

DataFrame Methods

FUNCTION	DESCRIPTION
index()	Method returns index (row labels) of the DataFrame
<u>insert()</u>	Method inserts a column into a DataFrame
<u>add()</u>	Method returns addition of dataframe and other, element-wise (binary operator add)
<u>sub()</u>	Method returns subtraction of dataframe and other, element-wise (binary operator sub)
<u>mul()</u>	Method returns multiplication of dataframe and other, element-wise (binary operator mul)
<u>div()</u>	Method returns floating division of dataframe and other, element-wise operator truediv)
unique()	Method extracts the unique values in the dataframe
<u>nunique()</u>	Method returns count of the unique values in the dataframe
value_counts()	Method counts the number of times each unique value occurs within the Series
columns()	Method returns the column labels of the DataFrame
axes()	Method returns a list representing the axes of the DataFrame
<u>isnull()</u>	Method creates a Boolean Series for extracting rows with null values
<u>notnull()</u>	Method creates a Boolean Series for extracting rows with non-null values
between()	Method extracts rows where a column value falls in between a predefined range

DataFrame Methods

<u>isin()</u>	Method extracts rows from a DataFrame where a column value exists in a predefined collection
dtypes()	Method returns a Series with the data type of each column. The result's index is the original DataFrame's columns
<u>astype()</u>	Method converts the data types in a Series
values()	Method returns a Numpy representation of the DataFrame i.e. only the values in the DataFrame will be returned, the axes labels will be removed
sort_values()- <u>Set1</u>, <u>Set2</u>	Method sorts a data frame in Ascending or Descending order of passed Column
<u>sort_index()</u>	Method sorts the values in a DataFrame based on their index positions or labels instead of their values but sometimes a data frame is made out of two or more data frames and, as a result, later index can be changed using this method
<u>loc[]</u>	Method retrieves rows based on index label
<u>iloc[]</u>	Method retrieves rows based on index position
<u>ix[]</u>	Method retrieves DataFrame rows based on either index label or index position. This method combines the best features of the .loc[] and .iloc[] methods
<u>rename()</u>	Method is called on a DataFrame to change the names of the index labels or column names
columns()	Method is an alternative attribute to change the column name
<u>drop()</u>	Method is used to delete rows or columns from a DataFrame
<u>pop()</u>	Method is used to delete rows or columns from a DataFrame
<u>sample()</u>	Method pulls out a random sample of rows or columns from a DataFrame
<u>nsmallest()</u>	Method pulls out the rows with the smallest values in a column

DataFrame Methods

<u>nlargest()</u>	Method pulls out the rows with the largest values in a column
<u>shape()</u>	Method returns a tuple representing the dimensionality of the DataFrame
<u>ndim()</u>	Method returns an 'int' representing the number of axes / array dimensions. Returns 1 if Series, otherwise returns 2 if DataFrame
<u>dropna()</u>	Method allows the user to analyze and drop Rows/Columns with Null values in different ways
<u>fillna()</u>	Method manages and let the user replace NaN values with some value of their own
<u>rank()</u>	Values in a Series can be ranked in order with this method
<u>query()</u>	Method is an alternate string-based syntax for extracting a subset from a DataFrame
<u>copy()</u>	Method creates an independent copy of a pandas object
<u>duplicated()</u>	Method creates a Boolean Series and uses it to extract rows that have duplicate values
<u>drop_duplicates()</u>	Method is an alternative option to identifying duplicate rows and removing them through filtering
<u>set_index()</u>	Method sets the DataFrame index (row labels) using one or more existing columns
<u>reset_index()</u>	Method resets index of a Data Frame. This method sets a list of integer ranging from 0 to length of data as index
<u>where()</u>	Method is used to check a Data Frame for one or more condition and return the result accordingly. By default, the rows not satisfying the condition are filled with NaN value