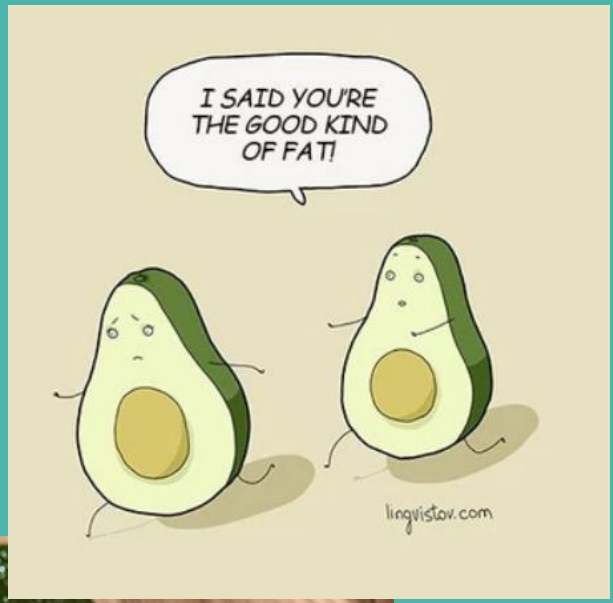
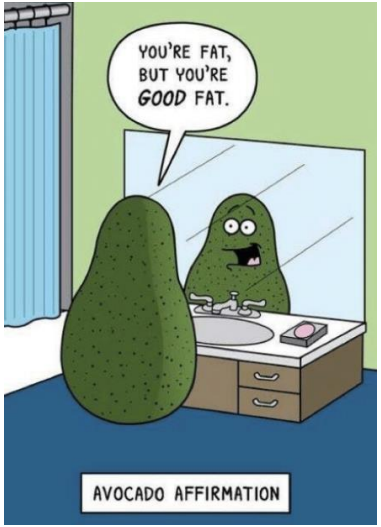
Analysis of Avocado Prices

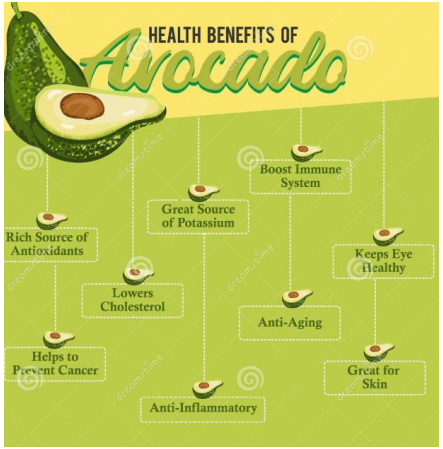
Nikhar Jain
nikhar.jain@utexas.edu
December 2019

