

Трансформировавшаяся комплексная статистика народонаселения и миграции: актуализация

Луиза Блэквелл

Национальное статистическое управление

ЕЭК ООН, сентябрь 2022 г.

Содержание

1. Более широкий контекст: трансформация Управления национальной статистики (УНС)
2. Ограничения имеющейся системы
3. Динамическая демографическая модель
4. Окно оценки
5. Преимущества и недостатки
6. Предварительная обработка данных и панель данных в реальном времени
7. Статистические наборы данных по населению и специальным группам населения
8. Оценки миграции на базе административных данных
9. Работа УНС над характеристиками административных данных

Выражение благодарности: Группа моделирования УНС и сотрудники Университета Саутгемптона: Питер Смит, Якуб Биджак, Джейсон Хилтон. Также научная консультативная группа лонгитюдных исследований.

Почему нужна трансформация?

Соответствие стратегии Управления статистики Великобритании (UKSA) — «Статистика на благо общества»

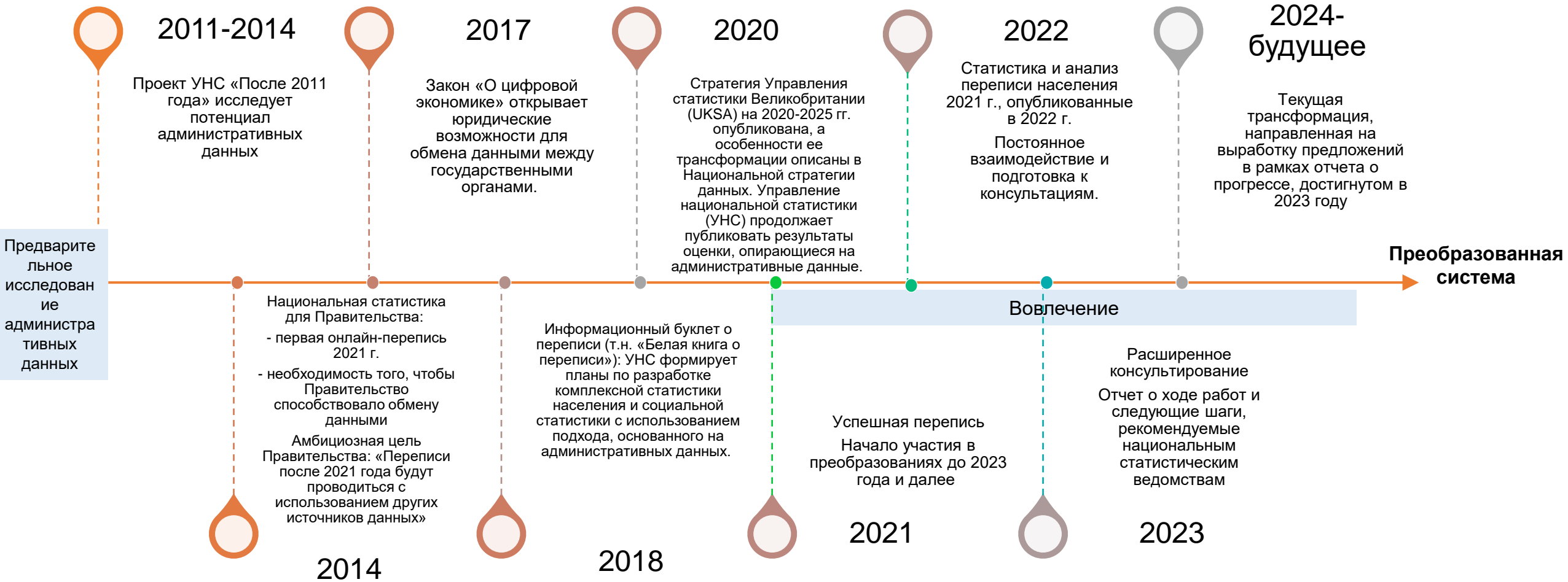
- **Устойчивость:**
 - Используется **наилучшая доступная информация**;
 - Поддержание стабильного качества во времени;
- **Фундаментальность:**
 - Понимание необходимости более частых расчетных оценок;
 - Принятие во внимание **радикальных изменений** системы;
- **Амбициозные цели:**
 - Стремление производить **более частые и своевременные оценки**.
- **Инклюзивность:**
 - Стремление производить **более инклюзивную/полную статистику** (с течением времени).



*«Пандемия произвела **революцию** в общедоступных данных, и назад пути нет»*

- Сэр Ян Даймонд

Сроки



Наши амбициозные цели

Предоставить наилучшую картину населения Великобритании в долгосрочной перспективе, двигаясь в направлении создания системы данных народонаселения и миграции, которая обеспечит:

- Более широкие, подробные и своевременные обновления данных переписи населения, чем когда бы то ни было ранее;
- Увеличение широты и детализации в ближайшие годы.

