# Verkehrserzeugung

# VORLESUNGSTEIL

# Einleitung

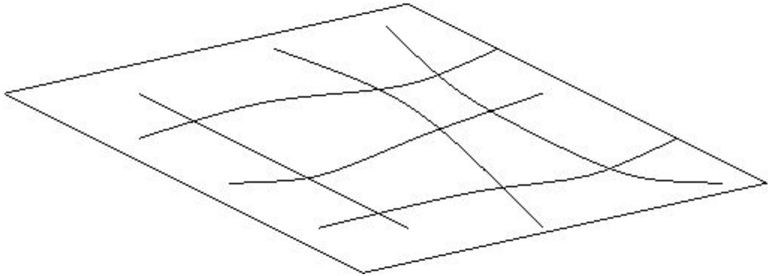
## Erinnerung

Zur Erinnerung: 4-Stufen-Verfahren:

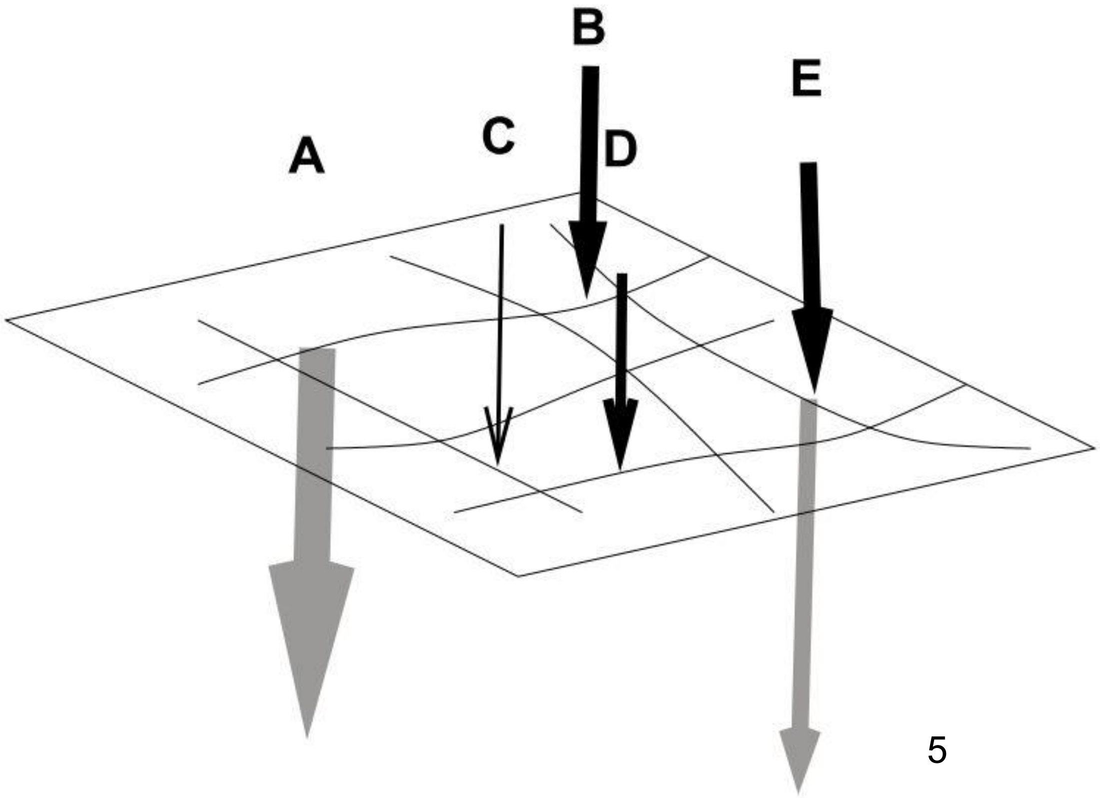
- Verkehrserzeugung
- Verkehrsverteilung
- Verkehrsmittelwahl
- Umlegung

Zur Erinnerung: Verkehrserzeugung = Quellen und Senken von Fahrten.

# Illustration



→

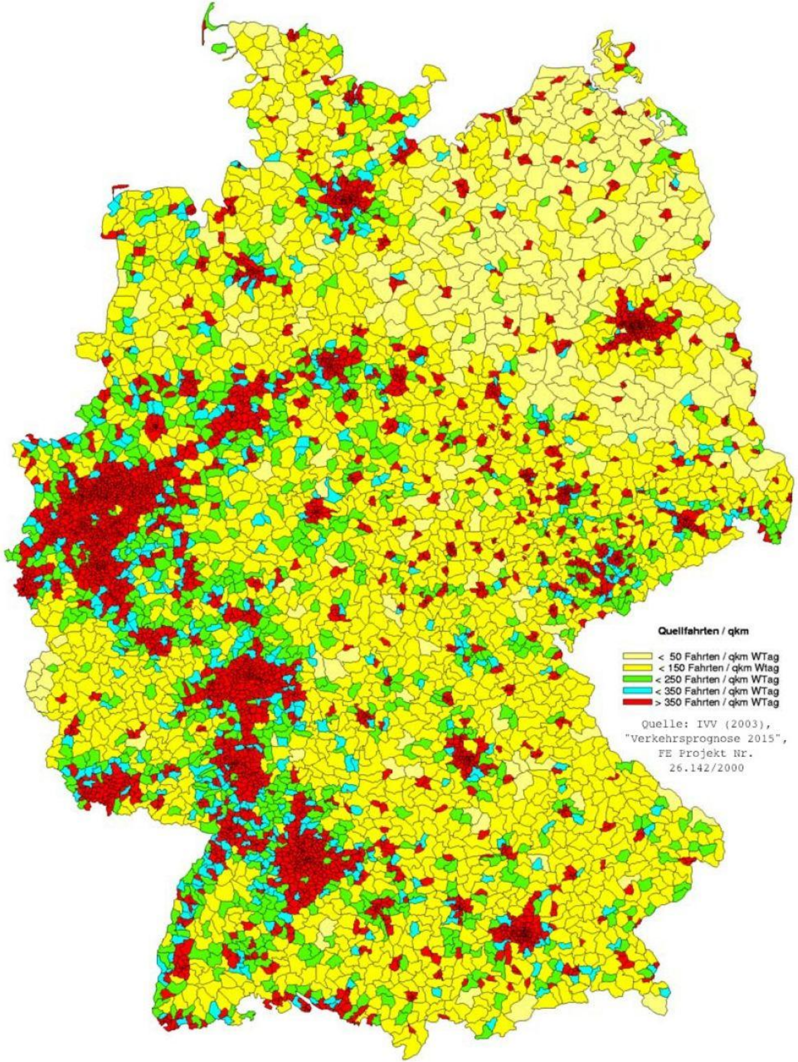


Vektoren der Quellen/Senken

$$\underline{O} = [0, 3, 1, 2, 3] \qquad \underline{D} = [6, 0, 0, 0, 3]$$

Als Randeinträge einer Matrix

| $N_{dp}$ \ $N_{ar}$ | A: 6 | B: 0 | C: 0 | D: 0 | E: 3 |
|---|---|---|---|---|---|
| A: 0 | | | | | |
| B: 3 | | | | | |
| C: 1 | | | | | |
| D: 2 | | | | | |
| E: 3 | | | | | |


Quellfahrten / qkm
< 50 Fahrten / qkm WTag
< 150 Fahrten / qkm Wtag
< 250 Fahrten / qkm WTag
< 350 Fahrten / qkm WTag
> 350 Fahrten / qkm WTag
Quelle: IVV (2003), "Verkehrsprognose 2015", FE Projekt Nr. 26.142/2000

# # Definitionen

Weg, Fahrt, Ortsveränderung (trip): Zwischen zwei Aktivitäten-Orten.

Etappe (leg, stage): Teile von Wegen/Fahrten, die mit dem gleichen Verkehrsmittel zurückgelegt werden.

Reise (journey): Kette von Wegen/Fahrten, die zu Hause beginnt und endet.

Tour (tour): Kette von Wegen/Fahrten, die am selben Ort beginnt und endet.

Route: Kette von Kanten, aus denen sich eine Etappe zusammensetzt.

*Leider kann man sich darauf nicht verlassen.* Obiges von Axhausen; für OW ist “journey” ein “one-way movement”; Lohse hat es wieder anders.

Auch Umgangssprache nicht eindeutig: “Trip to Rome” schließt Rückf. mit ein.

Also bitte immer nachfragen/nachforschen.

Hier bis auf weiteres nur Trip, und wir gehen davon aus, dass jeder Trip nur ein Verkehrsmittel verwendet (somit trip = leg).

Home-based trip: Trips, die zu Hause beginnen + Trips, die zu Hause enden

Generell werden wir sehen, dass die in diesem Kapitel vorgestellte Theorie eigentlich nur dann gut funktioniert, wenn man zwischen zwei Außer-Haus-Aktivitäten immer nach Hause zurückkehrt.

Auswege, später im Semester:

- sogenannte “Quelle-Ziel-Gruppen” (VISEVA)
- aktivitäten-basierte Nachfrageerzeugung

# # Klassifikation von Wegen

Zweck: Arbeit, Ausbildung, Einkaufen, Freizeit, ...

Tageszeit: AM-peak; mid-day; PM-peak; night

Personen-Attribute: Einkommen; Autobesitz; Größe/Struktur des Haushaltes

# Steigerungsfaktor-Modell (growth factor model)

Angenommen, ich habe, für jede Zone $i$, die Quell- und Senkterme $O_i$ und $D_j$ . Dann kann ich für die Vorhersage Steigerungsfaktoren annehmen:

$$O'_i = f_i\, O_i \quad , \quad D'_j = g_j\, D_j \,.$$

Frage ist nun, woher die $f_i$ und $g_j$ kommen sollen, und darauf gibt es in der Tat keine einfachen Antworten.

Andererseits ist die Annahme von Steigerungsfaktoren unaufwändig, und oft hat man Situationen, wo der Aufwand für deutlich aufwändigere Modelle nicht gerechtfertigt erscheint.

Oft gilt das auch für Teile eines Modells. Z.B. “Verkehr im Ausland”.

# Regressionsmodelle

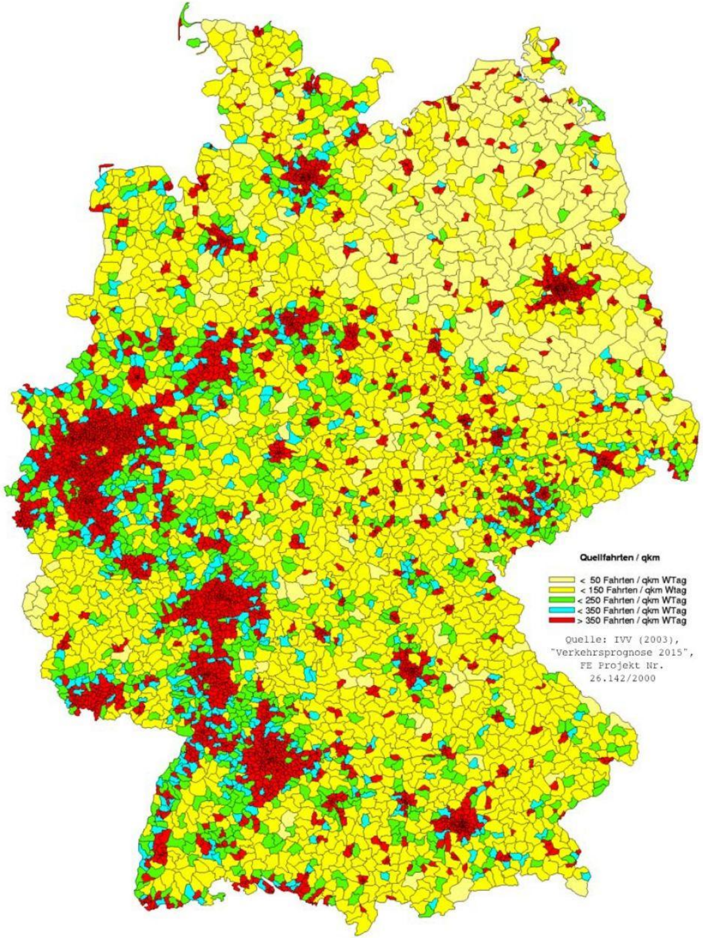
Man kann sich denken, dass die Verkehrserzeugung etwas mit der Einwohnerzahl zu tun haben sollte:

