

Into the wild

MCP 743

classes 1-3

numbers
strings
lists
tuples
for loops
while loops
if/else statements
files

class topics covered

class 5

dictionaries
functions
classes
scope

class 7

modules in the wild
database survey
define your project

class 4

files
dictionaries
functions/class

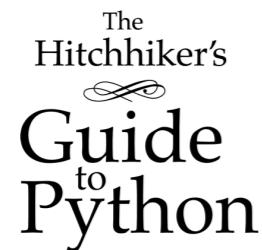
class 6

classes
modules
PYTHON PATH

class 8

Excel module
finish your project
visit from CCS

Hitchhiker's Guide to Python: a great resource



This opinionated guide exists to provide both novice and expert Python developers a best practice handbook to the installation, configuration, and usage of Python on a daily basis.

Stay Informed

Receive updates on new releases and upcoming projects.

[Follow @kennethreitz](#)

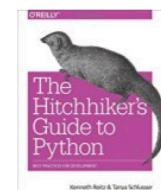
[Follow @kennethreitz](#)

[Join Mailing List](#)

[Say Thanks!](#)

O'Reilly Book

This guide is now available in tangible book form!



All proceeds being directly donated to the [DjangoGirls](#) organization.

Other Projects

More Kenneth Reitz projects:

[edmsynths.com](#)

[pipenv](#)

[pep8.org](#)

[Requests: HTTP for Humans](#)

[Maya: Datetimes for Humans](#)

[Records: SQL for Humans](#)

[Legit: Git for Humans](#)

[Tablib: Tabular Datasets](#)

Contributors

This guide is the result of the collaboration of [hundreds](#) of people around the world, and your

The Hitchhiker's Guide to Python!

Greetings, Earthling! Welcome to The Hitchhiker's Guide to Python.

This is a living, breathing guide. If you'd like to contribute, [fork us on GitHub!](#)

This handcrafted guide exists to provide both novice and expert Python developers a best practice handbook to the installation, configuration, and usage of Python on a daily basis.

This guide is **opinionated** in a way that is almost, but not quite, entirely *unlike* Python's official documentation. You won't find a list of every Python web framework available here. Rather, you'll find a nice concise list of highly recommended options.

Let's get started! But first, let's make sure you know where your towel is.

Getting Started with Python

New to Python? Let's properly setup up your Python environment.

- [Picking an Interpreter](#)
 - [The State of Python \(3 & 2\)](#)
 - [Recommendations](#)
 - [So.... 3?](#)
 - [Implementations](#)
- Properly Install Python
 - [Properly Installing Python](#)
 - [Installing Python on Mac OS X](#)
 - [Installing Python on Windows](#)
 - [Installing Python on Linux](#)

Writing Great Python Code

This part of the guide focuses on the best-practices for writing Python code.

- [Structuring Your Project](#)
 - [Structure of the Repository](#)
 - [Structure of Code is Key](#)
 - [Modules](#)
 - [Packages](#)
 - [Object-oriented programming](#)
 - [Decorators](#)
 - [Context Managers](#)
 - [Dynamic typing](#)
 - [Mutable and immutable types](#)
 - [Vendorizing Dependencies](#)
 - [Runners](#)
 - [Further Reading](#)
- [Code Style](#)
 - [General concepts](#)
 - [Idioms](#)
 - [Zen of Python](#)
 - [PEP 8](#)
 - [Conventions](#)

Fork me on GitHub

v: latest ▾

Excel modules

Working with Excel Files in Python

This site contains pointers to the best information available about working with [Excel](#) files in the [Python](#) programming language.

The Packages

There are python packages available to work with Excel files that will run on any Python platform and that do not require either Windows or Excel to be used. They are fast, reliable and open source:

openpyxl

The recommended package for reading and writing Excel 2010 files (ie: .xlsx)

[Download](#) | [Documentation](#) | [Bitbucket](#)

xlsxwriter

An alternative package for writing data, formatting information and, in particular, charts in the Excel 2010 format (ie: .xlsx)

[Download](#) | [Documentation](#) | [GitHub](#)

xlrd

This package is for reading data and formatting information from older Excel files (ie: .xls)

[Download](#) | [Documentation](#) | [GitHub](#)

xlwt

This package is for writing data and formatting information to older Excel files (ie: .xls)

[Download](#) | [Documentation](#) | [Examples](#) | [GitHub](#)

xlutils

This package collects utilities that require both xlrd and xlwt, including the ability to copy and modify or filter existing excel files.

NB: In general, these use cases are now covered by openpyxl!

[Download](#) | [Documentation](#) | [GitHub](#)

The Mailing List / Discussion Group

There is a [Google Group](#) dedicated to working with Excel files in Python, including the libraries listed above along with manipulating the Excel application via COM.

Commercial Development

The following companies can provide commercial software development and consultancy and are specialists in working with Excel files in Python:



Comments

You do not have permission to add comments.

[Sign in](#) | [Recent Site Activity](#) | [Report Abuse](#) | [Print Page](#) | Powered By [Google Sites](#)

Hitchhiker's Guide to Python: science modules



This opinionated guide exists to provide both novice and expert Python developers a best practice handbook to the installation, configuration, and usage of Python on a daily basis.

