Исходная база данных представляет собой более ста тысяч строк с различной информацией о государственных закупках в период с 2016 по 2020 годы. Для анализа и интерпретации будут использоваться библиотеки Pandas и Geopandas для работы с датафреймами, а также библиотеки Matplotlib и Seaborn для построения графических зависимостей. Перед началом анализа необходимо отформатировать данные, для чего исключим строки, в которых какие-либо параметры не определены, либо определены некорректно, например, строки с увеличением итоговой цены более чем на сто или менее чем на ноль процентов.

Представим распределение значений итогового снижения цены в виде диаграммы размаха, показанного на рисунке 1. Численные параметры, по которым строилась диаграмма, представлены в таблице 1. Вывод

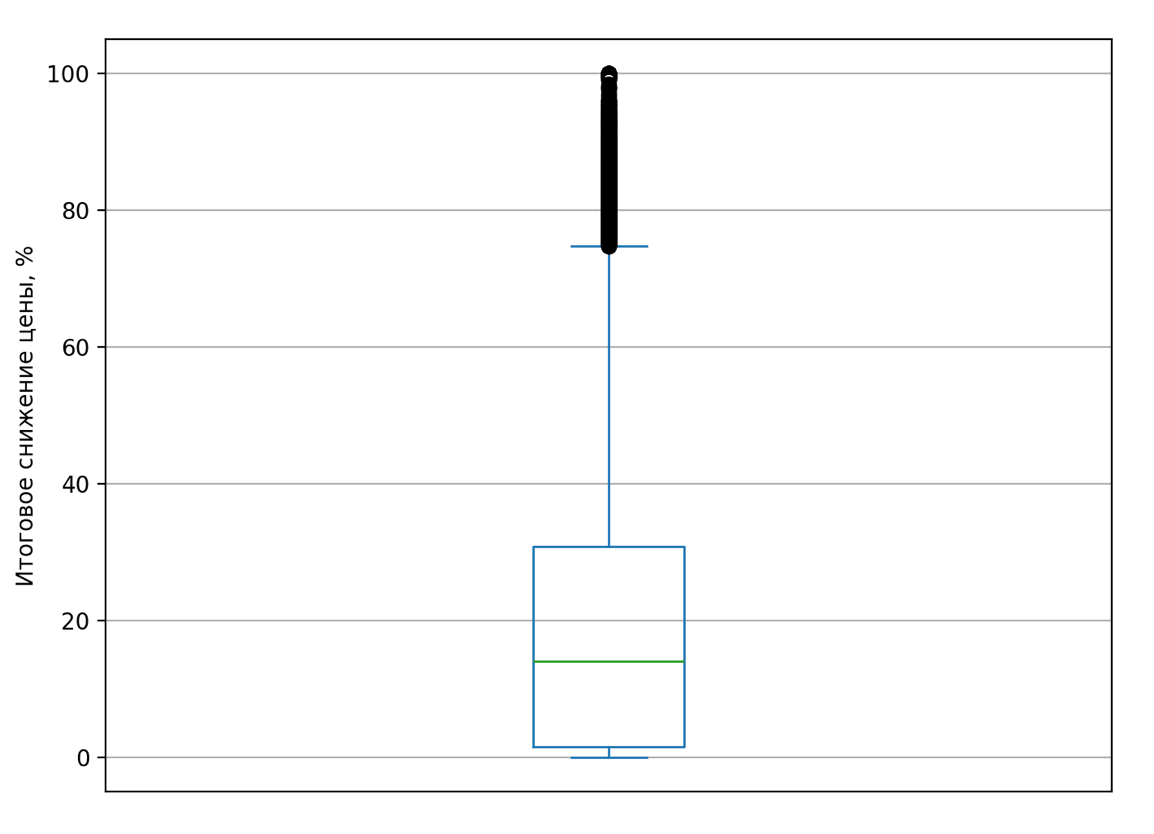


Рисунок 1. Распределение итогового снижения цены

Таблица 1. Значения параметров для построения диаграммы размаха итогового снижения цены

|  |  |
| --- | --- |
| Параметр | Значение |
| Размер выборки | 98153 |
| Минимум | 0,0 |
| Q1 | 1,5 |
| Медиана | 14,0 |
| Q3 | 30,8 |
| Максимум | 100,0 |
| Среднее | 19,7 |

