



ПРЕЗИДЕНТСКАЯ
АКАДЕМИЯ

2024

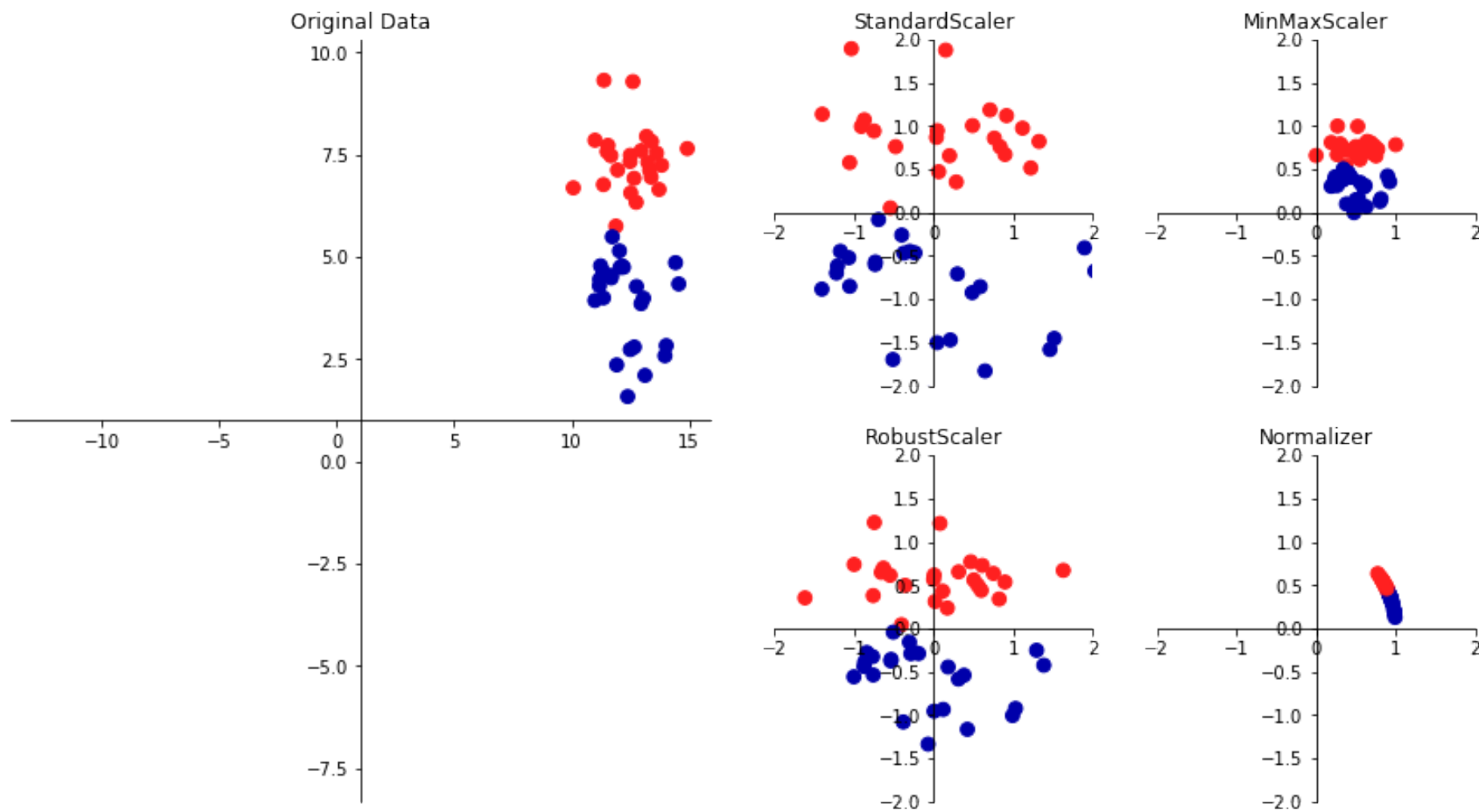
ВВЕДЕНИЕ В МАШИННОЕ ОБУЧЕНИЕ

О КРОСС-ВАЛИДАЦИИ, КОНТЕЙНЕРАХ И ПЕРЕМАСШТАБИРОВАНИИ
ДАННЫХ

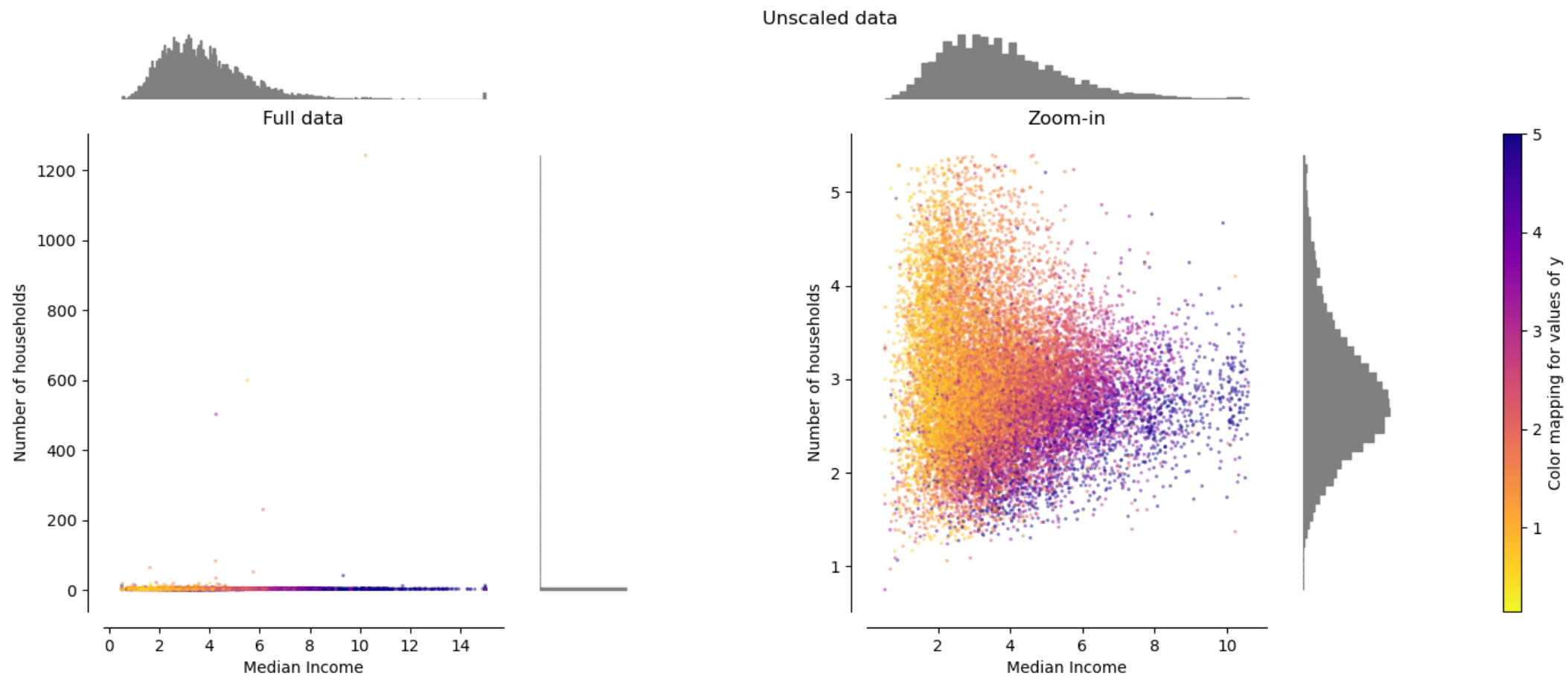
