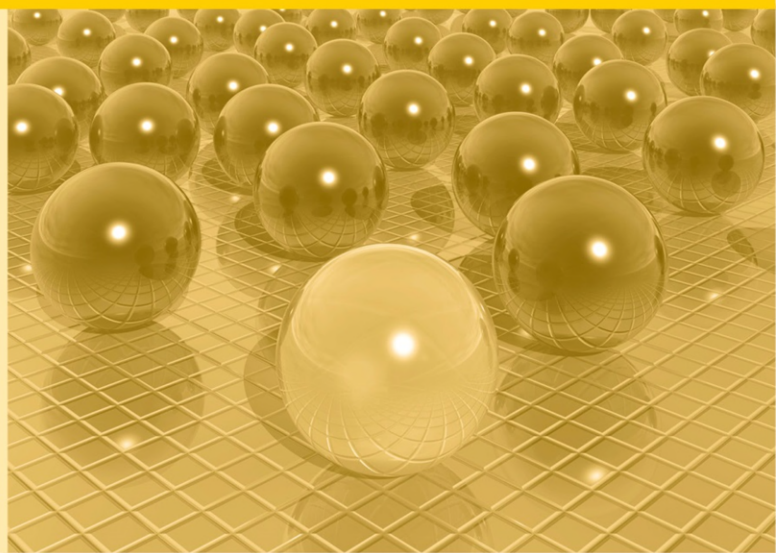
Metadatenreport



Teil II: Produktspezifische Informationen zur Nutzung der Einzeldaten des Verbraucherpreisindex 2025 (EVAS-Nummer: 61111) per Scientific Use File

DOI: 10.21242/61111.2025.00.00.3.1.0

Version 1

STATISTISCHE ÄMTER
DES BUNDES UND DER LÄNDER
FORSCHUNGSDATENZENTREN

Impressum

Herausgeber: Statistische Ämter des Bundes und der Länder
Herstellung: Information und Technik Nordrhein-Westfalen
Telefon 0211 9449-01 • Telefax 0211 9449-8000
Internet: www.forschungsdatenzentrum.de
E-Mail: forschungsdatenzentrum@it.nrw.de

Fachliche Informationen
zu dieser Veröffentlichung:

Forschungsdatenzentrum der
Statistischen Ämter der Länder
– Standort Hessen –
Tel.: 0611 3802-822
forschungsdatenzentrum@statistik.hessen.de

Informationen zum Datenangebot:

Statistisches Bundesamt
Forschungsdatenzentrum

Tel.: 0611 75-2420
Fax: 0611 72-3915
forschungsdatenzentrum@destatis.de

Forschungsdatenzentrum der
Statistischen Ämter der Länder
– Geschäftsstelle –
Tel.: 0211 9449-2883
Fax: 0211 9449-8087
forschungsdatenzentrum@it.nrw.de

Erscheinungsfolge: unregelmäßig
Erschienen im November 2026

Diese Publikation wird kostenlos als PDF-Datei zum Download unter www.forschungsdatenzentrum.de angeboten.

© Information und Technik Nordrhein-Westfalen, Düsseldorf, 2026
(im Auftrag der Herausgebergemeinschaft)

Vervielfältigung und Verbreitung, nur auszugsweise, mit Quellenangabe gestattet. Alle übrigen Rechte bleiben vorbehalten.

Fotorechte Umschlag: ©artSILENCEcom –Fotolia.com

Empfohlene Zitierung:

Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder: Metadatenreport. Teil II: Produktspezifische Informationen zur Nutzung der Einzeldaten des Verbraucherpreisindex 2025 (EVAS-Nummer: 61111) per Scientific Use File. Version 1. DOI: 10.21242/61111.2025.00.00.3.1.0. Wiesbaden 2026.

Metadatenreport

Teil II: Produktspezifische Informationen zur Nutzung der Einzeldaten des Verbraucherpreisindex 2025 (EVAS-Nummer: 61111) per Scientific Use File

DOI: 10.21242/61111.2025.00.00.3.1.0

Version 1

Inhalt

1 Datenaufbereitung in den FDZ 2
1.1 Datenaufbereitung 4
1.2 Anonymisierungsmaßnahmen 5
1.3 Methodik der Verknüpfung 7
2 Produkt 8
2.1 Merkmale und Merkmalsbeschreibung 8
2.2 Vergleichbarkeit der Merkmale über die Zeit 25
2.3 Eckwerte relevanter Merkmale und Merkmalskombinationen 26
2.4 Auswertbare regionale Ebene 26
3 Praktische Hinweise 27
3.1 Hinweise zur Geheimhaltung 27
3.1.1 Gesetzliche Grundlagen der statistischen Geheimhaltung 27
3.1.2 Geheimhaltung von Ergebnissen 28
3.1.3 Praktische Tipps zur Vermeidung von Geheimhaltungsfällen ... 29
3.2 FAQ 30
3.3 Verfügbare Tools 30
4 Literaturverzeichnis 30
Anhang 1 – Merkmalsübersicht 31
Anhang 2 – Anonymisierungskonzept der Feinbeschreibungsmerkmale (Auspraegung1 bis Auspraegung10) 34
Anhang 3 – Umschlüsselungstabelle der SEA 2013 auf SEA 2021 36
Anhang 4 – Umschlüsselungstabelle mit der Zuordnung der „gelöschten“ 10-Steller des Erhebungskatalogs 2020 (Basis 2020 = 100) zu den entsprechenden Fortschreibungspositionen des Erhebungskatalogs 2025 39

1 Datenaufbereitung in den FDZ

Wichtiger Hinweis zu den Einzeldaten des Berichtsjahres 2025

In Vorbereitung der Umstellung auf das Basisjahr 2025 im Januar 2028 erfolgt die Erhebung der Verbraucherpreise seit Dezember 2024 turnusgemäß nach einem neuen Erhebungskatalog. Dieser berücksichtigt insbesondere die neuen Klassifikationen: „Europäische Klassifikation der Verwendungszwecke des Individualkonsums“ („European Classification of Individual Consumption according to Purpose“ – ECOICOP Version 2, COICOP 2018) und „Systematik der Einnahmen und Ausgaben der privaten Haushalte, Ausgabe 2021“ (SEA 2021). Dadurch hat sich die Struktur des Erhebungskatalogs deutlich verändert.

Ein Teil der bisherigen 10-Steller (Erhebungspositionen) wurde im Zuge dieser Umstellung direkt umgeschlüsselt: Die bisherige Erhebungsposition existiert also weiterhin, hat jedoch einen neuen 10-Steller-Code. Um diesen „neuen“ 10-Steller-Code auf den für das aktuelle Basisjahr 2020 verwendeten 10-Steller-Code zurückzuschlüsseln, befindet sich im Anhang 3 eine entsprechende Tabelle. Für den vorliegenden SUF wurde die Umschlüsselung bereits am Datenmaterial vorgenommen.

Neben diesen direkten Umschlüsselungen ergeben sich durch Umsortierungen auch inhaltlich veränderte Erhebungspositionen. In der Folge werden 10-Steller des aktuellen Basisjahres (hier: 2020) auf der Erhebungsebene durch einen neuen 10-Steller abgelöst: Der bisherige 10-Steller existiert auf der Erhebungsebene also nicht weiter, sondern wird auf der Berechnungsebene mit der Entwicklung des neuen 10-Stellers (Fortschreibungsposition) fortgeschrie-

ben. Für die Fortschreibungspositionen (neue 10-Steller) werden dazu ab Januar 2025 neue Indexreihen mit dem Startmonat Dezember 2024 = 100 berechnet. Mit der Preisentwicklung dieser neuen Indexreihen werden die aktuellen („alten“) gelöschten 10-Steller bis zur Revision (hier: bis einschließlich Dezember 2027) auf das neue Basisjahr fortgeschrieben und wie bisher auch zur Berechnung der höheren Aggregate herangezogen.

Die Fortschreibungspositionen lassen sich auf der Einzeldaten-Ebene im Gegensatz zu den im Anhang 3 genannten 10-Stellern nicht immer eindeutig auf den Stand der SEA 2013 umschlüsseln, weswegen diese Positionen mit einem separaten Datensatz bereitgestellt werden. Das Datenmaterial des VPI 2025 besteht daher aus zwei Datensätzen: dem Datensatz nach Basisjahr 2020 und dem Datensatz mit den Fortschreibungspositionen des Erhebungskatalogs für das Basisjahr 2025. Eine Umschlüsselungstabelle der „gelöschten“ 10-Steller zu den Fortschreibungspositionen befindet sich in Anhang 4.

Bei den zurückliegenden Umstellungen der Basisjahre wurden die jeweils nur wenigen Fortschreibungspositionen nicht im FDZ-Datenmaterial bereitgestellt. Aufgrund der erheblichen Änderungen des Erhebungskatalogs in Vorbereitung auf das zukünftige Basisjahr 2025 hat fast die Hälfte der 10-Steller eine neue Codierung durch klassifikationsbedingte Umsortierungen erhalten, die eine Fortschreibung bedingen. Die gesonderte Bereitstellung der Fortschreibungspositionen ist daher erforderlich. Darüber hinaus beinhaltet der Datensatz neue Positionen, die zwar für den VPI erst mit der nächsten Revision indexrelevant werden, für den HVPI aber bereits im Jahr 2025 ein Gewicht erhalten haben.

1.1 Datenaufbereitung

Scientific Use Files (SUF) sind standardisierte Einzeldatensätze, die von den Statistischen Ämtern für wissenschaftliche Vorhaben bereitgestellt werden. Im Gegensatz zu Gastwissenschaftlerarbeitsplätzen oder der kontrollierten Datenfernverarbeitung können SUF außerhalb der geschützten Räumlichkeiten der amtlichen Statistik genutzt werden.

§ 16 Abs. 6 Bundesstatistikgesetz (BStatG) ermöglicht den Statistischen Ämtern des Bundes und der Länder Einzelangaben an Hochschulen und andere Einrichtungen der unabhängigen wissenschaftlichen Forschung zu übermitteln. Voraussetzung hierfür ist jedoch, dass die Einzelangaben nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft zugeordnet werden können. „Unverhältnismäßig“ bedeutet, dass der Aufwand für eine Reidentifikation höher ist, als der Nutzen, der daraus gezogen werden kann.

SUF haben daher stets zwei sehr gegensätzliche Anforderungen zu erfüllen. Auf der einen Seite sollen sie für einen möglichst großen Teil an Forschungsvorhaben ein Höchstmaß an Analysepotenzial bieten, zum anderen müssen sie den strengen Maßgaben der faktischen Anonymität genügen.

Den Anforderungen des BStatG zur Sicherstellung der faktischen Anonymität wird bei einem SUF zum einen dadurch entsprochen, dass mit dem Datennutzer ein Vertrag geschlossen wird, in dem ein De-Anonymisierungsverbot festgeschrieben ist und Sanktionen bei Nichteinhaltung dieses Verbotes festgesetzt werden. Zum anderen wird die faktische Anonymität durch datenverändernde Maßnahmen erreicht. Im Folgenden werden die datenverändernden Maßnahmen beschrieben.

Für detaillierte Informationen zum Verbraucherpreisindex für Deutschland wird an dieser Stelle auf die Metadaten zur Erhebung (Teil I) verwiesen.

1.2 Anonymisierungsmaßnahmen

Für den SUF des Verbraucherpreisindex für Deutschland (VPI) werden nur die Merkmale bereitgestellt, die sich für die Durchführung wissenschaftlicher Vorhaben auf Basis faktisch anonymer Daten eignen. Eine Übersicht über die im SUF enthaltenen Merkmale befindet sich in Anhang 2.

Neben der Löschung von Merkmalen sind weitere Maßnahmen erforderlich, um die faktische Anonymität der Daten sicherzustellen. Diese Maßnahmen stellen im Wesentlichen auf die Merkmale ab, deren Information sowohl in den vertraulichen Daten als auch in dem möglichen Zusatzwissen eines potentiellen Datenangreifers enthalten sein können (so genannte Überschneidungsmerkmale). Bei der Recherche des möglichen Zusatzwissens haben sich insbesondere die Regionalangabe, die Feinbeschreibungsmerkmale, die Geschäftstypengewichte und die Kombination aus den Merkmalen Monat, 10-Steller, Bundesland und Geschäftstyp als mögliche Überschneidungsmerkmale herausgestellt.

Im Folgenden werden die Maßnahmen beschrieben, die auf die einzelnen Merkmale angewendet wurden.

Gemeinde

Dieses Merkmal entfällt, sodass eine regionale Gliederung nur bis zur Ebene der Bundesländer möglich ist. Dadurch werden seltene und regional einmalig vorkommende Branchen vor Aufdeckung geschützt (z.B. Schornsteinfeger, Apotheken, Möbelhäuser, Instrumentenbauer etc.). Da für die Nachverfolgung der erhobenen Produkte über die Zeit die Gemeinde ein zentrales Merkmal ist, wird eine ID erstellt, die die erhobenen Produkte im Zeitverlauf identifiziert.

Berichtsstelle

Wie im On-Site-Material wird für die Berichtsstelle eine neue Nummer vergeben, die allerdings die Systematik der Klassifikation – dezentral, zentral und Bund – widerspiegelt. Die neue Berichtsstellennummer wird konsistent für alle SUF Berichtsjahre vergeben.

Auspraegung1 bis Auspraegung10 (Feinbeschreibungsmerkmale)

Die Inhalte der Merkmale werden vollständig pseudonymisiert[1]. Entgegen der On-Site-Nutzung, ist die Verwendung einer Whitelist, also die Darstellung bestimmter Begriffe im Klartext, nicht möglich. Die Pseudonyme werden konsistent für alle Berichtsjahre vergeben.

Entfernen von Erhebungspositionen

Durch eine gemeinsame Betrachtung der Merkmale Monat, 10-Steller, Bundesland und Geschäftstyp werden Erhebungspositionen identifiziert, die zusammen in weniger als drei verschiedenen dezentralen Berichtsstellen erhoben wurden. Da diese Kombination von Eigenschaften die Geheimhaltung der dezentralen Berichtsstellen gefährden kann, werden die zugrundeliegenden Erhebungspositionen (ca. 1,1 Prozent) aus dem Datensatz entfernt.

Geschäftstypengewichtung

Aufgrund der sensiblen Information in der Geschäftstypengewichtung, wird das Gewicht auf der 10-Steller-Ebene nur für Deutschland insgesamt (also durch

[1] Für detaillierte Informationen zum Verfahren der Pseudonymisierung und Vereinheitlichung siehe Kaukal und Peters 2019.

die Bundesländer gewichtete Durchschnitte der Geschäftstypengewichte) zur Verfügung gestellt.

1.3 Methodik der Verknüpfung

Der verfügbare Datensatz stellt einen ‚gestapelten' Datensatz aller Berichtsmonate eines Jahres dar und bildet somit ein Berichtsjahr ab. Die einzelnen Waren und Dienstleistungen erhalten im Original keine eigene ID, die über die Berichtsmonate konstant bleibt, da Produktwechsel innerhalb einer Güterklasse vorkommen können. Für die Bundesländer erhobene Preisreihen eines Berichtsmonats können über den On-Site-Zugang jedoch mit Hilfe der Merkmale „Gemeinde", „Berichtsstelle", „COICOP", „MB" (Meldebogennummer) und „Produktvariante" über die Berichtsmonate innerhalb eines Basisjahres hinweg identifiziert werden. Für das Bundesgebiet insgesamt gültige zentral erhobene Preisreihen (Bundeslandkennung „99") sind durch die Merkmale „Gemeinde", „Berichtsstelle", „COICOP" und „Produktvariante" eindeutig identifizierbar.

Da im vorliegenden SUF die Gemeinde zur Identifikation der Preisreihe nicht zur Verfügung steht, wird durch das FDZ eine systemfreie ID erstellt, die über die Berichtsjahre eines Basisjahrs hinweg konsistent vergeben wird. Beobachtungen derselben Produkte bzw. der entsprechenden Ersatzprodukte nach einem Erzeugniswechsel sind somit auch im SUF über die Zeit hinweg möglich.

2 Produkt

2.1 Merkmale und Merkmalsbeschreibung

Genannt sind zuerst der Merkmalsname und anschließend die Merkmalsbeschreibung.

ID – ID einer Erhebungsposition im Basisjahr

Da das Merkmal „Gemeinde“, das für die Identifikation eindeutiger Preisreihen notwendig ist, im SUF nicht zur Verfügung steht, wird seitens des FDZ eine neue ID erstellt. Diese ID wird konsistent innerhalb eines Basisjahres vergeben und ermöglicht somit die Nachverfolgung einzelner Erhebungspositionen über die Zeit.

Die ID wird anhand der Merkmale „Gemeinde“, „Berichtsstelle“, „MB“, „COICOP“ und „Produktvariante“ erstellt. Dies gilt allerdings nur für den Zeitraum zwischen zwei Meldebogenreformen. Bei einer Meldebogenreform wird der Erhebungskatalog an das Konsumverhalten der Bevölkerung angepasst. Eine Anpassung des Erhebungskatalogs findet alle fünf Jahre statt (zuletzt zum Dezember 2019).

Monat – Erhebungsmonat

Monat, in dem die Preiserhebung des Produkts durchgeführt wurde.

Jahr – Erhebungsjahr

Jahr, in dem die Preiserhebung des Produkts durchgeführt wurde.

Bundesland – Erhebungsbundesland

Amtlicher Schlüssel des Bundeslandes, in dem der Preis für das Produkt erhoben wurde.

Ausprägungen:

01 = Schleswig-Holstein
02 = Hamburg
03 = Niedersachsen
04 = Bremen
05 = Nordrhein-Westfalen
06 = Hessen
07 = Rheinland-Pfalz
08 = Baden-Württemberg
09 = Bayern
10 = Saarland
11 = Berlin
12 = Brandenburg
13 = Mecklenburg-Vorpommern
14 = Sachsen
15 = Sachsen-Anhalt
16 = Thüringen
99 = Bund

Preise, die dem Code 99 „Bund“ zugewiesen werden, werden durch das Statistische Bundesamt erhoben und besitzen keinen Bundeslandbezug. Dies betrifft bspw. die Angaben zu Pauschalreisen.

Gewicht_BL – Gewicht des Bundeslandes

Gewichte, mit denen im Rahmen der Berechnung des Verbraucherpreisindex die Teilindizes der Bundesländer pro 10-Steller zu Teilindizes pro 10-Steller der Bundesebene zusammengefasst werden.

Berichtsstelle – Identifizierungsnummer der Berichtsstelle

Identifizierungsnummer der Berichtsstelle für die Preiserhebung. Die Vergabe der Nummern entspricht folgender Systematik:

1 – 4999 = Diese Nummern sind den dezentralen Berichtstellen in den Ländern vorbehalten. Sie werden einmalig in Gemeinden, aber mehrfach im Bundesland vergeben.