Далее был построен линейный график зависимости среднего снижения цены от количества участников (рисунок 2). Средние значения были получены путем деления суммы всех значений снижения цены для определенного количества участников на количество таких значений. Из-за низкой частоты аукционов с большим количеством участников (рисунок 3) было принято решение обрезать линейный график среднего снижения цены по горизонтальной оси до значения в 14 участников. Вывод

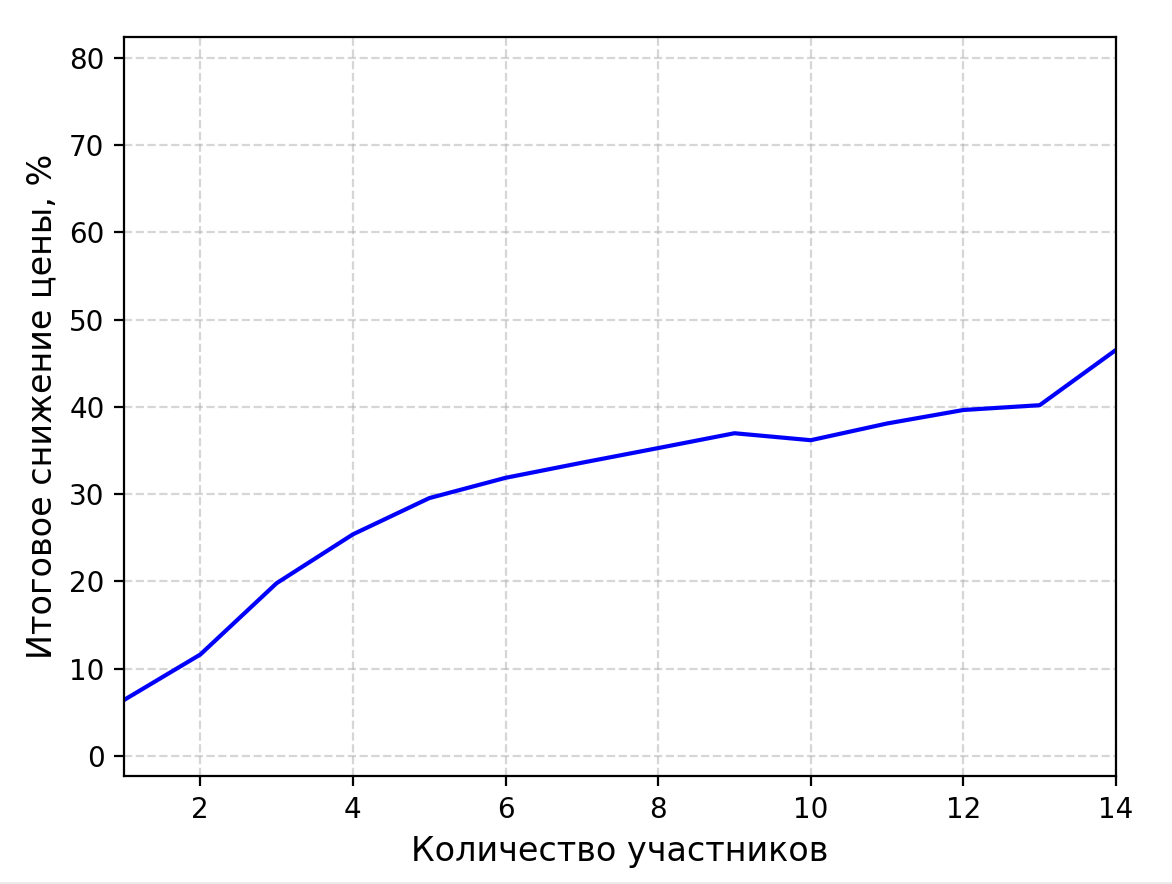


Рисунок 2. Линейный график зависимости среднего итогового снижения цены от количества участников

