

2.4 корр-ии

1 scatter alpha hexbin

2 corr coef - шкала

3 треб-я для корр-ии

4 pairplot

case

1 матрица диаграмм рассеяния по исходным данным vs по отфильтрованным(аном макс мин)

2 снизили влияние подозрительных данных на итоговый результат.

1 plot(, kind='scatter')

fix наложение точек scatter

1 **alpha** - прозрачность

2 **hexbin** - плотность точек = цвет шестиугольника

kind='hexbin'

ячеечная диаграмма

2 df.corr()

до 0,2 Очень слабая

до 0,5 Слабая

до 0,7 Средняя

до 0,9 Высокая

свыше 0,9 Очень высокая корреляция

**** 3**Коэффициент корреляции не рассчитывается:**

1 не лин

2 в данных имеется больше 1-го наблюдения по каждому случаю;

3 имеются выбросы,

4 данные содержат ярко выраженные подгруппы

4 pd.plotting.scatter_matrix(data

sns.pairplot

попарный коэфф

тоже .corr()