

Европейская Экономическая Комиссия
Организации Объединенных Наций

Статистика на основе регистров в Северо-европейских странах

Обзор передовых методик с
уделением основного внимания
на статистику населения и
социальной статистике



ОРГАНИЗАЦИЯ ОБЪЕДИНЕННЫХ НАЦИЙ

**ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ
ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ**

Статистика на основе регистров в Североевропейских странах

**Обзор передовых методик с уделением основного
внимания на статистику населения и социальной
статистике**



**ОРГАНИЗАЦИЯ ОБЪЕДИНЕННЫХ НАЦИЙ
НЬЮ-ЙОРК И ЖЕНЕВА, 2008**

ПРИМЕЧАНИЕ

Условные обозначения документов Организации Объединенных Наций состоят из прописных букв в сочетании с цифрами. Упоминание такого условного обозначения означает ссылку на документ Организации Объединенных Наций.

Обозначения, используемые в настоящем издании, и изложение материала не подразумевают выражения со стороны Секретариата Организации Объединенных Наций какого бы то ни было мнения в отношении правового статуса той или иной страны, территории, города или районов или их органов власти, или делимитации их границ.

UNITED NATIONS PUBLICATION

Sales No. E.07.II.E.11

ISSN 0069-8458

ISBN 978-92-1-116963-8

Авторское право принадлежит ООН, 2008
Все права защищены
Напечатано в ООН, Женева (Швейцария)

СОДЕРЖАНИЕ

Предисловие	v
Об отчете	1
Глава 1. ЗАЧЕМ НУЖНА СТАТИСТИКА НА ОСНОВЕ РЕГИСТРОВ?	3
Глава 2. ИСТОРИЧЕСКОЕ РАЗВИТИЕ СТАТИСТИКИ НА ОСНОВЕ РЕГИСТРОВ	5
Глава 3. ОБЩИЕ ПРЕДПОСЫЛКИ	7
3.1. Правовая база.....	7
3.2. Общественная поддержка.....	7
3.3. Единые системы идентификации.....	8
3.4. Исчерпывающие и достоверные системы регистров, созданные для административных нужд	9
3.5. Сотрудничество между органами управления	9
Глава 4. ИСПОЛЬЗОВАНИЕ АДМИНИСТРАТИВНЫХ ДАННЫХ В СОСТАВЛЕНИИ СТАТИСТИКИ	11
4.1. Сравнение выборочных обследований и регистров.....	11
4.2. Статистика только на основе регистров.....	12
4.3. Комбинированное использование выборочных наблюдений и данных регистров	12
Глава 5. РЕГИСТРЫ – ОБЩИЙ ОБЗОР	15
5.1. Что такое регистр?.....	15
5.2. Базовые и специализированные регистры	16
5.3. Регистры физических лиц.....	17
5.4. Регистрация собственности	18
5.5. Регистры хозяйствующих субъектов.....	19
5.6. Регистры рода занятий	20
5.7. Обновление регистров	20
Глава 6. ТРЕБОВАНИЯ К ДАННЫМ ИЗ АДМИНИСТРАТИВНЫХ ИСТОЧНИКОВ	23
Глава 7. ОТ АДМИНИСТРАТИВНЫХ ДАННЫХ К СТАТИСТИКЕ – ОБРАБОТКА ДАННЫХ	27
7.1. Редактирование.....	27
7.2. Использование множественных источников	28
7.3. Разработка статистической системы на основе регистров	30
Глава 8. РАСПРОСТРАНЕНИЕ И ПОТРЕБНОСТИ ПОЛЬЗОВАТЕЛЕЙ	35
8.1. Географические информационные системы и статистика на основе регистров.....	35
8.2. Микроданные для исследователей	36
Глава 9. КАЧЕСТВО И СТАНДАРТИЗАЦИЯ	39
9.1. Качество	39
9.2. Метаданные и документация	40
9.3. Согласование	42
Глава 10. ПЕРЕПИСИ НАСЕЛЕНИЯ И ЖИЛИЩНОГО ФОНДА НА ОСНОВЕ РЕГИСТРОВ	46
10.1. Разработка проведения переписей на основе регистров.....	46
10.2. Сопоставимость с международными рекомендациями	47
10.3. Преимущества и недостатки проведения переписи на основе регистров.....	48

Глава 11. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ.....	51
Термины, принятые в статистике на основе регистров	53
Использованная литература	55

Предисловие

Североевропейские страны имеют давнюю традицию использования административных регистров при составлении официальной статистики. У них также существует давняя традиция сотрудничества в сфере статистики. Кроме того, их объединяет опыт использования административного учета при переписи населения как последнего этапа процесса, который начинается с составления статистики в различных предметных областях в зависимости от типа доступных регистров. При составлении статистики по населению или занятости на основании административных данных, они изучили влияние, которое такие регистры оказывают на качество статистики. После периода опробования и доработки они поняли, что качество административных данных совместимо с качеством, рекомендуемым для переписи населения, и приняли решение использовать для переписи населения также и регистры. Время, которое проходит от создания административного регистра до получения качества данных, необходимого для переписи населения, может быть различным в различных предметных областях. Однако этот процесс в различных предметных областях и различных странах одинаков. То же самое, похоже, касается и создания статистической системы на основе информации из регистров.

В последние годы увеличивается число стран в регионе ЕЭК ООН, рассматривающих возможность составления статистики на основе административных регистров. В связи с этим Национальные статистические институты (НСИ) в Североевропейских странах (Дания, Исландия, Норвегия, Финляндия и Швеция) решили поделиться своим опытом и знаниями с международным статистическим сообществом путем создания обширной документации по своим передовым методикам. Целью данного издания является предоставление сотрудникам стратегических и плановых отделов НИС знаний о том, что такое статистика, создаваемая на основе регистров, в т. ч. необходимые технические и административные функции и возможное приложение этих методов для составления официальной статистики. Основное внимание в данном издании уделено использованию административных регистров при составлении демографической и социальной статистики.

Опубликовывая данное издание, Европейская экономическая комиссия Организации Объединенных Наций (ЕЭК ООН) хотела бы поддержать Североевропейские страны в передаче их опыта по данному вопросу с международным статистическим сообществом в целом. Издание представляет собой ценный инструмент для всех НСИ (как в регионе ЕЭК ООН, так и за его пределами), планирующих составление официальной статистики на основе административных регистров. Оно также поддерживает проведение в 2010 г. цикла переписи населения и жилищного фонда, поскольку растущее число стран за пределами Северной Европы планирует использовать административные источники в качестве дополнения к сбору традиционных данных на местах или их замены. Руководство по использованию регистров для переписи населения и жилищного фонда также включено в новый комплект «Рекомендации Конференции европейских статистиков по проведению переписи населения и жилищного фонда 2010 года»¹, подготовленный ЕЭК ООН совместно со Статистическим агентством Европейского Союза (ЕВРОСТАТ). Новые рекомендации, Конференции европейских статистиков (КЕС) по переписям, впервые содержат подробное описание методологических подходов, в т. ч. к переписям, проводимым только на основе административных регистров или в сочетании с другими источниками. Для каждого из этих подходов описаны необходимые условия, а также их относительные преимущества и недостатки.

Настоящий отчет подготовлен рабочей группой, учрежденной руководителями статистических служб Североевропейских стран. В группе работали: Юхан-Кристиан Тендер (Руководитель группы, Центральное статистическое бюро Норвегии), Финн Спикер (Центральное статистическое бюро Дании), Пекка Мюрскюля (Центральное статистическое бюро Финляндии) и

¹ См. веб-сайт Отдела статистики ЕЭК ООН: <http://www.unece.org/stats/publ.htm>

Клаус-Йёран Йельм (Центральное статистическое бюро Швеции). Харальд Утне (Центральное статистическое бюро Норвегии) являлся секретарем группы. Центральное статистическое бюро (ЦСБ) Исландии не смогло принимать участия в данном проекте.

ЕЭК ООН выражает благодарность национальным статистическим институтам Дании, Норвегии, Финляндии и Швеции, которые содействовали подготовке отчета, и вышеупомянутым экспертам, выполнявшим работу. ЕЭК ООН хотела бы также поблагодарить лиц, высказавших ценные комментарии к ранним проектам отчета, в особенности экспертам из НИС Латвии, Нидерландов, Словении, Соединенного Королевства и Швейцарии.

Женева, 2007г.



Марек Белька,
Исполнительный секретарь
ЕЭК ООН

Об отчете

1. С течением времени североевропейские НСИ представили ряд отчетов по составлению статистики на основе административных источников. Цель настоящего отчета - собрать весь основной опыт в *едином* документе. Опыт североевропейских стран очень похож, что позволяет описать некоторые передовые методики, одинаковые во всех странах. Однако там, где это уместно, даются примеры из отдельных стран и проводится сравнение между странами.
2. То, что работает в Северной Европе, не обязательно будет приложимо в других странах. Решение о роли статистики на основе регистров в какой-либо стране всегда должно приниматься в соответствии с национальными условиями. Однако мы надеемся, что описание североевропейского опыта окажется полезным для процесса принятия решений.
3. Авторы старались, чтобы отчет получился кратким. Это, конечно, означает, что многие детали были опущены. Перечень отчетов, содержащих дополнительную информацию, приводится в конце данного отчета.
4. В основу данного отчета легли несколько отчетов и документов, как изданных так и не изданных. Основными из них являются следующие:
 - «Статистика лиц в Дании – статистическая система на основе регистров» (Statistics on Persons in Denmark), Центральное статистическое бюро Дании, 1995;
 - «Использование регистров и источников административных данных в статистических целях – Передовые методики в Центральном статистическом бюро Финляндии» (Use of Registers and Administrative Data Sources for Statistical Purposes – Best practices in Statistics Finland), Центральное статистическое бюро Финляндии, 2004;
 - «Статистика на основе регистров – административные данные в статистических целях» (Register-based statistics – administrative data for statistical purposes) Валлгрэн и Валлгрэн, Центральное статистическое бюро Швеции, издано на шведском языке;
 - «Снижение расходов при проведении переписи населения в Норвегии путем использования административных регистров» (Reducing Costs of Censuses in Norway Through Use of Administrative Registers) Лонгва, Томсен и Северейде, Центральное статистическое бюро Норвегии, 1998.
5. Структура в основном взята из отчета Финляндии, но была изменена в соответствии с объемом данного отчета.

Глава 1. ЗАЧЕМ НУЖНА СТАТИСТИКА НА ОСНОВЕ РЕГИСТРОВ?

6. Все национальные статистические институты (НСИ) обязаны составлять официальную статистику максимально высокого качества, отвечающую требованиям:

- Значимости и завершенности;
- Своевременности и пунктуальности;
- Точности;
- Сопоставимости и согласованности;
- Доступности и ясности;
- Рентабельности;
- Короткого времени для ответов респондентов на вопросы обследования.

7. При планировании переписи или выборочного обследования НСИ имеют некоторое представление о качестве, когда идет определение единиц и переменных. Мы хотим использовать методы компоновки и редакции данных и для представления статистики на основе статистических целей и на идеалах, которые статистики изучали в университетах или узнали от коллег по НСИ. На практике данные приходится приспособлять. Возможно, нам не удастся провести интервью со всеми членами домашнего хозяйства и придется довольствоваться ответами одного из них (опрос доверенных лиц). Объяснить переменную, которую мы на самом деле хотим измерить, может быть настолько сложно, что нам придется ее упростить, и тем не менее людям, возможно, будет сложно ответить на вопросы. Кроме того, ограниченный бюджет и ограниченное время для ответов респондентов могут уменьшить большую перепись до размеров малой или перепись до выборочного наблюдения (или сочетание обоих методов). Тщательное редактирование вручную может быть сведено к приблизительному автоматическому. Возможно, нам удастся оценить последствия «срезания углов» и компенсировать его условными исчислениями на уровне статистических единиц или исправлением таблиц. Мы всегда должны подводить баланс между искомым качеством и практической и финансовой реальностью.

8. Административные данные составляются на основе некоторых административных процессов, а единицы и переменные определяются, исходя из административных правил и требований. Определения могут отличаться от нужд официальной статистики, но данные обычно получаются хорошего качества, достаточного для ее административных целей (подробнее об административных целях см. Главу 9).

9. Статистикам всегда приходится идти на компромиссы с идеальными условиями работы, чтобы получить практические результаты от сбора данных. В какой-то степени использование административных данных переменяло ситуацию: у нас есть продукт процесса сбора данных, и надо его сравнить с нашими требованиями по качеству к статистике, чтобы увидеть, приемлема ли разница. Административные определения целевого населения могут не соответствовать нашим нуждам (работники по найму, но не индивидуальные предприниматели в регистрах занятости), переменные не всегда определены так, как нам хочется (место проживания по прописке, а не фактическое в регистре населения), и временные ссылки не всегда настолько точны, как нам хотелось бы (время, указанное в регистре, может не совпадать со статистическим «днем проведения переписи»). Но поскольку эти данные почти бесплатны для использования НСИ, мы можем использовать ресурсы для дополнения информации, не вошедшей в административные данные, и вносить поправки точно так же, как и при использовании традиционных методов.

10. В итоге результат может быть таким, что при определенном уровне финансирования требования к качеству удовлетворяются лучше с помощью данных административных регистров, чем

при традиционном сборе данных с помощью анкет. Похоже, что большинству НСИ сокращают бюджет, не сообщая о приоритетах или указаниях, как сокращать статистическую программу или поддерживать уровень качества статистики. Вместо этого мы видим растущий интерес к новым статистическим объектам и новым сочетаниям переменных. Во многих странах спрос на статистику по небольшим группам населения выше, чем может обеспечить перепись населения, проводимая раз в 10 лет.

11. Административные ведомства профессионально формулируют прозрачные правила и процедуры и рассматривают заявки, распределяют средства, собирают налоги или оказывают другие административные услуги. Эти ведомства также работают под жестким финансовым давлением. В такой ситуации ведомства и НСИ должны сотрудничать. НСИ могут предлагать опыт и возможность по сбору, обработке и хранению данных. Сообща они могут разработать систему, удовлетворяющую как административные, так и статистические цели.

Пример: Норвежская статистика по образованию

12. Почти все образование в Норвегии регулируется одним или несколькими «законами об образовании». Статья, которая встречается во всех этих законах, дает право Министерству образования (или учреждению, уполномоченному Министерством) собирать данные от учебных заведений об учениках, преподавателях, зданиях, расходах, доходах и т. п. Министерству нужна эта информация, чтобы управлять системой образования, а также иногда, чтобы контролировать использование денег, выделяемых правительством учебным заведениям.

13. Но Министерство не может профессионально собирать, обрабатывать и представлять данные и, поэтому, сотрудничает с ЦСБ Норвегии. Ему разрешено от имени Министерства собирать данные от учебных заведений для административных и статистических целей. Министерство может получить копию «сырых» данных, потому что оно по закону является инстанцией, контролирующей отдельные заведения. Но в большинстве случаев Министерство желает, чтобы НСИ предоставляли статистику для целей правительства. Соответственно, Министерство может сосредоточить свои усилия на анализе статистики и формировании политики.

14. Даже если данные собраны на основании закона об образовании, ЦСБ Норвегии имеет через сотрудничество с Министерством возможность влиять на содержание и методы сбора данных от учебных заведений. Это означает, что НСИ может влиять на факторы, которые сказываются на качестве статистики. Разница между административными и статистическими данными сведена почти к нулю.

15. Эти данные можно назвать «административными данными» потому что они собраны из административных архивов учебных заведений для административных целей (управление и контроль). НСИ составляет подробную статистику для Министерства, а также официальную статистику для всей страны. Статистика составлена независимо от того, какую картину положения в образовании хочет представить Министерство, хотя она составлена на основе министерских данных. Поэтому статистика составляется в соответствии с Основными принципами официальной статистики ООН.

16. В упрощенном виде можно сделать такой вывод: используя традиционные методы сбора данных, НСИ спрашивает: «Мы задали наш вопрос, но получаем ли мы правильный ответ?» Используя данные на основе регистров (административные), НСИ спрашивает: «Конечно, мы получили правильный ответ, но как он соотносится с нашим вопросом?» НСИ пытается достичь той же самой цели, но из другой исходной точки.

Глава 2. ИСТОРИЧЕСКОЕ РАЗВИТИЕ СТАТИСТИКИ НА ОСНОВЕ РЕГИСТРОВ

17. Все североевропейские страны имеют длительный опыт в составлении статистики на основе регистров, которая сначала касалась социальной статистики, а с годами расширилась и в последнее время стала охватывать также и статистику хозяйствующих субъектов. Важно отметить, что административные данные никогда не смогут полностью заменить непосредственный сбор данных при проведении выборочных обследований, но эти два метода дополняют друг друга.

