CMSE201_ExampleProject_3

November 10, 2019

0.0.1 ORGANIC VS. CONVENTIONAL PRODUCTS

0.0.2 Background and Motivation

After doing much of my own personal research, it was easy for me to choose this as my semester project. I have decided to try and answer a few questions about the differences between organic and conventional food products; such as nutritional facts, ingredients, process, and prices. The specific questions I came up with were: 1. What are the nutritional differences? 2. What are the ingredient differences? 3. What are the cost differences? And does this have an effect on how much organic food is sold?

0.0.3 Methodology

[1]:
```
%matplotlib inline #include all imports
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
```

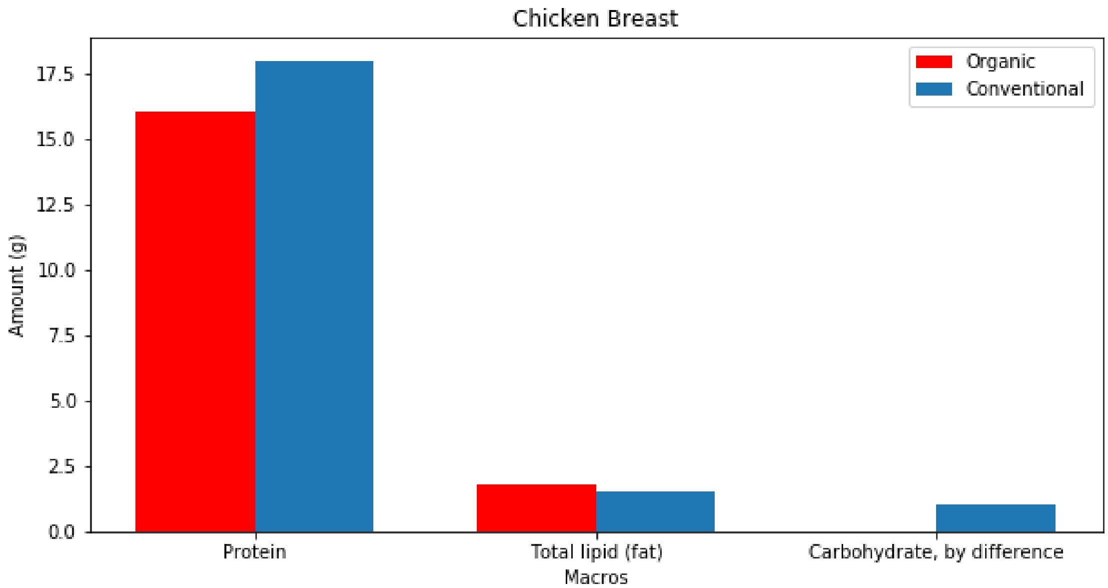
The following cells were used to create bar charts of different products that are popular items for people to choose as an organic option. The bar charts are comparing the macros of organic vs. conventional. I chose to only look at the macros because after doing a lot of research on nutrition, I have found macro counting to be more effective and more important than counting calories.

[4]:
```
organic_chicken = pd.read_csv("Organic chicken.csv") #open organic chicken file␣
↪that contains its nutritional facts
chicken = pd.read_csv("Chicken.csv") #do the same for regular chicken
```

[5]:
```
#going to find all the nutritional info for macros (protein,fats,carbs) because␣
↪this is what I look for on labels

organic_protein = organic_chicken.loc[2] #get row that contains organic chicken␣
↪protein info
the_organic_protein = organic_protein[-1] #get specific protein value

#These two steps are repeated for the organic chicken's fat and carb info

organic_fat = organic_chicken.loc[3]
the_organic_fat = organic_fat[-1]
```

```
organic_carbs = organic_chicken.loc[4]
the_organic_carbs = organic_carbs[-1]

protein = chicken.loc[2] #get regular chicken protein info & repeat for fats␣
 ↪and carbs
fat = chicken.loc[3]
carbs = chicken.loc[4]

nutrients = [organic_protein[0], organic_fat[0], organic_carbs[0]]
#this creates a list containing the words "protein","fat", "carbs" in order to␣
 ↪use it for the ticks of the x axis

link = [the_organic_protein, the_organic_fat, the_organic_carbs] #put the␣
 ↪organic values in a list
link2 = [protein[1], fat[1], carbs[1]] #put the regular values in a list
```

[6]:
```
plt.figure(figsize=(10,5)) #set figure size
n_groups = 3 #there are three different groups of bars that will be on the␣
 ↪chart
index = np.arange(n_groups) #the index and width came in handy when I had to␣
 ↪put the bars side by side
width = 0.35

plt.bar(index, link, width, label="Organic", color='r') #graph the organic info
plt.bar(index + width, link2, width, label = "Conventional") #graph the␣
 ↪conventional info

plt.xticks(index + width / 2, nutrients)
plt.xlabel("Macros")
plt.ylabel("Amount (g)")
plt.title("Chicken Breast")
plt.legend()
plt.show()

###### THIS PROCESS WAS REPREATED 2 MORE TIMES FOR THE FOLLOWING ITEMS: ICE␣
 ↪CREAM SANDWICH & EGG ######
```



The next following cells are printing out the ingredient lists of organic and conventional item. I decided to look at this because this is a very important step in understanding the differences between organic and conventional. Organic products have many less ingredients, and all ingredients tend to also be organic or natural. Whereas conventional products tend to have ongoing ingredient lists, usually full of stuff I have no idea what it is and can barely pronounce. They also have many artificial and processed ingredients. I know many people do not even bother looking at the ingredient list so I found this to be an interesting and useful step.

[23]:
```
my_string = organic_chicken["Nutrient"].loc[21] #Locate the row with organic␣
 ↪ingredient info
my_list = my_string.split(",") #split the string to get items separated by␣
 ↪commas
my_new_list = my_list[0:-1]
last_ingredient = my_list[-1]
new_last_ingredient = last_ingredient.split(".") #these lines were needed to␣
 ↪get rid of unnecessary info after the ingredients
my_new_list.append(new_last_ingredient[0])
for i, value in enumerate(my_new_list, 1): #print out the enumerated list of␣
 ↪ingredients
    print(i, value)
```

```
1 ORGANIC CHICKEN BREAST MEAT WITH RIB MEAT
2  WATER
3  SEA SALT
```

[24]:
```
my_string = chicken["Nutrient"].loc[21] #Locate the row with conventional␣
 ↪ingredient info
```

```
my_list = my_string.split(",") #split the string to get items separated by␣
↪commas
my_new_list = my_list[0:-1]
last_ingredient = my_list[-1]
new_last_ingredient = last_ingredient.split(".") #these lines were needed to␣
↪get rid of unnecessary info after the ingredients
my_new_list.append(new_last_ingredient[0])
for i, value in enumerate(my_new_list, 1): #print out the enumerated list of␣
↪ingredients
    print(i, value)

    ##### THIS PROCESS WAS REPEATED AGAIN TO GET THE INGREDIENTS FOR ICE CREAM␣
↪SANDWICH #####
```

