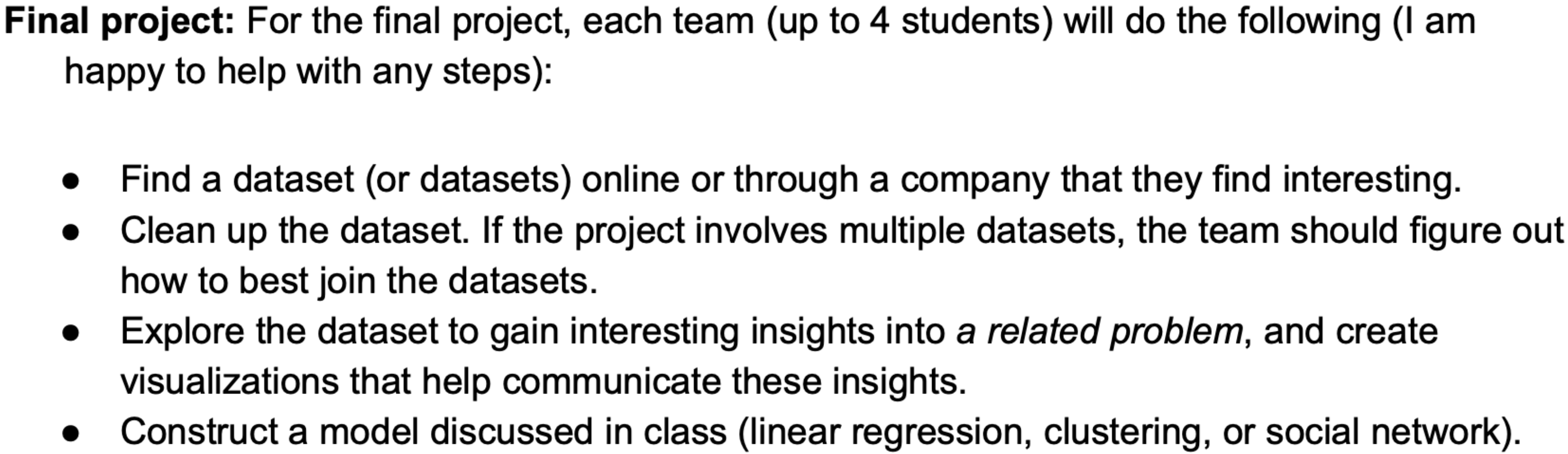
IE 0015: Introduction to Data Analytics

Lecture 6: More Python

Announcements

- Questions from last week?
- Final project
- Assignments
  - Lab 2 is due by midnight this Friday
  - Homework 1 will be posted ***today*** (due 2/19)
  - Midterm is 2/19

Announcements

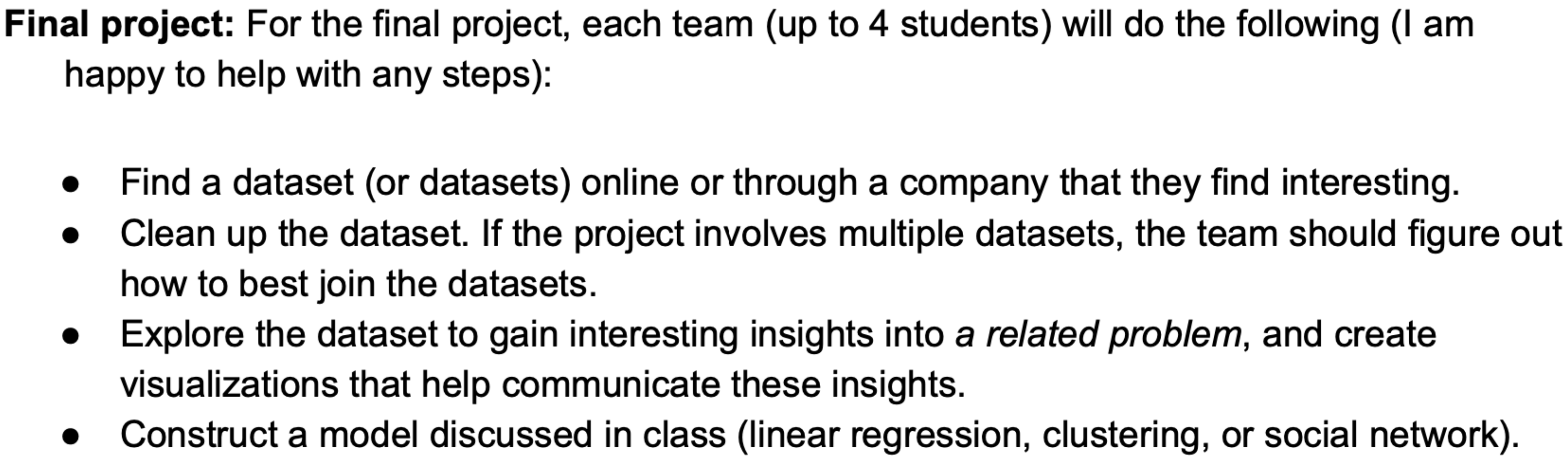


Final project: For the final project, each team (up to 4 students) will do the following (I am happy to help with any steps):

