

Университет ИТМО

Практическая работа №1
по дисциплине «Визуализация и моделирование»

Автор: Мигулаева Татьяна Алексеевна

Поток: ВИМ 1.2

Группа: К3222

Факультет: ИКТ

Преподаватель: Чернышева А.В.

Санкт-Петербург, 2021 2021г.

Датасет:

В данном датасете представлены еженедельные данные сканирования розничной торговли за 2018 год по продажам разрозненных сортов авокадо в различных городах. Данные сканирования розничной торговли поступают непосредственно из кассовых аппаратов розничных продавцов на основе фактических розничных продаж авокадо Хасс.

Данные датасета:

Название	Суть столбца	Тип данных	Тип шкалы
Date	Дата наблюдения	date	относительная
AveragePrice	средняя цена одного авокадо	float	относительная
Type	обычные или органические	string	номинальная
Year	год	integer	
Region	город или район наблюдения	string	номинальная
Total Volume	Общее количество проданных авокадо	integer	относительная
4046	Общее количество проданных авокадо с PLU 4046	integer	относительная
4225	Общее количество проданных авокадо с PLU 4225	integer	относительная
4770	Общее количество проданных авокадо с PLU 4770	integer	относительная
Small bags	Количество проданных стандартных упаковок с авокадо	integer	относительная
Large bags	Количество проданных больших упаковок с авокадо	integer	относительная
Total bags	Общее количество проданных упаковок с авокадо	integer	относительная

Задачи датасета:

Проанализировать потребность в авокадо за определенный промежуток времени. Проанализировать, зависит ли количество покупаемых авокадо от сорта. Вычислить самый популярный сорт авокадо, а также в чем потребность больше - в упаковках авокадо или

поштучной продаже. Проверить, обращают ли внимание покупатели на тип выращивания авокадо. Проверить, зависит ли популярность авокадо от района наблюдения.

Список диаграмм и гипотез

0.1 Гипотеза 1

В наше время при выращивании авокадо традиционный способ используется больше, чем органический, так как сейчас все виды производства переходят на фабрики и удобрения, ускоряющие рост растений, несмотря на их потерю полезных свойств. Как мы

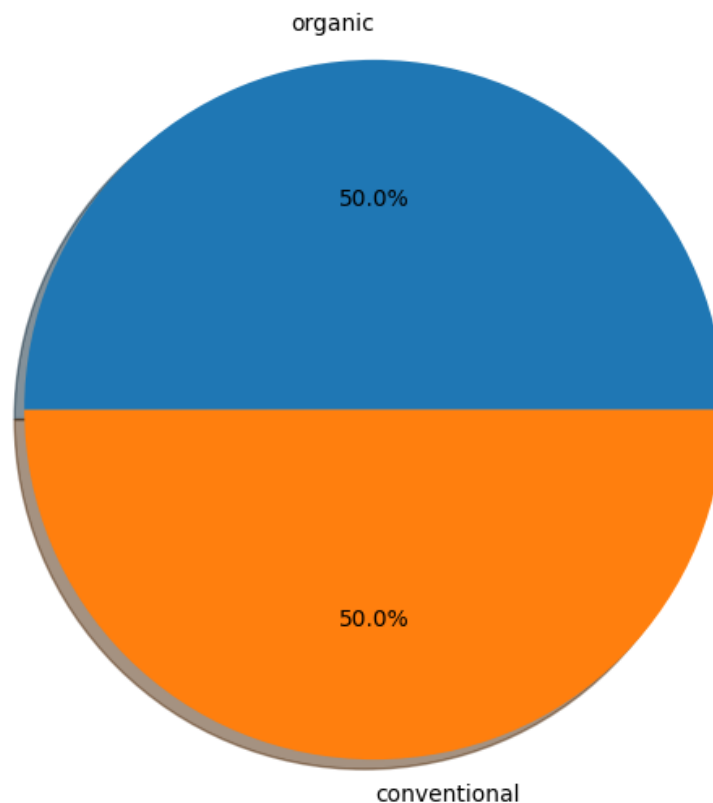


Рис. 1: Сравнение органических и традиционных авокадо

можем видеть – гипотеза не подтвердилась. По всей видимости, авокадо очень часто покупают те потребители, которые заинтересованы в наличии в нем полезных веществ и поэтому готовы немного переплатить за органическое производство

