Avocados Galore!

Avocado Demand in the United States

By
Radhika Satapathy, Timothy Quek and Marcial Nava

This report was prepared for American Fruit Distributors, Inc.
December, 2018

Introduction

Avocados have conquered the American palate. Avocados are consumed in virtually every corner of the country. They have become increasingly common in restaurants, food chains and supermarkets. A series of factors explain the increasing popularity of avocados in the country. First, since 1990 there have been no barriers to the imports of avocados coming from Mexico (the world's top producer). Second, avocados are largely consumed by the 40 million Hispanics of Mexican descent living in the United States. Third, a series of studies have highlighted the positive impact that avocados could have on human health. The taste and consistency of avocados makes them a great ingredient for numerous recipes. In this report, we explore trends in demand of avocados by region. Although there are multiple varieties of avocado, Hass avocados account for 95% of avocado sales in the United States. This is due to their thick skin and the fact that they ripen slowly, attributes that make them perfect for supermarket shelves. Our goal is to identify the regional markets that offer the best opportunities for the distribution of avocados.

Data & Data Sources

Avocado Sales Data: The team analyzed a database containing information generated by the Avocado Hass Board (AHB), a non-profit organization dedicated to promote the consumption of Hass Avocados in the United States. The dataset was obtained from Kaggle.com under the name of Avocado Prices: Historical Data on Avocado Prices and Sales Volume in Multiple U.S. Markets. It was built and maintained by Justin Kiggins. The dataset contains weekly records, running from the week of January 4, 2015 to the week of March, 25, 2018. The dataset comprises 13 columns and 18249 rows. The table below shows the variables contained in the dataset and their descriptions.

Variables contained in the Avocado Price Database

Variable	Description
Date	The date of the observation (week ending)
AveragePrice	The average price of a single avocado
type	Conventional or organic
year	The year
region	The city or region of the observation
Total Volume	Total number of avocados sold
4046	Total number of avocados with PLU 4046 sold (small)
4225	Total number of avocados with PLU 4225 sold (large)
4770	Total number of avocados with PLU 4770 sol (extra large)
Total Bags	Total number of avocados sold in bags
Small Bags	Total number of avocados sold in small bags
Large Bags	Total number of avocados sold in large bags
XLarge Bags	Total number of avocados sold in extra large bags

A distinctive feature of the avocado dataset is the specification of data based on the BEA (Bureau of Economic Analysis) regions as shown in the table below.

Population Data: Population data was also obtained for the same regions included in the avocado dataset. The population dataset was provided by Haver Analytics, a time series data provider that includes, among other series, population data from the Census Bureau. Population data was aggregated to match with the geographies included in the avocado dataset.

Geographies contained in the Avocado Price Database

CALIFORNIA Los Angeles Sacramento San Diego San Francisco	**WEST** Denver Phoenix Portland Salt Lake City Seattle West Texas/New Mexico	**PLAINS** St. Louis	**SOUTH CENTRAL** Dallas Houston New Orleans
GREAT LAKES Chicago Cincinnati Columbus Detroit Grand Rapids Indianapolis	**MIDSOUTH** Baltimore Charlotte Louisville Nashville Raleigh Richmond Roanoke	**SOUTHEAST** Atlanta Jacksonville Miami Orlando South Carolina Tampa	**NORTHEAST** Albany Boston Buffalo Harrisburg/Scranton Hartford/Springfield New England New York Philadelphia Pittsburgh Syracuse

Data Cleansing and Sanity Checks

Missing values: The two most important variables for our analysis, total volume and average price do not have missing values.

Consistency: Particular attention was paid to variables that express volume such as volume by size (“4046”: small, “4225”: large, “4770”: extralarge), volume by packaging type (“TotalBags”), and total volume (“TotalVolume”). Two issues emerged from the sanity check.

- Figures were not in whole numbers despite being expressed in units (avocados). This is because units were calculated from dividing sales by unit price. We corrected this by applying a function to round the data, specifically total volume. Variables indicating size and form of package were left unrounded since they are not relevant for the research question.
- For 209 rows, total volume was not equal to the sum of volumes by size (“4046”, “4225”, “4770”) and volumes by packaged (“TotalBags”. This was assumed to be due to an error in either data collection or

data entry. We therefore recalculated the “TotalVolume” column by adding the “4046”, “4225”, “4770” and “TotalBags” columns and rounded these volumes to the nearest whole number.