Table of Contents

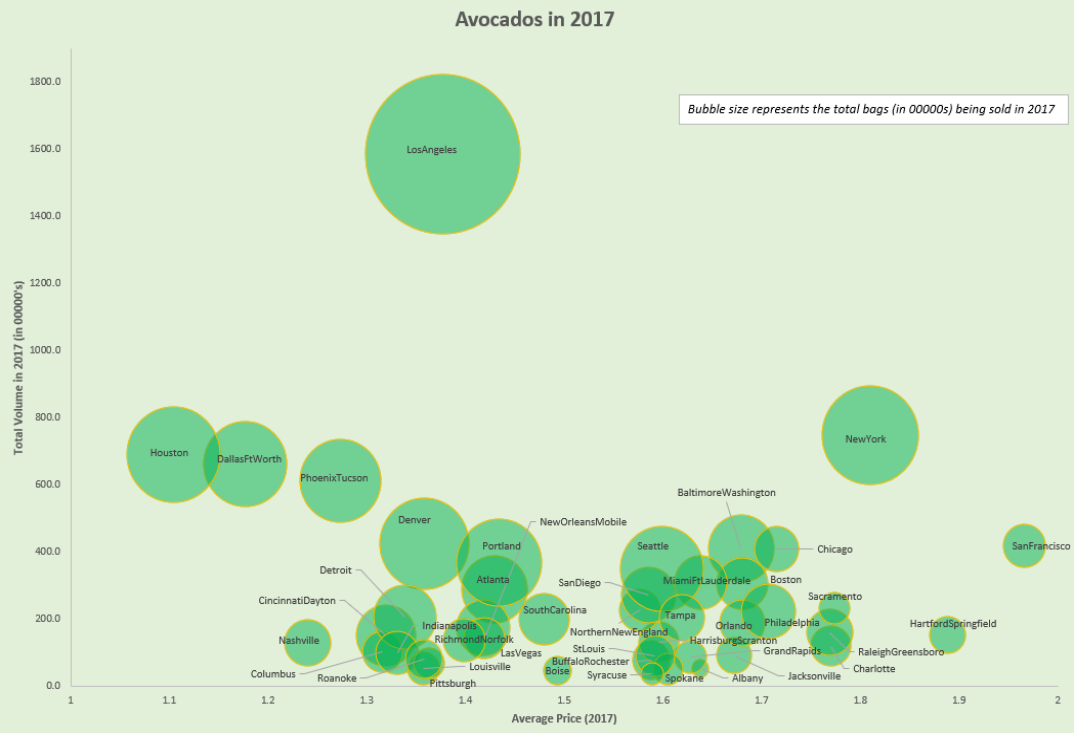


Why Avocados?





Summary



- Over 2015-2018, the average price of avocados has been rising due to its increasing popularity
- Conventional avocados dominate the overall avocado market; However, organic avocados are gaining traction
- In 2017, a nation-wide shortage of avocados led to an increase in the average prices
- Avocado sales are lower in winter months, typically due to its harvesting season
- Cities in the northeast and northwest regions have higher avocado prices

Notes: In the chart above, x-axis for average prices has been deliberately made to start with 1 instead of zero, as the chart was not giving a clear picture - this shows that regions have varied prices between $1-$2; The above chart does not include data for California, Great Lakes and Plains as these are regions and cannot be compared to city-level data; Since 2018 data was available for only few months, I used 2017 to show what was happening to avocados in 2017

The Data

Data Source

kaggle

The dataset includes 15,886 observations with 14 columns for the period 2015 - March 2018

Original Data Source

HASS AVOCADO BOARD

Name	Description
Date	Observation Date
AveragePrice	Average price of each avocado
Type	Type of avocado - conventional or organic
Year	Year in consideration
Region	Observed city or state
Total Volume	Total volume purchased
4046	Total number of avocados with PLU code 4046
4225	Total number of avocados with PLU code 4225
4770	Total number of avocados with PLU code 4770
Total Bags	Total number of bags sold
Small Bags	Total number of small bags sold
Large Bags	Total number of large bags sold
XLarge Bags	Total number of extra-large bags sold

The Data

What does the data look like?

```
data.head(5)
```

	Number	Date	AveragePrice	Total Volume	4046	4225	4770	Total Bags	Small Bags	Large Bags	XLarge Bags	type	year	region
0	0	12/27/2015	1.33	64236.62	1036.74	54454.85	48.16	8696.87	8603.62	93.25	0.0	conventional	2015	Albany
1	1	12/20/2015	1.35	54876.98	674.28	44638.81	58.33	9505.56	9408.07	97.49	0.0	conventional	2015	Albany
2	2	12/13/2015	0.93	118220.22	794.70	109149.67	130.50	8145.35	8042.21	103.14	0.0	conventional	2015	Albany
3	3	12/6/2015	1.08	78992.15	1132.00	71976.41	72.58	5811.16	5677.40	133.76	0.0	conventional	2015	Albany
4	4	11/29/2015	1.28	51039.60	941.48	43838.39	75.78	6183.95	5986.26	197.69	0.0	conventional	2015	Albany

