

Что такое  
**журналистика данных**

Алексей Смагин | ВШЭ 2021

**журналистика** **данных**

**Данные** — представление фактов, понятий или инструкций в форме, приемлемой для общения, интерпретации, или обработки человеком или с помощью автоматических средств

**Данные** – представление фактов, понятий или инструкций в форме, приемлемой для общения, интерпретации, или обработки человеком или с помощью автоматических средств

В общем, **данные** — это любая  
информация, имеющая хоть какой-  
то смысл

**маленький**

42

**или большой**

Высшая ценность человеческого существования, выраженная в форме конечной цели, которой подчинены все частные интересы и потребности человека. Среди многих подходов к понятию этой

а не 'цифры' или 'статистика'

Когда в данных можно выделить  
**объекты, обладающие общими**  
**свойствами**, информацию будет  
нести не только каждый объект в  
отдельности, но и их совокупность

# Объекты — сотрудники на сайте ВШЭ



**Абрамов Роман Сергеевич**

Методист, Преподаватель: [Факультет коммуникаций, медиа и дизайна / Департамент медиа](#)

Реклама и PR

медиакоммуникации

медиапроизводство

Web-дизайн

digital маркетинг

графический дизайн

видео



**Адемукова Надежда Владимировна**

Старший преподаватель: [Факультет коммуникаций, медиа и дизайна / Департамент медиа](#)

социология

медиакоммуникации

массовая коммуникация

журналистика

измерения аудитории тел...

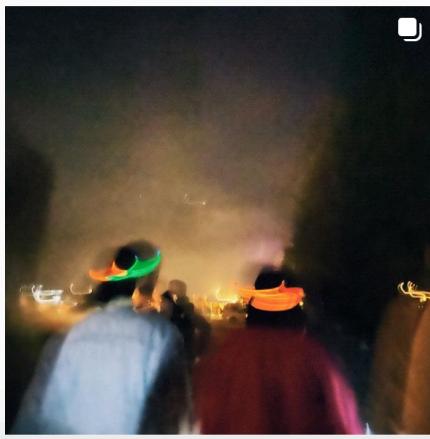
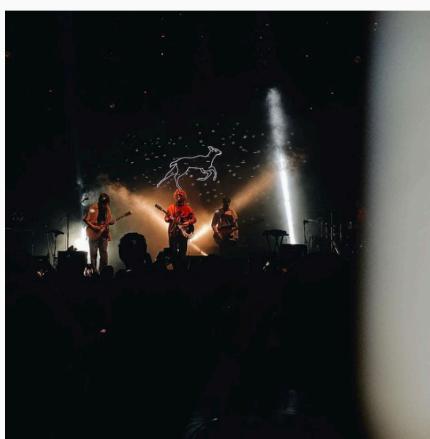
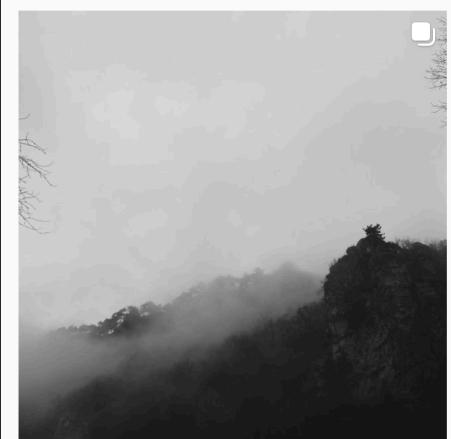
Можно узнать должность для каждого сотрудника в отдельности, а можно — сотрудников каких должностей на департаменте больше всего

# Объекты — файлы в папке Загрузки

Имя	Размер	Тип	Дата добавления
2. Сведения о доходах.docx	190 КБ	Micros... (.docx)	15 февр. 2020 г., 03:48
2.jpg	3,2 МБ	JPEG	18 марта 2020 г., 23:23
3 (1).jpg	4,9 МБ	JPEG	19 марта 2020 г., 01:32
3-2-3_2019 (1).doc	100 КБ	Micros...t (.doc)	10 июля 2020 г., 12:29
3-2-3_2019.doc	100 КБ	Micros...t (.doc)	24 июня 2020 г., 14:04
3-2-3-2017 (1).doc	87 КБ	Micros...t (.doc)	10 июля 2020 г., 12:28
3-2-3-2017.doc	87 КБ	Micros...t (.doc)	24 июня 2020 г., 14:10
3-4 (1).doc	439 КБ	Micros...t (.doc)	28 апр. 2020 г., 14:47
3-4.doc	439 КБ	Micros...t (.doc)	28 апр. 2020 г., 12:39
3.jpg	4,9 МБ	JPEG	18 марта 2020 г., 23:23
04-04.doc	49 КБ	Micros...t (.doc)	19 мая 2020 г., 17:44
05-05_2017-2018 (1).xls	34 КБ	Micros...ok (.xls)	10 июля 2020 г., 12:29

Файлов какого типа больше всего? Сколько места занимают все файлы в папке?

# Объекты — посты в Instagram



Как часто добавляются  
фото?  
Какой цвет на них  
преобладает?

# Объекты — предметы на моём столе



Сколько места это всё  
займёт, если я решу  
переехать?

