



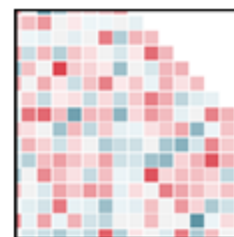
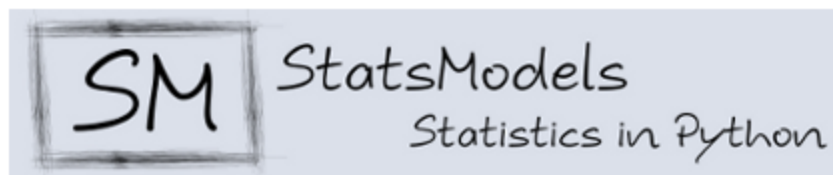
Python aplicado a la ciencia de datos

ANGELA DEVIA

ING.ANGELA.DEVIA@GMAIL.COM

¿Por qué Python?

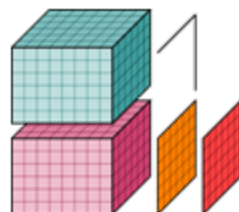
- ▶ Es un lenguaje de programación muy versátil
- ▶ Es intuitivo, diseñado para fácil lectura y escritura de código
- ▶ Es software libre -: Comunidades de código abierto



Seaborn

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



xarray



scikit
learn



scikit-image
image processing in python



NumPy



matplotlib



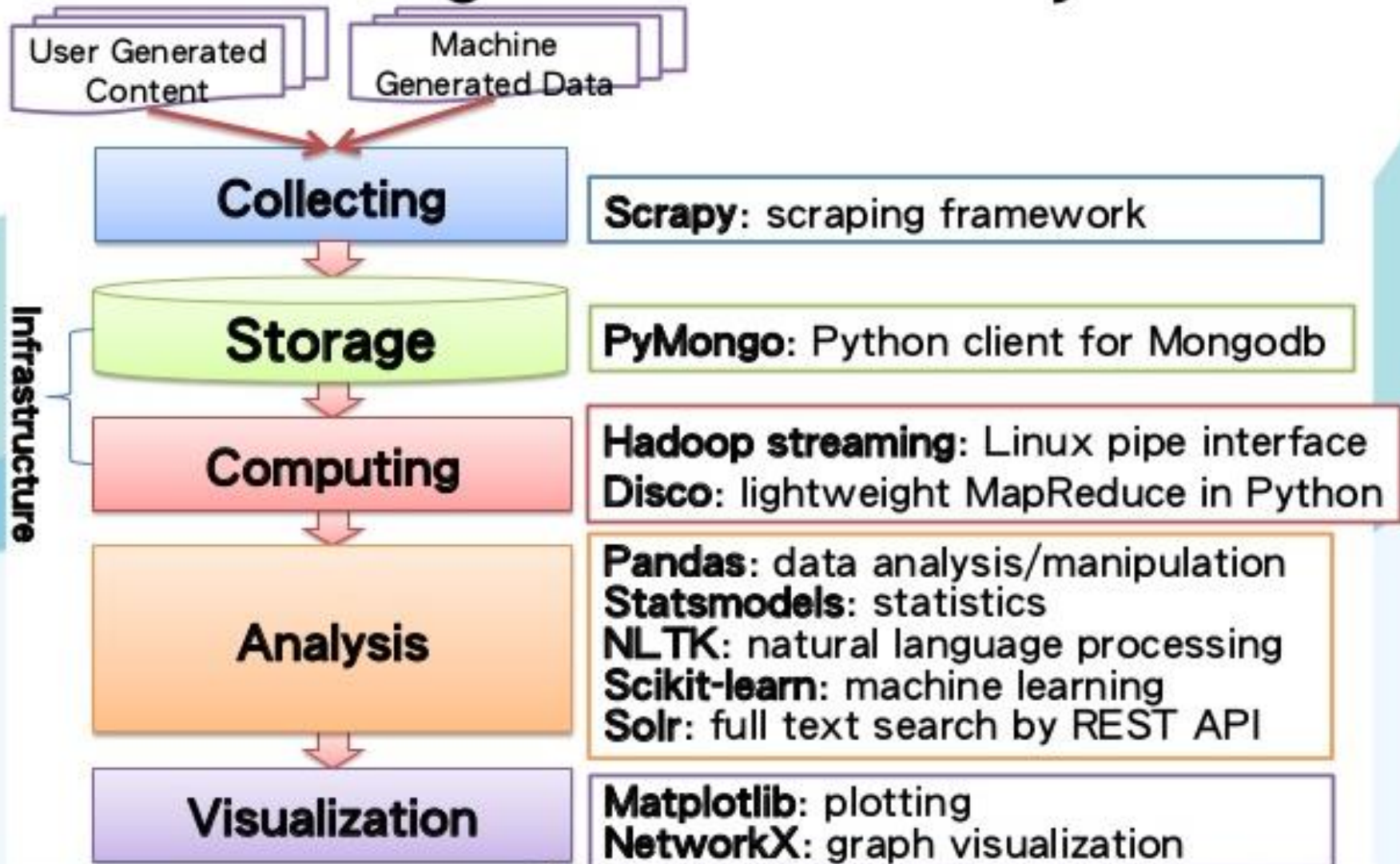
jupyter

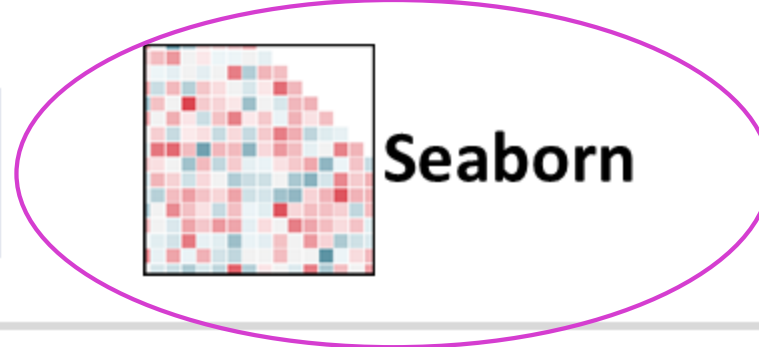
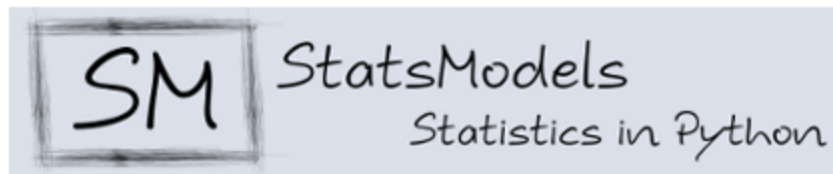


python™

IP[y]:
IPython

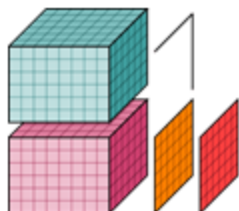
When Big Data meet Python





pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



xarray



scikit-learn



scikit-image
image processing in python



NumPy



matplotlib



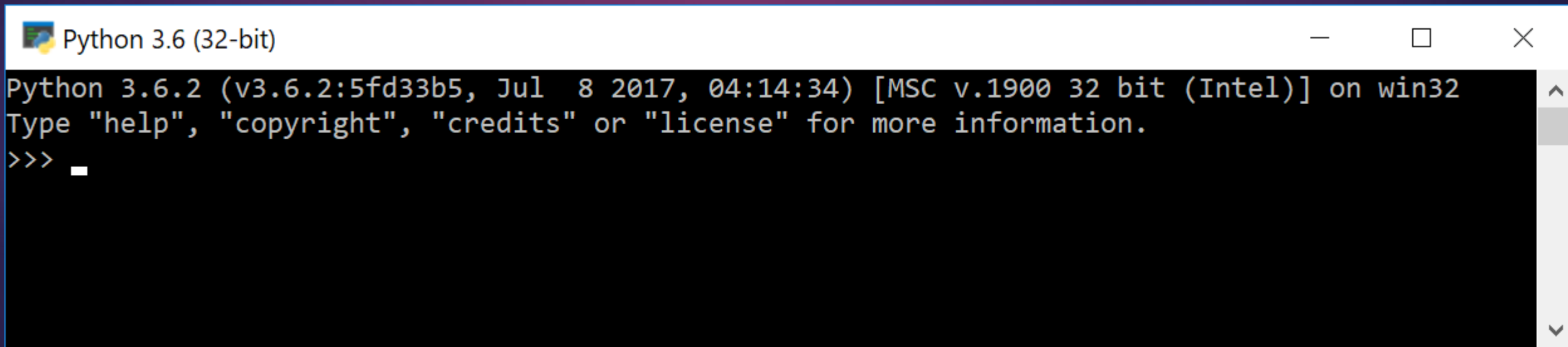
python™

IP[y]:
IPython

¿Cómo empezar a utilizar Python?

Usando consola de comandos

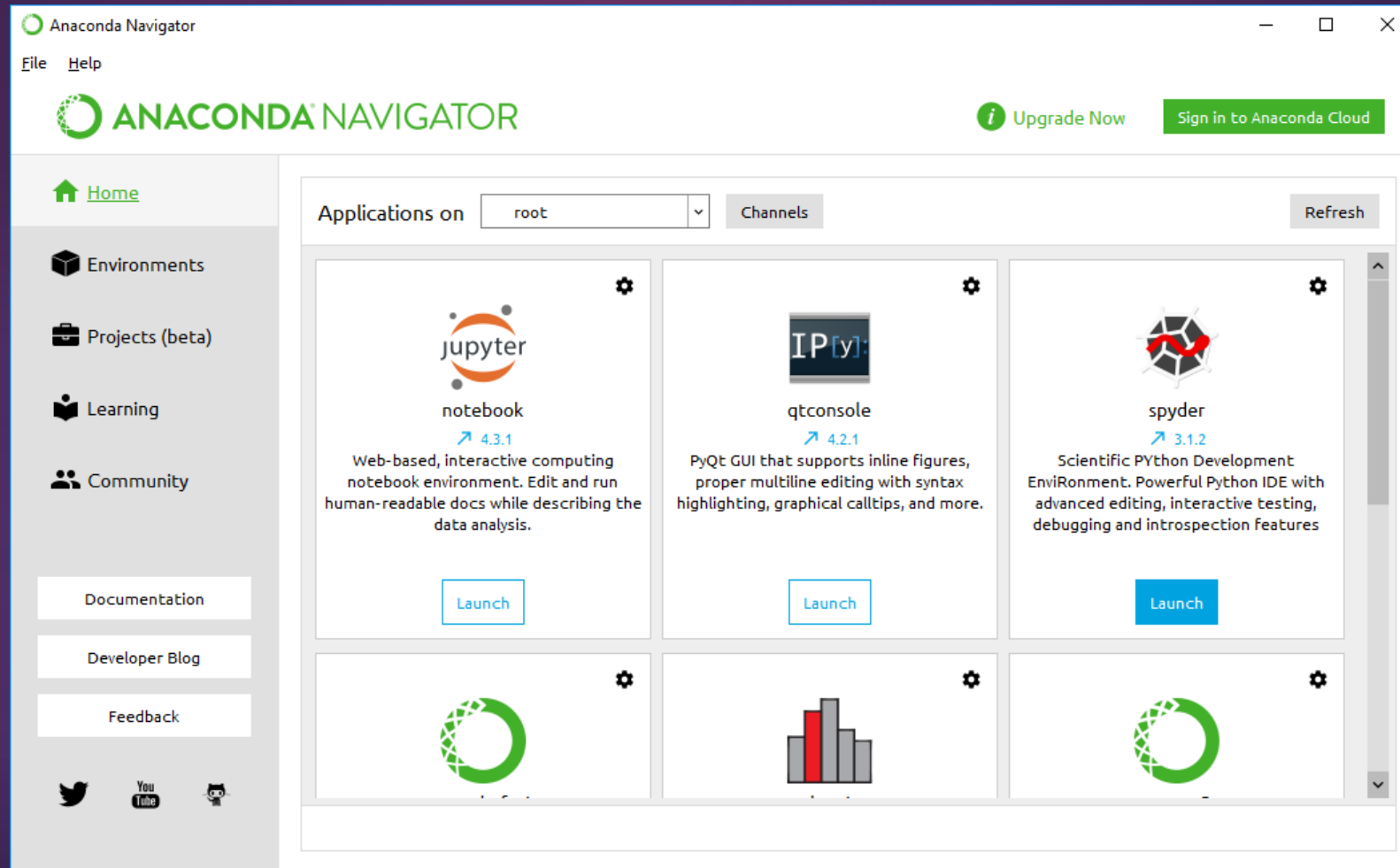
► www.python.org/download



```
Python 3.6 (32-bit)
Python 3.6.2 (v3.6.2:5fd33b5, Jul 8 2017, 04:14:34) [MSC v.1900 32 bit (Intel)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> _
```

