



# Leistungsdaten-Analysen in der PKV

## Grundlagen, Möglichkeiten und Grenzen

Nürnberg, 31.01.2017

# Inhaltsverzeichnis

1. Das Wesen der Leistungsdaten
2. Datennutzung - rechtlicher Rahmen
3. Projekt 1: Schadenprognosen
4. Projekt 2: Betrugserkennung
5. Projekt 3: Krankheitsprognosen
6. (Datennutzung - Kundensicht)

A wide-angle photograph of a modern university building complex. The architecture features large glass windows, metal facades, and balconies. In the foreground, there is a paved courtyard with a few people walking. To the right, a large, shallow pool of water reflects the sky and the building. Several white banners with the 'HDK-Coblenz' logo are visible in the middle ground. The sky is blue with a few clouds.

# Das Wesen der Leistungsdaten

# Unterschiede der KV zu anderen Versicherungsarten

- ▶ kontinuierlicher Informationsfluss (11 Belege / Versicherten und Jahr)
- ▶ komplexer Datensatz
- ▶ langfristige Abhängigkeit der Daten
- ▶ indirekte Datenerfassung

# Die ambulante Rechnung

<p>Herrn Patient Nebenstrasse 21 54321 Tral</p> <p>Sehr geehrter Herr Patient, für meine ärztliche Bemühungen erlaube ich mir zu berechnen: € 565,25</p>	<p>Dr. WHO Facharzt für Allgemeine Medizin Hauptstrasse 42 12345 Gallifrey</p> <p>Rechnungsdatum: 31.01.2017 Rechnungs-Nr.:1704 Patientennummer: 001 Arztnummer: 11235812</p>																												
<div style="border: 1px solid black; padding: 2px;"><b>Diagnose(n):</b> J10.00, Grippe durch saisonale nachgewiesene Influenzaviren</div>																													
<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th style="width: 10%;">Datum</th> <th style="width: 10%;">GOÄ-Ziffer</th> <th style="width: 30%;">Leistung</th> <th style="width: 10%;">Anzahl</th> <th style="width: 10%;">Faktor</th> <th style="width: 10%;">Einzel</th> <th style="width: 10%;">Gesamt</th> </tr> </thead> <tbody> <tr> <td>31.12.16</td> <td>1</td> <td>Beratung – auch mittels Fernsprecher</td> <td>1</td> <td>2,3</td> <td>10,72</td> <td>10,72</td> </tr> <tr> <td>31.12.16</td> <td>5</td> <td>Symptombezogene Untersuchung</td> <td>1</td> <td>2,3</td> <td>10,72</td> <td>10,72</td> </tr> <tr> <td colspan="6" style="text-align: right;"><b>Summe</b></td> <td>565,25</td> </tr> </tbody> </table>		Datum	GOÄ-Ziffer	Leistung	Anzahl	Faktor	Einzel	Gesamt	31.12.16	1	Beratung – auch mittels Fernsprecher	1	2,3	10,72	10,72	31.12.16	5	Symptombezogene Untersuchung	1	2,3	10,72	10,72	<b>Summe</b>						565,25
Datum	GOÄ-Ziffer	Leistung	Anzahl	Faktor	Einzel	Gesamt																							
31.12.16	1	Beratung – auch mittels Fernsprecher	1	2,3	10,72	10,72																							
31.12.16	5	Symptombezogene Untersuchung	1	2,3	10,72	10,72																							
<b>Summe</b>						565,25																							
<p>Bitte überweisen Sie den Rechnungsbetrag auf das Konto Meins bei Meiner-Bank (BLZ XXX).</p>																													

21,44

# ICD-Diagnosen

- ▶ ICD = Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme
- ▶ Form z.B. J10.00
- ▶ J10.00 -> Krankheitskapitel (Krankheiten des Atmungssystems) [22]
- ▶ J10.00 -> Krankheitsgruppe (Grippe und Pneumonie) [261]
- ▶ J10.00 -> Klasse (Grippe durch saisonale nachgewiesene Influenzaviren) [12.161]

# ICD-Diagnosen

## Probleme:

- ▶ Krankheitskapitel medizinisch heterogen
  - ▶ K = Krankheiten des Verdauungssystems
  - ▶ K02= Karies aber K72= Leberversagen
- ▶ Krankheitsbilder passen nicht immer 100% auf ICD
- ▶ Codierung der Ärzte:

J20.9 Akute Bronchitis, nicht näher bezeichnet

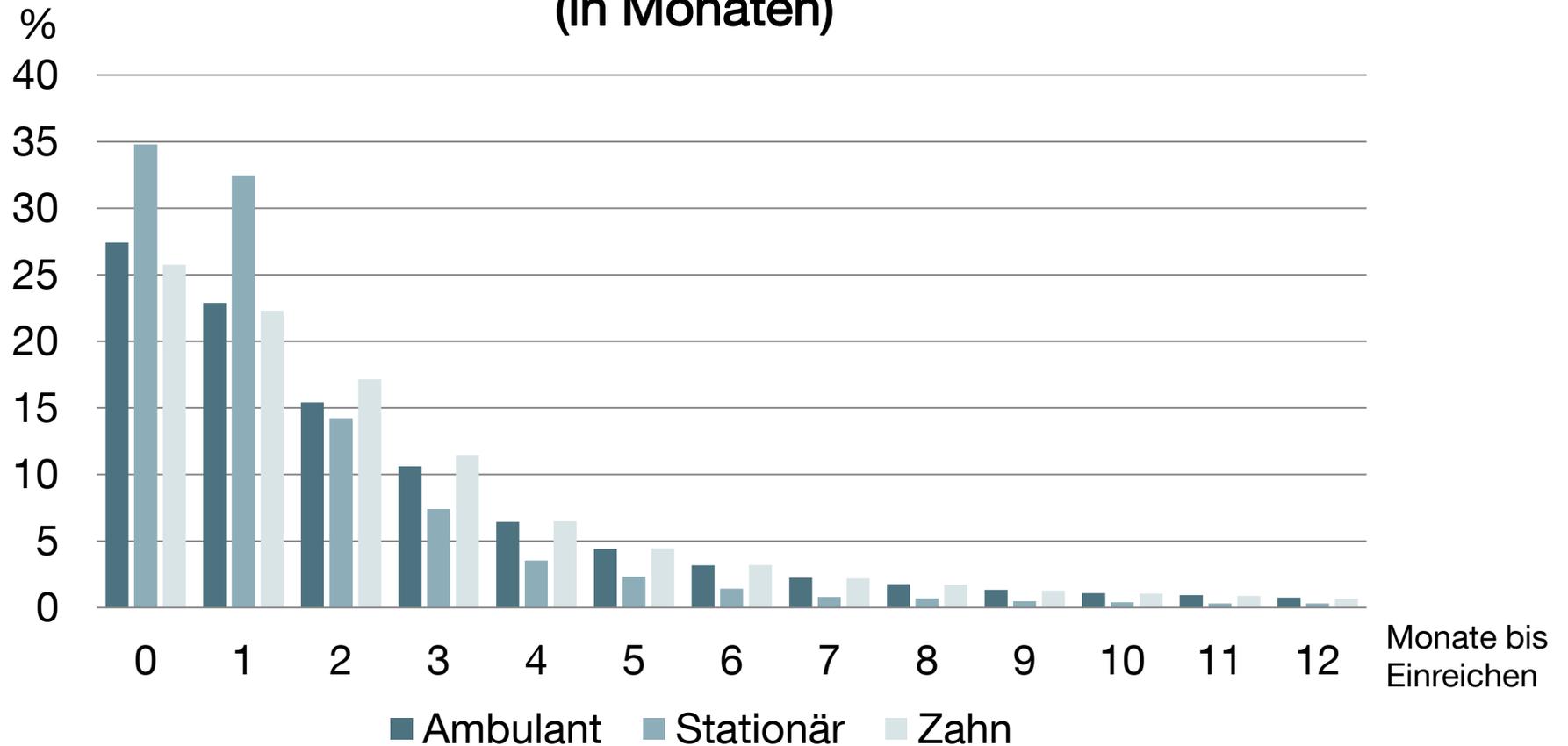
Diagnosen:  
Zustand nach Cholecystektomie; Sigmadivertikulose;  
Prostatahyperplasie ( BPE ) Stadium 1 nach Vahlensieck;  
Radikulopathie: Sakral- und Sakrokokzygealbereich; Segmentale und somatische Funktionsstörungen: Lumbalbereich [lumbosakral]; Segmentale und somatische Funktionsstörungen: Sakralbereich [sakrokokzygeal, sakroiliakal]; Skoliose, nicht näher bezeichnet: Lumbalbereich; Osteochondrose der Wirbelsäule, nicht näher bezeichnet: Lumbalbereich; Sonstige Spondylose mit Radikulopathie: Lumbalbereich; Bandscheibenschaden, nicht näher bezeichnet; Radikulopathie: Lumbosakralbereich; Radikulopathie: Lumbalbereich; Sonstige Spondylose mit Radikulopathie: Zervikalbereich; Sonstige und nicht näher bezeichnete Kyphose: Zervikalbereich; Skoliose, nicht näher bezeichnet: Zervikalbereich; Osteochondrose der Wirbelsäule, nicht näher bezeichnet: Zervikalbereich; Zervikaler Bandscheibenschaden, nicht näher bezeichnet; Radikulopathie: nicht näher bezeichnete Lokalisation; Radikulopathie; Fettleber 2; Coronarinsuffizienz; coronare 2 Gefäßerkrankung bei Zustand 2 fach Bypass Op - septale Hypokinesie - diastolische Relaxationsstörung; KHK; Hypertonus Stad. II WHO; Polyneuropathie; restless legs Syndrom; Burn out Syndrom; Diabetes mellitus Typ 2 B; Otalgie; Hyperbilirubinämie; Mikrohämaturie; Magenschmerzen und Übelkeit; GPT Erhöhung; Bandscheibenprotrusio C4/C5/C6; Bandscheibenprotrusio L4/L5/S1

## Weitere (Problem-) Faktoren

- ▶ Bestandsalterung
- ▶ medizinische Inflation
- ▶ Selektion
- ▶ Selbstbehalte und BRE
- ▶ „geringes“ Bestandsvolumen
- ▶ Daten-LAG

# Datensatz - Besonderheiten

## Dauer Behandlung bis Rechnungseingang (in Monaten)



A wide-angle photograph of a modern university building complex. The architecture features large glass windows and balconies. In the foreground, there is a paved courtyard with a swimming pool on the right side. Several people are walking in the courtyard. In the background, there are several tall poles with banners that read 'HDK-Coblenz'. The sky is blue with a few clouds.

