

Boolean Indexing with NumPy: Takeaways



by Dataquest Labs, Inc. - All rights reserved © 2020

Syntax

READING CSV FILES WITH NUMPY

- Reading in a CSV file:

```
import numpy as np

taxi = np.genfromtxt('nyctaxis.csv', delimiter=',', skip_header=1)
```

BOOLEAN ARRAYS

- Creating a Boolean array from filtering criteria:

```
np.array([2,4,6,8]) < 5
```

- Boolean filtering for 1D ndarray:

```
a = np.array([2,4,6,8])

filter = a < 5

a[filter]
```

- Boolean filtering for 2D ndarray:

```
tip_amount = taxi[:,12]

tip_bool = tip_amount > 50

top_tips = taxi[tip_bool, 5:14]
```

ASSIGNING VALUES

- Assigning values in a 2D ndarray using indices:

```
taxi[28214,5] = 1  
taxi[:,0] = 16  
taxi[1800:1802,7] = taxi[:,7].mean()
```

- Assigning values using Boolean arrays:

```
taxi[taxi[:, 5] == 2, 15] = 1
```

Concepts

- Selecting values from a ndarray using Boolean arrays is very powerful. Using Boolean arrays helps us think in terms of filters on the data, instead of specific index values (like we did when working with Python lists).

Resources

- [Reading a CSV file into NumPy](#)
- [Indexing and selecting data](#)



Takeaways by Dataquest Labs, Inc. - All rights reserved © 2020