

Rheinstraße 65
64295 Darmstadt
Germany

Fon: +49(0)6151/2904-0
Fax: +49(0)6151/2904-97

info@iwu.de
www.iwu.de

Datenerhebung Wohngebäudebestand 2016

Datenerhebung zu den energetischen Merkmalen und Modernisierungsraten im deutschen und hessischen Wohngebäudebestand

Anleitung zur Durchführung von Auswertungen mit der Auswertungsdatenbank



Darmstadt, 17.04.2018

Autoren: Dr. Holger Cischinsky
Dr. Nikolaus Diefenbach
Markus Rodenfels

Inhalt

1	Einführung	4
2	Merkmalsbeschreibung	4
3	Logische Ausdrücke in der Auswertungsdatei	6
4	Aufbau der Auswertungsdatei und Durchführung der Auswertungen.....	7
Anhang A:	Fragebogenvariablen.....	17
Anhang B:	Rohfassung der Auswertungsdatei.....	21

1 Einführung

Im Rahmen des Forschungsprojekts „Datenerhebung Wohngebäudebestand 2016 – Datenerhebung zu den energetischen Merkmalen und Modernisierungsraten im deutschen und hessischen Wohngebäudebestand“ hat das Institut Wohnen und Umwelt (IWU) eine deutschlandweite Repräsentativbefragung der Eigentümerinnen und Eigentümer von Gebäuden mit Wohnraum durchgeführt. Die aus dem Projekt entstandene Auswertungsdatenbank liegt am IWU vor. Sie soll auch Dritten für eigene Auswertungszwecke zugänglich gemacht werden, und zwar auf dem Wege der kontrollierten Datenfernverarbeitung. Die Rahmenbedingungen hierfür werden in einem mit dem IWU abzuschließenden Vertrag festgelegt (vgl. Mustervertragstext auf der Homepage www.wohngebaeuedaten2016.iwu.de).

Dem Vertragspartner, hier „Anwender“ genannt, wird auf der Homepage die Vorlage für eine Auswertungsdatei zur Verfügung gestellt. Diese hat das Format einer Textdatei (.txt). Der Anwender erstellt daraus und unter Beachtung der in der vorliegenden Anleitung dokumentierten Syntax seine Auswertungsdatei und sendet sie per E-Mail an das IWU (wohngebaeuedaten@iwu.de). Die Datei muss weiterhin das Format einer Textdatei haben. Andere Formate (z.B. Word, Excel) sind nicht zulässig.

Das IWU verwendet die Auswertungsdatei als Eingabedatei für eine mit der Statistik-Software R (<http://www.r-project.org>) durchgeführte Auswertung der Datenbank. Anschließend wird dem Anwender eine Ergebnisdatei zurückgesendet. Diese hat ebenfalls das Format einer Textdatei.

Die Methodik des Forschungsprojekts und detaillierte Informationen zu seiner Durchführung sowie zahlreiche Auswertungsergebnisse sind im Projektendbericht dokumentiert¹. Demgegenüber liefert dieser Bericht Informationen, die für eigenständige Auswertungen relevant sind, und stellt daher für den Anwender eine wichtige Arbeitsgrundlage zur Erstellung seiner Auswertungsroutinen dar. Fundierte Kenntnisse und Erfahrungen des Anwenders bei der statistischen Auswertung von Befragungsdaten werden dabei vorausgesetzt.

2 Merkmalsbeschreibung

Die Auswertungsdatenbank enthält 16.982 Datensätze, von denen jeder für ein anderes Gebäude mit Wohnraum steht, das als Untersuchungsgebäude in die Datenerhebung eingebunden war.

Bei den Merkmalen der Auswertungsdatenbank handelt es sich um aufbereitete² Fragebogenangaben („Fragebogenmerkmale“) und um zugespielte ergänzende Merkmale.

Die Fragebogenmerkmale mitsamt ihren Ausprägungen sind im Anhang A dargestellt. Beispielsweise steht das Merkmal *GebArt* für die Art des Gebäudes und die Ausprägung 1 für Wohngebäude.

Ließ der Befragte eine Frage unbeantwortet oder wurden seine Angaben im Zuge der Datenaufbereitung als unplausibel gewertet, weist das entsprechende Fragebogenmerkmal die Ausprägung 0 auf. Zu beachten ist dabei, dass der Grund für eine unbeantwortete Frage nicht notwendigerweise darin liegen muss, dass der Befragte eine eigentlich zu beantwortende Frage aus welchen Gründen auch immer unbeantwortet ließ (Item Nonresponse). Vielmehr kann eine Frage auch deshalb – richtigerweise – unbeantwortet sein, weil sie im Einklang mit der Fragebogenführung gar nicht beantwortet werden sollte.

Besondere Vorsicht ist in diesem Kontext bei metrischen Merkmalen geboten, und zwar dahingehend, dass die Ausprägung 0 im Regelfall nicht quantitativ interpretiert werden darf. Beispielsweise bedeutet die Ausprägung 0 beim Fragebogenmerkmal *Baujahr* nicht, dass das Gebäude im Jahr 0, also zu Christi Geburt, errichtet wurde, sondern dient ausschließlich zur Kennzeichnung fehlender oder unplausibler Fragebogenangaben. Auch bei nachgeordneten Feldern, in die der Befragte im Bedarfsfall eine Zahl einzutragen hatte, bedeutet die Ausprägung 0 zunächst einmal nur, dass keine bzw. keine plausible Eintragung vorliegt. Immer dann, wenn die Eintragung bedingt durch die Beantwortung der Vorschaltfrage richtigerweise unterblieb,

¹ Vgl. Cischinsky, H.; Diefenbach, D.: Datenerhebung Wohngebäudebestand 2016. Datenerhebung zu den energetischen Merkmalen und Modernisierungsraten im deutschen und hessischen Wohngebäudebestand, Darmstadt, 2018.

² Zum Verfahren der Datenaufbereitung vgl. Projektendbericht, Unterkapitel 2.3.3.

kann die Ausprägung 0 allerdings auch quantitativ, d. h. als Menge 0, interpretiert werden. Hat ein Befragter beispielsweise die Frage nach zu Wohnzwecken vermieteten (oder mietfrei überlassenen) Wohnungen im Untersuchungsgebäude verneint ($WhgMiet = 1$) und folgerichtig keine Angabe zur Anzahl der Mietwohnungen ($nWhgMiet$) gemacht, kann der im Zuge der Datenaufbereitung zugewiesene Wert 0 auch quantitativ interpretiert werden, d. h. als Zahl der Mietwohnungen im Gebäude, die Null beträgt. Der Umkehrschluss, also die Folgerung aus $nWhgMiet = 0$, dass es im Gebäude keine Mietwohnungen gibt, ist jedoch unzulässig, denn es ist nicht ausgeschlossen, dass ein Befragter überhaupt keine Angaben zu Mietwohnungen in seinem Gebäude gemacht hat.

Eine besondere Art von „Null-Ausprägungen“ kann bei Mehrfachantwortfeldern vorkommen, von denen es im Fragebogen insgesamt sechs gibt und die dort durch quadratische Ankreuzfelder gekennzeichnet sind³. Hier bedeutet die Merkmalsausprägung 0 lediglich, dass das Kästchen nicht angekreuzt wurde⁴, wohingegen die Merkmalsausprägung 1 bei angekreuztem Kästchen vergeben wurde.

Auf einen kleinen Teil der auf dem Fragebogen enthaltenen Informationen kann aus Datenschutzgründen nicht zu Auswertungszwecken zurückgegriffen werden. Dazu gehören die Erhebungsgruppe, der Kommunnennamen⁵ sowie handschriftliche Eintragungen unter „Sonstiges, und zwar“.

Über die Fragebogenmerkmale hinaus stehen folgende ergänzende Merkmale zur Verfügung:

Merkmalsbezeichnung	Beschreibung	Wertelabels
<i>Gebiet</i>	Gebietsaggregat auf Basis der Bundesländer*	1:alte Bundesländer Nord (Schleswig-Holstein, Niedersachsen, Bremen, Hamburg, Nordrhein-Westfalen) 2:alte Bundesländer Süd (Baden-Württemberg, Bayern, Hessen, Saarland, Rheinland-Pfalz) 3:neue Bundesländer (Berlin, Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, Thüringen)
<i>Hessen</i>	Bundesland Hessen	1:ja 2:nein (außerhalb Hessens)
<i>wachsindklss</i>	Wachstumsindikator der Kommune / des Grundsteuerstellengebiets**	1:nicht wachsend (schrumpfend oder keine Entwicklungsrichtung erkennbar) 2:wachsend
<i>bevklss</i>	Einwohnergrößenklasse der Kommune / des Grundsteuerstellengebiets am Jahresende 2013	1:weniger als 50.000 Einwohner 2:50.000 Einwohner oder mehr
<i>nWE</i>	Anzahl der Wohnungen im Gebäude (abgeleitet aus nWE1 und nWE2)	
<i>Geb</i>	Zählmerkmal: Für jeden Gebäudesatz gleich 1	1:Gebäudesatz

* Auf Nachfrage sind bei Vorliegen bestimmter Voraussetzungen auch bundeslandspezifische Auswertungen möglich.