5000 – 9999 = Hierbei handelt es sich um zentral erhobene Preise (siehe auch „Erfassungsart“). Das bedeutet, dass ein Statistisches Landesamt oder das Statistische Bundesamt stellvertretend den Preis bei einem überregionalen Anbieter erhebt (z. B. einer Handelskette) und dieser auf die übrigen Bundesländer angewendet wird. Des Weiteren ist es möglich, dass innerhalb eines Landes nur eine Berichtsstelle besucht wird und deren Preise auf die übrigen Gemeinden mit Filialen übertragen wird (z. B. eine Discounterkette).

>= 10 000 = Hierbei handelt es sich um Preise, die das Statistische Bundesamt bei zentralen Anbietern (bspw. Reisebüros, E-Commerce) erhebt und für das gesamte Bundesgebiet gültig sind.

Eine Berichtseinheit (Einzelhandels- und Dienstleistungseinheiten, einschließlich öffentlich-rechtlicher und staatlicher Anbieter) kann bei mehreren Erhebungspositionen verschiedene Berichtsstellennummern aufweisen.

Für die Nutzung im FDZ wird die Berichtsstellennummer in eine systemfreie Nummer überführt. Die Bereiche der drei Kategorien (0-4999 = dezentral; 5000-9999 = zentral; >=10000 = zentral, Statistisches Bundesamt) bleiben erhalten.

Berichtsstellenmultiplikator – Anzahl der Filialen

Gibt an, mit welcher expliziten Gewichtung alle Preise einer Berichtsstelle (Erhebungseinheit) in die Indexberechnung einfließen. Das Feld „Anzahl der Filialen“ ist in der Regel mit dem Wert „1“ belegt.

Multiplikatoren (auch: Vervielfacher) bieten die Möglichkeit, unterhalb der Geschäftstypengewichtung eine Wägungskomponente zu schaffen. Dadurch kann der Einfluss der einzelnen Preisreihe innerhalb des Durchschnittspreises der Elementarindexabgrenzung variiert und beeinflusst werden.

GKat – Geschäftskategorie

Bei der Geschäftskategorie wird zwischen folgenden Geschäftstypen unterschieden:

1 = Kaufhaus/Warenhaus

2 = Verbrauchermarkt/SB-Warenhaus

3 = Supermarkt

4 = Discounter/Fachmarkt

5 = Fachgeschäft

6 = sonstiger Einzelhandel

7 = Dienstleistungen/Miete

8 = Versandhandel/Internethandel

Gewicht_GKat – Gewicht der Geschäftstypen

Gewicht, mit denen die Elementarindizes jeder Elementarindexabgrenzung (pro Geschäftstyp) zu Teilindizes pro 10-Steller (siehe Merkmal „COICOP“) auf Bundeslandebene zusammengefasst werden.

Das Merkmal „Gewicht_GKat“ enthält nur für die einzelnen Positionen („COICOP-Typ“ = 10) ein Geschäftstypengewicht, bei Varianten eines 10-Stellers ist das Feld leer. Das Gewicht eines Geschäftstyps für einen 10-Steller mit Varianten kann stattdessen dem Merkmal „Gewicht_GKat_Oberposition“ entnommen werden.

Aufgrund der sensiblen Angaben des Geschäftstypengewichts, wird das Gewicht auf der 10-Steller-Ebene nur bezogen auf Deutschland insgesamt (also durch die Bundesländer gewichtete Durchschnitte der Geschäftstypengewichte) zur Verfügung gestellt.

Gewicht_GKat_Oberposition – Gewicht der Geschäftstypen eines 10-Stellers mit Varianten (übergeordnete Erhebungsposition)

Das Merkmal gibt das Gewicht des Geschäftstyps der Teilindizes für 10-Steller an, bei denen es sich um solche mit Varianten (Unterpositionen) handelt („COICOP_Typ“ = 12). Das bedeutet, dass nicht die einzelnen Varianten explizit gewichtet werden, sondern der aus den einzelnen Varianten zusammengesetzte 10-Steller. Die Varianten sind über das Merkmal „COICOP“ und die Meldebogennummer „MB“ identifizierbar.

Wie bei „Gewicht_GKat“ wird auch hier eine gewichtete Durchschnittsberechnung auf 10-Steller-Ebene vorgenommen.

Vermietertyp – Nummer des Vermietertyps

Es wird zwischen drei Vermietertypen unterschieden:

1 = Privater Kleinvermieter

2 = Öffentliche Trägerschaft, Wohnungsgenossenschaft

3 = Wohnungsunternehmen

Aufgrund von versehentlichen Eingaben bestehen teilweise auch Nennungen von Vermietertypen in COICOP Bereichen, die nicht der Vermietung zuzuordnen sind.

Erhebungsart – Bezeichnung der Erhebung

Hier wird die Form der Erhebung in Textform genannt. Mögliche Ausprägungen sind:

- „Begehung“,
- „Elektronisch“
- „Papier und elektronisch“.
- „Selbstausfüller“
- „Telefonisch“

MB – Identifizierungsnummer des Meldebogens

Beinhaltet die Identifizierungsnummer des Meldebogens für Preiserfassungen und ist geeignet, den COICOP weiter auszudifferenzieren. Zentral erfasste

Preise, die ausschließlich durch das Statistische Bundesamt erhoben werden, weisen hier einen fehlenden Wert auf[2].

COICOP – Klassifikation der Verwendungszwecke des Individualverbrauchs

Nummer der internationalen Klassifikation der Verwendungszwecke des Individualverbrauchs (COICOP = Classification of individual consumption according to purpose) auf 10-Steller Ebene.

https://www.destatis.de/DE/Methoden/Klassifikationen/Private-Haushalte/sea-2013.pdf?__blob=publicationFile&v=3

Für das Basisjahr 2020 ist die COICOP-Klassifikation von 1999 relevant. Sie bildet auch die Grundlage für die SEA 2013.

https://www.klassifikationsserver.de/klassService/thyme/variant/sea_2013

Für das Basisjahr 2025 findet die COICOP-Klassifikation von 2018 Anwendung, ebenfalls als Grundlage für die SEA 2021.

https://www.klassifikationsserver.de/klassService/thyme/variant/sea_2021

In den Veröffentlichungen der Statistischen Ämter werden einzelne 10-Steller zu einem neuen 10-Steller zusammengefasst (erkennbar anhand von drei Nullen am Ende des 10-Stellers). Diese zusammengefassten 10-Steller sind nicht im FDZ-Datensatz enthalten, deren zugrundeliegenden 10-Steller hingegen schon.

[2] Dies ist nicht gleichzusetzen mit der Bundeslandkennung „99“ aus dem Merkmal „Bundesland“. Vereinzelt erhebt das Statistische Bundesamt für die Bundesländer zentral Preise, die dann auf die einzelnen Bundesländer übertragen werden. Fehlende Meldebogennummern kommen daher auch bei Erhebungspositionen vor, die den Bundesländern zugeordnet sind.

Kurztext – Kurztext der COICOP-Klassifikation

Kurze Beschreibung der Erhebungsposition (10-Steller) in Worten, der das einzelne erhobene Gut zugeordnet ist.

Langtext – Langtext der COICOP-Klassifikation

Umfangreiche Güterbeschreibung der Erhebungsposition (10-Steller) in Worten, der das einzelne erhobene Gut zugeordnet ist. Die Güterbeschreibung ist meist weiter gefasst, um die Auswahl der passenden Güter zu erleichtern.

COICOP_Typ – Typ des 10-Stellers

Dieses Merkmal gibt an, ob es sich bei einem Produkt um das eines klassischen 10-Steller handelt oder um eine Variante (Unterposition) eines 10-Stellers unter dem zwei oder mehrere Varianten zusammengefasst sind:

- 10 = 10-Steller ohne Varianten
- 12 = Variante eines 10-Stellers

Die Gewichte des Wägungsschemas zur Berechnung des VPI liegen auf der untersten Ebene auf dem 10-Steller. Für bestimmte Erhebungspositionen existieren unterhalb dieser 10-Steller-Ebene zwei oder mehr Varianten (sog. Unterpositionen mit jeweils eigenen Meldebogennummern „MB“), die allein zur besseren Strukturierung der Erhebung dienen, z. B. zum Abbilden bestimmter Saisonstrukturen. Im Rahmen der Berechnung werden diese zu einem Elementarindex (je Geschäftskategorie) auf 10-Steller-Ebene zusammengefasst.

Kurztext_Oberposition – Kurztext des 10-Stellers mit Varianten (übergeordnete Erhebungsposition)

Wenn "COICOP_Typ" = 12 für den aktuellen Datensatz gesetzt ist, dann ist hier der Kurztext der übergeordneten Erhebungsposition eingetragen. Ansonsten ist das Feld leer.

Produktvariante – Nummer der Produktvariante

Sofern in derselben Berichtsstelle mehr als eine Variante eines Produkts innerhalb einer COICOP-Klassifikationsnummer erhoben wird, gibt dieses Merkmal die entsprechende Produktvariante als fortlaufende Nummer an.

Merkmal1 bis Merkmal10 – COICOP spezifische Benennung des Merkmals

Die sogenannten Feinbeschreibungsmerkmale enthalten weitere Details zu den Produkten, deren Preise erhoben werden. Sie dienen dazu, während der Preisermittlung sicherzustellen, dass monatlich immer dasselbe Produkt erfasst wird. Merkmal1 bis Merkmal10 benennt, welches Produktdetail analog in Auspraegung1 bis Auspraegung10 dokumentiert werden soll.

Auspraegung1 bis Auspraegung10 – Feinbeschreibungsmerkmale nach COICOP-Klassifikation

Jeder 10-Steller verfügt über bis zu zehn verschiedene Feinbeschreibungsmerkmale, die bei der ersten Preisermittlung von den Preisermittlerinnen und Preisermittlern ausgefüllt und bei Bedarf, z. B. im Rahmen von Produktwechseln, angepasst werden (siehe „Merkmal1“ bis „Merkmal10“).