Ограничения существующей системы оценки численности населения

1. Вручную, с весомым экспертным суждением

- Медленно
- Трудность повторного воспроизведения

2. Отсутствие гибкости

- Сложности с переходом на другие источники данных или выходные данные

3. **Критическая зависимость от переписи**

- **Растущая неопределенность и десятилетняя перебазировка**

Новые вызовы

- Спрос на более частые обновления данных
- Спрос на более детализированные результаты
- Возможный отказ от традиционной переписи
- Максимальное использование административных данных
- Изменения во входных данных
- Появление новых источников данных (например, сотовые телефоны)

Структура

Демографический счет: Джон Брайант и Хуанни Чжан

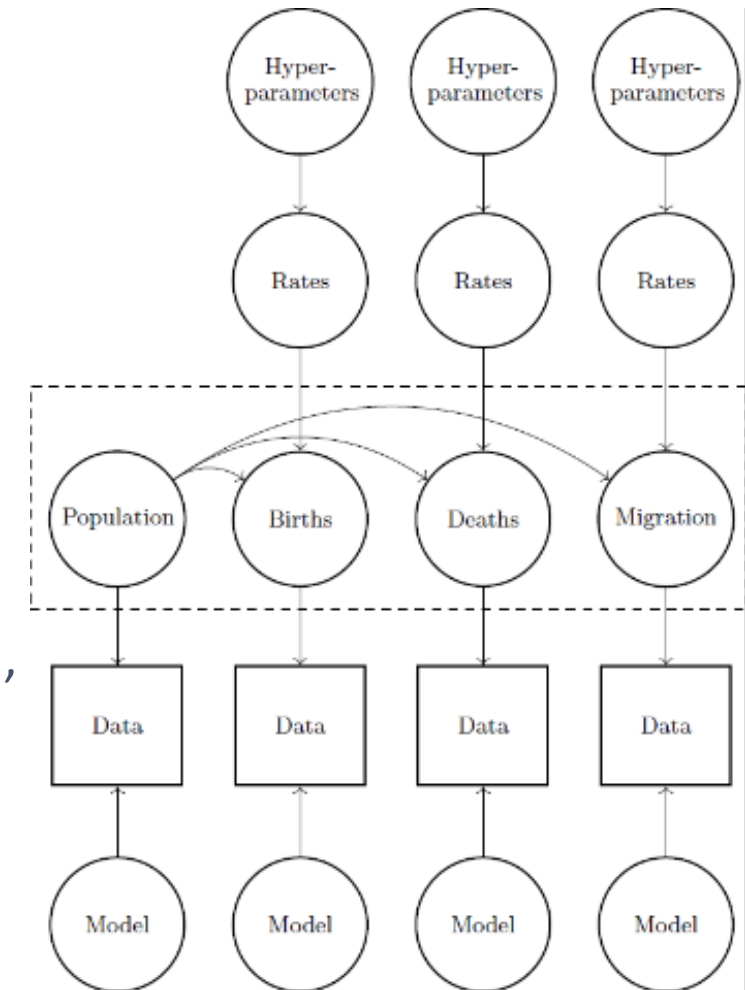
- Демографический эквивалент национального счета
- Стандартизированный формат
- Согласованность между запасами и потоками

Модели системы

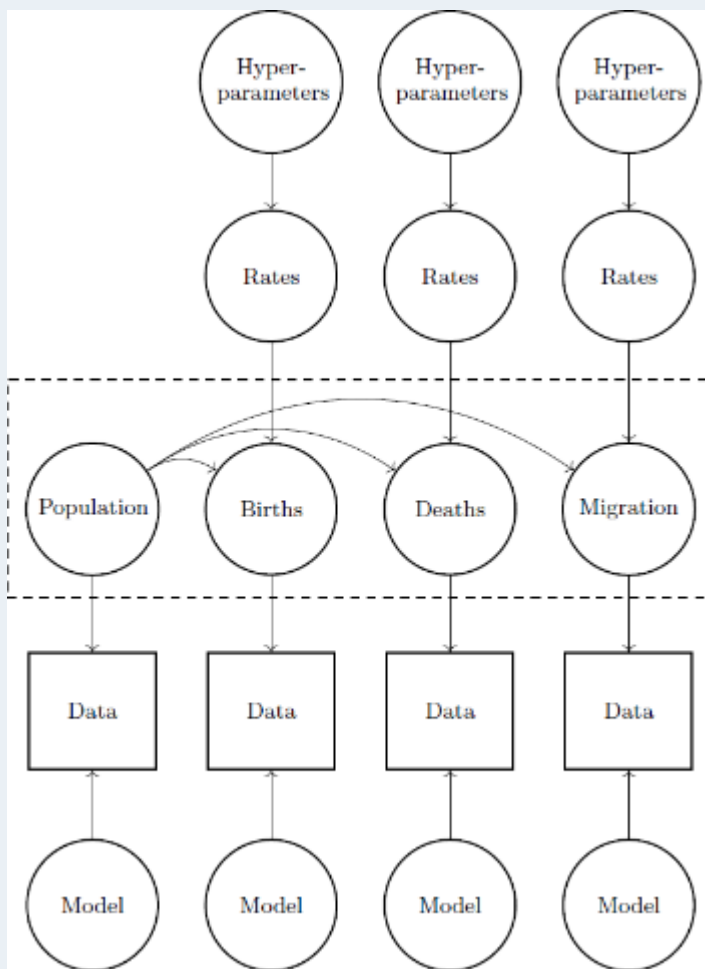
- Статистические модели закономерностей рождаемости, смертности, миграции
- Формальное представление о том, что опытный аналитик знает о демографических тенденциях

