

Qualifizierter Mietspiegel 2024 für Fulda

Dokumentation der Mietspiegelerstellung

Herausgeber: Magistrat der Stadt Fulda
Bürgerservice
Statistik
Schlossstr. 1
36037 Fulda



Autor: EMA-Institut für empirische Marktanalysen
Im Gewerbepark C 25
93059 Regensburg



Datum: 29.04.2024
Version: 1.0
Redaktionelle Bearbeitung: EMA-Institut für empirische Marktanalysen

Aus Gründen der besseren Lesbarkeit wird bei Personenbezeichnungen und personenbezogenen Hauptwörtern in diesem Dokument die männliche Form verwendet. Entsprechende Begriffe gelten im Sinne der Gleichbehandlung grundsätzlich für alle Geschlechter. Die verkürzte Sprachform hat nur redaktionelle Gründe und beinhaltet keine Wertung.

Die Mietspiegelerstellung erfolgte im Auftrag der Stadt Fulda. Es ist nicht gestattet, ohne ausdrückliche Genehmigung der Städte die Daten der Dokumentation oder Teile daraus zu vervielfältigen und in elektronischen Systemen zu speichern und anzubieten.

Inhaltsverzeichnis

1	Auftrag und zeitliche Abfolge der Mietspiegelerstellung	5
2	Gesetzliche Grundlagen	7
2.1	Zugrundelegung eines einheitlichen Mietbegriffs	8
2.2	Datenschutz	10
3	Grundgesamtheit	12
3.1	Geltungsbereich	12
4	Stichprobenziehung	14
5	Datenerhebung	16
5.1	Zahlen zum Wohnungsmarkt und Rücklaufstatistik der Datenerhebung	17
5.2	Übertragung der Daten und Kontrollmaßnahmen	19
5.3	Datenselektion	20
5.4	Gewichtung	21
5.4.1	Gewicht 1	21
5.5	Einzeldatenanalyse und Datenumkodierung	22
6	Deskriptive Statistik	24
6.1	Eigenschaften der Merkmale Nettomiete und Wohnfläche	24
7	Regression	26
7.1	Der Gesamtansatz und das gewählte Modell	26
7.2	Die Grundstruktur des gewählten Regressionsmodells	27
7.3	Auswahl der Merkmale	28
7.4	Die methodische Vorgehensweise bei der Schätzung	31
7.5	Ermittlung des durchschnittlichen Mietniveaus	32
7.5.1	Ermittlung des Einflusses der Wohnfläche	32
7.5.2	Darstellung des durchschnittlichen Mietniveaus	35
7.5.3	Ermittlung des Einflusses des Baujahres	36
7.6	Ermittlung von Zu- und Abschlägen für weitere Wohnwertmerkmale	38
7.6.1	Ergebnis und Übersicht der Regressionsanalyse Phase 2	39
7.6.2	Beschreibung der Variablen im Regressionsmodell - Phase 2	41
7.7	Behandlung von außergesetzlichen Merkmalen	48
7.8	Behandlung von Ausreißern	48
7.9	Ermittlung von Spannbreiten	49

8	Schlussbemerkung	52
9	Literaturverzeichnis	53
10	Anhang	55
10.1	Tabellen und Grafiken	55
10.1.1	Tabellen und Grafiken des Regressionsmodels Phase 1	55
10.1.2	Tabellen und Grafiken des Regressionsmodels Phase 2	59
10.2	Fragebogen	72

1 Auftrag und zeitliche Abfolge der Mietspiegelerstellung

Am 01. Juni 2022 wurde das EMA-Institut für empirische Marktanalysen beauftragt, einen qualifizierten Mietspiegel für die Stadt Fulda zu erstellen.

Am 23. Februar 2023 fand ein erster vorbereitender interner Austausch zwischen Vertretern der Stadt Fulda sowie Vertretern des EMA-Instituts statt. Hierbei wurde der zeitliche und inhaltliche Rahmen der Mietspiegelerstellung konkretisiert.

Am 02. Mai 2023 tagten die Mitglieder des Arbeitskreises Mietspiegel, bestehend aus Vertretern der Stadtverwaltungen, Interessensvertretern der Vermieter und Mieter sowie weiteren Vertretern mit Kenntnissen des lokalen Wohnungsmarktes und des EMA-Instituts. Die Beteiligung der Interessenvertreter von Mietern und Vermietern im Arbeitskreis Mietspiegel erhöht die Akzeptanz des Mietspiegels. Zudem verfügen die Vertreter über Kenntnisse des lokalen Wohnungsmarktes, die es bei Fragebogenerstellung und Auswertung der erhobenen Daten zu beachten gilt. In der ersten Sitzung wurde die grundsätzliche Konzeption und Vorgehensweise zur Mietspiegelerstellung festgelegt. Diese umfasste vornehmlich die Festlegung der Art der Datenerhebung sowie einen ersten, richtungsweisenden Fragebogenentwurf für die Datenerhebung. Für den Fragebogenentwurf hat der Arbeitskreis die einzelnen Wohnwertmerkmale gemeinsam erarbeitet und abgestimmt. Als Art der Datenerhebung wurde einvernehmlich die schriftliche Befragung per Brief in Kombination mit einer Antwortoption über das Internet gewählt.

Die Datenerhebung wurde im Zeitraum von Juni 2023 bis Juli 2023 bei Mietern und Vermietern durchgeführt. Die Briefsendungen wurden am 12. Juni 2023 an die Haushalte verschickt. Stichtag für die Erhebung der Mieten war der 1. Juli 2023. Am 12. Juli 2023 wurde ein Erinnerungsschreiben an die Befragten, welche sich noch nicht bei der ersten Befragungswelle zurückgemeldet hatten, verschickt. Die Datenerhebung des Erinnerungsschreibens wurde im Juli 2023 bei den Mietern durchgeführt. Während und kurz nach der Erhebungsphase fand die Erfassung bzw. Digitalisierung der Papierantworten auf elektronische Datenträger statt. Es wurden gleichzeitig Kontrollmaßnahmen hinsichtlich inhaltlicher Plausibilitäten für alle digitalisierten und online erfassten Datensätze durchgeführt. Bis Anfang August 2023 erfolgte die Plausibilisierung und Zusammenführung der Antwortdatensätze. Anschließend erfolgte die Auswertung der Daten. Am 21. September 2023 wurden die ersten Mietspiegelergebnisse der Stadtverwaltung übermittelt und das modifizierte Ergebnis am 5. September 2023 dem Arbeitskreis Mietspiegel präsentiert. Nach kleineren inhaltlichen Anpassungen konnten die Ergebnisse sowie das Mietspiegeldokument am 5. Oktober dem Arbeitskreis präsentiert werden. Der Mietspiegel wurde von dem Verband der Haus-, Wohnungs- und Grundeigentümer Fulda und Umgebung e.V. und dem Mieterverein Fulda und Umgebung e.V. in der Sitzung vom 5. Oktober als qualifizierter Mietspiegel anerkannt.

Der Mietspiegel der Stadt Fulda besitzt seine Gültigkeit als qualifizierter Mietspiegel vom 01.01.2024 an für den Zeitraum von zwei Jahren, d.h. bis zum 31.12.2025.

2 Gesetzliche Grundlagen

Ein Mietspiegel ist gemäß Mietspiegelreformgesetz (Bundesregierung 2021a) i. V. m. der Mietspiegelverordnung (Bundesregierung 2021b), sowie §§ 558c und 558d Bürgerliches Gesetzbuch (BGB) eine Übersicht über die gezahlten Mieten für nicht preisgebundenen Wohnraum vergleichbarer Art, Größe, Ausstattung, Beschaffenheit und Lage einschließlich der energetischen Ausstattung und Beschaffenheit. Die ortsübliche Vergleichsmiete setzt sich aus Mieten zusammen, die in den letzten sechs Jahren neu vereinbart oder, von Betriebskostenerhöhungen abgesehen, geändert worden sind („6-Jahres-Frist“). Es wird seit der Mietrechtsreform 2001 zwischen *qualifizierten* und *einfachen* Mietspiegeln unterschieden. An den qualifizierten Mietspiegel werden deutlich höhere Anforderungen gestellt als an den einfachen Mietspiegel und gleichzeitig auch weitreichendere Folgen geknüpft. Qualifiziert ist ein Mietspiegel gemäß § 558d Abs. 1 und 2 BGB dann, wenn er

1. nach anerkannten wissenschaftlichen Grundsätzen erstellt,
2. nach zwei Jahren per Stichprobe oder Preisindex fortgeschrieben bzw. nach vier Jahren neu erstellt wird und
3. von der nach Landesrecht zuständigen Behörde oder den Interessenvertretern von Mietern und Vermietern als qualifiziert anerkannt wird.

Ein qualifizierter Mietspiegel impliziert die Vermutungswirkung, dass er die ortsüblichen Vergleichsmieten richtig wiedergibt (§ 558d Abs. 3). Zudem muss ein Vermieter bei Mieterhöhungsverlangen auf einen qualifizierten Mietspiegel Bezug nehmen, sobald der Mietspiegel Werte für die entsprechende Wohnung enthält, auch wenn sich der Vermieter auf ein anderes Begründungsmittel nach § 558a Abs. 2 BGB beruft.

Mit der Einführung des Instruments „qualifizierter Mietspiegel“ hat der Gesetzgeber die Bedeutung eines Mietspiegels hervorgehoben und dessen Qualitätscharakter gegenüber Sachverständigengutachten und insbesondere der Heranziehung von drei Vergleichsmieten bei der Begründung eines Mieterhöhungsverlangens betont. Der Gesetzgeber stellt fest, dass qualifizierte Mietspiegel zur Ermittlung der ortsüblichen Vergleichsmiete das beste und verlässlichste Instrument sind (Begründung zum Kabinettsbeschluss, Abschnitt II. 2a)¹.

Bei der Erstellung des neuen Mietspiegels der Stadt Fulda wurden von Seiten der Mietspiegelersteller die Voraussetzungen für die Anerkennung zum qualifizierten Mietspiegel geschaffen, indem die Erstellung des Mietspiegels nach anerkannten wissenschaftlichen Methoden bei der Datenerhebung und der Datenauswertung erfolgt ist. Die Datenerhebung basierte auf einer Zufallsauswahl von Wohnungen, wodurch aus stichprobentheoretischer Sicht ein repräsentatives Abbild des Woh-

¹ Vgl. (Rips und Eisenschmid 2001, 372)

nungsmarktes gesichert wird. Aus dieser Auswahl wurden nur die gesetzlich vorgeschriebenen mietspiegelrelevanten Wohnungen berücksichtigt, sodass auch ein repräsentatives Abbild dieser Wohnungen für den mietspiegelrelevanten Mietwohnungsmarkt in Fulda zugrunde liegt. Die Daten wurden über eine schriftliche Befragung in Kombination mit einer Onlineantwortmöglichkeit abgefragt. Als Auswertungsmethodik wurde die Regressionsmethode verwendet, die in der oben erwähnten Begründung zum Kabinettsbeschluss als ein von der Wissenschaft anerkanntes statistisches Auswertungsverfahren genannt ist². Die einzelnen Arbeitsschritte im Rahmen der Mietspiegelerstellung sind in dieser Dokumentation detailliert festgehalten.

Durch Anerkennung des Mietspiegels entweder durch die nach Landesrecht zuständige Behörde oder durch die Interessenvertreter der Vermieter und der Mieter wird einem Mietspiegel der Status „qualifiziert“ zugewiesen. Haben die nach Landesrecht zuständige Behörde und Interessenvertreter der Vermieter und Mieter den Mietspiegel als qualifizierten Mietspiegel anerkannt, so wird vermutet, dass der Mietspiegel anerkannten wissenschaftlichen Grundsätzen entspricht (§558 d Abs 1 Satz 3). Qualifizierte Mietspiegel können gemäß nach zwei Jahren mittels Verbraucherpreisindex oder einer Stichprobe fortgeschrieben werden (§ 558d Abs. 2 Satz 3 BGB), wodurch der Status der Qualifizierung um weitere zwei Jahre verlängert wird. Vier Jahre nach einer erfolgten Mietspiegelneuerstellung muss ein Mietspiegel neu erstellt werden, um die Ausweisung von aktuellen ortsüblichen Vergleichsmieten im Mietspiegel zu gewährleisten und den Status „qualifiziert“ zu erhalten.

2.1 Zugrundelegung eines einheitlichen Mietbegriffs

Bei der Aufstellung eines Mietspiegels ist es notwendig, einen einheitlichen Mietbegriff zugrunde zu legen, um eine Vergleichbarkeit der Mieten zu gewährleisten. In Mietverträgen werden aber verschiedene Mietbegriffe verwendet, die sich aus der unterschiedlichen Handhabung der Nebenkosten ergeben:

1. Die **Nettomiete**, auch Nettokalt- oder Grundmiete genannt, stellt den eigentlichen Preis für die Überlassung einer Wohnung dar und enthält keine Betriebskosten gemäß Anlage 3 zu § 27 II. BV bzw. gemäß § 2 Betriebskostenverordnung, die seit dem 1. Januar 2004 gilt.
2. Die **Bruttomiete**, auch als Bruttowarm-, Pauschal- oder Inklusivmiete bezeichnet, enthält neben dem Preis für die Wohnungsüberlassung sämtliche Betriebskosten, einschließlich der Heiz- und Warmwasserkosten.

² Vgl. (Rips und Eisenschmid 2001, 422)

3. Die **Bruttokaltmiete** unterscheidet sich von der Bruttomiete dadurch, dass die Heiz- und Warmwasserkosten nicht im Mietzins enthalten sind, wohl aber die übrigen Betriebskosten.
4. Eine **Teilinklusiv- oder Teilpauschalmiete** liegt vor, wenn ein Teil der Nebenkosten wie etwa einzelne Betriebskosten, Küchen-, Stellplatz-/Garagenmieten, Zuschläge für Möblierung oder Untervermietung, Anteile für Schönheitsreparaturen im Mietpreis enthalten ist, ein anderer Teil jedoch getrennt abgerechnet wird. Die Bruttokaltmiete stellt einen Spezialfall der Teilinklusivmiete dar.

Von den in § 2 Betriebskostenverordnung aufgezählten **Betriebskosten** können die umlagefähigen Kosten für den Mieter als Nebenkosten in Frage kommen. Je nach Art der Abrechnung werden Betriebskosten als umgelegt (einzeln abgerechnet) oder nicht umgelegt (undifferenziert in der Vertragsmiete enthalten) bezeichnet. Die folgende Abbildung veranschaulicht die einzelnen Mietzinsbegriffe und ihre Zusammenhänge.

Abbildung 1: Mietzinsbegriffe und ihre Bestandteile

Nettomietzins	Nicht umgelegte Betriebskosten	umgelegte Betriebskosten		Zuschlagszahlungen (z.B. für Garage)
		Allgemeine Betriebskosten	Heiz- und Warmwasserkosten	
Nettomiete				
	Teilinklusivmiete			
		Bruttokaltmiete		
			Bruttomiete	
				Bruttomiete inkl. Zuschläge

In Mietspiegeln werden üblicherweise durchschnittliche Nettomietbeträge als ortsübliche Entgelte ausgewiesen. Dies erweist sich auch in diesem Fall aus mehreren Gründen als sinnvoll: Zum einen bildet die Nettomiete die Ausgangsbasis, um durch das Hinzuaddieren von Betriebskosten die individuelle Vertragsmiete berechnen zu können. Andererseits ist den meisten Miethaushalten, nämlich ca. 95 Prozent, die Höhe ihrer Nettomiete laut Mietspiegelerhebung bekannt.

Um bei der Auswertung aber nicht auf Fragebögen, in denen nur die monatliche Mietzahlung angegeben war, verzichten zu müssen, war eine Rückführung der monatlichen Mietzahlung auf die monatliche Nettomiete erforderlich. Dazu wurden im Fragebogen die folgenden mit dem Mietpreis zusammenhängenden Größen erfragt:

- Die gesamte monatliche Mietzahlung (einschließlich Nebenkosten und Mietanteilen)
- Die monatliche Nettomiete (ohne Nebenkosten und Mietanteile)

- Die Höhe des Betriebskostenabschlags
- Mietanteile
- Mietermäßigungen

Eine der beiden erstgenannten Positionen wurde immer beantwortet. Falls die Angabe der Nettomiete verfügbar war, so fand diese Verwendung. Sofern nur die gesamte monatliche Mietzahlung vorlag, musste mit Hilfe von Zusatzangaben auf die entsprechende Nettomiete umgerechnet werden. Für alle Fälle, in denen sowohl die gesamte monatliche Mietzahlung als auch die Nettomiete vorlagen, konnten die Differenzen berechnet werden. Im Rahmen einer Regressionsanalyse wurde diese Differenz mit Hilfe einer Vielzahl von erklärenden Merkmalen (Wohnfläche, Baujahr, Ausstattungsmerkmale, Aufzug, Heizungsart, Höhe der Nebenkosten usw.) beschrieben. Damit konnte dann auch in den Fällen mit fehlenden Angaben über die Nettomiete die geschätzte Differenz ermittelt und damit anhand von Durchschnittsangaben bei den Betriebskosten auf die unbekannte Nettomiete umgerechnet werden (siehe Abschnitt 5.3).

2.2 Datenschutz

Die gesamte Vorgehensweise war mit der Auftraggeberin über einen Auftragsverarbeitungsvertrag (Art. 28 Abs. 3 Datenschutz-Grundverordnung (DSGVO)) geregelt.

Durch Trennung der Adressangaben von den sonstigen Angaben zum Mietverhältnis wurde eine Anonymisierung aller Daten bei der Auswertung gewährleistet.

Das EMA-Institut erhielt von der Stadt Fulda die benötigten Meldedaten. Die Übertragung der Meldedaten erfolgte via verschlüsseltem Cloudlink und einem zugehörigen Passwort, welches telefonisch übergeben wurde. Nach der in Abschnitt 4 beschriebenen Bereinigung der Adressen wurde aus den Meldedaten eine Stichprobe per Zufallsauswahl gezogen.

Sodann erfolgte eine Pseudonymisierung der personenbezogenen Daten. Jedem Datensatz wurde eine Identifikationsnummer zugewiesen. Diese Identifikationsnummer diente nach Erhalt des ausgefüllten Fragebogens dazu, dass die darin enthaltenen Informationen ab dem Zeitpunkt der EDV-Erfassung einer Nummer zugeordnet werden konnten. Zentraler Punkt für die Gewährleistung der Pseudonymität der abgefragten Daten war, dass auf den Fragebögen keine personenbezogenen Daten waren, sondern nur die jeweilige Identifikationsnummer. Ab diesem Zeitpunkt war zwischen Identifikationsnummer und personenbezogenen Daten der Befragten keinerlei Beziehung mehr gegeben. Nach der Übertragung des Fragebogeninhalts auf elektronische Datenträger waren die Datensätze bei der Auswertung nur noch mit ihrer Identifikationsnummer, ohne Name und ohne Adresse,

enthalten. Alle weiteren Analysen fanden ausschließlich mit diesen nicht mehr personenbezogenen Daten statt. Die verbliebenen Datensätze, in denen die sogenannten Erhebungs- und Hilfsmerkmale enthalten waren, wurden bis Abschluss des Projekts gesondert aufbewahrt und danach unwiderruflich gelöscht.

3 Grundgesamtheit

Die **Grundgesamtheit** für die Mietspiegelerhebung bildet der vergleichsmietenrelevante Mietwohnungsmarkt (Cischinsky et al. 2014). Ein Mietspiegel soll die tatsächlichen Mieten von mietspiegelrelevanten Wohnungen als Teilmenge der Grundgesamtheit wiedergeben. Alle mietspiegelrelevanten Wohnungen und deren Mietpreise zu erfassen ist sowohl im Hinblick auf ein angemessenes Kosten-Nutzen-Verhältnis der Mietspiegelerstellung als auch vor dem Hintergrund der schier unerschöpflichen Anzahl der anzuschreibenden Haushalte kaum umsetzbar. Deshalb wird aus der Grundgesamtheit aller Wohnungen eine Stichprobe gezogen, deren Mieter bzw. Vermieter Angaben zu mietspiegelrelevanten Fragen machen sollen.

3.1 Geltungsbereich

Der qualifizierte Mietspiegel für die Stadt Fulda **gilt ausschließlich** für Mietwohnungen und vermietete Häuser auf dem nicht preisgebundenen Wohnungsmarkt im Wohnflächenbereich zwischen 25 m² und 150 m² im jeweiligen räumlichen Geltungsbereich.

Zur Mietspiegelrelevanz von Wohnungen werden im Bürgerlichen Gesetzbuch (BGB) nähere Ausführungen gemacht. Ein Mietspiegel gilt demnach *nicht* für:

- Wohnungen, bei denen es sich um selbstgenutztes Eigentum handelt
- Wohnungen, die Teil eines Wohnheimes oder einer Sammelunterkunft sind (z.B. Studenten-, Alten-, Pflege-, Personalwohnheim, vorläufige Unterbringung (Geflüchtete), Behinderteneinrichtung, „Betreutes Wohnen“)
- Wohnungen, für die eine gewerbliche Nutzung oder nur ein kurzfristiger Gebrauch vorgesehen ist (z.B. Ferienwohnung, maximal drei Monate pro Mieter, Hinweis: Das private Arbeitszimmer oder der private Home-Office-Bereich ist damit nicht gemeint!)
- Wohnungen, bei denen es sich um eine Werkswohnung handelt (Arbeits- und Mietverhältnis stehen miteinander in Verbindung)

Diese Mietverhältnisse werden *per Gesetz* von der Grundgesamtheit ausgeschlossen, da sie nicht als typische Mietverhältnisse gelten.

Aufgrund von **Plausibilitäts- und erhebungstechnischen Überlegungen** wurden einvernehmlich einzelne Sonderfälle von Wohnungen bzw. Wohnverhältnissen zusätzlich ausgeklammert. Dabei handelt es sich um:

- Wohnraum, der mietfrei oder verbilligt überlassen wird (z.B. Dienst- oder Werkswohnung, Wohnung gehört Verwandten)
- Wohnungen, die ganz- oder teilmöbliert vermietet werden (einzelne Möbelstücke sowie Einbauküche und Einbauschränke zählen nicht als Möblierung)

4 Stichprobenziehung

Zur Ermittlung der Grundgesamtheit muss auf zusätzliche Daten zurückgegriffen werden (Kauermann et al. 2020). Dabei wird die für den Zweck am besten geeignete und aktuelle Datengrundlage gewählt. Welche Daten das sind, kann von Gemeinde zu Gemeinde unterschiedlich sein. Es ist üblich Daten aus Einwohnermelderegistern, Grundsteuer- oder Stromzählerdateien zu verwenden. Bei der Stromzählersitzdatei kann davon ausgegangen werden, dass jede Wohnung die gleiche Wahrscheinlichkeit besitzt in die Stichprobe zu gelangen. Bei den Einwohnermelderegister- oder Grundsteuerdaten, bei welchen nicht Wohnungen, sondern Haushalte als zentrale Untersuchungsvariable ausweisen, ist dies nicht unmittelbar der Fall ist (Cischinsky et al. 2014), (Kauermann et al. 2020). Leider ist die Stromzählerdatei nicht in jeder Stadt zentral verfügbar. Daher werden in ca. 95% der Fälle die Einwohnermeldedaten als Auswahlrahmen für eine Zufallsstichprobe verwendet. Eine grobe Bereinigung der Eigentümer, welche ihren Wohnraum selbst nutzen, kann anhand der Grundsteuerdatei erfolgen.

Die Stadt Fulda übermittelte dem EMA-Institut für die Generierung einer Haushaltsdatei Abzüge der Einwohnermelde- und Grundsteuerdaten gemäß Artikel 2, Mietspiegelreformgesetz (MsRG). Die Übertragung erfolgte via verschlüsseltem Cloudlink und einem zugehörigen Passwort, welches telefonisch übergeben wurde.

Die Einwohnermeldedatei wurden durch die Verwaltung bereits vorab um nicht volljährige Personen bereinigt. Ebenfalls entfernt wurden bekanntermaßen geförderte Wohnungen oder Wohnraum in Heimen und Anstalten. Anschließend wurde eine Haushaltsgenerierung durchgeführt. Diese erfolgte auf Basis der Attribute Nachname, Straße, Hausnummer und Zusatz. Nach dieser Sortierung wurden alle mit diesem Attributenschlüssel vorhandenen Duplikate gelöscht. Somit war jeder Haushalt nur noch maximal einmal in der Datenbasis enthalten. Bei der späteren Stichprobenziehung wird somit jedem Stichprobenelement die gleiche Wahrscheinlichkeit, in die Zufallsstichprobe zu gelangen, zugeordnet. Auf eine Gewichtung der Stichprobe aufgrund von unterschiedlichen Ziehungswahrscheinlichkeiten pro Haushalt kann somit verzichtet werden. Das hier beschriebene Verfahren ähnelt dem Programm HHGen des KOSIS-Verbunds (www.staedtestatistik.de/arbeitsgemeinschaften/hhstat/hhgen). Wesentlicher Unterschied zu diesem Programm ist, dass volljährige Personen innerhalb eines Haushalts nicht als separate Haushalte generiert werden. Ein weiterer Vorteil ist wiederum, dass aus dieser generierten Haushalts-Liste einfache oder geschichtete Stichproben gezogen werden können. Ein Nachteil bei beiden Vorgehensweisen ist, dass Lebensgemeinschaften mit unterschiedlichen Nachnamen stets als getrennte Haushalte aufgefasst werden. D. h. durch diese theoretische Haushaltsgenerierung lassen sich aber nicht alle real existierenden Haushalte erzeugen (Kauermann et al. 2020).

In dem Auswahlrahmen der Einwohnermeldedaten befinden sich zudem auch die Informationen über die Wohnungsgeber, welche wiederum die Vermieterinformation als Teilmenge beinhaltet. Diese Datengrundlage kann bei ausreichend guter Datenqualität als weitere Stichprobengrundlage für die Auswahl von Vermieteradressen herangezogen werden. Fehlende Informationen können gegebenenfalls durch die Grundsteuerdatei ergänzt werden, sodass für 90% der gezogenen Elemente in der Stichprobe eine konsistente Adresse des Wohnungsgebers bzw. des Vermieters gefunden werden konnte. In Fulda ist dies als Spezialfall zu betrachten, da die dort vorliegenden Daten eine überdurchschnittliche Qualität im Vergleich zu anderen Städten aufwiesen.

Mit dieser bereinigten Haushaltsdatei wurde folgender Stichprobenplan umgesetzt:

Tabelle 1: Bruttostichprobe

Kommune	Bruttostichprobe
Fulda gesamt	4.000
Mieter	3.000
Vermieter/Wohnungsgeber	1.000

5 Datenerhebung

Zum Zwecke der Datenerhebung wurden folgende, gemäß des unterzeichneten Auftragsverarbeitungsvertrags (Art. 28 Abs. 3 DSGVO), die gemäß Art. 2 MsRG, Art. 238 EGBGB § 1 nach zugelassenen personenbezogenen Daten erhoben: Vorname/Rufname, Nachname, Straße, Hausnummer, Zusatz, Postleitzahl, Ort und Ortsteil. Die Adressdaten wurden vom Auftraggeber an den Auftragsverarbeiter im Sinne der DSGVO übermittelt und verarbeitet.

Anlaufadresse für die Mietspiegelerhebung waren private Mieterhaushalte. Die zufällig ausgewählten Haushalte in der Stichprobe wurden mit einem Anschreiben und einem Informationsblatt gemäß Artikel 13, 14 der Datenschutz-Grundverordnung und einem darin enthaltenen Fragebogen vom EMA-Institut im Namen der Stadt Heideberg angeschrieben.

Die zufällig ausgewählten Einwohner und Einwohnerinnen konnten den ausgefüllten Fragebogen mit einem beigefügten Rückantwortkuvert kostenlos an das EMA-Institut zurücksenden. Zudem bestand die Möglichkeit den Fragebogen online auszufüllen. Der Vorteil einer schriftlichen Erhebung gegenüber der mündlichen Befragung ist die hohe Flexibilität bezüglich des Ausfüllzeitraums. Denn der Bürger kann sich jederzeit mit dem Fragebogen in Papierform beschäftigen. Ein Nachteil ist, dass bei der Datenbereinigung und -aufbereitung eine höhere Anzahl an Datensätzen mit fehlenden oder nicht verwertbaren Angaben aussortiert werden muss, weshalb bei der Stichprobenziehung ein höherer Bruttostichprobenumfang zu berücksichtigen ist.

Der Datenerhebung lag ein Konzept mit zweiteiligem Aufbau zugrunde. Im ersten Teil wurde zunächst die Mietspiegelrelevanz der angeschriebenen Wohnung überprüft. Folgende Filterfragen wurden gestellt:

- Handelt es sich bei der Wohnung um selbstgenutztes Eigentum oder gehören Sie zum Haushalt des Eigentümers?
- Wird der Wohnraum mietfrei oder vergünstigt überlassen (z.B. wegen Verwandtschaftsverhältnis)?
- Ist die Wohnung Teil eines Wohnheimes oder einer Sammelunterkunft (z.B. Studenten-, Alten-, Pflege-, Personalwohnheim, vorläufige Unterbringung (Geflüchtete), Behinderteneinrichtung, „Betreutes Wohnen“,)?
- Wird die Wohnung ganz- oder teilmöbliert vermietet? (Einbauküche und/oder Einbauschränke zählen nicht als Möblierung)
- Ist für die Wohnung eine gewerbliche Nutzung oder nur ein kurzfristiger Gebrauch vorgesehen (z.B. Ferienwohnung, maximal drei Monate pro Mieter, Hinweis: Das private Arbeitszimmer oder der private Home-Office-Bereich ist damit nicht gemeint!)?

- Handelt es sich bei der Wohnung um eine Werkswohnung? (Arbeits- und Mietverhältnis stehen miteinander in Verbindung)

Die Zustimmung zu einer der vorangegangenen Filterfragen führte zum Ausschluss der Wohnung aus der Mietspiegelauswertung. Nur falls sich eine Wohnung als mietspiegelrelevant erwies, kam der Hauptfragebogen zum Einsatz. Der Hauptfragebogen (siehe Anlagen) enthielt Fragen zu folgenden Schwerpunkten:

- Mietpreis, Betriebskosten, Mietverhältnis
- Art und Alter des Gebäudes und der Wohnung
- Größe, Ausstattung und sonstige Beschaffenheit der Wohnung
- Nachträgliche Verbesserungen der Wohnungsqualität und energetischer Zustand

Für Rückfragen während der Datenerhebungsphase standen die Projektleitung des EMA-Instituts und die Stadtverwaltung allen Beteiligten telefonisch oder per E-Mail zur Verfügung.

5.1 Zahlen zum Wohnungsmarkt und Rücklaufstatistik der Datenerhebung

Tabelle 2: Wohnungskennzahlen lt. Zensus 2011

Kommune	Einwohner	Wohnungen gesamt	Von Eigentümern bewohnt	Zu Wohnzwecken vermietet (auch mietfrei) ³	Quote Eigentum/Miete
Fulda	64.414	31.787	10.543	19.950	31%/59%

³ Quelle: Zensusdatenbank 2011: <https://ergebnisse.zensus2011.de>

Die Rücklaufstatistik ist in nachfolgender Tabelle zu entnehmen.

Tabelle 3: Rücklaufstatistik der angeschriebenen Haushalte

Kommune	# Befragungen	Rücklauf ge- samt	Quote Stichpro- genausschöp- fung gesamt	Rücklauf abzü- gich aktiver Filter- fragen	Quote Stichpro- benausschöpfung abzüglich aktiver Filterfragen	Rücklauf abzü- gich aktiver Filterfragen und "6-Jahres- Frist"	Quote Stich- probenauss- schöpfung ab- züglich aktiver Filterfragen und "6-Jahres- Frist"
Fulda: Mieter	3.000	1.665	56%	1.173	39%	934	31%
Fulda: Vermieter	1.000	522	52%	412	41%	360	36%
Gesamt	4.000	2.187	55%	1.585	40%	1.294	32%

Der oben beschriebene verwertbare Rücklauf innerhalb der gesetzlichen „6-Jahres-Frist“ ergibt sich durch folgende Datenreduktionen:

Tabelle 4: Ausschuss nicht relevanter Daten

Rücklauf	nicht relevante Datensätze	verbleibende Datensätze
Rücklauf gesamt		2.187
Rücklauf abzgl. Filterfragen	602	1.585
Rücklauf abzgl. „6-Jahres-Frist“, sozial geförderter Wohnungen und Wohngemeinschaften	291	1.301
Rücklauf abzgl. „Ausreißer“	7	1.294

Die Gesamtanzahl der Antworten lag bei 2.187 Datensätzen, welche sowohl per Post (905 Datensätze) als auch über das Onlineportal (1.282 Datensätze) gesammelt wurden. Diese Antworten wurden um die in Tabelle 4 dargestellten Datensätze bereinigt. Es verblieben 1.294 Datensätze, mit denen die Auswertung fortgesetzt werden konnte.

5.2 Übertragung der Daten und Kontrollmaßnahmen

Nach Eingang der Fragebögen beim EMA-Institut wurde deren Inhalt datenschutzkonform auf elektronische Datenträger übertragen. Dies erfolgte mit Hilfe eines Hochleistungsscanners sowie der Datenerfassungssoftware „AbbyyFlexiCapture“, welche die weltweit führende Software zur Datenerfassung darstellt. Daraufhin lief innerhalb der Software eine eigens für den Mietspiegel der Stadt Fulda programmierte Plausibilitätsprüfung über die digitalisierten Daten. Dabei wurde geprüft, ob Daten vom Anwender falsch angegeben oder vom Scanner falsch ausgelesen wurden. Hierbei wurde insbesondere jede zahlenmäßige Angabe zunächst vom Programm geprüft und anschließend noch zweimal von einer eigens für diese Software geschulten Person manuell überprüft. Ein sehr einfaches Beispiel für einen derartigen Algorithmus sind sogenannte Prüfsummen, welche z.B. die Angaben der Bruttomiete, der Nettomiete sowie der Nebenkosten entsprechend der Logik, dass die Nettomiete niedriger sein muss als die Bruttomiete, oder die Bruttomiete die Summe aus Nettomiete und Nebenkosten bilden kann, untersucht und einen Fehler bei Verletzung dieser Logik berichtet.

Datensätze, in denen im Falle eines Fehlers keine Klärung herbeigeführt werden konnte, wurden von den weiteren Analysen ausgeschlossen.

Die statistische Analyse beinhaltete weitere Kontrollschritte hinsichtlich der thematischen Logik und Plausibilität. Ein Beispiel soll dies verdeutlichen: Wurde sowohl die Frage „kein Badezimmer in der Wohnung vorhanden?“ als auch die Frage

„Fußbodenheizung im Bad vorhanden?“ angekreuzt, lag eine nicht nachvollziehbare inhaltliche Plausibilität vor. Im Falle solcher logischer Widersprüche wurden Korrekturen vorgenommen, wenn diese aus anderen Angaben des Fragebogens abgeleitet werden konnten. Wenn die logischen Widersprüche nicht mit weiteren Angaben aus dem Fragebogen aufgelöst werden konnten, wurde der Fragebogen aufgrund fehlender Plausibilisierungsmöglichkeiten nicht weiterverwendet.

Für die Schätzung der ortsüblichen Vergleichsmiete und der Erstellung des zugehörigen Regressionsmodells wurde die aktuellste Version der Python-Module „statsmodels“ (Seabold und Perktold 2010) und „sklearn“ (Pedregosa et al. 2011) in Verbindung mit der aktuellsten Version der Entwicklungsumgebung „Spyder“ (Raybaut 2009) verwendet. Alle Programme sind frei zugänglich und werden zusammen in der Software Distribution „Anaconda“ zur Verfügung gestellt (Anaconda Software Distribution 2020).

5.3 Datenselektion

Bei einigen mietspiegelrelevanten Datensätzen fehlten teilweise Angaben zu Wohnwertmerkmalen. Die verwendeten statistischen Auswertungsverfahren erlauben die Berücksichtigung von Erhebungseinheiten mit teilweise fehlenden Werten, sodass zumindest die vorhandene Information genutzt werden kann. Für die Auswertung wesentlich fehlender Angaben wurden, dort wo es möglich war, Imputationsverfahren angewandt und die Datensätze somit vervollständigt. Imputationsverfahren können die Schätzungen der Zielvariable, hier die Nettomiete pro Quadratmeter, erheblich stabilisieren und verbessern (van Buuren 2019), (Little 2012), (Allison 2007). Auf die betreffenden Datensätze musste also nicht verzichtet werden, solange für die zu imputierende Variable die nötigen Informationen vorhanden waren. Diese Vorgehensweise lässt sich am einfachsten veranschaulichen, indem man die Abfrage des Baujahres als Beispiel verwendet. War bei der Datenerhebung das Baujahr nicht exakt bekannt, so konnte es in eine Baujahresklasse eingeordnet werden. Jede Baujahresklasse hat einen konkreten Mittelwert, welcher aus den Datensätzen errechnet werden kann. Dort wo nun das Baujahr als exakte Angabe fehlte, konnte dann z. B. der Mittelwert dieser Baujahresklasse als konkretes Baujahr („mean imputation“) oder als Regressions-Imputation („regression imputation“) imputiert werden. D. h. man schätzt das Baujahr einer Wohnung mit Hilfe der Regressionsanalyse anhand weiterer im Fragebogen angegebener Ausstattungsmerkmale und vergleicht das Ergebnis mit der zur vor beobachteten Klassenmitte. Wurde durch diese Vorgehensweise kein plausibles Ergebnis für einen Datensatz erzielt, wurde der Datensatz nicht weiterverwendet. So verblieb es bei der Reduzierung um die in Abschnitt 5.1 dargestellten Datensätze.

5.4 Gewichtung

5.4.1 Gewicht 1

In der in Abschnitt 4 vorgestellten Art der Stichprobenziehung gab es Überschneidungen der Stichprobenelemente. Konkret bedeutet das, es gab Wohnungen, zu welchen sowohl von Mieter- als auch institutioneller Vermieterseite eine Antwort vorlag. Die Stichprobenelemente der Mieter werden mit „Schicht A“ und die der institutionellen Vermieter mit „Schicht B“ bezeichnet. Diese Wohnungen tauchten demnach mit doppelter Häufigkeit in der Stichprobe auf, was zu einer Verzerrung der Schätzergebnisse führen kann (Horvitz und Thompson 1952). Um die auf Basis der unterschiedlichen Ziehungswahrscheinlichkeiten herbeigeführte Verzerrungen zu korrigieren, kamen Designgewichte zum Einsatz, welche nachfolgend beschrieben werden.

Die Gewichtung mit dem Gewicht 1 erfolgte auf Basis von Auswahl- bzw. Inklusionswahrscheinlichkeiten der Stichprobenelemente (Horvitz und Thompson 1952), (Lohr 2022). Insbesondere sollen Antworten gewichtet werden, welche sowohl von Mieter- und/oder auch Vermieterseite beantwortet wurden. Insgesamt befanden sich 78 doppelt bzw. dreifache vorhandene Stichprobenelemente im Datensatz. Nach Prüfung auf Antwortdiskrepanzen bei den vergleichbaren Merkmalen der jeweils doppelt vorliegenden Antworten konnten keine signifikanten Unterschiede festgestellt werden. In der Stichprobe wurde jeweils ein per Zufall gewähltes Element der Duplikate im Datensatz belassen.

Konkret setzt sich das Designgewicht aus den Wahrscheinlichkeiten zusammen, mit welcher ein Haushalt in die Zufallsstichprobe der Schicht A oder Schicht B gelangt. Weiter gilt es diejenige Auswahlwahrscheinlichkeit zu berücksichtigen, mit welcher ein Haushalt im Überlappungsbereich der beiden Zufallsstichproben von A und B liegt. Mathematisch bezeichnet man die Auswahlwahrscheinlichkeit für ein Element aus der Stichprobe von Schicht A mit π_i^A , für ein Element aus der Stichprobe von Schicht B mit π_j^B , bzw. eines Elements aus dem Überlappungsbereich von A und B mit $\pi_{ij}^{A \cap B}$. Wir notieren dies kurz mit $\pi_A, \pi_B, \pi_{A \cap B}$. Anschließend werden die Inversen ($d_i := 1/\pi_i$) als Gewichtungsfaktoren in die Schätzung des Mietpreismodells mit aufgenommen. Die Gewichte werden errechnet und auf die entsprechende Stichprobengröße normiert. Das normierte Gewicht d_i^* errechnet sich aus dem Quotienten des Gewichts und der Summe aller Einzelgewichte in der Stichprobe (Rücklauf, Tabelle 3, Spalte 5). Anschließend wird dieser Quotient mit der Stichprobengröße aller Antworten (Rücklauf, Tabelle 3, Spalte 7) multipliziert. Damit erfolgt eine Skalierung dieser Designgewichte, sodass diese in ihrer Summe der Stichprobengröße entsprechen und einen Mittelwert von „1“ aufweisen.

Tabelle 5: Übersicht Gewicht 1

Auswahlwahrscheinlichkeit	Wert	Gewicht d	d normiert
π_A	$\pi_A = \left(\frac{1665}{3000}\right) = 0,56$	$\frac{1}{0,56}$ $= 1,80$	$\frac{1,804}{4230} \cdot 2187$ $= 0,70$
π_B	$\pi_B = \left(\frac{522}{1000}\right) = 0,522$	$\frac{1}{0,522}$ $= 1,92$	$\frac{1,92}{4230} \cdot 2187$ $= 6,142$
$\pi_{A \cap B}$	$\pi_{A \cap B} = (0,56 + 0,522 - 0,56 \cdot 0,522) = 0,792$	$\frac{1}{0,792}$ $= 1,27$	$\frac{1,27}{4230} \cdot 2187$ $= 0,681$

5.5 Einzeldatenanalyse und Datenumkodierung

Um einen Eindruck von der Bedeutung und Häufigkeit einzelner Wohnwertmerkmale zu bekommen, wurden Häufigkeitsanalysen vorgenommen. Kreuztabellen und Korrelationsanalysen gewährten Einblicke in die Zusammenhangsstruktur einzelner Merkmale. Dies diente der Identifikation potenzieller Interaktionen einzelner Wohnwertmerkmale untereinander. Dieser Schritt ist unabdingbar, um über eine Entscheidungsgrundlage für die nachfolgende Merkmalsauswahl, Merkmalsumkodierungen und die Bildung komplexer Merkmalskombinationen zu verfügen. Grundsätzlich werden, nach Abschluss der Plausibilisierungsarbeiten des Datensatzes, alle gesammelten Wohnwertmerkmale in ihrer originären Form dahingehend untersucht, ob genügend Häufigkeiten vorhanden sind. Die allgemeine Faustregel liegt hier bei mindestens 30 Häufigkeiten pro Merkmal. Merkmale, welche unter 30 Häufigkeiten aufweisen, werden entweder nicht weiter für das eigentliche statistische Modell (siehe Abschnitt 7.4) verwendet und fallen somit in die Spannenregelung (siehe Abschnitt 7.9) oder werden zu übergeordneten Variablen sachlogisch zusammengefasst. Liegen beispielsweise für die Bodenbeläge Parkett- und Dielenholzboden 15 und 40 Häufigkeiten vor, so können diese Bodenbelagsarten zu einem neuen Merkmal zusammengefasst werden, welche als „hochwertiger Bodenbelag“ definiert werden kann. Ein weiteres Beispiel für das Zusammenfassen von Merkmalen bilden die Baujahresklassen (siehe Abschnitt 7.5.3). In einem weiteren Schritt werden die Merkmale dann hinsichtlich ihres Einflusses auf den Mietpreis untersucht. Je nach Verteilung und Skalierung der Variable kommen zwei bekannte Korrelationskoeffizienten zur Anwendung: Für normalverteilte Variablen berechnet man üblicherweise die Pearson-Korrelation. Bei zu starken Abweichungen von der Normalverteilung muss auf einen anderen Korrelationsbegriff zurückgegriffen werden. Hierfür eignet sich dann z. B. der Spearman'sche Rangkorrelationskoeffizient (Fahrmeir 2016). Der Korrelationskoeffizient ist eine Abbildung in das reelle Intervall $[-1,1]$. Liegt der Korrelationskoeffizient zweier Variablen signifikant zwischen dem Wert $-0,3$ und -1 oder $0,3$ und 1 ,

so kann im Mietspiegelkontext oft von einer mittleren bis starken Korrelation gesprochen werden. Liegt der Wert nahe bei 0, so korrelieren die Werte nur schwach oder gar nicht miteinander⁴. Für die weitere Betrachtung von Merkmalen im statistischen Modell wurden nur Variablen verwendet, welche eine starke Korrelation mit dem Mietpreis zeigten (vgl. Abbildung 6).