- Find a dataset (or datasets) online or through a company that they find interesting.
- Clean up the dataset. If the project involves multiple datasets, the team should figure out how to best join the datasets.
- Explore the dataset to gain interesting insights into *a related problem*, and create visualizations that help communicate these insights.
- Construct a model discussed in class (linear regression, clustering, or social network).

Announcements *(extra credit)*

+ 2 / 100 to final grade:



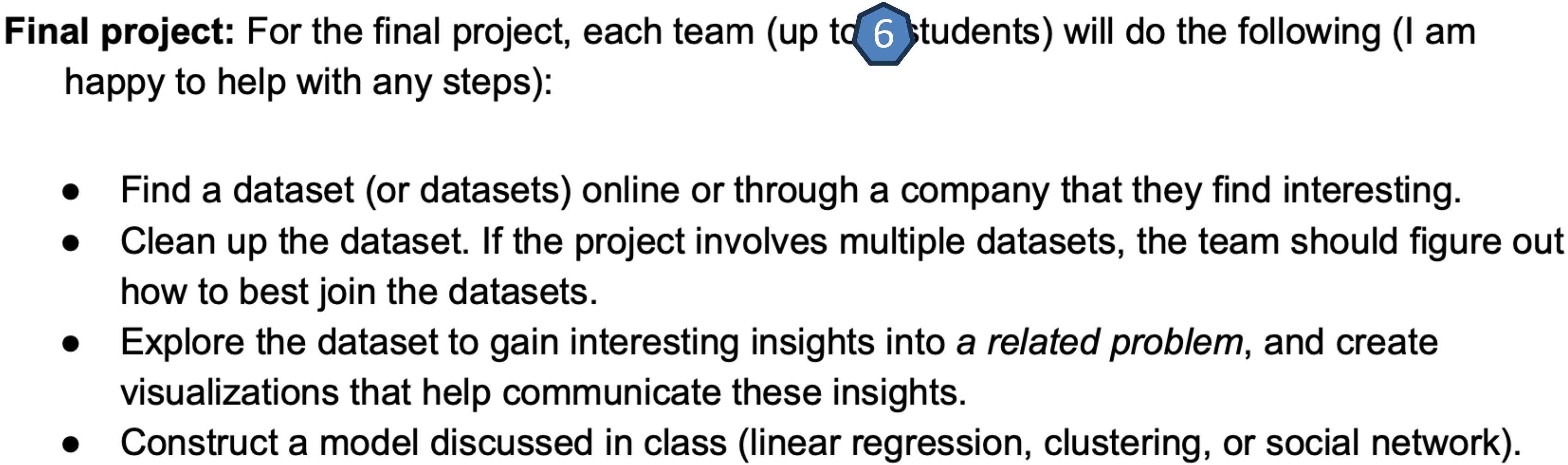
Final project: For the final project, each team (up to 4 students) will do the following (I am happy to help with any steps):

- Find a dataset (or datasets) online or through a company that they find interesting.
- Clean up the dataset. If the project involves multiple datasets, the team should figure out how to best join the datasets.
- Explore the dataset to gain interesting insights into *a related problem*, and create visualizations that help communicate these insights.
- Construct a model discussed in class (linear regression, clustering, or social network).

- Build something that works

Announcements

+ 2 / 100 to final grade:



Final project: For the final project, each team (up to 6 students) will do the following (I am happy to help with any steps):

- Find a dataset (or datasets) online or through a company that they find interesting.
- Clean up the dataset. If the project involves multiple datasets, the team should figure out how to best join the datasets.
- Explore the dataset to gain interesting insights into *a related problem*, and create visualizations that help communicate these insights.
- Construct a model discussed in class (linear regression, clustering, or social network).

- Build something that works

Announcements

Final project 1:

Deliverable 1 (3/17):

- Data exploration, hypothesis + abstract

Deliverable 2:

- Final report

Final project 2:

Deliverable 1 (3/17):

- Idea + plan, hypothesis + abstract

Deliverable 2:

- Final report + code (any attempt!)

