

Avocado Price Prediction

KuldeepPatel

17/02/2022

Contents

Introduction	2
Exploratory Data Analysis:	3
Type of Avocados:	3
Relationship between Prices and Total Volume:	6
Patterns	8
Type of Patterns	8
Analyzing Seasonal Patterns:	8
Regional Analysis:	16
Time Series Analysis	19
Using the Autoplot library for Time Series:	19
Seasonal Subseries Plot:	21
Stationary Time Series:	22
Auto-Correlations	23
Type Series Forecasting	24
Forecasting Methods:	24
Using Naive and Drift Methods:	25
Understanding Residuals	32
Understanding ARIMA Model:	33
Conclusion	35

Introduction



In this kernel we will forecast using the time series in R. In this project, I will brief regarding the terms used and steps taken for the Forecasting Project.

Reference:

- Forecasting: Principles and Practice by Rob J Hyndman and George Athanasopoulos

Terms to Know:

- **Explanatory model:** Is a model that adds variables into our model not only based on past historical prices (in this case Avocados prices), but uses other variables to make accurate predictions.
- **Time Series Model:** Is a model that only uses past information (in this case past avocado prices) to make accurate predictions.

Steps for our Forecasting Project:

- **Determine what is the problem:** In this case we want to have accurate forecast of Avocado prices.
- **Gathering information:** Understand what was the process that was used to gather the information and if the information is sufficient to have effective predictive models.
- **Implementing Exploratory Analysis:** Determine if there are any sort of patterns in our data before going into building the models.
- **Choosing predictive models:** This is the phase were we decide which model is the most appropriate to make our forecasting most effective.
- **Testing our model:** Analyze if our model is effective enough to make effective predictions.

```
## [1] "avocado.csv"
```

This dataset contains the details of all sale made by type of avocado, season it was sold in.

Exploratory Data Analysis:

In this phase of the project we will concentrate mainly on two aspects of the time series forecasting analysis which consists of:

- **Seasonal Patterns:** In this section we will focus on constant patterns that occur frequently from year to year and from month to month in both types of avocados conventional and organic.
- **Cyclical Patterns:** This section will require me to search deeper into what factors could have affected the price of avocados during those years (low supply, storms etc.) and will be interesting to see if any of those factors contributed to any significant price changes.

```
## [1] "conventional" "organic"
```

Table 1: Table for Volume trend by Avocado type

type	avg.vol	pct
conventional	1653212.90	97.189269
organic	47811.21	2.810731

Type of Avocados:

In this section we will analyze the different types of avocados that we have in this dataset. Basically, we have two types of avocados:

- **Conventional**
- **Organic**

```
## [1] "Date"
```

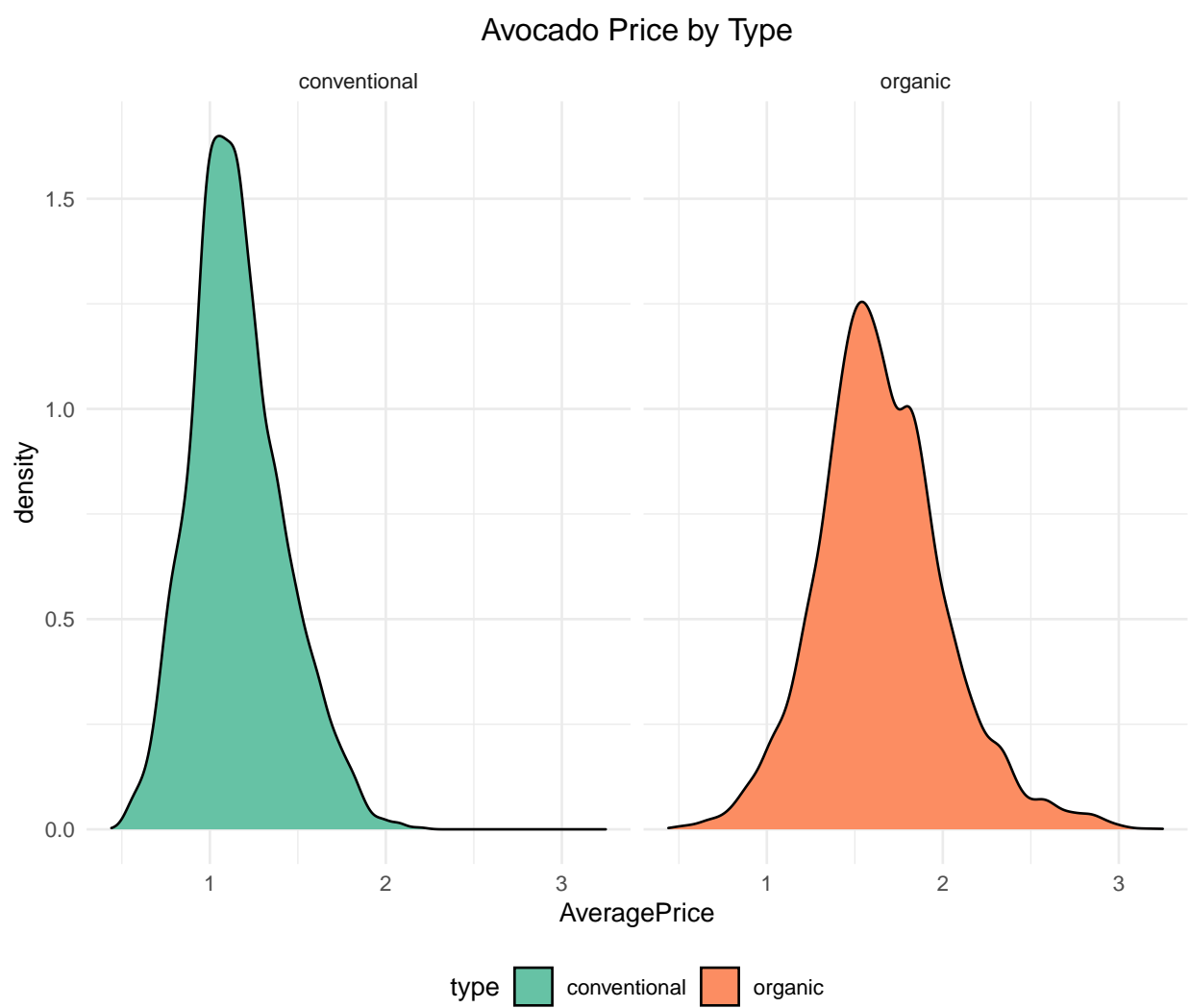
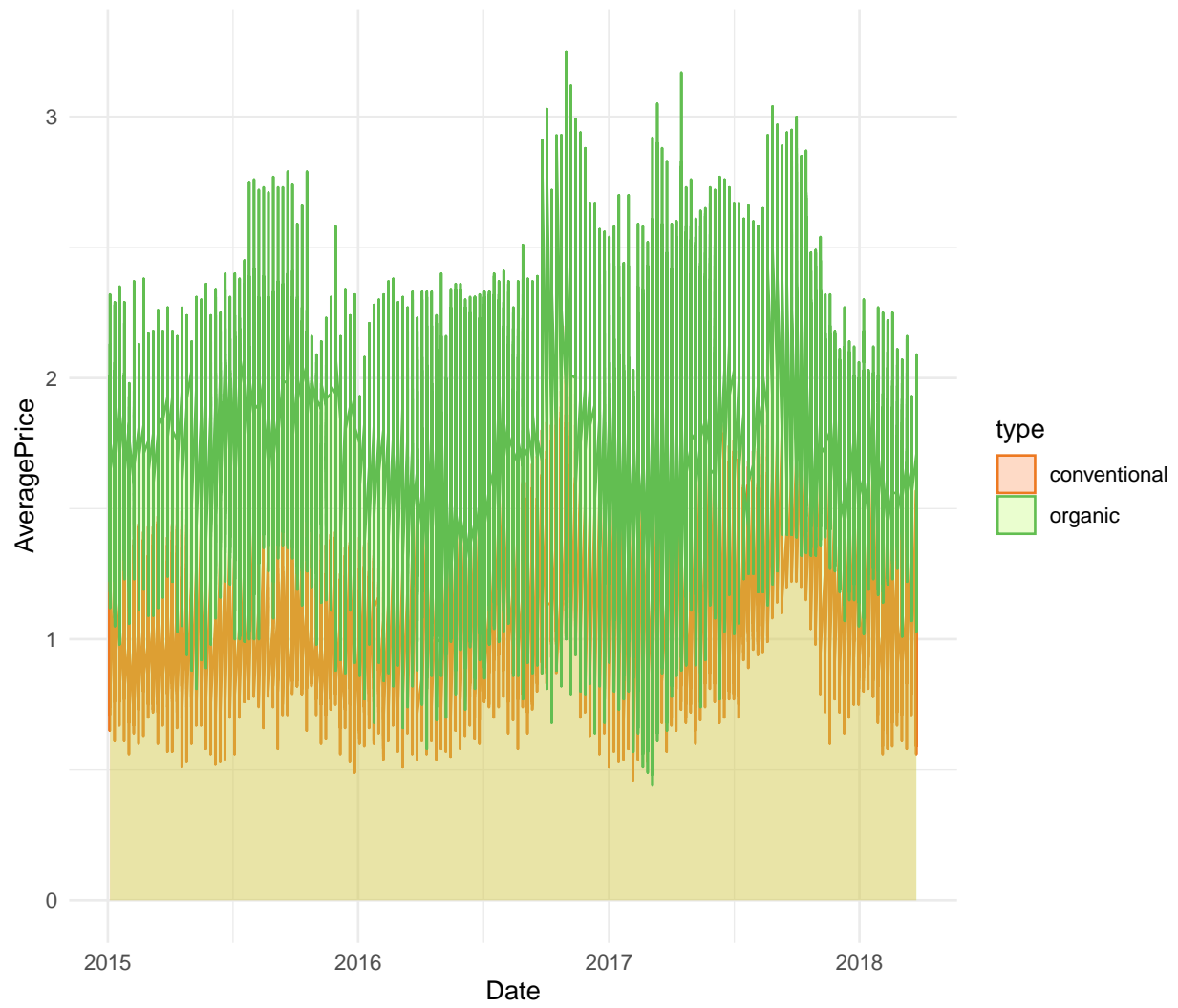
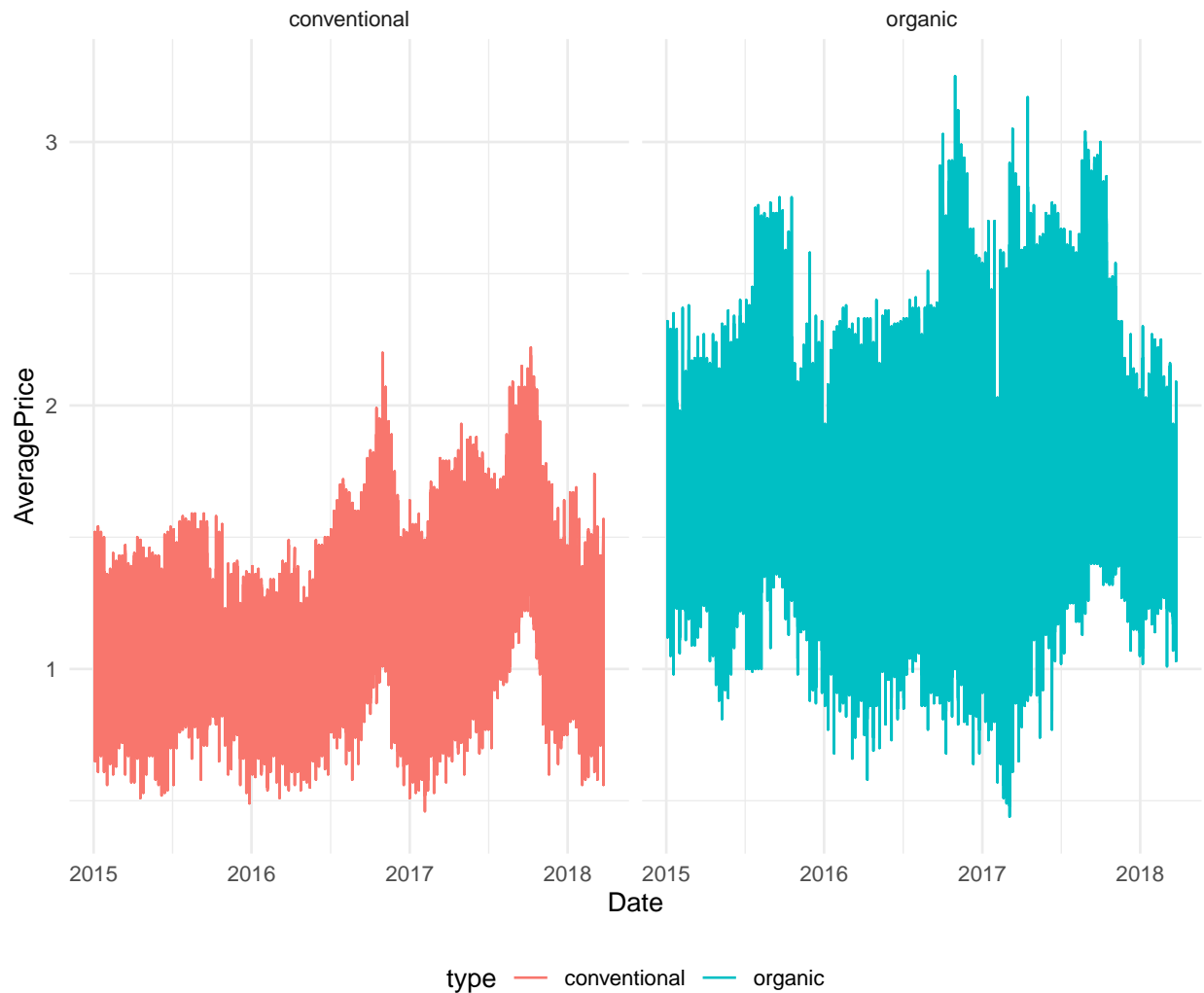


Figure 1: Plot of Avocado price by type





Summary:

- **Organic avocados:** Based on the price changes throughout time we can see that they are **more** expensive.
- **Conventional avocados:** Based on price changes throughout time we can see that they are **less** expensive.

