

todo

(noch) mehr Beispiele

Daten in der Verkehrsplanung (insbes. im Hinblick auf Modelle)

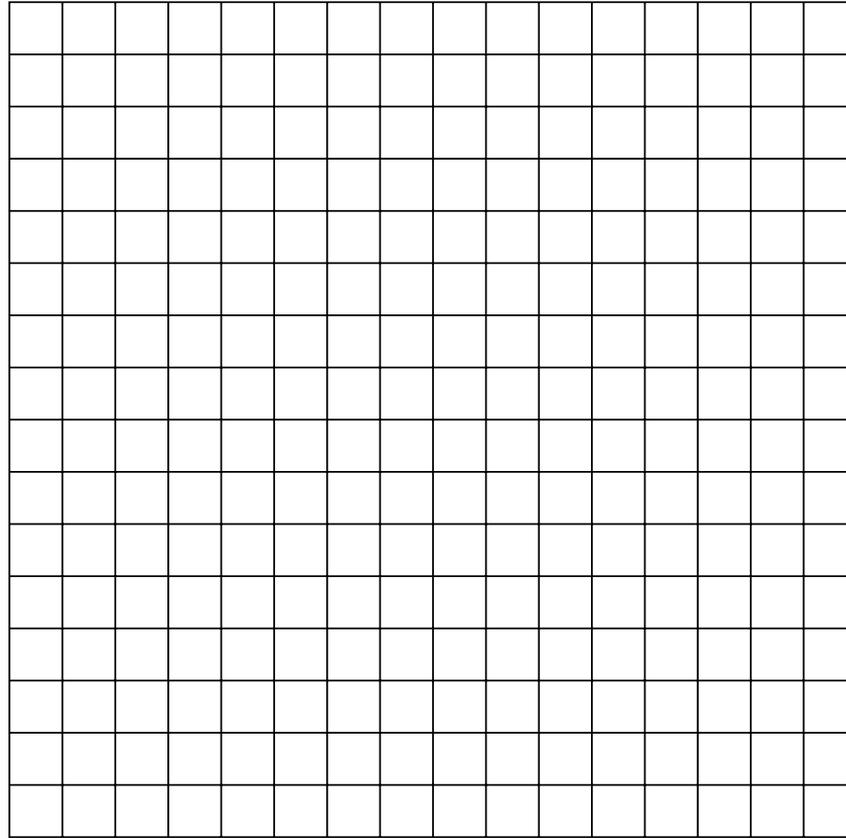
Daten in der Verkehrsplanung

- Räumliche Gliederung von Daten
- Kategorien von Daten
- Inhaltliche Differenzierung von Daten

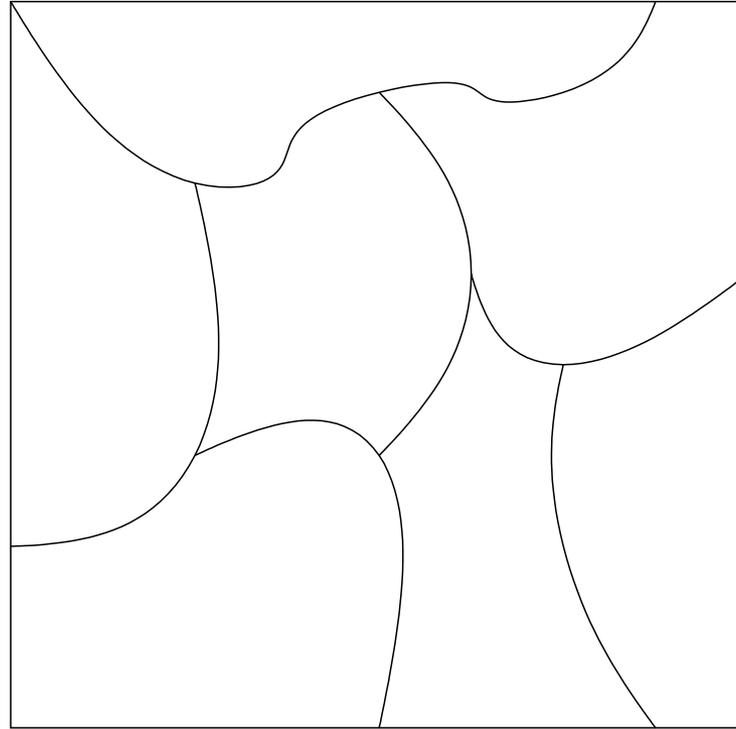
Daten sollten gut dokumentiert sein. Dazu gehört die Erhebung und "Aufbewahrung" von Metadaten. Beispiele für Metadaten sind:

- Räumliche Bezugsebene
- Bezugszeitraum
- Aufschlüsselung der Attributkodierung
- Ggf. Stichprobenauswahl

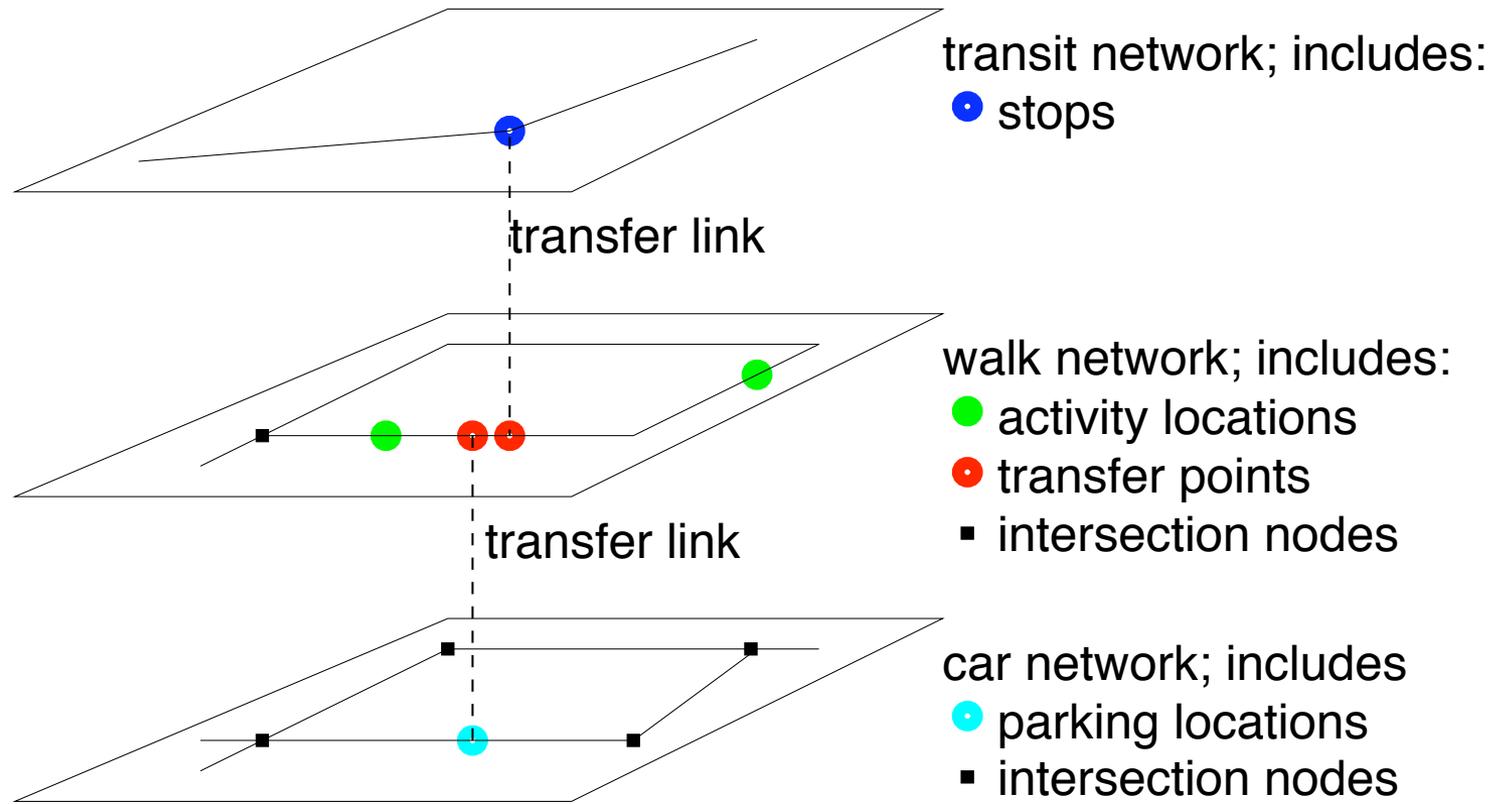
0. Datensysteme



- Raum unterteilt in Zellen gleicher Größe
- Mathematisch und für Computer einfach zu handhaben
- Probleme, wenn die Dichte von Strukturen sehr stark variiert (z.B. stark variierende Bevölkerungsdichte)