2. Авокадо бывают разных видов, каждый из которых имеет свой номер. В данном датасете анализируются авокадо под номерами 4046, 4225 и 4770. Проанализируем, насколько зависит возможность покупки авокадо от его вида

Как мы можем видеть, за 2015 год самым популярным было авокадо под номером 4225,

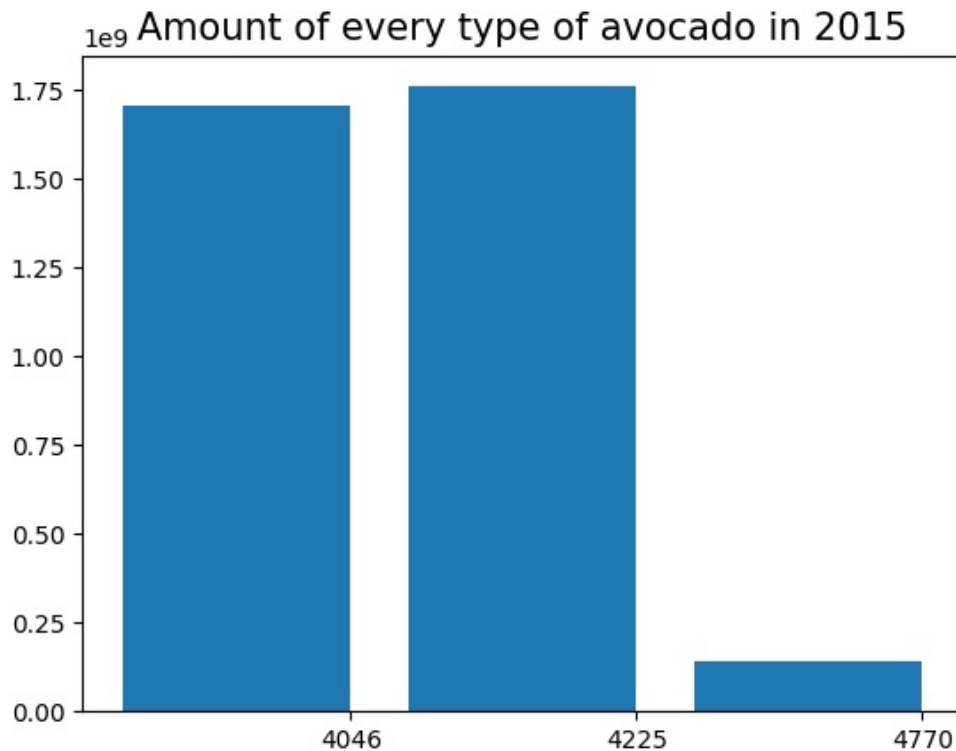


Рис. 2: Сравнение популярности видов авокадо

на втором месте – 4046, а авокадо 4770 пользуется наименьшим спросом.

3. Предположим, что так как популярность здорового образа жизни сейчас возрастает, то и потребление авокадо с каждым годом будет постепенно увеличиваться.

Как видим по диаграмме, гипотеза подтвердилась – в 2015 году куплено чуть меньше авокадо, чем в 2016 и 2017. В 2018 году доля покупок авокадо намного меньше, так как данные сняты в начале 2018 года. Поэтому в дальнейшем при подобном анализе будем исключать 2018 год из анализа.

4. Исходя из предыдущей диаграммы, узнаем, растет ли потребление авокадо, построив таким образом линейную диаграмму

Как можно по ней видеть – да, количество покупок авокадо растет с 2015 по 2017 год.

