```
import pandas as pd
```

```
data = {
    "Country": ["Belgium", "India", "Brazil"],
    "Capital": ["Brussels", "New Delhi", "Brasilia"],
    "Population": [11190846, 1303171035, 202747528]
}
```

```
df
```

	Country	Capital	Population
0	Belgium	Brussels	11190846
1	India	New Delhi	1303171035
2	Brazil	Brasilia	202747528

Next steps: Generate code with df New interactive sheet

```
df1 = pd.read_csv("BigBasket.csv")
```

```
df1
```

Show hidden output

Next steps: Generate code with df1 New interactive sheet Generate code with df1 New interactive sheet

```
df.head()
```

	Country	Capital	Population
0	Belgium	Brussels	11190846
1	India	New Delhi	1303171035
2	Brazil	Brasilia	202747528

Next steps: Generate code with df New interactive sheet

```
df.tail()
```

	Country	Capital	Population
0	Belgium	Brussels	11190846
1	India	New Delhi	1303171035
2	Brazil	Brasilia	202747528

```
df.sample()
```

	Country	Capital	Population
2	Brazil	Brasilia	202747528

```
df.shape
```

```
(3, 3)
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3 entries, 0 to 2
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype 
---  ------      --------------  ----- 
 0   Country     3 non-null      object
 1   Capital     3 non-null      object
 2   Population  3 non-null      int64 
dtypes: int64(1), object(2)
memory usage: 204.0+ bytes
```

```
df.describe()
```

```
         Population
count  3.000000e+00
 mean  5.057031e+08
  std  6.972372e+08
  min  1.119085e+07
  25%  1.069692e+08
  50%  2.027475e+08
  75%  7.529593e+08
  max  1.303171e+09
```

```
df.dtypes
```

```
                 0
   Country  object
   Capital  object
Population   int64

dtype: object
```

```
df.columns
```

```
Index(['Country', 'Capital', 'Population'], dtype='object')
```

```
df.index
```

```
RangeIndex(start=0, stop=3, step=1)
```

```
df1["Price"]
```

```
       Price
   0   69.75
   1  174.35
   2   34.87
   3   69.74
   4  174.37
 ...     ...
8203   49.00
8204  399.00
8205  465.00
8206  799.00
8207  250.00
8208 rows × 1 columns

dtype: float64
```

```
df1.isnull().sum()/len(df1)
```

```
                 0
 ProductName   0.0
       Brand   0.0
       Price   0.0
DiscountPrice  0.0
   Image_Url   0.0
    Quantity   0.0
    Category   0.0
 SubCategory   0.0
Absolute_Url   0.0

dtype: float64
```

```
df.notnull()
```

```
   Country  Capital  Population
0     True     True        True
1     True     True        True
2     True     True        True
```

```
df.dropna()
```

```
   Country    Capital  Population
0  Belgium   Brussels    11190846
1    India  New Delhi  1303171035
2   Brazil   Brasilia   202747528
```

```
df.fillna(value)
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
/tmp/ipython-input-654142765.py in <cell line: 0>()
----> 1 df.fillna(value)

NameError: name 'value' is not defined
```

Next steps: Explain error

```
df.replace(1 , "one")
```

```
   Country    Capital  Population
0  Belgium   Brussels    11190846
1    India  New Delhi  1303171035
2   Brazil   Brasilia   202747528
```

```
df.rename(columns={"old": "new"})
```

```
   Country    Capital  Population
0  Belgium   Brussels    11190846
1    India  New Delhi  1303171035
2   Brazil   Brasilia   202747528
```

```
df.drop_duplicates()
```

	Country	Capital	Population
0	Belgium	Brussels	11190846
1	India	New Delhi	1303171035
2	Brazil	Brasilia	202747528

```
df.groupby('Country').count()
```

	Capital	Population
Country		
Belgium	1	1
Brazil	1	1
India	1	1

```
Start coding or generate with AI.
```

```
df1.groupby('Brand').count()
```

