Workshop: **Introduction to Python**

Overview

Christian C. Luhmann
Stony Brook University

Where to Find these Slides

github.com/cluhmann/python-psych-workshop

Who am I?

- B.S. in Computer Science
- Ph.D. in Psychology
- Stony Brook University
- Decision-making, learning, methods (stats & “cognitive modeling”)
- Using Python since ~2002

Who are You?

- Faculty/students?
- Who has used…
  - Matlab?
  - R?
  - Some other programming language (e.g., Java, C)?
  - SPSS?
  - Eprime?
  - SAS?

Goals

- Appreciation of the **ends**
  - benefits of Python
  - functionality provided by Python and its ecosystem
  - how to integrate these tools into your existing workflow
- Non-goals of this workshop: **means**
  - Ability to program Python without further consultation
  - Encyclopedic knowledge of packages, APIs, etc.
- Think of this as a open house
  - If you’d like buy, you still need to move all your stuff

What I will assume of you…

- Not much
- You’re not terrified of programming
- You use data in your research
- You’re looking for tools to conduct efficient, flexible, reproducible (maybe sharable) analyses
- You (maybe) conduct laboratory experiments

Why?

- Why bother to learn another thing?
  - We already have Matlab, R, etc.
- Why Python?
- Python…
  - is general-purpose
  - is free and open source
  - is eminently readable (i.e., readily learned)
  - has an extensive, well-integrated ecosystem of tools
  - and more!
- This workshop, hopefully, is a more comprehensive answer

What is Python?

- Developed by Guido van Rossum in the early 1990s
- Python 2.0 was released October 16th, 2000
- Python 3.0 was released December 3rd, 2008

Python

- Free and open source
- Cross-platform
- Widely-used and well-supported
- Well-documented
- Multiple options for boosting performance
- **Highly readable**
- Substantial **standard library**
- Vibrant **third-party ecosystem**

Standard Library

```
>>> abs(-42)
42
```

```
>>> pow(2, 10)
1024
```

Standard Library

```
>>> min([1, 4, 12, 42])
1
```

```
>>> max([1, 4, 12, 42])
42
```

```
>>> len([1, 4, 12, 42])
4
```

```
>>> sum([1, 3, 5])
9
```

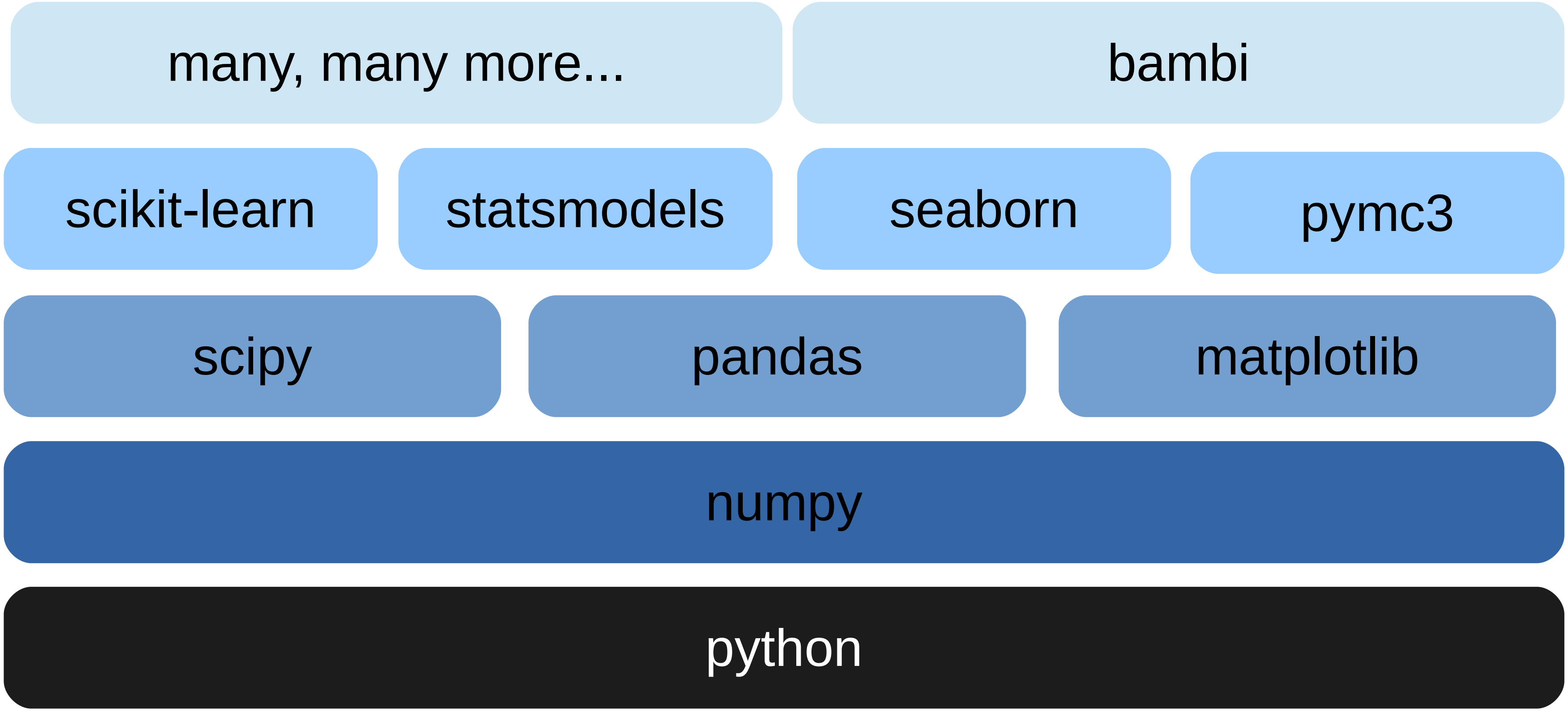
```
>>> sorted([2, 4, 6, 8, 1, 3, 5, 7, 9])
[1, 2, 3, 4, 5, 6, 7, 8, 9]
```