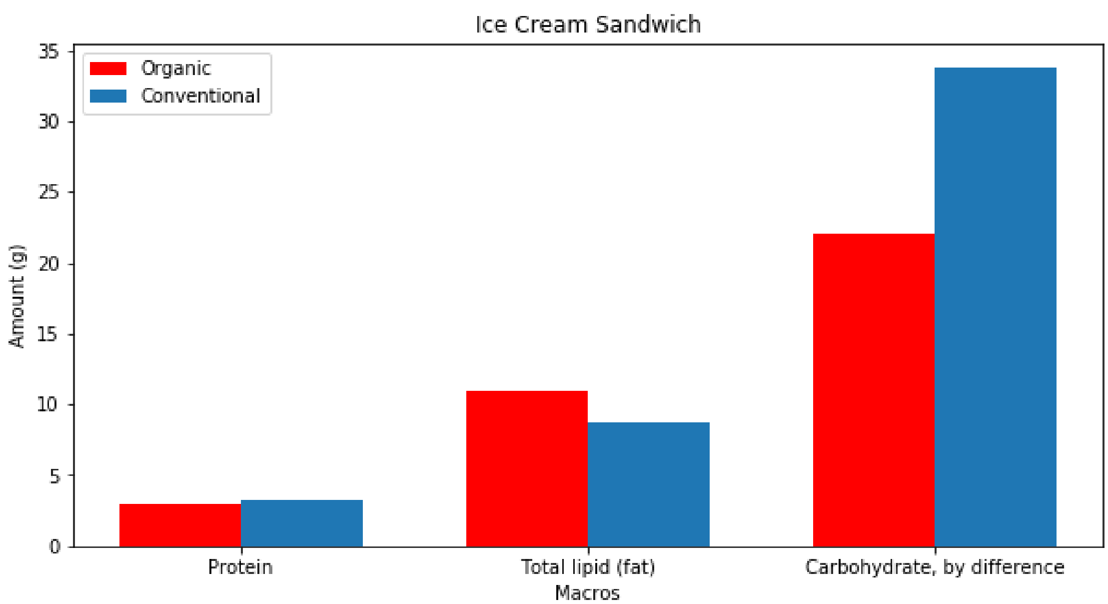
```
1 CHICKEN BREAST
2  WATER
3  SEASONING (DEHYDRATED GARLIC
4  SUGAR
5  DEHYDRATED ONION
6  SPICES
7  PAPRIKA
8  NATURAL FLAVOR [MALTODEXTRIN]
9  DEHYDRATED RED BELL PEPPER
10  DEHYDRATED SHALLOTS
11  CHICKEN BROTH
12  PARSLEY
13  OLIVE OIL)
14  CONTAINS 2% OR LESS OF SEA SALT
15  VINEGAR POWDER*
16  SEASONING (YEAST EXTRACT
17  NATURAL FLAVOR)
18  NATURAL FLAVORS
```

[25]:
```
organic_ice_cream = pd.read_csv("Organic Ice Cream Sandwich.csv")
ice_cream = pd.read_csv("Ice Cream Sandwich.csv")
```

[26]:
```
organic_protein = organic_ice_cream.loc[2]
organic_fat = organic_ice_cream.loc[3]
organic_carbs = organic_ice_cream.loc[4]


protein = ice_cream.loc[2]
fat = ice_cream.loc[3]
carbs = ice_cream.loc[4]


nutrients = [organic_protein[0], organic_fat[0], organic_carbs[0]]
```

```
link = [organic_protein[1], organic_fat[1], organic_carbs[1]]

link2 = [protein[1], fat[1], carbs[1]]
```

[27]:
```
plt.figure(figsize=(10,5))
n_groups = 3
index = np.arange(n_groups)
width = 0.35

plt.bar(index, link, width, label="Organic", color='r')
plt.bar(index + width, link2, width, label="Conventional")

plt.xticks(index + width / 2, nutrients)
plt.xlabel("Macros")
plt.ylabel("Amount (g)")
plt.title("Ice Cream Sandwich")
plt.legend()
plt.show()
```



[28]:
```
my_string = organic_ice_cream["Nutrient"].loc[21]
my_list = my_string.split(",")
my_new_list = my_list[0:-1]
last_ingredient = my_list[-1]
new_last_ingredient = last_ingredient.split(".")
my_new_list.append(new_last_ingredient[0])
for i, value in enumerate(my_new_list, 1):
    print(i, value)
```

```
1 ICE CREAM: ORGANIC CREAM
2  ORGANIC MILK
3  ORGANIC CANE SUGAR
4  ORGANIC EGG YOLKS
5  ORGANIC VANILLA EXTRACT
6  ORGANIC LOCUST BEAN GUM
7  ORGANIC GUAR GUM.
8  CHOCOLATE WAFER: ORGANIC WHEAT FLOUR
9  ORGANIC SUGAR
10  ORGANIC CANE SYRUP
11  ORGANIC PALM OIL
12  ORGANIC CARAMEL COLOR
13  BAKING SODA
14  SOY LECITHIN
15  ORGANIC COCOA
16  SALT
17  NATURAL VANILLA FLAVOR
```

[29]:

```
my_string = ice_cream["Nutrient"].loc[21]
my_list = my_string.split(",")
my_new_list = my_list[0:-1]
last_ingredient = my_list[-1]
new_last_ingredient = last_ingredient.split(".")
my_new_list.append(new_last_ingredient[0])
for i, value in enumerate(my_new_list, 1):
    print(i, value)
```

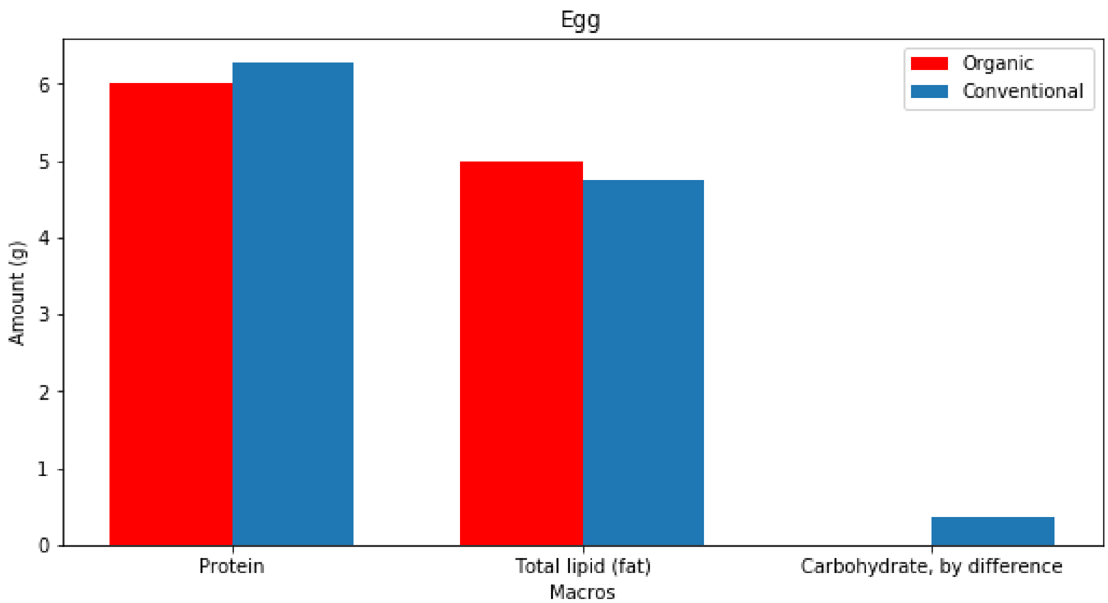
```
1 ICE CREAM - MILK
2  CREAM
3  CORN SYRUP
4  SUGAR
5  WHEY
6  BUTTERMILK
7  MALTODEXTRIN
8  CELLULOSE GEL
9  MONO AND DIGLYCERIDES
10  GUAR GUM
11  NATURAL AND ARTIFICIAL FLAVORS
12  LOCUST BEAN GUM
13  CELLULOSE GUM
14  POLYSORBATE 80
15  CARRAGEENAN
16  SUCRALOSE. WAFERS - BLEACHED WHEAT FLOUR
17  SUGAR
18  WHOLE WHEAT FLOUR
19  PALM OIL
20  DEXTROSE
```

```
21   MOLASSES
22   HIGH FRUCTOSE CORN STARCH
23   SALT
24   BAKING SODA
25   NATURAL FLAVOR
26   MONO & DIGLYCERIDES
27   CARAMEL COLOR
28   SOY LECITHIN. CHOCO CHIPS - SUGAR
29   COCONUT OIL
30   COCOA PROCESSES WITH ALKALI
31   PARTIALLY HYDROGENATED COCONUT OIL
32   COCOA
33   SALT
34   LECITHIN
35   NATURAL FLAVOR
```