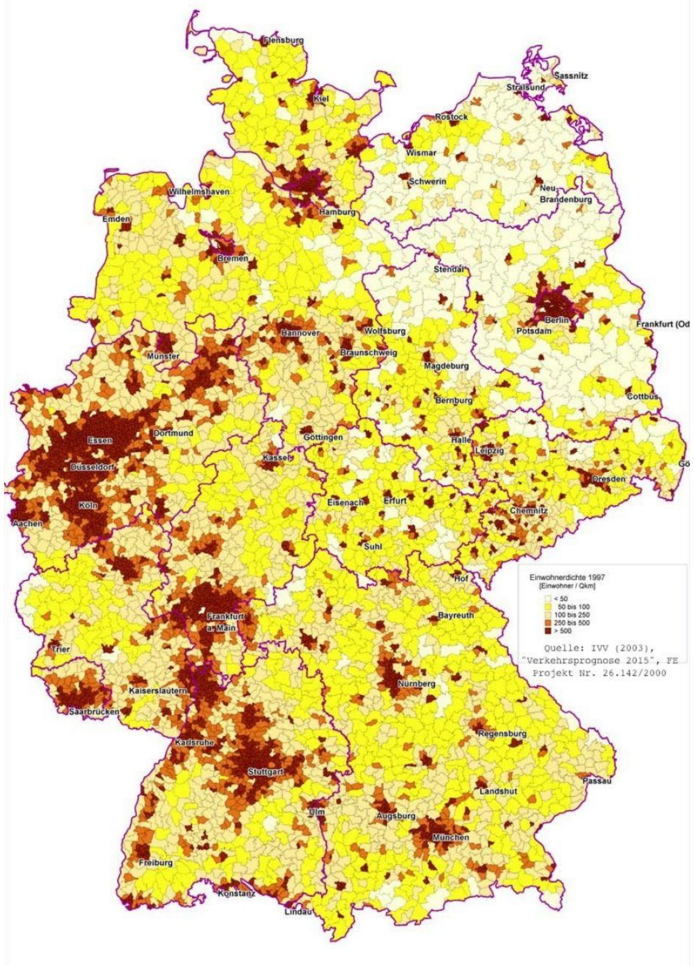
Tendenziell sollten doppelt so viele Einwohner in etwa doppelt so viele Fahrten erzeugen.

In einem zweiten Schritt wird man dann korrigieren wollen (z.B. erzeugen große HHe pro Person weniger Fahrten als kleine HHe).

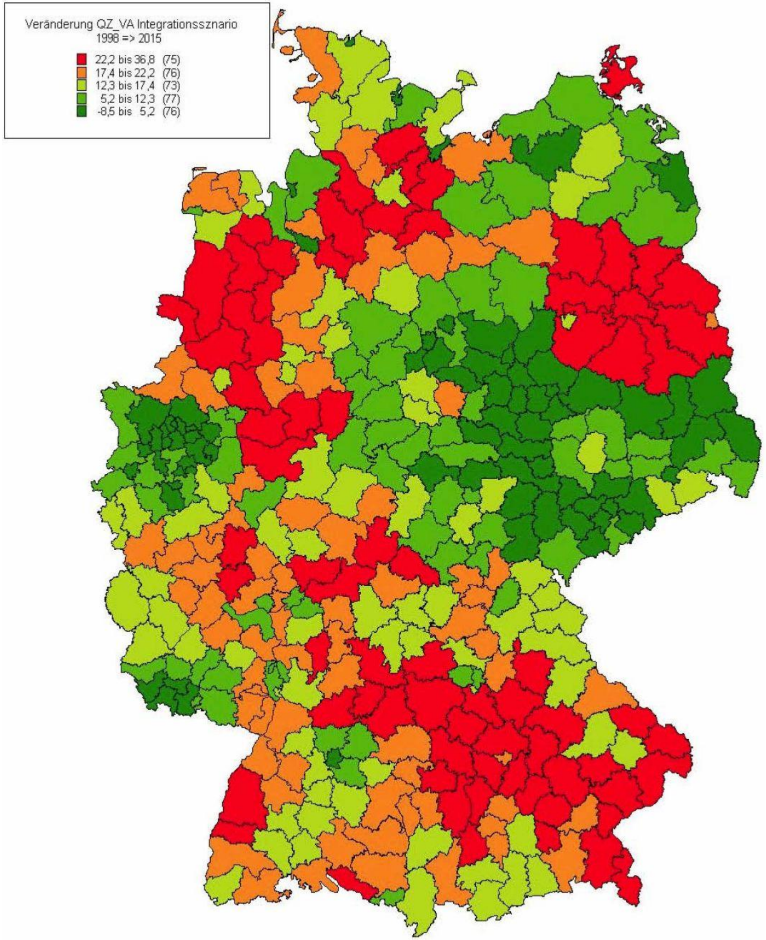
Bild ... links Quellverkehr ... rechts Einwohnerzahl.

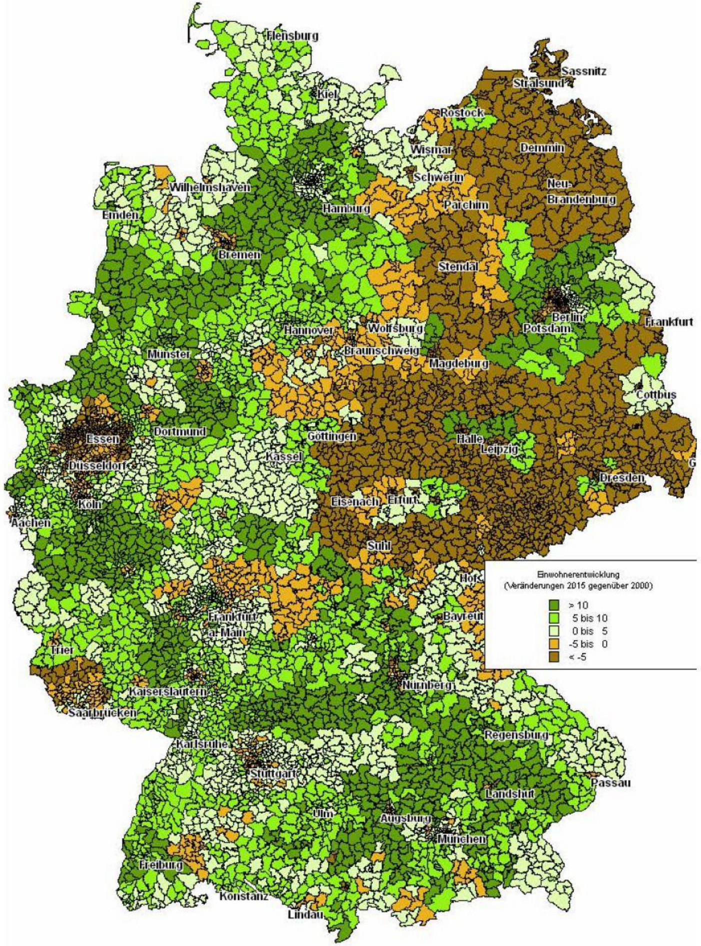
# Verkehrserzeugung (pro qkm) ← Einwohnerdichte



# # Veränd. V.erzeugung (pro qkm) ← Veränd. Einw.dichte



# Vorbereitung: Regressionsanalyse

(= Linearer Fit)

Ich gehe davon aus, dass Sie das schon einmal gesehen haben.

Gegeben: “Zigarre” von Punkten im x/y-Raum. [[draw]]

Gesucht: Eine Gerade $ax+c$, die die Zigarre möglichst gut approximiert.

(Technisch: Minimiere mittleren quadratischen Fehler zwischen Gerade und Punkten.)
[[draw]]

Irgendwie geht das, und heraus kommen:

- Schätzwerte $\hat{a}$ und $\hat{c}$ für die Geradengleichung
- Schätzwerte für die Fehler von $\hat{a}$ und $\hat{c}$ .
- $R^2$ als Messzahl der Qualität (zwischen 0 (worst) und 1 (best)).

