1. **Рассчитайте основные статистики (меры центра и меры разброса) по распределениям всех переменных, имеющихся в файле данных.**

|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| AddressCount | | |
| N | Валидные | 79 |
| Пропущенные | 0 |
| Среднее | | 1048,04 |
| Медиана | | 371,00 |
| Стандартная отклонения | | 1642,066 |
| Диапазон | | 9543 |
| Минимум | | 9 |
| Максимум | | 9552 |

|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| CallsCount | | |
| N | Валидные | 79 |
| Пропущенные | 0 |
| Среднее | | 3648,68 |
| Медиана | | 931,00 |
| Стандартная отклонения | | 8124,105 |
| Диапазон | | 48477 |
| Минимум | | 20 |
| Максимум | | 48497 |

|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| ClicksCount | | |
| N | Валидные | 79 |
| Пропущенные | 0 |
| Среднее | | 21826,01 |
| Медиана | | 6921,00 |
| Стандартная отклонения | | 32474,960 |
| Диапазон | | 166897 |
| Минимум | | 258 |
| Максимум | | 167155 |

|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| FirmsCount | | |
| N | Валидные | 79 |
| Пропущенные | 0 |
| Среднее | | 305,09 |
| Медиана | | 185,00 |
| Стандартная отклонения | | 382,052 |
| Диапазон | | 2365 |
| Минимум | | 14 |
| Максимум | | 2379 |

|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| GeoPart | | |
| N | Валидные | 79 |
| Пропущенные | 0 |
| Среднее | | ,342644600423168 |
| Медиана | | ,322341513292434 |
| Стандартная отклонения | | ,103607604865767 |
| Диапазон | | ,4632587854302933 |
| Минимум | | ,0929166666666667 |
| Максимум | | ,5561754520969600 |

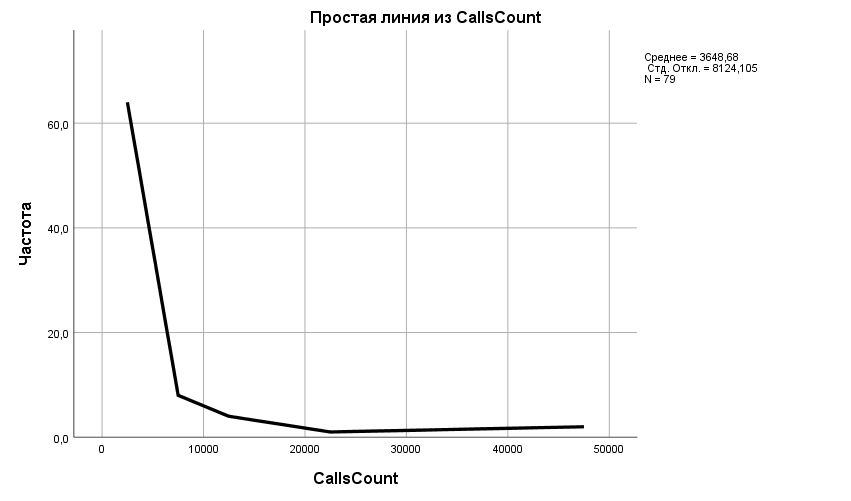
|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| MobilePart | | |
| N | Валидные | 79 |
| Пропущенные | 0 |
| Среднее | | ,445746310605206 |
| Медиана | | ,463743676222597 |
| Стандартная отклонения | | ,146130624734777 |
| Диапазон | | ,647288321970530 |
| Минимум | | ,090000000000000 |
| Максимум | | ,737288321970530 |

|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| UsersCount | | |
| N | Валидные | 79 |
| Пропущенные | 0 |
| Среднее | | 9753,13 |
| Медиана | | 2934,00 |
| Стандартная отклонения | | 13927,296 |
| Диапазон | | 60970 |
| Минимум | | 157 |
| Максимум | | 61127 |

|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| Distance | | |
| N | Валидные | 79 |
| Пропущенные | 0 |
| Среднее | | 2669,42635208869340 |
| Медиана | | 2586,50327354116000 |
| Стандартная отклонения | | 1427,785303810908800 |
| Диапазон | | 5577,420075181411 |
| Минимум | | 714,787236203679 |
| Максимум | | 6292,207311385090 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **IsGeo** | | | | | |
|  | | Частота | Проценты | Валидный процент | Накопленный процент |
| Валидные | 0 | 51 | 64,6 | 64,6 | 64,6 |
| 1 | 28 | 35,4 | 35,4 | 100,0 |
| Всего | 79 | 100,0 | 100,0 |  |