# Datennutzung - rechtlicher Rahmen

# Datensatz - Datenschutz

## Grundsätzliches

a. § 4 Abs. 1 BDSG: „Verbot mit Erlaubnisvorbehalt“

Personenbezogene Daten dürfen nur erhoben, verarbeitet oder genutzt werden, wenn eine Rechtsvorschrift oder die Einwilligung des Betroffenen dies gestattet.

b. „Grundsatz der Zweckbindung“ (implizit im BDSG)

Personenbezogene Daten dürfen nur zu dem Zweck genutzt/verarbeitet werden, zu dem sie erhoben wurden.

c. § 3 Abs. 9 BDSG: „Gesundheitsdaten sind besonders schützenswert“

Besondere Arten personenbezogener Daten sind Angaben über die rassische und ethnische Herkunft, politische Meinungen, [...] Gesundheit [...].

# Datensatz - Datenschutz

## Grundlage für Datenauswertung

### §28 (1) BDSG

Das Erheben, Speichern, Verändern oder Übermitteln personenbezogener Daten oder ihre Nutzung als Mittel für die Erfüllung eigener Geschäftszwecke ist zulässig...

2. soweit es zur Wahrung berechtigter Interessen der verantwortlichen Stelle erforderlich ist und kein Grund zu der Annahme besteht, dass das schutzwürdige Interesse des Betroffenen an dem Ausschluss der Verarbeitung oder Nutzung überwiegt...

**Immer Interessenabwägung zwischen Kunde und VU**



**Datenschützer im Unternehmen einbeziehen**



# Projekt 1: Schadenprognosen

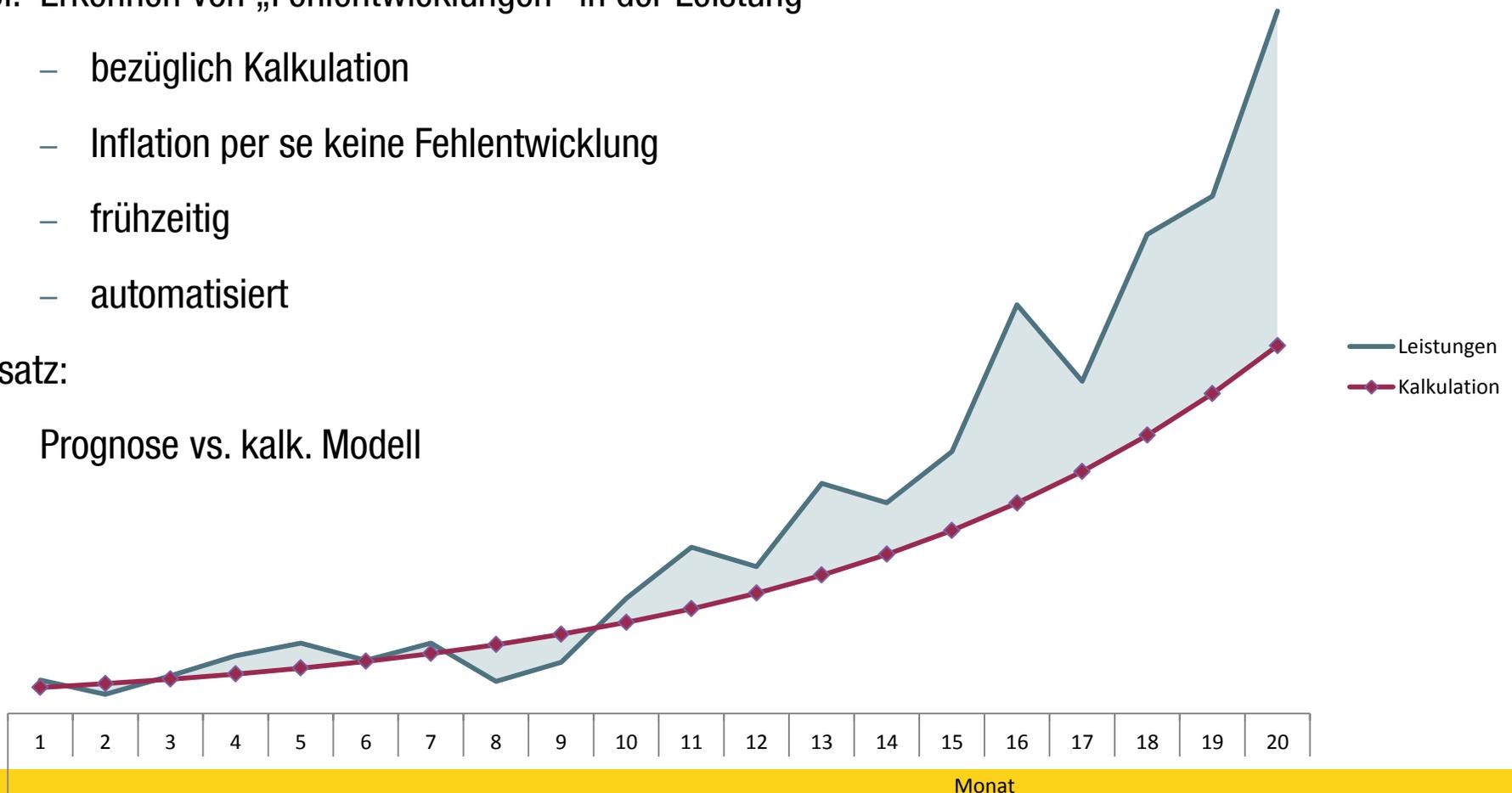
# Projekt 1: Risikomonitor

Ziel: Erkennen von „Fehlentwicklungen“ in der Leistung

- bezüglich Kalkulation
- Inflation per se keine Fehlentwicklung
- frühzeitig
- automatisiert

Ansatz:

Prognose vs. kalk. Modell



# Projekt 1: Risikomonitor

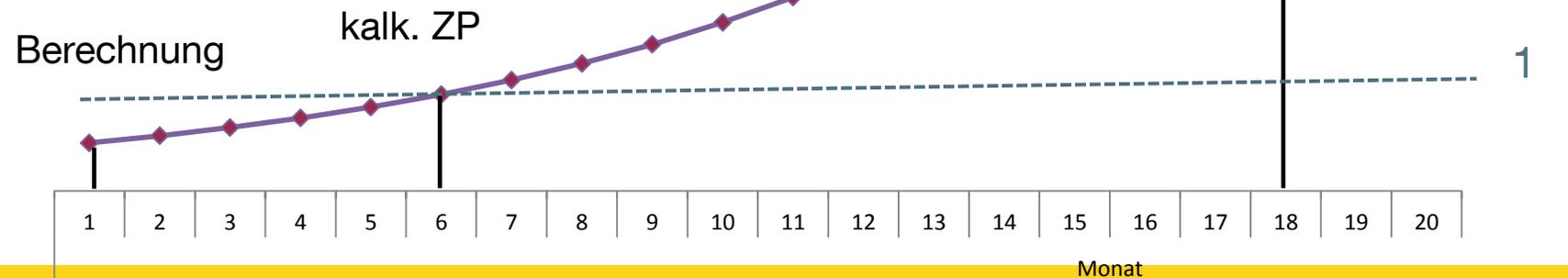
## Kalkulatorisches Modell

$$\text{Leistungsquote} = e^{a \cdot t/12}$$

$a = a(\text{Inflation, Selbstbehalt, RMS, Selektion})$

$t = \text{Monate seit kalk. ZP}$

## Einzelvertragliche Berechnung

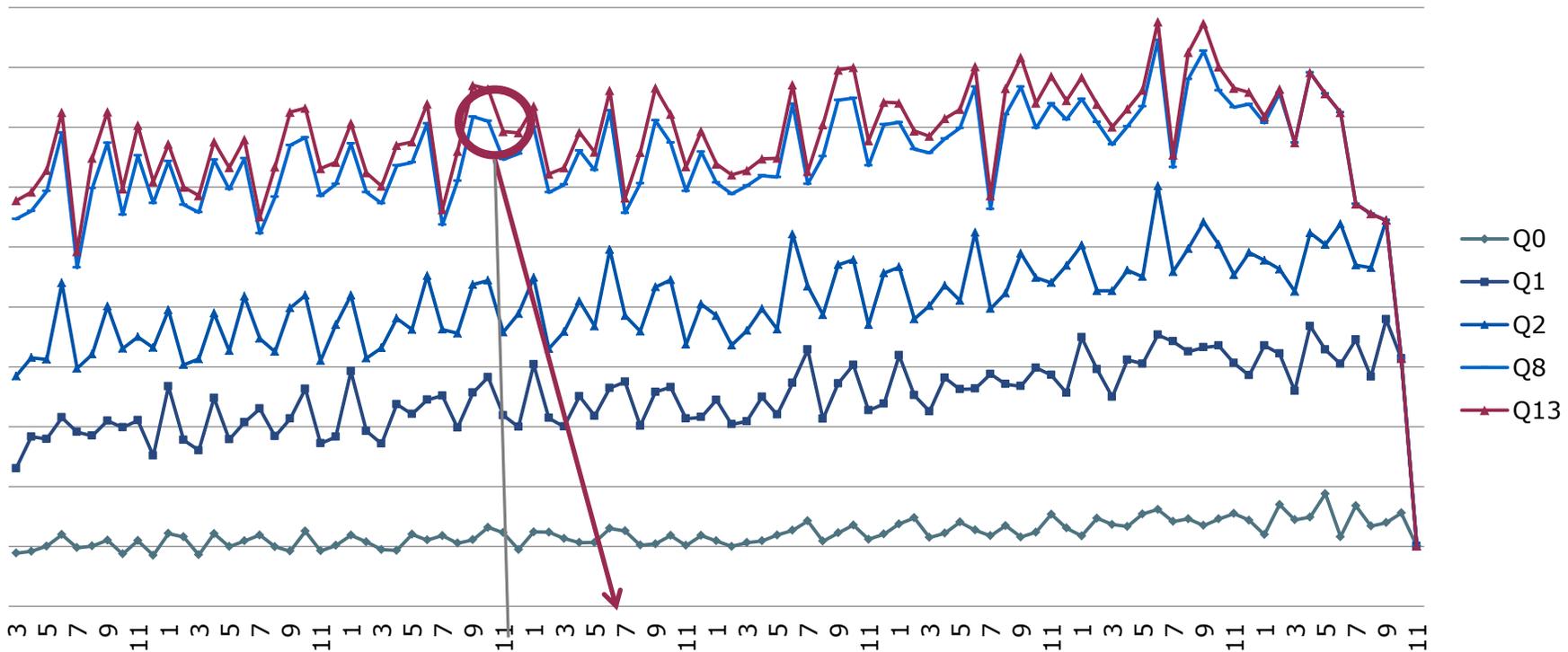


# Projekt 1: Risikomonitor

## Leistungsquote

- ▶ Aufteilen nach Entstehen und Einreichen
- ▶ Großteil noch unbekannt (IBNR)

$$Q = \frac{\text{Leistungen}}{\text{rechnungsmäßige Leistungen}}$$



# Projekt 1: Risikomonitor

ARIMA(X)- Ansatz:

▶ Lineares Modell:

▶ 
$$Y_q(t) = \alpha \cdot X_q(t - 12) + \beta \cdot X_q(t - 24) + \dots + \varepsilon(t) + \sum_{i=1}^{q-1} (\theta_i Y_i(t) + \varepsilon_i(t))$$

▶  $Y_q(t)$  prognosewerte zum ZP  $t$

▶  $X(t - 12)$  Istwert zum ZP  $t - 12$

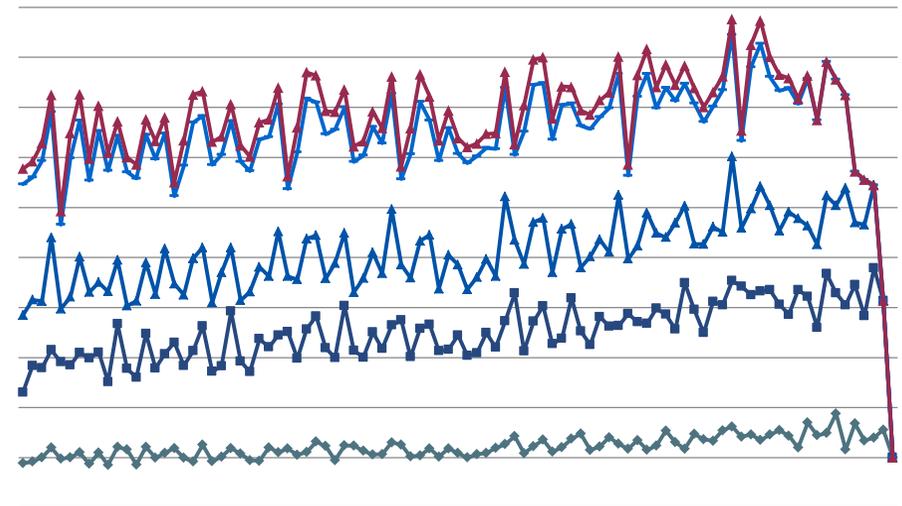
▶  $\varepsilon(t), \varepsilon_i(t)$  Fehlerterme

▶ Saisonalität

▶ „Mehrschritt-Prognose“-Verfahren

▶ Nutzen aller bis dato vorhandenen Informationen

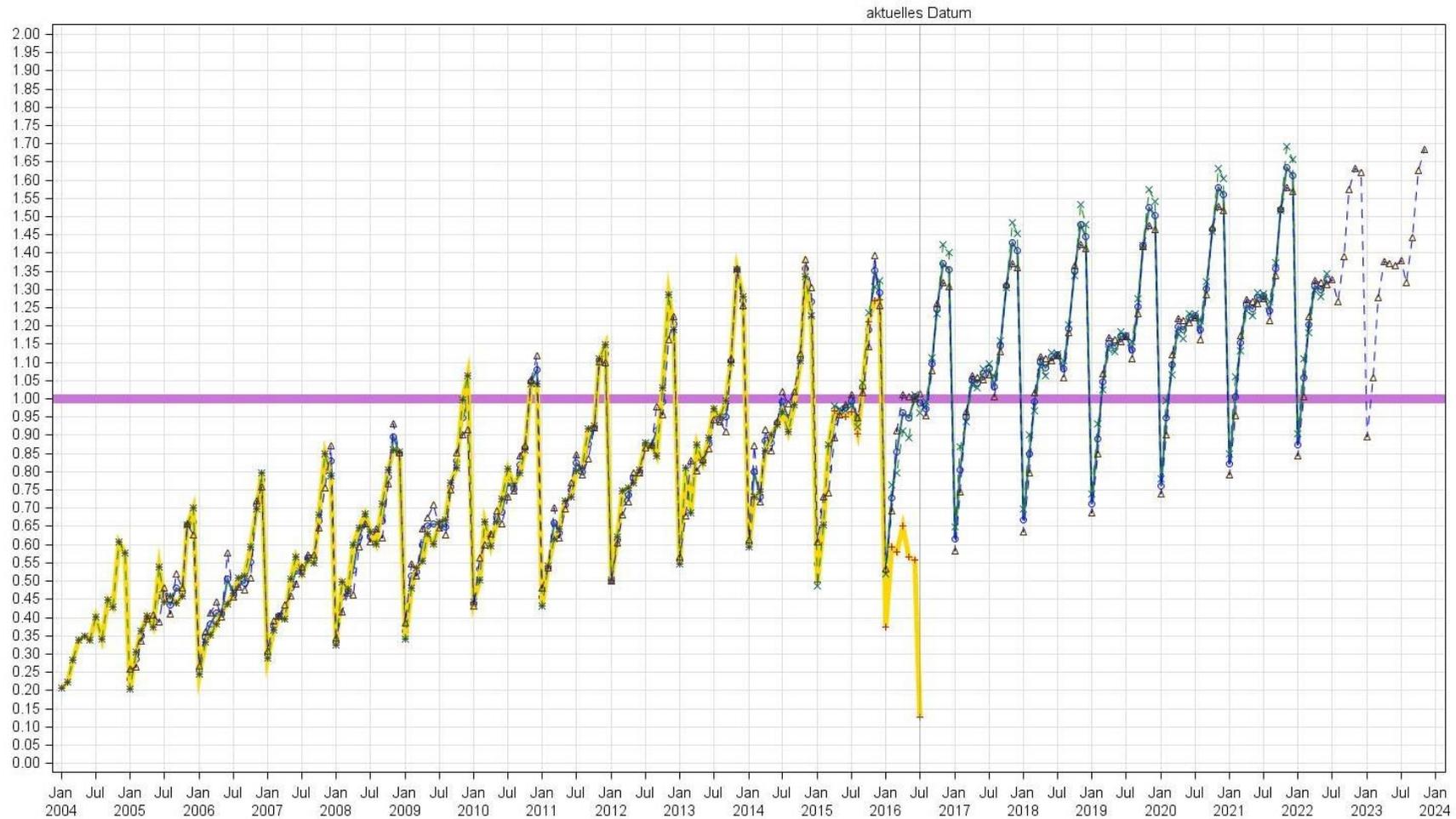
▶ vergleichsweise einfacher Ansatz



# Projekt 1: Risikomonitor

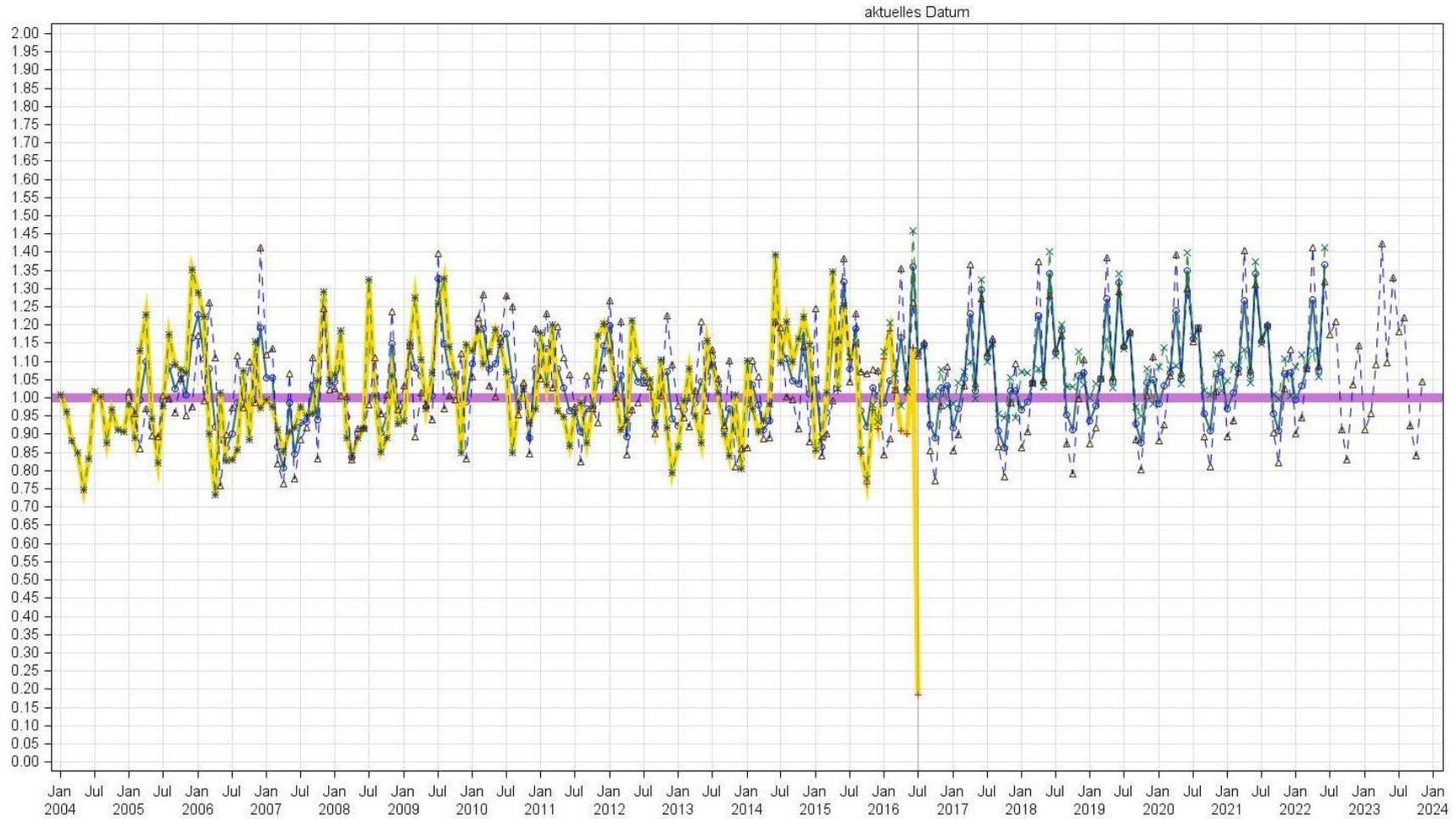
ARIMA(X)- Ansatz:

Bedeutung von Selbstbehalten



# Projekt 1: Risikomonitor

## ARIMA(X)- Ansatz:



# Projekt 1: Risikomonitor

Prognostiziertes Leistungs-Delta nach Inflationsbereinigung, rollierende Werte mit Hochkosten;  
positive Werte sind Fehlbeträge, negative Überschüsse





## Projekt 2: Betrugserkennung

## Projekt 2: Betrugserkennung

### Ziel:

- ▶ Erkennen von Betrugsfällen und Übermaßbehandlungen bei Arzneimitteln

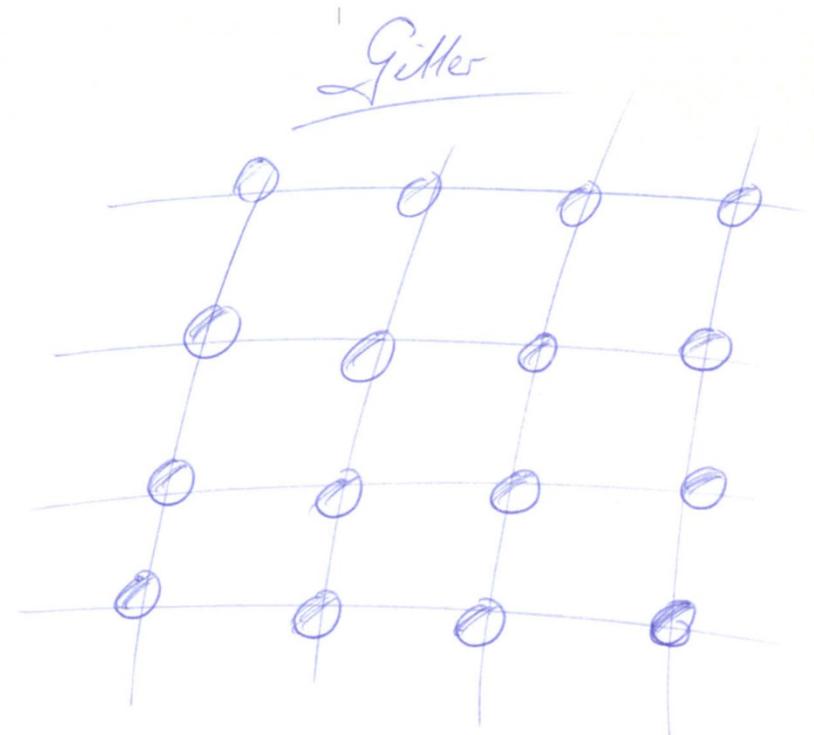
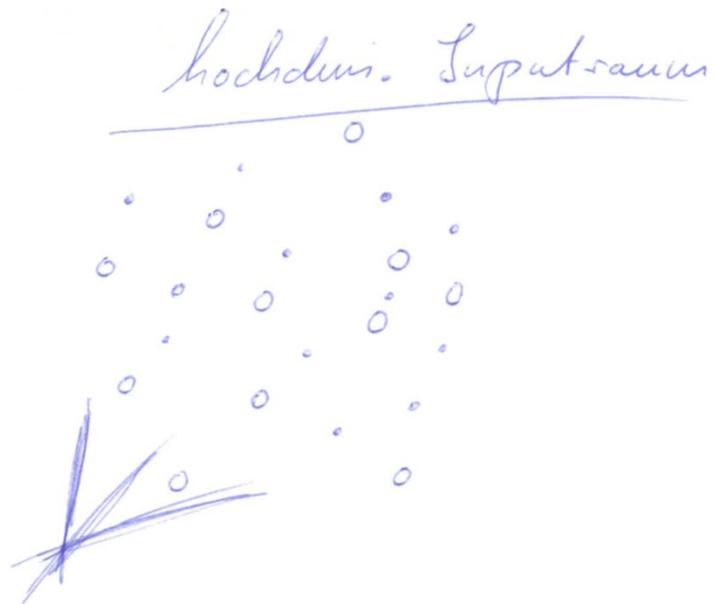
### Probleme:

- ▶ nur wenige, bekannte Betrugsfälle
- ▶ Betrugsmuster i.A. sehr unterschiedlich