Announcements

Final project 1:

Deliverable 1:

- Data exploration, hypothesis + abstract

Deliverable 2:

- Final report

Final project 2:

Deliverable 1:

- Idea + plan, hypothesis + abstract

Deliverable 2:

- Final report + code (any attempt!)

- Can I bring an existing project: ***YES***

Announcements

Final project 1:

Deliverable 1:

- Data exploration, hypothesis + abstract

Deliverable 2:

- Final report

Final project 2:

Deliverable 1:

- Idea + plan, hypothesis + abstract

Deliverable 2:

- Final report + code (any attempt!)

- Can I bring research: ***YES***

Announcements

Final project 1:

Deliverable 1:

- Data exploration, hypothesis + abstract

Deliverable 2:

- Final report

Final project 2:

Deliverable 1:

- Idea + plan, hypothesis + abstract

Deliverable 2:

- Final report + code (any attempt!)

- Can I work with my BFF: ***YES***

Announcements

Final project 1:

Deliverable 1:

- Data exploration, hypothesis + abstract

Deliverable 2:

- Final report

Final project 2:

Deliverable 1:

- Idea + plan, hypothesis + abstract

Deliverable 2:

- Final report + code (any attempt!)

- Will Luca help us: ***YES***

Announcements

Final project 1:

Deliverable 1:

- Data exploration, hypothesis + abstract

Deliverable 2:

- Final report

Final project 2:

Deliverable 1:

- Idea + plan, hypothesis + abstract

Deliverable 2:

- Final report + code (any attempt!)

- Will we lose points if the code doesn't work: ***NO***
- You should be ambitious, use ChatGPT, you'll get graded for data analytics and creativity

New poll ***(5 participation points)***

Announcements

- Today’s lecture
    - If statements

Refresher

Is python interpreted or compiled?

Refresher

What is a compiler?

a program that reads in a whole source code file, and translates it to assembly code…

What is an interpreter?

a program that reads in a source code file line by line,

running each line…

How does this affect my languages?

Python is almost always interpreted, so compilation and execution are hidden (by the interpreter), C and C++ are compiled.

Refresher

What is a variable?

Refresher

What is a variable?

A symbolic name for data, with an associated storage location in memory

Name + storage

Refresher

What is a type?

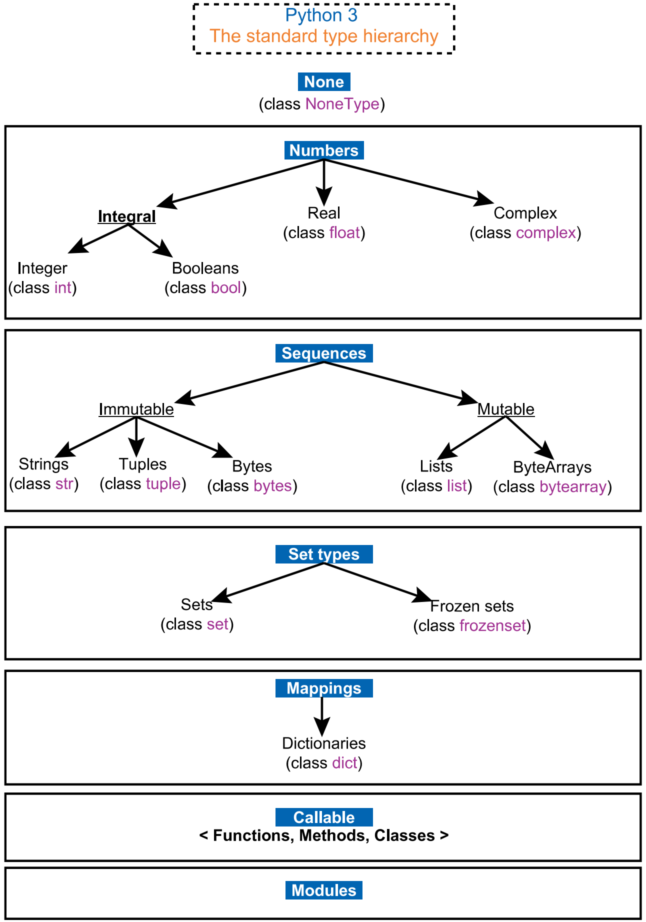
Refresher

What is a type?