18. История использования административных регистров в ведении статистики в североевропейских странах кратко представлена в таблице ниже. Она ограничена переменными, используемыми при переписи населения и жилищного фонда.

Таблица. Год создания регистров/использования впервые регистров в данных переписи по типам регистров и странам

Тип регистра	Дания		Финляндия		Норвегия		Швеция	
	Создан	Впервые использован	Создан	Впервые использован	Создан	Впервые использован	Создан	Впервые использован
Центральный регистр населения	1968	1981	1969	1970	1964	1970	1967	1975
Регистр хозяйствующих субъектов	1975	1981	1975	1980	1965	1980	1963	1975
Жилье	1977	1981	1980	1985	2001	2011	2008?	2011?
Жилищные условия	1977	1981	1980	1985	2001	2011	2008?	2011?
Образование	1971	1981	1970	1975	1970	1980	1985	1990
Занятость	1979	1981	1987	1990	1978	2001	1985	1985
Семья	1968	1981	1978	1980	1964	1980	1960	1975
Домохозяйство ^а	1968	1981	1970	1975	2001	2011	2011?	2011?
Доход	1970	1981	1969	1970	1967	1980	1968	1975
Перепись полностью на основе регистров		1981		1990		2011		2011?

^а Домохозяйство - жилищная единица, т. е. все лица, проживающие в одном жилище.

19. В период с 1964 по 1969 гг. центральные регистры населения были открыты во всех североевропейских странах, и была введена система уникальных персональных номеров для каждого человека. В последующие годы было создано еще несколько административных регистров. Административные данные как источник получения статистических данных стали использоваться с начала 1970-х гг. Сначала регистры использовались в некоторых предметных областях статистики, первыми из которых были статистические данные по населению и доходам. В дальнейшем статистика на основе регистров была разработана во всех странах.

20. Время, проходящее от создания административных регистров до момента, когда данные становятся удовлетворительными для переписи, может быть разным для разных предметных областей и стран. Примером этому может служить составление статистики по занятости. В Дании, Финляндии и Швеции этот процесс длился несколько лет, но в Норвегии ситуация была иной.

Пример: Составление статистике по занятости на основе регистров в Норвегии.

21. Норвежский регистр работающих по найму был создан в 1978 г. С 1983 г. статистика малых территориальных единиц по работникам по найму была опубликована как приложение к Обследованию рабочей силы. Однако статистика не отразила все значимые единицы (работающих по найму) и содержала только несколько переменных. В конце 1980-х и 1990-е гг. были созданы новые регистры, и объем статистических данных рынка рабочей силы на основе регистров постепенно увеличивался. В 2001 г. объем данных и качество стали достаточно удовлетворительными для опубликования переменных переписи занятости на основе регистров.

22. В 1981 г. Дания стала первой страной в мире, проведшей перепись полностью на основе регистров, за ней в 1990 г. последовала Финляндия. С 1980 г. переписи в Норвегии и Швеции частично основывались на регистрах, эти страны планируют провести свои первые переписи на основе регистров в 2011 г. Основным различием между странами является то, что Дания и Финляндия имеют регистры жилья с 1980 г., а в Норвегии такой регистр был создан в 2001 г. В Швеции планируется создать такой регистр в 2008 г.

23. Однако поэтапное развитие идет одинаково во всех странах: сначала статистика предметных областей была опробована и опубликована в различных областях. Затем, как только качество было сочтено удовлетворительным, переменные на основе регистров были использованы при переписи. Наконец когда статистика разработана для всех областей, регистрируемых при переписи, становится возможным проводить перепись, полностью основанную на регистрах.

24. Последние разработки в сфере информационных технологий позволяют сократить время, проходящее между созданием административных регистров и их использованием для официальной статистики. Но даже для регистров, создаваемых сегодня, этот процесс занимает определенное время. Необходимо вводить процедуры отчетности, которые обычно не работают безукоризненно с первого дня, особенно когда речь идет о сведениях, поступающих от населения.

Глава 3. ОБЩИЕ ПРЕДПОСЫЛКИ

25. В свете североευропейского опыта можно утверждать, что существуют определенные предварительные условия, которые способствуют широкому использованию административных источников данных для составления статистики.

3.1. Правовая база

26. Законодательство предоставляет основной фундамент для использования источников административных данных в статистических целях. Национальное законодательство должно отражать широко распространенное мнение о том, что целесообразнее использовать источники административных данных, а не собирать данные еще раз для статистических целей. Все североευропейские страны имеют *законы о статистике*, дающие НСИ право доступа к административным данным на уровне единиц отбора с опознавательными данными и право связывать их с другими административными регистрами для статистических целей. Кроме того, законы о статистике содержат детальное определение защиты данных.

27. В некоторых североευропейских странах составители статистики обязаны использовать в первую очередь данные, собранные в других контекстах. НСИ обязаны проверить наличие данных в административных регистрах перед началом непосредственного процесса сбора данных. Все данные должны собираться рентабельным способом, причинять минимум неудобств и расходов респондентам.

28. В дополнение к этому, нормативно-правовые акты ЕС согласованы с законами о статистике. *Глава 16 Регламента Совета по статистике сообщества* гласит, что в целях уменьшения трудностей для респондентов национальные органы власти и орган власти Сообщества должны иметь доступ к источникам административных данных, каждый в сфере своего общественного управления, в степени, в которой эти данные необходимы для составления статистики сообщества².

29. Все североευропейские страны имеют *закон об обработке личных данных*, содержащий требования к обработке личных данных. Цель закона – обеспечить такое создание и использование электронных регистров, содержащих личные данные, которое бы не нарушало законные права отдельных граждан на защиту и неприкосновенность информации о них.

30. В соответствии с этим законом, обработка данных для статистических целей разрешена даже, если она не являлась основной целью сбора данных. После обработки данных в НСИ они не должны использоваться для других целей, кроме статистических и исследовательских (принцип «одностороннего движения»). Статистическое агентство может позволить доступ к конфиденциальным данным для научных исследований или статистических обследований. Данные, собранные для статистических целей, являются конфиденциальными, независимо от источника. Данные, собранные из административных источников, являются конфиденциальными, когда ими распоряжаются статистические органы власти, даже если эти данные открыты, и когда ими распоряжаются административные органы власти. При обращении с личными или коммерческими данными, обычно исключается прямое и косвенное опознавание.

3.2. Общественная поддержка

31. Наличие все большего числа административных регистров в обществе, конечно, вызывает дискуссии о тайне личной жизни. Если мнение общественности станет отрицательным, политики

² Регламент Совета (ЕС) № 322/97 от 17 февраля 1997 г. о статистике Сообщества.

могут проявить нежелание создавать новые регистры или совершенствовать существующие. Статистическое использование административных данных обычно сопряжено со сведением данных из разных регистров, что может создать впечатление, будто НСИ знают «все» о каждом отдельном гражданине («синдром старшего брата»).

32. С другой стороны люди очень хорошо знают, что административные власти собирают те же данные, которые НСИ используют для статистических целей. Например, налоговые власти имеют информацию о работодателе и заработанных доходах каждого человека, органы, начисляющие пенсии, - о трудовых периодах работников по найму, биржи труда – обо всех безработных, отделы социальной защиты населения – о престарелых и других категориях пенсионеров. В такой ситуации очень трудно призывать людей предоставлять одни и те же данные для статистических целей.

33. Чрезвычайно важно, чтобы общественность высоко оценивала и понимала преимущества использования регистров в статистических целях и чтобы использование административных данных для составления статистики пользовалось широкой общественной поддержкой. Крайне необходимо, чтобы национальное законодательство о регистрах было современным, а работа органов власти, отвечающих за регистры, - открытой и прозрачной. Показателями состояния в отдельной стране может быть процент ответивших на обследования официальной статистики, отношение политиков к предложениям о создании и использовании регистров, извлеченные уроки от общественных дебатов.

34. С течением времени в разных странах развитие пошло по-своему. В Финляндии и Дании население, похоже, поверило, что статистическое использование данных регистров рационально, там есть большая вера в НСИ. В 1970-е гг. в Норвегии прошла широкая дискуссия об административных регистрах и статистике, но после этого ситуация стала более или менее такой же, как в Финляндии и Дании. В Швеции дискуссия продолжается примерно с 1970 г., то замирая, то возобновляясь. Однако общим для всех этих стран является то, что открытое обсуждение и дебаты, объясняющие смысл и преимущества использования регистров, всегда считались основным принципом.

35. С точки зрения населения использование административных данных для статистических целей имеет следующие преимущества:

- Рентабельность;
- Люди, домохозяйства и фирмы могут функционировать нормально, их не беспокоят излишними вопросами;
- Защита данных лучше, поскольку их обрабатывает меньшее число людей, как внутри НСИ, так и за их пределами;
- Только компьютер «видит» все данные, поскольку они существуют в электронном виде.

36. Похоже, что большинство населения считает эти преимущества вескими аргументами для статистического использования существующих административных данных. Однако важно, чтобы статистическое агентство никогда не теряло бдительности. Очень легко потерять доверие общественности, а восстановление его требует огромных усилий.

3.3. Единые системы идентификации

37. Одним из главных факторов, делающих возможным статистическое использование административных данных, является использование единых систем идентификации для разных источников. При отсутствии таких единых систем гораздо труднее связать между собой различные регистры, а в этом как раз заключается смысл составления статистики на основе регистров. Минимальным требованием является наличие единой идентификационной системы для базовых регистров.

38. Почти во всех регистрах, используемых для составления статистики в североевропейских странах, в настоящее время введены единые идентификационные индивидуальные коды (личные идентификационные номера). Для других базовых регистров – таких, как коммерческий, адресный, жилищный, - существуют почти единые идентификационные системы. Связывать различные регистры возможно даже без единых идентификационных кодов, но это, конечно, более трудоемкая работа, требующая много времени.

3.4. Исчерпывающие и достоверные системы регистров, созданные для административных нужд

39. Создание регистров административных данных было обусловлено потребностями для функционирования общества и развития управления. Оно было тесно связано с развитием общественной безопасности, системы налогообложения и других административных функций. В основном это системы, управляемые государством, поэтому регистры было необходимо создать на государственном уровне. Очень часто цели административных систем связаны, и поэтому различные учреждения обмениваются информацией регистров. Такие процессы вносят полезные корректировки в административные регистры, улучшая, таким образом, качество статистики на основе регистров.

40. Жители и органы власти североевропейских стран всегда проявляли большое доверие к точности источников регистров, которые являются основой многих личных прав и обязанностей. Например, официальное место жительства каждого жителя страны определяется на основании регистровой информации. Выписка из регистра населения служит основным документом для выдачи паспорта, оформления брака или развода, проведения похорон или распределения имущества. В интересах самого человека удостовериться, что все данные в административных регистровых системах действительно правильные. Хорошее качество данных для административных нужд, конечно, также важно для статистического использования, даже если «административное качество» автоматически не ведет к «статистическому качеству».

41. Одной из причин успешного развития административных регистров в североевропейских странах, возможно, является то, что все эти страны довольно маленькие и гомогенные. В более крупных и гетерогенных странах создание административных регистров на государственном уровне может быть более проблематичным из-за технических, организационных и других ограничений. Однако даже если явно предпочтительно создать систему регистров в масштабах всей страны, местные регистры также могут быть полезными при составлении статистики при условии, что они содержат необходимые переменные, что определения переменных делают их сопоставимыми для статистики, а НСИ имеет юридическое основание на доступ к этим регистрам.

3.5. Сотрудничество между органами управления

42. Совместные усилия при создании статистики на основе регистров требуют твердого и ясного обязательства на самом высоком уровне, а также тесное сотрудничество между органами власти.

43. Правительства североевропейских стран оказывают политическую поддержку усилиям НСИ в разработке статистических систем на основе регистров. Законы о статистике закрепляют такую политику. На практике это означает большую поддержку НСИ в их переговорах с административными органами власти по получению доступа к административным данным. Сотрудничество между НСИ и органами власти, владеющими тем или иным регистром, часто ведется в рамках национального форума держателей регистров, но также и в двустороннем порядке.

Глава 4. ИСПОЛЬЗОВАНИЕ АДМИНИСТРАТИВНЫХ ДАННЫХ В СОСТАВЛЕНИИ СТАТИСТИКИ

44. В этой главе мы сравним два основных типа источников данных, используемых в североевропейских странах: *выборочные обследования* и *данные на основе регистров*. Кроме того, мы обсудим различные пути использования административных данных при составлении статистики.

4.1. Сравнение выборочных обследований и регистров

Преимущества и недостатки метода выборочного обследования

45. Основным преимуществом данных обследования является прямой контроль, который можно проводить за содержанием данных, поскольку есть возможность задавать вопросы именно на те темы, информацию о которых необходимо получить. Кроме того, данные выборочных обследований можно в большинстве случаев обрабатывать и публиковать гораздо быстрее, чем данные регистров, если использовать современные методы – такие, как личные интервью с помощью компьютера (CAPI), компьютеризированные интервью по телефону (CATI) и т. п.

46. С другой стороны статистика на основе обследований является довольно дорогостоящим методом сбора данных. Еще одна проблема заключается в качестве данных. Отсутствие ответов может привести к смещению статистических результатов, а проинтервьюированные домохозяйства, лица и т. д. не всегда могут или хотят давать правильные ответы.

Преимущества и недостатки регистравого метода

47. Особым преимуществом статистики на основе регистров является то, что она в принципе обеспечивает полный охват, даже если затраты на сбор и обработку довольно низки. Используя регистры, мы можем составлять более подробную статистику, чем используя выборочные наблюдения, например, статистику для малых территориальных единиц или для очень подробных классификаций.

48. Данные регистров имеют большой потенциал, потому что можно связать различные регистры на основе точно определенных ключей (числа для обозначения людей, хозяйствующих субъектов, адресные коды и т. п.). Более того, данные отдельных административных регистров обычно согласованы и высококачественны, по крайней мере, для административных целей. Когда данные регистров используются для статистических целей и особенно когда комбинируются данные из нескольких регистров, могут обнаруживаться проблемы согласованности данных и их качества. Использование нескольких источников подробнее описывается в п. 7.2.

49. Регистравый метод может иметь ту же черту, что и отсутствие ответа в выборочных наблюдениях, - неполноту охвата. Вот пример из Норвегии: регистр самого высшего образования населения становится все менее полным, потому что иммигранты не обязаны указывать свое образование при въезде в страну. Поэтому каждые десять лет Центральное статистическое бюро Норвегии должно проводить особое обследование образования иммигрантов.

50. Другим недостатком статистики на основе регистров является то, что статистики в большой степени связаны определениями и практикой управления органов власти, которые отвечают за ведение регистров. Данные из административных источников можно рассматривать как «представление о мире органов власти». Более того, составитель статистики не очень близок к настоящему сбору данных, в результате чего он может не знать достаточно о точном содержании

данных, их обработке, качестве и т. д. Лучший способ решить проблему – тесно сотрудничать с владельцами регистров.

4.2. Статистика только на основе регистров

51. Некоторые статистические данные составляются только с использованием регистров. Например, статистика по населению, различная статистика хозяйствующих субъектов, статистика доходов и благосостояния, статистика преступлений и статистика жилья. В некоторых случаях данные из других источников, например, из выборочных обследований, могут быть использованы для оценки качества, но не для изменения данных регистров, использованных в статистике.

4.3. Комбинированное использование выборочных наблюдений и данных регистров

Данные регистров, дополняемые данными выборочных обследований

52. Часто запрашиваемые статистические переменные нельзя получить из одного административного источника и необходимо комбинировать данные из нескольких регистров (использование множественных источников описывается в п. 7.2). Примером этому может служить статистика занятости на основе регистров, где информация о работе собрана из нескольких административных регистров. Не все данные исчерпывающие, и могут возникать несоответствия между источниками данных. Разработаны методы построения модели или оптимальной группы правил принятия решений, чтобы установить значение параметра, который соответствует данной статистической концепции. Например, данные обследования рабочей силы могут быть использованы для калибровки модели, использованной в обследовании рынка труда на основе регистров. Для переменной *род занятий в настоящее время* используется целый ряд источников данных (в Финляндии, например, более 20). Таким образом, информация об одном и том же человеке доступна из различных административных источников. В моделях для выбора наиболее надежных источников данных можно вводить данные из выборочных наблюдений.

53. Восстановление недостающих данных для отсутствия ответа в данных регистров чаще всего исправляется использованием дополнительных данных из других регистров. Однако в некоторых случаях в условных исчислениях используются данные обследования. Примером этому могут служить рабочие часы индивидуальных предпринимателей в норвежской региональной статистике занятости, в которой использованы данные обследования рабочей силы на макроуровне.