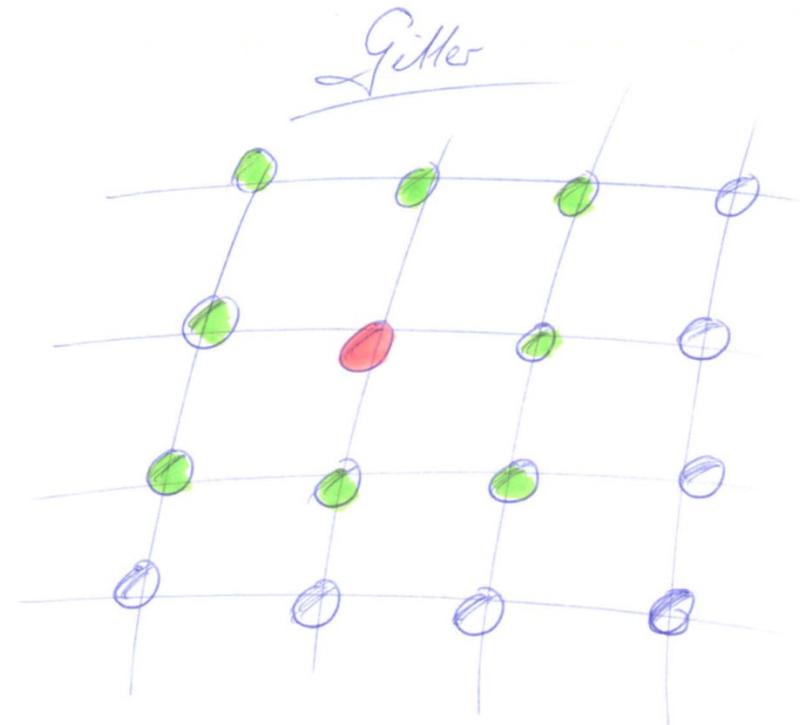
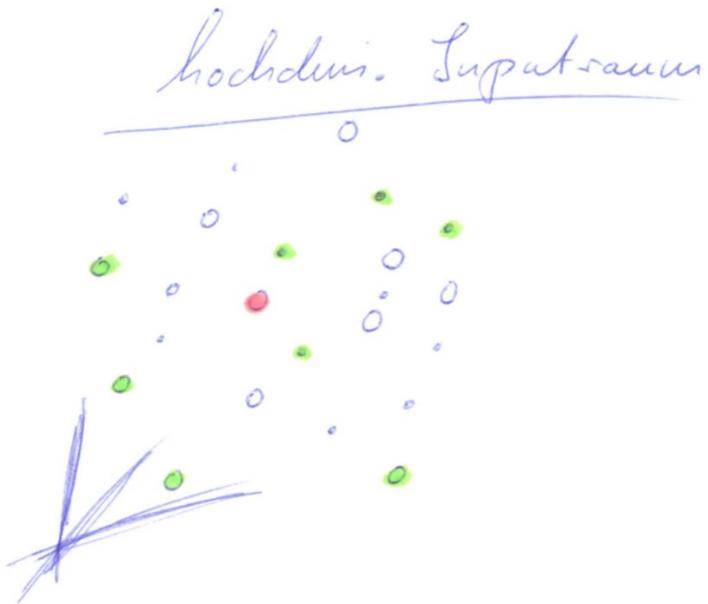
### Vorgehen:

- ▶ Zusammenarbeit mit externen Anbieter
- ▶ Konzentration auf „lohnende“ Bereiche (u.a. Betäubungsmittel)
- ▶ Einsatz eines Kohonen-Netzes

## Projekt 2: Kohonen-Skizze



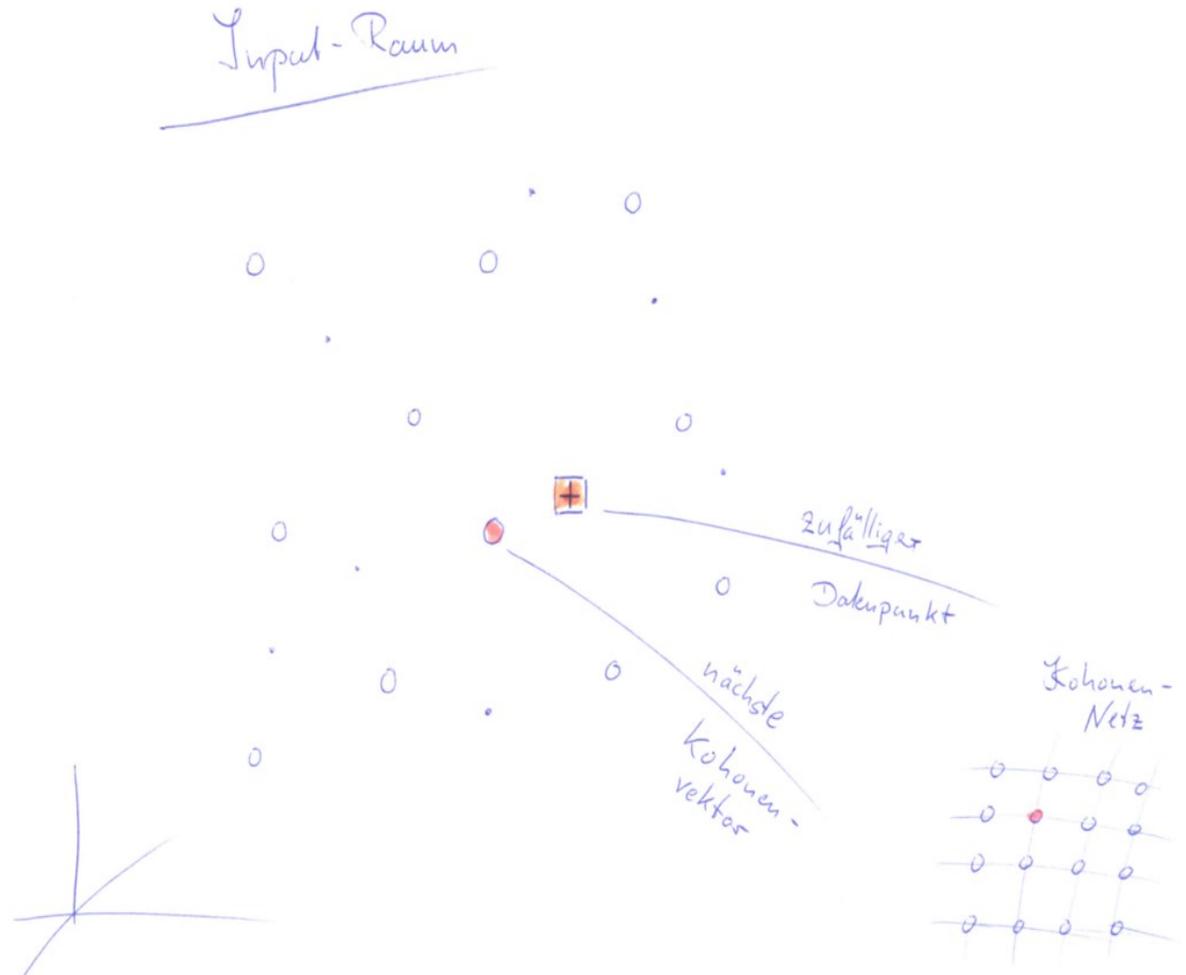
# Projekt 2: Kohonen-Skizze



## Projekt 2: Kohonen-Skizze

Algorithmus:

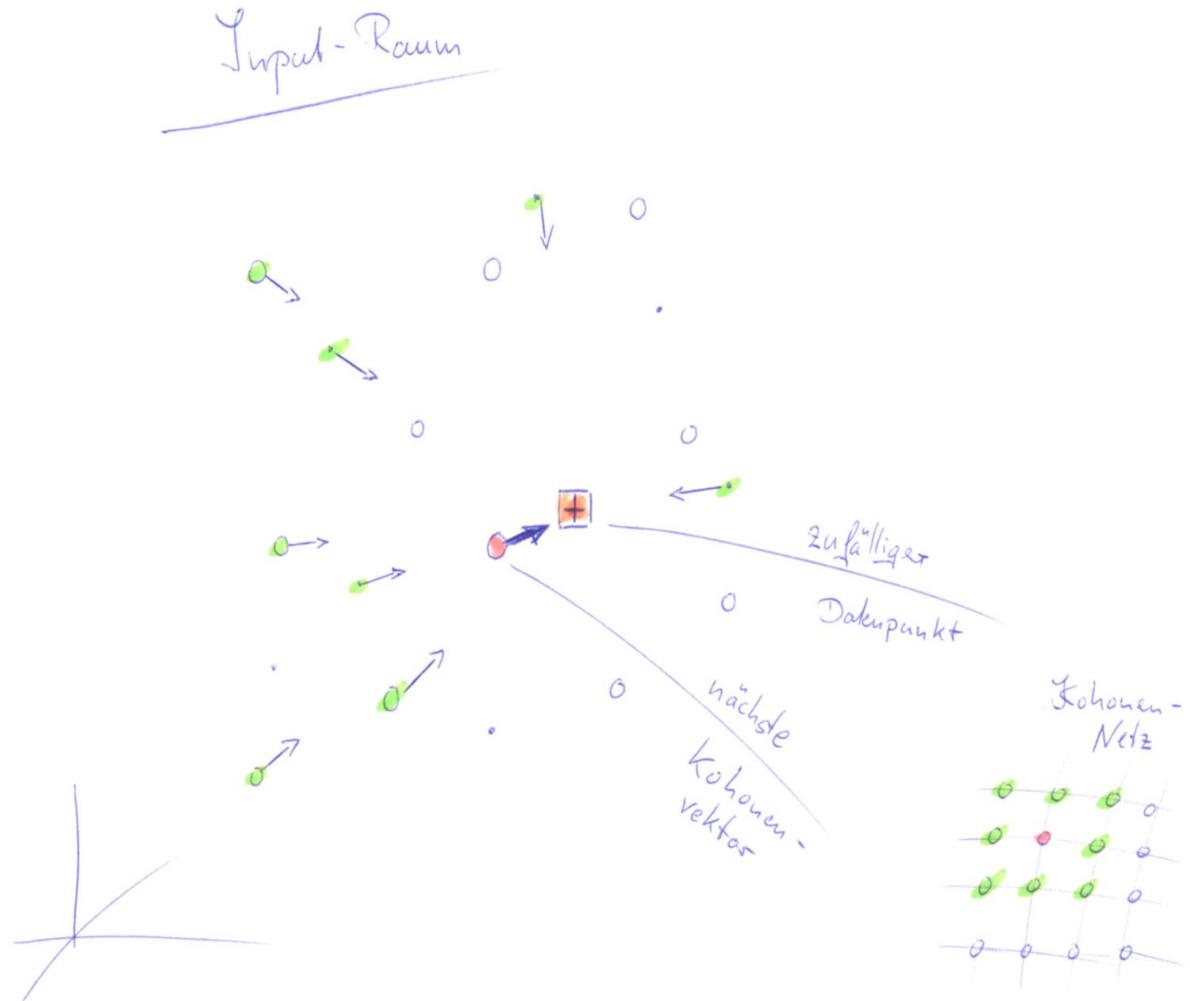
1. Wähle zufälligen Input-Vektor
2. Suche im Eingaberaum das dichteste „Bild-Neuron“
3. Aktualisiere die Bilder des dichtesten „Bild-Neurons“ und dessen Nachbarn