Москва, 2024

Материалы занятия <https://github.com/kshilin/machine-learning>

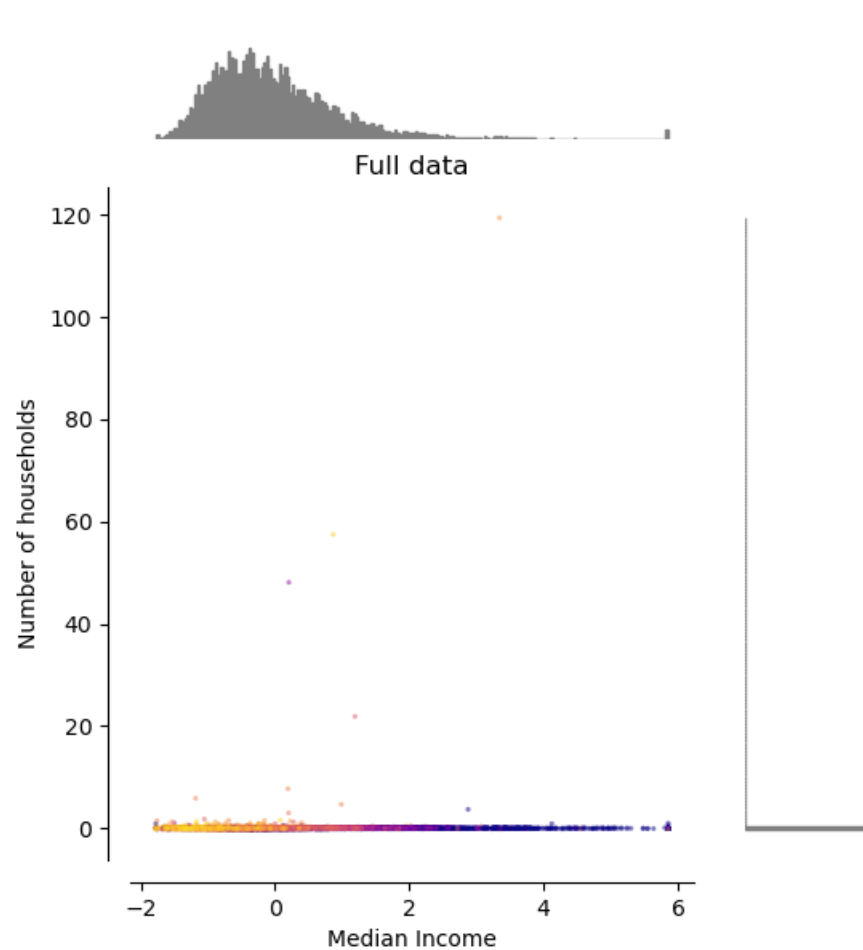
РАНХиГС



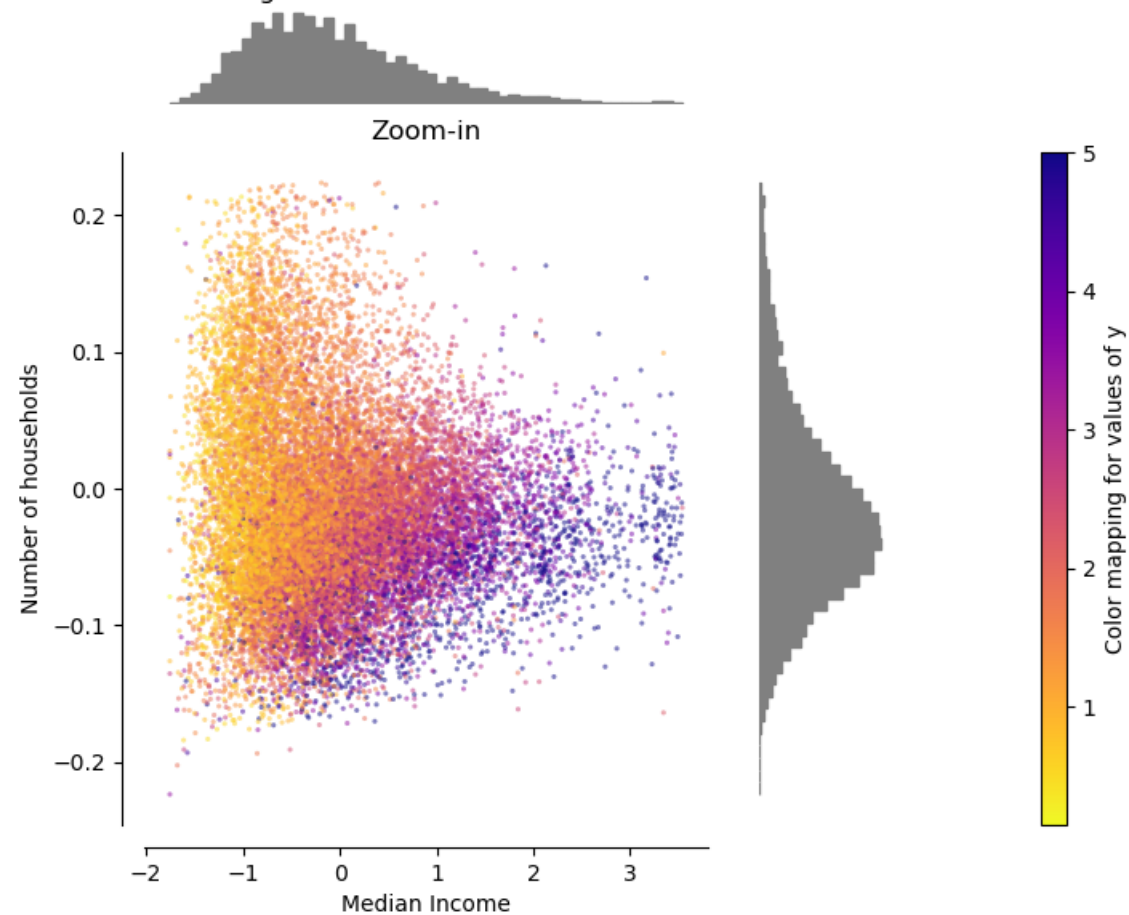
Исходные данные