Get Updates

Receive updates on new releases and upcoming projects.

[Follow @kennethreitz](#)

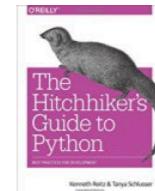
[Follow @kennethreitz](#)

[Say Thanks!](#)

[Join Mailing List](#)

O'Reilly Book

This guide is now available in tangible book form!



All proceeds are being directly donated to the [Django Girls](#) organization.

Other Projects

[More Kenneth Reitz projects](#):

[edmsynths.com](#)

[pipenv](#)

[pep8.org](#)

[Requests: HTTP for Humans](#)

[Maya: Datetimes for Humans](#)

[Records: SQL for Humans](#)

[Legit: Git for Humans](#)

[Tablib: Tabular Datasets](#)

Translations

[English](#)

[French](#)

[Chinese](#)

[Japanese](#)

[Korean](#)

Scientific Applications

Context

Python is frequently used for high-performance scientific applications. It is widely used in academia and scientific projects because it is easy to write and performs well.

Due to its high performance nature, scientific computing in Python often utilizes external libraries, typically written in faster languages (like C, or FORTRAN for matrix operations). The main libraries used are [NumPy](#), [SciPy](#) and [Matplotlib](#). Going into detail about these libraries is beyond the scope of the Python guide. However, a comprehensive introduction to the scientific Python ecosystem can be found in the [Python Scientific Lecture Notes](#)

Tools

IPython

[IPython](#) is an enhanced version of Python interpreter, which provides features of great interest to scientists. The *inline mode* allows graphics and plots to be displayed in the terminal (Qt based version). Moreover, the *notebook mode* supports literate programming and reproducible science generating a web-based Python notebook. This notebook allows you to store chunks of Python code along side the results and additional comments (HTML, LaTeX, Markdown). The notebook can then be shared and exported in various file formats.

Libraries

NumPy

[NumPy](#) is a low level library written in C (and FORTRAN) for high level mathematical functions. NumPy cleverly overcomes the problem of running slower algorithms on Python by using multidimensional arrays and functions that operate on arrays. Any algorithm can then be expressed as a function on arrays, allowing the algorithms to be run quickly.

NumPy is part of the SciPy project, and is released as a separate library so people who only need the basic requirements can use it without installing the rest of SciPy.

NumPy is compatible with Python versions 2.4 through to 2.7.2 and 3.1+.

Numba

[Numba](#) is a NumPy aware Python compiler (just-in-time (JIT) specializing compiler) which compiles annotated Python (and NumPy) code to LLVM (Low Level Virtual Machine) through special decorators. Briefly, Numba uses a system that compiles Python code with LLVM to code which can be natively executed at runtime.

SciPy

[SciPy](#) is a library that uses NumPy for more mathematical functions. SciPy uses NumPy arrays as the basic data structure, and comes with modules for various commonly used tasks in scientific programming, including linear algebra, integration (calculus), ordinary differential equation solving and signal processing.

Matplotlib

Fork me on GitHub

v: latest ▾

Hitchhiker's Guide to Python: science modules

[Japanese](#)
[Korean](#)

Table Of Contents

Scientific Applications

- [Context](#)
- [Tools](#)
 - [IPython](#)
- [Libraries](#)
 - [NumPy](#)
 - [Numba](#)
 - [SciPy](#)
 - [Matplotlib](#)
 - [Pandas](#)
 - [Rpy2](#)
 - [PsychoPy](#)
- [Resources](#)
 - [Unofficial Windows Binaries for Python Extension Packages](#)
 - [Anaconda](#)
 - [Canopy](#)

Quick search

Enter search terms or a module, class or function name.

Matplotlib

[Matplotlib](#) is a flexible plotting library for creating interactive 2D and 3D plots that can also be saved as manuscript-quality figures. The API in many ways reflects that of [MATLAB](#), easing transition of MATLAB users to Python. Many examples, along with the source code to re-create them, are available in the [matplotlib gallery](#).

Pandas

[Pandas](#) is data manipulation library based on Numpy which provides many useful functions for accessing, indexing, merging and grouping data easily. The main data structure (DataFrame) is close to what could be found in the R statistical package; that is, heterogeneous data tables with name indexing, time series operations and auto-alignment of data.

Rpy2

[Rpy2](#) is a Python binding for the R statistical package allowing the execution of R functions from Python and passing data back and forth between the two environments. Rpy2 is the object oriented implementation of the [Rpy](#) bindings.

PsychoPy

[PsychoPy](#) is a library for cognitive scientists allowing the creation of cognitive psychology and neuroscience experiments. The library handles presentation of stimuli, scripting of experimental design and data collection.

Resources

Installation of scientific Python packages can be troublesome, as many of these packages are implemented as Python C extensions which need to be compiled. This section lists various so-called scientific Python distributions which provide precompiled and easy-to-install collections of scientific Python packages.