Relationship between Prices and Total Volume:

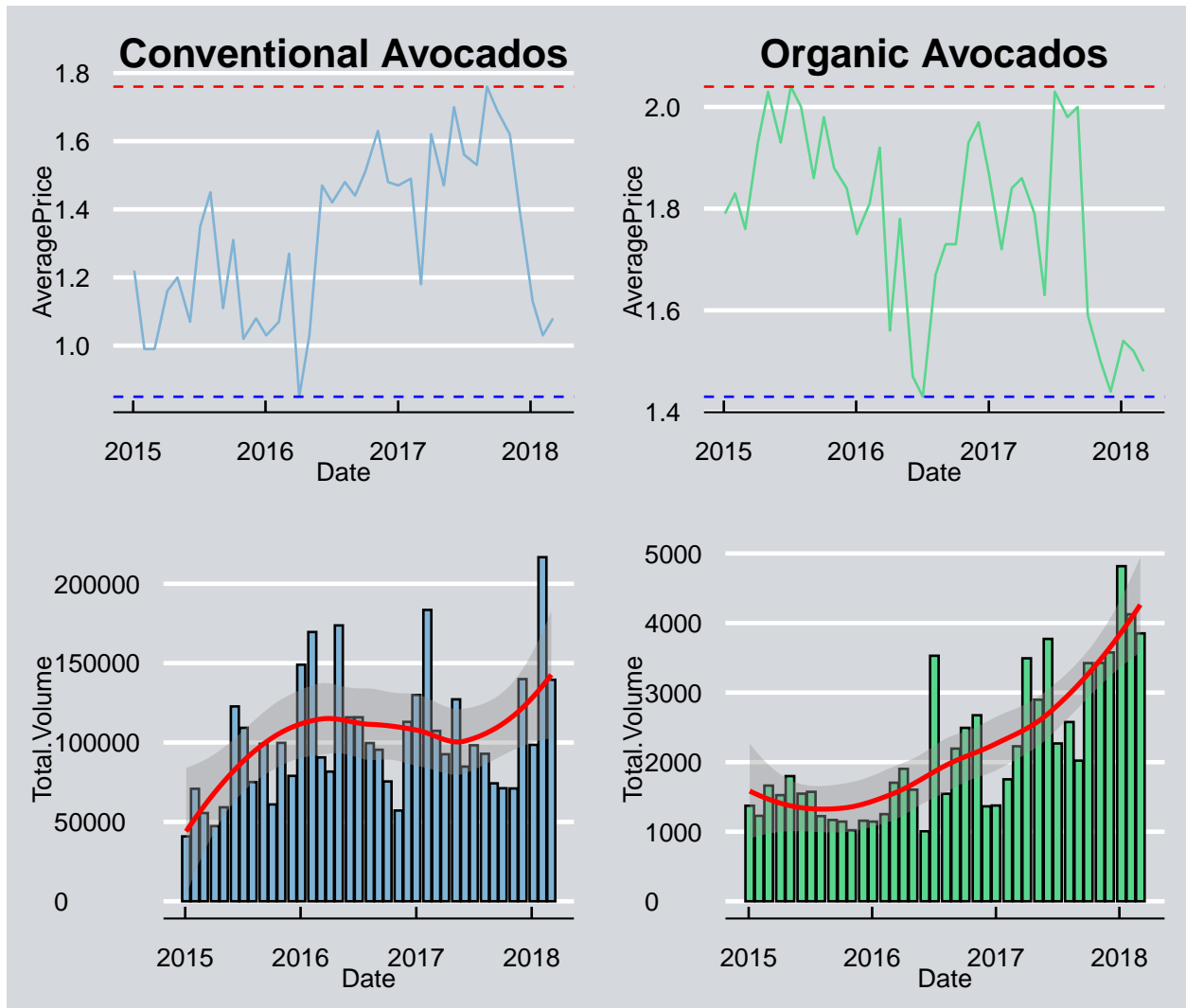
In this phase we will analyze the impact supply has on the price of avocados as a whole. Normally, there is an inverse relationship between supply and prices. When there is an overproduction of avocados they will have a negative impact on the market price of avocados. Let's see if this is the case for both conventional and organic avocados.

- **Conventional:** At the end of 2017 we can see a large drop in prices, at the same time there is an increasing amount of volume of avocados in the market.

- **Organic:** Same happens with organic avocados, at the end of 2017 there is a big drop and we can see a huge increase in volume.
- **Volume peaks:** Notice how each volume peak is a signal for an upcoming **drop** in avocado prices.

Table 2: Head of Conventional

Date	AveragePrice	type	Total.Volume
2015-01-04	1.22	conventional	40873.28
2015-02-01	0.99	conventional	70873.60
2015-03-01	0.99	conventional	55595.74
2015-04-05	1.16	conventional	47362.13
2015-05-03	1.20	conventional	59197.67
2015-06-07	1.07	conventional	122743.06



Patterns

Type of Patterns

- **Cyclical:** Fluctuations **do not happen** on a fixed frequency,
- **Seasonal:** Fluctuations **do happen** on a fixed frequency.
- **Trends:** This occurs when there is a **consistent increase or decrease** in our data.

Table 3: Head records for Organic

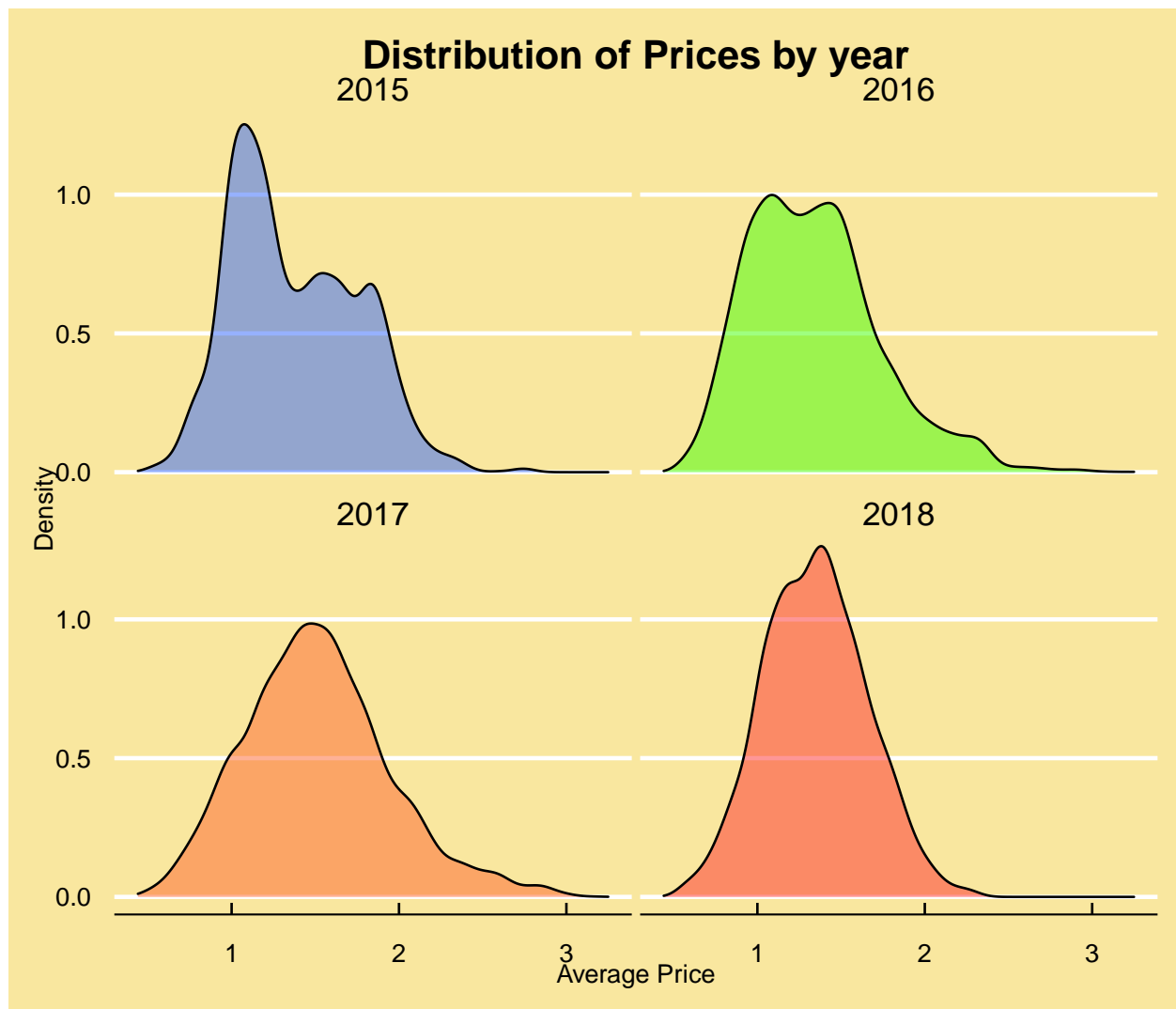
Date	AveragePrice	type	Total.Volume
2015-01-04	1.79	organic	1373.95
2015-02-01	1.83	organic	1228.51
2015-03-01	1.76	organic	1663.35
2015-04-05	1.93	organic	1526.10
2015-05-03	2.03	organic	1798.99
2015-06-07	1.93	organic	1547.03

Analyzing Seasonal Patterns:

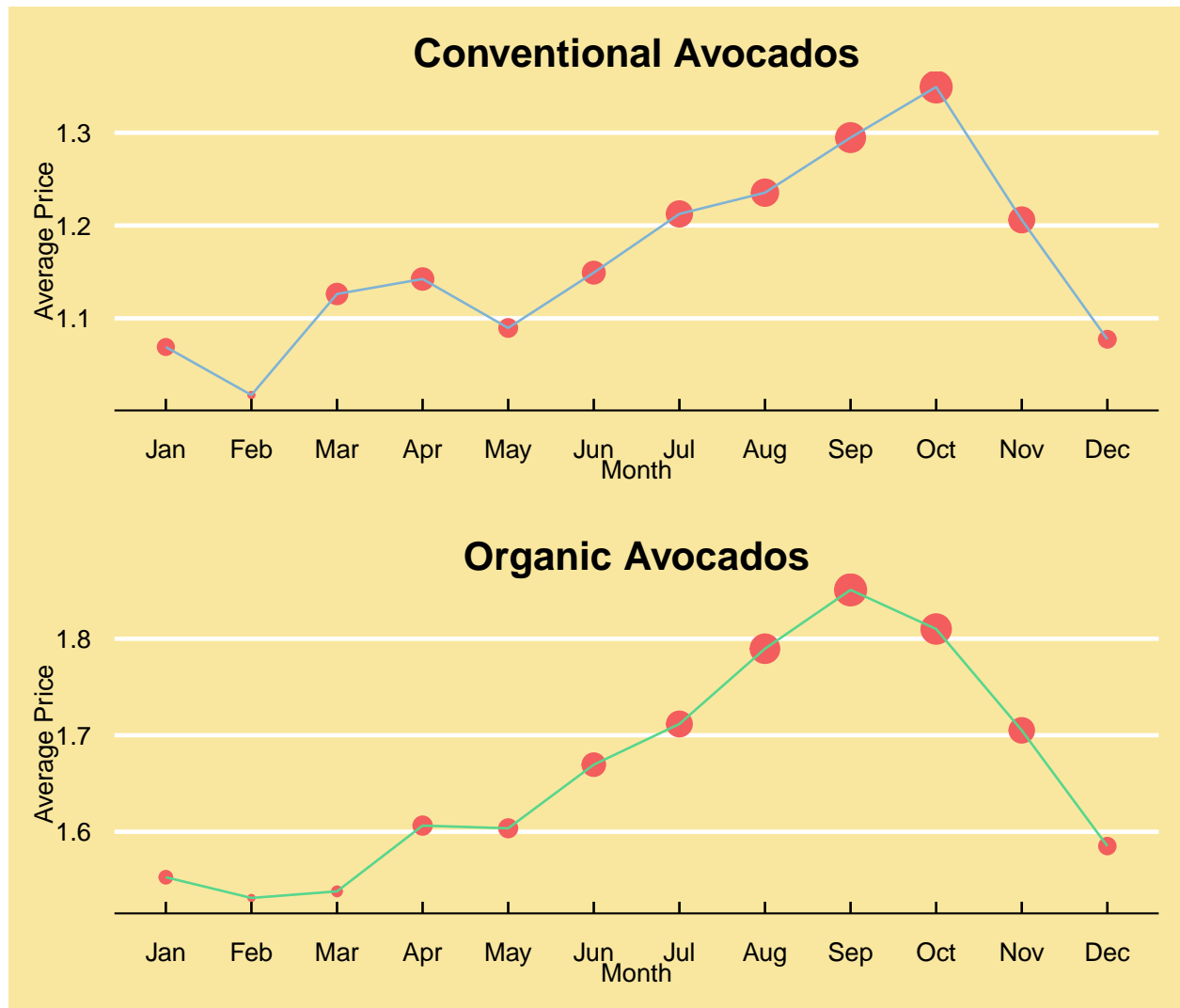
In this section we will try to find if there are any significant **reoccurring seasonal patterns**. By this I meant, if there are any repeating trends in which the avocado price tend to increase. For instance, in May of each year we see that avocado prices tend to increase for some specific reason.

Summary

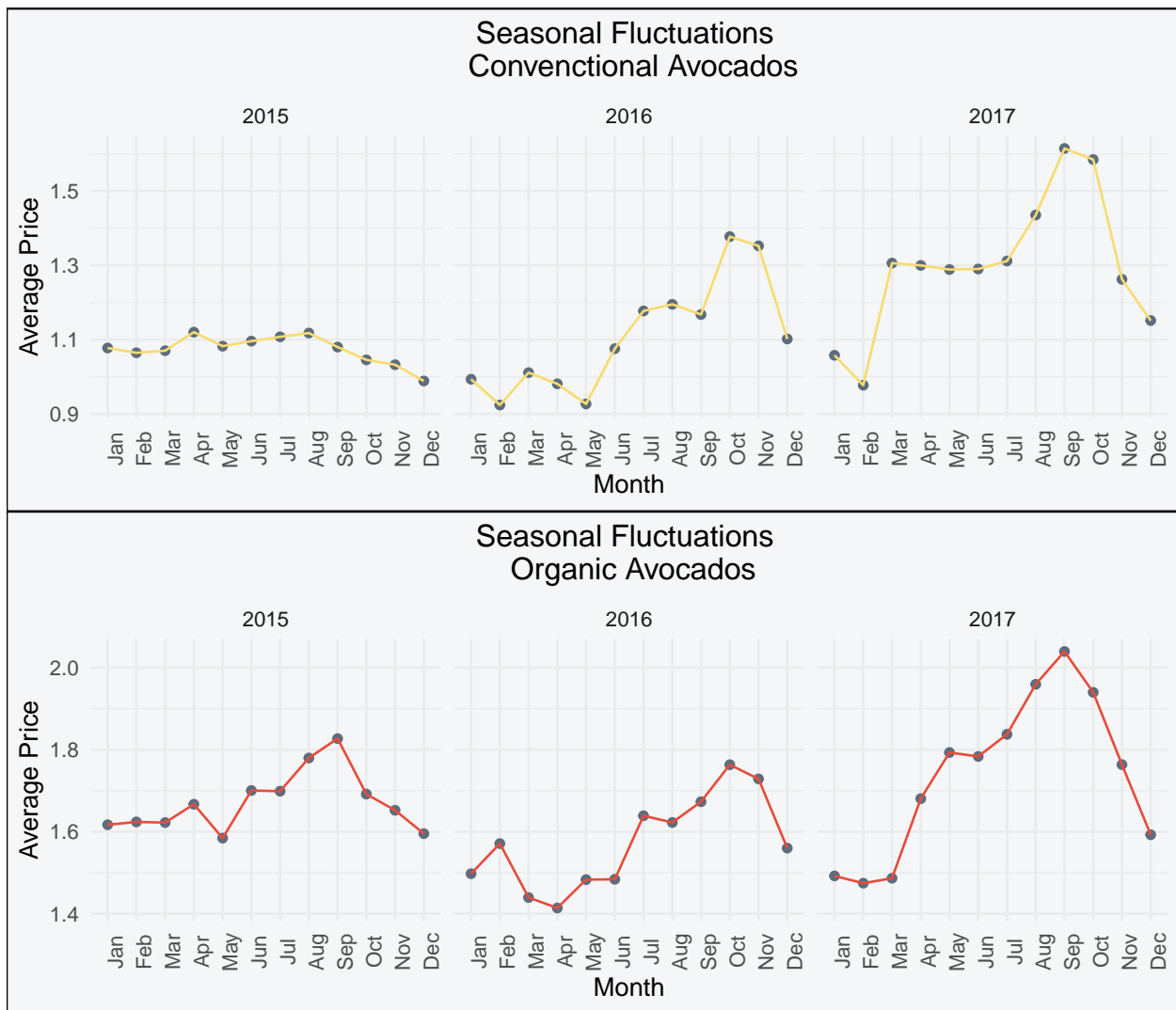
- **Distributions per year:** It looks that most of the prices in the year of 2015 were in the \$1.00 for conventional avocados. While for 2016 and 2017 the density of the prices were a little bit higher.