Модели данных

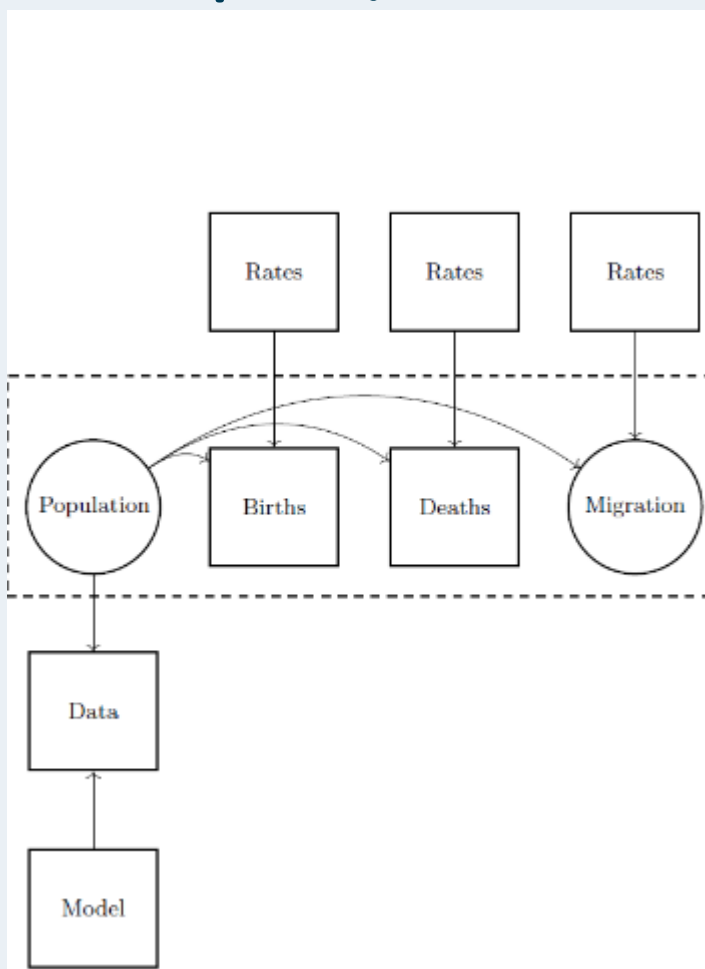
- Измерение статистических моделей, ошибки охвата
- Формальное представление того, что опытный аналитик знает о качестве данных



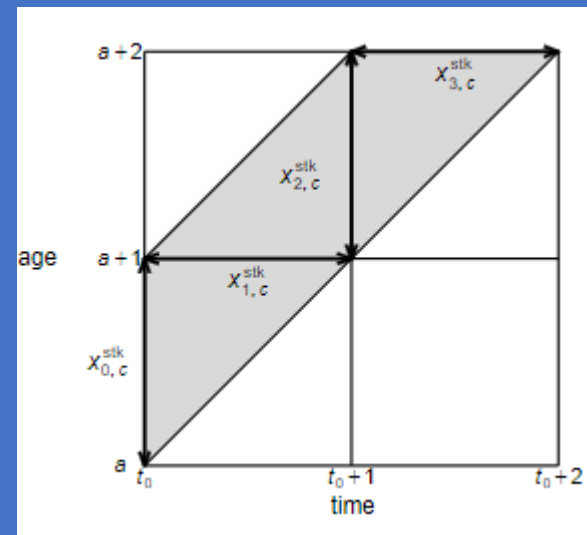
Полная модель



Упрощенная



Частотные фильтры в когортах



Окна оценки для динамической модели населения

Исходные предположения: 1) отставание в получении доступа к компонентам данных на 15 месяцев
(максимальное отставание в получении данных о международной миграции?)

2) ключевые предикторы (показатели, необходимые для прогнозирования) доступны ежемесячно

Demographic accounts		2021					2022												2023									
		Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Monthly Reporting	Estimates																											
	Jun-22	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
	Jul-22	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
	Aug-22	Blue	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
	Sep-22	Blue	Blue	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
	Oct-22	Blue	Blue	Blue	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
	Nov-22	Blue	Blue	Blue	Blue	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
	Dec-22	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
	Jan-23	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
	Feb-23	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange
Mar-23	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	
Apr-23	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	Orange	

Окончательные оценки с помощью модели динамики популяции (МДП) с использованием окончательных предикторов

Лучшие ежемесячные предварительные оценки с помощью модели динамики популяции (МДП) на основе неполных предикторов

Прогнозы на базе исторических временных рядов и мнений экспертов

Преимущества

Автоматизация

- *Возможность частого обновления, получения подробных выходных данных.*

Прозрачность

- *Воспроизводимость.*
- *Возможность экспериментировать.*

Гибкость

- *Легко изменять входные / выходные данные.*

Неопределенность

- *Формальные меры неопределенности.*

Своевременность

- *Используются ранние данные агрегированного уровня.*

Недостатки

Вычисление

Как и многие байесовские вычисления, современные методы работают медленно. Текущее программное обеспечение не масштабируется до 330 местных органов власти.

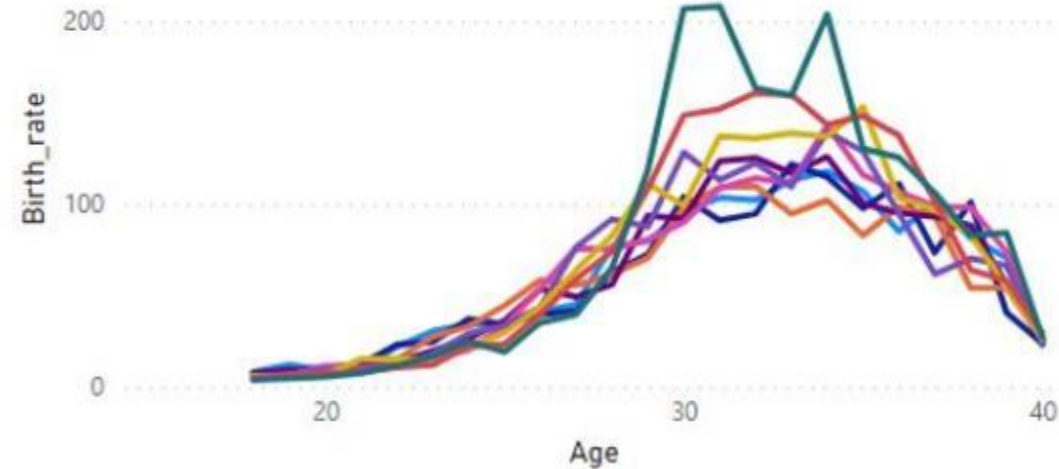
Высокотехнологичность

*Нужно думать о коммуникациях с пользователями
У Управления национальной статистики (УНС) должен быть специализированный потенциал статистических вычислений.*