- Zonen begrenzt durch "Polylinien"
- Halbwegs beliebige Formen
- Größe kann angepasst werden (z.B. Zonen so dass Anzahl Einwohner pro Zone ungefähr konstant)

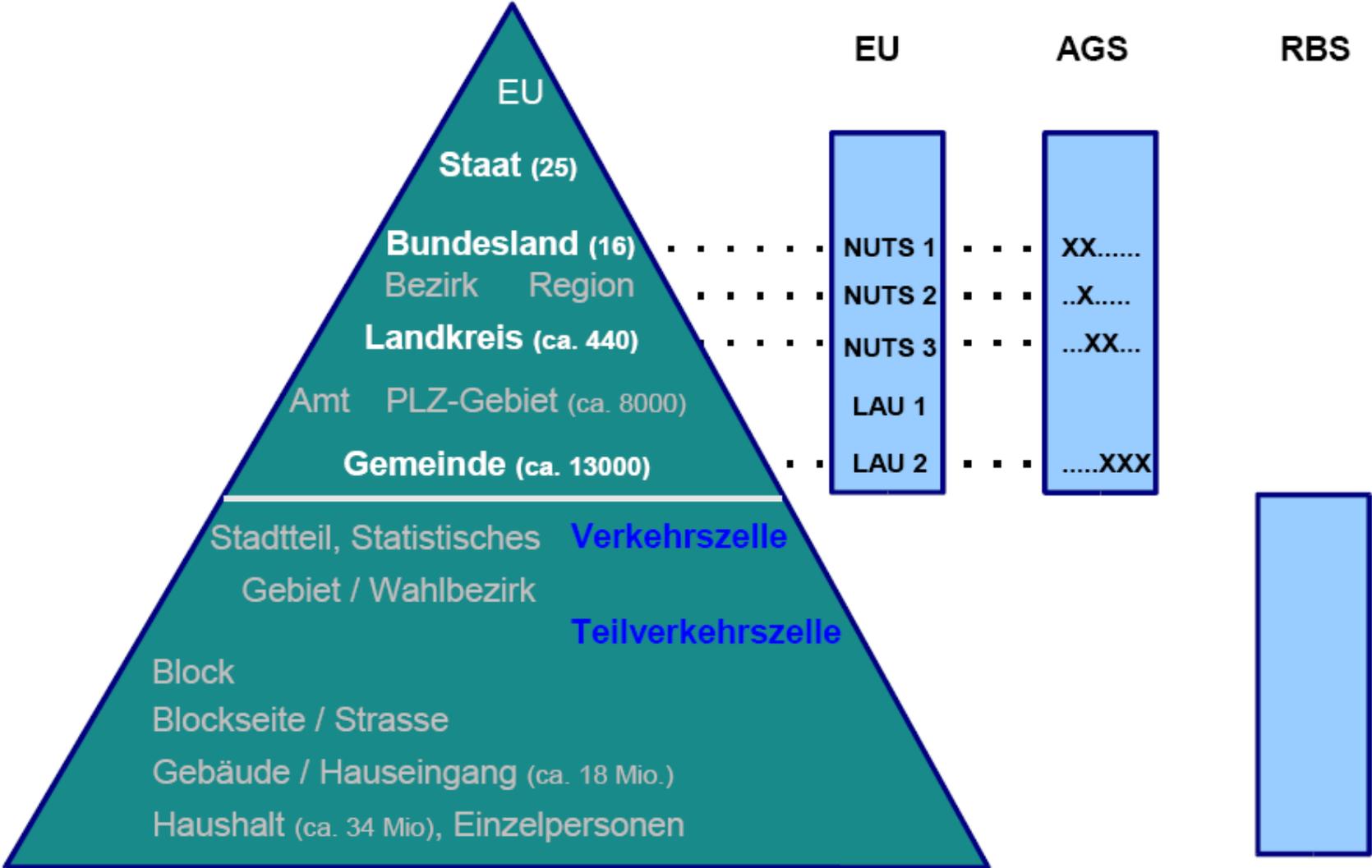


- Knoten und Kanten
- Verschiedene Netze übereinander (auch: Strom, Wasser, usw.)

A. Räumliche Gliederung von Daten

Gründe für die Verwendung von Gebietsgliederungen:

- Planungsverfahren spielen sich auf verschiedenen Ebenen ab.
- Datenschutz und Kosten der Datenerhebung erfordern Reduktion des Detaillierungsgrades.
- Daten werden auf verschiedene räumliche Ebenen aggregiert.
- Räumliche Ebenen bauen aufeinander auf.
- Hierarchische Schlüssel identifizieren die Gebiete und ordnen sie in die verschiedenen Ebenen ein.



EU: NUTS + LAU

Nomenclature des **u**nités **t**erritoriales **s**tatistiques

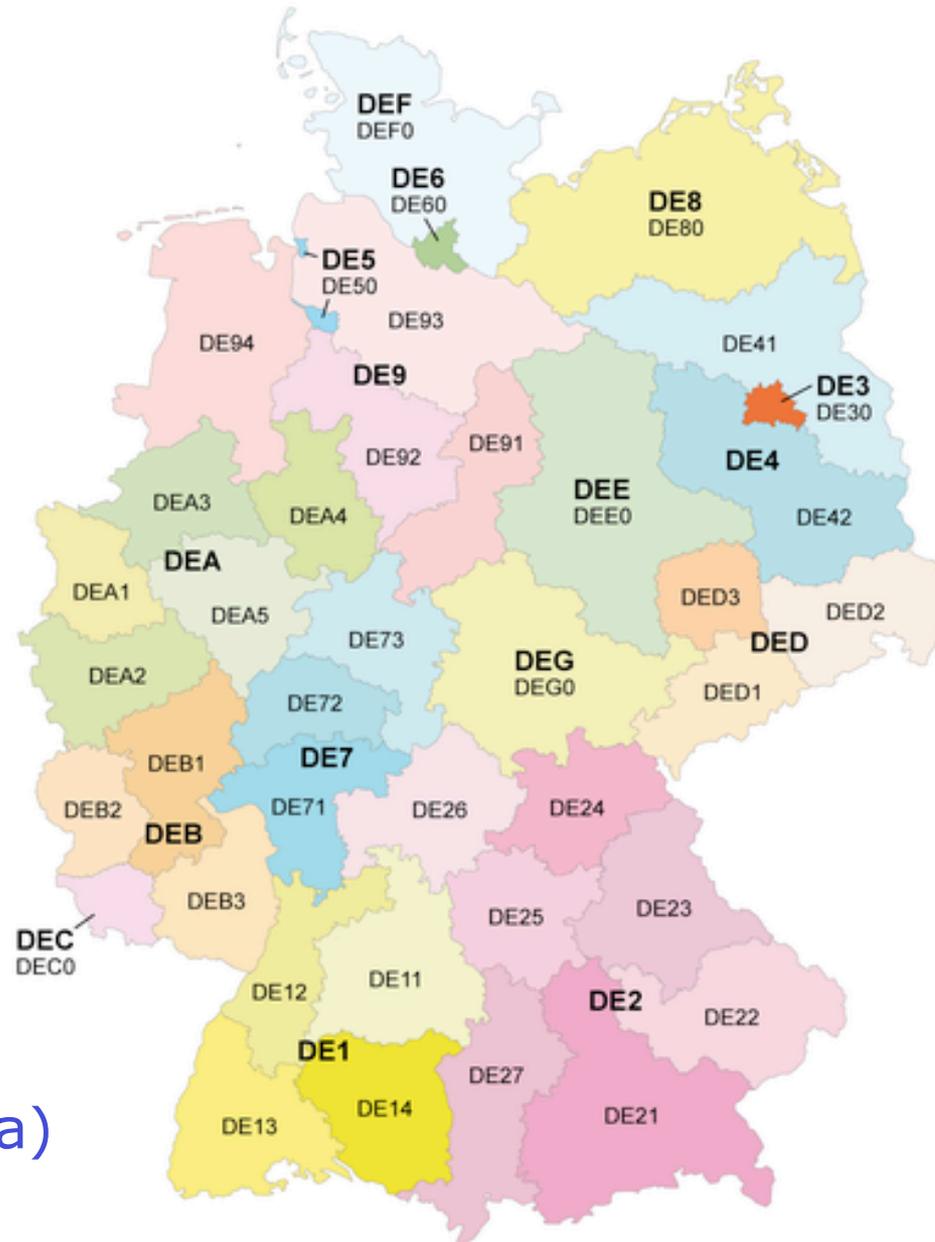
- NUTS0: Nationalstaaten
- NUTS1: Bundesländer
- NUTS2: "Regionen"
- NUTS3: Kreisebene