- A significant number of zeroes was noted in the data for ‘SmallBags’, ‘LargeBags’ and ‘XLargeBags’ variables. It was unclear whether these actually represented the fact that there were no bags of that size sold or whether this was missing data. Since this data to be unimportant to our subsequent analysis, it was not analyzed.
- It was also noted that for the year 2018, there was only 3 months worth of sales data. This subset would not have been useful to us for yearly comparisons. Although we considered extrapolating from previous months for the entire year, seasonality in demand prevented us from obtaining an accurate picture. Hence we decided to disregard 2018 for the purpose of our analysis of annual sales.
- Another caveat we had to deal with was making sure that we do not consider city level data with regional level data in the same plot. We had to make sure to extract region specific rows and city specific rows into different data frames, or at least filter the data prior to analysis.

Data Wrangling

Merging with Population Data

It was noted that it would be challenging to compare avocado sales volume data without some kind of normalization by population size. Hence, population data was obtained for each city and for the entire country (‘TotalUS’) and merged with the sales data. (Refer to see above for details of the population dataset).

Subsequently, we realized that it was challenging to obtain regional level population data on the sales regions. Hence, comparisons between different parts of the country were limited to comparisons between cities. The population data was merged with the avocado sales data by city (using the ‘region’ variable) and by year.

Creation of New Variables

The following variables were created after the merge:

(A) Total Sales Revenue (‘TotalSales’) by multiplying the total sales volume (‘TotalVolume’) for each row by the ‘AveragePrice’
(B) Volume Per Capita (‘Vol_Per_Cap’) by dividing the total sales volume (‘TotalVolume’) for each row by the population (in thousands)
(C) Sales Per Capita (‘Sales_Per_Cap’) by dividing the total sales revenue (‘TotalSales’) for each row by the population (in thousands)

Insight from Data Analysis

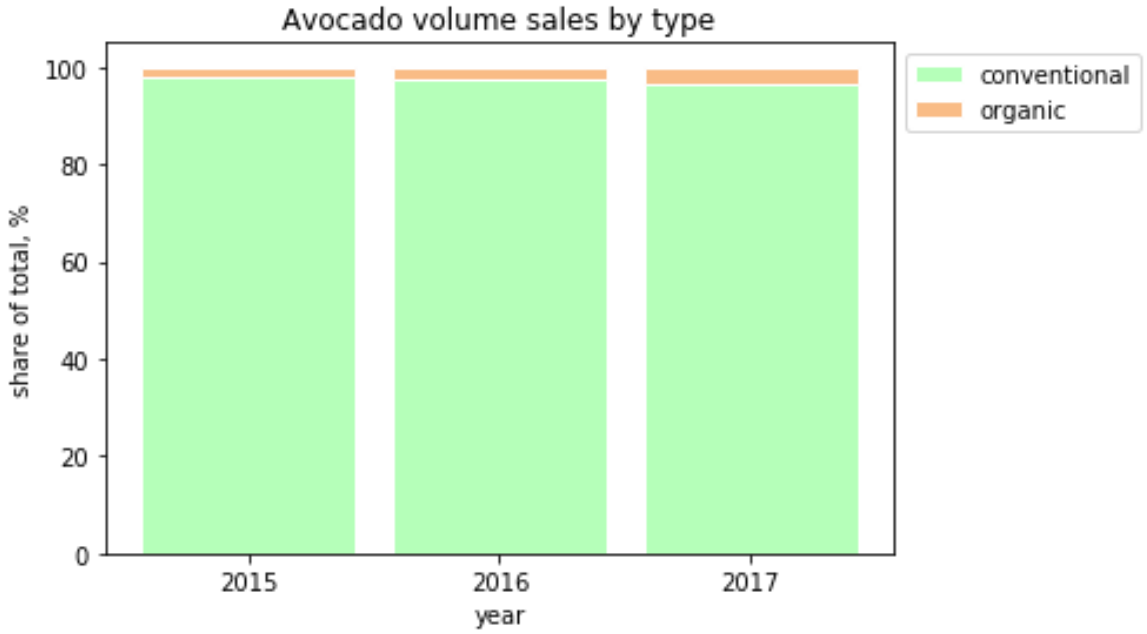
A. Nationwide Avocado Consumption by Avocado Type

Based on nationwide data, total avocado consumption has increased from 1.66 billion avocados per year in 2015 to 1.86 billion avocados a year in 2017, of which between 2-4% are organic avocados. Although avocado consumption continues to grow between 2015 and 2017, what is striking is the growth of organic avocado consumption. Between 2015 and 2016, organic avocado consumption grew 45.7%, and this continued to grow another 28.7% between 2016 and 2017.

The table and graph here show the nationwide conventional and organic avocado consumption between 2015 and 2018. Data for 2018 is only up till Mar 2018.

