

Выброс (Outlier)

Синонимы: Экстремальное значение, Extreme value

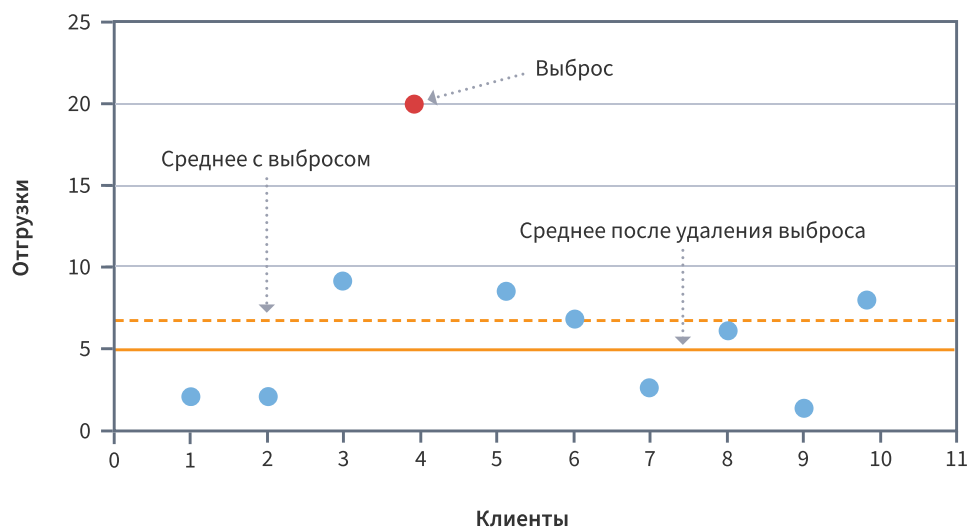
Выбросами называются значения, которые не укладываются в общую модель поведения анализируемого процесса. Они сильно отличаются от окружающих данных и могут быть вызваны как ошибками измерений, так и некорректным вводом данных, или являться результатом их сильной изменчивости.

Выбросы являются следствием:

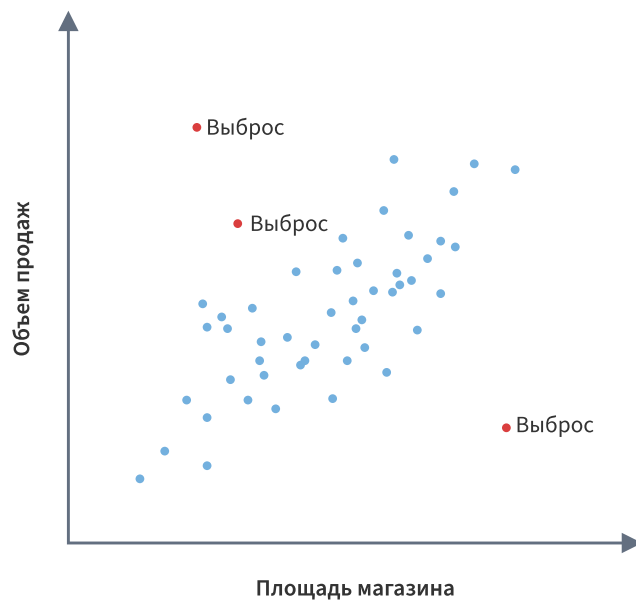
- ошибок в данных (погрешности измерений, неверная запись или считывание данных и т.п.);
- воздействия случайных, не поддающихся прогнозированию факторов (например, разовый наплыв клиентов из-за массового мероприятия).

Выбросы могут быть как одномерными, так и многомерными.

Пример одномерных выбросов:



Пример двумерных выбросов:



Перед выполнением аналитической обработки выбросы необходимо подавить или удалить, поскольку они могут вызвать некорректную работу алгоритмов и привести к искажению результатов анализа данных.

Степень устойчивости алгоритма к наличию в данных экстремальных значений называется робастностью.

Специализированный компонент редактирование выбросов предназначен для автоматической корректировки выбросов и экстремальных значений набора данных в Logiном.