College Students on a Budget for Avocados: Forecasting Avocado Average Price in the U.S.

Authors: Gregory King, Andrew Muliadi, Sunny Shen

ABSTRACT

Most college students live on a budget, where they apportion the money they have earned or have been given to various living necessities. One such necessity would have to be food, and specifically, the avocado. As such, the purpose of this study is to predict how avocado prices change to allow students to budget their grocery spending accordingly. Our econometric model concludes that avocado prices are highly predictive based on prices in previous months and seasonal trends, while the predictive power of types of avocados is not statistically significant.

INTRODUCTION

In general, most college students are living on a budget, with most of their attention and energy directed towards learning rather than a paying job. In general, the budget that most students have is used for living necessities such as paying rent, tuition, and groceries. For our paper, we focus specifically on the aspect of how much students should budget for their groceries on a monthly basis, further dialing in our study by discussing the approximate amount students should set aside to purchase avocados.

We decided to focus our study on avocados because there is a mutual love for the green berry (as defined by botanist). Given their relatively short shelf life as well as ability to spoil quickly, we believe it's important that purchases of this fruit are done strategically. As a result, students will be able to enjoy their avocados at their own leisure free from worry that they might have to sacrifice something else in their budget.

As such the overall goal of this paper is to effectively predict the prices of avocados in both the short and long term. (e.g. prices in future months and years). In addition to time, we would also like to account for cyclical and seasonal patterns to further visualize and predict price trends.

Our research draws inspiration from other prominent studies on the prices of avocados. For example, one piece we reference is "Forecasting Price Trends in the U.S. Avocado Market" by Edward Evans and Sikavas Nalampang. With changes in disposable income, aggressive promotion towards leading a healthier life, and changes in the supply of avocados, Evans and Nalampang were able to create a predictive pricing model that took into account these qualitative and quantitative factors that was surprisingly robust. Taking this into account, we have expanded our research to follow the suit of Evans and Nalampang to include various qualitative and quantitative variables to predict prices through the data set we procured and cleaned.

MATERIAL AND METHODS

We utilize this model framework to examine the average sales price of avocados. We build off of traditional demand analysis and utilize the inverse demand analysis to be able to predict average avocado pricing off of quantity demanded along with previous period avocado pricing. With Marshallian demand, we can use total volume sold at a given price in previous periods to see market clearing conditions (Huang 1988). We expect to see a negative relationship with the effects of total volume sold on average avocado prices as more people, such as college students on a budget, are willing to pay a lower price for avocados. On the other hand, we expect prices of previous time periods of our model to have a positive impact on the average price of avocados as demand for avocados has increased due to popularity over the years for a healthy option to consume. Our model aimed to predict the average price of avocados with the regression function written as:

$$\widehat{Avg\ P} = \beta_0 + \beta_1 Q + \beta_2 P_{t-1} + \beta_3 P_{t-2} + \textit{Entity Fixed Effects} + \textit{Time Fixed Effects}. \qquad (1)$$

The econometric specification of the average price forecasting model is as follows below:

$$Avg\_P = \beta_0 + \beta_1 Q + \beta_2 P_{t-1} + \beta_3 P_{t-2} + \mu_2 D2_i + ... + \mu_n DT_i + \delta_2 B2_t + u_{t,c} \qquad (2)$$

where

β_0	=	Intercept
Q	=	Quantity Sold
P_{t-1}	=	Lagged prices one month before year-month combo t
P_{t-2}	=	Lagged prices two months before year-month combo t
DT_i	=	Year-month fixed effects represented by {0, 1} indicator variables
$B2_t$	=	Avocado type fixed effects represented by {0, 1} indicator variables
$u_{t,c}$	=	Error term
c	=	Avocado type c subscript
t	=	Year-month combo t subscript.