Grouping or categorization of data values

Refresher ***(Python)***

- Python is dynamically-typed
- You don’t really know what this means
- The alternative is statically-typed



Python 3
The standard type hierarchy

None
(class NoneType)

Numbers

Integral
Integer
(class int)
Booleans
(class bool)

Real
(class float)

Complex
(class complex)

Sequences

Immutable
Strings
(class str)
Tuples
(class tuple)
Bytes
(class bytes)

Mutable
Lists
(class list)
ByteArrays
(class bytearray)

Set types

Sets
(class set)
Frozen sets
(class frozenset)

Mappings

Dictionaries
(class dict)

Callable
< Functions, Methods, Classes >

Modules

What are we going to learn next?

What might we learn next?

✓ memory and disk (manipulate + save data)

✓ libraries / packages (leverage others’ code)

…

What are we going to learn next?

What might we learn next?

- ✓ *memory and disk (manipulate + save data)*
- ✓ *libraries / packages (leverage others' code)*
- ✓ *control flow (take your program down different paths)*
- ✓ *loops and containers (efficient operations & storage)*
- ✓ *functions and classes (reusable work)*

What are we going to learn next?

control flow *(take your program down different paths)*

What are we going to learn next?

***control flow** (take your program down different paths)*

Control flow and expressions

What is an expression?

combination of operators and operands that is interpreted to produce some other value

Control flow and expressions

What is an expression?

combination of operators and operands that is interpreted to produce some other value

Expressions

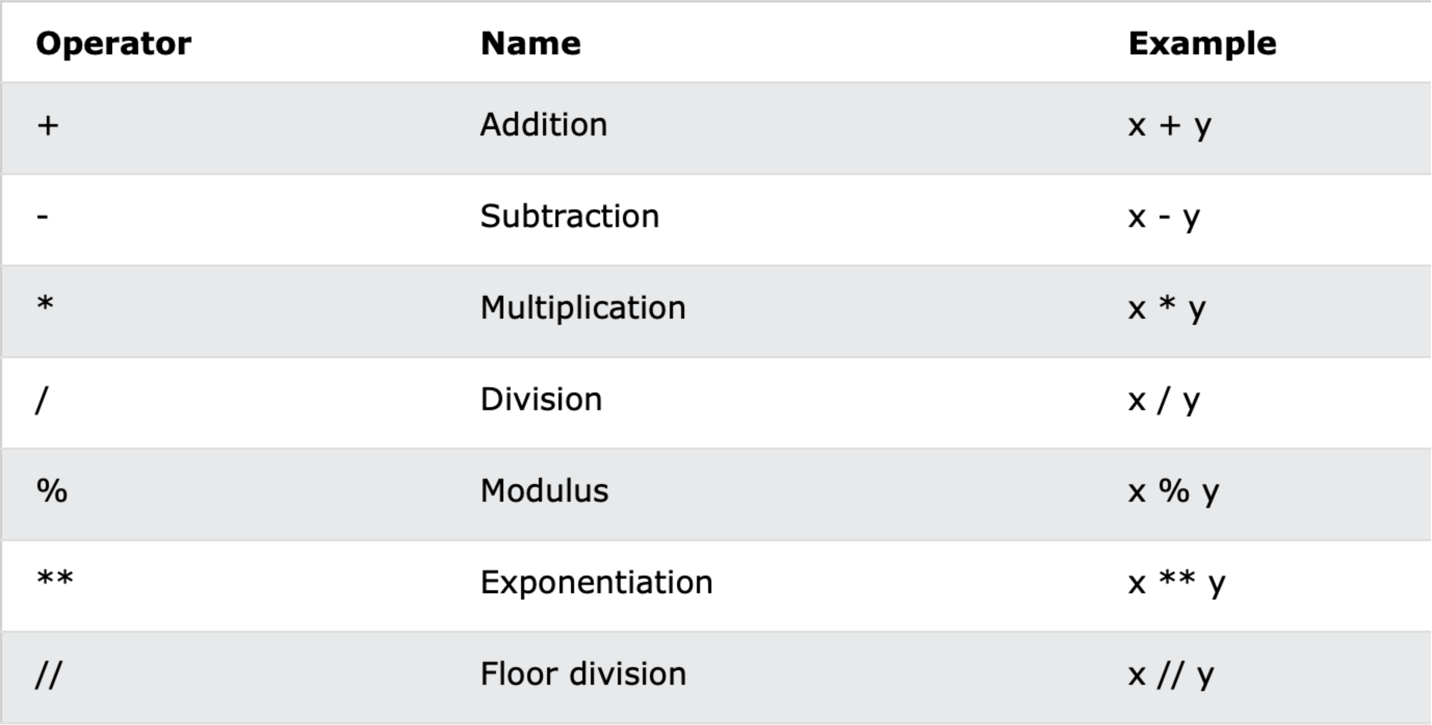
Operators

- Literally, operations
- Adding, subtracting, basic math
- Binary stuff – truth tables
- We’ll see the two examples in a second