Unofficial Windows Binaries for Python Extension Packages

Many people who do scientific computing are on Windows, yet many of the scientific computing packages are notoriously difficult to build and install on this platform. [Christoph Gohlke](#) however, has compiled a list of Windows binaries for many useful Python packages. The list of packages has grown from a mainly scientific Python resource to a more general list. If you're on Windows, you may want to check it out.

Anaconda

[Continuum Analytics](#) offers the [Anaconda Python Distribution](#) which includes all the common scientific Python packages as well as many packages related to data analytics and big data. Anaconda itself is free, and Continuum sells a number of proprietary add-ons. Free licenses for the add-ons are available for academics and researchers.

Canopy

[Canopy](#) is another scientific Python distribution, produced by [Enthought](#). A limited 'Canopy Express' variant is available for free, but Enthought charges for the full distribution. Free licenses are available for academics.

©2016. A [Kenneth Reitz](#) Project. CC BY-NC-SA 3.0.

SciPy



Install



Getting Started



Documentation



Report Bugs



SciPy Central



Blogs

SciPy (pronounced "Sigh Pie") is a Python-based ecosystem of open-source software for mathematics, science, and engineering. In particular, these are some of the core packages:



NumPy
Base N-dimensional array package



SciPy library
Fundamental library for scientific computing



Matplotlib
Comprehensive 2D Plotting



IPython
Enhanced Interactive Console



Sympy
Symbolic mathematics



pandas
Data structures & analysis

[More information...](#)

[About SciPy](#)
[Install](#)
[Getting Started](#)
[Documentation](#)
[Bug Reports](#)
[Topical Software](#)
[Citing](#)
[SciPy Central](#) ▾
[Cookbook](#) ▾
[SciPy Conferences](#) ▾
[Blogs](#) ▾
[NumFOCUS](#) ▾

CORE PACKAGES:
[NumPy](#) ▾
[SciPy library](#) ▾
[Matplotlib](#) ▾
[IPython](#) ▾
[Sympy](#) ▾
[Pandas](#) ▾

Search

News

- NumPy 1.12.0 released** See [Obtaining NumPy & SciPy libraries.](#)
(2017-01-15)
- NumPy 1.11.3 released** See [Obtaining NumPy & SciPy libraries.](#)
(2016-12-18)
- SciPy 0.18.1 released** See [Obtaining NumPy & SciPy libraries.](#)
(2016-09-19)
- SciPy 0.18.0 released** See [Obtaining NumPy & SciPy libraries.](#)
(2016-07-25)
- SciPy 0.17.1 released** See [Obtaining NumPy & SciPy libraries.](#)
(2016-05-12)
- SciPy 0.17.0 released** See [Obtaining NumPy & SciPy libraries.](#)
(2016-01-23)
- SciPy 0.16.1 released** See [Obtaining NumPy & SciPy libraries.](#)
(2015-10-24)
- SciPy 0.16.0 released** See [Obtaining NumPy & SciPy libraries.](#)
(2015-07-23)

[Past news...](#)

Scikit-image



Home Download Gallery Documentation Community Guidelines Source

Search documentation ...

Image processing in Python

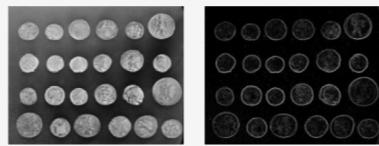
scikit-image is a collection of algorithms for image processing. It is available **free of charge and free of restriction**. We pride ourselves on high-quality, peer-reviewed code, written by an active **community of volunteers**.

[Download](#)

Getting Started

Filtering an image with `scikit-image` is easy! For more examples, please visit our [gallery](#).

```
from skimage import data, io, filters
image = data.coins()
# ... or any other NumPy array!
edges = filters.sobel(image)
io.imshow(edges)
io.show()
```



If you find this project useful, please cite:

[[BiTeX](#)]

Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu and the scikit-image contributors. **scikit-image: Image processing in Python.** PeerJ 2:e453 (2014) <http://dx.doi.org/10.7717/peerj.453>

Announcements

- Release! Version 0.12.0 06/03/2016
- Release! Version 0.11.0 04/03/2015
- Release! Version 0.10.0 27/05/2014
- Pre-print of the scikit-image paper: <https://peerj.com/preprints/336/>
- Release! Version 0.9.0 19/10/2013
- Release! Version 0.8.0 04/03/2013
- Release! Version 0.7.0 30/09/2012
- EuroSciPy Sprint, Belgium, August 2012
- SciPy 2012 Sprint, Austin, July 2012

Developers

- [Pull requests](#)
- [Bug reports](#)
- [Ohloh summary](#)

Stable (release notes)

x.y.z

[Download](#)

Development

pre-x.y.z

[Download](#)

[G+1](#) 296 [Star](#) 1,458

Links

- [Issue tracker](#)
[Mailing list](#)
[Test results](#)

Related Projects

- [OpenCV](#)
[Scikit-learn](#)
[Mahotas](#)
[SimpleCV](#)
[SimpleITK](#)
[Ilastik](#)