5. В датасете присутствует информация о продажах упаковок авокадо разного раз-

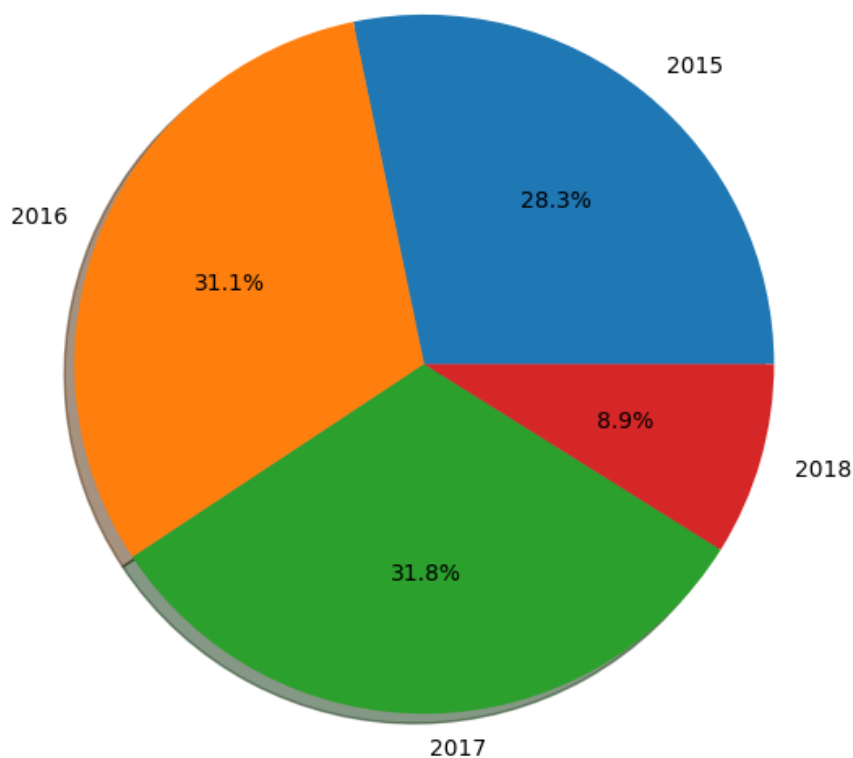


Рис. 3: Потребление авокадо за 4 года

мера – малых, средних и крупных. Предположим, что малых упаковок продается намного больше. Построим диаграмму, показывающую количество проданных упаковок каждого вида.

Гипотеза очевидно подтверждена – чем меньше упаковка, тем больше их продается.

6. Как все мы знаем – с каждым годом все вокруг дорожает. Предположим, что даже авокадо в США не является исключением, узнав изменение средней цены на него с 2015 по 2017 год.

Как видим, с 2015 по 2016 год цена немного опустилась, но к 2017 резко поднялась.

7. В датасете присутствует информация о продажах авокадо в каждом штате, но так как их слишком много, сделаем выборку самых больших штатов, наиболее отдаленных друг от друга. Узнаем, в каком штате жители больше всего любят авокадо!