Operands

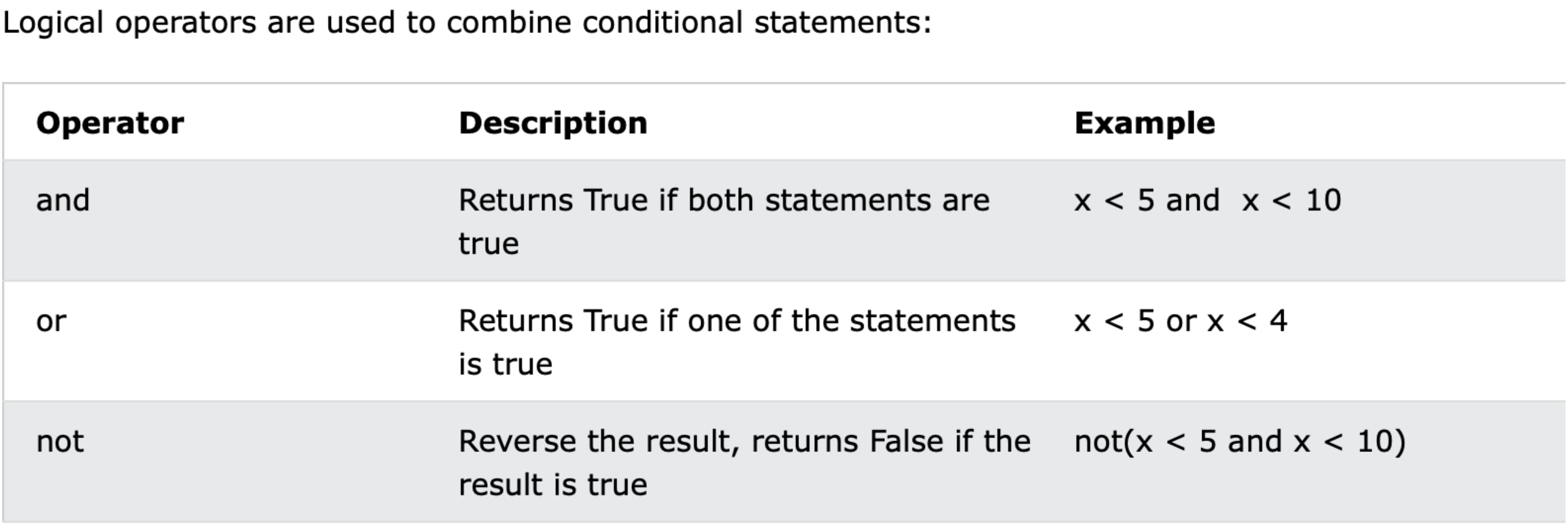
- Variables

Arithmetic operators



Operator	Name	Example
+	Addition	x + y
-	Subtraction	x - y
*	Multiplication	x * y
/	Division	x / y
%	Modulus	x % y
**	Exponentiation	x ** y
//	Floor division	x // y

