DATA 1202

Spring 2024

Lecture 4

Data Types

Announcements

- **HW 2** is due Wed 1/24 @ 11pm
- **Lab 3** is due Friday 1/26 @ 5pm

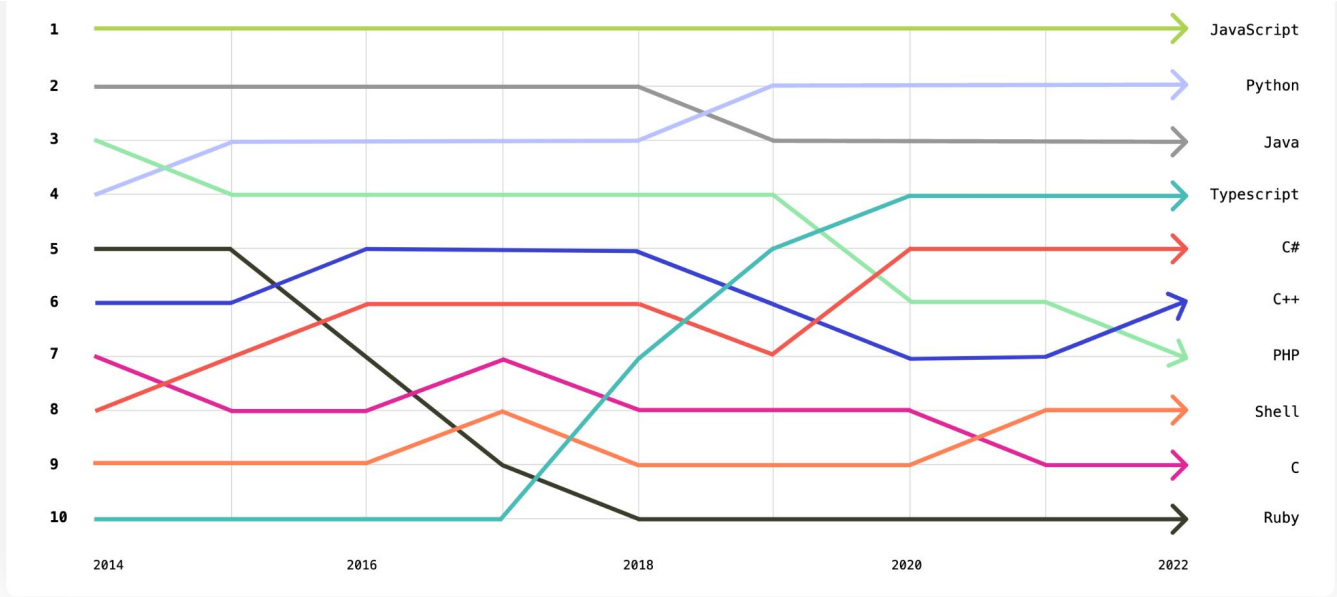
Weekly Goals

- Monday
  - Data Types

- Wednesday
  - Building Tables
  - Example of Table: Census Data

Recap of the last lecture

- Python: one of the most popular programming language for data science and machine learning



Github's Octoverse Survey 2022

Python

- Assignments and Expressions
- Functions
  - Function’s Anatomy: name, arguments, returned value
- Table structure
  - Sequence of labeled columns

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

(Demo)

Numbers

(Demo)

Ints and Floats

Python has two real number types

- `int`: an integer of **any** size
- `float`: a number with an optional fractional part

An `int` never has a decimal point; a **float** always does

A `float` might be printed using scientific notation

Three limitations of float values:

- They have limited size (but the limit is huge)
- They have limited precision of 15-16 decimal places
- After arithmetic, the final few decimal places can be wrong

Strings

(Demo)

Text and Strings

A string value is a snippet of text of any length

- `'a'`
- `'word'`
- `"there can be 2 sentences. Here's the second!"`

Strings consisting of numbers can be converted to numbers

- `int('12')`
- `float('1.2')`

Any value can be converted to a string

- `str(5)`

Discussion Question

Assume you have run the following statements:

```
x = 3
y = '4'
z = '5.6'
```

What's the source of the error in each example?

A. `x + y`
B. `x + int(y + z)`
C. `str(x) + int(y)`
D. `y + float(z)`

Types

(Demo)

Every value has a type

We've seen 5 types so far:

- `int: 2`
- `float: 2.2`
- `str: 'Identity theft is not a joke'`
- `builtin_function_or_method: abs`
- `Table`

The `type` function can tell you the type of a value

- `type(2)`
- `type(2 + 2)`

An expression's "type" is based on its value, not how it looks

- `x = 2`
- `type(x)`

Arrays

(Demo)

Arrays

An array contains a sequence of values

- All elements of an array should have the same type
- Arithmetic is applied to each element individually
- Adding two arrays adds the corresponding elements (but the arrays must be the same length!)
- A column of a table is an array