

Очистка данных

Поиск выбросов

Skillbox

образовательная платформа

Цель видео

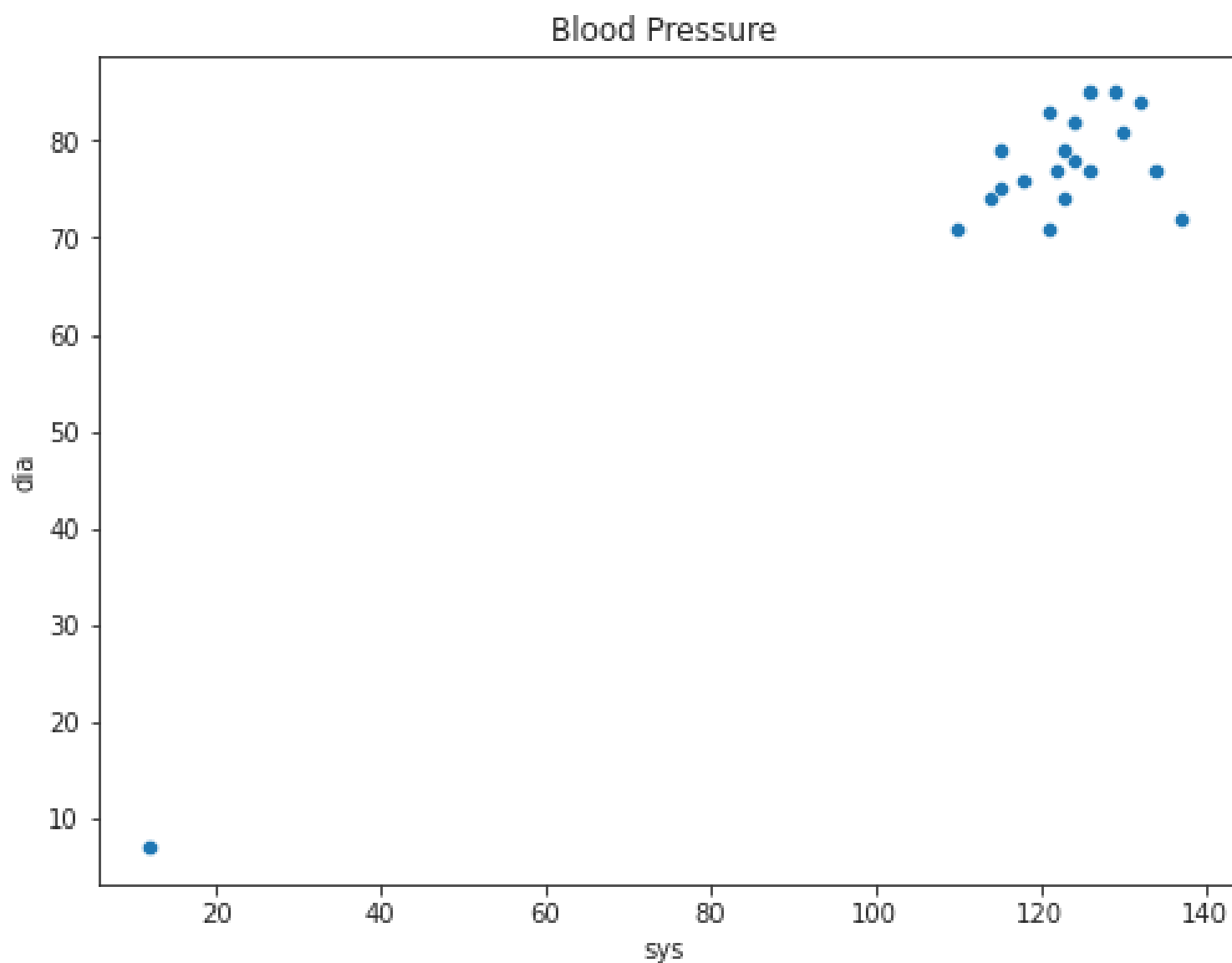
Изучить стратегии поиска и борьбы
с выбросами в данных.