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15886 entries, 0 to 15885
Data columns (total 14 columns):
Number          15886 non-null int64
Date            15886 non-null object
AveragePrice    15886 non-null float64
Total Volume    15886 non-null float64
4046            15886 non-null float64
4225            15886 non-null float64
4770            15886 non-null float64
Total Bags      15886 non-null float64
Small Bags      15886 non-null float64
Large Bags      15886 non-null float64
XLarge Bags     15886 non-null float64
type            15886 non-null object
year            15886 non-null int64
region          15886 non-null object
dtypes: float64(9), int64(2), object(3)
memory usage: 1.7+ MB
```

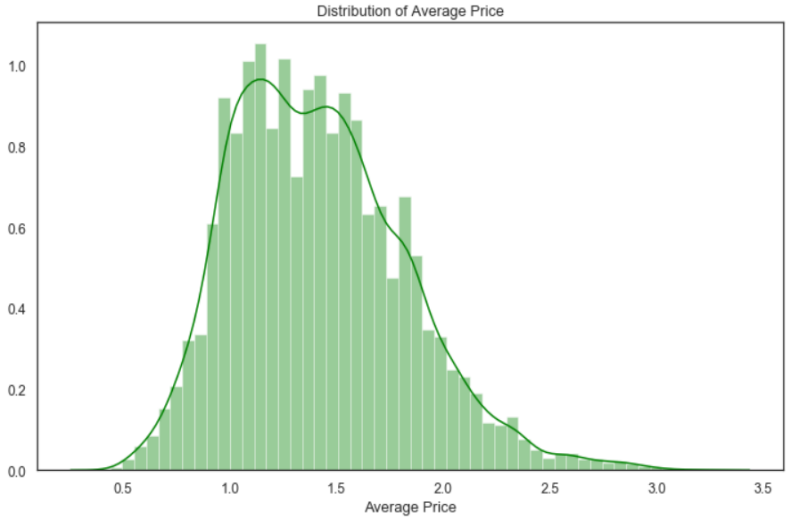
Data Preparation

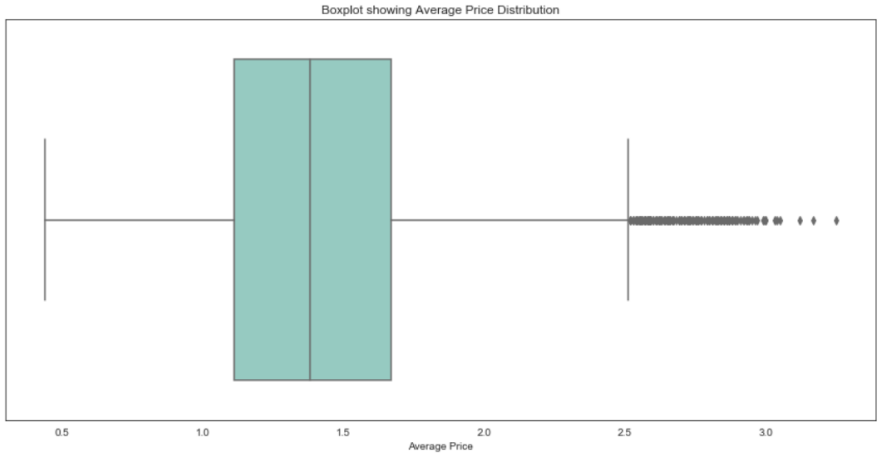