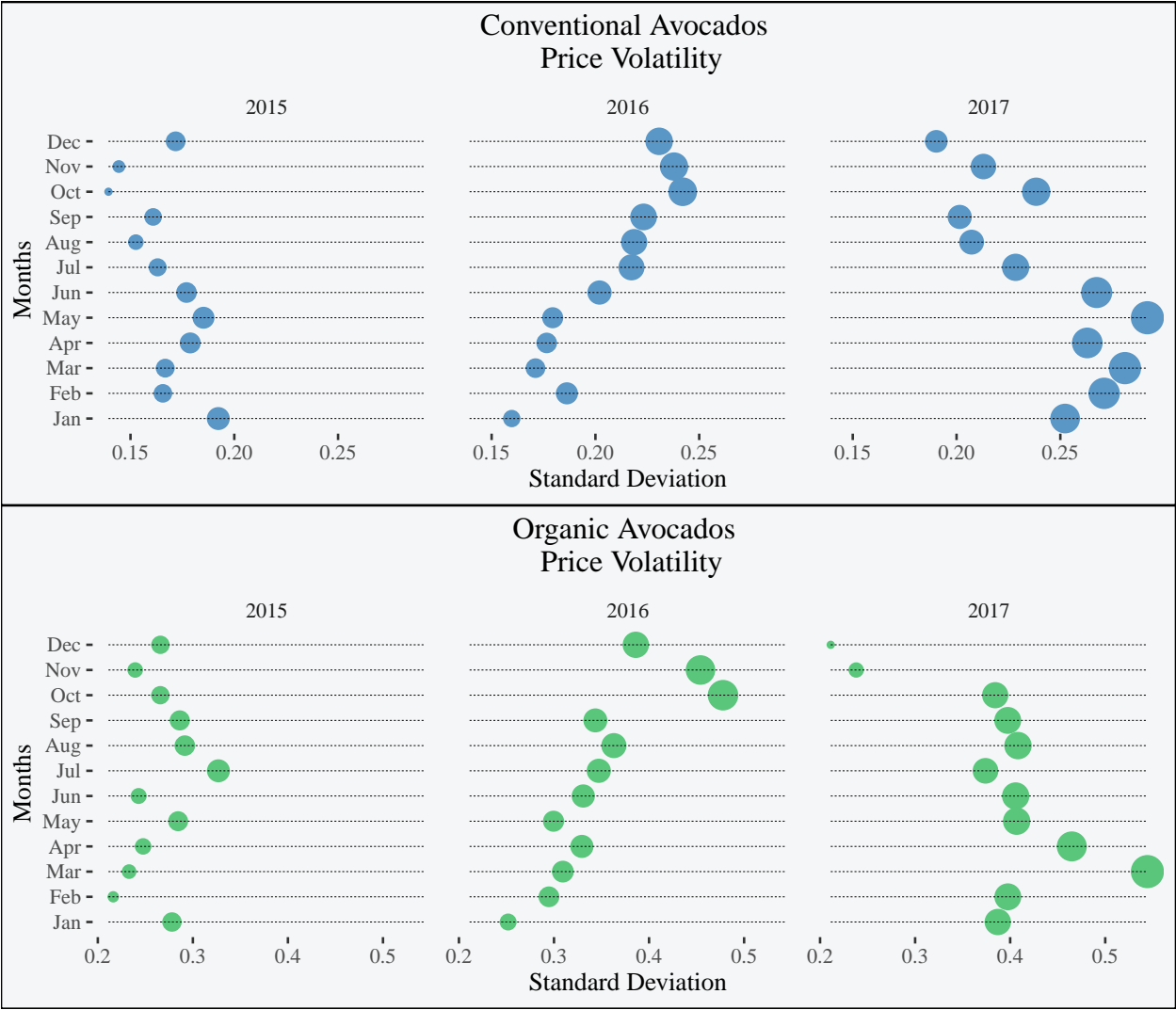
- **Price peaks per Month:** It looks that most price peaks occur for both conventional and organic avocados between the months of **September** and **October**.

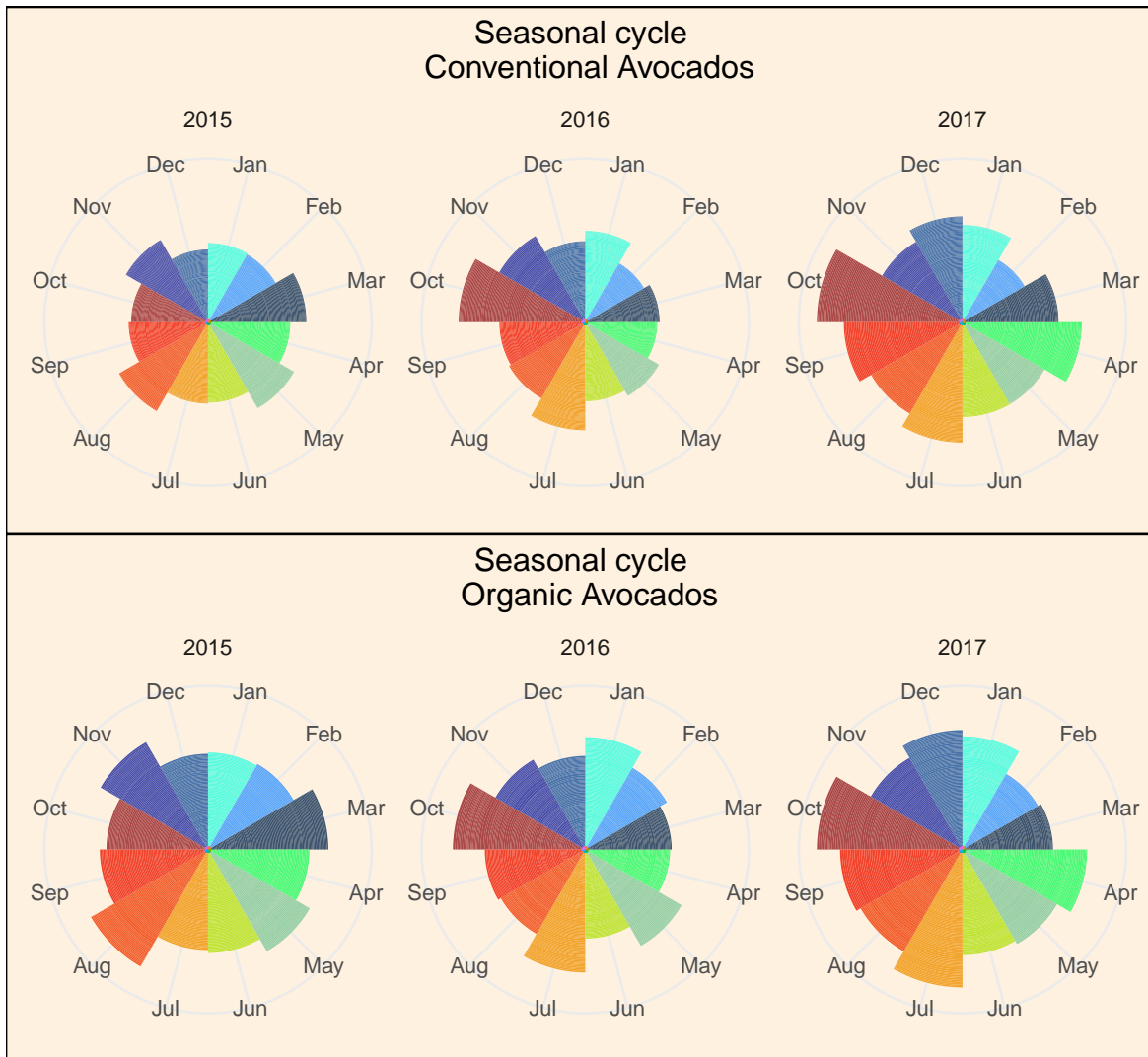


- **Major drop in prices at the end of the year:** Interesting enough we see that at the end of the year there is a major price drop in the price of avocados. I wonder why that would be? What possible reason could there be to have a lower price at the end of each year.

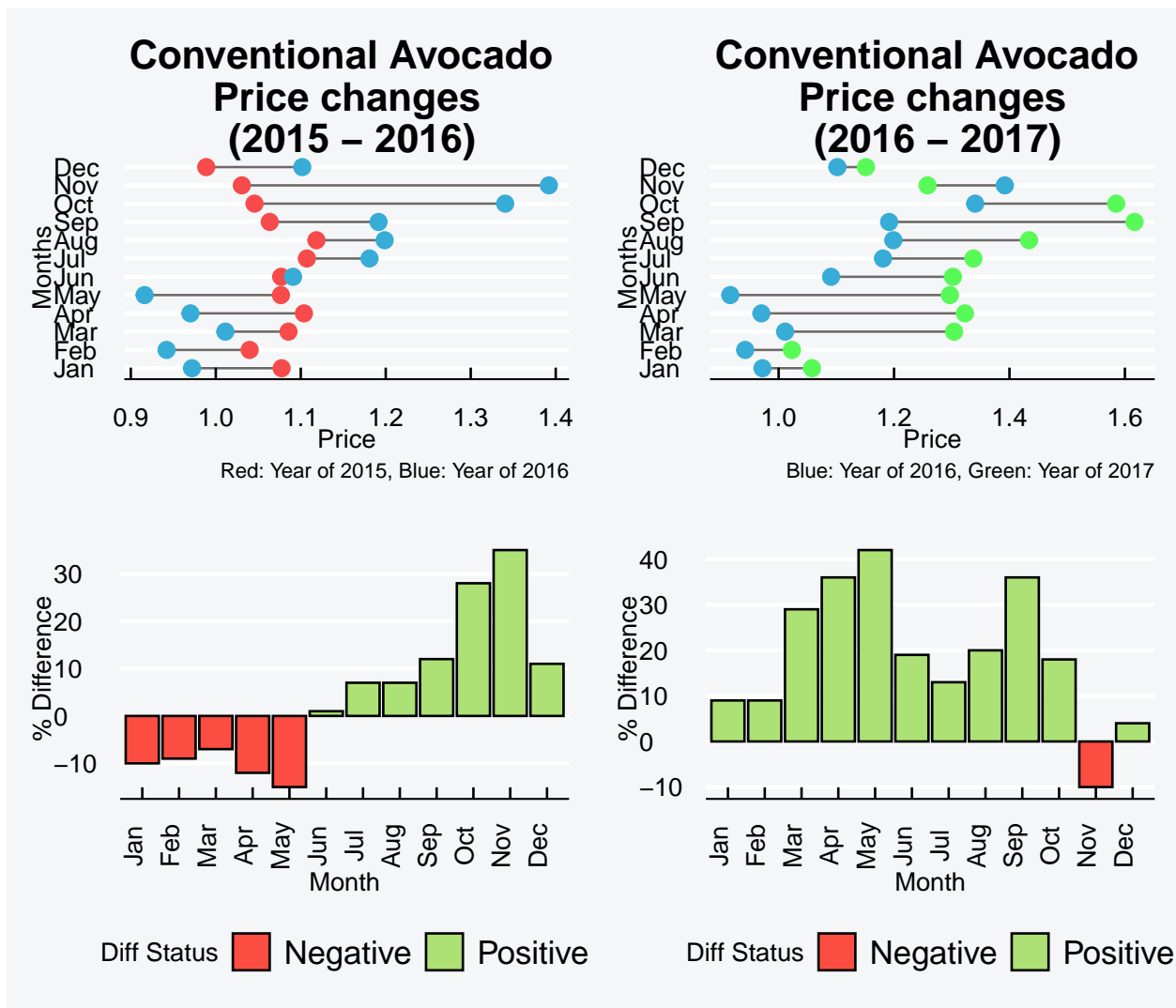


- Standard deviation as a measure of volatility:** Standard deviation is just the square root of a variance. We can see that during the year of 2017, the avocado market experienced the highest volatility for both conventional and organic avocados.

