ANALISIS DE DATASET AVOCADOS

MICHAEL VENTURA

Se requiere:

1. Obtener Cuantas filas y cuantas columnas tiene el conjunto de datos
2. Mostrar los primeros 100 registros
3. Mostrar los últimos 20 registros
4. Cual es el precio mínimo, máximo y promedio del aguacate en ese conjunto de datos
5. Generar un gráfico de tipo scatter usando para la x la variable 'year' y para y la variable 'AveragePrice' para 3 regiones

In [1]:
```
# importando dataset
import pandas as pd
import matplotlib.pyplot as plt
avocados = pd.read_csv('avocado.csv')
```

Filas y Columnas

In [2]:
```
columns = avocados.columns
totalcolumns = len(columns)
totalrows = len(avocados)
print(f'TOTAL FILAS: {totalrows}, TOTAL COLUMNAS {totalcolumns}' )
print('\nCOLUMNAS')
for column in columns:
    print(f'- {column}')
```

```
TOTAL FILAS: 18249, TOTAL COLUMNAS 14

COLUMNAS
- Unnamed: 0
- Date
- AveragePrice
- Total Volume
- 4046
- 4225
- 4770
- Total Bags
- Small Bags
- Large Bags
- XLarge Bags
- type
- year
- region
```

Primeros 100 registros

In [5]:
```
avocados.head(100)
```

Out[5]:

	Unnamed: 0	Date	AveragePrice	Total Volume	4046	4225	4770	Total Bags	
0	0	2015-12-27	1.33	64236.62	1036.74	54454.85	48.16	8696.87	86
1	1	2015-12-20	1.35	54876.98	674.28	44638.81	58.33	9505.56	94
2	2	2015-12-13	0.93	118220.22	794.70	109149.67	130.50	8145.35	80
3	3	2015-12-06	1.08	78992.15	1132.00	71976.41	72.58	5811.16	56
4	4	2015-11-29	1.28	51039.60	941.48	43838.39	75.78	6183.95	59
...	...	...	...	...	...	...	...	...	
95	43	2015-03-01	0.99	512532.44	441544.00	21183.70	347.78	49456.96	321
96	44	2015-02-22	1.10	431308.56	369316.29	18324.76	237.47	43430.04	334
97	45	2015-02-15	1.06	427391.28	364446.77	17420.41	638.18	44885.92	267
98	46	2015-02-08	1.03	433883.91	377256.38	17162.50	524.85	38940.18	180
99	47	2015-02-01	0.96	636771.37	553198.56	31583.38	294.94	51694.49	105

100 rows × 14 columns

Ultimos 20 registros

In [7]:
```
avocados.tail(20)
```

Out[7]:

	Unnamed: 0	Date	AveragePrice	Total Volume	4046	4225	4770	Total Bags
18229	4	2018-02-25	1.49	301985.61	34200.18	49139.34	85.58	218560.51
18230	5	2018-02-18	1.64	224798.60	30149.00	38800.64	123.13	155725.83
18231	6	2018-02-11	1.47	275248.53	24732.55	61713.53	243.00	188559.45
18232	7	2018-02-04	1.41	283378.47	22474.66	55360.49	133.41	205409.91
18233	8	2018-01-28	1.80	185974.53	22918.40	33051.14	93.52	129911.47
18234	9	2018-01-21	1.83	189317.99	27049.44	33561.32	439.47	128267.76
18235	10	2018-01-14	1.82	207999.67	33869.12	47435.14	433.52	126261.89
18236	11	2018-01-07	1.48	297190.60	34734.97	62967.74	157.77	199330.12
18237	0	2018-03-25	1.62	15303.40	2325.30	2171.66	0.00	10806.44
18238	1	2018-03-18	1.56	15896.38	2055.35	1499.55	0.00	12341.48
18239	2	2018-03-11	1.56	22128.42	2162.67	3194.25	8.93	16762.57
18240	3	2018-03-04	1.54	17393.30	1832.24	1905.57	0.00	13655.49
18241	4	2018-02-25	1.57	18421.24	1974.26	2482.65	0.00	13964.33
18242	5	2018-02-18	1.56	17597.12	1892.05	1928.36	0.00	13776.71
18243	6	2018-02-11	1.57	15986.17	1924.28	1368.32	0.00	12693.57
18244	7	2018-02-04	1.63	17074.83	2046.96	1529.20	0.00	13498.67
18245	8	2018-01-28	1.71	13888.04	1191.70	3431.50	0.00	9264.84
18246	9	2018-01-21	1.87	13766.76	1191.92	2452.79	727.94	9394.11
18247	10	2018-01-14	1.93	16205.22	1527.63	2981.04	727.01	10969.54