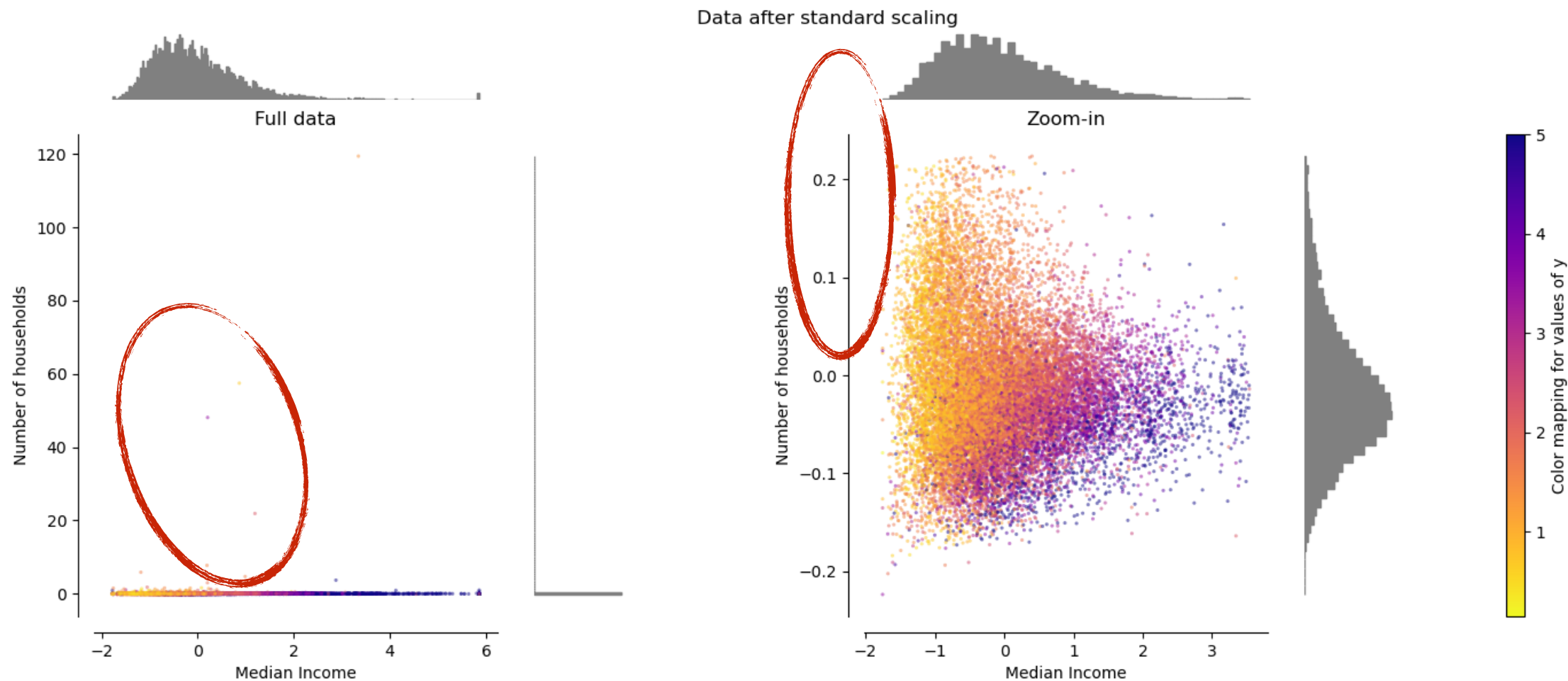
StandardScaler



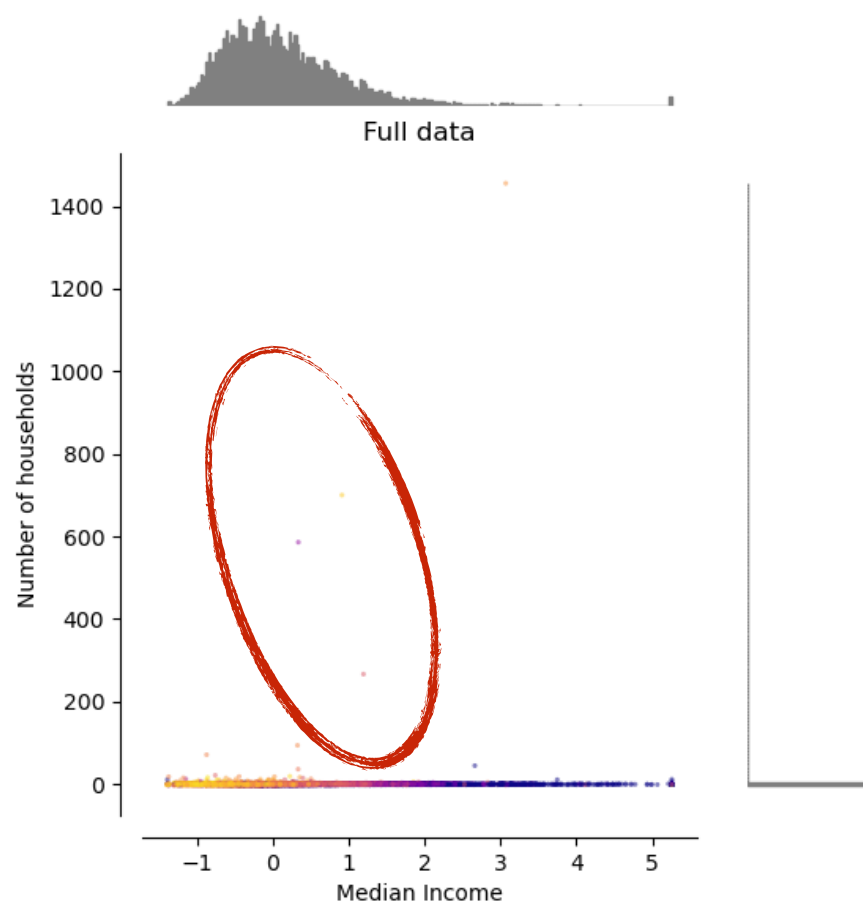
Data after standard scaling



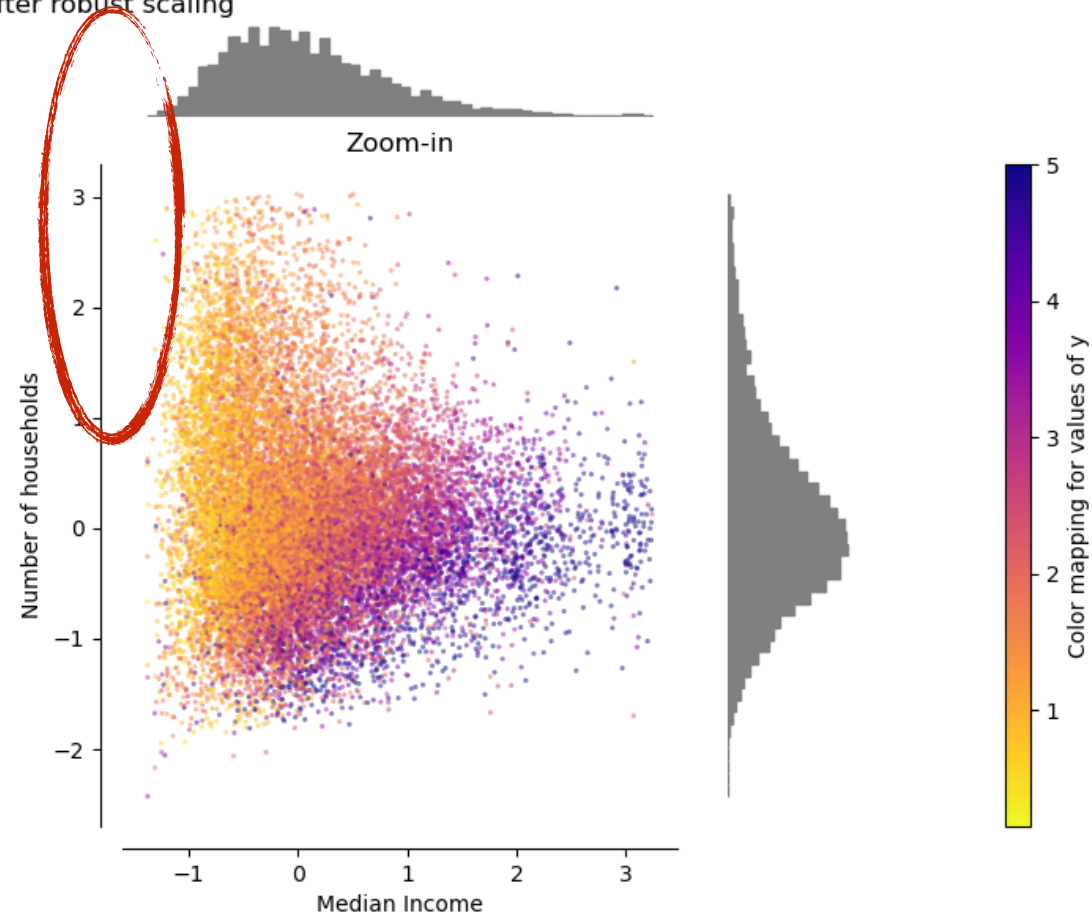
StandardScaler и выбросы