Данные регистров, используемые в выборочных обследованиях

Представление в наглядной форме / основа выборки

54. Регистры используются как основа для построения выборки обследований. Чаще всего используются регистр населения и регистр хозяйствующих субъектов. Регистр населения, например, содержит важные демографические данные - такие, как возраст, пол, семейное положение, местожительство, - которые используются для определения населения и составления выборки. Используя в дополнение данные из других регистров, можно производить выборки для желаемых подгрупп, например, работающих по найму или студентов. Для составления выборки по предприятиям и учреждениям можно, например, использовать информацию из регистров хозяйствующих субъектов об отрасли промышленности и местонахождении.

Совершенствование информации

55. Информация регистров используется для повышения качества данных от выборочных обследований. Ошибки, возникающие при сборе данных, можно искать путем сопоставления данных регистров и данных обследований на микроуровне (*редактирование*). Данные регистров можно

использовать при *присвоении кода*. Например, если название места работы и адрес (учреждения) получено в интервью, отраслевой код можно найти в регистре хозяйствующих субъектов. Единица «отсутствие ответа» или недостающая информация могут быть исправлены, если обратиться к данным регистров (*восстановление недостающих данных*).

56. При *оценке* данные регистров можно использовать для расчета веса, например, путем пост-стратификации. Распределение населения, получаемое из регистров, например, возрастное распределение, будет, таким образом, оценено правильно. Более того, погрешность оценки снизится, если это пристрастная выборка по отношению к населению, а надежность оценок может значительно повыситься, если есть прочная корреляция между данными регистров и переменными выборочного наблюдения.

Дополнительная информация

57. Данные регистров используются, чтобы дополнять объем информации обследований. Например, демографические данные и данные об образовании в обследованиях часто берутся из регистров. Одна из причин этого, конечно, - экономия времени при проведении интервью, поскольку время в этом контексте является прямым фактором, влияющим на стоимость. Кроме того, в ряде случаев данный метод предоставляет более точную информацию, например, легче обычно получить надежную информацию о зарплате и доходах из регистров, чем спрашивая людей напрямую.

Анализ отсутствия ответов

58. В регистрах содержится много информации о людях и других единицах обследования, которые отказались от интервью или по различным причинам не были охвачены. Надежность выборочных обследований можно оценить с помощью таких анализов отсутствия ответов.

Статистика, сочетающая выборочные обследования и данные регистров

59. В некоторых случаях для составления статистики необходимо сочетать данные из административных регистров и выборочных обследований. Примером этому может служить статистика доходов домохозяйств. В странах, где нет регистра домохозяйств, данные обследования состава домохозяйств комбинируются с данными регистров доходов. Например, в датском обследовании бюджета домохозяйств и в датской части статистики по доходам и условиям жизни ЕС (EU-SILC), данные регистров в большой степени используются как дополнение к данным обследования, полученных в ходе интервью, проводимых в домохозяйствах. Используемая информация из регистров в основном содержит данные о доходах и некоторые сведения о роде занятий, образовании и жилищных условиях.

60. При комбинации данных из регистров и выборочных обследований, одним из основных методов связности/взаимосогласованности на макроуровне является *задание веса*. Примером этому может служить система переписи населения и жилищного фонда в Нидерландах, где широко используются такие методы взвешивания.

61. Административные регистры не всегда содержат все необходимые переменные для составления статистики. В таких случаях дополнительную информацию можно собрать из выборочных обследований. *Восстановление недостающих данных* на микроуровне становится при этом альтернативой комбинированию статистики на макроуровне. Этот метод заключается во вводе оценок, основанных на информации соответствующих выборочных обследований. Могут быть использованы различные виды моделей, большинство из которых включают некоторую стратификацию и случайное распределение переменной на каждой ступени. Результат таких процедур восстановления недостающих данных в основном касается макроуровня.

Глава 5. РЕГИСТРЫ – ОБЩИЙ ОБЗОР

62. В данной главе рассматривается концепция регистров, и описываются различные их виды. Кроме того, приводится описание источников административных данных, используемых для составления статистики в североевропейских странах. Конечно, существующие регистры не абсолютно одинаковы в разных странах, но их виды очень похожи. Там, где это необходимо, объяснены различия между странами.

5.1 Что такое регистр?

63. Регистр здесь определен как систематический сбор *данных на уровне единиц*, организованный так, чтобы было возможно его *обновление*. Обновление – это обработка поддающейся распознаванию информации с целью создания, внесения новых данных, исправления или расширения регистра, т.е. отслеживания любых изменений данных, описывающих единицы и их свойства.

64. Как правило, регистр содержит информацию о полной группе единиц, *обследуемом населении* (например, лицах, зданиях, фирмах). Эти единицы определены точным набором правил (например, *постоянное население в стране*), а свойства обновляются в соответствии с изменениями, которые эти единицы претерпели. Более подробно об обновлении говорится в п. 5.7.

65. Приводимое здесь определение регистров указывает исключительно на рассматриваемое нами содержание сбора данных. Информацию в регистрах можно хранить несколькими способами. Однако следует отметить, что все административные регистры, используемые в статистических целях в североевропейских странах машинные, а данные существуют в машиночитаемой форме.

66. Основное требование заключается в том, чтобы каждая единица регистра имела уникальную идентификацию. Лучше всего это достигается при использовании системы *идентификационных кодов* (ключей), но идентификация также возможна без кода, если имеется достаточная информация о единицах (для лиц это - имя, адрес, дата рождения и т. д.).

67. В идеале для статистического использования регистры должны быть максимально *полными и точными*, но даже самые неполные регистры можно использовать для статистических целей.

68. В сущности, регистром может быть список всех единиц населения и ничего больше. На практике большинство регистров также содержат некоторые дополнительные свойства для каждой единицы (данные). Термин *регистр данных* используется для различения регистров с данными от просто списков единиц. В нашем отчете термин *регистр* используется в обоих значениях.

69. *Административные регистры* – это регистры, используемые в первую очередь в административных информационных системах. Это означает, что их используют для производства товаров и оказания услуг в государственных и частных учреждениях или компаниях, или что эта информация является результатом такого производства товаров и оказания услуг.

70. В североевропейских странах большинство административных регистров, используемых для статистических целей, охватывают всю страну, ведутся государственными органами или совместно с местными органами. Однако используются и частные регистры, например, регистры страховых компаний и организаций работодателей.

71. *Статистические регистры* создаются путем обработки данных из административных регистров. Они могут иметь основой отдельный административный регистр, но чаще всего сочетание данных из нескольких административных источников.

72. Файлы данных, составляемые при традиционных переписях и других полных сборах сведений о населении, не являются регистрами. Такие *статистические файлы данных* не являются предметами обновительного процесса, т. е. когда появляются изменения, нет намерения обновлять единицы населения или информацию о единицах. Необходимо собирать всю информацию в каждой базовой точке времени или базовом периоде (в день или в год переписи). Но когда установлены файлы данных, охватывающие разные области и основывающиеся либо на данных регистров, либо на данных традиционных переписей, способ составления статистики остается тем же.

73. Поскольку административные регистры являются частью статистической системы на основе регистров, их часто называют *первичными регистрами*. Однако некоторые первичные регистры ведутся НСИ, и в первую очередь - регистры по образованию. В этом случае не ведется центрального регистра для административных целей. Статистические регистры также называют *вторичными*.

74. Такая организация статистических систем на основе регистров описывается более подробно в Главе 7.

5.2. Базовые и специализированные регистры

75. В соответствии со своим предназначением административные регистры делятся на *базовые* и *специализированные*. Термин «базовые регистры» также используется для описания статистических регистров.

Базовые регистры

76. *Административные базовые регистры* хранятся как базовый ресурс для общественного управления. Их функция обычно - вести счет населению, т. е. записывать рождения и смерти единиц и отслеживать, какие единицы присутствуют в населении в любой заданный момент времени. Кроме того, базовые регистры также хранят идентификационную информацию, которая должна использоваться другими административными регистрами.

77. *Статистические базовые регистры* основаны на соответствующих административных регистрах. Они очень важны для всей регистровой системы. Их основные цели – определять важные совокупности и содержать ссылки на другие базовые регистры.

78. На практике административный базовый регистр будет содержать больше информации, чем просто идентификацию единиц, - как правило, данные, представляющие общий интерес для ряда ведомств. Центральный регистр населения (ЦРН), например, будет иметь информацию о поле, дате рождения, семейном положении, наличии детей и т. д. по каждому человеку.

79. Базовые регистры должны предоставлять данные о различных совокупностях. Например, основное население в ЦРН – это лица, проживающие в стране. Однако ЦРН должен выдавать информацию о людях, работающих в стране, но живущих за ее пределами. Кроме того, в североевропейских странах наш опыт показывает, что даже уже неактивные единицы должны храниться в базовых регистрах в статистических целях. Это означает, что ЦРН также должен охватывать умерших и эмигрантов.

80. Во всех североевропейских странах административными базовыми регистрами являются:

- Регистр лиц (регистр населения);

- Регистр собственности (недвижимости, зданий и жилья);
- Регистр хозяйствующих субъектов.

81. В каждой из стран есть все три регистра. Кроме того, в Швеции существует четвертый регистр – рода занятий. Регистры рода занятий (работа, другие виды деятельности, связанные с рынком труда, образовательная деятельность) есть и в других странах, но они не являются базовыми (См. п. 5.6).

82. Административные базовые регистры по-разному организованы в разных странах. Например, в Финляндии данные о жилье и зданиях входят в информационную систему по населению, а в Норвегии и Дании - в ту же систему, что и недвижимость и адреса.

83. Хотя в ежедневной работе эти различия важны, они не влияют на структуру статистических систем на основе реестров по той причине, что все базовые регистры североевропейских стран имеют ключи-связки со смежными регистрами и другими базовыми регистрами. Такими ключами-связками являются:

- Личные идентификационные номера;
- Числовые адреса;
- Идентификационные номера хозяйствующих субъектов.

Специализированные регистры

84. В отличие от административных базовых регистров, специализированные административные регистры служат одной определенной цели или точно определенной группе целей. Ведомство, которое ведет такие регистры, как правило, является главным пользователем информации, и этот факт естественным образом влияет на содержание данных. Специализированные регистры часто получают информацию о населении и некоторые базовые данные из базового регистра, но другие данные поставляют сами.

85. Почти все специализированные регистры североевропейских стран используют тот же ключ идентификации, что и соответствующий базовый регистр. Таким образом, обеспечивается прямое соединение. Но не во всех случаях ясно, какой регистр является соответствующим. Например, регистр автомобилей. Автомобили можно определить как базовую единицу, но регистр не считается настолько важным, что классифицируется как базовый. Поэтому данные об автомобилях частных лиц связаны с регистром населения, а корпоративные автомобили – с регистром хозяйствующих субъектов. В таком случае регистр автомобилей в зависимости от контекста считается регистром лиц или регистром хозяйствующих субъектов.

86. Что касается базовых регистров, специализированные статистические регистры созданы на основе одного или нескольких административных регистров. Реестры, перечисленные в пп. 5.3-5.6, - административные, но существуют и соответствующие статистические регистры.

5.3. Регистры физических лиц

Базовый регистр

87. Все североевропейские страны имеют Центральный реестр населения (ЦРН) с уникальным идентификатором (личным идентификационным номером). Регистр содержит некоторую базовую информацию обо всех лицах (пол, возраст, семейное положение, гражданство и т. д.). Кроме того, включены ссылки на родителей и супруг/ов, чтобы было возможно установить единицу *семьи*.

Другая важная переменная – местожительство. В странах, где есть регистр жилья³, установлена единица *домохозяйство* (жилье-домохозяйство) путем объединения лиц и жилья.

Другие регистры физических лиц

88. Во всех североевропейских странах существует множество административных регистров, используемых для статистических целей, все из них используют личный идентификационный номер из ЦРН. Их можно найти почти во всех сферах. Это:

- Налоговые регистры;
- Доходы и благосостояние;
- Заработная плата рабочих, управленческого и обслуживающего персонала;
- Регистры рынков труда;
- Место работы;
- Безработица, обучение на рынке труда;
- Регистры уровня образования;
- Регистры студентов;
- Регистры системы социального обеспечения;
- Регистры состояния здоровья;
- Регистры преступности;
- Регистры автомобильных средств.

89. Данный список содержит самые важные регистры физических лиц, но не является полным.

90. Регистры рынков труда и студентов можно также классифицировать как *регистры рода занятий* (см. п. 5.6).

5.4. Регистрация собственности

91. Административные регистры собственности несколько по-разному организованы в североевропейских странах, тем не менее, у них много схожих черт. Дания и Финляндия имеют регистры жилья, действующие уже много лет и используемые в статистических целях. В Норвегии полный регистр жилья был создан в 2001 г., и развивается статистика жилья на основе регистров. В Швеции создание регистра жилья планируется в настоящее время (2006 г.).

Регистры недвижимости

92. Эти регистры содержат информацию о недвижимости и земельных участках, а также о правах на собственность и залогах, местонахождении недвижимости и идентификации ее собственников.

Регистры зданий и жилья

93. Регистры зданий и жилья можно рассматривать как продолжение регистров недвижимости. В них указаны год постройки, площадь, устройство и использование зданий и жилья. Регистры обновляются за счет отчетов об изменениях, новых постройках и сносе.

Адресные регистры

94. Эти регистры содержат географические сведения обо всех адресах, например, указание административных единиц (муниципалитета и т. д.), а также другие географические регионы, к которым принадлежит адрес, расположение адреса (в густо- или малонаселенном пункте). Кроме того, в них находятся географические координаты адреса.

³ В 2006 г. все страны, кроме Швеции, где планируется создать регистр жилья.

Взаимосвязанность регистров

95. Регистры недвижимости, зданий и жилья и адресов *взаимосвязаны*. Таким образом, можно получать данные о точном местонахождении зданий и жилья, включая географические данные.

96. Числовой адрес является ключом-связкой между *реестром населения* и реестром зданий и жилья, а также между *реестром хозяйствующих субъектов* и реестром зданий. Таким образом, можно узнать точное местожительства и место работы.

5.5. Регистры хозяйствующих субъектов

Базовый регистр

97. Все североевропейские страны создали центральные регистры, охватывающие единицы всех типов деятельности, включая центральные и местные единицы государственного управления. Компании и учреждения являются основными видами единиц, для них в регистре созданы уникальные идентификационные номера. Эти номера используются в большинстве административных регистров хозяйствующих субъектов.

98. Центральные регистры хозяйствующих субъектов содержат основную информацию о единицах: форму собственности, где возможно – идентификацию собственника/ов, отрасль, местонахождение и размер.

99. В последние года произошли существенные улучшения в системе регистрации хозяйствующих субъектов. В качестве примера мы будем использовать регистры хозяйствующих субъектов Норвегии.

Пример: Регистрация хозяйствующих субъектов в Норвегии

100. До середины 1990-х гг. существовало несколько реестров хозяйствующих субъектов. Они не были согласованы с точки зрения единиц, идентификационных номеров и совокупностей и обновлялись по отдельности. В 1995 г. был создан Центральный координационный регистр юридических лиц. Он собирает всю основную информацию в единую систему. Всем единицам, независимо от своего предназначения, присваивается идентификационный номер. Это похоже на личные идентификационные номера, присваиваемые Центральным реестром населения. Единицы, зарегистрированные в Реестре работодателей, Реестре НДС, Корпоративном регистре, Губернаторском регистре фондов и Центральном регистре хозяйствующих субъектов ЦСБ Норвегии обязаны по закону зарегистрироваться в Центральном координационном регистре юридических лиц.

101. Использование единых номеров облегчает предоставление данных и регистрацию, поскольку отдельная единица должна сообщать о каждом изменении только в один регистр. Регистр, получающий информацию, обязан затем предоставить ее в Центральный координационный регистр юридических лиц для обновления.

Другие реестры хозяйствующих субъектов

102. Во всех североевропейских странах имеется несколько административных регистров хозяйствующих субъектов, самые важные из которых находятся в ведении налоговых властей (например, регистры налогообложения юридических лиц, регистры налога на добавочную стоимость, регистры вкладов работодателей и годовые контрольные данные работников по найму). Другой крупный источник данных – торговые регистры. Существует также ряд регистров для подгрупп совокупности для регистра хозяйствующих субъектов, например, регистр школ.

103. Существование уникальных идентификационных номеров облегчает связку с коммерческими данными различных регистров. Например, данные о товарообороте можно получить из регистра НДС, количество работников – из регистра работников по найму и регистров заработной платы.