Мы можем анализировать совокупность объектов с одинаковыми свойствами, чтобы получить новое знание

Это знание может быть полезным, важным, интересным

Журналистика данных — это способ создания материалов в медиа, при котором мы анализируем данные и получаем некоторое **новое знание**, интересное широкой аудитории

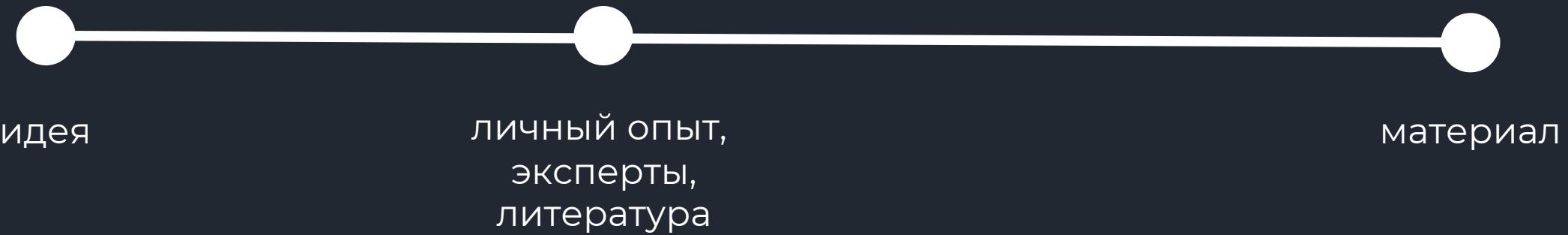
**большие данные**

**большие данные**

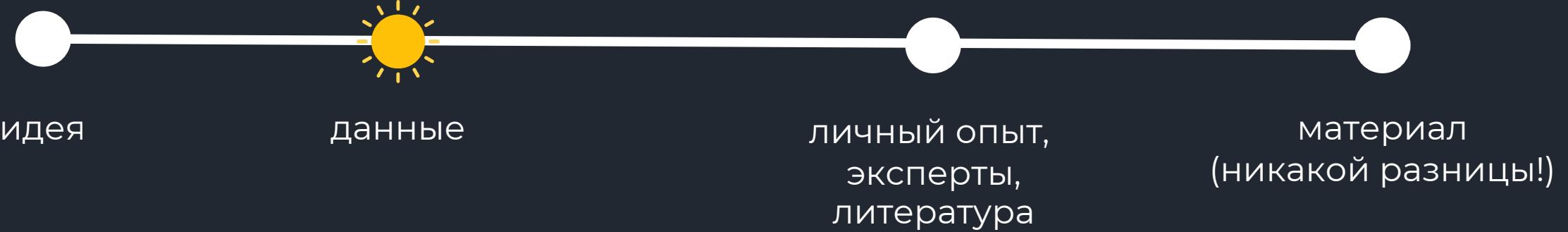
**Большие данные** — огромные потоки информации, поступающие регулярно (например, информация обо всех абонентах мобильной связи в стране, совершивших звонки за час)

В **журналистике данных** мы работаем чаще всего с сотнями или тысячами объектов. Реже — с миллионами (хотя это всё ещё не Big Data в строгом смысле)

# Как работает журналист



# Как работает **дата-журналист**



**журналистика  
данных  
в примерах**

**Методы работы дата-  
журналистов существовали  
задолго до развития  
компьютерной техники**

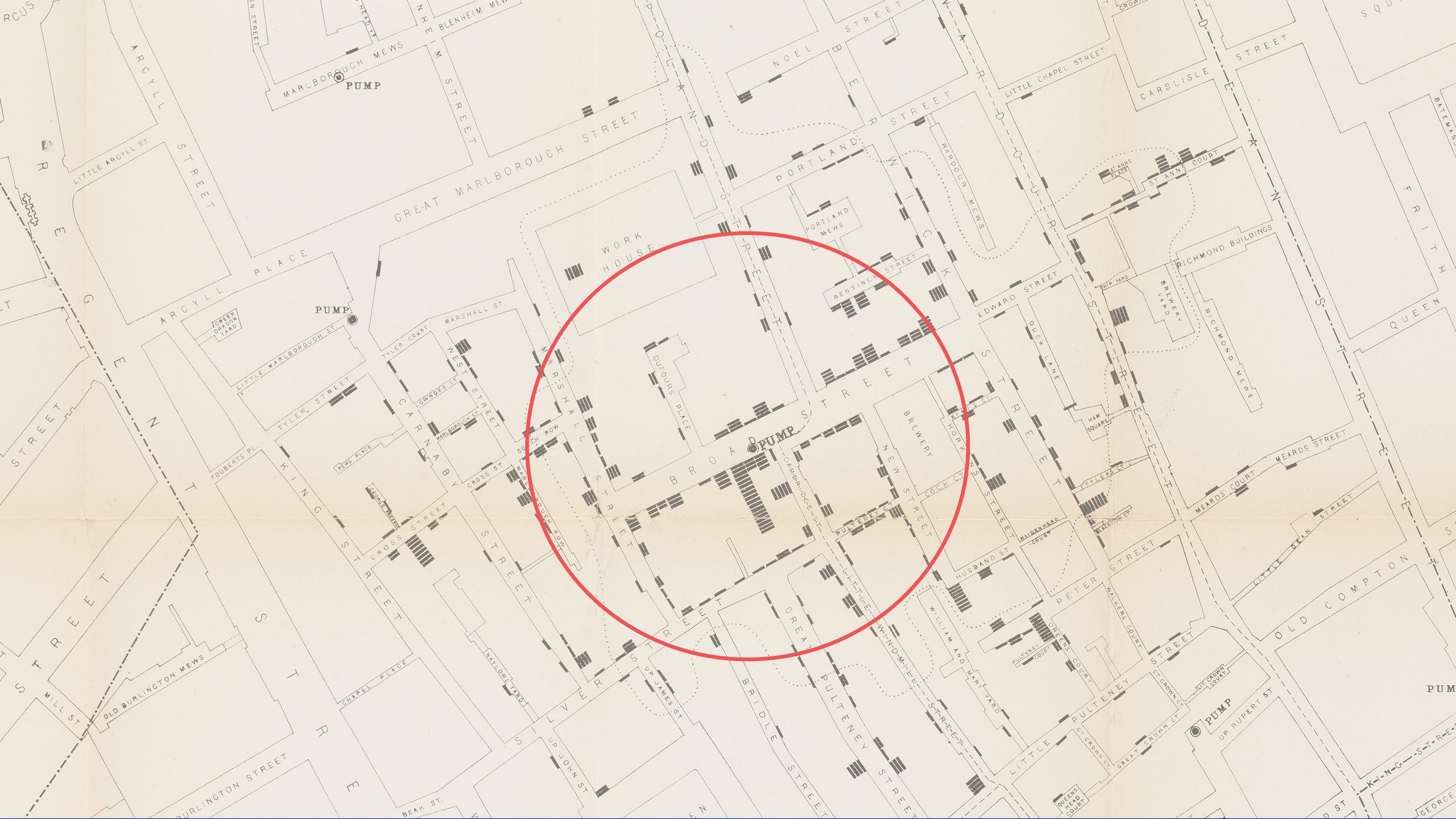
**1854 год.** В мире — вспышки холеры.  
умирают сотни тысяч людей в разных  
странах

популярна теория «миазмов», которая  
гласит, что люди заражаются, вдыхая  
нездоровы й воздух