** Für nähere Informationen zu diesem Merkmal vgl. Fußnote 59 im Projektendbericht

Bei wohngebäudespezifischen Auswertungen entlang der geografischen Merkmale *wachsindklss* und *bevklss* muss eine Übereinstimmung mit Wohngebäudebeständen, wie sie aus der amtlichen Bautätigkeitsstatistik abgeleitet werden können, nicht zwingend gegeben sein. Bei Auswertungen für Wohnheime und sonstige Gebäude mit Wohnraum gilt diese Aussage darüber hinaus auch für die Merkmale *Gebiet* und *Hessen*. Generell gilt, dass eine Übereinstimmung mit amtlichen Eckdaten nur in den der vorgenommenen Anpassungsrechnung zugrunde gelegten Strukturen sichergestellt ist⁶.

Grundsätzlich ist bei der Auswertung von (nicht notwendigerweise regional abgegrenzten) Teilmengen des Gebäudebestandes die Reduzierung der Fallzahlen, also die Verkleinerung der Zahl verwendbarer Gebäu-

³ Betroffen sind die Merkmale *FassPutzNeu*, *FassFarbNeu*, *VerteilHZNeu*, *VerteilWWNeu*, *PHFoerdKfW* und *PHFoerdSonst*.

⁴ Ggf. wurde im Zuge der Datenaufbereitung auch ein gesetztes Kreuz aufgrund Unplausibilität entfernt.

⁵ Im Musterfragebogen in Anhang A wurde als Platzhalter für den jeweiligen Kommunnennamen „Musterstadt“ eingesetzt.

⁶ Zur Anpassungsrechnung vgl. Projektendbericht, Unterkapitel 2.3.3, insbesondere Tabelle 10 (für Wohngebäude).

dedatensätze, zu beachten. Dieser Umstand äußert sich unter sonst gleichen Bedingungen in steigenden Standardfehlern und kann dazu führen, dass Auswertungen nicht mehr sinnvoll möglich sind – bis zum Extremfall, dass Auswertungsroutinen aus Datenschutzgründen erst gar nicht ausgeführt werden (sofern sie sich auf weniger als fünf Gebäudedatensätze stützen).

Zu beachten ist ferner, dass alle zur Verfügung stehenden Merkmale der Auswertungsdatenbank zunächst einmal nur auf Gebäude abstellen. In vielen Fällen lässt sich die entsprechende Merkmalsausprägung aber auch auf alle Wohnungen des jeweiligen Untersuchungsgebäudes übertragen, d. h. das Gebäudemerkmal lässt sich auch wohnungsbezogen interpretieren. Liegt ein Untersuchungsgebäude beispielsweise in einem der nördlichen alten Bundesländer, trifft dies in gleicher Weise auf die Wohnungen des betreffenden Untersuchungsgebäudes zu. Das betreffende Merkmal *Gebiet* lässt sich daher nicht nur gebäude-, sondern auch wohnungsbezogen interpretieren. Anders verhält es sich dagegen zum Beispiel mit der überwiegenden Beheizungsart im Gebäude (*HeizArt*). Hat der Befragte als überwiegende Beheizungsart beispielsweise „Zentralheizung“ angegeben, bedeutet dies im Fall eines Gebäudes mit mehr als einer Wohnung nicht notwendigerweise, dass alle Wohnungen des Gebäudes gleichermaßen zentralbeheizt sind, denn möglicherweise verfügt ein kleinerer Teil der Wohnungen auch über andere Beheizungsarten (insbesondere Etagenheizung).

3 Logische Ausdrücke in der Auswertungsdatei

Zur Erstellung der Auswertungsdatei greift der Anwender auf die zur Verfügung stehenden Auswertungsmerkmale zurück. Diese Merkmale kann er in Form logischer Ausdrücke miteinander verknüpfen. Logische Ausdrücke können den Wert TRUE (wahr) oder FALSE (falsch) annehmen.

Zur Definition der logischen Ausdrücke kann – entsprechend der Syntax des Statistik-Software R (<http://www.r-project.org>) – auf die folgenden Operatoren zurückgegriffen werden:

<u>Symbol</u>	<u>Bedeutung</u>
==	gleich (Achtung: Es ist das doppelte Gleichheitszeichen zu verwenden.)
!=	ungleich
>	größer als
>=	größer gleich
<	kleiner als
<=	kleiner gleich
!	nicht (Negation)
&	und
	oder (inklusives „oder“, d. h. nicht-ausschließende Disjunktion)

Bei komplizierteren Ausdrücken sind entsprechende Klammern zu setzen. Dabei werden runde Klammern verwendet: (). Bei den Merkmalsbezeichnungen ist die Groß- und Kleinschreibung zu beachten.

Beispiele:

Alle Datensätze mit einem angegebenen Gebäudebaujahr von 2005 oder später werden durch den folgenden logischen Ausdruck beschrieben:

Baujahr >= 2005

Der nachfolgende Ausdruck steht für Gebäude mit ein- oder zweischaligem Mauerwerk und ohne Solarthermieanlage, für die die Frage nach der Außenwanddämmung beantwortet wurde:

```
(AWTyp == 1 | AWTyp == 2) & SolTh == 1 & AWDaemmung > 0
```

An dieser Stelle sei darauf verwiesen, dass Merkmale, deren zugrunde liegende Fragen die beiden Antwortmöglichkeiten „nein“ und „ja“ vorsehen, im Fall „nein“ stets die Ausprägung 1 und im Fall „ja“ die Ausprägung 2 annehmen. Entgegen der Intuition beschreibt also der Ausdruck

```
SolTh == 1
```

nicht etwa Gebäude mit, sondern ohne Solarthermieanlage.

4 Aufbau der Auswertungsdatei und Durchführung der Auswertungen

Zur Erstellung der Auswertungsdatei ist die Syntax der Statistik-Software R (<http://www.r-project.org>) zu verwenden. Auswertungsbefehle werden zeilenweise eingegeben. Groß- und Kleinschreibung ist zu beachten. Kommentarzeilen, die beim Programmlauf ignoriert werden sollen, müssen mit dem Rautezeichen (#) eingeleitet werden.

Die Rohfassung der Auswertungsdatei, die als txt-Datei unter www.wohngebaeuedaten2016.iwu.de zum Download zur Verfügung steht, ist in Anhang B dokumentiert. Diese dient als Grundlage für die im Folgenden zur Erläuterung angeführten Beispielcodes.

Die Auswertungsdatei ist in mehrere Abschnitte unterteilt:

Titel

Die Auswertungsdatei trägt folgenden Titel, der vom Anwender nicht verändert werden darf:

```
#####
#
# Auswertungsdatei für die Durchführung eines Auswertungsgangs mit der
# Auswertungsdatenbank zum Projekt „Datenerhebung Wohngebäudebestand 2016“
#
#####
```

Header

Im auf den Titel folgenden Header macht der Anwender (im Beispiel: Max Mustermann) die geforderten Angaben:

```
#####
###
###Bitte folgende Angaben vervollständigen:
###
###Dateiname des Programmcodes: Musterprogramm.txt
###erstellt: 01.07.2018
###von: Max Mustermann, Musterinstitut GmbH
###E-Mail: max@mustermann.de
###Tel.: 01234/56789
###
#####
```

Teil A: Kopfzeile

Die Kopfzeile enthält den Pfadnamen zur Auswertungsdatenbank. Die Kopfzeile darf vom Anwender nicht verändert werden.

```
#####
###Teil A: Kopfzeile
###Diese Kopfzeile darf nicht verändert werden
#####
source("U:\\R\\headerkdfv.r")
```

Teil B: Festlegung von Teilmengen des Gebäudebestandes

Der Anwender legt fest, für welche Teilmenge(n) des Gebäudebestandes er Auswertungen vornehmen möchte. Für die Anzahl der definierten Teilmengen gibt es dabei keine Beschränkung.

