9/12/24, 2:37 PM

#### Comparison Operators

INTERMEDIATE PYTHON



Data Scientist at DataCamp Hugo Bowne-Anderson



## NumPy recap

```
# Code from Intro to Python for Data Science, Chapter 4
                                                                                                                                                 np_weight = np.array([65.4, 59.2, 63.6, 88.4, 68.7])
                                                                                                np_height = np.array([1.73, 1.68, 1.71, 1.89, 1.79])
                                                                                                                                                                                                                                                                                                                                                                                                     array([ 21.852, 20.975, 21.75 , 24.747, 21.441])
                                                                                                                                                                                                        bmi = np\_weight / np\_height ** 2
                                                      import numpy as np
```

```
bmi > 23
```

```
array([False, False, True, False], dtype=bool)
bmi[bmi > 23]
```

array([ 24.747])

Comparison operators: how Python values relate



# Numeric comparisons

2 < 3

3 <= 3

True

True

2 == 3

× = 2

×

× ×

True

2 <= 3

False

True



# Other comparisons

"carl" < "chris"

True

3 < "chris"

TypeError: unorderable types: int() < str()

3 < 4.1

True



# Other comparisons

bmi

array([21.852, 20.975, 21.75 , 24.747, 21.441])

bmi > 23

array([False, False, False, True, False], dtype=bool)

#### R datacamp

PDF.js viewer

#### 9/12/24, 2:37 PM

## Comparators

Comparator	Meaning
V	Strictly less than
II V	Less than or equal
٨	Strictly greater than
II ^	Greater than or equal
II	Equal
<u></u>	Not equal

### R datacamp

#### R datacamp

# Let's practice! INTERMEDIATE PYTHON

9/12/24, 2:37 PM

# **Boolean Operators**

INTERMEDIATE PYTHON



Hugo Bowne-Anderson Data Scientist at DataCamp



9/12/24, 2:37 PM

PDF is viewer

# **Boolean Operators**

and

0 P

not



#### and

False and True True and True

True and False False x > 5 and x < 15# True x = 12True

False

False and False

True

False



True or True

True

False or True

True

True or False

True

False

False or False

y = 5 y < 7 or y > 13

True

not True

False

not False

True

INTERMEDIATE PYTHON

R datacamp

bmi # calculation of bmi left out

array([21.852, 20.975, 21.75 , 24.747, 21.441])

bmi > 21

array([True, False, True, True], dtype=bool)

bmi < 22

array([True, True, True, False, True], dtype=bool)

bmi > 21 and bmi < 22

ValueError: The truth value of an array with more than one element is ambiguous. Use a.any() or a.all()

#### R datacamp

#### NumPy

- logical\_and()
- logical\_or()
- logical\_not()

```
np.logical_and(bmi > 21, bmi < 22)
```

array([True, False, True, False, True], dtype=bool)

bmi[np.logical\_and(bmi > 21, bmi < 22)]</pre>

array([21.852, 21.75, 21.441])



# Let's practice!

9/12/24, 2:37 PM

## if, elif, else

INTERMEDIATE PYTHON

#### P

#### Hugo Bowne-Anderson Data Scientist at DataCamp

#### Overview

Comparison Operators

Boolean Operators

Conditional Statements

#### R datacamp

#### 4

if condition :
 expression

control.py

z = 4
if z % 2 == 0 : # True
print("z is even")

z is even



#### <u>+</u>

if condition :
 expression

expression not part of if

control.py

```
z = 4
if z % 2 == 0 : # True
print("z is even")
```

z is even



```
if condition :
    expression
```

#### control.py

```
z = 4
if z % 2 == 0 :
    print("checking " + str(z))
    print("z is even")
```

```
checking 4
z is even
```

#### R datacamp

```
expression
if condition :
```

#### control.py

```
print("checking " + str(z))
                    # False
                                                             print("z is even")
                    if z % 2 == 0 :
S = S
```

#### R datacamp

```
if condition :
```

expression

else:

expression

control.py

```
z = 5
if z % 2 == 0 :  # False
print("z is even")
```

else :

print("z is odd")

z is odd

R datacamp

PDF, js viewer

9/12/24, 2:37 PM

#### elif

```
elif condition :
                 expression
                                               expression
                                                                              expression
if condition :
                                                               else:
```

#### control.py

```
print("z is neither divisible by 2 nor by 3")
                                                 # False
                                                                                                 # True
                                              print("z is divisible by 2")
                                                                                              print("z is divisible by 3")
                                                                       elif z % 3 == 0 :
                       if z % 2 == 0 :
z = 3
```

z is divisible by 3



#### elif

```
elif condition :
                 expression
                                                expression
                                                                               expression
if condition :
                                                                else :
```

