

```
import pandas as pd
```

```
data = pd.read_csv('avocado_full.csv')
```

```
las_vegas_data = data[data['region'] == 'LasVegas']
```

```
las_vegas_data
```

	Date	AveragePrice	Total Volume	4046	4225	4770	Total Bags	Small Bags
1092	2015-12-27	0.83	292707.14	135799.94	93317.46	4977.64	58612.10	34675.97
1093	2015-12-20	0.96	232294.71	96355.05	77086.11	4749.61	54103.94	34992.96
1094	2015-12-13	0.95	216119.35	101042.61	68490.51	4942.71	41643.52	33834.13
1095	2015-12-06	0.89	244024.83	94777.64	107758.80	10794.34	30694.05	30201.26
1096	2015-11-29	0.95	205476.61	88653.67	81160.21	5227.59	30435.14	29475.24
...	...	...	...	...	...	...	...	...
17860	2018-02-04	1.79	9503.13	1479.32	1512.78	0.00	6511.03	6465.39
17861	2018-01-28	1.78	8672.12	1005.73	2110.22	0.00	5556.17	5538.46

Primero se hace la región de Charlotte

```
Charlotte_data = data[data['region'] == 'Charlotte']
```

Se hace la separación para que solo se consideren los datos de esta región

```
produccion_total_Charlotte = Charlotte_data['Total Bags'].sum()
produccion_promedio_Charlotte = Charlotte_data['Total Bags'].mean()
produccion_desviacion_estandar_Charlotte = Charlotte_data['Total Bags'].std()
```

Se hace la suma total de producción en Charlotte, el promedio y la desviación estandar

```
small_bags_Charlotte = Charlotte_data['Small Bags'].sum()
large_bags_Charlotte = Charlotte_data['Large Bags'].sum()
xlarge_bags_Charlotte = Charlotte_data['XLarge Bags'].sum()
```

Se suman cada uno de los tamaños de las bolsas para sacar el porcentaje

```
porcentaje_small_Charlotte = (small_bags_Charlotte / produccion_total_Charlotte) * 100
porcentaje_large_Charlotte = (large_bags_Charlotte / produccion_total_Charlotte) * 100
porcentaje_xl_Charlotte = (xlarge_bags_Charlotte / produccion_total_Charlotte) * 100
```

Se saca el porcentaje de bolsas por tamaño

```
print("Producción total de Charlotte:", produccion_total_Charlotte)
print("Producción promedio de Charlotte:", produccion_promedio_Charlotte)
print("Desviación estándar de la producción de Charlotte:", produccion_desviacion_estandar_Charlotte)
print("Porcentaje por tamaño de bolsa chica de Charlotte:", porcentaje_small_Charlotte)
print("Porcentaje por tamaño de bolsa grande de Charlotte:", porcentaje_large_Charlotte)
print("Porcentaje por tamaño de bolsa extra grande de Charlotte:", porcentaje_xl_Charlotte)
```

```
Producción total de Charlotte: 11697942.41
Producción promedio de Charlotte: 34609.297071005916
Desviación estándar de la producción de Charlotte: 37032.501145198745
Porcentaje por tamaño de bolsa chica de Charlotte: 90.98909190133378
Porcentaje por tamaño de bolsa grande de Charlotte: 8.415900980658018
```

Porcentaje por tamaño de bolsa extra grande de Charlotte: 0.5950071180082002

```
NewYork_data = data[data['region'] == 'NewYork']

produccion_total_NewYork = NewYork_data['Total Bags'].sum()
produccion_promedio_NewYork = NewYork_data['Total Bags'].mean()
produccion_desviacion_estandar_NewYork = NewYork_data['Total Bags'].std()

small_bags_NewYork = NewYork_data['Small Bags'].sum()
large_bags_NewYork = NewYork_data['Large Bags'].sum()
xlarge_bags_NewYork = NewYork_data['XLarge Bags'].sum()

porcentaje_small_NewYork = (small_bags_NewYork / produccion_total_NewYork) * 100
porcentaje_large_NewYork = (large_bags_NewYork / produccion_total_NewYork) * 100
porcentaje_xl_NewYork = (xlarge_bags_NewYork / produccion_total_NewYork) * 100

print("Producción total de Nueva York:", produccion_total_NewYork)
print("Producción promedio de Nueva York:", produccion_promedio_NewYork)
print("Desviación estándar de la producción de Nueva York:", produccion_desviacion_estandar_NewYork)
print("Porcentaje por tamaño de bolsa chica de Nueva York:", porcentaje_small_NewYork)
print("Porcentaje por tamaño de bolsa grande de Nueva York:", porcentaje_large_NewYork)
print("Porcentaje por tamaño de bolsa extra grande de Nueva York:", porcentaje_xl_NewYork)

    Producción total de Nueva York: 67994156.75999999
    Producción promedio de Nueva York: 201166.144260355
    Desviación estándar de la producción de Nueva York: 197491.93271625863
    Porcentaje por tamaño de bolsa chica de Nueva York: 85.89220792330921
    Porcentaje por tamaño de bolsa grande de Nueva York: 13.579104470094174
    Porcentaje por tamaño de bolsa extra grande de Nueva York: 0.5286890773112369

California_data = data[data['region'] == 'California']

produccion_total_California = California_data['Total Bags'].sum()
produccion_promedio_California = California_data['Total Bags'].mean()
produccion_desviacion_estandar_California = California_data['Total Bags'].std()

small_bags_California = California_data['Small Bags'].sum()
large_bags_California = California_data['Large Bags'].sum()
xlarge_bags_California = California_data['XLarge Bags'].sum()

porcentaje_small_California = (small_bags_California / produccion_total_California) * 100
porcentaje_large_California = (large_bags_California / produccion_total_California) * 100
porcentaje_xl_California = (xlarge_bags_California / produccion_total_California) * 100

print("Producción total de California:", produccion_total_California)
print("Producción promedio de California:", produccion_promedio_California)
print("Desviación estándar de la producción de California:", produccion_desviacion_estandar_California)
print("Porcentaje por tamaño de bolsa chica de California:", porcentaje_small_California)
print("Porcentaje por tamaño de bolsa grande de California:", porcentaje_large_California)
print("Porcentaje por tamaño de bolsa extra grande de California:", porcentaje_xl_California)

    Producción total de California: 246682193.70999998
    Producción promedio de California: 729828.9754733727
    Desviación estándar de la producción de California: 813755.2335129493
    Porcentaje por tamaño de bolsa chica de California: 92.8715917612309
    Porcentaje por tamaño de bolsa grande de California: 4.991011197376464
    Porcentaje por tamaño de bolsa extra grande de California: 2.13739704139264
```

Se puede observar que California tiene la producción total más alta, después va Nueva York y finalmente Charlotte. Esto sugiere que California es líder en la producción de aguacates, con más información se podría decir si es por los cuidados que se les da o por el tipo de clima

Viendo la producción promedio se puede inferir que California tiene una producción total más alta que las otras dos regiones. Después sigue Nueva York y por último Charlotte.

California tiene la mayor desviación estándar, esto significa que son menos constantes al momento de su producción esto se puede dar por distintas variables. Nueva York tiene la desviación estandar menor, esto significa que tienen una menor variabilidad en comparación de las otras regiones.

En las tres regiones la mayor parte de la producción total se encuentra en bolsas pequeñas. Pero, en California, una parte significativa también se encuentra en las bolsas grandes y extra grandes.

California tiene la producción total más altas, pero también tienen una desviación estandar mayor.