Local **A**ministrative **U**nits

- LAU1: "Verwaltungsgemeinschaften"
- LAU2: Gemeinden
-

Deutschland

NUTS-Ebenen 1 und 2



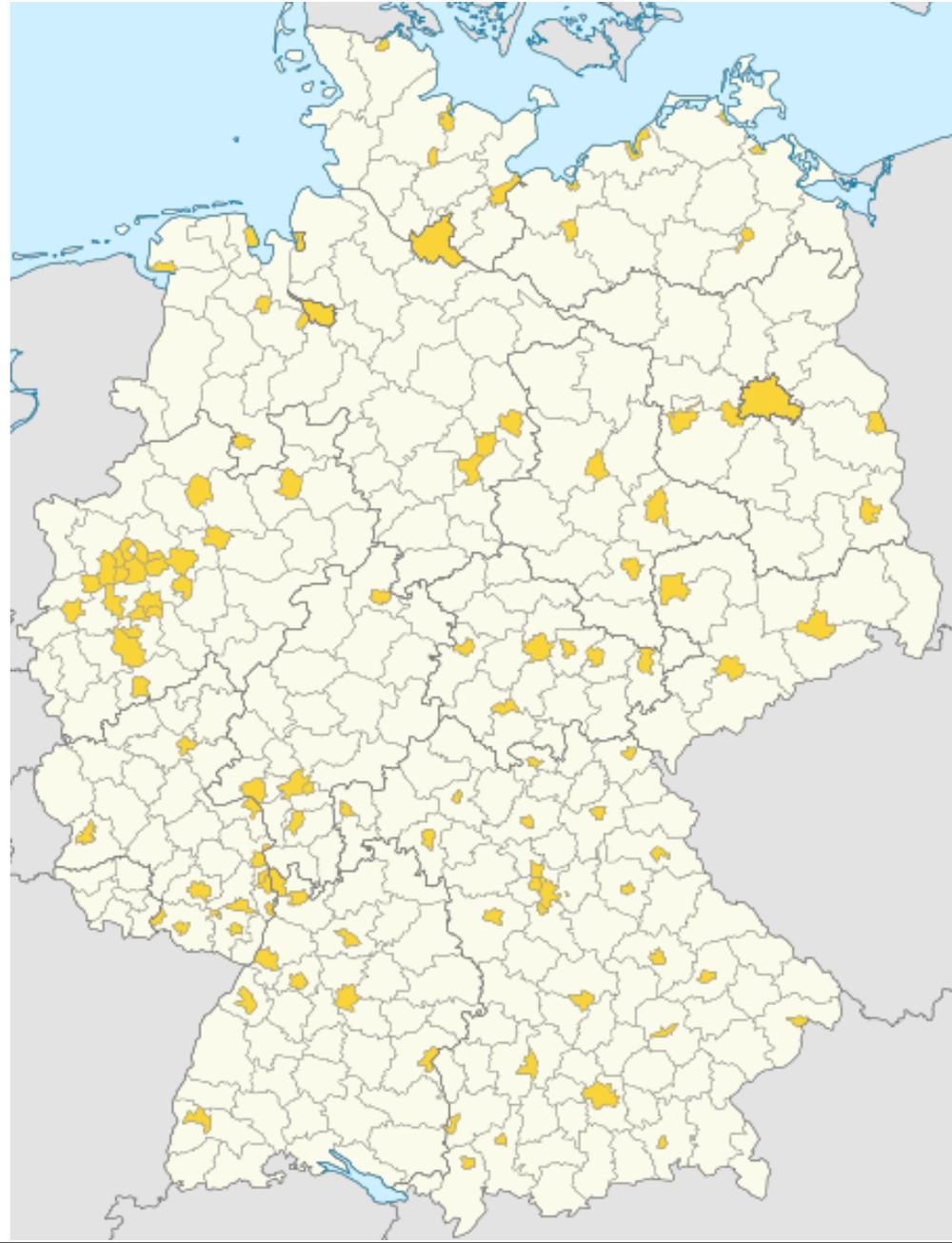
- DE1 Baden-Württemberg**
- DE11 Stuttgart
- DE12 Karlsruhe
- DE13 Freiburg
- DE14 Tübingen
- DE2 Bayern**
- DE21 Oberbayern
- DE22 Niederbayern
- DE23 Oberpfalz
- DE24 Oberfranken
- DE25 Mittelfranken
- DE26 Unterfranken
- DE27 Schwaben
- DE3 Berlin**
- DE30 Berlin
- DE4 Brandenburg**
- DE41 Brandenburg - Nordost
- DE42 Brandenburg - Südwest
- DE5 Bremen**
- DE50 Bremen
- DE6 Hamburg**
- DE60 Hamburg
- DE7 Hessen**
- DE71 Darmstadt
- DE72 Gießen
- DE73 Kassel
- DE8 Mecklenburg-Vorpommern**
- DE80 Mecklenburg-Vorpommern
- DE9 Niedersachsen**
- DE91 Braunschweig
- DE92 Hannover
- DE93 Lüneburg
- DE94 Weser-Ems
- DEA Nordrhein-Westfalen**
- DEA1 Düsseldorf
- DEA2 Köln
- DEA3 Münster
- DEA4 Detmold
- DEA5 Arnsberg
- DEB Rheinland-Pfalz**
- DEB1 Koblenz
- DEB2 Trier
- DEB3 Rheinhessen-Pfalz
- DEC Saarland**
- DEC0 Saarland
- DED Sachsen**
- DED1 Chemnitz
- DED2 Dresden
- DED3 Leipzig
- DEE Sachsen-Anhalt**
- DEE0 Sachsen-Anhalt
- DEF Schleswig-Holstein**
- DEF0 Schleswig-Holstein
- DEG Thüringen**
- DEG0 Thüringen

(Quelle: wikipedia)

Kreisebene (Landkreise, Kreise, kreisfreie Städte)

14

(= NUTS3)



(Quelle: wikipedia)

Deutschland:

- Amtlicher Gemeindeschlüssel (AGS) / Gemeindekennziffern (GKZ)
- 8-stelliger hierarchischer Schlüssel
- LLBKKGGG (L=Land, B=Bezirk, K=Kreis, G=Gemeinde)
- Berlin = 11 0 00 000, Schönfeld = 12 0 60 208