[30]:
```
organic_egg = pd.read_csv("Organic Egg.csv")
egg = pd.read_csv("Egg.csv")
```

[31]:
```
organic_protein = organic_egg.loc[2]
organic_fat = organic_egg.loc[3]
organic_carbs = organic_egg.loc[4]


protein = egg.loc[3]
fat = egg.loc[4]
carbs = egg.loc[5]


nutrients = [organic_protein[0], organic_fat[0], organic_carbs[0]]
link = [organic_protein[1], organic_fat[1], organic_carbs[1]]

link2 = [protein[2], fat[2], carbs[2]]
```

[32]:
```
plt.figure(figsize=(10,5))
n_groups = 3
index = np.arange(n_groups)
width = 0.35

plt.bar(index, link, width, label="Organic",color='r')
plt.bar(index + width, link2, width, label="Conventional")

plt.xticks(index + width / 2, nutrients)
plt.xlabel("Macros")
plt.ylabel("Amount (g)")
plt.title("Egg")
plt.legend()
plt.show()
```



Fruits and vegetables have always been at the top of organic food sales, and some of the most popular products people should always buy organic are apples, strawberries, and lettuce. These next cells show how the price of these items have changed over time, and I predicted the price of them for the next year.

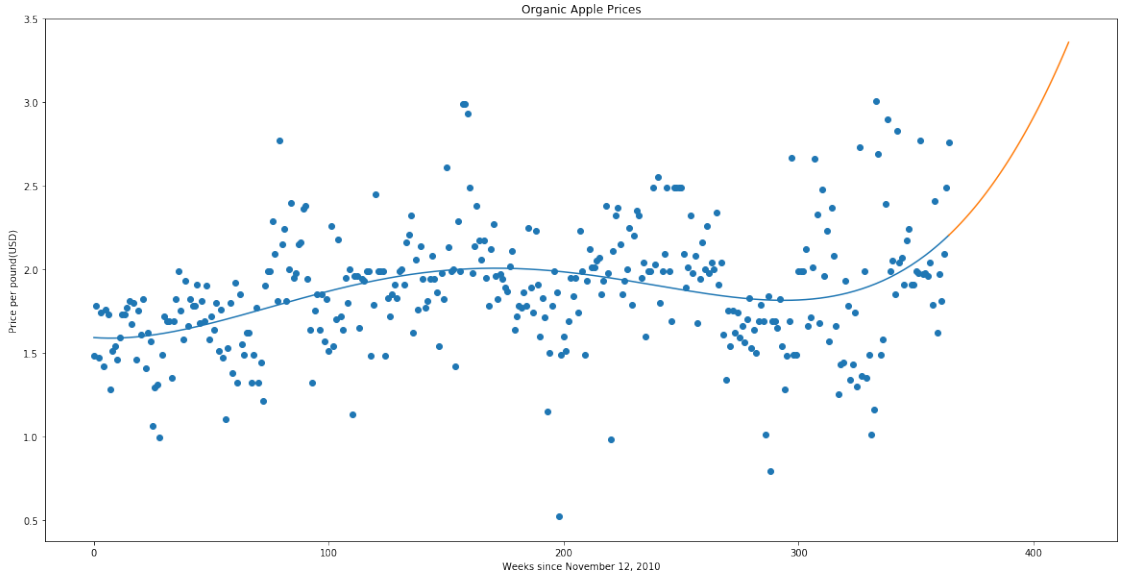
[57]:
```
apples = pd.read_csv("Apple prices.csv") #open file
apple_date = apples["REPORT_DATE"] #get specific columns
apple_price = apples["WEIGHTED_AVERAGE_PRICE"]
```

[58]:
```
date = np.arange(0,len(apple_date))
max_date = max(date)
extended_date = np.arange(max_date, max_date+52, 1) #create extended date array␣
 ↪- used 52 because there are 52 weeks in a year

parameters1 = np.polyfit(date, apple_price, 4) #use polyfit to create a␣
 ↪regression line for the price data I had
my_poly_function1 = np.poly1d(parameters1)
expected_y_poly1 = my_poly_function1(date)

parameters2 = np.polyfit(date, apple_price, 4) #use polyfit to create a␣
 ↪regression line predicting the values for the next year
my_poly_function2 = np.poly1d(parameters2)
expected_y_poly2 = my_poly_function1(extended_date)
```

[59]:
```
plt.figure(figsize=(20,10))
plt.xlabel("Weeks since November 12, 2010") #set labels
plt.ylabel("Price per pound(USD)")
plt.title("Organic Apple Prices")
plt.scatter(date, apple_price) #create plots and regression lines
```

```
plt.plot(date, expected_y_poly1)
plt.plot(extended_date, expected_y_poly2)

#### THIS PROCESS WAS REPEATED FOR STRAWBERRY AND ROMAINE LETTUCE PRICES ####
```

[59]: [<matplotlib.lines.Line2D at 0x11c3a6ba8>]

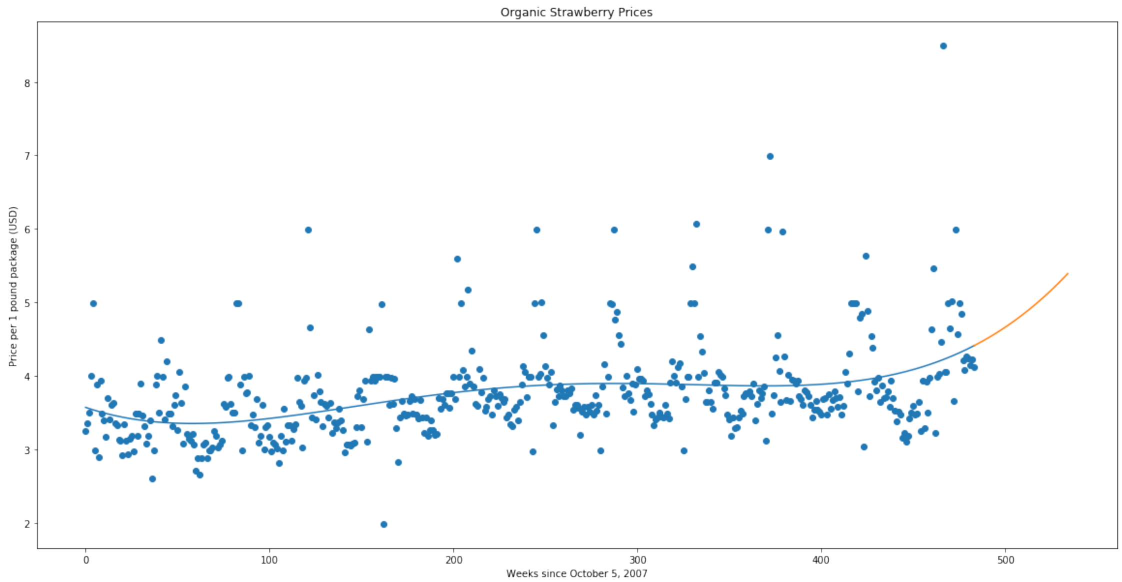


[61]:
```
strawberry = pd.read_csv("Strawberry prices.csv")
strawberry_date = strawberry["REPORT_DATE"]
strawberry_price = strawberry["WEIGHTED_AVERAGE_PRICE"]
```

[62]:
```
date = np.arange(0,len(strawberry_date))
max_date = max(date)
extended_date = np.arange(max_date, max_date+52, 1)

parameters1 = np.polyfit(date, strawberry_price, 4)
my_poly_function1 = np.poly1d(parameters1)
expected_y_poly1 = my_poly_function1(date)

parameters2 = np.polyfit(date, strawberry_price, 4)
my_poly_function2 = np.poly1d(parameters2)
expected_y_poly2 = my_poly_function1(extended_date)
```

[63]:
```
plt.figure(figsize=(20,10))
plt.scatter(date, strawberry_price)
plt.plot(date, expected_y_poly1)
plt.plot(extended_date, expected_y_poly2)
plt.title("Organic Strawberry Prices")
plt.xlabel("Weeks since October 5, 2007")
plt.ylabel("Price per 1 pound package (USD)")
```

[63]: Text(0,0.5,'Price per 1 pound package (USD)')

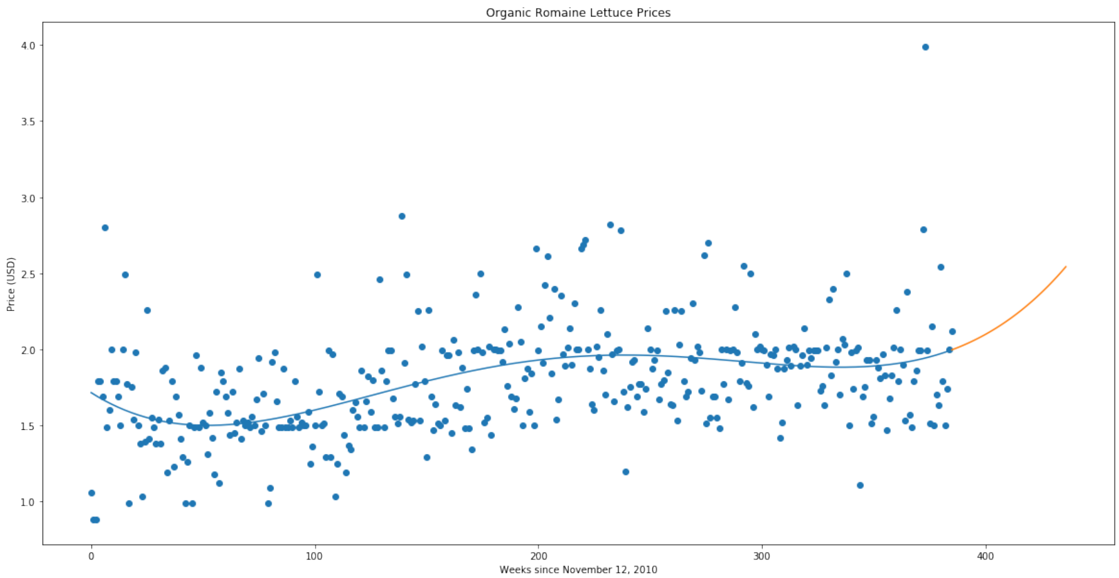


[65]:
```
lettuce = pd.read_csv("Lettuce prices.csv")
lettuce_date = lettuce["REPORT_DATE"]
lettuce_price = lettuce["WEIGHTED_AVERAGE_PRICE"]
```

[66]:
```
date = np.arange(0,len(lettuce_date))
max_date = max(date)
extended_date = np.arange(max_date, max_date+52, 1)

parameters1 = np.polyfit(date, lettuce_price, 4)
my_poly_function1 = np.poly1d(parameters1)
expected_y_poly1 = my_poly_function1(date)

parameters2 = np.polyfit(date, lettuce_price, 4)
my_poly_function2 = np.poly1d(parameters2)
expected_y_poly2 = my_poly_function1(extended_date)
```

[67]:
```
plt.figure(figsize=(20,10))
plt.scatter(date, lettuce_price)
plt.plot(date, expected_y_poly1)
plt.plot(extended_date, expected_y_poly2)
plt.title("Organic Romaine Lettuce Prices")
plt.xlabel("Weeks since November 12, 2010")
plt.ylabel("Price (USD)")
```

[67]: Text(0,0.5,'Price (USD)')



The following cells contains data comparing prices between conventional and organic products at four different supermarkets. I was able to use masking to find the percentage differences of the same food products I used above: apples, strawberries, lettuce. I also found the overall average of how much extra we pay for organic at these supermarkets.

[44]:
```
price_diff = pd.read_csv("Price Differences.csv")
price_diff
```

[44]:

	Table 2	Price Chopper	Safeway	Walmart	\
0	Apples (lb.)	NaN	NaN	NaN	
1	Regular	$1.00	$1.83	NaN	
2	Organic	$1.20	$2.20	NaN	
3	% difference	20%	20%	NaN	
4	NaN	NaN	NaN	NaN	
5	Bananas (lb.)	NaN	NaN	NaN	
6	Regular	59 cents	48 cents	58 cents	
7	Organic	79 cents	79 cents	78 cents	
8	% difference	34%	65%	34%	
9	NaN	NaN	NaN	NaN	
10	Beef (85% lean ground, lb.)	NaN	NaN	NaN	
11	Regular	NaN	$5.99	NaN	
12	Organic	NaN	$8.79	NaN	
13	% difference	NaN	47%	NaN	
14	NaN	NaN	NaN	NaN	
15	Butter (lb.)	NaN	NaN	NaN	
16	Regular	$2.99/lb.	NaN	$3.88/lb.	
17	Organic	$7.98/lb.	NaN	$6.48/lb.	
18	% difference	167%	NaN	67%	
19	NaN	NaN	NaN	NaN	
20	Carrots (baby, lb.)	NaN	NaN	NaN	

```
21                           Regular     $1.33     $2.19      $1.68
22                           Organic     $1.99     $2.19      $3.48
23                      % difference       50%        0%       107%
24                               NaN       NaN       NaN        NaN
25      Chicken, whole/cutup (lb.)       NaN       NaN        NaN
26                           Regular     $1.49     $1.99        NaN
27                           Organic     $3.49     $2.49        NaN
28                      % difference      134%       25%        NaN
29                               NaN       NaN       NaN        NaN
..                               ...       ...       ...        ...
45          Iceberg lettuce (head)       NaN       NaN        NaN
46                           Regular     $1.99     $2.79      $1.68
47                           Organic     $3.49     $3.29      $2.48
48                      % difference       75%       18%        48%
49                               NaN       NaN       NaN        NaN
50     Maple syrup (Grade A, pint)       NaN       NaN        NaN
51                           Regular    $11.99    $11.84     $10.21
52                           Organic    $10.65    $16.97     $11.84
53                      % difference      -11%       43%         6%
54                               NaN       NaN       NaN        NaN
55               Milk (half gallon)       NaN       NaN        NaN
56                           Regular     $2.99     $2.69      $2.20
57                           Organic     $3.99     $3.49      $3.88
58                      % difference       33%       30%        76%
59                               NaN       NaN       NaN        NaN
60  Olive oil (extra virgin, quart)       NaN       NaN        NaN
61                           Regular    $16.08     $8.64      $8.98
62                           Organic    $17.02    $13.44     $10.87
63                      % difference        6%       56%        21%
64                               NaN       NaN       NaN        NaN
65               Strawberries (lb.)       NaN       NaN        NaN
66                           Regular     $2.99     $4.39        NaN
67                           Organic     $4.99     $7.69        NaN
68                      % difference       67%       75%        NaN
69                               NaN       NaN       NaN        NaN
70                   Zucchini (lb.)       NaN       NaN        NaN
71                           Regular     $1.99  72 cents      $1.80
72                           Organic     $2.99     $1.12      $1.98
73                      % difference       50%       56%        10%
74       Average premium for organic       59%       34%        51%

   Whole Foods
0          NaN
1          NaN
2          NaN
3          NaN
4          NaN
```

```
5           NaN
6      79 cents
7      99 cents
8           25%
9           NaN
10          NaN
11        $6.99
12        $9.99
13          43%
14          NaN
15          NaN
16    $3.79/lb.
17    $4.39.lb.
18          16%
19          NaN
20          NaN
21          NaN
22          NaN
23          NaN
24          NaN
25          NaN
26        $2.49
27        $3.49
28          40%
29          NaN
..          ...
45          NaN
46          NaN
47          NaN
48          NaN
49          NaN
50          NaN
51       $11.99
52       $11.72
53          -2%
54          NaN
55          NaN
56        $2.39
57        $3.99
58          67%
59          NaN
60          NaN
61       $13.24
62       $13.24
63           0%
64          NaN
65          NaN
```

```
66       $4.99
67       $6.99
68         40%
69         NaN
70         NaN
71         NaN
72         NaN
73         NaN
74         24%

[75 rows x 5 columns]
```

[45]:
```
price_diff.loc[0:3]
```

[45]:
```
         Table 2 Price Chopper Safeway Walmart Whole Foods
0   Apples (lb.)           NaN     NaN     NaN         NaN
1        Regular         $1.00   $1.83     NaN         NaN
2        Organic         $1.20   $2.20     NaN         NaN
3   % difference           20%     20%     NaN         NaN
```

[46]:
```
price_diff.loc[65:68]
```

[46]:
```
                Table 2 Price Chopper Safeway Walmart Whole Foods
65  Strawberries (lb.)           NaN     NaN     NaN         NaN
66             Regular         $2.99   $4.39     NaN       $4.99
67             Organic         $4.99   $7.69     NaN       $6.99
68        % difference           67%     75%     NaN         40%
```

[47]:
```
price_diff.loc[45:48]
```

[47]:
```
                    Table 2 Price Chopper Safeway Walmart Whole Foods
45  Iceberg lettuce (head)           NaN     NaN     NaN         NaN
46                 Regular         $1.99   $2.79   $1.68         NaN
47                 Organic         $3.49   $3.29   $2.48         NaN
48            % difference           75%     18%     48%         NaN
```

[48]:
```
price_diff.loc[74]
```

[48]:
```
Table 2          Average premium for organic
Price Chopper                            59%
Safeway                                  34%
Walmart                                  51%
Whole Foods                              24%
Name: 74, dtype: object
```

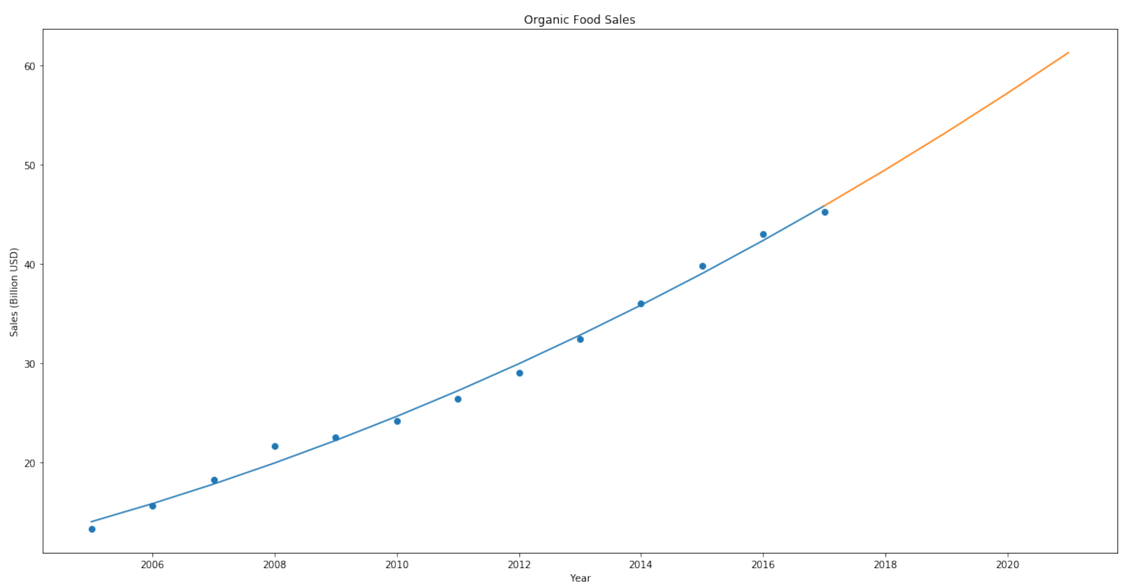
The following cells show how sales of organic food in the U.S. have changed over time. I was very interested in showing this after showing the price differences, because I wanted to show that sales are still increasing in spite of the costs of organic products also increasing.

[72]:
```
us_sales = pd.read_csv("US Organic Sales.csv", skiprows=2)
years = us_sales["Year"]
money = us_sales["Sales in billion U.S. dollars"]
```

[73]:
```
max_date = max(years)
extended_date = np.arange(max_date, max_date+5, 1)

parameters1 = np.polyfit(years, money, 2)
my_poly_function1 = np.poly1d(parameters1)
expected_y_poly1 = my_poly_function1(years)

parameters2 = np.polyfit(years, money, 2)
my_poly_function2 = np.poly1d(parameters2)
expected_y_poly2 = my_poly_function1(extended_date)
```

[52]:
```
plt.figure(figsize=(20,10))
plt.scatter(years, money)
plt.plot(years, expected_y_poly1)
plt.plot(extended_date, expected_y_poly2)
plt.title("Organic Food Sales")
plt.xlabel("Year")
plt.ylabel("Sales (Billion USD)")
```