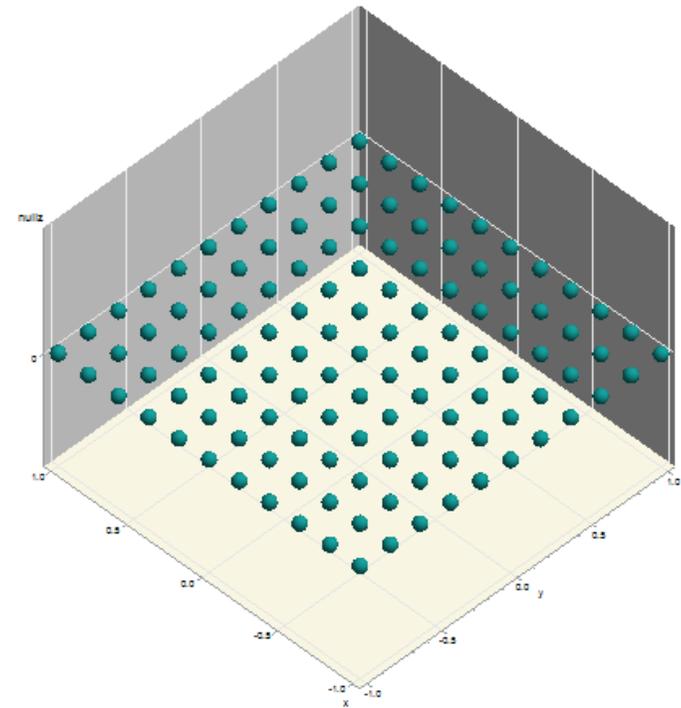
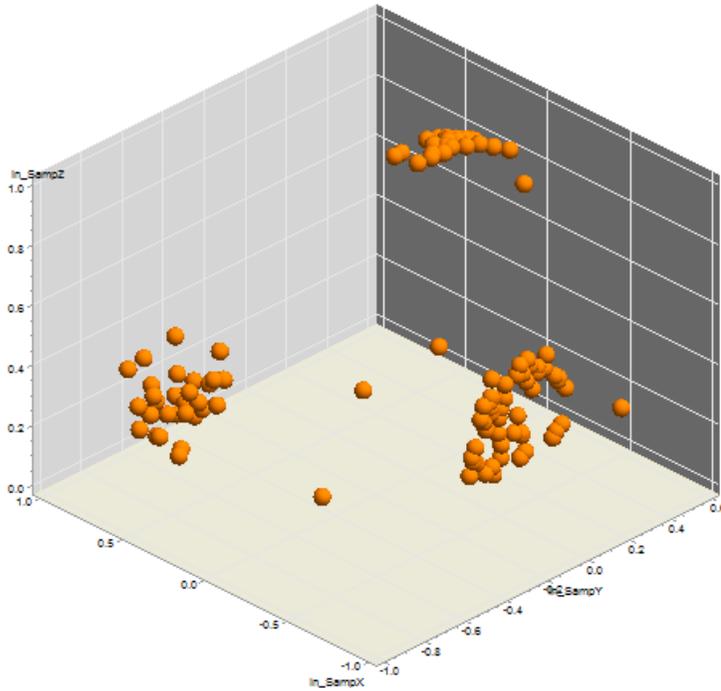
## Projekt 2: Kohonen-Skizze

Algorithmus:

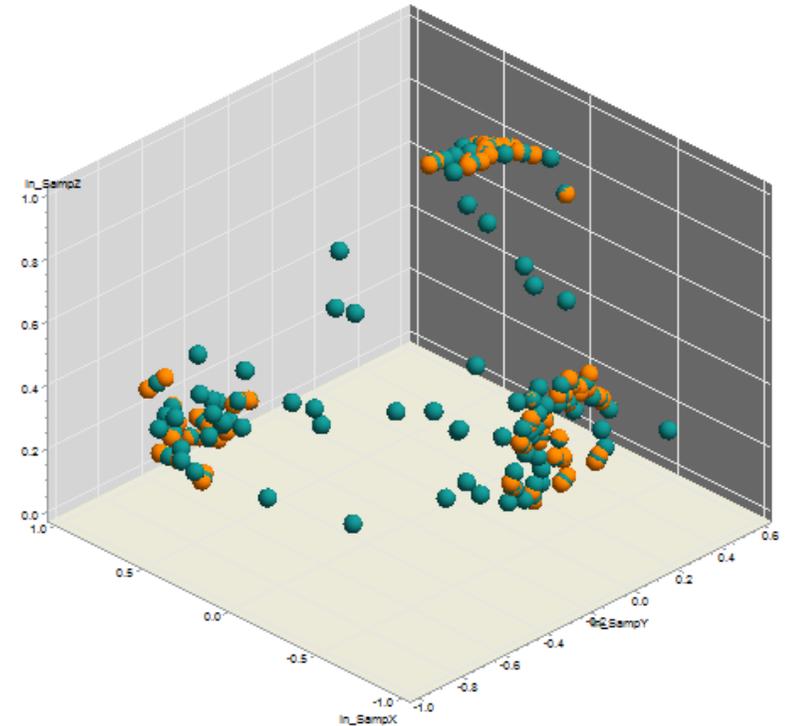
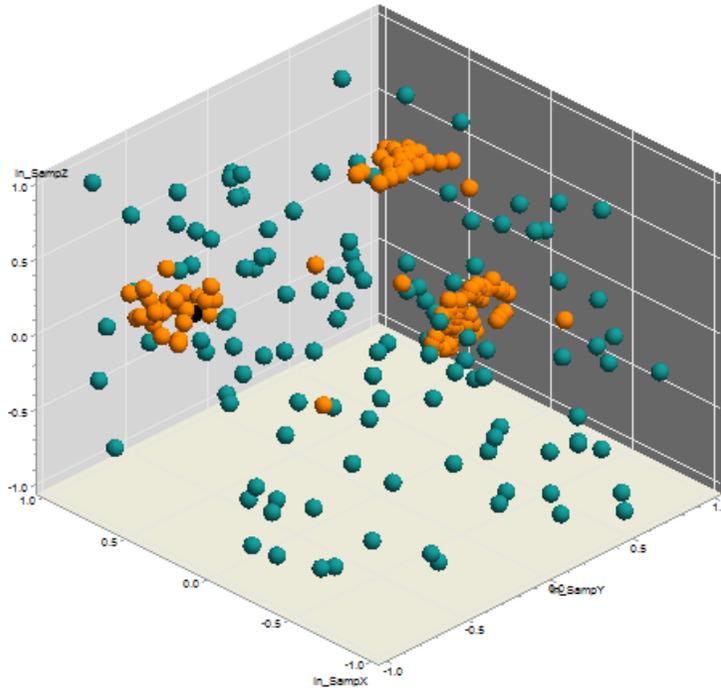
1. Wähle zufälligen Input-Vektor
2. Suche im Eingaberaum das dichteste „Bild-Neuron“
3. Aktualisiere die Bilder des dichtesten „Bild-Neurons“ und dessen Nachbarn