Как мы видим, в Калифорнии жители наиболее увлечены здоровым образом жизни и

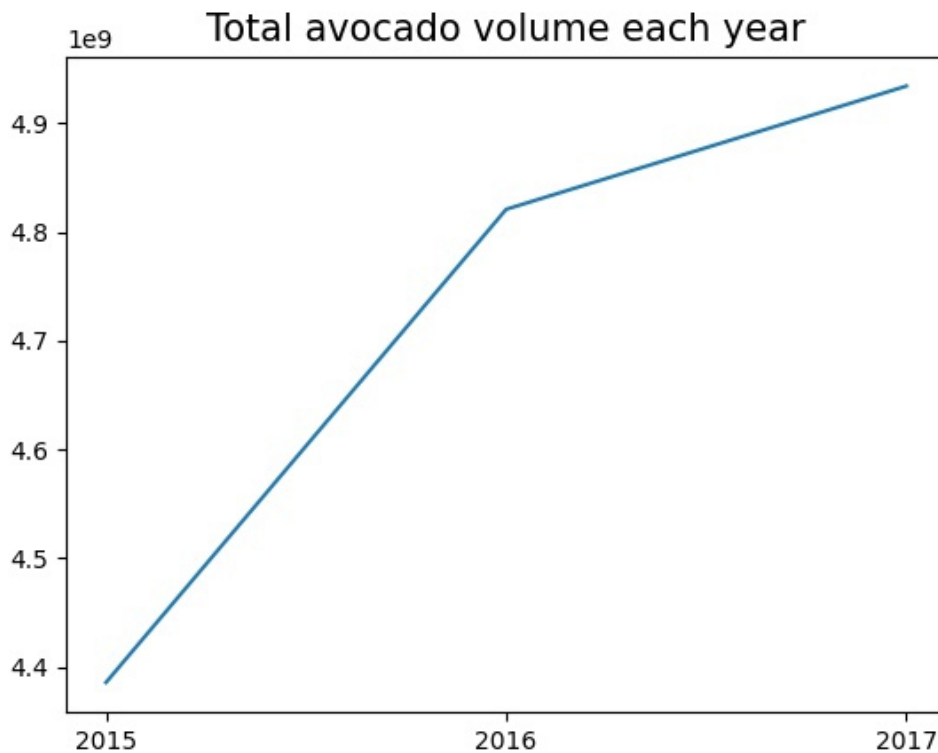


Рис. 4: Точное изменение потребления авокадо

правильным питанием, а наименее всего в Детройте. На втором месте по покупкам Лос-Анджелес, а Нью-Йорк только лишь на третьем.

8. Предположим, что авокадо в упаковках покупают магазины и рестораны, поэтому их продажи не должны значительно изменяться за несколько лет.

Предположение оказалось неверным – авокадо покупают упаковками с каждым годом только больше.

9. Проанализируем по датасету, будет ли зависеть цена на авокадо от штата. Для анализа возьмем те же штаты, что и в п. 7

Как можно заметить, цена может разниться максимум на 50 центов, и самое дорогое авокадо как раз в Нью-Йорке. Можно предположить, что поэтому несмотря на количество жителей, количество покупок в нем только на 3 месте.

10. Вернемся к видам авокадо, обозначенным номерами. Посмотрим, будет ли изменяться популярность какого-либо вида из года в год.

Гипотеза подтвердилась – за 3 год популярность вида 4225 стабильно понижается, а вида 4046 наоборот резко выросла за последний год. Вид 4770 все время пользуется наименьшей популярностью.

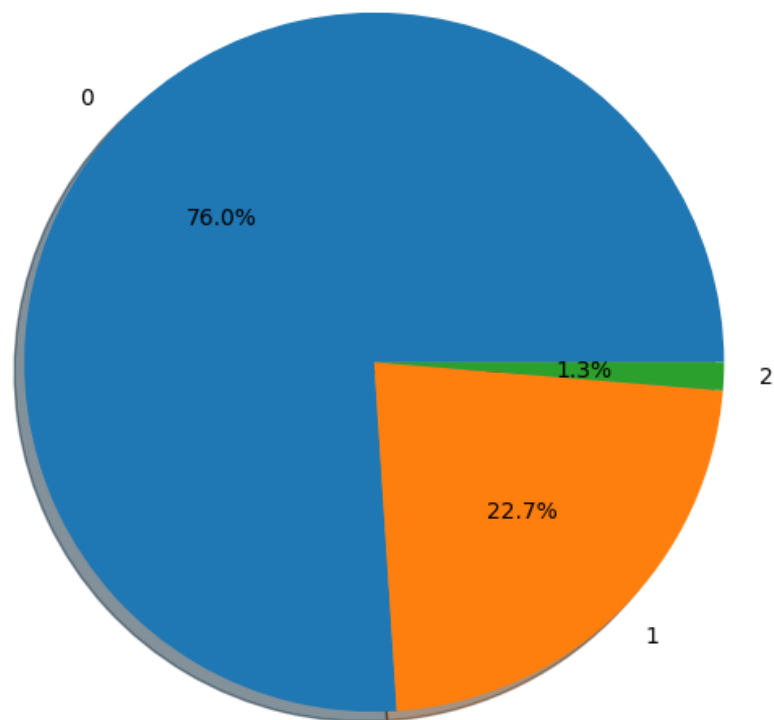


Рис. 5: Сравнение популярности упаковок авокадо

Вывод: Большинство гипотез, предположенных по датасету, подтвердились. Авокадо действительно пользуется популярностью среди жителей США, и эта популярность растет в течение нескольких лет.