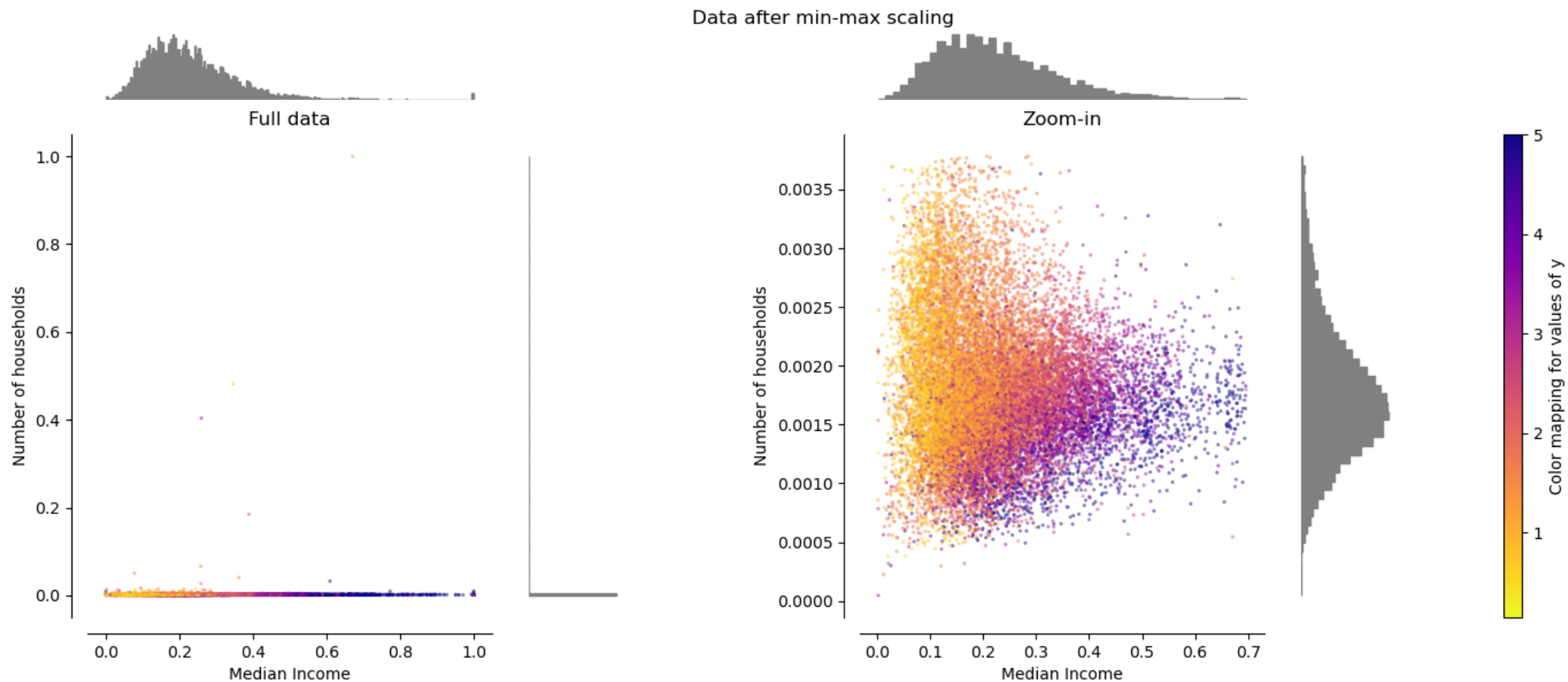
RobustScaler



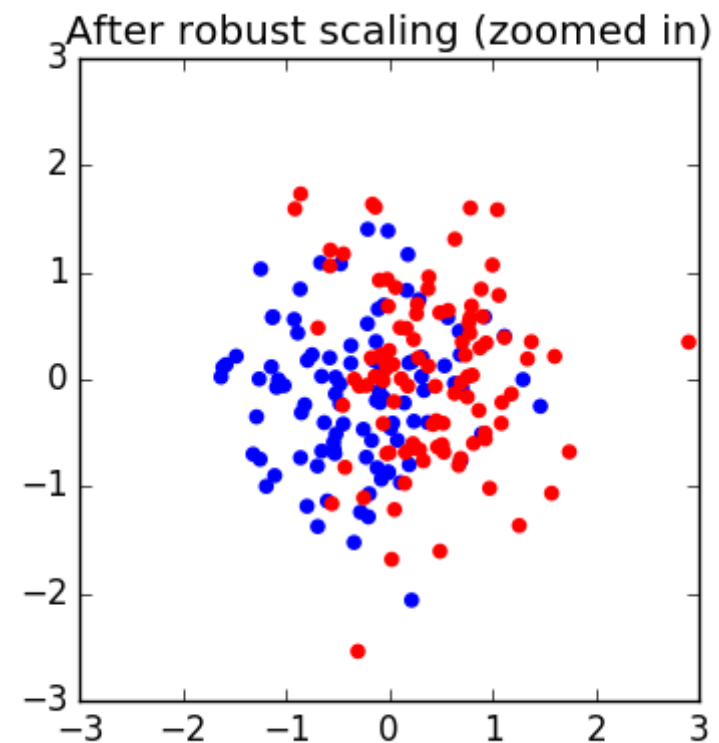
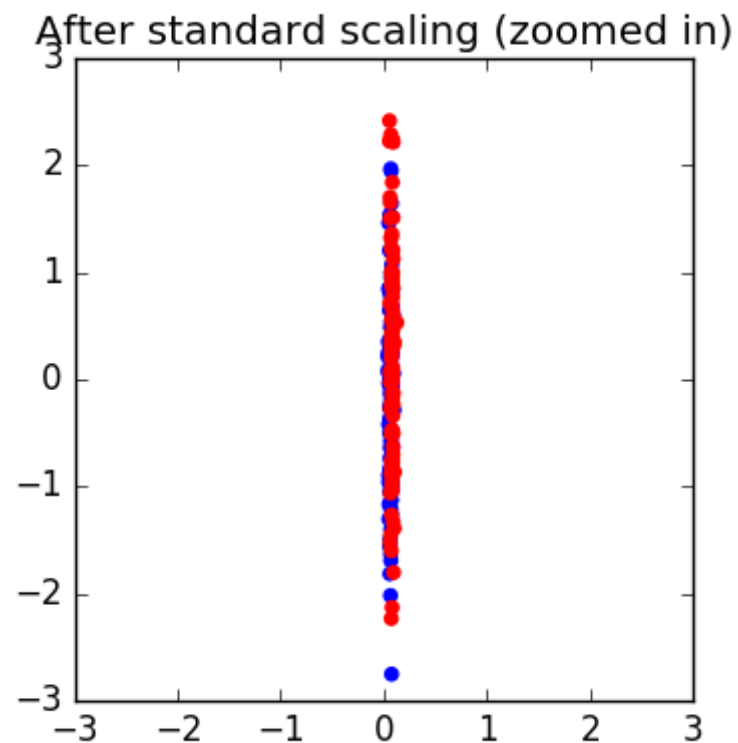
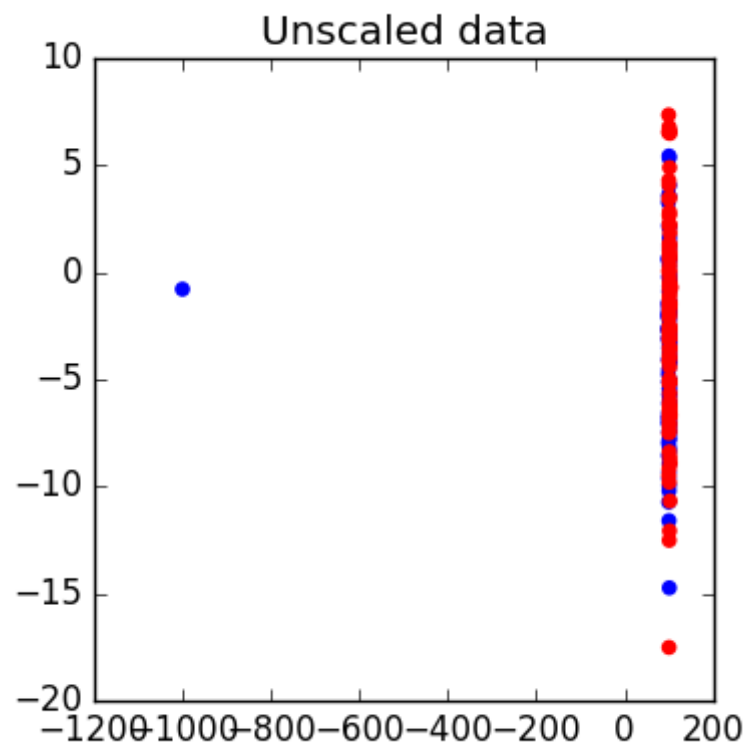
Data after robust scaling



