

Effective computation in physics : field guide to research with Python / Anthony Scopatz and Kathryn D. Huff. – First edition june 2015, first release. – Beijing ; Boston ; Farnham ; Sebastopol ; Tokyo, 2015

Spis treści

Foreword	xv
Preface	xvii
Part I. Getting Started	
1. Introduction to the Command Line	1
Navigating the Shell	1
The Shell Is a Programming Language	2
Paths and pwd	3
Home Directory (~)	5
Listing the Contents (ls)	6
Changing Directories (cd)	7
File Inspection (head and tail)	10
Manipulating Files and Directories	11
Creating Files (nano, emacs, vi, cat, >, and touch)	11
Copying and Renaming Files (cp and mv)	17
Making Directories (mkdir)	18
Deleting Files and Directories (rm)	18
Flags and Wildcards	20
Getting Help	21
Reading the Manual (man)	21
Finding the Right Hammer (apropos)	24
Combining Utilities with Redirection and Pipes (>, >>, and)	25
Permissions and Sharing	26
Seeing Permissions (ls -l)	26
Setting Ownership (chown)	28
Setting Permissions (chmod)	29
Creating Links (ln)	29
Connecting to Other Computers (ssh and scp)	30
The Environment	31
Saving Environment Variables (.bashrc)	33
Running Programs (PATH)	34
Nicknaming Commands (alias)	36
Scripting with Bash	36
Command Line Wrap-up	38

2. Programming Blastoff with Python	39
Running Python	40
Comments	41
Variables	42
Special Variables	44
Boolean Values	45
None Is Not Zero!	45
NotImplemented Is Not None!	45
Operators	46
Strings	49
String Indexing	50
String Concatenation	53
String Literals	54
String Methods	55
Modules	57
Importing Modules	58
Importing Variables from a Module	58
Aliasing Imports	59
Aliasing Variables on Import	59
Packages	60
The Standard Library and the Python Ecosystem	62
Python Wrap-up	63
3. Essential Containers	65
Lists	66
Tuples	70
Sets	71
Dictionaries	73
Containers Wrap-up	75
4. Flow Control and Logic	77
Conditionals	77
if-else Statements	80
if- elif- else Statements	81
if-else Expression	82
Exceptions	82
Raising Exceptions	84
Loops	85
while Loops	86
for Loops	88
Comprehensions	90
Flow Control and Logic Wrap-up	93
5. Operating with Functions	95
Functions in Python	96

Keyword Arguments	99
Variable Number of Arguments	101
Multiple Return Values	103
Scope	104
Recursion	107
Lambdas	108
Generators	109
Decorators	112
Function Wrap-up	116
6. Classes and Objects	117
Object Orientation	118
Objects	119
Classes	123
Class Variables	124
Instance Variables	126
Constructors	127
Methods	129
Static Methods	132
Duck Typing	133
Polymorphism	135
Decorators and Metaclasses	139
Object Orientation Wrap-up	141
Part II. Getting It Done	
7. Analysis and Visualization	145
Preparing Data	145
Experimental Data	149
Simulation Data	150
Metadata	151
Loading Data	151
NumPy	152
PyTables	153
Pandas	153
Blaze	155
Cleaning and Munging Data	155
Missing Data	158
Analysis	159
Model-Driven Analysis	160
Data-Driven Analysis	162
Visualization	162
Visualization Tools	164
Gnuplot	164
matplotlib	167

Bokeh	172
Inkscape	174
Analysis and Visualization Wrap-up	175
8. Regular Expressions	177
Messy Magnetism	178
Metacharacters on the Command Line	179
Listing Files with Simple Patterns	180
Globally Finding Filenames with Patterns (find)	182
grep, sed, and awk	187
Finding Patterns in Files (grep)	188
Finding and Replacing Patterns in Files (sed)	190
Finding and Replacing a Complex Pattern	192
sed Extras	193
Manipulating Columns of Data (awk)	195
Python Regular Expressions	197
Regular Expressions Wrap-up	199
9. NumPy: Thinking in Arrays	201
Arrays	202
dtypes	204
Slicing and Views	208
Arithmetic and Broadcasting	211
Fancy Indexing	215
Masking	217
Structured Arrays	220
Universal Functions	223
Other Valuable Functions	226
NumPy Wrap-up	228
10. Storing Data: Files and HDF5	229
Files in Python	230
An Aside About Computer Architecture	235
Big Ideas in HDF5	237
File Manipulations	239
Hierarchy Layout	242
Chunking	245
In-Core and Out-of-Core Operations	249
In-Core	249
Out-of-Core	250
Querying	252
Compression	252
HDF5 Utilities	254
Storing Data Wrap-up	255