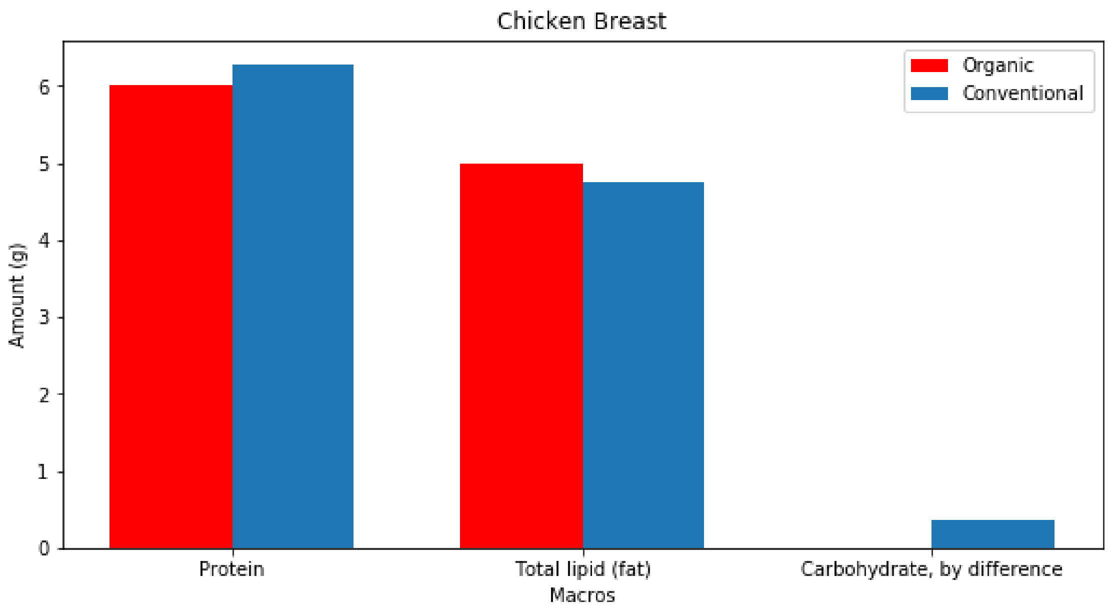
[52]: Text(0,0.5,'Sales (Billion USD)')



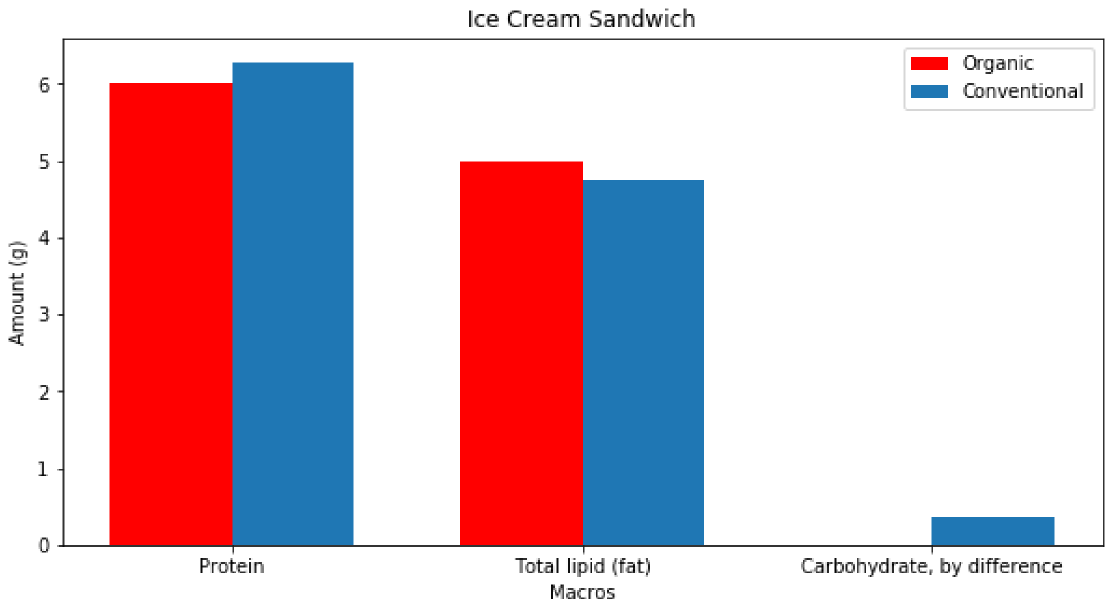
0.0.4 Results

[53]:
```
plt.figure(figsize=(10,5)) #set figure size
n_groups = 3 #there are three different groups of bars that will be on the␣
 ↪chart
index = np.arange(n_groups) #the index and width came in handy when I had to␣
 ↪put the bars side by side
width = 0.35
```

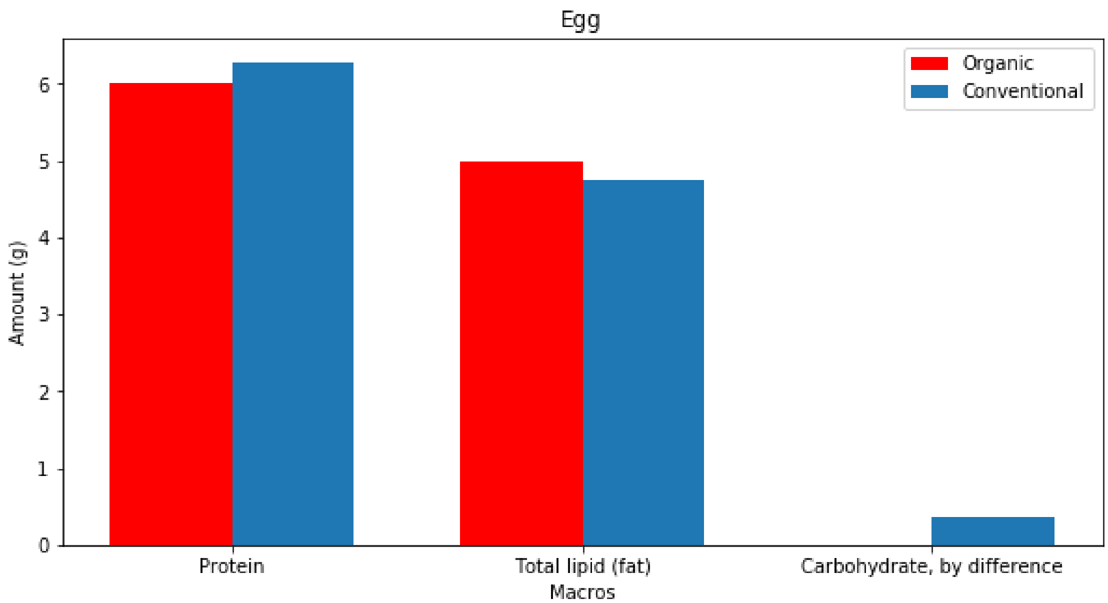
```
plt.bar(index, link, width, label="Organic", color='r') #graph the organic info
plt.bar(index + width, link2, width, label = "Conventional") #graph the␣
 ↪conventional info

plt.xticks(index + width / 2, nutrients)
plt.xlabel("Macros")
plt.ylabel("Amount (g)")
plt.title("Chicken Breast")
plt.legend()
plt.show()
```



[54]:
```
plt.figure(figsize=(10,5))
n_groups = 3
index = np.arange(n_groups)
width = 0.35

plt.bar(index, link, width, label="Organic", color='r')
plt.bar(index + width, link2, width, label="Conventional")

plt.xticks(index + width / 2, nutrients)
plt.xlabel("Macros")
plt.ylabel("Amount (g)")
plt.title("Ice Cream Sandwich")
plt.legend()
plt.show()
```