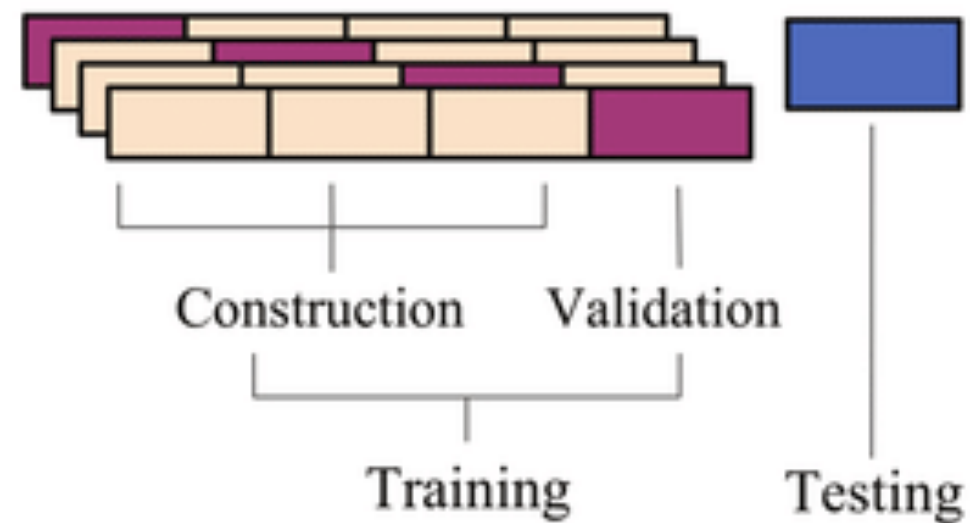
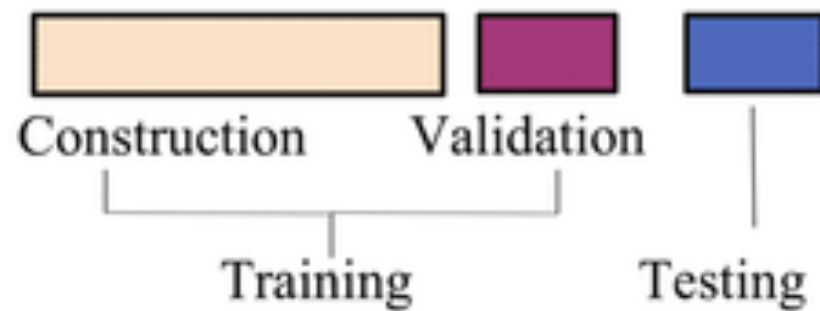
MinMaxScaler



