

SCANNERDATEN IN DER DEUTSCHEN PREISSTATISTIK

Konferenz „Messung der Preise“ am 7. und 8. Juni in Düsseldorf

Scannerdaten im VPI – Stufenplan zur Umsetzung in Deutschland

- 1. Stufe: Entwicklung von Methoden und Verfahren zur Berechnung von Preisindizes anhand von Marktforschungsdaten**
- 2. Stufe: Nutzung von Scannerdaten großer Einzelhandelsunternehmen zur Erprobung der Methodik**
- 3. Stufe: Anpassung der rechtlichen Grundlage**

1. Stufe: Entwicklung von Methoden und Verfahren

Kauf eines Scannerdatensatzes vom Marktforschungsinstitut Nielsen

- » Wöchentliche Daten zu 44 Produkten aus den Bereichen Lebensmittel und Getränke von Januar 2015 bis Dezember 2016
- » Informationen zu Absatz, Umsatz, Produktmerkmalen, GTIN, Region und Geschäftstyp

Auswertung des Datensatzes im Rahmen eines Eurostat-Projektes

- » Untersuchung verschiedener methodischer Fragestellungen
- » Berechnungen für Deutschland und die Bundesländer
- » Vergleiche mit den veröffentlichten Ergebnissen

1. Stufe: Entwicklung von Methoden und Verfahren

GTIN	Produktname	Hersteller	Marke	Umsatz (in €)	Absatz (in Stk)	Inhalt
NA	DIAMANT RAFFIN 1000 G	PFEIFER & LANGEN	DIAMANT	116266,95	175130,1	1000 G
SENSITIVE	EIGENMARKE 250 G BTL	HANDEL	EIGENMARKE	152224,25	175129,3	250 G
4002971012408	ALMIGHURT WALDBEER 3.8 PRZ 150 G BE	EHRMANN	ALMIGHURT	49342,48	175126,5	150 G
SENSITIVE	EIGENMARKE 3.5 PRZ 1.0 L TBRIK	HANDEL	EIGENMARKE	87492,57	175124	1.0 L
4008400290829	PINGUI SCHOKO 8 ST 240 G	FERRERO	PINGUI	338156,72	175122,6	240 G
4006544416006	SUEDZUCKER GELIER ZUCK 2+1 500 G	SUEDZUCKER	SUEDZUCKER	98064,59	175122,5	500 G

1. Stufe: Entwicklung von Methoden und Verfahren

Einschränkungen des Datensatzes

- » Scannerdaten nicht auf Filialebene verfügbar
- » Abbildung mancher Einzelhandelsunternehmen nur für Deutschland insgesamt
- » Regionale Abgrenzung bis maximal 36 Regionen

➔ **Mögliche Verzerrung der Ergebnisse**

1. Stufe: Entwicklung von Methoden und Verfahren

Beispiel: Verzerrung des Unit-Value (UV) durch Aggregation

Filiale	Absatz t1	Umsatz t1	UV t1	Absatz t2	Umsatz t2	UV t2
1	100	150	1,50 €	100	150	1,50 €
2	25	25	1,00 €	35	35	1,00 €
1+2	125	175	1,40 €	135	185	1,37 €

1. Stufe: Entwicklung von Methoden und Verfahren

Untersuchung der Länderstichprobe in Nordrhein-Westfalen und Hessen

- » Discounter zeigen erwartungsgemäß die höchste Preishomogenität
- » Ursache von Preisunterschieden nicht identifizierbar
 - » Regional / Zeitlich
 - » Freie Preissetzung? Identische Produkte? Andere Gründe?
- » Zahl der Beobachtungen zu gering für Auswertungen auf tiefer Aggregationsebene

1. Stufe: Entwicklung von Methoden und Verfahren

Berechnungen mit Statischem Warenkorb

» Traditioneller Ansatz

- » Dumping Filter
- » Top 70% der umsatzstärksten Produkte je COICOP-10-Steller und Geschäftstyp
- » Fortschreibung bei Produktausfall
- » Ersetzung bei mehrfachen, aufeinanderfolgenden Produktausfällen (länger als 1 Monat)