5.6. Регистры рода занятий

104. Основной тип регистров рода занятий – *регистры места работы*. Не существует одного административного регистра, охватывающего все виды работ, поэтому статистические регистры места работы зависят от различных административных источников в различных странах. Их главные типы – налоговые регистры, различные регистры работников по найму и регистры пенсионных систем. Дополнительную информацию об особых группах работников можно собрать из других административных регистров, например, регистра сотрудников центральных и местных органов власти, учителей и призывников на военную службу.

105. Единица *работа* имеет два идентификационных ключа: номер хозяйствующего субъекта (для учреждения или компании) и личный номер (для работников по найму). Характеристики работника по найму (пол, возраст, место жительства, образование и т.д.) и хозяйствующего субъекта (отрасль, местонахождение и т.д.) в этом случае заносятся в регистр места работы.

106. Статистика занятости основывается на единице *лицо*. В качестве характеристики работающего лица в этом случае выбирается основная работа, указанная в статистическом регистре рода занятий (в заданный момент времени или период).

107. Регистры места работы также используются в статистике хозяйствующих субъектов. Учреждения и компании характеризуются количеством работников по найму и другими сведениями о кадрах (например, пол, возраст и распределение по образованию).

108. *Регистры периодов незанятости, обучения на рынке труда, регистры студентов* – другие примеры регистров рода занятий. Регистры студентов содержат личный идентификационный номер, номер учебного заведения как хозяйствующего субъекта и дополнительную информацию об учебе (например, тип образования, полная/частичная учебная нагрузка).

109. Только в Швеции регистры рода занятий считаются *базовыми*, но они очень важны во всех североевропейских странах.

Пример: Четыре статистических базовых регистра в Швеции

110. ЦСБ Швеции разграничивает три административных базовых регистра и четыре статистических базовых регистра. Совокупность рода занятий и других смежных видов деятельности считается важной целевой совокупностью, которую следует проанализировать, например, демографическими методами. Это, а также факт, что множество административных данных связано с деятельностью, является причиной того, что Регистр рода занятий считается в Швеции статистическим базовым регистром.

5.7. Обновление регистров

111. Основным принципом регистров является возможность регулярно обновлять их, чтобы охватывать все значимые единицы и иметь самые последние данные, описывающие единицы. Поэтому в регистры регулярно добавляются новые единицы, входящие в целевую совокупность (новорожденные и иммигранты в ЦРН), а единицы, покидающие целевую совокупность, «стираются» (точнее классифицируются как «неактивные»). Данные, описывающие единицы, тоже могут

изменяться и поэтому тоже должны обновляться. Для лиц изменяться могут семейное положение, количество детей, уровень образования, профессия и т.д. Может измениться площадь здания, владелец автомобиля и т.д.

112. В зависимости от события, которое повлекло за собой изменение, обновление может происходить с большей или меньшей задержкой. В некоторых случаях оповещение может напрямую не совпадать с событием, так что достаточно выяснять точные данные через определенные промежутки времени. Данные, не поддающиеся систематическому обновлению, не считаются правильными данных регистров.

113. Основным принципом ведения регистров является сбор всех необходимых данных, когда они появляются, и только один раз и в связи с их появлением. Это эффективно как с экономической точки зрения и для недопущения дублирования в регистровой системе. Действительно, самое большое преимущество хорошей регистровой системы – это то, что обработка ограничивается теми единицами и теми свойствами, которые действительно *изменились*. Единицы и свойства, которые остались неизменными, не обрабатываются совсем. Чем более стабильны данные, содержащиеся в регистре, тем меньше работы требуется по его обслуживанию.

Пример: Порядок обновления информационной системы по населению Финляндии

114. В Информационную систему по населению входят данные обо всех постоянно проживающих в стране, т. е. каждом лице с постоянным местожительством в Финляндии, и обо всех зданиях и жилье. Информационную систему по населению обслуживает **Финляндский регистр населения** и **37 местных отделений**.

115. Как видно на схеме, обновлением Информационной системы по населению в основном занимаются несколько органов власти. Люди лишь должны уведомить один из 37 местных отделений о смене адреса в течение семи дней после переезда. Личные данные всегда обновляются при указании личного идентификационного номера. Данные о зданиях и жилье тоже обновляются на основании кода здания или жилого помещения.



Глава 6. ТРЕБОВАНИЯ К ДАННЫМ ИЗ АДМИНИСТРАТИВНЫХ ИСТОЧНИКОВ

116. Общие условия использования административных данных в статистических целях рассматриваются в Главе 3, использование административных данных и комбинированное использование регистров и обследований - в Главе 4. Данная Глава более подробно описывает некоторые требования, которые необходимо выполнять, чтобы официальная статистика в оптимальной мере использовала административные регистры.

Содержание

117. Прежде всего, административные регистры должны содержать данные об основных предметных сферах какой-либо статистической системы, чтобы проследить тенденции развития общества. Важным предварительным условием для статистики на основе регистров, имеющей полный охват, является то, чтобы данные были подробными и охватывали много переменных, связанных со значимыми единицами. Недостаточность данных либо делает необходимым дополнительный сбор информации с помощью традиционных методов, либо ограничивает содержание статистики.

Единицы и идентификаторы

118. Для структуры статистики три единицы представляют первоочередное значение: лица, фирмы/учреждения и жилье. Это отражено в организации соответствующих базовых регистров, о которой говорилось в Главе 5. Конечно, существует множество других единиц, которые важно регистрировать. Например, *события*, связанные с лицами, такие как демографические события (рождения, смерти, заключение брака и т.д.).

119. Для статистического использования важно, чтобы единицы были четко определены. Для таких «естественных» единиц, как лица или автотранспортные средства, это не проблема. Так же легко определить концепцию дома или жилья, хотя есть несколько пограничных случаев, в которых наличие дома и его характеристики можно определить только с помощью подробных правил. Некоторые единицы проявляются как сочетание естественных единиц, например, семей и домохозяйств.

120. Более проблематично рассмотрение других единиц – таких, как единиц хозяйствующих субъектов. Фирма или рабочее место – абстрактные понятия. Можно зафиксировать их существование и определить их сущность, только соблюдая точные правила. Например, исходя из административных данных, трудно заключить, открыта ли новая компания или претерпела крупные изменения уже существовавшая.

121. В идеале регистры должны содержать информацию обо всех единицах определенного типа в стране, т. е. должны быть полными. Неполный охват часто будет систематическим. Например, если не были охвачены некоторые муниципалитеты, составленная статистика будет немедленно подвергнута смещениям, которые, возможно, будет трудно компенсировать.

122. *Идентификаторы* играют весьма существенную роль в ведении административных регистров и их статистическом использовании, особенно при сведении информации из различных источников. В идеале идентификационные коды не должны меняться на всем протяжении существования единицы. Лучше всего достигнуть этого, используя коды, не содержащие информации. Когда личный идентификационный номер содержит дату рождения и пол, в течение жизни у его владельца могут произойти изменения. Чаще всего причина этому – исправление даты рождения, в основном иммигрантов, в ЦРН. Проблемы также могут возникнуть, когда при переезде в страну человек

получает сначала временный номер, а позднее – постоянный личный идентификационный. В таких случаях НСИ должен создать связь между идентификационными кодами, которые указывают на одно и то же лицо. Решением проблемы может стать замена административного идентификационного номера статистическим идентификатором в статистических регистрах.

Указание времени

123. Очень важную роль в статистике играет фактор времени, раскрывающий тенденции развития общества. Во всех областях необходимо проводить сравнения по времени. Поэтому для практичности использования данных совершенно необходимо верное указание времени в регистрах.

124. Важнее всего *даты изменений или событий*. Наибольший интерес представляют «рождения» и «смерти» единиц, но и другие изменения, касающиеся единиц, также важны. Мы должны знать реальный момент времени, когда произошло событие, повлекшее за собой изменение данных, например, дату закрытия или дату смены отрасли деятельности коммерческой компании.

125. На практике часто оказывается трудно определить точное время события. Вид деятельности компании – это переменная, которая может меняться постепенно, пока компания меняет свой профиль, и часто бывает невозможно сказать, когда новый вид деятельности стал основным. Но даже и для переменных, которые в принципе можно наблюдать в любое заданное время и изменение которых, следовательно, можно датировать достаточно точно, могут возникать практические проблемы в процессе регистрации. Это может происходить из-за того, что дата события не обязательно будет важна для управления использованием данными: все, что нужно знать, – это самое последнее значение в любое заданное время. В других случаях проблемы могут возникать из-за того, что невозможно или непрактично регистрировать события в режиме реального времени. Может быть, возможно будет выявить изменения данных на основании ответов обследований, например, раз в год в связи с подачей налоговых деклараций.

126. Кроме даты событий, важны *даты регистраций*, т.е. указание времени, когда соответствующие данные занесены в регистр. В идеале любая информация в административном регистре должна иметь две даты. На практике регистры часто не существенно отклоняются от этого идеала.

127. Если нет возможности точно зарегистрировать даты, конечно, остается возможность приближений, что куда лучше, чем отсутствие каких-либо данных вообще.

Данные о событиях

128. В некоторых случаях требуется иметь возможность рассмотреть изменения данных, зарегистрированных в регистре, как модель событий, описанных в статистике. Это касается, например, статистики по населению, в которой должна иметься возможность получать сообщения об изменениях (переезд, заключение брака и т.д.) из измененных данных в Центральном реестре населения (ЦРН). В целях статистики по миграции, например, должна иметься возможность различать действительные события (переезды) в регистре от исправлений и изменений данных более технического свойства.

129. Исправление ошибочно зарегистрированных данных является для статистики проблемой, даже если возможно отличить их от действительных событий. Во многих случаях исправления должны считаться модификацией ранее зарегистрированных событий: исправление адреса может быть исправлением сообщения о переезде. Исправление может также отменить событие, сообщенное ранее, что происходит в системе ЦРН особенно, когда человек уведомил регистр населения о своем намерении эмигрировать, но передумал и остался в стране.

Стабильность

130. Важной задачей статистики является описать процесс во времени, т. е. показать, как отдельная величина изменяется из месяца в месяц и из года в год. Поэтому чрезвычайно важно, чтобы понятия в административных регистрах оставались постоянными долгое время, иначе могут появиться крупные проблемы при переводе сопоставимых цифр из одного периода в следующий. В некоторых, хотя и далеко не во всех, случаях можно учитывать изменения с большей или меньшей точностью. Парадоксально, что проблема может появиться даже, если регистры стали более верными, потому что даже в этом случае мы можем получить разрыв данных. Большие проблемы могут также появляться, когда вносятся изменения в законы или нормативные документы, приводящие к изменениям в содержании данных административных регистров. С одной стороны оценить долгосрочную тенденцию отдельной величины может быть трудно или невозможно, если в базовом материале использованы различные определения. С другой – проблемы могут появляться при решении, какие изменения показателей данных должны рассматриваться как отражающие действительные события, а какие – представляют просто новые концепции или определения.

131. Статистические последствия зависят от типа прилагаемой статистики. Если это фактически *статистика мониторинга законодательства*, задачей которой является показать, как управление каким-либо законом влияет на обыкновенных людей, то статистика, конечно, просто будет учитывать и принимать понятия нового законодательства. *Общая статистика*, с другой стороны, стремится объяснить некоторые понятия, определения которым не даны в законодательстве, например, безработица – понятие, чье определение необходимо найти в международной конвенции. Если статистика касается выплаты социальных пособий, а правила выплаты изменились, может быть, будет трудно или невозможно компенсировать это изменение в статистике. Но нужно, по крайней мере, попытаться оценить значимость изменений, чтобы до определенной степени выстроить цепь из временных рядов.

132. Существуют также примеры административных регистров, претерпевающих изменения без каких-либо изменений основных правил. Типичными примерами служат изменения, вызванные техническими причинами или рационализацией. Может возникнуть проблема, если, например, принято решение о нецелесообразности дальнейшего использования для административных целей какого-либо определенного вида данных.

133. Ясно, что описанные изменения имеют серьезные последствия для статистики, если они осуществлены без достаточного учета требований. Поэтому важно, чтобы статистики участвовали в подготовке к изменениям, чтобы по крайней мере можно было оценить их последствия прежде, чем будут внесены изменения.

Качество

134. Требования к качеству, предъявляемые статистическим использованием, совпадают до определенной степени с требованиями, которые также необходимо соблюдать для достижения целей ведения регистров; информация должна быть верной и регистрироваться с достаточной точностью. Кроме того, данные должны представлять интерес для статистики. Аспекты качества рассматриваются в Главе 9.

Сотрудничество с хранителями регистров

135. При составлении статистики на основе административных данных невозможно осуществлять такой же контроль над содержанием базовых данных, как при составлении статистики на основе анкет. Мы не можем быть уверены, что регистры охватывают необходимые единицы с той же степенью точности или что данные определены в соответствии с потребностями пользователей.

136. Конечно, для статистиков желательно оказывать определенное влияние на содержание данных, но не следует забывать, что регистры ведутся для очень специфических административных целей и что задача хранителей регистров – служить этим целям как можно лучше. Поэтому они не могут уделять слишком много внимания потребностям, которые служат «только» статистическим целям. По этой причине статистики не должны составлять слишком широкие требования к дополнительным данным, различным определениям и т. п. и ожидать выполнения этих требований. Хранители регистров должны пристально следить за эффективностью своей собственной работы и использованием ресурсов. Статистики должны быть непритязательными и выставлять только главные требования, когда можно ожидать, что поправки регистра принесут очень значительную пользу обществу. В североевропейских странах известно очень немного случаев сбора дополнительных данных путем использования регистров исключительно в статистических целях.

137. Национальный закон о статистике является полезным инструментом для хорошего сотрудничества с владельцами регистров (см. 3.1).

Глава 7. ОТ АДМИНИСТРАТИВНЫХ ДАННЫХ К СТАТИСТИКЕ – ОБРАБОТКА ДАННЫХ

138. Перед тем, как административные данные можно будет использовать в качестве статистики, они должны пройти статистическую обработку, которая интерпретирует их в соответствии с целью. В этом процессе сведения упорядочиваются в *статистические регистры*, которые являются собранными данными с идентификационными ключами, в которых генеральная совокупность и содержание данных каждого из регистров соотносятся с тем, чтобы они удовлетворяли требованиям определенной статистической области. Данные в статистическом регистре могут быть получены из одного или нескольких административных регистров.

7.1. Редактирование

139. Одним из самых важных и трудоемких процессов в создании регистров является проверка внесенных исходных данных и исправление ошибок, также называемая редактированием. Обычно данные сначала проверяются ответственными за данный административный регистр органами власти и затем статистическим бюро.

140. Целью редактирования, производимой хранителями регистра, является улучшение качества данных для административного пользования. Следовательно, в центре внимания находятся переменные, являющиеся наиболее важными для работы администрации. Переменные, включенные в регистр большей частью для стратегических целей, могут не контролироваться так же тщательно. Чрезвычайно важно, чтобы статистикам было хорошо известно об этих методах редактирования и накладываемых ими ограничениях.

141. В редактировании, производимой НСИ, в центре внимания находится статистическое использование данных. Для устранения ошибок каждый регистр сначала проверяется при помощи внутренних логических проверок. В данном процессе не используются внешних данных. Проблемы охвата и прочие неточности также возможно обнаружить путем совмещения данных из нескольких регистров (подробнее об использовании множественных источников в п. 7.2).

142. Статистический регистр в идеале должен быть «чистым» и целостным, без противоречий между отдельными данными. Маловероятную информацию необходимо проверять. Нет ничего необычного в обнаружении системных ошибок в административных регистрах или выдержках, сделанных поставщиком до получения их статистическим бюро.

143. Ошибки могут состоять в неспособности информации соответствовать *формальным требованиям* для данного спорного типа информации. В подобных случаях, когда ошибки часто появляются группами, необходимо принять решение, стоит ли менять спорное значение так, чтобы информация соответствовала допустимым интервалам или исключить информацию из статистической базы, поскольку нет возможности внести обоснованное исправление.

144. Отдельные ошибки в определенных случаях можно обнаружить и исправить при непосредственной ссылке на исходные регистры. Регистры охватывают генеральную совокупность, так что возможности для связи с отдельными единицами весьма ограничены. Такое положение действительно имеет место, когда единицами являются частные лица. Когда мы имеем дело с регистрами предприятий, представляется возможным сбор дополнительной информации посредством связи с отдельными предприятиями, но это весьма трудоемкий процесс. Поэтому основным методом проверки и исправлений является автоматизированная логическая редактирования данных.