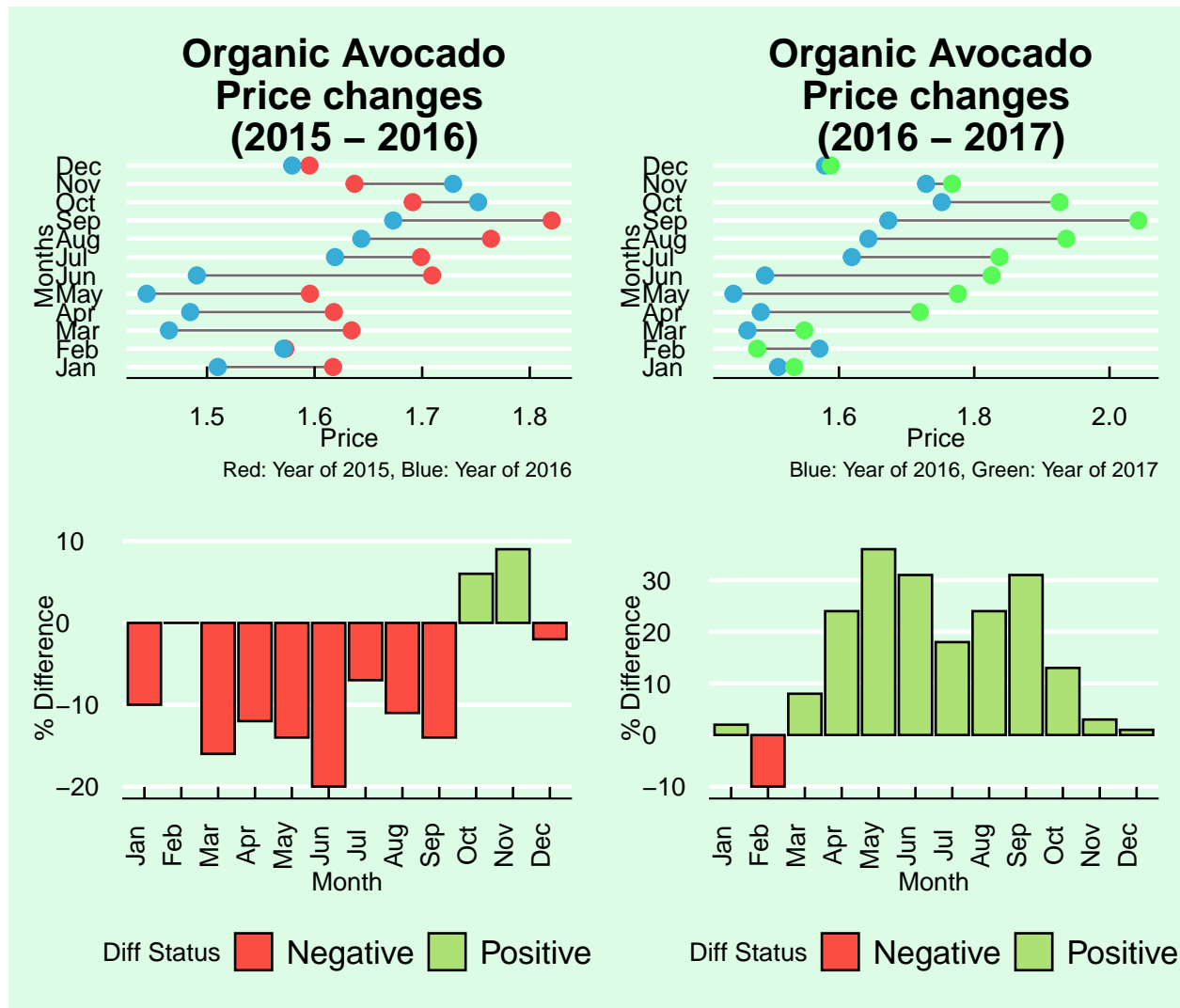




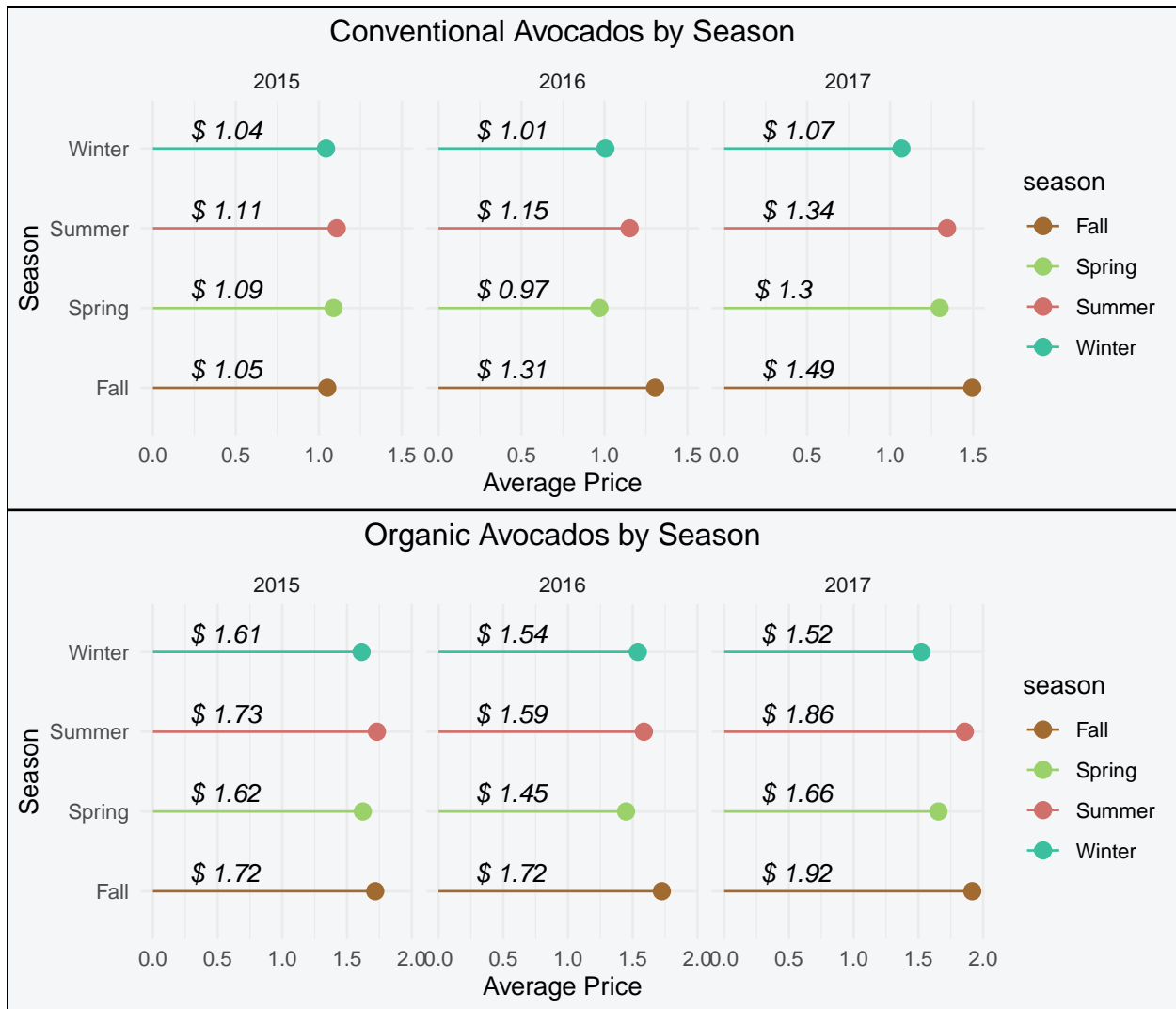
- Percent differences for conventional avocados:** Here we want to see if there is a percent increase from year to year when it comes to the average price per month of each type of avocado. From 2015 - 2016 we see that the conventional avocados had a lag in the first six months of the year, then it picked for the next six months. From 2016 - 2017 conventional avocados performed greatly in the year of 2017 as opposed to the year of 2016 when compared to the previous year which was 2016.



- **Percent differences for organic avocados:** Now for organic avocados is a totally different story at least for the years 2015 - 2016. In 2016 organic avocados performed worst in almost all months as opposed to the previous year of 2015. Then it performed better in the year of 2017 in almost all the months of 2017 as opposed to the previous year of 2016.



- **2017 a good year for avocados:** Based on this trend, we can see that 2017 was a good year for avocados, would our forecast (later in this project) predict that it will do so in the year of 2018?



Formulas used for this section:

- **Percent increase Difference:**

$$\frac{\text{EndingValue} - \text{BeginingValue}}{\text{BeginingValue}}$$

- **Standard Deviation:**

$$\sqrt{\frac{\sum |x - \bar{x}|^2}{n}}$$

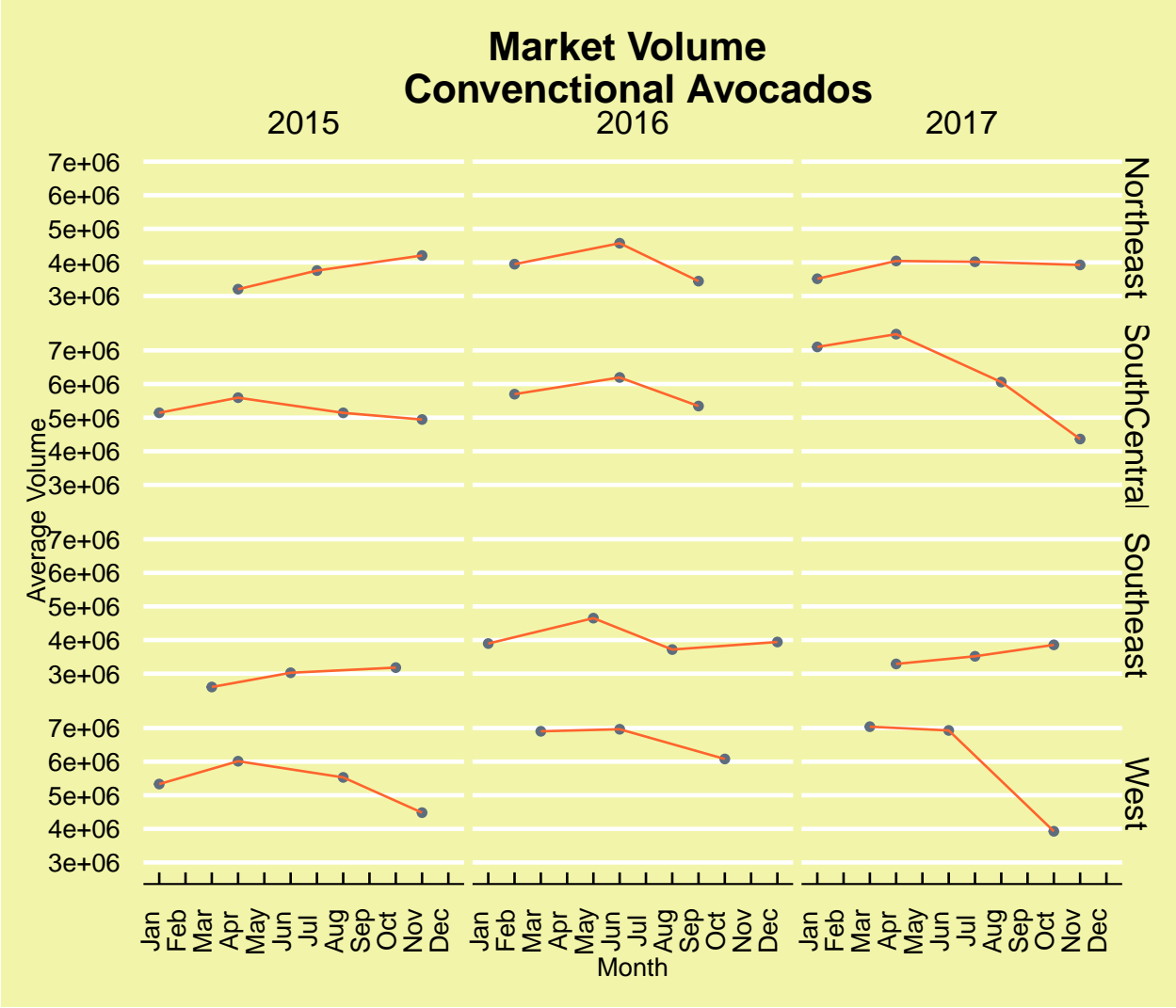
Regional Analysis:

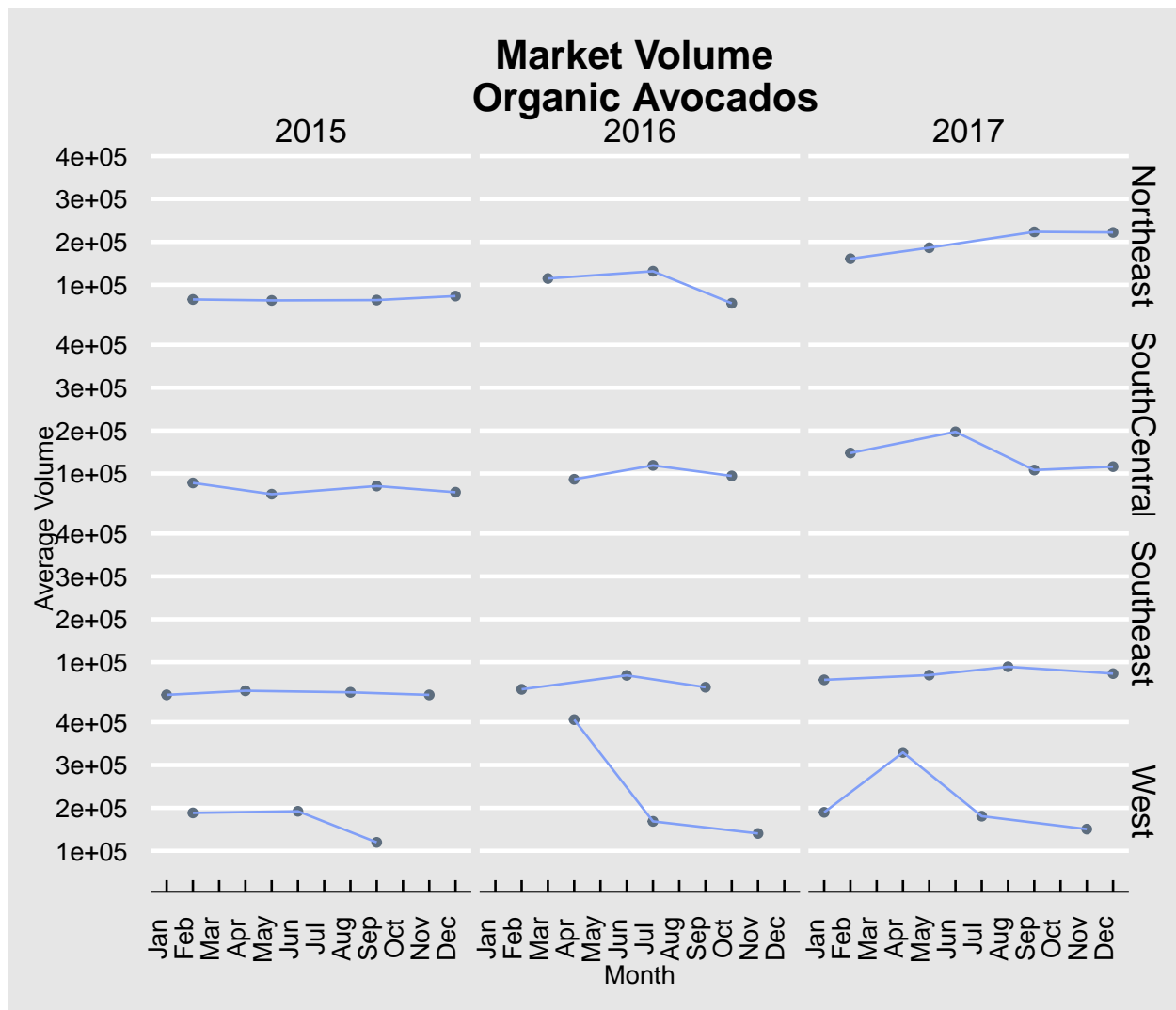
In this phase of the exploratory data analysis we will analyze two main things the volume of avocados and price of the different types of avocados in each region. We will see the inverse correlation there is between supply and demand and how it affects price. Since we have 54 regions, we will use just a few to make note of this inverse correlation.

