabei kann es auch sein, dass durch die Vielzahl an

⁵⁵ Vgl. Farkisch 2011, S. 60f.

Quelldaten mehrere Transformationen durchlaufen werden müssen. Während der einzelnen Transformation werden die Quelldaten teilweise manipuliert, neu formatiert oder auch aus den bestehenden Daten neue Daten abgeleitet.⁵⁶ Bei der Transformation im Allgemeinen können folgende Arten unterscheiden:

- Datenabgleich: Quelldaten werden mit einer Zielstruktur abgeglichen und in diese umgewandelt,
- Datenvalidierung: Quelldaten werden geprüft und identifizierte inkonsistente Daten werden bereinigt bzw. korrigiert,
- Datenfilterung: Quelldaten werden nach bestimmten Selektionskriterien gefiltert.⁵⁷

Innerhalb dieser Arten der Transformation können verschiedene Funktionen verwendet werden, welche die Umwandlung der Quelldaten in ein konsistentes und nützliches Format ermöglichen. Die folgende Tabelle stellt Beispiele für diese Funktionen vor.

Funktionen der Transformation	Beschreibung
Splitten / Joinen	Aufteilen oder Zusammenführen der Daten in logische Bestandteile.
Konvertierung	Standardisierung der Quelldaten und die Aufbereitung in ein lesbares Format (bspw. das Datumsformat YYYYMMdd in dd.MM.YYYY).
Anreicherung	Vereinfachung und Umgestaltung einzelner Felder.
Dekodierung von Daten	Bei verschiedenen Quellsystem kann es zu unterschiedlichen Kodierungen von Datensätzen kommen (bspw. Quellsystem 1 verwendet "€" und Quellsystem 2 verwendet "Euro").
Ableiten von Daten	Aus den Quelldaten werden neue Daten berechnet und abgeleitet.
Duplizieren von Daten	Bei der Zusammenführung der Quelldaten kann es sein, dass Duplikate benötigt werden (bspw. Datenmapping von Personen auf Gebäude)
Plausibilisierung	Die Quelldaten werden auf Widersprüchlichkeiten und Konflikte geprüft.

Tabelle 2: Beispiele für Funktionen der Transformation^{58 59}

⁵⁶ Vgl. Ponniah 2010, S. 295f.

⁵⁷ Vgl. Farkisch 2011, S. 62f.

⁵⁸ Vgl. Farkisch 2011, S. 64. Vgl. Ponniah 2010, S. 297f.

⁵⁹ Vgl. Ponniah 2010, S. 297f.

Wenn nun die Quelldaten durch den Transformationsprozess in die Zielstruktur überführt wurden, können die Daten an die Phase des Ladens übergeben werden.

In der letzten Phase des ETL-Prozesses werden die nun konsistenten und validierten Daten in das Zielsystem überführt und stehen dort für weiterreichende Analysen und Bedarfe bereit. Beim Laden der Daten sollte jedoch beachtet werden, dass das Zielsystem während des Prozesses offline sein muss, um die Daten zu migrieren. Eine strukturierte Planung des Zyklus ist somit unerlässlich, da das Laden nur dann stattfinden kann, wenn das Zielsystem offline sein kann.⁶⁰

Neben dem ETL-Prozess gibt es auch eine abgewandelte Form des Prozesses das ELT. Dieser verwendet ebenso die gleichen Schritte wie der ETL-Prozess, also Extraktion, Transformation und Laden. Jedoch werden hier die Rohdaten erst in einer Datenbank gespeichert und später direkt bei dem Laden der Daten, die Transformation durchgeführt. Bei diesem Ansatz befinden sich also sowohl die Rohdaten als auch die transformierten Daten auf dem Zielsystem und können für Analysen und andere Zwecke verwendet werden.⁶¹ Die großen Vorteile bei der Verwendung des ELT-Prozesses liegt darin, dass zum einen die Performance gesteigert werden kann, da die Daten direkt zur benötigten Zeit transformiert werden. Zum anderen wird aber auch die Flexibilität gesteigert, da nicht nur die transformierten Daten, sondern auch die Rohdaten für Analysen bereitstehen. Somit entstehen durch den ELT-Prozess noch andere Möglichkeiten die benötigten Daten bereitzustellen.⁶²

Die Daten, welche in das Zielsystem überführt wurden, können nun für verschiedensten Analysezwecke als auch Auswertungen herangezogen werden. Somit wird es ermöglicht, dass aus unterschiedlichen Daten neue Erkenntnisse bzw. Wissen generiert werden kann. Durch das Aufstellen eines ETL- bzw. ELT-Prozesses soll es später ermöglicht werden, sowohl die georeferenzierten Daten als auch die demographischen Daten von Kommunen auf einem Zielsystem zu vereinen. Der genaue Prozess wird im Kapitel zur Umsetzung vorgestellt und analysiert.

⁶⁰ Vgl. Ponniah 2010, S. 303.

⁶¹ Vgl. Keboola 2022, S. 13.

⁶² Vgl. Keboola 2022, 14f.

3 Datengrundlage am Beispiel Mücheln (Geiseltal)

In diesem Kapitel wird nun die zu untersuchende Kleinstadt Mücheln (Geiseltal) vorgestellt. Dabei sollen die zur Verfügung stehenden Daten aufgezeigt werden, welche dann später die Grundlage für die Ist-Analyse bilden. Dies ist vor allem wichtig, da die vorhandenen Daten ebenfalls aufzeigen, welche Daten noch benötigt werden. Vorstellung der Stadt

3.1 Vorstellung der Stadt Mücheln (Geiseltal)

Für diese Arbeit wurde die Stadt Mücheln (Geiseltal) im Saalekreis innerhalb von Sachsen-Anhalt als Untersuchungsobjekt ausgewählt. Die Stadt Mücheln lässt sich geografisch gesehen südwestlich des Geiseltalsees und somit auch südwestlich der Stadt Halle (Saale) verorten. Mit ca. 8500 Einwohnern und einer Fläche von 98,57 km² lässt sich Mücheln nach der Definition der BBSR als „kleine“ Kleinstadt einordnen.^{63,64} Die Stadt umfasst zusätzlich die fünf einzelne Ortschaften Branderoda, Gröst, Langeneichstädt, Oechlitz und Wünsch. Die folgende Abbildung zeigt die einzelnen Ortschaften der Stadt Mücheln.

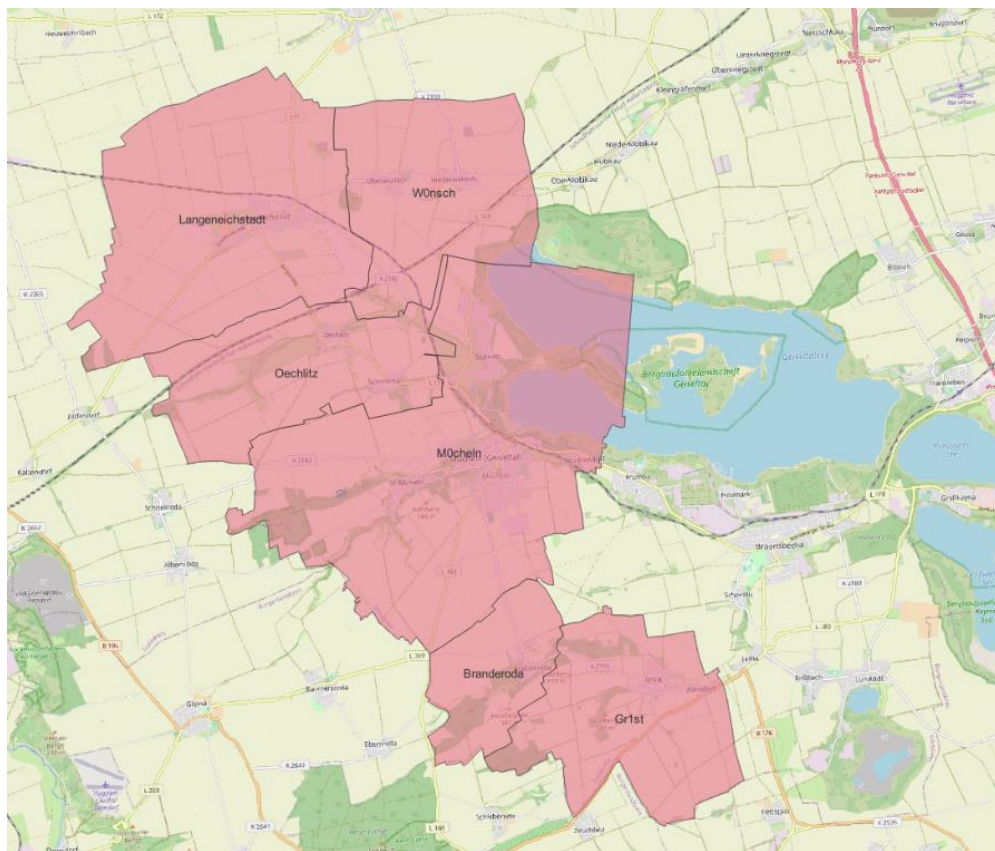


Abbildung 4: Gemarkungen der Stadt Muecheln (Geiseltal)
Datenquelle: eigene Darstellung im QGIS

⁶³ Vgl. BBSR.

⁶⁴ Vgl. Statistisches Landesamt Sachsen-Anhalt 2022.

Seit 1698 war Mücheln ein zentraler Treiber der Braunkohleförderung im Geiseltal und wurde dadurch wirtschaftlich geprägt. Als 1993 der Bergbau im Geiseltal eingestellt und das Abbaugbiet geflutet wurde, entstand der Geiseltalsee als neues touristisches Ziel in der Region.⁶⁵ Die Stadt Mücheln vollzog somit schon vor der Jahrtausendwende den ersten tiefgreifenden Strukturwandel von einem Industrie- bzw. Braunkohlestandort hin zu einem touristischen Zentrum am Geiseltal. Nun befindet sich die Stadt in dem nächsten Strukturwandel und wird stark durch die alternde Bevölkerung und die abgeschwächte Wirtschaft geprägt. Zudem kam es durch die Einstellung des Braunkohleabbaus bereits zu einem starken Bevölkerungsrückgang, u.a. durch Abwanderung überwiegend jüngerer Bevölkerung in attraktive Arbeitsmarktregionen. Insgesamt führten diese verschiedenen Entwicklungen zu einer rückläufigen Bevölkerungsanzahl und alternden Bevölkerungsstruktur. Nun erhofft sich die Stadt Mücheln durch den Einstieg in neuartige Technologien, wie Industrie 4.0, digitale Gesundheitsvorsorge sowie attraktive Wohn-, Kultur- und Erholungsangebote im Zuge des flächendeckenden Breitbandausbaus, eine Chance auch als ländliche Region wieder die wirtschaftliche Stärke zu erhöhen und ihre Bevölkerungsdichte zu verbessern. Auch die ansässigen Bewohner der Stadt Mücheln sollen weiterhin gebunden und in die Stadtentwicklung einbezogen werden. Die wichtigsten Themen rund um die Stadtentwicklung, mit welchen sich die Stadt Mücheln derzeit beschäftigt, sind:

- Wohnen (Sozialstruktur, Leerstandsmanagement, Modernisierung, usw.)
- Flächenmanagement
- Breitbandausbau
- Demografie.

Im Rahmen dieser Arbeit sollen die Themen Demografie, Wohnen und Flächenmanagement, durch die Nutzung georeferenzierter Daten sowie der Verwaltungsdaten der Stadt Mücheln analysiert werden. Zusätzlich sollen die Daten in Verbindung mit den Geodaten visualisiert und durch Algorithmen neue Erkenntnisse gewonnen werden, ob diese zu einer transparenteren, effizienteren und moderneren Stadtplanung beitragen können. Die Stadt Mücheln erhofft sich durch die Verwendung von Geodaten neue Erkenntnisse über die Strukturen der Stadt zu erlangen und dadurch neue Entscheidungsprozesse aufstellen zu können. Für die Umsetzung entschied sich die Stadt für die Open Source Software QGIS. Die Software wird derzeit schon für einzelne Analysen verwendet und wurde bereits in vorhergehenden Forschungsarbeiten eingesetzt. Der große Vorteil an der Verwendung von QGIS sind zum einen

⁶⁵ Vgl. mediaprint infoverlag gmbh 2020, S. 5.

die vielen Plugins, welche von der Community bereitgestellt werden und zum anderen auch die Möglichkeit eigene Plugins als Erweiterungen zu entwickeln. Somit können im QGIS eigene Lösungen implementieren, welche dann später produktiv genutzt werden können. Des Weiteren bietet QGIS auch verschiedene Möglichkeiten, die analysierten Daten in einer anschaulichen Weise zu visualisieren. Auch die Verbindung zu einer Datenbank stellt mit QGIS kein Problem dar und somit können beispielsweise Geodaten über QGIS auf den Datenbanken gespeichert oder von diesen abgefragt werden.

3.2 Untersuchungsobjekt

Da im Rahmen dieser Arbeit nicht alle Anwendungsfälle für georeferenzierte Daten im Bereich der kommunalen Verwaltung aufgezeigt werden können, wurde ein bestimmter Anwendungsfall ausgewählt. Dabei wurde ein aktuelles Thema gewählt, welches sich später auch auf andere Kleinstädte gut anwenden lassen soll. Der Anwendungsfall soll sich mit der Verteilung von Kindergärten in Kleinstädte beschäftigen und durch einen Planungsalgorithmus Gebäude und Standorte ermitteln, welche als Kindergarten genutzt werden können. Dieses Thema wurde gewählt, da es für Familien heutzutage einen großen Wert darstellt, einen Kindergartenplatz zu bekommen. Ein bedarfsgerechtes Betreuungsangebot gilt als grundlegende Voraussetzung, um Familie und Beruf zu verbinden.⁶⁶ Die folgende Abbildung zeigt, wie der Betreuungsbedarf in Deutschland immer weiter wächst.

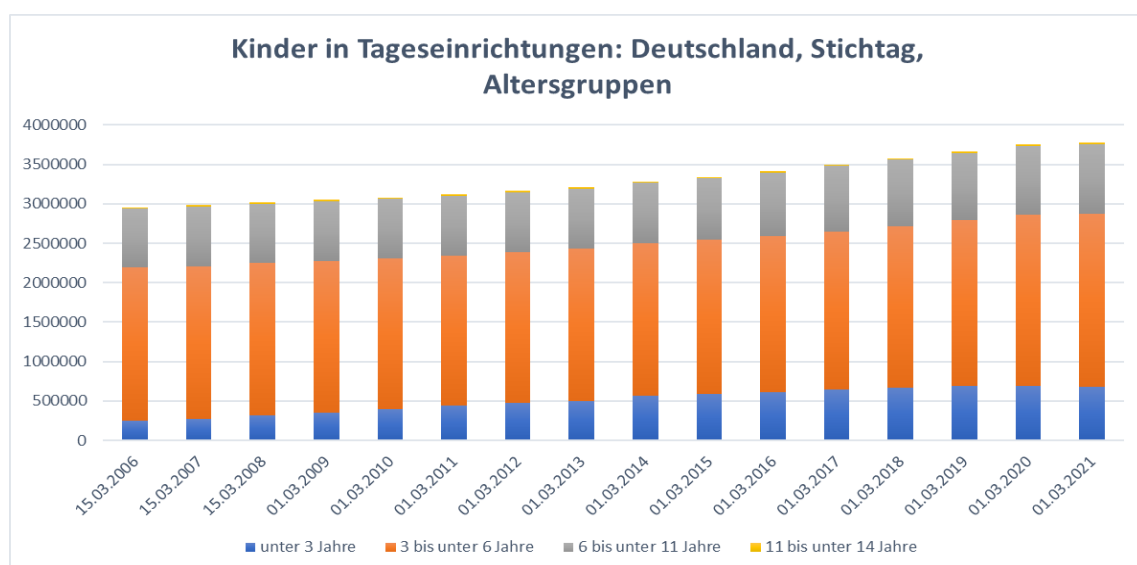


Abbildung 5: Kinder in Tageseinrichtungen: Deutschland, Stichtag, Altersgruppe
Datenquelle: Statistisches Bundesamt (2022)⁶⁷

⁶⁶ Vgl. Wido Geis-Thöne 2020, S. 1.

⁶⁷ Vgl. Statistisches Bundesamt 2022.


Doch bereits 2020 konnte ein Mangel an ca. 342.000 Kindergartenplätzen verzeichnet werden.⁶⁸ Es konnte auch festgestellt werden, dass gerade in den Altersgruppen der 1-3-jährigen Kinder die Betreuungsnachfrage immer weiter steigt. Diese Bedarfssteigerung lässt sich vor allem mit der Veränderung der Erwartung von Eltern an die Betreuung ihrer Kinder erklären. So wollen Eltern weiterhin erwerbstätig bleiben und Ihre Perspektiven am Arbeitsmarkt nicht gefährden.^{69,70} In dieser Situation des generellen Mangels an Betreuungsplätzen können Kleinstädte wie Mücheln etwas anbieten, was Großstädte meist nicht können, und zwar Platzangebote. In Mücheln lassen sich beispielsweise verschiedene Plätze ermitteln, welche derzeit leerstehen und für neue Zwecke bzw. Bedarfe verwendet werden könnten. Dann müssen diese Standorte vorerst identifiziert und auf Ihre Nutzbarkeit geprüft werden. Zu beachten ist, dass nicht nur die Ermittlung freier und geeigneter Plätze für Betreuungsangebote hier Abhilfe schaffen kann, sondern auch noch das qualifizierte Personal für die Betreuung vonnöten ist. Die vorliegende Arbeit wird sich jedoch nur auf die Analyse und Bestimmung von geeigneten Standorten durch georeferenzierte Daten beschränken und den aktuellen Arbeitsmarkt nicht berücksichtigen.

Wichtig für die Analyse der vorliegenden Daten ist dabei, dass die Altersstruktur der jeweiligen Stadt einbezogen wird und somit Hotspots hinsichtlich des Alters bestimmt werden können. Durch die Daten der Verwaltung können anhand einer deskriptiven und prädiktiven Analyse verschiedene Erkenntnisse zur Entwicklung der Stadt getroffen werden. In diesem Fall soll aufgezeigt werden, in welchem Bereich der Stadt vor allem jüngere Menschen leben. Diese Bestimmung ist wichtig, damit eine zeitliche Veränderung aufgezeigt werden kann. Durch das Wissen über die Veränderung der Altersstruktur über einen zeitlichen Bestimmungsraum lässt sich somit auch ermitteln, welcher Standort hinsichtlich der Altersverteilung am profitabelsten wäre. Die genaue Analyse der Daten wird im nächsten Kapitel durchgeführt und bildet eine Grundlage für den Planungsalgorithmus. Eine weitere Grundlage bildet die Verbindung von demografischen- und georeferenzierten Daten auf eine Datenbasis, sodass für ein Geoobjekt festgelegt werden kann, welche demografischen Daten an diesem zu verorten sind. In folgender Abbildung wird gezeigt, wie die demografischen Daten an einem Gebäude hinterlegt werden können. Hier kann man erkennen, dass an dem Gebäude mehrere Einträge hinterlegt sind, welche auf mehrere Personen an dem Gebäude hinweisen.

⁶⁸ Vgl. Wido Geis-Thöne 2020, S. 1.

⁶⁹ Vgl. Wido Geis-Thöne 2020, S. 1.

⁷⁰ Vgl. Peucker et al 2003, S. 9.



Objekt	Wert
LAND	Deutschland
GEMEINDE	Mücheln (Geiseltal), Stadt
STRASSE	Eptinger Rain
HAUSNUMMER	67
STR_HNR	Eptinger Rain 67
NAME	NULL
RELHOEHE	16,611
ABSH_BODEN	156,516
ABSH_FIRST	173,127
ABSH_TRAUF	171,083
DACHNEIG	68,226
DACHORIENT	-95,743
DACHFLAECH	145,785
DACHTYP	GableRoof
ELEMCLASS	Ground
BldgRoot	DESTLIKA0001gkcz
BldgPart	DESTLIKA0001gkcz
BldgParent	NULL
BldgPartx	0
str_name	Eptinger Rain
addr_hnr	67
addr_hnra	NULL
addr_plz	6249
ker_gebdat	19971221
ker_gebort	Merseburg
ker_gesch	M
ker_ein	20200928
ker_aus	NULL
ker_age	24
ker_years_living	1
ker_still_living	True
> NAME	NULL
> NAME	NULL
> NAME	NULL
> NAME	NULL
> NAME	NULL
> NAME	NULL
> NAME	NULL
> NAME	NULL

Abbildung 6: Geoobjekt mit demografischen Daten
Datenquelle: eigene Darstellung im QGIS

Die zuvor aufgezeigten Daten bilden die Grundlage für die Aufstellung von verschiedenen Planungsalgorithmen. Jedoch müssen auch noch weitere Faktoren für den spezifischen Anwendungsfall Kindergarten analysiert und in die Betrachtung einbezogen werden. Dafür wurde ein kurzes Interview mit der Kinderland Geiseltal GmbH durchgeführt, indem die Voraussetzungen für ein Gebäude, welches als Kindergarten genutzt werden soll, diskutiert wurden.

Bei diesem Interview betraf eine der entschiedensten Fragen die ideale Fläche eines Kindergartens. Diese Frage ließ sich so jedoch nicht beantworten, da es nicht die eine richtige Fläche für einen Kindergarten gibt, sondern die Fläche im Wesentlichen abhängig von der Anzahl der zu betreuenden Kinder ist. Somit konnte uns die Kinderland Geiseltal GmbH sagen, dass für ein Krippenkind im Allgemeinen mindesten 2,5-3,5 m², für ein Kindergartenkind mindesten 5 m² und für ein Hortkind mindestens 10 m² eingeplant werden müssen. Aus diesen Angaben müsste dann der ideale Flächenbedarf bestimmt werden. Dabei wurde angemerkt, dass es heutzutage nicht unüblich ist, dass die Krippenzeit für Kinder schon ab 3 Monaten nach der Geburt beginnen kann. Eine weitere Frage, welche im Rahmen des Interviews beantwortet wurde, war ob es Vorgaben für die Platzwahl eines Kindergartens gibt und dieser beispielsweise in einem verkehrsberuhigten Bereich liegen sollte. Hier war die

Antwort der Kinderland, dass es für die Verortung des Kindergartens in erster Linie keine Vorgaben gibt, jedoch die Auswahl der Lokation eine Rolle für das Thema des Kindergartens spielen kann. So wurde gesagt, dass beispielsweise auch ein Kindergarten in einem Gewerbegebiet ein interessante bzw. ansprechende Lokation darstellen kann. Die Begründung für diese Antwort war, dass die Kinder somit die Strukturen der heutigen Gesellschaft schon besser kennen- und verstehen lernen. Somit lassen sich zum Beispiel in einem Gewerbegebiet verschiedenste Unternehmungen planen, welche den Kindern hinsichtlich Ihrer Bildung einen Mehrwert bringen. Jedoch wurde hier auch betont, dass natürlich viele Faktoren wie z.B. die Lärmbelästigung eine große Rolle spielen. Somit lässt sich festhalten, dass erst einmal alle Gebäude für den Algorithmus relevant sind und bis auf wenige Ausnahmen alle Gebäude einbezogen werden können. Ausnahmen stellen dabei Gebäude dar, welche bereits für eine bestimmte Nutzung vorgesehen sind, wie Kirchen, das Rathaus und andere funktionspezifische Gebäude. Zu der Frage, welche sonstigen Parameter für den Kindergarten betrachtet werden sollten, gab es verschiedene Antworten der Kinderland. Zum einen wäre ein Flachbau mit nur einem Stockwerk anderen Gebäuden vorzuziehen, da diese aus Sicht des Brandschutzes die bessere Wahl darstelle. Zum anderen sollte nicht nur das Gebäude betrachtet werden, sondern auch die Fläche des Außenbereiches, welche eine wichtige Rolle spielt. Hier gibt es zwar keine genauen Vorgaben, welche zu berücksichtigen sind, jedoch sollte die Fläche genügend Platz für Spiel- bzw. Erkundungsflächen bieten. Somit lässt sich auch hier für den Algorithmus festhalten, dass dieser auch die Fläche des Flurstückes berücksichtigen muss, auf dem das Gebäude zu verorten ist und diese dann nach der Größe priorisiert. Letztlich gab es noch einen Faktor, der nicht zwingend berücksichtigt werden muss, jedoch für einen guten Kindergarten dennoch eine wichtige Rolle darstellt. Hierbei handelt es sich um die Erreichbarkeit bzw. eine einfache Verbindung zu Parkfläche für den Kindergarten.

3.3 Bereitgestellte Daten

Für die Umsetzung des geplanten Algorithmus wurden verschiedene Daten durch die Stadt Mücheln bereitgestellt. Zum einen waren dies die Daten des amtlichen Liegenschaftskastasterinformationssystems (ALKIS). Dabei handelt es sich um georeferenzierte Daten, welche von dem Landesamt für Vermessung und Geoinformation Sachsen-Anhalt (LVermGeo) bereitgestellt werden. Diese bieten verschiedene Geobasisdaten, wie 3D-Gebäudemodelle, Flurstücke, Hausumringe und vieles mehr an. Dabei hat die Stadt Mücheln durch eine Vereinbarung mit dem LVermGeo den LoD2-Datensatz für Ihre Gebiet erhalten. Hierbei handelt

es sich um die komplette Datenstruktur mit allen vorhandenen Daten und nicht nur um den kostenlosen Datensatz mit einer beschränkten Struktur. Der LoD2-Datensatz wurde dabei als Shapefile zur Verfügung gestellt. Ein Shapefile ist eine nichttopologische Geometrie für die räumlichen Merkmale eines Datensatzes. Die Geometrie des Objektes in einem Shapefile wird dabei als eine Menge von Vektorkoordinaten gespeichert.⁷¹ Neben den LoD2 Daten stehen auch eine Vielzahl von Open Data zur Verfügung, welche für die Umsetzung herangezogen werden können.

