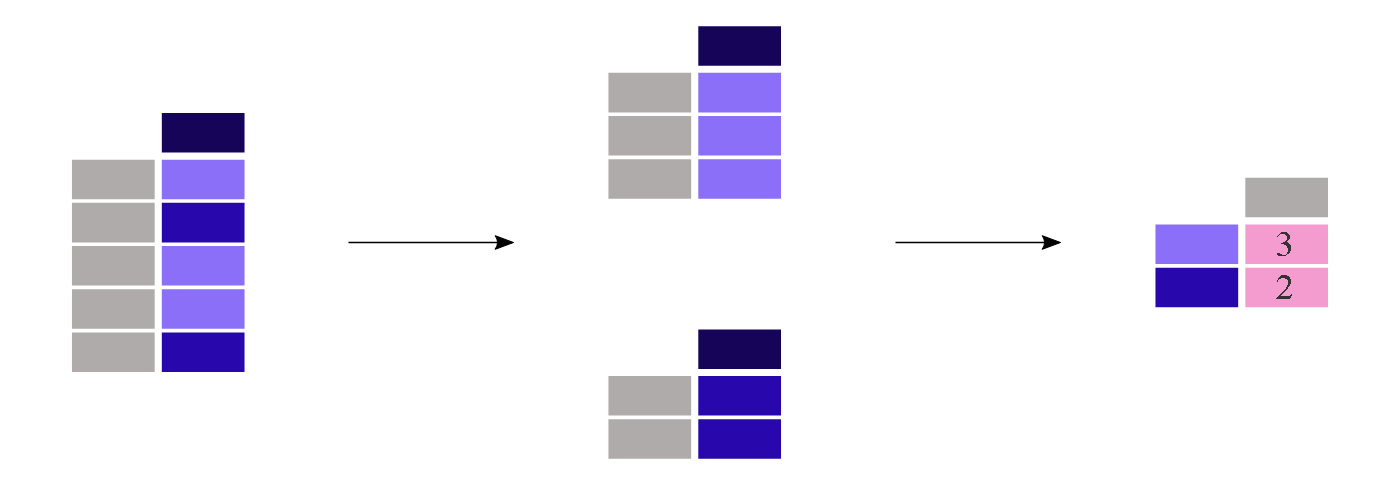
## Count number of records by category



The **[value\_counts()](https://pandas.pydata.org/docs/reference/api/pandas.Series.value_counts.html" \l "pandas.Series.value_counts" \o "pandas.Series.value_counts)** method counts the number of records for each category in a column.

## Use Pandas value\_counts()

You can use the **value\_counts()** function to count the frequency of unique values

This function uses the following basic syntax:

**series.value\_counts()**

### ****Count Frequency of Unique Values****

The following code shows how to count the occurrences of unique values in a pandas Series:

print('count the occurrences of unique values in a pandas Series:')

# import pandas

import pandas as pd

# create series

series1=pd.Series([2,2,3,3,3,3,4,4,111,2,33,11,222,333,222,111,222])

#count the occurance of unique values in the series

series1.value\_counts()

#this tells 3 occurs 4 times

#this tells 2 occurs 3 times

#this tells 222 occurs 3 times

#this tells 4 occurs 2 times

# and so on

Output

3 4

2 3

222 3

4 2

111 2

33 1

11 1

333 1

Name: count, dtype: int64

### ****Count Frequency of Unique Values (Including NaNs)****

By default, the **value\_counts()** function does not show the frequency of NaN values.

However, you can use the **dropna** argument to display the frequency of NaN values:

print('count the unique values from the serties but included null values also, that is included by numpy ')

# inport pandas

import pandas as pd

# import numpy

import numpy as np

# Create sertes with include null values

series2=pd.Series([2,2,3,3,3,4,4,4,4,6,6,6,5,5,5,43,34,44,44,44,np.nan,np.nan,55])

# count the occurance from series

series2.value\_counts() # this never shows the null values so we use dropna()

#show also occurance of null values

series2.value\_counts(dropna=False)

output

count the unique values from the serties but included null values also, that is included by numpy

4.0 4

3.0 3

5.0 3

6.0 3

44.0 3

2.0 2

NaN 2

43.0 1

34.0 1

55.0 1

Name: count, dtype: int64

### ****Count Relative Frequency of Unique Values****

The following code shows how to use the **normalize** argument to count the relative frequency of unique values in a pandas Series:

print('occurance of percentage use of normalize() function')

# import pandas

import pandas as pd

# create the series

series3=pd.Series([22,3,3,3,3,22,3,3,3,4,4,44,2,2,33,3,3,3,3,22,3,3,3,3,2])

#count the occurance of unique values in the series

series3.value\_counts(normalize=True)

#output shows

#The value 3 represents 60% of all values in the Series.

#The value 22 represents 12% of all values in the Series.

Output

occurance of percentage use of normalize() function

3 0.60

22 0.12

2 0.12

4 0.08

44 0.04

33 0.04

Name: proportion, dtype: float64

### ****Count Frequency in Bins****

The following code shows how to use the **bins** argument to count the frequency of values in a pandas Series that fall into equal-sized bins:

print('count frequency in bins')

# import pandas

import pandas as pd

# create the series

series4=pd.Series([22,3,3,3,3,22,3,3,3,4,4,44,2,2,33,3,3,3,3,22,3,3,3,3,2])

#count the occurance of unique values in the series

series4.value\_counts(bins=3)

#output display

#There are 20 values that fall in the range 1.957 to 16.0.

#There are 3 values that fall in the range 16.0 to 30.0.

Output

(1.957, 16.0] 20

(16.0, 30.0] 3

(30.0, 44.0] 2

Name: count, dtype: int64

**Count Frequency of Values in Pandas DataFrame**

We can also use the **value\_counts()** function to calculate the frequency of unique values in a specific column of a pandas DataFrame:

print('value counts in the dataframe')

#import pandas

import pandas as pd

#create dataframe

df=pd.DataFrame(

    {

        'points':[1,22,22,33,33,1,33,33,33,44,44,44,44],

        'scores':[200,200,300,30,300,30,300,22,22,22,22,22,22],

        'outruns':[2,3,2,3,2,3,3,3,4,4,4,2,2]

    }

)

df['points'].value\_counts()

output

points

33 5

44 4

22 2

1 2

Name: count, dtype: int64

### ****Count Frequency of Values in Pandas import any csv file****

print('value count function usage in file import')

#import pandas

import pandas as pd

#import file

df=pd.read\_csv('airlines\_flights\_data.csv')

df

#value count on a specified categoried column

df['class'].value\_counts()

output

class

Economy 206666

Business 93487

Name: count, dtype: int64