Probleme:

- Der deutsche Schlüssel hat eher nichts mit dem EU-Schlüssel zu tun.
- Berlin ist eine einzige Gemeinde; Brandenburg hat 419 Gemeinden. → Unterhalb NUTS2 wird die Auflösung sehr unbalanciert und oft zu schlecht (keine Binnenstruktur in Berlin)

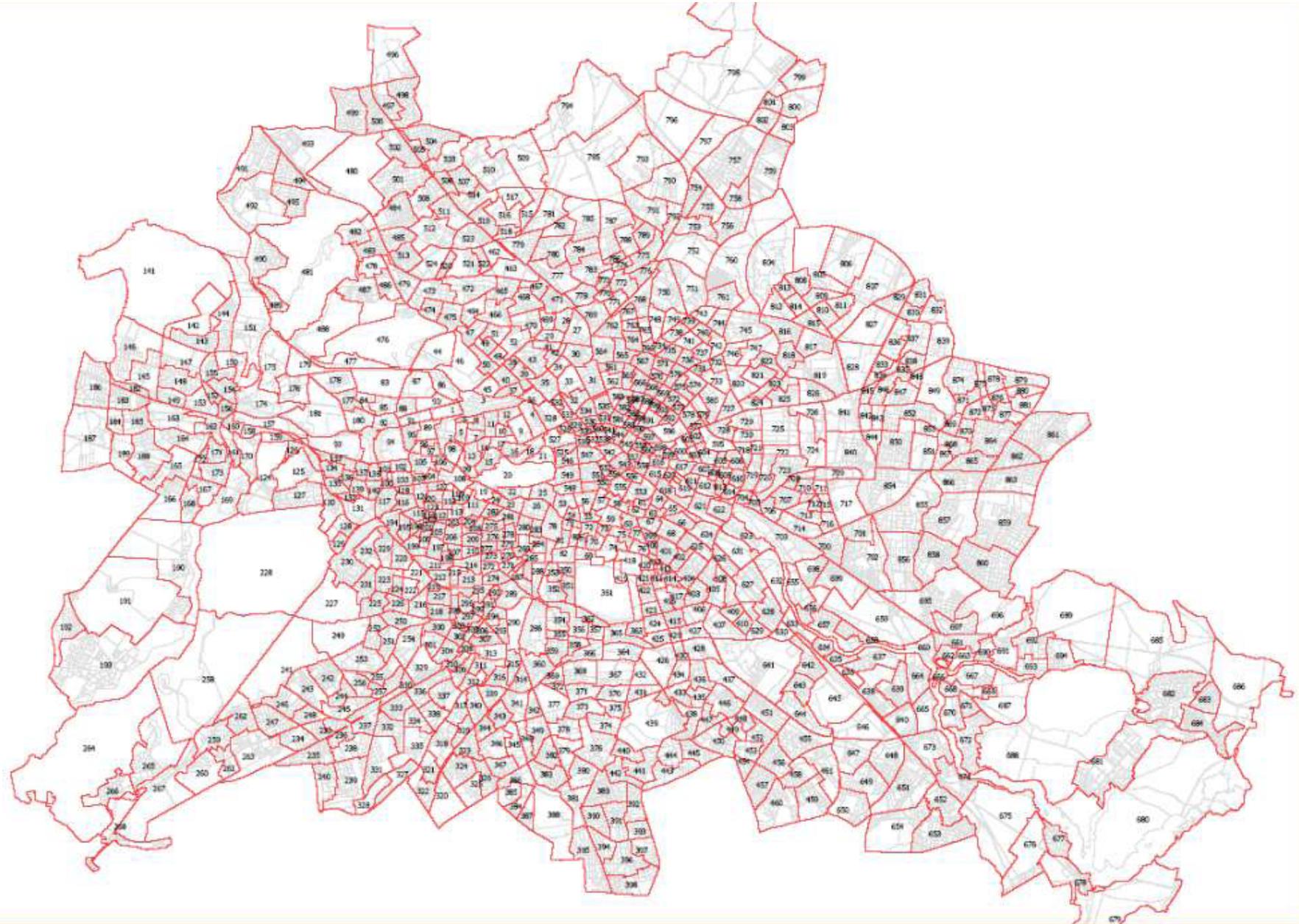
Berlin: Raumbezugssystem (RBS)

Weitere Untergliederung der Gemeindeebene

- Block (ca. 12000), Straße, Blockseite, (Adresse)
- Bezirk, Statistisches Gebiet, Verkehrszelle, Teilverkehrszelle, weitere Einteilungen...

D.h. Berlin hat für sich selber weitere Einteilungen ...

... aber die werden weder von EU noch vom Statistischen Bundesamt verwendet.



(nicht nur in Berlin)

Einteilung des Planungsgebietes in möglichst ‚homogene‘ Gebiete

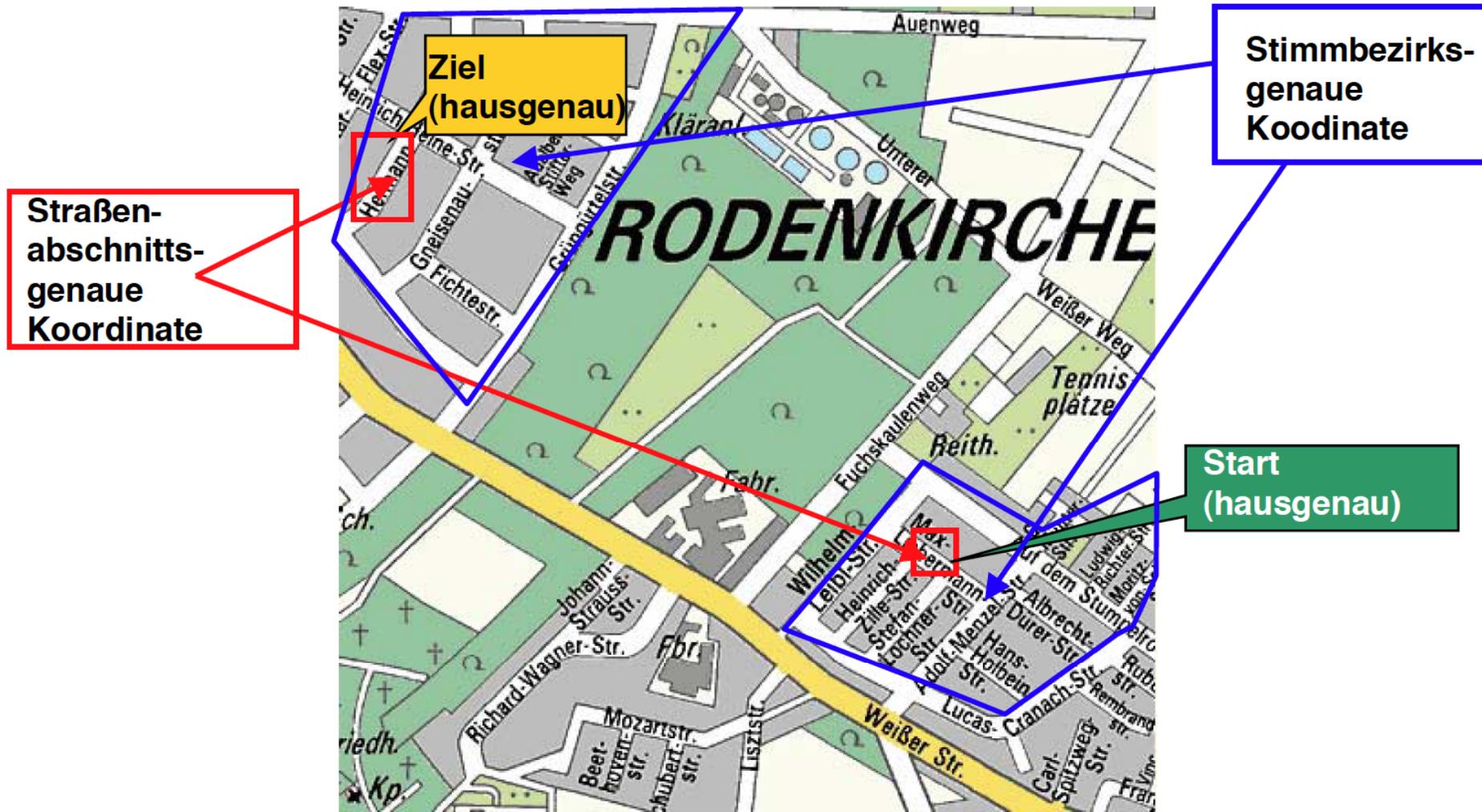
Grundlage für viele „disaggregierte“ Planungsmodelle (4-Stufen-Prozess)