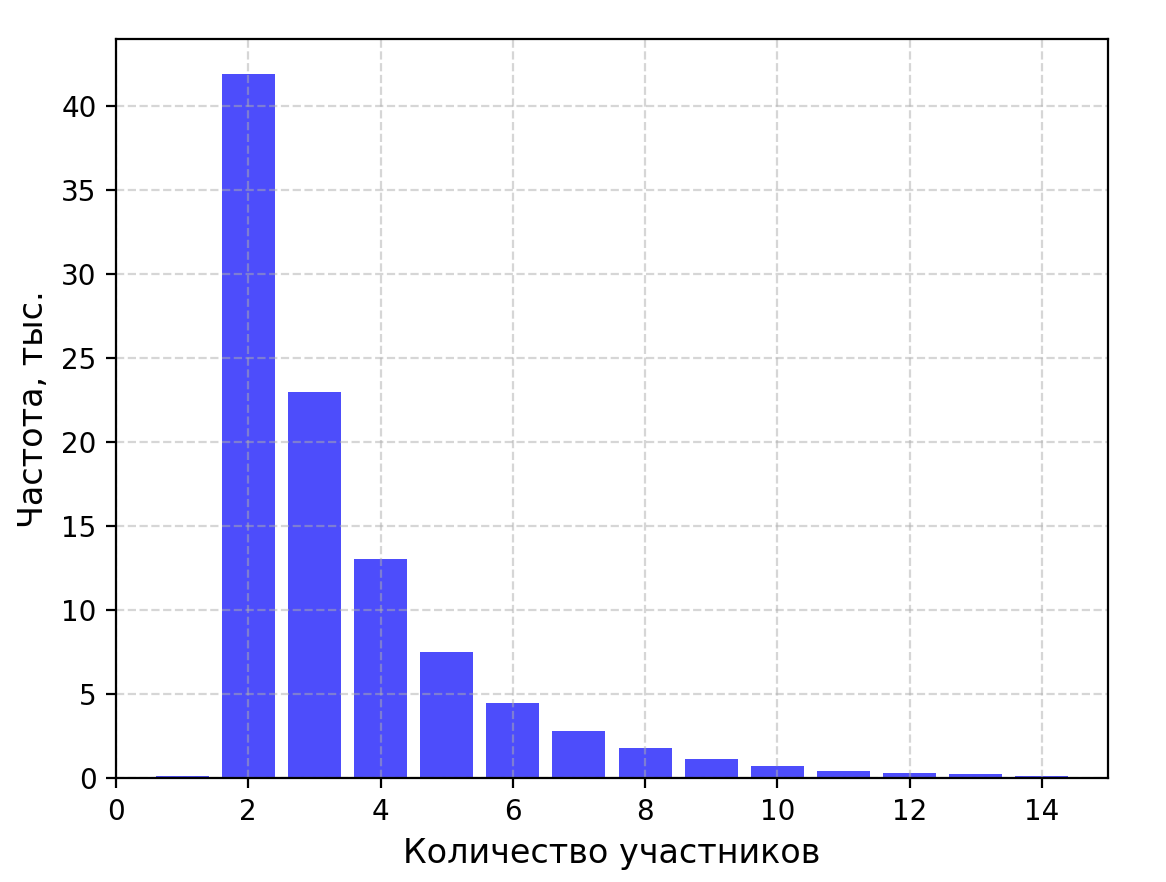


Рисунок 3. Гистограмма распределения количества участников аукционов

В отдельном окне были построены 2 графика зависимости снижения цены от количества участников: с фиксированной суммой шага и шагом-процентом от начальной максимальной цены (НМЦ) (рисунок 4). Вывод