Разработка данных

Birth_rate by Mother's Age and Year

Year ● 2012 ● 2013 ● 2014 ● 2015 ● 2016 ● 2017 ● 2018 ● 2019 ● 2020



- Информационная панель данных в реальном времени
 - Проверки входных данных
 - Проверки выходящих данных
 - Калибровка прогнозов
- Формальная статистическая база выявляет проблемы с данными
- Привлечение демографов к оценке входящих данных

Результативность модели

Мы тестируем производительность модель динамики популяции (МДП) с помощью моделирования.

Сюда входит моделирование демографического счета (население и его потоки) за 2011–2020 годы, чтобы взять его в качестве «золотого стандарта истины», которую мы пытаемся оценить.

Затем мы проверяем, как модель динамики популяции (МДП) работает в различных сценариях.

Сценарии следующие:

- ввод и вывод объективных данных о численности населения;
- комбинации различных типов ошибок в показателях потоков;

На данный момент мы установили:

- Модель динамики популяции (МДП) дает более точные оценки, если мы включаем в оценку непредвзятые данные о численности населения.
 - Например, если наши оценки миграции необъективны, то модель динамики популяции (МДП) с данными о численности населения вносит лучшие поправки, чем если бы данные о численности населения не учитывались.
 - Если и иммиграция, и эмиграция либо недооцениваются, либо переоцениваются, то маргинальное преимущество от включения данных о численности населения в оценку невелики.
 - Если один показатель переоценен, а другой недооценен, то маргинальное преимущество будет более существенным.

Пример: женщины в университетском городке

Когда оценки иммиграции и эмиграции завышены на 5%:

	Относительное смещение	
	без запаса	с запасом
Иммиграция	5,0%	4,8%
Эмиграция	5,1%	4,8%
Население	0,04%	0,02%

Когда оценка иммиграции завышена, а оценки эмиграция занижена на 5%:

	Относительное смещение	
	без запаса	с запасом
Иммиграция	4,8%	1,4%
Эмиграция	-3,5%	1,4%
Население	1,30%	0,02%

2012 г., 100 циклов

Разработка модели

- Переход на эндогенные показатели
- Разработка моделей данных
 - Вклад демографов
 - Связь между переписью и инфраструктурой данных
 - Оценки неопределенности
- Специальные корректировки данных о населении
- Распараллеливание

Публикации

Июль 2022 г.

Синтетическая местная власть

Сентябрь 2022 г.

14 тематических исследований LA,
2022 г.

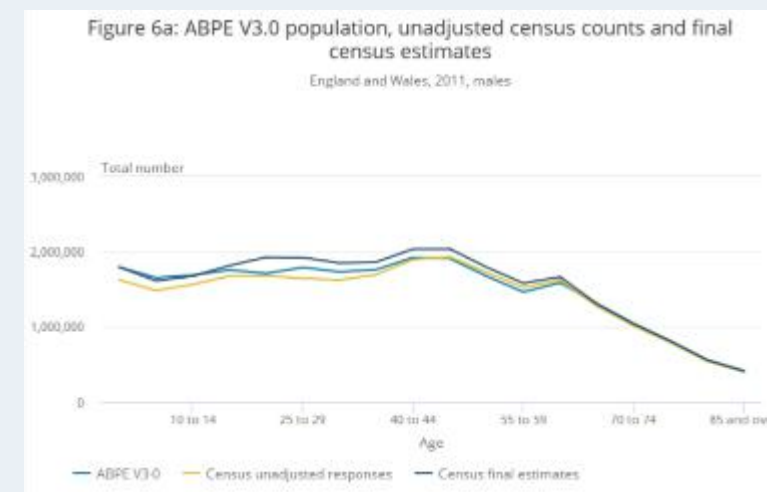
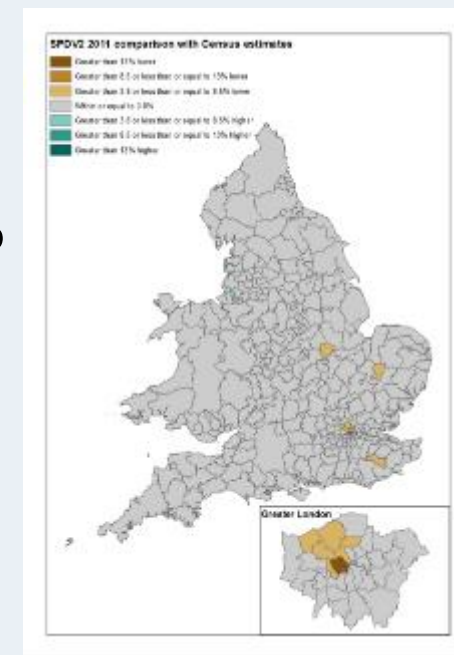
Декабрь 2022 г.

Все лонгитюдные оценки (ЛО)
2022 г.