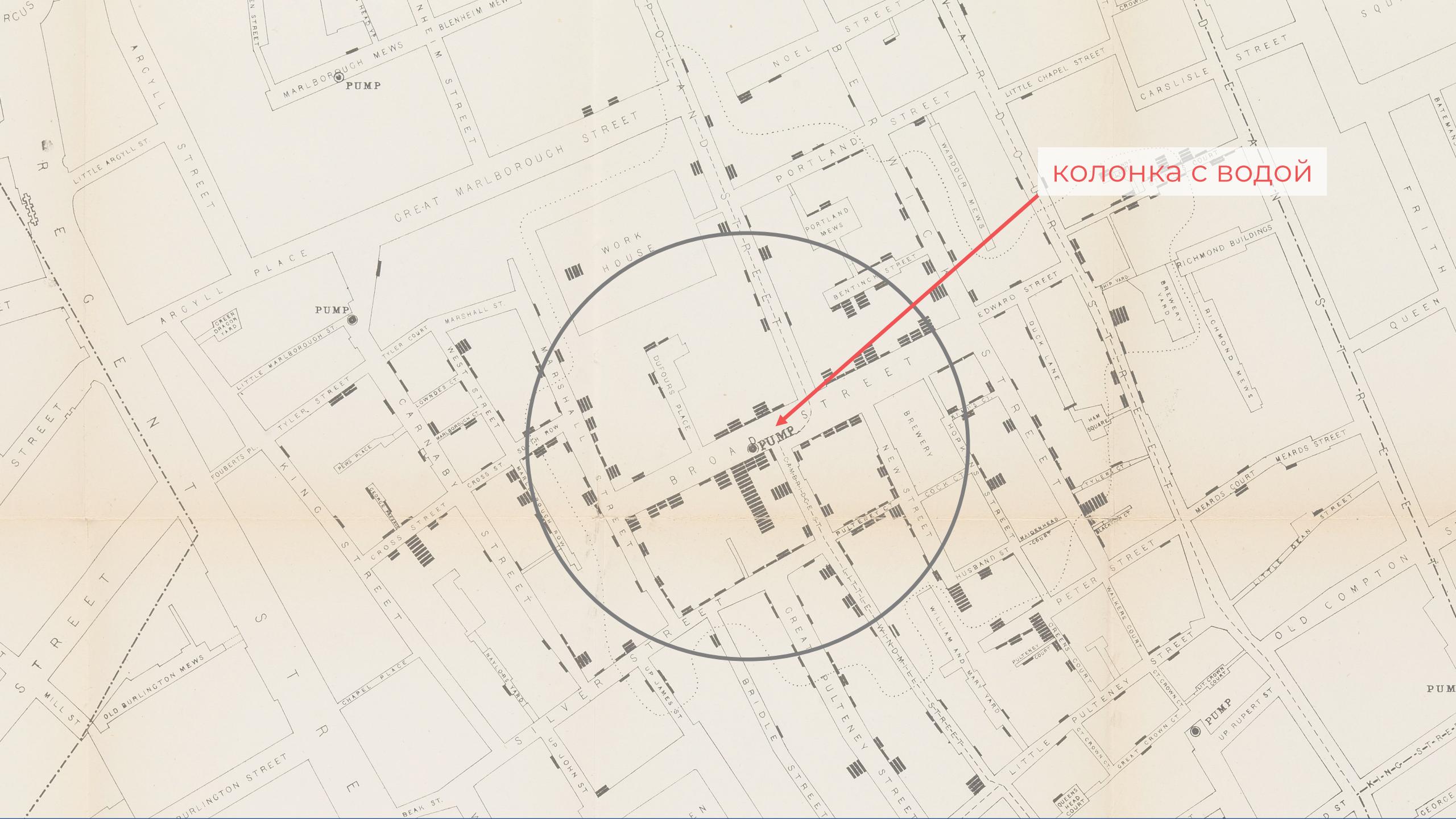


своё исследование проводит доктор Джон Сноу. он опрашивает местных жителей в Лондоне, отмечая на карте, в каких домах живут больные.





КОЛОНКА С ВОДОЙ

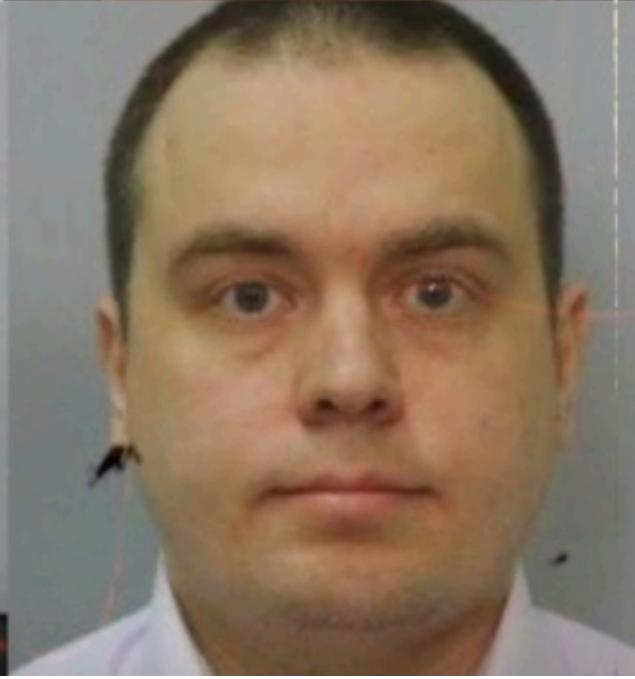
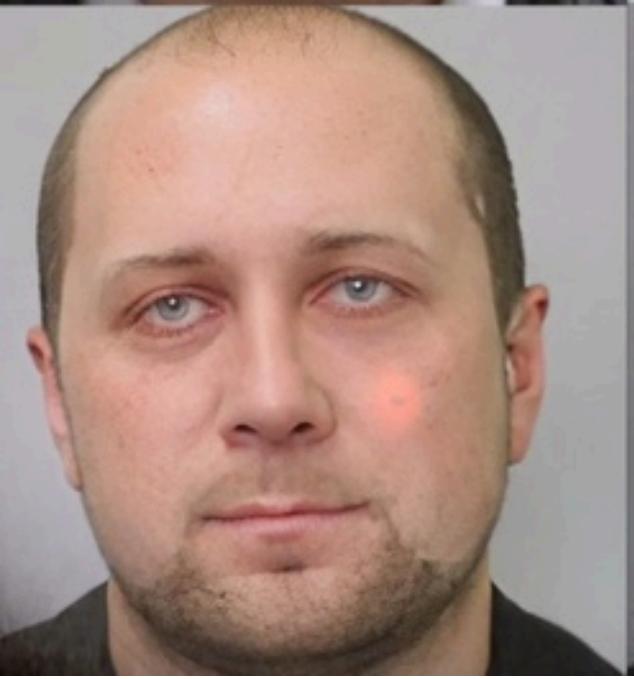




В монастыре рядом с  
пивоварней нет  
больных: монахи пили  
только пиво



**Работа с данными помогает  
узнавать вещи, которые не  
смогли бы найти обычные  
журналисты**



# **Методы журналистики данных в расследовании Bellingcat**

В расследовании Bellingcat использовались как традиционные методы работы с данными: ручная проверка сведений о сотрудниках ФСБ, чтение документов «глазами», так и методы **журналистики данных**

<https://ru.bellingcat.com/novosti/russia/2020/12/28/navalny-fsb/>

# **Методы журналистики данных в расследовании Bellingcat**



**Биллинги  
телефонных разговоров**



**Список пассажиров  
в самолётах**

## **Биллинги телефонных разговоров**

Расследователи «пробили» номера руководителей НЦ «Сигнал».

В биллингах содержится информация обо всех входящих и исходящих вызовах по данному номеру с указанием даты и времени.