In der Bereitstellung für die wissenschaftliche Nutzung in Form des SUF werden die Feinbeschreibungsmerkmale vollständig pseudonymisiert, sodass gewährleistet werden kann, dass sie keine sensiblen Informationen enthalten (siehe Anhang).

PreisErhoben – Erhobener Preis des Produkts

Erhobener Preis des jeweiligen Produkts. Da aus technischen Gründen die Eingabe des Wertes Null im Verbundprogramm des Verbraucherpreisindex nur begrenzt zulässig ist, wird von den Statistischen Landesämtern für Waren und Dienstleitungen, die kostenfrei sind, ein Preis von 0,01 € eingetragen. Wurde kein Preis erhoben, liegt ein systemfehlender Wert bzw. in Einzelfällen eine Null oder negativer Wert vor. Die Gründe für eine fehlende Preiseingabe lassen sich den Signierungen in „SigE“ und „SigB“ entnehmen.

Bestimmte Angaben stehen teilweise lediglich als bundes- oder länderspezifisch aggregierte Messzahlen zur Verfügung, die durch das Statistische Bundesamt berechnet werden. Die zugrundeliegenden Einzelpreise stehen zum jetzigen Zeitpunkt nicht zur Verfügung.

Wenn es sich bei einem Einzelwert um eine Messzahl statt eines Einzelpreises handelt, ist in dem Merkmal „Auspraegung1“ das Wort „Messzahl“ enthalten. Dieses ist von der Pseudonymisierung ausgenommen und damit lesbar (siehe auch „Auspraegung1-10“).

PreisBearbeitet – Bearbeiteter Preis des erhobenen Produkts

Der bearbeitete Preis eines Produkts wird auf Basis des erhobenen Preises und der Menge durch das Verbundprogramm berechnet. Veränderungen der

ursprünglich festgelegten Menge fließen in die Berechnung ein, um auch indirekte Preisänderungen zu berücksichtigen und die Vergleichbarkeit der Preise zu erhalten. Der erhobene Preis für die neue Menge des aktuellen Monats wird auf die Basismenge umgerechnet, sodass die implizite Gewichtung der in der Basisperiode festgelegten Stichprobe erhalten bleibt.

Bei Erzeugnis- oder Berichtsstellenausfällen und Zurückweisungen von Ersatzerzeugnissen oder -berichtsstellen (Signierung EV, E4, E5, BV, B3 oder B4) wird der bearbeitete Preis durch Fortschreibung des Normalpreises des Vormonats ermittelt, da kein erhobener Preis vorliegt.

Normalpreis – Normalpreis des erhobenen Produkts

Der Normalpreis spiegelt die Preisentwicklung ohne Sonderangebote wider. Im Normalfall entspricht der Normalpreis dem bearbeiteten Preis des aktuellen Berichtsmonats.

Qualitaetsaenderung – Qualitätsveränderung des Produktes seit dem Vormonat

Qualitätsveränderung des Produktes seit dem Vormonat. Hat sich die Qualität der Ware bei gleichem oder verändertem Preis verringert oder erhöht, ist das Ausmaß der Verteuerung / Verbilligung in Euro angegeben.

Menge – Menge des erhobenen Produkts

Menge des erhobenen Produkts. Die dazugehörige Maßeinheit wird unter „Mass“ angegeben.

Mass – Maßeinheit der Menge des erhobenen Produkts

Einheit, in der die Mengenangabe des erhobenen Produkts erfolgt ist.

Waegungsanteil – Gewicht des jeweiligen 10-Stellers

Das Wägungsschema beschreibt die Anteile der einzelnen Güterarten (10-Steller; siehe Merkmal „COICOP“) an den gesamten Konsumausgaben für Waren und Dienstleistungen privater Haushalte. Unter Verwendung dieses Wägungsschemas werden die Teilindizes für die einzelnen Güterarten zum Gesamtindex aggregiert.

Zur Ableitung der Gewichte auf den höheren Ebenen[3] wurden für die Basis 2020 = 100 nicht wie bisher die Ergebnisse der amtlichen Haushaltsbefragungen verwendet, sondern Ergebnisse privater Konsumausgaben aus den Volkswirtschaftlichen Gesamtrechnungen (VGR). Die Gewichtung auf den unteren Wägungsebenen basiert weiterhin auf Ergebnissen Wirtschaftsrechnungen der privaten Haushalte – der Einkommens- und Verbrauchsstichprobe (EVS) und den Laufenden Wirtschaftsrechnungen (LWR) – ergänzt durch weitere Quellen (z.B. den Steuerstatistiken). Das Wägungsschema ist für alle Bundesländer gleich.

Das Jahr 2020 war aufgrund des von der Corona-Pandemie deutlich beeinflussten Konsumverhaltens ein „untypisches“ Jahr und ist damit als Basisjahr für den Verbraucherpreisindex nicht optimal. Um die Besonderheiten des Jahres 2020 in der Wägung des VPI auszugleichen, wurde für die oberen Wägungsebenen ein Mittelwert der Jahre 2019 (Vor-Pandemie), 2020 (erstes

[3] Als höhere bzw. obere Ebenen werden die COICOP-Ebenen von den 2-Stellern bis hinunter zu einigen 3- und 4-Stellern bezeichnet, als untere Ebenen dann die COICOP-5-Steller bis hin zu den detaillierten10-Stellern.

Pandemiejahr) und 2021 (zweites Pandemiejahr) verwendet. Da aufgrund fehlender Informationen eine Durchschnittsberechnung für die unteren Wägungsebenen nicht möglich war, wurden dafür die Werte aus dem Jahr 2020 verwendet, die aufgrund der EVS 2018 erhoben und anhand der Ergebnisse der LWR 2019 und 2020 fortgeschrieben wurden.

Die Gewichte des Wägungsschemas zur Berechnung des VPI, die sogenannten „Wägungsanteile“, liegen auf der untersten Ebene auf dem 10-Steller. Für bestimmte Erhebungspositionen existieren unterhalb dieser 10-Steller-Ebene zwei oder mehr Varianten (sog. Unterpositionen mit jeweils eigenen Meldebogennummern „MB“), die allein zur besseren Strukturierung der Erhebung dienen, z. B. zum Abbilden bestimmter Saisonstrukturen. Im Rahmen der Berechnung werden diese zu einem Elementarindex (je Geschäftskategorie) auf 10-Steller-Ebene zusammengefasst.

Das Merkmal „Waegungsanteil“ enthält nur für die einzelnen Positionen („COICOP-Typ“ = 10) einen Wägungsanteil, bei Varianten eines 10-Stellers ist das Feld leer. Der Wägungsanteil für einen 10-Steller mit Varianten kann stattdessen dem Merkmal „Waegungsanteil_Oberposition“ entnommen werden.

Für mehr Details zur Bedeutung für den Verbraucherpreisindex siehe Teil I des Metadatenreports Abschnitt 2.4.

Waegungsanteil_Oberposition – Gewicht eines 10-Stellers mit Varianten (übergeordnete Erhebungsposition)

Das Merkmal gibt den Wägungsanteil für 10-Steller an, bei denen es sich um solche mit Varianten (Unterpositionen) handelt („COICOP_Typ“ = 12). Das bedeutet, dass nicht die einzelnen Varianten explizit gewichtet werden, sondern der aus den einzelnen Varianten zusammengesetzte 10-Steller. Die Varianten

sind über das Merkmal „COICOP“ und die Meldebogennummer „MB“ identifizierbar.

SigP – Preissignatur

Bei der Preissignatur wird zwischen folgenden Signierungen unterschieden:

00 = Preis ist gegenüber dem Vormonat unverändert

PA = Preisänderung gegenüber dem Vormonat

PA/PS = Preisänderung gegenüber dem Vormonat und Sonderangebot

PK = Korrigierter Preis

PS = Preis ist Sonderangebot

SigE – Erzeugnissignatur

Bei der Erzeugnissignatur wird zwischen folgenden Signierungen unterschieden:

00 = Erzeugnis gegenüber dem Vormonat unverändert

E1 = Ersetzung durch gleichwertiges Erzeugnis

E2 = Ersatzerzeugnis mit abweichender Qualität

E3 = Ersetzung durch unvergleichbares Erzeugnis

E4 = Auswahl eines Ersatzerzeugnisses nicht möglich

E5 = Ersatzerzeugnis wurde zurückgewiesen

EM = Veränderte Menge bei sonst unverändertem Erzeugnis

EV = Erzeugnis vorübergehend nicht verfügbar

ES = Saisonartikel nicht mehr im Angebot

EN = Aufnahme eines neuen Erzeugnisses

SigB – Berichtsstellensignatur

Bei der Berichtsstellensignatur wird zwischen folgenden Signierungen unterschieden:

00 = Berichtsstelle gegenüber dem Vormonat unverändert

B1 = Ersatzberichtsstelle innerhalb des Geschäftstyps

B2 = Ersatzberichtsstelle außerhalb des Geschäftstyps

B3 = Auswahl einer Ersatzberichtsstelle nicht möglich

B4 = Ersatzberichtsstelle wurde zurückgewiesen

BN = Aufnahme einer neuen Berichtsstelle

BV = Berichtsstelle vorübergehend geschlossen

Durchschnittspreis – Durchschnittspreis des 10-Stellers in der jeweiligen Geschäftskategorie und pro Bundesland (Elementarindexabgrenzung), nicht gerundet

Der ermittelte Durchschnittspreis bezieht sich auf den arithmetisch gemittelten Wert aller bearbeiteten Preise innerhalb einer Elementarindexabgrenzung.