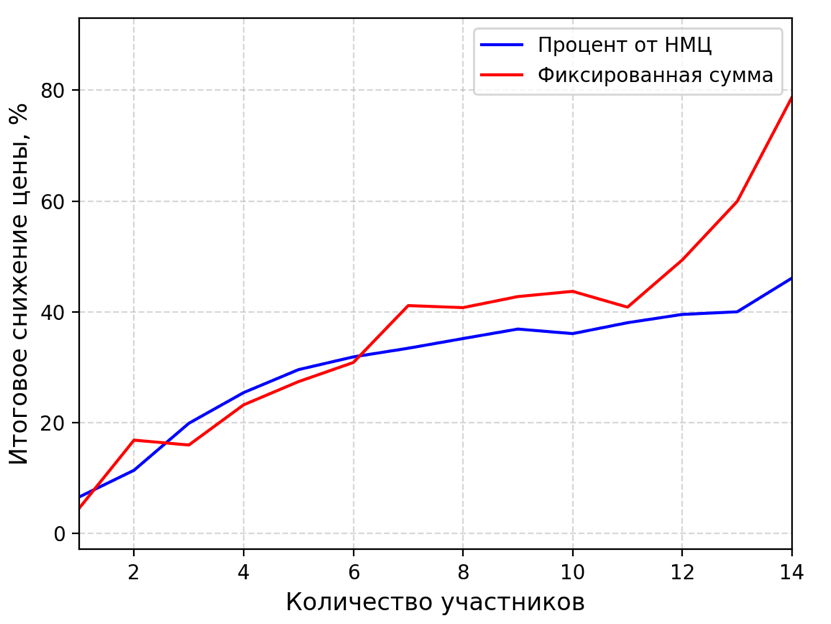


Рисунок 4. Линейный график зависимости среднего итогового снижения цены от количества участников и типа шага аукциона

После были построены диаграммы размаха, иллюстрирующие распределение снижения цены в зависимости от наличия или отсутствия второй фазы (рисунок 5). Численные значения, по которым были построены диаграммы, представлены в таблице 2. Вывод

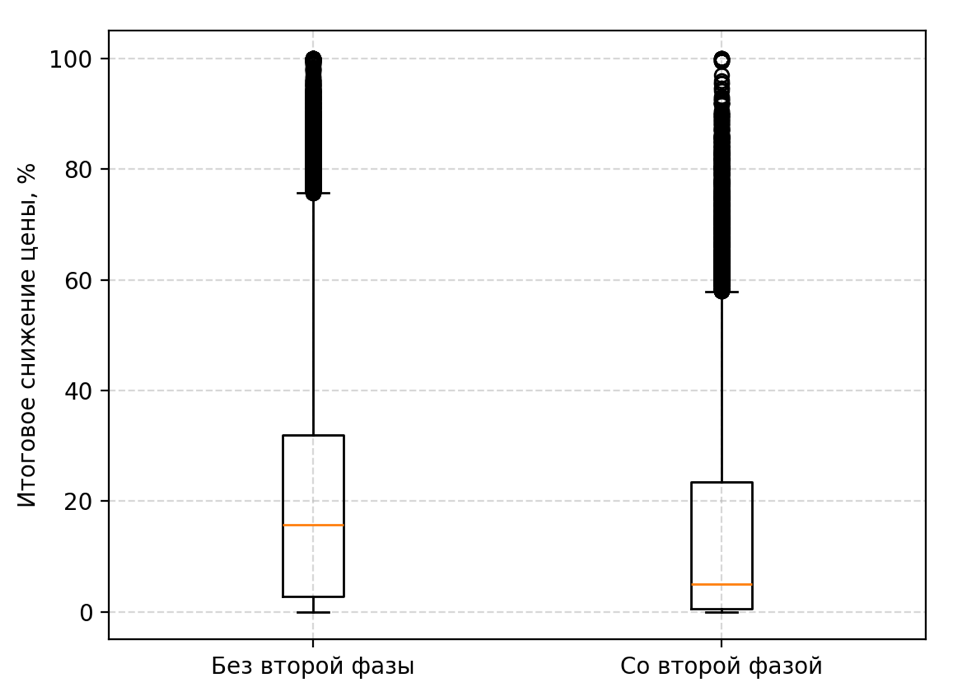


Рисунок 5. Распределение итогового снижения цены в зависимости от наличия или отсутствия второй фазы аукциона