Oben genanntes Schema wurde auf alle Variablen angewandt. In einem weiteren Schritt wurden sodann die verwendbaren Merkmale dort zusammengefasst bzw. kombiniert, wo es die statistische Sachlogik verlangte.

Erfahrungen aus der Erstellung früherer Mietspiegel in anderen Städten und die vorangegangene Datenanalyse zeigen, dass die Bildung einzelner neuer mietpreisdeterminierender Merkmale aus dem Primärmerkmalsbestand von Vorteil ist. Dies hat verschiedene Gründe. Zum einen stehen einzelne Merkmale oft für einen wesentlich komplexeren Sachverhalt. Zum anderen lassen sich Stellvertretervariablen identifizieren und das Problem der Multikollinearität besser überblicken.

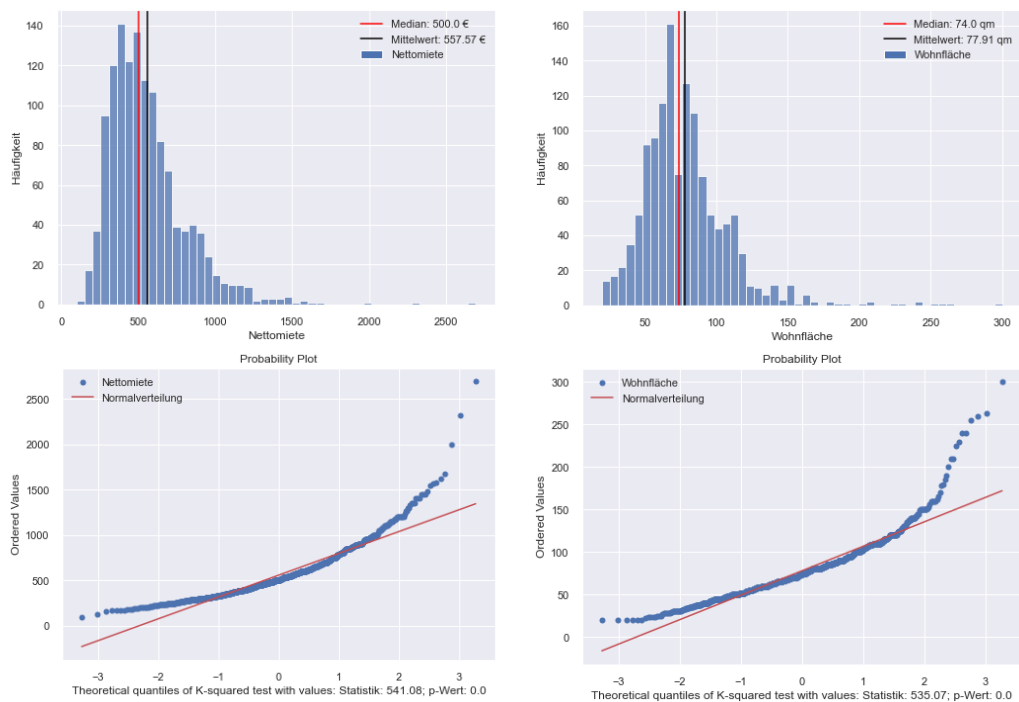
⁴ An dieser Stelle sei darauf hingewiesen, dass es in der Literatur verschiedene Definitionen von starker, mittlerer und schwacher Korrelation gibt.

6 Deskriptive Statistik

6.1 Eigenschaften der Merkmale Nettomiete und Wohnfläche

Nach der Datenaufbereitung ist zu prüfen, ob die wesentlichen Merkmale für das spätere Regressionsmodell geeignet und dessen zugehörigen statistischen Annahmen der Normalverteilung vorhanden sind. Die Variablen Nettomiete (nmneu) sowie Wohnfläche (wflneu) werden mit dem „K-Squared-Test“ nach D’Agostino auf Normalverteilung getestet (Ralph B. D’Agostino 1971). Die Testwerte, die Häufigkeitsverteilung sowie der Quantil-Quantil-Plot sind nachfolgend aufgeführt.

Tabelle 6: Normalverteilung von Nettomiete und Wohnfläche

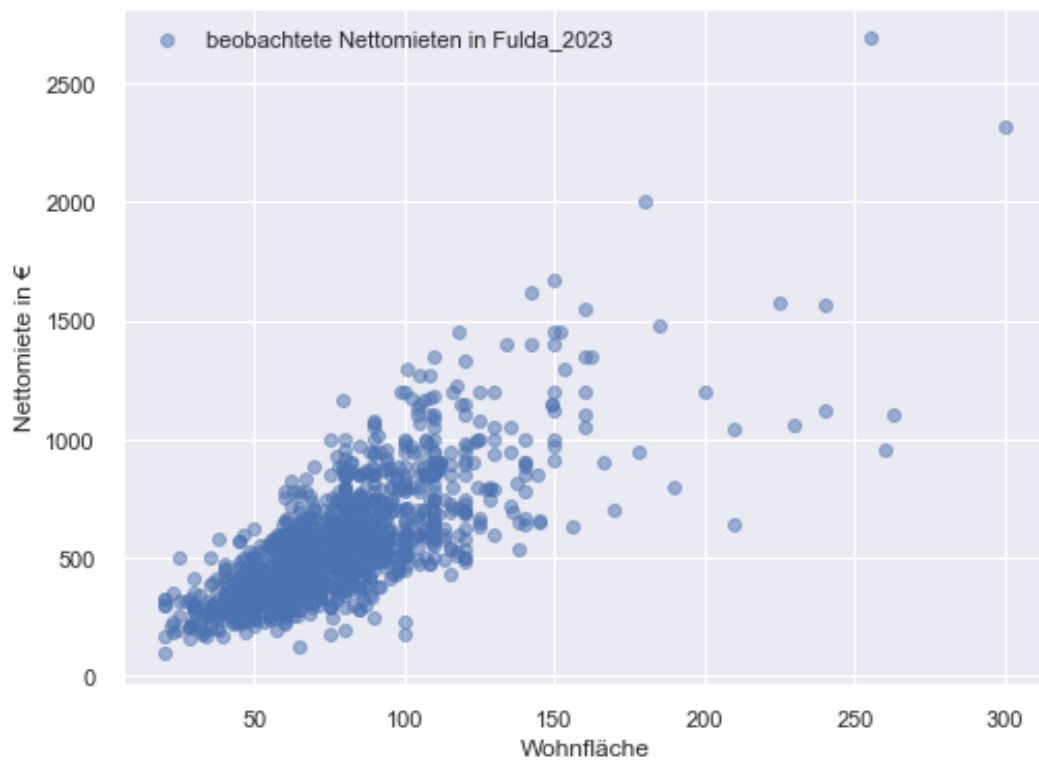


Für eine spätere Modellierung der Nettomiete (Zielvariable) anhand der Wohnfläche (Regressand) muss weiterhin eine Korrelation zwischen Zielvariable und Regressand nachgewiesen werden. Dieser Nachweis erfolgt mit einem Korrelationstest für normalverteilte Variablen nach Pearson (Fahrmeir 2016).

Tabelle 7: zweiseitiger Korrelationskoeffizient am Beispiel Fulda

	Nettomiete	Wohnfläche
Nettomiete	1,00	0,76
Wohnfläche	0,76	1,00

Abbildung 2: Streuung der Nettomieten entlang der Wohnfläche



7 Regression

7.1 Der Gesamtansatz und das gewählte Modell

Das Mietspiegelreformgesetz und das Bürgerliche Gesetzbuch (BGB) definiert die ortsübliche Vergleichsmiete als „übliche Entgelte, die in einer Gemeinde oder vergleichbaren Gemeinde für Wohnraum vergleichbarer Art, Größe, Ausstattung, Beschaffenheit und Lage einschließlich der energetischen Ausstattung und Beschaffenheit in den letzten **sechs** Jahren vereinbart oder, von Erhöhungen [der Betriebskosten] abgesehen, verändert worden sind“ (§ 558 Abs. 2 Satz 1 BGB). Damit werden durch den Gesetzgeber Festlegungen getroffen, die aber im Einzelfall noch viel Spielraum für weitere Präzisierungen lassen.

Die beiden Schlüsselbegriffe stellen die „*Vergleichbarkeit des Wohnraums*“ und die „*Üblichkeit der Entgelte*“ dar. Die Aufgabe eines Mietspiegels besteht darin, für vergleichbare Wohnungen einen ortsüblichen Mietpreis in einem näher definierten Wohnungsmarkt auszuweisen. Bei der Mietspiegelerstellung hat man im Rahmen vorgegebener äußerer Restriktionen sowohl die Vergleichbarkeit des Wohnraums als auch die Üblichkeit der Entgelte nach anerkannten Grundsätzen der Statistik zu quantifizieren. Dazu zählen z.B. finanzielle und zeitliche Ressourcen oder Diskrepanzen in den zur Verfügung stehenden Datengrundlagen.

Die eingangs des Kapitels definierte Aufgabenstellung ist eine typische Fragestellung der Regressionsanalyse. Fundamental ist die Zugrundelegung einer sinnvollen Abhängigkeitsstruktur von Wohnwertmerkmalen mit dem Mietpreis. Diesem Mietspiegel liegt ein Regressionsmodell zugrunde, das in seinen Grundzügen dem in der Literatur beschriebenen „Regensburger Modell“ gleicht (Aigner et al. 1993). Dieses Modell ermöglicht die Ermittlung des vorliegenden lokalen Mietniveaus und den davon ortsüblichen Abweichungen über ein System von Zu- und Abschlägen je nach dem Vorhanden- oder Nichtvorhandensein spezieller signifikanter Wohnwertmerkmale. Dieser Regressionsansatz wird in seiner Grundkonzeption häufig verwendet, z.B. in Augsburg, Erding, Erlangen, Esslingen, Freiburg, Friedrichshafen, Fürth, Heidelberg, Konstanz, Landshut, Münster, Nürnberg, Regensburg, Trier, Ulm und Villingen-Schwenningen.

Bis zur Fertigstellung des Mietspiegels waren im Rahmen der statistischen Auswertungen verschiedene Arbeitsschritte erforderlich:

1. Aufbereitung des erhobenen Datenmaterials für die Auswertung
2. Umrechnung aller ermittelten Mietpreise auf einen einheitlichen Mietbegriff
3. Ermittlung des durchschnittlichen Mietniveaus
4. Ermittlung von Zu- und Abschlägen für einzelne Wohnwertmerkmale

5. Ermittlung von Spannbreiten

6. Darstellung der ermittelten Vergleichsmieten im Mietspiegel.

Die Arbeitsschritte 1 und 2 wurden in den vorigen Kapiteln behandelt, die Arbeitsschritte 3 bis 6 und deren Resultate sind nachfolgend in der Dokumentation dargestellt.

7.2 Die Grundstruktur des gewählten Regressionsmodells

Seit Ende der 1980er Jahre wird für die Mietspiegelerstellung das multivariate statistische Verfahren der Regressionsanalyse angewendet, das als wissenschaftliches Berechnungsverfahren anerkannt ist. Von dem damaligen Lehrstuhlinhaber für Ökonometrie, Prof. Dr. W. Oberhofer der Universität Regensburg und dem EMA-Institut für empirische Marktanalysen wurde speziell für die Mietspiegelerstellung eine multiplikativ-lineare Regressionsvariante entwickelt, welches von der Form her einem nicht-linearen Regressionsmodell entspricht (Aigner et al. 1993).

Die gesuchte Variable im Regressionsansatz ist die „ortsübliche Vergleichsmiete“ in Form der Nettomiete. Der verwendete Modellansatz lautet

$$NM = (a + f(WFL)) \cdot (a_0 + a_1x_1 + a_2x_2 + \dots + a_px_p),$$

wobei insgesamt p Parameter, welche hier die Wohnwertmerkmale repräsentieren, $a_j, j = 1, \dots, p$ zu schätzen sind. Die Schreibweise a_jx_j ist eine Kurzschreibweise für $a_j \cdot x_j$, also der Multiplikation des Koeffizienten a_j (Zu/Abschlag) mit dem Merkmal x_j (Einflussfaktor bzw. Wohnwertmerkmal). Hier bezeichnet NM die Nettomiete, WFL die Wohnfläche, und x_1 bis x_p verschiedene weitere Wohnwertmerkmale. Die Größen (Koeffizienten) $a, a_0, \dots, a_n$ stellen zu schätzende Parameter und die Terme $f(WFL)$ sowie

$$(a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n)$$

noch näher zu spezifizierende Funktionen in Abhängigkeit der Variablen WFL und $x_1, x_2, \dots, x_n$ dar.

Dieser Ansatz impliziert, dass die Nettomiete aus zwei Faktoren gebildet wird: Einem ersten Faktor, der nur von der Wohnfläche abhängig ist und einem zweiten Faktor, der den Einfluss des Baujahres zusammen mit dem Einfluss weiterer Merkmale, abgeleitet aus dem Begriff der ortsüblichen Vergleichsmiete, erfasst. Die Wohnfläche liefert erfahrungsgemäß den größten Beitrag zur Erklärung der Nettomiete und interagiert sehr oft mit weiteren Variablen, die den Mietpreis bestimmen. Die Wohnfläche allein hat bei dieser Mietspiegelerstellung ein Bestimmtheitsmaß (entspricht dem quadrierten Korrelationskoeffizienten, der den

Gleichlauf von zwei Größen und damit den Zusammenhang misst) in Höhe von $R^2=0,646$.

Der erste Faktor bildet die „Basis-Nettomiete“, kurz die „Basismiete“. Die multiplikative Form des Ansatzes bedingt prozentuale Zu- oder Abschläge. Wenn z.B. x_1 für das Vorhandensein einer Einbauküche steht ($x_1=1$: Einbauküche vorhanden und $x_1=0$: keine Einbauküche vorhanden) und der zugehörige Koeffizient a_1 lautet 0,05, so bedeutet dies einen Zuschlag in Höhe von fünf Prozent für das Vorhandensein einer Einbauküche, bezogen auf die Basismiete für eine bestimmte Wohnfläche. So wird der Summand a_1x_1 aus oben genannter Gleichung gebildet. Alle anderen Summanden berechnen sich auf dieselbe Art und Weise. Der hier vorliegende Ansatz bedingt insbesondere Interaktionen zwischen der Größe der Wohnfläche und allen weiteren Merkmalen ($x_1, x_2, \dots, x_n$), da letztere einen von der Basismiete abhängigen Beitrag zur Nettomiete liefern.

7.3 Auswahl der Merkmale

Es besteht bei der Mietspiegelerstellung ein grundsätzlicher Unterschied zwischen Tabellenmethode und Regressionsmethode. Die Vorgabe für den Statistiker lautet in beiden Fällen: Für vergleichbare Wohnungen einen üblichen Mietpreis zu bestimmen. Die Konkretisierung der Vergleichbarkeit erfolgt beim Tabellenmietspiegel durch eine Auswahl geeigneter mietpreisbestimmender Merkmale, mit deren Hilfe Klassen oder Wohnungstypen gebildet werden. Wohnungen, die zu einer Klasse gehören, sind dann ex definitione vergleichbar. Es wird dann zu jedem Wohnungstyp ein durchschnittlicher Mietpreis berechnet, z.B. das arithmetische Mittel innerhalb der Klasse, und dies ist dann ex definitione der innerhalb der Klasse übliche Mietpreis.

Beim Regressionsmietspiegel werden keine Klassen gebildet. Im Prinzip wird davon ausgegangen, dass Wohnungen, die sich in nur einem Merkmalswert unterscheiden, auch nicht vergleichbar sind. Die Regressionsmethode unterstellt bei Wohnungen mit ähnlichen Merkmalskombinationen, die inhaltlich nahe beieinanderliegen, einen stetigen Übergang des Mietpreises. Im einfachsten Falle mit nur einer Variablen, z.B. der Wohnfläche, wird unterstellt, dass sich der Mietpreis einer Wohnung mit 40 m² Wohnfläche wenig vom Mietpreis einer Wohnung mit 41 m² Wohnfläche unterscheidet und letztere wenig vom Mietpreis einer Wohnung mit 42 m² Wohnfläche. Ein Regressionsmietspiegel modelliert genau diesen Übergang von einer Merkmalskombination zur nächsten. Im eben erwähnten Beispiel könnte beispielsweise unterstellt werden, dass sich die Wohnfläche kontinuierlich von 40 m² über 41 m² bis 42 m² ändert und bei entsprechenden Mietpreisen der Übergang linear mit der Wohnfläche erfolgt.

Während beim Tabellenmietspiegel nur eine geringe Zahl von Merkmalen zur Klassenbildung herangezogen werden kann, da bei einem begrenzten Stichprobenumfang die Anzahl der Klassen sehr beschränkt ist, steht bei der Anwendung des Regressionsmietspiegels eine weit größere Zahl an Merkmalen zur Verfügung. Demnach ist ein Regressionsmietspiegel im Vergleich zu einem Tabellenmietspiegel geeigneter auch komplexe Wohnwertmerkmalskombinationen abzubilden.

Im vorliegenden Falle steht aufgrund des ausführlichen Fragebogens eine Vielzahl von Wohnwertmerkmalen zur Auswahl (vgl. Anlage 10.2 Fragebogen zum Mietspiegel). Eine Auflistung aller im Datensatz vorhandenen Merkmale findet sich im Anhang,

Tabelle 15. Bezieht man alle im Fragebogen abgefragten Merkmale in ihrer originalen Form in die Auswertung mit ein, so würde das statistische Modell überladen werden. Das bedeutet, dass das zu lösende Gleichungssystem nicht mehr das eigentliche Problem schätzt, sondern ggf. andere vorliegende Zusammenhänge versucht zu schätzen. Dieses Problem ist in der Ökonometrie auch als sogenanntes Identifikationsproblem bekannt. Daher müssen iterativ Variablen ohne Korrelation und ohne Sachlogik zur Zielvariable aus der Schätzung ausgeschlossen werden.

Empirische Untersuchungen zeigen, dass die kritische Anzahl zu berücksichtigender Merkmale (bei einem Stichprobenumfang von etwa 1.000 Wohnungen) bei rund 20 liegt. Damit stellt sich das Problem der Auswahl von geeigneten Merkmalen. Hierbei kann man sich nicht auf Erkenntnisse aus der Wohnungsmarkttheorie stützen. Die Lösung des Problems wird zur Aufgabe des Statistikers (siehe Abschnitt 5.5).

Die vorzunehmende statistische Analyse muss explorativen Charakter haben. Dies bedeutet, dass anfangs eine Auswahl von geeigneten, d. h. der Sache entsprechenden Wohnwertmerkmalen getroffen wird, was in der Stadt Fulda in Zusammenarbeit mit dem Arbeitskreis Mietspiegel in Form eines Fragebogens geschehen ist, um mit dieser Auswahl einen Mietspiegel zu erstellen. Die Ergebnisse der Auswertung dieser Wohnwertmerkmale werden hinsichtlich sachlogischer Adäquatheit, Anpassungsgüte, richtiger Vorzeichen der Parameter und Signifikanz überprüft und wie bereits im vorletzten Absatz beschrieben, aus den Berechnungen ausgeschlossen, sofern diese Kriterien nicht erfüllt werden können.

Merkmale mit Parametern, die statistisch schlecht gesichert sind, werden nur begrenzt berücksichtigt. Dann wird der Auswahlprozess mit dem bereinigten Datensatz wiederholt usw.

In diesem Zusammenhang ist zu beachten, dass bei einem Mietspiegel der deskriptive Charakter eine große Rolle spielt. Es geht primär nicht darum, kausale Zusammenhänge zwischen mietpreisbestimmenden Merkmalen und Mietpreis zu finden, sondern zu gewährleisten, dass mit Hilfe der mietpreisbestimmenden Merkmale der Mietpreis gut getroffen wird. Bei Parametern von Merkmalen, die statistisch schlecht gesichert sind, ist ein Ausschluss nicht zwingend geboten. Hauptaugenmerk liegt auf der Güte der Erklärung des Merkmals.

Bei diesem explorativen Prozess spielt die Erklärungsgüte eine zentrale Rolle. In der Praxis tragen, abgesehen von der Wohnfläche und dem Baujahr, die einzelnen Merkmale relativ wenig zur Erhöhung der Erklärungsgüte bei. Es gibt auch den Fall, dass Merkmale, die im Vorhinein als eher unbedeutend betrachtet wurden, einen größeren Erklärungsbeitrag liefern als Merkmale, denen bereits bei der Auswahl

der Wohnwertmerkmale vor der statistischen Auswertung ein höherer Erklärungsbeitrag zum Mietpreis beigemessen wurde. Offensichtlich sind diese vorher als weniger bedeutend erachteten Merkmale Indikatoren für komplexe Sachverhalte. Hier bietet sich für einzelne Bereiche wie z.B. die Badausstattung an, komplexe Merkmalskombinationen zu bilden, d.h. die für einen Bereich relevanten Merkmale zu einem oder zwei Indikatoren zusammenzufassen und nur diese Indikatoren in die Regression einzubeziehen. Von dieser Möglichkeit wird regelmäßig Gebrauch gemacht.