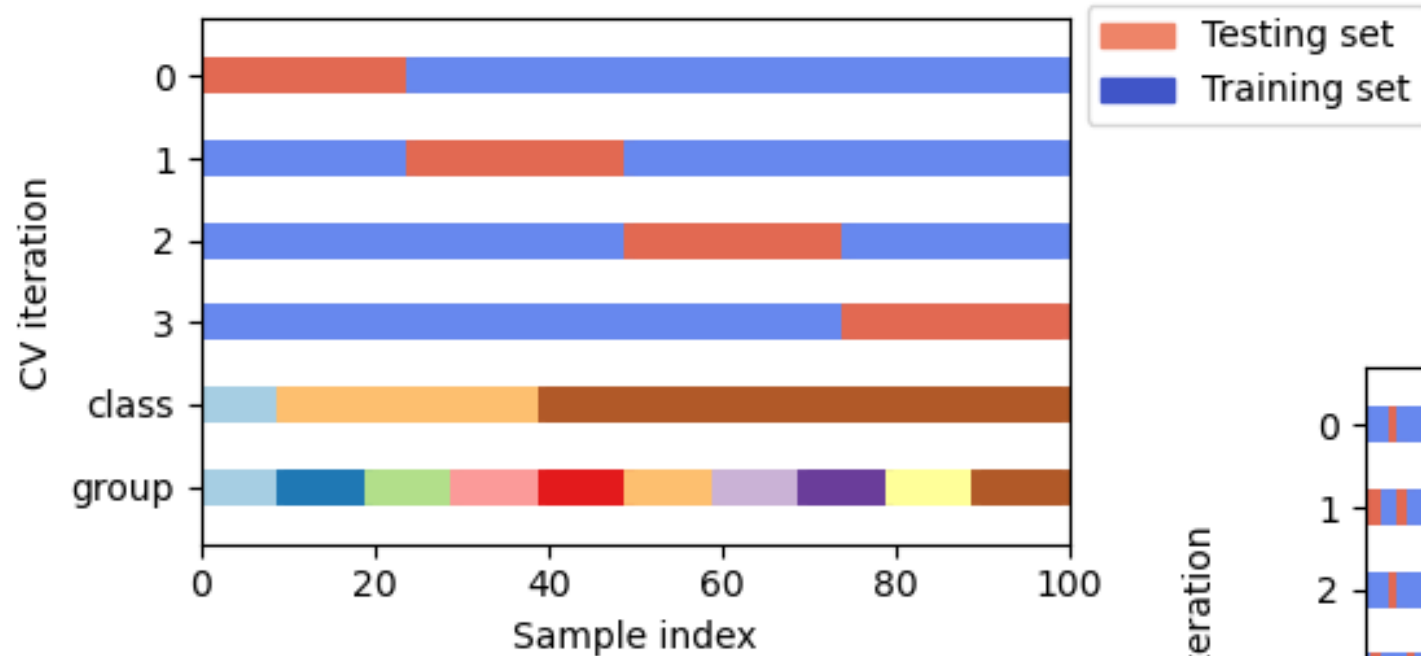
Влияние выбросов



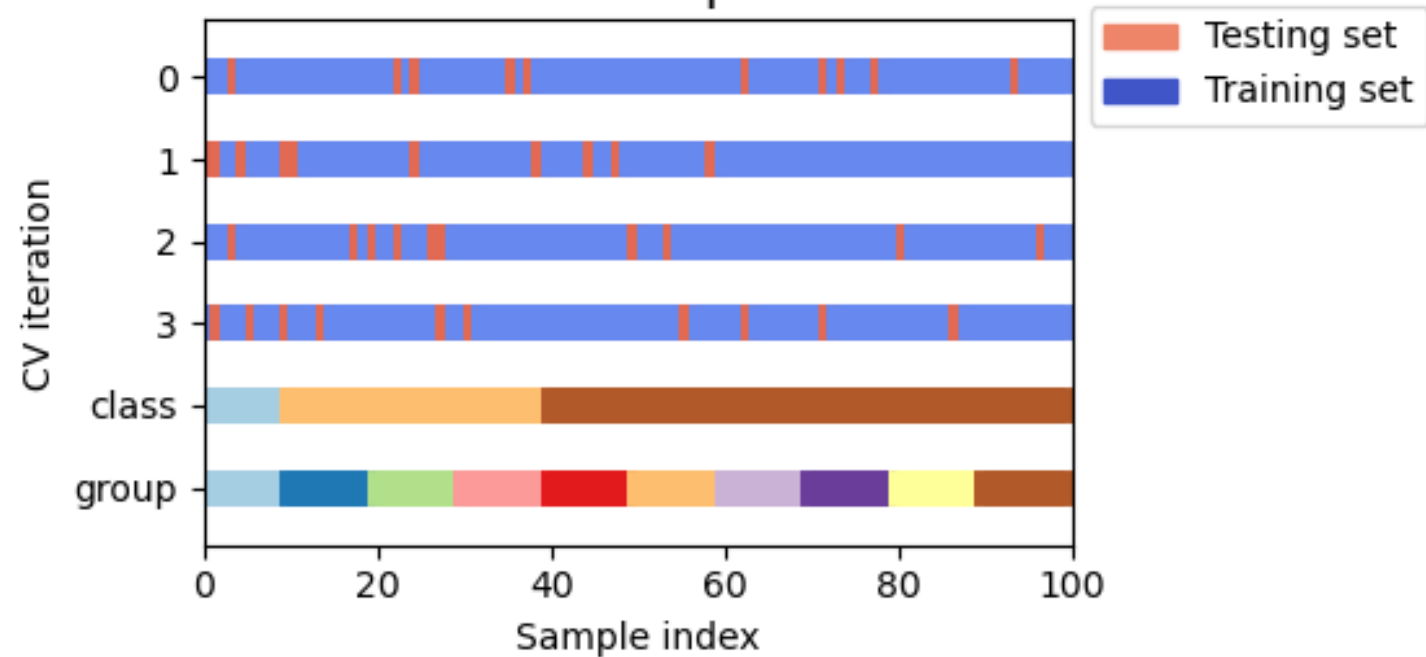
Кросс-валидация



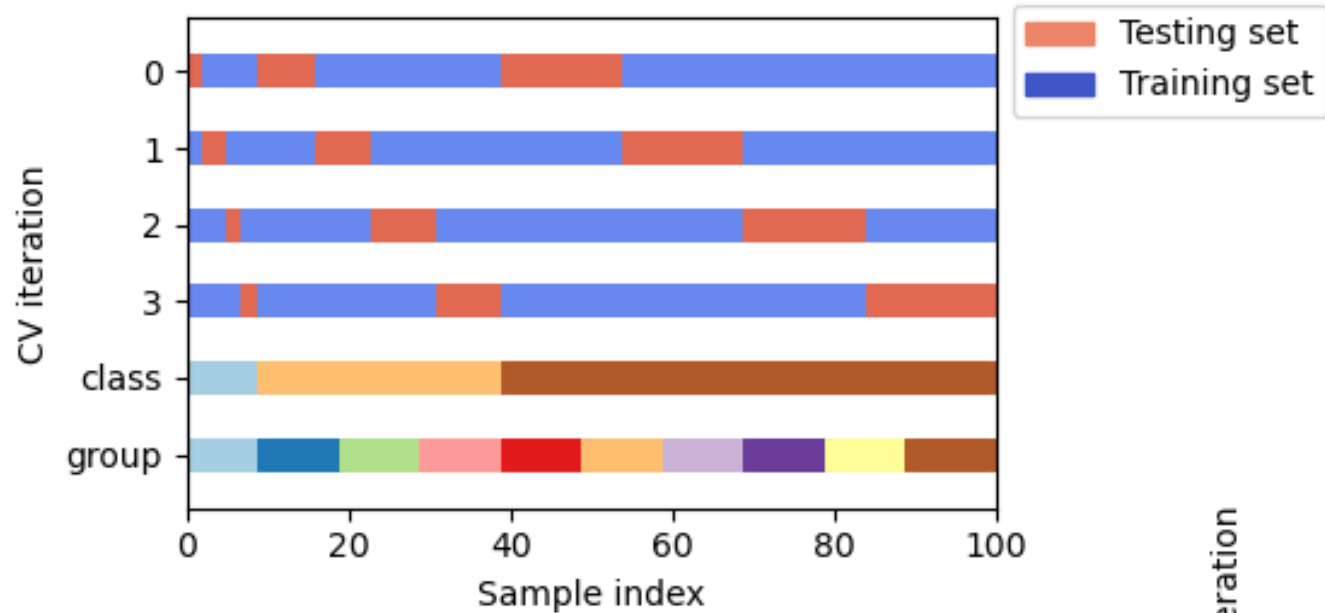
KFold



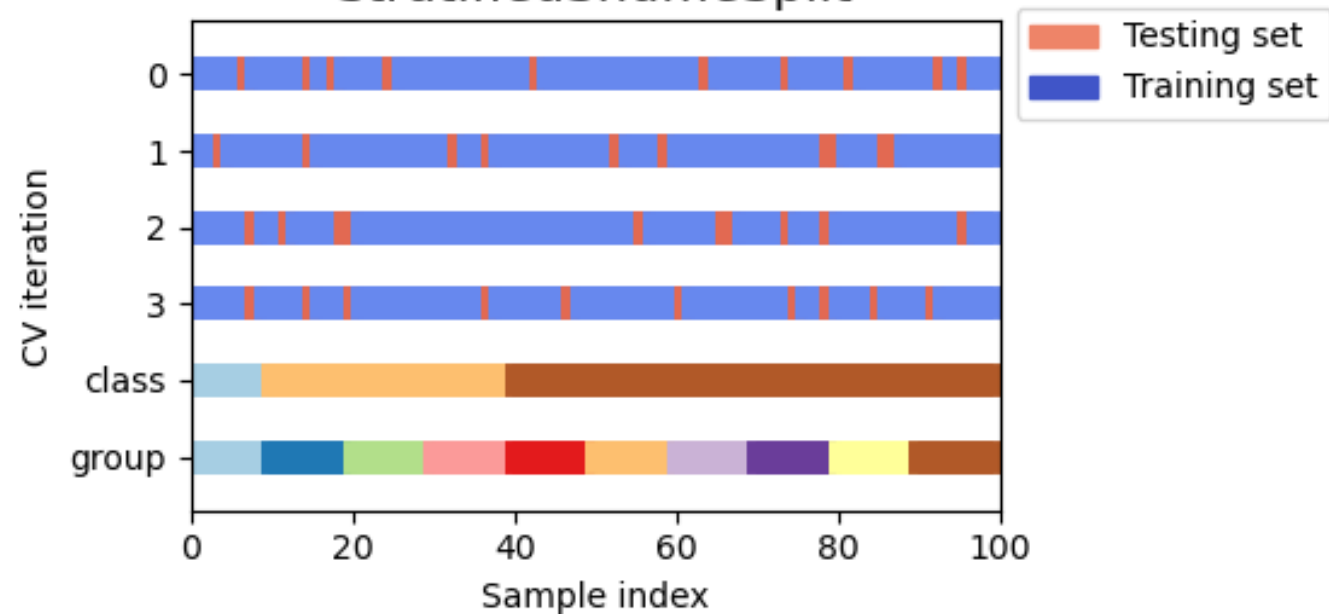