Таблица 2. Значения параметров для построения диаграммы размаха итогового снижения цены в зависимости от наличия или отсутствия второй фазы аукциона

|  |  |  |
| --- | --- | --- |
| Параметр | Значение | |
| Без второй фазы | Со второй фазой |
| Размер выборки | 79968 | 18185 |
| Минимум | 0,0 | 0,0 |
| Q1 | 2,8 | 0,5 |
| Медиана | 15,8 | 5,0 |
| Q3 | 32,0 | 23,5 |
| Максимум | 100,0 | 100,0 |
| Среднее | 20,6 | 15,7 |

По рассмотренным численным параметрам, которые могут влиять на снижение цены, была построена корреляционная матрица (рисунок 6). Вывод

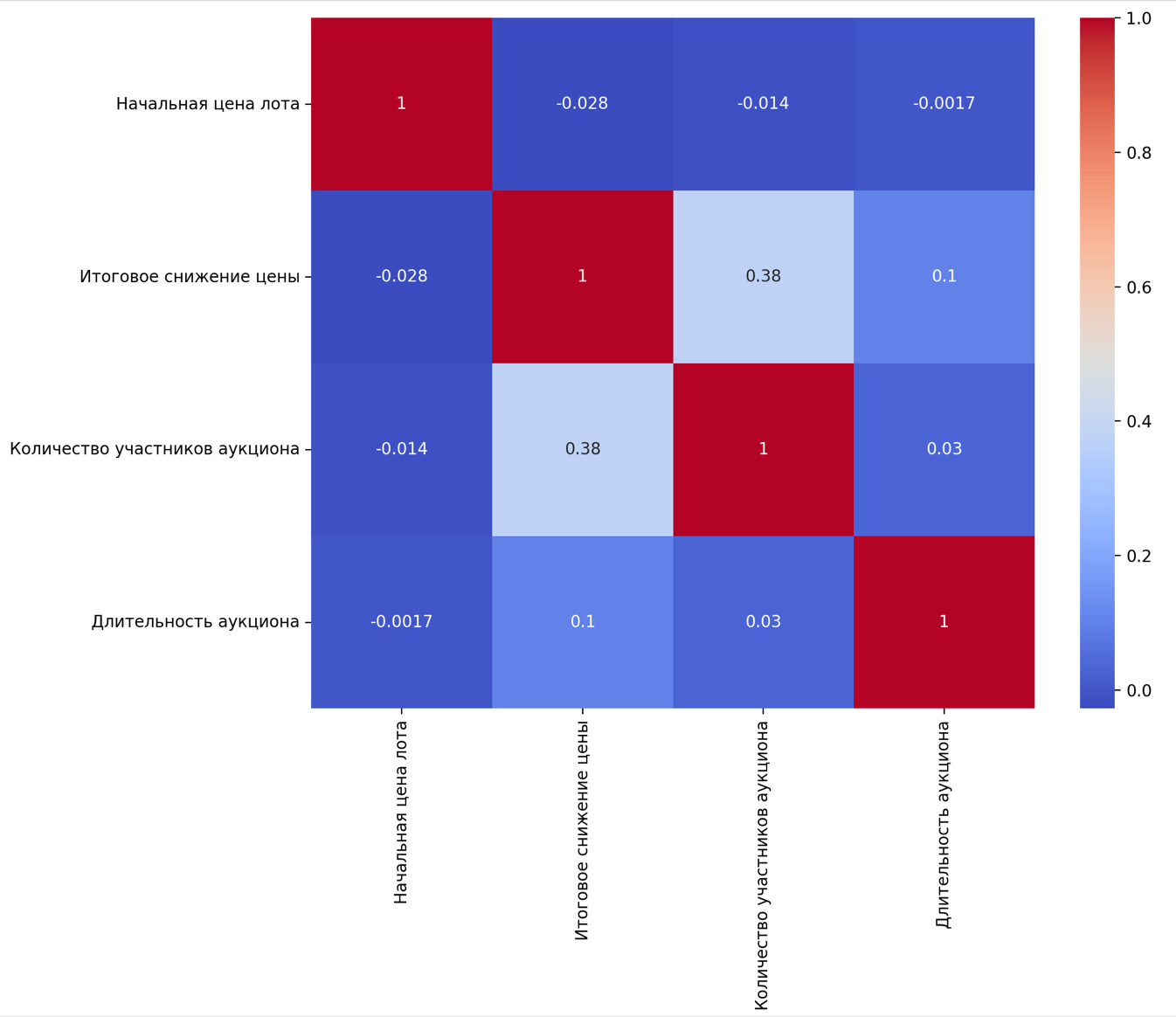


Рисунок 6. Корреляционная матрица численных параметров аукционов

В исходной базе данных были найдены юридические адреса организаторов аукционов, что дает возможность построения карты регионов России с цветовой кодировкой среднего снижения цены на торгах. Из открытых источников была найдена карта субъектов РФ в формате GeoJSON, после чего данные по аукционам были соотнесены с датафреймом, по которому строится карта. Так как в некоторых регионах среднее снижение цены незначительно отличается, было принято решение разбить цветовую шкалу на 5 интервалов по 20%. Полученная карта показана на рисунке 7.