Volume Price Difference = *Volume of Avocados Produced* – *Average Avocado Price*



Market Volume





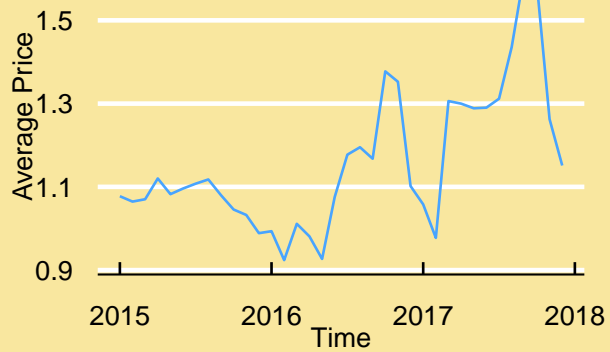
Time Series Analysis

Using the Autoplot library for Time Series:

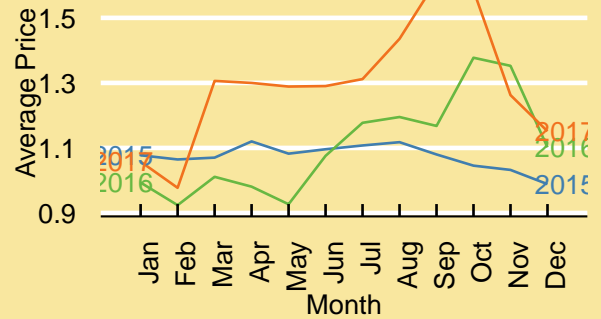
The auto-plot library allows us to see patterns between the different years. In the chart below we can see the following:

- Prices for 2017 were the highest overall.
- There is an increase in avocado prices for both organic and conventional types.
- The lowest price occurred in the year of 2015

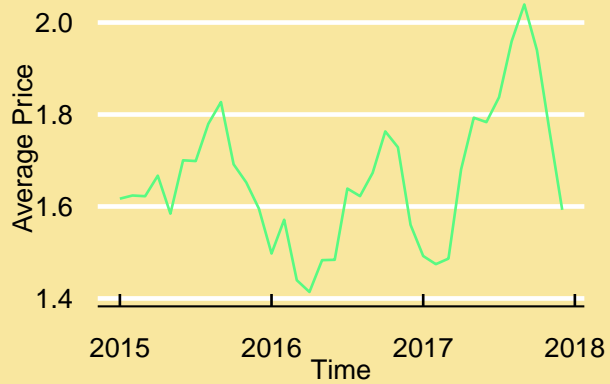
Average Price by Month Conventional Avocados



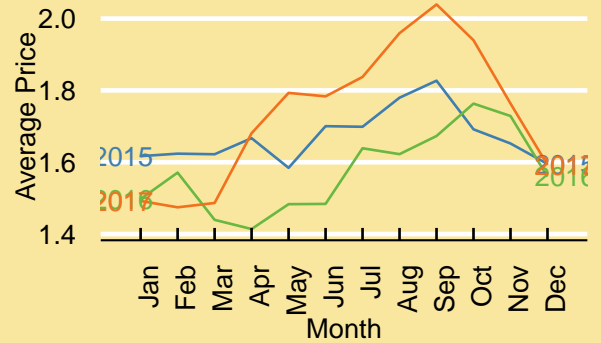
Average Conventional A. Price by Year for each month

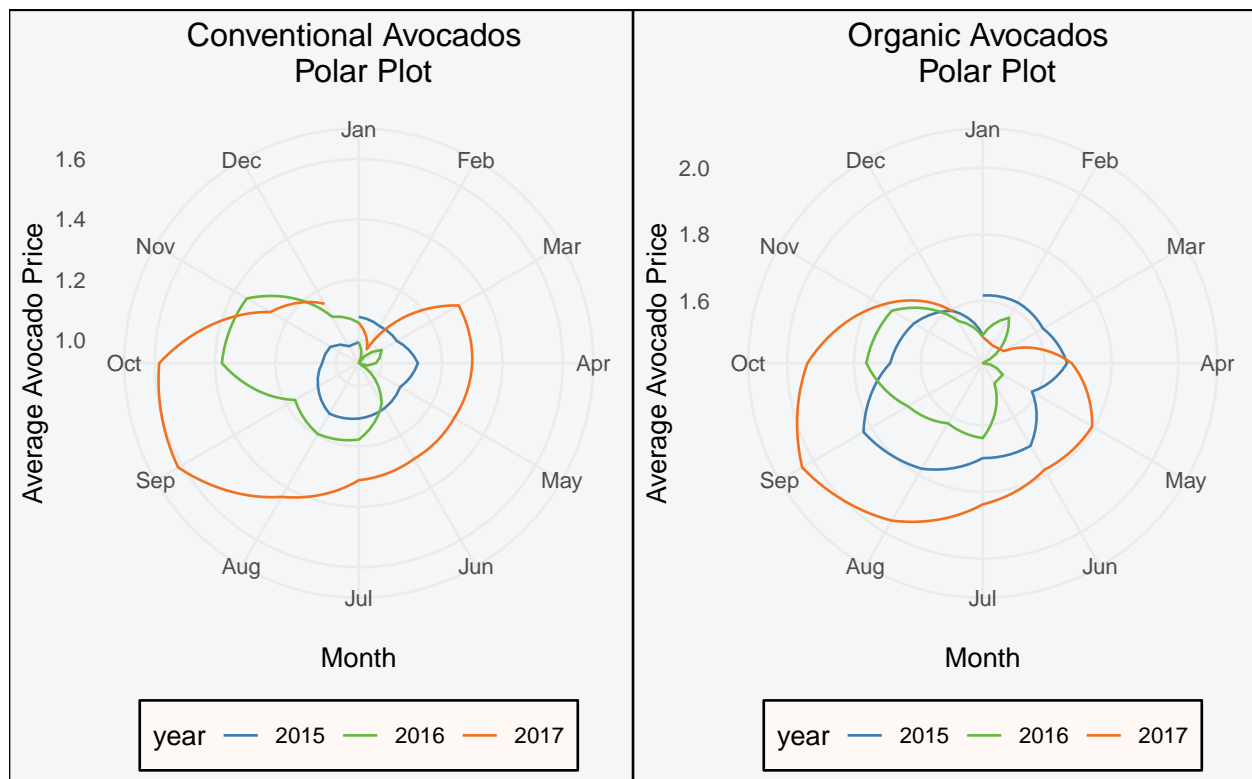


Average Price by Month Organic Avocados



Average Organic A. Price by Year for each month





Seasonal Subseries Plot:

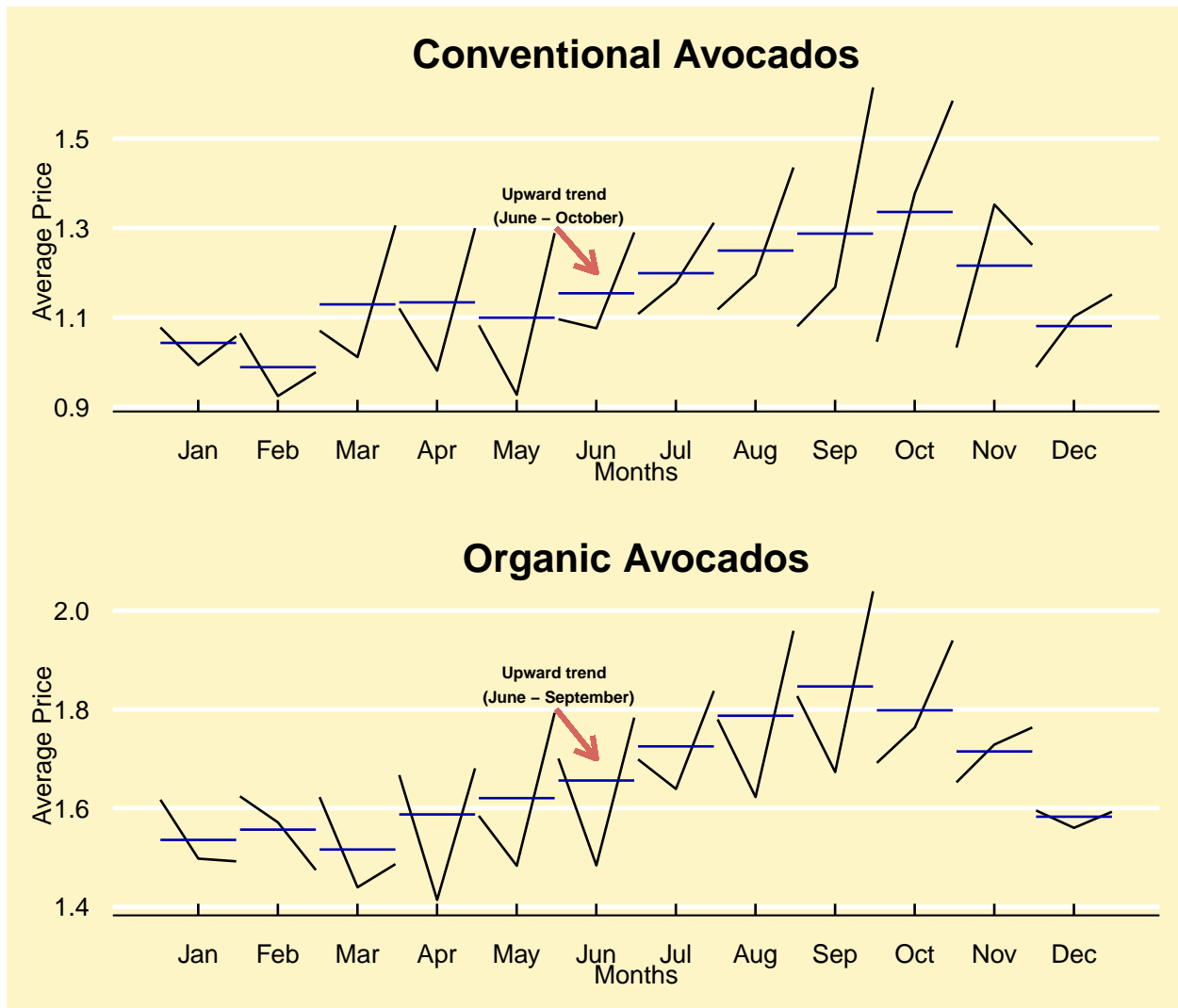
This is another form of analyzing the different trends per seasonality. Here is what you should know:

- **Blue line:** Is the **mean** of the avocado month for a specific month.
- **Black line:** It is a way to see the **fluctuations** throughout each month.

Summary

- **Upward Trend:** The most significant upward trend happens between June - September for organic avocados and June - October for conventional avocado types. Both types of avocados have a similar trend.
- **Most expensive months to buy avocados:** Like in the previous analysis, through this visualization we confirm that the months with the highest average price are September and October however, August is pretty expensive as well.

- **What does this graph tell us?** For each month we have three peaks. That represents the highest peak in each month of each year. Since we are using the prices of 2015, 2016, and 2017 we can see how the price of each year behave in all twelve month. This will help us see if there are any major seasonal patterns. Also, we can see how the average change behaves, allowing us to see if there are any major trends in each month.



Stationary Time Series:

Auto-Correlations:

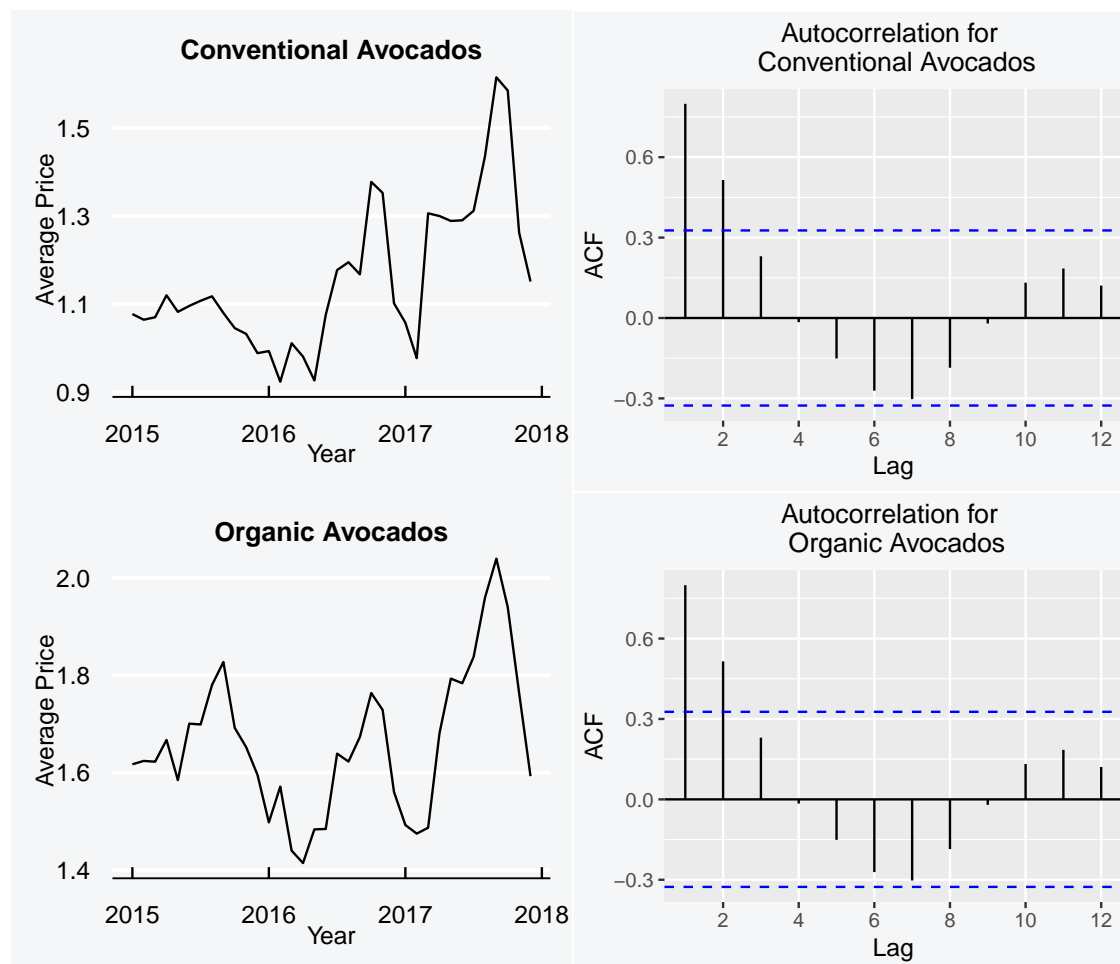
- **What does the autocorrelation graph tells us?** -What can we conclude about the autocorrelation?