Zum anderen wurden demografische und personalisierte Daten von der MESO-Datenbank der Stadt Mücheln zur Verfügung gestellt, welche durch das Löschen der Namen anonymisiert wurden. Durch ein Projekt, welches bereits durchgeführt wurde, sollten die Geodaten so aufbereitet werden, dass die Adresse mit den personenbezogenen Daten auf die richtige Geokodierung des Gebäudes aus dem LoD2-Datensatz überführt wird. Hierbei gab es jedoch einige Schwierigkeiten bei dem Mapping der richtigen Adressen. Diese gingen aus unterschiedlichen Schreibweisen und teilweise unterschiedlichen Strukturen der Daten hervor. Somit konnten die Daten nur durch ein zeitaufwendiges manuelles Vorgehen überführt werden. Da jedoch die Konsistenz und die Qualität der Daten eine wichtige Rolle für eine unverfälschte Analyse darstellen, wurde im Rahmen der Umsetzung zunächst ein Programm entwickelt, welches die verschiedenen Daten automatisiert kombiniert. Ziel war es dabei, dass ein Großteil der Daten richtig überführt wird, ohne manuelles Vorgehen und ohne Fehler bei der Vereinigung der Datensätze. Das genaue Vorgehen der Umsetzung wird in Kapitel 6 dieser Arbeit erläutert.

⁷¹ Vgl. ESRI 1998, S. 1.

4 Analyse der Datenbasis auf kommunaler Ebene

In diesem Kapitel werden nun die verschiedenen Datensätze, welche zur Verfügung stehen, analysiert. Die Analyse soll später die Grundlage für die Aufstellung des Algorithmus bilden und die Beziehungen der einzelnen Datensätze zueinander aufzeigen. Zusätzlich bildet die Analyse eine Möglichkeit, das Ergebnis des Algorithmus zu validieren.

4.1 Verwaltungsinterne Daten

Wie bereits im vorherigen Kapitel beschrieben, liefert die Stadt Mücheln verschiedene Datensätze für die Erarbeitung des Planungsalgorithmus. Somit wurde ein Zugang zur MESO-Datenbank der Stadt Mücheln bereitgestellt. Diese Datenbank enthält alle Informationen des Einwohnermeldeamtes und ermöglicht den Zugriff auf die demographischen Daten der Stadt. Die Datenbank umfasst verschiedene Tabellen mit unterschiedlichen Daten, die für die Bearbeitung benötigt werden. Die benötigten Daten für die Aufstellung des Algorithmus mussten somit vorerst durch eine Abfrage auf verschiedene Tabellen selektiert werden. Letztlich ergab sich folgende Struktur des Datensatzes.

Name	Beschreibung	Datentyp
addr_ort	ID des Ortes	Integer
ort_name	Name des Ortes	String
addr_str	ID der Straße	Integer
str_name	Name der Straße	String
addr_hnr	Nummer der Straße	Integer
addr_hnra	Zusatz der Straße	String
addr_hnra2	Zusatz der Straße	String
addr_plz	Postleitzahl	Integer
ker_gebdat	Geburtsdatum	Integer (Format: YYYYmmdd)
ker_gebort	Geburtsort	String
ker_gesch	Geschlecht der Person	String
ker_ein	Einzugsdatum	Integer (Format: YYYYmmdd)
ker_aus	Auszugsdatum	Integer (Format: YYYYmmdd)

Tabelle 3: Datenstruktur der demografischen Daten aus der MESO DB

Die aufgestellte Datenstruktur stellt die wichtigsten Daten für den Algorithmus bereit. Zum einen die Adresse bestehend aus dem Straßennamen, der Hausnummer und dem Zusatz, welche den Schlüssel zu jedem Eintrag darstellen und zum anderen die personenbezogenen Daten. Diese setzen sich zusammen aus dem Geburtsdatum der Person, dem Geburtsort, dem Geschlecht und letztlich dem Datum der Zuwanderung nach Mücheln und falls vorhanden

dem Datum der Abwanderung aus Mücheln. Wichtig ist noch einmal zu erwähnen, dass die Datensätze anonymisiert wurden, durch das Entfernen des Vor- und Nachnamens.

Am Ende konnten somit ca. 26500 Einträge selektiert werden, welche sich auf einen Zeitraum von 1900 - 2022 verteilen. Dabei sind in diesem Datensatz auch Einträge zu Bürgern, welche die Stadt bereits wieder verlassen haben, oder Bürgern, die bereits verstorben sind. Diese Einträge sind am Vorhandensein eines Auszugsdatums zu erkennen. Bei der Analyse des Datensatzes ist aufgefallen, dass bei verschiedenen Einträgen Eingabefehler vorliegen. Beispielsweise gibt es Einträge, bei denen das Datumsformat nicht eingehalten wurde und somit teilweise das Datum ein Format vorweist, welches nicht existiert. Ein Beispiel wäre die Angabe des Auszugsdatums in dem Format „19860000“. Diese fehlerhaften Daten müssen später von dem Algorithmus erkannt und als Mitteilung weggeschrieben werden.

Ein besonderes Augenmerk wurde bei der Analyse der Daten auf diejenigen Daten gelegt, welche für den Anwendungsfall des Planungsalgorithmus eine Rolle spielen. Somit wurde als erstes die aktuelle Altersstruktur der Stadt Mücheln analysiert, um einen allgemeinen Überblick zur aktuellen Lage zu bekommen. Dabei konnte ein durchschnittliches Alter von ca. 51 Jahren ermittelt werden, bei einer derzeitigen Bevölkerungsanzahl von ca. 9000 Menschen im Jahr 2022. Folgender Graph soll die Altersstruktur der Stadt Mücheln visualisieren und zeigt zugleich verschiedene Parameter auf.

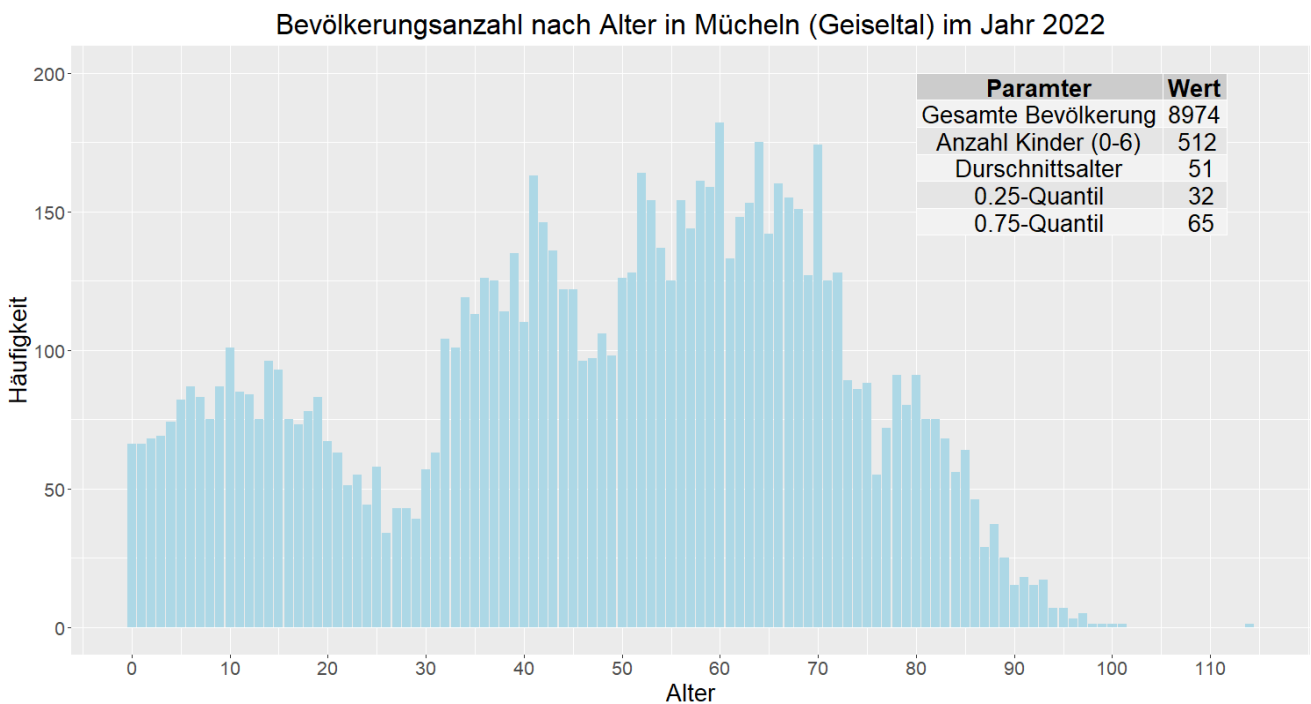


Abbildung 7: Bevölkerungsanzahl nach Alter in Mücheln (Geiseltal) im Jahr 2022
Datenquelle: eigene Darstellung

Zusätzlich zu der zuvor gezeigten Grafik findet sich noch eine weitere Abbildung im Anhang 2, welche die Altersstruktur in den einzelnen Ortsteilen aufschlüsselt. Anhand der Grafiken lässt sich gut erkennen, dass Mücheln als Stadt mitten im demografischen Wandel steckt. Einer der relevantesten Faktoren im demografischen Wandel ist per Definition die alternde Bevölkerung. Damit ist gemeint, dass die ältere Bevölkerung größer und auch schneller wächst als die Jüngere.⁷² Dies lässt sich auch in Mücheln gut erkennen. Wenn die Altersgruppe der 0-6-Jährigen mit der Altersgruppe 70-76 Jahre verglichen wird, dann kann erkannt werden, dass die ältere Bevölkerung um ca. 754 Menschen größer ist als die Gruppe der Jüngeren. Dennoch wurde versucht mit einer Analyse festzustellen, wie sich die Geburtenrate und somit die Anzahl an Kindern verändert hat. Dabei sollte festgestellt werden, ob ein generelles Wachstum der Anzahl an Kindern im Alter 0-6 vorliegt oder ob die Anzahl rückläufig ist. Schaut man sich nun folgenden Grafen an, kann erkannt werden, dass die Geburtenzahlen seit 2000 rückläufig sind und sich sogar teilweise halbiert haben.

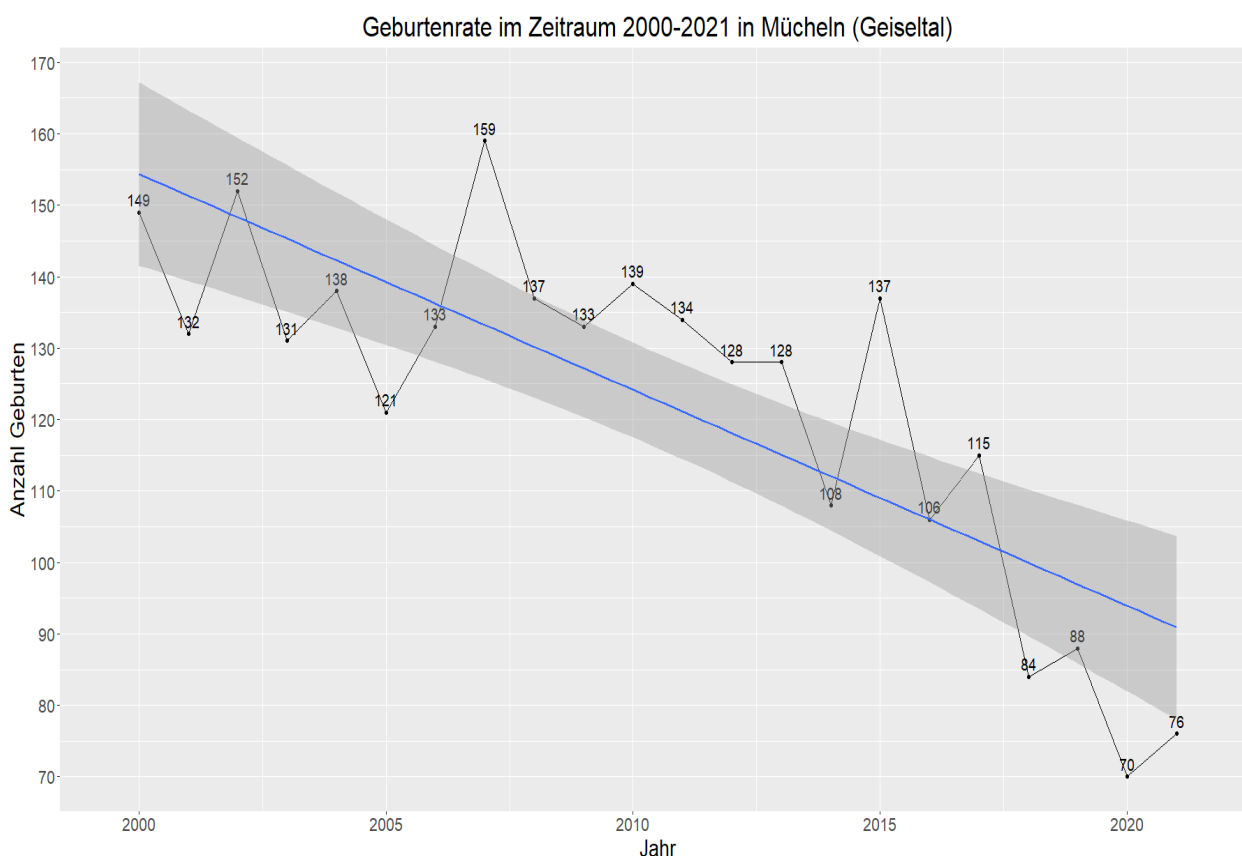


Abbildung 8: Geburtenrate im Zeitraum 2000-2021 in der Stadt Mücheln (Geiseltal)
Datenquelle: eigene Darstellung

⁷² Vgl. Bundeszentrale für politische Bildung 2022.

Doch allein die Betrachtung der Geburtenrate reicht nicht aus, um feststellen zu können, ob die jüngere Bevölkerung in Müheln wächst oder weiterhin rückläufig bleiben wird. Aus diesem Grund wurde zu den Geburten noch das Wachstum der Anzahl an Menschen analysiert, welche potenziell in einem Alter sind, Kinder zu bekommen. Die Altersspanne wurde dabei auf 15-49 Jahre festgelegt.⁷³ Diese Altersspanne ergibt sich aus einer Studie des Statistisches Bundesamtes, bei der das durchschnittliche Alter der Eltern bei der Geburt des Kindes untersucht wurde. Die Zeitspanne wurde dabei auf 10 Jahre in die Vergangenheit gelegt, also zwischen 2011 bis 2021. Um das Wachstum messen zu können, wurden nur die Einträge aus den Daten des Einwohnermeldeamtes genutzt, bei denen die Personen in den letzten 10 Jahren nach Müheln zugewandert oder aus Müheln abgewandert sind. Somit ergaben sich folgende Grafen, welche den Zusammenhang zwischen Jahr der Zu- oder Abwanderung und der Anzahl Personen je Jahr aufzeigt. Dabei lässt sich erkennen, dass für Müheln als Stadt ein positives Wachstum zu erkennen ist, da die Steigung der Geraden deutlich zu erkennen ist und es somit mehr Zuwanderungen gab als Abwanderungen. Schaut man sich z. B. Langeneichstädt als Ortsteil gesondert an, kann deutlich erkannt werden, dass hier die Abwanderung aus Langeneichstädt über die 10 Jahre rückläufig ist bei steigenden Zuwanderungen. Für die übrigen Ortsteile von Müheln befinden sich die Grafen ebenfalls im Anhang 2. Diese Erkenntnisse über das Wachstum der einzelnen Ortsteile soll später auch der Planungsalgorithmus erkennen und anhand eines Abgleichs mit den anderen Ortsteilen eine Entscheidung treffen, welcher Ortsteil den größten Zuwachs in den letzten Jahren erfahren hat. Dieses Thema wird später noch im Umsetzungskapitel abgehandelt.

⁷³ Vgl. Statistisches Bundesamt 2022.

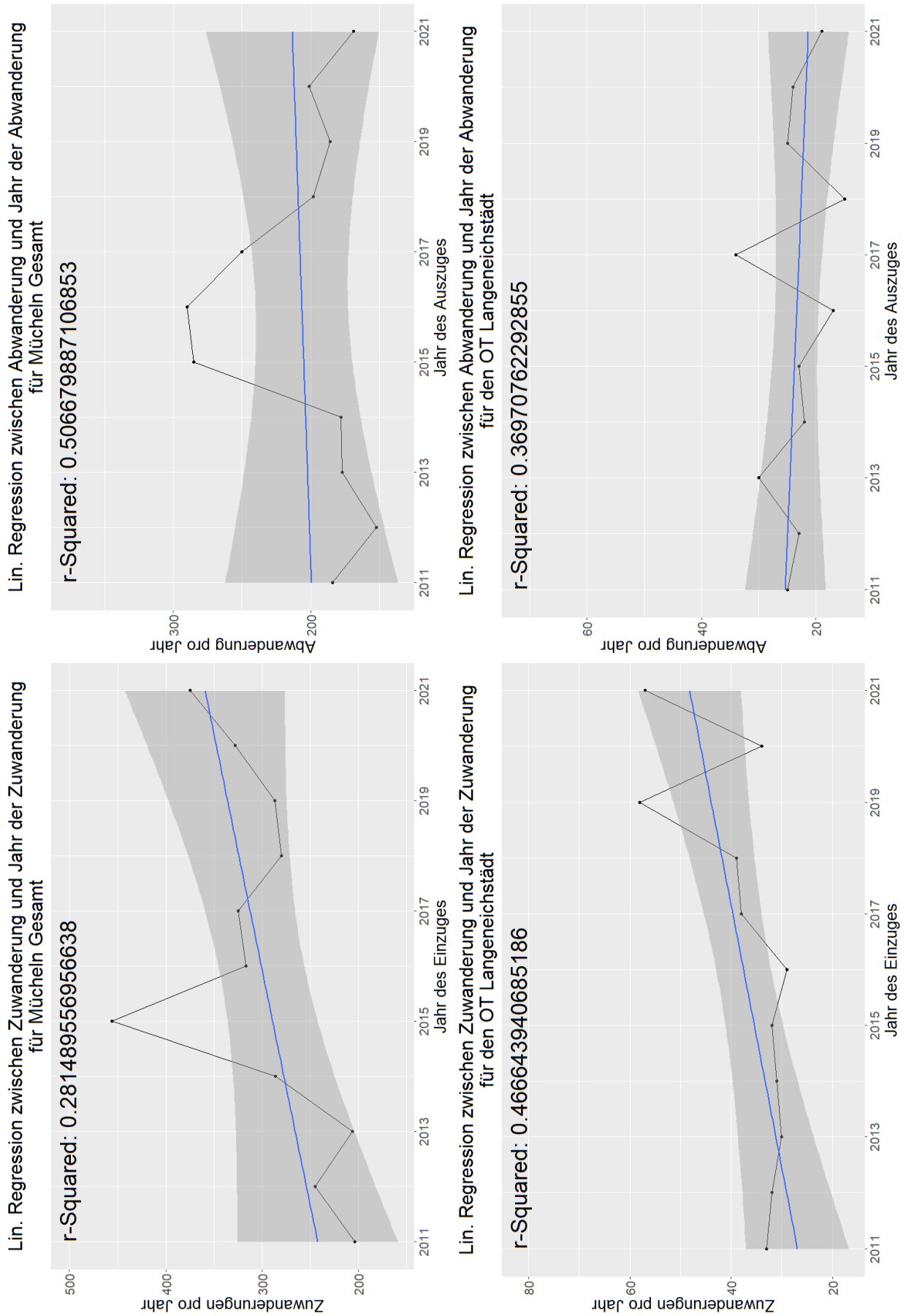


Abbildung 9: Lineare Regression für die Zu- und Abwanderung der Bevölkerung für Mueheln (Gesamt) und Langeneichstaedt
 Quelle: eigene Darstellung

4.2 Open Geodaten

Neben den Einwohnermeldedaten werden auch die georeferenzierten Daten durch die Stadt Mücheln bereitgestellt. Dabei handelt es sich um eine Menge an Datensätzen, welche einen geografischen Bezug zur Stadt Mücheln haben. Die wichtigste Rolle spielen dabei die LoD2-Daten, welche für jedes Gebäude der Stadt ein Geoobjekt liefern. Hierbei handelt es sich um ca. 156.000 Geoobjekte, welche die Gebäude abbilden. Der allgemeine Aufbau des Datensatzes wurde bereits im 2. Kapitel besprochen, nun soll noch der vorhandene Datensatz hinsichtlich des Fallbeispiels Kindergartenplanung untersucht werden. Für den Planungsalgorithmus spielt vor allem die hinterlegte Adresse eine wichtige Rolle, da diese - wie zuvor schon angesprochen - die Verbindung zwischen den Einwohnermeldedaten und einem Gebäude darstellt. Die Adresse teilt sich hierbei in zwei Felder auf, den Straßennamen und die Hausnummer, welche gleichzeitig auch den Hausnummernzusatz hält. Beim Abgleich des LoD2-Datensatzes mit den Daten des Einwohnermeldeamtes ist aufgefallen, dass der LoD2-Datensatz nicht auf aktuellem Stand ist. Dies lässt sich an verschiedenen Gebieten der Stadt Mücheln nachweisen, wo bereits neue Häuser erbaut wurden, diese jedoch noch nicht im LoD2-Datensatz enthalten sind. Im Folgenden sehen Sie eine Grafik, die am Beispiel der Straße Stöbnitzer Mitte aufzeigt, dass die Gebäude nicht hinterlegt und somit auch die Hausnummern 8a bis 8d nicht vorhanden sind. Dabei ist deutlich zu erkennen, dass Openstreetmap (OSM) diese Gebäude schon ermittelt hat.



*Abbildung 10: Aktualität des LoD2-Datensatzes am Beispiel Stöbnitzer Mitte
Datenquelle: eigene Darstellung*

Bei diesen Fällen muss darauf geachtet werden, dass diese bis zur Auslieferung der neuen Datensätze nicht vorhanden sind und somit auch bei Analysen nicht einbezogen werden. Des Weiteren spielt das Feld der Gebäudefunktion eine wichtige Rolle, wenn es darum geht, die bestehenden Kindergärten hinsichtlich des Planungsalgorithmus zu identifizieren. Jedoch musste hier festgestellt werden, dass die Gebäudefunktionen nicht die genau Funktion des

Gebäudes enthalten, sondern nur die Wertarten. So erscheint zum Beispiel ein Kindergarten als Gebäude für soziale Zwecke. Da das Ziel des Algorithmus ist, auch bestehende Kindergärten in die Berechnung einzubeziehen, stellt diese Verallgemeinerung der Gebäudefunktion ein großes Problem dar. Für die Beseitigung dieses Problems wurden zwei Methoden angewandt. Zum einen wurde versucht durch eine Abfrage über OSM, das Gebäude näher identifizieren zu können, zum anderen lässt sich die genaue Funktion auch manuell in den Datensatz korrigieren. Die genauere Beschreibung zum Vorgehen findet sich im Kapitel zur Umsetzung des Algorithmus. Anbei folgt noch eine Grafik, welche die Gebäudefunktion für einen Kindergarten in der Stadt Mücheln zeigt. Diese hat derzeit die ID „31001_3060“, welche für „Gebäude für soziale Zwecke“ steht. Zusätzlich zeigt die Grafik auch noch die per Abfrage des Gebäudes über die OSM-Schnittstelle ermittelten Daten.

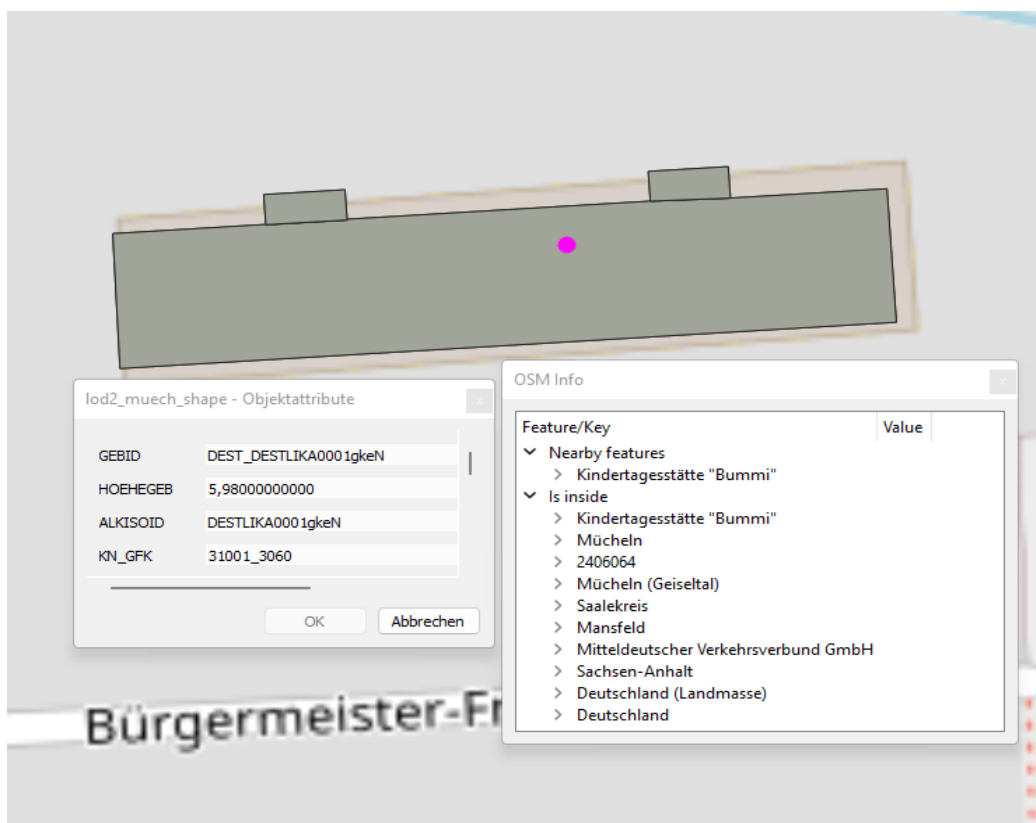


Abbildung 11: LoD2 Gebäudefunktion - „Gebäude für soziale Zwecke“ und OSM-Daten
Datenquelle: eigene Darstellung

Nach einer Absprache mit dem LVermGeo konnte festgestellt werden, dass die Gebäudefunktionen nicht mehr auf dem benötigten Detaillierungsgrad gepflegt werden. Es werden nur die „offiziellen Definition“ hinterlegt und somit werden Kindergärten mit der Gebäudefunktion „Gebäude für soziale Zwecke“ betitelt. Somit müssen bei der Umsetzung andere Wege verwendet werden, um dennoch die bestehenden Kindergärten identifizieren zu

können. Diese Möglichkeiten wären die Verwendung der OSM-Schnittstelle und das manuelle Pflegen der Funktionen im Datensatz. Des Weiteren stellte sich heraus, dass durch derzeitige technische Schwierigkeiten der LoD2-Datensatz für die Stadt Mücheln nicht aktualisiert wird. Aus diesem Grund werden die neu errichteten Gebäude in Mücheln erst später in dem LoD2-Datensatz ausgeliefert werden.

Zusätzlich zu den LoD2-Daten wurde für den Planungsalgorithmus der ALKIS-Datensatz benötigt, welcher die Nutzungsart für Fluren angibt. Dieser Datensatz soll beim Durchlauf des Algorithmus auch die Flächennutzung und den Besitzer der Fläche einbeziehen können. Dieser Faktor spielt eine Rolle, wenn es darum geht, ein Gebäude bzw. eine Fläche ermitteln zu können, welche die entsprechenden Voraussetzungen für einen Kindergarten erfüllen. Außerdem können über diese auch die Eigentümer bestimmt werden, sodass bestenfalls bereits im Besitz der Stadt liegende Fläche bzw. das Gebäude ermittelt werden können.

5 Experimentelle Umsetzung

Innerhalb dieses Kapitels soll aufgezeigt werden, wie die Geodaten und die Daten der Verwaltung in der Praxis genutzt werden können. Des Weiteren wird beschrieben, wie die Umsetzung einer Geodatenanwendung gelingt, sodass sich aus den Daten neue Erkenntnisse gewinnen lassen.

5.1 Vorbereitung der Umsetzung

5.1.1 Benötigte Software

Für die Analyse und die Gewinnung neuer Erkenntnisse bzw. Faktoren aus Geoinformationen, werden für den gewählten Ansatz verschiedene Softwareprodukte benötigt. Innerhalb dieses Abschnittes werden die benötigten Applikationen kurz aufgezeigt und auf deren Bedeutung hingewiesen. Die zentrale Applikation stellt dabei die Open Source Software QGIS dar. Dabei unterstützt die Software verschiedene Datenformate und Datenbanken, wie Shapefile, GeoTIFF, PostGIS und viele andere mehr.⁷⁴ QGIS bietet außerdem eigene Funktionen zum Analysieren von Geodaten und implementiert gleichzeitig noch andere Open-Source Geosoftware, wie SAGA, GDAL oder GRASS.⁷⁵ Durch diese Vielzahl an bereits vorhandenen Algorithmen, können durch die Verwendung von QGIS und den passenden Geodaten unterschiedlichste Analysen aufgestellt werden. Doch da es unzählige mögliche Varianten an Datensätzen, Analysen oder Prozessen gibt, welche durch Geodaten und QGIS verarbeitet werden könnten, gibt es nicht für jeden Anwendungsfall bereits einen integrierten Algorithmus oder Prozess. Aus diesem Grund bietet QGIS die Möglichkeit eigene Logiken oder Logiken aus der Community zu implementieren und diese über Plugins in die bestehenden Funktionen von QGIS zu integrieren.⁷⁶ Somit war es auch der erste Schritt eine eigene Erweiterung (Plugin) im QGIS zu erzeugen, bevor die eigentliche Umsetzung des Planungsalgorithmus beginnen konnte. Mit einer bereits existierenden Erweiterung von QGIS, dem „Plugin-Builder“ ist es möglich, ein eigenes Plugin zu erzeugen.⁷⁷ Dieses Plugin kann eigene Eigenschaften haben, wie einen Namen o.ä., und nach der Konfiguration und Generierung des Plugins wird automatisch eine Standardklasse für das Plugin erzeugt. Diese Klasse ermöglicht es, dass das Plugin im QGIS angezeigt wird und ausgeführt werden kann.

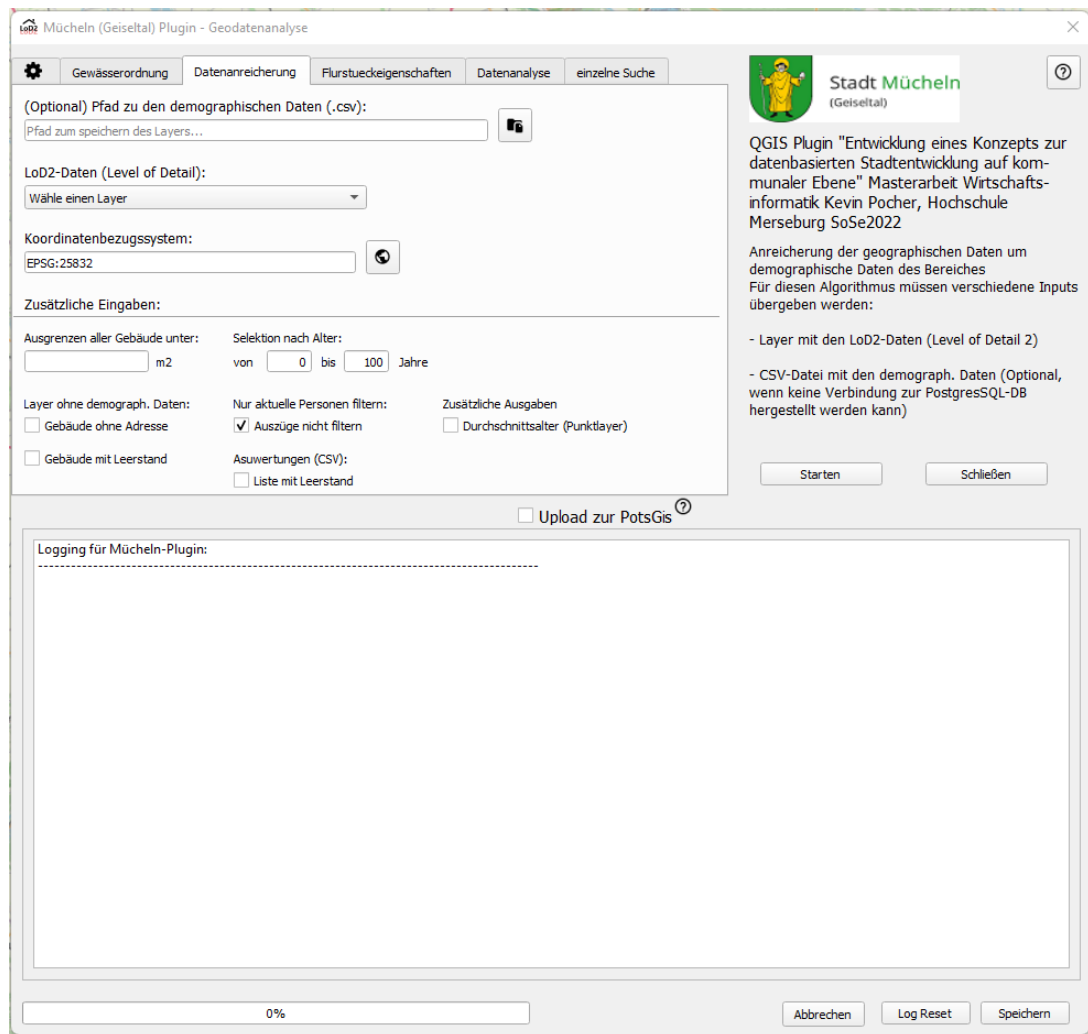
⁷⁴ Vgl. Stuckenholz 2020, S. 94.

⁷⁵ Vgl. Menke 2015, S. ix.

⁷⁶ Vgl. Menke 2015, S. 38.

⁷⁷ Vgl. Menke 2015, S. 324.

Nun können eigene Logiken in dieser Klasse implementiert und weitere Modifikationen nach den eigenen Anforderungen getätigt werden. Der große Vorteil von QGIS ist somit zum einen die bereits vorhandene Verarbeitungsoberfläche für Geodaten und zum anderen die Vielzahl an Möglichkeiten durch Open Source Software. Somit können problemlos Änderungen am Quellcode vorgenommen und dieser dann ohne Hindernisse weiterverbreitet werden.⁷⁸ Durch die Verwendung von QGIS können also kostenpflichtige Produkte vermieden und eine neue Entwicklung einer eigenständigen Softwarelösung verhindert werden.

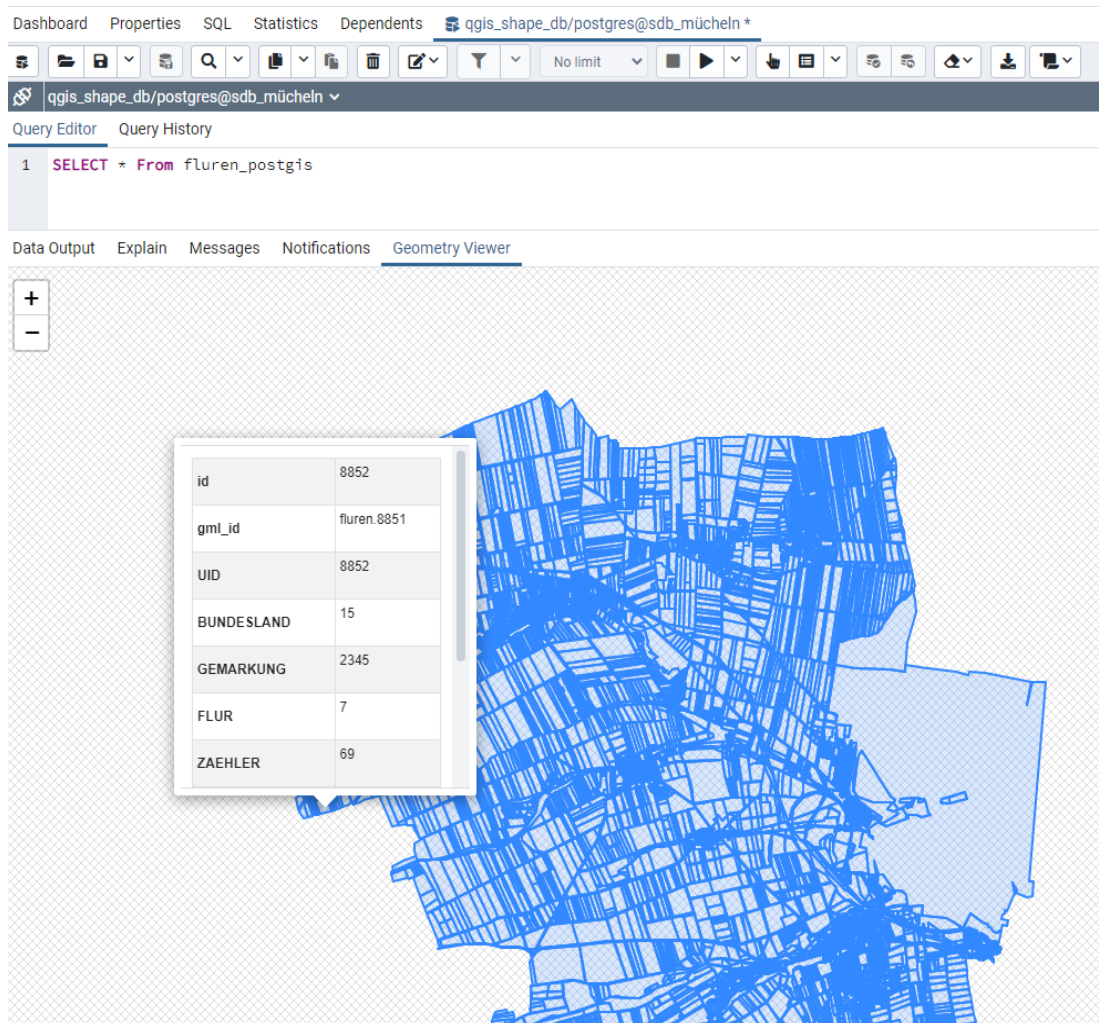


**Abbildung 12: Plugin für die Analysen im QGIS
Datenquelle: eigene Darstellung**

Für die Speicherung der georeferenzierten Daten wurden zwei mögliche Ansätze in Betracht gezogen. Zum einen können die Daten lokal gespeichert werden. Diese Variante kann jedoch schnell zu Problemen mit der Performance und zu fehlenden Speicherkapazitäten führen. Zum anderen wurde die Speicherung der Daten in einer Datenbank betrachtet. Für das

⁷⁸ Vgl. Bundesministerium für Wirtschaft und Technologie 2001, S. 8.

Speichern von Geodaten hat sich hier das Open Source Datenbanksystem PostgreSQL durchgesetzt, welches neben gewöhnlichen SQL-Tools auch eine Erweiterung für Geodaten anbietet, das PostGIS. Durch PostGIS werden neue Datentypen und Funktionen in das PostgreSQL eingebunden, welche dann die Verarbeitung von räumlichen Daten ermöglichen.⁷⁹ Folgende Abbildung zeigt die Flurstücke der Gemarkung Mücheln.



**Abbildung 13: Darstellung von Flurstücken der Stadt Mücheln im pgAdmin
Datenquelle: eigene Darstellung**

Die hier gespeicherten georeferenzierten Daten können nun von verschiedenen GIS als Datenquelle genutzt werden. Das GIS kann dann durch die Vielzahl an Funktionen zur Erfassung, Bearbeitung, Organisation und Präsentation die Daten weiterverarbeiten und Änderungen auf der Datenbank abspeichern.⁸⁰ Durch die Verbindung einer räumlichen Datenbank und den vielseitigen Funktionen eines Geoinformationssystems entsteht eine Vielzahl an

⁷⁹ Vgl. Stuckenholz 2020, S. 87.

⁸⁰ Vgl. Stuckenholz 2020, S. 93.

Auswertungs- und Visualisierungsmöglichkeiten.⁸¹ Zusätzlich wird durch die Einbindung einer PostgreSQL-DB im QGIS das Hochladen und Herunterladen von Daten stark vereinfacht, da diese nur in das jeweilige Projekt gezogen werden müssen. Die folgende Abbildung zeigt die Abfrage eines Datensatzes aus der PostGIS aus dem QGIS heraus.

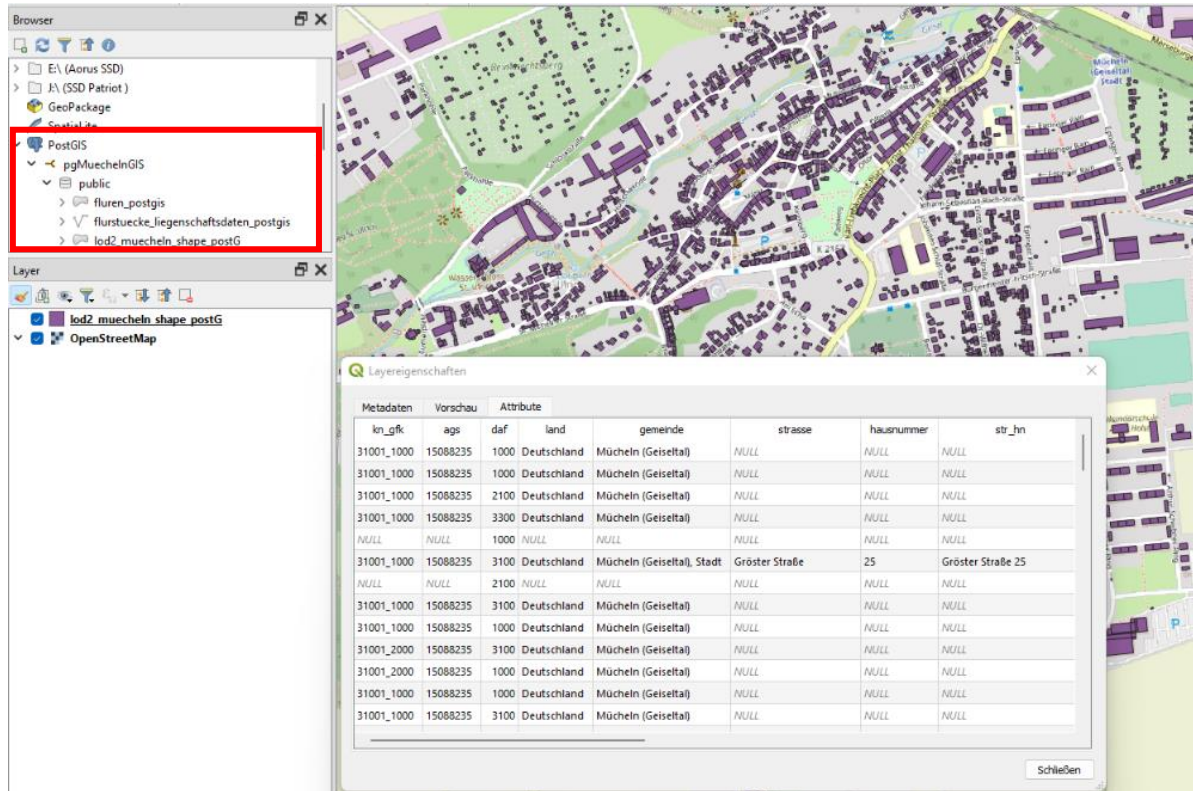


Abbildung 14: Darstellung des LoD2-Datensatzes aus der PostGIS innerhalb von QGIS
Datenquelle: eigene Darstellung

Doch neben der Verwendung der PostgreSQL als Datenbank für die räumlichen Daten wird diese auch als Datenbank für die Daten des Einwohnermeldeamtes verwendet. Somit stehen zwei Datenbanken zur Verfügung, je eine Datenbank für die Verwaltungsdaten und eine für die räumlichen Daten. Der Vorteil bei der Verwendung ist, dass keine lokalen Dateien mehr importiert werden müssen und somit zum einen weniger Aufwand beim Sammeln der Daten entsteht und zum anderen die Daten theoretisch von mehreren Systemen abrufbar sind.

5.1.2 Daten Mapping

Bevor nun jedoch die eigentliche Umsetzung des Planungsalgorithmus beginnen konnte, mussten erst die benötigten Daten gesammelt und auf ein Datenmodell abgebildet werden. Wie bereits zuvor erwähnt, wurde durch ein Teilprojekt der Stadt Mücheln bereits versucht,

⁸¹ Vgl. Stuckenholz 2020, S. 95.

ein Datenmodell zu erzeugen. Jedoch konnte während des Projekts nur durch einen hohen manuellen Aufwand ein solches Modell erzeugt werden, wobei die Verwaltungsdaten auf das richtige Gebäude überführt werden. Somit gab es noch keine Lösung, die Daten automatisiert zu mappen. Hierfür wurde nun ein eigenes Programm geschrieben, welches ein „Data Mapping“ vornimmt. Beim „Data Mapping“ handelt es sich um einen Prozess, der verschiedene Daten von unterschiedlichen Quellen in eine Zielstruktur überführt. Dieser Prozess ist vor allem in Hinblick auf ein einheitliches Format der Daten relevant.⁸²

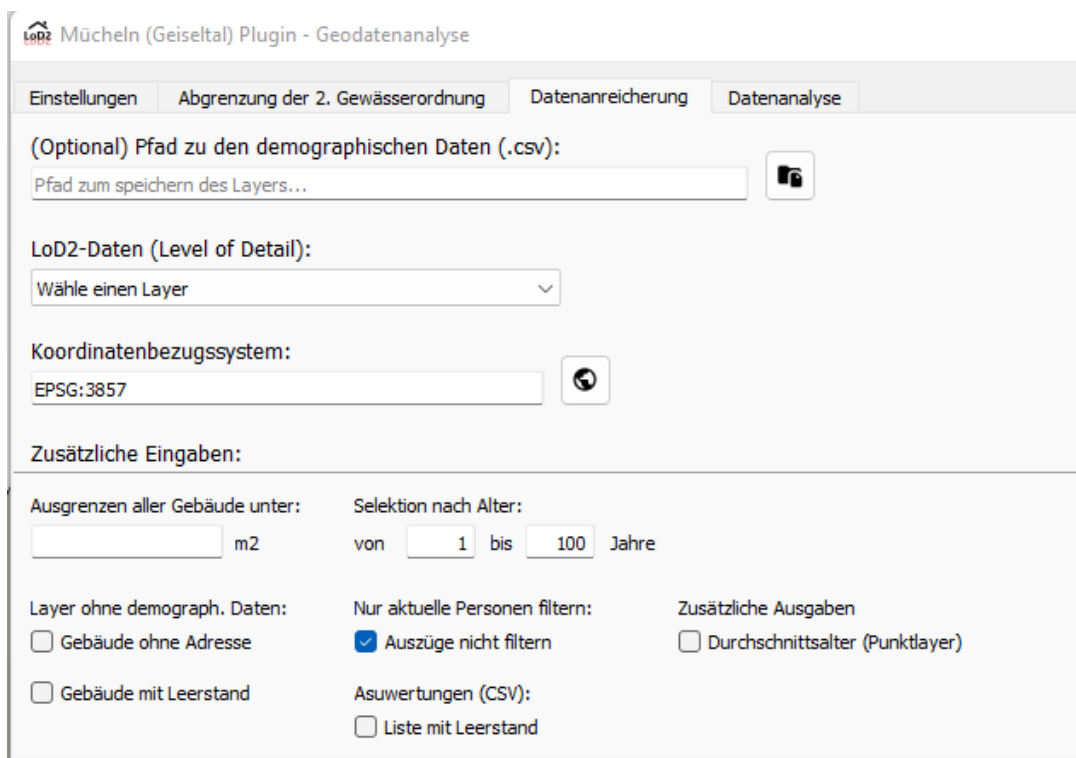
Bei der Umsetzung der Logik zum Vereinen der Daten trat die erste Hürde bei der Definition einer Identifikationsnummer auf. Die ID zum eindeutigen Verbinden der LoD2-Daten und der Verwaltungsdaten stellt somit die Adresse mit Straßennamen, Hausnummer und Hausnummernzusatz dar. Jedoch gibt es verschiedene Datensätze in den LoD2-Daten, welche keine Adresse enthalten. Nach einer genaueren Analyse des Datensatzes fiel auf, dass bei einer Gebäudegeometrie, welche aus mehreren Geobjekten besteht, nicht alle Objekte eine Adresse hinterlegt haben. Dennoch konnte erkannt werden, dass ein Elementtyp fast immer eine Adresse hinterlegt hat, dieser ist der Type „Ground“, also die Grundfläche des Gebäudes. Hierbei ist jedoch anzumerken, dass kleine Geometrien, wie beispielsweise ein Schuppen oder eine Garage keine Adressen zugewiesen bekommen, dafür aber mit der Spalte „BldgParent“ mit dem zugehörigen Gebäude verbunden werden. Somit wurde die Logik so aufgebaut, dass die Verwaltungsdaten nur an das Geobjekt mit dem Elementtypen „Ground“ hinterlegt werden. Hier könnte man nun noch die Daten über die Referenz zum Gebäude mit den anderen zugehörigen Geometrien verbinden, dies wurde jedoch in Hinblick auf die Übersichtlichkeit nicht umgesetzt. Folgend wird noch einmal der Datensatz zu einem Gebäude mit den verschiedenen Geobjekten aufgezeigt.

GEBID	HEG	KISOI	KN_GFK	AGS	DAF	LAND	MEIN	STRASSE	SNUM	TR_HP	NAME	LHOEI	H_BOI	SH_FH	H_TR	CHNI	HORI	HFLA	CHT	ELEMCLASS	BldgRoot	BldgPart	BldgParent
DEST_DESTLIKA...	4,25...	DES...	NULL	NULL	3200	NULL	NULL	NULL	NULL	NULL	NULL	4,25...	130...	134...	133...	66,4...	147...	43,4...	Hip...	Wall	DESTLIKA0001f...	GUID_14079404...	DESTLIKA0001f...
DEST_DESTLIKA...	8,98...	DES...	31001...	15088235	3100	Deu...	Mü...	Wenden	17	We...	NULL	8,98...	133...	142...	139...	48,0...	-11...	91,1...	Gab...	Wall	DESTLIKA0001f...	DESTLIKA0001f...	NULL
DEST_DESTLIKA...	8,98...	DES...	31001...	15088235	3100	Deu...	Mü...	Wenden	17	We...	NULL	8,98...	133...	142...	139...	48,0...	-11...	91,1...	Gab...	Wall	DESTLIKA0001f...	DESTLIKA0001f...	NULL
DEST_DESTLIKA...	8,98...	DES...	31001...	15088235	3100	Deu...	Mü...	Wenden	17	We...	NULL	8,98...	133...	142...	139...	48,0...	-11...	91,1...	Gab...	Wall	DESTLIKA0001f...	DESTLIKA0001f...	NULL
DEST_DESTLIKA...	8,98...	DES...	31001...	15088235	3100	Deu...	Mü...	Wenden	17	We...	NULL	8,98...	133...	142...	139...	48,0...	-11...	91,1...	Gab...	Wall	DESTLIKA0001f...	DESTLIKA0001f...	NULL
DEST_DESTLIKA...	8,98...	DES...	31001...	15088235	3100	Deu...	Mü...	Wenden	17	We...	NULL	8,98...	133...	142...	139...	48,0...	-11...	91,1...	Gab...	Wall	DESTLIKA0001f...	DESTLIKA0001f...	NULL
DEST_DESTLIKA...	8,98...	DES...	31001...	15088235	3100	Deu...	Mü...	Wenden	17	We...	NULL	8,98...	133...	142...	139...	48,0...	-11...	91,1...	Gab...	Wall	DESTLIKA0001f...	DESTLIKA0001f...	NULL
DEST_DESTLIKA...	5,50...	DES...	31001...	15088235	3200	Deu...	Mü...	Franken...	6	Fra...	NULL	5,50...	130...	135...	133...	66,1...	-32...	156...	Hip...	Roof	DESTLIKA0001f...	DESTLIKA0001f...	NULL

Abbildung 15: Beispielhafte Gebäudegeometrie aus dem LoD2-Datensatz
Datenquelle: eigene Darstellung

⁸² Vgl. Astera 2019, S. 2.

Die Logik ermittelt nun über den Elementtyp ‚Ground‘ und die Adresse als ID die zugehörigen Verwaltungsdaten. Dabei werden zusätzlich noch weitere Daten anhand der Verwaltungsdaten bestimmt, wie aktuelles Alter der Person, die Wohnjahre im aktuellen Gebäude und ein Kennzeichen zum aktuellen Wohnstatus. Das Alter lässt sich dabei aus dem Geburtsdatum berechnen und die Wohnjahre und der Status anhand des Einzugs- und Auszugsdatums. Für weiterführende Analysen wurde die Logik so aufgebaut, dass das Programm verschiedene Eingabekriterien liefert, welche eine genauere Bestimmung ermöglichen. Folgende Abbildung zeigt die Selektionsmaske mit den unterschiedlichen Eingabemöglichkeiten für das Programm. So kann neben der Fläche des Gebäudes auch festgelegt werden, ob nur Daten zu aktuell im Gebäude ansässigen Personen bezogen werden oder das Alter auf einen bestimmten Bereich begrenzt wird. Des Weiteren kann eine zusätzliche Ausgabe mit dem Durchschnittsalter erzeugt werden, wenn dieser in der Selektionsmaske ausgewählt wird.