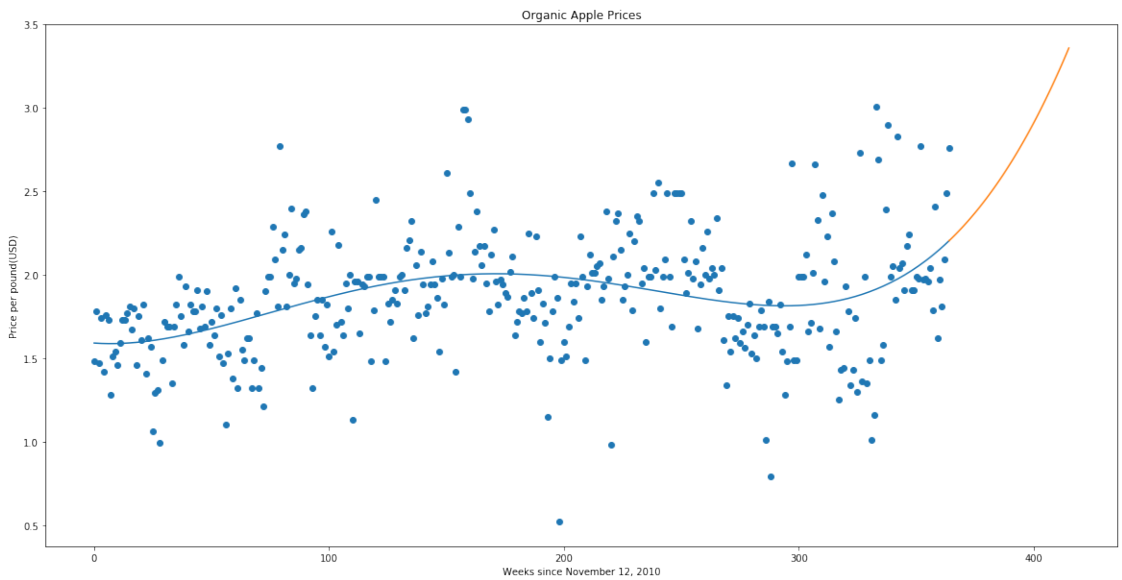


[55]:
```
plt.figure(figsize=(10,5))
n_groups = 3
index = np.arange(n_groups)
width = 0.35

plt.bar(index, link, width, label="Organic",color='r')
plt.bar(index + width, link2, width, label="Conventional")

plt.xticks(index + width / 2, nutrients)
plt.xlabel("Macros")
plt.ylabel("Amount (g)")
plt.title("Egg")
plt.legend()
plt.show()
```



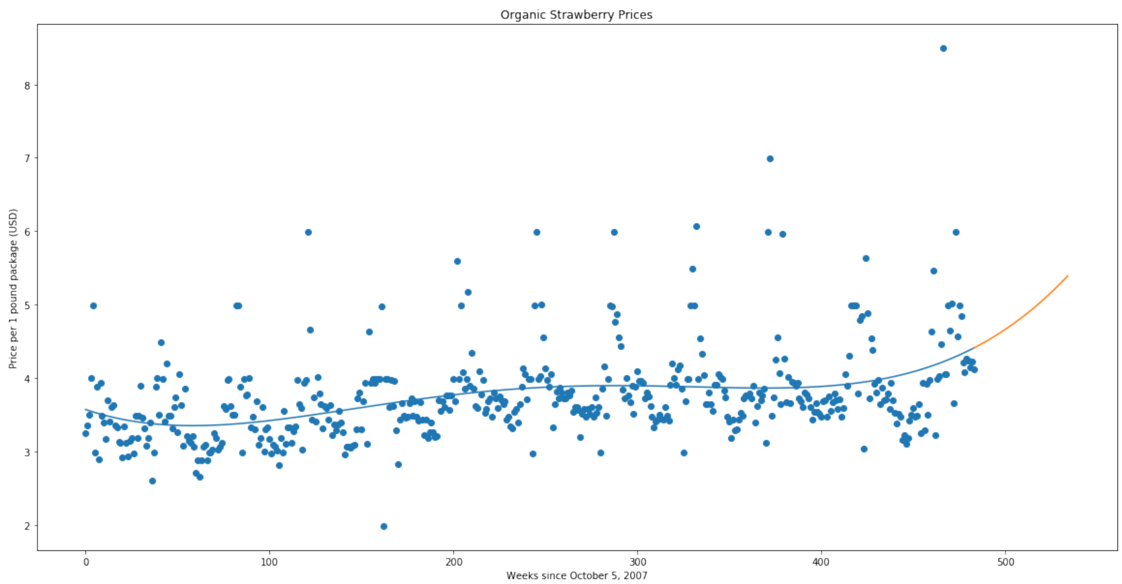
[60]:
```
plt.figure(figsize=(20,10))
plt.xlabel("Weeks since November 12, 2010") #set labels
plt.ylabel("Price per pound(USD)")
plt.title("Organic Apple Prices")
plt.scatter(date, apple_price) #create plots and regression lines
plt.plot(date, expected_y_poly1)
plt.plot(extended_date, expected_y_poly2)
```

[60]: [<matplotlib.lines.Line2D at 0x118f83f98>]



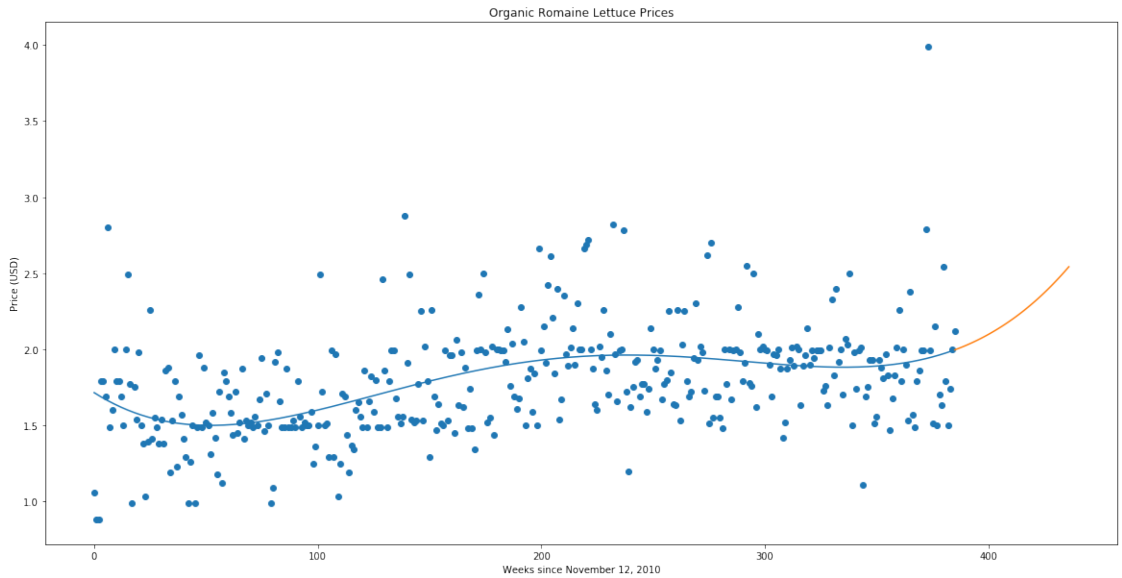
[64]:
```
plt.figure(figsize=(20,10))
plt.scatter(date, strawberry_price)
plt.plot(date, expected_y_poly1)
plt.plot(extended_date, expected_y_poly2)
plt.title("Organic Strawberry Prices")
plt.xlabel("Weeks since October 5, 2007")
plt.ylabel("Price per 1 pound package (USD)")
```