Выброс

Точка данных, которая:

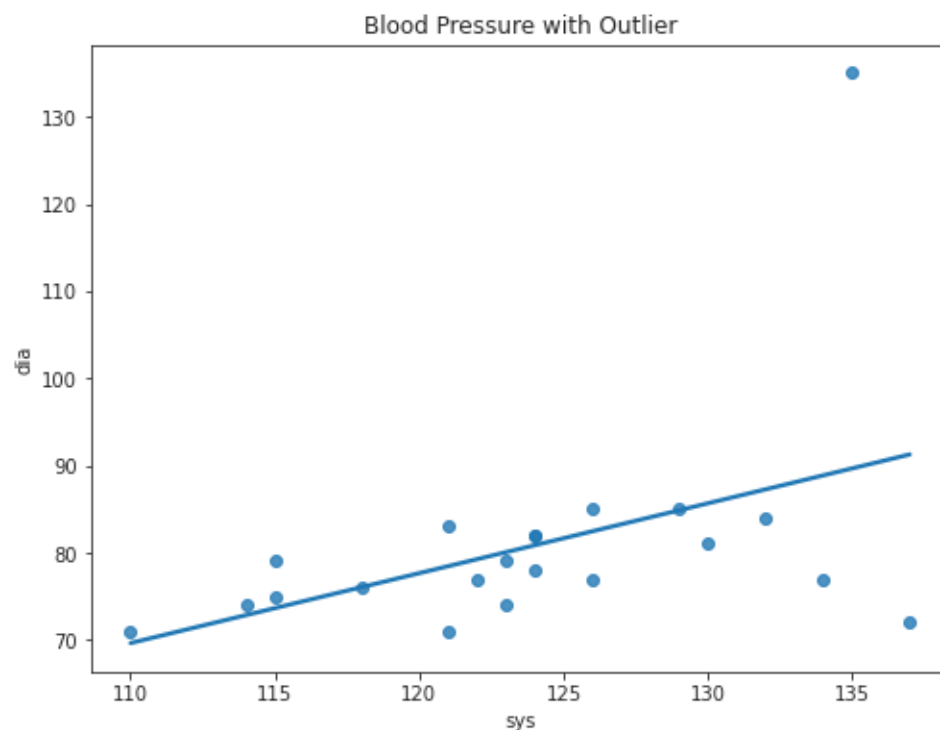
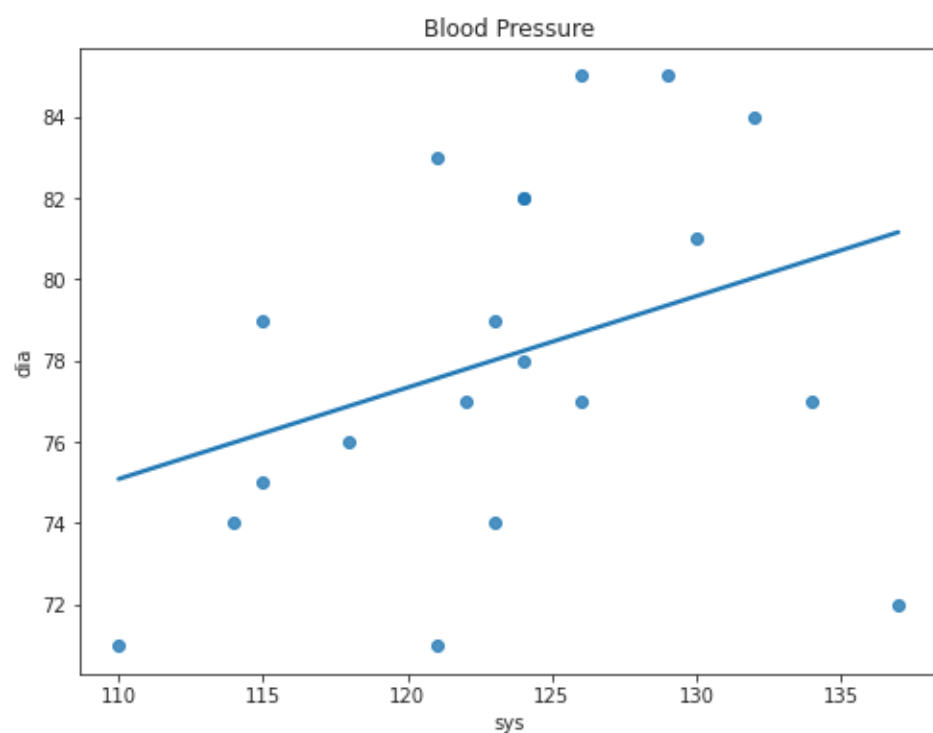
- сильно отклоняется от основного набора данных
- не вписывается в природу этих данных

Примеры выбросов



Изображение: работа Андрея Мещерякова

Примеры выбросов



Примеры выбросов



Изображение: работа Андрея Мещерякова

Аномалия

Точка данных, которая:

- сильно отклоняется от основного набора данных
- но вписывается в природу этих данных

Аномалия

Как отличить от выброса?

Почти всегда — только с помощью вашего понимания природы и особенностей данных.

Поиск выбросов

Вывод

Вы изучили основные стратегии поиска и борьбы с выбросами в данных.