Autocorrelation Formula:

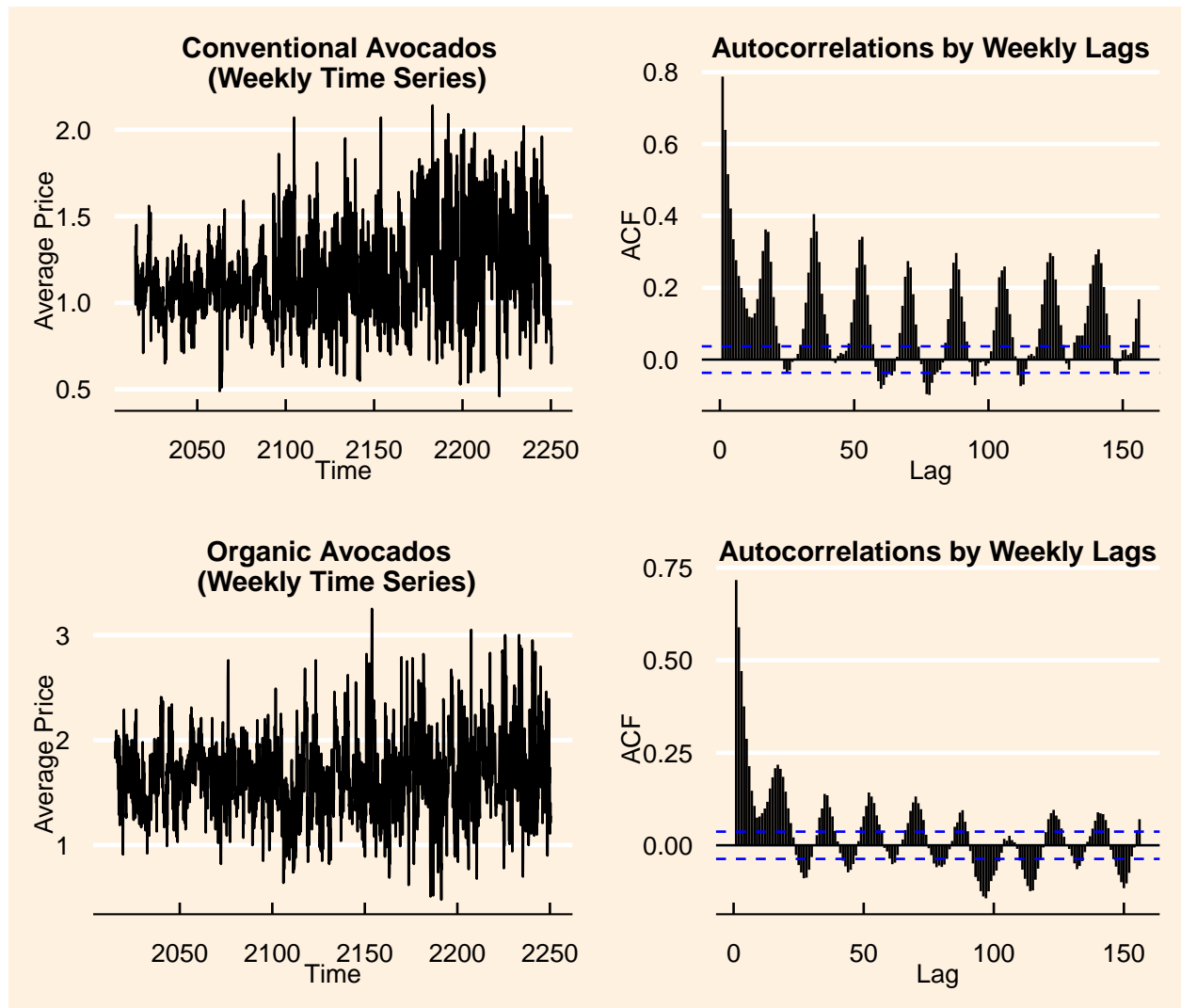
$$r_k = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=k+1}^T (y_t - \bar{y})^2}$$

Auto-Correlations

Auto-correlations helps us determine if there are certain patterns in our data. So what is the main purpose of using Auto-correlations and how could it help us determine certain patterns in our time series dataset? First thing, the word “Auto” gives us a hint that we are planning to use an internal correlation. What does internal correlation mean? We basically are taking a specific date (in the example below we use January) and we want to see whether the price of avocados have some sort of correlation with the price of January. In the first visualization, lags are a representation of months. So we can see that the correlation is higher the closer the month is from January/ This means that the price of February are similar to the avocado prices of January. As the months pass by, the correlation as opposed to the month of January gets lower indicating to us that there are no specific patterns to the month of January. Price movements in those months except for February and March do not have highly correlated prices compared to the prices of January.



Weekly Auto-Correlations:



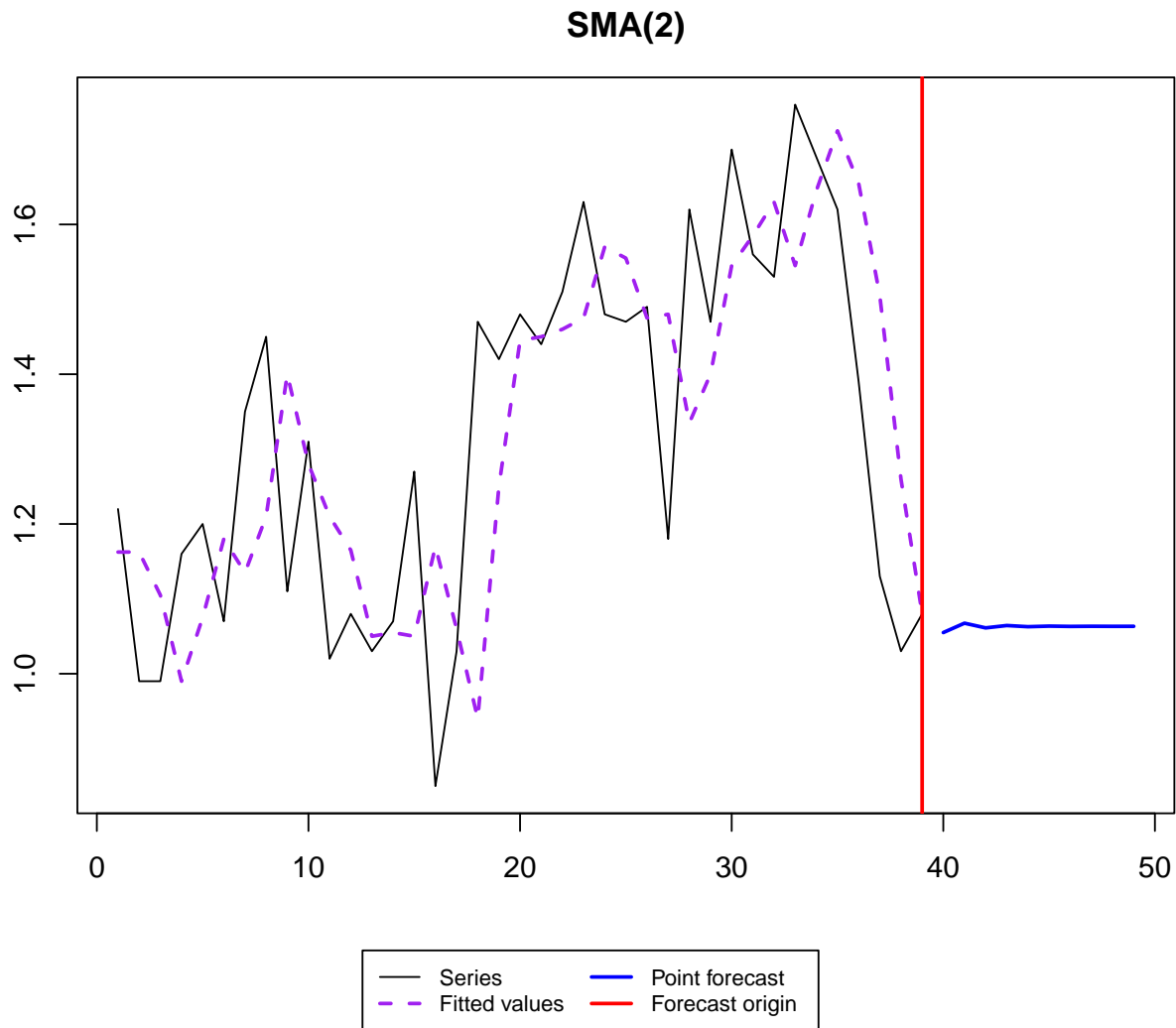
Type Series Forecasting

Forecasting Methods:

In this section we will use various models to come up with accurate predictions in order to see what the upcoming patterns will be for avocado prices.

The list of models we will use include:

- Smoothing Moving Average
- Seasonal Naive Method
- Drift Method
- ARIMA

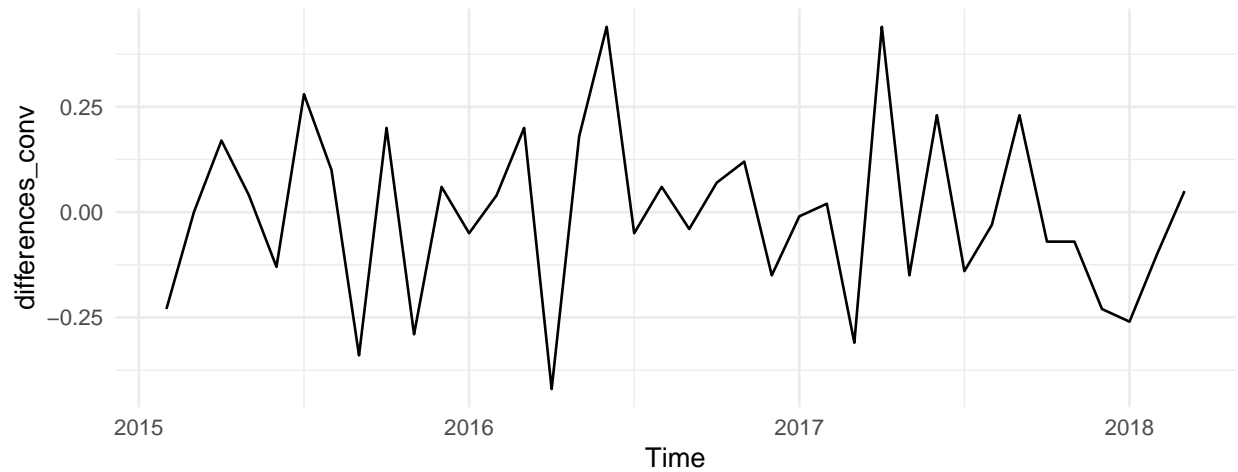


Using Naive and Drift Methods:

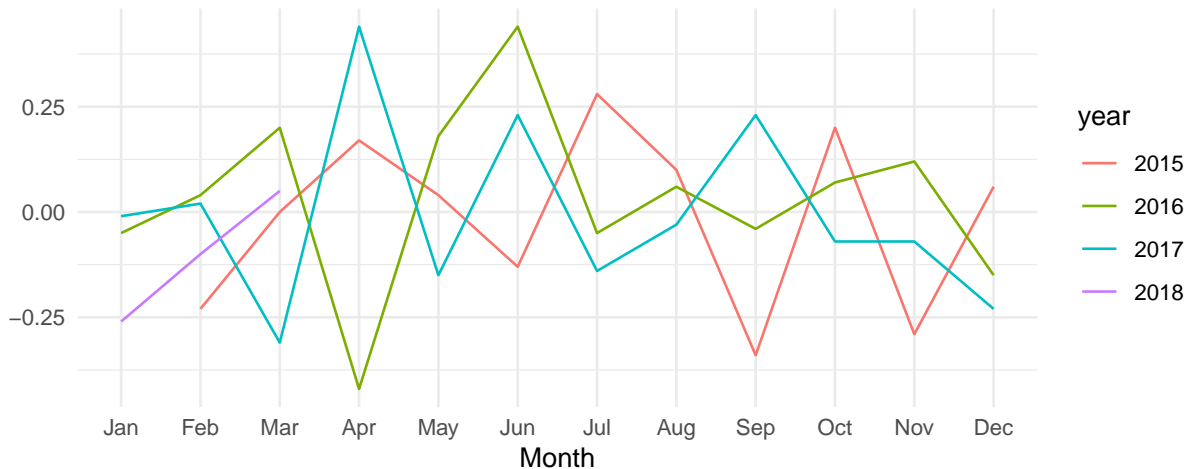
Before explaining how these types of forecasting models work there are a couple of things that we must understand:

- **Average Method:** This is just the average line of the last three years that we are analyzing. These years are 2015, 2016 and 2017. This will give us a fix value however, this value is important. When we talk about technical analysis, whenever prices go below the average line we are “expected” to see a downward trend.
- **Naive Method:** Is just setting all forecast values to the last value we saw. You can see that the blue line is at the same level of the last price in December. That explain why it is so low! Since the avocado price was so low for both conventional and organic avocados, this could have a similar functionality to a support line. If we see prices going below this line, it could indicate a new downward trend.

- **Seasonal Naive Method:** This will equal to the last value of the past year. So for instance if we are using monthly data, our forecast in january will be equivalent to the last January value of the year of 2018.
- **Drift Method:** This will better detect the pattern of prices since it will use the average change of the historical data as a point of reference when it comes to forecasting. It is most likely, that this forecast will understand the trend however, it will ignore any seasonality patterns.



Seasonal plot: differences_conv



```
##
## ARIMA(0,1,0)(0,1,0)[12] : 17.51894
## ARIMA(0,1,0)(0,1,1)[12] : Inf
## ARIMA(0,1,0)(1,1,0)[12] : 13.65754
## ARIMA(0,1,0)(1,1,1)[12] : Inf
## ARIMA(0,1,1)(0,1,0)[12] : 11.44504
## ARIMA(0,1,1)(0,1,1)[12] : Inf
```

```

## ARIMA(0,1,1)(1,1,0)[12] : 7.926653
## ARIMA(0,1,1)(1,1,1)[12] : Inf
## ARIMA(0,1,2)(0,1,0)[12] : 13.85906
## ARIMA(0,1,2)(0,1,1)[12] : Inf
## ARIMA(0,1,2)(1,1,0)[12] : 10.60394
## ARIMA(0,1,2)(1,1,1)[12] : Inf
## ARIMA(0,1,3)(0,1,0)[12] : Inf
## ARIMA(0,1,3)(0,1,1)[12] : Inf
## ARIMA(0,1,3)(1,1,0)[12] : Inf
## ARIMA(0,1,3)(1,1,1)[12] : Inf
## ARIMA(0,1,4)(0,1,0)[12] : Inf
## ARIMA(0,1,4)(0,1,1)[12] : Inf
## ARIMA(0,1,4)(1,1,0)[12] : Inf
## ARIMA(0,1,5)(0,1,0)[12] : Inf
## ARIMA(1,1,0)(0,1,0)[12] : 15.40448
## ARIMA(1,1,0)(0,1,1)[12] : Inf
## ARIMA(1,1,0)(1,1,0)[12] : 10.08229
## ARIMA(1,1,0)(1,1,1)[12] : Inf
## ARIMA(1,1,1)(0,1,0)[12] : 13.96476
## ARIMA(1,1,1)(0,1,1)[12] : Inf
## ARIMA(1,1,1)(1,1,0)[12] : 10.64378
## ARIMA(1,1,1)(1,1,1)[12] : Inf
## ARIMA(1,1,2)(0,1,0)[12] : Inf
## ARIMA(1,1,2)(0,1,1)[12] : Inf
## ARIMA(1,1,2)(1,1,0)[12] : Inf
## ARIMA(1,1,2)(1,1,1)[12] : Inf
## ARIMA(1,1,3)(0,1,0)[12] : Inf
## ARIMA(1,1,3)(0,1,1)[12] : Inf
## ARIMA(1,1,3)(1,1,0)[12] : Inf
## ARIMA(1,1,4)(0,1,0)[12] : Inf
## ARIMA(2,1,0)(0,1,0)[12] : 12.20614
## ARIMA(2,1,0)(0,1,1)[12] : Inf
## ARIMA(2,1,0)(1,1,0)[12] : 10.50862
## ARIMA(2,1,0)(1,1,1)[12] : Inf
## ARIMA(2,1,1)(0,1,0)[12] : 14.37757
## ARIMA(2,1,1)(0,1,1)[12] : Inf
## ARIMA(2,1,1)(1,1,0)[12] : 13.47267
## ARIMA(2,1,1)(1,1,1)[12] : Inf
## ARIMA(2,1,2)(0,1,0)[12] : 15.41755
## ARIMA(2,1,2)(0,1,1)[12] : Inf
## ARIMA(2,1,2)(1,1,0)[12] : Inf
## ARIMA(2,1,3)(0,1,0)[12] : Inf
## ARIMA(3,1,0)(0,1,0)[12] : 15.01092
## ARIMA(3,1,0)(0,1,1)[12] : Inf
## ARIMA(3,1,0)(1,1,0)[12] : 13.58839

