In [13]:

```
import  pandas as pd #importar bibliotecas
import os
print(os.getcwd())   #verificar la ruta de donde va el archivo
```

```
C:\Users\Royco
```

In [18]:

```
data = pd.read_csv("avocado.csv")  # definimos la variable  para leer la tabla del archivo excel
pd.read_csv("avocado.csv")
```

Out[18]:

	Unnamed: 0	Date	AveragePrice	Total Volume	4046	4225	4770	Total Bags	Small Bags	Large Bags	XLarge Bags	
0	0	2015-12-27	1.33	64236.62	1036.74	54454.85	48.16	8696.87	8603.62	93.25	0.0	co
1	1	2015-12-20	1.35	54876.98	674.28	44638.81	58.33	9505.56	9408.07	97.49	0.0	co
2	2	2015-12-13	0.93	118220.22	794.70	109149.67	130.50	8145.35	8042.21	103.14	0.0	co
3	3	2015-12-06	1.08	78992.15	1132.00	71976.41	72.58	5811.16	5677.40	133.76	0.0	co
4	4	2015-11-29	1.28	51039.60	941.48	43838.39	75.78	6183.95	5986.26	197.69	0.0	co
...	...	...	...	...	...	...	...	...	...	...	...	...

In [17]:

```
data.head() #se ponen los nombres de cada columna con solo los primeros 5 datos
```

Out[17]:

	Unnamed: 0	Date	AveragePrice	Total Volume	4046	4225	4770	Total Bags	Small Bags	Large Bags	XLarg Bag
0	0	2015-12-27	1.33	64236.62	1036.74	54454.85	48.16	8696.87	8603.62	93.25	0.
1	1	2015-12-20	1.35	54876.98	674.28	44638.81	58.33	9505.56	9408.07	97.49	0.
2	2	2015-12-13	0.93	118220.22	794.70	109149.67	130.50	8145.35	8042.21	103.14	0.
3	3	2015-12-06	1.08	78992.15	1132.00	71976.41	72.58	5811.16	5677.40	133.76	0.
4	4	2015-11-29	1.28	51039.60	941.48	43838.39	75.78	6183.95	5986.26	197.69	0.

In [32]:

```
#Cantidad de datos  en la tabla
data.loc[:,:]
```

Out[32]:

	Unnamed: 0	Date	AveragePrice	Total Volume	4046	4225	4770	Total Bags	Small Bags	Large Bags
0	0	2015-12-27	1.33	64236.62	1036.74	54454.85	48.16	8696.87	8603.62	93.25
1	1	2015-12-20	1.35	54876.98	674.28	44638.81	58.33	9505.56	9408.07	97.49
2	2	2015-12-13	0.93	118220.22	794.70	109149.67	130.50	8145.35	8042.21	103.14
3	3	2015-12-06	1.08	78992.15	1132.00	71976.41	72.58	5811.16	5677.40	133.76
4	4	2015-11-29	1.28	51039.60	941.48	43838.39	75.78	6183.95	5986.26	197.69
...	...	...	...	...	...	...	...	...	...	...
18244	7	2018-02-04	1.63	17074.83	2046.96	1529.20	0.00	13498.67	13066.82	431.85
18245	8	2018-01-28	1.71	13888.04	1191.70	3431.50	0.00	9264.84	8940.04	324.80
18246	9	2018-01-21	1.87	13766.76	1191.92	2452.79	727.94	9394.11	9351.80	42.31
18247	10	2018-01-14	1.93	16205.22	1527.63	2981.04	727.01	10969.54	10919.54	50.00
18248	11	2018-01-07	1.62	17489.58	2894.77	2356.13	224.53	12014.15	11988.14	26.01

18249 rows × 14 columns

In [33]:

```
data.columns
```

Out[33]:

```
Index(['Unnamed: 0', 'Date', 'AveragePrice', 'Total Volume', '4046', '4225',
       '4770', 'Total Bags', 'Small Bags', 'Large Bags', 'XLarge Bags', 'type',
       'year', 'region'],
      dtype='object')
```

In [37]:

```
#Tipos de datos
data.dtypes
```

Out[37]:

```
Unnamed: 0        int64
Date             object
AveragePrice    float64
Total Volume    float64
4046            float64
4225            float64
4770            float64
Total Bags      float64
Small Bags      float64
Large Bags      float64
XLarge Bags     float64
type             object
year              int64
region           object
dtype: object
```

In [40]:

```
data.columns
```

Out[40]:

```
Index(['Unnamed: 0', 'Date', 'AveragePrice', 'Total Volume', '4046', '4225',
       '4770', 'Total Bags', 'Small Bags', 'Large Bags', 'XLarge Bags', 'type',
       'year', 'region'],
      dtype='object')
```

In [44]:

```
data.AveragePrice.mean #media
```

Out[44]:

```
<bound method NDFrame._add_numeric_operations.<locals>.mean of 0        1.33
1        1.35
2        0.93
3        1.08
4        1.28
         ...
18244    1.63
18245    1.71
18246    1.87
18247    1.93
18248    1.62
Name: AveragePrice, Length: 18249, dtype: float64>
```

In [43]:

```
data.AveragePrice.std #desviación estándar
```

Out[43]:

```
<bound method NDFrame._add_numeric_operations.<locals>.std of 0        1.33
1        1.35
2        0.93
3        1.08
4        1.28
         ...
18244    1.63
18245    1.71
18246    1.87
18247    1.93
18248    1.62
Name: AveragePrice, Length: 18249, dtype: float64>
```

In [45]:

```
data.AveragePrice.median() #mediana
```

Out[45]:

```
1.37
```

In [46]:

```
data.AveragePrice.max() #Max
```

Out[46]:

```
3.25
```

In [47]:

```
data.AveragePrice.min()#Min
```

Out[47]:

```
0.44
```

Después de ver los datos con su media, mediana y desviaciación estándar. Se puede ver que la diferencia entre cada uno de los datos es mínima son decimales los que cambian entre cada uno. Y si analizamos la gráfica de la desviación estándar nos daremos cuenta que los datos no estan tan dispersos. Por lo tanto no hay una gran variación entre los datos analizados y su diferencia es muy pequeña.