Nationwide Avocado Consumption (Conventional vs. Organic)

Year	Conventional Avocado Volume (% growth from previous year)	Organic Avocado Volume (% growth from previous year)	Total Avocado Volume (% growth from previous year)	% Organic Avocados
2015	1,623,685,890	33,569,270	1,657,255,160	2.03
2016	1,770,259,362 (+9.0%)	48,899,727 (+45.7%)	1,819,159,089 (+9.8%)	2.69
2017	1,801,769,849 (+1.8%)	62,912,327 (+28.7%)	1,864,682,176 (+2.5%)	3.37
2018 (up to Mar)	505,506,395	18,120,774	523,627,169	3.46



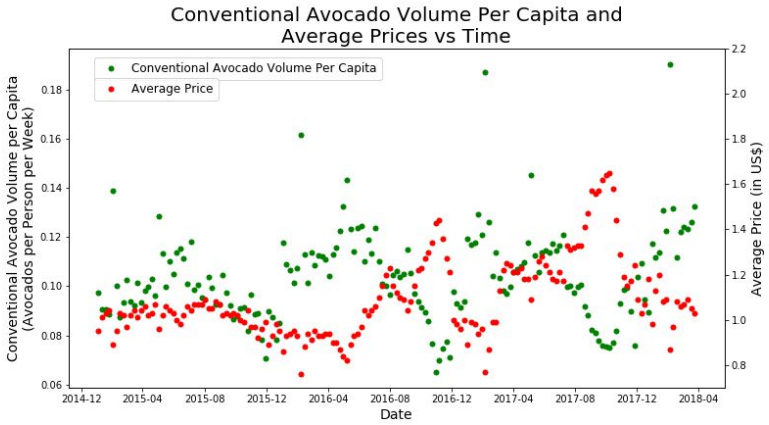
B. Trend in Nationwide Consumption for Avocados over Time (Conventional and Organic)

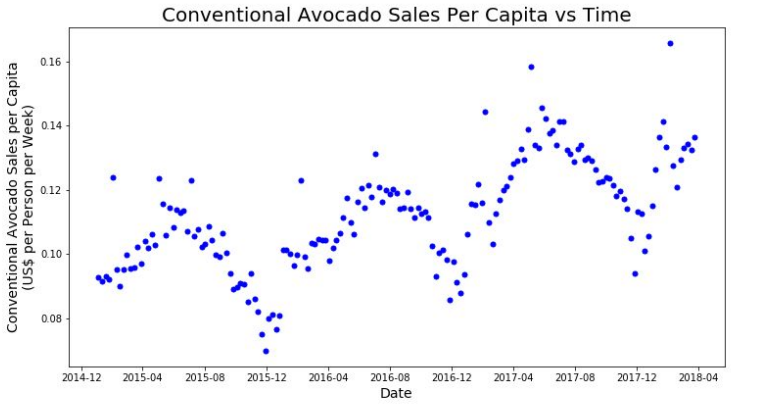
The graphs on the right show the trend in nationwide volume, prices and sales per capita for conventional and organic avocados over time.

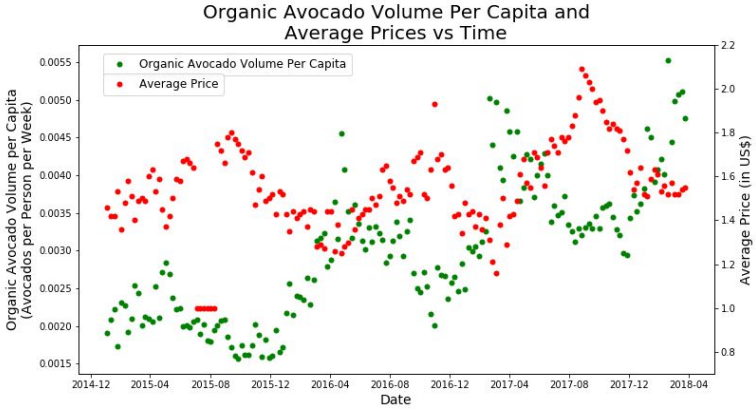
Clear seasonality is noted in the consumption of both conventional and organic avocados over time, with the volume of avocados consumed increasing in the summer (the middle of the year) and decreasing in the winter. The price of avocados shows the opposite trend, decreasing in the summer and increasing in the winter.

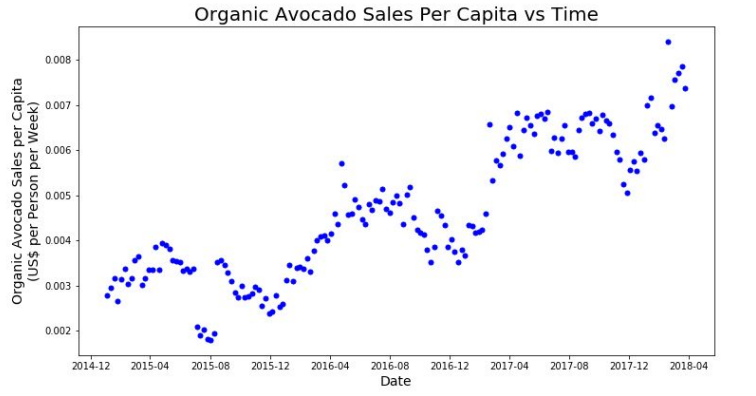
These trends can be explained by the fact that the source of avocados in the USA is largely from 2 sources, locally grown in California, as well as imported from Mexico[2]. These sources provide a large supply of avocados in summer, which explains the larger volumes and low prices. In the winter however, avocados have to be imported from countries in the southern hemisphere such as Chile and New Zealand[3,4]. The supply of these avocados is limited, resulting in higher prices and lower volumes.

Another phenomenon that can be seen is the significant volume and sales growth for organic avocados over time. Despite the lower volumes, the significant growth in demand provides a great opportunity for an avocado investor or grower to enter the market.



