



Статистический анализ датасета “Avocado prices”

Kosimov M.N.



1) Сравнение цен на обычные и органические авокадо

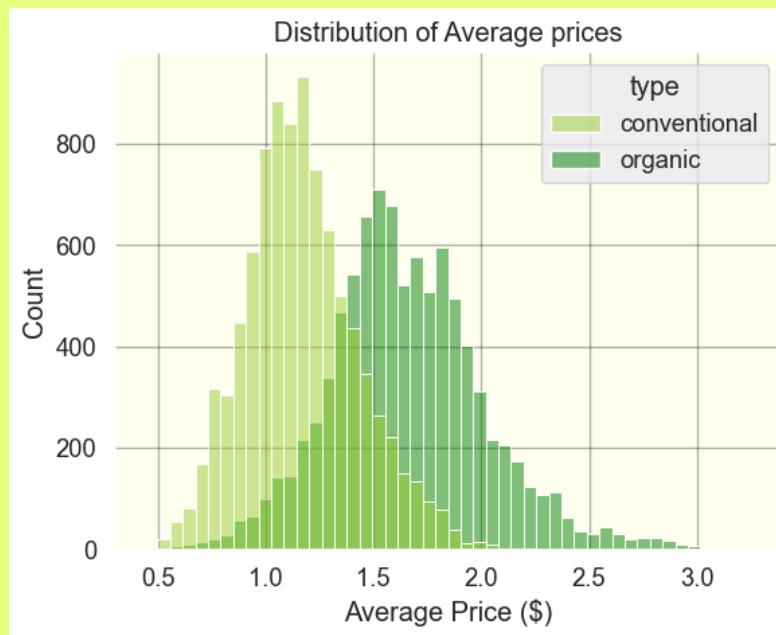


Рис 1. Гистограмма распределения средних цен на органические и обычные авокадо.

Применяемый тест:

двуихвборочный z-тест

Нулевая гипотеза: $\mu_1 - \mu_2 = 0$

(органические = обычные) **Z-статистика -** 105.587

Альт-ая гипотеза: $\mu_1 - \mu_2 > 0$

Результаты :

p-value - 0.0.

2) Построение 95% доверительных интервалов для средних цен на органические и обычные авокадо.

Для средней цены на орг-ое авокадо: [1.6465, 1.6614]

Для средней цены на обычное авокадо: [1.1598, 1.1705]

Интерпретация: при многократном построении 95% доверительных интервалов по случайным выборкам в 95% случаев он покрывает истинное значение средней цены на авокадо, а в 5% случае – нет.

3) Проверка наличия зависимости между ценой на авокадо и количеством проданных авокадо

$$\text{Уравнение: } Y = 14.69 - 2.74X$$

Коэф-нт детерминации: 0.299

Коэф-нт корреляции Пирсона: -0.547 (p-value=0.0)

slope significance:

t=-46.835, p-value=0.0

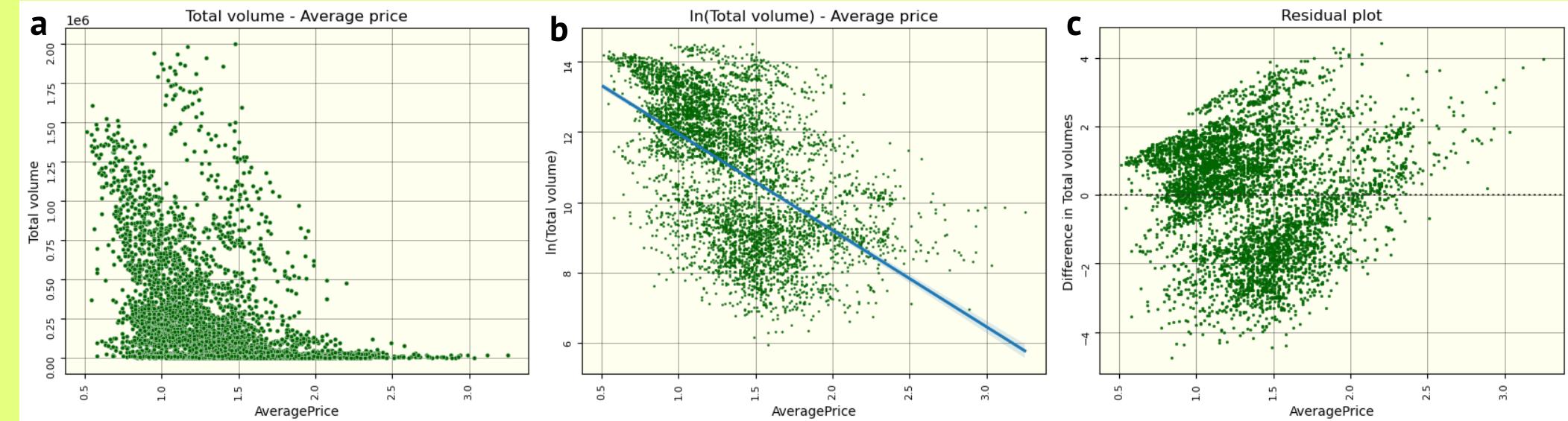


Рис 2. а) scatterplot средней цены и кол-ва проданных авокадо б) график зависимости ln(кол-ва проданных авокадо) от средней цены, $Y = 14.69 - 2.74X$ с) график остатков регрессии

4) Сравнение цен на авокадо в регионах

Применяемый тест: Однофакторная ANOVA

Нулевая гипотеза: равенство средних

Альт-ая гипотеза: неравенство средних

Результаты : F-статистика - 0.8085
p-value - 0.4891

Выводы

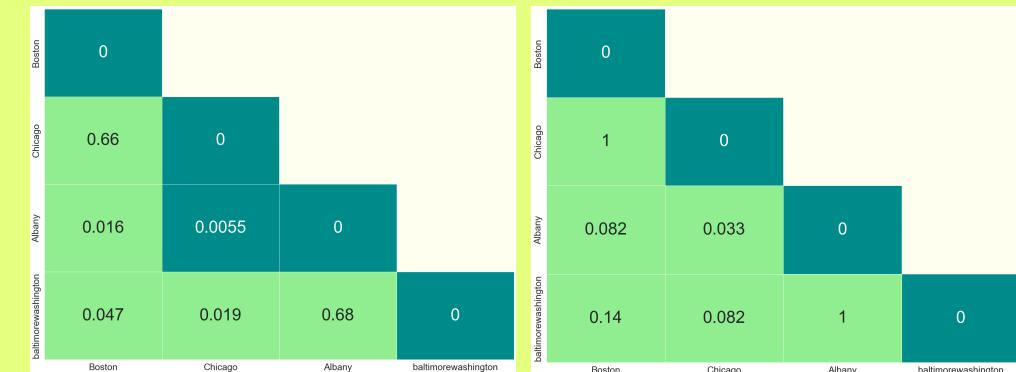


Рис 3. Матрица P-value, полученных в результате попарного F-теста (слева) и после применения поправки Холма-Бонферрони (справа)

- Средняя цена на органическое авокадо выше, чем средняя цена на обычное авокадо.
- Мы на 95% процентов уверены, что настоящая средняя цена на органическое авокадо лежит в интервале [1.6465, 1.6614], а на обычное авокадо в интервале [1.1598, 1.1705].
- Между средней ценой на авокадо и ln(количество проданных авокадо) имеется статистически значимая корреляция. С увеличением цены на авокадо ln(кол-во проданных авокадо) уменьшается.
- Средняя цена на авокадо одинакова в Бостоне, Чикаго, Олбани и Балтиморе-Вашингтоне.