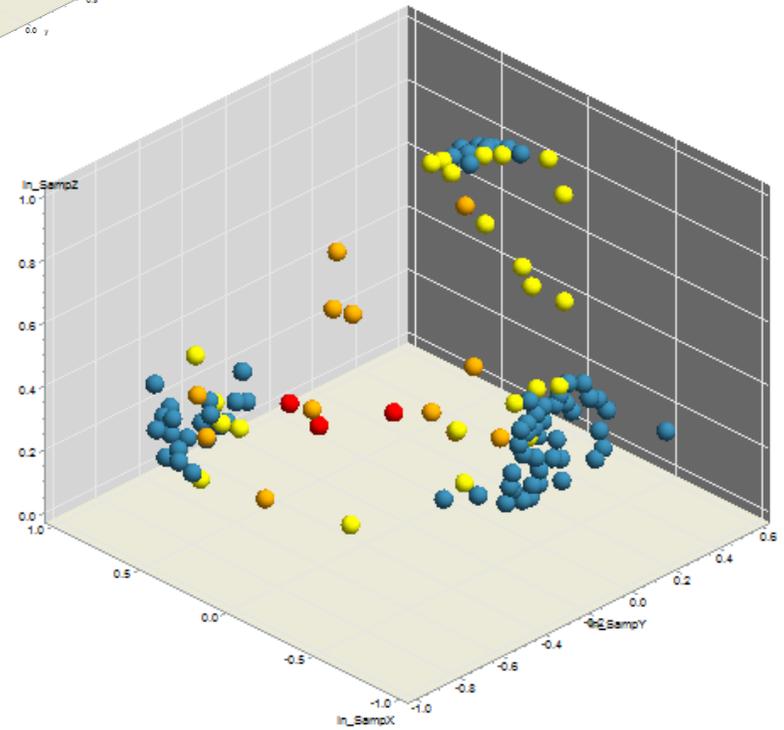
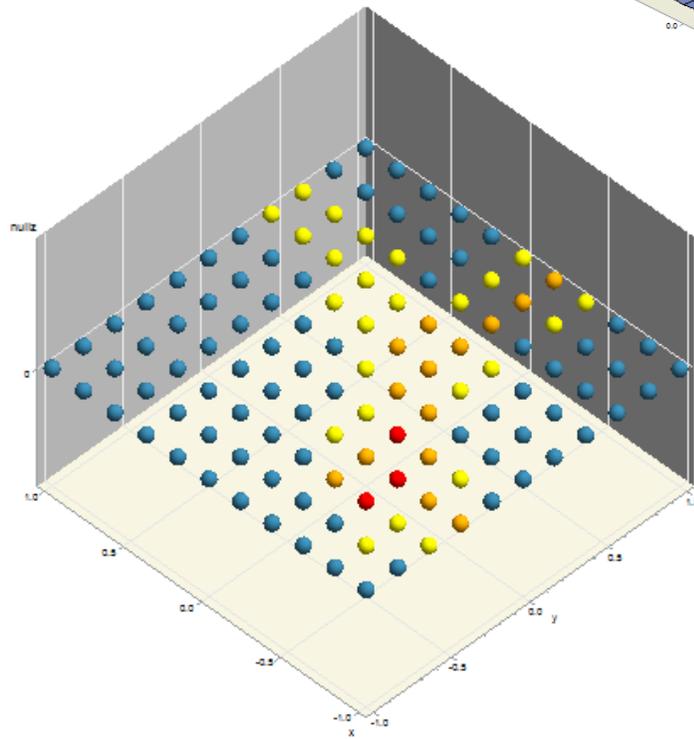
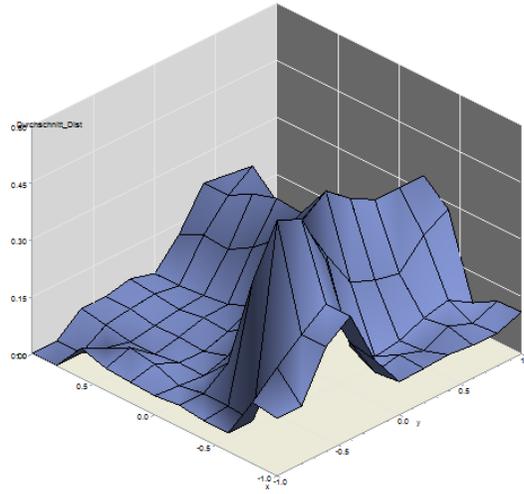
# Projekt 2: Kohonen-SOM



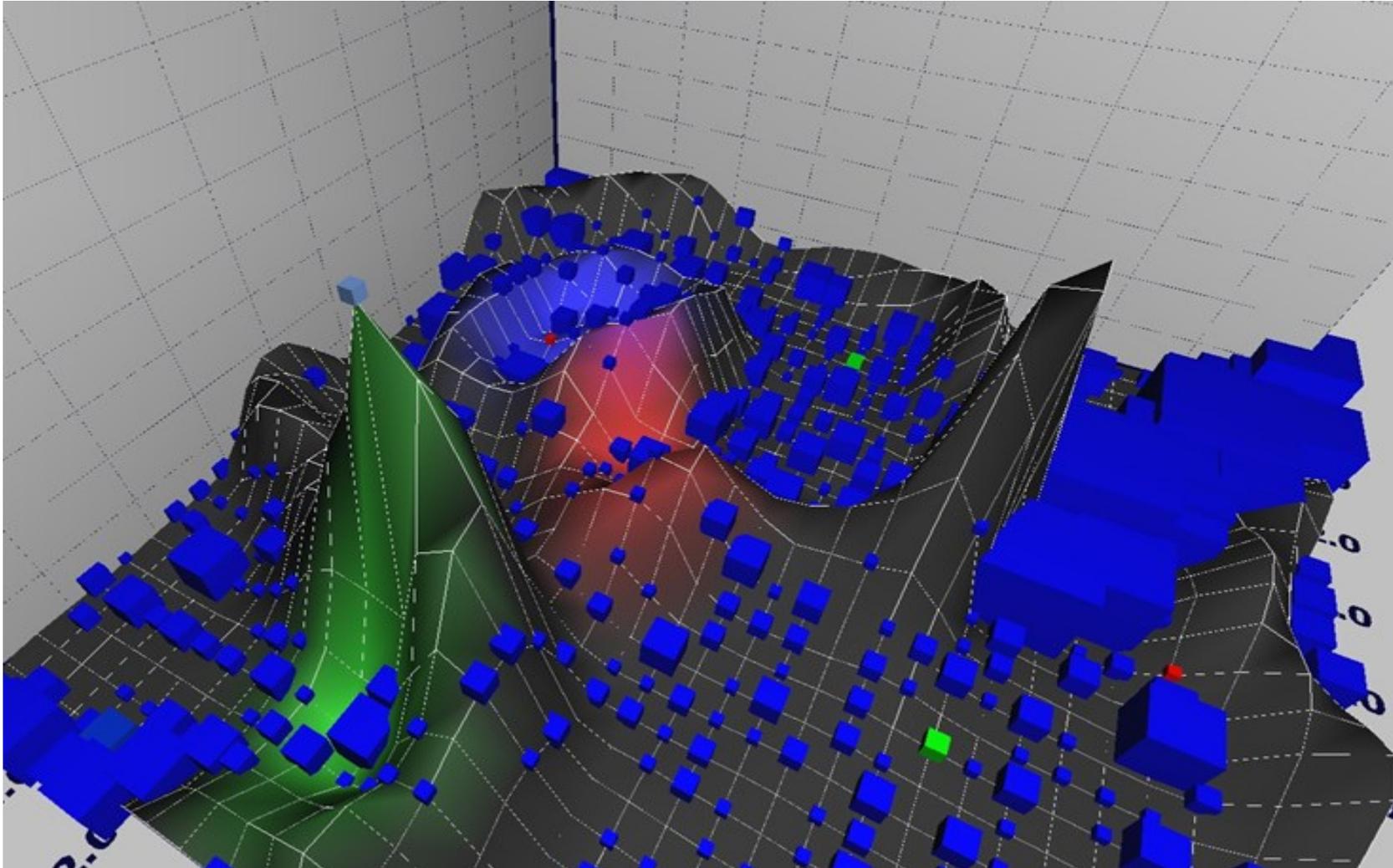
# Projekt 2: Kohonen-SOM



# Projekt 2: Kohonen-SOM



## Projekt 2: Kohonen-SOM



## Projekt 2: Kohonen-SOM

### Pro

- + Reduktion komplexer Zusammenhänge
  - Erhalt (von Teilen) der Topologie
- + i.A. robust
- + unüberwachtes Lernen
- + relativ universell einsetzbar

### Contra

- z.T. Blackbox
- Rechenzeit
- manuelle Inspektion der Ergebnisse nötig

## Projekt 2: Ergebnisse

### Ergebnisse der Analyse:

- ▶ sehr viele Übermaßbehandlungen gefunden
- ▶ mehrere Betrugsverdachtsfälle
- ▶ Staatsanwaltliche Ermittlung bei einem Fall (Betäubungsmittelsmissbrauch)
- ▶ mehrere 100.000 € Schadenssumme pro Jahr

### Aufwand:

- ▶ Aufwendige Datenvorbereitung
- ▶ z.T. hoher Nachbearbeitungsaufwand

### Fazit:

- ▶ vielversprechendes Verfahren für Erkennung auffälliger Muster
- ▶ aber kein Allheilmittel
- ▶ gezielte Datenaufbereitung könnte Trefferquote noch stark verbessern

A wide-angle photograph of a modern university building complex. The architecture features large glass windows and metal facades. In the foreground, there is a paved courtyard with a reflecting pool on the right side. Several people are walking in the courtyard. In the background, there are several tall poles with white banners that read 'HDK-Coblenz'. The sky is blue with a few clouds.

## Projekt 3: Krankheitsprognosen

## Projekt 3: Leberstiftung

### Krankheitsbild Hepatitis C:

- ▶ Virusinfektion, i.A. Übertragung durch Blut, Sexualkontakte
- ▶ weltweit ca. 150 Mio. Menschen infiziert
- ▶ ca. 5.000 Neuinfektionen pro Jahr in Deutschland
- ▶ i.A. keine Symptome bei Infektion
- ▶ ABER: chronischer Verlauf schädigt Leber (85% verlaufen chronisch)

### Ausgangspunkt:

- ▶ Forschungsprojekt der Deutschen Leberstiftung (Projektträger und Initiator)
- ▶ Datenlieferung und Vorbereitung (u.a.) von der HUK-COBURG

### Ziel der Analyse:

- ▶ Identifikation von Prädiktoren bei unentdeckter Hepatitis C
- ▶ Veröffentlichung der Ergebnisse

## Projekt 3: Leberstiftung

### Stand des Projekts:

- ▶ Analyse abgeschlossen
- ▶ mehrere Prädiktoren gefunden
- ▶ erste Ergebnisse scheinen vielversprechend, Validierung (in der HUK) steht noch aus