7.4 Die methodische Vorgehensweise bei der Schätzung

Der zugrundeliegende Regressionsansatz

$$NM = ((a + f(WFL)) \cdot (a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n)) + \varepsilon.$$

ist seiner Struktur nach multiplikativ-additiv. Um die unbekannten Parameter zu schätzen, wird eine zweistufige Schätzung durchgeführt, welche auf den einzelnen Phasen linearen Charakter hat. Damit kann die numerische Stabilität und damit die Genauigkeit der Schätzungen besser kontrolliert werden. Bei den Variablen x_1 bis x_n handelt es sich sowohl um einfache als auch um komplexe Wohnwertmerkmale.

Phase 1: Es wird zuerst die Nettomiete nur gegen die Wohnfläche mittels klassischer Regression geschätzt (Regression 1):

$$\widehat{NM} = (a + f(WFL)).$$

Damit resultiert eine rudimentäre Schätzung der Nettomiete nur in Abhängigkeit der Wohnfläche. Anschließend wird der Einfluss der Wohnfläche auf die Nettomiete durch Division der tatsächlich gezahlten Nettomiete NM durch die anhand der Wohnungsgröße geschätzten Nettomiete $\widehat{NM}$ herausgerechnet:

$$NMF = NM / \widehat{NM}.$$

Man erhält die sogenannten Nettomietfaktoren NMF , die um den Wert 1 streuen.

Phase 2: Auf der zweiten Phase wird der Einfluss aller restlichen Merkmale auf den Mietpreis unabhängig von der Wohnfläche regressionsanalytisch durch folgendes Prognosemodell der allgemeinen linearen Regression geschätzt (Fahrmeir et al. 2022):

$$\widehat{NMF} = (a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n),$$

Eine detaillierte Darstellung der Phase 2 erfolgt in den nächsten Abschnitten. Bei der Durchführung der Schätzung ist darauf zu achten, dass α_0 den Wert 1 annimmt.

Man kann festhalten, dass bei dem phasenweisen Vorgehen die Identifikationsproblematik abgemildert wird (Aigner et al. 1993), (Wooldridge 2013). Dies liegt daran, dass der vorliegende Zuordnungskonflikt zu Gunsten von Merkmalen der jeweils vorhergehenden Phase gelöst wird (hier im Falle des Einflusses der Wohnfläche, siehe Phase 1).

7.5 Ermittlung des durchschnittlichen Mietniveaus

Neben der Wohnfläche ist das Baujahr des Gebäudes, in dem sich die Wohnung befindet, von großem Einfluss auf den Mietpreis. Dieser Tatsache wird Rechnung getragen, indem die Wohnfläche und das Baujahr die Grundlage für die Mietniveauermittlung in allen Kommunen bilden.

7.5.1 Ermittlung des Einflusses der Wohnfläche

Die konkrete Spezifizierung des Zusammenhangs zwischen Nettomiete und Wohnfläche (Phase 1) erfolgt im Rahmen einer explorativen Datenanalyse. Trägt man für alle mietspiegelrelevanten Wohnungen die Nettomiete gegen die Wohnfläche ab, so ergibt sich grafisch eine Punktwolke, aus der die Art des funktionalen Zusammenhangs nicht ersichtlich ist.

Mit Hilfe des Bestimmtheitsmaßes lässt sich eine stetige, optimal durch die Punktwolke der erhobenen Nettomieten laufende Regressionsfunktion legen. Anschließend können die Werte der Funktion für jede beliebige Wohnfläche zur Bestimmung einer durchschnittlichen Miete nur in Abhängigkeit von der Wohnfläche verwendet werden. Abbildung 2 zeigt diese Regressionsfunktionen für alle in die Auswertung eingegangen Wohnungen im Untersuchungsgebiet.

Für die Schätzung der Nettomiete in Abhängigkeit der Wohnfläche wird ein Polynom ersten Grades betrachtet. Höhere polynomielle Funktionen zeigten kein besseres Schätzverhalten.

Der Verlauf der Regressionsfunktionen der ersten Phase lautet:

$$\begin{aligned} \text{Fulda:} \quad \widehat{NM}(wfl) = & 201,2818407 + 1,049114773 \cdot wfl + 0,057112662 \\ & \cdot wfl^2 - 0,00017 \cdot wfl^3 \end{aligned}$$

Abbildung 3 zeigt die Regressionskurve der durchschnittlichen Nettomiete in Euro in Abhängigkeit der Wohnfläche durch die Punktwolke aller tatsächlich erhobenen, mietspiegel-relevanten Nettomieten.

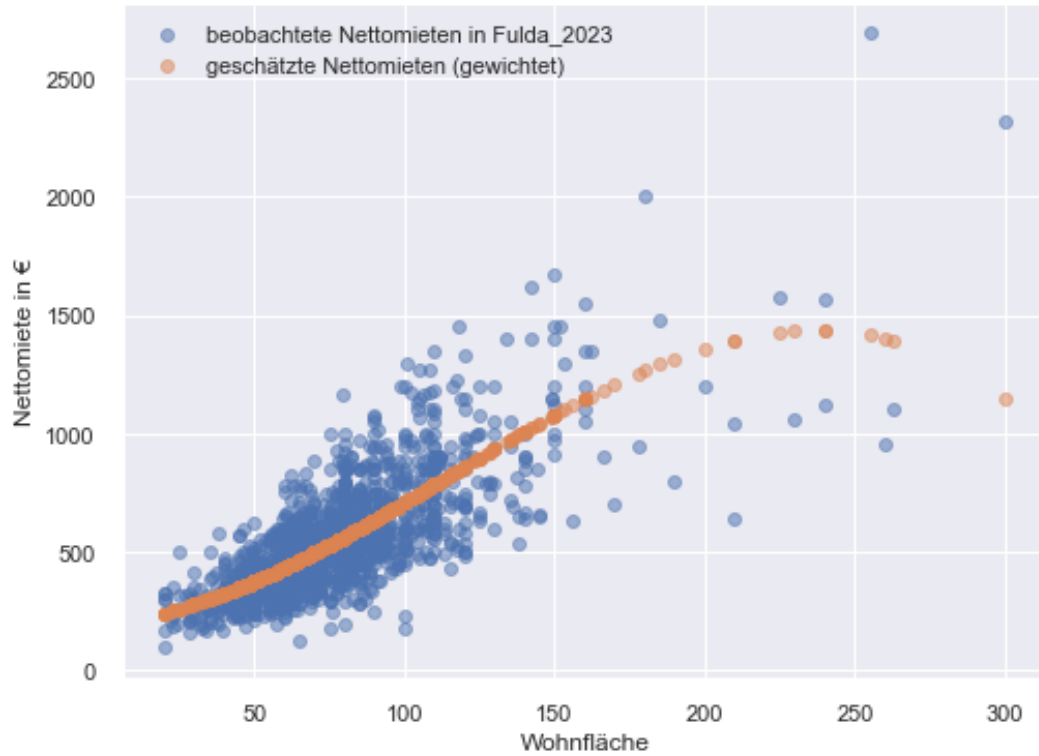
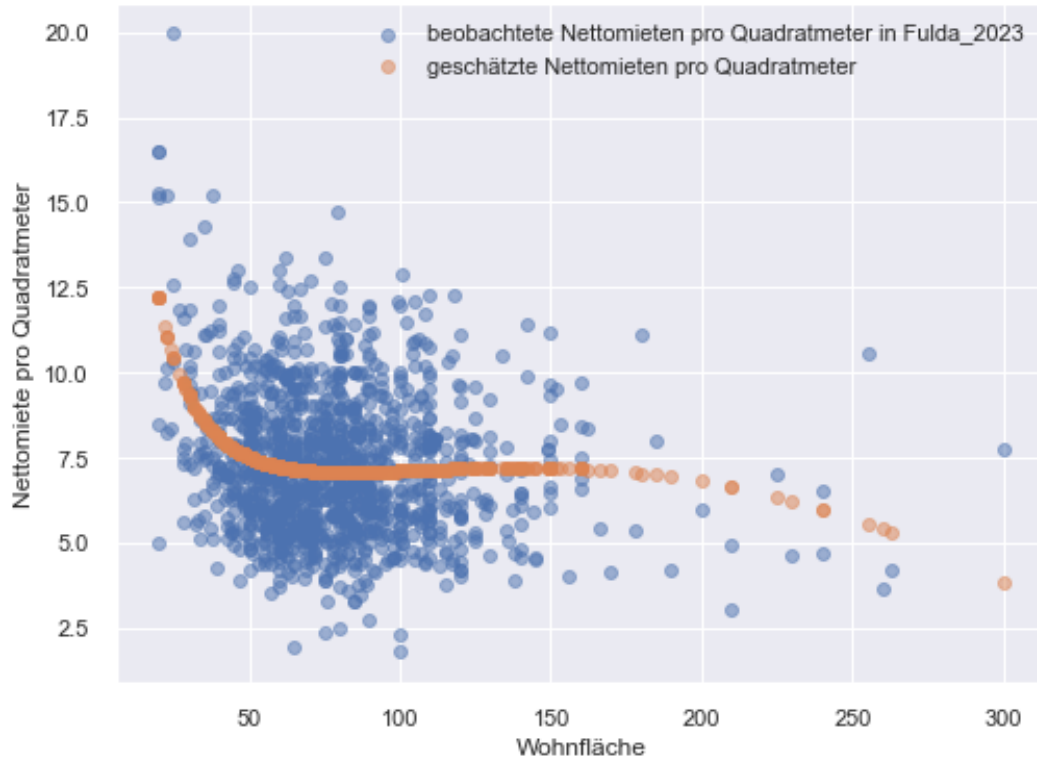


Abbildung 4 zeigt die Regressionskurve der durchschnittlichen Nettomiete in Euro pro Quadratmeter in Abhängigkeit der Wohnfläche durch die Punktwolke aller tatsächlich erhobenen, mietspiegel-relevanten Nettomieten.



Die Miethöhen streuen innerhalb einzelner Wohnflächenklassen unterschiedlich und sind damit nicht homogen, sondern heterogen. Meist streut die Miethöhe bei kleineren Wohnungen weniger im Vergleich zu großen Wohnungen. Statistische Tests belegten die nicht gleichbleibende Streuung und damit eine nicht einheitliche Varianz. Nach Schätzung der Originalmieten gegen die Wohnfläche erfolgte daher eine Varianzanalyse, welche mit Hilfe einer Kleinste-Quadrate-Schätzung durchgeführt wurde (Fahrmeir et al. 2022): Die Varianzen wurden korrigiert, indem der Abstand (Betragsfunktion) der geschätzten Nettomieten zu den tatsächlichen gezahlten Nettomieten nochmals gegen die Wohnfläche regressiert wurde. Die Schätzergebnisse flossen dann als reziprokes Gewicht mit in die Ausgangsregression ein. Eine graphische Übersicht über die Güte des zur Varianzanalyse verwendeten Regressionsmodells findet sich in Abschnitt 10.

Die Regression wird originär mit den Größen „Wohnfläche“ und „absolute monatliche Nettomiete“ durchgeführt. Im Mietspiegel werden dagegen die Ergebnisse in der üblichen Einheit Euro/m² ausgewiesen. Deshalb werden die Ergebnisse grafisch in Euro/m² veranschaulicht (vgl. Abbildung 5).

Die Verwendung einer Regressionsfunktion hat gegenüber der Tabellenmethode mehrere Vorteile:

- Durch die Verwendung dieser Funktion werden Sprünge, die zwischen den Mittelwerten benachbarter Wohnflächenklassen auftreten und u.a. auf Zufälligkeiten innerhalb der Stichprobe beruhen können, geglättet.
- Die Regressionsfunktion bildet die Basis für die nachfolgende mathematische Ermittlung der Zu- und Abschläge anhand weiterer Wohnwertmerkmale.
- Die Verwendung der Regressionsfunktion erleichtert die Fortschreibung des Mietspiegels in den nächsten Jahren, da damit die zeitliche Veränderung des Mietniveaus auf einfache Weise festgestellt werden kann und die Werte im Mietspiegel entsprechend angepasst werden können.

Weitere Validierungsinformationen des o. g. Regressionsansatzes in der **ersten Phase** sind in Abbildung 12 und Abbildung 13 bzw. **Fehler! Verweisquelle konnte nicht gefunden werden.** und **Fehler! Verweisquelle konnte nicht gefunden werden.** dargestellt.

7.5.2 Darstellung des durchschnittlichen Mietniveaus

Als Ergebnis dieser Berechnungen erhält man für die Kommune im Mietspiegel die Basisnettomiettablette in der üblichen Dimension Euro/m². In der nachfolgenden Tabelle sind die Basisnettomieten dargestellt.

Tabelle 8: monatliche Basis-Nettomiete in Abhängigkeit von Wohnflächen und Baujahr

Wohnfläche	Baujahr								
	bis 1918	1919 - 1945	1946 - 1977	1978 - 1984	1985 - 1995	1996 - 2004	2005 - 2012	2013 - 2018	2019 - 2023
25 - < 30	10,24	9,90	9,69	10,00	10,17	10,60	10,96	11,31	11,61
30 - < 40	8,93	8,63	8,45	8,72	8,87	9,24	9,56	9,86	10,12
40 - < 45	8,20	7,92	7,76	8,01	8,14	8,48	8,78	9,05	9,29
45 - < 50	7,89	7,63	7,47	7,71	7,84	8,16	8,45	8,71	8,94
50 - < 55	7,68	7,42	7,26	7,50	7,62	7,94	8,21	8,47	8,70
55 - < 60	7,53	7,27	7,12	7,35	7,48	7,79	8,05	8,31	8,53
60 - < 70	7,39	7,14	6,99	7,22	7,34	7,64	7,91	8,16	8,37
70 - < 80	7,31	7,06	6,91	7,14	7,26	7,56	7,82	8,07	8,28
80 - < 90	7,29	7,05	6,90	7,12	7,24	7,54	7,80	8,05	8,26
90 - < 115	7,34	7,09	6,95	7,17	7,29	7,59	7,85	8,10	8,32
115 - < 130	7,42	7,17	7,02	7,24	7,36	7,67	7,93	8,19	8,40
130 - 150	7,44	7,19	7,04	7,27	7,39	7,70	7,96	8,22	8,44

Die durchschnittliche⁵ Nettomiete pro m² über alle in Fulda erfassten, mietspiegelrelevanten Nettomieten pro m², unabhängig von allen Wohnwertmerkmalen, beträgt zum Zeitpunkt der Datenerhebung 7,29 Euro/m². Eine Differenzierung der Nettomiete pro m² nach den Mietpreis beeinflussenden Wohnwertmerkmalen kann mit Hilfe der Tabellen 1 und 2 im Mietspiegeldokument durchgeführt werden.

7.5.3 Ermittlung des Einflusses des Baujahres

Das Baujahr stellt für die Bestimmung der Nettomiete nach der Wohnfläche einen weiteren wichtigen Einflussfaktor dar.

Beim Baujahr sind zwei Besonderheiten zu beachten: Erstens sind im Fragebogen teilweise nur eine Baujahressklasse und kein exaktes Baujahr angegeben und zweitens ist der Einfluss des Bau auf den Mietpreis nicht durchgehend monoton. Letzteres bedeutet beispielsweise, dass eine Wohnung, die im Jahre 1940 gebaut wurde, gegenüber einer Wohnung aus dem Jahre 1900, auch wenn beide sich hinsichtlich der anderen berücksichtigten Merkmale nicht unterscheiden, im Schnitt einen niedrigeren Mietpreis pro m² aufweist.

⁵ wobei mit durchschnittlicher Nettomiete pro m² das arithmetische Mittel über alle Nettomieten pro m² gemeint ist, welche für die Mietpreisschätzungen herangezogen wurden.

Das erste Problem wurde in folgender Weise gelöst: Falls das genaue Baujahr vorlag, wurde dieses verwendet. War nur eine Baujahresklasse angegeben, fand die Klassenmitte Verwendung.

Tabelle 9: Baujahresklassen original lt. Fragebogen

#	Baujahresklassen	Häufigkeiten	relative Häufigkeiten
1	bis 1918	99	7,7%
2	1919 - 1945	194	15,0%
3	1946 - 1977	498	38,5%
4	1978 - 1984	133	10,3%
5	1985 - 1995	99	7,7%
6	1996 - 2004	78	6,0%
7	2005 - 2012	49	3,8%
8	2013 - 2018	68	5,3%
9	2019 - 2023	76	5,9%
	Summe	1.294	100%

Damit war jeder Wohnung ein konkretes Baujahr, kurz **bj**, zugeordnet.

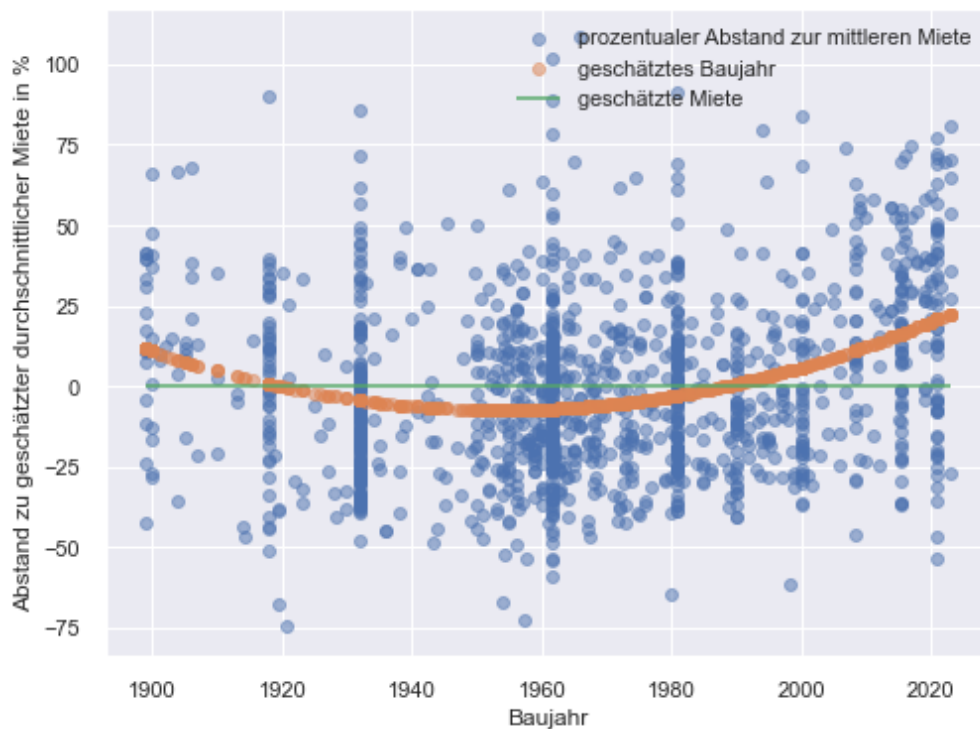
Um die Abhängigkeit des Mietpreises vom Baujahr exakt zu spezifizieren, kann ein Baujahrindikator „**bind**“⁶ gebildet werden, der den nichtmonotonen Verlauf des Baujahreseinflusses auf den Mietpreis modelliert. Analog zu der Vorgehensweise bei der Ermittlung des Einflusses der Wohnfläche auf den Mietpreis wurden dazu Baujahresklassen gebildet. Es resultiert ein Verlauf, der plausibel ist: Bis nach dem zweiten Weltkrieg nimmt die „*Bauqualität*“ einer Wohnung, die allein dem Baujahr zuzuschreiben ist, und damit der Mietpreis ab und danach kontinuierlich wieder zu. Mit Hilfe einer Regressionskurve kann der funktionale Verlauf dieser Zu- und Abschläge aufgrund des Baujahres beschrieben werden. Diese Funktion wird ebenfalls durch ein kubisches Polynom beschrieben und ist gegeben durch:

$$bind = (240,8607971 - 0,245586829 \cdot bj + 0,0000628429 \cdot bj^2)$$

Durch diesen funktionalen Verlauf kann jeder Wohnung über das entsprechende Baujahr **bj** ein Wert des Baujahrindikators „**bind**“ zugeordnet werden. Während Wohnungen in Gebäuden mit Baujahr zwischen 1900 und 1980 unterhalb der durchschnittlichen Miete liegen, liegen jüngere Wohnungen preislich über dem Schnitt. Damit kann jeder Baujahresklasse auch ein konkreter Zu- oder Abschlag zugewiesen werden.

Abbildung 5: Verlauf des Baujahresindikators

⁶ Die Variable „**bind**“ nimmt in dem in Abschnitt 7.2 beschriebenen Regressionsansatz die Rolle eines der $x_i, i \in \{1, \dots, n\}$ ein.



7.6 Ermittlung von Zu- und Abschlägen für weitere Wohnwertmerkmale

Neben der Wohnfläche gibt es noch zahlreiche weitere mietpreisrelevante Merkmale. Die Auswahl dieser Merkmale erfolgt ebenfalls auf der zweiten Phase des Regressionsverfahrens. Im Rahmen eines intensiven iterativen Auswertungsprozesses wurde eine auf inhaltlichen und statistischen Aspekten beruhende Merkmalauswahl getroffen. Bei der Auswahl kamen verschiedene Gesichtspunkte zum Tragen. Vorab konnte bei der Erstellung des Fragebogens (und damit bei der Vorauswahl der Merkmale) auf Erfahrungen aus früheren Mietspiegelerstellungen anderer Städte sowie auf die Erfahrung der im Arbeitskreis Mietspiegel vertretenen Mitgliedern zurückgegriffen werden. Ferner wurden Erkenntnisse über einzelne Merkmale aus den deskriptiven statistischen Analysen zur Merkmalsselektion verwendet (z.B. Häufigkeit des Auftretens). Zum Dritten wurden im Rahmen des beschriebenen explorativen und iterativen Auswertungsprozesses verschiedene Merkmalskombinationen untersucht und verglichen. Ebenfalls untersucht wurden die wichtigsten Interaktionsmöglichkeiten von Variablen.

Bei dieser Analyse wurden die jeweiligen Ergebnisse auch hinsichtlich sachlogischer Adäquatheit, Korrelation der Merkmale, Anpassungs- und Prognosegüte sowie statistischer Signifikanz der Merkmale kontrolliert.