# Биллинги телефонных разговоров

— Но вы же просто видите, что люди созываются, — как вы при таких ограниченных данных понимаете, что это был особенно важный рабочий разговор? Они же могли просто, например, задержки зарплат обсуждать?

— Здесь важно искать исключения. **Всплеск звонков**, в котором вдруг появляются **новые номера**, — это уже что-то. Например, в одном таком «пике» мы увидели номер [главы Центра специальной техники ФСБ генерал-майора Владимира] Богданова, их большого шефа, который вообще редко появлялся в звонках, — и мы проверили, кому он позвонил сразу после. И оказалось, что он звонил в «Сигнал». А потом еще двое участников этих переговоров звонят в «Сигнал». И мы понимаем, что этот всплеск должен быть как-то связан с «Новичком».

<https://meduza.io/feature/2020/12/16/ne-tak-to-prosto-otravit-cheloveka-novichkom>

# **Списки пассажиров в самолётах**

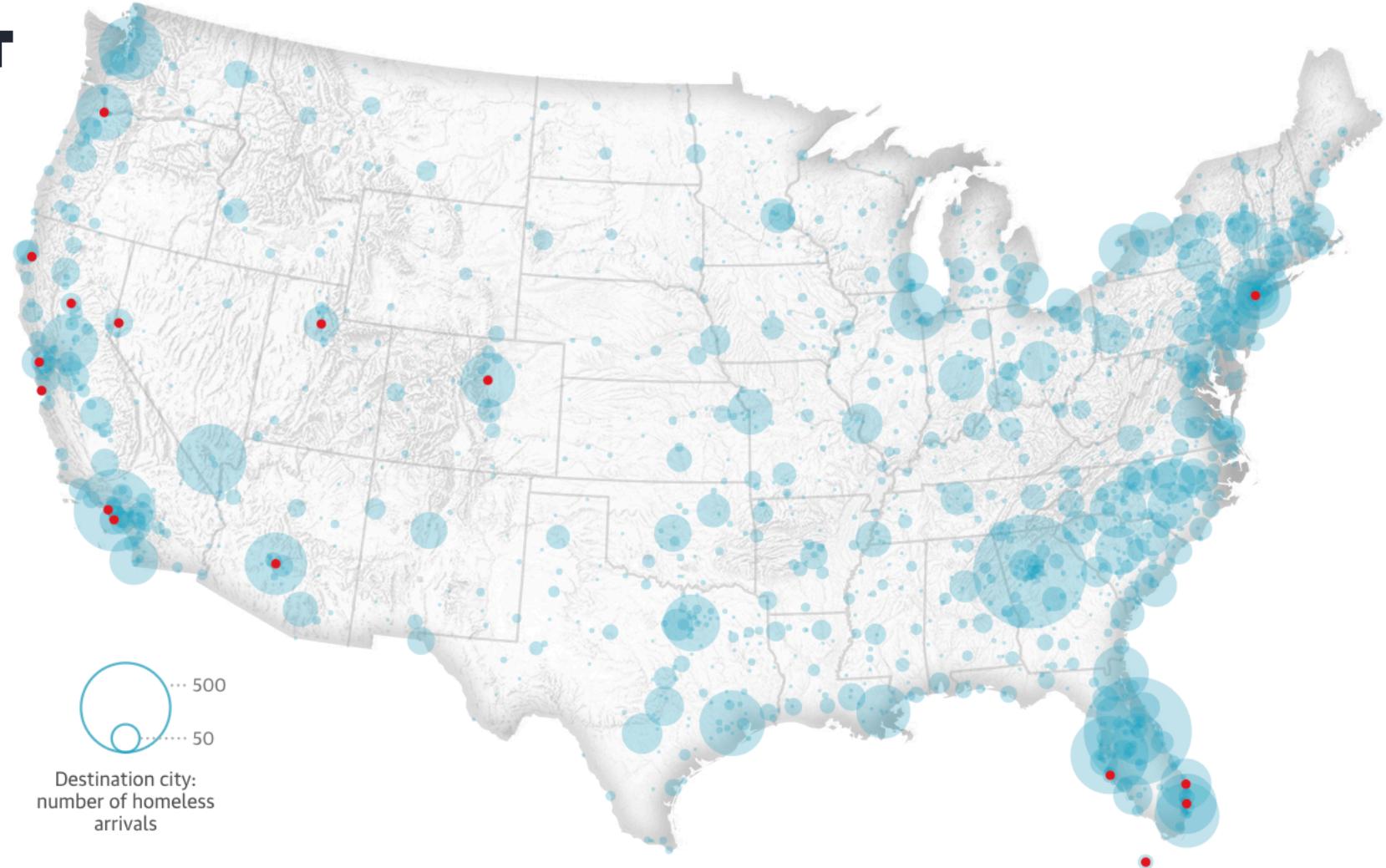
**Гипотеза:** сотрудники ФСБ летают вслед за Навальным в тот же день или с разницей в один день.

Расследователи покупают списки пассажиров на все рейсы в даты полетов Навального с разницей  $\pm$  день

Они проверяют совпадения в две даты: дату прилёта и дату отлёта. Так они выходят на группу из трёх сотрудников ФСБ, которые летали за Навальным в его роковую поездку.