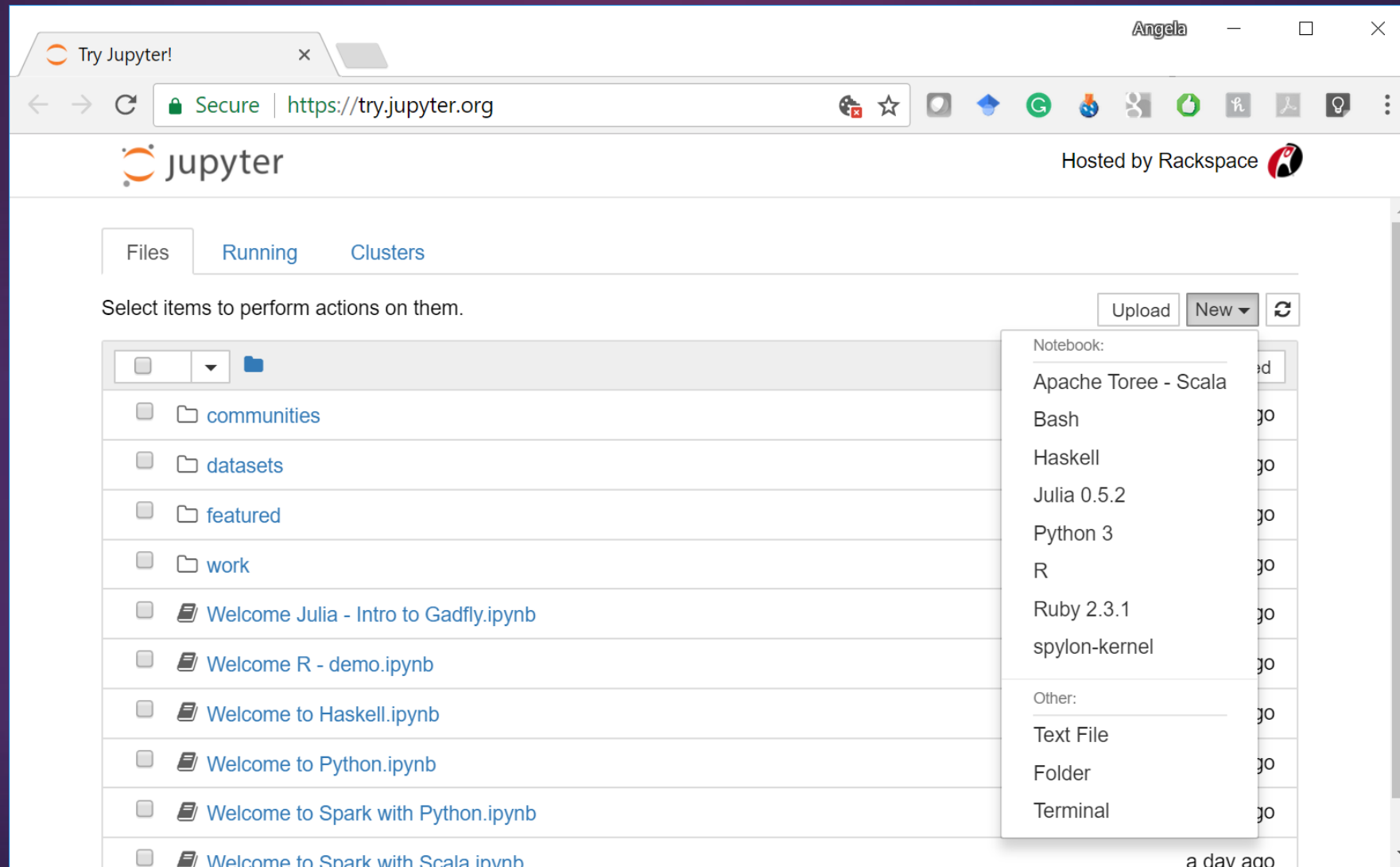

Usando interfaz gráfica

► <https://www.anaconda.com/download/>