```

```

## ARIMA(3,1,0)(1,1,1)[12] : Inf
## ARIMA(3,1,1)(0,1,0)[12] : 18.00893
## ARIMA(3,1,1)(0,1,1)[12] : Inf
## ARIMA(3,1,1)(1,1,0)[12] : 16.8641
## ARIMA(3,1,2)(0,1,0)[12] : Inf
## ARIMA(4,1,0)(0,1,0)[12] : 16.62101
## ARIMA(4,1,0)(0,1,1)[12] : Inf
## ARIMA(4,1,0)(1,1,0)[12] : 15.52644
## ARIMA(4,1,1)(0,1,0)[12] : 18.49889
## ARIMA(5,1,0)(0,1,0)[12] : 18.45644
##
##
##
## Best model: ARIMA(0,1,1)(1,1,0)[12]

##
## ARIMA(0,1,0)(0,1,0)[12] : 2.521273
## ARIMA(0,1,0)(0,1,1)[12] : Inf
## ARIMA(0,1,0)(1,1,0)[12] : -2.28856
## ARIMA(0,1,0)(1,1,1)[12] : Inf
## ARIMA(0,1,1)(0,1,0)[12] : 4.237894
## ARIMA(0,1,1)(0,1,1)[12] : Inf
## ARIMA(0,1,1)(1,1,0)[12] : -0.9972336
## ARIMA(0,1,1)(1,1,1)[12] : Inf
## ARIMA(0,1,2)(0,1,0)[12] : 6.536184
## ARIMA(0,1,2)(0,1,1)[12] : Inf
## ARIMA(0,1,2)(1,1,0)[12] : 1.305819
## ARIMA(0,1,2)(1,1,1)[12] : Inf
## ARIMA(0,1,3)(0,1,0)[12] : 9.325564
## ARIMA(0,1,3)(0,1,1)[12] : Inf
## ARIMA(0,1,3)(1,1,0)[12] : 3.595351
## ARIMA(0,1,3)(1,1,1)[12] : Inf
## ARIMA(0,1,4)(0,1,0)[12] : Inf
## ARIMA(0,1,4)(0,1,1)[12] : Inf
## ARIMA(0,1,4)(1,1,0)[12] : 7.006226
## ARIMA(0,1,5)(0,1,0)[12] : Inf
## ARIMA(1,1,0)(0,1,0)[12] : 4.34258
## ARIMA(1,1,0)(0,1,1)[12] : Inf
## ARIMA(1,1,0)(1,1,0)[12] : -0.7000006
## ARIMA(1,1,0)(1,1,1)[12] : Inf
## ARIMA(1,1,1)(0,1,0)[12] : Inf
## ARIMA(1,1,1)(0,1,1)[12] : Inf
## ARIMA(1,1,1)(1,1,0)[12] : Inf
## ARIMA(1,1,1)(1,1,1)[12] : Inf
## ARIMA(1,1,2)(0,1,0)[12] : 8.908865

```

```

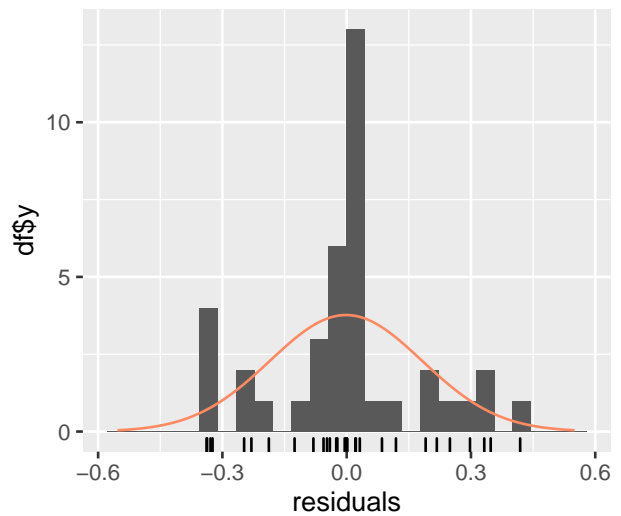
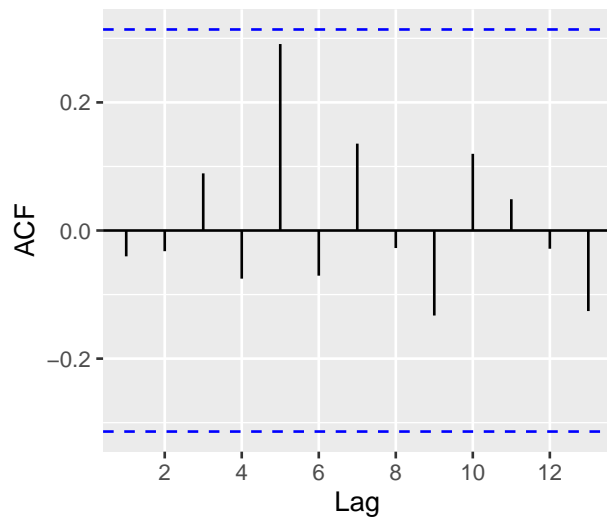
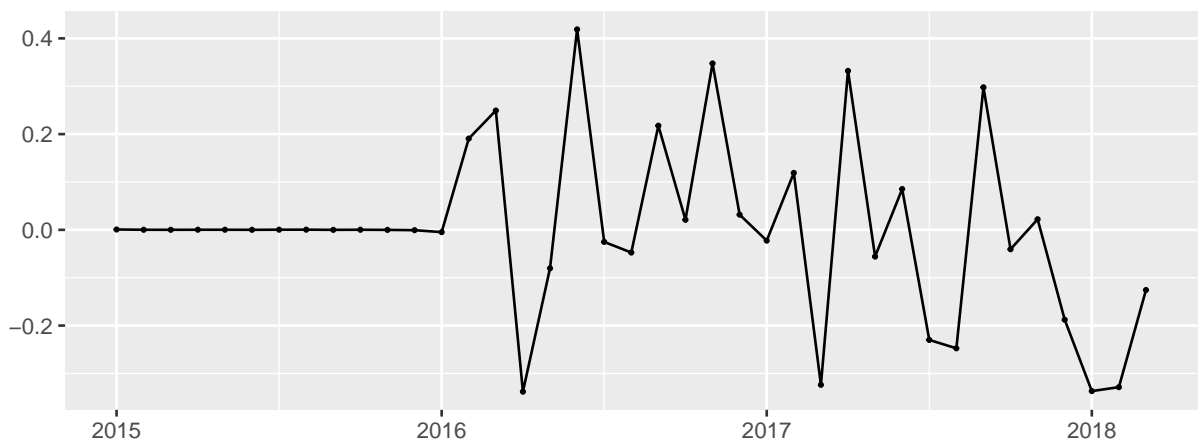
## ARIMA(1,1,2)(0,1,1)[12] : Inf
## ARIMA(1,1,2)(1,1,0)[12] : Inf
## ARIMA(1,1,2)(1,1,1)[12] : Inf
## ARIMA(1,1,3)(0,1,0)[12] : 11.89595
## ARIMA(1,1,3)(0,1,1)[12] : Inf
## ARIMA(1,1,3)(1,1,0)[12] : 7.010278
## ARIMA(1,1,4)(0,1,0)[12] : Inf
## ARIMA(2,1,0)(0,1,0)[12] : 6.613931
## ARIMA(2,1,0)(0,1,1)[12] : Inf
## ARIMA(2,1,0)(1,1,0)[12] : 1.897622
## ARIMA(2,1,0)(1,1,1)[12] : Inf
## ARIMA(2,1,1)(0,1,0)[12] : 8.612471
## ARIMA(2,1,1)(0,1,1)[12] : Inf
## ARIMA(2,1,1)(1,1,0)[12] : Inf
## ARIMA(2,1,1)(1,1,1)[12] : Inf
## ARIMA(2,1,2)(0,1,0)[12] : Inf
## ARIMA(2,1,2)(0,1,1)[12] : Inf
## ARIMA(2,1,2)(1,1,0)[12] : Inf
## ARIMA(2,1,3)(0,1,0)[12] : 14.56648
## ARIMA(3,1,0)(0,1,0)[12] : 9.038767
## ARIMA(3,1,0)(0,1,1)[12] : Inf
## ARIMA(3,1,0)(1,1,0)[12] : Inf
## ARIMA(3,1,0)(1,1,1)[12] : Inf
## ARIMA(3,1,1)(0,1,0)[12] : 11.40005
## ARIMA(3,1,1)(0,1,1)[12] : Inf
## ARIMA(3,1,1)(1,1,0)[12] : Inf
## ARIMA(3,1,2)(0,1,0)[12] : Inf
## ARIMA(4,1,0)(0,1,0)[12] : 10.26243
## ARIMA(4,1,0)(0,1,1)[12] : Inf
## ARIMA(4,1,0)(1,1,0)[12] : Inf
## ARIMA(4,1,1)(0,1,0)[12] : Inf
## ARIMA(5,1,0)(0,1,0)[12] : 12.89994
##
##
##
## Best model: ARIMA(0,1,0)(1,1,0)[12]

## Series: conv_ts
## ARIMA(0,1,1)(1,1,0)[12]
##
## Coefficients:
##          ma1          sar1
##      -0.5852   -0.5753
## s.e.    0.1547    0.1753
##

```

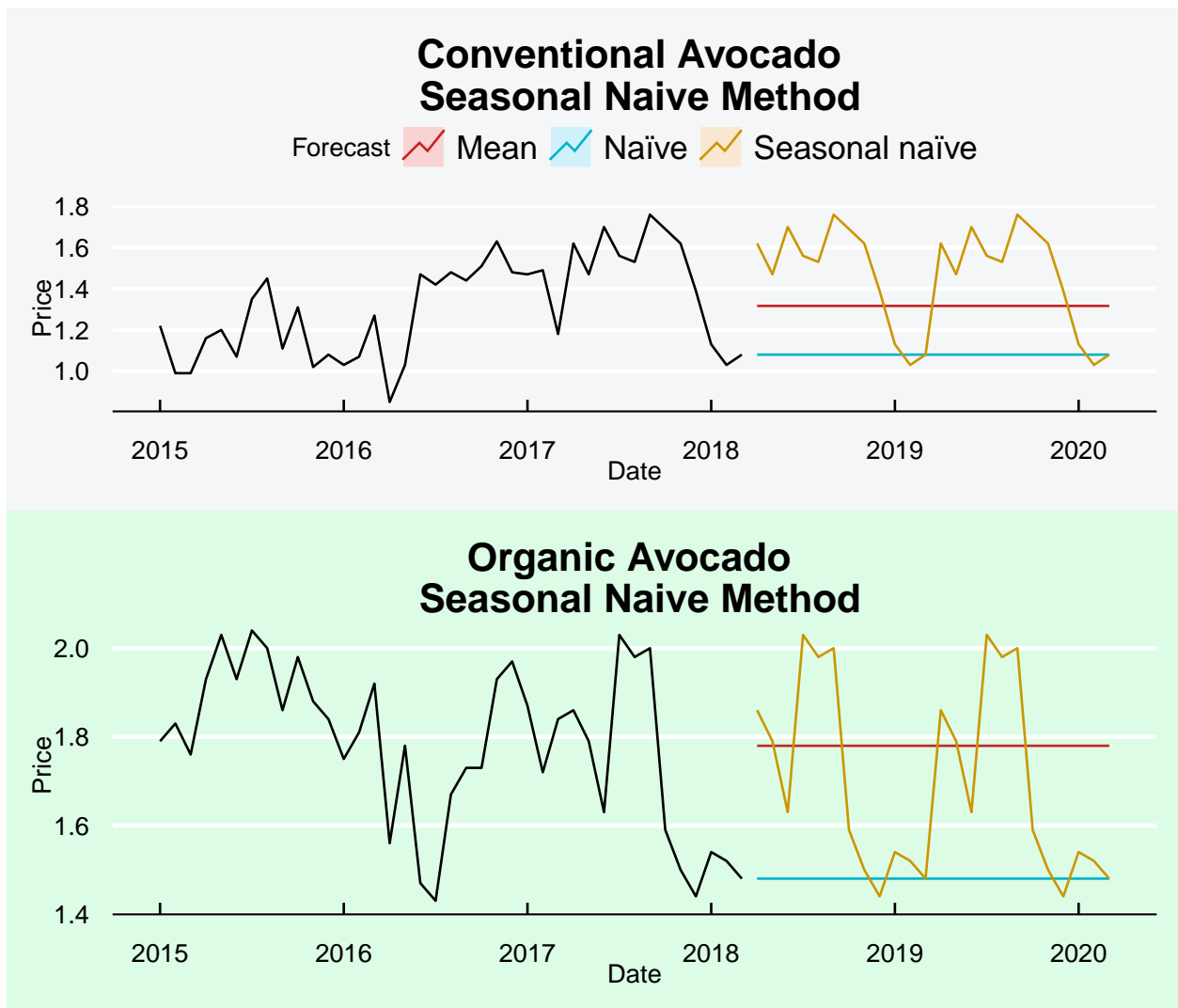
```
## sigma^2 = 0.05354: log likelihood = -0.42
## AIC=6.84 AICc=7.93 BIC=10.61
##
## Training set error measures:
##               ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -0.001571837 0.181512 0.1213104 -1.236511 9.331712 0.4580951
##               ACF1
## Training set -0.04030292
```