145. Временные ссылки также необходимо обрабатывать для обеспечения наилучшей статистической модели для совокупности единиц и данных о ней в искомые моменты или периоды времени. Административные регистры обычно не соответствуют подобной модели, так как их основной целью обычно является обеспечение свежих данных по всем сферам деятельности.

146. Решающим является исправление всех систематических типов ошибок, встречающихся в административных регистрах, данные из которых извлекаются для статистических целей. Достичь этого можно только при тесном контакте с теми органами власти, которые несут ответственность за данный регистр. Известно много случаев, когда типы ошибок, обнаруженные статистиками, были впоследствии исправлены в административных регистрах, что благотворно сказывалось и на регистре, и на статистике.

7.2. Использование множественных источников

147. При сборе данных для статистических целей традиционными методами вопросы и рекомендации по ответам на них сформулированы таким образом, чтобы полученная информация соответствовала используемым в рассматриваемой статистической области определениям.

148. Использование административных регистров в качестве основы для статистических данных не обеспечивает одинаковый объем для контроля. Можно осуществлять поиск в доступной информации и выбирать те элементы, которые лучше всего подходят для намеченной цели. Однако часто даже самая подходящая информация не удовлетворяет непосредственным нуждам в достаточной степени. Причиной этому могут служить расхождения в определении, ограниченный охват, ошибки или неполноценное обновление. В такой ситуации одним из решений является использование информации из нескольких административных источников. По этой причине создание статистических регистров часто включает в себя компоновку исходных данных на элементарном уровне. Основываясь на различных данных, необходимо попытаться найти наилучшие оценочные функции для статистической переменной.

149. Там, где невозможно удовлетворить потребность в определенном виде информации об одной или нескольких статистических функций непосредственным обращением к имеющейся в распоряжении информации, задачу можно решить построением *производных переменных*. Переменная подобного вида получается формированием цепочки из различных единиц информации, что обеспечивает необходимую информацию, если не с абсолютной точностью, то, по крайней мере, с достаточно хорошим приближением.

150. При введении производной переменной, отправным пунктом является формулировка понятия, которое желают использовать в статистике. Затем осуществляется поиск административной информации, которая может способствовать удовлетворению сформулированных требований. Можно выделить три типа создания переменной: *создание новых переменных, установление приоритета информации для заданной переменной и уточнение или корректировка информации*.

Создание новых переменных

151. Создание новой переменной может быть возможным решением, если мы располагаем необходимой информацией в достаточной мере, но сама она не содержится в отдельной единице административных данных. Несколько единиц информации, собранных и обработанных в соответствии с административными правилами, могут способствовать обеспечению комплексного описания отдельных лиц по отношению к требуемой статистической концепции.

152. При создании новой переменной, существенным является качественное наличие необходимых компонентов для всех единиц с точки зрения намеченной цели.

Примеры производных переменных в датской статистике

153. *Пример 1:* Информацию о владении предприятием можно найти только для лиц, включенных в Центральный регистр хозяйствующих субъектов. Это работодатели и плательщики НДС. Небольшая группа владельцев предприятий не входит в данную совокупность, например, лица, занятые в транспортной сфере деятельности и поставщики определенных услуг в сфере здравоохранения. В этом случае используются сведения, собранные в соответствии с обязанностью предоставлять бюджет, и большее значение придается доходам или убыткам, зарегистрированным предприятием.

154. *Пример 2:* Отделение заработной платы от прочих источников дохода, облагаемых точно так же, как и заработная плата, находится в некоторой степени под влиянием того факта, что для налоговых органов данное разделение не имеет большого значения при оценке суммы обложения. Однако можно сделать вывод что, несмотря на указанные недостатки, результат хорошо сочетается со сформулированным определением.

Установление приоритета информации из множественных источников

155. Существуют ситуации, когда необходимую информацию можно найти в разных административных регистрах с неравномерным охватом и/или переменным качеством. Неравномерный охват возникает вследствие того, что определенные регистры созданы для выполнения специфических административных функций, связанных с менее четко определенной частью населения. Качество информации зависит от ее значимости по отношению к той административной цели, которой она служит.

156. Например, переменную *текущий род занятий* (имеет работу, безработный, не относится к трудоспособному населению) невозможно собрать из одного регистра. Для лиц, осуществляющих более одной деятельности одновременно, занятость по статистическому определению будет иметь приоритет. Однако если сведения по безработице считать более достоверными, по крайней мере, там, где рассматриваются временные ссылки, тогда лиц, зарегистрированных как имеющие работу и безработных одновременно, следует относить к безработным.

Исправление информации

157. Исправление информации в регистре производится по мере возможности при возникновении ошибок или дефектов. В известной степени, подобные исправления можно внести, исходя из предполагаемых связей между единицами информации одного источника, однако может возникнуть необходимость использовать сведения из других источников.

158. Внесение исправлений на основе дополнительной информации из других источников может обеспечить полезной информацией в тех областях, где доступный материал предрасположен к неточностям или является неполным в других отношениях. Однако вносить систематические исправления, не рассмотрев внимательно ограничения применяемого метода, может быть опасно.

Использование совмещенных данных на протяжении времени

159. Использование совмещенных данных часто может потребовать очень сложного набора правил для обработки информации. Может оказаться сложно поддерживать полную картину происходящего, что может привести к тому, что важности отдельных элементов будет уделяться недостаточно внимания. Изменение исходной информации, которое при обособленном рассмотрении кажется относительно несущественным, может иметь серьезные последствия при обработке совмещенной информации.

160. Важно, чтобы набор правил для обработки был обстоятельно описан с четким указанием того, какие сведения нужно использовать, как их необходимо использовать и при каких условиях. Кроме того, важно вводить поправки на возможные изменения в содержании информации для того, чтобы правила продолжали действовать при возникновении подобных изменений.

161. Узнать заранее обо всех существенных изменениях может быть сложно, поэтому крайне важно тестировать обработку в процессе, например, при помощи проверки того, что различные условия влияют на конечный результат в той же мере, как и ранее. Отклонения сами по себе могут и не указывать на какие-либо изменения, однако они указывают на возможную необходимость более тщательного исследования.

162. Использование производных переменных, созданных на основе информации из различных источников чрезвычайно чувствительно к административным переменам. Использование множественных источников повысит риск возникновения изменений соответствия статистической цели. Об изменениях может быть известно, однако ввести на них поправку не всегда возможно, например, если пропущена единица информации. В этом случае может возникнуть необходимость скорректировать концептуальное описание для того, чтобы привести его в соответствие с существующими обстоятельствами, что означает отклонение от ранее использованного описания и, как следствие, нарушение целостности данных.

163. Административные изменения также могут иметь положительные стороны в том смысле, что возникает новая информация или уже существующая информация уточняется для того, чтобы соответствовать статистическим требованиям. Конечно, улучшения вызовут неоднородность данных, однако выгода от усовершенствованной базы может более чем компенсировать неудобство при нарушении целостности данных.

Подача статистики, основанной на множественных источниках

164. Действующие определения, связанные с производными понятиями часто оказываются довольно сложными, так как они отражают набор применяемых правил. Поскольку это может ввести в растерянность лиц, пользующихся статистикой, сведения предоставляются в упрощенной форме, как и в случае с традиционными совокупностями данных. При использовании информации из множественных источников можно получить более дифференцированную картину. Например, регистр налогообложения может предоставить сведения о заработной плате отдельного лица, а регистр хозяйствующих субъектов может показать, что тот же самый человек выступал и в качестве владельца предприятия. Набор правил должен определить способы взвешивания информации, стоит ли относить такого человека к наемным работникам или индивидуальным предпринимателям, исходя из его статуса занятости. Для наглядности основная деятельность лица указывается так же, как и его статус занятости, указанный в регистре. Конечно, если для особых исследований нужен анализ разграничения между различными группами, можно предоставить подробности.

7.3. Разработка статистической системы на основе регистров

165. Во всех североевропейских странах была разработана информационная система на основе регистров. Следовательно, все статистические регистры в принципе рассматриваются в качестве части одной и той же системы, а не обособленно. Такой подход отражается на всех стадиях статистического производства: при сборе и обработке данных, контроле качества и распространении. Краеугольными камнями являются наличие базовых регистров, отслеживающих целевую группу и возможность взаимосвязывания всех основных регистров. Таким образом, различные предметы изучения статистики можно частично основывать на одних и тех же источниках данных. Например, демографические переменные создаются в статистическом регистре населения и используются во всех видах социальной статистики.

166. На практике же это не означает, что все переменные связаны со всеми другими доступными в статистической системе переменными. При использовании специализированных регистров для статистических целей они всегда связаны с соответствующим регистром или регистрами. Связывание специализированных регистров осуществляется при целесообразности данной процедуры. Например, если переменные никогда не будут публиковаться или использоваться иным способом в сочетании, в связывании нет необходимости. Однако, в общем, комплексное использование данных регистров улучшит качество всех привлеченных данных.

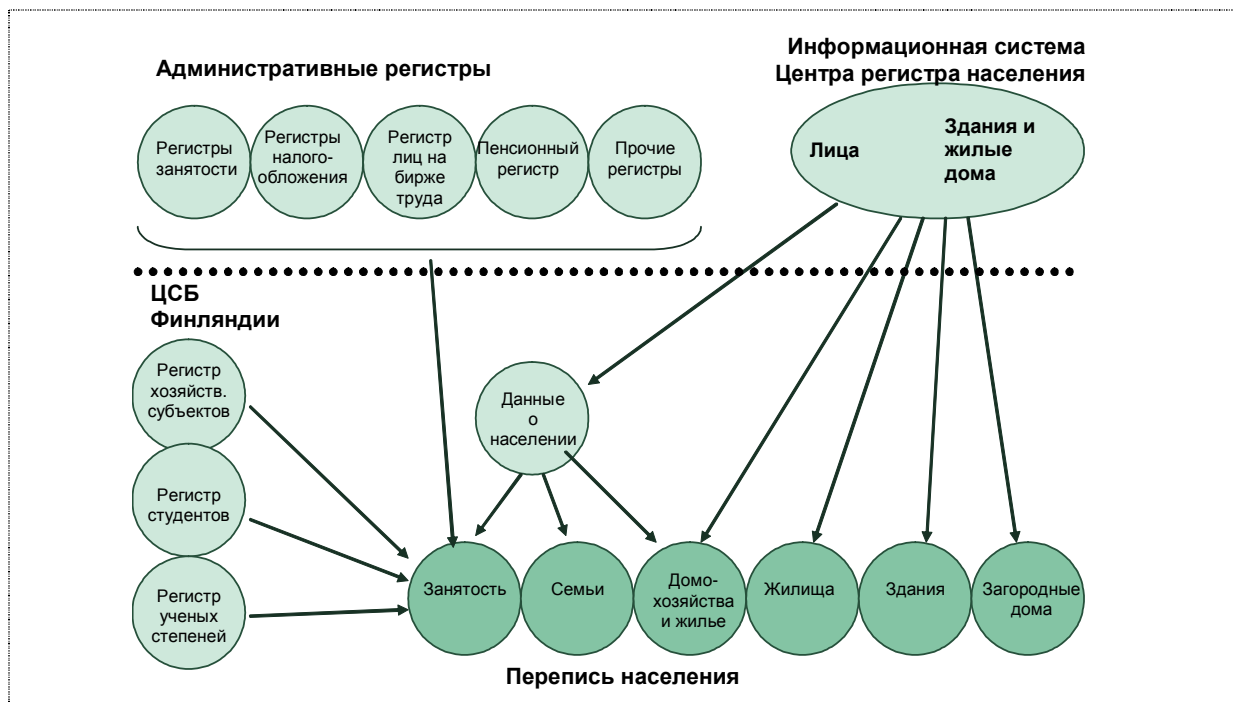
167. Как указывается в Главе 2, разработка системы данных на основе регистров во всех североевропейских странах велась в течение длительного времени. Статистические регистры созданы в отдельных областях, и к 2011 году во всех североевропейских странах, по крайней мере, согласно государственному плану, будет введена система переписи населения и жилищного фонда, полностью основанная на регистрах. И даже если мы назовем ее системой переписи, те же самые источники используются в соответствующем предмете статистики.

168. Статистическую систему, основанную на регистрах, никогда не следует считать завершенной. По мере возникновения новых потребностей у пользователей статистики и формирования новых административных регистров свежую информацию следует интегрировать в систему.

Система переписи населения и жилищного фонда, основанная на регистрах

169. Перепись населения и жилищного фонда представляет собой лучший пример использования регистрационных данных в производстве статистики. Здесь в качестве основного примера мы используем систему Финляндии.

Пример: Система переписи населения и жилищного фонда, основанная на регистрах в Финляндии



170. Базовые регистры охватывают относящиеся к переписи статистические единицы: *лиц*, проживающих в стране, *здания и жилье* в стране, а также предприятия и их учреждения (регистр хозяйствующих субъектов). Все статистические единицы можно связать друг с другом посредством систем идентификации: *лица* можно связать с *семьями* и *домохозяйствами*, с *жилым домом* или *зданием*, в котором они живут, и с *работодателем*, на которого они работают. Таким же образом можно установить местоположение всех единиц на картах при помощи географических координат, потому что все здания имеют координаты.

Данные переписи и статистика предметных областей из тех же самых источников данных

171. Система переписи населения на основе регистров содержит в себе несколько статистических подсистем, включая следующие:

- Демографическая статистика;
- Структура населения;
- Изменения в населении;
- Статистика по семьям и домохозяйствам;
- Региональная статистика занятости;
- Статистика доходов;
- Образовательная статистика;
- Статистика законченного образования и ученых степеней;
- Студенты;
- Статистика жилищного фонда;
- Статистика жилищных условий.

172. Это означает, что пользователи получают ежегодное (а в некоторых сферах даже ежемесячное или ежеквартальное) обновление большинства переменных, включенных в статистические материалы переписи.

173. В данном разделе мы более подробно остановимся на двух подсистемах: региональной статистики занятости и статистике зданий и жилых домов. Оба примера касаются Финляндии.

174. Статистика занятости, основанная на регистрах, принадлежит к числу самых важных статистических систем во всех североевропейских странах. Основной целью данных статистик является предоставление ежегодных региональных данных об экономической активности населения: они включают такие пункты как основной вид занятий, состояние занятости, промышленность, местонахождение рабочего места и профессия.

175. Первый набор статистических данных по занятости, основанный на данной системе был произведен в 1980 году в Дании, а в течение 80-х гг. — в Финляндии, Норвегии и Швеции. При составлении годовой статистики занятости используется ряд различных регистров. Большинство из них нужны для определения основного вида деятельности в контрольный день (например, конец года) или в течение контрольной недели, которая в действительности является первоочередной задачей статистики занятости.

Пример: Основанная на регистрах статистика занятости в Финляндии

176. Система Финляндии выбрана в качестве примера, потому что системы североевропейских стран отличаются друг от друга лишь в незначительной степени.

177. Данные о рабочих местах получают из ряда различных источников. Существует три основных типа трудовых отношений: трудовые отношения в частном секторе (включая индивидуальных предпринимателей), государственном секторе и органах местного самоуправления (включая муниципальные объединения).

178. При сборе сведений из различных административных источников, составляется отдельный регистр всех рабочих мест текущего года. Кроме того, к регистру хозяйствующих субъектов и другим регистрам учреждений государственного сектора обращаются за сведениями об отрасли промышленности, форме собственности и т.д. Для людей, работающих на предприятиях с несколькими хозяйствами, сведения о месте работы касаются только последней недели года, основываясь на отдельном исследовании, проводимом ЦСБ Финляндии.

179. Сведения о деятельности, отличной от обычной работы по найму, собираются из нескольких административных регистров: регистра безработных и участвующих в программах по трудоустройству, регистра призванных на военную службу и пенсионного регистра. ЦСБ Финляндии также формирует регистр студентов из множества источников.

Пример: Статистика жилья и домохозяйств Финляндии

180. Данные о строительстве и жилье используются для формирования статистик *строительного производства, строительства, жилищного фонда и жилищных условий*. Кроме того, данные регистра используются в качестве основы выборки для различных исследований касательно жилых помещений и домашних хозяйств, таких как исследования уровня арендной платы и бюджета домохозяйств.

181. [Illegible text]

182. Человек и жилье связаны с помощью *кода места жительства*, который состоит из кадастрового номера, номера здания и кода жилья. Единицы «домохозяйство-жилье», сформированные СЦБ Финляндии, связаны с жилищными данными на основе регистра через код места жительства. Данные налогового учета для единицы «домохозяйство-жилье» связаны с набором данных о ней.

183. *Состояние населенности жилья* обновляется на основе уведомлений о смене адреса проживания, а также на основе регистра гербового сбора, взимаемого при покупке акций жилищно-строительных кооперативов.