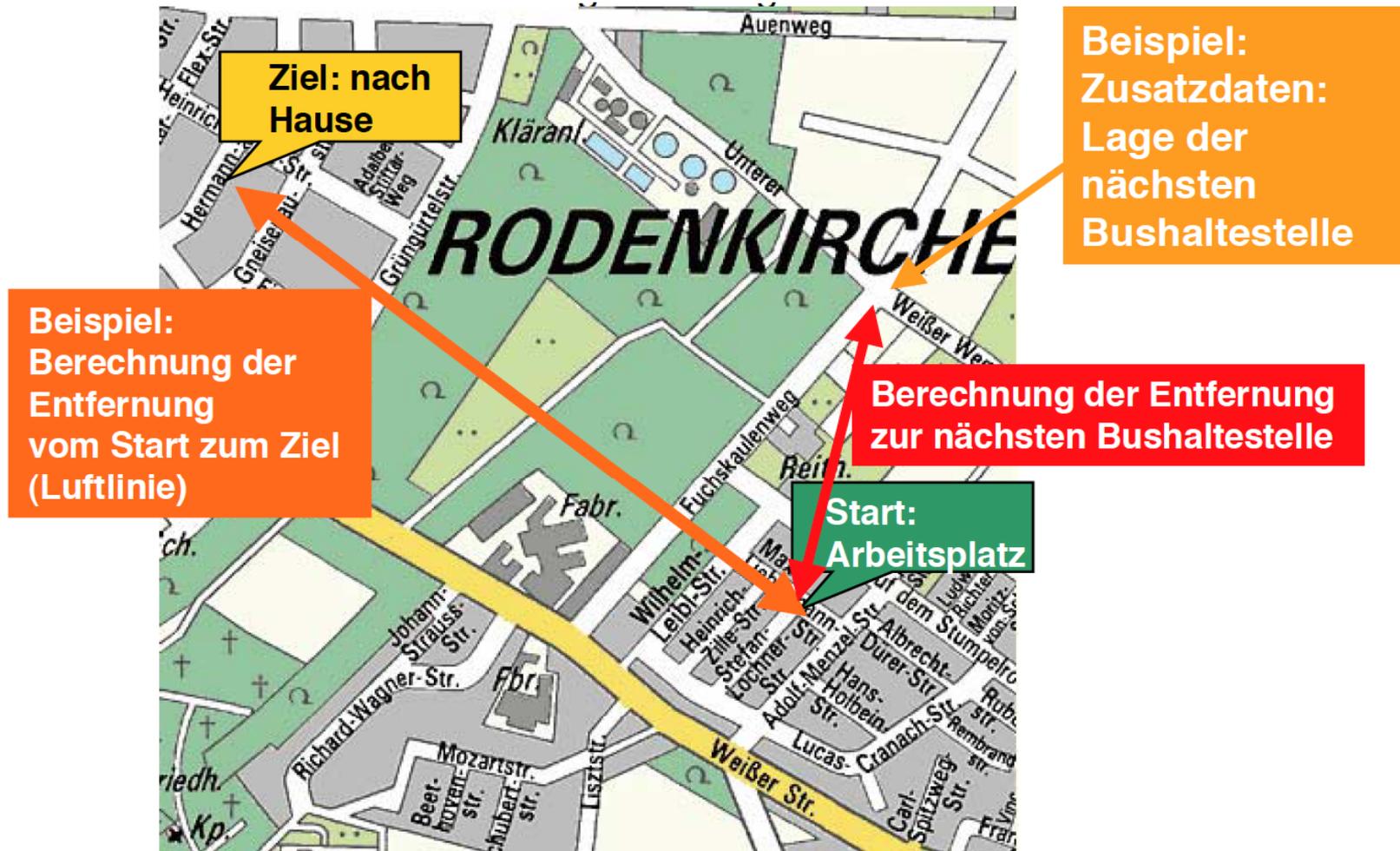
Verkehrsnachfrage wird als Ströme zwischen diesen Gebieten betrachtet (-> Fahrtenmatrizen, OD-Matrizen)

Richtgröße: 3000 – 10000 Einwohner pro Verkehrszelle (Lohse)

Berlin:

- ca 350 Verkehrszellen inkl. Umland
- weiter untergliedert in ca. 1000 Teilverkehrszellen





Entf. zur Bushaltestelle benötigt genaue Geokodierung.

B. Kategorien von Daten

1. Amtliche Register
2. Amtliche Statistiken
3. Öffentliche Erhebungen / Erhebungen von Verbänden
4. Kommunale Daten
5. Daten privater Anbieter

- Register zur Durchführung staatlicher Exekutivverfahren
- Zugang nicht oder nur sehr beschränkt möglich
- *Häufig Grundlage zur Stichprobenziehung für andere Erhebungen*

→ z.B. Einwohnerregister, KFZ-Register, Steuer

- Bundes- bzw. EU-weite Erhebung
- Verantwortlich: Bundesländer
- Feste Themen, Klassifikationen, Erhebungszyklen
- Daten stehen z.T. im WWW zur Verfügung, z.T. nur auf Bestellung / kostenpflichtig

→ z.B. Bevölkerung, Energieversorgung, Übernachtungen

- www.statistik-berlin-brandenburg.de/
- www.destatis.de

- i.d.R. bundesweit
- häufig im Auftrag von Ministerien
- Durchführung / Auswertung durch Forschungsinstitute bzw. private Unternehmen
- Zugang unterschiedlich

→ z.B. MiD (Kontiv), Mobilitätspanel

- www.dlr.de/cs ("Clearingstelle des DLR", notfalls per google)
- www.mobilitaet-in-deutschland.de/
- mobilitaetspanel.ifv.uni-karlsruhe.de/

- Verschiedenste Themen, Datengrundlagen
- Daten liegen nur für begrenzten Raum vor

→ z.B. Verkehrszählungen, Planungsnetze, Gewerbestruktur, Schülerzahlen

- www.stadtentwicklung.berlin.de/planen/basisdaten_stadtentwicklung/
- <http://gvp2025.net/> (Berlin)
- <http://www.lvp.nrw.de/>

(Prognosen und Modellresultate sind keine "Daten", aber die die Webseiten enthalten auch Hinweise auf Eingangsdaten.)

- Bieten unterschiedlichste Daten an
- Themen, die in der amtlichen Statistik nicht vorhanden sind,...
- teilweise erheblich feinere räumliche Auflösungen / Merkmalsdifferenzierungen als die amtliche Statistik
- i.d.R. sehr kostenintensiv

→ z.B. Straßennetze: Navteq, TeleAtlas

Demographische Daten: GfK, macrom, Deutsche Post Direkt

C. Inhaltliche Differenzierung von Daten

1. Allgemeine Daten zum Verkehr
2. Verkehrliche Infrastruktur
3. Raumstruktur, Einrichtungen (Ziele)
4. Demographische Daten
5. Verhaltensdaten
6. Zähldaten (Status quo oder Zeitreihen)

"Verkehr in Zahlen"

- Jährlich vom DIW im Auftrag BMVBS
- Bevölkerung, Erwerbstätige, private Haushalte
- Kfz-Dichte, Straßenbelastung
- Verkehrsausgaben, Transportbilanz
- Kfz-Bestand, Neuzulassung, Fahrleist.
- Grenzüberschreitender Verkehr
- Straßenverkehr nach Ländern

[[pdf: DIW2000...]]



