

Отчет по заданию 4
Разведовательный анализ данных

«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ИМЕНИ М.В. ЛОМОНОСОВА

Липатова Александра, Разумова Вера, Смирнов Михаил

Постановка задачи

Необходимо провести разведывательный анализ данных с целью ответа на вопрос: "С каким из поставщиков стали следует заключить договор?". То есть задача сводится к транспортной задаче по поиску максимально выгодного поставщика.



Основные понятия

Разведочный анализ данных (англ. exploratory data analysis, EDA) — анализ основных свойств данных, нахождение в них общих закономерностей, распределений и аномалий, построение начальных моделей, зачастую с использованием инструментов визуализации.

Понятие введено математиком Джоном Тьюки, который сформулировал цели такого анализа следующим образом:

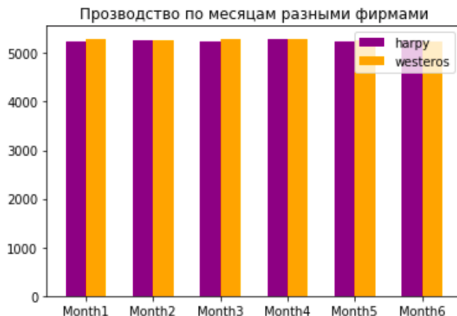
- максимальное «проникновение» в данные
- выявление основных структур
- выбор наиболее важных переменных
- обнаружение отклонений и аномалий
- проверка основных гипотез
- разработка начальных моделей.

Ход работы

1. Получить данные из CSV таблицы.
2. Провести анализ количества продукции, произведенного компаниями Harpy и Westeros каждый месяц.
3. Для большей наглядности изобразим разницу количества дефектов на графике.
4. Рассмотреть плотность распределения усредненных по месяцам дефектов к произведенному товару.
5. Построим графики количества дефектов продукции компании для каждого месяца.

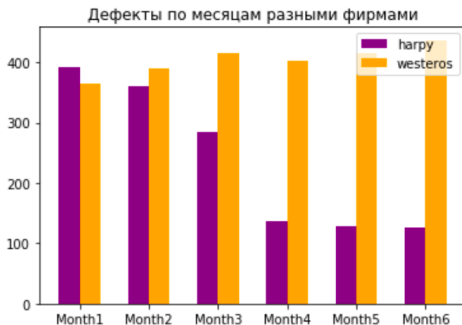
Описание программы

1. Построим диаграмму количества продукции по месяцам.



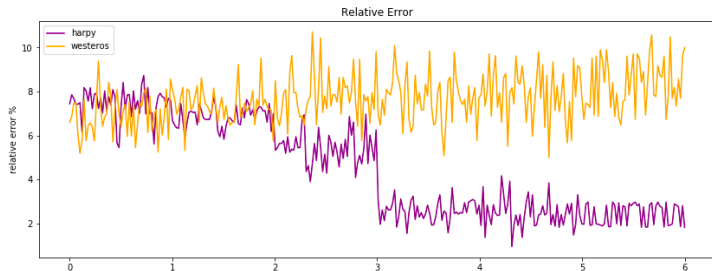
Описание программы

2. Построим диаграмму количества дефектов по месяцам.



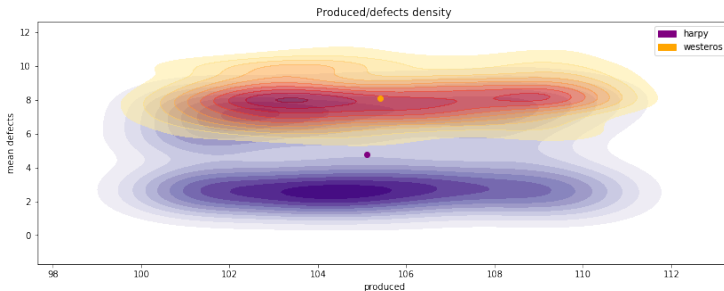
Описание программы

3. Построим график относительной ошибки.



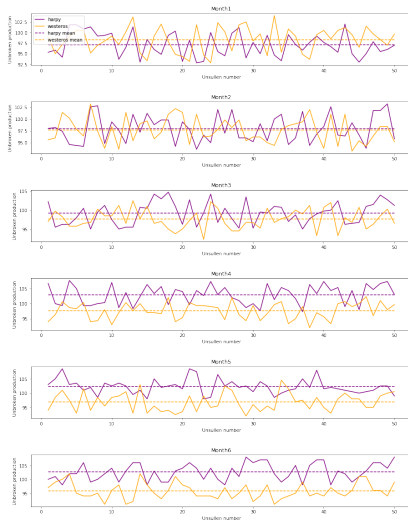
Описание программы

4. Построим диаграмму плотности распределения усредненных по месяцам дефектов к произведенному товару



Описание программы

5. Построим графики количества дефектов по всем месяцам.



Выводы

Видно, что при примерно одинаковом количестве произведенной продукции среднее количество дефектов компании Harpy ниже, чем компании Westeros. На основании проделанного анализа делаем вывод, что работать стоит с компанией Harpy.