Die unter diesen Vorgaben durchgeführte explorative Vorgehensweise führte zur Auswahl von mietpreisbeeinflussenden Wohnwertmerkmalen in Form einer Mischung aus komplexen und einfachen Wohnwertkombinationen. Die einzelnen

Merkmale stellen eine Konkretisierung der in § 558 Abs. 2 Satz 1 BGB genannten Wohnwertmerkmale Art, Beschaffenheit, Ausstattung und Lage der Wohnung dar. Das Ergebnis der Regression der zweiten Phase ist im nächsten Abschnitt abgebildet.

7.6.1 Ergebnis und Übersicht der Regressionsanalyse Phase 2

Nachfolgend wird das Ergebnis der Schätzung zweiten Phase der Regression geliefert. Anschließend werden die Koeffizienten benannt und erläutert.

Tabelle 10: Variablen im Regressionsmodell (Phase 2)

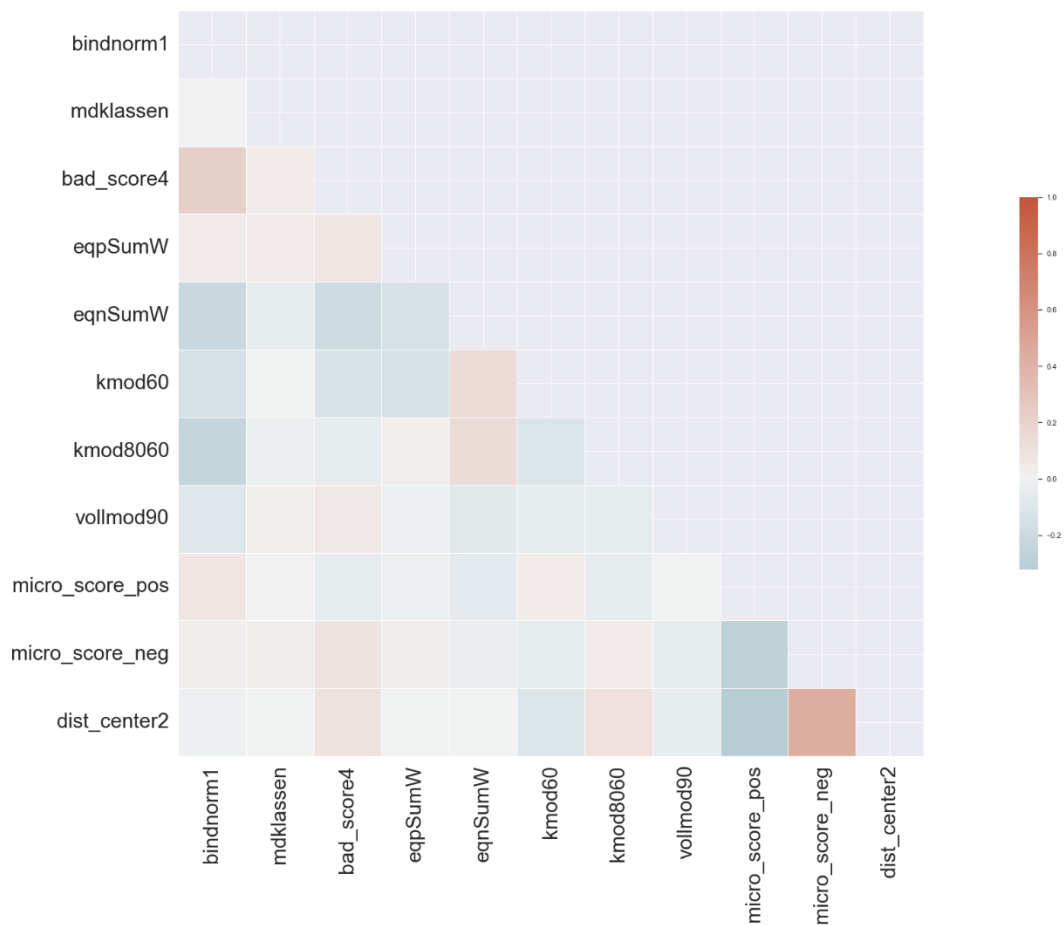
#	Merkmal	Kurzbeschreibung	Min.	Max.	Häufigkeit	N	Mittelwert
1	nmf	Nettomietfaktoren	0,26	2,08	-	1294	1,0
2	bindnorm1	Baujahresindikator	-0,03	0,27	-	1294	0,04
3	mdklassen	außergesetzliche Merkmale	1	6	-	1294	2,48
4	bad_score4	gehobene Sanitärausstattung	0	1	{1: 627, 0: 667}	1294	-
5	eqpSumW	Score: positive Ausstattungsmerkmale	0	7	-	1294	1,47
6	eqnSumW	Score: negative Ausstattungsmerkmale	0	5	-	1294	0,95
7	kmod60	Keinerlei Modernisierung an Wohnung oder Gebäude mit Baujahr vor 1960 durchgeführt seit 2013	0	1	{0: 1131, 1: 163}	1294	-
8	kmod8060	Keinerlei Modernisierung an Wohnung oder Gebäude mit Baujahr zwischen 1960 und 1980 durchgeführt seit 2013	0	1	{0: 1190, 1: 104}	1294	-
9	vollmod90	Vollsanierung der Wohnung oder des Gebäudes mit Baujahr vor 1990 durchgeführt seit 2013	0	1	{0: 1235, 1: 59}	1294	-
10	micro_score_pos	Score: positive Mikrolage	0	5	-	1294	1,15
11	micro_score_neg	Score: negative Mikrolage	0	4	-	1294	0,26
12	dist_center2	Distanz zum Stadtzentrum > 1.000m entfernt	0	1	{1: 513, 0: 781}	1294	-

Tabelle 11: Koeffizienten und Kennzahlen des Regressionsmodells (Phase 2)

Results: Weighted least squares						
Model:	WLS	Adj. R-squared:	0.199			
Dependent Variable:	nmf	AIC:	61.8329			
Date:	2024-01-14 23:22	BIC:	123.8188			
No. Observations:	1294	Log-Likelihood:	-18.916			
Df Model:	11	F-statistic:	30.19			
Df Residuals:	1282	Prob (F-statistic):	5.11e-57			
R-squared:	0.206	Scale:	0.059155			
	Coef.	Std.Err.	t	P> t	[0.025	0.975]
const	1.0002	0.0219	45.7228	0.0000	0.9573	1.0431
bindnorm1	0.6840	0.0879	7.7788	0.0000	0.5115	0.8564
mdklassen	-0.0148	0.0043	-3.4142	0.0007	-0.0233	-0.0063
bad_score4	0.0653	0.0143	4.5520	0.0000	0.0372	0.0934
eqpSumW	0.0190	0.0054	3.5298	0.0004	0.0085	0.0296
eqnSumW	-0.0341	0.0065	-5.2375	0.0000	-0.0469	-0.0213
kmod60	-0.0776	0.0217	-3.5771	0.0004	-0.1201	-0.0350
kmod8060	-0.0502	0.0263	-1.9099	0.0564	-0.1018	0.0014
vollmod90	0.1104	0.0334	3.2996	0.0010	0.0447	0.1760
micro_score_pos	0.0169	0.0063	2.7036	0.0070	0.0046	0.0292
micro_score_neg	-0.0405	0.0110	-3.6800	0.0002	-0.0621	-0.0189
dist_center2	-0.0407	0.0160	-2.5408	0.0112	-0.0722	-0.0093
Omnibus:	112.965	Durbin-Watson:	1.976			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	210.979			
Skew:	0.582	Prob(JB):	0.000			
Kurtosis:	4.599	Condition No.:	48			

Um einen Wert nahe 1 der Konstante „const“ im Mietpreismodell zu erreichen, wurde der Einfluss des Baujahres um durchschnittlich -4,0% gemindert.

Abbildung 6: Nicht-parametrische Korrelationen der Variablen im Regressionsmodell (Phase 2). Auffällig ist die mittlere Korrelation der Ausstattungsgegenstände mit dem Baujahr.



Die genannten Zu-/Abschläge wurden systematisch in vier Kategorien, Baujahr, Wohnungsausstattung- bzw. Beschaffenheit, Modernisierungsmaßnahmen sowie Wohnlage eingeordnet. Der Wert eines Regressionskoeffizienten a aus obiger Tabelle, welcher jedem der aufgeführten Variablen zugewiesen ist, repräsentiert den prozentualen Zu- bzw. Abschlag dieses Merkmals der in der Basis-Nettomiettable ausgewiesenen Preisangaben (€/m²) und muss mit dem Faktor 100 multipliziert werden. Zum Beispiel entspricht der Koeffizient $a=0,01$ dabei einem Wert eines Prozents, ein Koeffizient von $a=0,02$ einem Wert von zwei Prozent usw.

7.6.2 Beschreibung der Variablen im Regressionsmodell - Phase 2

Von den ca. 100 im Fragebogen abgefragten verschiedenen Wohnungsausstattungs- und Beschaffenheitsmerkmalen erwiesen sich, neben dem Baujahr und der Wohnfläche, die in Tabelle 11 genannten Merkmale als ausreichend signifikant mietpreisbeeinflussend.

Tabelle 12: Auf den Mietspreis signifikant wirkende Wohnwertmerkmale

Merkmal	in %	
	Zuschlag	Abschlag
Vollmodernisierung, durchgeführt seit 2013		
Vollsanierung der Wohnung oder des Gebäudes mit Baujahr vor 1990	11	
Teilmodernisierung, durchgeführt seit 2013		
Mindestens 4 der nachfolgend genannten Modernisierungsmaßnahmen: • Sanitärbereich (mind. Fliesen, Wanne, WC) erneuert • Elektroinstallation (zeitgemäß) erneuert • Heizanlage/Warmwasserversorgung erneuert • Schallschutz eingebaut • Fußböden erneuert • Fenster-/Rahmenerneuerung • Innen- und Wohnungstüren erneuert • Treppenhaus, Eingangsbereich erneuert • Grundriss verbessert • barrierearme Ausstattung geschaffen (Mindestvoraussetzung: schwellenfrei (max. 4cm Höhe), stufenloser Zugang, bodengleiche Dusche)	4	
fehlende Modernisierung, durchgeführt seit 2013		
Keinerlei Modernisierung an Wohnung oder Gebäude mit Baujahr zwischen 1960 und 1980 durchgeführt		5
Keinerlei Modernisierung an Wohnung oder Gebäude mit Baujahr vor 1960 durchgeführt		8
Ausstattung und Beschaffenheit Hinweis: Alle Ausstattungskriterien einer Wohnung müssen vom Vermieter zur Verfügung gestellt werden		
Mindestens 4 der nachfolgend genannten Sanitärausstattungskriterien: • zwei oder mehr abgeschlossene Badezimmer in der Wohnung vorhanden • zweites WC/Gäste-WC vorhanden • Badewanne	4	

Merkmal	in %	
	Zuschlag	Abschlag
• separate Einzeldusche • Fußbodenheizung • WC im Badezimmer • Boden ist gefliest • Belüftung(sanlage) • separater WC-Raum vorhanden • Handtuchheizkörper • zweites Waschbecken		
Einbauküche mit mindestens zwei Elektroeinbaugeräten (z. B. Herd inkl. Ofen, Gefrier-schrank/-truhe, Kühlschrank, Geschirrspülmaschine) wird vom Vermieter ohne zusätzlichen Mietzuschlag gestellt.	4	
barrierearme Wohnung (Mindestvoraussetzung: schwellenfrei ⁷ , stufenloser Zugang, bodengleiche Dusche)	4	
Terrasse oder Dachterrasse	2	
Installationsleitungen (z.B. Elektro, Wasser, Gas) freiliegend sichtbar über Putz		3
weder Keller- noch Dachspeicheranteil vorhanden		3
dezentrale Heizungsversorgung, d. h. Einzelöfen (Öl, Gas, Holz, Kohle, Stromspeicher u.ä.)		3
überwiegend Einfachverglasung der Fenster		3
Teppichboden, PVC- oder Linoleum-Boden im überwiegenden Teil des Wohn-/ Schlafbereichs, abgesehen von Flur/Bad verbaut, gilt nur für Baujahre vor 2000 ohne Sanierung des Fußbodenbelages		6
Lage (Distanzen als Luftlinie in m gemessen)		
Distanz zur nächsten Schule <= 150m	2	
Distanz zur nächsten Kita <=200m	2	
Distanz zum nächsten Supermarkt <= 200m	2	

⁷ *Schwellen bis max. 4cm Höhe

Merkmal	in %	
	Zuschlag	Abschlag
Distanz zur nächsten Freizeiteinrichtung <=150m	2	
hohe Bebauungsdichte ⁸	2	
Distanz zur nächsten Schule > 800m		2
Distanz zum nächsten Supermarkt > 1500m		4
Distanz zur nächsten Freizeiteinrichtung > 600m		2
Distanz zum Stadtzentrum > 1.000m entfernt (Definition: Bürgerbüro Fulda: https://www.openstreetmap.org/relation/454863#map=15/50.5519/9.6788)		4
Punktsumme der Zuschläge:		
Punktsumme der Abschläge:		

⁸ mehr als 35% der umliegenden 15.000 m² sind durch Gebäude bebaut

7.6.2.1 Lage

Der Einfluss der Makro- und Mikrolage wurde ausschließlich anhand von georeferenziertem Datenmaterial untersucht und dargestellt. Damit wurden die gesetzlichen Anforderungen der Mietspiegelreform umgesetzt (§ 19 MSV). Das Datenmaterial wurde von der Stadtverwaltung Fulda zur Verfügung gestellt. Eine Übersicht über die verwendeten Datenattribute sowie deren Herkunft findet sich in Tabelle 14. Die dort aufgeführten Merkmale wurden verschiedenen Analysen hinsichtlich ihres Mittelwertes und ihres Einflusses auf die Nettokaltmiete sowie anderen Merkmalen unterzogen. Alle dort enthaltenen Distanzangaben sind als Luftliniendistanzen gemessen. Diese Untersuchung führte zu verschiedenen Darstellungen, Trennungen und Zusammenfassung der genannten Lagemerkmale. Z. B. wurden Bodenrichtwertzonen an Stadtteile und der Bebauungsdichte gekoppelt oder Mikrolagemerkmale wie beispielsweise die Luftliniendistanz zur nächsten Bushaltestelle mit der Luftliniendistanz zum nächsten Supermarkt kombiniert. Insgesamt wurden knapp 50 Lagemerkmale gebildet und untersucht (vgl.

Tabelle 15). Schlussendlich zeigte sich neben der Zentrumsnähe einige weitere Mikrolagermerkmale wie Freizeiteinrichtung, Bebauungsdichte, Schule, Supermarkt usw. (vgl. Tabelle 12).

Abbildung 7: Ausschnitt aus der Stichprobenverteilung auf das Stadtgebiet.

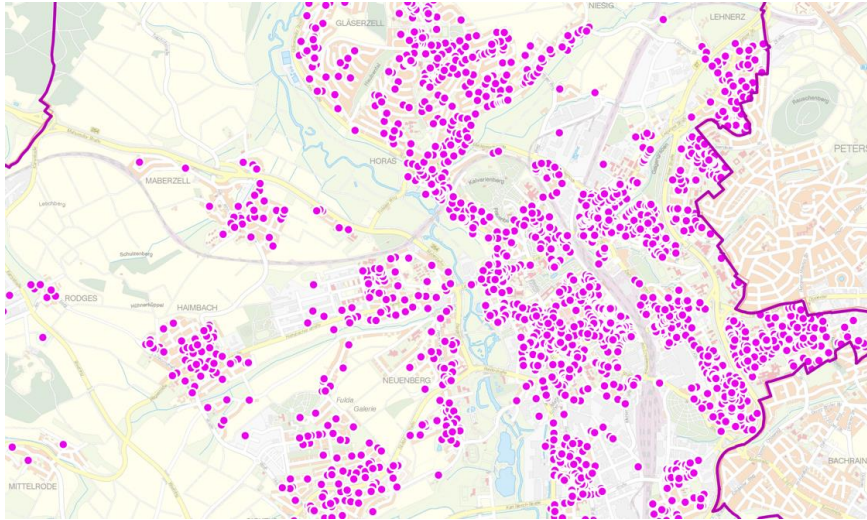


Abbildung 8: Ausschnitt aus der Grafik „Bebauungsdichte“ des Stadtgebietes. Je dunkler die Farbhinterlegung, desto höher ist die Bebauungsdichte.

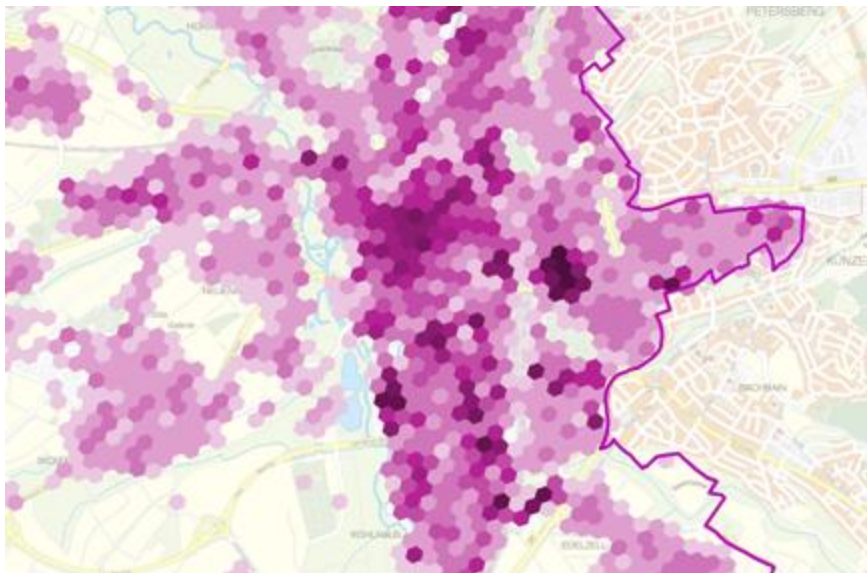


Abbildung 9: Ausschnitt aus der Grafik „Freizeiteinrichtungen“ des Stadtgebietes.

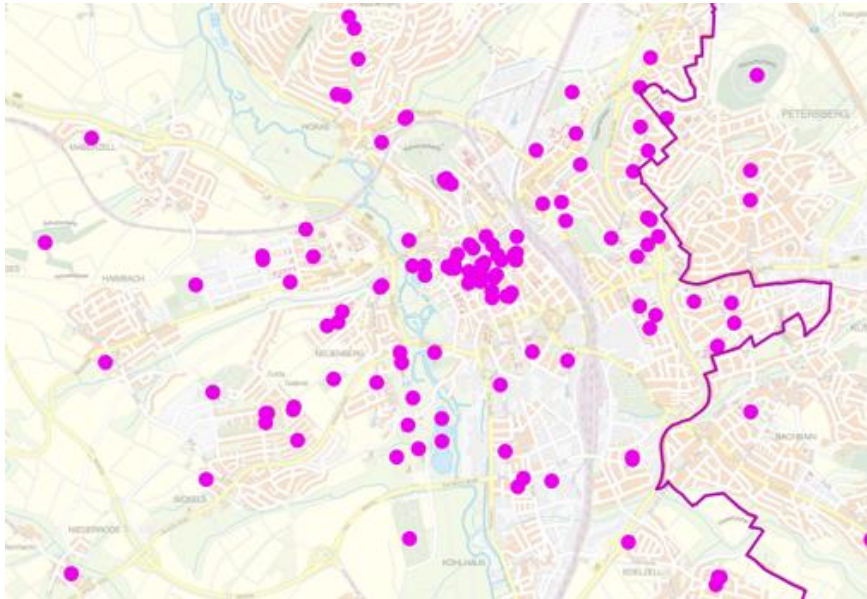
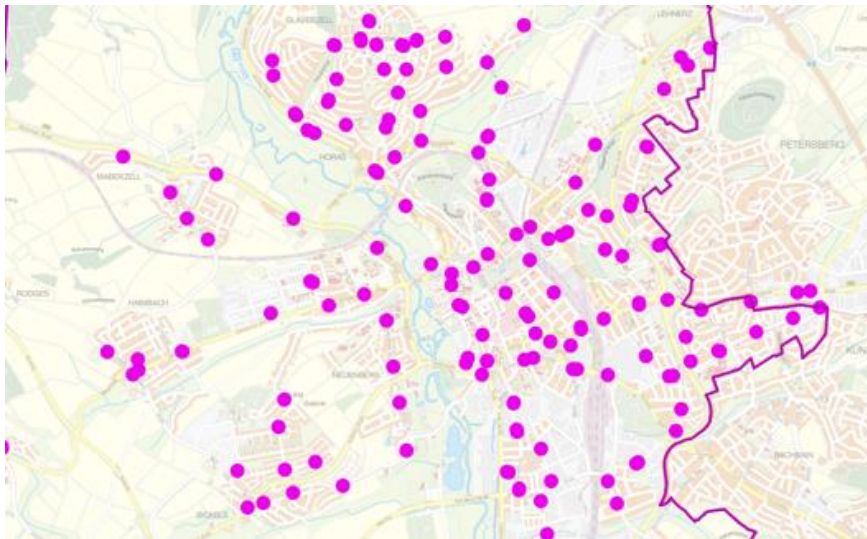


Abbildung 10: Ausschnitt aus der Grafik „ÖPNV-Haltestellen“ des Stadtgebietes.



7.7 Behandlung von außergesetzlichen Merkmalen

Außergesetzliche Merkmale sind Merkmale in Bezug auf die Wohnung oder das Mietverhältnis, die in § 558 Absatz 2 Satz 1 des Bürgerlichen Gesetzbuchs nicht genannt sind, aber dennoch für die Mietpreisbildung relevant sind oder im Erstellungsstadium des Mietspiegels relevant sein können. Außergesetzliche Merkmale können insbesondere zur Wahl des Regressionsmodells und bei der Bemessung von Spannen nach § 16 Absatz 3 MSV herangezogen werden (Bundesregierung 2021b). Als konkrete Zu- bzw. Abschläge im Mietspiegel dürfen außergesetzliche Merkmale nicht herangezogen werden.