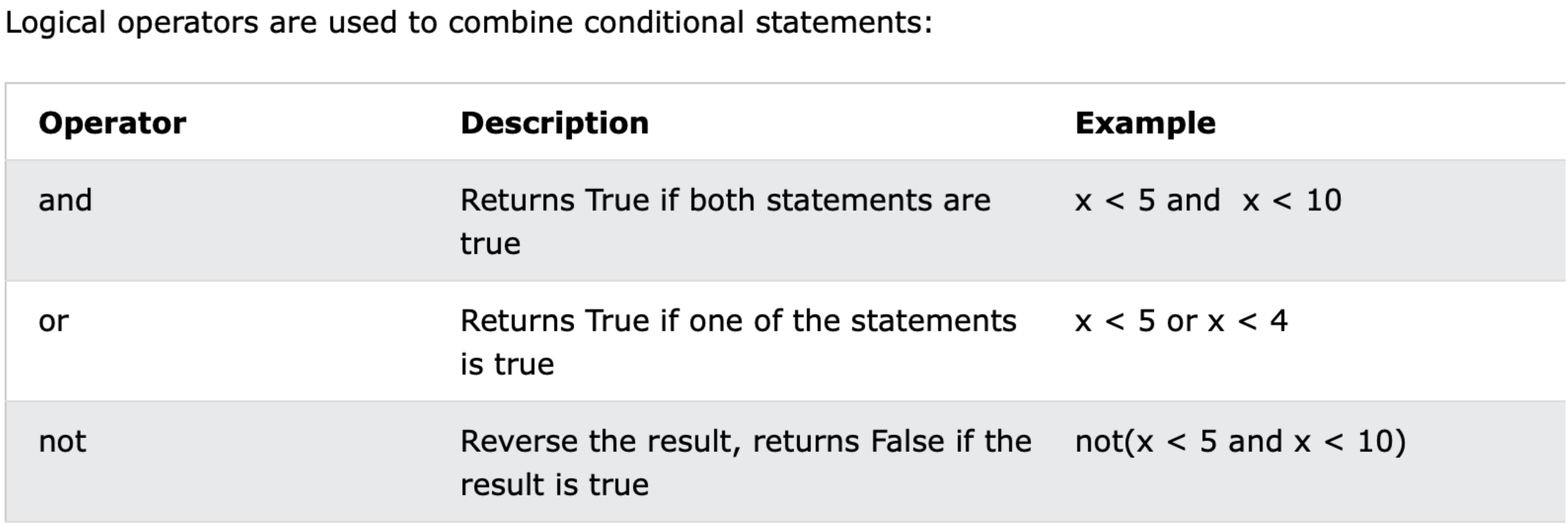
Binary / logical operators



Logical operators are used to combine conditional statements:

Operator	Description	Example
and	Returns True if both statements are true	x < 5 and x < 10
or	Returns True if one of the statements is true	x < 5 or x < 4
not	Reverse the result, returns False if the result is true	not(x < 5 and x < 10)

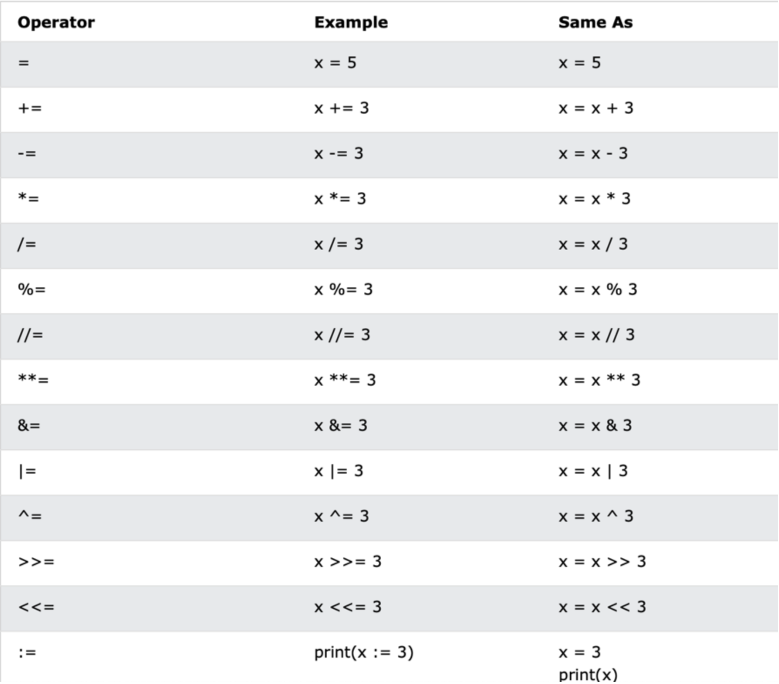
Binary / logical operators



Logical operators are used to combine conditional statements:

Operator	Description	Example
and	Returns True if both statements are true	x < 5 and x < 10
or	Returns True if one of the statements is true	x < 5 or x < 4
not	Reverse the result, returns False if the result is true	not(x < 5 and x < 10)

Assignment operators



Operator	Example	Same As
=	x = 5	x = 5
+=	x += 3	x = x + 3
-=	x -= 3	x = x - 3
*=	x *= 3	x = x * 3
/=	x /= 3	x = x / 3
%=	x %= 3	x = x % 3
//=	x //= 3	x = x // 3
**=	x **= 3	x = x ** 3
&=	x &= 3	x = x & 3
|=	x |= 3	x = x | 3
^=	x ^= 3	x = x ^ 3
>>=	x >>= 3	x = x >> 3
<<=	x <<= 3	x = x << 3

:=	print(x := 3)	x = 3
		print(x)

https://www.w3schools.com/python/python_operators.asp

Summary

Is an expression a variable?