Preissätze mit B2-Signierung fließen in die Berechnung der neuen Elementarindexabgrenzung ein. Sätze mit EN-Signierungen werden nicht berücksichtigt. Für Preissätze mit E4- oder B3-Signierungen wird der fortgeschriebene Normalpreis berücksichtigt.

AbweichungDurchschnitt – Prozentuale Abweichung des bearbeiteten Preises vom Durchschnittspreis

Rev – Revidierter Wert

Der Eintrag bezieht sich auf den jeweiligen Durchschnittspreis eines 10-Stellers pro Geschäftskategorie, nicht auf den einzelnen korrigierten Preis. Erfolgt eine Fehlerkorrektur mit anschließender Neuberechnung in bereits abgeschlossenen Berichtsmonaten werden die gegenüber den ursprünglichen Ergebnissen abweichenden Durchschnittspreise mit einem "r" gekennzeichnet. Bitte beachten: Bei den genannten Korrekturen handelt es sich nicht um solche, die im Rahmen der regulären Revision des Verbraucherpreisindex vorgenommen werden!

DatPreisermittlung – Datum der Preiserhebung

DatSonderangebot – Datum, seitdem das Produkt im Sonderangebot ist

DatNichtVerfuegbar – Datum, seitdem das Produkt nicht verfügbar ist

Ursprüngliches angedachtes Erhebungsdatum, an dem ein Produkt jedoch nicht verfügbar war.

DatPreisUnveraendert – Letzte Änderung des Preises

Datum, seit dem der Preis des erhobenen Produkts unverändert geblieben ist.

Saisongut

Sofern es sich bei einer 10-Steller-Position um ein Saisongut handelt, ist das Merkmal mit 1 codiert. Dies betrifft allerdings nur Produkte der Abteilung 1 und 3 des COICOP Systems.

Als saisonale Güter werden Waren und Dienstleistungen bezeichnet, die nur saisonal verfügbar sind, d. h. die zu bestimmten Zeiten des Jahres nicht oder nur in geringen (vernachlässigbaren) Mengen angeboten werden: frischer Fisch, frisches Obst, frisches Gemüse, Bekleidungsartikel (Sommer- und Winterbekleidung).

Berechnung – Indexrelevanz des Preissatzes im gewählten Berichtsmonat

In der Erhebung werden u. U. nicht nur die Preise erhoben, die direkt in den Verbraucherpreisindex eingehen, sondern auch Waren und Dienstleistungen, die erst zukünftig zur Indexberechnung herangezogen werden sollen. Dies kann z. B. der Fall sein, wenn eine neue Erhebungsmethode getestet wird.

„Ja“ = geht aktuell in die Berechnung ein

„Zukunft“ = geht in Zukunft in die Berechnung ein

„Januar“ = geht ab kommendem Januar in die Berechnung ein

„E4“ = geht ab kommendem Januar nicht mehr in die Berechnung ein

Produktmultiplikator

Multiplikatoren (auch: Vervielfacher) bieten die Möglichkeit, unterhalb der Geschäftstypengewichtung eine Wägungskomponente zu schaffen. Dadurch

kann der Einfluss der einzelnen Preisreihe innerhalb des Durchschnittspreises der Elementarindexabgrenzung variiert und beeinflusst werden.

Regionenvervielfacher

Multiplikatoren (auch: Vervielfacher) bieten die Möglichkeit, unterhalb der Geschäftstypengewichtung eine Wägungskomponente zu schaffen. Dadurch kann der Einfluss der einzelnen Preisreihe innerhalb des Durchschnittspreises der Elementarindexabgrenzung variiert und beeinflusst werden.

2.2 Vergleichbarkeit der Merkmale über die Zeit

Die zeitliche Vergleichbarkeit der einzelnen Positionen über die Berichtsmonate hinweg ist grundsätzlich gegeben. Durch Erzeugnis- und Berichtsstellenwegfälle kann es jedoch zu Erzeugnisänderungen und Berichtsstellenwechseln kommen. Die entsprechenden Signierungen (SigP, SigE, SigB) sind daher zu beachten. Über die Zeit vergleichbar ist letztlich der bearbeitete Preis (Preis-Bearbeitet).

Darüber hinaus wird der Erhebungskatalog alle fünf Jahre angepasst, sodass er das Konsumverhalten der Bevölkerung möglichst repräsentativ abbildet. Für das Berichtsjahr 2025 war dies zuletzt der Dezember 2019. Dies ist bei der Bildung von Zeitreihen zu berücksichtigen.

Das Berichtsjahr 2025 entspricht in der vorliegenden Version dem Warenkorb des Basisjahres 2020.

2.3 Eckwerte relevanter Merkmale und Merkmalskombinationen

	2025
Januar	738.081
Februar	738.578
März	739.064
April	739.173
Mai	748.823
Juni	767.746
Juli	784.167
August	803.202
September	842.530
Oktober	873.930
November	896.212
Dezember	900.465
Gesamt	**9.571.971**

2.4 Auswertbare regionale Ebene

Im SUF steht eine Gliederung bis zur Bundeslandebene zur Verfügung.

3 Praktische Hinweise

3.1 Hinweise zur Geheimhaltung

3.1.1 Gesetzliche Grundlagen der statistischen Geheimhaltung

Unter Geheimhaltung versteht man das Herstellen der absoluten Anonymität der Ergebnisse statistischer Analysen. Konkret bedeutet das, dass im Rahmen der Geheimhaltung sichergestellt wird, dass mit den veröffentlichten Ergebnissen keine Rückschlüsse auf einen Einzelfall (z.B. Person, Betrieb, Einrichtung) gezogen werden können. Statistische Geheimhaltung wird überall dort angewendet, wo statistische Ergebnisse oder Einzeldaten die geschützten Räume der amtlichen Statistik verlassen.

Die Geheimhaltung in der amtlichen Statistik ist in § 16 Bundesstatistikgesetz (BStatG) geregelt und beinhaltet, dass Einzelangaben über persönliche und sachliche Verhältnisse, die für eine Bundesstatistik angegeben werden, von den jeweils durchführenden statistischen Stellen geheim zu halten sind, soweit es keine anderslautenden Bestimmungen gibt. Dies wird auch als Statistikgeheimnis bezeichnet. Das Statistikgeheimnis verpflichtet die amtliche Statistik, die erhaltenen Informationen zu schützen, d.h. sie in einer Form zu anonymisieren, die keine Rückschlüsse mehr auf die betreffende Person und den dargelegten Sachverhalt enthält. Die Geheimhaltung ist auch im Hinblick auf die informationelle Selbstbestimmung von besonderem Interesse: Viele Erhebungen der amtlichen Statistik unterliegen der Auskunftspflicht, somit steht es den Befragten nicht frei, selbst zu entscheiden, ob sie eine Information weitergegeben möchten. Die amtliche Statistik muss deshalb sicherstellen, dass die erhobenen Daten keinem Befragten zugeordnet werden können.

Das BStatG sieht jedoch auch Fälle vor, in denen das Statistikgeheimnis nicht gilt. In § 16 BStatG sind die Ausnahmen von der Geheimhaltungspflicht dargelegt. Unter anderem wird dort festgelegt, unter welchen Umständen die Daten der amtlichen Statistik für die Wissenschaft zugänglich gemacht werden dürfen und welche Regeln dabei einzuhalten sind.

Die Ausführungen in den folgenden Abschnitten 3.1.2 und 3.1.3 beziehen sich auf Ergebnisse, die am Gastwissenschaftlerarbeitsplatz oder per Kontrollierter Datenfernverarbeitung erzeugt wurden (On-Site-Nutzung). Dort werden lediglich formal-anonyme Einzeldaten ausgewertet, so dass die Ergebnisse vor Herausgabe einer sorgfältigen Geheimhaltungsprüfung unterzogen werden müssen.

Die Einzeldaten eines Scientific Use Files (SUF) sind hingegen stärker anonymisiert (faktisch-anonym). Eine Geheimhaltungsprüfung der Ergebnisse durch das FDZ-Personal erfolgt hier nicht. Die Nutzerinnen und Nutzer sind allerdings ihrerseits durch entsprechende Regelungen in den Nutzungsverträgen zur Einhaltung der statistischen Geheimhaltung verpflichtet.

3.1.2 Geheimhaltung von Ergebnissen

Um die gesetzlich vorgeschriebene Geheimhaltung von Einzelfällen in den Daten sicherzustellen, müssen alle Ergebnisse, die am Gastwissenschaftlerarbeitsplatz oder per Kontrollierter Datenfernverarbeitung erzeugt werden, vor ihrer Freigabe an den Nutzer von den FDZ einer Geheimhaltungsprüfung unterzogen werden. Dabei stellen die FDZ sicher, dass die Ergebnisse absolut anonym sind und eine Reidentifikation einzelner Befragter nach menschlichem Ermessen ausgeschlossen werden kann. Entsprechend handeln auch die Fachabteilungen der Statistischen Ämter vor der Veröffentlichung von Ergebnissen.

Zur Sicherstellung der Geheimhaltung wenden die FDZ verschiedene Geheimhaltungsregeln an, die jeweils individuell auf die jeweilige Statistik zugeschnitten sind. In der Broschüre „Regelungen zur Auswertung von Mikrodaten in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder“ werden die gebräuchlichsten Regeln zur primären Geheimhaltung dargestellt. Diese Regeln werden in den FDZ im Grunde auf alle Statistiken angewendet. Die Anlage dieser Broschüre enthält Informationen darüber, welche Geheimhaltungsregeln auf welche Statistiken anzuwenden sind.