Zur Definition der Teilmengen ist auf die in Kapitel 3 aufgeführten logischen Ausdrücke zurückzugreifen.

Jede Teilmenge wird vom Anwender durch einen festzulegenden Variablennamen definiert. Als Variablennamen dürfen keine Bezeichnungen von Fragebogenmerkmalen sowie von ergänzenden Merkmalen (vgl. Tabelle in Kapitel 2) verwendet werden. Der Variablenname darf ferner nicht mit einer Zahl beginnen und außer dem Unterstrich („_“) keine Sonderzeichen enthalten. Für die Zuweisung des Variablennamens ist folgender Ausdruck zu verwenden: <-

Nach der Zuweisung des Variablennamens steht der Befehl *subset*. In der sich anschließenden Klammer muss zunächst die Obermenge spezifiziert werden. Danach folgt der logische Ausdruck, durch den die betreffende Teilmenge festgelegt wird. Was die Obermenge angeht, muss bei der als erstes festgelegten Teilmenge zwingend *gebaeude2016* stehen, wodurch die Obermenge aller 16.982 Gebäudedatensätze der Auswertungsdatenbank angesprochen wird. Bei danach festgelegten Teilmengen kann anstelle von *gebaeude2016* auch eine andere, zuvor definierte Teilmenge als Obermenge herangezogen werden. Möchte der Anwender stets auf alle 16.982 Gebäudedatensätze zurückgreifen und daher nur Auswertungen für die Gesamtheit aller Gebäude mit Wohnraum in Deutschland durchführen, ist die Festlegung von Teilmengen nicht erforderlich. Bei der Durchführung von Auswertungen (vgl. Teil C) spricht er alle 16.982 Gebäudedatensätze mit *gebaeude2016* an.

Im nachfolgenden Beispiel legt der Anwender fünf Teilmengen fest:

```
#####
###Teil B: Festlegung von Teilmengen des Gebäudebestandes
#####

wogeb <- subset(gebaeude2016, GebArt == 1)
wogebefzh <- subset(wogeb, nWE1 == 1 | nWE1 == 2)
Altbau78 <- subset(gebaeude2016, Baujahr > 0 & Baujahr <= 1978)
wogebheizart <- subset(wogeb, HeizArt > 0)
wogebfoerdermittel <- subset(wogeb, PHFoerdKfW == 1 | PHFoerdSonst == 1)
```

Für die erste Teilmenge, die alle Wohngebäude (*GebArt == 1*) umfasst, hat der Anwender im Beispiel den Variablennamen *wogeb* gewählt.

Zur zweiten Teilmenge (*wogebefzh*) gehören nur Wohngebäude mit einer oder zwei Wohnungen. Die zweite Teilmenge ist somit eine Teilmenge von *wogeb*.

Für die dritte Teilmenge wurde die Bezeichnung *Altbau78* verwendet. Zu dieser Teilmenge gehören alle Gebäude mit Wohnraum, die dem Befragten zufolge vor 1979 gebaut wurden (*Baujahr <= 1978*). Die dritte Teilmenge ist keine Teilmenge von *wogeb*, da sie auch sonstige Gebäude mit Wohnraum und Wohnheime

einschließt. Zu beachten ist ferner, dass sich Auswertungen für diese Teilmenge sinnvollerweise nur auf solche Untersuchungsgebäude stützen, die verwertbare Angaben zum Baujahr aufweisen. Hierzu dient die Bedingung $Baujahr > 0$, die alle Untersuchungsgebäude ohne verwertbare Baujahresangabe ausschließt, denn für diese Gebäude wurde im Zuge der Datenaufbereitung das Baujahr auf 0 gesetzt (vgl. auch Kapitel 2).

Die vierte Teilmenge (*wogebheizart*) stellt auf Wohngebäude ab, für die verwertbare Angaben zur überwiegenden Beheizungsart (*HeizArt*) vorliegen.

Bei der fünften und letzten Beispiel-Teilmenge handelt es sich erneut um eine Teilmenge der Wohngebäude, und zwar um solche Wohngebäude, die nach dem 1.1.2010 errichtet wurden sowie einen KfW-Effizienzhaus- oder Passivhausstandart erreichen, wofür Fördermittel (aus einem KfW-Programm und/oder andere Fördermittel) in Anspruch genommen wurden. Zu beachten ist an dieser Stelle, dass die Eigenschaften „nach dem 1.1.2010 errichtet“ ($Baujahr \geq 2010$), „Erreichung eines KfW-Effizienzhaus- oder Passivhausstandards“ ($PH == 2$) sowie „Inanspruchnahme von Fördermitteln“ ($PHFoerd == 2$) nicht explizit vorgegeben werden müssen, da dank der Datenaufbereitung sichergestellt ist, dass nur solche Untersuchungsgebäude $PHFoerdKfW == 1$ und/oder $PHFoerdSonst == 1$ aufweisen können, die diese drei Eigenschaften erfüllen.

Durch die entsprechende Verknüpfung und Verschachtelung mittels logischer Operatoren lassen sich nicht nur beliebige Teilmengen, sondern prinzipiell auch beliebig kleine Teilmengen definieren. Die Reduzierung der Fallzahlen geht unter sonst gleichen Bedingungen jedoch mit einer Vergrößerung der Standardfehler einher, was die Aussagefähigkeit der für diese Teilmengen abgeleiteten Auswertungsergebnisse erheblich einschränken kann. Werden Teilmengen mit weniger als fünf Gebäudedatensätzen gebildet, erhält der Anwender aus Datenschutzgründen keine Auswertungsergebnisse.

Teil C: Durchführung der Auswertungen

Es dürfen bis zu 25 Auswertungen vorgenommen werden. Weiterhin gilt die im Vorabsatz bereits erwähnte Einschränkung, dass Teilmengen, für die weniger als fünf Gebäudedatensätze zur Verfügung stehen, aus Datenschutzgründen nicht ausgewertet werden; in diesen Fällen wird das Auswertungsergebnis auf die Zahl „-999“ gesetzt. Zur Generierung der Auswertungsergebnisse werden die jeweils eingebundenen Gebäudedatensätze „automatisch“ (d. h. ohne Zutun des Anwenders) entsprechend ihres jeweiligen Hochrechnungsfaktors gewichtet. Die Auswertungsergebnisse sind daher repräsentativ für die jeweils zugrunde liegende Teilmenge der Grundgesamtheit. Das Verfahren zur Bestimmung der Hochrechnungsfaktoren ist im Projektendbericht beschrieben (vgl. Unterkapitel 2.3.3 bzw. Anhang A).

Eine Auswertung besteht aus einer Programmzeile mit Anwendung jeweils eines der vier Befehle *MW*, *Anz*, *MW_W* oder *Anz_W*.

MW liefert einen (geschätzten) Mittelwert über alle Gebäude der zu analysierenden Teilmenge und *Anz* die (hochgerechnete) Anzahl der zur Teilmenge gehörenden Gebäude. *MW_W* und *Anz_W* sind die wohnungsbezogenen Pendanten zu *MW* und zu *Anz*, generieren also Ergebnisse für die zweite Art von Untersuchungseinheiten der Stichprobenerhebung, nämlich die in den Gebäuden mit Wohnraum liegenden Wohnungen⁷.

Die Syntax ist für alle vier Funktionen gleich und wird im Folgenden anhand von Beispielen erläutert.

Im ersten Beispiel wird die mittlere Anzahl der Wohnungen für die Beispiel-Teilmenge der Ein-/Zweifamilienhäuser (nur Wohngebäude), die zuvor in Teil B definiert wurde (siehe oben), bestimmt. Dabei ist zu beachten, dass unmittelbar vor das auszuwertende Merkmal (hier: *nWE*) stets das Tildezeichen (~) zu setzen ist.

MW(~nWE, wogebefzh)

⁷ Vgl. hierzu insbesondere Unterkapitel 2.1 des Projektendberichts.

Das Ergebnis wird im R-Format ausgegeben und lautet hier:

```
mean SE
nWE 1.24604 0.00933
```

Der gesuchte Mittelwert des Fragebogenmerkmals *nWE* beträgt demnach 1,24604 und der (geschätzte) Standardfehler (SE) 0,00933. Zur Interpretation des Standardfehlers wird auf den Projektendbericht (insbesondere Unterkapitel 2.3.3) sowie auf die einschlägige Fachliteratur verwiesen.