Außergesetzliche Merkmale können die Vorhersagegüte und den Bias bei der Schätzung der Koeffizienten des Regressionsmodells beeinflussen. Eine generelle statistische Empfehlung, in welcher Form die außergesetzlichen Merkmale genutzt werden sollen, kann kaum gegeben werden (Kauermann und Windmann 2023).

Im Rahmen der Mietspiegelneuerstellung wurden erstmals auf Grundlage des § 2, Abs 1 i. V. m. § 14, Abs. 1 MSV auch außergesetzliche Merkmale bei der hier durchgeführten Mietspiegelerstellung untersucht. Hierbei stand als einziges Merkmal die „Mietdauer“ zur Verfügung. Weitere außergesetzliche Merkmale wie z. B. Geschlecht, sexuelle Orientierung, Einkommen usw. konnten aufgrund von datenschutzrechtlichen Aspekten nicht erhoben werden.

Bei der Analyse zeigte sich nur ein geringer Einfluss auf die Wohnfläche, das Baujahr sowie die Lageparameter im Modell. Ein signifikanter Einfluss dieser Merkmale auf Ausstattungskriterien wie z. B. Fußbodenheizung, Bodenbelag, Modernisierungszustand usw. konnte gemessen werden. Bei der Aufnahme in das Regressionsmodell konnte ein Anstieg des (korrigierten) Bestimmtheitsmaßes beobachtet werden. Darüber hinaus wurden weitere Gütekriterien wie z. B. die Streuung der Residuen oder die Veränderung verschiedener Qualitätsmaße der Schätzung geprüft. Weiterhin wurde der mittlere quadratische Fehler sowie die Standardabweichung, welche für die Spannenbildung herangezogen wird, untersucht. Auf dem Gesamtdatensatz, sowie auf Trainings- und Testdaten zeigte sich eine Verbesserung verschiedener Werte bei Hinzunahme der außergesetzlichen Merkmale. Daher fiel die Wahl auf ein Regressionsmodell, welches die Mietdauer als außergesetzliches Merkmal berücksichtigt.

7.8 Behandlung von Ausreißern

Um potenzielle Datensätze zu identifizieren, welche eine große Hebelwirkung (engl. „leverage“) auf die Schätzgenauigkeit und Güte des statistischen Modells haben, wird der Cook-Abstand berechnet und mit den standardisierten Residuen

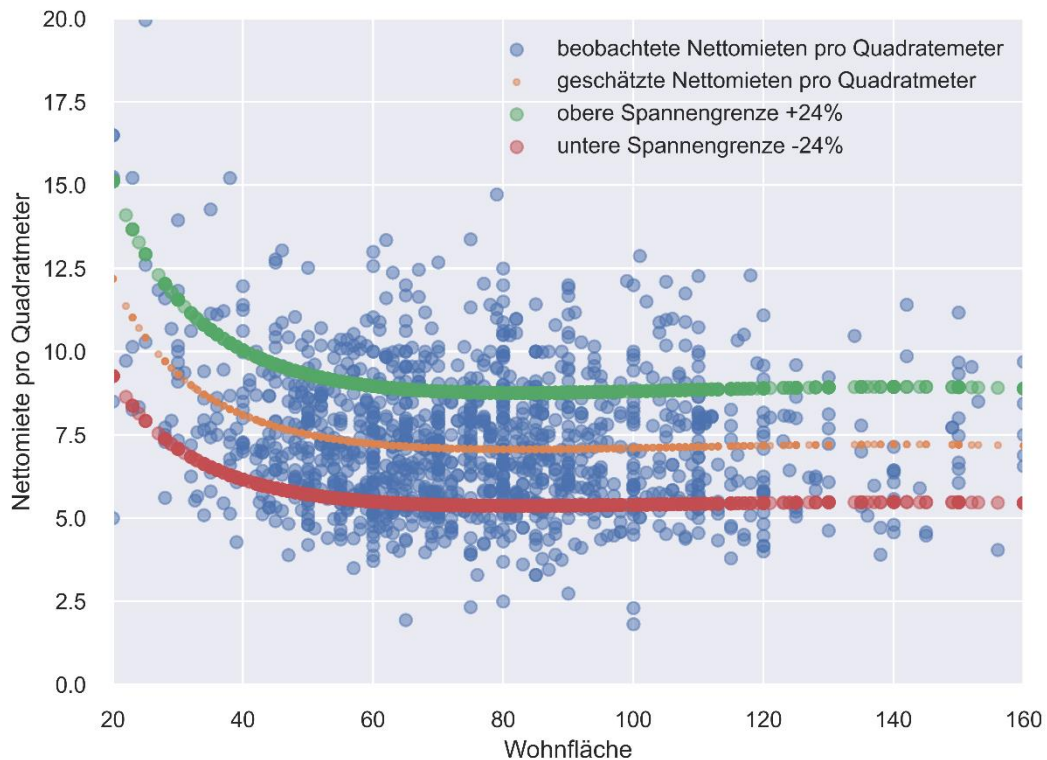
verglichen (James et al. 2017). Beobachtungen, welche größer als der zugehörige Cook-Abstand sind, werden als potenzielle Ausreißer betrachtet und können die Prädiktion („fit“) negativ beeinflussen (vgl. Abbildung 18). In der Praxis werden solche Beobachtungen als potenzielle Ausreißer identifiziert, welche einen Cook-Abstand größer $4/n$ aufweisen, wobei n die Gesamtanzahl aller Beobachtungen bezeichnet. Datensätze, welche eine große Hebelwirkung lt. Cook-Abstand aufweisen, werden nicht automatisch gelöscht. Es bleibt immer eine Einzelfallentscheidung, die zusammen mit der Plausibilität der erfassten Daten einhergeht.

7.9 Ermittlung von Spannbreiten

Mietspiegel sollen die örtlichen Wohnungsmarktstrukturen möglichst realitätsnah wiedergeben. Da die erhobenen Mieten auch innerhalb einer sehr genau definierten Wohnungsklasse streuen, wird zur Orientierung in vielen Mietspiegeln eine Spanne ausgewiesen, innerhalb der eine bestimmte Anzahl vergleichbarer Wohnungen liegt. Konventionell werden dafür sogenannte 2/3-Spannen verwendet, die auch vom Bundesbauministerium angegeben werden (Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) 2020). Das bedeutet, dass jeweils unter und über der ermittelten mittleren durchschnittlichen ortsüblichen Vergleichsmiete ein Drittel aller Mieten dieser Wohnungsklasse liegen sollen.

Dieser Spannenbildung entspricht die Berechnung eines $1-\sigma$ -Sicherheitsintervalls bei der hier gewählten regressionsanalytischen Vorgehensweise. Unter Verwendung aller genannten Wohnwertkriterien ergibt sich dabei eine durchschnittliche 2/3-Spanne in Höhe von ± 24 Prozent. Dies bedeutet mit anderen Worten: Bei Differenzierung nach verschiedenen Wohnwertmerkmalen liegen zwei Drittel aller Wohnungen dieser bestimmten Wohnungskategorie innerhalb der genannten Spannbreite.

Abbildung 11: Grafische Darstellung der 2/3-Spanne



Die Spannbreite beruht auf Mietpreisunterschieden, die durch den freien Markt (unterschiedliche Mieten für Mietobjekte mit gleichen Wohnwertmerkmalen) sowie subjektive (z.B. Wohndauer, freundschaftliche Beziehung zwischen Mieter und Vermieter) bzw. nicht erfasste objektive Wohnwertmerkmale (z.B. Besonderheiten wie Sauna) bedingt sind.

Abweichungen nach oben oder unten von der in diesem Mietspiegel errechneten durchschnittlichen ortsüblichen Vergleichsmiete innerhalb der Spannbreite sind gemäß BGH - VIII ZR 227/10 - zu begründen. Zur Begründung können insbesondere nicht im Mietspiegel ausgewiesene Merkmale herangezogen werden. Es ist zu beachten, dass bei der Mietspiegelerstellung viele Wohnwertmerkmale erhoben und auf deren Mietpreiseinfluss analysiert wurden. Wohnwertmerkmale mit eindeutig nachweisbarem signifikantem Einfluss auf den Mietpreis sind in den Tabellen 1 und 2 des Mietspiegels jeweils mit ihrem durchschnittlichen Wert enthalten.

Im Zuge der Datenerhebung zu dem hier vorliegenden Mietspiegel wurden auch Merkmale abgefragt, welche keinen korrelativen bzw. signifikanten Einfluss auf die Nettokaltmiete hatten.

Tabelle 13: nicht signifikante Merkmale

- Gebäudetyp ist ein Einfamilienhaus, Doppelhaus, Einliegerwohnung, Mehrfamilienhaus
- zentrale oder dezentrale Warmwasserversorgung
- überwiegender Teil des Wohn-/Schlafbereichs, abgesehen von Flur/Bad ist mit Parkett-, Dielenholz-, Laminat-, Fliesen-, Kacheln- oder Vinylboden ausgestattet
- mindestens ein Wohnraum, Küche oder Bad ohne fest installierte Heizung
- Mietvertrag schließt Nutzung eines Gartens bzw. eines Gartenanteils durch eine oder mehrere Hausparteien mit ein
- keine zeitgemäße Elektroinstallation (z.B. nur eine Sicherung für Beleuchtung/Steckdosen bzw. Elektroherd, max. zwei Steckdosen pro Raum, keine FI-Schalter)
- Glasfaseranbindung vorhanden

Diese Merkmale sind bei einer Anwendung der 2/3-Spannbreite nicht zu berücksichtigen.

8 Schlussbemerkung

Die vorliegende Dokumentation beschreibt die Methodik und die Ergebnisse der Mietspiegelerstellung der Stadt Fulda zum Mietspiegel 2024. Der Mietspiegel beruht auf einer repräsentativen empirischen Erhebung, welche eigens zum Zweck der Mietspiegelerstellung durchgeführt werden. Durch das Offenlegen der einzelnen Arbeitsschritte der Erstellung und der statistischen Methodik werden die (statistischen) Anforderungen und damit die Wissenschaftlichkeit der Ergebnisse dargestellt.

Der Hauptzweck von Mietspiegeln liegt in der Ermittlung der ortsüblichen Vergleichsmiete. Er dient sowohl Vermietern als auch Mietern, deren Interessensverbänden, Wohnungsunternehmen, Maklern, der städtischen Verwaltung und nicht zuletzt den Gerichten und Sachverständigen, indem er eine zuverlässige, unverzerrte Übersicht über den Mietwohnungsmarkt vermittelt. Die größte Wirkung entfaltet ein Mietspiegel im vorprozessualen Bereich, indem er Anhaltspunkte für eine außergerichtliche Einigung zwischen den Mietvertragsparteien liefert. Durch diese Orientierungshilfe zur Mietpreisfestsetzung für alle am Wohnungsmarkt Interessierten werden viele gerichtliche Mietstreitigkeiten verhindert.

9 Literaturverzeichnis

Aigner, Konrad; Walter Oberhofer; Bernhard Schmidt (1993): Eine neue Methode zur Erstellung eines Mietspiegels am Beispiel der Stadt Regensburg. In: *Wohnungswirtschaft und Mietrecht WM* (1/2/93), S. 16–21.

Allison, Paul D. (2007): Missing data. [Nachdr.]. Thousand Oaks, Calif: Sage Publ (Sage university papers 07, Quantitative applications in the social sciences, 136).

Anaconda Software Distribution (2020): Anaconda Inc. In: *Anaconda Documentation*. Online verfügbar unter <https://docs.anaconda.com/>.

Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) (Hg.) (2020): Hinweise zur Erstellung von Mietspiegeln. *BBSR Sonderveröffentlichung*. Bonn.

Bundesregierung (2021a): Gesetz zur Reform des Mietspiegelrechts. Mietspiegelreformgesetz - MsRG. In: *Bundesgesetzblatt* (Teil 1, Nr. 53).

Bundesregierung (2021b): Verordnung über den Inhalt und das Verfahren zur Erstellung und zur Anpassung von Mietspiegeln sowie zur Konkretisierung der Grundsätze für qualifizierte Mietspiegel. Mietspiegelverordnung - MsV. In: *Bundesgesetzblatt*.

Cischinsky, Holger; Malottki, Christian von; Rodenfels, Markus (2014): „Repräsentativität“ im Mietspiegel – Stichprobenmethodische Anforderungen an qualifizierte und grundsicherungsrelevante Mietspiegel 67.

Fahrmeir, Ludwig (2016): Statistik. Der Weg zur Datenanalyse. 8. Aufl. 2016. Berlin, Heidelberg: Springer Spektrum (SpringerLink Bücher).

Fahrmeir, Ludwig; Kneib, Thomas; Lang, Stefan; Marx, Brian D. (2022): Regression. Models, methods and applications. Second edition. Berlin, Heidelberg: Springer (Springer eBook Collection).

Horvitz, D. G.; Thompson, D. J. (1952): A Generalization of Sampling Without Replacement From a Finite Universe. In: *Journal of the American Statistical Association* 47 (260), S. 663. DOI: 10.2307/2280784.

James, Gareth; Witten, Daniela; Hastie, Trevor; Tibshirani, Robert (2017): An introduction to statistical learning. With applications in R. Corrected at 8th printing. New York, Heidelberg, Dordrecht, London: Springer (Springer texts in statistics).

Kauermann; Windmann (2023): Die Berücksichtigung von außergesetzlichen Merkmalen bei der Mietspiegelerstellung - Kausalität versus Vorhersage. In: *Allgemeines statistisches Archiv : AStA : journal of the German Statistical Society*.

Kauermann, Göran; Windmann, Michael; Münnich, Ralf (2020): Datenerhebung bei Mietspiegeln: Überblick und Einordnung aus Sicht der Statistik. In: *Wirtschafts- und sozialstatistisches Archiv* 14 (2), S. 145–162. DOI: 10.1007/s11943-020-00272-x.

Little, Roderick J. A. (2012): *Statistical analysis with missing data*. 3. rev. ed. Chichester, West Sussex: Wiley Blackwell.

Lohr, Sharon L. (2022): *Sampling. Design and analysis*. Third edition. Boca Raton, London, New York: CRC Press Taylor & Francis Group (Chapman & Hall/CRC texts in statistical science).

Pedregosa, Fabian; Varoquaux, Gaël; Gramfort, Alexandre; Michel, Vincent; Thirion, Bertrand; Grisel, Olivier et al. (2011): Scikit-learn: Machine learning in Python. In: *Journal of machine learning research* 12 (Oct), S. 2825–2830.

Ralph B. D’Agostino (1971): An Omnibus Test of Normality for Moderate and Large Size Samples. In: *Biometrika* 58 (2), S. 341–348. Online verfügbar unter <http://www.jstor.org/stable/2334522>, zuletzt geprüft am 01.09.2022.

Raybaut, Pierre (2009): Spyder-documentation. In: *Available online at: python-hosted.org*.

Seabold, Skipper; Perktold, Josef (2010): statsmodels: Econometric and statistical modeling with python. In: 9th Python in Science Conference.

van Buuren, Stef (2019): *Flexible Imputation of Missing Data*, Second Edition. 2nd ed. Milton: CRC Press LLC (Chapman and Hall/CRC Interdisciplinary Statistics Ser). Online verfügbar unter <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5455460>.

Wooldridge, Jeffrey M. (2013): *Introductory econometrics. A modern approach*. 5th ed. Mason, Ohio: South-Western Cengage Learning. Online verfügbar unter <http://www.loc.gov/catdir/enhancements/fy1403/2012945120-b.html>.

10 Anhang

10.1 Tabellen und Grafiken

10.1.1 Tabellen und Grafiken des Regressionsmodells Phase 1

Abbildung 12: Modellprognose in der ersten Phase vor Varianz Anpassung.

WLS Regression Results						
=====						
Dep. Variable:	nmneu	R-squared:	0.571			
Model:	WLS	Adj. R-squared:	0.570			
Method:	Least Squares	F-statistic:	572.0			
Date:	Mon, 23 Oct 2023	Prob (F-statistic):	2.36e-236			
Time:	13:20:45	Log-Likelihood:	-8464.7			
No. Observations:	1294	AIC:	1.694e+04			
Df Residuals:	1290	BIC:	1.696e+04			
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

intercept	68.2282	42.319	1.612	0.107	-14.794	151.250
wflneu	5.8512	1.269	4.610	0.000	3.361	8.341
wflneu2	0.0081	0.011	0.720	0.472	-0.014	0.030
wflneu3	-2.967e-05	2.85e-05	-1.040	0.299	-8.57e-05	2.63e-05
=====						
Omnibus:	199.226	Durbin-Watson:	1.906			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	616.524			
Skew:	0.768	Prob(JB):	1.33e-134			
Kurtosis:	6.013	Cond. No.	1.51e+07			
=====						

Abbildung 13: Oben links zeigt die Normalverteilung der Residuen der Schätzung in Phase 1. Unten links zeigt die Streuung der Residuen. Oben rechts zeigt die Abweichung zwischen der erwarteten vs. der beobachteten kumulativen Wahrscheinlichkeit der Residuen aus der Schätzung der 1. Phase.

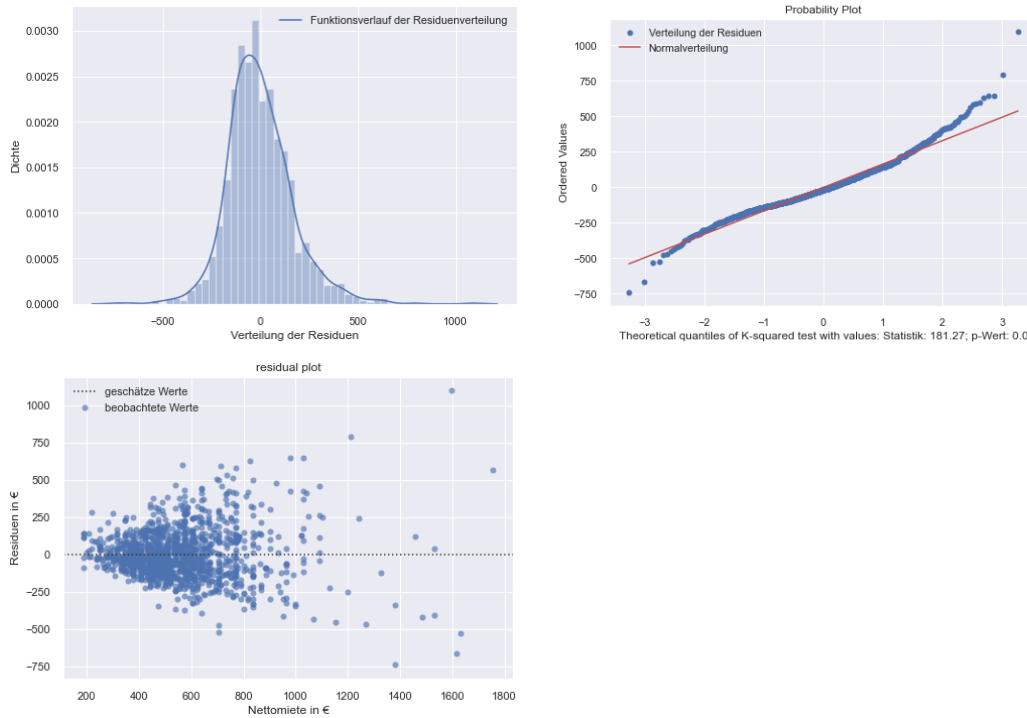


Abbildung 14: Modell der Varianzanpassung

Results: Weighted least squares						
=====						
Model:	WLS	Adj. R-squared:	0.023			
Dependent Variable:	nmst	AIC:	4233.1549			
Date:	2024-01-15 12:05	BIC:	4253.8169			
No. Observations:	1294	Log-Likelihood:	-2112.6			
Df Model:	3	F-statistic:	11.03			
Df Residuals:	1290	Prob (F-statistic):	3.74e-07			
R-squared:	0.025	Scale:	1.5379			

	Coef.	Std.Err.	t	P> t	[0.025	0.975]

abst	195.6804	28.8488	6.7830	0.0000	139.0847	252.2761
wflst	1.2321	1.1348	1.0857	0.2778	-0.9942	3.4584
wflst2	0.0550	0.0133	4.1501	0.0000	0.0290	0.0811
wflst3	-0.0002	0.0000	-3.5500	0.0004	-0.0003	-0.0001

Omnibus:	66.799	Durbin-Watson:	1.883			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	75.958			
Skew:	0.580	Prob(JB):	0.000			
Kurtosis:	3.248	Condition No.:	4733003			
=====						

Abbildung 15: Oben links zeigt die Normalverteilung der Schätzung für die Varianzkorrektur. Unten links zeigt die Streuung der Residuen für die Varianzkorrektur. Oben rechts zeigt die Abweichung zwischen der erwarteten vs. der beobachteten kumulativen Wahrscheinlichkeit der Residuen aus der Schätzung für die Varianzkorrektur sowie die Normalverteilung im Quantil-Quantil-Plot.