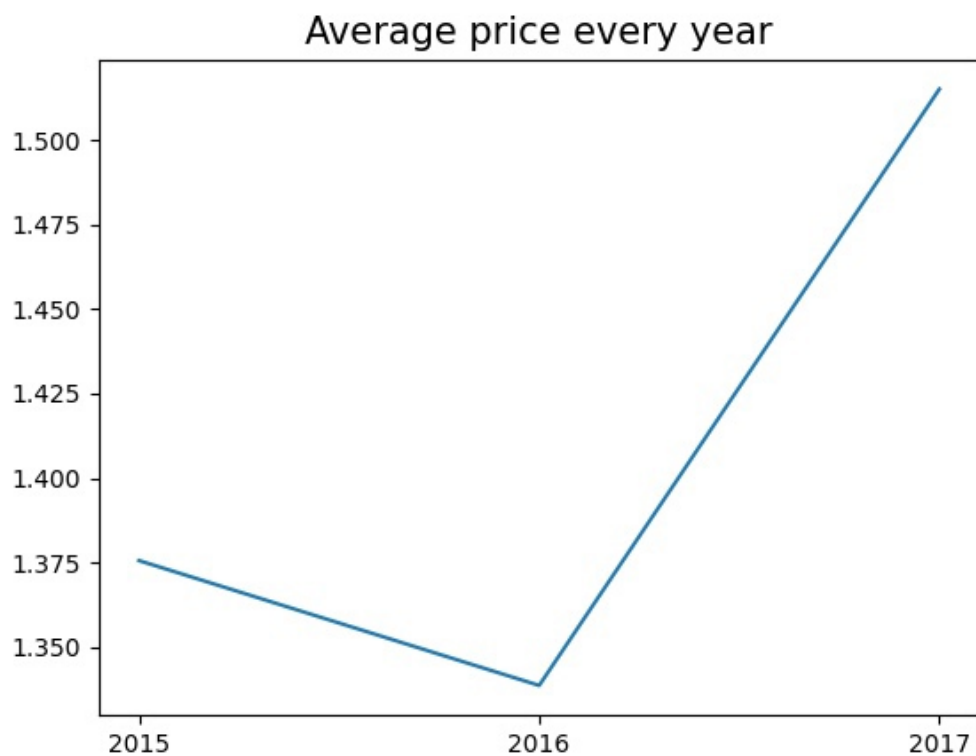


Рис. 6: Изменение цены на авокадо за несколько лет

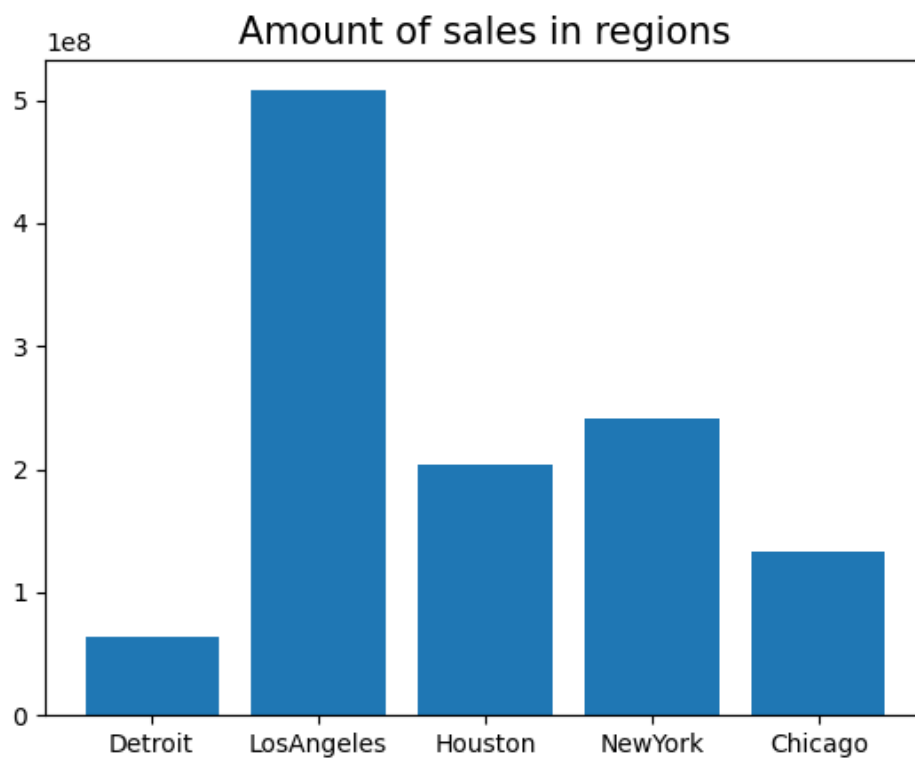