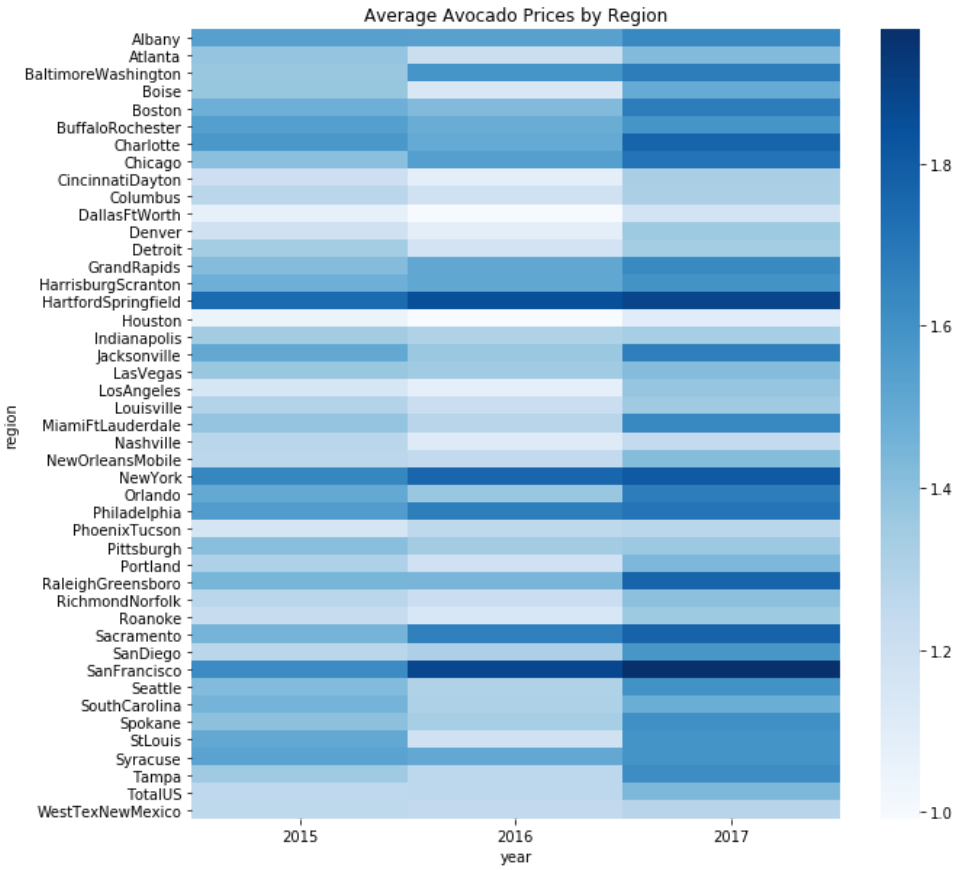




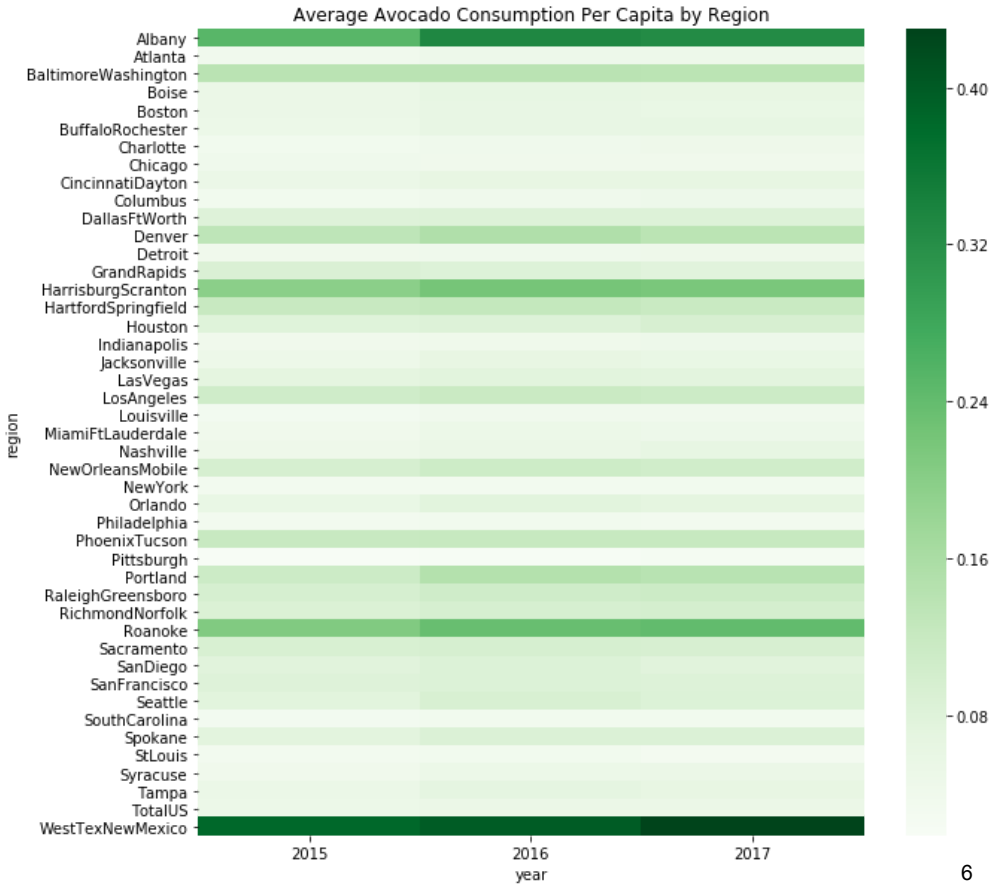


C. Trends in Avocado Prices and Sales per Capita across Cities

A preliminary inspection of prices by metropolitan area reveals significant disparities across regions. There are regions in which prices are closer to $2 per unit, and regions where prices are closer to $1 per unit. Two extreme examples are illustrated by San Francisco and Houston. The former show some of the highest prices in the sample, while the latter shows some of the lowest prices. When we look at the evolution of prices over the years, the heatmap shows that prices increased for the most part between 2016 and 2017.

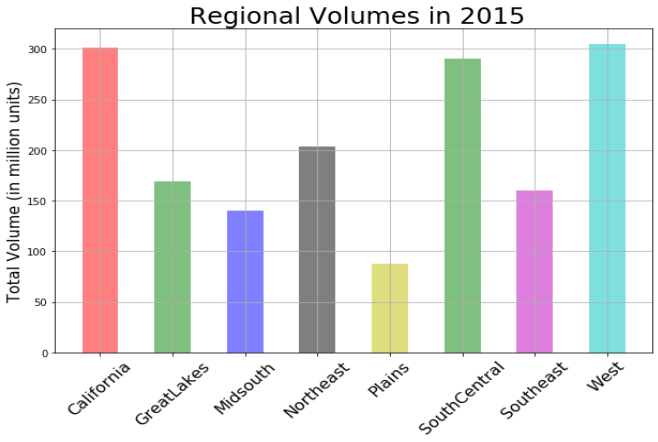


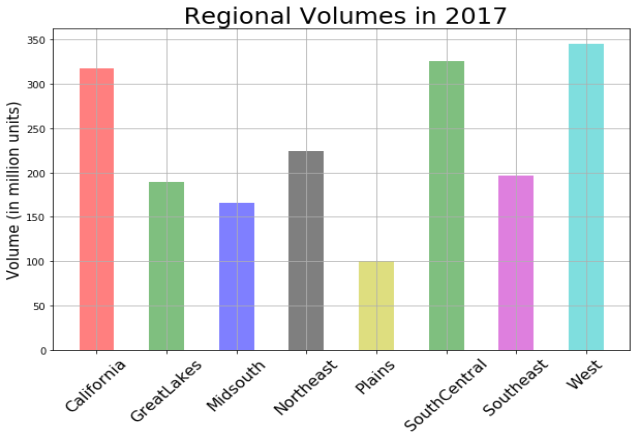
A heatmap of sales per capita shows a different pattern than prices. Sales per capita tend to be very stable across cities, with only a few exceptions where they appear to have increased from one year to the other. That is the case of Albany, Denver, Portland, and Roanoke, among others.