Verkehrsnetze

- Straßen, Kreuzungen, Schienenwege, Wasserstraßen
 - Länge
 - zugelassene Verkehrsmittel
 - Durchfahrtsbeschränkungen
 - Geschwindigkeiten
 - Kapazitäten
- Haltestellen

Ruhender Verkehr

- Parkhäuser, Stellflächen, Stellplätze, ...

Verkehrsmittel / Betriebsdaten

- Fahrzeugeigenschaften

- Transportkapazität
- Geschwindigkeit

- Fahrzeugbestand

- Fahrplandaten

⇒ *Nötig für die Ermittlung von Erreichbarkeiten, Raumwiderständen und Verkehrsbelastungen (Umlegung)*

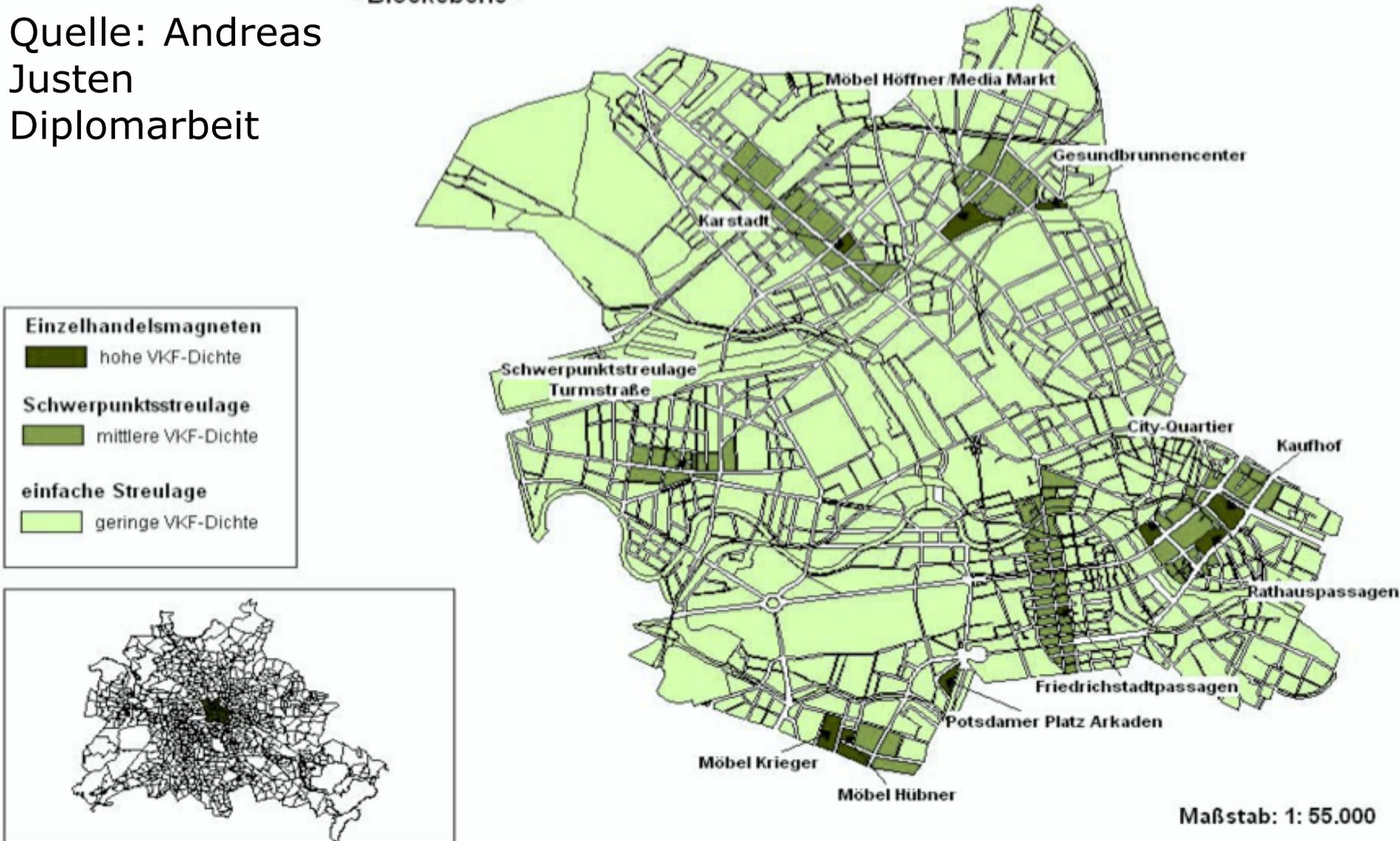
Standorte und Größe von

- Öffentliche Einrichtungen
 - Gewerbe (Arbeitsstätten)
 - Schulen, Kindertagesstätten, Universitäten
 - Einkaufszentren und Einzelhandel
 - Sehenswürdigkeiten
- ⇒ *Verwendung für die Ermittlung von Zielpotenzialen und Verkehrsströmen*

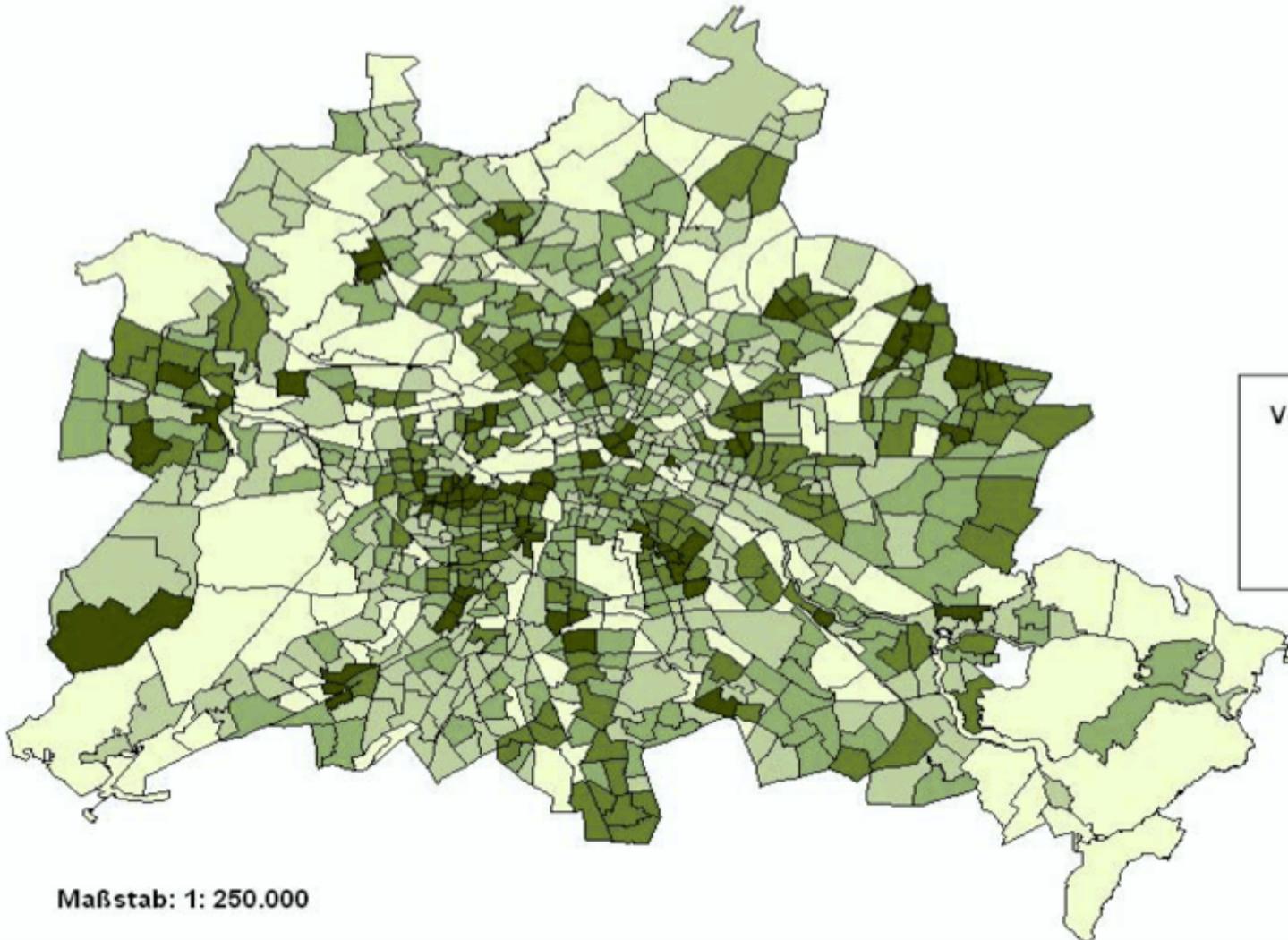
Verkaufsraumflächen im Bezirk Berlin-Mitte