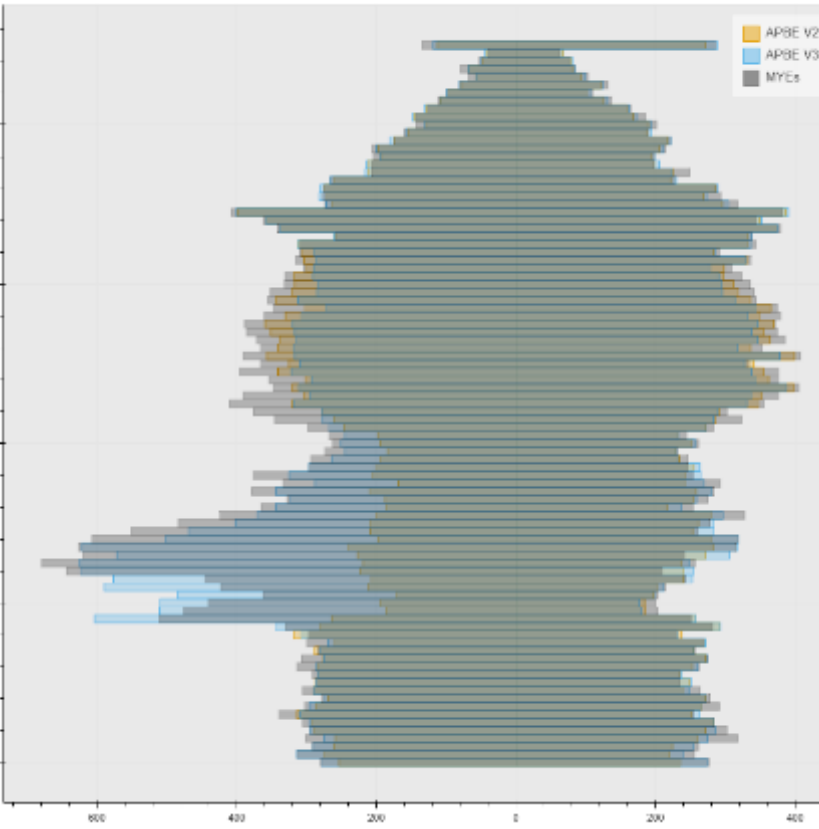
Преобразованная статистика населения: прогресс, достигнутый на сегодняшний день

- Оценки численности **населения**

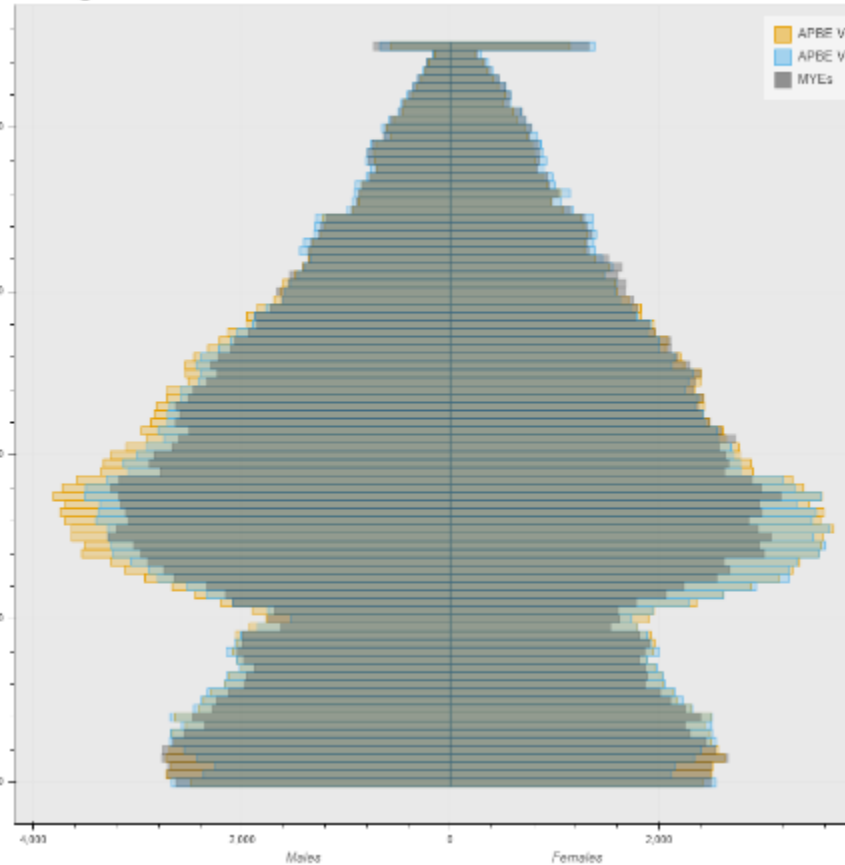
- Наборы статистических данных о населении (НСДН), созданные путем применения правил к связанным источникам данных для оптимизации охвата (т.е. для обеспечения точности данных).
- Наши НСДН использовались во время работы по обеспечению качества переписи, и от местных органов власти были получены положительные отзывы.



