# **How to Get Unique Values from a Column in Pandas Data Frame?**

January 31, 2018 by [cmdline](https://cmdlinetips.com/author/cmdline/)

Often while working with a big data frame in pandas, you might have a column with string/characters and you want to find the number of unique elements present in the column. Pandas library in Python easily let you find the unique values.

Let us get started with some examples from a real world data set.

#### **Load gapminder data set**

| 1  2  3  4  5  6 | # import pandas as pd  import pandas as pd  # software carpentry url for gapminder data  gapminder\_csv\_url ='<http://bit.ly/2cLzoxH>'  # load the data with pd.read\_csv  gapminder = pd.read\_csv(gapminder\_csv\_url) |
| --- | --- |

Let us check the basic information of the data frame. We can see that the variables ‘continent’ and ‘country’ are objects/strings and we can find the number of unique values for them.

| 1  2 | # check the data frame info  print(gapminder.info()) |
| --- | --- |

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 1704 entries, 0 to 1703

Data columns (total 6 columns):

country 1704 non-null object

year 1704 non-null int64

pop 1704 non-null float64

continent 1704 non-null object

lifeExp 1704 non-null float64

gdpPercap 1704 non-null float64

dtypes: float64(3), int64(1), object(2)

memory usage: 79.9+ KB

### **Unique values of the column “continent”**

Let us say we want to find the unique values of column ‘continent’ in the data frame. We can use pandas’ function **unique** on the column of interest. It will return NumPy array with unique values of the column.

| 1  2 | >gapminder['continent'].unique()  array(['Asia', 'Europe', 'Africa', 'Americas', 'Oceania'], dtype=object) |
| --- | --- |

If we want the the unique values of the column in pandas data frame as a list, we can easily apply the function tolist() by chaining it to the previous command.

| 1  2 | >gapminder['continent'].unique().tolist()  ['Asia', 'Europe', 'Africa', 'Americas', 'Oceania'] |
| --- | --- |

If we try the unique function on the ‘country’ column from the dataframe, the result will be a big numpy array.

| 1 | >gapminder['country'].unique() |
| --- | --- |

Instead, we can simply count the number of unique values in the country column and find that there are 142 countries in the data set.

| 1  2 | >len(gapminder['country'].unique().tolist())  142 |
| --- | --- |