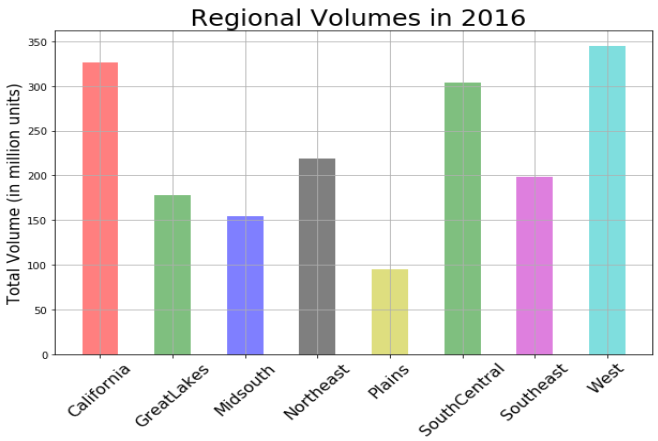


D. Best Market for Avocados

California, South Central and West regions were found to be the biggest markets for avocados in terms of volume, in excess of 300 million avocados sold per year.

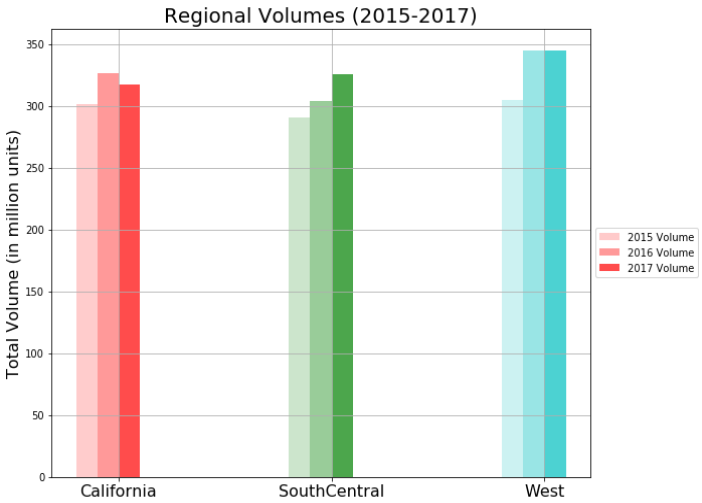


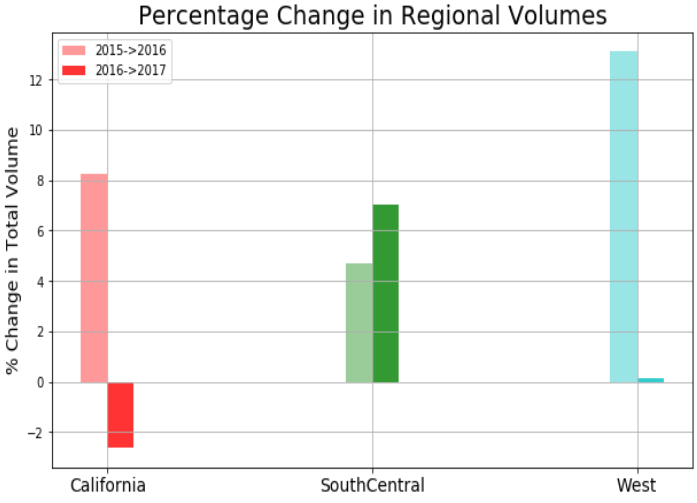




Examining volumes of these top 3 markets across the time period revealed that although there was a significant increase in total sales from 2015 to 2016, the data for 2016 to 2017 was quite different:

California - market slumped, showing a decline of more than 2%
South Central - market continued to show strong growth of 7%
West - market seemed to flatten out





In order to gain a deeper understanding of the situation, we explored sales data for the constituent cities in each of these regions.

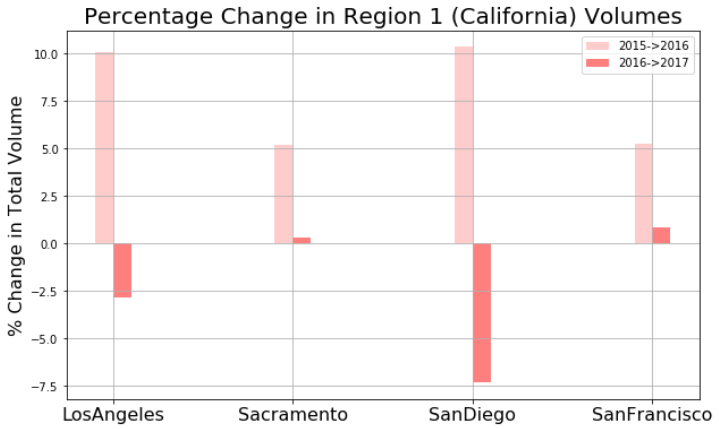
City level data analysis revealed that presence of strong markets in South Central. Houston with a 6% increase over 2015-16 and almost an 18% increase in 2016-17 seemed to be the main driver in the region. Houston was perhaps the only city across the 11 cities examined that showed incredible sales growth.