[64]:
```
Text(0,0.5,'Price per 1 pound package (USD)')
```



[68]:
```
plt.figure(figsize=(20,10))
plt.scatter(date, lettuce_price)
plt.plot(date, expected_y_poly1)
plt.plot(extended_date, expected_y_poly2)
plt.title("Organic Romaine Lettuce Prices")
plt.xlabel("Weeks since November 12, 2010")
plt.ylabel("Price (USD)")
```

[68]:
```
Text(0,0.5,'Price (USD)')
```



[69]:
```
price_diff.loc[0:3]
```

[69]:
```
       Table 2 Price Chopper Safeway Walmart Whole Foods
0  Apples (lb.)          NaN     NaN     NaN         NaN
1       Regular        $1.00   $1.83     NaN         NaN
2       Organic        $1.20   $2.20     NaN         NaN
3  % difference          20%     20%     NaN         NaN
```

[70]:
```
price_diff.loc[65:68]
```

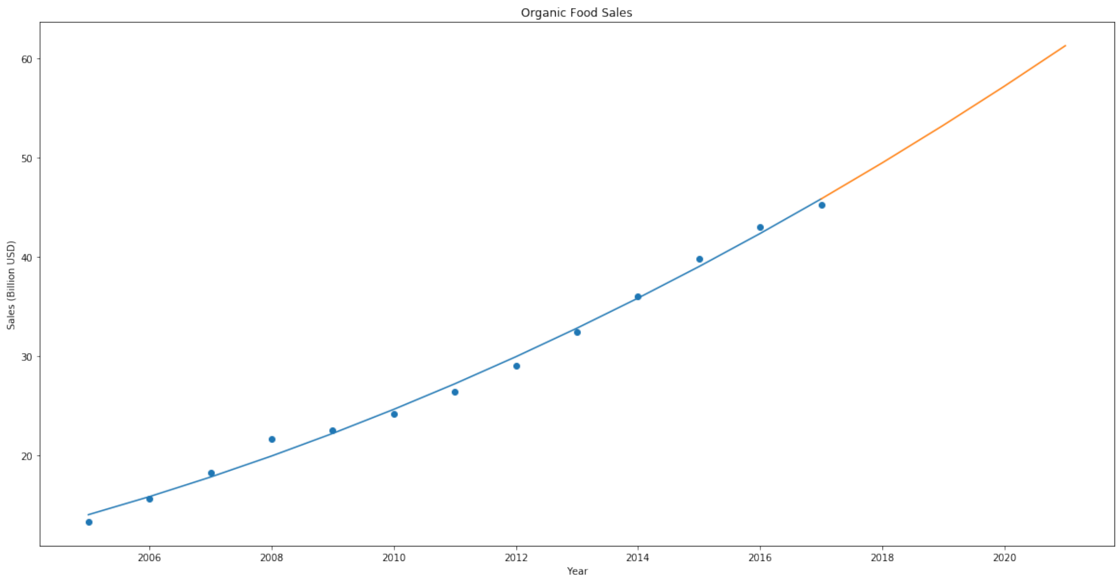
[70]:
```
              Table 2 Price Chopper Safeway Walmart Whole Foods
65  Strawberries (lb.)          NaN     NaN     NaN         NaN
66             Regular        $2.99   $4.39     NaN       $4.99
67             Organic        $4.99   $7.69     NaN       $6.99
68        % difference          67%     75%     NaN         40%
```

[71]:
```
price_diff.loc[45:48]
```

[71]:
```
                  Table 2 Price Chopper Safeway Walmart Whole Foods
45  Iceberg lettuce (head)          NaN     NaN     NaN         NaN
46                 Regular        $1.99   $2.79   $1.68         NaN
47                 Organic        $3.49   $3.29   $2.48         NaN
48            % difference          75%     18%     48%         NaN
```

[74]:
```
plt.figure(figsize=(20,10))
plt.scatter(years, money)
plt.plot(years, expected_y_poly1)
plt.plot(extended_date, expected_y_poly2)
plt.title("Organic Food Sales")
plt.xlabel("Year")
plt.ylabel("Sales (Billion USD)")
```

[74]: Text(0,0.5,'Sales (Billion USD)')



0.0.5 Discussion and Conclusion

Some obstacles that I ran into with this project were: 1. I originally wanted to model the differences between how organic and conventional foods are grown, and te difference between the pesticides used. However, this became such a complex and long task because there is so much information about these and every farm grows their things differently. So, unfortunately I was unable to include this in my project. 2. It was really hard finding data on this at first, because organic products have really only become popular within the last 5-10 years.

Some things I would have done differently are maybe start everything sooner, so I am not working up until the very last minute on this. I also wish I was able to find data sets on the demographics of what types of people purchase organic.

From my results, it can be found that there are essentially no nutritional difference in a products macros (carbs, fats, proteins) between organic and conventional. There is usually a 1-2g difference between the two. Again, I only chose to look at the macros because that is what I personally look for on a label to help me determine if it is "healthy" enough to include in my diet.

Even though there are little to no nutritional differences, there are very big differences when it comes to ingredients. Conventional products have twice or more times of ingredients than organic, and the ingredients in conventional include a lot of processed or artificial things, most of which I do not know what it is nor can even pronounce. It is clear to see that people do not buy organic because of its additional nutrients, since there are none, they buy it for the things it doesn't have; such as artificial/processed ingredients and pesticides.

The last thing I decided to compare were the prices of these two food types. Fruits and vegetables are the #1 organic product bought, and so I decided to look at a few of the prices for the top fruits/vegetables that people by organic, which are apples, strawberries, and lettuce. From the graphs it seemed that they all steadily increased in prices up until 2011-2012 and then they were all able to dip a little in prices. However, within the last 1-2 years the prices have begun to increase, and will continually increase this next year. Within this next year apples will increase by more than $1, strawberries will increase by $1, and lettuce will increase by around 50 cents.

While these may not seem like huge price increases, when we compare these to the prices of the same conventional products it is shown that we are paying an extra 20% for organic apples, an extra 40-75% for organic strawberries, and an extra 18-75% for lettuce (these depend on which supermarket you shop at). On average people will pay an extra 40-50% for organic products.

I wanted to see if organic sales were going to decrease due to the expected increase in prices, however the complete opposite happened. Organic sales have been steadily increasing since 2005, and are expected to continue to increase by $3 billion per year, for the next 5 years.

Seeing that sales in organic products are increasing drastically in spite of their prices rising, it is viable to assume that people are realizing the health benefits of organic products and believe they are worth the cost.

0.0.6 References

Food Composition Databases Show Foods List, ndb.nal.usda.gov/ndb/search/list.

"Fruits and Vegetables Top Organic Food Sales." USDA ERS - Chart Detail, www.ers.usda.gov/data-products/chartgallery/gallery/chart- detail/?chartId=87354.

"Market News Data Download." Market News, marketnews.usda.gov/mnp/dataDownload.

"Organic Food Sales in the U.S. 2017 | Statistic." Statista, www.statista.com/statistics/196952/organic-food-sales-in-the-us- since-2000/.