

Actividad Evaluable: Obtención de estadísticas descriptivas

Eduardo Rodríguez Gil - A01274913, Jose Manuel Neri Villeda - A01706450, Héctor Javier Calderón González - A01067542

```
In [1]: import pandas as pd
        from matplotlib import pyplot as plt

        data = pd.read_csv('Data Bitcoin.csv')
```

```
In [38]: data
```

```
Out[38]:
```

	Date	Price	Open	High	Low	Volume	Change
0	Apr 25, 2021	49561.9	50088.2	50438.8	49226.5	66.26K	-1.05%
1	Apr 24, 2021	50088.9	51140.8	51183.0	48775.2	82.25K	-2.06%
2	Apr 23, 2021	51143.6	51707.1	52099.9	47659.4	214.46K	-1.13%
3	Apr 22, 2021	51729.5	53821.3	55408.4	50590.9	168.13K	-3.88%
4	Apr 21, 2021	53820.2	56479.5	56764.4	53657.6	100.26K	-4.71%
...
3930	Jul 22, 2010	0.1	0.1	0.1	0.1	2.16K	0.00%
3931	Jul 21, 2010	0.1	0.1	0.1	0.1	0.58K	0.00%
3932	Jul 20, 2010	0.1	0.1	0.1	0.1	0.26K	0.00%
3933	Jul 19, 2010	0.1	0.1	0.1	0.1	0.57K	0.00%
3934	Jul 18, 2010	0.1	0.0	0.1	0.1	0.08K	0.00%

3935 rows × 7 columns

```
In [39]: filas=len(data)
        filas
```

```
Out[39]: 3935
```

En esta parte de la documentación, mostramos los datos cargados mediante el uso de la librería de Pandas, por lo que al revisar esta tabla se muestra que es lo que contiene nuestra base de datos que cuenta con 3935 filas y 7 columnas.

```
In [40]: data.columns
```

```
Out[40]: Index(['Date', 'Price', 'Open', 'High', 'Low', 'Volume', 'Change'], dtype='object')
```

```
In [41]: type(data)
```

```
Out[41]: pandas.core.frame.DataFrame
```

```
In [42]: type(data.Date)
```

```
Out[42]: pandas.core.series.Series
```

```
In [43]: type(data.Price)
```

```
Out[43]: pandas.core.series.Series
```

```
In [44]: type(data.Open)
```

```
Out[44]: pandas.core.series.Series
```

```
In [45]: type(data.High)
```

```
Out[45]: pandas.core.series.Series
```

```
In [46]: type(data.Low)
```

```
Out[46]: pandas.core.series.Series
```

```
In [47]: type(data.Volume)
```

```
Out[47]: pandas.core.series.Series
```

```
In [48]: type(data.Change)
```

```
Out[48]: pandas.core.series.Series
```

Por medio de estos comandos, encontramos las 7 columnas con sus respectivos tipos de datos como lo es DataFrame y series.

```
In [49]: pd.unique(data['Price']) #como se comporta el dato
```

```
Out[49]: array([5.04388e+04, 5.11830e+04, 5.20999e+04, ..., 3.00000e-01,
                2.00000e-01, 1.00000e-01])
```

```
In [54]: data['Price'].describe()
```

```
Out[54]: count      3935.000000
         mean       4334.106912
         std        8888.395159
         min         0.100000
         25%        72.550000
         50%        567.600000
         75%        6583.550000
         max       63540.900000
         Name: Price, dtype: float64
```

```
In [53]: data['Price'].std()
```

```
Out[53]: 8888.395158961683
```

```
In [52]: data['Price'].median()
```

```
Out[52]: 567.6
```

```
In [56]: data['Price'].mean()
```

```
Out[56]: 4334.1069123253
```

Al analizar estos datos tenemos una desviación estándar muy alta por lo que me lleva a deducir que su precio es muy cambiante respecto a esta criptomoneda, como media se tiene que el valor medio es de 567.6. Por último, se tiene un promedio de sus 3935 datos que se tiene de registro desde julio de 2010 dando un valor de 4334.1069.