

Cheat Sheet - Day 2

Introduction to Data Analysis with Python

<https://github.com/janekfleper/Workshop-Konstanz-2024>

Basics

<code>ord("a")</code>	A <i>function</i> is called with
<code>len([0, 1, 2])</code>	(optional) arguments
<code>y.count("a")</code>	A <i>method</i> is called on a value
<code>"abc".upper()</code>	or on a variable
<code># a comment</code>	A <code>#</code> starts a <i>comment</i> that will not be evaluated
<code>type(x)</code>	Get the <i>type</i> of a value or of a variable
<code>def square(x):</code> <code> return x**2</code>	Define a function with a <i>parameter</i> <code>x</code> and a <i>return</i> value
<code>sorted(x)</code>	Get a <i>sorted</i> copy of an iterable (list, string, etc...)

Data types

<code>"hello", 'abc', "0.9", str(123)</code>	A <i>string</i> is a sequence of characters in quotation marks
<code>12, -4, int("5")</code>	An <i>integer</i> is a number without a decimal part
<code>0.9, -3.1415, float("-0.1")</code>	A <i>float</i> is a number with a decimal part
<code>True, False, bool(0), x < 1</code>	A <i>boolean</i> can only take the values <code>True</code> or <code>False</code>
<code>[0, "abc", 0.1]</code> <code>list("hello")</code>	A <i>list</i> is a mutable, sorted collection of values
<code>{"a": 1, "b": 2}</code> <code>dict(a=1, b=2)</code>	A <i>dictionary</i> is a mutable collection of key-value pairs
<code>(0, "0.9", True)</code> <code>tuple([0, 1, 2])</code>	A <i>tuple</i> is an immutable, sorted collection of values

Files

<code>with open("a.txt") as f:</code> <code> # do something</code>	Open a file for reading
<code>for line in f:</code> <code> print(line)</code>	Iterate over all lines in a file
<code>f.read()</code>	Read a file as a string
<code>f.readline()</code>	Read a single line as a string
<code>f.readlines()</code>	Read a file as a list of strings

Strings

<code>s1 = "hello"</code>	Initialize a regular string
<code>s2 = f"x = {x}"</code>	Initialize an f-string
<code>s1.strip("ho")</code>	Strip characters from a string
<code>s2.split("=")</code>	Split a string at a substring
<code>bool(s1)</code>	Only an empty string is <code>False</code>

Arrays

<code>import numpy as np</code>	
<code>n = [1, 1, 2, 3]</code> <code>a = np.array(n)</code>	Initialize an array
<code>np.arange(1, 9)</code>	Get a sequence of integers
<code>a[0], a[0:-3]</code>	Get values from an array
<code>a[0] = 4</code> <code>a[0:-3] = 0</code>	Set values in an array
<code>3 * a, a + 3.14,</code> <code>a / 4, a - a**2</code>	Compute element-wise arithmetic operations
<code>a[a == 1]</code>	Filter data in an array
<code>a.sum(), a.prod()</code> <code>a.mean(), a.std()</code> <code>a.min(), a.max()</code>	Run computations on all values in an array
<code>np.bincount(a)</code>	Get the histogram distribution of an array

<code>a.shape</code>	Get the shape of an array
<code>a.dtype</code>	Get the data type of an array
<code>a.astype(t)</code>	Get a copy of an array with a specific data type <code>t</code>
<code>np.sort(a)</code>	Get a sorted copy of an array
<code>np.argsort(a)</code>	Get the indices that would sort an array

Plots

<code>import matplotlib.pyplot as plt</code>	
<code>plt.plot(y)</code> <code>plt.plot(x, y)</code>	Create a simple line plot
<code>plt.bar(x, y)</code>	Create a (vertical) bar plot
<code>plt.xlabel("x")</code> <code>plt.ylabel("y")</code>	Add axis labels to a plot
<code>plt.title("plot")</code>	Add a title to a plot
<code>plt.legend()</code>	Add a legend to a plot

Jupyter shortcuts

<code>Enter</code> / <code>Esc</code>	Start/exit the edit mode
<code>Shift</code> + <code>Enter</code>	Run cell(s) and select next
<code>A</code> / <code>B</code>	Insert new cell above/below
<code>↑</code> , <code>K</code> / <code>↓</code> , <code>J</code>	Select cell above/below
<code>X</code> , <code>C</code> , <code>V</code>	Cut, copy or paste cell(s)
<code>D</code> + <code>D</code>	Delete cell(s)
<code>Z</code> / <code>Shift</code> + <code>Z</code>	Undo/redo cell operation
<code>Shift</code> + <code>Tab</code>	Open the documentation