



Forschungsdatenbank Nichtwohngebäude

(ENOB:dataNWG)

Forschungsprojekt im Förderbereich
Energieoptimierte Gebäude und Quartiere im
6. Energieforschungsprogramm der Bundesregierung
Förderkennzeichen 03ET1315

Stichprobendesign

Dr. Holger Cischinsky
Institut Wohnen und Umwelt
Expertenbeirat, 20.09.2016

0. Phase: Geodatenanalyse

- Geokoordinate
- Gebäudefunktion
- Ca. 40 weitere Attribute
- Stichprobe ziehen

1. Phase: Screening

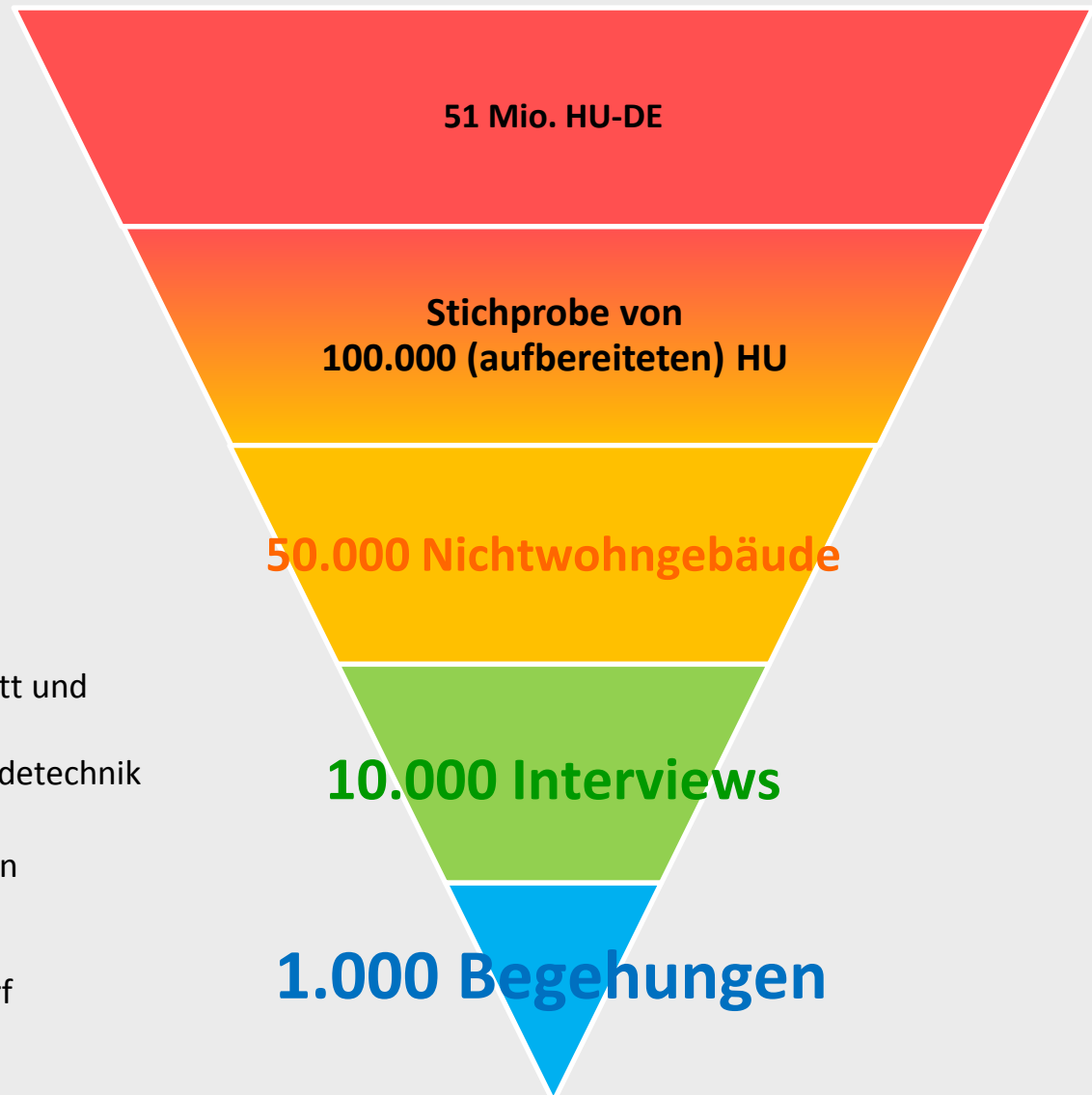
- Relevanz
- Gebäudebasismerkmale
- Bebauungssituation
- Adressdaten

