DS 100 – Intro to Data Science

Lecture 10– Iteration

02/18/2025

Adam Poliak

BRYN MAWR COLLEGE

Announcements

Lab04 (Simulations) due Friday

HW04 – Applying Functions & Iteration
- Due Wednesday (02/26)
- Short assignment

Checkpoint/Project 1:
- Paired assignment that covers the previous section of the course material
- Due Friday 02/28

Comparisons

Comparison Operators

Operator	Table predicate
==	are.equal_to
!=	are.not_equal_to
>	are.above
>=	are.above_or_equal_to
<	are.below
<=	are.below_or_equal_to

The result of a comparison expression is a bool value:

True, False

Comparison Operators

The result of a comparison expression is a bool value

x = 2   y = 3

Assignment Statements

x > 1   x > y   y >= 3
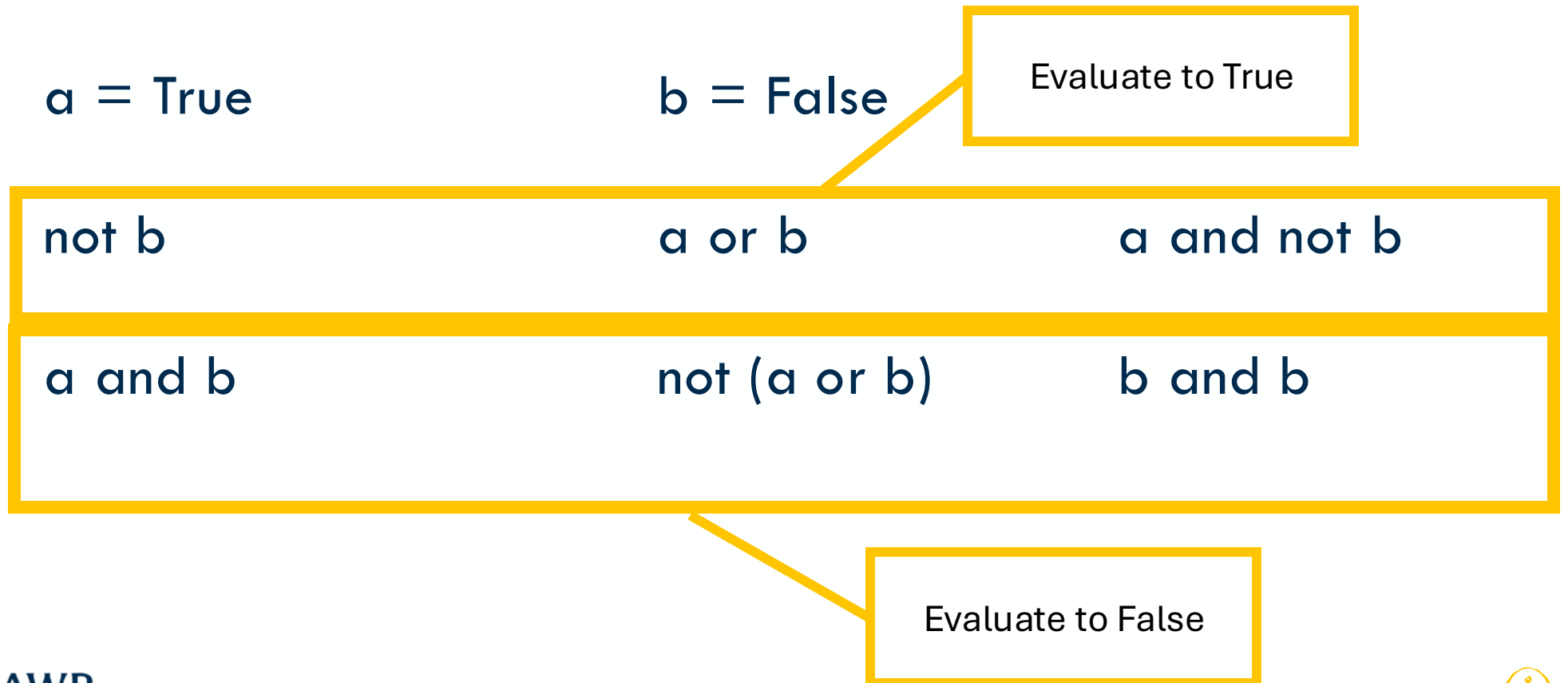
x == y   x != 2   2 < x < 5

Comparison Expressions

Combining Comparisons

The result of a comparison expression is a bool value

a = True   b = False   Evaluate to True

not b   a or b   a and not b

a and b   not (a or b)   b and b

Evaluate to False

BRYN MAWR
COLLEGE

Aggregating Comparisons

Summing an array or list of bool values count the number of True values

```
1  + 0  + 1
True + False + True
sum([1    , 0    , 1])
sum([True, False, True)]
```

Control Statements

brynmawr.edu

Control Statements

These statements *control* the sequence of computations that are performed

The keywords if and for begin control statements

The purpose of **if** is to define functions that choose different behavior based on their arguments

Control & Treatment

Treatment group

Control group

How do we determine causaility?