By getting a concise summary of the dataset, that there are no missing values in this dataset

Split the date column to get different columns for day, month and year. By doing so, we see that these observations are weekly – they are recorded for Sunday of every week over 2015-2018

Introduced an additional column of **Total Sales** (Average Price * Volume) to get the dollar value sales

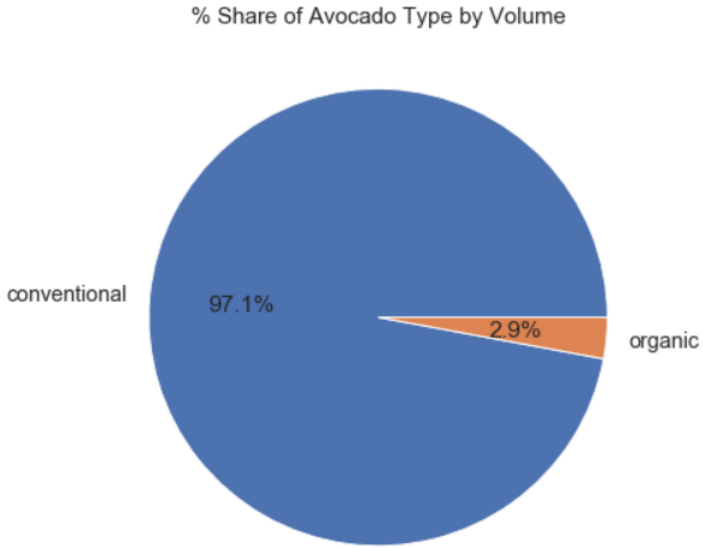
Analysis – Distribution of Average Price





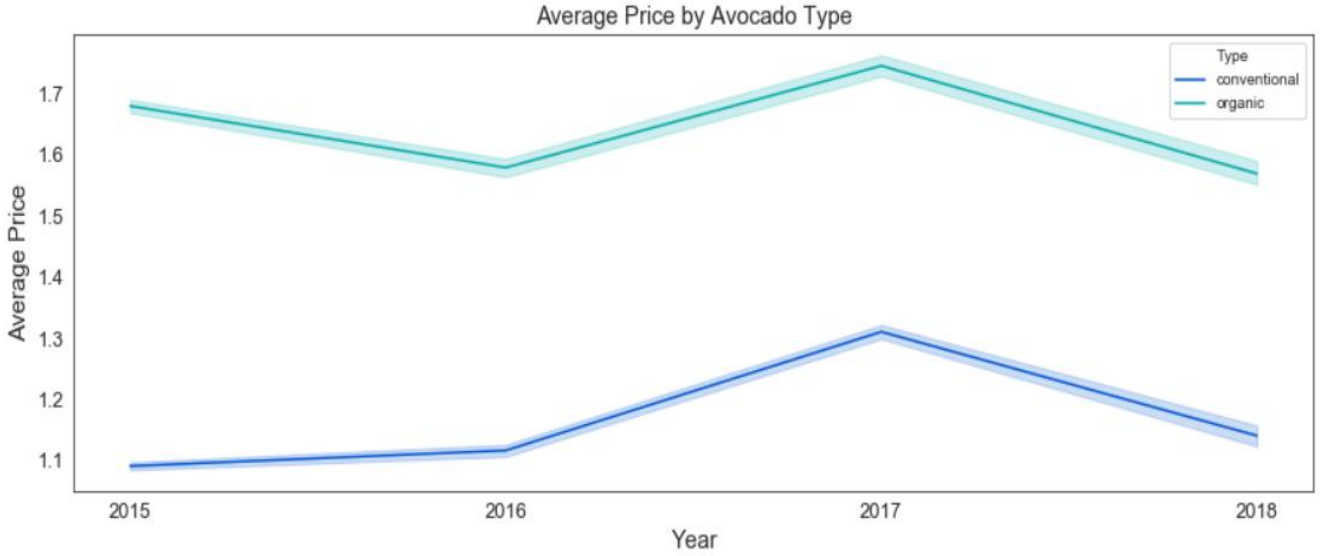
The average avocado price is nearly normally distributed with prices ranging between $1.1-$1.6 per avocado. However, there have been some instances when the price increased to more $3.0. The boxplot shows the spread of these prices

Analysis – Type of Avocados



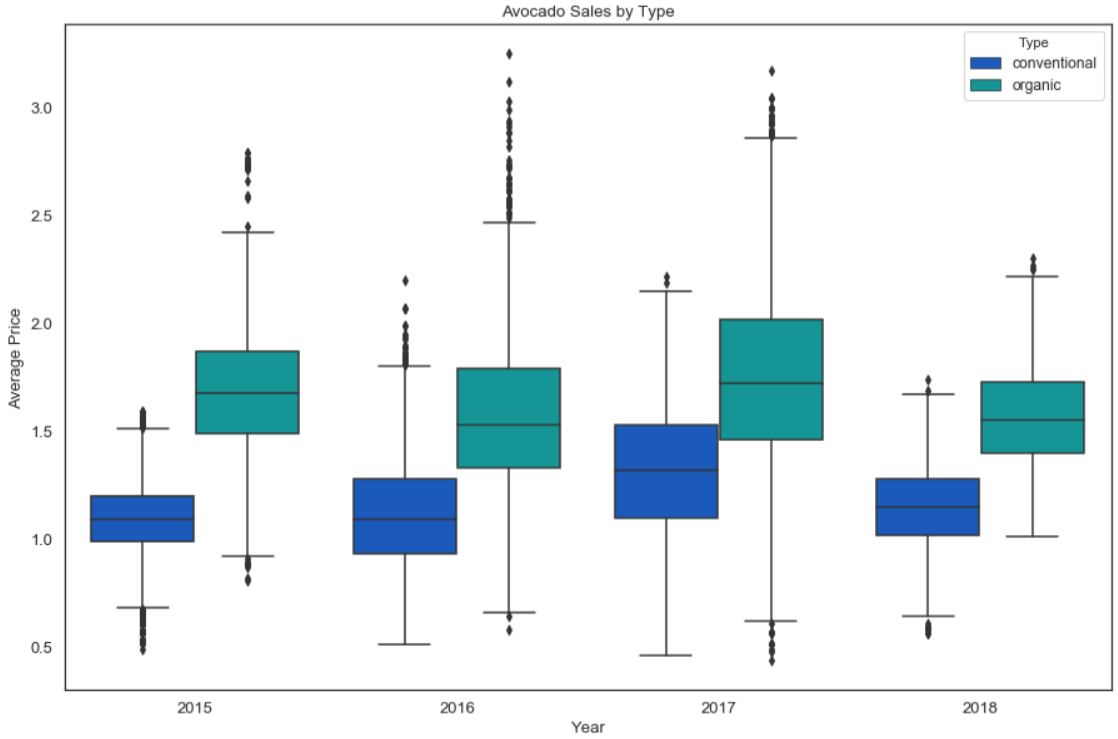