Рис. 7: Популярность авокадо в разных штатах

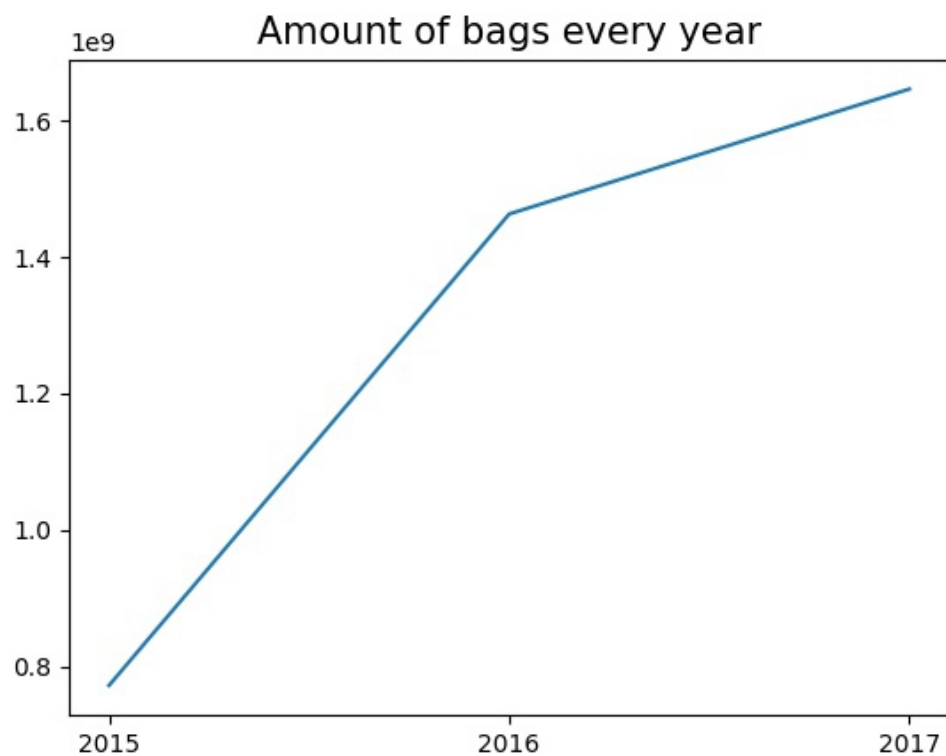


Рис. 8: Покупки упаковок авокадо

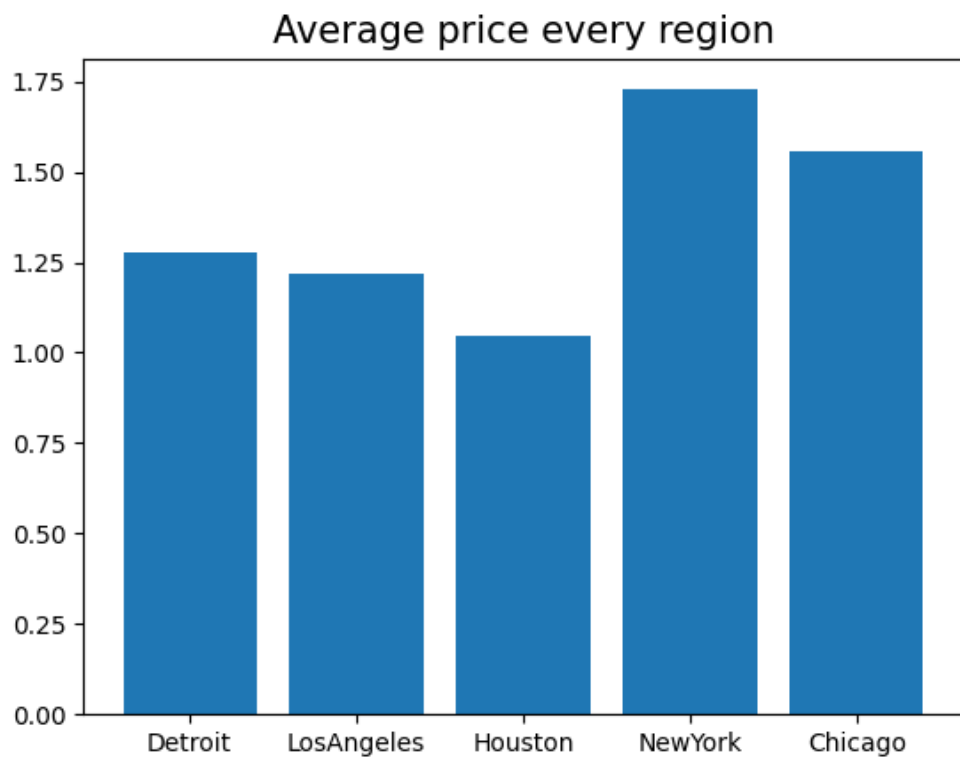


