

Потребительские цены: объединенные панельные данные по средним ценам и индексам цен за 2000-2022 гг.

Лицензия: CC BY-SA

Доступ: Открытый

Дата обновления: 27.06.2022

Теги:

экономика, цены

Описание

Набор содержит сведения о средних ценах и индексах средних цен по наблюдаемым Росстатом группам товаров, в Российской Федерации и субъектах. Набор состоит из двух основных таблиц: `prices_and_indexes_2000-2022.csv` (10318896 наблюдение по 16 атрибутам за 2000-2022 гг.); `structure_2012-2022.csv` (548663 наблюдения по 5 атрибутам за 2012-2022 гг.). В дополнение к ним используется четыре вспомогательных справочника: `okato.csv` (430 значений по 3 атрибутам); `kipc_prices.csv` (1461 значения по 2 атрибутам); `kipc_structure.csv` (868 значения по 2 атрибутам); `type.csv` (16 значений по 2 атрибутам). Подробная [методология сбора данных о потребительских ценах](#) приводится Росстатом, в частности цены фиксируются у поставщиков в магазинах. Примеры использования набора данных со средними ценами и структурой расходов представлены в [wiki](#)

Характеристики датасета

Использование

Лицензия: [CC BY-SA](#)

Доступ: Открытый

Предоставление данных

Формат:

CSV

Характеристики набора

Временной период: 2000–2022гг.

Количество наблюдений: 10867559

Обновления

Размещено: 27.06.2022

Обновлено: 27.06.2022

Частота обновления: не указано

Дата следующего обновления: не указано

Происхождение

Источники данных: Федеральная служба государственной статистики

Ответственные за набор данных: Тихонов Сергей Владимирович

Структура набора данных

Потребительские цены в 2000-2022 гг. (prices_and_indexes_2000-2022)

| Атрибут | Описание | Количество пропусков (NaN) (шт/%) | Единица измерения | ОКЕИ | Формат |
|---------|-----------------------|-----------------------------------|-------------------|------|---------|
| DEC | Значения за декабрь | 222597 | | | numeric |
| NOV | Значения за ноябрь | 232388 | | | numeric |
| OCT | Значения за октябрь | 228488 | | | numeric |
| SEP | Значения за сентябрь | 230514 | | | numeric |
| AUG | Значения за август | 218575 | | | numeric |
| JUL | Значения за июль | 211067 | | | numeric |
| JUN | Значения за июнь | 216114 | | | numeric |
| MAY | Значения за май | 218477 | | | numeric |
| APR | Значения за апрель | 218549 | | | numeric |
| MAR | Значения за март | 166077 | | | numeric |
| FEB | Значения за февраль | 16656 | | | numeric |
| JAN | Значения за январь | 21608 | | | numeric |
| kipc | Код товара или услуги | 0 | | | integer |
| okato | Код ОКАТО | 0 | | | integer |
| type | Тип показателя | 0 | | | integer |
| year | Год измерения | 0 | | | integer |

Структура потребительских расходов в 2012-2022 гг. (structure_2012-2022)

| Атрибут | Описание | Количество пропусков (NaN) (шт/%) | Единица измерения | ОКЕИ | Формат |
|---------|-----------------------|-----------------------------------|-------------------|------|---------|
| value | Значение | 92 | | | numeric |
| type | Тип показателя | 0 | | | integer |
| kipc | Код товара или услуги | 0 | | | integer |
| okato | Код ОКАТО | 0 | | | integer |
| year | Год измерения | 0 | | | integer |

Расшифровка типов товаров и услуг для наборов данных о потребительских ценах (kipc_prices)

| Атрибут | Описание | Количество пропусков (NaN) (шт/%) | Единица измерения | ОКЕИ | Формат |
|-------------|-----------------------|-----------------------------------|-------------------|------|---------|
| description | Расшифровка кода | 0 | | | text |
| kipc | Код товара или услуги | 0 | | | integer |

Расшифровка типов товаров и услуг для наборов данных о структуре расходов населения (kipc_structure)

| Атрибут | Описание | Количество пропусков (NaN) (шт/%) | Единица измерения | ОКЕИ | Формат |
|-------------|-----------------------|-----------------------------------|-------------------|------|---------|
| description | Расшифровка кода | 0 | | | text |
| kipc | Код товара или услуги | 0 | | | integer |

Расшифровка кодов ОКАТО (okato)

| Атрибут | Описание | Количество пропусков (NaN) (шт/%) | Единица измерения | ОКЕИ | Формат |
|-------------|--|-----------------------------------|-------------------|------|---------|
| okato_level | Уровень ОКАТО: 0 - Россия целиком, 1 - группы регионов, 2 - регионы, 3 - внутри регионов | 0 | | | integer |
| okato_name | Расшифровка кода | 0 | | | text |
| okato | Код ОКАТО | 0 | | | integer |

Расшифровка типов цен, ИПЦ и групп населения (types)

| Атрибут | Описание | Количество пропусков (NaN) (шт/%) | Единица измерения | ОКЕИ | Формат |
|-----------|------------------------|-----------------------------------|-------------------|------|---------|
| type_name | Расшифровка показателя | 0 | | | text |
| type | Тип показателя | 0 | | | integer |

Структура данных

Датасет состоит из двух ключевых наборов данных: `prices_and_indexes_2000-2022` и `structure_2012-2022`, а также четырех словарей: `okato`, `kipc_prices`, `kipc_structure` и `types`.

Потребительские цены в 2000-2022 гг. (`prices_and_indexes_2000-2022`)

Набор данных представляет собой информацию о средних ценах и индексах потребительских цен для различных типов товаров и услуг по месяцам с 2000 по 2022 годы в территориальном разрезе с различной детализацией. Средние потребительские цены на товары и услуги (на конец месяца) доступны по субъектам РФ с декабря 2000 г. по февраль 2022 г., включая города с 2009 г.

Индексы потребительских цен (ИПЦ) на товары и услуги в группировке классификатора индивидуального потребления по целям (КИПЦ) доступны по месяцам с 2004 г. по февраль 2022 г. Единицей наблюдения является цена или ИПЦ для конкретного типа товаров и услуг в определенный год в определенной территориальной единице.

Структура потребительских расходов в 2012-2022 гг. (`structure_2012-2022`)

Набор данных представляет собой информацию о структуре потребления определенной группы населения по годам с 2012 по 2022 в территориальном разрезе с различной детализацией. Единицей наблюдения является доля определенного типа товаров или услуг в потреблении населения целиком или отдельной децильной группы в определенный год в определенной территориальной единице.

Соответственно, суммарное потребление должно суммироваться к 1 по колонкам: год, `okato`, тип (группа населения).

Расшифровка кодов ОКАТО (`okato`)

Набор данных является словарем, который ставит в соответствие каждому коду `okato` название территориальной единицы и ее уровень. Есть четыре уровня: 0 – Российская Федерация целиком, 1 – группа субъектов РФ (федеральный округ или экономический район), 2 – отдельный субъект РФ, 3 – населенные пункты (за исключением городов федерального значения) или иные территориальные образования внутри субъектов РФ.

Расшифровка типов товаров и услуг (`kipc_prices` и `kipc_structure`)

Наборы данных содержат расшифровку типов товаров и услуг для наборов данных о ценах и структуре расходов населения, соответственно, по классификатору индивидуального потребления по целям (КИПЦ).

Расшифровка типов цен, ИПЦ и групп населения (`types`)

Набор данных содержит расшифровку типов цен, ИПЦ и групп населения.

Примеры использования

Набор данных `prices_and_indexes_2000-2022` представляет интерес для исследований динамики инфляции и цен в территориальном разрезе. К примеру, можно отслеживать изменения цены как на конкретные товары или услуги, так и на полноценную потребительскую корзину или определенный набор. Помимо этого, можно посмотреть на стоимость в разрезе регионов и населенных пунктов, что может быть полезно при исследованиях российских регионов, к примеру в контексте неравенства или доступности тех или иных услуг. Также набор данных позволяет исследовать влияние шоков или иных внешних факторов на динамику цен. Данные по индексам потребительских цен представляют интерес в контексте исследований инфляции, в том числе региональную вариацию. К тому же детальные данные по ИПЦ позволяют отследить за счет изменения цен на какие конкретные товары и услуги менялась инфляция. Соответственно, для упрощения работы стоит отфильтровать даты, территориальные единицы, и группы товаров и услуг, которые представляют интерес для исследователя воспользовавшись словарями. Например, чтобы получить данные о ценах на соль в среднем по стране в 2004 году, надо отфильтровать значения с `кпрс == 2002` (Соль поваренная пищевая, кг), `type == 10` (средняя цена), `okato == 643` (Российская Федерация) и `year == 2004`.

Набор данных `structure_2012-2022` в свою очередь может быть интересен для исследования временной динамики потребления и имущественного неравенства в территориальном разрезе. В частности, можно напрямую отследить какую долю своего дохода граждане тратят на определенный вид товаров или услуг, например высокая доля расходов на еду может служить индикатором низких доходов. К тому же этот набор данных дополняет информацию о ценах, позволяя исследовать не просто абсолютные, но и относительные (относительно дохода определенной группы населения) значения. Аналогичным образом, стоит пользоваться словарями для того, чтобы отфильтровать интересующие исследователя временные промежутки, территориальные единицы, группы населения и типы товаров и услуг. Например, чтобы получить данные о доле расходов жителей Хабаровского края на сырки творожные, глазированные шоколадом необходимо отфильтровать значения с `кпрс == 1126` и `okato == 8000000`.

Известные ограничения данных и внесенные в набор изменения по сравнению с источниками

Из-за постепенного добавления новых товаров, услуг и регионов для показателя «Средние потребительские цены на товары и услуги по субъектам РФ на конец месяца» на начало периода доступно всего 84 КИПЦ (из возможных 2646) и только 88 ОКАТО.

Структура потребления может быть равна 2, если добавлен тип товаров или услуг «все

товары и услуги», который соответствует 100% потребления. Также структура потребления может быть меньше 1, если данные о доле расходов населения на определенную группу товаров или услуг не известна.

Кодировки КИПЦ частично не совпадают для наборов данных о ценах и о структуре потребления, поэтому необходимо выбирать правильный словарь. Если необходимо совместить два набора данных (о ценах и о структуре потребления), мы советуем сверять не только код КИПЦ, но и его расшифровку.

Относительно исходных данных в опубликованных было внесено два изменения:

- из словарей были исключены коды КИПЦ, содержащие буквенные добавления и не использующиеся в данных;
- в словарь ОКATO была добавлена переменная, группирующая значения по территориальным уровням.