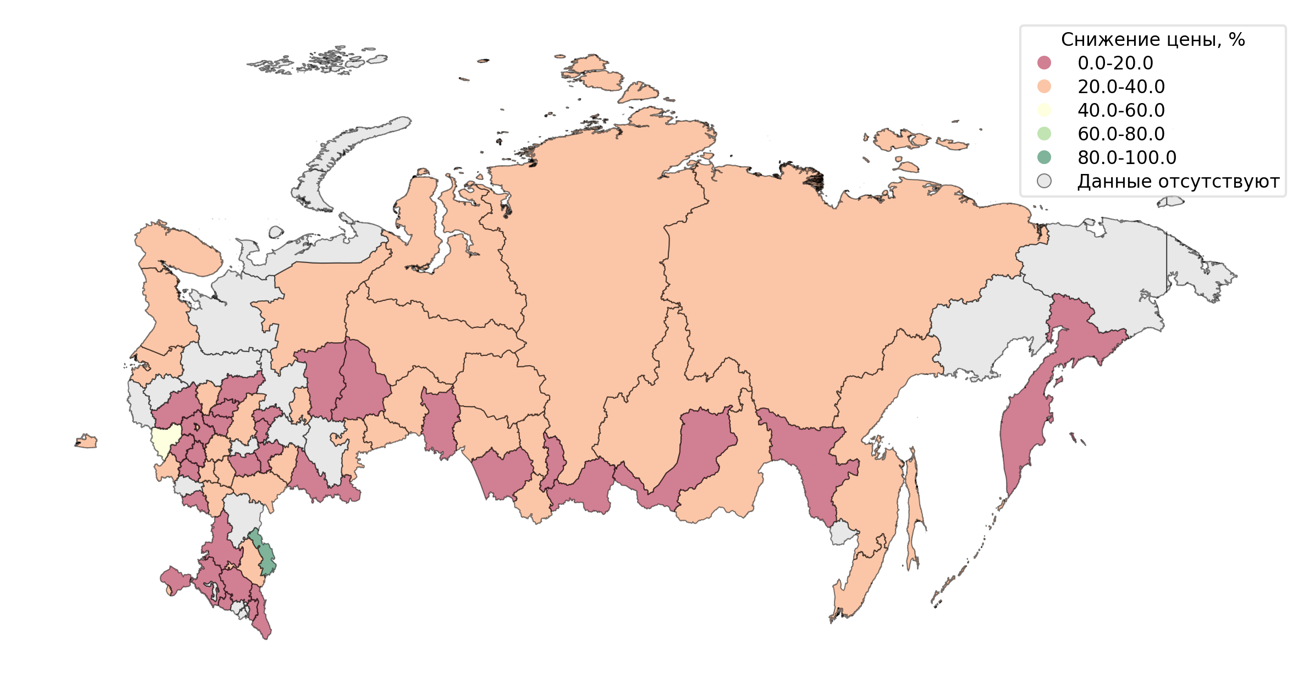


Рисунок 7. Карта среднего снижения цены по регионам

При интерпретации важно учитывать не только показатель снижения цены регионе, но и размер выборки для каждого региона, так как небольшие