184. Выборочная совокупность для *статистики жилищных условий* состоит из всех лиц, постоянно проживающих в собственном жилье. Лица, проживающие в учреждениях и пансионатах, не принадлежат к данной совокупности. Лицам, переезжающим в пансион или учреждение, а также тем, кто становится бездомным, присваивается код места жительства в зависимости от муниципалитета, в котором они проживают, который также включает отдельные коды для сельских и городских районов. Эти лица удаляются из файлов «домохозяйство-жилье».

Глава 8. РАСПРОСТРАНЕНИЕ И ПОТРЕБНОСТИ ПОЛЬЗОВАТЕЛЕЙ

185. Преимуществом данных на основе регистров является предоставление информации обо всей изучаемой совокупности на регулярной основе (ежегодно, а в некоторых случаях даже ежемесячно или ежеквартально). Следовательно, данные на основе регистров очень хорошо подходят для производства статистики малых групп, как *малых социальных групп*, так и *малых территориальных единиц*. В этой главе мы приведем примеры из статистики малых территориальных единиц.

186. При *текущем учете* сведения об одном том же лице за последовательные промежутки времени связаны воедино (лонгитюдные данные), что позволяет следить за статистическими единицами во времени. Примером этому может служить статистика, описывающая переход от получения образования к трудовой жизни. Регистры особенно удобно использовать в качестве источников для текущего учета, поскольку они содержат все население в любой момент времени, например, ежегодно.

187. Лонгитюдные данные также очень ценны для *исследовательских целей*. Ряд регистров, например регистр населения, содержит даты изменений и событий. Таким образом, возможно установить точные даты переходов или продолжительность состояний и деятельности. Во всех североевропейских странах сформированные на основе регистров микроданные доступны исследователям.

8.1. Географические информационные системы и статистика на основе регистров

188. Самыми важными региональными единицами в статистике традиционно были и остаются административные единицы. Система зонального кодирования административных единиц явилась основой для региональной статистики. Впрочем, управление динамично по своей сути и не перестает изменяться и, следовательно, границы административных единиц также подвержены постоянным изменениям. Кроме того, многим пользователям нужны сведения о более мелких территориальных единицах или районах, пересекающих границы муниципалитета.

189. Все североевропейские страны разработали *точечную* статистическую систему, в которой расположение зданий определяется географическими координатами, указывающими на точное местоположение каждой статистической единицы при помощи разработанной системы увязки. Данная система кодирования зданий с ее координатами обеспечила солидную основу надежного и удобного определения статистических зон.

190. Принятие географических координат для зданий также проложило дорогу более гибкому определению статистических зон. При помощи самых обычных программных комплексов географических информационных систем (ГИС) можно производить анализ принадлежности точки к полигону на основе данных, расположенных в точках (зданиях) и полигонах (статистических зонах).

191. Наиболее часто используемые в статистике территориальные единицы, не подпадающие под схему административного деления:

- Статистическое разделение на городские и сельские районы⁴;
- Населенные пункты (городские поселения);
- Муниципальные подзоны;
- Почтовые зоны;
- Квадраты координатной сетки.

⁴ Фактическая классификация в разных странах может различаться.

192. Например, можно сделать привязку всех переменных переписи населения и жилищного фонда к географическим координатам и тем самым предоставить сведения лицам, использующим статистику территориальных единиц, не подпадающих под схему административного деления. Однако вследствие высокой детальности сведений малых территориальных единиц следует обращать особое внимание на вопросы защиты информации и соблюдения конфиденциальности.

8.2. Микроданные для исследователей

193. Участвующее использование микроданных требует совершенствованных возможностей для предоставления лучших данных, удовлетворяющих требованиям пользователей. Чрезвычайно важно, чтобы исследователи и уполномоченные пользователи могли в полной мере использовать все изобилие накопленных данных. Говоря в общем, доступ к микроданным означает высокую отдачу от вложений в официальную статистику.

194. Технологические новшества в аппаратном и программном обеспечении, документировании данных и Интернет в большой мере увеличили возможности улучшения доступа к микроданным. Однако, так как вопрос конфиденциальности и защиты личных данных тесно связан с использованием микроданных, необходимо достичь равновесия между использованием данных и конфиденциальностью. Исходным положением является наличие у исследователей ограниченного доступа к необходимым для конкретного проекта данным - доступ не предоставляется для общих целей.

195. Микроданные для исследователей должны быть стандартизированными сведениями высокого качества. В североевропейских странах составлен ряд комплексных статистических регистров, применимых для анализа и исследовательских целей. Шведский регистр лонгитюдных данных *Луиза* содержит анонимизированные микроданные о лицах и семьях относительно их образования, дохода и занятости и может служить примером. Данный регистр включает в себя ежегодные данные обо всех взрослых в Швеции с 1990 и обновляется каждый год. В Финляндии был создан Массив лонгитюдных данных переписи, включающий в себя все переписи с 1970 г. (каждые 5 лет), а ежегодные Массивы лонгитюдных данных занятости создаются с 1987 г. Выборки для исследователей можно делать из таких объединенных массивов данных, которые предлагают широкие возможности по проведению различных анализов.

196. В течение ряда лет НСИ в Норвегии, Швеции и Финляндии предоставляют обезличенные микроданные большому числу исследовательских заведений и органам власти на магнитных лентах, компакт-дисках, DVD и в прочих форматах. Объем увеличился одновременно с увеличением числа версий/заданий. Ранее в Дании доступ к микроданным предоставлялся лишь на сайте датского ЦСБ.

Пример: Исследовательские базы данных в Дании

197. У исследователей возрастает потребность в анализе многомерных проблем, в ходе которого связываются данные ряда статистических регистров или/и множество периодических версий одного и того же регистра. Это приводит к высоким издержкам на начальном этапе исследовательского проекта.

198. Однако было установлено, что разным проектам иногда необходимы одинаковые наборы биографических данных, которые каждый раз приходится формировать заново. Например, это касается данных рынка труда, здоровья и образования. В этих случаях огромная работа по определению первичных данных производится раз и навсегда для того, чтобы исследователи могли начинать работу над своими проектами без необходимости проведения подготовительной работы. Эти базы данных постоянно обновляются.

199. Для усиления проводимых на основе регистров исследований, ЦСБ Дании предоставляет доступ к нестандартным наборам данных, скомпилированных из различных регистров, включая вышеупомянутые исследовательские базы данных. Уполномоченные научно-исследовательские учреждения и аналитические институты могут получить прямой доступ к не поддающимся идентификации микроданным, т.е. данным с удаленными идентификационными номерами.

Пример: Доступ к микроданным в Швеции

200. В 2005 году ЦСБ Швеции разработало новую систему удаленного доступа к микроданным (MONA). Данная система предоставляет пользователям безопасный доступ к базам данных ЦСБ Швеции практически отовсюду, где есть доступ к сети Интернет. Данные обрабатываются и анализируются посредством ряда приложений, а результирующие совокупности данных автоматически отсылаются пользователям. Основными задачами системы MONA являются:

- Расширение доступности микроданных для внешних пользователей при одновременном усилении безопасности и секретности;
- Хранение всех типов микроданных для исследований на сайте ЦСБ Швеции, тем самым обеспечивая контроль над тем, где, когда, кем и как используются данные;
- Мгновенное обновление данных в случае необходимости при отсутствии каких-либо требований по производству новых комплектов дисков или лент для повторного распространения;
- Предоставление конечным пользователям удобного в эксплуатации интерфейса, основывающегося на известных стандартных компонентах и методах, таких, как например, использование серверных вычислений;
- Предоставление полной системы с мощными серверами и широким набором приложений без необходимости в дорогостоящем оборудовании и затратах на программное обеспечение для конечных пользователей.

201. Несомненно, что североевропейские страны связаны обязательствами по улучшению доступа к высококачественным микроданным. Важной задачей для последующего развития является составление нескольких новых тематических регистров, специально созданных для более полного удовлетворения потребностей исследовательского сообщества. Для осуществления этого необходимо проделать значительную работу с привлечением как методологов, так и отраслевых специалистов. Другим перспективным направлением является развитие методов, позволяющих связывать данные из различных источников, как в рамках североевропейских НСИ, так и за их пределами. Кроме того, некоторые из североевропейских НСИ проектируют хранилища статистических данных, которые дадут им возможность создавать сгруппированные регистры и кубы методами, позволяющими осуществлять непрерывное обновление данных. Само собой разумеется, все упомянутые направления развития обязаны соблюдать правовые нормы о неразглашении.

Глава 9. КАЧЕСТВО И СТАНДАРТИЗАЦИЯ

9.1. Качество

202. Насколько статистика на основе регистров соответствует перечисленным в Главе 1 требованиям качества? Среди основных причин введения статистики на основе регистров можно выделить *снижение затрат* и *удержание усилий по заполнению анкет* на наиболее низком уровне.

203. В системе на основе регистров главным образом уделяется внимание *сопоставимости* и *связности* на микроуровне. Кроме того, североевропейские НСИ приложили усилия по согласованию данных статистики на основе регистров со статистическими данными на основе других источников.

204. *Доступность* и *доходчивость* лишь в малой мере зависят от используемых источников. Однако, поскольку системы данных на основе регистров довольно сложны и используются в ряде статистических материалов, метаданные обязаны быть исчерпывающими и упорядоченными. Говоря в общем, требования качества, налагаемые статистическим использованием данных регистров, в некоторой степени совпадают с теми требованиями, которые налагает на них основная цель регистра. Для оценки качества регистра необходимы достаточные метаданные от его владельцев (см. Главу 6).

Релевантность

205. Важным критерием соответствия первичных данных требованиям качества является их *релевантность* по отношению к статистике. Это означает, что данные должны быть связаны с теми понятиями, которые желательно осветить в статистике. Намного проще проверить релевантность в статистике на основе данных, специально собранных для этой цели. Когда статистические материалы основываются на данных, собранных для иной (административной) цели, обычно должно хватать любой необходимой или искомой для этой цели информации.

206. Однако в наиболее благоприятных случаях статистику удается даже оказывать влияние на содержание регистров.

207. Также возможно позволить разным источникам дополнять друг друга посредством увязки с тем, чтобы искомая информация была предоставлена при помощи сравнения данных различных регистров, что документирует предмет в относительно полной мере.

208. Статистическая система, всецело или в значительной степени основанная на регистрах органов государственной власти, может описывать картину мира глазами этих властей. Несуществующие в разрезе налоговых властей или социального управления понятия или подсовокупности также могут отсутствовать в статистических материалах. Во избежание этого важно, чтобы существовали определенные базовые регистры, такие как Центральный реестр населения (ЦРН), чьей задачей является запись всех единиц без привязки к определенной административной цели. Нам также достоверно известно, что большинство сведений различных органов власти можно увязать между собой, чтобы предоставить статистику, не ограниченную отдельным административным разрезом. В то же время важно учитывать, что регистры не могут предоставить полной картины общества. По причине этого необходимо собирать данные, рассматривающие аспекты, не рассматриваемые в регистрах. Сбор этих дополнительных данных осуществляется посредством опросов с помощью проведения интервью или анкетирования, причем обследовании в идеале должны включать в себя идентификаторы, которые позволят их сравнивать с данными регистров.

Точность

209. Другое важное требование, которому должны удовлетворять исходные статистические данные, касается точности. Это относится как к данным, полученным из опросных листов, так и к данным, полученным из регистров. Необходима полная уверенность в точности соответствия регистрируемых данных исследуемым обстоятельствам.

210. Конечно, на это можно возразить, что если неточности не носят систематического характера, они не приведут к смещению распределений, поскольку ошибки в противоположных направлениях имеют тенденцию взаимно погашать друг друга. Тем не менее, при рассмотрении статистических *связей* между несколькими переменными, подобные ошибки создают затруднения, поскольку они могут исказить оценку этих связей.

211. Если ошибки систематично проявляются в определенном направлении, проблемы незамедлительно увеличиваются, однако планомерность ошибок может облегчить их исправление.

212. И наконец, важно, чтобы данные регистрировались с необходимой для статистики *степенью точности*, т.е. используемый масштаб должен быть достаточно подробен. Например, степень детализации профессиональных данных должна быть довольно высока, если информация будет использоваться для эпидемиологических исследований, где риск заболевания рассматривается исходя из влияния профессии на индивида.

Своевременность и пунктуальность

213. Задержки в обновлении административных регистров во многих случаях вызывают необходимость в публикации нескольких изданий количественных данных: *предварительных данных*, которые следует издать из необходимости в текущих данных, которым, однако, не хватает точности по причине недостаточной актуализации, и более полных *окончательных данных*, которые выходят позже. Иногда может существовать целый ряд временных версий. Так как наличие нескольких версий данных может создать путаницу и причинить неудобство пользователям, его следует по возможности избегать. Несомненно, эта проблема также относится к статистике, сформированной из других источников, однако она особенно касается статистики на основе регистров, где актуализация существенно задерживается.

Изучение качества

214. Совмещение данных обследований и регистров предлагает большие возможности по изучению качества. Сравнения производятся как на макро-, так и на микроуровне. Например, в переписи населения и жилищного фонда, используется релевантное выборочное обследование данных (например, из обследования рабочей силы и опросов по жилищным условиям и обеспечению жильем) для оценки качества вызывающих сомнение данных.

215. Во времена, когда статистика на основе регистров была чем-то новым, было сделано предположение, что данные выборочных обследований необходимо использовать для измерения качества данных регистров. Мнение до сих пор распространено, однако сегодня данные регистров также считаются пригодными для оценки качества данных выборочного обследований.

9.2. Метаданные и документация

216. «Статистические метаданные являются описательной информацией или документацией статистических данных, то есть микроданных, макроданных и иных метаданных. Статистические

метаданные облегчают совместное использование, запрос и понимание статистических данных за все время существования данных»⁵

217. В последние десятилетия потребность в метаданных в процессе формирования статистики проявлялась все в большей мере. Большинство статистических бюро стараются внедрить системы метаданных или улучшить уже имеющиеся.

218. Характер метаданных значительно различается от данных на основе регистров и анкет с их собственным сбором сведений. Также следует отличать документацию *регистров* от документации *статистики на основе регистров*. При создании новых регистров на основе уже имеющихся документация имеет ключевое значение. Данный тип документации характеризуется следующими факторами:

- Объем метаданных может быть весьма значительным;
- Каждый административный источник подлежит документированию;
- Изменения в административной системе также документируются;
- Из-за сложности переменных документация должна быть точной;
- Производится большая работа по обработке данных для создания единиц и переменных, которая тоже должны быть документированы

219. Следовательно, система метаданных должна удовлетворять требованиям, предъявляемым самой системой регистров и статистикой на их основе.

Документирование административных источников

220. Поставщики данных (владельцы регистра) предоставляют на рассмотрение *описание записей* и т.п., указывающее на структуру и содержание поставляемых данных. Кроме того, у НСИ должен быть электронный доступ к *опросным листам*, включая *инструкции*, для их хранения в системе метаданных.

221. Все изменения необходимо тщательно отмечать и сохранять с целью облегчения оценки сопоставимости в будущем. Следовательно, система метаданных должна включать в себя *календарь*, т.е. формализованную систему метаданных, в которой возможно производить поиск информации, имеющей отношение ко *времени, регистру* и *переменной*.

Документирование источников в рамках системы регистров

222. При создании нового статистического регистра мы часто используем переменные других статистических регистров, уже находящихся в распоряжении в регистровой системе. В этом случае существенным является наличие простого доступа к уже существующим метаданным для возможности поиска и выбора подходящих переменных. Важно иметь возможность перемещать метаданные из существующих источников в документацию нового регистра. Чтобы избежать двойной работы и вместо этого с удобством использовать метаданные повторно, документация должна быть строго формализована в соответствии с общими правилами, и храниться в универсально-доступном виде.

Документирование изменений во времени

223. Четыре типа событий могут влиять на статистику на основе регистров. Во избежание неверных интерпретаций временного ряда из опросов на основе регистров нам необходимо знать следующее:

⁵ Семинар ЕЭК ООН по статистическим метаданным [METIS, 1999], Рабочий документ № 2

- Произошли ли изменения в административном устройстве, создающем источники, вследствие чего административным понятиям были присвоены новые определения?
- Произошли ли изменения в способе формирования регистра, например, появление новых источников или методов оценки?
- Произошли ли изменения в классификациях, используемых в регистре?
- Могли ли какие-либо внешние перемены косвенным образом повлиять на статистику?

224. Следовательно, календарь событий должен существовать там, где необходимо документировать и хранить в одном месте события, могущие повлиять на различные регистры. Это облегчит положение дел как для тех, кто работает над созданием регистра, так и для пользователей регистра. Тем самым снизится риск ошибочного толкования данных.

