

Отчет по заданию «Закономерности в ascott group»

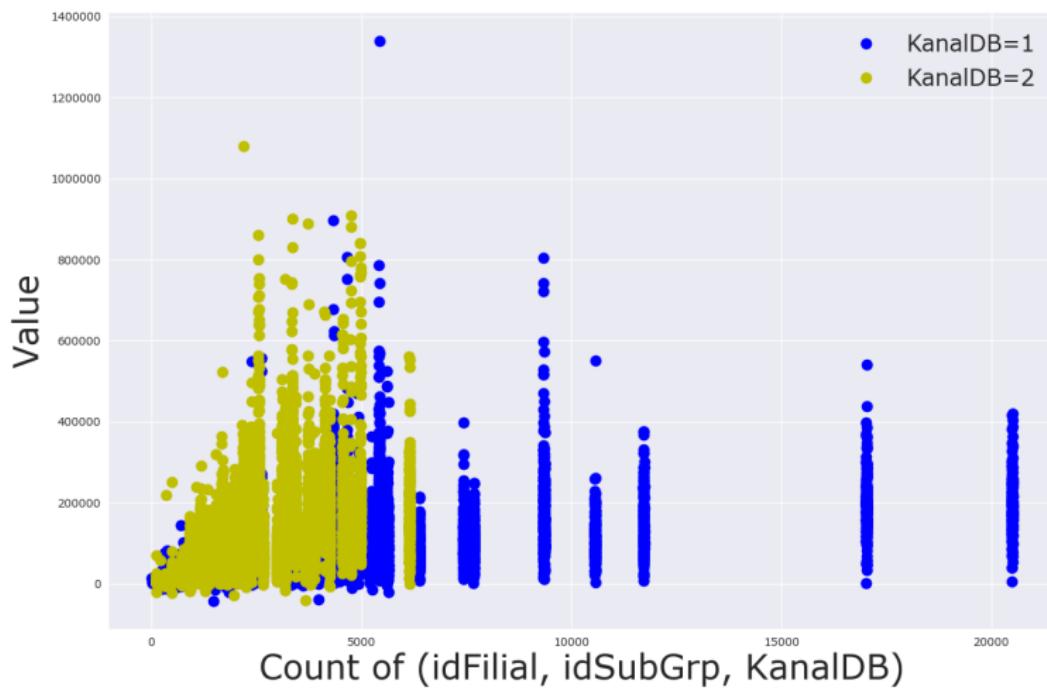
Таскынов Ануар, 517 группа

Курс «Прикладные задачи анализа данных»

14 декабря 2017 г.

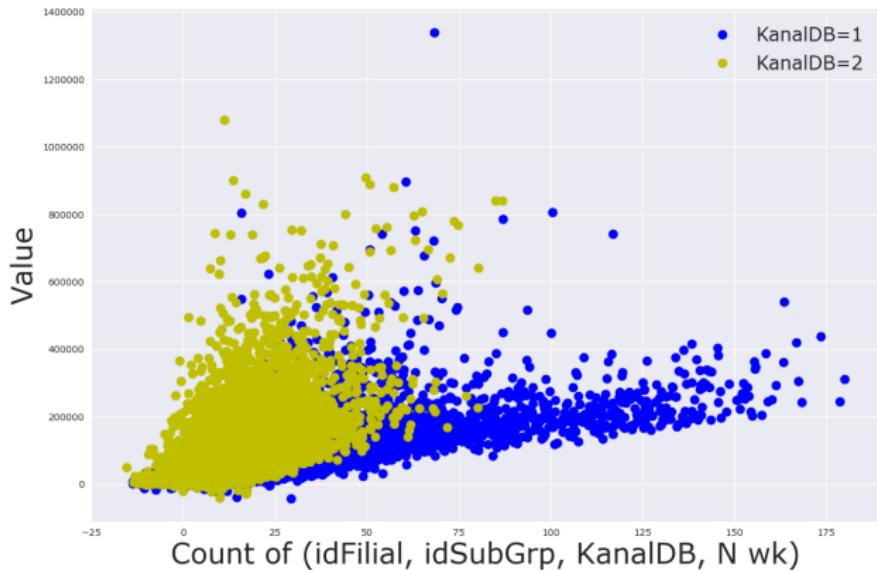
Количество (idFilial, idSubGrp, KanalDB)

Подсчитаем количество троек (idFilial, idSubGrp, KanalDB) в train_set_days и рассмотрим зависимость между значением целевого вектора и количеством троек.



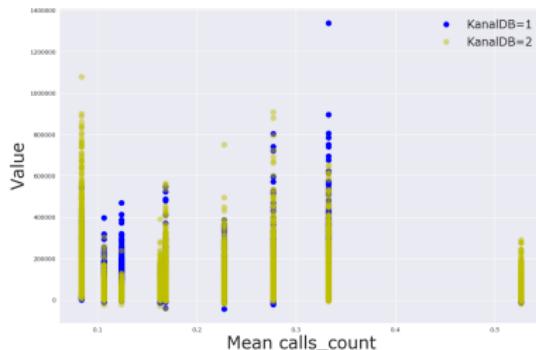
Количество (idFilial, idSubGrp, KanalDB, N wk)

По-хорошему нужно бы сгруппировать по неделям...

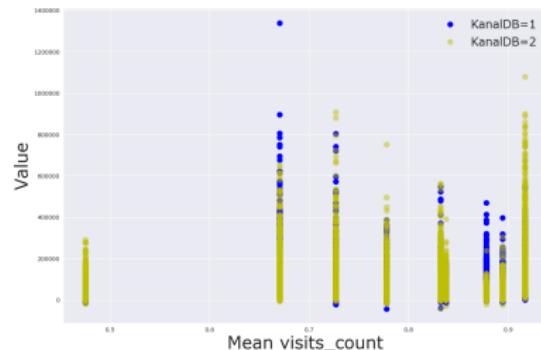


Тут прослеживается явная закономерность, однако пока неясно как внести такой признак в тестовую выборку.

Среднее число visits_count, calls_count



calls_count



visits_count

- Добавление пар (idFilial, idSubGrp, KanalDB) улучшает качество на лидерборде.
- Добавление visits_count, calls_count сильно ухудшает качество.
- Так как в тестовой выборке заранее известны недели, где нет покупок, то можно добавить признак: процент недель для пары (idFilial, idSubGrp, KanalDB), когда делались покупки.