#### control.py

```
# Never reached
                                                                                                                                                    print("z is neither divisible by 2 nor by 3")
                                                  # True
                                                print("z is divisible by 2")
                                                                                                  print("z is divisible by 3")
                                                                          elif z % 3 == 0 :
                       if z % 2 == 0 :
9 = 2
```

z is divisible by 2



# Let's practice!

9/12/24, 2:37 PM

### Filtering pandas DataFrames

INTERMEDIATE PYTHON



Hugo Bowne-Anderson Data Scientist at DataCamp



#### brics

brics = pd.read\_csv("path/to/brics.csv", index\_col = 0) import pandas as pd brics

	country	capital	area	population
BR	Brazil	Brasilia	8.516	200.40
RU	Russia	Moscow	17.100	143.50
NI	India	New Delhi	3.286	1252.00
НЭ	China	Beijing	9.597	1357.00
SA	SA South Africa	Pretoria	1.221	52.98



PDF.js viewer

9/12/24, 2:37 PM

#### Goal

population	200.40	143.50	1252.00	1357.00	52.98
area	8.516	17.100	3.286	9.597	1.221
capital	Brasilia	Moscow	India New Delhi	Beijing	Pretoria
country	Brazil	Russia	India	China	South Africa
	BR	RU	NI	СН	SA

- Select countries with area over 8 million km2
- 3 steps
- Select the area column
- Do comparison on area column 0
- Use result to select countries



## Step 1: Get column

```
143.50
                                              52.98
area population
         200.40
                           1252.00
                                     1357.00
                  Moscow 17.100
                                     9.597
                            3.286
                                              1.221
                           New Delhi
         Brasilia
                                     Beijing
capital
                                               Pretoria
                                   CH China
SA South Africa
                            India
                   Russia
         Brazil
country
                         NI
                 RU
```

```
brics["area"]
```

```
# - Need Pandas Series
                                              Name: area, dtype: float64
           3.286
                     9.597
17.100
                                 1.221
                      H
 RU
```

#### Alternatives:

```
brics.loc[:,"area"]
brics.iloc[:,2]
```



## Step 2: Compare

```
brics["area"]
```

```
Name: area, dtype: float64
       17.100
3.286
9.597
1.221
8.516
```

```
brics["area"] > 8
```

```
Name: area, dtype: bool
True
          False
                     True
                             False
```

```
is_huge = brics["area"] > 8
```



## Step 3: Subset DF

is\_huge

```
Name: area, dtype: bool
True
True
False
                          True
                                  False
```

brics[is\_huge]

population	200.4	143.5	1357.0	
area	8.516	17.100	9.597	
capital	Brasilia	Moscow	Beijing 9.597	
country	BR Brazil	RU Russia	CH China	
	BR	RN	Н	



PDF.js viewer

9/12/24, 2:37 PM

### Summary

```
52.988
               143.50
                                 1357.00
                         1252.00
        200.40
population
        8.516
                Moscow 17.100
                         3.286
                                 9.597
                                          1.221
       Brasilia
                         New Delhi
                                 Beijing
                                           Pretoria
capital
               Russia
                         India
       Brazil
                                 China
                                          SA South Africa
country
               R
                         IN IN
        BR
```

```
is_huge = brics["area"] > 8
                           brics[is_huge]
```

```
1357.0
         200.4
                   143.5
area population
         8.516
                   Moscow 17.100
                             9.597
         Brasilia
                            Beijing
capital
         BR Brazil
RU Russia
country
                             China
                            끙
```

```
brics[brics["area"] > 8]
```

population	200.4	143.5	1357.0
area	8.516	17.100	9.597
capital	Brasilia	Moscow	Beijing 9.597
country	BR Brazil	Russia	China
	BR	R	품 -



PDF.js viewer

## 9/12/24, 2:37 PM

## **Boolean operators**

```
143.50
200.40
                     1252.00
                               1357.00
8.516
                               9.597
           Moscow 17.100
                     3.286
Brasilia
                     India New Delhi
                                          Pretoria
                               Beijing
BR Brazil
RU Russia
IN India N
CH China
SA South Africa
```

```
np.logical_and(brics["area"] > 8, brics["area"] < 10)
import numpy as np
```

```
Name: area, dtype: bool
            False
                        True
False
```

```
brics[np.logical_and(brics["area"] > 8, brics["area"] < 10)]
```

```
country capital area population
                   BR Brazil Brasilia 8.516
CH China Beiing 9.597
```



#### R datacamp

# Let's practice! INTERMEDIATE PYTHON