Gegenstand des zweiten Beispiels ist eine wohnungsbezogene Auswertung. Bestimmt werden soll das mittlere Gebäudebualter von Wohnungen, die sich in bis 1978 gebauten Gebäuden mit Wohnraum befinden.

```
MW_W(~Baujahr, Altbau78)
```

Als Ergebnis erhält man:

```
mean SE
Baujahr 1941.85 1.23844
```

Zieht man für jede Wohnung, die sich in bis 1978 gebauten Gebäuden mit Wohnraum befinden, das Gebäudebualter heran und mittelt über alle Wohnungen, erhält man demnach als Mittelwert in etwa das Jahr 1942 (genau: 1941,85) bei einer Standardabweichung (SE) von 1,23844. Zu beachten ist an dieser Stelle, dass der erwähnte Mittelwert nicht notwendigerweise dem mittleren Bualter von Wohnungen entspricht, die sich in bis 1978 gebauten Gebäuden mit Wohnraum befinden. Ursächlich hierfür ist der Umstand, dass das Fragebogenmerkmal *Baujahr* ein gebäudebezogenes Merkmal darstellt und somit alle Wohnungen in einem Untersuchungsgebäude dieselbe Merkmalsausprägung aufweisen, also auch solche Wohnungen, die durch eine nachträgliche Aufstockung erst Jahre nach der Gebäudeerrichtung entstanden und somit jüngeren Datums sind⁸.

Mithilfe der Befehle *MW* und *MW_W* lassen sich für nicht-metrische, also nominal- oder ordinalskalierte Merkmale auch Anteilsschätzungen durchführen. Für dichotome Merkmale, die nur die beiden Ausprägungen 1 für „zutreffend“ und 0 für „nicht zutreffend“ aufweisen, ist eine Merkmalsanteilsschätzungen unter Verwendung der bisher vorgestellten Syntax unmittelbar möglich. Ein Beispiel für ein solches dichotomes Merkmal ist *PHFoerdSonst*. Dieses Merkmal stellt darauf ab, ob bei Neubauten mit Baujahr ab 2010 für die Erreichung eines KfW-Effizienzhaus- oder Passivhausstandards Fördermittel in Anspruch genommen wurden, die nicht bzw. nicht ausschließlich aus einem KfW-Programm stammen. Die Ausprägung 1 wird angenommen, wenn solche anderen Fördermittel eingesetzt wurden. Andernfalls nimmt das Merkmal die Ausprägung 0 an, wobei diese Ausprägung auch vergeben wird, wenn die Frage nach der Art der Fördermittel überhaupt nicht relevant war (z. B. weil das Gebäude vor 2010 erstellt wurde) oder der Befragte die Frage unbeantwortet ließ, obwohl er die Vorschaltfrage nach der Inanspruchnahme von Fördermitteln bejaht hat. Beschränkt man sich auf Wohngebäude mit verwertbaren Angaben zur Inanspruchnahme von Fördermitteln (*wogebfoerdermittel*) und will wissen, wie groß unter den Neubauten mit Baujahr ab 2010, die einen geförderten KfW-Effizienzhaus- oder Passivhausstandard aufweisen, der Anteil derjenigen Gebäude ist, die – ungeachtet des Einsatzes von KfW-Fördermittel – andere Fördermittel in Anspruch genommen haben, lautet der Befehl wie folgt:

⁸ Vgl. hierzu auch die Ausführungen am Ende von Kapitel 2 dieser Anleitung.

```
MW(~PHFoerdSonst, wogebfoerdermittel)
```

Als Ergebnis erhält man:

```

      mean  SE
PHFoerdSonst 0.148085 0.1034

```

Demnach liegt der gesuchte Anteil bei (ca.) 14,8 %.

Mit Ausnahme der Mehrfachantwortfelder (wozu auch die Variable *PHFoerdSonst* gehört) weisen die nominal- oder ordinalskalierten Merkmale der Auswertungsdatenbank nicht die Ausprägungen 1 und 0 auf, sind also nicht in dieser Form dichotom. Um in solchen Fällen Merkmalsanteilsschätzungen durchführen zu können, ist der „factor“-Befehl anzuwenden. Am Beispiel des nominalskalierten Fragebogenmerkmals *HeizArt* soll der Aufbau des Befehls sowie der sich einstellende Output erläutert werden. Die Untersuchungsmenge soll sich dabei aus allen Wohngebäuden (*wogeb*) zusammensetzen.

```
MW(~factor(HeizArt), wogeb)
```

Der Ergebnisoutput stellt sich wie folgt dar:

```

      mean  SE
factor(HeizArt)0 0.0220456 0.00233
factor(HeizArt)1 0.0635432 0.00574
factor(HeizArt)2 0.0174849 0.00256
factor(HeizArt)3 0.7906222 0.00816
factor(HeizArt)4 0.0558230 0.00410
factor(HeizArt)5 0.0504811 0.00428

```

Für 2,2 % der Wohngebäude liegt demnach keine verwertbare Angabe zur überwiegenden Beheizungsart vor. 6,4 % der Wohngebäude sind (überwiegend) fernwärmebeheizt, 1,7 % haben eine Block- und 79,1 % eine Zentralheizung. 5,6 % der Wohngebäude sind überwiegend mit Etagenheizungen ausgestattet, während 5,0 % der Wohngebäude überwiegend raumweise beheizt werden.

Der Ausdruck *~factor(HeizArt)* hat somit eine Zerlegung (Faktorisierung) des Fragebogenmerkmals *HeizArt* in seine sechs Ausprägungen erzeugt und dadurch ermöglicht, dass für jede dieser Ausprägungen der geschätzte Anteil ausgegeben wird.

Nun ist „keine verwertbare Angabe“ ganz offensichtlich keine sinnvolle Ausprägung zur Beschreibung der Beheizungsstruktur von Wohngebäuden. Es ist deshalb vernünftig, diese Ausprägung aus der Auswertung herauszuhalten. Unter der Annahme, dass sich die Beheizungsstruktur derjenigen (hochgerechneten) Untersuchungsgebäude, die keine verwertbare Angaben aufweisen, nicht von derjenigen der Kompletärmenge unterscheidet, bietet es sich an, die beschriebene Auswertung nur auf Wohngebäude anzuwenden, bei denen verwertbare Angaben zur überwiegenden Beheizungsart vorhanden sind. Dies führt zu folgendem Befehl, bei dem gegenüber dem vorangegangenen Befehl die zu analysierende Teilmenge nicht mehr *wogeb* ist, sondern *wogebheizart*:

```
MW(~factor(HeizArt), wogebheizart)
```

Die Ausführung des Befehls führt zu folgendem Ergebnis:

	mean	SE
factor(HeizArt)1	0.0649756	0.00588
factor(HeizArt)2	0.0178791	0.00261
factor(HeizArt)3	0.8084449	0.00791
factor(HeizArt)4	0.0570814	0.00421
factor(HeizArt)5	0.0516190	0.00438

Demnach stellt sich die Beheizungsstruktur der Wohngebäude wie folgt dar: 6,5 % der Wohngebäude sind (überwiegend) fernwärmebeheizt, 1,8 % haben eine Block- und 80,8 % eine Zentralheizung. 5,7 % der Wohngebäude sind überwiegend mit Etagenheizungen ausgestattet, während 5,2 % der Wohngebäude überwiegend raumweise beheizt werden. Alle Anteile sind verglichen mit der vorangegangenen Auswertung gestiegen, und zwar zusammengenommen um den Anteil der Wohngebäude ohne verwertbare Angaben zur Beheizungsart (2,2 %).

Alternativ lässt sich diese Struktur auch dadurch ableiten, dass man den Anteilswert jeder „gültigen“ Beheizungsart-Ausprägung durch die Summe aller Anteilswerte der „gültigen“ Beheizungsart-Ausprägungen dividiert, womit man ebenfalls eine proportionale Aufteilung des Anteilswerts für Wohngebäude ohne verwertbare Angaben zur Beheizungsart (2,2 %) auf die „gültigen“ Beheizungsart-Ausprägungen erreicht. Der Standardfehler lässt sich dabei in sehr guter Näherung durch die Anwendung des Gaußschen Fehlerfortpflanzungsgesetzes schätzen⁹.

An dieser Stelle sei nochmals auf die beiden Vorgehensweisen innewohnende Annahme verwiesen, wonach sich die Beheizungsstruktur der (hochgerechneten) Gebäude ohne verwertbare Angaben nicht von derjenigen der Gebäude mit gültigen Ausprägungen unterscheidet. Ob der Anwender einer solchen Annahme folgt oder stattdessen eine andere Annahme trifft, liegt in seinem Ermessen.