- Blockebene -

Quelle: Andreas
Justen
Diplomarbeit



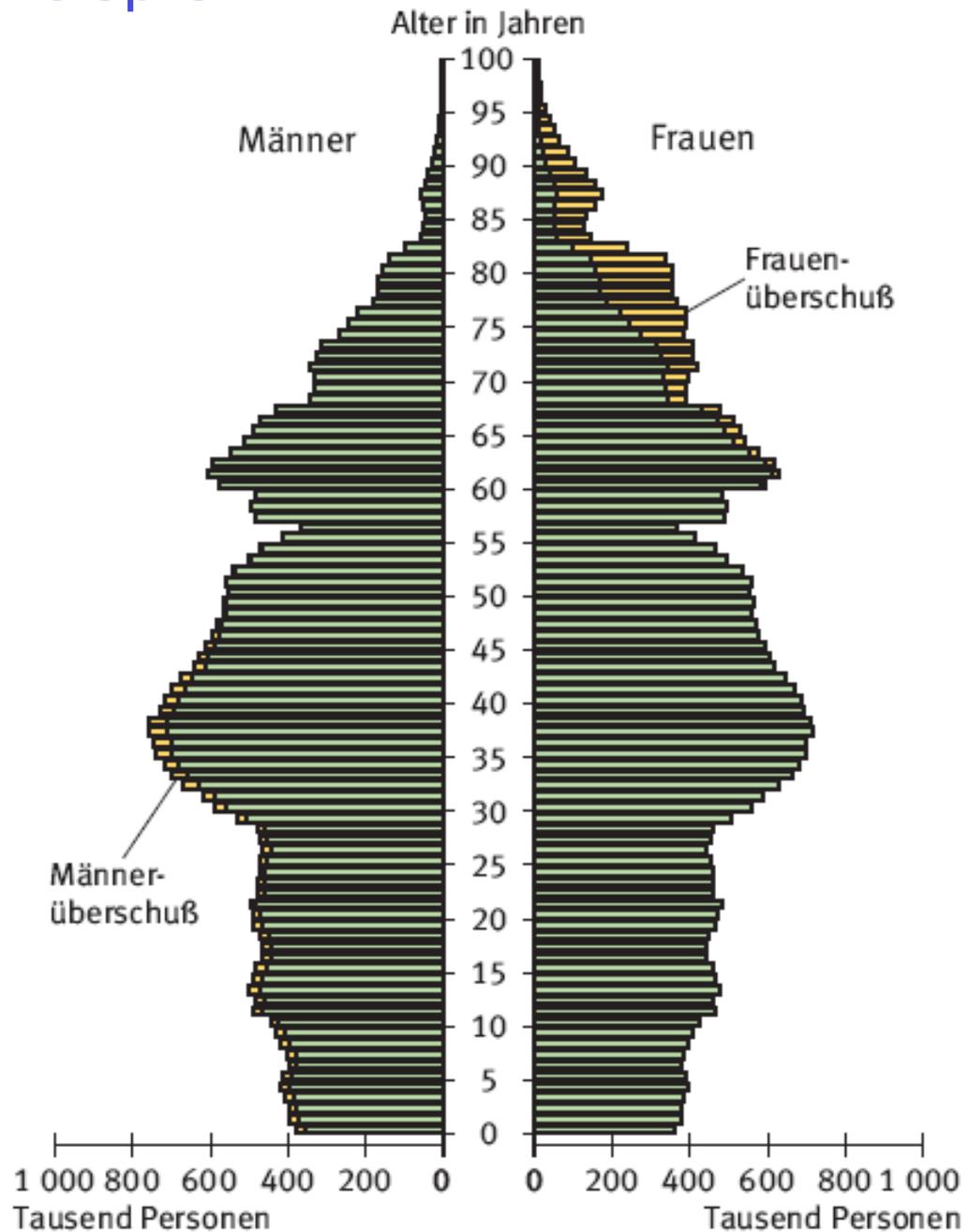
Verkaufsraumflächen in Berlin
- Teilverkehrszellenebene -



Maßstab: 1: 250.000

Quelle: Andreas
Justen
Diplomarbeit

- Amtliche Statistik (EU, Bund, Land) / Daten privater Anbieter
 - Teilweise kostenlos, teilweise kostenpflichtig
 - Immer auf Gebiete aggregiert
 - Je feiner aufgelöst, desto teurer
 - Nicht auf beliebig feiner Ebene verfügbar
- ⇒ *Erforderlich für die Ermittlung von Quellpotenzialen und Verkehrsströmen*



Quelle: Statistisches Bundesamt, Hrsg. Bevölkerung Deutschlands bis 2050. 10. koordinierte Bevölkerungsvorausberechnung., Wiesbaden, 2003, S. 30

<http://www.destatis.de/bevoelkerungspyramide/>

Bevölkerung:

- Einwohner, Altersstruktur, Geschlecht

Haushalte:

- Haushaltsstruktur, Alter d. Haushaltsvorstandes, Kaufkraft

Erwerbstätige:

- Arbeitnehmer, Arbeitslose, Arbeiter / Beamte / Angestellte / Selbstständige

Ausbildung:

- Höchster Schulabschluss u.ä.

Wirtschaft:

- Betriebsgrößen, Branchenstruktur

Flächennutzung:

- bebaute zu unbebauter Fläche, Gebäudetypen, Nutzungsarten

Aufwendige Erhebungen (Befragungen) mit repräsentativer Stichprobe (als Einmalerhebungen oder Panel)

Siehe Vorlesungsunterlagen Erhebungsmethoden

Verwendung:

- Verkehrsursachenforschung
- Grundlage Hochrechnungen für aggregierte Verkehrsdaten (Verkehrsleistung etc.)
- Bestimmung von Verhaltensvariablen (z.B. Verkehrsmittelwahl)
- Gewinnung von Modellparametern (z.B. zu Verkehrsmittelnutzung, Reiseweiten, Preis-Elastizitäten, Aktivitäten, Wege)

	A	B	C	D	E	F	M	N	U	V	AC	AD	
1	caseid	h01	h001c	h02	h1sex	h2sex	h1alter	h2alter	h1beruf	h2beruf	h1besch	h2besch	h3
2	709	2	1.00	3	2.00	1.00	35.00	38.00	2.00	1.00	-7.00	1.00	
3	711	2	1.00	2	2.00	1.00	46.00	44.00	1.00	1.00	2.00	1.00	
4	713	1	2.00	1	1.00	-21.00	71.00	-21.00	-6.00	-21.00	-6.00	-21.00	
5	720	1	1.00	1	1.00	-21.00	44.00	-21.00	1.00	-21.00	1.00	-21.00	
6	722	2	1.00	5	2.00	1.00	35.00	35.00	1.00	1.00	2.00	1.00	
7	726	2	1.00	2	1.00	2.00	60.00	50.00	1.00	1.00	1.00	2.00	

