



Módulo 1: Software Engineering and Data Science

Python, Data Science and Software Engineering

Modularity

Programar se vuelve menos complejo cuando el código es dividido en unidades funcionales pequeñas. Esta es la idea detrás de la modularidad.

- Mejora la legibilidad.
- Mejora la mantenibilidad.
- Reutilización.

Documentation

- Muestra cómo usar el proyecto
- Previene la confusión de los colaboradores.
- Previene la frustración de nuestro futuro yo.

Automated Testing

- Ahorra tiempo frente al testing manual.
- Encuentra y soluciones más bugs.
- Ejecutar tests en cualquier momento y en cualquier lugar.

Introduction to Packages and Documentation

PyPi

Nos da una plataforma sencilla de donde tomar paquetes públicos.

pip

Gracias a que los paquetes son modulares, podemos instalarlos desde PyPi usando una herramienta llamada pip.

Reading documentation with help()

```
help(numpy.busday_count)
```

```
busday_count(begindates, enddates)
    Counts the number of valid days between `begindates` and
    `enddates`, not including the day of `enddates`.

    Parameters
    -----
    begindates : the first dates for counting.
    enddates : the end dates for counting (excluded from the count)

    Returns
    -----
    out : the number of valid days between the begin and end dates.

    Examples
    -----
    >>> # Number of weekdays in 2011
    ... np.busday_count('2011', '2012')
```

Reading documentation with help()

```
import numpy as np
help(np)
```

Provides

1. An array object of arbitrary homogeneous items
2. Fast mathematical operations over arrays
3. Linear Algebra, Fourier Transforms, Random Number Generation

```
help(42)
```

```
class int(object)
| int(x=0) -> integer
| int(x, base=10) -> integer
|
| Convert a number or string to an integer, or return 0 if no arguments
| are given. If x is a number, return x.__int__(). For floating point
| numbers, this truncates towards zero.
```

Conventions and PEP 8

Python Enhancement Protocol 8 - PEP 8

De facto Style Guide para el código de Python. Cómo formatear nuestro código para que sea lo más legible posible.

"Code is read much more often than it is written"

Violating PEP 8

```
#define our data
my_dict ={
    'a' : 10,
    'b': 3,
    'c' : 4,
    'd': 7}

#import needed package
import numpy as np
#helper function
def DictToArray(d):
    """Convert dictionary values to numpy array"""
    #extract values and convert
    x=np.array(d.values())
    return x
print(DictToArray(my_dict))
```

```
array([10, 4, 3, 7])
```

Following PEP 8

```
# Import needed package
import numpy as np

# Define our data
my_dict = {'a': 10, 'b': 3, 'c': 4, 'd': 7}

# Helper function
def dict_to_array(d):
    """Convert dictionary values to numpy array"""
    # Extract values and convert
    x = np.array(d.values())
    return x

print(dict_to_array(my_dict))
```

```
array([10, 4, 3, 7])
```

PEP 8 Tools

pycodestyle

Using pycodestyle

```
datacamp@server:~$ pip install pycodestyle
datacamp@server:~$ pycodestyle dict_to_array.py
```

```
dict_to_array.py:5:9: E203 whitespace before ':'
dict_to_array.py:6:14: E131 continuation line unaligned for hanging indent
dict_to_array.py:8:1: E265 block comment should start with '# '
dict_to_array.py:9:1: E402 module level import not at top of file
dict_to_array.py:11:1: E302 expected 2 blank lines, found 0
dict_to_array.py:13:15: E111 indentation is not a multiple of four
```

Output from pycodestyle

The diagram shows a single line of text representing a pycodestyle error output: `dict_to_array.py:9:1: E402 module level import not at top of file`. This line is enclosed in a red rectangular border. Four labels with arrows point to specific parts of the line: 'File' points to `dict_to_array.py`, 'Line number' points to `9`, 'Column number' points to `1`, and 'Error code' points to `E402`. The text `module level import not at top of file` is the error description.

File	Line number	Column number	Error code	Error description
<code>dict_to_array.py</code>	<code>9</code>	<code>1</code>	<code>E402</code>	<code>module level import not at top of file</code>