- Conventional avocados dominate the overall avocado sales accounting for ~97% of the total volume
- Organic avocados, which form only ~3% of the volume sales, are slowly gaining traction as organic agriculture has been on a rise since 2015

Analysis – Type of Avocados



- Organic avocados are expensive than the conventional ones
- Irrespective of the type, there was a peak in the avocado prices in 2017 and In 2018, these prices are seen to be reaching their normal levels

Analysis – Type of Avocados



- Price increase in 2017 was due to the surging global demand and reduced harvests from major avocado producing regions including Mexico and California
- Farmers strike in Mexico and a major drought in California in 2016 led to a severe supply crunch, elevating avocado prices in 2017

Analysis – Regional Analysis

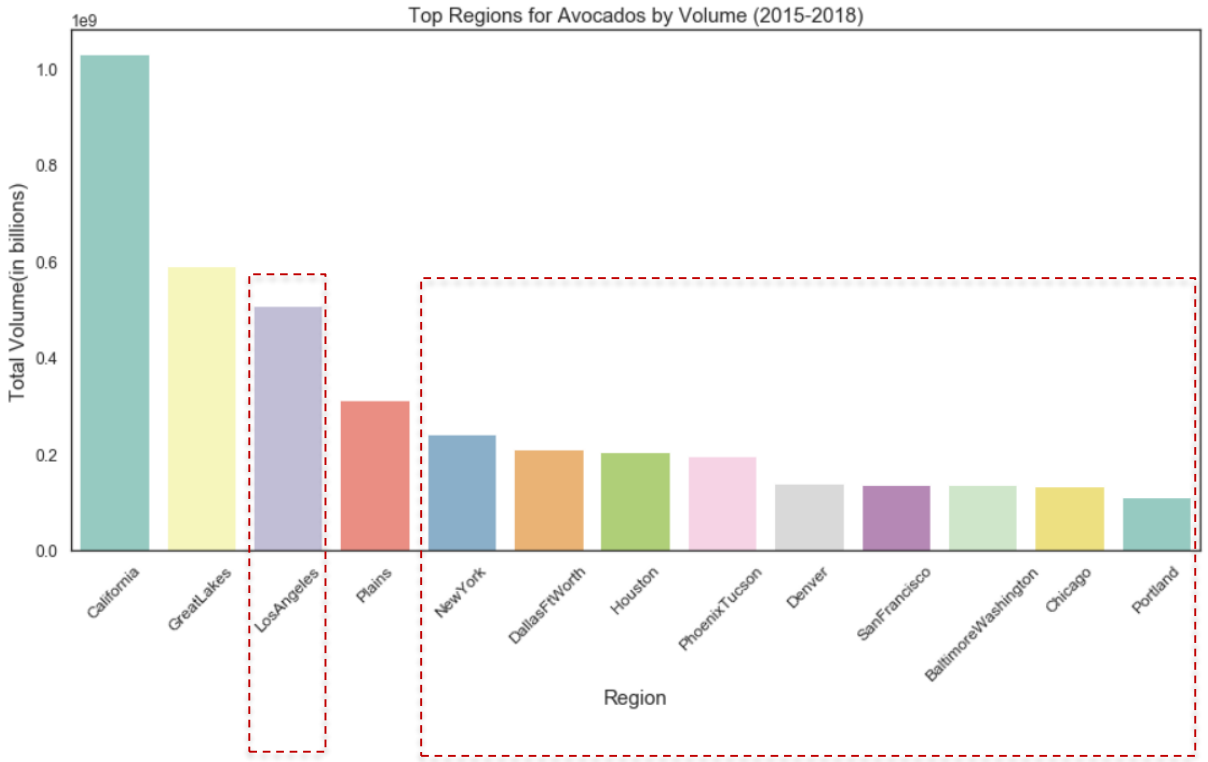
Top 3 Avocado Loving Cities