» Formel (nach Laspeyres):
$$I_t^L = \frac{\sum_{i=1}^N p_t^i \cdot q_0^i}{\sum_{i=1}^N p_0^i \cdot q_0^i}$$

1. Stufe: Entwicklung von Methoden und Verfahren

Berechnungen mit Dynamischem Warenkorb

- » Quality Adjusted Unit Value (QUAV) Index
 - » Kein Dumping Filter
 - » Zeitspanne pro Indexberechnung: 13 Monate
 - » Art der Verkettung: Movement Splice

» Formel (nach Chessa):
$$I_t^{QUAV} = \frac{\sum_{i=1}^{N_t} p_t^i \cdot q_t^i}{\sum_{i=1}^{N_t} v^i \cdot q_t^i} / \frac{\sum_{i=1}^{N_0} p_0^i \cdot q_0^i}{\sum_{i=1}^{N_0} v^i \cdot q_0^i}$$

mit
$$v^i = \sum_{z \in T} \varphi_z^i \frac{p_z^i}{I_z^{QUAV}} \text{ und } \varphi_z^i = \frac{q_z^i}{\sum_{s \in T} q_s^i}$$

1. Stufe: Entwicklung von Methoden und Verfahren

Berechnungen mit Dynamischem Warenkorb

» Gini-Eltetö-Köves-Szulc (GEKS) Index

» Dumping Filter

» Zeitspanne pro Indexberechnung: 13 Monate

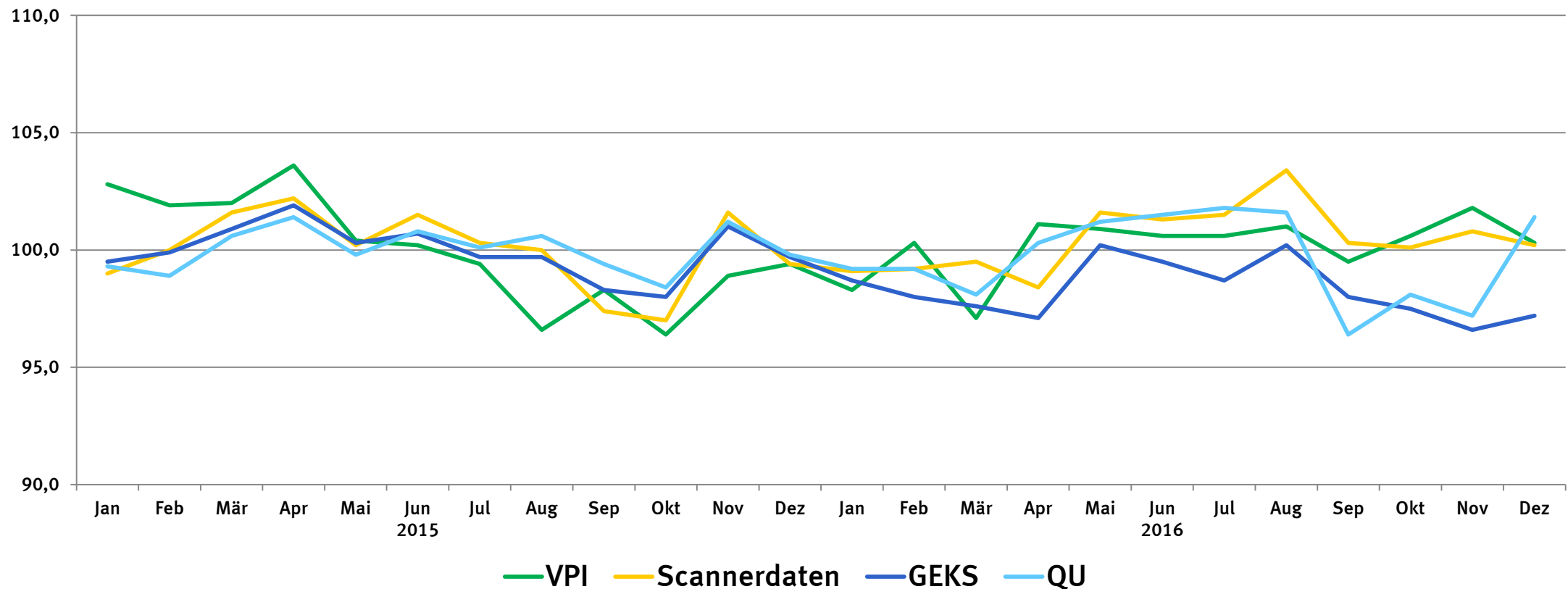
» Art der Verkettung: Movement Splice

» Formel (nach Ivancic, Diewert, Fox): $I_t^{GEKS} = \prod_{z \in T} \left(\frac{I_{tz}^{FISH}}{I_{0z}^{FISH}} \right)^{\frac{1}{|T|}}$

$$\text{mit } I_{tz}^{FISH} = \sqrt{I_{tz}^L * I_{tz}^{Pa}} \text{ wobei } I_{tz}^L = \frac{\sum_{i=1}^N p_t^i \cdot q_z^i}{\sum_{i=1}^N p_z^i \cdot q_z^i} \text{ und } I_{tz}^{Pa} = \frac{\sum_{i=1}^N p_t^i \cdot q_t^i}{\sum_{i=1}^N p_z^i \cdot q_t^i}$$

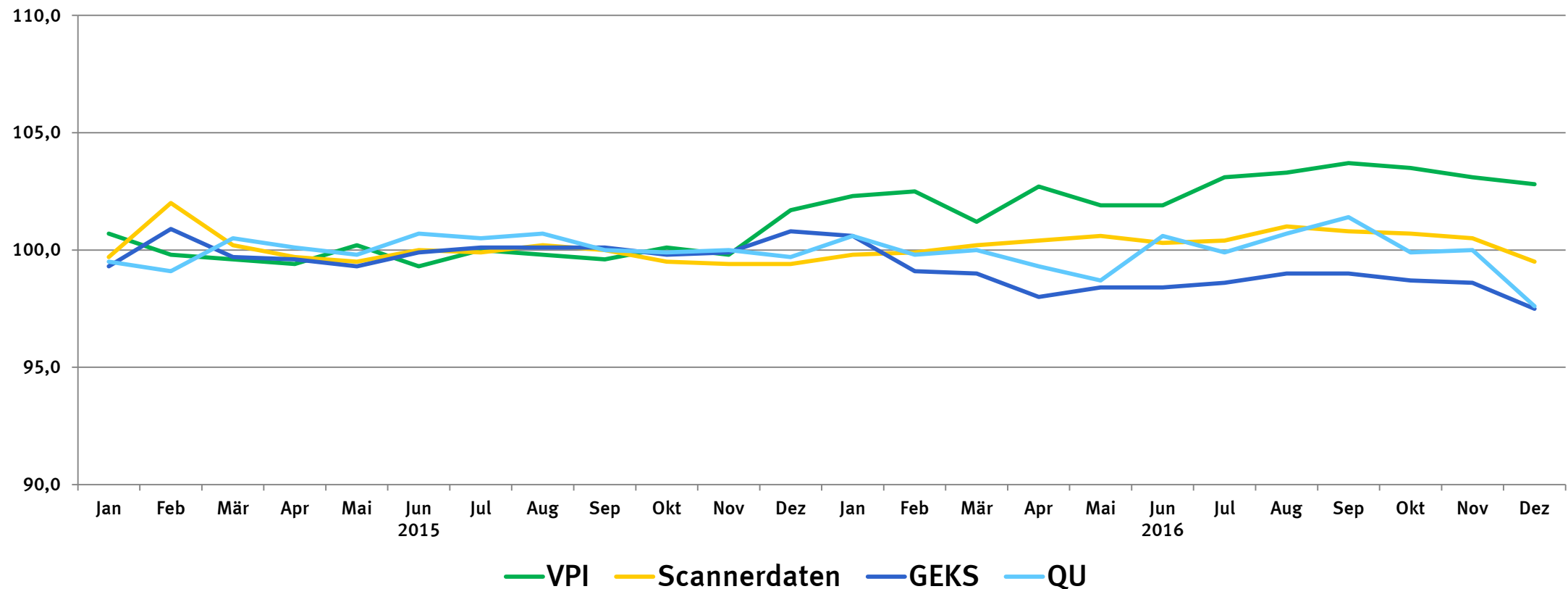
1. Stufe: Entwicklung von Methoden und Verfahren – Erste Berechnungen