Die Broschüre finden Sie hier:

https://www.forschungsdatenzentrum.de/de/geheimhaltung.

3.1.3 Praktische Tipps zur Vermeidung von Geheimhaltungsfällen

Treten in den erstellten Analysen Geheimhaltungsfälle auf, werden diese Werte von den FDZ zur Sicherstellung der Geheimhaltung durch ein Sperrmuster ersetzt. Gerade in Kreuztabellen entstehen so durch die notwendige Sekundärsperrung schnell viele „Löcher“ in den Auswertungen. Da eine einmal zur Sekundärsperrung herangezogene Tabellenzelle auch in allen folgenden Analysen gesperrt werden muss (tabellenübergreifende Geheimhaltung) – auch, wenn es in der neu erstellten Tabelle nicht nötig wäre – ist es sinnvoll, bei jeder Ergebniserstellung darauf zu achten, dass möglichst keine Geheimhaltungsfälle erzeugt werden. Treten in einem Output Geheimhaltungsfälle auf, steht es dem betreuenden FDZ frei, die Prüfung und Freigabe des Outputs abzulehnen.

Um Geheimhaltungsfälle in den Analysen zu vermeiden, sollte immer darauf geachtet werden, dass die erstellten Analysen auf ausreichend großen Fallzahlen beruhen. Bei geringen Fallzahlen empfiehlt es sich, Variablenausprägungen zusammen zu fassen und damit größere Fallzahlen zu erzielen.

3.2 FAQ

Bitte wenden Sie sich bei auftretenden Fragen an den im Impressum für fachliche Informationen genannten FDZ-Standort.

3.3 Verfügbare Tools

Für dieses Produkt werden seitens der Forschungsdatenzentren keine weiterführenden Tools angeboten.

4 Literaturverzeichnis

Kaukal, Malte und Peters, Normen, 2019: Vom Wort zur Zahl: Wie mit Hilfe automatisierter Verfahren Produktbeschreibungen in der Verbraucherpreisstatistik für das Forschungsdatenzentrum effizient bereitgestellt werden können – Ein Werkstattbericht. Hessisches Statistisches Landesamt.

Anhang 1 – Merkmalsübersicht

		Basisjahr 2005	Basisjahr 2010	Basisjahr 2015	Basisjahr 2020
Merkmal	**Bezeichnung**	**2005-2009**	**2010-2014**	**2015-2019**	**2020-2025**
ID	ID einer Erhebungsposition im Basisjahr	+	+	+	+
Monat	Erhebungsmonat	+	+	+	+
Jahr	Erhebungsjahr	+	+	+	+
Bundesland	Erhebungsbundesland	+	+	+	+
Gewicht_BL	Bundeslandgewichtung	+	+	+	+
Berichtsstelle	Identifizierungsnummer der Berichtsstelle	+	+	+	+
Berichtsstellen-multiplikator	Anzahl Filialen	+	+	+	+
GKat	Geschäftskategorie	+	+	+	+
Gewicht_GKat	Gewicht der Geschäftstypen	+	+	+	+
Gewicht_GKat_Oberposition	Gewicht der Geschäftstypen eines 10-Stellers mit Varianten (übergeordnete Erhebungsposition)	+	+	+	+
Vermietertyp	Nummer des Vermietertyps		+	+	+
Erhebungsart	Bezeichnung der Erhebung	+	+	+	+
MB	Identifizierungsnummer des Meldebogens	+	+	+	+
COICOP	Klassifikation der Verwendungszwecke des Individualverbrauchs	+	+	+	+
Kurztext	Kurztext der COICOP Klassifikation	+	+	+	+
Langtext	Langtext der COICOP Klassifikation	+	+	+	+
COICOP_Typ	Typ des 10-Stellers	+	+	+	+
Kurztext_Oberposition	Kurztext des 10-Stellers mit Varianten (übergeordnete Erhebungsposition)	+	+	+	+
Produktvariante	Nummer der Produktvariante	+	+	+	+
Merkmal1	COICOP spezifische Benennung des Merkmals	+	+	+	+
Auspraegung1	Feinbeschreibungsmerkmal	+	+	+	+
Merkmal2	COICOP spezifische Benennung des Merkmals	+	+	+	+
Auspraegung2	Feinbeschreibungsmerkmal	+	+	+	+
Merkmal3	COICOP spezifische Benennung des Merkmals	+	+	+	+
Auspraegung3	Feinbeschreibungsmerkmal	+	+	+	+
Merkmal4	COICOP spezifische Benennung des Merkmals	+	+	+	+

		Basisjahr 2005	Basisjahr 2010	Basisjahr 2015	Basisjahr 2020
Merkmal	**Bezeichnung**	**2005-2009**	**2010-2014**	**2015-2019**	**2020-2025**
Auspraegung4	Feinbeschreibungsmerkmal	+	+	+	+
Merkmal5	COICOP spezifische Benennung des Merkmals	+	+	+	+
Auspraegung5	Feinbeschreibungsmerkmal	+	+	+	+
Merkmal6	COICOP spezifische Benennung des Merkmals	+	+	+	+
Auspraegung6	Feinbeschreibungsmerkmal	+	+	+	+
Merkmal7	COICOP spezifische Benennung des Merkmals	+	+	+	+
Auspraegung7	Feinbeschreibungsmerkmal	+	+	+	+
Merkmal8	COICOP spezifische Benennung des Merkmals	+	+	+	+
Auspraegung8	Feinbeschreibungsmerkmal	+	+	+	+
Merkmal9	COICOP spezifische Benennung des Merkmals	+	+	+	+
Auspraegung9	Feinbeschreibungsmerkmal	+	+	+	+
Merkmal10	COICOP spezifische Benennung des Merkmals	+	+	+	+
Auspraegung10	Feinbeschreibungsmerkmal	+	+	+	+
PreisErhoben	Erhobener Preis des Produkts	+	+	+	+
PreisBearbeitet	Bearbeiteter Preis des erhobenen Produkts	+	+	+	+
Normalpreis	Normalpreis des erhobenen Produkts	+	+	+	+
Qualitaetsaenderung	Qualitätsveränderung des Produktes seit dem Vormonat	+	+	+	+
Menge	Menge des erhobenen Produkts	+	+	+	+
Mass	referenzierte Kurzbezeichnung der Maßeinheit	+	+	+	+
Waegungsanteil	Gewicht des jeweiligen 10-Stellers	+	+	+	+
Waegungsanteil_Oberposition	Gewicht eines 10-Stellers mit Varianten (übergeordnete Erhebungsposition)	+	+	+	+
SigP	Preissignatur	+	+	+	+
SigE	Erzeugnissignatur	+	+	+	+
SigB	Berichtsstellensignatur	+	+	+	+
Durchschnittspreis	Durchschnittspreis des 10-Stellers in der jeweiligen Geschäftskategorie und pro Bundesland (Elementarindexabgrenzung), nicht gerundet	+	+	+	+
Abweichung-Durchschnitt	Prozentuale Abweichung des bearbeiteten Preises vom Durchschnittspreis	+	+	+	+

		Basisjahr 2005	Basisjahr 2010	Basisjahr 2015	Basisjahr 2020
Merkmal	**Bezeichnung**	**2005-2009**	**2010-2014**	**2015-2019**	**2020-2025**
Rev	Kennzeichnung revidierter Werte mit „r“	+	+	+	+
DatPreisermitt-lung	Datum der Preiserhebung	+	+	+	+
DatSonderange-bot	Datum, seitdem das Produkt im Sonderangebot ist	+	+	+	+
DatNichtVerfueg-bar	Datum, seitdem das Produkt nicht verfügbar ist	+	+	+	+
DatPreisUnvera-endert	Letzte Änderung des Preises	+	+	+	+
Saisongut		+	+	+	+
Berechnung	Indexrelevanz des Preissatzes im gewählten Berichtsmonat	+	+	+	+
Produktmultiplika-tor		+	+	+	+
Regionenverviel-facher		+	+	+	+

Anhang 2 – Anonymisierungskonzept der Feinbeschreibungsmerkmale (Auspraegung1 bis Auspraegung10)

Die Feinbeschreibungsmerkmale stellen unstrukturierte Daten dar, die vereinzelt und unsystematisch sensible Daten enthalten können. Daher werden die u. U. vereinheitlichten Feinbeschreibungsmerkmale grundsätzlich vollständig pseudonymisiert. Die Pseudonyme werden über die zehn Feinbeschreibungsmerkmale und die beantragten Berichtsjahre konsistent vergeben. Sollten spezifische Begriffe für das Forschungsprojekt unerlässlich sein, besteht die Möglichkeit bei einer On-Site-Nutzung (Zugang über einen Gastwissenschaftsarbeitsplatz (GWAP) oder die Kontrollierte Datenfernverarbeitung (KDFV)) eine Liste von Begriffen zu definieren, die von der Pseudonymisierung ausgenommen werden. Die Liste wird durch die Mitarbeiter des FDZ Standort Hessen geprüft und freigegeben, sofern sie keine identifizierenden Inhalte aufweist. Die Klartexte der übrigen Feinbeschreibungsmerkmale werden durch die pseudonymisierten Begriffe ersetzt.

Für die Nutzung des SUF besteht die Möglichkeit der Verwendung einer Whitelist nicht.

In Hinblick auf die vorzunehmende Anonymisierung durch Pseudonymisierung der Feinbeschreibungsmerkmale, soll durch eine Vereinheitlichung der Schreibweisen vermieden werden, dass unterschiedliche Schreibweisen zu mehreren Pseudonymen führen und der Nutzer fälschlicherweise von unterschiedlichen Feinbeschreibungen ausgehen muss.

Um zu entscheiden, welche Texteinträge zu einem gemeinsamen Eintrag vereinheitlicht werden sollen, beruht der Entscheidungsprozess auf einer Kombination aus deterministischen Entscheidungsregeln und der Anwendung maschinellen Lernens.

Die Entscheidungsregeln geben vor, welche Begriffe sich so sehr ähneln, dass sie potenziell zusammengefügt werden können. Sie wurden selbst aufgestellt, um unterschiedliche Begriffe mit fälschlicherweise hohen Ähnlichkeiten auszuschließen (Regel eins bis drei) bzw. sie wurden aus der Evaluation erster Probeläufe abgeleitet (Regel vier und fünf). Dies ist notwendig, um die grundsätzliche Wahrscheinlichkeit einer falschen Vereinheitlichung gering zu halten:

1. Begriffe werden nur innerhalb einer Produktkategorie (10-Steller) vereinheitlicht.
2. Die Begriffe müssen über denselben Anfangsbuchstaben verfügen.
3. Das Ähnlichkeitsmaß muss mindestens 90 Prozent betragen.
4. Texteinträge, in denen eine Zeichenfolge (bis zu drei Zeichen, bestehend aus Buchstaben, Ziffern oder Satzzeichen) durch Leerzeichen losgelöst vom übrigen Text steht, werden ignoriert (z. B. „Pflegestufe 1“).
5. Texteinträge, in denen zwei oder mehr Ziffernfolgen durch einen Punkt getrennt sind, werden ignoriert (z. B. „Version 1.1“, „Art-Nr. 1515.4546.5465.“ „04.07.2012“).

Die Arbeitsbelastung durch eine Prüfung jeder Vereinheitlichung soll möglichst geringgehalten werden und nur die Vereinheitlichungen geprüft werden, die vorher durch einen Algorithmus als falsch prognostiziert wurden. Zur Prognose mit möglichst hoher Genauigkeit wird hierfür ein Verfahren des vollüberwachten Maschinellen Lernens verwendet. Die als falsch prognostizierten Vereinheitlichungen werden geprüft und ggf. korrigiert.

Anhang 3 – Umschlüsselungstabelle der SEA 2013 auf SEA 2021

COICOP SEA 2013	COICOP SEA 2021
0111203100	0111205110
0112200200	0112220110
0112200300	0112220120
0112200500	0112220130
0113100100	0113101110
0113400100	0113402110
0114110100	0114100110
0114210100	0114200110
0114300100	0114301110
0114507100	0114506110
0114507200	0114506120
0114606100	0114605110
0115500100	0115900110
0116120100	0116102110
0116170200	0116109110
0117111100	0117101110
0119930100	0119905110
0119940100	0119904110
0119990200	0119907110
0220100100	0230100110
0220200100	0230200110
0220301100	0230902110
0220301200	0230904110
0312110300	0312116110
0312110400	0312117110
0312121100	0312122110
0312125200	0312128200
0312126100	0312129110
0312150100	0312155110
0312150200	0312155120
0312163200	0312162110
0312210200	0312216110
0312227200	0312223110
0312260100	0312265110
0312260200	0312265120
0312351300	0312308110
0312352100	0312303110
0313111200	0313101110
0313115100	0313104110
0313120200	0313107110
0314210200	0314200110
0421011200	0421000110
0421014300	0421000120

COICOP SEA 2013	COICOP SEA 2021
0421021400	0421000130
0421024300	0421000140
0421024600	0421000150
0441000100	0441100110
0441000300	0441100120
0441000400	0441100130
0443000200	0443100110
0443000300	0443100120
0444109100	0444108110
0451010200	0451015110
0451010300	0451015120
0451010400	0451015130
0451010500	0451015140
0453001100	0453002110
0511141100	0511140110
0532100100	0532102110
0540101100	0540102110
0540101200	0540102120
0540323300	0540322110
0552102200	0552104120
0552102600	0552104130
0552103100	0552105110
0552103500	0552105120
0552201200	0552202110
0552209300	0552208110
0561101100	0561111110
0561103100	0561113110
0561105100	0561115110
0561105200	0561115120
0561105300	0561115130
0561107100	0561117110
0562900200	0562909110
0630010100	0631001110
0630020300	0631002110
0712003100	0712007110
0712004100	0712007120
0712005100	0712008110
0713000100	0713003110
0721211100	0721201110
0721212200	0721202110
0721220100	0721208110
0721311100	0721301110
0721319100	0721305110
0721320100	0721308110
0723014100	0723015110
0724101200	0724102110

COICOP SEA 2013	COICOP SEA 2021
0724201100	0724200110
0724329100	0724328110
0921402100	0912402110
1111201100	1111203110
1111201200	1111203120
1111201300	1111203130
1111201400	1111203140
1111201500	1111202110
1111201600	1111202120
1111203100	1111205110
1111203300	1111208110
1111203400	1111205120
1111203500	1111204110
1111203700	1111204120
1111205100	1111207110
1120110100	1120101110
1120120100	1120102110
1120120200	1120102120
1120201100	1120202110
1120201300	1120202120
1120203100	1120204110
1120309100	1120300110
1262100100	1222000110

Anhang 4 – Umschlüsselungstabelle mit der Zuordnung der „gelöschten“ 10-Steller des Erhebungskatalogs 2020 (Basis 2020 = 100) zu den entsprechenden Fortschreibungspositionen des Erhebungskatalogs 2025

COICOP (BJ2020)	Fortschreibungsposition
0111109100	0119110110
0111320300	0111320500
0111410200	0111320110
0111423100	0111343110
0111424300	0111344110
0111431200	0111351110
0111433100	0111352110
0111442100	0111361110
0111444200	0111362110
0111450100	0111370110
0111500100	0119140110
0111610100	0111500110
0111621200	0119110110
0111701100	0111403110
0111703100	0111403110
0111801100	0111901110
0111809100	0111908110
0112101100	0112211110
0112107100	0112217110
0112107200	0112217120
0112107300	0112217130
0112109100	0112219110
0112200100	0112220130
0112300100	0112230110
0112410100	0112240110
0112410200	0112240120
0112500100	0112250110
0112600100	0112400110
0112710200	0112310110
0112710300	0112310120
0112721100	0112502110
0112721200	0112502120
0112721300	0112502130
0112722100	0112503100
0112723100	0112504110
0112801100	0112505110
0112805100	0112506110

COICOP (BJ2020)	Fortschreibungsposition
0112806100	0112507110
0112807200	0112508110
0112808100	0112509110
0112808200	0112509120
0113100200	0113101110
0113100300	0113101110
0113200100	0113102110
0113300200	0113401110
0113500100	0113200110
0113601200	0113301110
0113602100	0113301110
0113602200	0113302110
0113603100	0113303110
0113608200	0113303110
0114400200	0114601110
0114601100	0114302110
0114604100	0114704110
0114609100	0114400110
0114701100	0114801110
0115100100	0115200110
0115201100	0115300110
0115300100	0115105110
0115400100	0115109110
0116111100	0116201110
0116112100	0116202110
0116115100	0116202110
0116117100	0116207110
0116130200	0116301110
0116140100	0116302110
0116150200	0116109110
0116161100	0116401110
0116165100	0116501110
0116170300	0116502110
0116200200	0116600110
0116301100	0116700110
0116303100	0116901110
0116309100	0116909110
0116401100	0116904110
0116402100	0116905110
0116403100	0116906110
0117119200	0117409110
0117131100	0117201110

COICOP (BJ2020)	Fortschreibungsposition
0117133100	0117203110
0117134100	0117204110
0117139200	0117209110
0117139300	0117209110
0117141100	0117401110
0117142100	0117402110
0117145100	0117403110
0117149200	0117300110
0117201100	0117809110
0117209100	0117929120
0117310200	0117600110
0117321100	0117911110
0117323100	0117913110
0117324100	0117914110
0117325100	0117914110
0117328400	0117919110
0117329100	0117928110
0117329200	0117929110
0117401300	0117501110
0117402100	0117801110
0117403100	0117926110
0117500200	0117925110
0118201100	0118302110
0118203100	0118306110
0118205100	0118505110
0118301100	0118503110
0118309100	0118509110
0118401100	0118502110
0118405100	0118903110
0118409100	0118909110
0118500100	0118600110
0118601100	0118200110
0119101100	0119302110
0119102100	0119302110
0119103200	0119304110
0119103300	0119304120
0119103400	0119304130
0119104100	0119305110
0119201100	0119301110
0119202100	0119401110
0119203100	0119402110
0119302100	0119202110

COICOP (BJ2020)	Fortschreibungsposition
0119303100	0119203110
0119405100	0119124110
0119406100	0119125110
0119407100	0119150110
0119911100	0119130110
0119913100	0119130110
0119990300	0119905110
0119990400	0119907110
0121110200	0122001110
0121110300	0122001120
0121121100	0122002110
0121201100	0123001110
0121203100	0123003110
0121300100	0118501110
0122100100	0125000110
0122100200	0125000120
0122211100	0126001110
0122219100	0126002110
0122230100	0129000110
0122311100	0121001110
0122312200	0121002110
0122315100	0121003110
0122320300	0121003110
0211110100	0211010110
0211120100	0211020110
0211130100	0211020110
0211140100	0211040110
0211200100	0212165110
0212300100	0212150110
0212400100	0212165110
0213100100	0213001110
0213200100	0213001120
0213300100	0213003110
0213400100	0219010110
0312231300	0312232110
0312311300	0312182110
0312311400	0312292110
0312312400	0312183110
0312312500	0312293110
0312314500	0312185110
0312314600	0312295110
0312323200	0312294110