	Unnamed: 0	Date	AveragePrice	Total Volume	4046	4225	4770	Total Bags
18248	11	2018-01-07	1.62	17489.58	2894.77	2356.13	224.53	12014.15

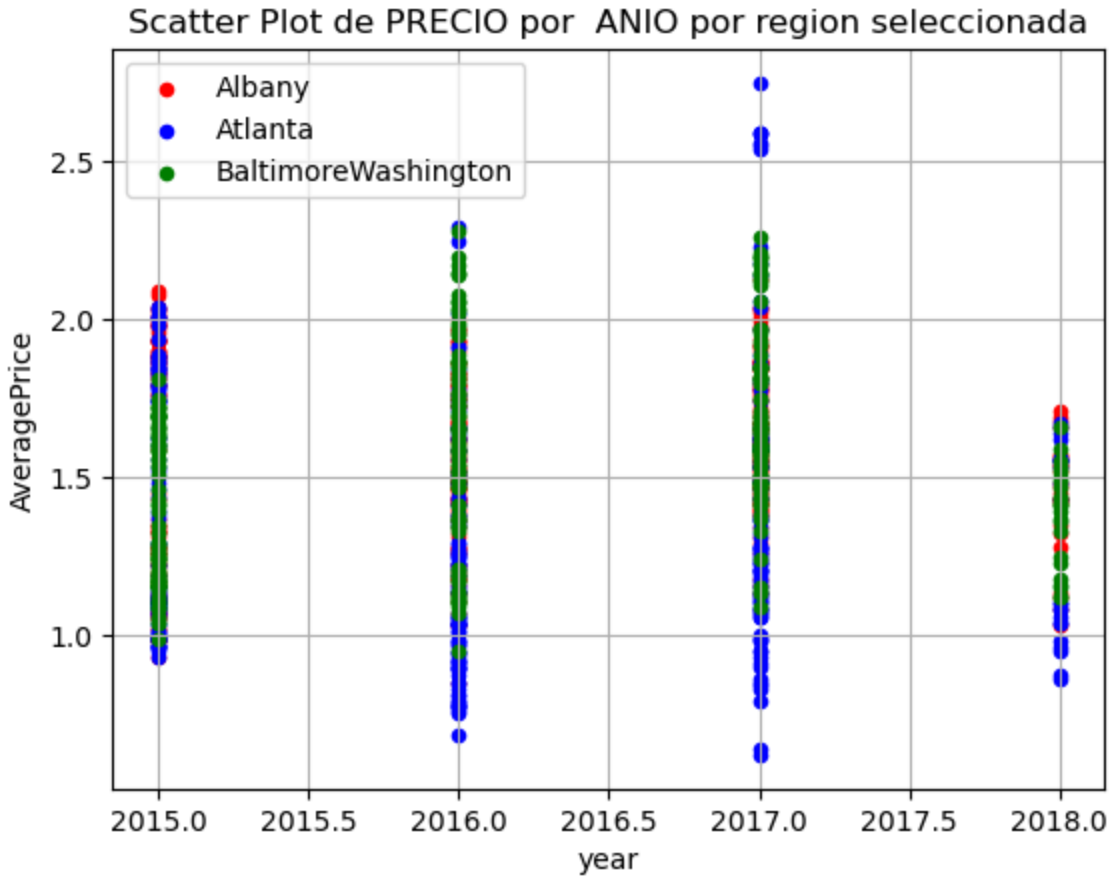
Precios de Aguacates

In [10]:
```
min = avocados['AveragePrice'].min()
max = avocados['AveragePrice'].max()
prom = round(avocados['AveragePrice'].mean(),2)

print(f'PRECIO MINIMO: {min} | PRECIO MAXIMO: {max} | PRECIO PROMEDIO: {prom}')
```

```
PRECIO MINIMO: 0.44 | PRECIO MAXIMO: 3.25 | PRECIO PROMEDIO: 1.41
```

Graficos

In [24]:
```
ax = None
colors = ['red', 'blue', 'green']
for region,color in zip(avocados['region'].unique()[0:3],colors):
    subset = avocados[avocados['region'] == region]
    ax = subset.plot.scatter(x='year', y='AveragePrice', label=region, color=color,

plt.title('Scatter Plot de PRECIO por  ANIO por region seleccionada')
plt.grid(True)
plt.show()
```



HECHO POR MICHAEL VENTURA

In [ ]: