

# **А что дальше?**

Алексей Смагин | ВШЭ 2021

# 1. Программирование для анализа данных: Python или R

# Зачем программировать?

Полная свобода. Вы не зависите от типов данных и программ, с которыми работаете

## Примеры:

- Информация о каждом регионе находится в собственном xls-файле, нужно их соединить
- Работа с большими объёмами данных
- Веб-скрейпинг Twitter
- Информацию с сайта можно получить только при помощи API

# Зачем программировать?

Автоматизация любых процессов.

## Примеры:

- Рейтинг кандидатов в президенты
- Рассылка госконтрактов «Важных историй»



# Зачем программировать?

Продвинутый текстовый анализ

## Примеры:

- Русский рэп как набор слов
- Фальсификация наркотических дел
- Женщины в СМИ

# Зачем программировать?

Машинное обучение

- Предсказание значений
- Классификация объектов

**Примеры:**

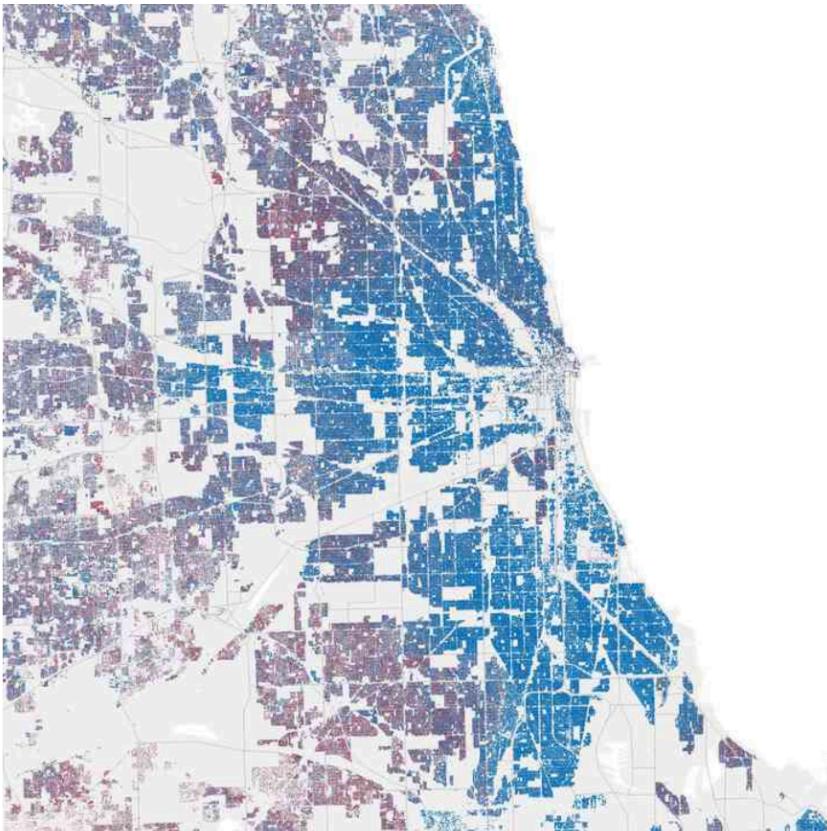
- Масштаб нелегальной добычи янтаря на Украине
- Ошибочно классифицированные преступления в ЛА

## **2. Визуализация данных и дизайн**

# Чему учиться в визуализации?

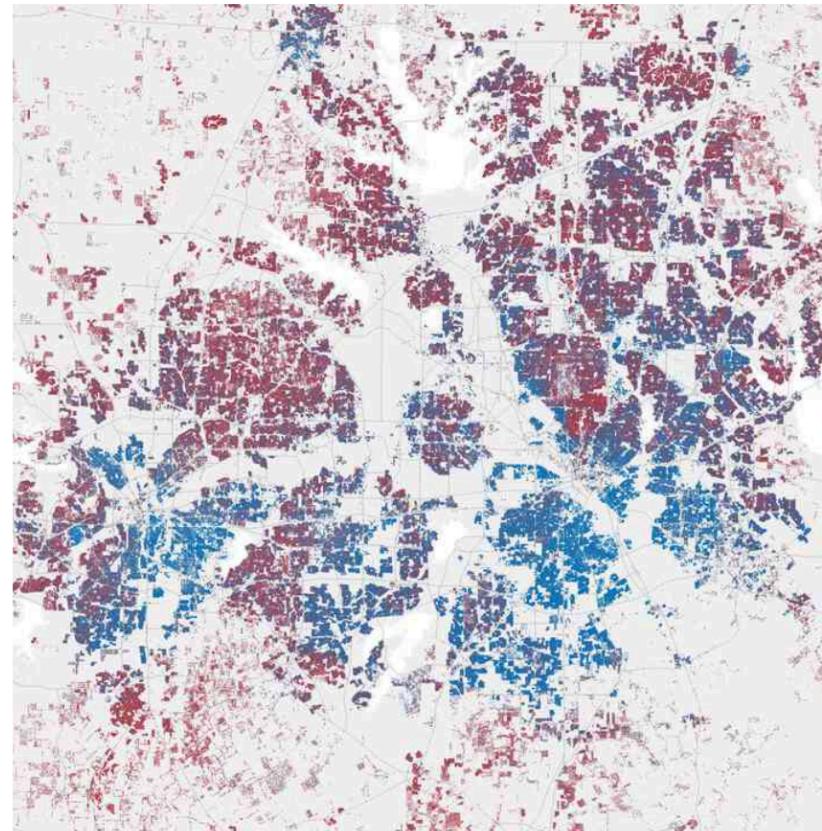
- Виды визуализаций, как их использовать и строить
- Дизайн и типографика
- Интерактивная визуализация (программирование на JavaScript)
- Картография
- Работа с видео, 3D-моделирование, иллюстрация...