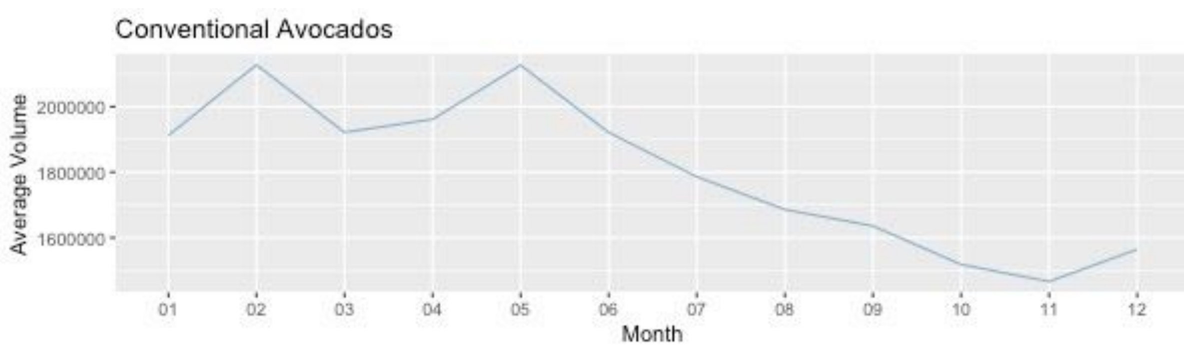
The estimated parameter of β_1 measures the association of the effects of Q on average avocado price (Avg_P) holding all else constant, β_2 measures the association of the effects of P_{t-1} on Avg_P holding all else constant, and β_3 measures the association of the effects of P_{t-2} on Avg_P holding all else constant. As for μ_2 - μ_n, these are the measures of the entity fixed effects while δ_2 measures the time fixed effect.

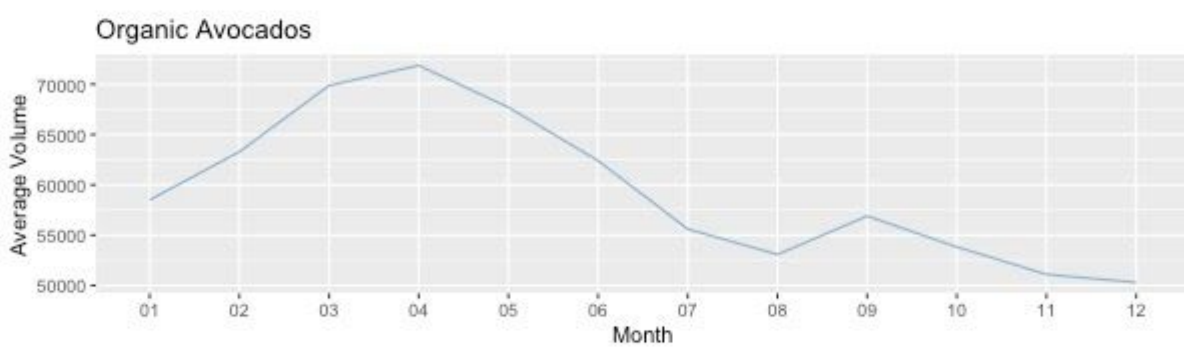
Ordinary least squares (OLS) regression was used to estimate the coefficients of the average avocado price of the forecasting model from 2015 to mid-2020 using RStudio 1.3.959®. The validity of the parameter estimates were then checked by looking at the residuals of the predicted average avocado prices through the given time period to look for heteroskedasticity. Heteroskedasticity is when the variances of the data points are not constant throughout the set of points under investigation. With heteroskedasticity present, this would cause the OLS estimates' standard errors to be inefficient and thus cause the hypothesis tests and confidence intervals to be incorrect. To check for heteroskedasticity, the Breusch-Pagan Test was used on our estimates.

The data on U.S. avocado prices between 2015 to mid-2020 comes from the Hass Avocado Board. The important variables in this dataset that were used were: Date, Average Price, Type, and Total Volume. This panel data of the average avocado prices led to the opportunity to obtain the year-month pair by parsing through the Date variable. Additionally, the Average Price column was used to get the lagged prices of avocados from a month and two months prior. The Type variable was used to see time fixed effects on the regression. Lastly, the Total Volume variable was used directly in the model as its own variable.

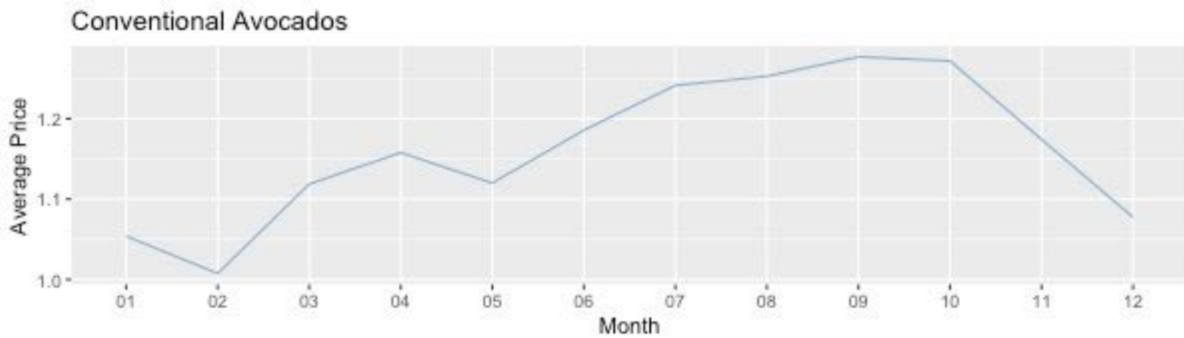
Before working on the regression model, exploratory data analysis was done on the Hass Avocado Board's dataset to see some trends in the data. First off, we plotted a time series of the average prices over the five year timespan that we have access to from the data for both conventional avocados and organic avocados.

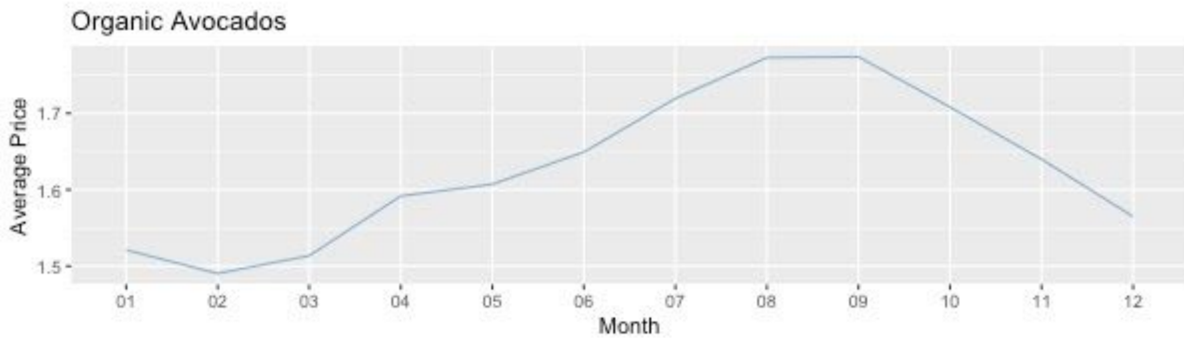
Next, we took a look at the average volume of avocados sold across the months over the five year timespan to us to see when there is the most avocados sold during the year to identify any seasonal patterns in the purchase of avocados. We see seasonal patterns in the average quantity sold for avocados, so we would like to add time fixed-effect in our econometric model.