Statt auf Mittelwerte kann man auch auf die Gebäudeanzahlen hochrechnen (sog. Merkmalssummenschätzung). So liefert die Auswertung

```
Anz(~factor(HeizArt), wogebheizart)
```

das Ergebnis

	total	SE
factor(HeizArt)1	1195794	111768.1
factor(HeizArt)2	329043	48807.4
factor(HeizArt)3	14878412	242972.5
factor(HeizArt)4	1050511	79346.2
factor(HeizArt)5	949984	79365.2

Zu denselben Anzahlen gelangt man allerdings auch, wenn man anstelle *wogebheizart* als zu analysierende Teilmenge *wogeb* wählt, wobei man im zweiten Fall noch die (hochgerechnete) Anzahl der Wohngebäude ohne verwertbare Angaben erhält, die bei 414.868 liegt und die man in geeigneter Weise (z. B. proportional) auf die gültigen Merkmalsausprägungen umlegen muss.

⁹ Vgl. hierzu Unterkapitel A.5 im Anhang A des Projektendberichts.

	total	SE
factor(HeizArt)0	414868	44174.1
factor(HeizArt)1	1195794	111768.1
factor(HeizArt)2	329043	48807.4
factor(HeizArt)3	14878412	242972.5
factor(HeizArt)4	1050511	79346.2
factor(HeizArt)5	949984	79365.2

Die Einbindung der *factor*-Anweisung in Merkmalssummenschätzungen ist bei nicht-dichotomen nominal- oder ordinalskalierten Merkmalen zwingend erforderlich, denn andernfalls werden die – letztlich willkürlich vergebenen – Zahlencodes der Ausprägungen aufsummiert, was im Normalfall zu nicht sinnvoll zu interpretierenden Ergebnissen führt. Bei dichotomen Merkmalen wie dem bereits erwähnten Fragebogenmerkmal *PHFoerdSonst*, das darauf abstellt, ob bei Neubauten mit Baujahr ab 2010 für die Erreichung eines KfW-Effizienzhaus- oder Passivhausstandards Fördermittel in Anspruch genommen wurden, die nicht bzw. nicht ausschließlich aus einem KfW-Programm stammen, ist diese Faktorisierung allerdings nicht erforderlich. So liefert der Befehl

```
Anz(~PHFoerdSonst, wogebfoerdermittel)
```

folgendes Ergebnis:

	total	SE
PHFoerdSonst	51669.1	41563.7

Auf den ersten Blick könnte man dieses Ergebnis dahingehend interpretieren, dass die (geschätzte) Anzahl von neugebauten Wohngebäuden (Baujahr ab 2010), die einen KfW-Effizienzhaus- oder Passivhausstandard erreichen, wofür – entweder ausschließlich oder zusätzlich – Fördermittel aus einem anderen Fördertopf als dem der KfW eingesetzt wurden, bei 51.669 Gebäuden liegt. Eine solche Interpretation ignoriert allerdings den Umstand, dass es in der Auswertungsdatenbank Untersuchungsgebäude geben kann (hier konkret solche mit Baujahr ab 2010), die bezüglich des interessierenden Fragekomplexes keine verwertbaren Angaben aufweisen, die realiter aber dennoch zur Gruppe derjenigen neugebauten Wohngebäude (Baujahr ab 2010) gehören, die einen geförderten KfW-Effizienzhaus- oder Passivhausstandard erreichen und für deren Errichtung ausschließlich oder zusätzlich „andere Fördermittel“ geflossen sind. Solche Gebäude wurden in der obigen Auswertung jedoch nicht berücksichtigt, wodurch der obige Befehl tendenziell zu einer Unterschätzung der gesuchten Gebäudezahl führt. Gelöst werden kann dieses Problem wieder auf verschiedenen Wegen, wobei jede Vorgehensweise implizit oder explizit eine Annahme darüber erfordert, ob und wenn ja wie sich die Untersuchungsgebäude mit fehlenden Angaben von denjenigen mit verwertbaren Angaben unterscheiden. So besteht eine Herangehensweise, den (hochgerechneten) Anteil der Gebäude mit fehlenden Angaben zu bestimmen, ihn (wenn er von 0 verschieden sein sollte, es also entsprechende Untersuchungsgebäude in der Auswertungsdatenbank gibt) durch den Anteil der Gebäude mit verwertbaren Angaben zu dividieren, den Quotienten um 1 (bzw. 100 %) zu erhöhen und diesen Faktor mit der oben geschätzten Gebäudezahl (51.669) zu multiplizieren¹⁰. Zur Berechnung des (geschätzten) Standardfehlers bietet sich wiederum das Gaußsche Fehlerfortpflanzungsgesetz an.

Auch bei bestimmten metrischen Merkmalen wie zum Beispiel derjenigen zur Zahl der zu Wohnzwecken vermieteten (oder mietfrei überlassenen) Wohnungen im Untersuchungsgebäude kann eine Merkmals-

¹⁰ Dieser Vorgehensweise liegt implizit die Annahme zugrunde, dass sich Untersuchungsgebäude mit fehlenden Angaben nicht von denjenigen mit verwertbaren Angaben unterscheiden.

summenschätzung sinnvoll sein, beispielsweise, wenn die Zahl der Mietwohnungen, etwa für ein bestimmtes Gebäudesegment, geschätzt werden soll. Aufgrund der hierarchischen Abfrage der Mietwohnungszahl, der eine Frage zum generellen Vorhandensein von Mietwohnungen im Gebäude vorangeht, und der Möglichkeit von Item Nonresponse auf beiden Hierarchiestufen sind entsprechende Auswertungen jedoch anspruchsvoller und bedingen mehrere Anweisungen. Ein alleiniger Befehl der Art „Anz(~nWhgMiet, ·)“ wird hier nicht zum Erfolg führen.

Möchte man die Besetzungsstärke einer Teilmenge, also die Gesamtzahl der Gebäude in einem Gebäudesegment, schätzen¹¹, bietet sich ein Rückgriff auf das ergänzende Merkmal *Geb* an, das für jedes der 16.982 Untersuchungsgebäude den Wert 1 annimmt. Auch hier ist jedoch wiederum die Möglichkeit fehlender Angaben zu berücksichtigen. So liefert die Anweisung

```
Anz(~Geb, Altbau78)
```

nicht die (hochgerechnete) Zahl der Gebäude mit Wohnraum, die vor 1979 errichtet wurden, sondern stattdessen die Zahl der Gebäude mit Wohnraum mit verwertbaren Angaben zum Baualter, die vor 1979 errichtet wurden. Um den Gebäudebestand bis Baujahr 1978 zu schätzen, ist es wiederum erforderlich, eine Annahme dahingehend zu treffen, wie sich die Baualtersstruktur der Gebäude ohne verwertbare Angaben darstellt, insbesondere ob sie derjenigen entspricht, die für die Gebäude mit Baualtersangabe abgeleitet werden konnte.

Ein Auswertungsergebnis besteht nicht unbedingt nur aus einem einzigen Schätzwert (incl. Standardfehler), sondern kann auch eine Kreuztabelle sein. Hierfür ist die Anweisung *interaction* vorgesehen. An einem Beispiel soll die Funktionsweise dieser Anweisung erläutert werden. Die Anweisung

```
MW(~interaction(FassSchutz,DenkmalSchutz),wogeb)
```

liefert folgendes Ergebnis

	mean	SE
interaction(FassSchutz, DenkmalSchutz)0.0	0.005413748	0.00107
interaction(FassSchutz, DenkmalSchutz)1.0	0.004694132	0.00205
interaction(FassSchutz, DenkmalSchutz)2.0	0.000445412	0.00025
interaction(FassSchutz, DenkmalSchutz)0.1	0.026663869	0.00401
interaction(FassSchutz, DenkmalSchutz)1.1	0.837207500	0.00823
interaction(FassSchutz, DenkmalSchutz)2.1	0.090985174	0.00560
interaction(FassSchutz, DenkmalSchutz)0.2	0.001429019	0.00076
interaction(FassSchutz, DenkmalSchutz)1.2	0.002416837	0.00075
interaction(FassSchutz, DenkmalSchutz)2.2	0.030744309	0.00424

Durch *interaction* werden die beiden Fragebogenmerkmale¹² *FassSchutz* (Vorhandensein besonders erhaltenswerter Gebäudefassaden: ja/nein) und *DenkmalSchutz* (Gebäude unter Denkmalschutz: ja/nein) faktorisiert, also in ihre möglichen Ausprägungen (hier jeweils 0, 1 und 2) zerlegt. Durch den Befehl *MW* wird für

¹¹ In methodischer Hinsicht handelt es sich um die Schätzung des Umfang einer Domain, vgl. z. B. Särndal, C.-E.; Swensson, B.; Wretman, J.: *Model Assisted Survey Sampling*, New York et al., 1992, S. 69 f..