What is the type of an expression?

Summary

Is an expression a variable?

No, but its result is. i.e., the expression itself is never stored in memory, but it produces a variable.

What is the type of an expression?

Summary

Is an expression a variable?

No, but its result is. i.e., the expression itself is never stored in memory, but it produces a variable.

What is the type of an expression?

Expression is evaluated at runtime by the interpreter, we refer to "its type" and the type of its return value interchangeably (though this is slightly informal).

Summary

Quite python specific!!!

Is an expression a variable?

No, but its result is. i.e., the expression itself is never stored in memory, but it produces a variable.

What is the type of an expression?

Expression is evaluated at runtime by the interpreter, we refer to "its type" and the type of its return value interchangeably (though this is slightly informal).

On to live coding

Let’s go see some expressions…

Recap

- Expressions are combinations of operands and operations
- Expressions are not variables, but ***can*** return variables
- Expressions are everywhere

If statements

```
if <boolean expression>:
        <some code>

<some more code>
```

If statements

```
if <boolean expression>:
        <some code>

<some more code>
```

if expression is True

If statements

```
if <boolean expression>:
          <some code>
else:
          <some other code>

<some more code>
```

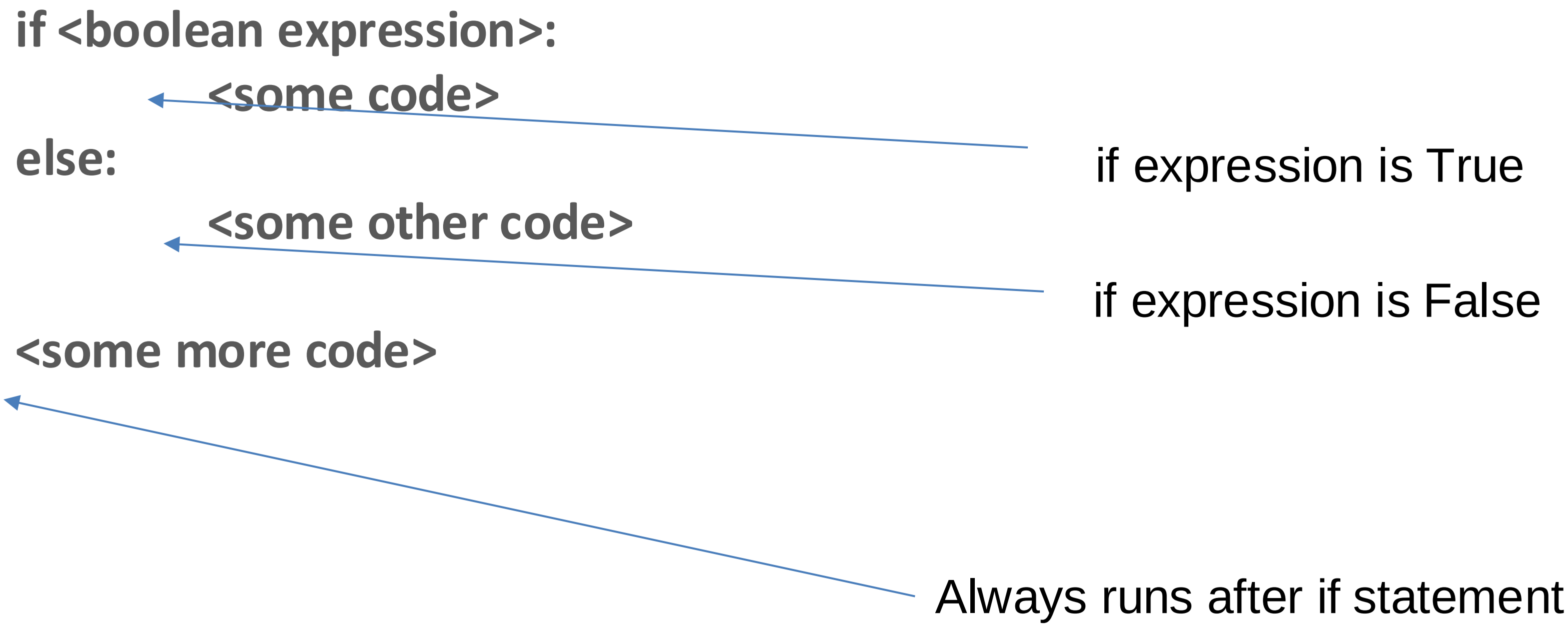
if expression is True

If statements

```
if <boolean expression>:
        <some code>
else:
        <some other code>

<some more code>
```

```
if <boolean expression>:
        <some code>

<some more code>
```

