

Разведывательный анализ данных

Иртуганов Мансур 312



Задача - какого поставщика выбрать



Переменные

1. **unsullen.id** - номер кузнеца, диапазон от 1 до 100.
2. **production.date** - дата изготовления, диапазон от 1 до 6.
3. **report.date** - дата отчета для каждого production.date сообщения о сломанном оружии, диапазон от 1 до 7.
4. **produced** - количество созданных мечей для каждого нового месяца у каждого кузнеца, диапазон от 0(101) до 110.
5. **defects** - количество сломанного оружия на каждый новый месяц для каждого момента изготовления, диапазон от 0 до 20
6. **supplier** - поставщик, их два(по условию) и по 1350 записей в таблице, с 1 по 50 - 'harpy.co', с 51 по 100 - 'westeros.inc'(это важно)



Экстремальные значения/выбросы

- Пропущенных значений не обнаружено
- В столбце “defects” было обнаружено значение -1 в единичном количестве, на общую картину не влияло, могло быть выброшено, но тогда бы нарушался порядок по 27 записей на каждого кузнеца.

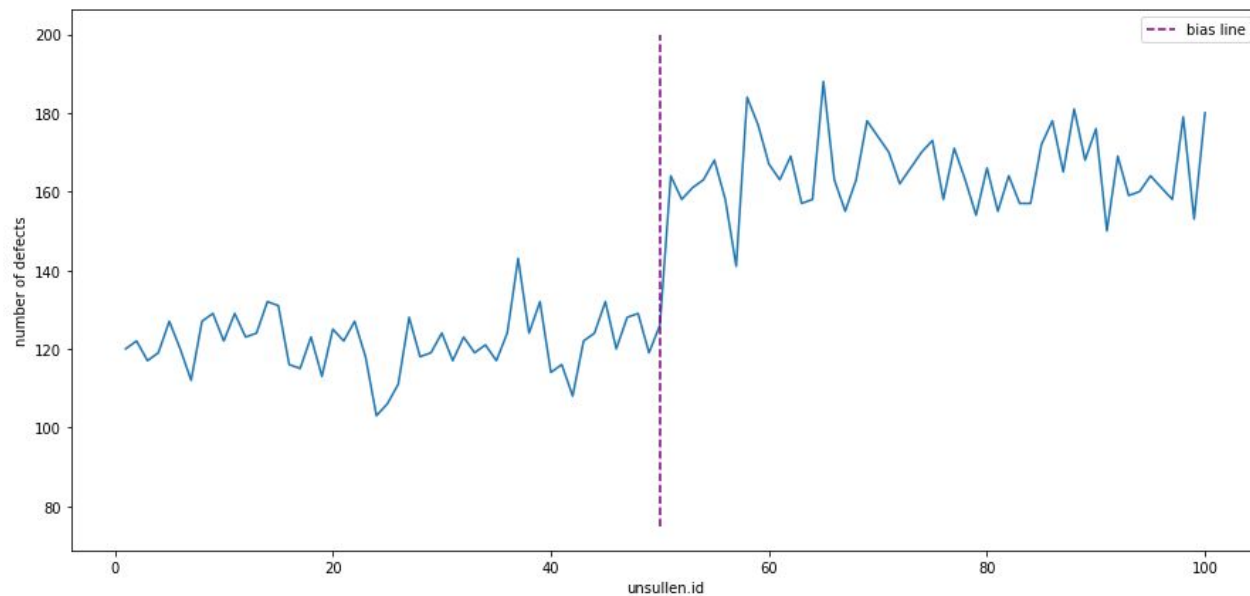


Идея

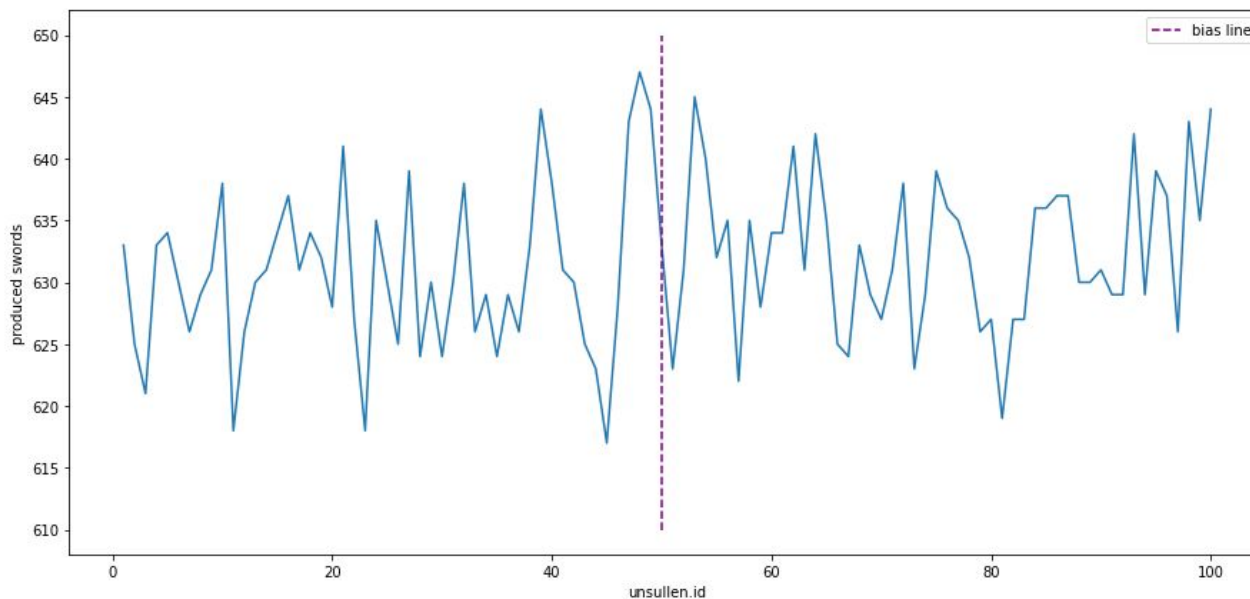
Так как контракты эксклюзивные, то есть не нужно искать оптимальное соотношения между поставщиками, то найдем отношение сломанных мечей к общему количеству для каждого из производителей.

Соответственно, у кого процент брака будет меньше, того и выберем.

Визуальный анализ - количество сломанных



Визуальный анализ - количество произведенных





Визуальный анализ - итоги

На первом графике видно, что в `westeros.inc` количество сломанных резко возросло в отличие от `harpy.co`

На втором графике можно заметить, что количество продукции резко не отличается.



Прямой расчёт

На почти 7% процент брака у **westeroc.inc** выше, чем у **harpy.co**, при почти равном количестве произведенных мечей

Отсюда делаем вывод, что заключить договор надо с достойными господами из Миэрина.