Los Angeles, CA

New York, NY

Dallas, TX

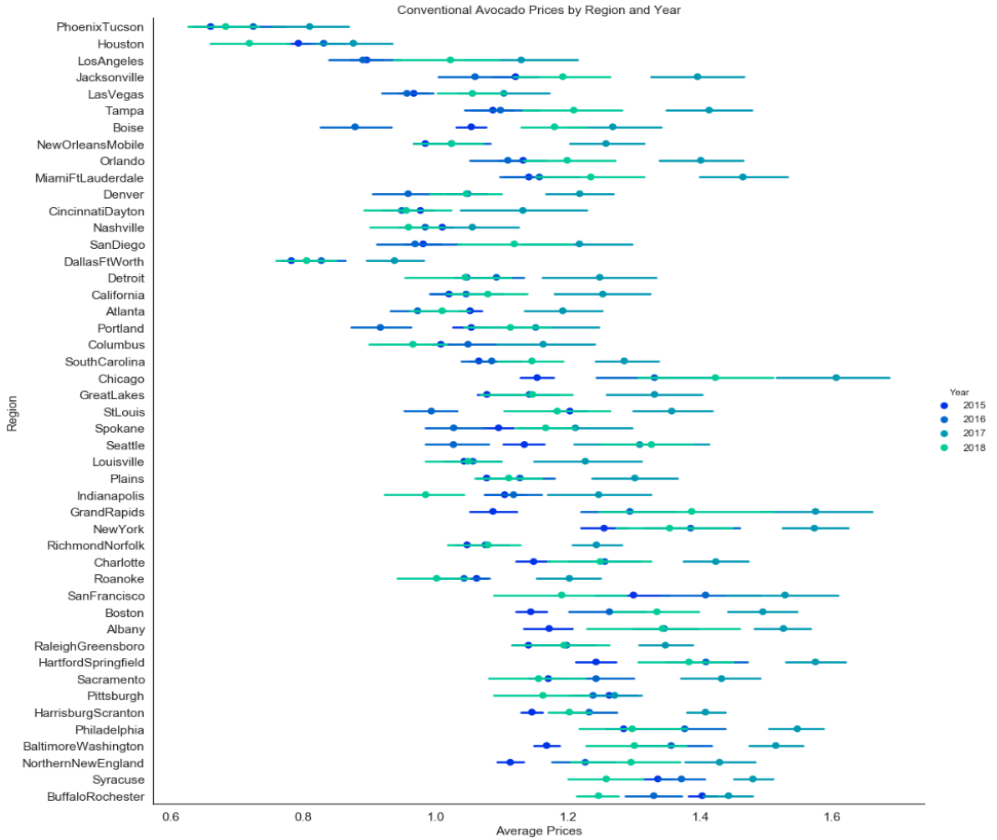
Analysis – Regional Analysis



- California is the leading producer of Hass avocados and home to ~90% of the nation's crop, which makes Los Angeles as the most avocado consuming city

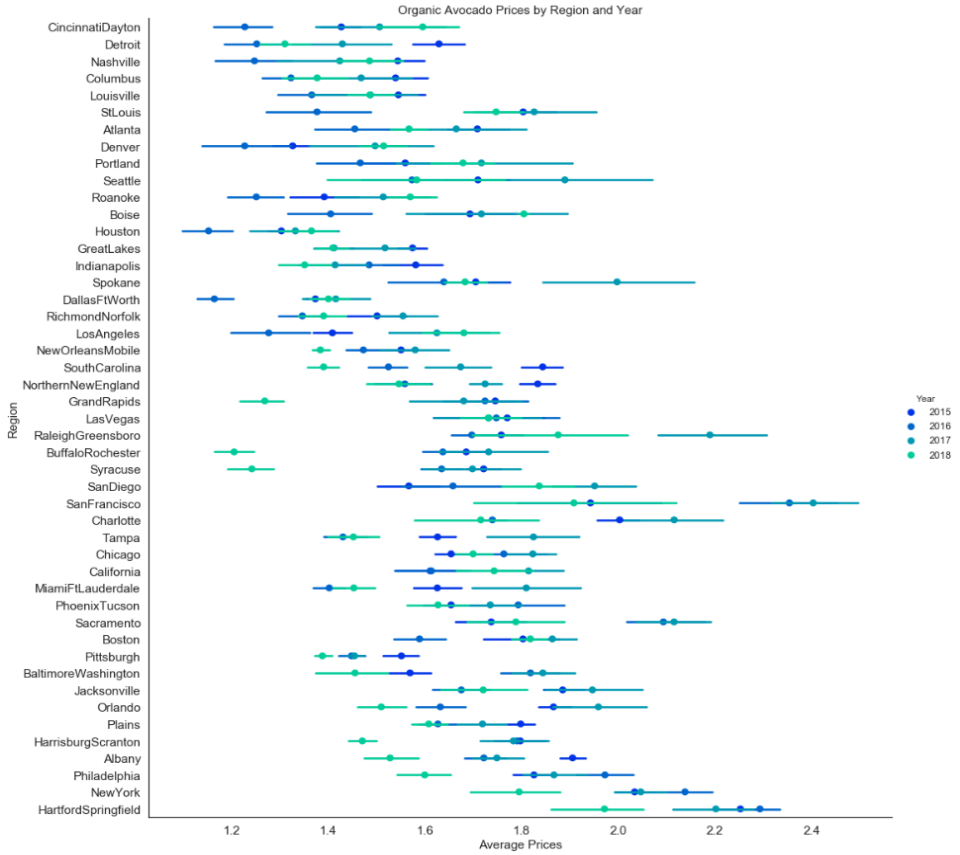
- An interesting thing to note here is that 2 major cities of California – Los Angeles and San Francisco, are among the top 10 cities that love avocados

Analysis – Regional Analysis



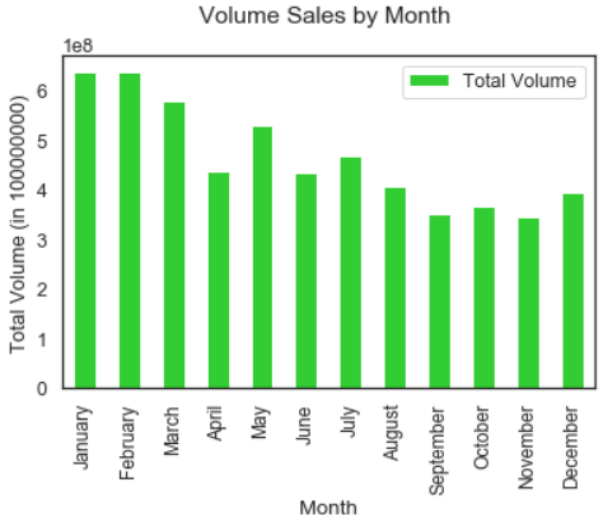
- PhoenixTucson has the cheapest conventional avocados
- BuffaloRochester has the most AveragePrice of conventional avocados
- Grand Rapids region had the most variance in AveragePrice through 2015 to 2016, followed by Chicago

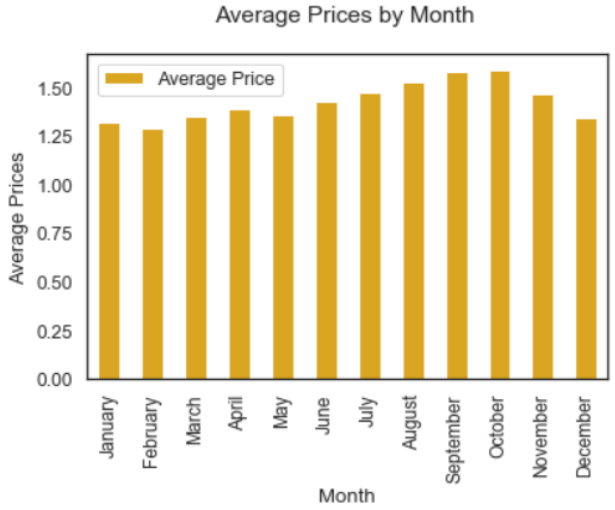
Analysis – Regional Analysis