2. Phase: Breitenerhebung

- typologische Merkmale
- Modernisierungsfortschritt und Modernisierungsrate bei Wärmeschutz und Gebäudetechnik
- Eigentümerstrukturen
- Bewirtschaftungsverhalten

3. Phase: Tieferhebung

- Energieverbrauch /-bedarf
- Abgleich
- Szenarien



Ziel: Aufgrund des Ausscheidens einer Vollerhebung sollen auf Basis einer Stichprobe Rückschlüsse (im Sinne von datenorientierten Aussagen) auf die Grund- bzw. Untersuchungsgesamtheit aller EnEV-relevanten Nichtwohngebäude (NWG) (= Untersuchungseinheiten) in Deutschland gezogen werden.

Problem: Ein solcher Rückschluss ist **stets** mit Unsicherheit behaftet.

Aufgabe: Es ist dafür Sorge zu tragen, dass die stichprobenimmanente Unsicherheit

- **unsystematisch** („Erwartungstreue“)
- **quantifizierbar** und
- unter den gegebenen Randbedingungen (insbes. Einhaltung eines Kostenrahmens) **möglichst gering** ist.

Postulat der Stichprobentheorie:

1. Entscheidet über die Aufnahme eines EnEV-relevanten NWG in die Stichprobe allein der **Zufall**,
2. sind die **Aufnahmewahrscheinlichkeiten** („Inklusionswahrscheinlichkeiten 1. Ordnung“) aller Stichprobengebäude **bekannt** und
3. hat grundsätzlich **jedes EnEV-relevante NWG** in Deutschland eine **Chance auf Aufnahme in die Stichprobe** (positive Auswahlwahrscheinlichkeiten),
wird **Repräsentativität** im Sinne von Erwartungstreue erreicht.

Sind darüber hinaus

4. die **kombinierten Auswahlwahrscheinlichkeiten** („Inklusionswahrscheinlichkeiten 2. Ordnung“) zweier beliebiger EnEV-relevanter NWG **positiv und bekannt**, lässt sich die Ergebnisunsicherheit in Gestalt von **Stichprobenfehlern** quantifizieren.

Bsp.: Wohngebäudeerhebung „Datenbasis Gebäudebestand“ (IWU/BEI, 2010)

Mittlere jährliche Modernisierungsrate der Außenwand älterer Wohngebäude (Bj. bis 1978) im Zeitraum 2005 bis 2008: 1,06% +/- 0,10%

Interpretation: Die wahre, aber unbekannte Modernisierungsrate **in der Grundgesamtheit** liegt mit einer Wahrscheinlichkeit von 68% zwischen 0,96% und 1,16%.

Stichprobenfehler (in Prozentpunkten) bei einer Anteilsschätzung in Abhängigkeit des geschätzten Anteils und des Stichprobenumfangs (netto)