If statements



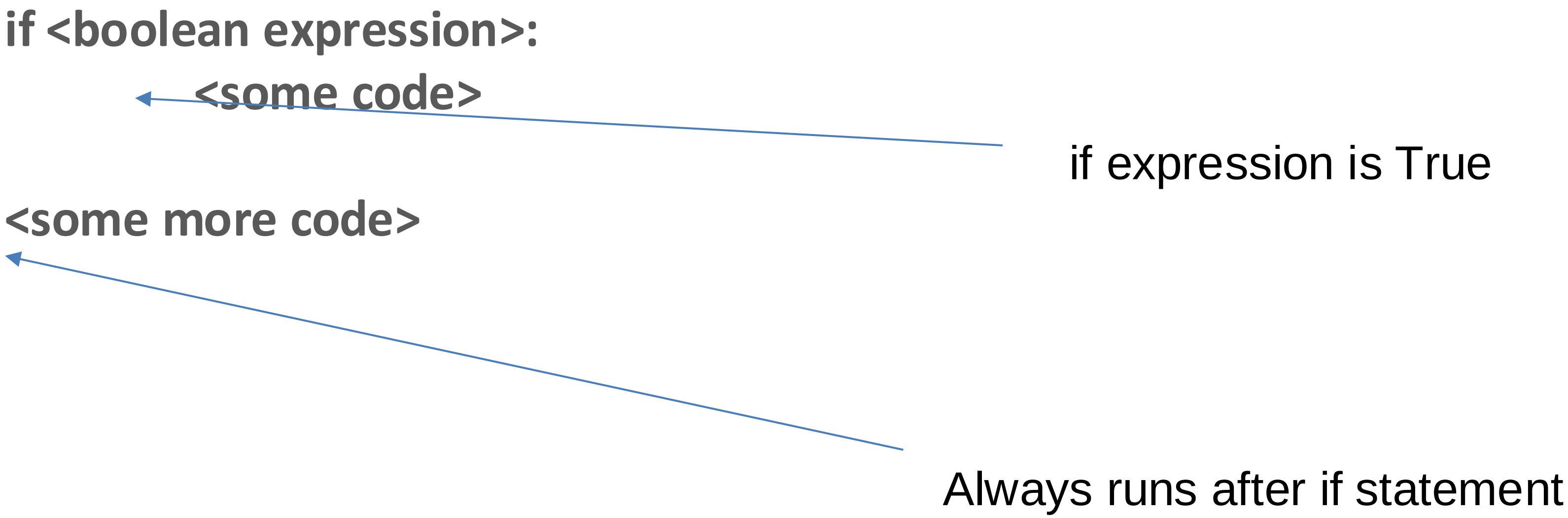
```
if <boolean expression>:
        <some code>
else:
        <some other code>

<some more code>
```

if expression is True

if expression is False

Always runs after if statement

If statements



```
if <boolean expression>:
        <some code>

<some more code>
```

if expression is True

Always runs after if statement

If statements

```
if <boolean expression1>:
        <some code>

elif <boolean expression2>:
        <some other code>

elif <boolean expression3>:
        <some new code>

<some more code>
```

Will only run if expression 1 is false, and expression 2 is true

If statements

```
if <boolean expression1>:
        <some code>

elif <boolean expression2>:
        <some other code>

elif <boolean expression3>:
        <some new code>

<some more code>
```

Will only run if expression 1 is false, and expression 2 is true

Will only run if expression 1 is false, and expression 2 is false, and expression 3 is true

If statements

```
if <boolean expression1>:
        <some code>

elif <boolean expression2>:
        <some other code>

elif <boolean expression3>:
        <some new code>

else:
        <some crazyyy code>

<some more code>
```

Will only run if expression 1 is false, and expression 2 is true

Will only run if expression 1 is false, and expression 2 is false, and expression 3 is false

if - conceptually

What does the keyword **“if”** mean to the interpreter?

if - conceptually

What does the keyword **“if”** mean to the interpreter?

…HINT: you are the interpreter. You start reading a new line to execute. You read the first word of the line; it is **“if”.**

You know you will have to execute some special instructions that are *related to the next lines…*

for loops

What is an iterable?

an object capable of returning its members one at a time

for loops

What is an iterable?

an object capable of returning its members one at a time

- Lists
- Dictionaries
- Pandas series

for loops

```
for item in sequence:
    print(item)
```

for loops

```
for item in <sequence>:
      print(item)
```

Sequence must be an iterable object

What is “item”?