Z.B.

$$n_{tripsPerDay} = 1 + \frac{1}{10kEu} annualIncome$$

Hier ist

- $\hat{c} = 1$
- $\hat{a} = \frac{1}{10kEu}$

# # Weitere Vorber.: Multiple Regression

Weitere “erklärende/unabhängige Variablen” (explanatory/independent variables)

Z.B.:

$$y = b + a_1\, x_1 + a_2\, x_2$$

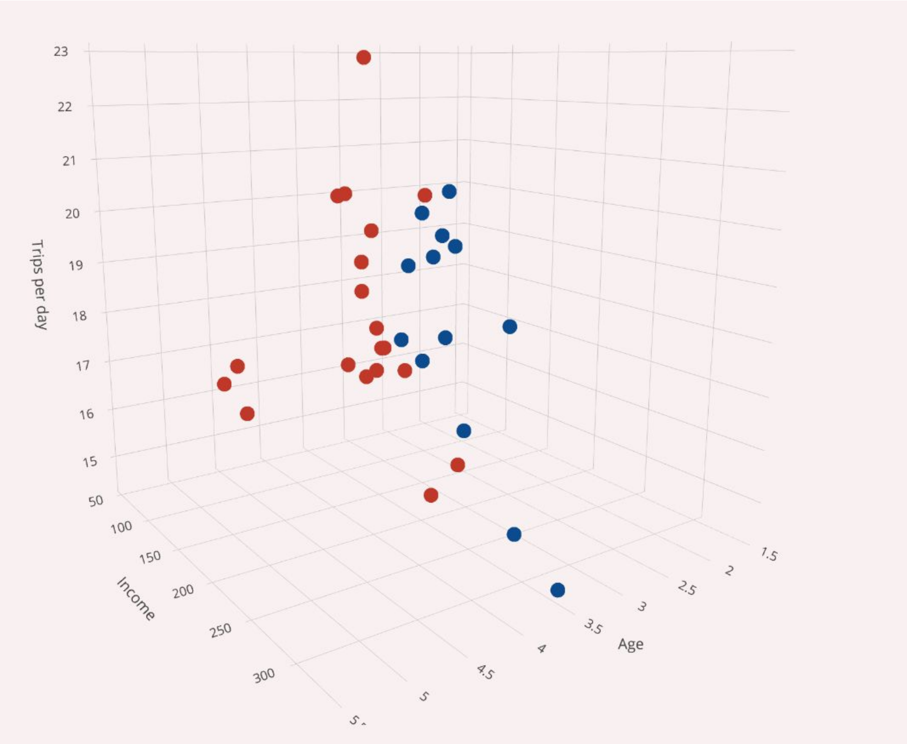
$$n_{trips\ per\ day} = 1 + \frac{1}{10\ kEu}\,\mathrm{annual\_income} + 0.9\, n_{children}\ .$$

[[Siehe maple.]]

Bem.:

- Anschaulich ist das nun das Problem, eine Ebene durch eine 3D-Punktewolke zu fitten.
- Jede zusätzliche erklärende Variable sollte das $R^2$ erhöhen.
- Die Mathematik wird schwieriger, wenn die erklärenden Variablen miteinander korrelieren (z.B. Einkommen und Wohnungsgröße).

Beispiel für solch eine “3D-Punktewolke”



# # Vorschau

Praktisch alles, was nun folgt, sind solche Regressionen. Sie unterschieden sich z.B. in folgenden Punkten:

- Zonen. Macht man eine Regression für die gesamte Region, oder separat für jede Zone?
- Klassifikationen. Aufteilung der Bevölkerung in Klassen (z.B. nach Alter) und separate Regressionen.
- Verwendung von “Kategorien” statt quantifizierbarer Größen (z.B. Altersklassen statt Alter als unabhängige Variable). Vorteil: Braucht nicht linear zu sein (z.B. wenige Trips wenn jung/alt, viele Trips wenn mittelalt).
- Generelles Problem: Je mehr separate Regressionen man macht, desto schlechter ist die Datenlage. Z.B.:
    - 1 Million Einwohner; 1000 Zonen, also im Mittel 1000 Einwohner pro Zone. Regression pro Zone möglich.
    - Annahme 3 Altersklassen, 3 Einkommensklassen, damit im Mittel 1000/9 = 111 Einwohner pro Zone und Klasse. Wird knapp ...

# # Regression basierend auf Haushalten

Z.B.

$$n_{trips\ in\ hh} = b + a_1 \times n_{persons\ in\ hh} + a_2 \times I + a_3 \times n_{cars}\ .$$

$I$: Einkommen

Bem:

- Summiere Wege aller Haushalte in jeder Zone, um zonale Werte zu erhalten.
- Warum Haushalte und nicht Personen ... Haushalte funktioniert in Praxis besser.

# # Nichtlinearitäten; dummy variables

Ein Problem gibt es offensichtlich, wenn die Anzahl der Wege nicht linear in der Anzahl der untersuchten Größe ist.

Ein (hypothetisches) Bsp. wäre:

- Personen mit wenig Einkommen machen wenige Wege, weil sie sich nicht soviel leisten können.
- Personen mit mittlerem Einkommen machen viele Wege.
- Personen mit hohem Einkommen machen wenige Wege, weil sie sich alles bringen lassen.

Mögliche Auswege:

- Nichtlineare Funktion, z.B.

$$n_{trips\ per\ day} = a_0 - a_I \left(I - I_0\right)^2 + \ldots$$

- 0/1-Variablen (dummy variables). *Wichtig; kommen öfter vor!* Z.B.:

$$n_{tripsPerDay} = b + a_{low.inc}\, x_{low.inc} + a_{high.inc}\, x_{high.inc}$$

wobei

$$x_{low.inc} = \begin{cases} 1 & \text{bei Personen mit niedrigem Einkommen} \\ 0 & \text{sonst} \end{cases},$$

$$x_{high.inc} = \begin{cases} 1 & \text{bei Personen mit hohem Einkommen} \\ 0 & \text{sonst} \end{cases},$$

sowie $a_{low.inc}, a_{high.inc}$ negativ für unser hypothetisches Beispiel.

Sieht vielleicht nicht so aus, aber kann via ganz normale Regression geschätzt werden.

Bem.: $(x_{low.inc}, x_{high.inc}) = (0, 0)$ entspricht dann “mittlerem Einkommen”.

$(x_{low.inc}, x_{high.inc}) = (1, 1)$ kann zwar mathematisch-formal vorkommen, macht aber keinen Sinn.