11. Important Data Structures in Physics	747
Hash Tables	258
Resizing	259
Collisions	261
Data Frames	263
Series	264
The Data Frame Structure	266
B-Trees	269
K-D Trees	272
Data Structures Wrap-up	277
12. Performing in Parallel	279
Scale and Scalability	280
Problem Classification	282
Example: N-Body Problem	284
No Parallelism	285
Threads	290
Multiprocessing	296
MPI	300
Parallelism Wrap-up	307
13. Deploying Software	309
Deploying the Software Itself	311
pip	312
Conda	316
Virtual Machines	319
Docker	321
Deploying to the Cloud	325
Deploying to Supercomputers	327
Deployment Wrap-up	329
Part III. Getting It Right	
14. Building Pipelines and Software	333
make	334
Running make	337
Makefiles	337
Targets	338
Special Targets	340
Building and Installing Software	341
Configuration of the Makefile	343
Compilation	345
Installation	346
Building Software and Pipelines Wrap-up	346

15. Local Version Control	349
What Is Version Control?	349
The Lab Notebook of Computational Physics	350
Version Control Tool Types	351
Getting Started with Git	352
Installing Git	352
Getting Help (git -help)	352
Control the Behavior of Git (git config)	354
Local Version Control with Git	355
Creating a Local Repository (git init)	355
Staging Files (git add)	357
Checking the Status of Your Local Copy (git status)	357
Saving a Snapshot (git commit)	358
git log: Viewing the History	361
Viewing the Differences (git diff)	362
Unstaging or Reverting a File (git reset)	363
Discard Revisions (git revert)	364
Listing, Creating, and Deleting Branches (git branch)	365
Switching Between Branches (git checkout)	366
Merging Branches (git merge)	367
Dealing with Conflicts	369
Version Control Wrap-Up	369
16. Remote Version Control	371
Repository Hosting (github.com)	371
Creating a Repository on GitHub	373
Declaring a Remote (git remote)	373
Sending Commits to Remote Repositories (git push)	374
Downloading a Repository (git clone)	375
Fetching the Contents of a Remote (git fetch)	379
Merging the Contents of a Remote (git merge)	380
Pull = Fetch and Merge (git pull)	380
Conflicts	381
Resolving Conflicts	382
Remote Version Control Wrap-up	384
17. Debugging	385
Encountering a Bug	386
Print Statements	387
Interactive Debugging	389
Debugging in Python (pdb)	390
Setting the Trace	391
Stepping Forward	392
Querying Variables	393
Setting the State	393

Running Functions and Methods	394
Continuing the Execution	394
Breakpoints	395
Profiling	396
Viewing the Profile with pstats	396
Viewing the Profile Graphically	397
Line Profiling with Kernprof	400
Linting	401
Debugging Wrap-up	402

18. Testing **403**

Why Do We Test?	404
When Should We Test?	405
Where Should We Write Tests?	405
What and How to Test?	406
Running Tests	409
Edge Cases	409
Corner Cases	410
Unit Tests	412
Integration Tests	414
Regression Tests	416
Test Generators	417
Test Coverage	418
Test-Driven Development	419
Testing Wrap-up	422

Part IV. Getting It Out There

19. Documentation **427**

Why Prioritize Documentation?	427
Documentation Is Very Valuable	428
Documentation Is Easier Than You Think	429
Types of Documentation	429
Theory Manuals	430
User and Developer Guides	431
Readme Files	431
Comments	432
Self-Documenting Code	434
Docstrings	435
Automation	436
Sphinx	437
Documentation Wrap-up	440

20. Publication **441**

Document Processing	441
---------------------	-----

Separation of Content from Formatting	442
Tracking Changes	443
Text Editors	443
Markup Languages	444
LaTeX	445
Bibliographies	456
Publication Wrap-up	459
21. Collaboration	461
Ticketing Systems	462
Workflow Overview	462
Creating an Issue	464
Assigning an Issue	466
Discussing an Issue	467
Closing an Issue	468
Pull Requests and Code Reviews	468
Submitting a Pull Request	469
Reviewing a Pull Request	469
Merging a Pull Request	470
Collaboration Wrap-up	470
22. Licenses, Ownership, and Copyright	471
What Is Copyrightable?	472
Right of First Publication	473
What Is the Public Domain?	473
Choosing a Software License	474
Berkeley Software Distribution (BSD) License	475
GNU General Public License (GPL)	477
Creative Commons (CC)	479
Other Licenses	480
Changing the License	482
Copyright Is Not Everything	483
Licensing Wrap-up	485
23. Further Musings on Computational Physics	487
Where to Go from Here	487
Glossary	493
Bibliography	499
Index	503