

## Base Types

integer, float, boolean, string, bytes

```
int 783 0 -192 0b010 0o642 0xFF3
      zero binary octal hexa
float 9.23 0.0 -1.7e-6
bool True False
str "One\nTwo"
  escaped new line
  'I\'m'
  escaped '
bytes b"toto\xfe\775"
      hexadecimal octal
```

Multiline string:  
"""X\tY\tZ  
1\t2\t3"""  
escaped tab

⚡ immutables

## Container Types

- ordered sequences, fast index access, repeatable values
  - list** [1,5,9] ["x",11,8.9] ["mot"]
  - tuple** (1,5,9) 11,"y",7.4 ("mot",)
- Non modifiable values (immutables) ⚡ expression with only comas → **tuple** (ordered sequences of chars / bytes)
- key containers, no a priori order, fast key access, each key is unique
  - dict** {"key": "value"} dict(a=3,b=4,k="v")
  - (key/value associations) {1: "one", 3: "three", 2: "two", 3.14: "pi"}
  - set** {"key1", "key2"} {1,9,3,0} **set** {}
  - ⚡ keys=hashable values (base types, immutables...) **frozenset** immutable set empty

## Identifiers

for variables, functions, modules, classes... names

**a...zA...Z\_** followed by **a...zA...Z\_0...9**

- diacritics allowed but should be avoided
- language keywords forbidden
- lower/UPPER case discrimination

Ⓢ a toto x7 y\_max BigOne  
Ⓢ 8y and for

## Variables assignment

⚡ assignment ⇔ **binding** of a name with a value

- evaluation of right side expression value
- assignment in order with left side names

```
x=1.2+8+sin(y)
a=b=c=0 assignment to same value
y,z,r=9.2,-7.6,0 multiple assignments
a,b=b,a values swap
a,*b=seq unpacking of sequence in
*a,b=seq item and list
x+=3 increment ⇔ x=x+3
x-=2 decrement ⇔ x=x-2
x=None « undefined » constant value
del x remove name x
```

## Conversions

**type** (expression)

can specify integer number base in 2<sup>nd</sup> parameter

```
int("15") → 15
int("3f",16) → 63
int(15.56) → 15 truncate decimal part
float("-11.24e8") → -1124000000.0
round(15.56,1) → 15.6 rounding to 1 decimal (0 decimal → integer number)
bool(x) False for null x, empty container x, None or False x; True for other x
str(x) → "..." representation string of x for display (cf. formatting on the back)
chr(64) → '@' ord('@') → 64 code → char
repr(x) → "..." literal representation string of x
bytes([72,9,64]) → b'H\t@'
list("abc") → ['a','b','c']
dict([(3,"three"),(1,"one")]) → {1:'one',3:'three'}
set(["one","two"]) → {'one','two'}
```

separator **str** and sequence of **str** → assembled **str**

```
':'.join(['toto','12','pswd']) → 'toto:12:pswd'
```

**str** splitted on whitespaces → **list** of **str**

```
"words with spaces".split() → ['words','with','spaces']
```

**str** splitted on separator **str** → **list** of **str**

```
"1,4,8,2".split(",") → ['1','4','8','2']
```

sequence of one type → **list** of another type (via list comprehension)

```
[int(x) for x in ('1','29','-3')] → [1,29,-3]
```

## Sequence Containers Indexing

for lists, tuples, strings, bytes...

negative index	-5	-4	-3	-2	-1
positive index	0	1	2	3	4

```
lst=[10,20,30,40,50]
```

positive slice	0	1	2	3	4	5
negative slice	-5	-4	-3	-2	-1	

**Items count**  
**len(lst) → 5**  
⚡ index from 0 (here from 0 to 4)

Individual access to **items** via **lst[index]**

```
lst[0] → 10 ⇒ first one
lst[1] → 20
lst[-1] → 50 ⇒ last one
lst[-2] → 40
```

On mutable sequences (**list**), remove with **del lst[3]** and modify with assignment **lst[4]=25**

Access to **sub-sequences** via **lst[start slice: end slice: step]**

```
lst[: -1] → [10,20,30,40]
lst[1: -1] → [20,30,40]
lst[: 2] → [10,30,50]
lst[: -1] → [50,40,30,20,10]
lst[: -2] → [50,30,10]
lst[: ] → [10,20,30,40,50] shallow copy of sequence
```

Missing slice indication → from start / up to end.

On mutable sequences (**list**), remove with **del lst[3:5]** and modify with assignment **lst[1:4]=[15,25]**

## Boolean Logic

Comparisons : < > <= >= == != (boolean results)

**a and b** logical and both simultaneously

**a or b** logical or one or other or both

⚡ pitfall : **and** and **or** return **value** of **a** or of **b** (under shortcut evaluation).  
⇒ ensure that **a** and **b** are booleans.

**not a** logical not

**True**  
**False** } True and False constants

## Statements Blocks

```
parent statement:
┌ statement block 1...
│ ...
└ parent statement:
  ┌ statement block 2...
  │ ...
  └ next statement after block 1
```

⚡ configure editor to insert 4 spaces in place of an indentation tab.

## Modules/NAMES Imports

module **truc** ⇔ file **truc.py**

```
from monmod import nom1,nom2 as fct
  → direct access to names, renaming with as
import monmod
  → access via monmod.nom1 ...
```

⚡ modules and packages searched in python path (cf **sys.path**)

## Conditional Statement

statement block executed only if a condition is true

**if logical condition:**  
→ statements block

Can go with several **elif**, **elif...** and only one final **else**. Only the block of first true condition is executed.

```
if age <= 18:
    state="Kid"
elif age > 65:
    state="Retired"
else:
    state="Active"
```

⚡ with a var **x**:  
**if bool(x)==True:** ⇔ **if x:**  
**if bool(x)==False:** ⇔ **if not x:**

## Maths

floating numbers... approximated values

Operators: + - \* / // % \*\*

Priority (...)

integer ÷ ÷ remainder

@ → matrix × python3.5+numpy

```
(1+5.3)*2 → 12.6
abs(-3.2) → 3.2
round(3.57,1) → 3.6
pow(4,3) → 64.0
```

⚡ usual order of operations

angles in radians

```
from math import sin,pi...
sin(pi/4) → 0.707...
cos(2*pi/3) → -0.4999...
sqrt(81) → 9.0
log(e**2) → 2.0
ceil(12.5) → 13
floor(12.5) → 12
```

modules **math**, **statistics**, **random**, **decimal**, **fractions**, **numpy**, etc. (cf. doc)

## Exceptions on Errors

Signaling an error:  
**raise ExcClass(...)**

Errors processing:  
**try:**  
→ normal processing block  
**except Exception as e:**  
→ error processing block

⚡ **finally** block for final processing in all cases.

```
normal
raise X()
error processing
error
raise
```



# Python For Data Science Cheat Sheet

## NumPy Basics

Learn Python for Data Science Interactively at [www.DataCamp.com](https://www.datacamp.com)



### NumPy

The NumPy library is the core library for scientific computing in Python. It provides a high-performance multidimensional array object, and tools for working with these arrays.

Use the following import convention:



```
>>> import numpy as np
```

### NumPy Arrays

#### 1D array

```
1 2 3
```

#### 2D array

axis 1  
axis 0

```
1.5 2 3  
4 5 6
```

#### 3D array

axis 2  
axis 1  
axis 0

### Creating Arrays

```
>>> a = np.array([1,2,3])  
>>> b = np.array([(1.5,2,3), (4,5,6)], dtype = float)  
>>> c = np.array([(1.5,2,3), (4,5,6)], [(3,2,1), (4,5,6)]],  
                 dtype = float)
```

### Initial Placeholders

```
>>> np.zeros((3,4))  
>>> np.ones((2,3,4),dtype=np.int16)  
>>> d = np.arange(10,25,5)  
  
>>> np.linspace(0,2,9)  
  
>>> e = np.full((2,2),7)  
>>> f = np.eye(2)  
>>> np.random.random((2,2))  
>>> np.empty((3,2))
```

Create an array of zeros  
Create an array of ones  
Create an array of evenly spaced values (step value)  
Create an array of evenly spaced values (number of samples)  
Create a constant array  
Create a 2X2 identity matrix  
Create an array with random values  
Create an empty array

### I/O

#### Saving & Loading On Disk

```
>>> np.save('my_array', a)  
>>> np.savez('array.npz', a, b)  
>>> np.load('my_array.npy')
```

#### Saving & Loading Text Files

```
>>> np.loadtxt("myfile.txt")  
>>> np.genfromtxt("my_file.csv", delimiter=',')  
>>> np.savetxt("myarray.txt", a, delimiter=" ")
```

### Data Types

```
>>> np.int64  
>>> np.float32  
>>> np.complex  
>>> np.bool  
>>> np.object  
>>> np.string_  
>>> np.unicode_
```

Signed 64-bit integer types  
Standard double-precision floating point  
Complex numbers represented by 128 floats  
Boolean type storing TRUE and FALSE values  
Python object type  
Fixed-length string type  
Fixed-length unicode type

### Inspecting Your Array

```
>>> a.shape  
>>> len(a)  
>>> b.ndim  
>>> e.size  
>>> b.dtype  
>>> b.dtype.name  
>>> b.astype(int)
```

Array dimensions  
Length of array  
Number of array dimensions  
Number of array elements  
Data type of array elements  
Name of data type  
Convert an array to a different type

### Asking For Help

```
>>> np.info(np.ndarray.dtype)
```

### Array Mathematics

#### Arithmetic Operations

```
>>> g = a - b  
array([[ -0.5,  0. ,  0. ],  
       [ -3. , -3. , -3. ]])  
>>> np.subtract(a,b)  
>>> b + a  
array([[ 2.5,  4. ,  6. ],  
       [ 5. ,  7. ,  9. ]])  
>>> np.add(b,a)  
>>> a / b  
array([[ 0.66666667,  1. ,  1. ],  
       [ 0.25 ,  0.4 ,  0.5 ]])  
>>> np.divide(a,b)  
>>> a * b  
array([[ 1.5,  4. ,  9. ],  
       [ 4. , 10. , 18. ]])  
>>> np.multiply(a,b)  
>>> np.exp(b)  
>>> np.sqrt(b)  
>>> np.sin(a)  
>>> np.cos(b)  
>>> np.log(a)  
>>> e.dot(f)  
array([[ 7. ,  7. ],  
       [ 7. ,  7.]])
```

Subtraction  
Subtraction  
Addition  
Addition  
Division  
Division  
Multiplication  
Multiplication  
Exponentiation  
Square root  
Print sines of an array  
Element-wise cosine  
Element-wise natural logarithm  
Dot product

#### Comparison

```
>>> a == b  
array([[False,  True,  True],  
       [False, False, False]], dtype=bool)  
>>> a < 2  
array([[True, False, False], dtype=bool)  
>>> np.array_equal(a, b)
```

Element-wise comparison  
Element-wise comparison  
Array-wise comparison

#### Aggregate Functions

```
>>> a.sum()  
>>> a.min()  
>>> b.max(axis=0)  
>>> b.cumsum(axis=1)  
>>> a.mean()  
>>> b.median()  
>>> a.corrcoef()  
>>> np.std(b)
```

Array-wise sum  
Array-wise minimum value  
Maximum value of an array row  
Cumulative sum of the elements  
Mean  
Median  
Correlation coefficient  
Standard deviation

### Copying Arrays

```
>>> h = a.view()  
>>> np.copy(a)  
>>> h = a.copy()
```

Create a view of the array with the same data  
Create a copy of the array  
Create a deep copy of the array

### Sorting Arrays

```
>>> a.sort()  
>>> c.sort(axis=0)
```

Sort an array  
Sort the elements of an array's axis

### Subsetting, Slicing, Indexing

Also see Lists

#### Subsetting

```
>>> a[2]  
3  
>>> b[1,2]  
6.0
```

Select the element at the 2nd index  
Select the element at row 1 column 2 (equivalent to b[1][2])

#### Slicing

```
>>> a[0:2]  
array([1, 2])  
>>> b[0:2,1]  
array([ 2.,  5.])
```

Select items at index 0 and 1  
Select items at rows 0 and 1 in column 1

```
>>> b[:1]  
array([[1.5, 2., 3.]])  
>>> c[1,...]  
array([[ 3.,  2.,  1.],  
       [ 4.,  5.,  6.]])
```

Select all items at row 0 (equivalent to b[0:1, :])  
Same as [1, :, :]

```
>>> a[: :-1]  
array([3, 2, 1])
```

Reversed array a

#### Boolean Indexing

```
>>> a[a<2]  
array([1])
```

Select elements from a less than 2

#### Fancy Indexing

```
>>> b[[1, 0, 1, 0], [0, 1, 2, 0]]  
array([ 4. ,  2. ,  6. , 1.5])  
>>> b[[1, 0, 1, 0]][:, [0,1,2,0]]  
array([[ 4. ,  5. ,  6. ,  4. ],  
       [ 1.5,  2. ,  3. , 1.5],  
       [ 4. ,  5. ,  6. ,  4. ],  
       [ 1.5,  2. ,  3. , 1.5]])
```

Select elements (1,0), (0,1), (1,2) and (0,0)  
Select a subset of the matrix's rows and columns

### Array Manipulation

#### Transposing Array

```
>>> i = np.transpose(b)  
>>> i.T
```

Permute array dimensions  
Permute array dimensions

#### Changing Array Shape

```
>>> b.ravel()  
>>> g.reshape(3,-2)
```

Flatten the array  
Reshape, but don't change data

#### Adding/Removing Elements

```
>>> h.resize((2,6))  
>>> np.append(h,g)  
>>> np.insert(a, 1, 5)  
>>> np.delete(a, [1])
```

Return a new array with shape (2,6)  
Append items to an array  
Insert items in an array  
Delete items from an array

#### Combining Arrays

```
>>> np.concatenate((a,d),axis=0)  
array([ 1,  2,  3, 10, 15, 20])  
>>> np.vstack((a,b))  
array([[ 1. ,  2. ,  3. ],  
       [ 1.5,  2. ,  3. ],  
       [ 4. ,  5. ,  6. ]])  
>>> np.r_[e,f]  
>>> np.hstack((e,f))  
array([[ 7.,  7.,  1.,  0.],  
       [ 7.,  7.,  0.,  1.]])  
>>> np.column_stack((a,d))  
array([[ 1, 10],  
       [ 2, 15],  
       [ 3, 20]])  
>>> np.c_[a,d]
```

Concatenate arrays  
Stack arrays vertically (row-wise)  
Stack arrays vertically (row-wise)  
Stack arrays horizontally (column-wise)  
Create stacked column-wise arrays  
Create stacked column-wise arrays

#### Splitting Arrays

```
>>> np.hsplit(a,3)  
[array([1]),array([2]),array([3])]  
>>> np.vsplit(c,2)  
[array([[ 1.5,  2. ,  1. ],  
       [ 4. ,  5. ,  6. ]]),  
 array([[ 3.,  2.,  3.],  
       [ 4. ,  5. ,  6.]])]
```

Split the array horizontally at the 3rd index  
Split the array vertically at the 2nd index

DataCamp

Learn Python for Data Science Interactively





# Python For Data Science Cheat Sheet

## SciPy - Linear Algebra

Learn More Python for Data Science [Interactively](https://www.datacamp.com) at [www.datacamp.com](https://www.datacamp.com)



### SciPy

The **SciPy** library is one of the core packages for scientific computing that provides mathematical algorithms and convenience functions built on the NumPy extension of Python.



### Interacting With NumPy

[Also see NumPy](#)

```
>>> import numpy as np
>>> a = np.array([1,2,3])
>>> b = np.array([(1+5j,2j,3j), (4j,5j,6j)])
>>> c = np.array([(1.5,2,3), (4,5,6)], [(3,2,1), (4,5,6)])
```

#### Index Tricks

<pre>&gt;&gt;&gt; np.mgrid[0:5,0:5] &gt;&gt;&gt; np.ogrid[0:2,0:2] &gt;&gt;&gt; np.r_[3,[0]*5,-1:1:10j] &gt;&gt;&gt; np.c_[b,c]</pre>	Create a dense meshgrid Create an open meshgrid Stack arrays vertically (row-wise) Create stacked column-wise arrays
---	---

#### Shape Manipulation

<pre>&gt;&gt;&gt; np.transpose(b) &gt;&gt;&gt; b.flatten() &gt;&gt;&gt; np.hstack((b,c)) &gt;&gt;&gt; np.vstack((a,b)) &gt;&gt;&gt; np.hsplit(c,2) &gt;&gt;&gt; np.vsplit(d,2)</pre>	Permute array dimensions Flatten the array Stack arrays horizontally (column-wise) Stack arrays vertically (row-wise) Split the array horizontally at the 2nd index Split the array vertically at the 2nd index
--	--

#### Polynomials

<pre>&gt;&gt;&gt; from numpy import polyld &gt;&gt;&gt; p = polyld([3,4,5])</pre>	Create a polynomial object
---	----------------------------

#### Vectorizing Functions

<pre>&gt;&gt;&gt; def myfunc(a):     if a &lt; 0:         return a*2     else:         return a/2 &gt;&gt;&gt; np.vectorize(myfunc)</pre>	Vectorize functions
---	---------------------

#### Type Handling

<pre>&gt;&gt;&gt; np.real(b) &gt;&gt;&gt; np.imag(b) &gt;&gt;&gt; np.real_if_close(c,tol=1000) &gt;&gt;&gt; np.cast['f'](np.pi)</pre>	Return the real part of the array elements Return the imaginary part of the array elements Return a real array if complex parts close to 0 Cast object to a data type
---	--

#### Other Useful Functions

<pre>&gt;&gt;&gt; np.angle(b,deg=True) &gt;&gt;&gt; g = np.linspace(0,np.pi,num=5) &gt;&gt;&gt; g[3:] += np.pi &gt;&gt;&gt; np.unwrap(g) &gt;&gt;&gt; np.logspace(0,10,3) &gt;&gt;&gt; np.select([c&lt;4],[c*2])  &gt;&gt;&gt; misc.factorial(a) &gt;&gt;&gt; misc.comb(10,3,exact=True) &gt;&gt;&gt; misc.central_diff_weights(3) &gt;&gt;&gt; misc.derivative(myfunc,1.0)</pre>	Return the angle of the complex argument Create an array of evenly spaced values (number of samples) Unwrap Create an array of evenly spaced values (log scale) Return values from a list of arrays depending on conditions Factorial Combine N things taken at k time Weights for Np-point central derivative Find the n-th derivative of a function at a point
---	--

## Linear Algebra

You'll use the `linalg` and `sparse` modules. Note that `scipy.linalg` contains and expands on `numpy.linalg`.

```
>>> from scipy import linalg, sparse
```

### Creating Matrices

```
>>> A = np.matrix(np.random.random((2,2)))
>>> B = np.asmatrix(b)
>>> C = np.mat(np.random.random((10,5)))
>>> D = np.mat([[3,4], [5,6]])
```

### Basic Matrix Routines

#### Inverse

```
>>> A.I
>>> linalg.inv(A)
```

Inverse  
Inverse

#### Transposition

```
>>> A.T
>>> A.H
```

Tranpose matrix  
Conjugate transposition

#### Trace

```
>>> np.trace(A)
```

Trace

#### Norm

```
>>> linalg.norm(A)
>>> linalg.norm(A,1)
>>> linalg.norm(A,np.inf)
```

Frobenius norm  
L1 norm (max column sum)  
L inf norm (max row sum)

#### Rank

```
>>> np.linalg.matrix_rank(C)
```

Matrix rank

#### Determinant

```
>>> linalg.det(A)
```

Determinant

#### Solving linear problems

```
>>> linalg.solve(A,b)
>>> E = np.mat(a).T
>>> linalg.lstsq(F,E)
```

Solver for dense matrices  
Solver for dense matrices  
Least-squares solution to linear matrix equation

#### Generalized inverse

```
>>> linalg.pinv(C)
```

Compute the pseudo-inverse of a matrix (least-squares solver)  
Compute the pseudo-inverse of a matrix (SVD)

```
>>> linalg.pinv2(C)
```

### Creating Sparse Matrices

<pre>&gt;&gt;&gt; F = np.eye(3, k=1) &gt;&gt;&gt; G = np.mat(np.identity(2)) &gt;&gt;&gt; C[C &gt; 0.5] = 0 &gt;&gt;&gt; H = sparse.csr_matrix(C) &gt;&gt;&gt; I = sparse.csc_matrix(D) &gt;&gt;&gt; J = sparse.dok_matrix(A) &gt;&gt;&gt; E.todense() &gt;&gt;&gt; sparse.isspmatrix_csc(A)</pre>	Create a 2X2 identity matrix Create a 2x2 identity matrix  Compressed Sparse Row matrix Compressed Sparse Column matrix Dictionary Of Keys matrix Sparse matrix to full matrix Identify sparse matrix
--	--

### Sparse Matrix Routines

#### Inverse

```
>>> sparse.linalg.inv(I)
```

Inverse

#### Norm

```
>>> sparse.linalg.norm(I)
```

Norm

#### Solving linear problems

```
>>> sparse.linalg.spsolve(H,I)
```

Solver for sparse matrices

### Sparse Matrix Functions

<pre>&gt;&gt;&gt; sparse.linalg.expm(I)</pre>	Sparse matrix exponential
---	---------------------------

### Asking For Help

```
>>> help(scipy.linalg.diagsvd)
>>> np.info(np.matrix)
```

[Also see NumPy](#)

### Matrix Functions

#### Addition

```
>>> np.add(A,D)
```

Addition

#### Subtraction

```
>>> np.subtract(A,D)
```

Subtraction

#### Division

```
>>> np.divide(A,D)
```

Division

#### Multiplication

```
>>> A @ D
```

Multiplication operator (Python 3)

```
>>> np.multiply(D,A)
```

Multiplication

```
>>> np.dot(A,D)
```

Dot product

```
>>> np.vdot(A,D)
```

Vector dot product

```
>>> np.inner(A,D)
```

Inner product

```
>>> np.outer(A,D)
```

Outer product

```
>>> np.tensordot(A,D)
```

Tensor dot product

```
>>> np.kron(A,D)
```

Kronecker product

#### Exponential Functions

```
>>> linalg.expm(A)
```

Matrix exponential

```
>>> linalg.expm2(A)
```

Matrix exponential (Taylor Series)

```
>>> linalg.expm3(D)
```

Matrix exponential (eigenvalue decomposition)

#### Logarithm Function

```
>>> linalg.logm(A)
```

Matrix logarithm

#### Trigonometric Functions

```
>>> linalg.sinm(D)
```

Matrix sine

```
>>> linalg.cosm(D)
```

Matrix cosine

```
>>> linalg.tanm(A)
```

Matrix tangent

#### Hyperbolic Trigonometric Functions

```
>>> linalg.sinhm(D)
```

Hypperbolic matrix sine

```
>>> linalg.coshm(D)
```

Hyperbolic matrix cosine

```
>>> linalg.tanhm(A)
```

Hyperbolic matrix tangent

#### Matrix Sign Function

```
>>> np.signm(A)
```

Matrix sign function

#### Matrix Square Root

```
>>> linalg.sqrtm(A)
```

Matrix square root

#### Arbitrary Functions

```
>>> linalg.funm(A, lambda x: x*x)
```

Evaluate matrix function

### Decompositions

#### Eigenvalues and Eigenvectors

```
>>> la, v = linalg.eig(A)
```

Solve ordinary or generalized eigenvalue problem for square matrix  
Unpack eigenvalues  
First eigenvector  
Second eigenvector  
Unpack eigenvalues

```
>>> l1, l2 = la
```

```
>>> v[:,0]
```

```
>>> v[:,1]
```

```
>>> linalg.eigvals(A)
```

#### Singular Value Decomposition

```
>>> U,s,Vh = linalg.svd(B)
```

Singular Value Decomposition (SVD)

```
>>> M,N = B.shape
```

```
>>> Sig = linalg.diagsvd(s,M,N)
```

Construct sigma matrix in SVD

#### LU Decomposition

```
>>> P,L,U = linalg.lu(C)
```

LU Decomposition

### Sparse Matrix Decompositions

```
>>> la, v = sparse.linalg.eigs(F,1)
```

Eigenvalues and eigenvectors

```
>>> sparse.linalg.svds(H, 2)
```

SVD

DataCamp

Learn Python for Data Science [Interactively](#)