выборки не могут соответствовать достоверным значениям. Для этого была также построена карта РФ с размером выборки для каждого региона (рисунок 8).

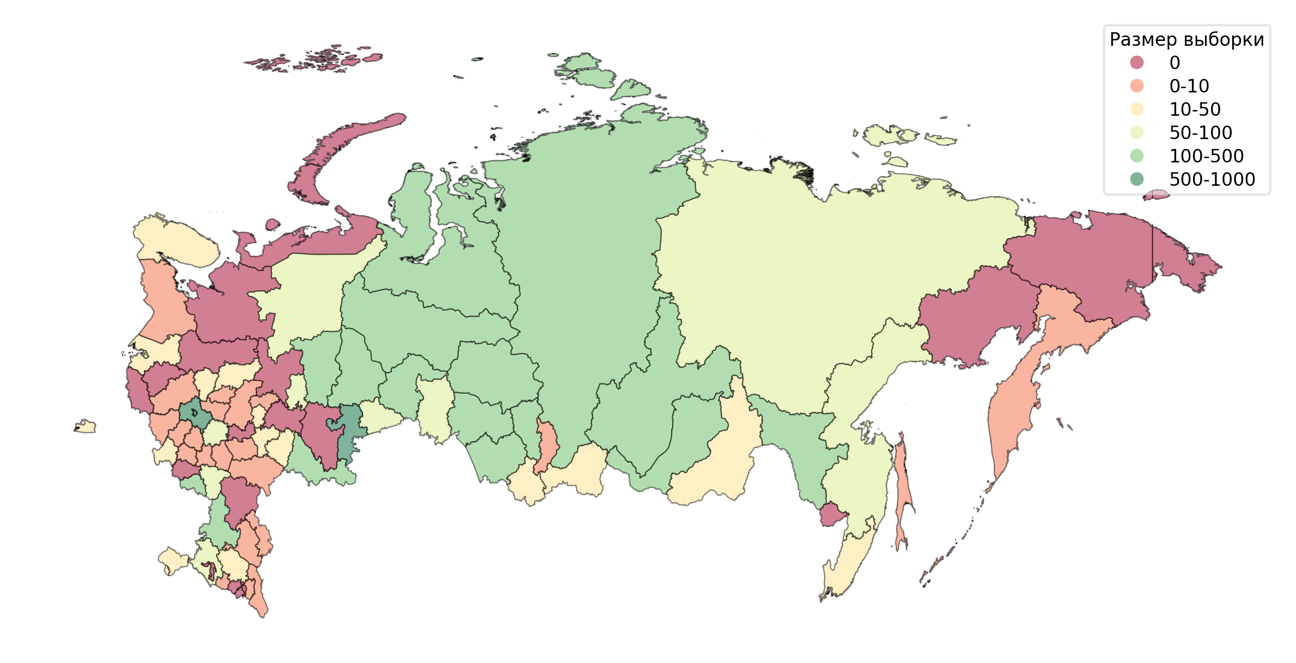


Рисунок 8. Карта размера выборки по регионам