**Abbildung 16: Selektionsmaske des Programmes zum Daten Mapping
Datenquelle: eigene Darstellung**

Als Ergebnis des Mappings der Daten liefert der Prozess mit Bezug auf die Eingabekriterien eine Ausgabe im QGIS, welche zu jedem gefundenen Datensatz aus den Verwaltungsdaten eine Georeferenzierung zugeordnet bekommt. Dabei werden alle Daten mit derselben Adresse eindeutig zugeordnet, jedoch gab es auch hier die Herausforderung Adressen mit unterschiedlichen Schreibweisen oder Strukturen zu verbinden. Ein Beispiel für eine solche

Adresse ist die „Walter-Rathenau-Straße“ in dem Mühelner Ortsteil Neu Biendorf, welcher im LoD2-Datensatz eine andere Schreibweise als in den Verwaltungsdaten aufweist. So ist der Straßename dort als „Walther-Rathenau-Straße“ hinterlegt und wird mit „th“ und nicht nur mit „t“ geschrieben. Solche Datensätze lassen sich nicht eindeutig zusammenführen und das Programm konnte somit keine Verbindung aufbauen. Da es sich hier jedoch um einen erheblichen Teil der Daten gehandelt hat, wurde eine Technik der Linguistik verwendet, die N-Gramm-Analyse. Hierbei handelt es sich um ein Vorgehen, bei dem Texte in einzelne Segmente zerlegt und in eine Trefferliste überführt werden. Diese Liste an Segmenten wird dann mit einer anderen Liste von Segmenten aus einem Text verglichen und es entsteht eine Trefferliste aus der Schnittmenge der beiden Texte.⁸³ Dabei gilt folgende Formel für die Trefferliste (T) aus zwei Texten:

$$T(st) = T(text1) \cap T(text2)^{84}$$

Ein N-Gramm kann dabei einen bestimmten Index zugewiesen bekommen, der die Länge des Segmentes bestimmt. Für das Programm wurde die Länge von drei Zeichen gewählt, welche dann einem 3-Gramm oder auch Trigramm entspricht.⁸⁵ Das 3-Gramm wurde ausgewählt, da es für die Ermittlung der Adresse eine ausreichende Übereinstimmungsmenge ergibt und somit die Performance durch große Trefferlisten nicht beeinträchtigt wird. Mit den Trefferlisten der Texte und der Trefferliste aus der Schnittmenge der Texte lässt sich dann der Dice-Koeffizient berechnen. Dieser Koeffizient ermöglicht es, eine Segmentierungsanalyse zu bewerten, indem die verschiedenen Segmente über folgende Formel verglichen werden...

$$d(text1, text2) = \frac{2 * |T(text1) \cap T(text2)|}{|T(text1)| + |T(text2)|}$$

⁸³ Vgl. Selbach 2011, S. 38.

⁸⁴ Vgl. Selbach 2011, S. 38.

⁸⁵ Vgl. Selbach 2011, S. 40f.

Der Dice-Koeffizient gibt dann das Maß der Überlappung der Texte an und ermittelt so den prozentualen Anteil an N-Grammen der Texte.⁸⁶ Folgend finden Sie noch die praktische Anwendung des N-Gramm's in Verbindung mit dem Dice Koeffizienten am Beispiel der „Walter-Rathenau-Straße“.

```
streetName1: 'Walter-Rathenau-Straße'  
streetName2: 'Walther-Rathenau-Straße'  
> trigramm_Street1.trigrammHashMap: {'##W': '##W', '#WA': '#WA', 'WAL': 'WAL', 'ALT': 'ALT', 'LTE': 'LTE', 'TER': 'T  
> trigramm_Street2.trigrammHashMap: {'##W': '##W', '#WA': '#WA', 'WAL': 'WAL', 'ALT': 'ALT', 'LTH': 'LTH', 'THE': 'T  
> result: ['##W', '#WA', 'WAL', 'ALT', 'ER-', 'R-R', '-RA', 'RAT', 'ATH', 'THE', 'HEN', 'ENA', 'NAU', 'AU-', ...]  
diceCoefficient: 0.92
```

*Abbildung 17: N-Gramm und Dice Koeffizient am Beispiel Straßename
Datenquelle: eigene Darstellung*

Wie in der obigen Abbildung zu erkennen, konnte eine Übereinstimmung von 92% der N-Gramme ermittelt werden. Hier stellt sich nun die Frage, ab wann ein Straßename mit einem anderen Namen übereinstimmt und wann dies nicht der Fall ist. In Absprache mit der Stadt Mücheln wurde hier vorerst ein vereinheitlichter Wert von 80% als Maßstab für eine Übereinstimmung festgelegt. Dieser Maßstab müsste jedoch hinsichtlich der Anzahl der Fragmente zur Laufzeit bestimmt werden, da es sonst bei einer kleinen Anzahl an Fragmenten zu Fehlleistungen kommen kann.

Um die Masse an Datensätzen, welche nicht zugeordnet werden konnten, so speichern zu können, dass eine performante Suche gewährleistet werden kann, wurden die Daten in einer Hash Tabelle gespeichert. In einer Hash Tabelle oder auch Streuwerttabelle werden Datensätze mit einem Schlüssel und einem Wert abgespeichert.⁸⁷ Durch ein Hashverfahren wird bei diesem Vorgehen versucht eine Berechnung des Schlüssels mit dem zugehörigen Datensatz durchzuführen. Dabei liefert der Schlüssel die genaue Position des Datensatzes, ohne dass ein Schlüsselvergleich, wie bei anderen Suchfunktionen, durchgeführt werden muss. Somit ist auch die benötigte Zeit für die Suche eines Datensatzes meist nicht abhängig von der Länge der Liste.⁸⁸ Der Schlüssel kann dabei eine Zeichenkette oder auch eine Zahl sein, welcher durch eine Hashfunktion in einen Hashwert umgewandelt wird. Bei einer Zeichenkette wird meist die Summe der einzelnen Zeichen der Kette über deren Unicode-Wert bestimmt und dieser dann mit der Hashfunktion umgewandelt.⁸⁹ Ein Hashverfahren setzt sich grundlegend aus drei auszuführenden Funktionen zusammen, dem Suchen, Einfügen und

⁸⁶ Vgl. Shamir, et al. 2018, S. 1.

⁸⁷ Vgl. Solymosi; Grude 2017, S. 88.

⁸⁸ Vgl. Ottmann; Widmayer 2012, S. 189f.

⁸⁹ Vgl. Solymosi; Grude 2017, S. 89.

Löschen von Schlüsseln und deren Datensatz. Damit diese Funktionen mit möglichst hoher Performance bereitgestellt werden können, ist es wichtig, dass die Hashfunktion möglichst leicht und schnell zu berechnen ist. Für die Umsetzung des Daten Mappings wurde eine Hashfunktion mit der Divisions-Rest-Methode gebildet, bei der aus der Division des Schlüssels k und der Größe der Tabelle m der Rest als Schlüssel verwendet wird. Somit ergibt sich die folgende Formel für die Hashfunktion⁹⁰ :

$$h(k) = k \bmod m$$

Folgend ein Beispiel für eine Hashtabelle, welche als Schlüssel Straßennamen und als Datensatz Trigramme verwendet. In dieser kann erkannt werden, dass für die Straßennamen ein Schlüsselwert berechnet wurde und dieser den Speicherplatz des Datensatzes innerhalb der Tabelle abbildet.

```

key: 'Hauptstr'
> value: <__main__.Trigramm object at 0x0000023CDEC29120>
key_hash: 2
> key_value: ['Hauptstr', <__main__.Trigramm o...CDEC29120>]
self.map: [[...], None, [...], None, None, None, None]
> special variables
> function variables
> 0: [['Pfeiffers Garten', <__main__.Trigramm o...CDEC29D50>]]
1: None
> 2: [['Hauptstr', <__main__.Trigramm o...CDEC29120>]]
3: None

key: 'Pfeiffers Garten'
> value: <__main__.Trigramm object at 0x0000023CDEC29D50>
key_hash: 0
> key_value: ['Pfeiffers Garten', <__main__.Trigramm o...CDEC29D50>]
self.map: [[...], None, None, None, None, None, None]
> special variables
> function variables
> 0: [['Pfeiffers Garten', <__main__.Trigramm o...CDEC29D50>]]
1: None
2: None
3: None
    
```

Abbildung 18: Beispiel für eine Hashtabelle
Datenquelle: eigene Darstellung

Jedoch kann es bei jeder Hashtabelle vorkommen, dass trotz gut gewählter Hashfunktionen mehrere Einträge denselben Schlüssel haben und somit eine Adresskollision auftritt. In dem Fall ist der Platz $h(k)$ bereits besetzt und der Wert muss an einem anderen Platz gespeichert werden.⁹¹ Um solchen Kollisionen entgegenzuwirken, wurde bei der Umsetzung das Verfahren der direkten Verkettung (engl. chaining) angewandt. Hierbei ist jeder Datensatz der Hashtabelle ein Zeiger auf eine Liste, welche die Daten speichert. Jedoch muss bei diesem Verfahren darauf geachtet werden, dass die Performance bei der Verkettung beeinträchtigt werden kann. Dies geschieht dann, wenn mehrere Datensätze in eine Liste der Hashtabelle angehängt werden. Da diese Liste nur mit einem Schlüsselvergleich durchsucht wird kann bei einer hohen Anzahl an Einträgen die Suchzeit steigen.⁹² Für die Umsetzung wurden

⁹⁰ Vgl. Ottmann; Widmayer 2012, S. 191.

⁹¹ Vgl. Ottmann; Widmayer 2012, S. 196.

⁹² Vgl. Ottmann; Widmayer 2012, S. 198ff.

deswegen verschiedene Szenarien geprüft und es konnte festgestellt werden, dass durch eine möglichst große Wahl der Tabellengröße die Kollisionen minimiert werden können. Da es sich bei den Daten nur um die Straßennamen der jeweiligen Stadt handelt, ist die Anzahl der zu speichernden Werte nicht kritisch. Folgend findet sich noch ein Beispiel für eine Kollision von unterschiedlichen Straßennamen, welche denselben Schlüssel zugewiesen bekommen haben.

```
hashMap.map: [[...], [...], [...], None, [...], None, [...], None, [...], None, None, [...], [...]]
> special variables
> function variables
> 000: [['öltor', <__main__.Trigramm o...28FE84A90>]]
> 001: [['Schulstr', <__main__.Trigramm o...28FE85330>], ['Am Echo', <__main__.Trigramm o...28FE86DA0>]]
002: None
> 003: [['Auf dem Gerade', <__main__.Trigramm o...28FE87190>]]
004: None
> 005: [['Veg Stöbnitz', <__main__.Trigramm o...28FE62D40>]]
006: None
> 007: [['Untere Freyburger Str', <__main__.Trigramm o...28FE872B0>]]
008: None
009: None
> 010: [['Springberg', <__main__.Trigramm o...28FE879D0>], ['Leibaer Str', <__main__.Trigramm o...28FE843A0>]]
```

Abbildung 19: Verkettung bei einer Kollision in einer Hashtabelle
Datenquelle: eigene Darstellung

Die Hashtabelle bietet sich in diesem Fall an, da bei der Prüfung der Straßennamen die Liste an nicht zuordenbaren Daten mehrfach abgefragt werden muss und somit die Suchzeit verringert werden kann. Bei einer sortierten Liste müsste hier jeweils die komplette Liste durchlaufen werden, bis die Schlüssel übereinstimmen und somit wäre die Performance davon beeinträchtigt.

Nachdem das Mapping der Daten für den Anwendungsfall der Stadt Mücheln durchgeführt wurde und alle Einwohnerdaten versucht wurden zuzuordnen, konnte für die Stadt Mücheln eine Übereinstimmung von 100% der Daten erreicht werden. Dies lässt sich Nachweisen, indem zuerst über eine Datenbankabfrage geprüft wird, wie viele Bürger derzeit in Mücheln leben und dann geschaut wird, wie viele Gebäude verbunden werden konnten. Dabei konnten insgesamt 8970 Menschen in Mücheln identifiziert werden und es wurden dazu ebenso 8970 Gebäude angelegt. Der Hintergrund war, dass jede Person eine eigene Georeferenzierung auf das jeweilige Gebäude erhält.

5.2 Ermittlung von Gebäuden für Kindergärten

Nachdem nun die Grundlage für die Umsetzung und Durchführung des Planungsalgorithmus aufgestellt wurde, kann dessen Aufbau und Logik analysiert werden. Doch zuvor soll noch einmal kurz geklärt werden, welches Ziel der Algorithmus verfolgt und welche Ergebnisse zu erwarten sind. Der Algorithmus soll durch die Verbindung von georeferenzierten Daten und demografischen Daten den Leerstand der Gebäude innerhalb der Stadt Mücheln identifizieren und anhand einer Analyse der Altersstruktur, Gebäude vorschlagen, die für einen Faktor relevant wären. Als Fallbeispiel wurde hier der Faktor „Kindergarten“ gewählt und es sollen folglich Gebäude vorgeschlagen werden, die für einen Kindergarten geeignet wären. Dabei soll auch die Wahl des Standortes anhand der Altersstruktur getroffen werden. Dafür soll der Algorithmus die Berechnung für jeden Ortsteil der Stadt durchführen, so dass später erkannt werden kann, wo der meiste Bedarf besteht. Als Ergebnis soll eine Ausgabe in Form einer PDF-Datei geliefert werden, welche die wichtigsten Informationen enthält. Zusätzlich soll im QGIS eine visuelle Ausgabe der ermittelten Gebäude aufgebaut werden, um die Ergebnisse auf einen Blick sichtbar zu gestalten. Wie zuvor schon angesprochen werden dafür die LoD2-Daten, die demografischen Daten und die Verbindung dieser beiden Datensätze für eine Stadt benötigt. Zusätzlich kommt auch noch die OSM-Schnittstelle und andere Open Data zum Einsatz, um bestimmte Geodaten zu ermitteln. Um nun den Prozess des Algorithmus aufzeigen zu können, folgt eine Prozessmodellierung, welche die einzelnen Schritte darstellen soll.

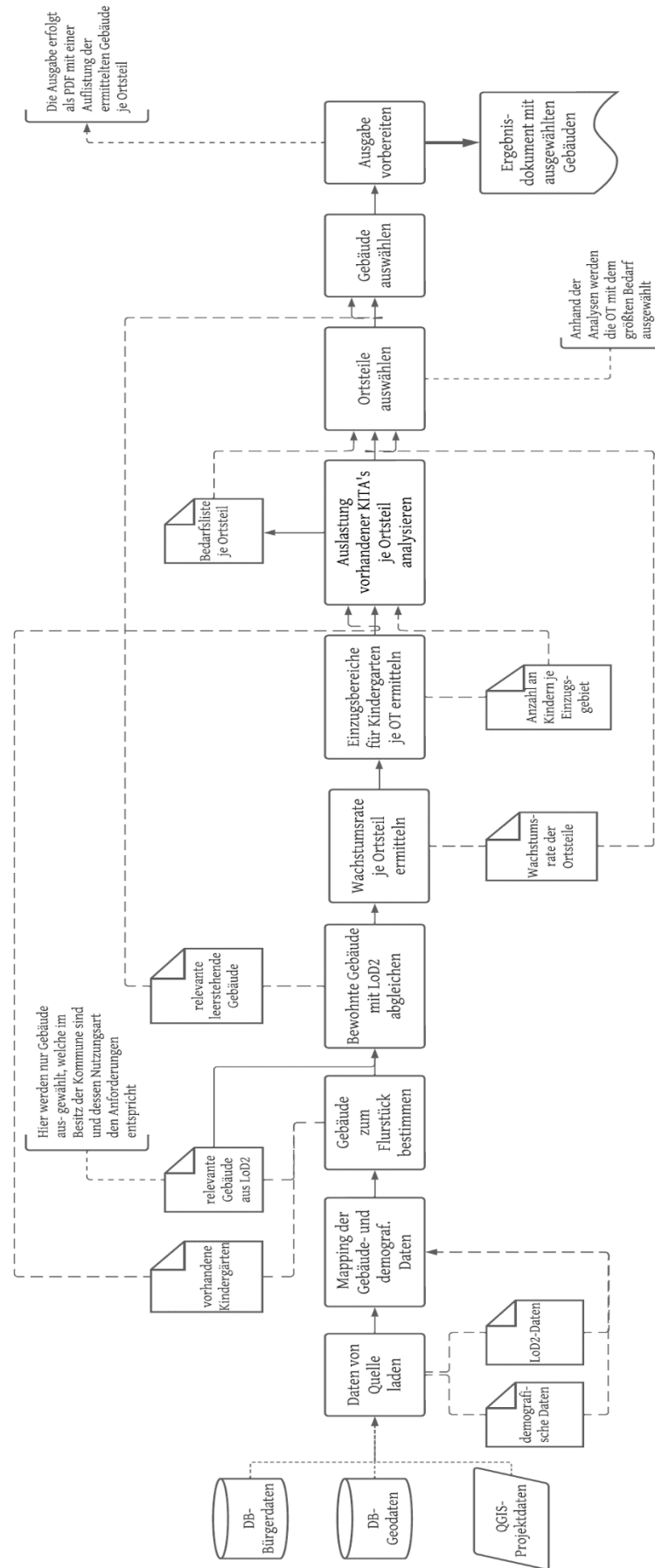


Abbildung 20: Prozessablauf des Algorithmus zur Bestimmung der Gebäude für Kindergärten
 Datenquelle: eigene Darstellung

Wie aus der obigen Darstellung zu erkennen ist, startet der Prozess mit einer Auslösung über das Plugin innerhalb von QGIS. Dabei müssen die zuvor genannten Daten mitgegeben und der Prozesswert „Kindergarten“ ausgewählt werden. Der Prozesswert bestimmt dabei die Wahl des Algorithmus, d.h. hier können später noch weitere Algorithmen integriert werden. Nach dem Start beginnt der Prozess mit einer Abfrage der demografischen Daten von der Datenbank und bereitet die übrigen Daten so auf, dass diese innerhalb des Prozesses verwendet werden können. Nach der Datensammlung werden die Analysen gestartet. Bevor jedoch die demografischen Daten analysiert werden können, bedarf es zwei einleitender Prozesse, welche prozessrelevante Daten liefern. Zum einen wird, wie im Algorithmus zum Mappen der Daten, eine Verbindung zwischen der Gebäudegeometrie und den Adressaten aufgebaut. Das genaue Vorgehen wurde bereits im vorherigen Kapitel aufgezeigt. Jedoch werden hier die Daten nicht als Layer im QGIS visualisiert, sondern es werden zwei Hash Tabellen erzeugt, eine für die bewohnten Gebäude und eine für die unbewohnten Gebäude. Diese werden später benötigt, um bestimmen zu können, welche Gebäude zur Verfügung stehen. Zum anderen wird ein Prozess durchlaufen, welcher die Eigenschaften der Flurstücke der Stadt ermittelt und anschließend die Flurstücke und dessen Gebäude identifiziert, welche bestimmten Anforderungen aufweisen. Die zu ermittelnden Eigenschaften stellen dabei der Eigentümer und die Nutzungsart des Flurstückes dar. Dabei wurde Eigentümer auf „kommunal“ und „privat“ verallgemeinert. Somit können die Gebäude o.a. Flächen, welche der Stadt bereits gehören identifiziert werden. Die Nutzungsart sollte bestimmt werden, damit der Algorithmus später beschreiben kann, ob es sich um ein, dem Anwendungsfall entsprechendes Gebiet handelt. Ziel war es somit, dass automatisiert erkannt wird, ob sich das Gebäude beispielsweise in einem Industriegebiet oder in einem Wohngebiet befindet. Jedoch konnte der Flächennutzungsplan, welcher die Daten zur Nutzungsart enthält, nicht rechtzeitig von der Stadt Müheln ausgeliefert werden. Grund dafür war, dass dieser nur als PDF zur Verfügung steht und die Umwandlung in eine Shapefile derzeit nicht möglich war. Neben den Eigenschaften werden aber auch die Anforderungen an das Gebäude und Flurstück überprüft. Dabei werden die Gebäude, welche nicht den Anforderungen entsprechen nicht weiter betrachtet. Bei den Anforderungen handelt es sich derzeit um eine maximale Höhe des Hauses von 8 m, einer Grundfläche von mindestens 80 m² für das Gebäude und einer Grundfläche des Flurstückes von mindestens 300 m². Diese Werte wurden aus dem Interview mit der Kinderland GmbH abgeleitet. Hier wurde darauf hingewiesen, dass vor allem Flachbauten sehr empfehlenswert sind. Ein weiterer wichtiger Aspekt ist der Brandschutz, welcher bei Kindergärten eine große Rolle spielt. Außerdem wurde eine

ausreichende Fläche für Spielraum und entsprechende Geräte im Außenbereich empfohlen. Nachdem dieser Prozess abgeschlossen ist, liegen die grundlegenden Geodaten in einer aufbereiteten Form bereit.

Nach der Vorbereitung der Daten wird nun anhand der demografischen Daten das Wachstum (in Bezug auf die Bevölkerungsentwicklung) der einzelnen Ortsteile der Stadt bestimmt. Diese Bestimmung soll es später ermöglichen, dass eine Auswahl an leerstehenden Gebäuden je Ortsteil ausgeben werden kann. Gleichzeitig kann dadurch auch für die einzelnen Ortsteile bestimmt werden, ob die Anschaffung eines weiteren Kindergartens für den Ortsteil notwendig ist. Um die einzelnen Wachstumsraten für die Ortsteile bestimmen zu können, muss zuvor die Nettozuwanderung der Ortsteile bestimmt werden. Dafür wurden als erstes alle Bürger der Stadt Mücheln identifiziert, welche zum einem in dem Alter von 15-45 Jahren sind und zum anderen in den letzten 10 Jahren entweder nach Mücheln gezogen oder weggezogen sind. Wie bereits im vorherigen Kapitel angesprochen, wurde diese Altersspanne anhand einer Studie des Statistisches Bundesamtes ermittelt, bei der das durchschnittliche Alter der Eltern bei der Geburt des Kindes analysiert wurde.⁹³ Dann werden die Auszüge und Einzüge in die Stadt Mücheln für die identifizierten Bürger bestimmt. Dabei ergibt sich nun die Nettozuwanderung aus der Zuwanderung pro Jahr minus die Abwanderung pro Jahr. Zur Berechnung des Wachstums pro Ortsteil wurde die Formel zur durchschnittlichen Wachstumsrate verwendet. Bei dieser Formel wird anhand einer Periode, einem Anfangswert und einem Endwert das Wachstum bestimmt. Dabei gilt folgende Formel für die Berechnung:⁹⁴

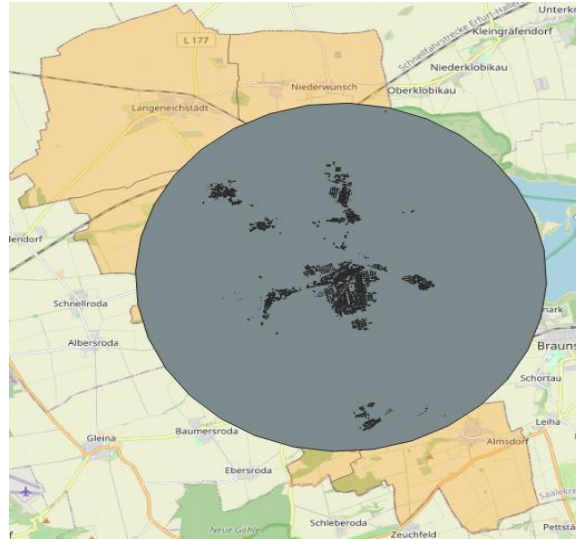
$$\text{durschnittliches Wachstum} = \text{Periode} \sqrt{\frac{\text{Endwert}}{\text{Anfangswert}}} - 1$$

Wie der Abbildung zum Prozess zu entnehmen ist, findet nun eine Berechnung der Auslastung der bestehenden Kindergärten statt. Diese soll aufzeigen, ob der jeweilige Ortsteil einen Bedarf für einen weiteren Kindergarten hat oder ob dieser Ortsteil nicht weiter betrachtet werden soll. Dafür muss als erstes das Einzugsgebiet für die Ortsteile ermittelt werden. Um diese Gebiete bestimmen zu können, wird eine Toleranzgrenze für jeden Ortsteil festgelegt. Diese Toleranz wird standardmäßig auf einen Umkreis von 5 km gesetzt und zeigt das Einzugsgebiet für den jeweiligen Ortsteil an. Wenn die Toleranzgrenze ermittelt wurde, findet eine Verschneidung aller Gebäude mit den demografischen Daten statt. Diese erzeugt einen

⁹³ Vgl. Statistisches Bundesamt 2022.

⁹⁴ Vgl. Prof. Dr. Rolf Hüpen 2022, S. 3-4.

neuen Datensatz, der nur die Gebäude bzw. Bürger enthält, welche innerhalb der ermittelten Grenze liegen. Folgende Abbildung zeigt, wie die Toleranz und die Verschneidung für die Kernstadt Mücheln aussehen.



**Abbildung 21: QGIS Berechnung des Einzugsgebietes und Verschneidung der Gebäude
Datenquelle: eigene Darstellung**

Anhand dieser Geodaten, welche sich nun aus den Gebäude-Geometrien und den demografischen Daten zusammensetzen, kann nun für jeden Ortsteil und dessen Einzugsgebiet eine Anzahl an Kindern für die entsprechende Altersspanne von 0-6 Jahren ermittelt werden. Nun kann der Algorithmus bestimmen, ob der jeweilige Ortsteil einen Bedarf an einem neuen Kindergarten aufweist. Dafür werden die Grundflächen der verschnittenen Gebäude addiert, so dass eine Gesamtfläche als Ist-Wert berechnet werden kann. Dieser Ist-Wert stellt die derzeitige Fläche für die Nutzung von Kindergartenplätzen dar. Nun wird noch ein Soll-Wert berechnet, welcher den Bedarf an Grundfläche für den Ortsteil darstellt. Für die Berechnung des Soll-Wertes werden ebenso die genannten Parameter aus dem Interview mit der Kinderland GmbH verwendet. Hier wurde gesagt, dass für jedes Krippenkind ein Platz von $2,5\text{m}^2$ und für jedes Kindergartenkind 5m^2 eingeplant werden sollte. Um den Rahmen dieser Arbeit nicht zu übersteigen, wurde hier der Mittelwert von $3,75\text{m}^2$ für alle Kinder im Alter von 0-6 Jahren verwendet. Nun ergibt sich der Soll-Wert aus der Multiplikation der Mindestfläche für jedes Kind und der Anzahl Kinder für das Einzugsgebiet des Ortsteils. Wenn nun der Soll-Wert größer ist als der Ist-Wert, dann besteht ein Bedarf an neuen Flächen für Kindergärten innerhalb des Ortsteils. Folgende Tabelle zeigt die Berechnung am Beispiel des Ortsteils Gröst von Mücheln. Dabei besitzt Gröst nur einen Kindergarten und es leben 86 Kinder im Einzugsgebiet.