COICOP (BJ2020)	Fortschreibungsposition
0312330300	0312181110
0312330400	0312291110
0312341300	0312188110
0312341400	0312298110
0312343200	0312188120
0312343300	0312298120
0312347400	0312188130
0312347500	0312298130
0313114100	0721303110
0313209100	0313201100
0314110300	0314110400
0431001100	0431101110
0431001200	0431101120
0431002100	0431102110
0431002500	0431102120
0431006100	0431103110
0431007100	0431104110
0431009100	0431108110
0431009400	0431108120
0432100400	0432001110
0432200200	0432002110
0432300100	0432003110
0432400100	0432004110
0432500100	0432005110
0432900100	0432009110
0454900100	0454200110
0511113300	0511113400
0511114200	0511114300
0511125200	0511125300
0511901100	0511150110
0511903200	0511160110
0511904100	0511170110
0512111200	0511405110
0512120100	0431105110
0512200200	0431106110
0512300100	0432006110
0513000200	0512000110
0520101100	0521101110
0520101200	0521101120
0520201200	0521201110
0520202100	0521202110
0520203100	0521205110

COICOP (BJ2020)	Fortschreibungsposition
0520205100	0521204110
0520205200	0521204110
0520301100	0521301110
0520302200	0521302110
0520901100	0521901110
0520902200	0521902110
0531202100	0531104110
0531303100	0531107110
0531304100	0531108110
0531420100	0531306110
0531431100	0531109110
0531500100	0531400110
0532300200	0532901110
0532400100	0532101110
0532900200	0532102120
0540321200	0532102130
0540324200	0532902110
0551101100	0551001110
0551102100	0551002120
0551102200	0551002130
0551200100	0553011110
0552102400	0552104120
0552205100	0431108140
0552205500	0431200110
0552209200	0431108130
0561101200	0561111110
0561101300	0561111110
0561211100	0561902110
0561211200	0561902120
0561241100	0561903110
0561243100	0561905110
0561249200	0561909110
0561291100	0561121110
0561292100	0561122110
0562200100	0562901110
0611220100	0611120110
0611390100	0611220110
0612100100	0612200110
0612903100	0612303110
0612905100	0612102110
0613300100	0622901130
0613300200	0622904130

COICOP (BJ2020)	Fortschreibungsposition
0613906100	0612102120
0613909100	0613306110
0621105100	0623105110
0621205100	0623105110
0621205200	0621905110
0622014100	0622901110
0622014200	0622901120
0622015100	0622904110
0622015200	0622904120
0622016100	0622905110
0622019100	0622906110
0623102100	0641002120
0623201100	0623101110
0623203100	0623105120
0623901100	0623201110
0623902100	0623105130
0623903100	0623203110
0712001100	0713001110
0721216100	0721307110
0724105100	0724400110
0724105200	0724400120
0724205100	0724105110
0724205200	0724105120
0724329200	0724328110
0734100100	0734001110
0734200100	0734002110
0736200100	0749100110
0810100200	0741100110
0810900100	0741200110
0820100300	0811001110
0820200200	0812000110
0820300100	0819100110
0820400100	0835001110
0830100200	0831000110
0830200100	0832000110
0830300100	0833000110
0830400100	0834000110
0911112100	0814011110
0911114100	0814011110
0911141100	0814015110
0911143100	0814016110
0911150100	0814018110

COICOP (BJ2020)	Fortschreibungsposition
0911210200	0814021110
0911220100	0814022110
0911220200	0814022110
0911240200	0814015110
0911250100	0814025110
0911300100	0814091110
0911901100	0814091110
0911909100	0819100120
0912113100	0911100110
0912120100	0911100110
0912210300	0911201110
0912310200	0911300110
0913110100	0813101110
0913120100	0813103110
0913221300	0813201110
0913229300	0813204110
0913230400	0813203110
0913309100	0820000110
0914110100	0952001110
0914110200	0971909110
0914120100	0952002110
0914140100	0952004110
0914150100	0839204110
0914210100	0815002110
0914220100	0815002110
0914910100	0815004110
0914920100	0815004110
0914990100	0911207110
0915000100	0835001120
0921101100	0912101110
0921102100	0912101120
0921103100	0912103110
0921302100	0912303110
0921500100	0912302110
0922111100	0951002110
0922115100	0951003110
0922116100	0951004110
0922117100	0951005110
0922120100	0951002120
0923001100	0945001110
0931101100	0921201110
0931102100	0921100110

COICOP (BJ2020)	Fortschreibungsposition
0931102200	0921100120
0931102300	0921100130
0931201100	0921202110
0931202100	0921203110
0931203100	0921204110
0931204100	0921205110
0931204200	0921205110
0931206100	0921206110
0931206200	0921206120
0931207200	0921303110
0932111100	0922111110
0932112100	0922112110
0932113100	0922113110
0932118100	0922117110
0932118200	0922117120
0932119100	0922119110
0932124100	0922124110
0932201100	0922201110
0932202200	0922202110
0932204100	0922204110
0932205100	0922205110
0932205200	0922205120
0933101100	0931101110
0933102200	0931102110
0933103100	0931103110
0933200100	0931212110
0933200200	0931226110
0933229100	0749200110
0934100100	0932100110
0934201200	0932201110
0934201400	0932201120
0934202100	0932202110
0934208100	0932204110
0934209100	0932209110
0935001100	0945002110
0941101100	0946300110
0941102100	0946101110
0941210100	0946201110
0941220100	0946202110
0941251100	0946205110
0941259100	0946207110
0941261100	0969000110

COICOP (BJ2020)	Fortschreibungsposition
0942101100	0961001110
0942102100	0961002110
0942102200	0961002120
0942103100	0961003110
0942103200	0961003120
0942200100	0962000110
0942310200	0839100110
0942321100	0839201110
0942323100	0839203110
0942500200	0963000110
0942500300	0963000120
0951101210	0971901110
0951101220	0971901110
0951102410	0971902110
0951102420	0971902110
0951200100	0971100110
0951301310	0971903110
0951301320	0971903110
0951304100	0971906110
0951309100	0971907110
0951402100	0971908110
0952111100	0972101110
0952111200	0972101110
0952112100	0972101110
0952121100	0972104110
0952121200	0972104110
0952121300	0972107110
0952121400	0972107110
0952122100	0972104110
0952122200	0972107110
0952201100	0972201110
0952201200	0972201110
0952201300	0972201120
0952201400	0972201120
0953001100	0973001110
0953002100	0973002110
0953003100	0973003110
0954101100	0974001110
0954105100	0974002110
0954109100	0974003110
0954921100	0974004110
0954923100	0974005110

COICOP (BJ2020)	Fortschreibungsposition
0954925100	0974006110
0954925200	0974006120
0954925300	0813207110
0954941100	0974008110
1050010200	1050900110
1050020100	1050100110
1111111100	1111111200
1111112100	1111112200
1111113100	1111113200
1111114100	1111114200
1111116100	1111116200
1111120100	1111120200
1111141100	1111141200
1111143100	1111143200
1111143300	1111143400
1111144100	1111144200
1111151100	1111151200
1111152100	1111152200
1111153100	1111153200
1111154100	1111154200
1111203200	1111205110
1112000100	1112900110
1112000200	1112100110
1112000300	1112100120
1211101100	1313111110
1211102100	1313112110
1211201100	1313121110
1211202100	1313129110
1211203100	1313129110
1211301100	1313201110
1211303100	1313204110
1211304100	1313204110
1212101100	1311101110
1212102100	1311102110
1212103100	1311103110
1213102100	1312102110
1213105100	1312105110
1213105200	1312105120
1213105300	1312101110
1213211100	1312211110
1213212100	1312212110
1213212200	1312212110

COICOP (BJ2020)	Fortschreibungsposition
1213213100	1312213110
1213214100	1312214110
1213215100	1312216110
1213216100	1312216110
1213217100	1312217110
1213221100	1312221110
1213222100	1312222110
1213223200	1312223110
1213229100	1312229110
1213229300	0613306120
1213231100	1312231110
1213232100	1312232110
1213232200	1312240120
1213232300	1312232120
1213232400	1312232130
1213233100	1312234110
1213233200	1312234110
1213233300	1312234120
1213240100	1312240110
1231101100	1321111110
1231102100	1321112110
1231201100	1321120110
1231209100	0511408110
1231300100	1321200110
1232101100	1329111110
1232102100	1329112110
1232103100	1329113110
1232105100	1329115110
1232106100	1329116110
1232109100	1329119110
1232201100	1329120110
1232202100	0721302110
1232901100	1329131110
1232902100	1329132110
1232902200	1329132110
1232904100	1329139110
1232905100	1329139110
1232907200	1329137110
1240103100	1330102110
1240201100	1330200110
1240203100	0632002110
1240203200	0632001110

COICOP (BJ2020)	Fortschreibungsposition
1240300100	0623203120
1240300200	0623201120
1240300300	1330300110
1252000100	1213000110
1253200100	1212002110
1253200200	1211003110
1254100100	1214100110
1254100200	1214100120
1254200100	1212002120
1255000100	1219000110
1255000200	1219000130
1262203100	1390904110
1270101100	1390901110
1270102100	1390902110
1270200100	1390903110
1270300100	1390905110
1270402100	1390906110
1270409100	1390909110

Statistische Ämter des Bundes und der Länder,
Metadatenreport – Teil II: Produktspezifische Informationen zur Nutzung der Einzeldaten des Verbraucherpreisindex 2025 (EVAS-Nummer: 61111) per Scientific Use File

Fotorechte Umschlag: ©artSILENCEcom –Fotolia.com