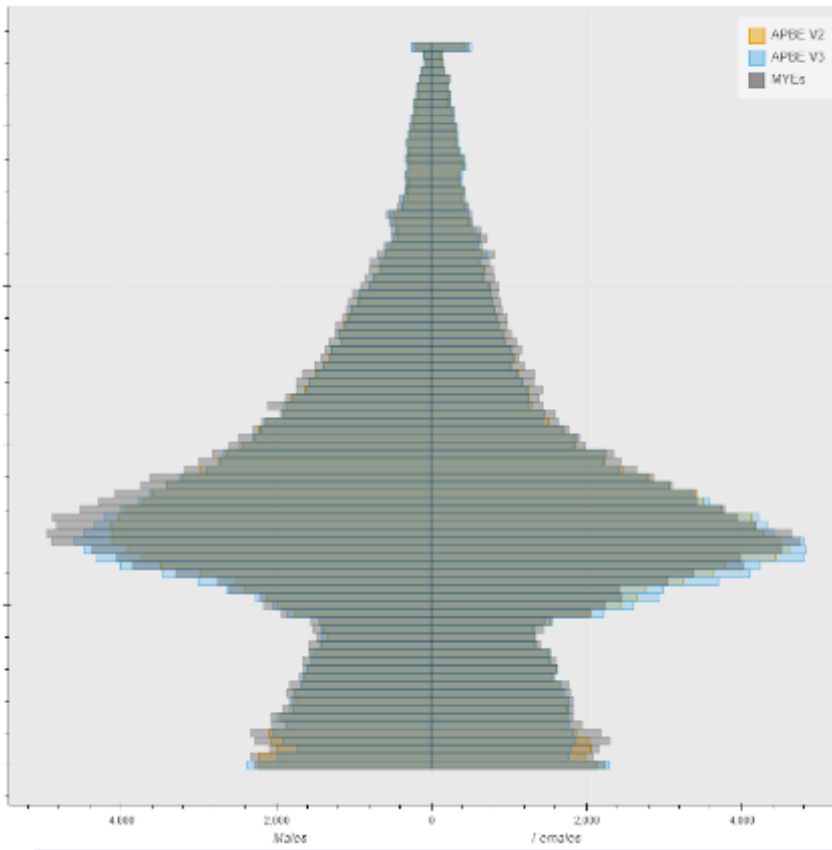
Richmondshire mid-2016

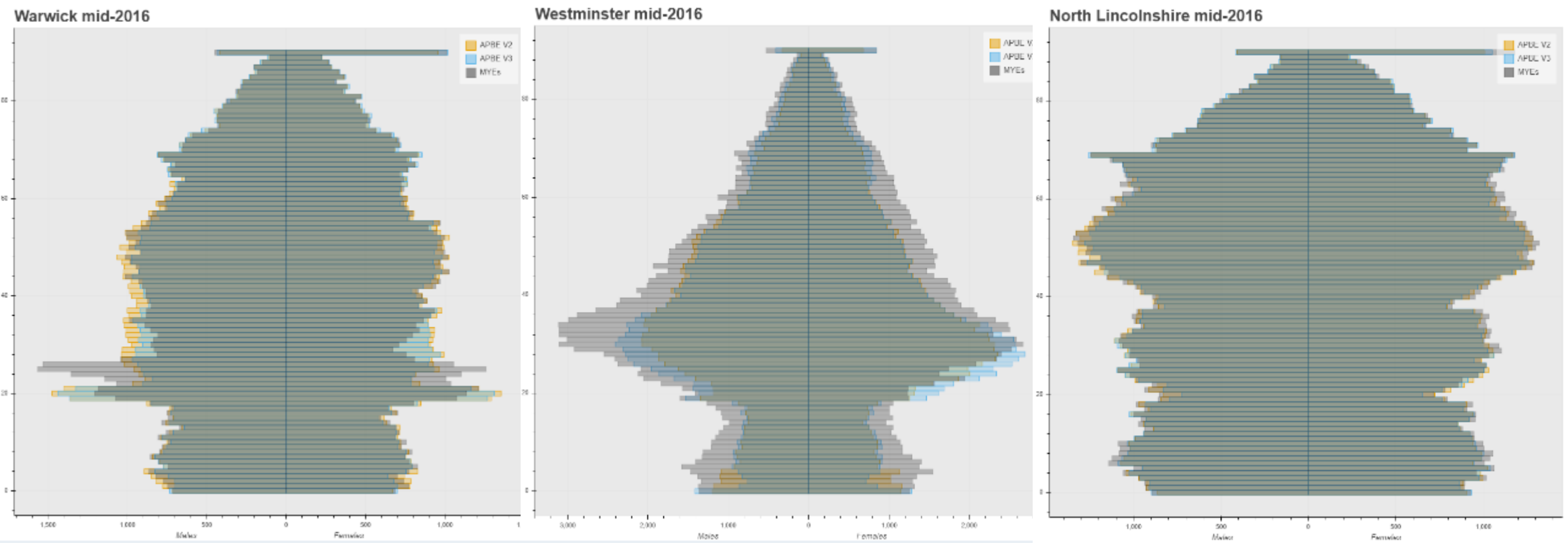


Ealing mid-2016



Tower Hamlets mid-2016





Серые полосы — официальные полугодовые оценки
 Желтые полосы — ABPE v2 (административная оценка численности населения, версия 2)
 Синие полосы — ABPE v3 (административная оценка численности населения, версия 3)

Выводы: это не универсальный подход
 Следующие этапы — использование того, что мы узнали, для разработки SPDv4

Исторический подход к оценке миграции

Международный опрос пассажиров (МОП)

Оценки на основе намерений мигрантов



Миграционные потоки
Северной Ирландии
(NISRA)

Просители убежища
(Управление внутренних
дел)

Адаптация к людям,
меняющим свои
намерения



Долгосрочная
международная
миграция (ДСММ)



Migration Statistics Quarterly Report: August 2020

A summary of the latest long-term international migration estimates for the UK for the year ending March 2020. Data from the Home Office, Department for Work and Pensions (DWP) and Higher Education Statistics Agency (HESA) are also included.