Margarine - Preisindizes für 2015 und 2016 (2015=100)



1. Stufe: Entwicklung von Methoden und Verfahren – Erste Berechnungen

Koffeinfreie Limonade - Preisindizes für 2015 und 2016 (2015=100)



1. Stufe: Entwicklung von Methoden und Verfahren

Weitere Berechnungen in Planung

- » Testrechnungen mit variierenden Parametern
 - » Ausreißerbereinigung
 - » Dumping Filter
 - » Umsatzabdeckung
 - » Ersetzungskriterien
 - » Zeitspanne pro Indexberechnung
 - » Art der Verkettung

1. Stufe: Entwicklung von Methoden und Verfahren

Mapping der Artikel zur COICOP

- » Problem:
 - » (Einmalige) Klassifizierung von Tausenden verschiedener Artikel zu COICOP-10-Stellern notwendig
 - » Monatliche Anpassungen erforderlich (Neue Produkte)
- » Lösungsansatz: Mapping der Artikel mit Hilfe von Maschine Learning



1. Stufe: Entwicklung von Methoden und Verfahren

Softwareentwicklung im Auftrag von Eurostat:

- » Proof of Concept
- » Mapping über Produktbeschreibung
- » Verwendung von Trigrammen der Produktbeschreibung (Name) als Input-Variable
- » Verfahren des Maschine Learning zur Auswahl:
Support Vector Maschine, Random Forest, Logistische Regression

Adaptierung der Software zur Erprobung bei Destatis

- » Anpassung der Klassifikation (von 5- auf 10-Steller)
- » Exportschnittstelle schaffen

2. Stufe: Nutzung von Scannerdaten großer Einzelhandelsunternehmen

Erforderlich: Große Einzelhandelsunternehmen liefern auf freiwilliger Basis Scannerdaten zu Tests unter realen Bedingungen

- » Informationsgespräch mit dem BMWi und dem Handelsverband Deutschland am 10.04.2018
- » Vier Einzelhandelsunternehmen sind zu freiwilligen Scannerdatenlieferungen bereit
- » Erste Scannerdatenlieferung kurz vor dem Abschluss
- » Erster Untersuchungsgegenstand: Wie groß ist der Aufwand zur Datenaufbereitung?

3. Stufe: Anpassung der rechtlichen Grundlage

Europäische Grundlage

- » Artikel 5 Absatz 4 der Verordnung (EU) 2016/792:
„Auf Verlangen der nationalen Stellen, ... übermitteln die statistischen Einheiten soweit verfügbar elektronische Aufzeichnungen von Transaktionen, z.B. Scannerdaten...“

Anpassung der nationalen Gesetzesgrundlage

- » Arbeiten von Destatis sind 2018 gestartet
- » Prüfung inwieweit Anpassung der nationalen Gesetzesgrundlage erforderlich ist

VIELEN DANK FÜR IHRE AUFMERKSAMKEIT !

Matthias Bieg

Telefon: +49/(0) 611 / 75 39 93

matthias.bieg@destatis.de

www.destatis.de