		Stichprobenumfang (netto)						
		500	1.000	2.500	5.000	10.000	15.000	20.000
geschätzter Anteil	0,5 %	0,32%	0,22%	0,14%	0,10%	0,07%	0,06%	0,05%
	1 %	0,44%	0,31%	0,20%	0,14%	0,10%	0,08%	0,07%
	2 %	0,63%	0,44%	0,28%	0,20%	0,14%	0,11%	0,10%
	5 %	0,97%	0,69%	0,44%	0,31%	0,22%	0,18%	0,15%
	10 %	1,34%	0,95%	0,60%	0,42%	0,30%	0,24%	0,21%
	20 %	1,79%	1,26%	0,80%	0,57%	0,40%	0,33%	0,28%
	30 %	2,05%	1,45%	0,92%	0,65%	0,46%	0,37%	0,32%
	40 %	2,19%	1,55%	0,98%	0,69%	0,49%	0,40%	0,35%
	50 %	2,24%	1,58%	1,00%	0,71%	0,50%	0,41%	0,35%
	60 %	2,19%	1,55%	0,98%	0,69%	0,49%	0,40%	0,35%
	70 %	2,05%	1,45%	0,92%	0,65%	0,46%	0,37%	0,32%
	80 %	1,79%	1,26%	0,80%	0,57%	0,40%	0,33%	0,28%
90 %	1,34%	0,95%	0,60%	0,42%	0,30%	0,24%	0,21%	

- Angestrebter Stichprobenumfang (netto) für **Strukturaussagen** und zur Angabe von **energetischen Modernisierungsraten**
- Angestrebter Stichprobenumfang (netto) für Aussagen im Hinblick auf **Energiebedarf und -verbrauch**



Stichprobendesign: Auswahlgrundlage

- Ideal: Liste aller EnEV-relevanten NWG in Deutschland (Untersuchungsgesamtheit) mitsamt
- Objektanschrift
 - kompetenten Ansprechpartnern und
 - fundamentalen Gebäudeeigenschaften

Objektanschrift	Ansprechpartner	Gebäudebeschreibung
Musterstraße 4 (Haus 1) 12345 Musterhausen	Max Mustermann (Verwalter) Tel.: 0123456789 12345 Musterhausen	Bürogebäude 4 Vollgeschosse Baujahr 2000
Musterweg 17 54321 Musterdorf	Mustermann AG Frau Mustermann 99999 Musterstadt	Werkstatt 1 Vollgeschoss Baujahr 1960
Musterpfad 6 99999 Musterstadt	Autohaus Mustermann Herr Mustermann 99999 Musterstadt	Autohaus 1 Vollgeschoss Baujahr 1995
(...)	(...)	(...)

- Ersatzweise aufbereitete **Datenbank amtlicher Hausumringe (HU-DE)** als Auswahlgrundlage für die Stichprobenziehung

Amtlicher Hausumring (HU-DE): 2D-Gebäudegrundriss (Teilauszug aus der amtlichen Liegenschaftskarte)

Bsp.: Gebäudekomplex Rheinstraße 63-67 in Darmstadt (Hausnummer 65: IWU-Institutsgebäude)



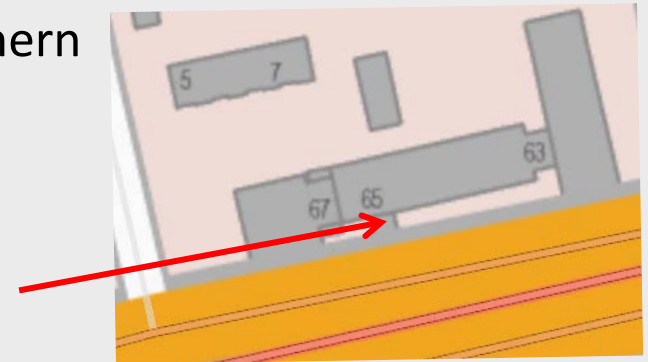
Quelle: Bundesamt für Kartographie und Geodäsie (DOP-Viewer)

aber: **Datenbank** der 51 Mio. amtlichen Hausumringe (HU-DE) aus einer Reihe von Gründen **noch nicht als Auswahlgrundlage** nutzbar.