- DallasFtWorth has the cheapest organic avocados
- San Francisco has the most AveragePrice of organic avocados
- Further, San Francisco also had the most variance in AveragePrice through 2015 to 2018, followed by BaltimoreWashington

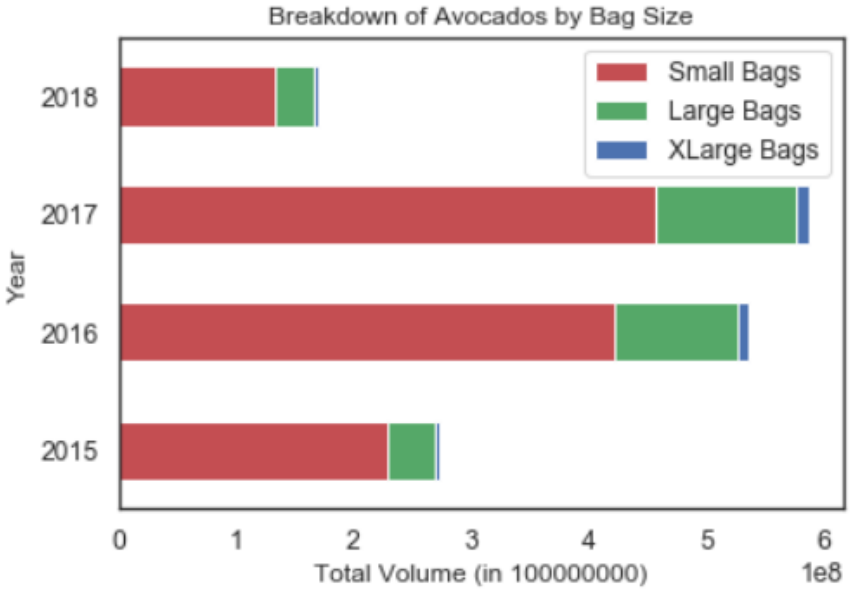
Analysis – Average Price vs Volume





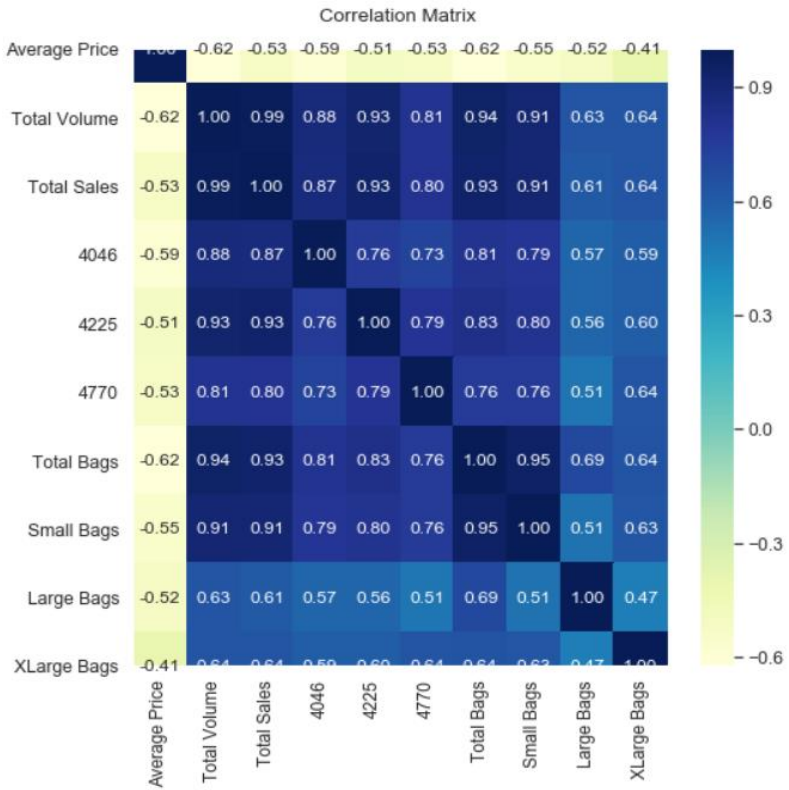
- In the above chart depicting total volume over 2015-2018, the months of January, February and May witnessed the highest sales whereas the ending months from September - December witnessed the least sales
- On the other hand, average price per avocado has been the highest during the ending months of September – December. This shows that there is an inverse relationship between average prices and the sales volume

Analysis – Bag Size Analysis



- Consumers prefer buying Small Bags of avocados as compared to Large and XLarge Bags
  - However, as avocados gained popularity amongst millennials, the share of Large bags has also increased
- From 2015 to 2016, the number of bags sold almost doubled, with a similar growth being registered by the Large bags

Analysis – Correlation Matrix



- There is a negative correlation of -0.62 between average prices and total volume. This means that as prices increase, the demand for avocados or the volume of avocados sold decreases
- Total Volume is also strongly related to the PLU 4225 (large Hass avocados) as compared to other variants of Hass avocados
- There is a linear relationship between total volume and Small Bags, Large Bags and XLarge bags. However, the correlation for Small bags is relatively more strong as compared to Large and XLarge bags

Thank You

Enjoy your guacamole!