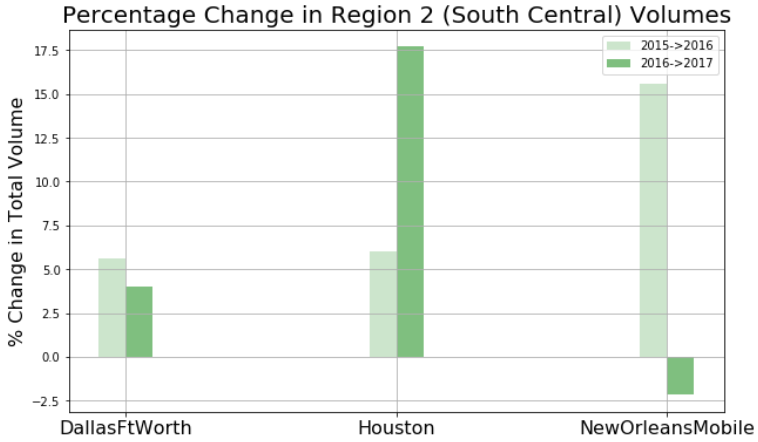
Clearly from 2015 to 2016, all these cities saw a significant increase in avocado sales, so what may be the reason that some cities were experiencing growth, despite a slowdown, whereas some others were declining? Given the proximity to Mexico as well as the high percentage of hispanic population, it was surprising to note sales plummeting in LA and San Diego.

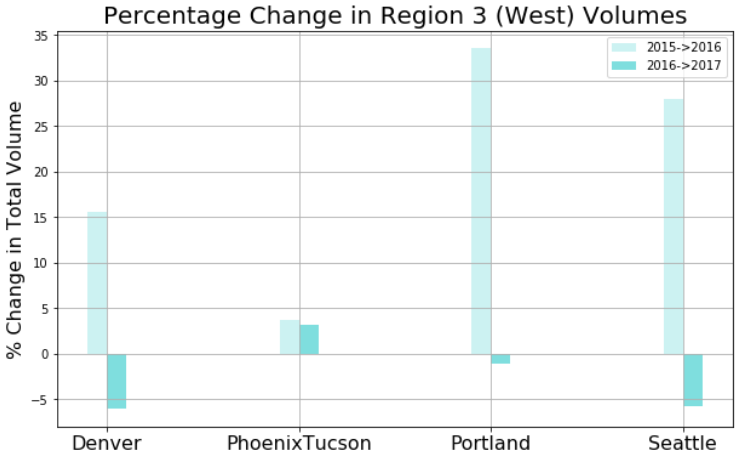
Therefore we followed through with a comparative assessment of average avocado prices in these markets over time and split cities up into two categories.

Growing markets were defined as cities which showed a positive percentage increase in sales from 2015 to 2016 as well as from 2016 to 2017.

Slowing markets were defined as cities that noted a positive percentage growth in sales from 2015 to 2016, but which showed a reverse trend in the subsequent time period.







<u>Growing Markets:</u>

- Houston
- Dallas
- Phoenix-Tucson
- San Francisco
- Sacramento

<u>Slowing Markets:</u>

- Los Angeles
- San Diego
- Seattle
- Denver
- New Orleans - Mobile

And what we found was extremely interesting. While all of the cities displaying positive growth in 2017 managed to keep the price increases to under the national average of 1.2%, almost all of the slumping markets exhibited almost double that increase in average prices. It's no wonder that avocadoes did not sell as much in these regions in 2017.