Gründe:

- Vorhandensein von Kleinstpolygonen (z.B. Bauteilen), bei denen es sich nicht um eigenständige EnEV-relevante NWG handeln kann
- Topologische Inkonsistenzen (Überlappungen)
- Fehlende Informationen zur Identifizierung von EnEV-relevanten NWG
- Mutmaßlich größter Teil der amtlichen Hausumringe gehört nicht zu EnEV-relevanten NWG („irrelevante Einheiten“)
- Fehlende Adressinformationen
- Fehlende Informationen zu Gebäudeansprechpartnern
- Nicht notwendigerweise 1:1-Äquivalenz zwischen Hausumring und NWG

Vordach des IWU-Institutsgebäudes
ist eigener Hausumring !!!



➤ umfangreiche Datensatzaufbereitung erforderlich

- Beseitigung von Überlappungen, Kleinstpolygonen und atypischen Grundrissen

Ziel: möglichst gute Entsprechung von Umring und realem Gebäudegrundriss

- Anreicherung der aufbereiteten Hausumringe mit numerischen und semantischen Merkmalen unter Rückgriff auf diverse Datenquellen

Ziel: Rückschluss auf NWG-Eigenschaft bzw. EnEV-Relevanz

- Adressanreicherung



aber: Datensatzaufbereitung noch nicht ausreichend

- Von knapp 20 Mio. der 49 Mio. aufbereiteten Hausumringe liegen keine Informationen zur Untersuchungsrelevanz (EnEV-relevantes NWG) vor, wobei der Großteil dieser Hausumringe vermutlich irrelevant ist.

- Nach wie vor keine Informationen zu Gebäudeansprechpartnern

- 1:1-Äquivalenz zwischen aufbereiteten Hausumringen und NWG ist nicht sichergestellt.

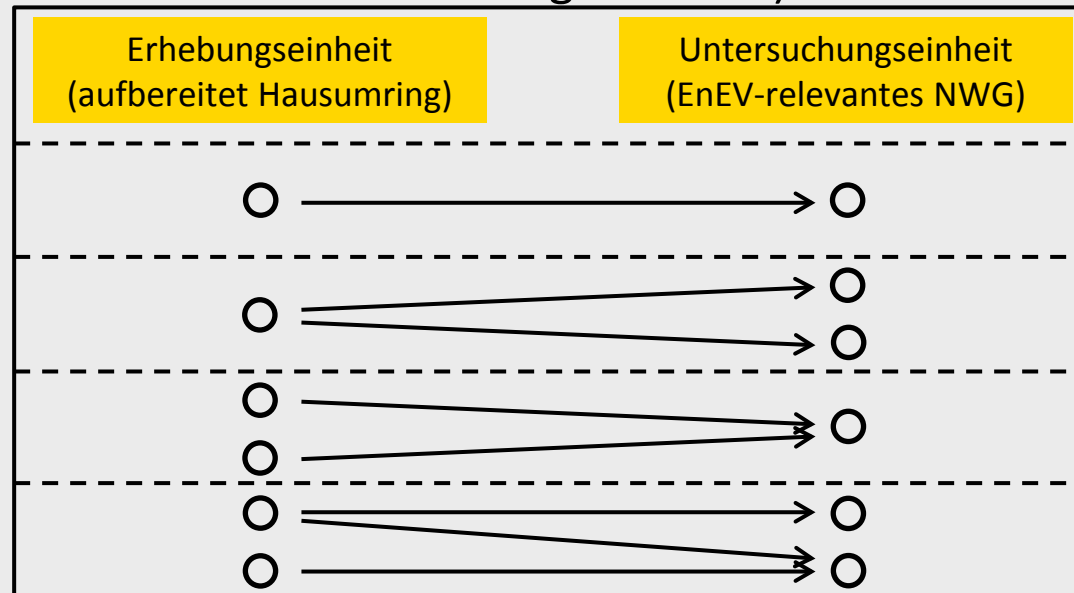
- Adressanreicherung unvollständig

- Erfordernis einer Vor-Ort-Begehung („Screening“) einer ausreichend großen Stichprobe aufbereiteter Hausumringe



Hauptnutzen:

- **Relevanzfeststellung** mit dem Ziel der Identifizierung von mind. 50.000 EnEV-relevanten NWG
- **Klärung Bebauungssituation** (Beziehungsgeflecht aufbereiteter Hausumring vs. NWG)



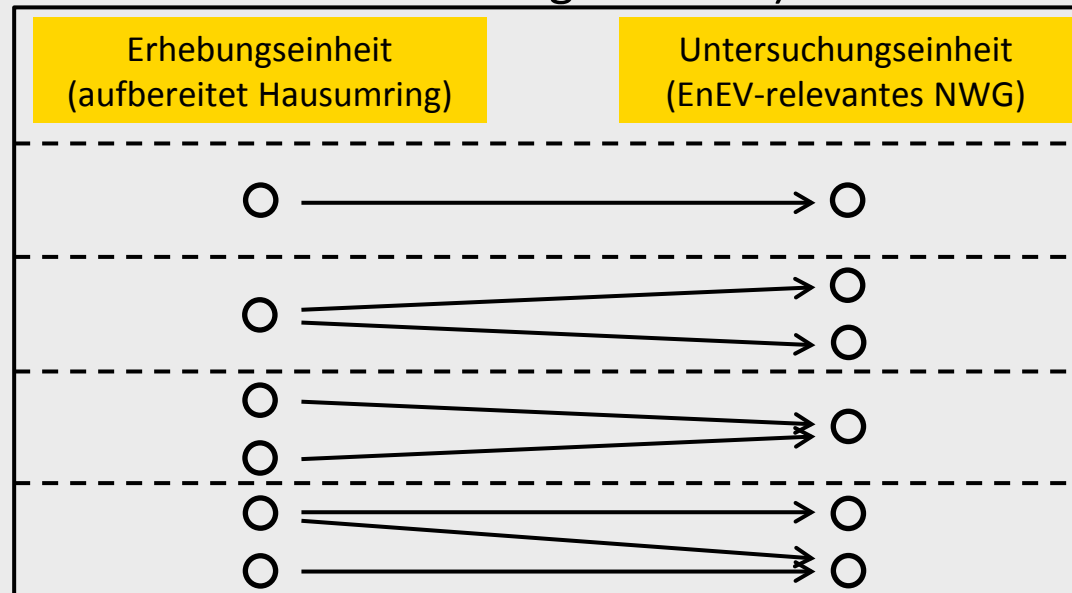
$$\pi_{k|i} = 1 - \prod_{h=1}^H \frac{\binom{N_i(h) - M_{k|i}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}}$$

$$\pi_{kl|i} = 1 - \prod_{h=1}^H \frac{\binom{N_i(h) - M_{k|i}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}} + \prod_{h=1}^H \frac{\binom{N_i(h) - M_{l|i}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}} - \prod_{h=1}^H \frac{\binom{N_i(h) - M_{kl|i}(h)}{n_i(h)}}{\binom{N_i(h)}{n_i(h)}}$$

- Erfordernis einer Vor-Ort-Begehung („Screening“) einer ausreichend großen Stichprobe aufbereiteter Hausumringe



- Hauptnutzen:**
- **Relevanzfeststellung** mit dem Ziel der Identifizierung von mind. 50.000 EnEV-relevanten NWG
 - **Klärung Bebauungssituation** (Beziehungsgeflecht aufbereiteter Hausumring vs. NWG)



- Informationen zu **potentiellen Gebäudeansprechpartnern**
- **Adressfeststellung**

- Zusatznutzen:**
- Erhebung von **Gebäudebasismerkmalen** vor Ort
 - ⇒ bereits das Screening liefert auf sehr breiter Stichprobenbasis **wesentliche Strukturmerkmale zum NWG-Bestand**
 - ⇒ groß angelegter **Praxistest zur Validierung der geoinformatorischen Algorithmen** des IÖR zur Erkennung der NWG-Eigenschaft
 - ⇒ **Identifizierung und Korrektur systematischer Gebäudeausfälle** („Unit Nonresponse“) bei der späteren Breitenerhebung