¹² Bei der Auswertung der Auswertungsdatenbank ist die Verwendung der *interaction*-Anweisung auf maximal zwei Merkmale beschränkt.

jede mögliche Kombination (hier: 9 Kombinationen) deren anteiliges Gewicht berechnet, und zwar im vorliegenden Fall in Bezug auf die Teilmenge der Wohngebäude (*wogeb*). Die Summe aller Anteile beträgt dabei 100 %. Zu berücksichtigen ist wiederum das Vorhandensein fehlender Angaben. So weisen (hochgerechnet) beispielsweise 0,5 % aller Wohngebäude weder eine verwertbare Angabe zum Vorhandensein besonders erhaltenswerter Gebäudefassaden noch eine verwertbare Angabe dazu auf, ob das Gebäude unter Denkmalschutz steht. Zusammengenommen macht der Anteil der Kombinationen, die in Bezug auf wenigstens eine der beiden Merkmale keine verwertbaren Angaben aufweisen, 3,9 % aus. Es ist wiederum Sache des Anwenders zu entscheiden, wie er mit dem Phänomen fehlender Angaben umgeht. Eine Möglichkeit bestünde darin, das anteilige Gewicht der Kombinationen mit fehlenden Angaben proportional auf die anderen Anteile zu verteilen.

Schließlich sei noch die Möglichkeit erwähnt, statt Variablen auch logische Ausdrücke als Argumente der Funktionen *MW*, *MW_W*, *Anz* und *Anz_W* zu verwenden. Die Werte logischer Ausdrücke sind entweder FALSE („falsch“) oder TRUE („wahr“).

So liefert die Auswertung

```
MW(~factor(HeizArt == 3), wogeb)
```

das folgende Ergebnis

	mean	SE
factor(HeizArt == 3)FALSE	0.209378	0.00816
factor(HeizArt == 3)TRUE	0.790622	0.00816

Der logische Ausdruck *HeizArt == 3* wird in seine beiden möglichen Werte FALSE und TRUE zerlegt und die Mittelwertbildung erfolgt wie schon zuvor beschrieben. Konkret erhält man den (hochgerechneten) Anteil der Wohngebäude, für die der Befragte als überwiegende Gebäudebeheizungsart „Zentralheizung“ angegeben hat (hier: 79,1 %), sowie als Gegenstück dazu der Anteil der Wohngebäude, für die der Befragte als überwiegende Gebäudebeheizungsart eben nicht „Zentralheizung“ angegeben hat (hier: 20,9 %). Da sich im letztgenannten Anteil auch Gebäude mit fehlenden Angaben verbergen, kann nicht gefolgert werden, dass 79,1 % der Wohngebäude in Deutschland überwiegend zentral beheizt sind. Erneut ist eine Entscheidung zur Handhabung dieser fehlenden Werte erforderlich, wofür allerdings erst einmal berechnet werden muss, ob und wenn ja welches anteilige Gewicht den Wohngebäuden mit fehlenden Angaben zukommt. Dieses anteilige Gewicht berechnet man am besten mit der Anweisung

```
MW(~factor(HeizArt == 0), wogeb)
```

Der gesuchte Anteil beläuft sich auf 2,2 % (siehe nächste Seite) – ein Ergebnis, das bereits an früherer Stelle auf anderem Wege ermittelt wurde.


	mean	SE
factor(HeizArt == 0)FALSE	0.9779544	0.00233
factor(HeizArt == 0)TRUE	0.0220456	0.00233

Abschließend werden nochmals die zentralen Hinweise zur Durchführung der Auswertungen zusammengefasst:

- Es ist die oben dargestellte Syntax zu verwenden. Die Verwendung anderer Auswertungsbefehle des Programms R ist nicht zulässig.
- Innerhalb des Teils C dürfen maximal 25 Auswertungen durchgeführt werden.
- Die Funktion *interaction* darf, wie in dem gegebenen Beispiel, mit maximal zwei Variablen verwendet werden.
- Besondere Aufmerksamkeit ist auf fehlende Angaben (Variablenausprägung = 0) zu richten. Werden fehlende Angaben bei den Ausprägungen nicht adäquat berücksichtigt, kann dies zu falschen Auswertungsergebnissen führen.

Anhang A: Fragebogenvariablen

Erhebung Gebäude mit Wohnraum



Institut
Wohnen und
Umwelt

Erhebungsgruppe (A-C): – Musterstadt

Wie viele Gebäude mit Wohnraum gibt es unter der im Anschreiben genannten **Adresse**?

Bewohnte bzw. bewohnbare Unterkünfte (z. B. Gartenlauben, Wohncontainer) zählen nicht als Gebäude.

1 ein Gebäude nGeb1 nGeb2

2 mehr als eines, nämlich → Bitte ein Gebäude für die Befragung auswählen.

3 kein Gebäude mit Wohnraum → Die Befragung ist hier zu Ende. Den Fragebogen bitte dennoch zurücksenden.

Was ist das Baujahr des Gebäudes? (Zeitpunkt der Fertigstellung)

Ursprünglicher Gebäudeteil ohne spätere Anbauten. Falls das Gebäude früher weitgehend zerstört war: Jahr des Wiederaufbaus.

Baujahr Wenn Sie das Jahr nicht genau wissen, bitte schätzen.
Baujahr

Um welche **Art von Gebäude** handelt es sich?

1 Wohngebäude GebArt
Mindestens die Hälfte der Gesamtnutzfläche wird für Wohnzwecke genutzt.

2 Sonstiges Gebäude mit Wohnraum
Sonstige Nutzfläche ist größer als Wohnfläche, z. B. Geschäftshaus mit Wohnung.

3 Wohnheim
mit eigener Haushaltsführung der Bewohner/-innen (ohne eigene Haushaltsführung: kein Gebäude mit Wohnraum)

Wie viele Wohnungen befinden sich im Gebäude?

Auch leer stehende Wohnungen, Einliegerwohnungen sowie Ferien- und Freizeitwohnungen. Auch Wohnungen, die gewerblich genutzt werden (z. B. als Arztpraxis, Büro). Ohne Ladenlokale.

1 nWE1 2 2 3 3 4 4 5 mehr als 4, nämlich nWE2

Sind **darunter** auch folgende Wohnungen? nWhgMiet

- zu Wohnzwecken vermietete (oder mietfrei überlassene) Wohnungen: 1 nein 2 ja, Anzahl:

WhgMiet

- rein gewerblich genutzte Wohnungen: 1 nein 2 ja, Anzahl:

WhgGewerb

Was die **Bebauungssituation** angeht, handelt es sich um ... nWhgGewerb

1 ein freistehendes Gebäude

2 eine Doppelhaushälfte Bebau1

3 ein Gebäude in geschlossener Bebauung (mit drei oder mehr Gebäuden), und zwar:

→ 1 mit einem direkt angrenzenden Nachbargebäude
z. B. Endhaus in einer Häuserzeile Bebau2

2 mit zwei (oder mehr) direkt angrenzenden Nachbargebäuden
z. B. Mittelhaus in einer Häuserzeile oder einem Häuserblock

Wer ist **Eigentümer/-in** des Gebäudes?

1 Wohnungseigentümergeinschaft *Immer ankreuzen, wenn das Gebäude in Eigentumswohnungen aufgeteilt ist.*

2 Privatperson(en) Eigentuemer

3 Wohnungsgenossenschaft, privatwirtschaftliches oder kommunales Wohnungsunternehmen, Kommune

4 Anderes privatwirtschaftliches Unternehmen, Bund, Land, Organisation ohne Erwerbszweck (z. B. Kirche)

Art der Außenwände (ohne Berücksichtigung von Dämmstoffen)

überwiegende Bauart AWTyp

1 einschaliges Mauerwerk

2 zweischaliges Mauerwerk
(innere Mauerwerkswand und äußere Vormauerschale, durch Luftschicht oder Dämmstoff getrennt)

3 Fachwerkwand

4 Holz-Fertigbauteile oder sonstige Holzbauweise
(ohne Fachwerk)

5 Betonfertigteile, Großtafelbauweise, Plattenbau

6 Sonstiges, und zwar:

Seite 1

17

Erhebungsgruppe (A-C): – Musterstadt

Sind die **Außenwände** des Gebäudes ganz oder teilweise **mit Wärmedämmschichten versehen**?

z. B. Styropor, Mineralwolle, Dämmschüttungen oder Einblasdämmungen in Zwischenräumen
Nicht berücksichtigen: gut dämmende Mauerwerksteine

1 nein **AWDaemmung**

2 ja → Welcher **Anteil** der Außenwandfläche ist ungefähr gedämmt?