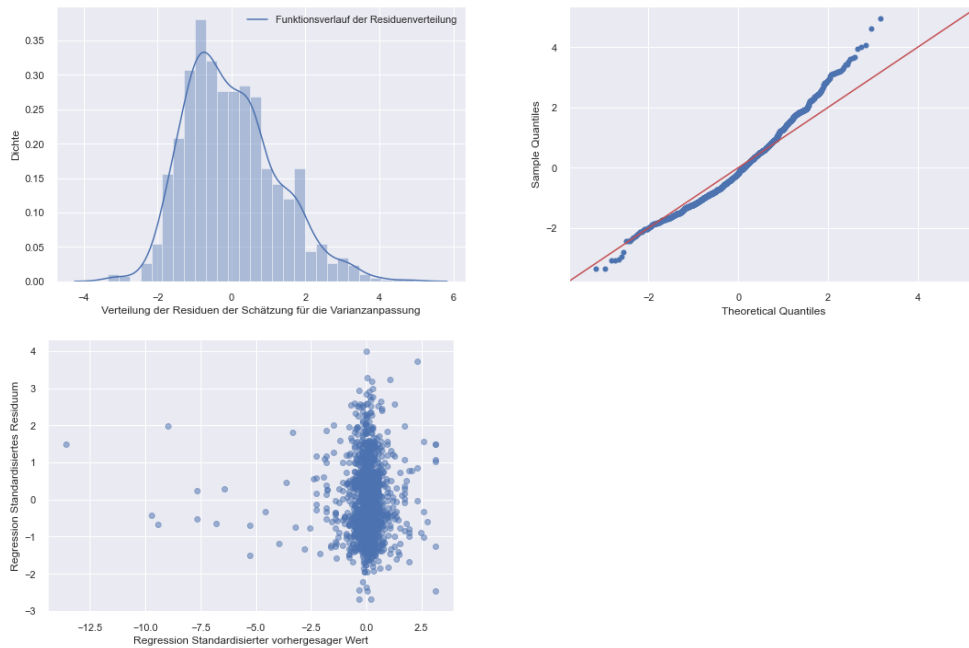


Abbildung 16: Die Analogon Plots der Nettomiete nach der Varianzkorrektur.

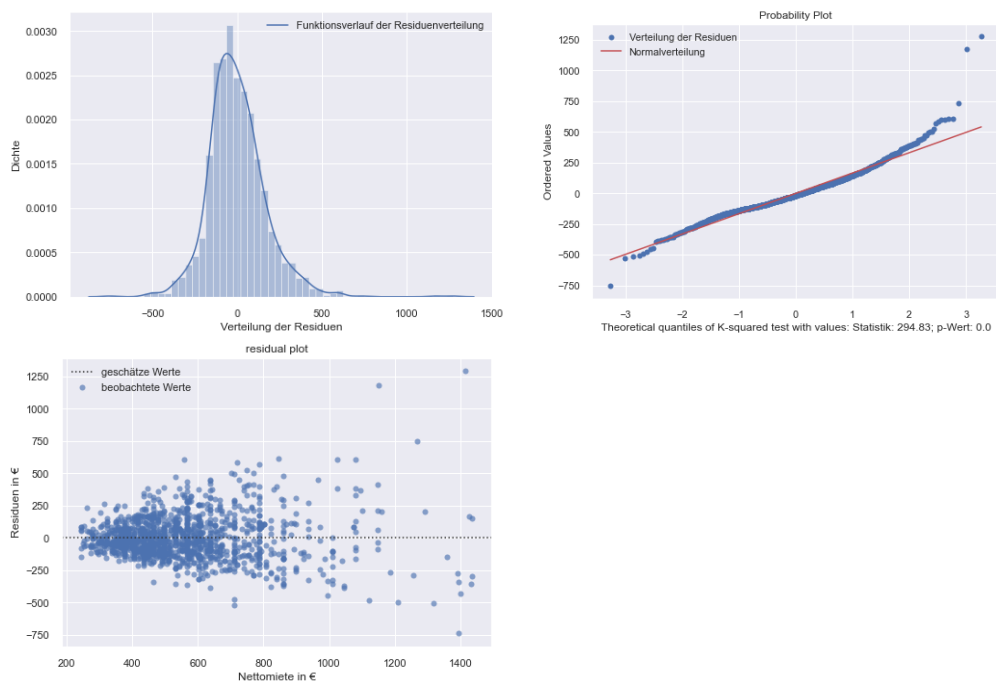
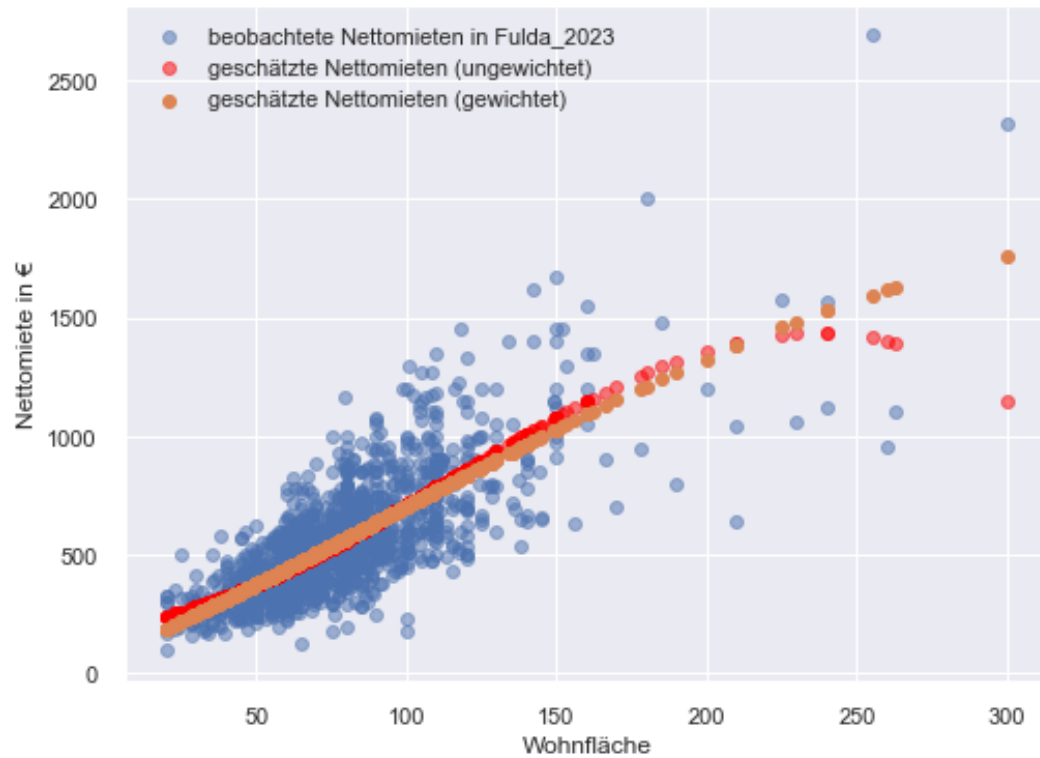


Abbildung 17: Vergleich der Schätzung vor und nach der Varianz Anpassung.



10.1.2 Tabellen und Grafiken des Regressionsmodells Phase 2

Weitere Validierungsinformationen des o. g. Regressionsansatzes in der zweiten Phase sind in nachfolgenden Grafiken dargestellt.

Abbildung 18: Die Grafiken auf der linken Seite zeigen die Verteilung der Residuen. Auf der rechten Seite oben zeigt der Quantil-Quantil-Plot die Normalverteilung der standardisierten Residuen. Unten rechts wird der Cook-Abstand berechnet, um potenzielle Datensätze zu identifizieren, welche eine große Hebelwirkung (engl. „leverage“) auf die Gesamtschätzung haben. In allen Grafiken sind die drei Datensätze markiert, welche demnach die höchste Hebelwirkung haben.

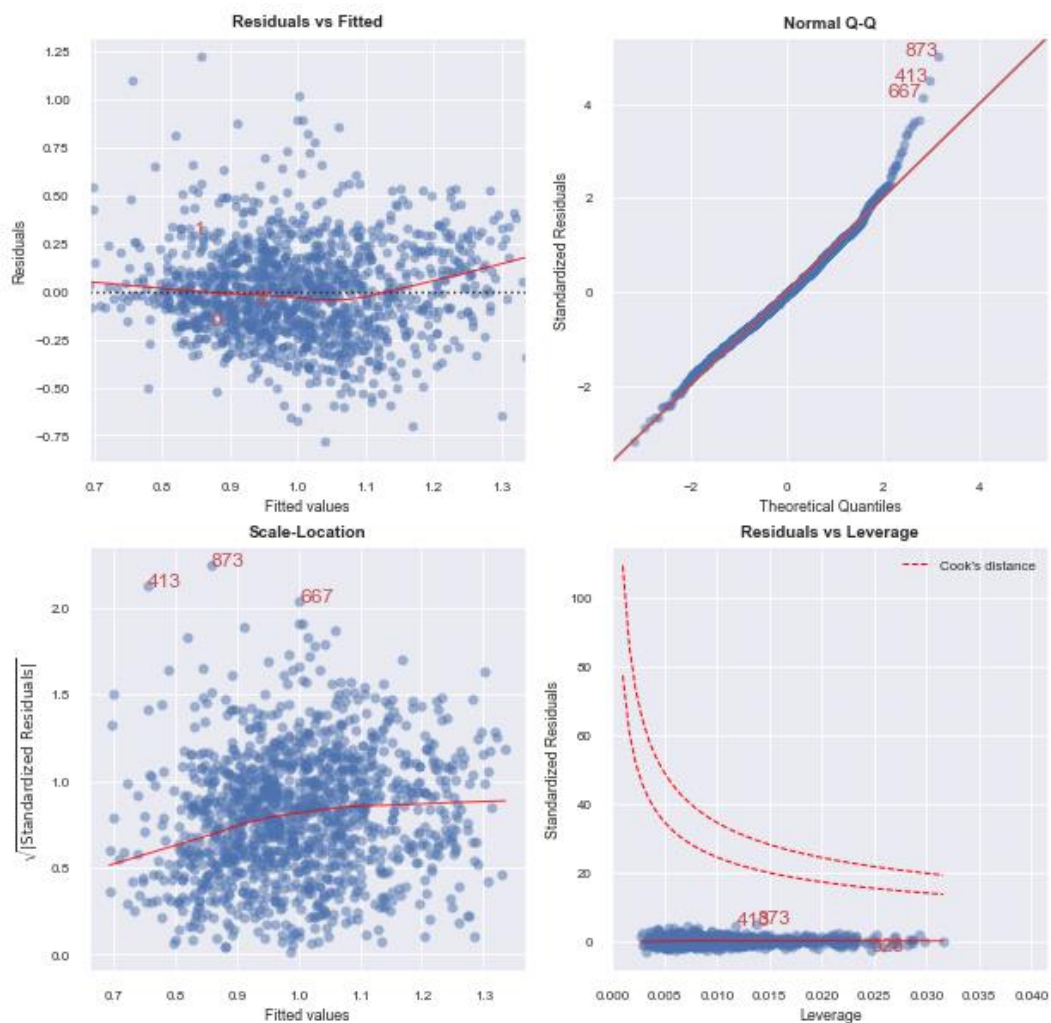


Tabelle 14: verwendete Geodatenattribute

Variablenname	Kurzbeschreibung	Beschreibung (Distanzen als Luftlinie gemessen)
Id_ema	Identifikator	Vom EMA-Institut bereitgestellte ID
commercial	Gewerbegebiet	Adresse ist nicht weiter als 100m von einem Gewerbegebiet entfernt
industry	Industriegebiet	Adresse ist nicht weiter als 100m von einem Industriegebiet entfernt
nature	Grünfläche	Adresse ist nicht weiter als 100m von einer größeren Grünfläche entfernt
dist_center	Distanz Stadtzentrum (Marktplatz) LU	Distanz zum Stadtzentrum in Metern LU
dist_playground	Distanz Spielplatz	Distanz zum nächsten Spielplatz in Metern
dist_school	Distanz Schule	Distanz zur nächsten Schule in Metern
dist_nursery	Distanz Kita	Distanz zur nächsten Kindertagesstätte in Metern
dist_supermarket	Distanz Supermarkt	Distanz zum nächsten Supermarkt in Metern
dist_leisure	Distanz Freizeiteinrichtung	Distanz zur nächsten Freizeiteinrichtung in Metern
dist_public_transport	Distanz ÖPNV	Distanz zur nächsten ÖPNV-Haltestelle in Metern
rel_building_area	Anteil Gebäudefläche	Anteil bebauter Fläche in der näheren Nachbarschaft (H3-Hexagon, ca. 15.000m ² Fläche)
brw	Bodenrichtwert	Bodenrichtwert in EUR
noise_rail_day	Lärm, Schienenverkehr, Tag	Lärm durch Schienenverkehr in Dezibel, tagsüber
noise_rail_night	Lärm, Schienenverkehr, Nacht	Lärm durch Schienenverkehr in Dezibel, nachts
noise_road_day	Lärm, Straßenverkehr, Tag	Lärm durch Straßenverkehr in Dezibel, tagsüber
noise_road_night	Lärm, Straßenverkehr, Nacht	Lärm durch Straßenverkehr in Dezibel, nachts

Tabelle 15: untersuchte Merkmale

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
1	nmneu	Nettomiete pro Monat	250	2690	-	1294	557,57
2	nmqm	Nettomiete pro Quadratmeter	1,81	19,96	-	1294	7,29
3	wflneu	Wohnfläche	20	300	-	1294	77,91
4	bjneue	Baujahr	1899	2023	-	1294	1967,88
5	B2a1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1247, 1: 47}	1294	-
6	B2a2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1224, 1: 70}	1294	-
7	B2a3	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 28, 0: 1266}	1294	-
8	B2a4	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 212, 1: 1082}	1294	-
9	B2b1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 913, 1: 381}	1294	-
10	B2b2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 768, 1: 526}	1294	-
11	B2b3	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1113, 1: 181}	1294	-
12	B2c1	Ausstattung: Merkmal lt. Fragebogen	0	2023	-	1294	914,17
13	B2c2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 814, 1.0: 480}	1294	-
23	B2e1	Ausstattung: Merkmal lt. Fragebogen	0	21	-	1294	3,64
24	B2f1	Ausstattung: Merkmal lt. Fragebogen	-1	14	-	1294	1,56
25	B2f2	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 59, 0.0: 1235}	1294	-
26	B2f3	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1081, 1.0: 213}	1294	-
27	B2f4	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1281, 1.0: 13}	1294	-
28	B3a2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 475, 1: 819}	1294	-
29	B3a3	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 264, 0: 1030}	1294	-
30	B3b1	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 957, 0: 337}	1294	-
31	B3b2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1077, 1: 217}	1294	-
32	B3c1	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 718, 0: 576}	1294	-

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
33	B3c2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 750, 1: 544}	1294	-
34	B3d1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1163, 1: 131}	1294	-
35	B3d2	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 985, 0: 309}	1294	-
36	B3e1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1280, 1.0: 14}	1294	-
37	B3e2	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 1153, 0.0: 141}	1294	-
38	B3e3	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1204, 1.0: 90}	1294	-
39	B3e4	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1045, 1.0: 249}	1294	-
40	B3f1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 497, 1.0: 797}	1294	-
41	B3f2	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 627, 0.0: 667}	1294	-
42	B3f3	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1130, 1.0: 164}	1294	-
43	B3f4	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 1053, 0.0: 241}	1294	-
44	B3f5	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 989, 0.0: 305}	1294	-
45	B3f6	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1065, 1.0: 229}	1294	-
46	B3f7	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1027, 1.0: 267}	1294	-
47	B3f8	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1263, 1.0: 31}	1294	-
48	B3f9	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1119, 1.0: 175}	1294	-
49	B3f10	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 353, 0.0: 941}	1294	-
50	B3f11	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1170, 1.0: 124}	1294	-
51	B3g1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1141, 1: 153}	1294	-
52	B3g2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1225, 1: 69}	1294	-
53	B3g3	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1276, 1: 18}	1294	-
54	B3g4	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1252, 1: 42}	1294	-
55	B3g5	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 902, 1: 392}	1294	-
56	B3g6	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 346, 0: 948}	1294	-

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
57	B3g7	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1173, 1: 121}	1294	-
58	B3g8	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1170, 1: 124}	1294	-
59	B3g9	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1256, 1: 38}	1294	-
60	B3h1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 582, 1.0: 712}	1294	-
61	B3h2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1077, 1.0: 217}	1294	-
62	B3h3	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1197, 1.0: 97}	1294	-
63	B3h4	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1212, 1.0: 82}	1294	-
64	B3h5	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1208, 1.0: 86}	1294	-
65	B3h6	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 104, 0.0: 1190}	1294	-
66	B3h7	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1171, 1.0: 123}	1294	-
67	B3h8	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1099, 1.0: 195}	1294	-
68	B3h9	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1039, 1.0: 255}	1294	-
69	B3h10	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 776, 0.0: 518}	1294	-
70	B3h11	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1207, 1.0: 87}	1294	-
71	B3h12	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 648, 1: 646}	1294	-
72	B3h13	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 451, 0: 843}	1294	-
73	B3i1	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 637, 0: 657}	1294	-
74	B3i2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 665, 1: 629}	1294	-
75	B3j1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1036, 1: 258}	1294	-
76	B3j2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 803, 1: 491}	1294	-
77	B3j3	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 421, 0: 873}	1294	-
78	B3j4	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1121, 1.0: 173}	1294	-
79	B3j5	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1119, 1.0: 175}	1294	-
80	B3j6	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1171, 1.0: 123}	1294	-

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
81	B3k1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 959, 1: 335}	1294	-
82	B3k2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 836, 1: 458}	1294	-
83	B3k3	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 335, 0: 959}	1294	-
84	B3l1	Ausstattung: Merkmal lt. Fragebogen	0	1	{0: 1217, 1: 77}	1294	-
85	B3l2	Ausstattung: Merkmal lt. Fragebogen	0	1	{1: 193, 0: 1101}	1294	-
86	B3m1	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 150, 0.0: 1144}	1294	-
87	B3m2	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1193, 1.0: 101}	1294	-
88	B3m3	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1147, 1.0: 147}	1294	-
89	B3m4	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1279, 1.0: 15}	1294	-
90	B3m5	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1135, 1.0: 159}	1294	-
91	B3m6	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1153, 1.0: 141}	1294	-
92	B3m7	Ausstattung: Merkmal lt. Fragebogen	0	1	{1.0: 86, 0.0: 1208}	1294	-
93	B3m8	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1204, 1.0: 90}	1294	-
94	B3m9	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1283, 1.0: 11}	1294	-
95	B3m10	Ausstattung: Merkmal lt. Fragebogen	0	1	{0.0: 1278, 1.0: 16}	1294	-
96	dist_public_transport	Lage: Geodaten	0	762	-	1294	143
97	dist_leisure	Lage: Geodaten	0	1214	-	1294	257,91
98	industry	Lage: Geodaten	0	384	-	1294	95,24
99	nature	Lage: Geodaten	0	532	-	1294	112,47
100	dist_center	Lage: Geodaten	0	6085	-	1294	1095,25
101	dist_nursery	Lage: Geodaten	0	3886	-	1294	361,29
102	dist_school	Lage: Geodaten	0	4058	-	1294	483,91
103	dist_playground	Lage: Geodaten	0	716	-	1294	231,12
104	dist_supermarket	Lage: Geodaten	0	4115	-	1294	428,58

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
105	rel_building_area	Lage: Geodaten	0	0,66	-	1294	0,22
106	brw	Lage: Geodaten	0	2700	-	1294	274,19
107	noise_road_lden_max	Lage: Geodaten	0	84	-	1294	46,33
108	noise_road_lngt_max	Lage: Geodaten	0	74	-	1294	41,82
109	bjklass	Baujahr: Baujahresklassen	1	9	-	1294	3,9
110	bj18	Baujahr: Baujahresklassen	0	1	{0: 1195, 1: 99}	1294	-
111	bj45	Baujahr: Baujahresklassen	0	1	{0: 1100, 1: 194}	1294	-
112	bj77	Baujahr: Baujahresklassen	0	1	{0: 796, 1: 498}	1294	-
113	bj84	Baujahr: Baujahresklassen	0	1	{1: 133, 0: 1161}	1294	-
114	bj95	Baujahr: Baujahresklassen	0	1	{0: 1195, 1: 99}	1294	-
115	bj04	Baujahr: Baujahresklassen	0	1	{0: 1216, 1: 78}	1294	-
116	bj12	Baujahr: Baujahresklassen	0	1	{0: 1245, 1: 49}	1294	-
117	bj018	Baujahr: Baujahresklassen	0	1	{0: 1226, 1: 68}	1294	-
118	bj023	Baujahr: Baujahresklassen	0	1	{0: 1218, 1: 76}	1294	-
119	dist_center0	Lage: Geodaten	0	1	{0: 1079, 1: 215}	1294	-
120	dist_center1	Lage: Geodaten	0	1	{0: 988, 1: 306}	1294	-
121	dist_center2	Lage: Geodaten	0	1	{1: 513, 0: 781}	1294	-
122	dist_playground1	Lage: Geodaten	0	1	{0: 915, 1: 379}	1294	-
123	dist_playground2	Lage: Geodaten	0	1	{1: 772, 0: 522}	1294	-
124	dist_playground3	Lage: Geodaten	0	1	{0: 1151, 1: 143}	1294	-
125	dist_school1	Lage: Geodaten	0	1	{0: 1134, 1: 160}	1294	-
126	dist_school2	Lage: Geodaten	0	1	{1: 859, 0: 435}	1294	-
127	dist_school3	Lage: Geodaten	0	1	{0: 1131, 1: 163}	1294	-
128	dist_nursery1	Lage: Geodaten	0	1	{0: 927, 1: 367}	1294	-

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
129	dist_nursery2	Lage: Geodaten	0	1	{1: 871, 0: 423}	1294	-
130	dist_nursery3	Lage: Geodaten	0	1	{0: 1238, 1: 56}	1294	-
131	dist_supermarket1	Lage: Geodaten	0	1	{0: 892, 1: 402}	1294	-
132	dist_supermarket2	Lage: Geodaten	0	1	{1: 837, 0: 457}	1294	-
133	dist_supermarket3	Lage: Geodaten	0	1	{0: 1239, 1: 55}	1294	-
134	dist_leisure1	Lage: Geodaten	0	1	{0: 916, 1: 378}	1294	-
135	dist_leisure2	Lage: Geodaten	0	1	{0: 441, 1: 853}	1294	-
136	dist_leisure3	Lage: Geodaten	0	1	{1: 63, 0: 1231}	1294	-
137	dist_public_transport1	Lage: Geodaten	0	1	{0: 869, 1: 425}	1294	-
138	dist_public_transport2	Lage: Geodaten	0	1	{1: 812, 0: 482}	1294	-
139	dist_public_transport3	Lage: Geodaten	0	1	{0: 1237, 1: 57}	1294	-
140	rel_building_area1	Lage: Geodaten	0	1	{0: 1195, 1: 99}	1294	-
141	rel_building_area2	Lage: Geodaten	0	1	{1: 1004, 0: 290}	1294	-
142	rel_building_area3	Lage: Geodaten	0	1	{0: 1115, 1: 179}	1294	-
143	wflneu2	Analyse Nettomiete	400	90000	-	1294	6993,02
144	wflneu3	Analyse Nettomiete	8000	27000000	-	1294	738797,67
145	nmd_r2_train	Analyse Nettomiete	239,73	1196,79	-	1294	559,01
146	nmd_mse_train	Analyse Nettomiete	211,71	1621,96	-	1294	550,24
147	nmd_r2_test	Analyse Nettomiete	255,59	2492,51	-	1294	567,34
148	nmd_mse_test	Analyse Nettomiete	129,94	1894,39	-	1294	564,03
149	intercept	Analyse Nettomiete	1	1	{1: 1294}	1294	-
150	nmd_wls	Analyse Nettomiete	188,27	1755,32	-	1294	559,09
151	nmqd_mse_train	Analyse Nettomiete	5,41	10,59	-	1294	7,2
152	nmqd_r2_train	Analyse Nettomiete	2,33	11,99	-	1294	7,38