Data science in Python

Safari File Edit View History Bookmarks Window Help

analyticsvidhya.com

Home Blog Jobs Trainings Learning Paths Discuss DataHack f t g+ in

Analytics Vidhya Learn Everything About Analytics

DATAFEST 2017 25+ Live Events \$10,000+ Cash Prizes Analytics Vidhya Presents DATAFEST 2017 20+ Speakers 4+ Meetups Register Now

HOME BLOG JOBS TRAININGS DISCUSS LEARNING PATHS DATAHACK WRITE FOR US DATAFEST 2017 CONTACT US

Home > Learning paths – Data Science, Business Analytics, Business Intelligence and Big Data > Comprehensive learning path – Data Science in Python

Comprehensive learning path – Data Science in Python

CERTIFICATION COURSES IN BIG DATA & DATA SCIENCE ATTEND FREE TRIAL CLASSES HADOOP SPARK MONGODB DATA ANALYTICS MACHINE LEARNING TABLEAU upx Move up in life

Journey from a Python noob to a Kaggle on Python

So, you want to become a data scientist or may be you are already one and want to expand your tool repository. You have landed at the right place. The aim of this page is to provide a comprehensive learning path to people new to python for data analysis. This path provides a comprehensive overview of steps you need to learn to use Python for data analysis. If you already have some background, or don't need all the components, feel free to adapt your own paths and let us know how you made changes in the path.

You can also check the mini version of this learning path -> [Infographic: Quick Guide to learn Data Science in Python](#)

Step 0: Warming up

Before starting your journey, the first question to answer is:

Why use Python?

or

How would Python be useful?

Watch the first 30 minutes of this [talk from Jeremy](#), Founder of DataRobot at PyCon 2014, Ukraine to get an idea of how useful Python could be.

Step 1: Setting up your machine

Now that you have made up your mind, it is time to set up your machine. The easiest way to proceed is to just [download Anaconda](#) from Continuum.io . It comes packaged with most of the things you will need ever. The major downside of taking this route is that you will need to wait for Continuum to update their packages, even when there might be an update available to the underlying libraries. If you are a starter, that should hardly matter.

If you face any challenges in installing, you can find more [detailed instructions for various OS](#) here

TOP AV USERS

Rank	Name	Points
1	SRK	6105
2	Aayushmnit	5182
3	vopani	5150
4	Nalin Pasricha	4417
5	binga	3644

More Rankings

GREAT LAKES INSTITUTE OF MANAGEMENT Global Mindset - Indian Roots Bangalore Batch Starts - March 2017 Know more

G Suite GET STARTED Gmail Sheets Google Drive Google Calendar

POPULAR POSTS

A Complete Tutorial to Learn Data Science

<https://www.analyticsvidhya.com/learning-paths-data-science-business-analytics-business-intelligence-big-data/learning-path-data-science-python/>

Exemplary databases

general informatics resources

NCBI Resources How To Sign in to NCBI

NCBI National Center for Biotechnology Information

All Databases Search

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News](#) | [Blog](#)

Submit Deposit data or manuscripts into NCBI databases 

Download Transfer NCBI data to your computer 

Learn Find help documents, attend a class or watch a tutorial 

Popular Resources

- [PubMed](#)
- [Bookshelf](#)
- [PubMed Central](#)
- [PubMed Health](#)
- [BLAST](#)
- [Nucleotide](#)
- [Genome](#)
- [SNP](#)
- [Gene](#)
- [Protein](#)
- [PubChem](#)

NCBI Announcements

[NCBI Insights | PubMed Citations: A New, Faster Process for Correcting Errors](#) 28 Feb 2017

[The latest blog post on NCBI Insights](#)

[March 1st NCBI Minute: Setting up new data alerts with MyNCBI](#) 23 Feb 2017

[Next Wednesday, March 1, 2017, NCBI will present a short webinar that will show](#)

[Bottlenose dolphin annotation release 101](#) 22 Feb 2017

[Annotation Release 101 for the bottlenose dolphin \(*Tursiops truncatus*\) is](#)

[More...](#)

You are here: NCBI > National Center for Biotechnology Information Write to the Help Desk

GETTING STARTED

- [NCBI Education](#)
- [NCBI Help Manual](#)
- [NCBI Handbook](#)
- [Training & Tutorials](#)
- [Submit Data](#)

RESOURCES

- [Chemicals & Bioassays](#)
- [Data & Software](#)
- [DNA & RNA](#)
- [Domains & Structures](#)
- [Genes & Expression](#)
- [Genetics & Medicine](#)
- [Genomes & Maps](#)
- [Homology](#)
- [Literature](#)
- [Proteins](#)
- [Sequence Analysis](#)
- [Taxonomy](#)
- [Variation](#)

POPULAR

- [PubMed](#)
- [Bookshelf](#)
- [PubMed Central](#)
- [PubMed Health](#)
- [BLAST](#)
- [Nucleotide](#)
- [Genome](#)
- [SNP](#)
- [Gene](#)
- [Protein](#)
- [PubChem](#)