**Журналистика данных тесно  
связана с традиционной  
журналистикой**

# Как переезжают бездомные?



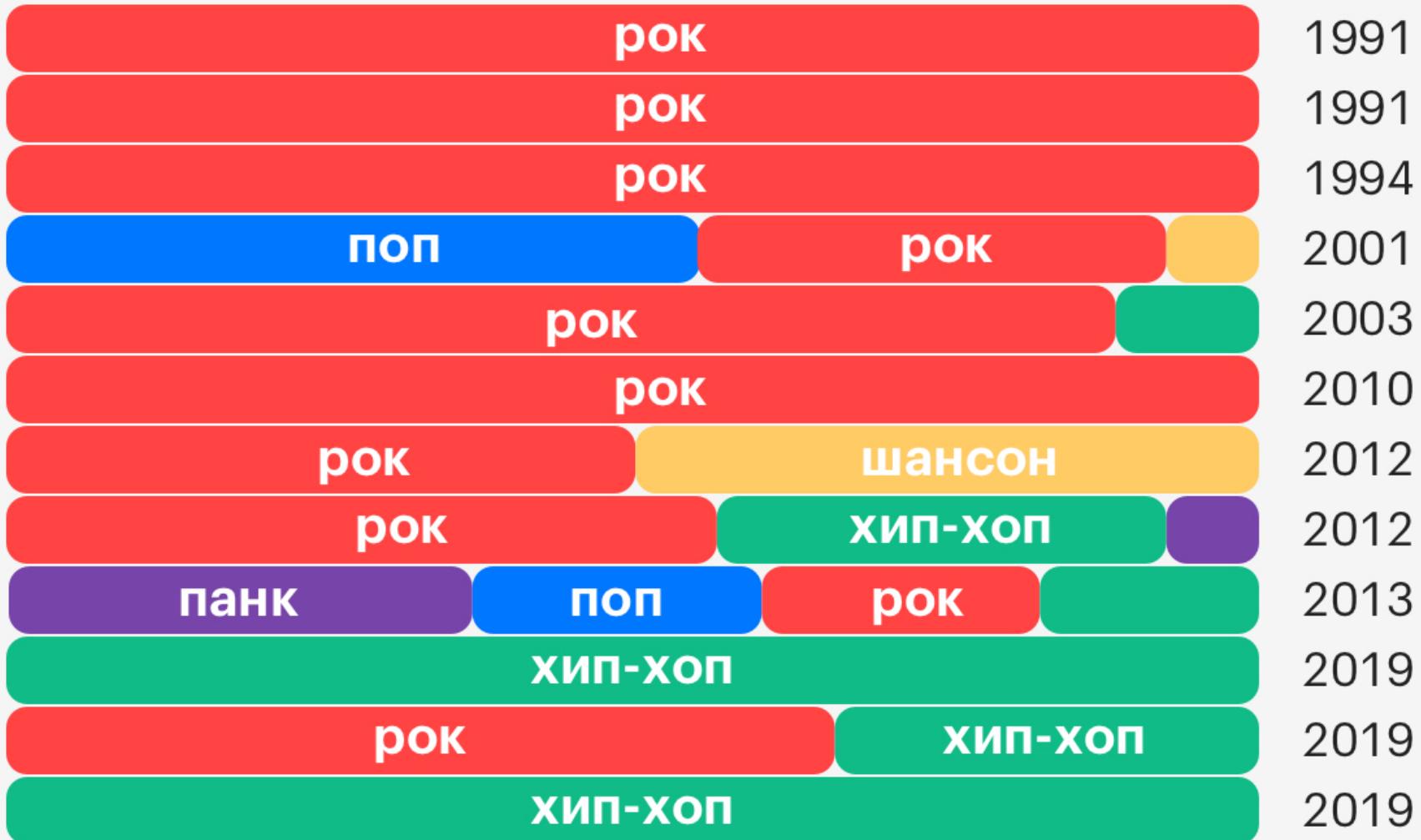
Homeless bus relocation journeys to destinations in the mainland US

**21,400**

Источник: The Guardian

**Журналистика данных — это  
не обязательно сложно**

# Кто выступал на российских митингах?



Источник: РБК

**Данные помогают понять  
масштаб проблем**



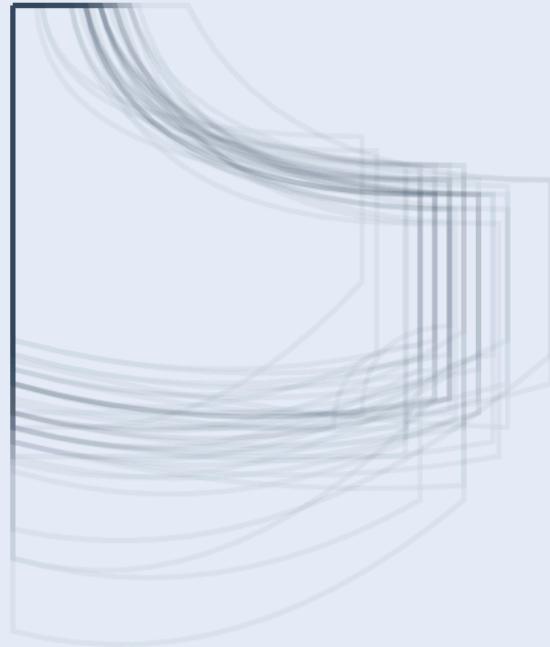
# Как избивали людей на митингах в Минске?