	ProductName	Price	DiscountPrice	Image_Url	Quantity	Category	SubCategory	Absolute_Url
Brand								
109°F	1	1	1	1	1	1	1	1
24 Mantra	5	5	5	5	5	5	5	5
5:15PM	7	7	7	7	7	7	7	7
A-1 Chips	1	1	1	1	1	1	1	1
AEROFIL	1	1	1	1	1	1	1	1
...	...	...	...	...	...	...	...	...
nature's eco	1	1	1	1	1	1	1	1
s.Oliver	6	6	6	6	6	6	6	6
sumeru	2	2	2	2	2	2	2	2
suny	6	6	6	6	6	6	6	6
tagz	1	1	1	1	1	1	1	1

842 rows × 8 columns

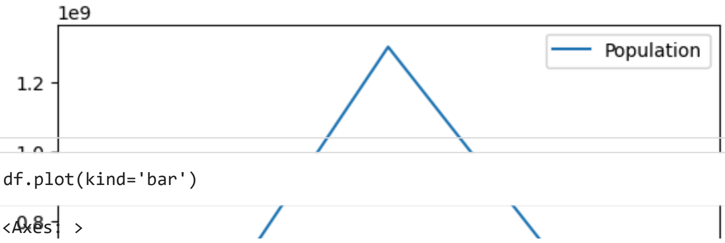
```
df['Country'].value_counts()
```

	count
Country	
Belgium	1
India	1
Brazil	1

dtype: int64

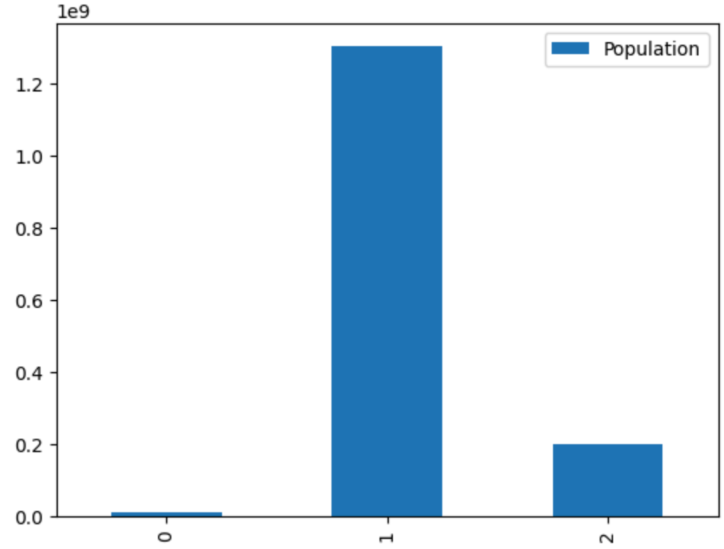
```
df.plot(kind='line')
```

```
<Axes: >
```



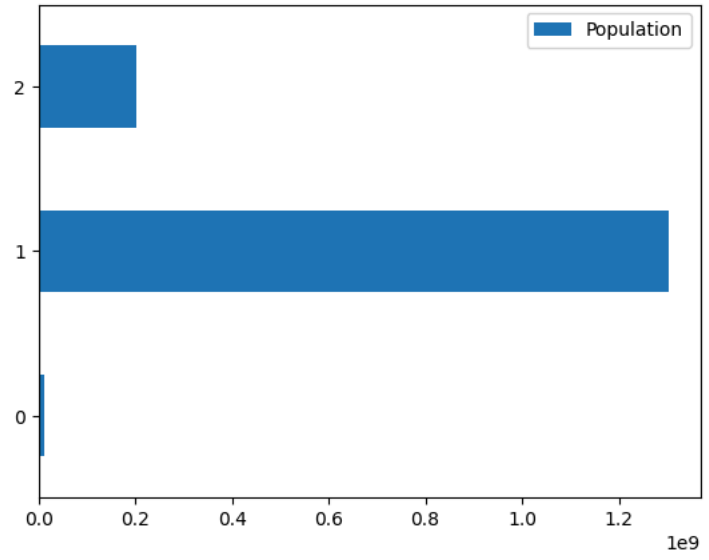
```
df.plot(kind='bar')
```

```
<Axes: >
```