Aufgabe:

Bemessung des Umfangs der Stichprobe der aufbereiteten Hausumringe und Strukturierung dieser Stichprobe, so dass bei möglichst geringem Stichprobenumfang mindestens 50.000 EnEV-relevante NWG identifiziert werden können

- Herausforderung, da unter den 20 Mio. aufbereiteten Hausumringen ohne abschließende Informationen zur Untersuchungsrelevanz nur ein relativ kleiner, aber in absoluter Hinsicht nicht zu vernachlässigender Anteil zu EnEV-relevanten NWG gehört

Lösung:

Den knapp 20 Mio. aufbereiteten Hausumringen ohne Informationen zur Untersuchungsrelevanz (EnEV-relevantes NWG) werden **Wahrscheinlichkeiten für Untersuchungsrelevanz** zugespielt.

Relevanz	Merkmal 1	Merkmal 2	Merkmals 3	(...)	prob (Relevanz)
1	a	908	1,8		1,00
0	b	160	3,1		0,00
0	c	253	2,3		0,00
0	c	91	5,2		0,00
1	a	476	1,2		1,00
1	b	806	1,9		1,00
0	c	321	5,3		0,00
0	c	102	3,9		0,00
0	b	387	4,1		0,00
	b	163	1,8		0,18
	c	240	3,9		0,08
	c	175	3,1		0,09
	a	614	3,4		0,82
	a	490	1,4		0,98
	c	361	4,8		0,62
	b	198	3,1		0,32
	a	956	1,9		0,91

Binär-logistische
Regressionsgleichung

$\Sigma \text{prob(Relevanz)} = 4,0$

➤ Im Erwartungswert
sind 4 der 8 Haus-
umringe relevant.

- Ziehung einer (geschichteten) Stichprobe von 100.000 aufbereiteten Hausumringen, unter denen im Erwartungswert mindestens 50.000 zu EnEV-relevanten NWG gehören

Aufgabe: Berücksichtigung der eingeschränkten räumlichen Mobilität des Erhebungspersonals / Begrenzung der Wegekosten

Lösung:

- Bildung von **Erhebungsbezirken** und vorgelagerte Ziehung einer regional geschichteten Stichprobe von 500 Erhebungsbezirken
 - 1. Ziehungsstufe: Ziehung von 500 Erhebungsbezirken („primäre Erhebungseinheiten“)
 - 2. Ziehungsstufe: Ziehung von 200 aufbereiteten Hausumringen („sekundäre Erhebungseinheiten“) je ausgewähltem Erhebungsbezirk
- aufwändige Zerlegung Deutschlands in Erhebungsbezirke unter Beachtung diverser methodischer Vorgaben
- **Optimiertes Routing**



- Zweistufiges Ziehungsverfahren (1. Stufe: Erhebungsbezirke, 2. Stufe: aufbereitete Hausumringe), wobei auf der 2. Ziehungsstufe dreiphasige Auswahl erfolgt:

1. Phase: **Screening**

100.000 Hausumringe (500 Erhebungsbezirke)



2. Phase: **Breitenerhebung**

10.000 NWG



3. Phase: **Tiefenerhebung**

1.000
NWG



0. Phase: Geodatenanalyse

- Geokoordinate
- Gebäudefunktion
- Ca. 40 weitere Attribute
- Stichprobe ziehen

1. Phase: Screening

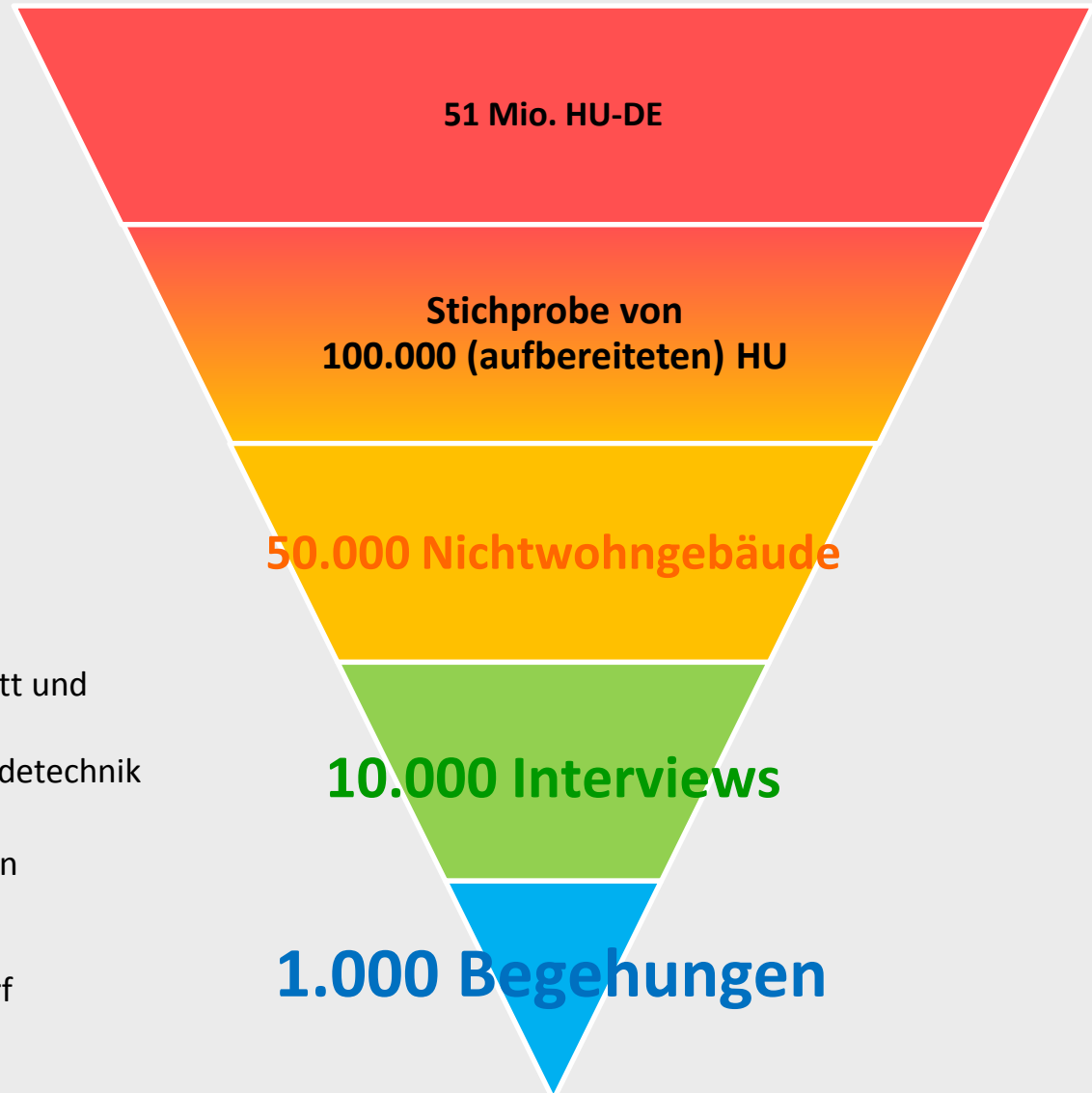
- Relevanz
- Gebäudebasismerkmale
- Bebauungssituation
- Adressdaten

2. Phase: Breitenerhebung

- typologische Merkmale
- Modernisierungsfortschritt und Modernisierungsrate bei Wärmeschutz und Gebäudetechnik
- Eigentümerstrukturen
- Bewirtschaftungsverhalten

3. Phase: Tiefenerhebung

- Energieverbrauch /-bedarf
- Abgleich
- Szenarien



Vielen Dank für Ihre Aufmerksamkeit !!!