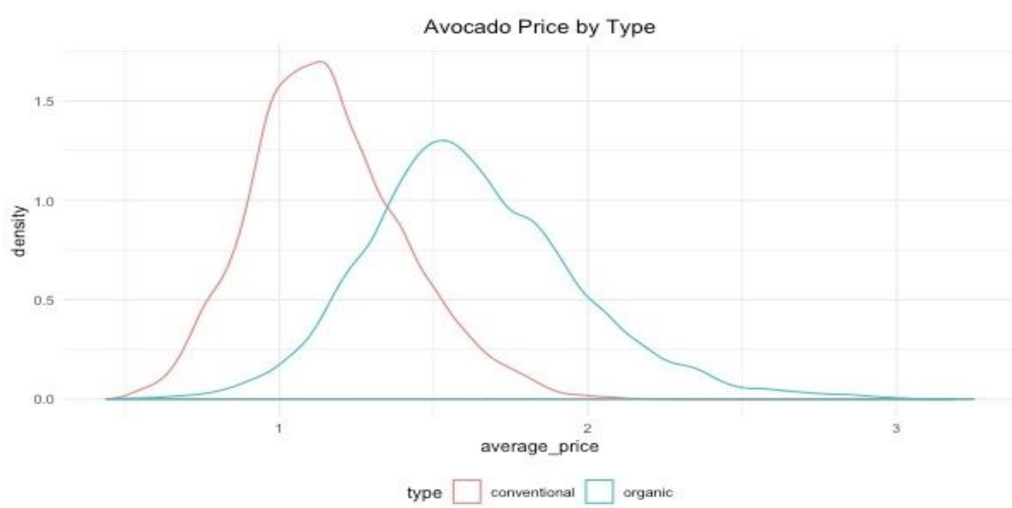


We then observed average prices of avocados across every month of the year, over the five years, to see which months average avocado prices spiked in the time series. We see seasonal patterns in the average price for avocados, so this further motivates adding time fixed-effect in our econometric model.





Lastly, we checked to see if average prices differed for conventional versus organic prices. The original assumption was that organic avocados would be more expensive, on average, as they are more labor intensive and there is a lower supply of organic avocados relative to conventional avocados produced. From the graph below, our assumptions were confirmed as the peak of the density curve for organic avocados is slightly higher than the peak of the density curve for the conventional avocado. Therefore, we would like to add entity fixed-effect to our econometric model.



RESULTS AND DISCUSSIONS

Regression Results

	Dependent variable:					
	Average Price					
	OLS	OLS	OLS	FE	FE	FE
	(1)	(2)	(3)	(4)	(5)	(6)
Quantity	-0.00000***	-0.00000***	-0.00000***	-0.00000**	-0.00000**	-0.000
	(0.000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Average price(t-1)		0.742***	1.017***	0.616***	1.017***	0.568***
		(0.057)	(0.084)	(0.124)	(0.085)	(0.130)
Average price(t-2)			-0.349***	0.220*	-0.356***	0.144
			(0.082)	(0.120)	(0.084)	(0.135)
Constant	1.634***	0.423***	0.542***			
	(0.018)	(0.094)	(0.093)			
Type Dummies	No	No	No	No	Yes	Yes
Year-Month Dummies	No	No	No	Yes	No	Yes
RSS	2.574	1.089	0.95	0.144	0.947	0.141
Observations	126	126	126	126	126	126
R^2	0.729	0.885	0.900	0.979	0.656	0.434

Note: *Significant at the 10% Level
**Significant at the 5% Level
***Significant at the 1% Level

The results for the econometric estimation are shown in the table on page 5. Our model 1 regresses avocado price on total quantity; model 2 regresses avocado price on total quantity and price from t-1, which is the previous month; model 3 regress average price on total quantity, price from t-1, and price from t-2, which is two months prior; model 4 includes time fixed effect on the top of model 3; model 5 includes entity fixed effects in addition to time fixed effects added to model 3.

We perform the following hypothesis tests to see if adding fixed effects is necessary:

1. Compare model 3 and model 4

- H_0: Introducing time fixed effects does not change the association between price of avocado and quantity sold, price from the previous month, and price from 2 months ago.
- H_1: Introducing time fixed effects changes the association between price of avocado and quantity sold, price from the previous month, and price from 2 months ago.
- Results: our F test gives us a p-value of 2.879e-10, which is statistically significant at 1% level and therefore we reject the null hypothesis. We conclude that adding time fixed effects changes the relationship between price of avocado and quantity sold, price from the previous month, and price from 2 months ago.
- Discussion: This is expected based on our exploratory data analysis. We expect to see seasonal changes of avocado prices and also changes across different years.