	Fläche in m ²
Kindergarten 'Pustebume' Gröst	210,04
Soll-Wert (bei 86 Kindern)	322,5
Bedarf	112,46

Tabelle 4: Berechnung des Bedarfs an Kindergartenfläche pro Ortsteil

Im nun letzten Schritt des Algorithmus werden die gesammelten Parameter ausgewertet und bestimmt, welche Ortsteile mehr Flächen für Kindertagesstätten benötigen. Dafür werden zum einen die Wachstumsrate und der aktuelle Bedarf ausgewertet. Nachdem die betroffenen Ortsteile ermittelt wurden, können nun die Gebäude zugeordnet werden. Dabei werden nochmals die identifizierten Gebäude nach der mitgelieferten Gebäudefunktion sortiert und gegebenenfalls ausgeschlossen. Am Ende werden die Gebäude und die zusätzlichen Informationen über den Standort und das Flurstück zum einen als Visualisierung im QGIS dargestellt und zum anderen wird eine PDF erzeugt, welche die Daten nach den Ortsteilen auflistet. Somit können die Sacharbeiter der Stadt genau nachvollziehen, welche Gebäude derzeit zur Verfügung stehen und geeignet sind. In dem Kapitel zur Diskussion der Ergebnisse werden noch Ausschnitte aus dem Ergebnisdokument vorgestellt.

5.3 ETL-Prozess Stadt Mücheln

Damit die benötigten Daten immer bereitgestellt werden können, wurde für die Umsetzung ein ELT-Prozess erarbeitet. Dieser soll es ermöglichen, dass die demografischen Daten und die georeferenzierten Daten immer in der benötigten Qualität und Aktualität bereitstehen. Da diese Daten derzeit bei der Stadt Mücheln noch auf getrennte Weise abgelegt werden, müssen diese vorerst auf eine gemeinsame und abrufbare Datenbank überführt werden. Wie bereits zuvor aufgezeigt, werden die benötigten Daten auf einer PostgreSQL-DB abgespeichert und können von dieser abgerufen werden. Somit müssen die Daten aus einem Quellsystem extrahiert und in ein Zielsystem überführt werden. Dafür könnte ein externes ETL-Tool verwendet werden. Jedoch wurde für diese Arbeit ein eigenes Tool entwickelt, welches die Ansammlung der benötigten Daten erlaubt. Der Grund dafür sind zum einen die begrenzten Möglichkeiten zum Aufbauen einer Verbindung zu den Servern bzw. Datenbanken der Stadt Mücheln. Und zum anderen die Vielfalt an unterschiedlichen Daten, welche für die Zwecke dieser Arbeit benötigt werden. Dabei stellen vor allem die georeferenzierten Daten eine Herausforderung dar, da für diese Art der Daten nur wenige ETL-Tools als Open-Source-Software zur Verfügung stehen. Außerdem werden viele Aspekte des ETL-Prozesses bereits durch QGIS und dessen eingebaute Funktionen abgehandelt, sodass sich eine eigene

Lösung für die Extraktion der Daten anbietet. Durch die Verwendung einer eigenen Lösung kann zusätzlich auch die Komplexität verringert werden, da diese direkt auf den Anwendungsfall zugeschnitten wird. Dennoch können die Funktionen des Tools bei Bedarf einfach durch neue Anwendungsfälle erweitert werden. Das Tool zur Extraktion der Daten bietet bis jetzt die Möglichkeit die Einwohnermeldedaten und die spezifischen Geodaten in ein Zielsystem zu laden und zu speichern. Bei den relevanten Geodaten handelt es sich derzeit noch um die LoD2-Daten, die Daten zu den Flurstücken und die Gemarkung der jeweiligen Stadt. Von dem Zielsystem aus können die Daten dann für die Analysen herangezogen werden und beispielsweise für den Planungsalgorithmus „Kindergärten“ genutzt werden.

Für das Extraktions-Tool wurde ebenfalls ein Programm in Python entwickelt, da Python spezielle Bibliotheken für die Verarbeitung von Geodaten anbietet, wie beispielsweise Geopandas oder auch OWSLIB. Geopandas bietet die Möglichkeit verschiedene Geodatendateiformate in Python einzulesen und zu verarbeiten.⁹⁵ Diese Funktionen werden vor allem für die LoD2-Daten benötigt, welche durch das „LVermGeo“ an die Kommunen in Sachsen-Anhalt ausgeliefert werden. Da diese somit nicht durch einen Web-Service abgefragt oder auf anderem Wege beschafft werden können, bietet Geopandas die Möglichkeit, diese Daten direkt als Datei zu verarbeiten. Zusätzlich können auch noch weitere Geodaten, wie Flurstückdaten als Shape-Datei eingelesen werden. OWSLIB hingegen bietet Funktionen an, mit denen WMS oder auch WFS-Services angefragt werden können. Diese Funktionen sind relevant, um beispielsweise die WFS-Services des „LVermGeo“ zu nutzen und somit die benötigten Geodaten abzufragen. Der Vorteil bei der Verwendung von WFS oder WMS ist, dass die Daten automatisiert abgefragt werden können. Aus Zeitgründen wurde jedoch nur Geopandas verwendet und somit die benötigten Geodaten manuell in das Extraktions-Tool eingeführt. Später sollte jedoch das Tool um WFS-Services erweitert werden.

Folgende Abbildung zeigt den ELT-Prozess für die Sammlung, das Laden und die Transformation der Daten.

⁹⁵ Vgl. Kelsey Jordahl 2016, S. 3.

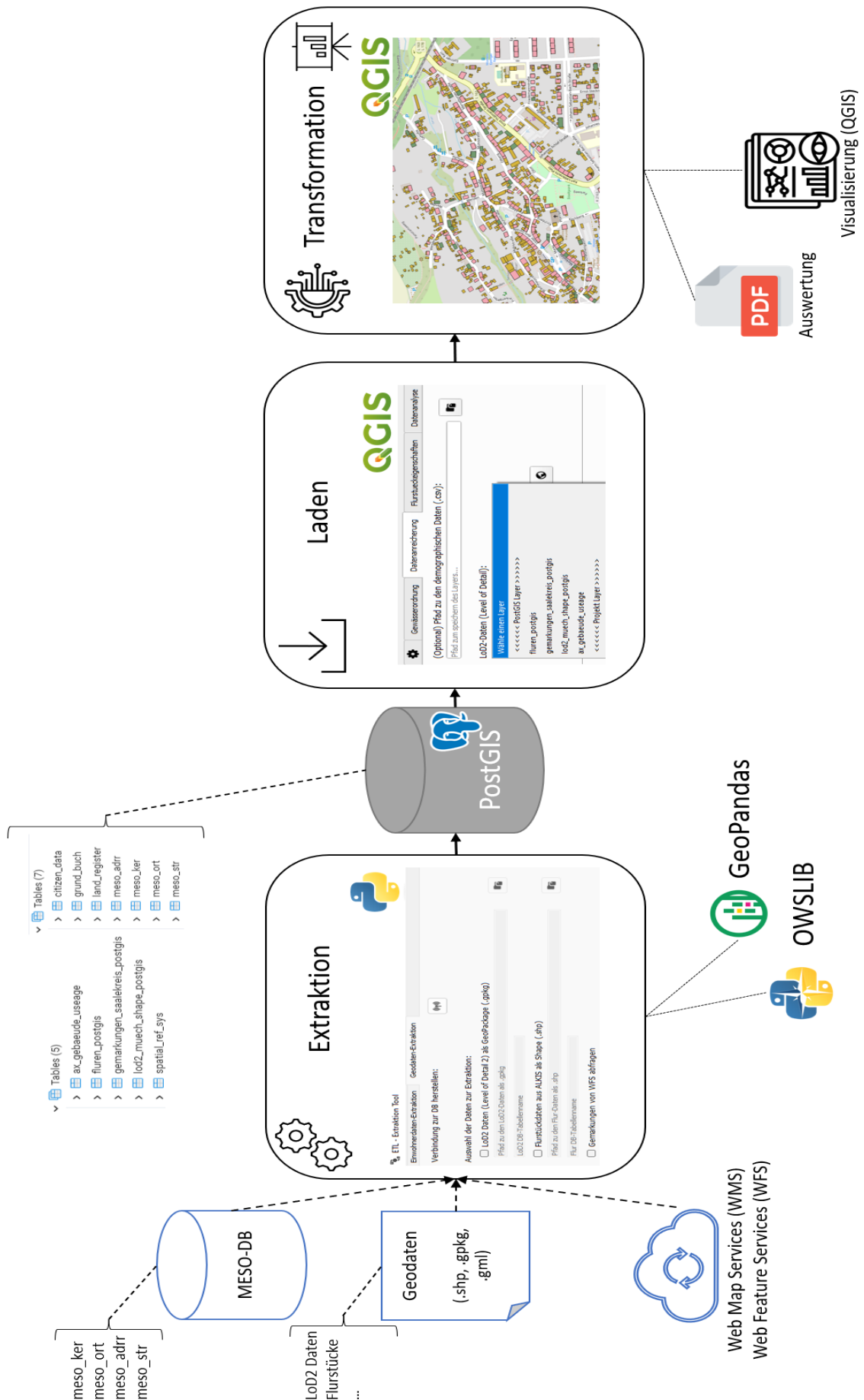


Abbildung 22: ELT-Prozess für die Zusammenführung und Analyse der Geodaten und demografischen Daten
 Datenquelle: eigene Darstellung

Wie bereits aufgezeigt handelt es sich bei diesem Prozess um einen ELT-Prozess und nicht einen ETL-Prozess, es wird also erst eine Transformation bei dem Laden der Daten durchgeführt. Der Grund dafür ist, dass zum einen immer die Rohdaten bei dem Laden bereitstehen und somit theoretisch verschiedene Analysen vollzogen werden können, so dass eine höhere Flexibilität vorliegt. Zum anderen bietet sich ein ELT-Prozess an, da die Extraktion und das Laden von verschiedenen Programmen durchgeführt werden. So werden die Daten mit dem speziell entwickelten Extraktions-Tool in die Zieldatenbank gespeichert und später dann innerhalb von QGIS mit dem Plugin geladen. Durch die Vielzahl an integrierten Funktionen von QGIS stehen somit bei dem Laden der Daten eine große Menge an Möglichkeiten, die Geodaten als auch die demographischen Daten zu transformieren, zur Verfügung.

Wie in der vorherigen Abbildung schon zu erkennen war, werden verschiedene Datensätze in das Zielsystem überführt. Dabei handelt es sich um die Daten des Einwohnermeldeamtes bzw. die demografischen Daten und die Geodaten, welche sich aus Open Data und Daten des Vermessungsamtes zusammensetzen. Da jedoch eine direkte Verbindung zu den Datenbanken der Stadt Müheln nicht ohne weiteres zugelassen wird, wurden die benötigten demografischen Daten exemplarisch in einer Datenbank gespeichert, welche für diese Forschungsarbeit genutzt werden kann. Durch das Extraktions-Tool wird es ermöglicht die demografischen- und die Geodaten in das Zielsystem zu überführen. Dabei stellt das Zielsystem einen PostgreSQL-Server dar, welcher zwei Datenbanken besitzt. Eine Datenbank für die demografischen Daten und eine Datenbank für die Geodaten. Die DB für die Geodaten wurde dabei so erweitert, dass diese auch georeferenzierte Daten verarbeiten kann. Wie bereits aufgezeigt, verwendet das Extraktions-Tool verschiedene Technologien, welche durch Erweiterung zu Python bereitgestellt werden. So können die LoD2-Daten und die Flurstückdaten durch „Geopandas“ im Programm gelesen und in der Datenbank gespeichert werden. Auch besteht die Möglichkeit durch WFS-Dienste Geodaten als Service abzufragen und diese dann in der Datenbank zu speichern. Hier besteht auch die Möglichkeit das Extraktions-Tool noch zu erweitern, falls noch andere „Open Geodata“ für die Analysen benötigt werden. Wenn die Daten dann in den entsprechenden Datenbanken abgelegt wurden, können diese nun über das Plugin im QGIS geladen werden. Die Daten können für die verschiedenen Analysezwecke geladen werden, welche im vorherigen Kapitel aufgezeigt wurden. Bei dem Laden der Daten findet nun auch gleichzeitig die Transformation dieser statt. Bei der Transformation werden die Daten so verändert, manipuliert oder auch angereichert, dass diese für die speziellen Analysezwecke verwendet werden können. Ein Beispiel wäre hier die Transformation bei dem Mapping der Daten, bei dem verschiedene Datentypen angepasst oder

auch bestimmt Format überprüft werden. Nach der Transformation können die Daten für die Analysen herangezogen werden. Diese werden dann zum einen über QGIS visualisiert und zum anderen auch als PDF ausgegeben.

Für den ETL- bzw. ELT-Prozess spielt es auch eine Rolle in welchen Zyklen dieser durchgeführt werden muss, also wann die Daten auf den neusten Stand gebracht werden müssen. Diese Planung ist ein wichtiger Teil des Prozesses, da eine gute Datenbasis über die Qualität der Analysen entscheidet. Nach einer Absprache mit der Stadt Müheln, wäre es ratsam die demografischen Daten entweder monatlich oder alle zwei Monate erneut zu extrahieren. Dieses Intervall wurde empfohlen, da innerhalb von 1-2 Monaten sich die Population innerhalb einer Stadt verändern kann. Bei den Geodaten sollte man zwischen den einzelnen Daten unterscheiden, da diese in unterschiedlichen Intervallen ausgeliefert werden. Die LoD2-Daten werden jährlich auf den neusten Stand gebracht und den Städten ausgeliefert, die Flurstückdaten als auch die Daten zu den Gemarkungen werden meist tagesaktuell ausgeliefert. Somit wäre es also sinnvoll die Geodaten, soweit es möglich ist, mit den demografischen Daten zusammen zu extrahieren, also in einem monatlichen Intervall. Die LoD2-Daten können dann natürlich nur zum jeweiligen Auslieferungstermin extrahiert werden.

5.4 Weitere Funktionen

Neben dem Algorithmus und dem ETL-Prozess wurden auch noch weitere Funktionen im Rahmen dieser Forschungsarbeit entwickelt. So verfügt das Plugin im QGIS auch über die folgenden Möglichkeiten:

- Vektorlayer in die PostGis zu laden (aus dem Algorithmus oder aus dem Projekt heraus),
- Suchfunktionen für Gebäude an bestimmten Adressen und die Suche von Flurstücken,
- Bestimmung zusätzlicher Daten zum Flurstück aus dem Grundbuch
- Bestimmung der Gewässer 2. Ordnung (speziell für Müheln),
- Zuwanderungsstatistik zur Stadt.

Die erste genannte Funktion ermöglicht es, dass Vektorlayer in die angegebene PostGis hochgeladen werden. Dabei können Layer entweder direkt aus dem Algorithmus heraus oder durch eine zusätzliche Eingabe hochgeladen werden. Um den Layer aus dem Algorithmus heraus hochzuladen, muss das dafür vorgesehene Kontrollkästchen aktiviert werden. Um jedoch einen Layer manuell in die PostGis hochzuladen, muss in dem Plugin die integrierte

Funktion „Upload“ ausgeführt werden. Der Vorteil des Hochladens der Daten liegt darin, dass die aus den Analysen entstandenen Vektorlayer in der PostGis abgespeichert werden und somit flexibel abrufbar sind. Somit können auch unterschiedliche Systeme auf dieselben Daten zugreifen und mit diesen arbeiten. Im Anhang 3 ist jeweils eine Abbildung für das Hochladen aus dem Prozess heraus und das Hochladen über die integrierte Funktion dargestellt.

Eine weitere Funktion ist die Suche von spezifischen Gebäuden über die Adresse oder die Suche von Flurstücken über die jeweiligen ID's. Diese Funktion wurde auf Wunsch der Stadt Mücheln integriert, da diese bis jetzt solche Suchfunktionen nur über die ALKIS-Software nutzen konnte, diese jedoch große Performanceprobleme aufweist. Gleichzeitig wurde auch noch ein Programm geschrieben, welches die Eigenschaften eines Flurstückes ermittelt. Bei diesen Eigenschaften handelt es sich um Daten aus dem Grundbuch der Stadt und ermöglicht eine Bestimmung des Besitzers des jeweiligen Flurstücks. Im Anhang 3 können ebenfalls Abbildungen zu diesen Funktionen gefunden werden.

Des Weiteren wurde ein Algorithmus entwickelt, welche für die Stadt Mücheln die Gebiete der Gewässer der 2.Ordnung identifiziert und diese dann auf die zugrundeliegenden Flurstücke aufteilt. Hintergrund für diesen Prozess ist, dass die Kommunen eine Gewässersteuer für alle Gebiete der Gewässer der 1. Ordnung zahlen müssen, die Gewässer 2. Ordnung jedoch von den Besitzern des jeweiligen Flurstückes bezahlt werden muss. Bisher wurde die Identifikation mit viel manuellen Aufwand über Excel-Tabellen betrieben. Nun können die entsprechenden Geodaten über das QGIS-Plugin verarbeitet werden. Dabei wird als Ergebnis des Algorithmus eine Excel-Tabelle erstellt, welche zu jedem Flurstück der 2. Ordnung, den Anteil des Gebietes aufzeigt. Somit kann die Stadt Mücheln nun ohne großen manuellen Aufwand die Gewässersteuer auf die Anlieger aufteilen. Im Anhang 3 befinden sich Abbildungen zum Ergebnis des Algorithmus zur Gewässerordnung.

Letztlich wurde noch eine Funktion integriert, die es erlaubt anhand der demografischen Daten die Zuwanderung der Bevölkerung innerhalb von Mücheln aufzuzeigen. Dabei wird über OSM und den Geburtsort der Person eine visuelle Ausgabe erzeugt. Diese kann dabei helfen, die kulturellen Unterschiede der Bevölkerung auf einen Blick zu analysieren und als Stadt entsprechende Maßnahmen einzuleiten. Ein Beispiel hierfür wäre die Zuwanderung von ausländischen Bürgern. Durch die Visualisierung kann dies schnell erkannt werden und beispielsweise mit entsprechenden Sprachkursen und dem Organisieren von Dolmetscher reagiert werden. Auch hierzu befindet sich eine Abbildung im Anhang, welche das Ergebnis als Visualisierung im QGIS aufzeigt.

6 Diskussion der Ergebnisse

Innerhalb dieses Kapitels werden die Ergebnisse der Umsetzung zusammengefasst und evaluiert. Dabei wird auf die Grenzen der Umsetzung eingegangen und werden weitere Probleme diskutiert. Im zweiten Teil des Kapitels werden noch weiterführende Möglichkeiten bei der Verwendung von Geodaten im Rahmen der Stadtentwicklung aufgezeigt.

6.1 Auswertung der Ergebnisse

Bei der Ausführung des Algorithmus zur Bestimmung des Leerstandes zur Ermittlung von Gebäuden für die Nutzung von Kindertagesstätten, konnten verschiedene Parameter bestimmt werden. Durch die Aufbereitung und Verbindung der Gebäudedaten des Lod2-Datensatzes und der demografischen Daten konnten im ersten Durchlauf insgesamt 481 Gebäude bestimmt werden, welche derzeit nicht bewohnt sind und auf die, die Gebädefunktion „Wohngebäude“ zutrifft. Nachdem jedoch der zweite Durchlauf der Verbindung mit der Verwendung von Trigrammen durchgeführt wurde, konnten nur noch 306 leerstehende Gebäude identifiziert werden. Diese sinkende Anzahl an Gebäuden hängt damit zusammen, dass durch den Einsatz der Trigramme mehr bewohnte Gebäude ermittelt werden können. Weiterhin konnten 14709 Gebäude identifiziert werden, welche keine Adresse hinterlegt haben. Bei diesen Gebäuden handelt es sich entweder um unbewohnte Objekte, wie ein Carport oder ein Schuppen oder um beispielsweise gewerblich genutzte Gebäude. Da diese nicht als Wohngebäude gemeldet sind, wurde diesen keine Adresse im LoD2-Datensatz zugewiesen. Da es auch keine Informationen vom Einwohnermeldeamt über gewerblich genutzte Gebäude gab, konnten diese nicht bestimmt werden. Hier zeigt sich das erste Hindernis bei der Auswertung der Daten, denn solche Gebäude können nicht ohne weiteres als „bewohnt“ bzw. als nicht leerstehend ermittelt werden. Für dieses Problem gab es zwei Lösungswege, welche derzeit angewendet werden können. Der Algorithmus könnte diese Gebäude aus der Ergebnisliste entfernen oder er belässt diese Gebäude weiterhin als unbewohnte Objekte. Im Sinne einer qualitativ höheren Auswertung wurde der zweite Weg gewählt und die Gebäude weiterhin als unbewohnt betrachtet. Das Resultat dieser Entscheidung ist, dass die Ergebnisliste auch unbewohnten Gebäude zeigt, welche derzeit für gewerbliche Zwecke genutzt werden. Somit müssten die Gebäude mit der entsprechenden Funktion nochmals geprüft werden oder der Algorithmus später so erweitert werden, dass dieser auch gewerblich genutzte Gebäude erkennen kann. Dafür müsste eine Liste an Daten bereitgestellt werden, welche neben dem Namen des Gewerbes auch die Anschrift

übermittelt. Durch die Verwendung aller Gebäude des LoD2-Datensatzes ohne Adresse und die gefundenen leerstehenden Gebäude konnten am Ende insgesamt 608 Gebäude ermittelt werden, welche für eine Kindertagesstätte infrage kommen würden. Dabei entsprechen die 608 Gebäude den Parametern, welche von der Kinderland GmbH vorgeschlagen wurden. Die genaue Zusammensetzung wurde bereits im Kapitel 5.2 erklärt. Neben dem Problem mit der Gebäudefunktion konnte auch erkannt werden, dass einige der Daten des Einwohnermeldeamtes nicht richtig gepflegt wurden. Wie bereits im Kapitel zur Analyse der Datenbasis aufgezeigt, liegen beispielsweise falsche Datumsformate vor. Diese führen während des Prozesses dazu, dass der betroffene Eintrag aus den Einwohnermeldedaten nicht berücksichtigt wird. Durch eine Ermittlung über den Algorithmus konnten insgesamt 36 falsche Datumsformate erkannt werden. Diese Anzahl an Einträgen ist noch nicht kritisch in Bezug auf eine relevante Verfälschung des Prozesses, jedoch wäre eine Bereinigung dieser Fehler empfehlenswert. Bei der Analyse der Altersstruktur wurde anhand der Nettozuwanderung das Wachstum pro Ortsteil ermittelt. Die Wachstumsrate bezieht sich dabei auf den Durchschnitt der letzten 10 Jahre und es wurden nur Bürger der Altersspanne 15-45 Jahre in der Analyse berücksichtigt. Dabei konnten festgestellt werden, dass Gröst und Langeneichstädt das größte Wachstum aufweisen. Hingegen zeigt Wünsch eine negative Wachstumsrate über die letzten 10 Jahre auf, d.h. es gibt eine höhere Abwanderung als Zuwanderung in diesem Ortsteil. Branderoda, Oechlitz und die Kern Stadt Mücheln weisen dabei im Vergleich eine sehr kleine bis neutrale Wachstumsrate vor. Aus der Analyse lässt sich nun also ableiten, dass für Gröst und Langeneichstädt, Oechlitz, als auch für die Kernstadt Mücheln ein Potential besteht, dass neue Flächen für Kindertagesstätten benötigt werden. Doch neben der Wachstumsrate wurden auch noch die Anzahl an Kindern in dem jeweiligen Ortsteil und dessen Einzugsgebiet ermittelt, so dass erkannt werden kann wie viele Kinder potenziell einen Kindergartenplatz benötigen. Folgende Tabelle zeigt die Anzahl von Kindern mit einem Einzugsgebiet von 5 km zum Ortsteil.

Ortsteil	Anzahl an Kindern (Umkreis 5 km)
Mücheln	188
Langeneichstädt	104
Branderoda	169
Gröst	86
Oechlitz	253
Wünsch	117

Tabelle 5: Anzahl der Kinder je Ortsteil mit Berücksichtigung eines Einzugsgebiets

Wie der Tabelle zu entnehmen ist, gibt es eine potenzielle Anzahl an Kindern von 253 in dem Ortsteil Oechlitz. Dieser Wert fällt gerade für diesen Ortsteil sehr hoch aus. Der Grund dafür ist die getätigte Annahme, dass ein Einzugsgebiet von 5 km vorliegt. Schaut man sich nun die Lage von Oechlitz an fällt auf, dass dieser Ortsteil sehr nah an andere Ortsteile der Stadt Mücheln grenzt. Hier entsteht ein weiteres Problem, welches es noch zu lösen gilt. Zwar ist das Einbeziehen eines Einzugsgebietes je Ortsteil aus Sicht einer Kindergartenplanung sehr wichtig, jedoch können dadurch auch die Zahlen für einzelne Ortsteile stark von einem realen Wert abweichen. Die berechneten Zahlen der Kinder je Ortsteil wurden nun verwendet, um einen Soll-Wert zu generieren, der die benötigte Grundfläche an Kindergartenbetreuungsplätzen aufzeigt. Neben dem Soll-Wert konnte auch ein Ist-Wert für die Ortsteile der Stadt Mücheln abgeleitet werden. Dieser setzt sich aus den Grundflächen der bereits vorhandenen Kindergärten je Ortsteil zusammen. Somit konnten folgende Werte ermittelt werden.

Ortsteil	Ist-Wert in m ²	Soll-Wert in m ²
Mücheln	1717,17	705
Branderoda	735,23	633,75
Gröst	210,04	322,5
Langeneichstädt	401,96	390
Oechlitz	573,12	948,75
Wünsch	259,54	438,75

Tabelle 6: Ist- und Soll-Wert an Flächen für Kindertagesstätten in Mücheln

Aus der obigen Tabelle ist zu erkennen, dass Gröst, Oechlitz und Wünsch einen Bedarf an neuen Flächen haben. Jedoch muss auch hier wieder gesagt werden, dass durch die Verwendung des Einzugsgebietes der Soll-Wert für Oechlitz beispielsweise sehr hoch ausfällt. Die tatsächlichen Auslastungen könnten jedoch abweichen. Werden nun die beiden Analysen der Altersstruktur zusammengefasst, lässt sich für jeden Ortsteil bestimmen, ob ein Bedarf vorliegt. Am Beispiel des Ortsteils Gröst ist zu erkennen, dass zum einen ein positives Wachstum der Bevölkerung zu verzeichnen ist und gleichzeitig auch der Soll-Wert größer ist als der Ist-Wert für benötigte Flächen. Somit kann für Gröst ein Bedarf an neuen Flächen erkannt werden. Aus diesem Grund werden die leerstehenden Gebäude innerhalb von Gröst am Ende auch in das Ergebnisdokument einfließen. Bei der Analyse des Ortsteils Gröst konnten insgesamt 69 Gebäude identifiziert werden, welche derzeit leerstehen und als potenzielle Fläche für Kindertagesstätten verwendet werden können. Derzeit werden noch alle

ermittelten Gebäude in die Ergebnisliste übernommen. Jedoch sollte später noch eine Verfeinerung der Auswahl stattfinden. Diese kann durch die Einbindung des Flächennutzungsplanes oder durch die Erweiterung der Datenbasis um Daten zu Gewerbeobjekten möglich werden. Durch die Einbeziehung des Flächennutzungsplanes können die Daten um ein Vielfaches besser bestimmt werden, da sich durch die tatsächliche Nutzung einer Fläche die Gebäude besser zuordnen lassen. Somit könnte später die Ausgabe noch verkürzt und die Qualität der Ausgabe verbessert werden. Folgende Abbildung zeigt einen Ausschnitt des Ergebnisdokuments am Beispiel der Stadt Mücheln. Zusätzlich sind im Anhang 4 weitere Abbildungen zum Thema Ergebnisdokument zu finden.

Ortsteile mit Bedarf an neuen Kindertagesstätten:

Ortsteil: Mücheln (Geiseltal) OT Gröst

Ergebnisse der Analyse zur Alterstruktur

Bedarfsermittlung	Wachstum des OT
IST-Grundfläche: 210.04m ²	Wachstumsrate: 0.23
Soll-Grundfläche: 322.5m ²	

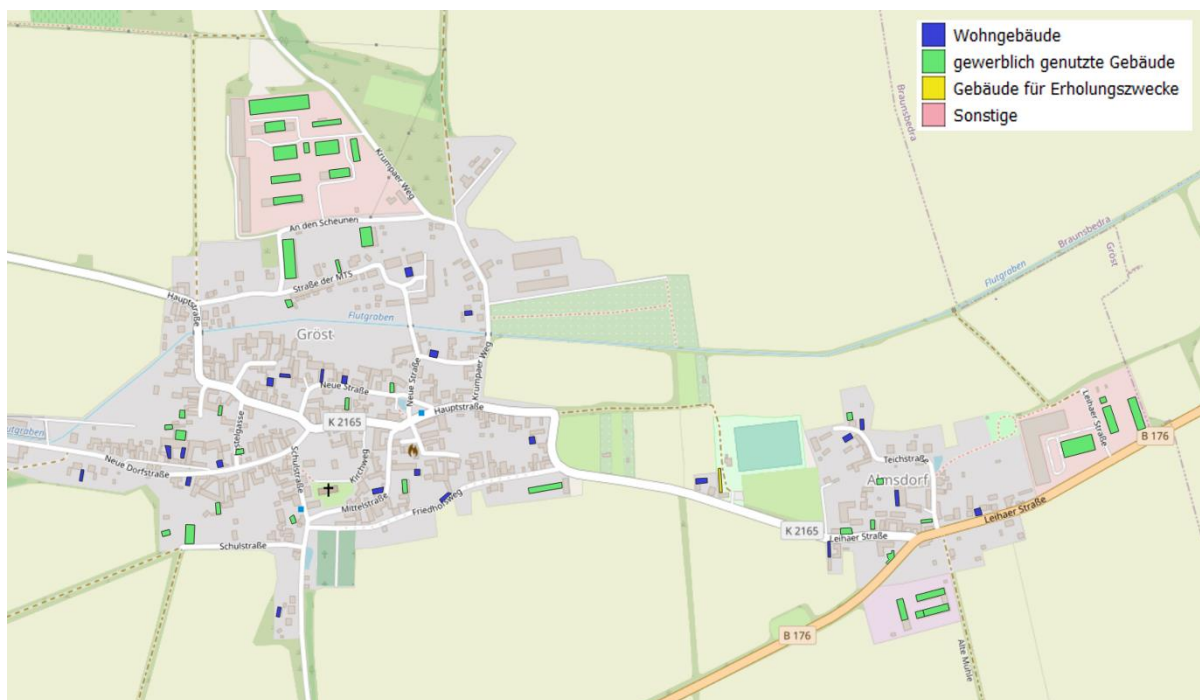
Für den Ortsteil Mücheln (Geiseltal) OT Gröst konnte Ein Bedarf an neuen Flächen für Kindergärten ermittelt werden. Folgende Gebäude würden theoretisch zur Verfügung stehen.

Gebäudedaten (LoD2)	Flurstückdaten
Gebäude ID: DEST_DESTLIKA0001e12M	Gemarkung: 2282
Gebäudefunktion: 31001_1000	Flur: 4
Straße/Hausnr.: Neue Dorfstraße 7	Zähler: 505
Dachtyp: PolyFlatRoof	Nenner: 0
Dachfläche: 90.0m ²	Besitzer: privat
Dachform: 1000	Lage: Neue Dorfstraße 7
Grundfläche: 98.01m ²	Grundfläche: 754.0m ²

Seite 2/71

Abbildung 23: Ausschnitt des Ergebnisdokument des Algorithmus für den OT Gröst
Datenquelle: eigene Darstellung

Wie der Abbildung zu entnehmen ist werden zum Ortsteil einerseits allgemeine Informationen zu den Analysen aufgezeigt und andererseits werden die ermittelten Gebäude und zusätzliche Daten zum Flurstück ausgegeben. Bis jetzt wird nur ein Teil der LoD2-Daten ausgegeben, da sonst die Übersichtlichkeit leiden würde, später könnte jedoch die Ausgabe um weitere Daten ergänzt werden. Dafür würden nicht nur die Daten des Flächennutzungsplanes in Frage kommen, sondern auch Geodaten zu den anliegenden Straßen oder auch die Analyse der Umgebung. Durch die Analyse der anliegenden Umgebung kann es ermöglicht werden, dass zusätzlich noch weitere Parameter berücksichtigt werden können, wie beispielsweise Parkflächen, Bus- und Bahnanbindungen oder auch Verkehrssicherheit. Diese angesprochenen Themen könnten später noch auf ihre Umsetzbarkeit geprüft und weiter erforscht werden. Abschließend gibt es noch die Visualisierung und Aufbereitung im QGIS, welche vor allem bei der Interpretation und Präsentation unterstützen soll. Dabei wird für jeden Ortsteil ein Vektorlayer in das Projekt eingefügt, welcher die Daten zu den Gebäuden pro Ortsteil besitzt. Zusätzlich werden die Layer farblich so gestaltet, dass die Gebäudefunktion erkannt werden kann. Folgende Abbildung zeigt die gefundenen Gebäude je Ortsteil.



**Abbildung 24: Ausschnitt des Ergebnisses des Algorithmus als Visualisierung im QGIS
Datenquelle: eigene Darstellung**

Durch die Resultate des Algorithmus konnte festgestellt werden, dass die Bereitstellung und die Verarbeitung von demografischen Daten und Geodaten im Zusammenhang möglich sind. Die Ergebnisse können von der Stadt Mücheln nun herangezogen werden, um neue Bedarfe für Kindertagesstätten zu ermitteln und geeignete Flächen und Gebäude dafür zu

identifizieren. Dabei kann dieser Anwendungsfall zur Stadtentwicklung beitragen, indem nicht verwendete Flächen aufbereitet werden und die Attraktivität der Stadt gesteigert werden kann, indem vor allem Familien mit kleinen Kindern ein Kindergartenplatz garantiert werden kann. Dabei sollte jedoch beachtet werden, dass die Aussagekraft der Ergebnisse nur so stark sein kann, wie die Qualität der Rohdaten ist. Sollten die Geodaten oder die demografischen Daten qualitativ schlecht aufgestellt sein, dann kann die Auswertung zu fehlerhaften Ergebnissen führen.

6.2 Weitere Möglichkeiten für die Stadt Mücheln

Neben den aufgezeigten Ansatz des Leerstandmanagements für Kindertagesstätten mit Hilfe von georeferenzierten Daten und demografischen Daten, gibt es auch noch viele andere Anwendungsgebiete für Geodaten im Bereich der Stadtentwicklung. Durch die mitgelieferten Gebädefunktionen lässt sich der Algorithmus auf fast jede Funktion übertragen und somit könnte der Algorithmus auch für die Bestimmung von Schulen, Supermärkten, Freizeitangeboten und vieles mehr verwendet werden. Doch neben dem Leerstandsmanagement lassen sich noch viele andere Möglichkeiten für die Verwendung von Geodaten in Bezug auf die Stadtentwicklung finden. Allein der LoD2-Datensatz bietet noch genügend Potential für weitere Forschungsthemen. Ein aktuelles Beispiel wäre die Verwendung der Gebäudegeometrien für die Berechnung der Solarkapazität für ein Gebäude. Ein Solarkataster bietet die Möglichkeit anhand der Dacheigenschaften, also Form, Typ oder auch Fläche, die Eignung des Daches für Solaranlagen zu bestimmen. Ebenso lassen sich anhand der Dachfläche verschiedene Parameter berechnen, wie beispielweise die potenziell erzeugbare Leistung der Solaranlagen.⁹⁶ Ein aktuelles Beispiel für ein bereits umgesetztes Solarkataster ist für das Bundesland Hessen zu finden. Dieses kann unter der URL ‚www.solarkataster.hessen.de‘ erreicht werden. Für den Aufbau eines Solarkatasters bedarf es jedoch neben den 3D Gebäudemodellen auch einer Analyse des Solarpotentials pro Gebäude. Dieses Potential kann anhand von der Dachneigung, Dachausrichtung, Verschattung durch andere Objekte und die Sonneneinstrahlung berechnet werden.⁹⁷ Somit müsste die Stadt Mücheln neben den 3D Gebäudemodellen auch noch die Geodaten zur städtischen Vegetation bereitstellen und eine geeignete Potentialanalyse aufstellen. Würden auch die Flächennutzungspläne und die Flurstückdaten mit einbezogen, lassen sich eine Vielzahl an weiteren Anwendungsfällen aufstellen. Ein weiteres Beispiel wäre die Analyse von freien Flächen für den Aufbau von

⁹⁶ Vgl. Hessisches Ministerium für Wirtschaft, Energie, Verkehr und Landesentwicklung 2017, S. 4.

⁹⁷ Vgl. Hessisches Ministerium für Wirtschaft, Energie, Verkehr und Landesentwicklung 2017, S. 6.

Windparks. Dafür müssten die freien Flächen hinsichtlich ihrer Nutzung analysiert und ihren Standort hin geprüft werden. Dadurch könnten Standorte gefunden werden, welche zum einen den benötigten Platz bieten und zum anderen, unter Berücksichtigung der anliegenden Bevölkerung, den richtigen zur Wohnbebauung berechnen. Durch eine solche Anwendung können auch Kleinstädte wie Mücheln einen großen Beitrag zum Ziel der Klimaneutralität leisten. Auch für die eigene Bevölkerung oder die weitere Entwicklung des Tourismus in dieser Geiseltalregion kann die Anwendung von Geodaten eine große Hilfe darstellen. Beispielsweise könnte ein Geoportal für die Stadt Mücheln aufgebaut werden, welches den Bürgern verschiedene Werkzeuge zur Verfügung stellt (bspw. Bürgerinformationsschnittstelle, Einbindung in Bauvorhaben, etc.) oder auch für Touristen die Möglichkeit bietet, einen guten Überblick über die Stadt zu bekommen. Eine Möglichkeit wäre die Verwendung des Open Source Kartenclient QGIS Webclient oder auch QWC. Dieser ermöglicht es Daten aus QGIS nach Außen bereitzustellen. Letztlich kann festgestellt werden, dass es noch eine Vielzahl an Forschungsansätzen im Bereich der Verarbeitung von Geodaten für die kommunale Stadtentwicklung gibt.

7 Fazit und Ausblick

In dieser Forschungsarbeit wurde am Beispiel der Stadt Mücheln (Geiseltal) analysiert, welchen Einfluss die Verwendung demografischer Daten und Geodaten im Hinblick auf die kommunale Stadtentwicklung haben kann. Dabei wurden zuerst die theoretischen Grundlagen behandelt, welche zum einen das derzeitige Problem des demografischen Wandels und der damit verbundenen Abwanderung aus Kleinstädten aufzeigen soll und zum anderen wurden in das Thema der georeferenzierten Daten eingeleitet. Ein wichtiger Punkt war hier die Beschreibung des LoD2-Datensatzes, welcher dann eine Grundlage für die Aufstellung des Planungsalgorithmus darstellte. Nachdem die Grundlagen erläutert wurden, erfolgte der Bezug zur Stadt Mücheln und dabei wurde insbesondere auf die bereitgestellten Daten eingegangen. Hierbei handelte es sich um Daten des Einwohnermeldeamtes, welche Informationen über die aktuelle Altersstruktur und Wohnsituation liefern. Außerdem wurden verschiedene Geodaten bereitgestellt, wie die LoD2-Daten und die Flurstückdaten aus dem ALKIS. Während der Umsetzung des Algorithmus wurde erkannt, dass es noch zusätzlicher Daten bedarf, welche als Open Daten durch die OSM-Schnittstelle bereitgestellt werden müssen. Im Anschluss wurde in das Experiment eingeleitet und die Zielsetzung des Algorithmus vorgestellt. Ziel sollte es sein anhand der 3D Gebäudemodelle und den demografischen Daten den Leerstand der Stadt zu identifizieren und anhand der Altersstruktur Aussagen zum Bedarf an neuen Nutzungsflächen zu treffen. Zur näheren Beschreibung eines möglichen Faktors (hier Kindertagesstätten) wurde ein Interview mit der Kinderland GmbH durchgeführt. Die Informationen aus dem Interview wurden dann für die Analyse der Daten und die Erstellung der Logik für den Planungsalgorithmus herangezogen. Zu den wichtigsten Informationen zählten dabei die allgemeinen Parameter, welche ein Kindergarten aufweisen muss. Anschließend wurde eine Datenanalyse durchgeführt, welche die Altersstruktur aufschlüsselt und den Nutzen des Planungsalgorithmus aufzeigen sollte. Bei der Analyse konnte festgestellt werden, dass über die letzten 10 Jahre deutlich mehr ältere Menschen in Mücheln leben bzw. gelebt haben als jüngere. Somit konnte ein Faktor des demografischen Wandels nachgewiesen werden, indem aufgezeigt wurde, dass gleichzeitig die Anzahl an Kindern im Alter von 0-6 Jahren seit 10 Jahren rückläufig ist. Parallel dazu wurde aber auch ein Wachstum der Bevölkerungsanzahl in der Altersgruppe 15-45 in einzelnen Ortsteilen nachgewiesen. Damit wurden die Grundlagen für den Planungsalgorithmus bereitgestellt und nachgewiesen, dass dieser für die Kommune einen Nutzen haben kann. Nach der Analyse folgt das Kapitel zur Umsetzung, welches den Aufbau und die technischen Bausteine bzw. die Logik

der Anwendung aufzeigt. Während der Umsetzung wurden verschiedenste Anwendungsbereiche mit den Geodaten ausgetestet und implementiert. Dabei wurde nicht nur der Planungsalgorithmus aufgestellt, sondern es wurden auch noch weitere hilfreiche Anwendungen entworfen, welche durch Geodaten und andere prozessrelevante Daten Auswertungen erzeugen. Diese Anwendungen wurden im Kapitel 5.4 aufgezeigt und auch deren Nutzen erläutert. Bei der Entwicklung dieser Anwendungen und dem Algorithmus konnte das große Potential der Geodaten erkannt werden. Dabei bieten die Geodaten nicht nur die Möglichkeit eine gute Visualisierung zu erstellen, sondern sie bieten auch die Möglichkeit komplexe Analysen anhand der Georeferenz durchzuführen. Ein Beispiel ist die Bestimmung der Lage von einzelnen Gebäuden über die Koordinaten oder auch die Verwendung von Grundflächen, die aus den Koordinaten ermittelt werden können. Noch größer wird das Potential, wenn die Geodaten mit anderen Fachdaten, wie beispielsweise den demografischen Daten, verbunden werden. Dadurch entstehen neue Möglichkeiten der Visualisierung von komplexen Fragestellungen bzw. Herausforderungen von Kommunen. Das Resultat einer guten Implementierung und Verwendung der eigenen Geodaten und Fachdaten könnte zu einer moderneren Stadtentwicklung, auch auf kommunaler (ländlicher) Ebene führen. Jedoch musste auch erkannt werden, dass die Qualität der Ergebnisse der Anwendungen stark von der Genauigkeit der Daten abhängt. Leider wiesen verschiedene Datensätze des Einwohnermeldeamtes Fehler auf, die dazu führten, dass ungenaue Ergebnisse auftraten. Bei der Beschaffung der benötigten Geodaten und Fachdaten sollte auch darauf geachtet werden, dass diese im richtigen Format und der richtigen Qualität vorliegen. Aus diesem Grund wurde, für die Datenlage der Kommune Mücheln, auch ein ELT-Prozess entworfen, welcher sicherstellen soll, dass die Daten zu jeder Zeit qualitativ hochwertig zur Verfügung stehen. Dieser ELT-Prozess beinhaltet ein Extraktions-Tool, welches die Rohdaten aus den verschiedenen Quellsystemen sammelt und diese in eine Zieldatenbank überführt. Der restliche Prozess findet überwiegend direkt in dem entwickelten Plugin im QGIS statt. Somit werden die Daten durch das Plugin geladen und transformiert. Dadurch kann die Aktualität der Daten gewahrt und gleichzeitig die Qualität erhöht werden.

Abschließend kann gesagt werden, dass die Verwendung von georeferenzierten Daten in der kommunalen Stadtentwicklung eine große Unterstützung für Entscheidungsprozesse darstellen kann. Diese aufbereiteten Daten ermöglichen eine sehr gute Visualisierung von Fachthemen, welche durch die Ergänzung von Fachdaten noch spezifiziert werden können. Gerade in der Stadtentwicklung sind solche Visualisierungsmodelle sehr wichtig und helfen den Beteiligten dabei, ein besseres Verständnis zu entwickeln. Die Forschung rundum die

Verwendung von Geodaten für die Stadtentwicklung steht noch ganz am Anfang und wird mit der fortschreitenden Digitalisierung immer weiter vorangetrieben. So werden die Geodaten, welche vom Vermessungsamt geliefert werden, mindestens jährlich erneuert und bei jeder Aktualisierung werden die Datensätze weiter ausgebaut. Somit können die Kommunen mit aktuellen Daten ihre jeweiligen Fragestellungen bearbeiten. Im Kapitel 6.2 wurde aufgezeigt, dass neben Stadtentwicklungsfragen auch noch weitere Themenbereiche als Forschungsgebiete einbezogen werden können.

Für die vorliegende Untersuchung wäre es sinnvoll noch folgende Thematik zu bearbeiten. Um die Genauigkeit des Ergebnisdokuments zu verbessern, sollte der Flächennutzungsplan noch in den Algorithmus integriert werden. So könnten Gebäude ausgeschlossen werden, welche nicht für den jeweiligen Faktor zu verwenden sind. Dafür müsste der Flächennutzungsplan mit den 3D Gebäudemodellen des LoD2-Datensatzes und den Flurstücken abgeglichen werden, um eine Verbindung zwischen den Daten aufbauen zu können. Des Weiteren sollten noch die Daten zu gewerblich genutzten Gebäuden ausgewertet werden und ebenfalls in die Analyse integriert werden. Somit könnten diese Gebäude, analog den bewohnten Gebäuden, aus der Analyse ausgeschlossen werden, da bereits genutzte Gebäude nicht für das Leerstandsmanagement herangezogen werden können. Auch könnten die Berechnungen für die Analyse der Altersstruktur noch verfeinert werden, so dass diese präzisiert wird und die Qualität der Prozessdaten verbessert werden könnte. Wichtig wäre es auch, speziell für die Stadt Mücheln noch eine direkte Verbindung zwischen den Quelldaten und der Zieldatenbank herzustellen, so dass das Extraktions-Tool die Daten direkt von der Datenbank der Stadt Mücheln beziehen kann und kein manueller Prozess von Nöten ist. Letztlich sollte auch noch ein Plan erarbeitet werden, welcher es ermöglicht die entwickelten Analysen in das Tagesgeschäft der Stadt einzubinden. Hierzu stellt die gezielte Schulung von Mitarbeitern eine Voraussetzung dar.

Schlussendlich kann gesagt werden, dass die Verwendung von Geodaten und Geoinformationssystemen dazu beitragen kann, den manuellen Aufwand im Bereich der Stadtentwicklung bzw. Planung für Kleinstädte zu minimieren und gleichzeitig neue Möglichkeiten eröffnet werden konkrete Sachverhalte zu analysieren und auszuwerten. Dies trägt zur Reduzierung von Kosten und zur Erhaltung wertvoller Ressourcen von Kommunen bei.

8 Literatur

Alan R. Hevner, et al. 2004: Design Science in Information Systems Research. . URL: https://wise.vub.ac.be/sites/default/files/thesis_info/design_science.pdf - Download vom 16.07.2022.

Antonia Milbert, L. P. 2021: Kleinstädte in Deutschland. . URL: https://www.bbsr.bund.de/BBSR/DE/veroeffentlichungen/sonderveroeffentlichungen/2021/kleinstaedte-in-deutschland-dl.pdf?__blob=publicationFile&v=1 - Download vom 10.04.2022.

Astera 2019: Data Mapping: Definition, Tools, and Techniques. . URL: <https://www.astera.com/wp-content/uploads/2019/07/DataMapping.pdf> - Download vom 17.07.2022.

Bayerische Vermessungsverwaltung 2018: Kundeninformation LoD2 Gebäudemodelle. . URL: https://www.ldbv.bayern.de/file/pdf/6723/Kundeninformation_LoD2.pdf - Download vom 17.07.2022.

BBSR: Laufende Stadtbeobachtung - Raumabgrenzungen. Stadt- und Gemeindetypen in Deutschland. . URL: <https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/gemeinden/StadtGemeindetyp/StadtGemeindetyp.html?nn=2544954> - Download vom 11.07.2022.

BBSR 2016: Potenziale von Kleinstädten in peripheren Lagen. . URL: https://www.bbsr.bund.de/BBSR/DE/veroeffentlichungen/exwost/50/exwost-50-1.pdf?__blob=publicationFile&v=2 - Download vom 20.06.2022.

Benjamin Seibel 2016: Open Data in der Praxis. . URL: file:///C:/Users/kevin/Downloads/160128_tsb_opendataberlin.pdf - Download vom 17.07.2022.

BMI 2021: Open-Data-Strategie der Bundesregierung. . URL: - Download vom 17.07.2022.

BMI 2021: Open-Data-Strategie der Bundesregierung. . URL: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/moderne-verwaltung/open-data-strategie-der-bundesregierung.pdf?__blob=publicationFile&v=4 - Download vom 22.06.2022.

- Borrmann, A., et al. (2015): Building Information Modeling, Springer Fachmedien Wiesbaden, Wiesbaden.
- Bundesamt für Kartographie und Geodäsie 2022: Dokumentation 3D Gebäudemodell Deutschland - LoD2. . URL: https://sg.geodatenzentrum.de/web_public/gdz/dokumentation/deu/LoD2-DE.pdf - Download vom 17.07.2022.
- Bundesministerium des Innern 2012: Open Government Data Deutschland. Eine Studie zu Open Government in Deutschland im Auftrag des Bundesministerium des Innern. Letzte Aktualisierung: 17.07.2022. URL: https://www.verwaltung-innovativ.de/SharedDocs/Publikationen/eGovernment/open_government_data_deutschland_langfassung.pdf?__blob=publicationFile&v=5 - Download vom .
- Bundesministerium für Wirtschaft und Technologie 2001: Open-Source-Software. . URL: <https://nats-www.informatik.uni-hamburg.de/pub/OSS2004/PaperCollection/Open-Source-Software.pdf> - Download vom 17.07.2022.
- Bundeszentrale für politische Bildung 2022: Informationen zur politischen Bildung – Demografischer Wandel. . URL: <https://www.bpb.de/shop/zeitschriften/izpb/507591/demografischer-wandel/> - Download vom 27.07.2022.
- ESRI 1998: ESRI Shapefile Technical Description. . URL: <https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/library/whitepapers/pdfs/shapefile.pdf> - Download vom 17.07.2022.
- Farkisch, K. 2011: Data-Warehouse-Systeme kompakt. . URL: - Download vom .
- Ferstl, O. K.; Sinz, E. J. (2013): Grundlagen der Wirtschaftsinformatik. 7. Auflage, Oldenbourg, München.
- GeoZG (2009): Gesetz über den Zugang zu digitalen Geodaten. GeoZG.
- Gerold Noack (2019): Geodäsie für Bauingenieure und Architekten. Grundlagen und Praxiswissen, Carl Hanser Verlag GmbH & Co. KG.
- Harwardt, M. (Hrsg.) (2022): Management der digitalen Transformation. Eine praxisorientierte Einführung. 2. Auflage, Springer Fachmedien Wiesbaden, Wiesbaden.
- Hessisches Ministerium für Wirtschaft, Energie, Verkehr und Landesentwicklung 2017: Solar-Kataster Hessen. LEITFADEN zur Nutzung des Solar-Katasters. . URL:

Literatur

https://www.energieland.hessen.de/mm/Leitfaden_Solar-Kataster-Hessen_201709.pdf - Download vom 28.08.2022.

Heuermann, R.; Tomenendal, M.; Bressemer, C. (Hrsg.) (2018): Digitalisierung in Bund, Ländern und Gemeinden. IT-Organisation, Management und Empfehlungen, Springer Berlin Heidelberg, Berlin, Heidelberg.

Hintze, P.; Lakes, T. 2009: Geographically Referenced Data in Social Science: A Service Paper for SOEP Data Users. . URL: https://www.econstor.eu/bitstream/10419/129257/1/diw_datadoc_2009-046.pdf - Download vom 18.07.2022.

https://www.lvermgeo.sachsen-anhalt.de/de/download_lod2.html: Kostenfreies 3D-Gebäudemodell - LoD2. . URL: https://www.lvermgeo.sachsen-anhalt.de/de/download_lod2.html - Download vom 17.07.2022.

Jörn von Lucke (2017): Technische Innovation – Potenziale von Open Government, offenen Daten und intelligenten Städten, in: Kersting, N. (Hrsg.) (2017): *Urbane Innovation*, Springer Fachmedien Wiesbaden, Wiesbaden.

Jörn von Lucke und Christian Geiger 2010: Open Government Data - Frei verfügbare Daten des öffentlichen Sektors. Gutachten für die Deutsche Telekom AG zur T-City Friedrichshafen. Letzte Aktualisierung: 17.07.2022. URL: <https://www.zu.de/institute/togi/assets/pdf/TICC-101203-OpenGovernmentData-V1.pdf> - Download vom .

Kay Ruge 2014: »Geodaten sinnvoll nutzen« – durch kommunale Kooperation zum Erfolg. . URL: <https://bayern.dvw.de/sites/default/files/landesverband/bayern/anhang/beitrag-skontext/2014/ruge.pdf> - Download vom 17.7.2022.

Keboola 2022: ETL vs ELT data pipelines. A guide to ETL vs ELT data pipelines. . URL: [https://assets.website-files.com/5e6f9b297ef3941db2593ba1/621f6c6d7f4f1731447cf416_Ebook%20-%20ETL%20vs%20ELT%20\(1\).pdf](https://assets.website-files.com/5e6f9b297ef3941db2593ba1/621f6c6d7f4f1731447cf416_Ebook%20-%20ETL%20vs%20ELT%20(1).pdf) - Download vom 26.08.2022.

Kelsey Jordahl 2016: GeoPandas Documentation. . URL: <http://sethc23.github.io/wiki/Python/geopandas.pdf> - Download vom 26.08.2022.

Klenk, T.; Nullmeier, F.; Wewer, G. (Hrsg.) (2020): Handbuch Digitalisierung in Staat und Verwaltung, Springer Fachmedien Wiesbaden, Wiesbaden.

Klenk, T.; Nullmeier, F.; Wewer, G. (Hrsg.) (2020): Handbuch Digitalisierung in Staat und Verwaltung, Springer Fachmedien Wiesbaden, Wiesbaden.

Literatur

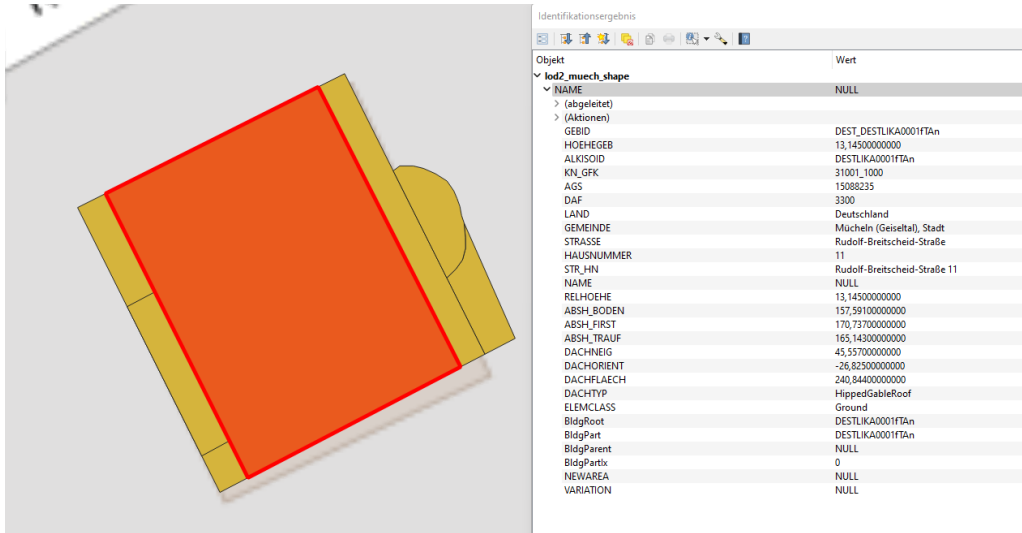
- Kubicek, H.; Jarke, J. (2020): Offene Daten (Open Data),
- Lange, N. de (2020): Geoinformatik in Theorie und Praxis. 4. Auflage, Springer Berlin Heidelberg, Berlin, Heidelberg.
- LVErmGeo 2021: Das Amtliche Raumbezugssystem. Letzte Aktualisierung: 10.07.2022.
URL: https://www.lvermgeo.sachsen-anhalt.de/de/datei/download/id/4016,501/afis_broschuere.pdf - Download vom .
- mediaprint infoverlag gmbh 2020: Stadt Muecheln - historisch, kulturell, se(e)henswert. .
URL: https://www.total-lokal.de/city/muecheln/data/06249_50_01_20/5/index.html#zoom=z - Download vom 11.07.2022.
- Menke, K. 2015: Mastering QGIS. . URL: - Download vom .
- Niedersächsische Vermessungs- und Katasterverwaltung 2010: Basiswissen ALKIS / ETRS 89. . URL: https://www.lgln.niedersachsen.de/download/126790/Basiswissen_ALKIS_ETRS89_Schulungsmaterial_Stand_12.04.2010.pdf - Download vom 15.07.2022.
- Ostrau; Kreis Lippe (Dr. Stefan): Einsatz von Geoinformationen in den Kommunen. Ergebnisse der Umfrage Good Practice Beispiele Handlungsempfehlungen. . URL: file:///C:/Users/kevin/Downloads/einsatz_geoinformationen_in_kommunen_studie_2013.pdf - Download vom 17.07.2022.
- Ottmann, T.; Widmayer, P. (2012): Algorithmen und Datenstrukturen. 5. Auflage, Spektrum Akademischer Verlag, Heidelberg.
- Peucker et al 2003: Familie, Kinder, Beruf. Familienunterstützende Kinderbetreuungsangebote in der Praxis. . URL: - Download vom 17.07.2022.
- Ponniah, P. (2010): Data Warehousing Fundamentals for IT Professionals, Second Edition. 2. Auflage.
- Porsche, L.; Steinführer, A.; Sondermann, M. 2019: Kleinstadtforschung in Deutschland. Stand, Perspektiven und Empfehlungen = Small town research in Germany : status quo, perspectives and recommendations. . URL: - Download vom .
- Prof. Dr. Rolf Hüpen 2022: Zur Berechnung von Wachstumsraten in diskreter Zeit. . URL: <https://www.ruhr-uni-bochum.de/wista/download/Beilagen/Wachstumsraten.pdf> - Download vom 26.08.2022.

Literatur

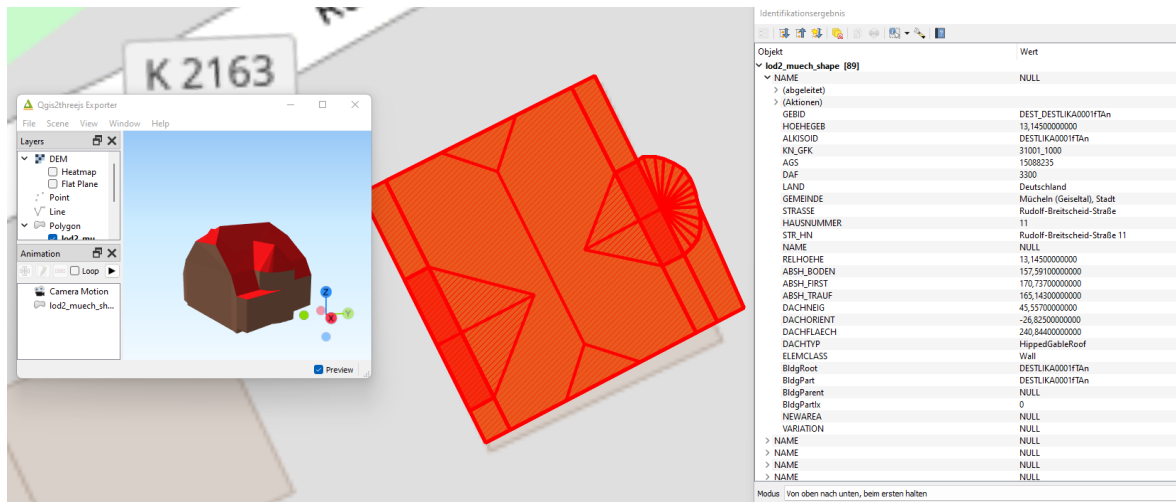
- Ralf Bill 2018: Handwörterbuch der Stadt- und Raumentwicklung. . URL: <https://www.econstor.eu/bitstream/10419/225734/1/HWB-SRE-0787-0792.pdf> - Download vom 15.07.2022.
- Selbach, S. 2011: Hybride bitparallele Volltextsuche. Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der Julius-Maximilians-Universität Würzburg. . URL: https://opus.bibliothek.uni-wuerzburg.de/opus4-wuerzburg/frontdoor/deliver/index/docId/5389/file/selbach_diss.pdf - Download vom 17.07.2022.
- Shamir, R. R., et al. 2018: Continuous Dice Coefficient: a Method for Evaluating Probabilistic Segmentations. . URL: - Download vom 20.07.2022.
- Solymosi, A.; Grude, U. (2017): Grundkurs Algorithmen und Datenstrukturen in JAVA, Springer Fachmedien Wiesbaden, Wiesbaden.
- Stadt Leipzig: Open Data-Portal der Stadt Leipzig. . URL: <https://opendata.leipzig.de/dataset> - Download vom 17.07.2022.
- Statistisches Bundesamt 2022: Durchschnittliches Alter der Eltern bei der Geburt des Kindes 2020. . URL: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Geburten/Tabellen/geburten-eltern-biologischesalter.html> - Download vom 27.07.2022.
- Statistisches Bundesamt 2022: Kinder in Tageseinrichtungen: Deutschland, Stichtag, Altersgruppen. . URL: <https://www-genesis.destatis.de/genesis/online?sequenz=tabelleErgebnis&selectionname=22541-0001#abreadcrumb> - Download vom .
- Statistisches Landesamt Sachsen-Anhalt 2022: Bericht Bevölkerung der Gemeinden Stand: 31.12.2021. . URL: - Download vom 11.07.2022.
- Stuckenholz, A. 2020: Basiswissen Energieinformatik. Ein Lehr- und Arbeitsbuch für Studierende und Anwender. . URL: - Download vom .
- Volker Galle 2014: Kommunale Geodaten im Kontext von Geodateninfrastruktur und Vernetzung. . URL: https://www.lvermgeo.sachsen-anhalt.de/datei/anzeigen/id/4130,501/10_Galle.pdf - Download vom 17.07.2022.
- Wido Geis-Thöne 2020: Kinderbetreuung: Über 340.000 Plätze für unter Dreijährige fehlen. . URL: https://www.iwkoeln.de/fileadmin/user_upload/Studien/Kurzberichte/PDF/2020/IW-Kurzbericht_2020-_Kinderbetreuung.pdf - Download vom 17.07.2022.

Anhangsverzeichnis

Anhang 1: LoD2-Daten Beispiele



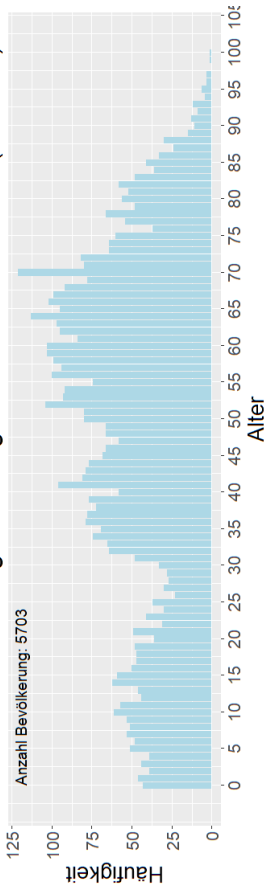
Digitales Gebäudemodell mit der Detaillierung LoD1



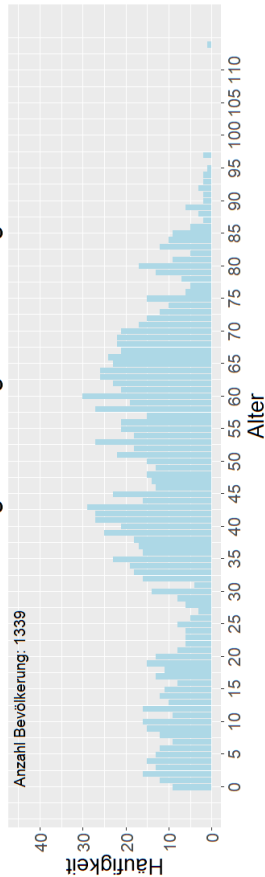
Digitales 3D Gebäudemodell mit der Detaillierung LoD2

Anhang 2: Durchgeführte statistische Analysen

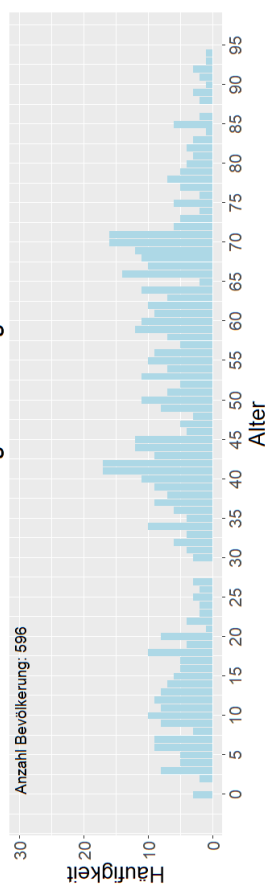
Aktuelle Bevölkerungsverteilung der Kernstadt Müheln (Geiseltal)



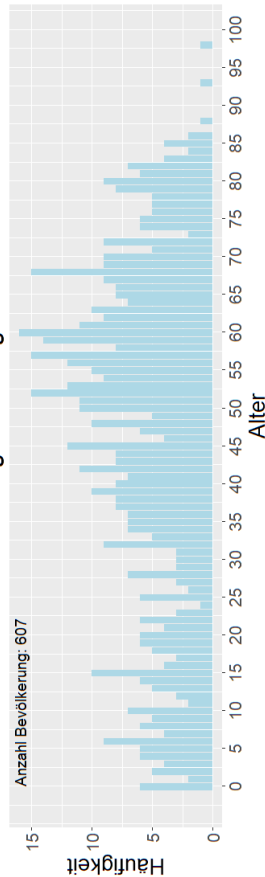
Aktuelle Bevölkerungsverteilung Ortsteil Langeneichstädt



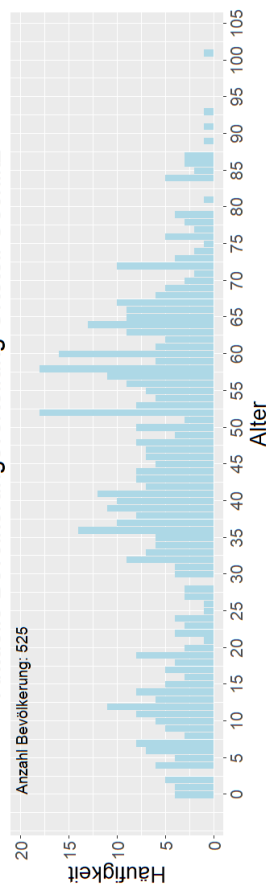
Aktuelle Bevölkerungsverteilung Ortsteil Gröst



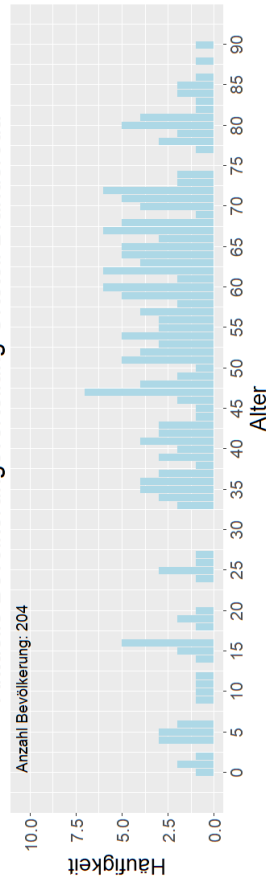
Aktuelle Bevölkerungsverteilung Ortsteil Wünsch



Aktuelle Bevölkerungsverteilung Ortsteil Oechlitz



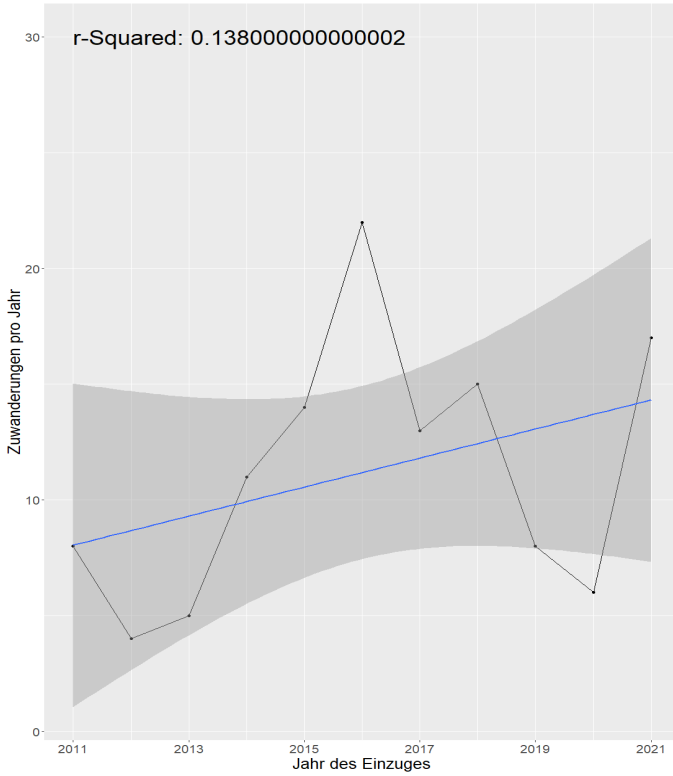
Aktuelle Bevölkerungsverteilung Ortsteil Branderoda



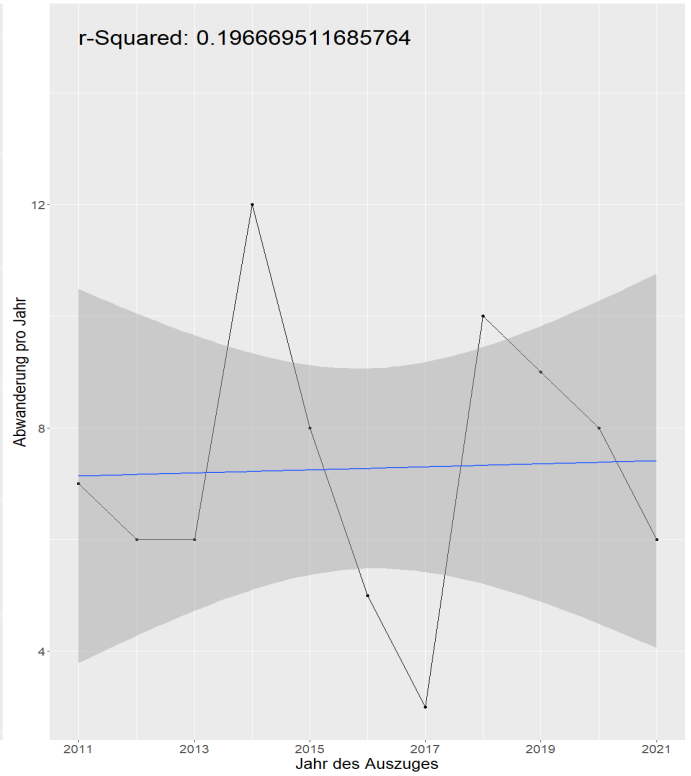
	Anzahl Bevölkerung	Anzahl Kinder	Durchschnittliches Alter	Verhältnis zur Gesamtanzahl
Müheln Gesamt	8974	512	47.22	100.00
Müheln	5703	310	47.86	63.55
Langeneichstädt	1339	90	46.18	14.92
Gröst	596	32	45.94	6.64
Wünsch	607	38	46.66	6.76
Oechlitz	525	30	43.56	5.85
Branderoda	204	12	51.22	2.27

Bevölkerungsverteilung nach Ortsteilen der Stadt Müheln (Geiseltal) im Jahr 2021

Lin. Regression zwischen Zuwanderung und Jahr der Zuwanderung für den OT Gröst

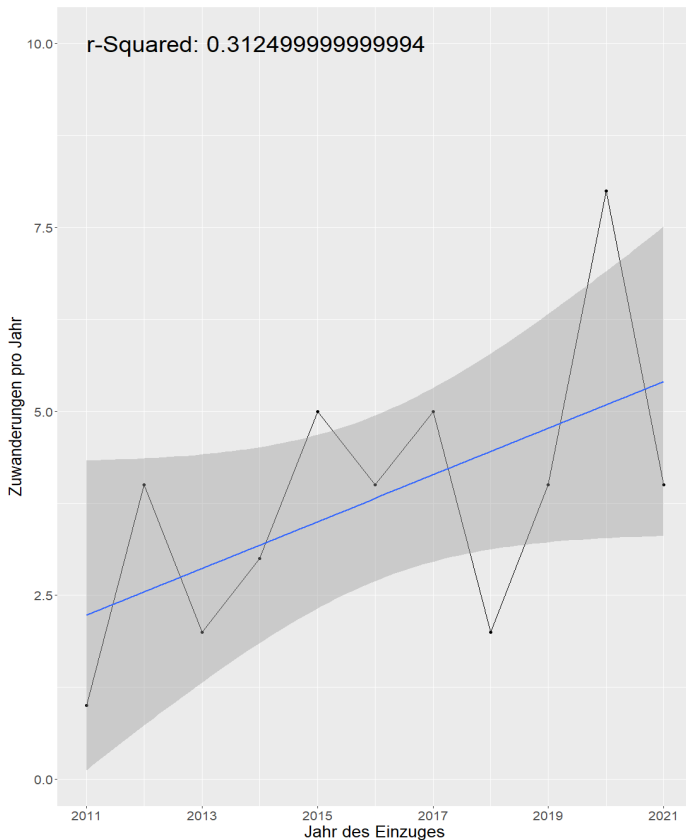


Lin. Regression zwischen Abwanderung und Jahr der Abwanderung für den OT Gröst

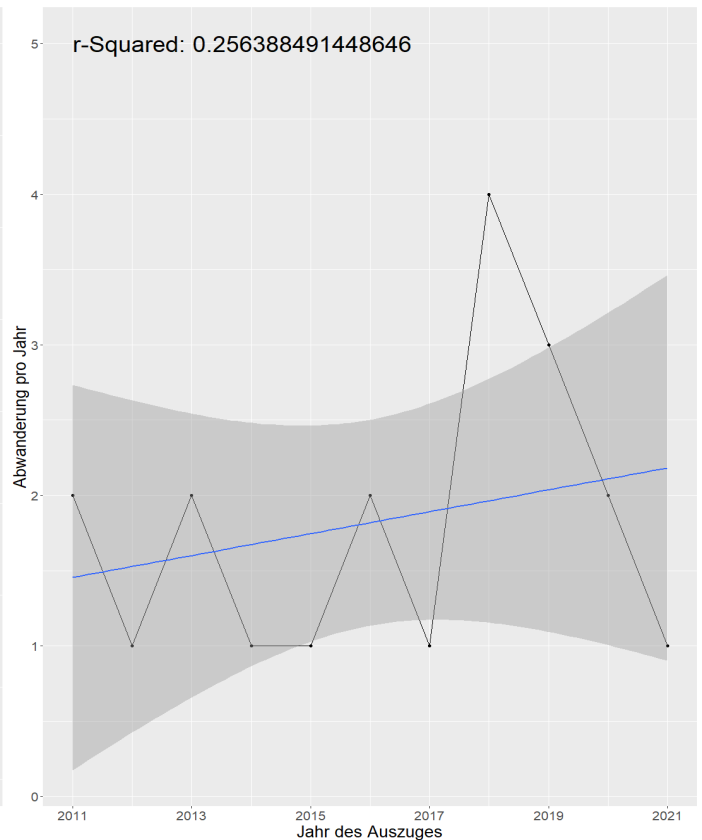


Lineare Regression für die Zu- und Abwanderung der Bevölkerung für den Ortsteil Gröst

Lin. Regression zwischen Zuwanderung und Jahr der Zuwanderung für den OT Branderoda

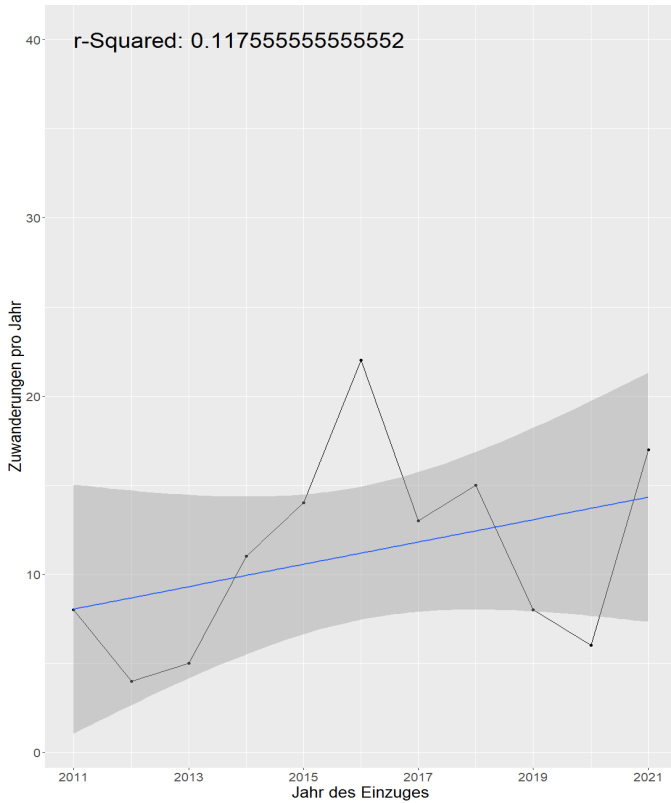


Lin. Regression zwischen Abwanderung und Jahr der Abwanderung für den OT Branderoda