ShuffleSplit



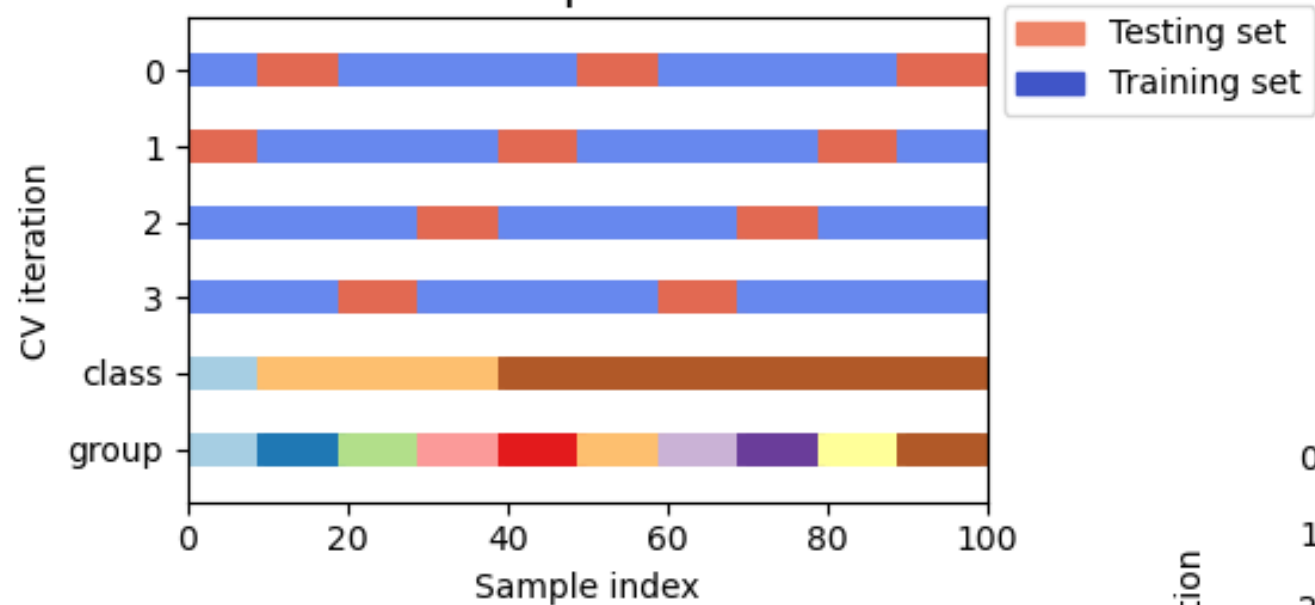
StratifiedKFold



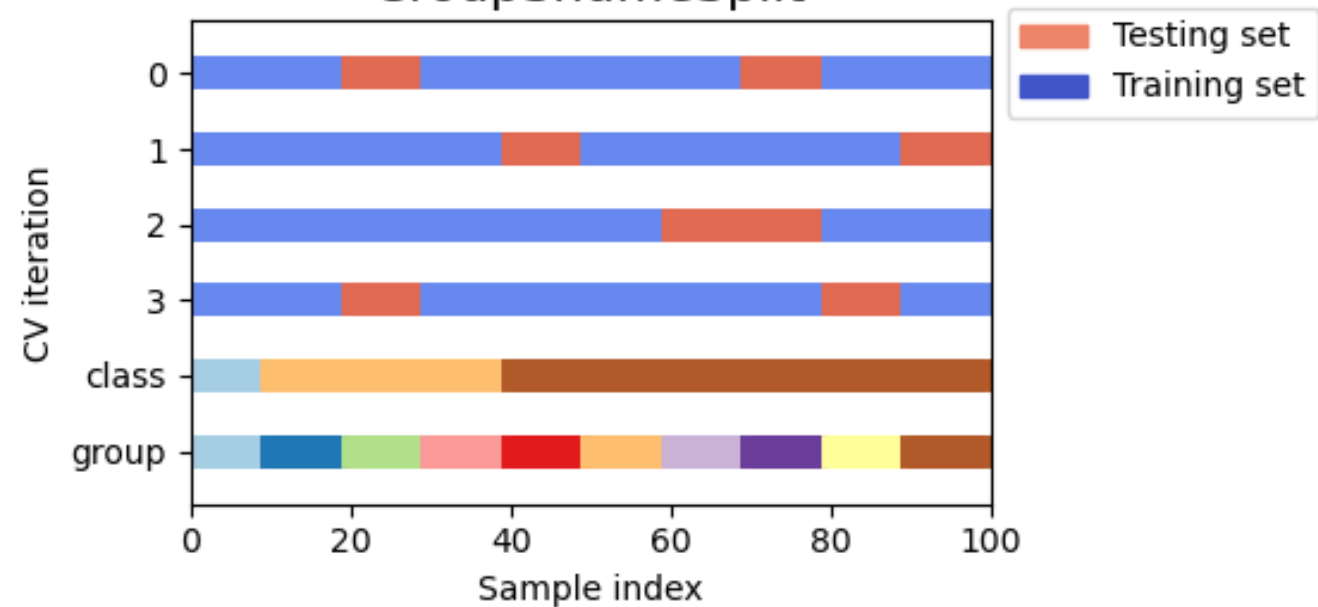
StratifiedShuffleSplit

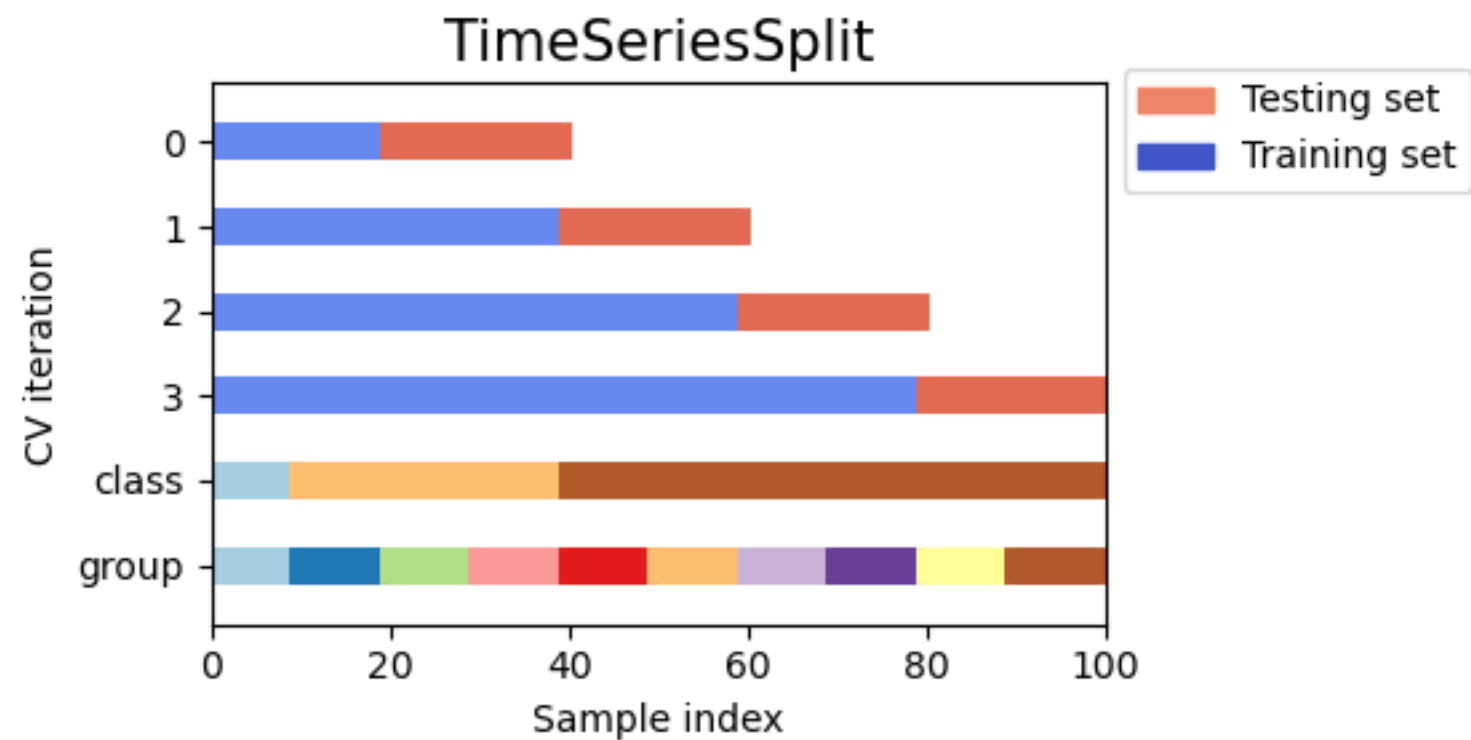