2. Compare model 4 and model 5

- H_0: the association between price of avocado and quantity sold, price from the previous month, and price from 2 months ago does not change when we add entity fixed effects in addition to time fixed effects.
- H_1: the association between price of avocado and quantity sold, price from the previous month, and price from 2 months ago does not change when we add entity fixed effects in addition to time fixed effects.
- Result: our F test gives us a p-value of 0.2342, which is not statistically significant 1%, 5%, or 10% level. Therefore, we conclude that adding entity fixed effects in addition to time fixed effects does not change the relationship between price of avocado and quantity sold, price from the previous month, and price from 2 months ago.
- Discussion: This is not expected based on our exploratory data analysis because we see two slightly different distributions of conventional vs. organic avocado, and organic avocado tend to have higher prices. However, our hypothesis testing concludes that although there is a difference, it is not statistically significant.

Based on the analysis above, time fixed effect model (model 4) is the most appropriate model for our purpose as the second hypothesis testing suggests that it is not necessary to add entity fixed effects to the model. Our time fixed-effect model has an R^2 of 97.9%, which means

97.9% of the variation in the price of avocado is explained by the variables selected when fixing the time that the price is recorded. The coefficient estimate of avocados price from the previous month is statistically significant at 1% level, and the coefficient estimate of price of avocados from 2 months ago is statistically significant at 10%. The estimated coefficient of 0.616 for price from the previous month indicates that if the price of avocado in the previous month increases by $1, the current price of avocado is expected to increase by $0.616. The estimated coefficient of 0.220 for price from 2 months ago indicates that if the price of avocado 2 months ago increases by $1, the price of avocado is expected to increase by $0.220. Interestingly, coefficients for quantity sold is approximately 0, which means that quantity sold for avocados is not helpful in predicting avocado prices.

We performed RESET (Regression Specification Error Test), which checks whether there is a nonlinear relationship between our independent variable (quantity sold, price from t-1, and price from t-2) and our dependent variable average price.

$$\widehat{Avg\,P} = \beta_0 + \beta_1 Q + \beta_2 P_{t-1} + \beta_3 P_{t-2} + \mu_2 D2_i + \ldots + \mu_n DT_i + \delta_2 B2_t + u_{t,c} + \lambda_1 \widehat{Avg\,P}^2 + \lambda_2 \widehat{Avg\,P}^3 \tag{3}$$

- H_0: $\lambda_1 = \lambda_2 = 0$. In other words, there is no nonlinear relationship between avocado prices and our independent variables.
- H_1: at least one $\lambda_j \neq 0$, for j = 1, 2. In order words, there exists some nonlinear relationships between avocado prices and our independent variables.
- Results: We got a p-value of 0.6138, which is not statistically significant, and therefore we fail to reject the null hypothesis that there is a nonlinear relationship between average price and our dependent variables.

Furthermore, We performed the Breusch-Pagan Test to check for heteroskedasticity.

- H_0: The data is homoscedastic so the variance of residuals is the same across different explanatory variables.
- H_1: The data is heteroscedastic so the variance of residuals is not the same across different explanatory variables.
- Results: We got a p-value of 0.4755 > 0.1, which is not statistically significant, and therefore we fail to reject the null hypothesis of homoscedasticity.

SUMMARY AND CONCLUSIONS

In summary, this paper intends to construct an econometric model to predict the price of avocados. We obtained the data on U.S. avocado prices between 2015 to mid-2020 from the Hass Avocado Board, and we would like to examine the relationship between avocado prices and quantity sold, time, and prices from previous months. Based on our exploratory data analysis, we noticed seasonal variation of average avocado prices, different average prices across years, and different price distribution between conventional and organic avocados. Therefore, we include quantity sold, price from the previous month, price from 2 months ago, and both entity fixed effects based on avocado types and time fixed effects based on year-month combination in our econometric model.

Based on our model outputs, we concluded that the average prices of avocados can be predicted by the increase in average prices of avocados from one and two months prior. As such, we recommend that students keep in mind what the prices of avocados are today to predict what prices will be on average in the future. We also observe a seasonal trend that avocados tend to be more expensive during summer months and tend to peak around September or October. In our regression, introducing time fixed effects is statistically significant and changes the relationship between avocado prices and our independent variables. However, although we observe a difference in the price distribution of conventional vs organic avocados, our regression results suggest that the entity-fixed effect based on avocado type is not statistically significant. This model will provide a baseline for the amount student's will need to save in advance to ensure they can always afford as much avocado as they desire.