Caseid Identifikationsnummer des Haushalts

Druckformat: F11

Schreibformat: F11

H01 Personen im Haushalt

Druckformat: F4

Schreibformat: F4

Wert Label

- 1 Lebe allein
- 2 Lebe mit anderen Personen im Haushalt
- 3 Lebe (allein) in einem Wohnheim
- 7 Verweigert
- 9 Keine Angabe

H1BERUF Berufstätigkeit der Person 1
Druckformat: F8.2
Schreibformat: F8.2
Missing-Bereich: -80,00 bis -5,00

Wert	Label
-80,00 M	In Papi nicht erhoben (PAPI)
-50,00 M	In Papi nicht erhoben (Switcher)
-6,00 M	CATI: Befragte über 69 Jahre werden nicht befragt
-5,00 M	CATI: Befragte unter 15 Jahre werden nicht befragt
1,00	Ja
2,00	Nein
7,00	Verweigert
8,00	Weiß nicht
9,00	Keine Angabe

H1BESCH Beschäftigungsart der Person 1
Druckformat: F8.2
Schreibformat: F8.2
Missing-Bereich: -8,00 bis -5,00

Wert	Label
-8,00 M	CATI: Verweigert/Weiß nicht zur Berufstätigkeit
-7,00 M	CATI: Befragte nicht berufstätig
-6,00 M	CATI: Befragte über 69 Jahre werden nicht befragt
-5,00 M	CATI: Befragte unter 15 Jahre werden nicht befragt
1,00	Vollzeit
2,00	Teilzeit
3,00	Auszubildende/r
4,00	Nicht berufstätig
6,00	Doppelnennung
7,00	Verweigert
8,00	Weiß nicht
9,00	Keine Angabe

	A	B	C	D	E	F	G	H
1	caseid	pid	p01	p02	p021	p022	p033	p031
2	709	1	1.00	2.00	-21.00	-21.00	5.00	1.00
3	709	3	-400.00	1.00	-400.00	-400.00	-400.00	-400.00
4	711	1	6.00	2.00	-21.00	-21.00	5.00	2.00
5	711	2	1.00	2.00	-21.00	-21.00	3.00	1.00
6	713	1	2.00	1.00	1.00	1.00	5.00	2.00
7	720	1	1.00	1.00	1.00	1.00	1.00	2.00
8	722	1	1.00	1.00	1.00	1.00	2.00	1.00
9	722	3	-400.00	1.00	-400.00	-400.00	-400.00	-400.00
10	722	4	-400.00	1.00	-400.00	-400.00	-400.00	-400.00

P01 Wie oft können Sie als Fahrer/in über ein Auto verfügen?
 Druckformat: F8.2
 Schreibformat: F8.2

Wert	Label
-400,00	Proxy kleiner gleich 13 Jahre werden nicht befragt
-300,00	Selbst kleiner gleich 13 Jahre werden nicht befragt
-200,00	Proxy ab 14 Jahre werden nicht befragt
-150,00	Im PAPI-Kinder-Fragebogen nicht erhoben
-20,00	CATI:Personen < 18(Alter aus HH-Matrix) werden nicht befragt
1,00	Jederzeit
2,00	Gelegentlich
3,00	Ausnahmsweise
4,00	Gar nicht
6,00	habe keinen Führerschein
7,00	Verweigert
8,00	Weiß nicht
9,00	Keine Angabe

P021 Wo steht Ihr hauptsächlich genutztes Fahrrad üblicherweise ü
Druckformat: F8.2
Schreibformat: F8.2
Missing-Bereich: -400,00 bis -21,00

Wert	Label
-400,00 M	Proxy kleiner gleich 13 Jahre werden nicht befragt
-300,00 M	Selbst kleiner gleich 13 Jahre werden nicht befragt
-200,00 M	Proxy ab 14 Jahre werden nicht befragt
-150,00 M	Im PAPI-Kinder-Fragebogen nicht erhoben
-100,00 M	Im PAPI-Erwachsenen-Fragebogen nicht erhoben
-22,00 M	Verweigert/Weiß nicht zum Besitz eines Fahrrads
-21,00 M	Personen ohne Fahrrad werden nicht befragt
1,00	In ein abgeschlossenem Raum unmittelbar im Haus
2,00	In einem abgeschlossenem Raum entfernt vom Haus
3,00	In nicht abgeschlossenem Raum
4,00	Unterschiedlich
5,00	Anderes
7,00	Verweigert
8,00	Weiß nicht

◇	A	B	C	D	E	F	G	H	v
1	caseid	pid	weg	w01	w111	w112	w04	w041	v
2	709		1	1.00	1.00	-73.00	1.00	5.00	-1.00
3	709		1	2.00	-39.00	-73.00	-72.00	4.00	501.00
4	709		1	3.00	-39.00	-73.00	-72.00	8.00	-1.00
5	709		3	1.00	1.00	-73.00	1.00	31.00	-1.00
6	709		3	2.00	-39.00	-73.00	-72.00	8.00	-1.00
7	711		1	1.00	1.00	-73.00	1.00	1.00	-1.00
8	711		1	2.00	-39.00	-73.00	-72.00	8.00	-1.00

CASEID Identifikationsnummer des Haushalts
 Druckformat: F8
 Schreibformat: F8

PID Identifikationsnummer der Person im Haushalt
 Druckformat: F11
 Schreibformat: F11

WEG Identifikationsnummer des Wegs
 Druckformat: F8.2
 Schreibformat: F8.2

Wegzweck
Druckformat: F8.2
Schreibformat: F8.2

Wert	Label
1,00	Erreichen des Arbeitsplatzes
2,00	Dienstlich/Geschäftlich
3,00	Erreichen der Ausbildungsstätte
4,00	Einkauf
5,00	Private Erledigungen
6,00	Bringen/Holen von Personen
7,00	Freizeitaktivität
8,00	Nach Hause
9,00	Rückweg vom vorherigen Weg
10,00	Anderes
11,00	Begleitung Erwachsener
31,00	(Vor-) Schule
32,00	Kita/Kindergarten
40,00	PAPI: Sport
41,00	PAPI: Freunde besuchen/treffen
42,00	PAPI: Unterricht (nicht Schule)
97,00	Verweigert
98,00	Weiß nicht
99,00	Keine Angabe

W041

Hauptzweck des Einkaufswegs
Druckformat: F8.2
Schreibformat: F8.2

Wert	Label
-150,00	Im PAPI-Kinder-Fragebogen nicht erhoben
-100,00	Im PAPI-Erwachsenen-Fragebogen nicht erhoben
-1,00	Kein Einkaufsweg
501,00	Täglicher Bedarf
502,00	Sonstige Waren
503,00	Allgemeiner Einkaufsbummel
504,00	Dienstleistungen
505,00	Bringen/Holen von Personen
506,00	sonstiger Einkaufszweck
999,00	Keine Angabe

Siehe Beispiel letzte Woche

Siehe Beispiele letzte Woche

D. Datenschutz

- Keine Veröffentlichung von Daten, die Rückschlüsse auf Einzelfälle (Individuen / Betriebe) zulassen.
- I.d.R. mindestens 7 Fälle pro Raumeinheit / Merkmalsausprägung.
- Null-Werte heißen nicht immer auch „nicht vorhanden“ !!!
- Problematisch vor allem bei kleinen Raumeinheiten / Merkmalen mit vielen Ausprägungen.

Konsequenz z.B.: MiD Datensatz wird nur ohne Geokodierung herausgegeben.

Rauslösen der Geolokalisierung ist erhebliches Problem für Modellierung und Simulation.

→ benötigen Methoden der automatischen und zertifizierten Anonymisierung

<http://www.geopkdd.eu/>