Рис. 9: Зависимость цены авокадо от штата

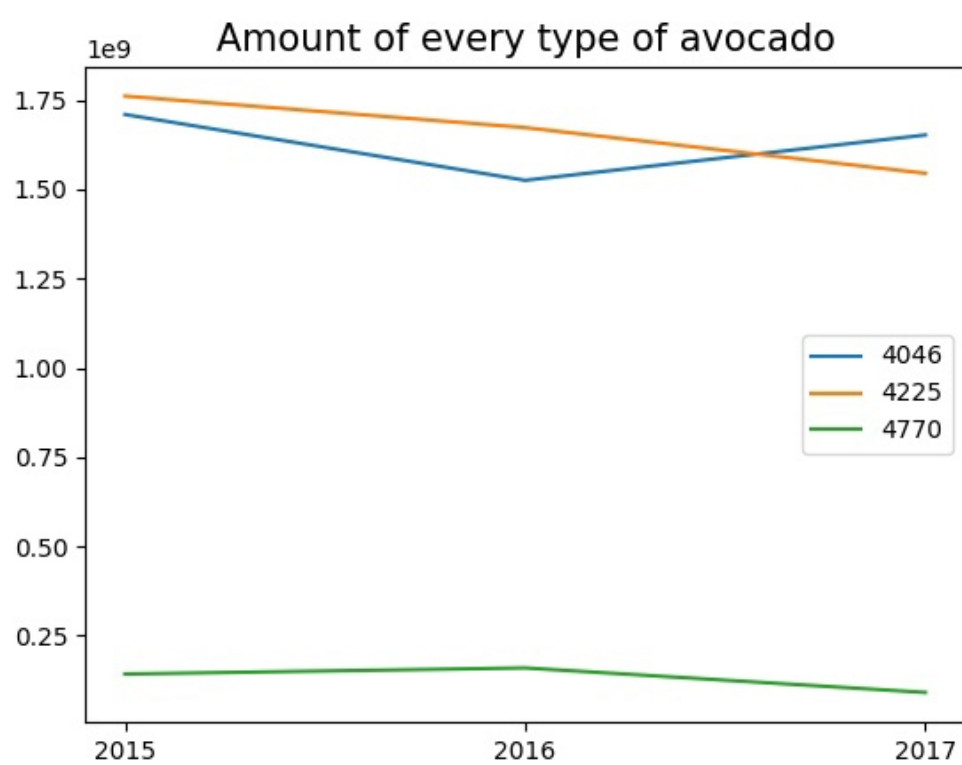


Рис. 10: Популярность видов авокадо