Interestingly, our regression results suggest that quantity sold is not helpful in predicting avocado prices because we have a coefficient of approximately 0 in our regression. This is probably because the data we obtained represents the market equilibrium prices and quantities rather than the prices and quantities on the demand/supply curve. Future research may consider introducing instrumental variables that will shift the supply curve so that we may be able to regress the demand function for avocados.

REFERENCES

Evans, Edward A. & Nalampang, Sikavas, 2009. "Forecasting Price Trends in the U.S. Avocado (Persea americana Mill.) Market," Journal of Food Distribution Research, Food Distribution Research Society, vol. 40(2), pages 1-10, July.

Huang, K. (1988). An Inverse Demand System for U.S. Composite Foods. American Journal of Agricultural Economics, 70(4), 902-909. doi:10.2307/1241932

In [3]:
```
library(tidyverse)
library(dplyr)
library(ggplot2)
library(lubridate)
library(lmtest)
library(stargazer)
library(plm)
library(cowplot)
```

In [4]:
```
avocado = read.csv('avocado.csv')
```

In [5]:
```
tail(avocado)
```

	date	average_price	total_volume	X4046	X4225	X4770	total_bags	small_b
30016	2020-05-17	1.16	51690121	15951219.7	9221698.7	728025.52	25788840.1	168964
30017	2020-05-17	1.58	2271254	150100.0	198457.0	5429.00	1917250.0	11216
30018	2020-05-17	1.09	8667913	2081824.0	1020965.1	33410.85	5531562.9	25808
30019	2020-05-17	1.71	384158	23455.0	39738.0	1034.00	319932.0	1300
30020	2020-05-17	0.89	1240709	430203.1	126497.3	21104.42	662904.2	3959
30021	2020-05-17	1.58	36881	1147.0	1243.0	2645.00	31846.0	256

EDA

In [6]:
```
avocado$year = as.factor(avocado$year)
avocado$date = as.Date(avocado$date)
avocado$month  = factor(months(avocado$date), levels = month.name)
```

1. Avocado Price by Type

In [16]:
```
jpeg(file="price_by_type.jpeg", width=600, height=350)

options(repr.plot.width=6, repr.plot.height=3)

ggplot(avocado, aes(x=average_price, group=type, color=type)) + geom_den
sity()+ theme_minimal() +
theme(plot.title=element_text(hjust=0.5), legend.position="bottom") + la
bs(title="Avocado Price by Type")
dev.off()
```

pdf: 2

In [8]:
```
seasonal_avo = avocado
seasonal_avo$month_year <- format(as.Date(avocado$date), "%Y-%m")
seasonal_avo$month <- format(as.Date(avocado$date), "%m")
seasonal_avo$year <- format(as.Date(avocado$date), "%Y")
```

2. Avocado Price by Month

In [15]:
```
jpeg(file="price_by_month.jpeg", width=600, height=350)

options(repr.plot.width=6, repr.plot.height=5)

conv_patterns <- seasonal_avo %>% select(month, average_price, type) %>%
filter(type == "conventional") %>%
group_by(month) %>% summarize(avg=mean(average_price)) %>%
ggplot(aes(x=month, y=avg)) + geom_line(group=1, color="#7FB3D5") +
labs(title="Conventional Avocados", x="Month", y="Average Price")

org_patterns <- seasonal_avo %>% select(month, average_price, type) %>%
filter(type == "organic") %>%
group_by(month) %>% summarize(avg=mean(average_price)) %>%
ggplot(aes(x=month, y=avg)) + geom_line(group=1, color="#7FB3D5") +
labs(title="Organic Avocados", x="Month", y="Average Price")

plot_grid(conv_patterns, org_patterns, nrow=2)

dev.off()
```