FEATURED

- [Genetic Testing Registry](#)
- [PubMed Health](#)
- [GenBank](#)
- [Reference Sequences](#)
- [Gene Expression Omnibus](#)
- [Map Viewer](#)
- [Human Genome](#)
- [Mouse Genome](#)
- [Influenza Virus](#)
- [Primer-BLAST](#)
- [Sequence Read Archive](#)

NCBI INFORMATION

- [About NCBI](#)
- [Research at NCBI](#)
- [NCBI News](#)
- [NCBI FTP Site](#)
- [NCBI on Facebook](#)
- [NCBI on Twitter](#)
- [NCBI on YouTube](#)

National Center for Biotechnology Information, U.S. National Library of Medicine
8600 Rockville Pike, Bethesda MD, 20894 USA

[Policies and Guidelines](#) | [Contact](#)

<https://www.ncbi.nlm.nih.gov>

general informatics resources

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

UniProtKB
UniProt Knowledgebase
Swiss-Prot (550,740)
Manually annotated and reviewed.
Records with information extracted from literature and curator-evaluated computational analysis.

UniRef
The UniProt Reference Clusters (UniRef) provide clustered sets of sequences from the UniProt Knowledgebase (including isoforms) and selected UniParc records.

UniParc
UniParc is a comprehensive and non-redundant database that contains most of the publicly available protein sequences in the world.

Proteomes
A proteome is the set of proteins thought to be expressed by an organism. UniProt provides proteomes for species with completely sequenced genomes.

Supporting data
Literature citations
Cross-ref. databases
Taxonomy
Diseases
Subcellular locations
Keywords

News
Forthcoming changes
Planned changes for UniProt
UniProt release 2016_03
From the Zika forest to the Amazon, news from a viral wanderer | Cross-references to EPD and TopDownProteomics
UniProt release 2016_02
Another one (antibiotic) bites the dust | Cross-references to SwissPalm and Gramene | Removal of cross-references to GeneFarm and GenoLi...
News archive

*my personal favorite

protein structure: Protein Data Bank (PDB)

RCSB PDB Deposit ▾ Search ▾ Visualize ▾ Analyze ▾ Download ▾ Learn ▾ More ▾ MyPDB Login ▾

RCSB PDB PROTEIN DATA BANK An Information Portal to 117022 Biological Macromolecular Structures

PDB-101 WORLDWIDE wWPDB PROTEIN DATA BANK EMDDataBank Nucleic Acid Database StructuralBiology Knowledgebase

Search by PDB ID, author, macromolecule, sequence, or ligands Go Advanced Search | Browse by Annotations

f t y a d q

Welcome

Deposit

Search

Visualize

Analyze

Download

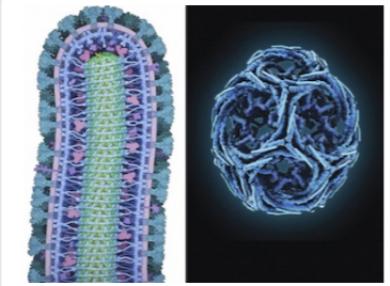
Learn

A Structural View of Biology

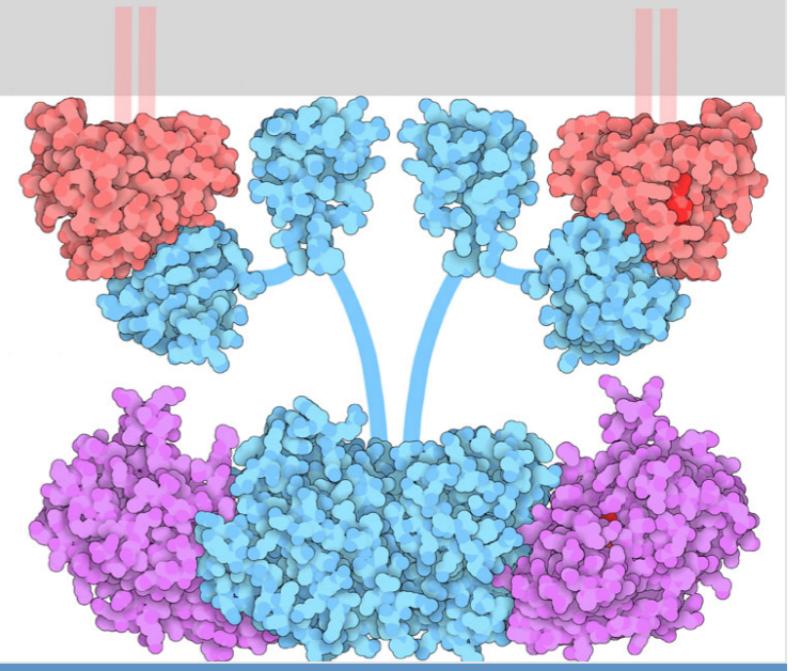
This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data. The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

RCSB PDB illustrations recognized at Wellcome Trust Image Awards



March Molecule of the Month



RAF Protein Kinases

protein structure: assorted databases

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

- UniProtKB**: UniProt Knowledgebase. Swiss-Prot (550,740) Manually annotated and reviewed. TrEMBL (63,039,659) Automatically annotated and not reviewed. Records with information extracted from literature and curator-evaluated computational analysis.
- UniRef**: The UniProt Reference Clusters (UniRef) provide clustered sets of sequences from the UniProt Knowledgebase (including isoforms) and selected UniParc records.
- UniParc**: UniParc is a comprehensive and non-redundant database that contains most of the publicly available protein sequences in the world.
- Proteomes**: A proteome is the set of proteins thought to be expressed by an organism. UniProt provides proteomes for species with completely sequenced genomes.
- Supporting data**: Literature citations, Taxonomy, Diseases, Subcellular locations, Keywords.
- News**: News items, Forthcoming changes, Uniprot release 2016_03, Uniprot release 2016_02.

