

УИК России: координаты, численность избирателей и оценка численности населения в крупных городах

Лицензия: CC BY-SA

Доступ: Открытый

Дата обновления: 27.04.2022

Теги:

Описание

Набор данных содержит сведения об участковых избирательных комиссиях России (УИК). В частности, в данных присутствует информация о географическом расположении комиссий (географические координаты) и численности избирателей (на момент проведения Общероссийского голосования о поправках в Конституцию в июле 2020 года). Кроме того, для населенных пунктов с численностью населения более 10 000 тысяч человек на основе данных об общей численности населения в населенном пункте была рассчитана примерная численность населения, проживающего по адресам, приписанным к каждому УИК, располагающемуся в крупном населенном пункте. Такие данные могут быть полезны исследователям, занимающимся пространственным анализом, т.к. дают возможность работать с более равномерным распределением населения.

Единицей наблюдения является УИК. К важным атрибутам относятся географические координаты (lat, lon), информация о количестве избирателей (voters_include) и оценка численности населения (population).

Всего датасет содержит 93 786 наблюдений по 14 атрибутам. Данные относятся к 2020 году. Набор доступен для работы в формате CSV (кодировка: «UTF-8», разделитель: «;»).

Характеристики датасета

Использование

Лицензия: [CC BY-SA](#)

Доступ: Открытый

Предоставление данных

Формат:

CSV

Характеристики набора

Временной период: 2020г.

Количество наблюдений: 93786

Обновления

Размещено: 29.04.2022

Обновлено: 27.04.2022

Частота обновления: Иное

Дата следующего обновления: Не обновляется

Происхождение

Источники данных: Центральная избирательная комиссия РФ

Ответственные за набор данных: Казанцев Кирилл Ильич

Структура набора данных

Координаты, численность избирателей и оценка численности населения в крупных городах (tik_uik_joined)

Атрибут	Описание	Количество пропусков (NaN) (шт/%)	Единица измерения	ОКЕИ	Формат
uik_population	Оценка общей численности населения, проживающего рядом с УИК	50686			integer
uik_weight	Доля избирателей, приписанных к УИК, в общем количестве избирателей в городе	50686			numeric
voters_include	Количество избирателей на голосовании по поправкам в Конституцию (июль 2020 года)	3360			integer
lon	Долгота	911			numeric
lat	Широта	911			numeric
geocoding_precision	Точность геокодирования	0			cat
geocoding_kind	Тип объекта, на который попал геокодер	911			cat
uik_geocoding_address	Адрес расположения комнаты для голосования УИК, полученный в ходе геокодирования	911			text
uik_voteroom_address	Адрес расположения комнаты для голосования УИК	0			text
uik_address	Адрес расположения УИК	0			text
uik_url	Ссылка УИК на сайте ЦИК РФ	0			text
uik	Наименование участковой избирательной комиссии	0			text
tik	Наименование территориальной избирательной комиссии	0			cat
region	Наименование региона	0			cat

Внесенные в набор данных изменения по сравнению с источниками

- Данные с разрозненных страниц на сайте ЦИК РФ и его региональных отделений собраны в единую таблицу
- Данные дополнены информацией о географических координатах УИК. Геокодирование производилось с помощью сервиса геолокации Яндекс API.
- На основе численности избирателей в крупных населенных пунктах (с населением более 10 000 человек) оценена примерная численность всего населения, проживающего по адресам, приписанным к УИК. Для этого, для каждого населенного пункта с численностью населения более 10 000 человек был определен перечень УИКов, относящихся к этому населенному пункту. Численность избирателей на этих УИК складывалась. После этого рассчитывался относительный "вес" УИК - доля избирателей населенного пункта, приписанная к конкретному УИК (`uik_weight`). На основе этого веса и информации о численности всего населения (т.е. включая граждан РФ до 18 лет) была рассчитана численность населения, проживающего рядом с УИК (`uik_population = uik_weight * Население города`).

Известные ограничения данных

В данных присутствуют пропуски по некоторым переменным.

Пропуски в информации о географических координатах УИК (`lat` и `lon`) присутствуют в связи с тем, что сервис геокодирования не смог определить координаты, исходя из адреса указанного на сайте УИК. Кроме того, для некоторых УИК указанные координаты показывают неточное расположение здания, где располагается УИК, а только улицу на которое находится здание или населенный пункт. Это связано с отсутствием информации о всех помещениях в базе сервисов геолокации и чаще всего встречается в малых или удаленных населенных пунктах.

Пропуски по численности избирателей на участках связаны с тем, что информация о расположении УИК собиралась в октябре-ноябре 2020 года, а данные по численности избирателей соответствуют голосованию по поправкам в Конституцию в июле 2020 года. За этот период некоторые УИК могли быть распущены или могли изменить свой номер, что привело к невозможности соединить данные между собой. Такие пропуски наблюдаются в 3.5% случаев. Пропуски в информации об оценке численности населения, проживающего рядом с УИК, связаны с тем, что подобная оценка проводилась только для населенных пунктов с населением более 10 000 тысяч человек