

Команда

Богданова Елена, Чистяков Иван, Раева Анастасия. Все этапы делались совместно

Постановка задачи

После оглушительного успеха в освобождении Астапора, Миэрина и Юнкая от власти работорговцев Дейенерис Бурерожденная открыла себе доступ к Летнему морю, а следовательно -- путь в Вестерос.

Для ведения войны с Семью Королевствами нужно оружие, а для оружия нужна сталь. Нет никаких сомнений в кузнечном искусстве Безупречных, однако поставщики стали не столь надежны. Два основных поставщика стали -- это Westeros Inc. и Harpy & Co. На протяжении нескольких месяцев мы закупаем сталь у обеих компаний, и каждая из них предлагает ощутимую скидку при заключении эксклюзивного договора на поставку.

Советник королевы Тирион Ланнистер знает о твоём умении принимать взвешенные рациональные решения и просит помощи в объективном решении вопроса о том, с какой из компаний следует заключить эксклюзивный договор на поставку стали.

У Тириона есть записи о производстве мечей каждым из кузнецов-безупречных, а также данные о количестве сломанных мечей в каждый из месяцев ведения боевых действий.

Данные имеют вид

	unsullen.id	production.date	report.date	produced	defects	supplier
0	1.0	1	1	103.0	0.0	harpy.co
1	1.0	1	2	0.0	2.0	harpy.co
2	1.0	1	3	0.0	4.0	harpy.co

Ход решения

Чтобы выбрать правильного поставщика были построены различные графики:

- 1) Произведено всего товара
- 2) Всего дефектов товара
- 3) Отношение дефектов к произведенным товарам
- 4) Производство товара по месяцам
- 5) Дефекты товара по месяцам
- 6) Качество продукции по месяцам(процент поломок)
- 7) Плотность поломок в месяц
- 8) Среднее количество поломок в месяц

Программа

- 1) all_period(df), где df - исходная таблица. Исследуем весь период, за который имеются данные.

Строим новые таблицы - одна для Harpy, другая для Westeros. Для этого группируем (groupby) исходное по поставщикам, из полученного извлекаем части

таблицы, относящиеся к разным поставщикам (`get_group`) и сопоставляем каждому поставщику суммарное(`sum`) количество дефектов и продукции за весь период. По полученному с помощью библиотеки `matplotlib` строим графики произведенной продукции за весь период, дефектов, отношения дефектов к произведенным товарам

2) `every_month (df)` Исследуем каждый месяц отдельно

Для этого аналогично предыдущему строим отдельные таблицы для `Harpy` и `Westeros`. Добавляем работу с месяцами. Для каждой таблицы:

Группируем по дате. Заводим 2 массива из шести элементов (столько месяцев), чтобы хранить количество товара и дефектов для каждого месяца. Циклом проходим по таблице для каждого месяца и считаем суммарное количество дефектов и товара.

Считаем процент поломок для каждого производителя в каждый месяц

Выводим графики по месяцам - производство товара, дефекты, качество(процент поломок)

3) `density_damage (df)` Исследуем плотность поломок в месяц

Отбрасываем (`drop`) ненужные столбцы - `Report.date` и `unsullen.id`. Группируем по поставщикам и дате, суммируем -> получаем суммарное количество дефектов и товара в каждый месяц. Делим число дефектов на число товаров и делим получившееся на число месяцев, прошедших с момента производства и * 100. По результатам строим график плотности поломок в месяц

4) `mean_after_month (df)`

Вычитаем из даты, когда сообщили о поломке (`report.date`), дату производства (`production.date`) и переименовываем полученное(`rename`) в `time_to_death` (время жизни = срок службы). Отбрасываем ненужные столбцы `unsullen.id` и `produced`. Затем группируем по дате производства, поставщику и сроку службы, суммируем. Отбрасываем дату производства, снова регрессируем по поставщику и сроку службы. Ищем среднее значение из полученного (`mean`)

По результатам строим график среднего количества дефектов в месяц.