[Источник: Медиазона](#)

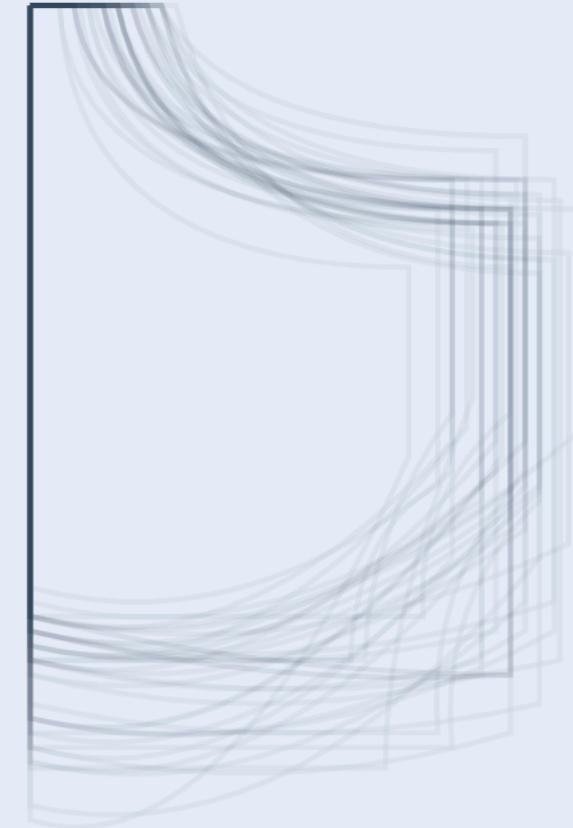
**... или наглядно их отразить**

# Карманы в мужских и женских джинсах

WOMEN



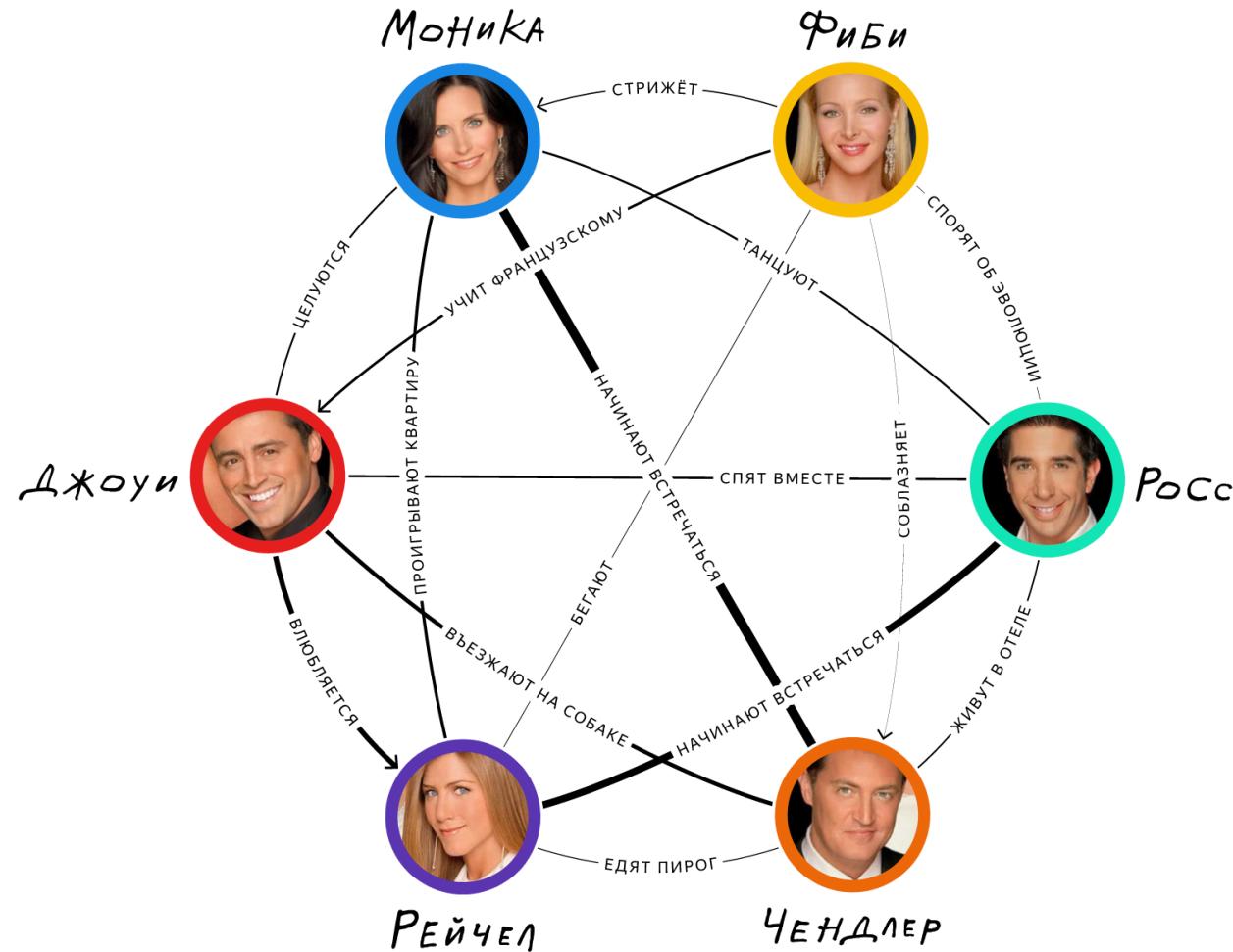
MEN



[Источник: The Pudding](#)