Prozentwert bezogen auf die gesamte Außenwandfläche des Gebäudes (ohne Fenster, ohne Haustrennwände)

1 ca.100% 2 ca. 75% 3 ca. 50% 4 ca. 25% 5 ca. 10% **AWDaeF**

Wurden die Wärmedämmschichten (überwiegend) **schon bei Errichtung des Gebäudes** oder **später** angebracht?

1 schon bei der Errichtung des Gebäudes **AWDaeWann**

2 später, und zwar in folgendem Zeitraum: **AWDaeBJ**

→ 1 bis 2009 2 2010 - 2012 3 2013 4 2014 5 2015 6 2016 / 2017

Wo ist die Wärmedämmung (überwiegend) angebracht?

1 außen auf der Wand **Außendämmung** 3 im Zwischenraum zwischen zwei Mauerwerksschalen **AWDaeOrt**

2 innen auf der Wand **Innendämmung** 4 im Bauteil selbst *z. B. bei Fertigbauteilen, eventuell auch bei Fachwerkwand*

Welche **Dämmstoffdicke in cm** liegt überwiegend vor? *wenn nötig: runden*

1 bis 2 2 3-5 3 6-9 4 10-13 5 14-17 6 18-21 7 22 cm oder mehr **AWDaeD**

Wurde nach dem 1.1.2010 die Fassade ganz oder teilweise **renoviert**, ohne dass gleichzeitig gedämmt wurde?

1 nein 2 ja → und zwar: Putzerneruerung oder neue Fassadenverkleidung neuer Anstrich

FassRenoOhneDae *zwei Angaben möglich* **FassPutzNeu** **FassFarbNeu**

Sind die **Fassaden** des Gebäudes ganz oder teilweise **besonders erhaltenswert**?


auch, aber nicht nur Denkmalschutz, z. B. das Stadtbild prägende Fassaden


1 nein 2 ja **FassSchutz**


Steht das Gebäude ganz oder teilweise unter **Denkmalschutz**?

1 nein 2 ja **DenkmalSchutz**

Ist das Dachgeschoss beheizt oder unbeheizt?

1 Dachgeschoss unbeheizt  **DaBeheizt**

2 Dachgeschoss voll beheizt *bzw. kein Dachgeschoss vorhanden* 

3 Dachgeschoss teilweise beheizt 

Ist der beheizte Bereich des Gebäudes **nach oben hin ganz oder teilweise gedämmt** (Dämmung der Dachfläche und / oder der obersten Geschossdecke zu unbeheiztem Dachraum)?

z. B. Styropor, Mineralwolle, Dämmschüttungen oder Einblasdämmungen in Zwischenräumen

1 nein **DaDaemmung**

2 ja → Wurden die Wärmedämmschichten (überwiegend) **schon bei Errichtung des Gebäudes** oder **später** angebracht?

1 schon bei der Errichtung des Gebäudes **DaDaeWann**

2 später, und zwar in folgendem Zeitraum: **DaDaeBJ**

→ 1 bis 2009 2 2010 - 2012 3 2013 4 2014 5 2015 6 2016 / 2017




Wo ist die Wärmedämmung (überwiegend) angebracht?

1 im Dach 2 auf oder in der obersten Geschossdecke **DaDaeWo**

Welche **Dämmstoffdicke in cm** liegt überwiegend vor? *wenn nötig: runden*

DaDaeD 1 bis 5 2 6-9 3 10-13 4 14-17 5 18-21 6 22-25 7 26 cm oder mehr

Erhebungsgruppe (A-C): – Musterstadt

Wurde <u>nach dem 1.1.2010</u> das Dachgeschoss ganz oder teilweise ausgebaut ? <i>als Wohnraum oder beheizte Nutzfläche</i> 1 <input type="radio"/> nein 2 <input type="radio"/> ja AusbauDG											
Wurde <u>nach dem 1.1.2010</u> die Dachhaut erneuert ? <i>z. B. Neueindeckung des Daches, Abdichtung Flachdach</i> 1 <input type="radio"/> nein 2 <input type="radio"/> ja DachhautNeu											
Welche Aussage hinsichtlich des Kellers trifft auf das Gebäude zu? 1 <input type="radio"/> nicht unterkellert KGBheizt 3 <input type="radio"/> teilweise beheizter Keller  2 <input type="radio"/> unbeheizter Keller  4 <input type="radio"/> voll beheizter Keller 											
Ist der Fußboden des Gebäudes zum Keller oder Erdreich ganz oder teilweise mit Wärmedämmschichten versehen (auch Dämmschichten an der Kellerdecke)? <i>z. B. Styropor, Mineralwolle, Dämmschüttungen oder Einblasdämmungen in Zwischenräumen</i> 1 <input type="radio"/> nein FBDaemmung 2 <input type="radio"/> ja → <table border="1" style="width: 100%;"> <tr> <td colspan="2"> Wurden die Wärmedämmschichten (überwiegend) schon bei Errichtung des Gebäudes oder später angebracht? </td> </tr> <tr> <td>1 <input type="radio"/> schon bei der Errichtung des Gebäudes FBDaeWann</td> <td>2 <input type="radio"/> später, und zwar in folgendem Zeitraum: FBDaeBJ</td> </tr> <tr> <td colspan="2" style="text-align: center;"> → 1 <input type="radio"/> bis 2009 2 <input type="radio"/> 2010 - 2012 3 <input type="radio"/> 2013 4 <input type="radio"/> 2014 5 <input type="radio"/> 2015 6 <input type="radio"/> 2016 / 2017 </td> </tr> <tr> <td colspan="2"> Welche Dämmstoffdicke in cm liegt überwiegend vor? <i>wenn nötig: runden</i> </td> </tr> <tr> <td colspan="2" style="text-align: center;"> 1 <input type="radio"/> bis 2 2 <input type="radio"/> 3-5 3 <input type="radio"/> 6-9 4 <input type="radio"/> 10-13 5 <input type="radio"/> 14-17 6 <input type="radio"/> 18 cm oder mehr FBDaeD </td> </tr> </table>		Wurden die Wärmedämmschichten (überwiegend) schon bei Errichtung des Gebäudes oder später angebracht?		1 <input type="radio"/> schon bei der Errichtung des Gebäudes FBDaeWann	2 <input type="radio"/> später, und zwar in folgendem Zeitraum: FBDaeBJ	→ 1 <input type="radio"/> bis 2009 2 <input type="radio"/> 2010 - 2012 3 <input type="radio"/> 2013 4 <input type="radio"/> 2014 5 <input type="radio"/> 2015 6 <input type="radio"/> 2016 / 2017		Welche Dämmstoffdicke in cm liegt überwiegend vor? <i>wenn nötig: runden</i>		1 <input type="radio"/> bis 2 2 <input type="radio"/> 3-5 3 <input type="radio"/> 6-9 4 <input type="radio"/> 10-13 5 <input type="radio"/> 14-17 6 <input type="radio"/> 18 cm oder mehr FBDaeD	
Wurden die Wärmedämmschichten (überwiegend) schon bei Errichtung des Gebäudes oder später angebracht?											
1 <input type="radio"/> schon bei der Errichtung des Gebäudes FBDaeWann	2 <input type="radio"/> später, und zwar in folgendem Zeitraum: FBDaeBJ										
→ 1 <input type="radio"/> bis 2009 2 <input type="radio"/> 2010 - 2012 3 <input type="radio"/> 2013 4 <input type="radio"/> 2014 5 <input type="radio"/> 2015 6 <input type="radio"/> 2016 / 2017											
Welche Dämmstoffdicke in cm liegt überwiegend vor? <i>wenn nötig: runden</i>											
1 <input type="radio"/> bis 2 2 <input type="radio"/> 3-5 3 <input type="radio"/> 6-9 4 <input type="radio"/> 10-13 5 <input type="radio"/> 14-17 6 <input type="radio"/> 18 cm oder mehr FBDaeD											
Wann wurden die Fenster eingebaut ? <i>überwiegender Anteil der Fensterfläche</i> 1 <input type="radio"/> bis Ende 1994 2 <input type="radio"/> ab 1995 BJFenster											
Welche Art der Verglasung liegt überwiegend vor? 1 <input type="radio"/> Ein-Scheiben-Verglasung 2 <input type="radio"/> Zwei-Scheiben-Verglasung 3 <input type="radio"/> Drei-Scheiben-Verglasung Scheiben <i>auch Kastenfenster</i>											
Wurden <u>nach dem 1.1.2010</u> Fenster oder Verglasungen am Gebäude erneuert ? 1 <input type="radio"/> nein FenNeu FenNeuWann 2 <input type="radio"/> ja → und zwar (überwiegend): 1 <input type="radio"/> 2010 - 2012 2 <input type="radio"/> 2013 3 <input type="radio"/> 2014 4 <input type="radio"/> 2015 5 <input type="radio"/> 2016 / 2017 Welcher Anteil der gesamten Fensterfläche wurde <u>nach dem 1.1.2010</u> ungefähr erneuert? 1 <input type="radio"/> ca. 100 % 2 <input type="radio"/> ca. 75 % 3 <input type="radio"/> ca. 50 % 4 <input type="radio"/> ca. 25 % 5 <input type="radio"/> ca. 10 % FenNeuAnteil											
Sind das Gebäude oder einzelne Wohnungen mit einer maschinellen Wohnungslüftungsanlage ausgestattet? <i>Nicht berücksichtigen: Dunstabzugshauben in der Küche oder Ventilatoren nur für Bad / WC</i> 1 <input type="radio"/> nein Lueftung LueftTyp 2 <input type="radio"/> ja → und zwar (überwiegend): 1 <input type="radio"/> ohne Wärmerückgewinnung 2 <input type="radio"/> mit Wärmerückgewinnung											
Welche Beheizungsart liegt in dem Gebäude überwiegend vor? <u>Auch hier nur ein Feld ankreuzen</u> 1 <input type="radio"/> Fernwärme <i>auch kleinere Fernwärmenetze (Nahwärme)</i> 2 <input type="radio"/> Blockheizung <i>gemeinsame Versorgung eines Häuserblocks oder einer Häuserzeile</i> 3 <input type="radio"/> Zentralheizung <i>zentrale Versorgung des Gebäudes, ohne dass andere Gebäude mitversorgt werden</i> 4 <input type="radio"/> Etagenheizung <i>im Mehrfamilienhaus, z. B. Gas-Etagenheizung</i> HeizArt 5 <input type="radio"/> Einzel- oder Mehrraumöfen, Raumheizgeräte, Nachtspeicherheizung											