This is the latest release. View previous releases

Correction

Contact:
Migrants

Release date:
27 July 2021

Next release:
To be announced

Table of contents

1. Of the migration statistics for the year ending March 2020
2. Long-term international migration statistics
3. Net migration
4. Of the migration statistics for the year ending March 2020
5. Contact and migration overview
6. Of the migration statistics for the year ending March 2020
7. International migration data
8. Of the migration statistics for the year ending March 2020
9. Of the migration statistics for the year ending March 2020
10. Of the migration statistics for the year ending March 2020
11. Of the migration statistics for the year ending March 2020

View the full report 

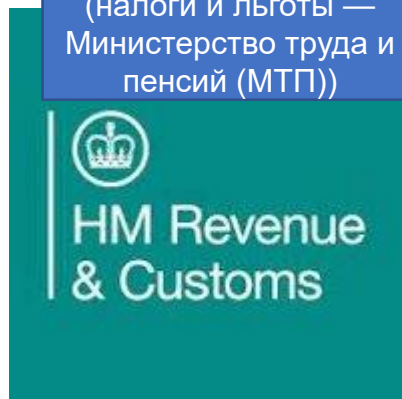
Download PDF 

Статистика миграции на базе административных данных — май 2022 г.

Страны, не входящие в ЕС = визовые данные
Управления внутренних дел



ЕС = База данных регистрации и взаимодействия населения (БДРВН/RAPID) (налоги и льготы — Министерство труда и пенсий (МТП))



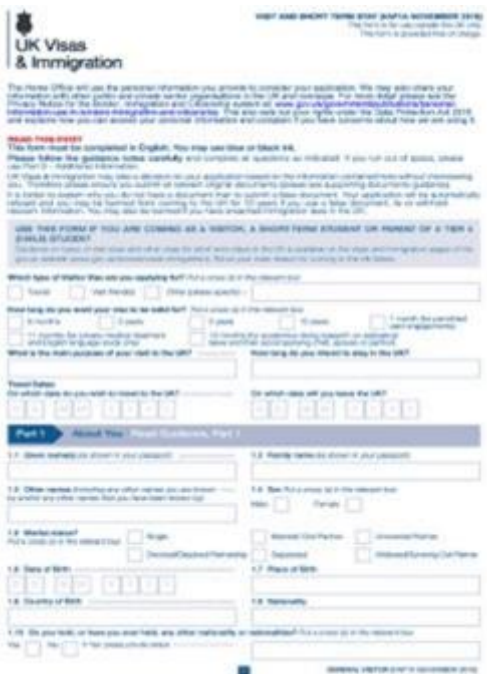
Department for Work & Pensions



Граждане Великобритании = предыдущие (на основе Международных опросов пассажиров (МОП)) смоделированные оценки



Иммиграция Эмиграция
Чистая миграция
ЕС, не ЕС и
Великобритания



Реализация нашего видения трансформации статистики миграции и населения

«Использовать наилучшую информацию для получения более частых, своевременных и всеобъемлющих статистических данных, которые являются согласующимися и гибкими для того, чтобы подстраиваться под меняющиеся потребности пользователей статистики»

Подкрепление методов новыми источниками

Своевременность

Согласованность

Причина миграции

Альтернативные определения

Социальная статистика – в первую очередь административные данные

- **Цель:** трансформировать метод, применяемый в УНС для получения социальной статистики по ряду основных направлений (характеристики населения / домохозяйств), чтобы обеспечить уровень детализации и географическую разбивку, которые требуются пользователям статистики.
- **Приоритеты на 2022 год:** представить результаты двух многовариантных тематических исследований, проведенных на базе административных данных, в качестве информационной основы для выработки рекомендаций в отношении переписи 2023 года, которые бы охватывали следующие аспекты:
 - Доход в разбивке по этнической принадлежности;
 - Жилье в разбивке по этнической принадлежности.

Успешность использования административных данных: жилой фонд

Характеристики жилого фонда, взятые из административных данных

Тип недвижимости, количество комнат (спален), общая площадь

- Изучение возможности использования административных данных о жилом фонде вместо сбора этой информации в ходе переписей и обследований.
- Сравнение данных о видах недвижимости, зарегистрированных Агентством по оценке недвижимости (АОН), с данными, собранными в ходе переписи 2011 года в Англии и Уэльсе.
- Точное соответствие между переменной вида недвижимости согласно АОН и переменной типа жилья согласно переписи установлено у 86% связанных адресов.

Agreement between census accommodation type and VOA property type



Source: Admin-based statistics for property type, feasibility research: England and Wales

Office for National Statistics

Успешность использования административных данных: этническая принадлежность, уровень дохода, рынок труда

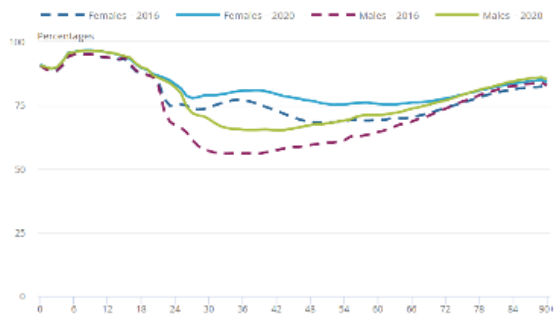
Статистика этнической принадлежности на основе административных данных (ABES)

Информация об этнической принадлежности, взятая из Обследования занятости домохозяйств (ОЗД/НЕС), Института дополнительного профессионального обучения (ИДПО/ИАРТ), Статистики неотложной помощи (СНП/ECDS), Статистического агентства высшего образования (CAVO/HESA), уведомлений о рождении, связанная с административной оценкой численности населения версии 3 от 2016 года (2016 ABPE V3) и с статистическими данными об этнической принадлежности на основе административных данных.