Random Selection

brynmawr.edu

Random Selection

np.random.choice

- Selects at random

- With replacement

- From an array

- A specific number of times

np.random.choice(some_array, sample_size)

Appending Arrays

brynmawr.edu

A longer array

`np.append(array_1, value)`:

- new array with value appended to array_1
- value has to be of the same type as elements of array_1

`np.append(array_1, array_2)`:

- new array with array_2 appended to array_1
- Elements of array_2 have to be of the same type as elements of array_1

Iteration

for statements

for is a keyword that begins a control statement

The purpose of for is to perform a computation for every element in a list or array

for *name* **in** *sequence:*
          BODY where we use the value in the name

Control Statements

These statements *control* the sequence of computations that are performed

- The keywords **if** and **for** begin control statements
- The purpose of **if** is to define functions that choose different behavior based on their arguments
- The purpose of **for** is to perform a computation for every element in a list or array

for *name* **in** *sequence:*
        BODY where we use the value in the name

BRYN MAWR
COLLEGE

Experimentation & Simulation

BRYN MAWR COLLEGE

brynmawr.edu

Experimentation

Why do we want to run experiments?
    To test a hypothesis

What could a hypothesis be in our previous (coin flipping) experiments?
    Is a coin fair or not?
    …

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study
      Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement
      Textbook lingo: figure how out to simulate the statistic

Step 3: Choose how many times to simulate the statistic
      Textbook lingo: Number of Repetitions

Step 4: Do it!
      Textbook lingo: simulate multiple values

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study
      Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement
      Textbook lingo: figure how out to simulate the statistic

Step 3: Choose how many times to simulate the statistic
      Textbook lingo: Number of Repetitions

Step 4: Do it!
      Textbook lingo: simulate multiple values

func   for   np.append()

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study
   Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement   func
   Textbook lingo: figure how out to simulate the statistic

Step 3: Choose how many times to simulate the statistic
   Textbook lingo: Number of Repetitions

Step 4: Do it!
   Textbook lingo: simulate multiple values

                              for   np.append()

BRYN MAWR
COLLEGE

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study
      Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement
      Textbook lingo: figure how out to simulate the statistic   func

Step 3: Choose how many times to simulate the statistic   for
      Textbook lingo: Number of Repetitions

Step 4: Do it!
      Textbook lingo: simulate multiple values

np.append()

BRYN MAWR
COLLEGE

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study
        Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement   func
        Textbook lingo: figure how out to simulate the statistic

Step 3: Choose how many times to simulate the statistic   for
        Textbook lingo: Number of Repetitions

Step 4: Do it!   np.append()
        Textbook lingo: simulate multiple values

BRYN MAWR
COLLEGE

How to simulate multiple values

Collection array
> empty array to store the simulated values/statistic
> make_array()

Create a "repetitions sequence"
> A sequence as long as the number of iterations
> For *n* repetitions, use the sequence np.arange(n)

Create a for loop. For each element:
> Simulate *one* value by using the function you wrote in Step 2
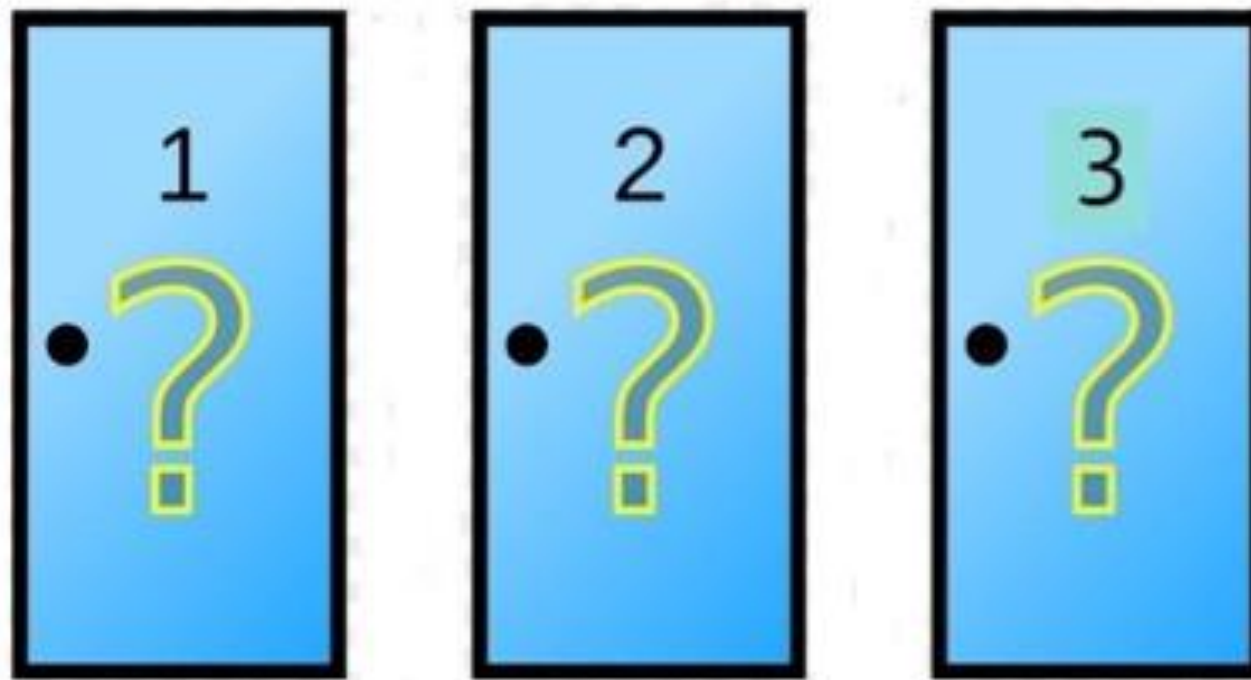> Augment the collection array with this simulated value

Monty Hall Problem

Monty Hall Problem

Monty Hall Problem