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
153	nmqd_mse_test	Analyse Nettomiete	6,31	7,39	-	1294	7,28
154	nmqd_r2_test	Analyse Nettomiete	7,02	12,78	-	1294	7,36
155	nmqd_wls	Analyse Nettomiete	5,85	9,41	-	1294	7,29
156	nmd_wfl_bj_smooth	Analyse Nettomiete	100	2690	-	1294	557,57
157	sig2	Analyse Nettomiete	46,88	640,38	-	1294	124,82
158	nmst	Analyse Nettomiete	1,15	9,64	-	1294	4,55
159	abst	Analyse Nettomiete	0	0,02	-	1294	0,01
160	wflst	Analyse Nettomiete	0,43	0,65	-	1294	0,63
161	wflst2	Analyse Nettomiete	8,53	140,54	-	1294	48,76
162	wflst3	Analyse Nettomiete	170,65	42162,2	-	1294	4320,88
163	errs	Analyse Nettomiete	-754,88	1274,38	-	1294	-1,55
164	nmdh	Analyse Nettomiete	243,77	1435,94	-	1294	559,12
165	nmdhw	Analyse Nettomiete	243,77	1435,94	-	1294	559,12
166	nmqdh	Analyse Nettomiete	3,84	12,19	-	1294	7,31
167	nmf	Analyse Nettomiete	0,26	2,08	-	1294	1
168	nmfnorm	Analyse Nettomiete	-0,74	1,08	-	1294	0
169	nmf0	Analyse Nettomiete	-74,49	108,3	-	1294	-0,25
170	bj1neu	Analyse Baujahr	1899	2023	-	1294	1967,88
171	bj2neu	Analyse Baujahr	3606201	4092529	-	1294	3873517,74
172	bj3neu	Analyse Baujahr	6,85E+09	8,28E+09	-	1294	7626365997
173	bind2	Analyse Baujahr	0,93	1,23	-	1294	1
174	bindnorm2	Analyse Baujahr	-7,43	22,52	-	1294	-0,25
175	bindnorm21	Analyse Baujahr	-0,07	0,23	-	1294	0
176	bind3	Analyse Baujahr	0,93	1,23	-	1294	1

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
177	bindnorm3	Analyse Baujahr	-7,36	22,76	-	1294	-0,25
178	bindnorm31	Analyse Baujahr	-0,07	0,23	-	1294	0
179	bind	Analyse Baujahr	0,97	1,27	-	1294	1,04
180	bindnorm	Analyse Baujahr	-3,43	26,52	-	1294	3,75
181	bindnorm1	Analyse Baujahr	-0,03	0,27	-	1294	0,04
182	nmfbjklass	Analyse Baujahr	-7,58	21,26	-	1294	-0,25
183	nmqmd_wfl_bj_smooth	Analyse Baujahr	1,81	19,96	-	1294	7,29
184	md1	außergesetzliches Merkmal	0	1	{1: 686, 0: 608}	1294	-
185	md2	außergesetzliches Merkmal	0	1	{0: 993, 1: 301}	1294	-
186	md3	außergesetzliches Merkmal	0	1	{0: 1114, 1: 180}	1294	-
187	md4	außergesetzliches Merkmal	0	1	{1: 170, 0: 1124}	1294	-
188	md5	außergesetzliches Merkmal	0	1	{0: 1178, 1: 116}	1294	-
189	md6	außergesetzliches Merkmal	0	1	{0: 1223, 1: 71}	1294	-
190	mdklassen	außergesetzliches Merkmal	1	6	-	1294	2,48
191	efhs	Ausstattung: Interaktionsterm	0	1	{0: 1225, 1: 69}	1294	-
192	efh	Ausstattung: Interaktionsterm	0	1	{0: 1263, 1: 31}	1294	-
193	einlieger	Ausstattung: Interaktionsterm	0	1	{1: 28, 0: 1266}	1294	-
194	nmfefh	Ausstattung: Interaktionsterm	-0,53	0,7	-	1294	0
195	mfh	Ausstattung: Interaktionsterm	0	1	{0: 212, 1: 1082}	1294	-
196	egmfh	Ausstattung: Interaktionsterm	0	1	{1: 358, 0: 936}	1294	-
197	ug	Ausstattung: Interaktionsterm	0	1	{1: 53, 0: 1241}	1294	-
198	dg	Ausstattung: Interaktionsterm	0	1	{0: 1084, 1: 210}	1294	-
199	hg	Ausstattung: Interaktionsterm	0	1	{0: 1283, 1: 11}	1294	-
200	zh	Ausstattung: Interaktionsterm	0	1	{1: 957, 0: 337}	1294	-

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
201	dzh	Ausstattung: Interaktionsterm	0	1	{0: 1077, 1: 217}	1294	-
202	wwz	Ausstattung: Interaktionsterm	0	1	{1: 718, 0: 576}	1294	-
203	wwdz	Ausstattung: Interaktionsterm	0	1	{0: 750, 1: 544}	1294	-
204	fenein	Ausstattung: Interaktionsterm	0	1	{0: 1163, 1: 131}	1294	-
205	fenmehr	Ausstattung: Interaktionsterm	0	1	{1: 985, 0: 309}	1294	-
206	kbad	Ausstattung: Interaktionsterm	0	1	{0: 1280, 1: 14}	1294	-
207	bad1	Ausstattung: Interaktionsterm	0	1	{1: 1107, 0: 187}	1294	-
208	bad2	Ausstattung: Interaktionsterm	0	1	{0: 1226, 1: 68}	1294	-
209	wc2	Ausstattung: Interaktionsterm	0	1	{0: 1087, 1: 207}	1294	-
210	badu	Ausstattung: Interaktionsterm	0	1	{0: 1012, 1: 282}	1294	-
211	baodu	Ausstattung: Interaktionsterm	0	1	{1: 1103, 0: 191}	1294	-
212	bad_score	Ausstattung: Interaktionsterm	-1	10	-	1294	3,53
213	bad_score4	Ausstattung: Interaktionsterm	0	1	{1: 627, 0: 667}	1294	-
214	bad_score1	Ausstattung: Interaktionsterm	0	1	{0: 1130, 1: 164}	1294	-
215	bodschl20	Ausstattung: Interaktionsterm	0	1	{1: 347, 0: 947}	1294	-
216	ebk	Ausstattung: Interaktionsterm	0	1	{1: 637, 0: 657}	1294	-
217	kebk	Ausstattung: Interaktionsterm	0	1	{0: 665, 1: 629}	1294	-
218	balk	Ausstattung: Interaktionsterm	0	1	{0: 582, 1: 712}	1294	-
219	kbalk	Ausstattung: Interaktionsterm	0	1	{1: 582, 0: 712}	1294	-
220	terr	Ausstattung: Interaktionsterm	0	1	{0: 1077, 1: 217}	1294	-
221	balkuterr	Ausstattung: Interaktionsterm	0	1	{0: 1228, 1: 66}	1294	-
222	balkoterr	Ausstattung: Interaktionsterm	0	1	{0: 431, 1: 863}	1294	-
223	aufzuog4vor13	Ausstattung: Interaktionsterm	0	1	{0: 1255, 1: 39}	1294	-
224	aufzuog4ab13	Ausstattung: Interaktionsterm	0	1	{0: 1243, 1: 51}	1294	-

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
225	aufzuog4	Ausstattung: Interaktionsterm	0	1	{0: 1204, 1: 90}	1294	-
226	barrierearm	Ausstattung: Interaktionsterm	0	1	{0: 1171, 1: 123}	1294	-
227	garagemv	Ausstattung: Interaktionsterm	0	1	{0: 1176, 1: 118}	1294	-
228	stellplatzmv	Ausstattung: Interaktionsterm	0	1	{0: 1251, 1: 43}	1294	-
229	fbhab13	Ausstattung: Interaktionsterm	0	1	{0: 1250, 1: 44}	1294	-
230	fbhvor13	Ausstattung: Interaktionsterm	0	1	{0: 1256, 1: 38}	1294	-
231	fbh	Ausstattung: Interaktionsterm	0	1	{0: 1212, 1: 82}	1294	-
232	gartenmv	Ausstattung: Interaktionsterm	0	1	{0: 1039, 1: 255}	1294	-
233	raeume	Ausstattung: Interaktionsterm	0	1	{1: 776, 0: 518}	1294	-
234	kbalkoterr	Ausstattung: Interaktionsterm	0	1	{1: 431, 0: 863}	1294	-
235	kheizraum	Ausstattung: Interaktionsterm	0	1	{0: 1197, 1: 97}	1294	-
236	install	Ausstattung: Interaktionsterm	0	1	{0: 1208, 1: 86}	1294	-
237	kkeller	Ausstattung: Interaktionsterm	0	1	{1: 104, 0: 1190}	1294	-
238	kelekt	Ausstattung: Interaktionsterm	0	1	{0: 1207, 1: 87}	1294	-
239	mod_score	Ausstattung: Interaktionsterm	0	8	-	1294	0,71
240	kmod60	Ausstattung: Interaktionsterm	0	1	{0: 1131, 1: 163}	1294	-
241	kmod8060	Ausstattung: Interaktionsterm	0	1	{0: 1190, 1: 104}	1294	-
242	mod_score490	Ausstattung: Interaktionsterm	0	1	{0: 1214, 1: 80}	1294	-
243	vollmod90	Ausstattung: Interaktionsterm	0	1	{0: 1235, 1: 59}	1294	-
244	vollteilmod90	Ausstattung: Interaktionsterm	0	1	{1: 234, 0: 1060}	1294	-
245	teilmod90	Ausstattung: Interaktionsterm	0	1	{1: 175, 0: 1119}	1294	-
246	vollteilmod90sumw	Ausstattung: Interaktionsterm	0	0,75	-	1294	0,07
247	equip_pos_sum1	Ausstattung: Interaktionsterm	0	3	-	1294	0,23
248	eqpSumW	Ausstattung: Interaktionsterm	0	7	-	1294	1,47

#	Variablenname	Kurzbeschreibung, Kategorie	Minimum	Maximum	Häufigkeit	Anzahl	Mittelwert
249	eqnSumW	Ausstattung: Interaktionsterm	0	5	-	1294	0,95
250	micro_sum_pos	Ausstattung: Interaktionsterm	0	7	-	1294	1,86
251	micro_sum_neg	Ausstattung: Interaktionsterm	0	2	-	1294	0,12
252	micro_score_pos	Ausstattung: Interaktionsterm	0	5	-	1294	1,15
253	micro_score_neg	Ausstattung: Interaktionsterm	0	4	-	1294	0,26
254	brw500	Ausstattung: Interaktionsterm	0	1	{0: 1169, 1: 125}	1294	-

10.2 Fragebogen

Fragebogen zur Erstellung des qualifizierten Mietspiegels 2023
für die Stadt **Fulda**

Id: _____

Fragebogen für Mieter¹

Zugangsschlüssel: _____

Bitte beachten Sie:

- Gemäß § 2 Abs. 1 Mietspiegelreformgesetz (MsRG) sind Sie zur Ausfüllung verpflichtet.
- Geben Sie den Fragebogen an den **Hauptmieter** der angeschriebenen Wohnung weiter.
- Vergleichen Sie die Angaben mit den Mietunterlagen (Mietvertrag/Betriebskostenabrechnung).
- Streichen Sie keine Kästchen durch, lassen Sie sie LEER.
- Senden Sie den Fragebogen mit beigefügtem **Freiumschlag kostenlos** zurück oder nutzen Sie alternativ auch die Online-Antwortmöglichkeit unter:

<https://survey.ema-institut.de/index.php/951398>



Bitte den Fragebogen entweder auf Papier **oder** über das Internet beantworten.

Rücksendung / Beantwortung bitte bis: 10. Juli 2023

Bei Fragen wenden Sie sich an die Telefonhotline 0941 / 38 07 10

oder per E-Mail an: support@ema-institut.de

A	FILTERFRAGEBOGEN	
→ Hinweis: Nachfolgende Fragen (A1 bis A7) stellen die Mietspiegelrelevanz fest.		
A1	Handelt es sich bei der angeschriebenen Wohnung um selbstgenutztes Eigentum?	☐ ja (Ende) ☐ nein
A2	Ist die Wohnung Teil eines Wohnheimes, einer sozialen Einrichtung oder einer Sammelunterkunft (z.B. Studenten-, Jugend-, Alten-, Pflege-, Personalwohnheim, vorläufige Unterbringung/Anschlussunterbringung (Geflüchtete), Behinderteneinrichtung, „Betreutes Wohnen“, soziale Wohngruppe)?	☐ ja (Ende) ☐ nein
A3	Wird der Wohnraum mietfrei oder verbilligt überlassen (z.B. Dienst- oder Werkswohnung, Wohnung gehört Verwandten)?	☐ ja (Ende) ☐ nein
A4	Handelt es sich bei der Wohnung um Wohnraum, der öffentlich gefördert ist oder anderen Preisbindungen unterliegt (z.B. Sozialwohnung mit Wohnberechtigungsschein)?	☐ ja (Ende) ☐ nein
A5	Wird die Wohnung ganz oder überwiegend möbliert vermietet (einzelne Möbelstücke sowie Einbauküche und Einbauschränke zählen nicht als Möblierung)?	☐ ja (Ende) ☐ nein
A6	Ist die Wohnung ganz oder teilweise gewerblich genutzt oder nur zu vorübergehendem Gebrauch vermietet (max. drei Monate, z.B. Ferienwohnung)?	☐ ja (Ende) ☐ nein
A7	Handelt es sich bei der Wohnung um eine nicht abgeschlossene Wohnung oder um ein Einzelzimmer, das Teil einer kompletten Wohnung ist?	☐ ja (Ende) ☐ nein
→ Hinweis: Bitte nur weiter ausfüllen, wenn <u>alle</u> Fragen (A1 bis A7) mit „<u>nein</u>“ beantwortet wurden. Falls Sie <u>eine</u> der Filterfragen mit „<u>ja</u>“ beantwortet haben, ist die Befragung für Sie an dieser Stelle beendet. Bitte senden Sie den Fragebogen im beigefügten, voradressierten Freiumschlag kostenlos zurück.		

¹ Aus Gründen der besseren Lesbarkeit wird bei Personenbezeichnungen und personenbezogenen Hauptwörtern in diesem Dokument die männliche Form verwendet. Entsprechende Begriffe gelten im Sinne der Gleichbehandlung grundsätzlich für alle Geschlechter. Die verkürzte Sprachform hat nur redaktionelle Gründe und beinhaltet keine Wertung.

B		HAUPTFRAGEBOGEN	
B1		Mietpreis, Betriebskosten, Mietverhältnis	
B1a	Ist die Aufteilung der monatlichen Mietzahlung zum 01.06.2023 bekannt?	1 ☐ Ja (bitte nachfolgend entsprechend aufteilen)	
B1b		1 , € monatliche Bruttogesamtmiete (gesamte Mietzahlung inkl. Betriebskosten und Zuschläge, Pauschalbetrag)	
		2 , € monatliche Nettokaltemiete (ohne Nebenkosten und ohne Zuschläge)	
		3 , € monatliche Nebenkosten/Betriebskosten (-vorauszahlung)	
		4 , € monatlicher Zuschlag (z. B. Einbauküche, Untervermietung usw.)	
		5 , € monatlicher Mietanteil für (Tief-)Garage	
B1c		separater Mietvertrag? 1 ☐ ja 2 ☐ nein	
B1d		6 , € monatlicher Mietanteil für Pkw-Abstellplatz separater Mietvertrag? 1 ☐ ja 2 ☐ nein	
		2 ☐ Nein (bitte die Bruttogesamtmiete nachfolgend angeben)	
		7 , € monatliche Bruttogesamtmiete (gesamte Mietzahlung inkl. Betriebskosten und Zuschläge, Pauschalbetrag)	
B1e	Wann ist der Haushalt in die Wohnung eingezogen?	1 Monat Jahr 2	
B1f	Wann ist die Nettokaltemiete das letzte Mal geändert bzw. angepasst worden (z.B. wegen Mieterhöhung)?	1 Monat Jahr 2	
B2		Art und Alter des Gebäudes und der Wohnung	
B2a	In welchem Gebäudetyp wohnen Sie?	1 ☐ Einfamilienhaus 2 ☐ Doppelhaushälfte, Reihenhäuser 3 ☐ Einliegerwohnung 4 ☐ Mehrfamilienhaus, mit wie vielen Wohnungen pro Hauseingang (siehe Anzahl der Klingeln am Hauseingang): → 1 ☐ 2-4 2 ☐ 5-10 3 ☐ ab 11	
B2b			
B2c	Wann wurde das Gebäude, in dem die Wohnung liegt, ursprünglich errichtet?	1 Baujahr 2 ☐ unbekannt → HINWEIS: Maßgebend ist das Jahr der Bezugsfertigkeit. War die Wohnung im Krieg zerstört, gilt das Jahr des Wiederaufbaus. Liegt die Wohnung in einem aufgestockten oder angebauten Gebäudeteil, gilt das Jahr der Aufstockung bzw. des Anbaus; bei Ausbau des Dachgeschosses gilt das Jahr des Ausbaus. Modernisierung/Sanierung beeinflussen nicht das Baujahr	
B2d	EINE ANTWORT: Falls Ihnen das Baujahr unbekannt ist, ordnen Sie es bitte in eine der folgenden Zeiträume ein:	1 ☐ bis 1918 4 ☐ 1978 - 1984 7 ☐ 2005 - 2012 2 ☐ 1919 - 1945 5 ☐ 1985 - 1995 8 ☐ 2013 - 2018 3 ☐ 1946 - 1977 6 ☐ 1996 - 2004 9 ☐ 2019 - 2023	
B2e	Wie viele Wohnebenen hat das Gebäude? → HINWEIS: Erdgeschoss zählt als eine Ebene!	1 Anzahl der Wohnebenen	
B2f	In welcher Ebene liegt Ihre Wohnung? → HINWEIS: Erdgeschoss = 0, erste Etage = 1, zweite Etage = 2, usw.	1 Ebene 2 ☐ Untergeschoss/Souterrain 3 ☐ Dachgeschosswohnung 4 ☐ Hanggeschoss	
B3		Größe, Ausstattung und sonstige Beschaffenheit der Wohnung	
B3a	Wie groß ist die Wohnfläche der Wohnung? → HINWEIS: Zur Wohnfläche zählen wie im Mietvertrag angegeben alle Hauptwohnräume, Küche, Bad/WC, Abstellraum in der Wohnung. Balkone, Loggien, Terrassen zählen mit einem Viertel bis der Hälfte ihrer Grundfläche, unbeheizbare Wintergärten, Schwimmbäder zur Hälfte.	1 , m ² , gemäß 2 ☐ schriftlicher Unterlagen (z.B. Mietvertrag, Betriebskostenabrechnung) 3 ☐ Angabe ohne schriftlichen Beleg	

B3j	Wurde die Wohnung bzw. das Gebäude durch bauliche Maßnahmen von Seiten des Vermieters nachträglich gedämmt?	1 ☐ ja 2 ☐ nein 3 ☐ unbekannt	Falls ja: 4 ☐ Dämmung Innen- und/oder Außenwand 5 ☐ Dämmung Dach/oberstes Geschoss 6 ☐ Dämmung Kellerdecke
B3k	Wurde die Wohnung bzw. das Gebäude seit 2013 durch weitere bauliche Maßnahmen von Seiten des Vermieters modernisiert/saniert, die zu einer wesentlichen Gebrauchs-werterhöhung im Vergleich zum ursprünglichen Zustand der Wohnung führten? (Nicht gemeint sind die üblichen Instandhaltungs- und Renovierungsarbeiten)	1 ☐ ja 2 ☐ nein 3 ☐ unbekannt	(weiter mit Frage B3l) (Ende) (Ende)
B3l	Falls ja: Welche der folgenden Modernisierungsmaßnahmen wurden seit 2013 durchgeführt? → Hinweis: *Schwellen bis max. 4cm Höhe	1 ☐ Vollsanierung (mit einem Neubau vergleichbaren Zustand der Wohnung zum Modernisierungszeitpunkt)	
B3m		2 ☐ einzelne Modernisierungsmaßnahmen (neuwertiger Zustand zum Modernisierungszeitpunkt), nämlich:	
		1 ☐ Sanitärbereich (mind. Fliesen, Wanne, WC) erneuert 2 ☐ Elektroinstallation (zeitgemäß) erneuert 3 ☐ Heizanlage/Warmwasserversorgung erneuert 4 ☐ Schallschutz eingebaut 5 ☐ Fußböden erneuert	6 ☐ Fenster-/Rahmenerneuerung 7 ☐ Innen- und Wohnungstüren erneuert 8 ☐ Treppenhaus, Eingangsbereich erneuert 9 ☐ Grundriss verbessert 10 ☐ barrierearme Ausstattung geschaffen (Mindestvoraussetzung: schwellenfrei*, stufenloser Zugang, bodengleiche Dusche)