

# Разведочный Анализ Данных

Пилипенко А.О., Банников Д.А.

ВМК МГУ

Апрель 2019

## Постановка задачи

Требуются закупки стали. Поставщики - ненадежны. Два основных поставщика стали – это Westeros Inc. и Harpy. Co. На протяжении нескольких месяцев мы закупаем сталь у обеих компаний, и есть ощутимая скидка при заключении эксклюзивного договора. Советник просит помощи выбора компании для заключения эксклюзивного договора на поставки. Есть записи о производстве мечей, а также данные о количестве сломанных мечей в каждый из месяцев ведения боевых действий. Цель - ответ на вопрос: С каким из поставщиков стали следует заключить договор?

# Используемые методы

## 1. Дефектность в среднем

- ▶ Каждая компания характеризуется скоростью поломки ее продукции. Так как в явном виде мы не можем определить скорость становления продукции дефектной, мы по имеющимся данным можем построить статистику поломки продукции после каждого месяца эксплуатации.

## 2. Совокупный срок службы

- ▶ Одним из параметров продукции является ожидаемое время целостности продукции. При закупке продукции в больших объемах часто используют совокупный срок службы для получения более точных оценок реального объема получаемого товара.

## 3. Продукция без дефектов

- ▶ Объединенные данные об объемах закупленной продукции, еще не ставшей дефектной, на каждый месяц эксплуатации.

# Используемые графики

Для визуализации описанных выше методов мы используем два вида графиков:

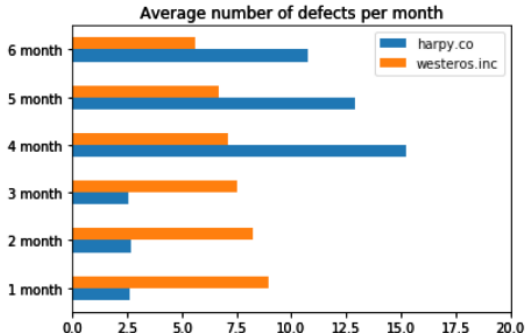
## 1. Столбчатая диаграмма (Bar Graph)

- ▶ Диаграмма, представленная прямоугольными зонами, длины которых пропорциональны величинам, которые они отображают.

## 2. Ящик с усами (Box Plot)

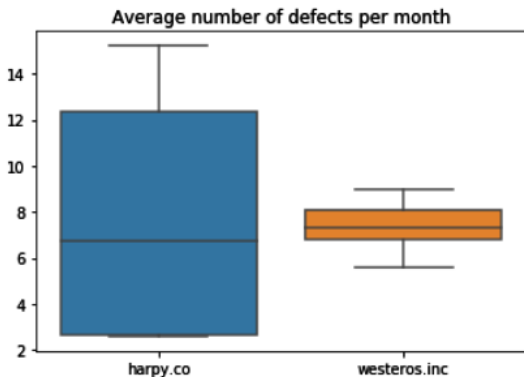
- ▶ График, использующийся в описательной статистике, компактно изображающий одномерное распределение вероятностей. В удобной форме показывает медиану, нижний и верхний квартили, минимальное и максимальное значение выборки и выбросы.

## Дефектность в среднем



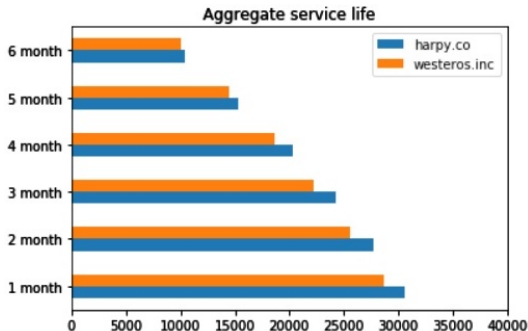
На данном графике видно, что первые три месяца эксплуатации продукция компании Harpy.Co. имеет малое число дефектов. Однако, начиная с четвертого месяца, объем поломок становится в два раза выше, чем у их конкурента в среднем. Здесь трудно определить, которая из компаний имеет худшие показатели, но продукция компании Harpy.Co. имеет тенденцию к большому числу дефектов в дальнейшем.