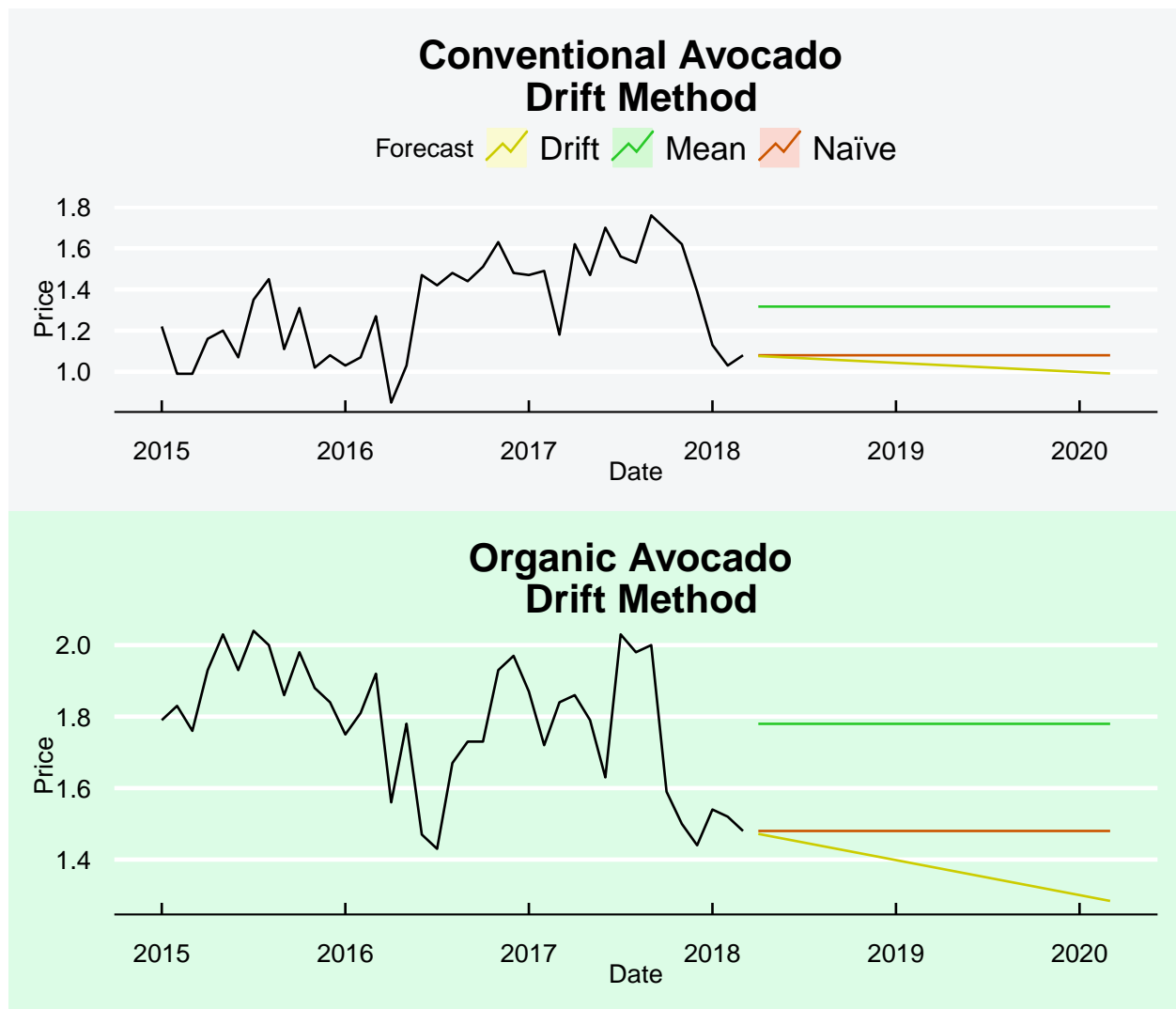
Residuals from ARIMA(0,1,1)(1,1,0)[12]



```
##
## Ljung-Box test
##
## data: Residuals from ARIMA(0,1,1)(1,1,0)[12]
## Q* = 5.909, df = 6, p-value = 0.4335
##
## Model df: 2. Total lags used: 8
```

NULL

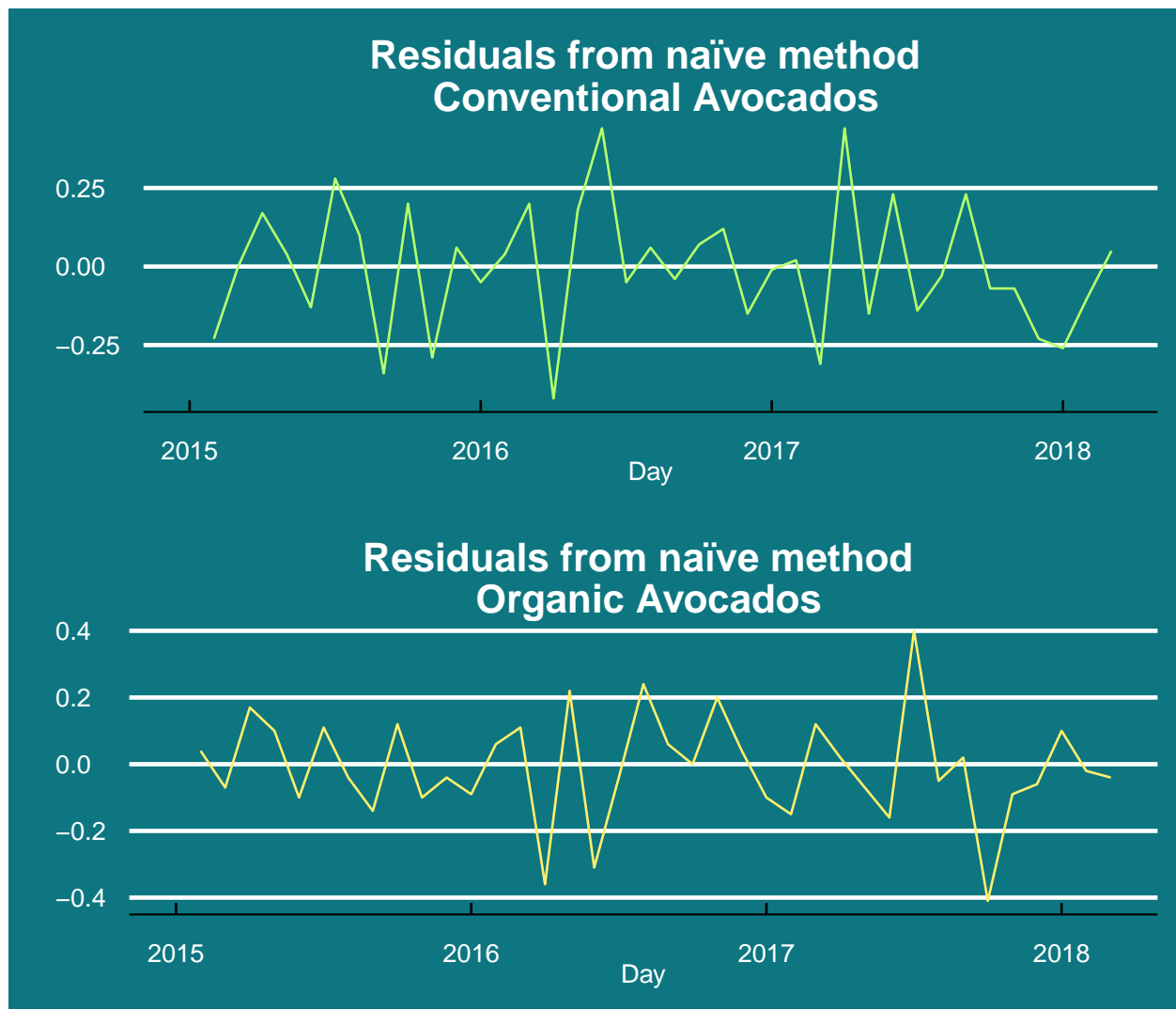




Understanding Residuals

Residuals is not only used to understand the difference between our forecast model and actual values, it also tells us some potential for “abnormal” movements during certain period. Let’s explore the charts below for both conventional and organic avocados. Here are my findings:

- **Conventional Avocados:** There are two abnormal peaks in the conventional avocados chart. This might indicate abnormal movements in prices that could indicate that this was due to some extraordinary event.
- **Organic Avocados:** We see some peaks in organic avocados but not to the extent of conventional avocados. Nevertheless, there are some abnormal peaks that should be further evaluated.

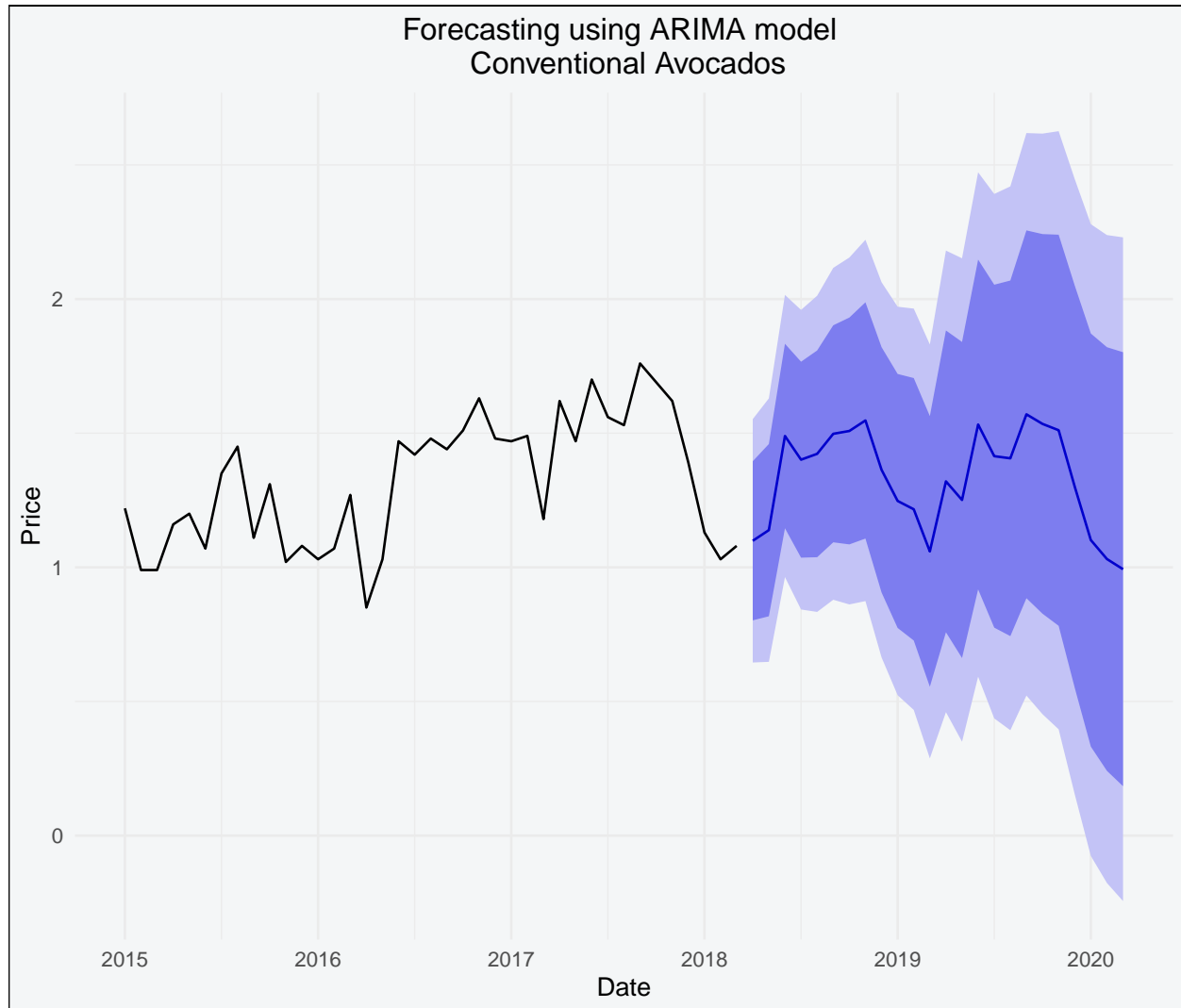


Understanding ARIMA Model:

- **Auto Regressive (AR):** Means that past time points could have a certain degree of current and future time observations. The ARIMA model takes into account lagged observations in order to come up with forecast observations. A weight is added to past observations however, the weight can vary on how recent the past observations are. The more recent, the more weight is added to the most recent past observation.
- **Integrated (I):** If there are consistent trends in the movement of past prices, it is most likely to be non-stationary meaning that seasonality persists in past movement of prices. Integrated removes the seasonality phase of our dataset in case there are consistent patterns that show that this is the case. The degree of differencing available in ARIMA models eliminates the seasonality trend issue.
- **Moving Average (MA):** Moving average helps remove the effect of random movements of avocado prices in our case. If there was an extraordinary event that led to a surge in avocado prices, moving average will help us “smooth” things up and our time

series model will not be prone to these fluctuations.

```
## [1] 0.2313871
```



```
##
## Forecast method: ARIMA(0,1,1)(1,1,0)[12]
##
## Model Information:
## Series: conv_ts
## ARIMA(0,1,1)(1,1,0)[12]
##
## Coefficients:
##          ma1      sar1
##      -0.5852  -0.5753
## s.e.   0.1547   0.1753
```

```
##
## sigma^2 = 0.05354: log likelihood = -0.42
## AIC=6.84 AICc=7.93 BIC=10.61
##
## Error measures:
##           ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -0.001571837 0.181512 0.1213104 -1.236511 9.331712 0.4580951
##           ACF1
## Training set -0.04030292
##
## Forecasts:
##           Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
## Apr 2018      1.0988344 0.8023047 1.395364 0.64533132 1.552337
## May 2018      1.1386782 0.8176499 1.459706 0.64770779 1.629649
## Jun 2018      1.4894879 1.1457024 1.833273 0.96371337 2.015262
## Jul 2018      1.4012635 1.0361364 1.766391 0.84284987 1.959677
## Aug 2018      1.4230391 1.0377508 1.808327 0.83379152 2.012287
## Sep 2018      1.4977123 1.0932666 1.902158 0.87916596 2.116259
## Oct 2018      1.5082521 1.0855162 1.930988 0.86173336 2.154771
## Nov 2018      1.5475561 1.1072892 1.987823 0.87422604 2.220886
## Dec 2018      1.3635789 0.9064528 1.820705 0.66446492 2.062693
## Jan 2019      1.2473999 0.7740148 1.720785 0.52341981 1.971380
## Feb 2019      1.2164340 0.7273300 1.705538 0.46841391 1.964454
## Mar 2019      1.0593317 0.5549984 1.563665 0.28802046 1.830643
## Apr 2019      1.3204560 0.7580290 1.882883 0.46029805 2.180614
## May 2019      1.2510857 0.6619899 1.840182 0.35014138 2.152030
## Jun 2019      1.5323955 0.9177870 2.147004 0.59243296 2.472358
## Jul 2019      1.4143854 0.7752819 2.053489 0.43696101 2.391810
## Aug 2019      1.4063753 0.7436816 2.069069 0.39287280 2.419878
## Sep 2019      1.5704056 0.8849331 2.255878 0.52206585 2.618745
## Oct 2019      1.5346121 0.8270937 2.242130 0.45255616 2.616668
## Nov 2019      1.5110352 0.7821375 2.239933 0.39628237 2.625788
## Dec 2019      1.3005818 0.5509142 2.050249 0.15406412 2.447099
## Jan 2020      1.1016649 0.3317875 1.871542 -0.07576099 2.279091
## Feb 2020      1.0309847 0.2414147 1.820555 -0.17655844 2.238528
## Mar 2020      0.9930251 0.1842418 1.801808 -0.24390228 2.229953
```

Conclusion

When it comes to the analysis we have implemented there are several observations we have observed:

- **Expensive Organic Avocados:** As expected, we have noticed that organic avocados are much more expensive than conventional avocados.

- **Similar patterns by type:** There are some distinguishable patterns between the two types of avocados however, we see that most patterns are similar between these two types of avocados.
- **Year 2017:** The year of 2017 was the best year of avocados. One of the factors that could have contributed was the well being of the economy as a whole but there could also be other hidden factors that could affect avocado prices in the market.
- **Volatility:** Although, 2017 was the year with the highest prices, it was also the year with the highest volatility. This means that price fluctuations were higher in this year as oppose to the past two years.
- **Buy avocados before fall!:** We see a consistent pattern of avocado prices increasing when fall is coming. This applies to both conventional and organic avocados.
- **Downward trend in the long run:** Based on our ARIMA model, for both types of avocados we expect a downward trend in prices at least in the long run. There could be some upward pressure in the short run however, our model predicts that both avocado prices will go downwards 2 years from now.