

**Европейская экономическая комиссия****Конференция европейских статистиков****Группа экспертов по переписям населения  
и жилищного фонда**

Двадцать первое совещание

Женева, 18–20 сентября 2019 года

Пункт 4 предварительной повестки дня

**Исследования по вопросам использования  
административных данных в целях переписи****Использование административных данных  
для составления списка населения, участвующего  
в переписи****Записка Статистического управления Польши\****Резюме*

В настоящем документе описывается методология составления списка населения, охватываемого переписью, на основе данных из административных источников. Кроме того, представлен новый передовой подход, который позволит, помимо определения численности охватываемого переписью населения, формировать конкретный диапазон информации с учетом любых категорий населения и самого низкого возможного территориального деления (включая 1 км<sup>2</sup>). Этот подход учитывает специфику реестров данных, в том числе их сложность и качество. В документе представлены как методологические, так и технические вопросы, связанные с логикой составления списков населения. Обсуждаются источники данных, а также методы их использования наряду с созданием качественных отчетов.

---

\* Автор: Павел Муравски.



## I. Введение

1. Перепись является единственным обследованием, из которого можно получить подробную информацию о населении, его территориальном распределении, демографической и социальной структуре и профессиональной миграции (в частности, наемных работников), домохозяйствах и семьях, а также о количественном и качественном состоянии жилищных и строительных ресурсов.
2. Эта информация собирается на всех уровнях территориального деления страны: общенациональном, региональном и местном. Благодаря такому подходу, основанному на информации, собранной в ходе переписи, можно запечатлеть образ общества, создать своеобразную фотографию текущей обстановки и состояния общества. В связи с реализацией мероприятий, направленных на расширение использования данных из административных источников, список населения будет формироваться только на их основе. В данном документе представлен метод составления списка, рассмотрены источники данных, а также методологические и технические вопросы, связанные с логикой создания списка населения.

## II. Составление списка населения

3. Целью составления списка населения является создание широкого и полного описания населения для переписи населения 2021 года и ежегодных переписей.
4. Список населения составляется на основе данных, интегрированных из многих разнородных источников, для получения согласованного представления о собранных данных.
5. На данном этапе работы было решено, что "ключом связи" для объединения источников данных станет персональный идентификационный номер (PESEL), поскольку он является основным идентификатором, встречающимся в большинстве административных источников.
6. В таблице 1 представлены административные источники, которые были отобраны для составления списка населения.

Таблица 1

### Описание административных источников данных, используемых для составления списков населения

<i>№</i>	<i>Обозначение вида данных</i>	<i>Поставщик/источник данных</i>	<i>Содержание</i>
1	PESEL	Министерство внедрения цифровых технологий/Реестр личных идентификационных номеров	Лица, зарегистрированные в качестве постоянных или временных жителей Польши
2	KEP	Министерство финансов/ Национальный реестр налогоплательщиков	Физические лица (налогоплательщики)  Физические лица, осуществляющие хозяйственную деятельность (налогоплательщики)
3	ZUS	Управление социального страхования Польши/ Центральный реестр застрахованных лиц	Физические лица (застрахованные лица)

<i>№</i>	<i>Обозначение вида данных</i>	<i>Поставщик/источник данных</i>	<i>Содержание</i>
		Управление социального страхования Польши/ Центральный реестр плательщиков взносов	Физические лица (плательщики взносов)
4	KRUS	Сельскохозяйственный фонд социального страхования	Фермеры, их супруги и члены домохозяйств, работающие с ними (застрахованные)
		Сельскохозяйственный фонд социального страхования	Лица, получающие пенсии и пособия по инвалидности для сельхозработников от Сельскохозяйственного фонда социального страхования
5	NFZ	Национальный фонд здравоохранения	Физические лица (бенефициары)
6	ARIMR	Агентство по реструктуризации и модернизации сельского хозяйства	Физические лица (фермеры)
7	PESEL_ WYMELD	Министерство внедрения цифровых технологий/Реестр личных идентификационных номеров	Лица, выбывшие из реестра без указания пункта назначения, которые до выбытия имели постоянное место жительства в Польше

#### **A. Создание набора данных о лицах с уникальными персональными идентификационными номерами (PESEL)**

7. Первым шагом при составлении списка населения является сбор и сведение воедино уникальных номеров PESEL из административных регистров, приведенных в таблице 1. Все персональные идентификационные номера, присвоенные лицам, проходят проверку на правильность.

8. Подтверждение номера PESEL заключается в проверке количества знаков в идентификаторе, соответствия контрольной цифры и возможности наступления даты рождения в календаре, кодируемой в идентификационном номере. Алгоритм проверки номера PESEL включает автоматическую проверку как прошлых, так и будущих дат, включая високосные годы.

9. В списке лиц, имеющих персональный идентификационный номер, требуется рассчитать такие переменные, как пол и возраст, в связи с необходимостью дальнейшего подтверждения для целей составления списка.

#### **B. Проверка отдельных групп лиц**

10. Вторым шагом в формировании списка населения является проверка следующих групп лиц.

##### **1. Умершие лица**

11. Включение в список умерших лиц проверяется на основании полученной информации за текущий и предыдущие годы. В результате проделанной работы все умершие лица, включенные в список, четко идентифицируются.

## **2. Люди в возрасте 90 лет и старше**

12. На основе рассчитанной переменной WIEK (возраст) проводится анализ лиц в возрасте 90 лет и старше (набор данных о лицах, в отношении которых мало вероятно, что их смерть может быть подтверждена свидетельством о смерти). Проверка таких лиц была проведена правильно, если их персональный идентификационный номер (PESEL) указан в наборах данных PESEL или WYMELD и, по крайней мере, в одном из других наборов данных.

## **3. Лица, покинувшие Польшу**

13. Для идентификации лиц, покинувших Польшу, используются данные, собранные в рамках текущего обследования: постоянные жители других стран, зарегистрированные в Польше для целей временного проживания, заявившие о выезде из Польши или чей заявленный срок пребывания в Польше истек.

## **4. Резиденты и получатели пенсии за пределами Польши**

14. На основании страховых кодов Центрального реестра застрахованных лиц можно идентифицировать записи о лицах, проживающих и получающих пособия в государстве – члене Европейского союза помимо Польши или государстве – члене Европейской ассоциации свободной торговли (ЕАСТ). Такие случаи рассматриваются в контексте пребывания или отсутствия на территории Польши.

## **5. Лица, включенные только в один реестр**

15. Лица, чей персональный идентификационный номер (или его часть) указан(а) только в одном реестре.

## **6. Лица с адресом проживания за пределами Польши**

16. На основании данных Национального реестра налогоплательщиков лицам с иностранным адресом присваивается код страны проживания.

17. Проверка отдельных групп лиц в списке предполагает выделение рассмотренных случаев (пункты 1–6 выше), что позволяет идентифицировать записи и совершенствовать процесс их тщательной проверки в переписи, а также позволяет вносить необходимые корректировки в процесс подсчета численности населения.

# **III. Добавление личных данных и определение адреса**

18. Третий шаг при составлении списка включает процесс добавления имен и фамилий лиц и определения их адресов. Это важный элемент создания полного списка для целей переписи в контексте персонализации данных и проведения переписи с помощью вопросника (аутентификация). Кроме того, персональные данные могут служить одним из альтернативных ключей связи.

## **A. Имя и фамилия**

19. Присвоению фамилии и имени предшествует поиск случаев использования разных фамилий и имен для одного и того же идентификационного номера, включая выверку на предмет повторяющихся значений в отдельных административных реестрах и расхождений между данными разных реестров. Отдельные реестры могут содержать такие ошибки (в основном это касается неправильного написания отдельных букв), которые легко проверить по ссылке на другой реестр.

## **B. Определение адреса**

20. Определение адреса состоит в добавлении информации об адресе к лицам, входящим в список, и проверке правильности решений и алгоритмов по выбору адреса проживания.

21. При определении адреса включенных в список лиц учитываются адреса четырех типов: постоянное место жительства, временное место жительства, адрес регистрации и почтовый адрес, которые сводятся в реестры, указанные в таблице 2.

Таблица 2

**Порядок добавления адресов**

<i>№</i>	<i>Источник</i>	<i>Вид адреса</i>
1	KEP	адрес места жительства (адрес регистрации на последнем этапе)
2	ZUS	адрес места жительства (адрес регистрации, почтовый адрес на последнем этапе)
3	PESEL	временный адрес
4	PESEL	постоянный адрес
5	KRUS	адрес места жительства (почтовый адрес на последнем этапе)
6	ARiMR	адрес места жительства (почтовый адрес на последнем этапе)
7	NFZ	адрес места жительства

22. Актуальность данных и ссылок с точки зрения адресных данных определяет иерархию выбора адреса из определенного набора данных. Уровень соответствия учитывается для всех адресов.

Таблица 3

**Уровень соответствия адресных данных**

<i>№</i>	<i>Уровень соответствия</i>
1	соответствие на уровне улицы
2	соответствие на уровне места (города)
3	соответствие на уровне общины
4	отсутствие несоответствия

23. Процесс добавления адресов в соответствии с алгоритмом был разделен на четыре этапа:

a) во-первых, последовательный выбор адресов с соответствием на уровне улицы из наборов данных KEP, ZUS и PESEL;

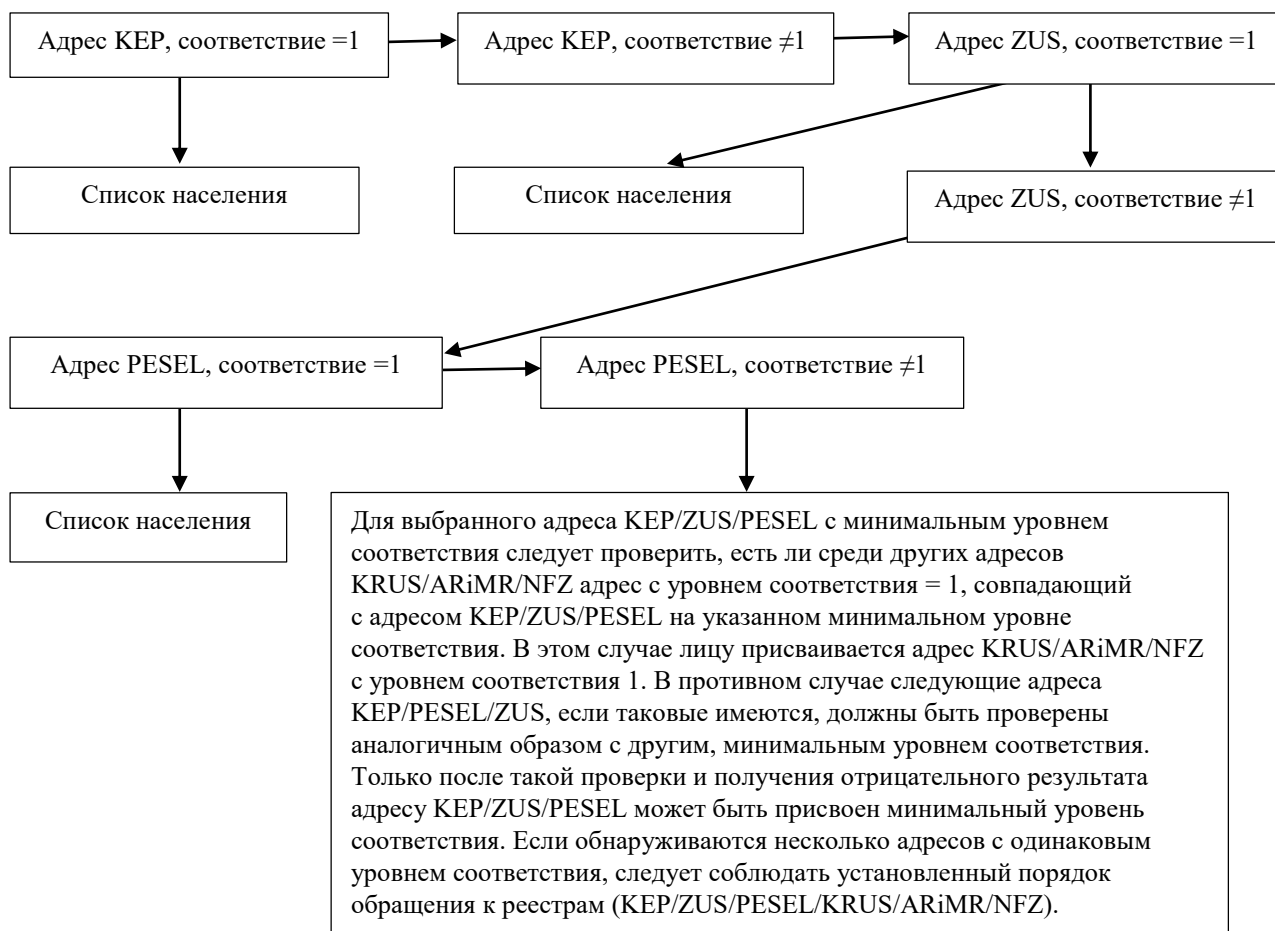
b) во-вторых, выбор адресов из наборов данных KEP, ZUS и PESEL с минимальным уровнем соответствия (отличным от 1) и корректировка/дополнение их с помощью сведений из наборов данных KRUS, ARiMR и NFZ, согласующихся с данными из KEP, ZUS и PESEL на определенном минимальном уровне соответствия. Единицы жилья, по которым невозможно получить более полную информацию о проживании, вносят в список населения;

c) в-третьих, добавление адреса с минимальным уровнем соответствия из наборов данных KRUS, ARiMR и NFZ;

d) в-четвертых, добавление так называемых дополнительных адресов.

Диаграмма 1

## Упрощенная диаграмма алгоритма выбора адреса



#### IV. Заключение

24. Представленный метод составления списка населения исключительно на основе данных из административных источников позволит охватить всю совокупность лиц, которые должны быть обследованы в ходе переписи. Благодаря этому возникает возможность добиться полной интеграции с другими предметными областями, обеспечить высокое качество данных, а процесс сделать полностью автоматизированным и воспроизводимым.