# Kennwertmodell

- Oben hatten wir z.B. so etwas wie $n_{trips\ per\ day} = b + a_A\ A$, mit $A$ Alter.
- Wir hatten dann gesagt, dass dieser lineare Ansatz ggf nicht gut funktioniert. Z.B. würde man erwarten, dass Menschen mittleren Alters die meisten Trips machen, und man sowohl mit niedrigem als auch mit hohem Alter darunter liegt.
- Als Ausweg wurden nichtlineare Funktionen sowie 0/1-Variablen kurz diskutiert.

Wenn man das zu Ende denkt, dann kommt man auf die Idee, auf eine funktionale Form komplett zu verzichten und die Menschheit gleich in Klassen/Kategorien einzuteilen. Z.B.: niedr./mittl./hohes Einkommen; niedr./mittl./hohes Alter. Kennwertmodell basierend auf Einkommen x Alter.

# Ein Beispiel (Quelle: Lohse)

**Tab. 10-15** Personengruppen und spezifisches Verkehrsaufkommen (SrV 1991, Dresden [28])

| **Spezifisches Verkehrsaufkommen** in [Ortsveränderungen pro Person und Tag] | **Personenalter** <18 | 18-R | | >R | | **Summe** |
|---|---|---|---|---|---|---|
| | NB | NB | B | NB | B | |
| Ohne Pkw-Verfügbarkeit | 2,89 | 2,90 | 3,33 | 2,15 | 3,00 | 2,89 |
| Fußgängerverkehr | 1,74 | 1,21 | 1,20 | 1,16 | 0,75 | 1,35 |
| Radverkehr | 0,25 | 0,19 | 0,23 | 0,09 | 0,25 | 0,20 |
| Öffentl. Personenverkehr | 0,43 | 0,90 | 1,17 | 0,67 | 1,25 | 0,80 |
| Motor. Individualverkehr | 0,48 | 0,60 | 0,73 | 0,23 | 0,75 | 0,53 |
| Mit Pkw-Verfügbarkeit | - | 3,65 | 3,80 | 4,28 | 4,96 | 3,82 |
| Fußgängerverkehr | - | 0,91 | 0,42 | 2,04 | 4,96 | 0,60 |
| Radverkehr | - | 0,12 | 0,19 | 0,08 | - | 0,17 |
| Öffentl. Personenverkehr | - | 0,72 | 0,27 | 0,93 | - | 0,37 |
| Motor. Individualverkehr | - | 1,90 | 2,92 | 1,22 | - | 2,67 |
| Gesamt | 2,89 | 3,08 | 3,56 | 2,42 | 3,10 | 3,15 |
| Fußgängerverkehr | 1,74 | 1,13 | 0,82 | 1,27 | 0,96 | 1,14 |
| Radverkehr | 0,25 | 0,17 | 0,21 | 0,09 | 0,24 | 0,19 |
| Öffentl. Personenverkehr | 0,43 | 0,86 | 0,73 | 0,70 | 1,19 | 0,68 |
| Motor. Individualverkehr | 0,48 | 0,92 | 1,80 | 0,36 | 0,71 | 1,14 |
| **Personenanteile** [%] in den Gruppen | **Personenalter** <18 | 18-R | | >R | | Summe |
| | NB | NB | B | NB | B | |
| Ohne Pkw-Verfügbarkeit | 21,96 | 10,40 | 23,76 | 14,72 | 0,83 | 71,66 |
| Mit Pkw-Verfügbarkeit | - | 3,35 | 22,84 | 2,11 | 0,04 | 28,34 |
| Summe | 21,96 | 13,74 | 46,60 | 16,83 | 0,87 | 100,00 |

NB = nicht berufstätig  B = berufstätig  R = Rentenalter

# # Person-category approach

Da man bei Kennwert-Modellen meistens zu wenig Daten hat, macht es Sinn, statt der Haushalte doch die Personen als Basis zu nehmen, denn davon hat man mehr.

# # Gegenüberstellung 0/1-Variablen vs Kennwertmodell

Das Kennwertmodell ist nicht identisch mit 0/1-Variablen. 0/1-Variablen nehmen immer noch an, dass der verkehrliche Unterschied zw. den Einkommensklassen für alle Altersklassen gleich ist:

$$n_{tripsPerDay} = b + a_{low.age}\, x_{low.age} + a_{hi.age}\, x_{hi.age} + a_{low.inc}\, x_{low.inc} + a_{hi.inc}\, x_{hi.inc}$$

hat immer den gleichen Unterschied z.B. zwischen $(x_{low.inc}, x_{hi.inc}) = (0,1)$ und $(0,0)$, egal, was das Alter sagt.

Folgendes *ist* darstellbar durch dummy variables:

- Differenzen zwischen low und med sowie med und high age überall gleich (4 bzw. –2).
- Differenzen zwischen low und med sowie med und high income überall gleich (9 bzw. –1)

| | low | med | high age |
|---|---|---|---|
| low | 3 | 7 | 5 |
| med | 12 | 16 | 14 |
| high income | 11 | 15 | 13 |

Dies ist der wesentliche Vorteil des cross-classification-Modelles:

- keinerlei Annahme einer funktionalen Form an irgendeiner Stelle (wie gerade gesehen, funktionieren dummy variables nur bei einer bestimmten Struktur der Daten).

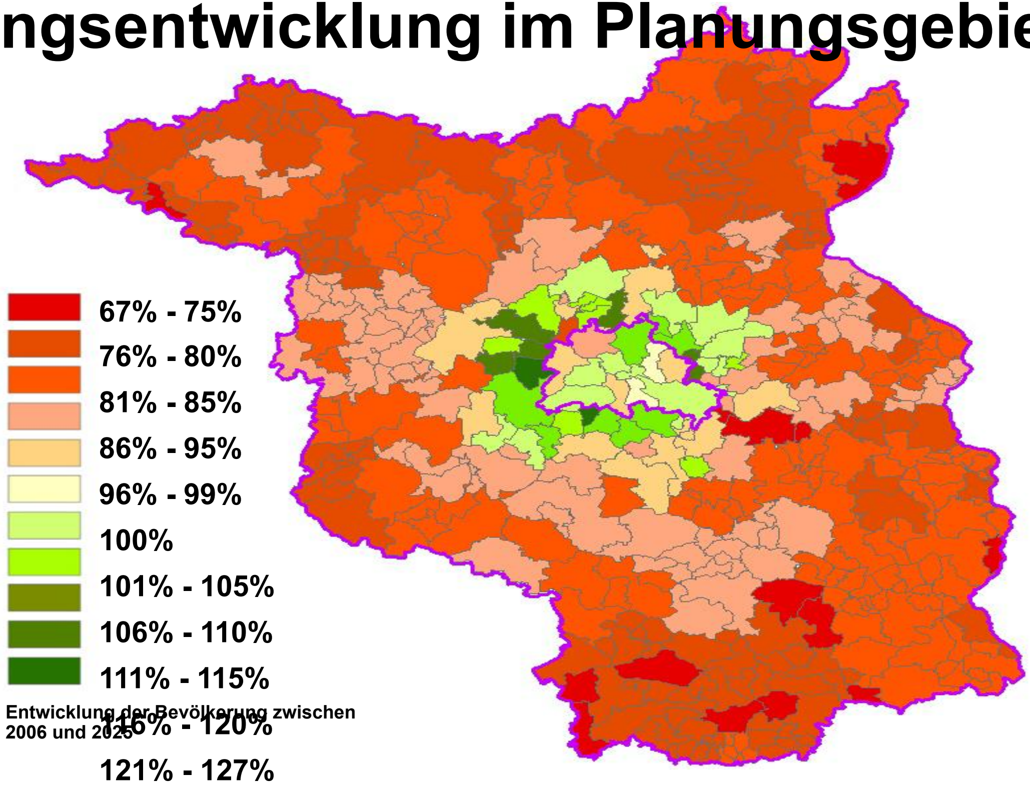
Nachteile:

- Wahl der Klassen benötigt Intuition.
- Keine Extrapolation möglich (da keine unterliegende mathematische Funktion).
- Datenhungrig.
  Bei Regression entspricht die Anzahl der Freiheitsgrade der Anzahl der Tabellen-“Richtungen”; beim Kennwertmodell hat man einen Freiheitsgrad pro Zelle.
  Ausserdem sind meistens noch die Haushalte sehr ungleichmäßig auf die Zellen verteilt.

# Verkehrserzeugung im Verkehrsmodell für Berlin-Brandenburg

# # Ausgewählte Einflussfaktoren des Verkehrs

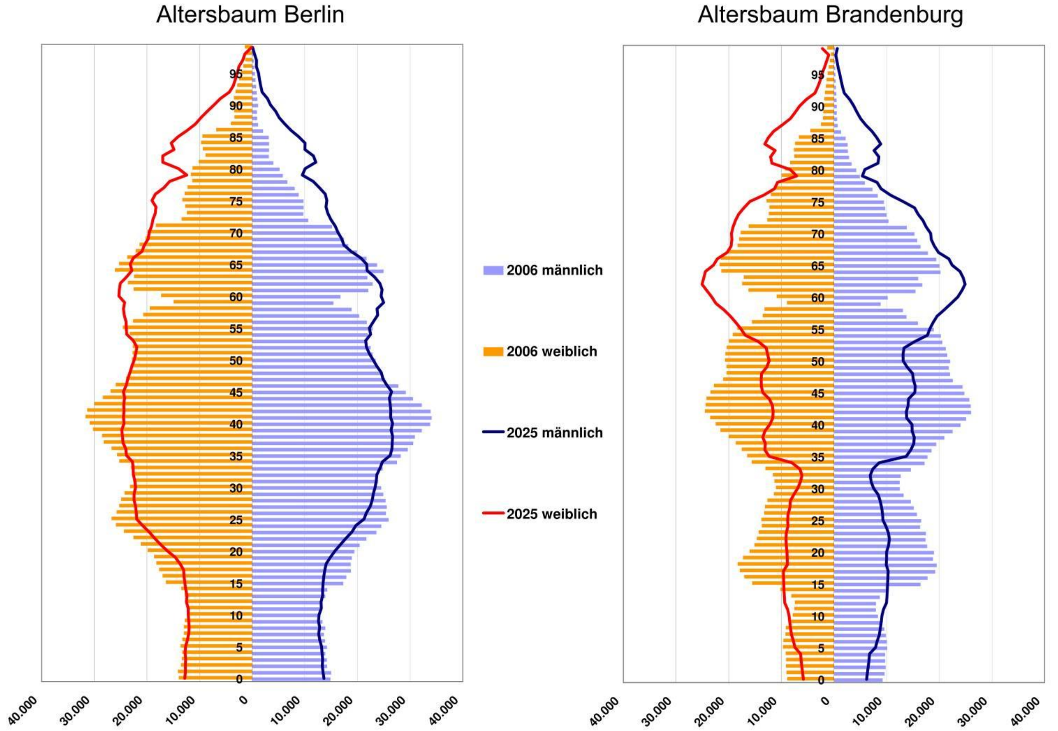
## ## Bevölkerungsentwicklung im Planungsgebiet




Entwicklung der Bevölkerung zwischen 2006 und 2025

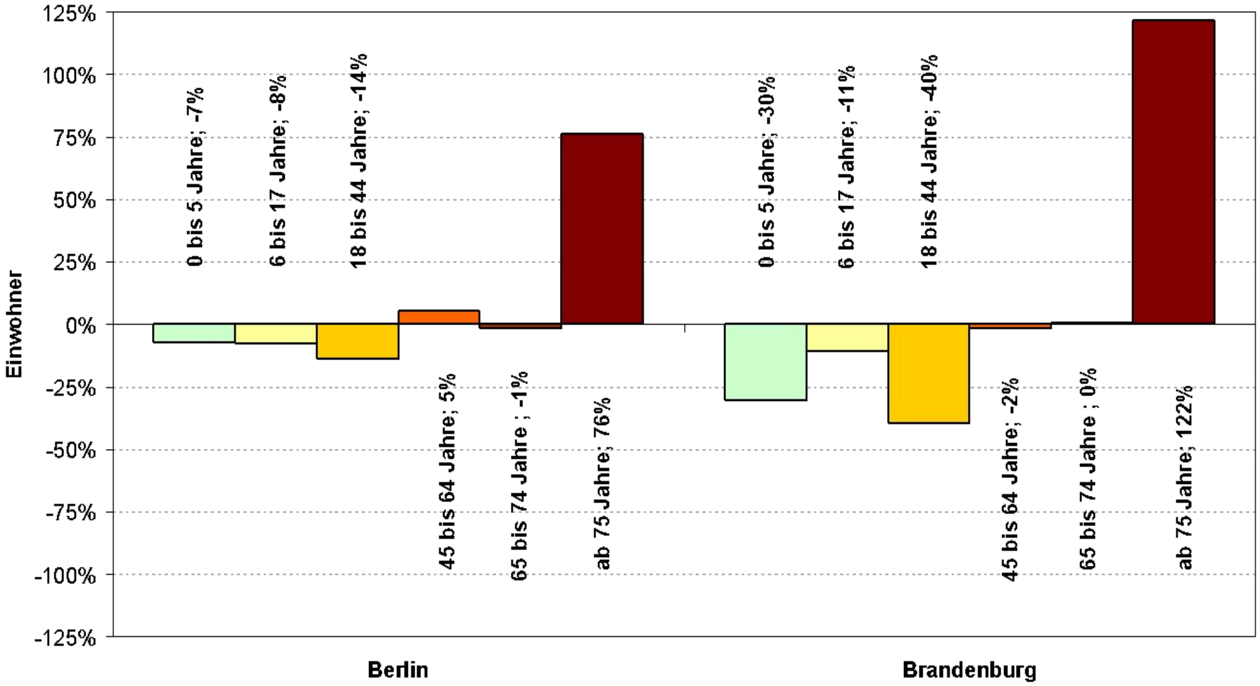
Quelle u.a.:
Landesbetrieb für Datenverarbeitung und Statistik, Land Brandenburg, Senatsverwaltung für Stadtentwicklung Berlin Abt. I