Standard Library

```
>>> print(‘Six times nine is ‘ + str(6*9))
Six times nine is 54
```

```
>>> file = open(‘myfile.txt’, ‘r’)
>>> contents = file.read()
>>> print(contents)
First line of my file.
Second line of my file.
Last line of my file.
```

Python's Ecosystem



many, many more...

bambi

scikit-learn

statsmodels

seaborn

pymc3

scipy

pandas

matplotlib

numpy

python

Installing Python

- **Anaconda**
- Enthought’s Canopy
- WinPython (Windows only)
- Each of these provides:
  - Python
  - Packages
  - Package manager
  - Editor (IDE)
  - Other tools

Anaconda

www.anaconda.com/download

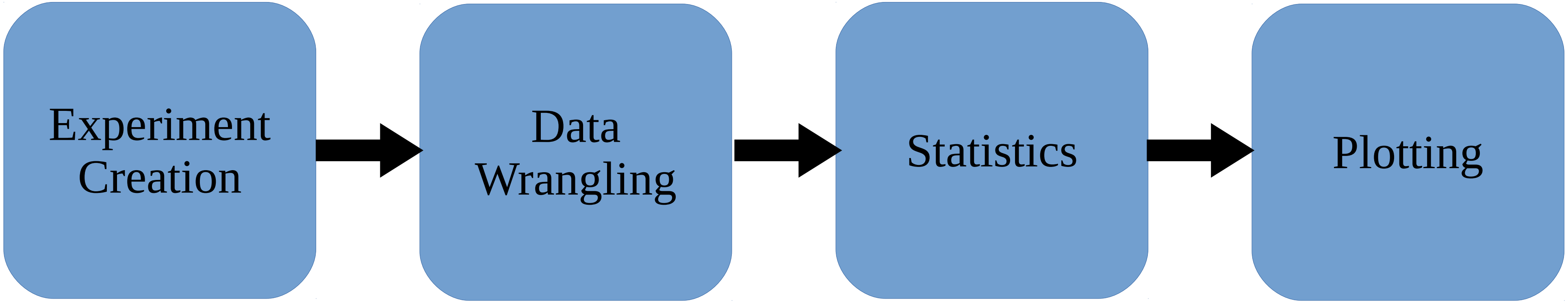
Installing Python

- Python 2.x or 3.x?

- Python 2.7's end-of-life initially 2015, but postponed to 2020
  - concern that much existing code could not easily be ported to Python 3

- Python 3.x is recommended
  - But you may see lots of references to 2.x (e.g., on StackOverflow)

The Pipeline



Experiment Creation

Data Wrangling

Statistics

Plotting

Outline

1. Overview
2. Ways of using Python
3. Python basics
4. Data set overview
5. Data wrangling
6. Statistics
7. Plotting
8. Experiment creation