



DATA 1202
Spring 2024

Lecture 3

Tables

Announcements

- HW 1 is due tonight, Wednesday (1/17) @ 11pm
- Lab 2 is Thursday (1/18)
 - Lab 2 is due Friday (1/19)

Weekly Goals

- Monday
 - No Class
- Wednesday
 - Tables

Python

Python

- Python is popular both for data science & general software development
- Mastering the language fundamentals is critical
- Learn through practice:
 - See some examples & learn the rules
 - Try out variants of those examples yourself
 - Write new code that solves new problems

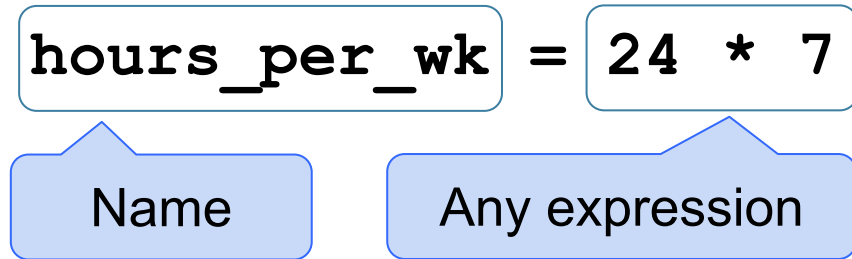
(Demo)

Review of Python Concepts

- An expression evaluates to a value
- Values can be numbers or strings (text); we'll see lots of other kinds of values soon
- The syntax (format) of the language is very rigid — even an extra space can cause a syntax error
- There is particular behavior associated with built-in operators that you need to learn (e.g., dividing produces 8.0 instead of 8)

Names

Assignment Statements



- An assignment statement changes the meaning of the name to the left of the = symbol
- The name is bound to the value of the expression to the right of the = symbol (its current value; not the equation)

(Demo)

Functions

Anatomy of a Call Expression

What
function
to call

Argument to the
function

f(27)

"Call f on 27."

Anatomy of a Call Expression

What
function
to call

First argument

Second
argument

`max` (`15`, `27`)

(Demo)

Review of Function Concepts

- Some functions require a particular number of arguments (e.g., **abs** must be called on one value)
- Arguments can be named in the call expression:
round(number=12.34)
But the names must match the documentation
- Type a ? after a function name to see its documentation

Tables

Table Structure

- A Table is a sequence of labeled columns
- Each row represents one individual
- Data within a column represents one attribute of the individuals

The diagram illustrates a table structure with three columns: Name, Code, and Area (m2). The first row contains 'California', 'CA', and '163696'. The second row contains 'Nevada', 'NV', and '110567'. Annotations include a 'Label' bubble pointing to the 'Code' header, a 'Row' bubble pointing to the 'Nevada' cell, and a 'Column' bubble pointing to the 'NV' cell. A blue box highlights the 'Nevada' row, and another blue box highlights the 'Code' column. The text '(Demo)' is located at the bottom right of the table area.

Name	Code	Area (m2)
California	CA	163696
Nevada	NV	110567

(Demo)

Some Table Operations

- `t.select(label)` - constructs a new table with just the specified columns
- `t.drop(label)` - constructs a new table in which the specified columns are omitted
- `t.sort(label)` - constructs a new table with rows sorted by the specified column
- `t.where(label, condition)` - constructs a new table with just the rows that match the condition