## Altersbäume für Berlin und Brandenburg




Quelle u.a.:
Landesbetrieb für Daten-verarbeitung und Statistik,
Land Brandenburg,
Senatsverwaltung für Stadt-entwicklung Berlin Abt. I

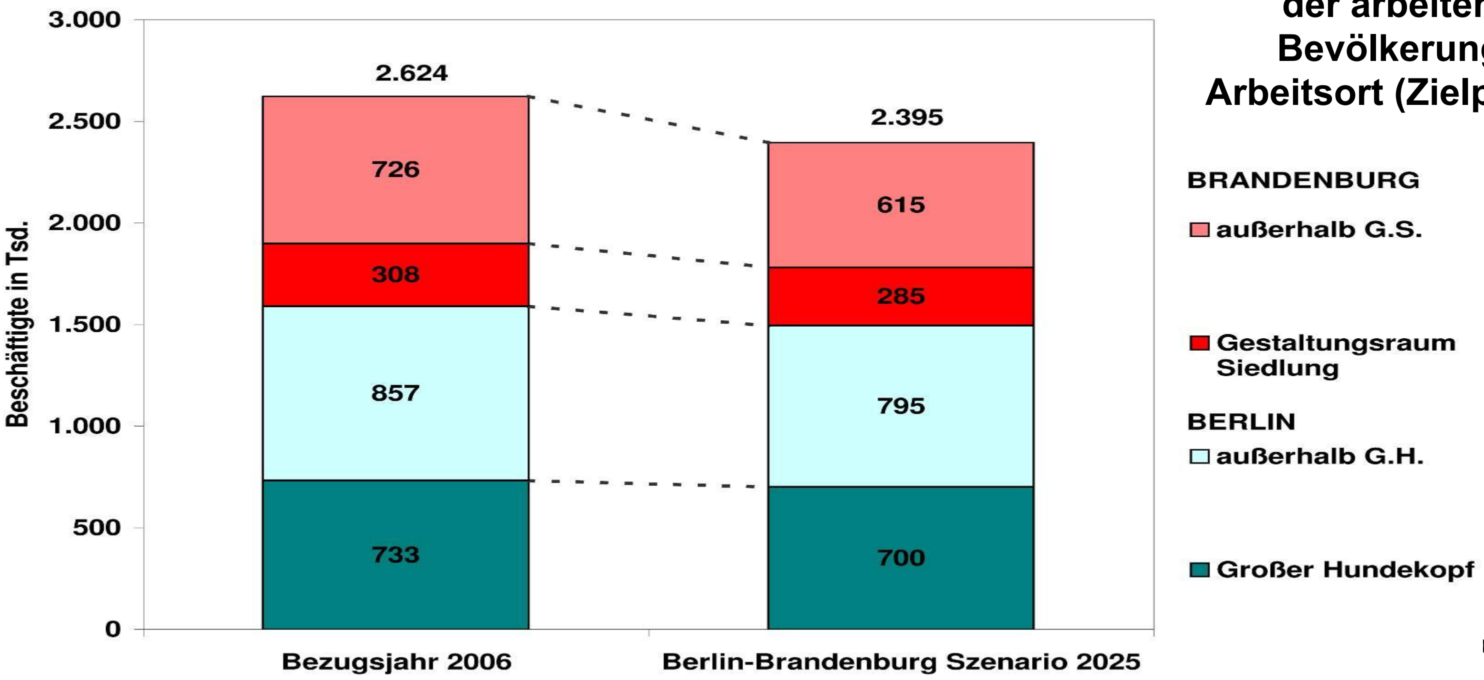
# ## Veränderungsraten der Einwohner nach Altersgruppen




**in Berlin und Brandenburg zwischen 2006 und 2025**

## ## Entwicklung der Beschäftigten im Planungsgebiet

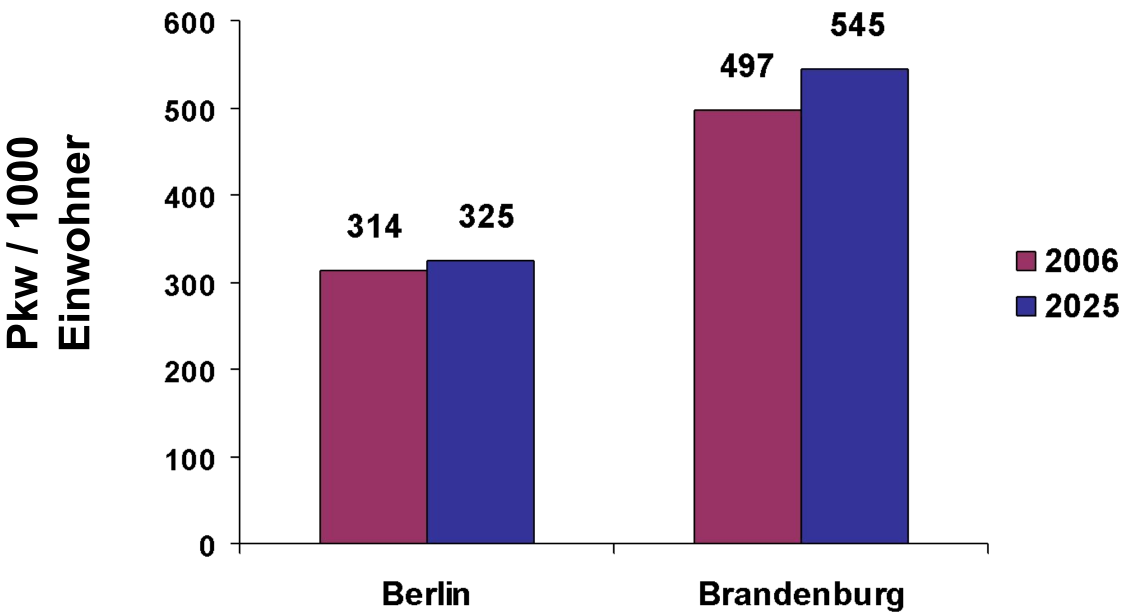
**Beschäftigte = Menge der arbeitenden Bevölkerung am Arbeitsort (Zielpotential)**