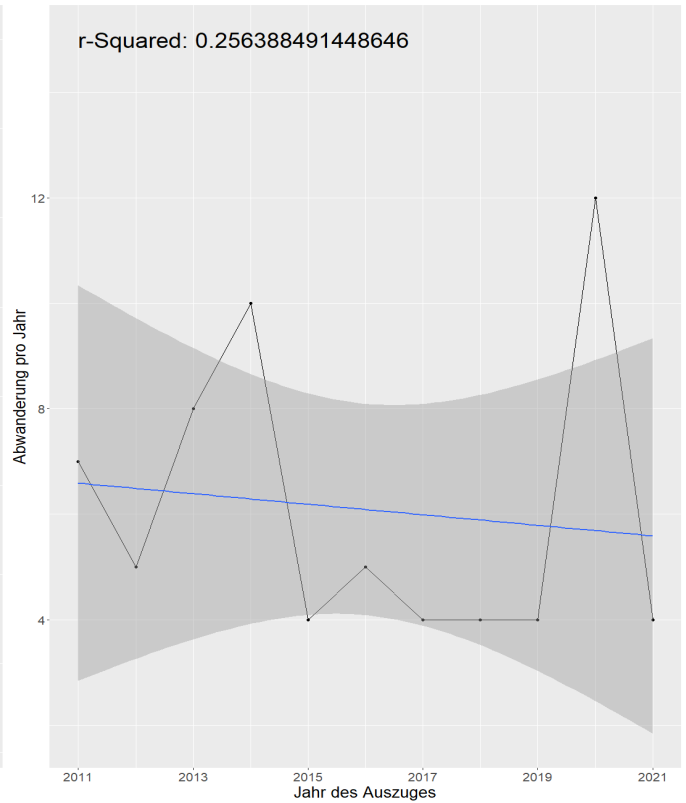


Lineare Regression für die Zu- und Abwanderung der Bevölkerung für den Ortsteil Branderoda

Lin. Regression zwischen Zuwanderung und Jahr der Zuwanderung für den OT Öchlitz

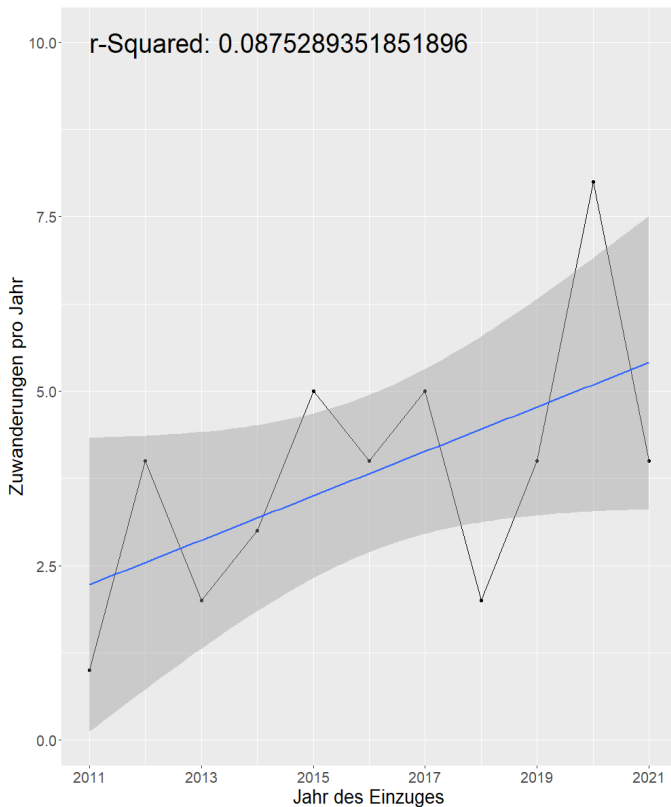


Lin. Regression zwischen Abwanderung und Jahr der Abwanderung für den OT Öchlitz

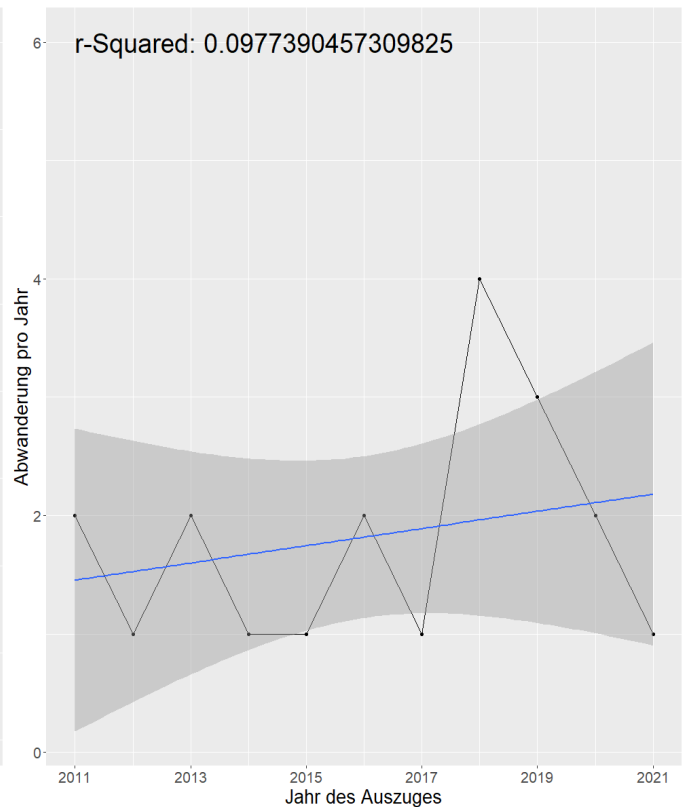


Lineare Regression für die Zu- und Abwanderung der Bevölkerung für den Ortsteil Öchlitz

Lin. Regression zwischen Zuwanderung und Jahr der Zuwanderung für den OT Wünsch

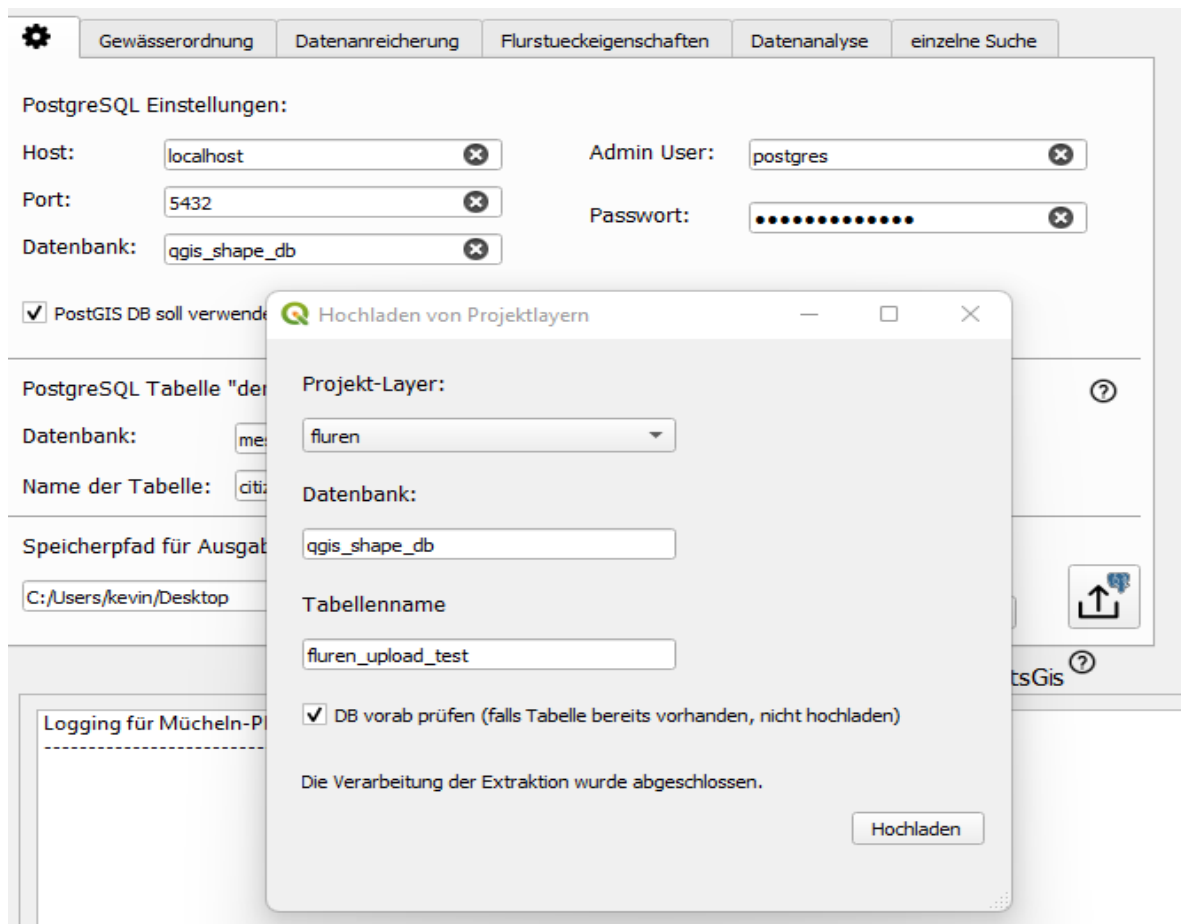


Lin. Regression zwischen Abwanderung und Jahr der Abwanderung für den OT Wünsch



Lineare Regression für die Zu- und Abwanderung der Bevölkerung für den Ortsteil Wünsch

Anhang 3: Beispiele zur Umsetzung und für Visualisierungen in QGIS



QGIS Plugin Funktion: Hochladen von Projektlayern

Mücheln (Geiseltal) Plugin - Geodatenanalyse

Gewässerordnung | Datenanreicherung | Flurstueckseigenschaften | Datenanalyse | **einzelne Suche**

Suchstring (Bitte jedes Element mit einem Semikolon trennen):

Wähle einen Modus für die Suche:
 Adresssuche Flurstück

Layer der durchsucht werden soll:

Upload zur PotsGis

Starten Schließen

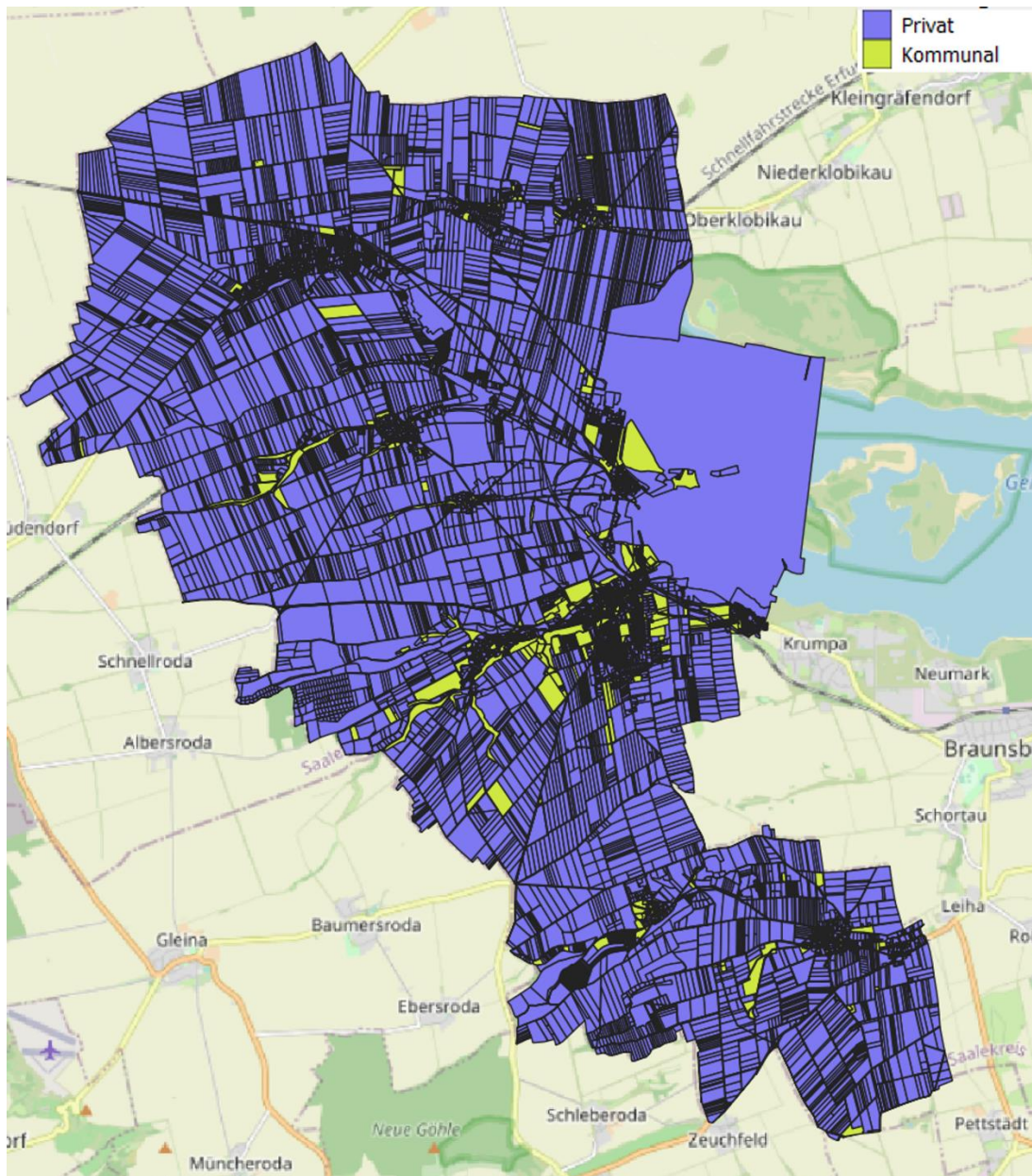
Stadt Mücheln (Geiseltal)
 QGIS Plugin "Entwicklung eines Konzepts zur datenbasierten Stadtentwicklung auf kommunaler Ebene" Masterarbeit Wirtschaftsinformatik Kevin Pocher, Hochschule Merseburg SoSe2022
 Suche TODO

Logging für Mücheln-Plugin:

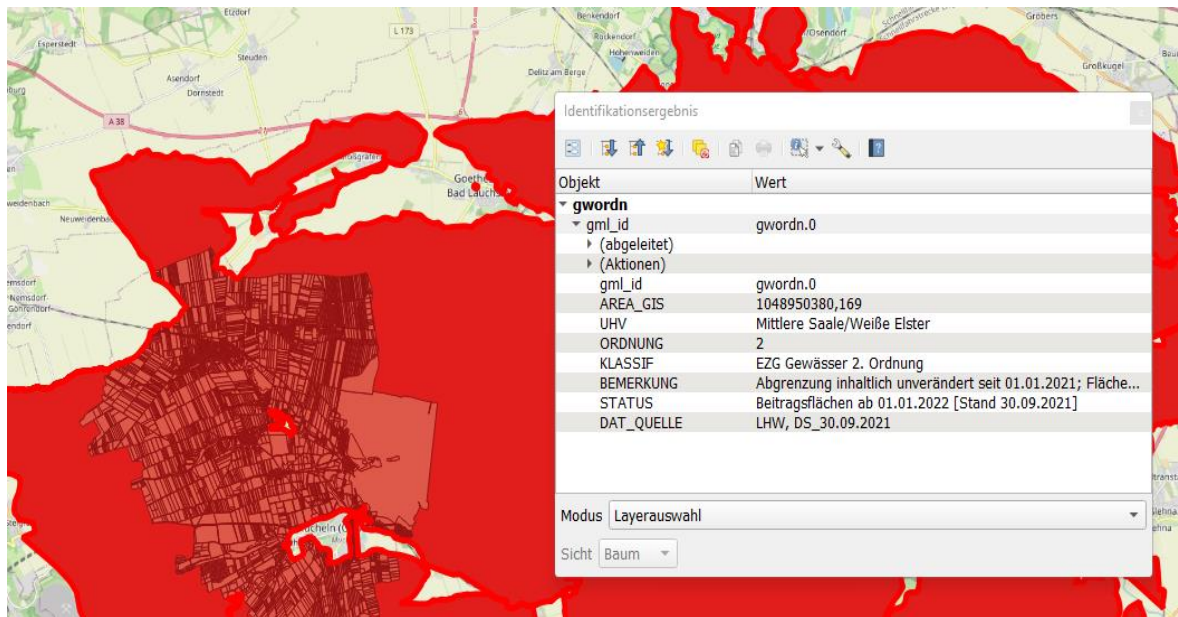
 <16:42:42> Suche mit dem Eingabeparameter Mühlstraße;3;* wurde gestartet...
 <16:42:44> Für die Suche wurde die folgende Adresse ermittelt:
 -Straße: Mühlstraße
 -Hausnummer: 3
 -Zusatz: *
 Für diese suche wurde ein Suchbereich mit 36397 Einträgen ermittelt.
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:44> GEBID: DEST_DESTLIKA0001FLF1; ADRESSE: Mühlstraße; 3
 <16:42:45> Es wurden 11 Ergebnisse gefunden.
 <16:42:45> Die Verarbeitung der Gewässerordnung wurde abgeschlossen.
 <16:42:45> Die Daten wurden in --- 2.3817379474639893 Sekunden --- übertragen.

Der Prozess wurde komplett durchlaufen!
 100% Suche abgeschlossen
 Abbrechen Log Reset Speichern

QGIS Plugin Funktion: Suche Gebäude zur Adresse oder Flurstück



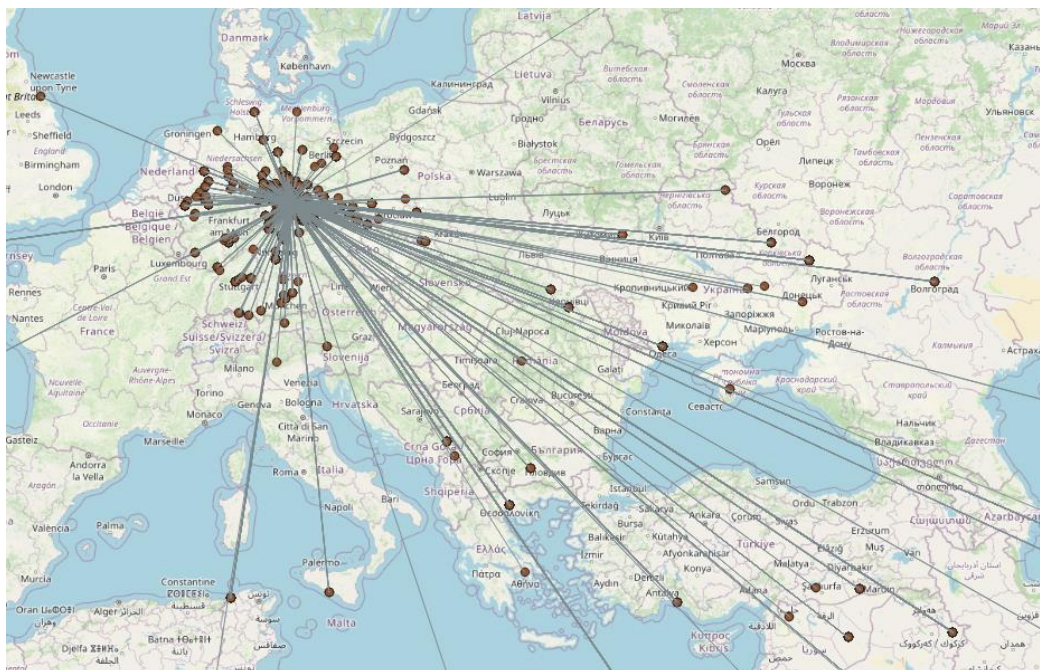
QGIS Visualisierung der Flurstücke farblich nach Besitzer unterteilt



QGIS Visualisierung der Gebiete mit Gewässer 2.Ordnung und enthaltene Flurstücke

UID	Bundesla	Gemarkur	Flur	Zaehler	Nenner	GISFLAEC	NEWAREA	VARIATIO	Klassif	UHV
5632	15	2300	11	450	0	61305,95	10692,8	17,45	EZG GewÄs	Mittlere Saal
5631	15	2300	11	394	0	25850,38	13586,34	52,56	EZG GewÄs	Mittlere Saal
5408	15	2300	11	119	1	18349,88	11246,4	61,29	EZG GewÄs	Mittlere Saal
5228	15	2300	11	660	0	32310,72	20280,77	62,77	EZG GewÄs	Mittlere Saal
4723	15	2300	11	152	38	24760,93	17394,05	70,25	EZG GewÄs	Mittlere Saal
5506	15	2300	11	33	1	13716,16	13693,69	99,84	EZG GewÄs	Mittlere Saal
4651	15	2300	11	160	1	23440,96	23440,96	100	EZG GewÄs	Mittlere Saal
4688	15	2300	11	149	1	11563,22	11563,22	100	EZG GewÄs	Mittlere Saal
4698	15	2300	11	147	1	12296,5	12296,5	100	EZG GewÄs	Mittlere Saal
4713	15	2300	11	161	0	21627,18	21627,18	100	EZG GewÄs	Mittlere Saal

Ergebnisse zur Bestimmung des Anteils des Flurstücks zum Gebiet der Gewässer 2.Ordnung



QGIS Visualisierung der Zuwanderungsstatistik für die letzten 3 Jahre für die Stadt Muehlen

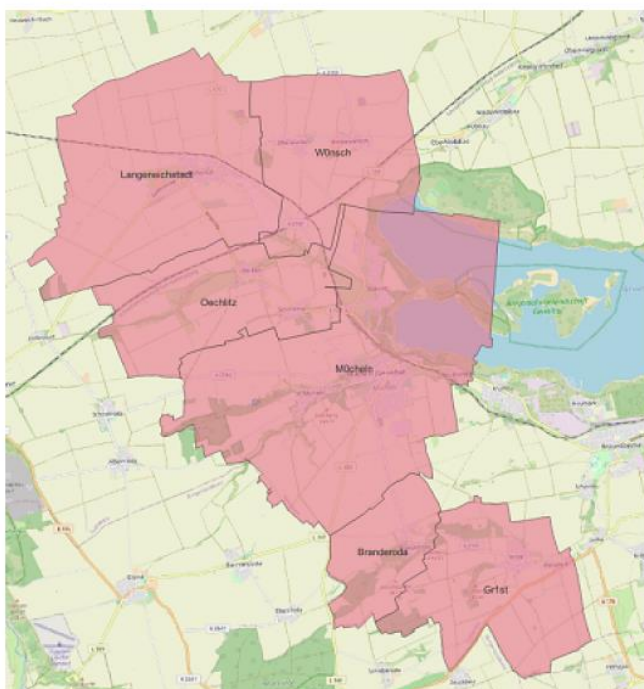
Anhang 4: Ergebnisse der Untersuchung für Mücheln



Stadt Mücheln
(Geseetal)

Auswertung des Leerstandsmanagements für Kindertagesstätten

Ergebnisliste aus dem Prozess zur Ermittlung von Gebäuden für Kindertagesstätten



Aufbau des Dokuments:

Dieses Dokument liefert die Ergebnisse des Algorithmus zur Ermittlung von passenden und freien Flächen bzw. Gebäuden für die Nutzung von Kindertagesstätten in der Stadt Mueheln. Die Ergebnisse werden für jeden Ortsteil der Stadt einzeln ausgegeben. Somit kann für jeden Ortsteil der Bedarf eingesehen und die Gebäude für Entscheidungsprozesse herangezogen werden. Die ermittelten Gebäude werden in Tabellenform aufgelistet und es werden Daten zu dem Gebäude, dem zugrundeliegenden Flurstück und zu den Ergebnisparametern der Auswertung aufgezeigt.

Seite 1/71

Ausschnitt des Ergebnisdokument zur Einleitung

Gebäude ID: DEST_DESTLIKA0001e14q	Gemarkung: 2282
Gebäudefunktion: 31001_2000	Flur: 4
Straße/Hausnr.: NULL	Zähler: 372
Dachtyp: PolyFlatRoof	Nenner: 0
Dachfläche: 90.0m2	Besitzer: privat
Dachform: 1000	Lage: Neue Dorfstraße 9
Grundfläche: 86.06m2	Grundfläche: 2173.0m2

Ortsteil: Mücheln (Geiseltal) OT Langeneichstädt

Ergebnisse der Analyse zur Alterstruktur

Bedarfsermittlung	Wachstum des OT
IST-Grundfläche: 401.96m2	Wachstumsrate: 0.2
Soll-Grundfläche: 390.0m2	

Für den Ortsteil Mücheln (Geiseltal) OT Langeneichstädt konnte Ein Bedarf an neuen Flächen für Kindergärten

Seite 25/71



Stadt Mücheln
(Geiseltal)

Auswertung des Leerstandsmanagements für Kindertagesstätten

ermittelt werden. Folgende Gebäude würden theoretisch zur Verfügung stehen.

Gebäudedaten (LoD2)	Flurstückdaten
Gebäude ID: DEST_DESTLIKA0001sBic	Gemarkung: 2345
Gebäudefunktion: 31001_2000	Flur: 17
Straße/Hausnr.: NULL	Zähler: 55
Dachtyp: PolyFlatRoof	Nenner: 2
Dachfläche: 90.0m2	Besitzer: privat
Dachform: 1000	Lage: Bahnhofsiedlung
Grundfläche: 172.76m2	Grundfläche: 2269.0m2

Ausschnitt des Ergebnisdokument des Algorithmus für die Stadt Mücheln für Langeneichstädt

Eidesstattliche Erklärung

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne unerlaubter Hilfe Dritter verfasst und keine anderen als die angegeben Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen stammen, sind als solche kenntlich gemacht. Diese Arbeit lag in gleicher oder ähnlicher Weise noch keiner Prüfungsbehörde vor und wurde bisher noch nicht veröffentlicht.