Welcome to SCOP2!

News

Citation

Description of the SCOP2 database

Search Browser

Search Graph

<http://www.uniprot.org>

<http://scop2.mrc-lmb.cam.ac.uk>

CATH / Gene3D

26 million protein domains classified into 2,738 superfamilies

Browse » Search » Download » Take the Tour »

What is CATH?

CATH is a classification of protein structures downloaded from the Protein Data Bank. We group protein domains into superfamilies when there is sufficient evidence they have diverged from a common ancestor.

- Search CATH by text, ID or keyword
- Search CATH by protein sequence (FASTA)
- Search CATH by PDB structure

- Browse CATH Hierarchy
- CATH Release Statistics
- CATH Tutorials

Example pages

- PDB "2bop"
- Domain "1cukA01"
- Relatives of "1cukA01"
- Superfamily "HUPs"

- Functional Family
- FunFam Alignment
- Search for "enolase"
- Superfamily Comparison

<http://www.cathdb.info>

SMART

SMART MODE: NORMAL GENOMIC

Select your default SMART mode

Different color schemes are used to easily identify the mode you're in.

Normal mode	Genomic mode
SMART MODE: NORMAL GENOMIC	SMART MODE: NORMAL GENOMIC
Simple Modular Architecture Research Tool	Simple Modular Architecture Research Tool

<http://smart.embl-heidelberg.de>

small molecules: assorted databases

Databases > Upload Services > Help more > Today's Statistics >



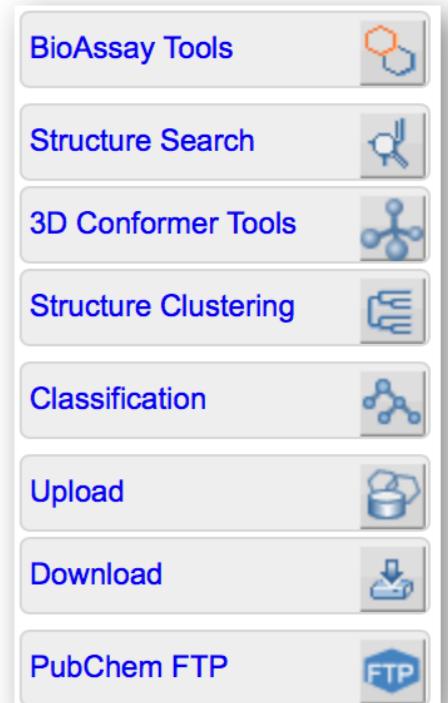
Go Limits Advanced

Try the new [PubChem Search](#)

New PubChem presents at the 251st American Chemical Society National Meeting in San Diego (March 13-17, 2016). Read more at <http://1.usa.gov/1QBp0aE>

New A new article about the PubChem Compound and Substance databases is available.
[Read more...](#)

[more ...](#)



[Write to Helpdesk](#) | [Disclaimer](#) | [Privacy Statement](#) | [Accessibility](#) | [Data Citation Guidelines](#)

National Center for Biotechnology Information

NLM | NIH | HHS

<https://pubchem.ncbi.nlm.nih.gov>

small molecules: assorted databases

Databases > Upload Services > Help more > Today's Statistics >

Compounds: 88,931,139
Substances: 219,157,086
BioAssays: 1,154,426
Tested Compounds: 2,101,163
Tested Substances: 3,320,600
RNAi BioAssays: 89
BioActivities: 229,903,012
Protein Targets: 9,960
Gene Targets: 19,260

BioAssay Compound Substance

Go Limits Advanced

Try the new [PubChem Search](#)

New PubChem presents at the 251st American Chemical Society National Meeting in San Diego (March 13-17, 2016). Read more at <http://1.usa.gov/1QBp0aE>

New A new article about the PubChem Compound and Substance databases is available.
[Read more...](#)

more ... 



