

2.2 pivot+plot, проверка.max.min..

- 1 проверка max min на аномалии
 - 2 баг-репорт
 - 3 пометить срез индикацией
 - 4 Округлять время и переводить его в другие часовые пояса;
 - 5 plot() с группировкой если много точек
-

1 case

проверка max min на аномалии

Определить границу, после которой можно считать заправку «слишком долгой»;

-тупо на глаз отсекали)

-без чека процентов/какой это перцентиль

Найдите аномально быстрые заправки

как нормально:

1. Чекаем max и min в .describe
2. Чекаем **boxplot/violin**
или
топы меньших и больших
3. Удалить выбросы если есть
4. Если есть аномалии - смотрим строчки полностью
5. Чек доли аномалий - если мало - можно выкинуть
6. Чек .describe еще раз

2 баг-репорт - суть ошибки - большая доля слишком коротких, особенно на нектрой азс

query и операторы | & ~

если без скобок то порядок:

сначала not, 2 and 3 or.

3 можно пометить срез

(индикация)

boolean столбцом

4

сдвиг времени

`df += pd.Timedelta(hours=10)`

округление времени

`.dt.round('3H')`

floor ceil

5

точек много на графике -

группировка делает визуально понятнее

`pivot_table().plot()`

увидеть выбросы - по ср

скрыть - по медиане(но не всегда скроются)

vs scatter

выбросы в группах

С помощью графиков с группировкой можно обнаружить выбросы, которые до этого не были видны

немного базы:

`mean(boolean)` = доля

`sum(1)/count(all)` = `count(1)/count(all)`