### Erstes Fazit:

- ▶ Win-Win Situation für Kunde und VU
- ▶ Problem der Kundenansprache



# Zusammenfassung

# Zusammenfassung

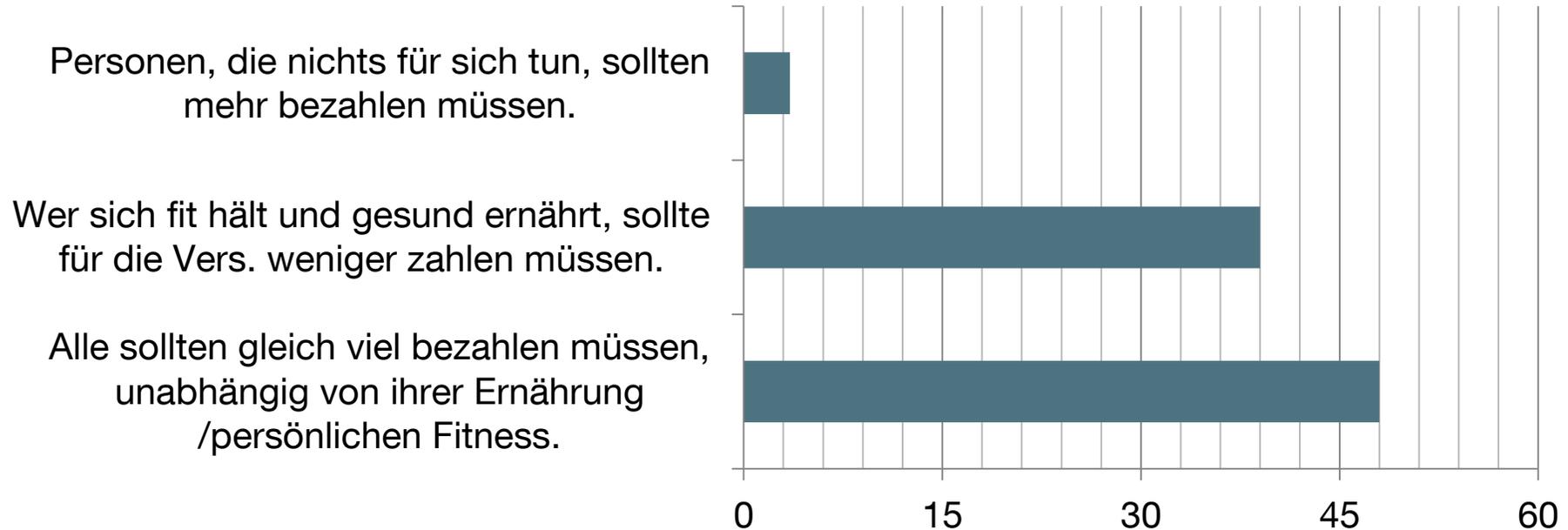
(Leistungs-) Datenanalysen in der PKV:

- ▶ können vor für die Kunden großen Nutzen bringen
- ▶ sollten immer auch unter Datenschutzaspekten bewertet werden
- ▶ können eine hohe Komplexität aufweisen
- ▶ benötigen i.d.R. eine gute (Daten-)Vorbereitung
- ▶ Kosten viel Zeit
- ▶ sollten nicht überschätzt werden
- ▶ können auch über Unternehmensgrenzen hinweg Nutzen stiften



# Datennutzung - Kundensicht

## Datennutzung - Kundensicht

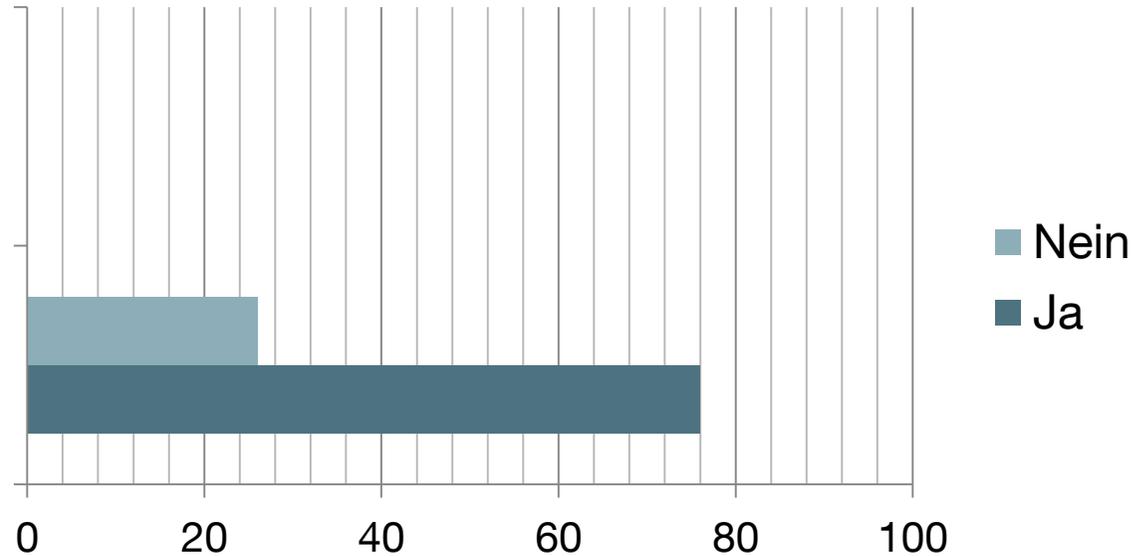


GDV Studie zur Datenweitergabe

[http://www.gdv.de/wp-content/uploads/2016/10/Praesentation\\_GfK-Befragung\\_Digitalisierung.pdf](http://www.gdv.de/wp-content/uploads/2016/10/Praesentation_GfK-Befragung_Digitalisierung.pdf)

# Datennutzung - Kundensicht

Mehrkosten für Raucher in RL/BU ist gerecht,  
n=772 1\*)

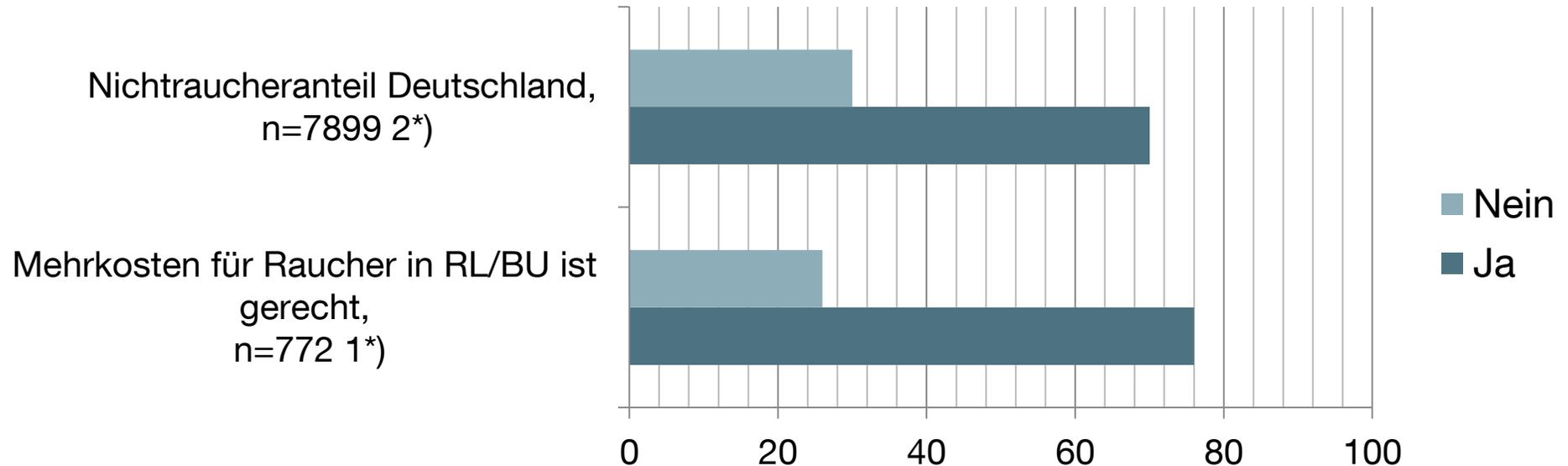


1\*) GDV Studie zur Datenweitergabe

„In der RL oder in der BU bezahlen Raucher heute bereits mehr als Nichtraucher, weil Raucher ein höheres Risiko haben. Finden Sie das gerecht?“

[http://www.gdv.de/wp-content/uploads/2016/10/Praesentation\\_GfK-Befragung\\_Digitalisierung.pdf](http://www.gdv.de/wp-content/uploads/2016/10/Praesentation_GfK-Befragung_Digitalisierung.pdf)

# Datennutzung - Kundensicht



1\*) GDV Studie zur Datenweitergabe

„In der RL oder in der BU bezahlen Raucher heute bereits mehr als Nichtraucher, weil Raucher ein höheres Risiko haben. Finden Sie das gerecht?“

[http://www.gdv.de/wp-content/uploads/2016/10/Praesentation\\_GfK-Befragung\\_Digitalisierung.pdf](http://www.gdv.de/wp-content/uploads/2016/10/Praesentation_GfK-Befragung_Digitalisierung.pdf)

2\*) Studie zur Gesundheit Erwachsener in Deutschland (DEGS1)“

# Datennutzung - Kundensicht

Fazit:

- ▶ Möglichkeiten der Datenanalyse in Öffentlichkeit oftmals auf Beitragskalkulation reduziert
- ▶ Angst der Kunden vor steigenden Beiträge
- ▶ Bereitschaft zur Weitergabe gegen Bonus z.T. vorhanden
- ▶ Beitragssteigerungen für Andere sind aber o.k.
  
- ▶ Vorteile der Datenanalysen für Kunden sollten besser kommuniziert werden