## Дефектность в среднем



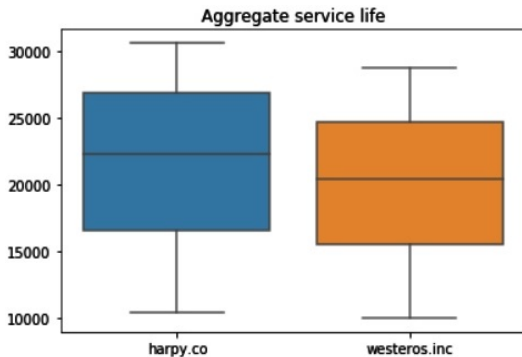
Наблюдается широкий разброс среднего числа дефектов в месяц у продукции компании Harpy.Co., однако медианное число дефектов всё же не превышает значения их конкурентов Westeros Inc.

# Совокупный срок службы



На графике наблюдаются схожие картины, компания Harpy.Co. единогласно опережает своего конкурента, однако различия менее 10%, которые к шестому месяцу сходят к нулю.

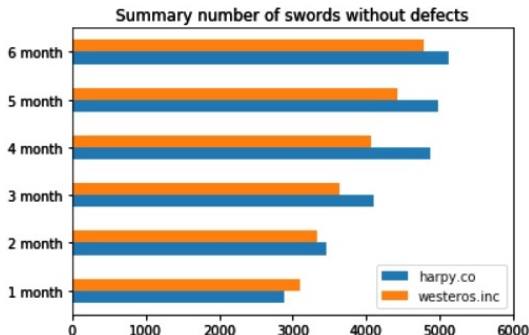
## Совокупный срок службы



Данный график является менее показательным в данном методе, поскольку не отражает тенденции, наблюдаемые на предыдущем графике. Объединив данные результаты с предыдущими мы можем вывести гипотезу о скорейшем спаде совокупного срока службы продукции компании Harpy.Co., чем у компании Westeros Inc. за пределами собранных данных.

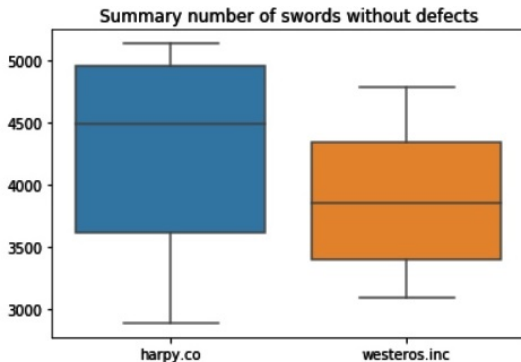


# Продукция без дефектов



В данной ситуации наблюдается полное преимущество компании Harpy.Co. над своим оппонентом. Но, полученные ранее результаты отражены в наблюдаемой здесь тенденции - скачок, возникший за первые четыре месяца, начинает терять свою силу из-за высокой скорости дефекции на поздних стадиях эксплуатации.

# Продукция без дефектов



Наблюдаемое полное доминирование компании Harpy.Co. получило ложную преувеличенность в данном графике. Данный график является непоказательным, так как абсолютно не отражает тенденций продукции. Однако, компания Harpy.Co. с радостью покажет данный график в своей рекламной компании.

# Выводы

В итоге Разведочного Анализа Данных были получены следующие результаты:

- ▶ Многие методы анализа работают эффективно только в совокупности, однако некоторые из них содержат частичное отражение тенденций из других методов.
- ▶ Объем собранных данных напрямую влияет на точность определения гипотез о дальнейших изменениях процессов.
- ▶ Советнику стоит заключить договор с компанией Westeros Inc. из-за производства компанией Harpy.Co. стали с задержанным эффектом разрушаемости.