pdf: 2

3. Avocado Price by Year

In [17]:
```
jpeg(file="price_by_year.jpeg", width=600, height=350)

options(repr.plot.width=6, repr.plot.height=5)

conv_patterns <- seasonal_avo %>% select(year, average_price, type) %>%
filter(type == "conventional") %>%
group_by(year) %>% summarize(avg=mean(average_price)) %>%
ggplot(aes(x=year, y=avg)) + geom_line(group=1, color="#7FB3D5") +
labs(title="Conventional Avocados", x="year", y="Average Price")

org_patterns <- seasonal_avo %>% select(year, average_price, type) %>% f
ilter(type == "organic") %>%
group_by(year) %>% summarize(avg=mean(average_price)) %>%
ggplot(aes(x=year, y=avg)) + geom_line(group=1, color="#7FB3D5") +
labs(title="Organic Avocados", x="year", y="Average Price")

plot_grid(conv_patterns, org_patterns, nrow=2)

dev.off()
```

pdf: 2

4. Quantities by Month

In [18]:
```
jpeg(file="quantities_by_month.jpeg", width=600, height=350)

options(repr.plot.width=6, repr.plot.height=5)

conv_patterns_vol <- seasonal_avo %>% select(month, total_volume, type)
%>% filter(type == "conventional") %>%
group_by(month) %>% summarize(avg=mean(total_volume)) %>%
ggplot(aes(x=month, y=avg)) + geom_line(group=1, color="#7FB3D5") +
labs(title="Conventional Avocados", x="Month", y="Average Volume")

org_patterns_vol <- seasonal_avo %>% select(month, total_volume, type) %
>% filter(type == "organic") %>%
group_by(month) %>% summarize(avg=mean(total_volume)) %>%
ggplot(aes(x=month, y=avg)) + geom_line(group=1, color="#7FB3D5") +
labs(title="Organic Avocados", x="Month", y="Average Volume")

plot_grid(conv_patterns_vol, org_patterns_vol, nrow=2)
dev.off()
```

pdf: 2

Modeling

In [20]:
```
seasonal_avo = seasonal_avo %>% select('date', 'average_price', 'total_v
olume', 'type', 'year', 'geography',
                                     #'average_price_lag', 'average_pric
e_lag_2',
                                      'month_year', 'month')
head(seasonal_avo)
```

date	average_price	total_volume	type	year	geography	month_year	month
2015-01-04	1.22	40873.28	conventional	2015	Albany	2015-01	01
2015-01-04	1.79	1373.95	organic	2015	Albany	2015-01	01
2015-01-04	1.00	435021.49	conventional	2015	Atlanta	2015-01	01
2015-01-04	1.76	3846.69	organic	2015	Atlanta	2015-01	01
2015-01-04	1.08	788025.06	conventional	2015	Baltimore/Washington	2015-01	01
2015-01-04	1.29	19137.28	organic	2015	Baltimore/Washington	2015-01	01

In [21]:
```
grouped = seasonal_avo %>%
    group_by(type, month_year) %>%
    summarise(average_price = mean(average_price),
              total_volume=mean(total_volume))  %>%
    group_by(type) %>%
    mutate(average_price_lag = dplyr::lag(average_price, n = 1, default
= NA)) %>%
    group_by(type) %>%
    mutate(average_price_lag_2 = dplyr::lag(average_price_lag, n = 1, de
fault = NA)) %>% ungroup()
# head(grouped)
```

In [23]:
```
pdata = na.omit(pdata.frame(grouped, index = c("type", "month_year"), st
ringsAsFactors = FALSE))
```

In [24]:
```
model1 = lm(average_price ~ total_volume, pdata)
model2 = lm(average_price ~ total_volume + average_price_lag, pdata)
model3 = lm(average_price ~ total_volume + average_price_lag + average_p
rice_lag_2, pdata)
model4 = plm(average_price ~ total_volume + average_price_lag+ average_p
rice_lag_2, model='within', effect='time', pdata)
model5 = plm(average_price ~ total_volume + average_price_lag+ average_p
rice_lag_2, model='within', effect='twoways', pdata)
```

In [26]:
```
#compare different lm without fixed effects
# stargazer(model1, model2, model3, type='text')
```

In [27]:
```
#compare the fixed effect models
# stargazer(model4, model5, type='text')
```

1 Check nonlinearity - no nonlinearity found

In [28]:
```
resettest(average_price ~ total_volume , power=2:11, type= c("fitted"), pdata)
```

```
	RESET test

data:  average_price ~ total_volume
RESET = 1.8188, df1 = 10, df2 = 114, p-value = 0.06489
```