BioAssay Tools 

Structure Search 

3D Conformer Tools 

Structure Clustering 

Classification 

Upload 

Download 

PubChem FTP 

small molecules: assorted databases

UCSF University of California, San Francisco | About UCSF | Search UCSF | UCSF Medical Center

Shoichet Laboratory [docking.org](#)

Not Authenticated – sign in

Active cart: Temporary Cart (0 items)

ZINC 12

About Search Subsets Help Social G+1 { 81 } Quick Search Bar... Go

Please consider switching to [ZINC15](#), which is superior to ZINC12 in most ways. If you prefer ZINC12 after trying ZINC15, we would like to know why so that we can get you to make the switch. Read more (coming soon)

Welcome to ZINC, a free database of commercially-available compounds for virtual screening. ZINC contains over 35 million purchasable compounds in ready-to-dock, 3D formats. ZINC is provided by the [Irwin](#) and [Shoichet](#) Laboratories in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF). To cite ZINC, please reference: Irwin, Sterling, Mysinger, Bolstad and Coleman, *J. Chem. Inf. Model.* 2012 DOI: [10.1021/ci3001277](https://doi.org/10.1021/ci3001277). The original publication is Irwin and Shoichet, *J. Chem. Inf. Model.* 2005;45(1):177-82 [PDF](#), [DOI](#). We thank [NIGMS](#) for financial support (GM71896).

ZINC ID, Drug Name, SMILES, Catalog, Vendor Code, Target & more Go

Structure/Draw Physical Properties Catalogs & Vendors ZINC IDS Targets Rings Combination

What's NEW? Feedback Like us
@chem4biology Blog RSS
Video Walkthroughs

ZINC Database
Like Page 940 likes

Quick Links

Download	Search
Target focused	Thanks
Natural Products	Special Subsets
Search By Target	PBCs
Rings	Carts

Your Carts

Create an account or [login](#) to have multiple carts.

10-Special Subsets in ...

Bioinformatics and Chemical Informatics Research Center (BCIRC) [Terms of use](#) [Privacy policy](#) [Questions, Discussion, Bug reports, Feature requests](#) Thank you NIGMS! [GM71896](#) [Go Secure](#)

small molecules: assorted databases

EMBL-EBI

ChEMBL

wellcome trust

EBI > Databases > Small Molecules > ChEMBL Database > Home

Search ChEMBL... Compounds Targets Assays Documents Cells Tissues Exact Match [Activity Source Filter](#)

Ligand Search Target Search Browse Targets Browse Drugs Browse Drug Targets Browse Drug Indications About

ChEMBL Statistics

- DB: ChEMBL_22
- Targets: 11,224
- Compound records: 2,036,512
- Distinct compounds: 1,686,695
- Activities: 14,371,197
- Publications: 65,213
- [Release Notes](#)

ChEMBL Blog

- Position to work on tractability in Open Targets
- [ChEMBL Webinars](#)

Marvin JS

ChemAxon

List Search

SMILES Search ChEMBL ID Search Keyword Search

Please enter a list of Compound IDs, keywords, or SMILES separated by newlines

Fetch Compounds

Biological Blast Search

Run BLAST

Compound Sketcher: Please select....

Substructure Search 100%

Fetch Compounds

Services

By topic
By name (A-Z)
Help & Support

Research

Overview
Publications
Research groups
Postdocs & PhDs

Training

Overview
Train at EBI
Train outside EBI
Train online
Contact organisers

Industry

Overview
Members Area
Workshops
SME Forum
Contact Industry programme

About us

Overview
Leadership
Funding
Background
Collaboration
Jobs
People & groups
News
Events
Visit us
Contact us

EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK +44 (0)1223 49 44 44

Copyright © EMBL-EBI 2017 | EBI is an outstation of the European Molecular Biology Laboratory | Terms of use

<https://www.ebi.ac.uk/chembl/>

cancer genomics



Home ▾ Resources ▾ Curation ▾ Tools ▾ Data ▾ News ▾ Help ▾ About ▾ Search COSMIC... Login ▾

COSMIC v80, released 13-FEB-17

COSMIC, the Catalogue Of Somatic Mutations In Cancer, is the world's largest and most comprehensive resource for exploring the impact of somatic mutations in human cancer.

Start using COSMIC by searching for a gene, cancer type, mutation, etc. below, or by browsing a region of the human genome using the map to the right.

eg: *Braf, COLO-829, Carcinoma, V600E, BRCA-UK, Campbell* **SEARCH**

R Resources

Key COSMIC resources

- Cell Lines Project
- COSMIC
- Cancer Gene Census
- Drug Sensitivity ↗
- Mutational Signatures
- GRCh37 Cancer Archive ↗

T Tools

Additional tools to explore COSMIC

- Cancer Browser
- Genome Browser
- GA4GH Beacon
- COSMIC-3D Beta
- COSMIC In BigQuery ↗ New
- CONAN

C Expert Curation

High quality curation by expert postdoctoral scientists

- Drug Resistance
- Cancer Gene Census
- Curated Genes
- Gene Fusions
- Genome-Wide Screens

D Data

Further details on using COSMIC's content

- Downloads
- License
- Submission
- Genome Annotation
- Datasheets
- Help
- FAQ

Browse the genomic landscape of cancer

COSMIC in BigQuery hosted by ISB-CGC

Institute for Systems Biology

Cancer Genomics Cloud

20th February 2017

A new agreement with the U.S.-based Institute For Systems Biology ↗ (ISB) has seen COSMIC data embedded within their Cancer Genomics Cloud platform (CGC).