uelle u.a.:
rognos AG
Baasner, Möller, Langwald, 2007

## Entwicklung der Motorisierung

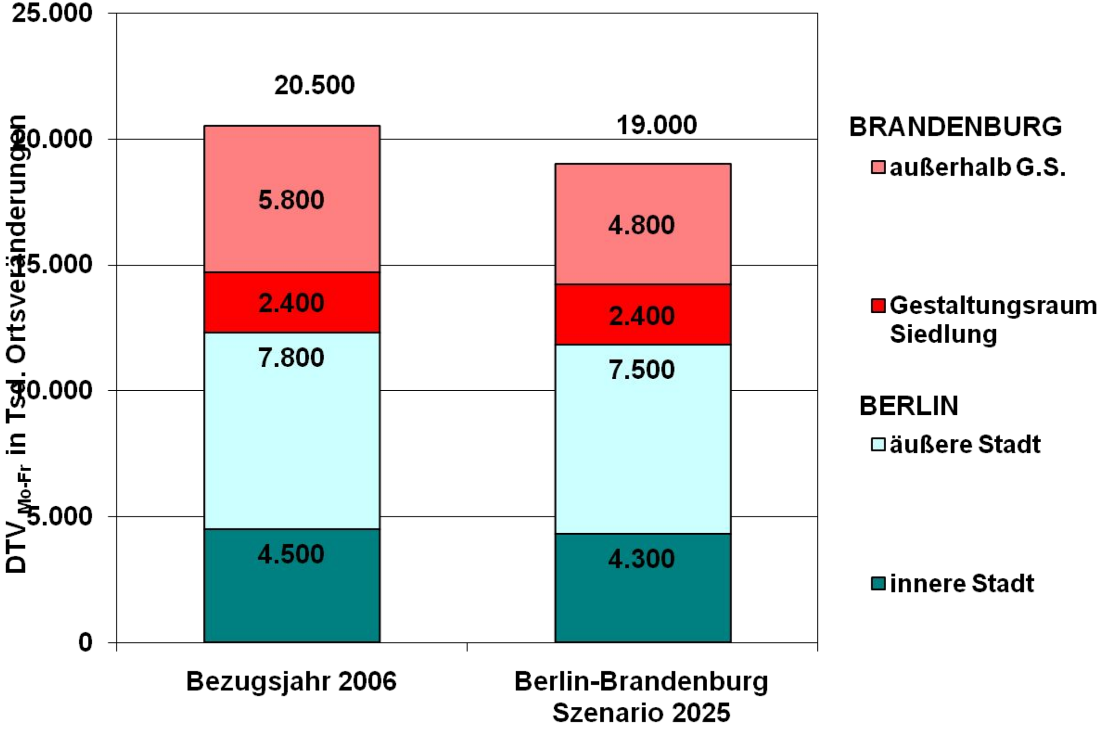
**Analyse – Prognose**




Quelle u.a.:
BVPI, Shell-Prognosen, KBA, Zulassungsstellen der Länder

# # Ergebnisse wichtiger Eckwerte der Verkehrsprognose

# ## Verkehrsaufkommen Personenverkehr




**im sogenannten Berlin-Brandenburg-Szenario**

Werte gerundet

# Dies und das

# # $\sum$ Quellterme = $\sum$ Senkenterme

Die meisten Modelle (Regression oder anders) wendet man separat für zonale Quellterme, $O_i$, und zonale Senkenterme, $D_i$ , an.

Es ist halbwegs klar, dass nicht automatisch

$$\sum_i O_i = \sum_j D_j$$

Das braucht man aber.

Lösung: Da normalerweise die Quellterme die besseren Schätzungen sind, werden die Senkenterme reskaliert:

$$D'_j = D_j \frac{\sum_i O_i}{\sum_j D_j} \; .$$

# # Zeitliche Stabilität

Bei Infrastruktur-Planung wünscht man Vorhersagen z.B. 20 Jahre in die Zukunft.

Dafür braucht man, bzgl. unserer Variablen $n_{tripsPerDay}$:

- Die Personen-/Haushaltsstruktur in 20 Jahren für jede Zone.
- Stabilität der Modelle über 20 Jahre.

Beides ist offensichtlich nicht so einfach zu erreichen:

- Personen-/Haushaltsstruktur innerhalb von Zonen ändert sich, Viertel entwickeln sich, etc. (Analog für Trip-Ziele.)
- Modelle/Koeffizienten können sich ändern, z.B. weil: die Wertesysteme sich ändern; die Kraftstoff -Preise sich ändern; etc.

Alternative Ansätze: später im Semester (activity-based demand generation). Aus theoretischer Sicht sind diese besser; der praktische Nutzen ist noch unklar.

# # Geographische Stabilität

Kann ein trip generation Modell von einer geographischen Region auf eine andere übertragen werden?

OW sagt, und ich höre dies auch immer wieder, dass das i.a. eher nicht geht.

Meine persönliche Meinung ist, dass wir dann mit unseren Modellen etwas falsch machen. Insbesondere: Wenn die für eine Region kalibrierten Modelle nicht auf einer andere Region übertragbar sind, wie sehr glauben wir dann den Vorhersagen?

Daraus folgt wohl, dass man “state-of-practice” trip generation Modellen in Bezug auf zeitliche Vorhersagen eher skeptisch gegenüberstehen sollte, wenn sie mehr als Trends vorhersagen.

Wissenschaftlich auswertbare Vergleiche eher selten.

Unsere eigenen Erfahrungen/Intuitionen:

- Die (NKV[1] -)*Reihenfolge* von Projekten hängt nicht sehr stark vom *allgemeinen* Verkehrswachstum ab.

  [1]Nutzen-Kosten-Verhältnis

- Eine Rolle spielen hingegen regional differenzierte Vorhersagen (z.B. starkes Wachstum in Bayern/BW; schwaches Wachstum in Mecklenburg-Vorpommern).

  U.E. stark getrieben von Bevölkerungsentwicklung.

# Die wichtigsten Punkte

- Es geht darum, Quellen und Senken für Fahrten festzustellen. Wir haben uns aber fast nur mit Modellen für die Quellen beschäftigt.
- Wir haben drei generelle Ansätze kennengelernt:
  - Steigerungsfaktor. $n_{tripsPerDay}$ wächst jedes Jahr um einen konstanten Faktor. Kann von Zone zu Zone unterschiedlich sein, aber dann ist endgültig unklar, wo diese Information herkommen soll.
  - Regressionsanalyse. $n_{tripsPerDay}$ ist linear in den Variablen. Ggf werden 0/1-Variablen verwendet.
  - Cross-classification/category analysis/Kennwertmodell. Z.B. $m$ Klassen für's Alter, $n$ für's Einkommen, $p$ für Autobesitz, etc. Resultiert in $m \times n \times p$ Zellen, und für jede Zelle wird $n_{tripsPerDay}$ separat geschätzt.
- Die geographische Übertragbarkeit der Modelle ist nicht sehr gut, was auch Zweifel an der zeitlichen Vorhersagestabilität weckt.