In [29]:
```
resettest(average_price ~ total_volume +average_price_lag , power=2:3, type= c("fitted"), pdata)
```

```
	RESET test

data:  average_price ~ total_volume + average_price_lag
RESET = 0.28792, df1 = 2, df2 = 121, p-value = 0.7503
```

In [30]:
```
resettest(average_price ~total_volume +average_price_lag +average_price_lag_2, power=2:4, type= c("fitted"), pdata)
```

```
	RESET test

data:  average_price ~ total_volume + average_price_lag + average_price
_lag_2
RESET = 0.60385, df1 = 3, df2 = 119, p-value = 0.6138
```

2. Test heteroskedasticity

In [33]:
```
residuals = model5$residuals
```

In [34]:
```
pdata$residualis = residuals
```

In [36]:
```
bptest(model4)
```

```
	studentized Breusch-Pagan test

data:  model4
BP = 2.4988, df = 3, p-value = 0.4755
```

In [37]:
```
bptest(model5)
```

```
	studentized Breusch-Pagan test

data:  model5
BP = 2.4988, df = 3, p-value = 0.4755
```

In [39]:
```
coeftest(model4, vcov.=vcovHC, type='HC4')
```

```
t test of coefficients:

                       Estimate  Std. Error t value  Pr(>|t|)    
total_volume        -4.1042e-08  5.3189e-10 -77.162 < 2.2e-16 ***
average_price_lag    6.1596e-01  1.1421e-03 539.322 < 2.2e-16 ***
average_price_lag_2  2.2001e-01  2.5844e-04 851.317 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

In [40]:
```
coeftest(model5, vcov.=vcovHC, type='HC4')
```

```
t test of coefficients:

                       Estimate  Std. Error t value  Pr(>|t|)    
total_volume        -1.0743e-08  8.7979e-10 -12.211 < 2.2e-16 ***
average_price_lag    5.6825e-01  2.8222e-03 201.349 < 2.2e-16 ***
average_price_lag_2  1.4383e-01  6.9877e-03  20.584 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

hypothesis testing on whether we should add fixed effects

In [41]:
```
pFtest(model4, model3)
```

```
        F test for time effects

data:  average_price ~ total_volume + average_price_lag + average_price
_lag_2
F = 5.4004, df1 = 62, df2 = 60, p-value = 2.879e-10
alternative hypothesis: significant effects
```

In [42]:
```
pFtest(model5, model4)
```

```
        F test for twoways effects

data:  average_price ~ total_volume + average_price_lag + average_price
_lag_2
F = 1.4445, df1 = 1, df2 = 59, p-value = 0.2342
alternative hypothesis: significant effects
```

In [52]:
```
stargazer(model1, model2, model3, model4, model5, model6,
          #se = list(NULL, NULL, NULL, robust_se_4, robust_se_5, robust_
se_6),
          type = "html",
          out = "Term Proj.doc",
          title = "Regression Results",
          align = TRUE,
          notes.append = FALSE,
          notes = c(" *Significant at the 10% Level",
                    " **Significant at the 5% Level",
                    " ***Significant at the 1% Level"),
          omit.stat = c("f","adj.rsq","ser"),
          add.lines = list(
                         c("Standard Errors"," "," "," ","Clustered",
"Clustered","Clustered"),
                         c("Type Dummies","No","No","No","No","Yes","Y
es"),
                         c("Year-Month Dummies","No","No","No","Yes",
"No","Yes"),
                         c("RSS",
                           round(sum(residuals(model1)^2), digits = 3
),
                           round(sum(residuals(model2)^2), digits = 3
),
                           round(sum(residuals(model3)^2), digits = 3
),
                           round(sum(residuals(model4)^2), digits = 3
),
                           round(sum(residuals(model5)^2), digits = 3
),
                           round(sum(residuals(model6)^2), digits = 3
))),
          dep.var.labels = c("Average Price"),
          covariate.labels = c("Quantity", "Average price(t-1)", "Averag
e price(t-2)"),
          column.labels = c("OLS", "OLS", "OLS", "FE", "FE", "FE"))
```