CHICAGO



5 miles

DALLAS-FORT WORTH

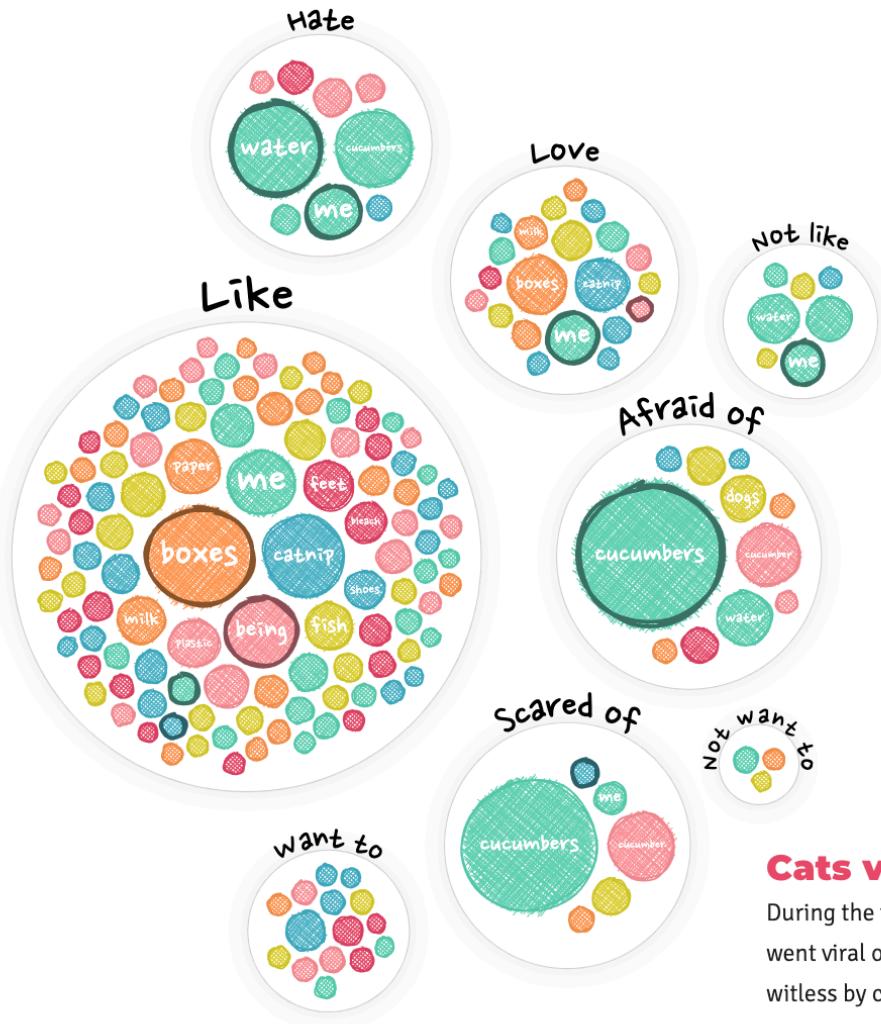


5 mi.

[New York Times](#)



Shirley Wu



Nadieh Bremer

### Cats vs. Cucumbers

During the final months of 2015 a meme went viral of videos showing cats scared witless by cucumbers ↗

### **3. Чему ещё можно учиться?**

# Чему ещё учиться?

- Изучить основы статистики
- Углубиться в расследования
- Подтянуть насыщенность и знание источников данных

**Где учиться, и что делать?**

# Где учиться, и что делать?

- Отдельные курсы по нужным направлениям (статистика, визуализация на JavaScript, курсы по Python, машинное обучение....)
- Магистратура ВШЭ по журналистике данных
- Мастерская «Важных историй»
- Участуйте в хакатонах и конференциях (Open Data Day, LAMPA)
- Вступите в профессиональные сообщества
- Приходите постажироваться в издания (Новая газета, Медиазона, Важные истории, РБК, Тинькофф-журнал...)
- Делайте собственные исследования и предлагайте их редакциям

# Полезные ссылки

Инфокультура — следит за всем, связанным с открытыми данными

Теплица социальных технологий — хакатоны, семинары и инструменты для журналистов

Блог Ивана Бегтина — много интересных заметок про данные  
data.csv — канал про журналистику данных: работы, инструменты, возможности

Мастерская «Важных историй» — уроки по Python и инструменты для журналистов

Чаты: Открытые данные в России, Датавиз-чат, чат дата-журналистов