**2. Выберите наиболее интересный для вас количественный признак и охарактеризуйте его распределение при помощи соответствующих описательных статистик и графиков:**



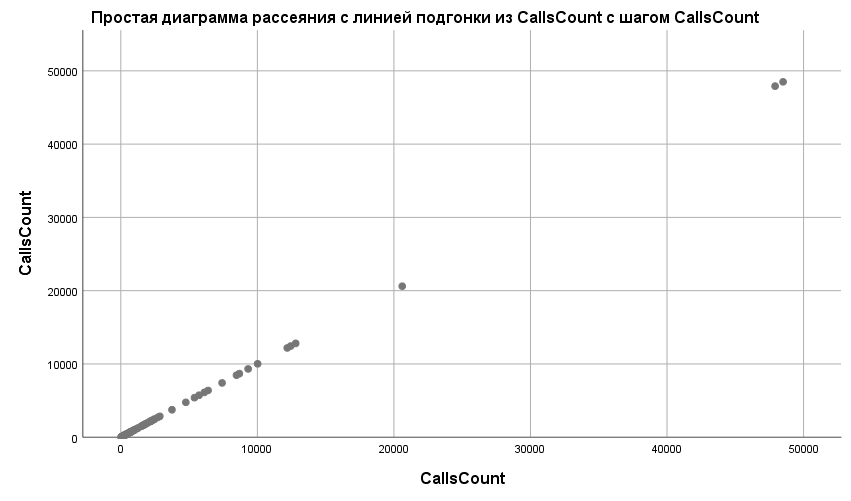
**2.1. Какова форма распределения признака?**

Распределение Фишера

**2.2. Можно ли говорить о том, что распределение признака согласуется с каким-либо теоретическим законом распределения?**

**2.3. Есть ли в данных «аномалии» / «выбросы»?**

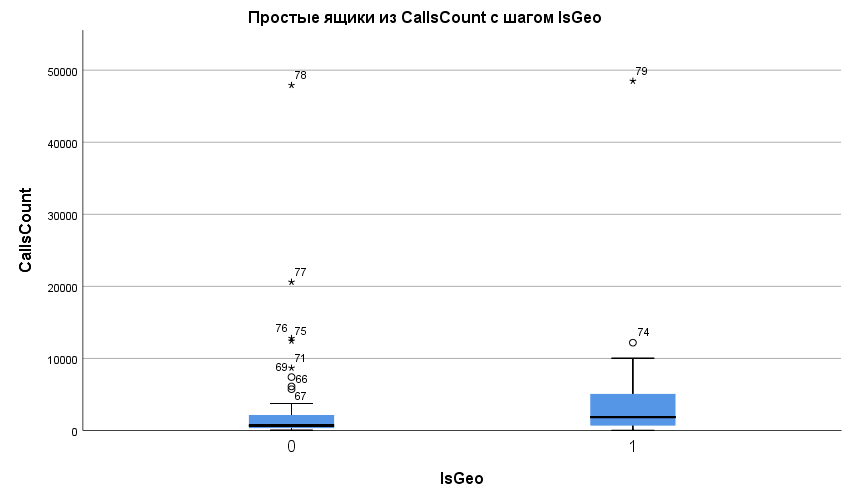
Да



**2.4. Какие меры центра и вариативности подходят для описания распределений лучше всего? Почему?**

Все меры которые не включают среднего при расчете: медиана, мода, межквартильный размах, несмещенная дисперсия. Т.к. они не подвержены влиянию выбросов.

**Сравните геозависимые и геонезависимые сферы (экспертная разметка, переменная IsGeo) по выбранному для анализа признаку. Есть ли отличия? В чем они состоят? (для ответа на вопрос используйте статистические и графические инструменты).**



Для isGEO = 1 Для isGeo = 0

|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| CallsCount | | |
| N | Валидные | 28 |
| Пропущенные | 0 |
| Среднее | | 4585,79 |
| Медиана | | 1850,50 |
| Стандартная отклонения | | 9242,558 |
| Диапазон | | 48477 |
| Минимум | | 20 |
| Максимум | | 48497 |

|  |  |  |
| --- | --- | --- |
| **Статистика** | | |
| CallsCount | | |
| N | Валидные | 51 |
| Пропущенные | 0 |
| Среднее | | 3134,20 |
| Медиана | | 698,00 |
| Стандартная отклонения | | 7488,024 |
| Диапазон | | 47875 |
| Минимум | | 37 |
| Максимум | | 47912 |

Гораздо больше данных для isGeo = 0.

Выбросов для isGeo = 0 в разы больше.

Медиана для isGeo = 1 в разы больше.

Среднее для isGeo = 1 больше в полтора раза.