Возможность установить этническую принадлежность 74,7% лиц, внесенных в административную базу данных населения 2016 года.

Figure 2: The proportion of people with a stated ethnicity increased for all ages between 2016 and 2020

Proportion of people in the ABPE with a stated ethnicity by age and sex, 2016 and 2020 admin-based ethnicity statistics, England



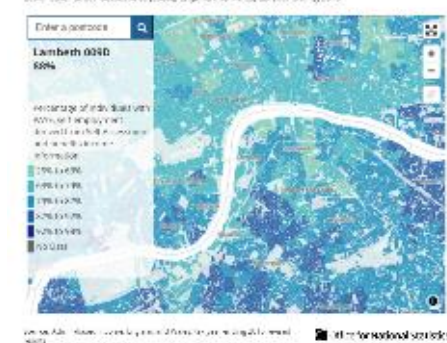
Статистика доходов на основе административных данных (ABIS)

Объединение данных Королевской налоговой-таможенной службы (HMRC) и данных Министерства труда и пенсий (DWP), увязка этих данных оценками численности населения и домохозяйств на базе административных данных.

Оценки чистого дохода отдельных лиц и домохозяйств в рамках Лонгитюдного исследования старения населения (LSOA) (оценки ниже, чем данные УНС, основанные на результатах обследований).

98,4% занятых адресов (домохозяйств) располагают информацией о доходах из одного источника.

Percentage of individuals with PAYE, self-employment derived from Self-Assessment and benefits income information



Административные данные о статистике рынка труда (ABLMS)

Охват высокий: 95% лиц в возрасте от 16 лет и старше (по нашей базе населения) присваивается статус на рынке труда на базе административных данных (ABLMS).

Figure 5: Percentage of individuals assigned an ABLMS remains at or above 90% from age 20

Percentage of individuals with an assigned ABLMS, England and Wales, tax year ending 2016



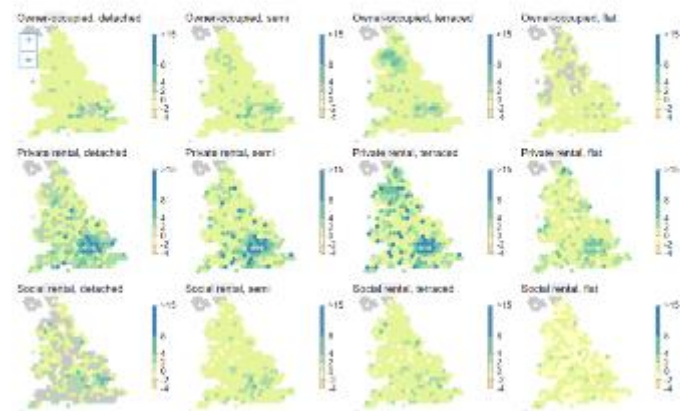
Успешность использования административных данных: перенаселенность, квалификации

Научное исследование целесообразности перенаселенности на базе административных данных

Оценка влияния замены вопроса переписи 2021 года о количестве комнат путем сравнения субрегиональных уровней перенаселенности.

Percentage point (pp) differences in the levels of overcrowding at LA level for 2011 Census number of bedrooms and VOA number of bedrooms using the bedroom standard by tenure and accommodation type, England & Wales

Households are identified as overcrowded if they have a bedroom occupancy rating of "1 or less" using the bedroom standard.



Source: Admin-based levels of overcrowding (using the bedroom standard and VOA, based on Office for National Statistics, feasibility research: England and Wales, January 2021)

Office for National Statistics

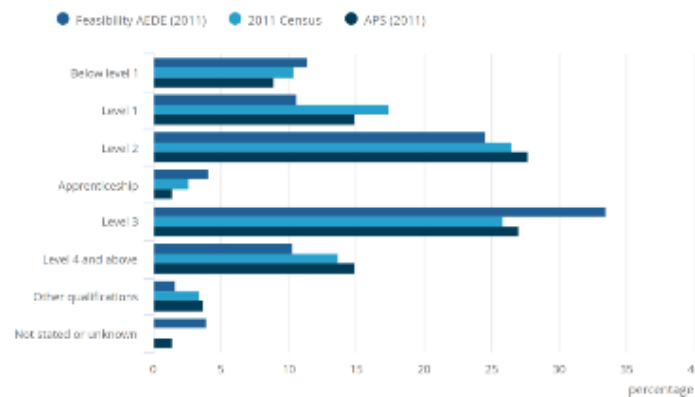
Научное исследование целесообразности высочайшего уровня квалификаций на базе административных данных

Сравнение административных данных и данных переписи 2011 года показало, что у 57% людей самый высокий уровень квалификации имел одинаковое значение у обоих источников данных.

У 84% людей наивысший квалификационный уровень по административным данным либо совпадал, либо находился в пределах того же уровня, что и уровень, зафиксированный переписью.

Administrative data recorded a lower percentage of individuals with "Level 1" and "Level 4 and above" qualifications but a higher percentage with "Level 3"

Highest level of qualification in 2011 for usual residents aged 16 to 24 years, from the feasibility AEDE, 2011 Census and APS (percent distribution), England



Source: Admin-based qualification statistics, feasibility research: England

Office for National Statistics