**Данные — это не скучно**

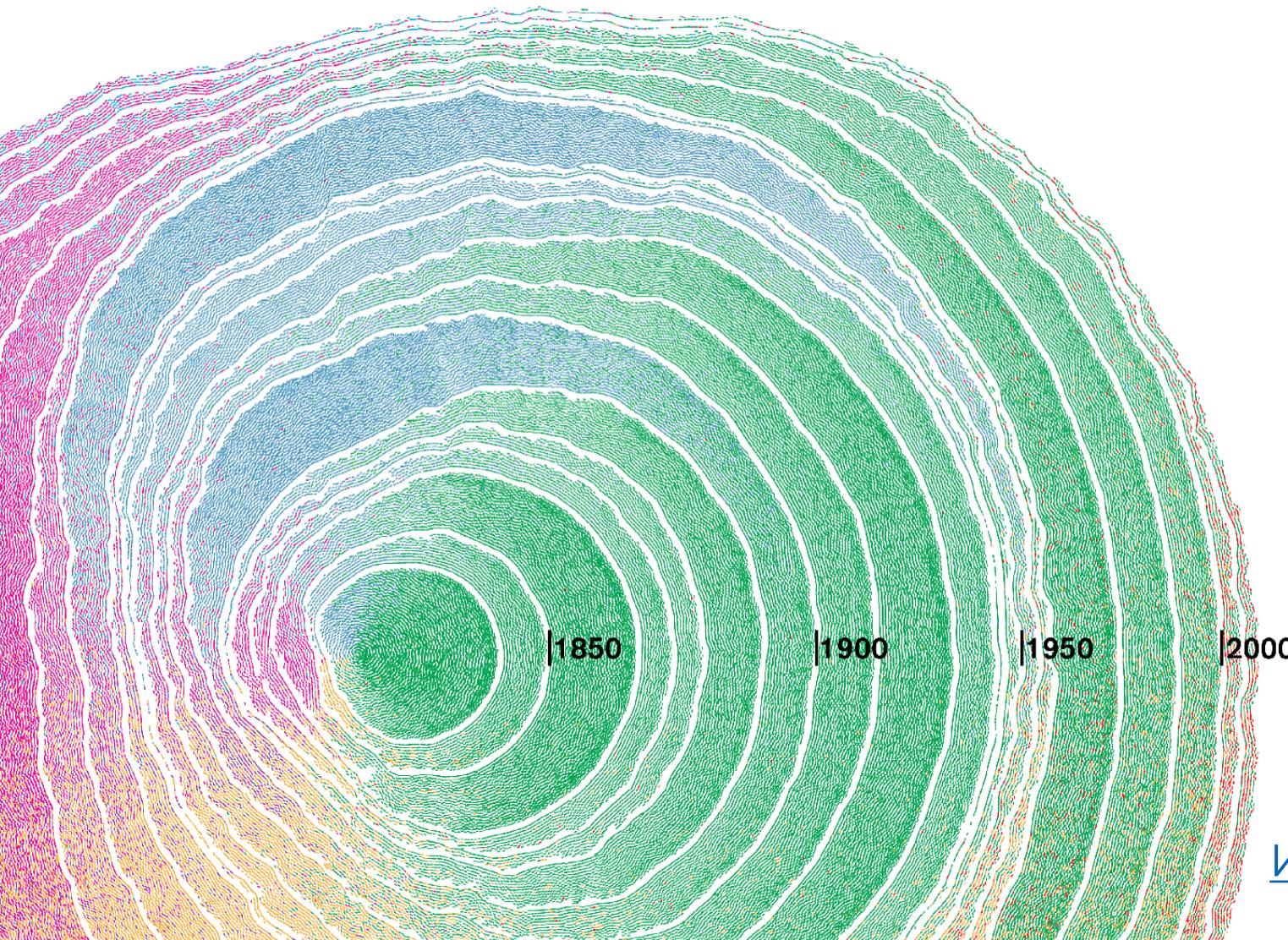
# Сериал «Друзья» в поисковых запросах



Источник: Исследования Яндекса

**В конце-концов, это  
прекрасно**

# История американской иммиграции



Источник: [Northeastern University](#)

# **Какие навыки нужны?**

## **Журналиста**

придумывать истории,  
писать текст, искать  
информацию

## **Аналитика**

работать с данными

## **Дизайнера**

делать инфографику и  
макеты сложных  
проектов

## **Программиста**

для сложной работы с  
данными и  
интерактивных  
проектов

**в добрый  
путь!**