База данных классификаций и определений (Классификатор)

225. Промышленная классификация, категория продукта, образование, профессия и коды зон являются примерами важных статистических *стандартов* и *классификаций*. Административные источники содержат данные об этих иерархических классификациях, которые используются для создания переменных в системе регистра. Настоящие классификации регулярно сменяются. Так как существует большое количество наборов значений (наборы всех кодов/категорий), необходим классификатор для управления всеми кодами и переходными ключами между различными версиями. Такой классификатор является важным ресурсом при условии документирования переменных в регистре.

Полная система метаданных

226. Должна существовать система, объединяющая все формализованные метаданные, включая календарь и базу данных классификаций и определений. Кроме этого, необходима система управления документами, содержащие иные типы метаданных.

9.3. Согласование

227. Важной частью методологической работы со статистикой на основе регистров является структурирование и совершенствование всей системы. Координация направлена в первую очередь на сведение к минимуму двойной работы, упрощение совместного использования данных и повышение взаимосвязанности между разными регистрами. Можно выделить 4 типа согласования:

- Организационное;
- Техническое;
- Связанное с содержанием;
- Методологическое.

Организационное согласование

228. Исследованиям на основе регистров нужна организация, содействующая обширному сотрудничеству между лицами, ответственными за разные части системы регистров. Создание соответствующей формальной организации важно, однако не является достаточным. Также необходимо, чтобы в организационной культуре неотъемлемой ее частью считалось сотрудничество, выходящее за рамки формального.

229. В североевропейских странах организация систем данных на основе регистров различна. В качестве примера - шведская модель.

Пример: Организационное согласование в Швеции

230. В ЦСБ Швеции существует центральное подразделение, которое несет ответственность за систему регистров. Данное подразделение работает над структурированием и согласованием регистров в единую функционирующую систему.

231. Работу с отдельными статистическими регистрами в рамках существующей системы и со статистикой на основе статистических регистров выполняют местные подразделения. Обязанности в НСИ должны быть четко регламентированы между центральным и местными подразделениями.

232. Полученные из административных источников сведения сравниваются с данными опросов, имеющих отношение к тем же единицам. Совместное редактирование такого типа является важной задачей, повышающей качество статистики. Во избежание двойной работы, ее необходимо выполнять в тесном сотрудничестве со всеми, вовлеченными в работу над статистикой. Совместное редактирование можно организовать в виде общих проектов или при помощи создания особых подразделений.

233. Развитие новых интегрированных (статистических) регистров означает, что на первый план выступают определенные организационные вопросы. Во избежание конфликта необходимо наличие постоянного сотрудничества между вовлеченными подразделениями.

Двумя основными факторами в наличии культурного перехода от «королевской собственности» к тематическим сведениям и статистическим продуктам к «общей собственности» являются высказывания руководителей НСИ и потребность в комплексных статистических данных, выраженная исследователями и прочими влиятельными пользователями. Однако НСИ также должно признать необходимость в отраслевых специалистах для обработки административных данных для статистического использования и для взаимодействия с хранителями реестров.

Техническое согласование

234. Наиболее важным отдельным фактором для технического согласования является надежно работающая единая нормированная *система метаданных* (см. п. 9.2).

235. *Соглашение о нормированном назначении имен* и одинаковые *форматы данных* для ключевых переменных, таких как переменные увязки, содействуют согласованности. Между названием переменной и ее определением в рамках системы регистра должна существовать однозначная связь. Название переменной, появляющееся в ряде регистров, всегда должно использовать одно и то же определение. В противоположность этому, переменные с одинаковым определением, но разными названиями в различных регистрах, никогда не должны иметь места.

236. Непротиворечивость требует скоординированного *физического доступа* к данным во взаимодействующих регистрах. Минимальным требованием является совместимость формирующих систему регистров с широко применяемыми средствами обработки и анализа для обеспечения возможного использования указанных средств для совместной обработки регистров. Это является преимуществом, если хотя бы часть совместной обработки можно выполнить внутри соответствующих баз данных.

Согласование, связанное с содержанием

237. Согласование содержания основывается на следующем:

- Стандартизированных совокупностях;

- Стандартизированных переменных;
- Стандартизированных группировках важных классификаций.

238. Все совокупности необходимо создавать при помощи статистических базовых регистров. Важным использованием базовых регистров является создание *стандартизированных совокупностей*, которые затем используются для создания совокупностей регистра в остальных статистических регистрах системы.

239. Стандартизированные совокупности должны быть высокого качества в отношении охвата и важных табулированных переменных. По мере возможности, необходимо получить и обработать всю административную информацию прежде, чем определять совокупность. Стандартизированные совокупности в таком случае не должны подвергаться изменениям, а другие статистические регистры должны использовать эти стандартизированные совокупности регистров.

240. Система регистров должна основываться на двух основных типах стандартизированных совокупностей регистров:

- Совокупность на конец года – единицы, существующие на конец года;
- Совокупность календарного года – единицы, существовавшие в некоторые моменты времени в течение года.

241. Другой существенной проблемой является *координация переменных*. Важным инструментом координации является центральный банк классификаторов, содержащий статистические классификаторы, используемые в рамках статистической системы. В нем должны содержаться классификаторы и переходные ключи между различными версиями классификатора, установленного в качестве общегосударственного или международного стандарта.

242. Переменные, неоднократно используемые в ряде регистров, являются релевантными для стандартизации. Необходимо прояснить, какое подразделение в НСИ несет ответственность за конкретные переменные. Обычно это подразделение, получающее переменную. В его ответственность входит обработка, присвоение имен и документирование вышеупомянутых переменных. Соответственно, подразделения, создающие важные производные переменные несут ответственность за присвоение им имен и их документирование.

243. Множество переменных необходимо сгруппировать или разбить на интервалы группирования. Данные группы и интервалы группирования необходимо сделать в соответствии с общими правилами, иначе таблицы из разных источников будут несопоставимы. Категории возраста и дохода, например, следует формировать одинаковым образом в различных регистрах и выборочных наблюдениях. Большие классификаторы, такие как классификаторы отраслей, образования и профессий, необходимо сгруппировать до того, как можно будет сформировать статистические таблицы. Для группировки этих переменных необходимо их хорошо знать. Подразделениям, несущим ответственность за эти переменные в рамках организации, следует создать ряд стандартизированных группировок, которые можно приспособить для разных нужд. Для различных целей необходимы как агрегированные, так и детализированные деления. Создаваемым группам нужно также присвоить стандартизированные названия.

Согласование статистической методологии

244. Важными для развития статистики на основе регистров является общий подход и общая терминология.

245. «*Базовые регистры следует использовать при определении совокупностей*» и «*каждый должен оказывать поддержку базовым регистрам*» являются двумя основополагающими

принципами, содействующими непротиворечивости совокупностей системы. Всякий, занимающийся увязкой и согласованием ряда регистров должен иметь уверенность в том, что каждый регистр был обработан, отредактирован и документирован для упрощения обработки и обеспечения качества.

246. Выкладки могут нуждаться в координации для обеспечения состоятельности оценок различных регистров. Если, например, у имеющейся в нескольких регистрах переменной частично отсутствует ответ, то отсутствием ответа нужно заниматься таким же образом, как и в рассматриваемых регистрах.

Глава 10. ПЕРЕПИСИ НАСЕЛЕНИЯ И ЖИЛИЩНОГО ФОНДА НА ОСНОВЕ РЕГИСТРОВ

247. Основной вывод, сделанный во всех североевропейских странах, заключается в том, что разработка проведения переписи на основе регистров является поэтапным процессом. Сначала статистику предметных областей опробовали, одновременно были опубликованы ее результаты в разных сферах. Как только качество было признано удовлетворительным, были введены переменные на основе регистров. Когда будет разработана статистика для всех сфер, включаемых в перепись, можно будет проводить перепись, полностью основанную на использовании регистров.

248. Другими словами перепись на практике является неотъемлемой частью системы, в которой информация черпается из регистров. Те же статистические регистры служат источниками для переписей и статистики предметных областей. Это означает, что почти все выводы, сделанные в данном отчете по статистике на основе регистров, в большинстве случаев также касаются переписей на основе регистров. В данной Главе мы несколько более подробно рассмотрим вопросы, касающиеся проведения переписей.

10.1. Разработка проведения переписей на основе регистров

Что такое перепись на основе регистров?

249. На одном конце линии у нас традиционные переписи, собирающие данные с помощью счетчиков и анкет, без использования какой-либо информации из регистров; на другом – переписи, полностью проводимые на основе регистров. В некоторых странах используется комбинированный режим сбора данных и использования данных из регистров и анкет (как полный счет или выборочное наблюдение). Но даже в тех странах, где в основном проводятся традиционные переписи, можно до определенной степени использовать информацию из регистров, например, список адресов лиц или домохозяйств для рассылки опросных бланков, заполняемых при переписи.

250. В некоторых странах, в частности, североевропейских, была поставлена цель разработать порядок проведения переписей полностью на основе регистров. В других странах целью может быть использование регистров в лишь некоторой степени, даже если отсутствует возможность или желание проводить перепись полностью на основе регистров. В третьих же странах вопрос об использовании регистров при переписи может вообще не стоять.

251. Какие требования должны быть выполнены, чтобы перепись считалась проведенной на основе регистров? Основное требование заключается в наличии регистра населения и регистра жилищного фонда. Все-таки постоянно проживающие лица и единицы жилья являются базовыми единицами при проведении переписей населения и жилищного фонда. В такой же степени важна увязка между лицами и их жильем, в т. ч. с единицей домохозяйства. Эти требования являются краеугольными камнями при разработке порядка проведения переписи на основе регистров.

Оценка качества и описательные процессы

252. Когда считать статистический регистр или систему регистров достаточно хорошей для использования в качестве основного источника при переписи или, более широко, для статистических целей? В качестве примера мы рассмотрим опыт Норвегии.

Пример: Разработка порядка проведения переписи на основе регистров в Норвегии

253. Центральный реестр населения (ЦРН) был создан в 1964 г. на основании данных переписи населения 1960 г. Перепись в 1970 г. проводилась для проверки и обновления ЦРН, а также для создания регистра образования. При проведении переписи в 1980 г. административные регистры использовались для рационализации сбора данных. Тогда впервые был использован сбор данных только с помощью почтовых отправок. Анкеты стали гораздо короче, их разослали только лицам старше 16 лет. Информация о демографии, образовании и доходах стала заноситься в регистры, которые также стали использоваться во время кодирования и обработки данных. Затем результаты переписи были использованы для обновления регистра образования.

254. В нескольких научных работах рассматривалось проведение переписи населения 1990 г. только на основе регистров. В качестве эксперимента существовавшие регистры были использованы в 1980 г. для создания переписной статистики, их результаты были сопоставлены с результатами переписи 1980 г. Был сделан вывод, что даже если в существующих регистрах содержатся данные по рынку труда, их качество не подходит для целей переписи. Регистры не содержали информации о домохозяйствах и жилье или содержали ее очень мало, поэтому было принято решение провести сбор данных о рынке труда и домохозяйствах с помощью анкетирования выборки населения. Информация этого выборочного наблюдения была сведена вместе с информацией различных административных и статистических регистров, что привело к созданию высококачественного статистического регистра.

255. При планировании переписи 2001 г. стало ясно, что все данные о лицах можно собрать из регистров. Однако на тот момент не существовало регистра жилищного фонда, поэтому данные о домохозяйствах и жилье необходимо было собрать, проведя полное обследование с анкетированием каждой занятой единицы жилья. В соответствии с решением Парламента, был реализован проект по введению в «Регистр недвижимости, адресов и зданий» данных о жилом фонде. Кроме того, официальные адреса, используемые ЦРН, получили точные жилищные номера, если по ним проживали лица в домах с несколькими жилыми единицами. Информация о жилье и связи между жилыми единицами и проживающими в них лицами были установлены в ходе проведения переписи 2001 г. и затем занесены в регистры. Таким образом, были созданы все необходимые регистры для проведения в будущем переписей полностью на основе регистров.

256. Для измерения качества данных регистров и оценки, является ли качество удовлетворительным, можно использовать несколько методов. Один из них в особой степени касается переписей и заключается в проведении альтернативной «переписи» данных на основе регистров и ее сопоставлении с последней проводившейся переписью. Это позволяет увидеть, какие части регистровой системы имеют удовлетворительное качество, а какие следует улучшить с тем, чтобы проводить переписи на основе регистров.

257. Еще одной отличительной чертой североевропейских стран является взаимосвязь между регистрами и переписями. Переписи использовались для создания регистров, а данные регистров, соответственно, используются как источник для переписей. Чтобы вводить данные переписей в административные регистры, было необходимо использовать в дополнение к Закону о статистике правовую базу задействованных административных регистров.

10.2. Сопоставимость с международными рекомендациями

Специфические черты проведения переписи населения и жилищного фонда

258. Данные переписи населения на основе регистров в большой степени соответствуют рекомендациям по проведению переписи населения и жилищного фонда Европейской экономической

комиссии Организации Объединенных Наций и статистическими бюро стран-членов Европейского сообщества. Данные рекомендации включают в себя пять специфических черт.

259. *Индивидуальная регистрация*: Регистр является систематическим сбором данных на уровне единиц, организованный таким образом, чтобы было возможно его обновление. Регистры, используемые при переписи, охватывают все лица и весь жилищный фонд, так что соблюдается условие индивидуальной регистрации.

260. *Универсальность в рамках установленной территории*: Перепись населения на основе регистров касается всех лиц, зарегистрированных в качестве проживающих в стране в день проведения переписи. Перепись жилищного фонда касается всех традиционных жилищ жилья и учреждений, но не включает нетрадиционные жилища.

261. *Одновременность*: Поскольку все регистры регулярно обновляются, можно зарегистрировать все единицы в день проведения переписи.

262. *Данные по малым районам*: Данные на основе регистров хорошо подходят для составления статистики малых территориальных районов.

263. *Установленная периодичность*: Большинство переменных переписи публикуются ежегодно. Полные переписи можно проводить по необходимости, например, каждые 10 лет.

Единицы и переменные

264. В соответствии с рекомендациями, данные следует собирать о следующих статистических единицах: *лицах, домохозяйствах, семейных ячейках, жилых помещениях и зданиях*. Можно составлять данные по всем этим единицам на основе регистров. Используется понятие *жилье-домохозяйство*, оно означает всех лиц, проживающих в одном и том же жилье, принадлежащем тому же домохозяйству.

265. Что касается переменных *переписи*, ситуация в разных странах может различаться. Однако, говоря в общем, если в наличии есть полностью созданная система проведения переписи на основе регистров, можно составить все *основные признаки*.

266. При этом некоторые *дополнительные признаки* нельзя составить по данным регистров. Например, сюда не внесена *этническая принадлежность*. Это – переменная, которая безусловно имеет субъективное измерение и должна записываться со слов индивида. Информация подобного рода обычно не включается в административные регистры⁶. Другими примерами являются *компьютерная грамотность, вид транспорта, используемый для проезда на работу, и потребительские товары длительного использования, принадлежащие домохозяйству*. Некоторые переменные есть в одной стране, но отсутствуют в другой, например, *вероисповедание, язык и обычные рабочие часы*.

10.3. Преимущества и недостатки проведения переписи на основе регистров

267. Преимущества и недостатки проведения переписи на основе регистров, конечно, те же самые, что и для *статистики* на основе регистров в целом. Однако есть ряд аспектов, касающихся переписной статистики в особенной степени.

⁶ Страну рождения и страну рождения родителей обычно можно установить из административных источников. В сочетании, например, с гражданством и вероисповеданием эти переменные могут предоставлять определенную информацию об этнической группе, но не покрывают полностью содержание переменной *этническая принадлежность*.

Преимущества

268. *Снижение расходов* очень актуально при проведении переписи. При проведении же переписи на основе регистров уже не нужно создавать и тестировать переписные бланки, посылать заказ на их печать в типографию, предварительно их заполнять, рассылать, рассылать потом напоминания, кодировать и сохранять данные. НСИ больше не нужно нанимать большое количество людей тщательно анализировать возвращенные бланки. Например, при проведении переписи населения в Финляндии в 1980 г. было принято на работу и обучено около 2500 человек.

269. Другим значительным преимуществом является *короткое время*, которое уходит у респондентов для ответов на вопросы обследования. Например, при проведении переписи населения в Финляндии в 1980 г. населению пришлось в общей сложности заполнить 8,5 миллионов бланков и ответить на около 100 миллионов вопросов, хотя административные данные уже имелись и широко использовались. Если чтение инструкций, заполнение бланка и отсылка анкеты заняли в среднем у каждого человека 10 минут, то количество затраченного труда у населения (при довольно консервативной оценке) составит более двух миллионов часов. Если взять за почасовую оплату 10 евро, то общие расходы по заполнению бланка составят 15 миллионов евро.