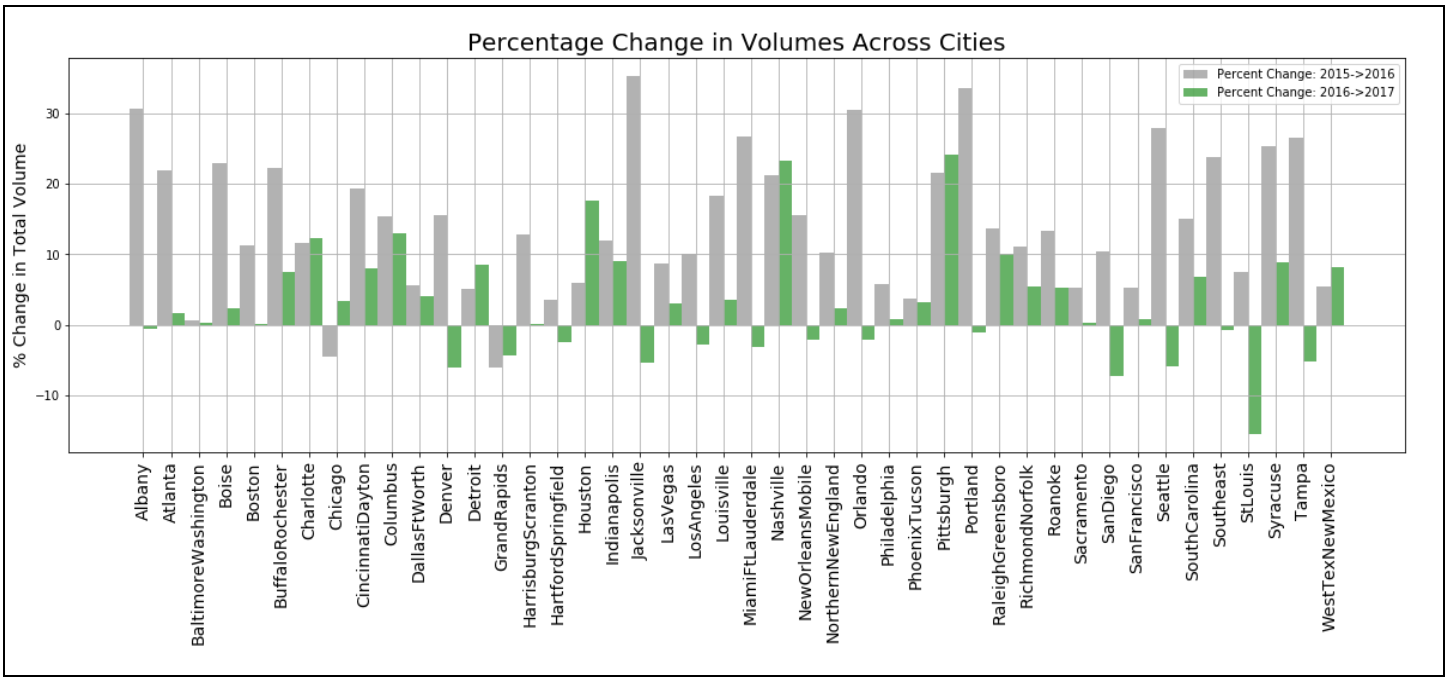
Trends in Average Unit Price of Avocados

GROWING MARKETS	2015	2016	2017	Share of organic avocado sales in 2017
Dallas	1.08	1.08 (+0.0%)	1.18 (+9.3%)	1.9%
Houston	1.05	0.99 (-5.7%)	1.10 (+11.1%)	2.6%
Phoenix-Tucson	1.16	1.26 (+8.6%)	1.27 (+0.8%)	1.2%
Sacramento	1.45	1.67 (+15.7%)	1.77 (+6.0%)	1.6%
San Francisco	1.42	1.88 (+32.4%)	1.97 (+4.8%)	2.6%
Total US	1.26	1.26 (+0.0%)	1.44 (+12.5%)	3.5%

SLOWING MARKETS	2015	2016	2017	Share of organic avocado sales in 2017
Denver	1.19	1.09 (-9.7%)	1.36 (+19.9%)	4.6%
Los Angeles	1.15	1.08 (-6.5%)	1.38 (+21.7%)	3.2%
New Orleans - Mobile	1.27	1.25 (-1.6%)	1.42 (+12.0%)	1.6%
San Diego	1.27	1.31 (+3.1%)	1.58 (+17.1%)	3.0%
Seattle	1.42	1.30 (-9.2%)	1.60 (+18.8%)	7.6%

As seen in the plot below, avocado sales seemed to have spiraled down in 2017 in almost 50% of the US markets. One could attribute it to fluctuations in average prices, however further analysis is warranted before we can make a strong conclusion.

Percentage Change in Avocado Sales Across Cities



Final Recommendation

Therefore, considering the analysis of avocado sales for the past 3 years, we would recommend the following:

1. The South Central region with 2 strong city markets presents itself to be the best region for market entry.
2. Despite the growing demand, the demand of avocados is quite price sensitive. Price setting would be a critical step and must be accomplished keeping in view that it might be optimal to maintain the price at or below the national average for a greater chance of success.
3. The growth of organic avocado consumption poses an opportunity despite them comprising only a small proportion of market share today.

Next Steps

The following types of analyses are recommended for the future.

1. Weekly price variations over time
2. Percentage increase in share of organic avocados versus price of conventional avocados
3. Proximity of cities versus avocado prices
4. Per capita consumption rates of organic avocados by city

References

1. https://en.wikipedia.org/wiki/IRI_(company)
2. https://en.wikipedia.org/wiki/Avocado
3. https://www.bonappetit.com/test-kitchen/ingredients/article/why-avocados-always-in-season
4. https://fruitsfromchile.com/fruit/avocados/