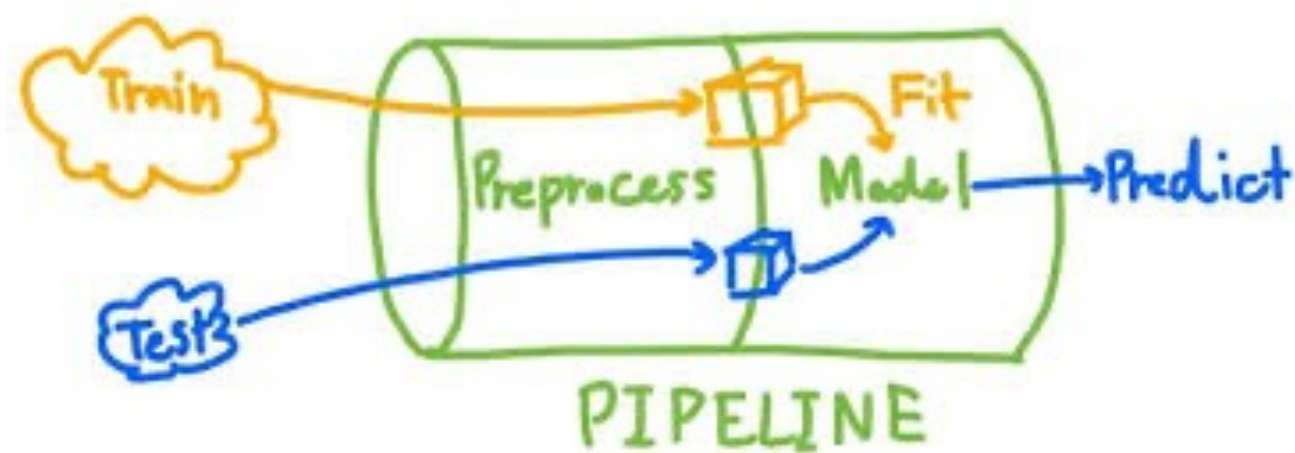
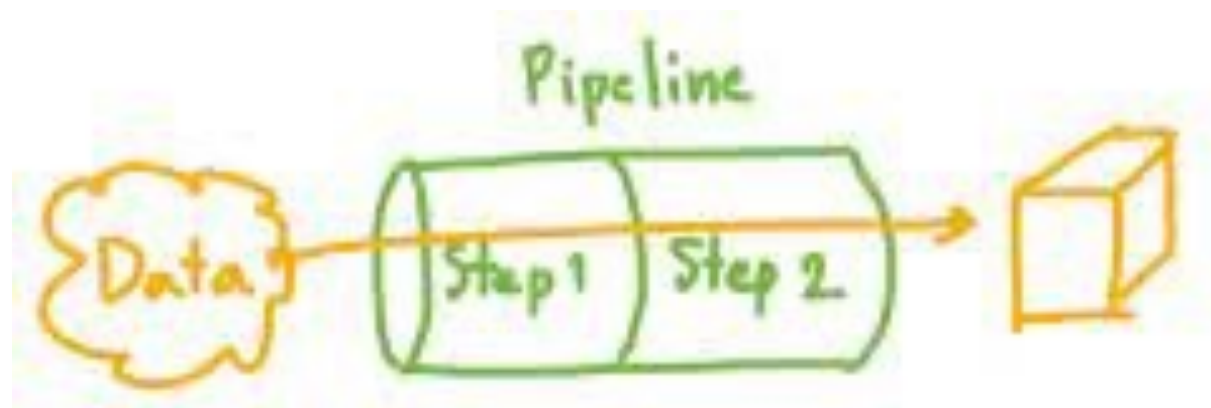
GroupKFold



GroupShuffleSplit









ПРЕЗИДЕНТСКАЯ
АКАДЕМИЯ

СПАСИБО ЗА ВНИМАНИЕ!

Москва

2024

РАНХиГС