This platform is based on Google BigQuery technology and brings unparalleled compute power to researchers around the world. COSMIC users have free access to run queries on the ISB CGC for (at least) the first 6 months. Signing up is easy and Documentation ↗ including full usage details and example queries are available to get users up and running quickly... [More ↗]

● ● ○ ○ ○ ○ ○

genomics

ExAC Browser (Beta) | Exome Aggregation Consortium

Search for a gene or variant or region

Examples - Gene: PCSK9, Transcript: ENST00000407236, Variant: 22-46615880-T-C, Multi-allelic variant: rs1800234, Region: 22:46615715-46615880

About ExAC

The [Exome Aggregation Consortium](#) (ExAC) is a coalition of investigators seeking to aggregate and harmonize exome sequencing data from a wide variety of large-scale sequencing projects, and to make summary data available for the wider scientific community.

The data set provided on this website spans 60,706 unrelated individuals sequenced as part of various disease-specific and population genetic studies. The ExAC Principal Investigators and groups that have contributed data to the current release are listed [here](#).

All data here are released under a [Fort Lauderdale Agreement](#) for the benefit of the wider biomedical community - see the terms of use [here](#).

Sign up for our mailing list for future release announcements [here](#).

Recent News

August 8, 2016

- CNV calls are now available on the ExAC browser

March 14, 2016

- Version 0.3.1 ExAC data and browser (beta) is released! ([Release notes](#))

January 13, 2015

- Version 0.3 ExAC data and browser (beta) is released! ([Release notes](#))

October 29, 2014

- Version 0.2 ExAC data and browser (beta) is released! Sign up for our mailing list for future release announcements [here](#).

October 20, 2014

- Public release of ExAC Browser (beta) at ASHG!

October 15, 2014

- Internal release to consortium now available!

gnomAD browser | genome Aggregation Database

Search for a gene or variant or region

Example - Gene: PCSK9, Variant: 1-55516888-G-GA

About gnomAD

The [Genome Aggregation Database](#) (gnomAD) is a resource developed by an international coalition of investigators, with the goal of aggregating and harmonizing both exome and genome sequencing data from a wide variety of large-scale sequencing projects, and making summary data available for the wider scientific community.

The data set provided on this website spans 123,136 exome sequences and 15,496 whole-genome sequences from unrelated individuals sequenced as part of various disease-specific and population genetic studies. The gnomAD Principal Investigators and groups that have contributed data to the current release are listed [here](#).

All data here are released for the benefit of the wider biomedical community, without restriction on use - see the terms of use [here](#).

Sign up for our mailing list for future release announcements [here](#).

Recent News

February 27, 2017

Official gnomAD release (version 2.0) with browser updates and data available for [download](#).

October 19, 2016

Public release of gnomAD Browser (beta) at ASHG!

<http://exac.broadinstitute.org>

<http://gnomad.broadinstitute.org>

 NHLBI Exome Sequencing Project (ESP)
Exome Variant Server

[Home](#) [Data Browser](#) [Data Usage and Release](#) [How to Use](#) [What's New](#) [Contact and FAQ](#) [Downloads](#)

The goal of the [NHLBI GO Exome Sequencing Project \(ESP\)](#) is to discover novel genes and mechanisms contributing to heart, lung and blood disorders by pioneering the application of next-generation sequencing of the protein coding regions of the human genome across diverse, richly-phenotyped populations and to share these datasets and findings with the scientific community to extend and enrich the diagnosis, management and treatment of heart, lung and blood disorders.

The groups participating and collaborating in the NHLBI GO ESP include:

- Seattle GO - University of Washington, Seattle, WA
- Broad GO - Broad Institute of MIT and Harvard, Cambridge, MA
- WHISp GO - Ohio State University Medical Center, Columbus, OH
- Lung GO - University of Washington, Seattle, WA
- WashU GO - Washington University, St. Louis, MO
- Heart GO - University of Virginia Health System, Charlottesville, VA
- ChargeS GO - University of Texas Health Sciences Center at Houston

The group includes some of the largest well-phenotyped populations in the United States, representing more than 200,000 individuals altogether from the:

- Women's Health Initiative (WHI)
- Framingham Heart Study (FHS)
- Jackson Heart Study (JHS)
- Multi-Ethnic Study of Atherosclerosis (MESA)
- Atherosclerosis Risk in Communities (ARIC)
- Coronary Artery Risk Development in Young Adults (CARDIA)
- Cardiovascular Health Study (CHS)
- Lung Health Study (LHS)
- COPD Genetic Epidemiology (COPDGene)
- Severe Asthma Research Project (SARP)
- Pulmonary Arterial Hypertension population (PAH)
- Acute Lung Injury cohort (ALI)
- Cystic Fibrosis cohort (CF)
- PennCATH
- Cleveland Clinic Genbank
- Massachusetts General Hospital Premature Coronary Artery Disease Study (MGH PCAD)
- Heart Attack Risk in Puget Sound (HARPS)
- Translational Research Underlying Disparities in Myocardial Infarction Patients' Health Status (TRIUMPH)

<http://evs.gs.washington.edu/EVS/>