



Кейс «Имя ребенка 1880-2018»

проф. кафедры Эконометрики и математической экономики ЭФ
д.т.н. Шилин Кирилл Юрьевич

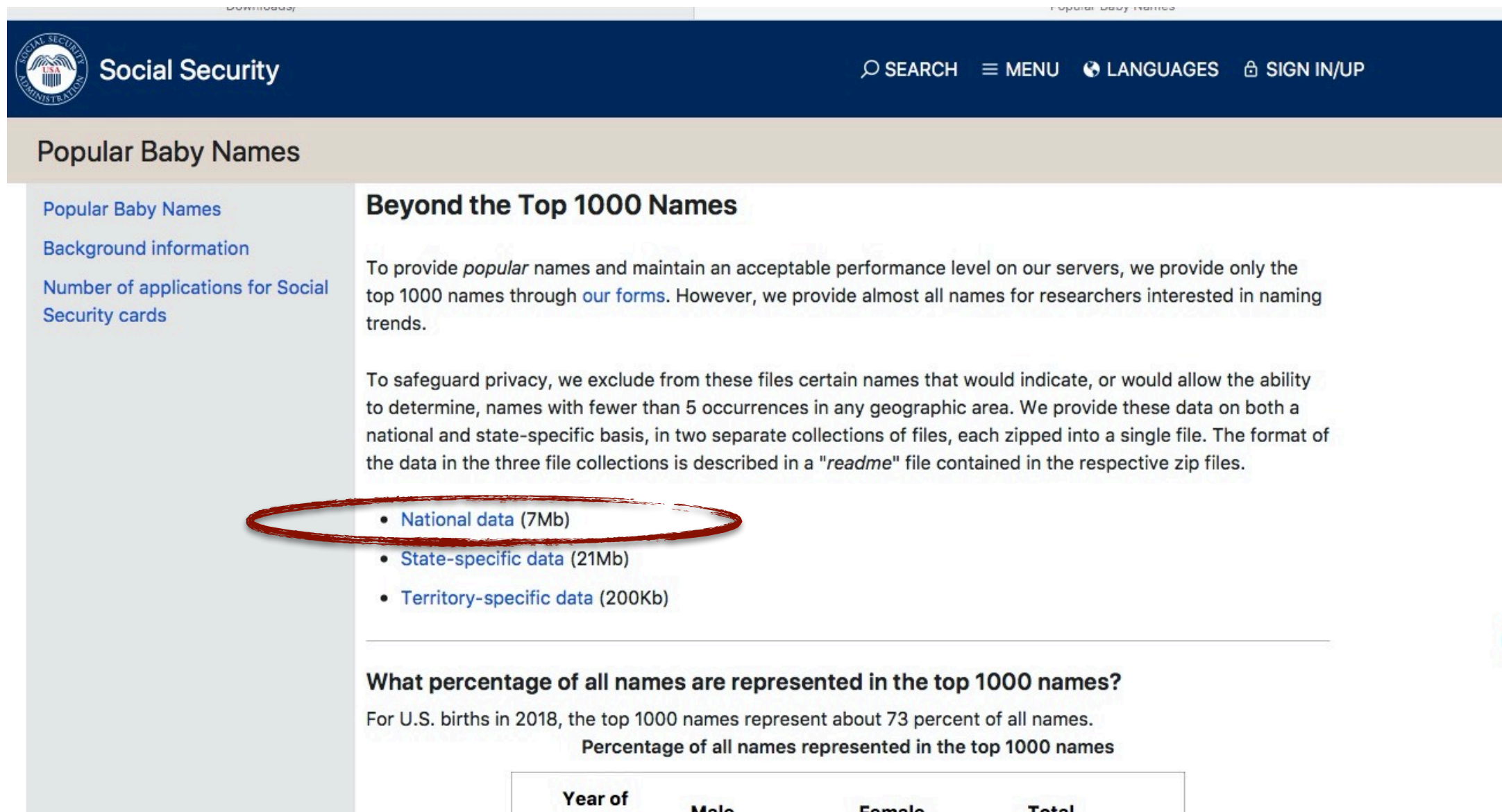
РАНХиГС каб. 403/4

email для вопросов: kshilin@ranepa.ru

email ДЗ: kirill.shilin@gmail.com

Качаем данные по именам

Ссылка <https://www.ssa.gov/oact/babynames/limits.html>



Social Security SEARCH MENU LANGUAGES SIGN IN/UP

Popular Baby Names

[Popular Baby Names](#)
[Background information](#)
[Number of applications for Social Security cards](#)

Beyond the Top 1000 Names

To provide *popular* names and maintain an acceptable performance level on our servers, we provide only the top 1000 names through [our forms](#). However, we provide almost all names for researchers interested in naming trends.

To safeguard privacy, we exclude from these files certain names that would indicate, or would allow the ability to determine, names with fewer than 5 occurrences in any geographic area. We provide these data on both a national and state-specific basis, in two separate collections of files, each zipped into a single file. The format of the data in the three file collections is described in a "*readme*" file contained in the respective zip files.

- [National data \(7Mb\)](#)
- [State-specific data \(21Mb\)](#)
- [Territory-specific data \(200Kb\)](#)

What percentage of all names are represented in the top 1000 names?

For U.S. births in 2018, the top 1000 names represent about 73 percent of all names.

Percentage of all names represented in the top 1000 names

| Year of | Male | Female | Total |
|---------|------|--------|-------|
|---------|------|--------|-------|

Работаем с именами 1

Часть 1.

1. Читаем файл `yob1880.txt` даем имена колонкам.
2. Сколько имен мальчиков, а сколько девочек?
3. Сколько родилось мальчиков, а сколько девочек?

Часть 2.

1. Собираем единый `DataFrame`
2. Сводная таблица по годам для количества рожденных мальчиков и девочек
3. Рисуем график рождаемости

Часть 3.

1. Найдем для каждого имени долю детей с данным именем относительно общего количества родившихся детей в зависимости от года и пола.
2. Отберем 1000 имен для каждой комбинации пол / год.

Работаем с именами 2

Часть 4.

Строим тренды для 2-х самых популярных мужских и женских имен.

Часть 5.

Строим тренд - разнообразия имен

Часть 6.

Строим тренд - количество популярных имен в топ 50%

Работаем с именами 3

Часть 7.

Революция «последней» буквы

Часть 8.

Смена половой принадлежности имени с корнем lesl