

Краткосрочное прогнозирование демографических тенденций на основе данных Google trends

И. Е. Калабихина¹, Е. П. Банин^{2}, В. Н. Архангельский¹, И. А. Абдуселимова¹,
Г. А. Клименко¹, А. В. Колотуша¹, У. Г. Николаева¹, В. Ш. Шамсутдинова¹*

¹ МГУ имени М. В. Ломоносова, экономический факультет, Москва, Россия

² МГТУ им. Н. Э. Баумана, Москва, Россия

* *evg.banin@gmail.com*

Аннотация. Демографические показатели являются важными параметрами государственных программ по развитию России, а оперативный мониторинг демографического развития – залог успешной реализации программ. Очень часто данные государственной статистики публикуются с задержкой, что не позволяет использовать их при оперативном мониторинге и планировании. В настоящей работе рассматривается подход, позволяющий оперативно оценивать демографические процессы в области формирования семьи и прогнозировать демографические тенденции в краткосрочном будущем на основе данных статистики запросов из Google Trends. Взаимосвязи между данными поисковых запросов и демографическими показателями анализируются с помощью корреляции Пирсона. Для анализа используются годовые (коэффициенты рождаемости, количество аборт на 100 рождений, количество аборт на 1000 женщин, количество браков и разводов на 1000 населения) и месячные данные (число родившихся, количество браков и разводов) по рожденьям, бракам и абортам с лагом и без. Анализ проводится на данных по России в целом и для восьми самых населенных регионов: Москва, Московская область, Краснодарский край, Санкт-Петербург, Ростовская область, Свердловская область, Республика Татарстан, Республика Башкортостан. Используя временные ряды, доступные в Google Trends с 2004 года, авторы демонстрируют, что некоторые демографические показатели могут быть спрогнозированы на основе данных родственных запросов к поисковому алгоритму Google с использованием модели ARIMA. Таким образом, возможно использовать данные поисковых запросов в качестве хорошего дополнения к данным демографических показателей, при построении моделей множественной регрессии для демографических расчетов или использовать в качестве прокси-переменных.

Ключевые слова: демографические прогнозы, рождаемость, брачность, временные ряды, большие данные, Google trends, ARIMA-модель

Для цитирования: *Калабихина И. Е., Банин Е. П., Архангельский В. Н., Абдуселимова И. А., Клименко Г. А., Колотуша А. В., Николаева У. Г., Шамсутдинова В. Ш.* Краткосрочное прогнозирование демографических тенденций на основе данных Google Trends // Прикладная информатика. 2020. Т. 15. № 6. С. 91–118. DOI: 10.37791/2687-0649-2020-15-6-91-118