Seite 3

Erhebungsgruppe (A-C): – Musterstadt

Welcher Energieträger wird überwiegend für die Beheizung verwendet? bei Fernwärme: keine Angabe
Auch hier nur ein Feld ankreuzen
 1 Erdgas 4 Biomasse (außer Holz), Biogas 5 elektrischer Strom
 auch Biogas aus dem Erdgasnetz, wenn *auch bei elektr. Wärmepumpe*
 2 Heizöl **BS** *Anteil laut Liefervertrag > 50 %*
 3 Holz, Holzpellets 6 Flüssiggas 7 Kohle

Was ist der Haupt-Wärmeerzeuger der Heizung? bei Fernwärme: keine Angabe
Auch hier nur ein Feld ankreuzen
 1 Heizkessel *auch Therme* 4 Wärmepumpe **Erzeuger**
 mit Brennstoff betrieben *Nutzung von Abwärme oder Umweltwärme (Abluft, Außenluft, Erdreich, Grundwasser)*
 2 Ofen/Öfen 5 Blockheizkraftwerk
 3 elektr. Heizgerät(e) *z. B. Nachtspeicherheizung* 6 Sonstiges, und zwar:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Wurde der Haupt-Wärmeerzeuger der Heizung nach dem 1.1.2010 neu eingebaut oder das Gebäude nach dem 1.1.2010 erstmalig an Fernwärme angeschlossen?
 1 nein **ErzNeu** **ErzBJ**
 2 ja → und zwar: 1 2010 - 2012 2 2013 3 2014 4 2015 5 2016 / 2017

Womit wurde die Heizwärme vorher (überwiegend) erzeugt?
Auch ankreuzen, wenn Typ gleichgeblieben.
ErzUr 1 Fernwärme 2 Gas-Heizkessel 3 Öl-Heizkessel 4 Biomasse-Heizkessel
 z. B. Holz-Heizkessel
 5 Wärmepumpe 7 elektr. Heizgerät(e) 6 Ofen/Öfen 8 Sonstiges oder nicht zutreffend
 z. B. Nachtspeicherheizung *z. B. Neubau*

Ist auf bzw. an dem Gebäude eine Solarwärmelanlage (Solarthermie) installiert?
Auch auf dazugehöriger Garage, Schuppen
 1 nein **SolTh**
 2 ja → **Wozu dient die Solarwärmelanlage?**
 1 nur zur Warmwasserversorgung **SolThNutz**
 2 zur Warmwasserversorgung und Heizungsunterstützung
Wann wurde die Solarwärmelanlage installiert? **SolThWann**
 1 bis 2009 2 2010 - 2012 3 2013 4 2014 5 2015 6 2016 / 2017

Ist auf bzw. an dem Gebäude eine Solarstromanlage (Photovoltaik) installiert?
Auch auf dazugehöriger Garage, Schuppen
 1 nein **PhV**
 2 ja → **Wann wurde die Solarstromanlage installiert?** **PhVWann**
 1 bis 2009 2 2010 - 2012 3 2013 4 2014 5 2015 6 2016 / 2017

Wurden nach dem 1.1.2010 die Verteilrohre der Heizung und / oder des Warmwassersystems neu eingebaut oder wärmegeklärt? *ganz oder teilweise*
 1 nein **VerteilNeu** **VerteilWWNeu**
 2 ja → und zwar: Heizungsrohre Warmwasserrohre
 zwei Angaben möglich **VerteilHZNeu** **VerteilBJ**
Zeitpunkt der Fertigstellung: 1 2010 - 2012 2 2013 3 2014 4 2015 5 2016 / 2017

Falls das Gebäude nach dem 1.1.2010 errichtet wurde:
Wird ein KfW-Effizienzhausstandard oder der Passivhausstandard erreicht?
KfW-Effizienzhaus 70 / 55 / 40 oder Passivhaus mit Nachweis durch einen Fachmann
 1 nein **PH**
 2 ja → **Wurde für diesen Standard eine Förderung (Zuschuss oder Förderkredit) in Anspruch genommen?**
 1 nein (Der Standard wurde erreicht, aber nicht gefördert.) **PHFoerdSonst**
 2 ja → **PHFoerd** **PHFoerdKfW**
 und zwar: aus einem KfW-Programm andere Fördermittel
 zwei Angaben möglich

Anhang B: Rohfassung der Auswertungsdatei

```
#####  
#  
# Auswertungsdatei für die Durchführung eines Auswertungsgangs mit der  
# Auswertungsdatenbank zum Projekt „Datenerhebung Wohngebäudebestand 2016“  
#  
#####  
###  
###Bitte folgende Angaben vervollständigen:  
###  
###Dateiname des Programmcodes: Musterprogramm.txt  
###erstellt: <Datum>  
###von: <Name>  
###E-Mail: <E-Mail-Adresse>  
###Tel.: <Telefonnummer>  
###  
#####  
  
#####  
###Teil A: Kopfzeile  
###Diese Kopfzeile darf nicht verändert werden  
#####  
source("U:\\R\\headerkdfv.r")  
  
#####  
###Teil B: Festlegung von Teilmengen des Gebäudebestandes  
#####  
  
#####  
###Teil D: Durchführung der Auswertungen  
###bis zu 25 Auswertungen sind zulässig  
#####  
  
##Auswertung 1  
  
##Auswertung 2  
  
##Auswertung 3
```

##Auswertung 4

##Auswertung 5

##Auswertung 6

##Auswertung 7

##Auswertung 8

##Auswertung 9

##Auswertung 10

##Auswertung 11

##Auswertung 12

##Auswertung 13

##Auswertung 14

##Auswertung 15

##Auswertung 16

##Auswertung 17

##Auswertung 18

##Auswertung 19

##Auswertung 20

##Auswertung 21

##Auswertung 22

##Auswertung 23

##Auswertung 24

##Auswertung 25