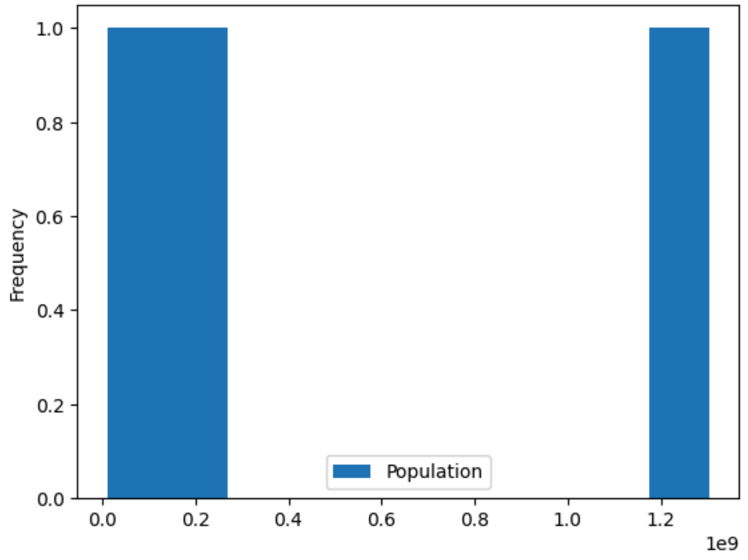
```
df.plot(kind='barh')
```

```
<Axes: >
```



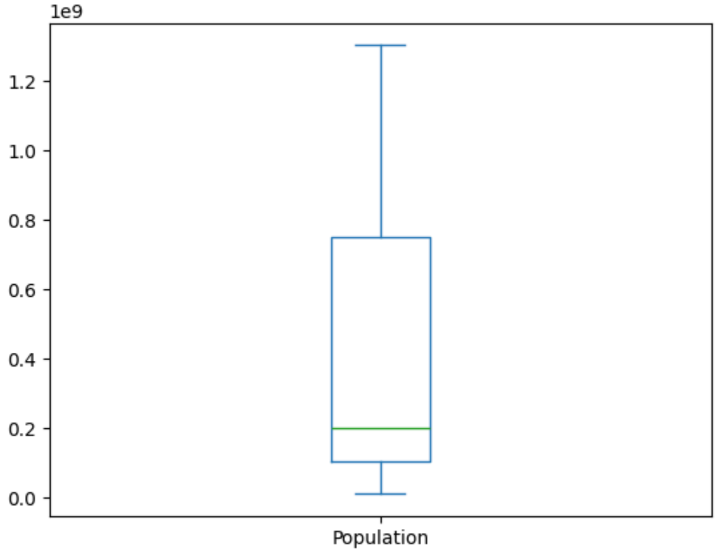
```
df.plot(kind='hist')
```

```
<Axes: ylabel='Frequency'>
```



```
df.plot(kind='box')
```

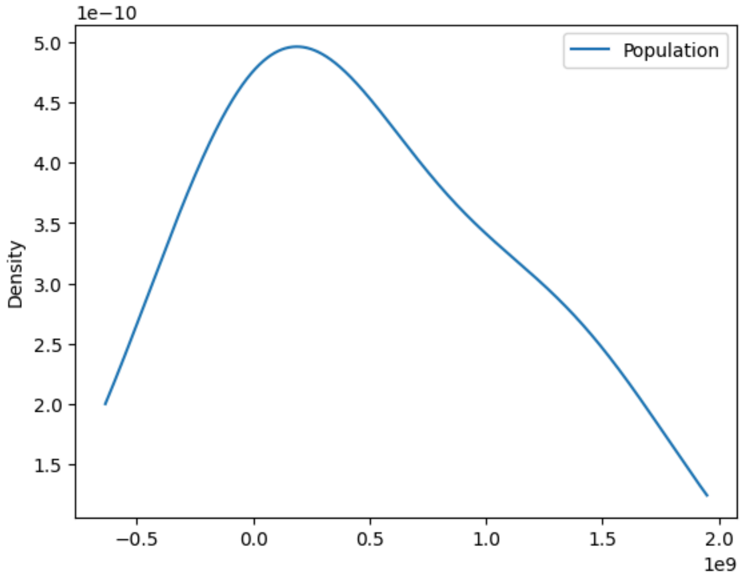
```
<Axes: >
```



```
df.plot(kind='line')
```

```
df.plot(kind='kde')
```

```
<Axes: ylabel='Density'>
```



```
df.plot(kind='pie', y='Population', labels=df['Country'], autopct='%1.1f%%')
```

```
<Axes: ylabel='Population'>
```