270. Вся статистика по переписи населения доступна *ежегодно*, т. е. пользователям больше не приходится ждать 10 лет обновленных данных *переписи*. Или говоря другими словами, можно напрямую сопоставлять данные переписи и ежегодную статистику предметных областей.

Недостатки

271. Недостатком является полная зависимость описания на основе регистров от сведений, которые можно создать на основе имеющихся регистров. Это налагает некоторые ограничения в отношении *явлений*, которые имеются для описания, т. е. *некоторые темы нужно убрать* из систем переписи населения на основе регистров (см. п. 10.2). Выборочные обследования можно использовать для сбора данных, которых нельзя получить из регистров. Способы восстановления недостающих данных можно применять для введения этих данных в картотеку переписи.

272. Кроме того, некоторые ограничения наложены на определения единиц и переменных. Частное домохозяйство определяется как все лица, проживающие в одном жилище (*жилище домохозяйство*). *Домохозяйственные единицы* (лица, проживающие в одном жилище с общим управлением) невозможно создать, используя регистры. При подсчете населения *обычно* берется *зарегистрированное* место жительства (в соответствии с регистром населения), а не *фактическое*.

273. Использование административных регистров как источника при переписи населения может иметь эффект своевременности. Время, необходимое для создания переписных переменных (в т. ч. для обновления административных регистров), может быть разным для разных предметов. Например, в Норвегии демографические переменные готовы к изданию уже через три месяца после проведения переписи, а время создания переменной доходы занимает почти 18 месяцев. Такая разница объясняется различным порядком работ по обновлению соответствующих административных регистров. Поэтому для переписи на основе регистров статистика для некоторых переменных будет более своевременной, а для некоторых – менее по сравнению с традиционной переписью.

274. После того, как данные переписи собраны из регистров, перепись *перестает быть инструментом сбора сведений для удовлетворения возникающих в обществе потребностей в информации*. В тех странах, где проводится традиционная перепись населения, в анкеты могут вводиться новые переменные для сбора новых данных, которые сочтены очень важными. Эта гибкость исчезает, когда данные перестают собирать с помощью анкет. Однако и при традиционной переписи эта возможность очень ограничена, поскольку введение в бланк новых вопросов – очень

дорогостоящая операция. Поэтому независимо от вида переписи, альтернативой для сбора дополнительных данных для большинства стран будет проведение выборочных обследований.

Глава 11. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Одинаковая статистика на основе другого типа данных

275. Составление статистики из административных источников предполагает, что сбор, обработка и другие виды обращения с данными выполняются с помощью иных методов, чем традиционных. При этом статистический продукт остается тем же самым. Вместо того чтобы задавать вопросы лицам, домохозяйствам или компаниям для получения информации, мы обращаемся к владельцам административных регистров. Вместо проверки качества данных, полученных от физических лиц, мы модифицируем административные данные с учетом нашего знания о различиях в терминах и охвате между административными источниками и потребностями статистики. В обоих случаях потребуются человеческие ресурсы, но их квалификация может быть различной. После создания регистров и выполнения требований к минимальному качеству информации статистика на основе регистров обходится обычно дешевле, чем традиционная.

Необходимость в некоторых базовых регистрах

276. Для статистической системы, в основном базирующейся на административных источниках, необходимо наличие трех базовых регистров: населения, хозяйствующих субъектов (компаний и учреждений) и регистр адресов, зданий и жилья. Необходимы также уникальный идентификационный номер каждой единицы базовых регистров и связи между ними. Уникальный идентификационный номер нужно, кроме того, использовать и в других административных регистрах данных единиц (регистры образования, налогообложения и т. п.).

277. Статистические системы на основе регистров можно создать без упомянутых трех базовых регистров, но наш опыт показывает, что их отсутствие создает проблемы с обеспечением качества производимой статистики.

Статистики должны оказывать активную поддержку базовым регистрам

278. Административные регистры обычно создаются и управляются соответствующими органами власти, но часто НСИ оказывают активную поддержку при учреждении и организации базовых регистров. В отдельных случаях НСИ было даже необходимо взять на себя задачу создания и управления как базовых регистров так и других административных регистров. В таких случаях как только регистр для административного использования создан и требования к нормальному качеству и функционированию работы выполнены, ответственность за ведение регистра должна перейти к другому официальному квалифицированному органу. Однако даже тогда НСИ должен оказывать активную техническую (информируя о проблемах качества) и политическую поддержку (привлекать необходимые ресурсы) регистра.

Чем больше пользователей у регистра, тем лучше качество статистики

279. Создание административных регистров и их управление очень дорогостоящи. Если многие пользователи (государственные и частные) утверждают, что могут более рационально достичь своих целей с помощью административного регистра, правительство, возможно, выделит на него финансирование.

280. Даже административные пользователи могут способствовать повышению качества регистра. Если, например, какой-либо пользователь сообщит, что необходимая единица отсутствует в регистре, включение такой новой единицы повысит качество на благо всем остальным пользователям, в т. ч. НСИ. НСИ может уведомить хранителей регистра о проблемах с качеством, хотя обычно не сообщает об ошибках, касающихся отдельных единиц регистров. Оказывая активную поддержку

регистру, НСИ должны удостовериться, что не нарушают государственного законодательства о статистике или основные принципы ООН по ведению официальной статистики.

Требования статистиков к данным регистров

281. Единицам и переменным следует давать точное определение, при чем его должно быть легко понять человеку, предоставляющему регистру информацию (частному лицу или служителю, задающему вопросы и записывающему ответы). Для каждого предмета предпочтительно иметь только один регистр, охватывающий все существующие единицы со всей необходимой информацией. На практике, однако, нам приходится мириться с тем, что информацию по одному предмету нужно собирать из более одного источника. Интеграция данных из двух или больше источников представляет для статистиков проблему. Если трудно дать определение статистической единицы из административных источников, мы должны создать процедуры по проведению оценки, иначе данные могут стать бесполезными для статистических целей. Указание времени (момент времени или период) должно быть точно определено, а в идеале о событии нужно сообщить как дату, когда оно произошло, так и дату, когда информация о нем была занесена в регистр.

282. Административные регистры должны служить административным целям, но правила и порядок работы для данного регистра могут меняться со временем. Необходимо учитывать влияние этих изменений на качество статистики в целом, и особенно на временные ряды, когда регистр корректируется в соответствии с новыми административными правилами. Лучший способ защитить статистические потребности заключается в поддержании хороших отношений между теми, кто отвечает за ведение регистров, и теми, кто отвечает за составление статистики. Такие контакты и опыт нужно документировать и знакомить с ними новых сотрудников в регистровых органах власти и НСИ, когда изменяется должность сотрудника.

Данные регистров - «общее благо» внутри НСИ

283. Данные из какого-либо административного источника могут быть прямым вводом в более одного вида официальной статистики. Такой источник можно также использовать при контроле качества других видов статистики, при этом, чем больше используется административных данных, тем выше будет качество статистики.

284. Все статистические данные, независимо от того, на основе ли регистров они составлены или нет, являются частью статистической системы, и все данные могут использоваться в предметных областях статистики или в крупных системах статистических данных - таких, как системах переписи населения и жилищного фонда. Поэтому данные являются «общим благом». Отношения между НСИ и административным регистром могут быть организованы в рамках одного отдела в НСИ или (на практике) ограничиваться одним сотрудником, но главное – чтобы ответственный отдел или сотрудник выступал от лица всего НСИ и по-настоящему сознательно относился к такой внутренней и внешней коммуникации.

Термины, принятые в статистике на основе регистров

Источник административных данных: В принципе все виды источников, используемых для административных целей. В данном отчете: все упоминаемые источники административных данных являются регистрами.

Административный регистр: Регистр, используемый прежде всего в административной информационной системе. Это означает, что регистры используются для производства товаров и оказания услуг в государственных или частных учреждениях или компаниях или что информация является результатом такого производства товаров или оказания услуг. Административные регистры, используемые для статистических целей, обычно находятся в ведении государства или местных органов власти, кроме того, используются и регистры частных организаций.

Базовый регистр: административные базовые регистры ведутся как базовый ресурс для управления на государственном или частном уровне. Его задача – вести учет генеральной совокупности и сохранять идентификационную информацию. *Статистические* базовые регистры основаны на соответствующих административных регистрах. Их основные задачи – определять важные совокупности и содержать связи с другими базовыми регистрами.

Производная переменная: Новая переменная, образованная путем использования существующих переменных.

Связь: Одна или несколько связующих переменных, которые идентифицируют отдельные единицы. Связи (или *ключи*) используются, когда сличаются несколько регистров.

Первичный регистр: Чаще всего идентичен административным регистрам, но также используется и для статистических регистров в областях, где не существует центрального административного регистра.

Регистр: Систематический сбор данных на уровне единиц, организованный таким образом, чтобы была возможность его обновлять. Обновление – это обработка поддающейся идентификации информации с целью создания, обновления, корректировки или расширения регистра.

Перепись на основе регистров: Сбор всех данных из статистических регистров мы называем (полной) переписью населения и жилищного фонда на основе регистров. Перепись, основанная на сочетании данных регистров и анкет, называется частичной переписью на основе регистров.

Статистическая система на основе регистров: Статистические регистры, входящие в общую согласованную систему.

Статистика на основе регистров: Статистика, созданная путем использования только данных регистров. Данные из других источников (выборочных обследований) могут использоваться опосредованно, например, для восстановления недостающих данных, калибровки моделей или качественного анализа.

Хранитель регистра: См. *владелец регистра*.

Владелец регистра: Орган власти, несущий ответственность за ведение административного регистра. Также может называться *хранителем регистра*.

Вторичный регистр: См. *статистический регистр*.

Специализированный регистр: Регистр, который в отличие от базовых регистров служит одной определенной цели или ясно определенной группе целей. Специализированные регистры часто получают информацию о совокупности и некоторые базовые данные из базового регистра, а другие данные предоставляют самостоятельно.

Статистический регистр: Регистр, созданный для статистических целей. Статистический регистр может иметь своей основой один или несколько административных регистров. Статистические регистры также могут называться *вторичными*.

Данные обследования: Данные, собранные непосредственно для статистических целей. Обследования могут охватывать всю совокупность, но чаще – выборку.

Единица: Совокупность, состоящая из нескольких единиц (или объектов).

Использованная литература

Хубирс М. (2004). О базе социальных статистических данных и унифицированных оценках в ЦСБ Нидерландов. «Журнал официальной статистики», том 20, № 1, 2004, с. 55-75. Houbiers, M., 2004. Towards a Social Statistical Database and unified estimates at Statistics Netherlands, in: Journal of Official Statistics, Volume 20, No. 1, 2004, 55-75.

Лонгва С., Томсен И. и Северейде П. И. (1998). Снижение расходов при проведении переписи населения в Норвегии путем использования административных регистров». Перепечатка из «Международного статистического обзора», том 66, с. 223-234. Longva S., Thomsen I. and Severeide P.I. (1998): Reducing Costs of Censuses in Norway through Use of Administrative Registers. Reprint from International Statistical Review, Vol. 66, 223-234.

Шулте Нордхолт Е., Хартгерс М. и Гиркур Р. (ред.), 2004. Нидерландская виртуальная перепись 2001 г., анализ и методология. (Форбург/Херлен, 2004). (<http://www.cbs.nl/en-GB/menu/themas/dossiers/volkstellingen/publicaties/2001-b57-pub.htm>). Schulte Nordholt, E., M. Hartgers and R. Gircour (eds.), 2004. The Dutch Virtual Census of 2001, Analysis and methodology. (Voorburg/Heerlen, 2004). (<http://www.cbs.nl/en-GB/menu/themas/dossiers/volkstellingen/publicaties/2001-b57-pub.htm>)

Шулте Нордхолт Е., 2005. Нидерландская виртуальная перепись 2001 г.: Новый подход путем совмещения различных источников. «Статистический журнал Европейской экономической комиссии Организации Объединенных Наций», том 22, № 1, 2005, с. 25-37. Schulte Nordholt, E., 2005. The Dutch virtual Census 2001: A new approach by combining different sources, in: Statistical Journal of the United Nations Economic Commission for Europe, Volume 22, Number 1, 2005, 25-37.

ЦСБ Дании (1995): Статистика физических лиц в Дании – статистическая система на основе регистров. Statistics Denmark (1995): Statistics on Persons in Denmark – a register-based statistical system.

ЦСБ Дании (2005): Новые разработки в датской системе доступа к микроданным. Рабочий документ совместной рабочей сессии ЕЭК ООН/Евростат по конфиденциальности статистических данных в Женеве 9-11 ноября 2005 г. Statistics Denmark (2005): New developments in the Danish system for access to micro data. Working paper for the joint UNECE/Eurostat Work Session on statistical data confidentiality in Geneva 9.11 November 2005.

ЦСБ Финляндии (2004): Использование источников регистров и административных данных для статистических целей – Передовые методики ЦСБ Финляндии. Statistics Finland (2004): Use of Registers and Administrative Data Sources for Statistical Purposes – Best practices in Statistics Finland.

ЦСБ Норвегии (2005): Обработка и восстановление недостающих данных для создания связанного микрофайла из базовых регистров и административных данных. Рабочий документ совместной рабочей сессии ЕЭК ООН/Евростат по обработке статистических данных в Оттаве 16-18 мая 2005 г. (<http://www.unece.org/stats/documents/2005/05/sde/wp.8.e.pdf>). Statistics Norway (2005): Editing and imputation for the creation of a linked micro file from base registers and other administrative data. Working paper for the UNECE Work session on statistical data editing in Ottawa 16-18 May 2005. (<http://www.unece.org/stats/documents/2005/05/sde/wp.8.e.pdf>).

ЦСБ Норвегии (2006): Роль переписей в стране, обладающей статистической системой на основе регистров: Опыт и планы Норвегии. Доклад для пленарной сессии КЕС в Париже 13-15 июня 2006 г.

(<http://www.unece.org/stats/documents/ece/ces/2006/32.e.pdf>). Statistics Norway (2006): The role of censuses in a country with a register-based statistical system: Norwegian experiences and plans. Paper for the CES Plenary Session in Paris 13-15 June 2006. (<http://www.unece.org/stats/documents/ece/ces/2006/32.e.pdf>)

Валлгрэн А. и Валлгрэн Б. (2006): Статистика на основе регистров – административные данные в статистических целях. «Джон Уили и сыновья, Лтд.» Wallgren A. and Wallgren B. (2006): Register-Based Statistics – Administrative Data for Statistical Purpose.. John Wiley & Sons, Ltd.

Статистика на основе регистров в Северо-европейских странах

Обзор передовых методик с уделением основного внимания на статистику населения и социальной статистике

Североевропейские страны (Дания, Финляндия, Исландия, Норвегия, и Швеция) имеют давнюю традицию использования административных регистров при составлении официальной статистики. У них также существует давняя традиция сотрудничества в сфере статистики. Национальные статистические институты (НСИ) в Североевропейских странах решили поделиться своим опытом и знаниями с международным статистическим сообществом путем создания обширной документации по своим передовым методикам.

В последние годы увеличивается число стран в регионе ЕЭК ООН, рассматривающих возможность составления статистики на основе административных регистров. Настоящее издание предоставляет сотрудникам стратегических и плановых отделов Национальных статистических институтов понятие о том, что такое статистика, создаваемая на основе регистров, в т. ч. необходимые технические и административные функции и возможное приложение этих методов для составления официальной статистики. Основное внимание в данном издании уделено использованию административных регистров при составлении демографической и социальной статистики.

Опубликовывая данное издание, Европейская экономическая комиссия Организации Объединенных Наций (ЕЭК ООН) хотела бы поддержать Североевропейские страны в передаче их опыта по данному вопросу с международным статистическим сообществом в целом. Издание представляет собой ценный инструмент для всех НСИ, планирующих составление официальной статистики на основе административных регистров. Оно также поддерживает проведение в 2010 г. цикла переписи населения и жилищного фонда, поскольку растущее число стран за пределами Северной Европы планирует использовать административные источники в качестве дополнения к сбору традиционных данных на местах или их замены. Руководство по использованию регистров для переписи населения и жилищного фонда также включено в новый комплект «Рекомендации Конференции европейских статистиков по проведению переписи населения и жилищного фонда 2010 года», подготовленный ЕЭК ООН совместно со Статистическим агентством Европейского Союза (ЕВРОСТАТ).