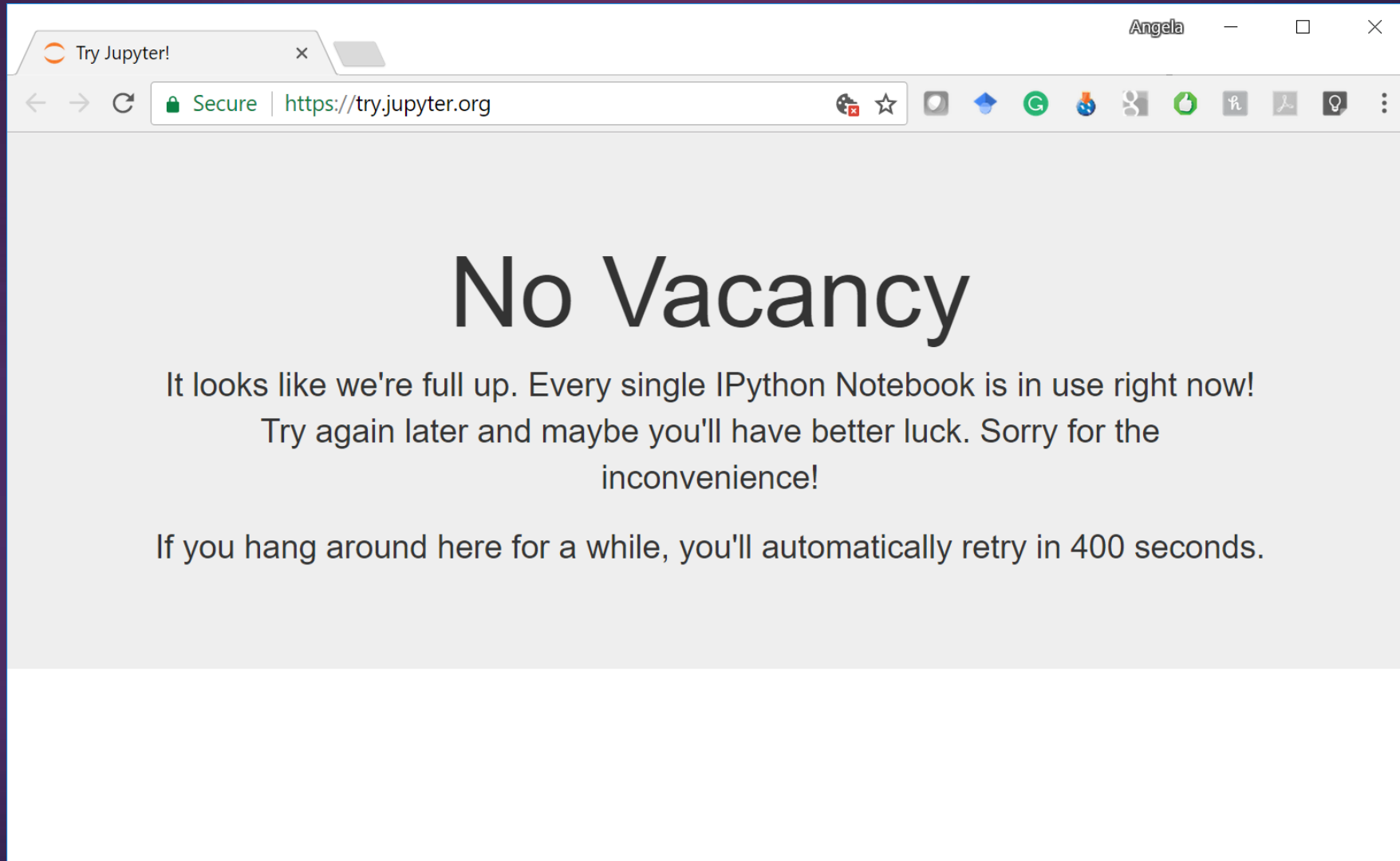
Usando una aplicación web

► <http://jupyter.org/>

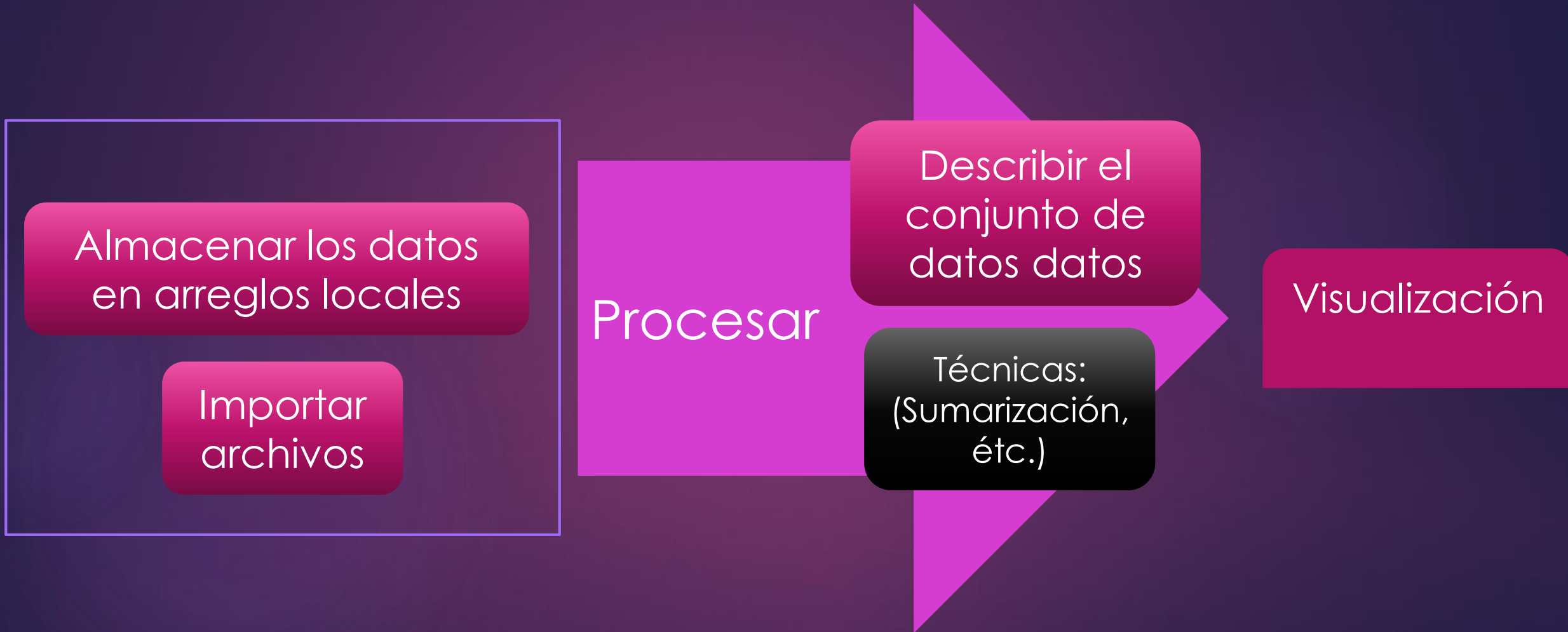


Usando una aplicación web

► <http://jupyter.org/>



¿Cómo se hace?



4.1. Creating NumPy data arrays

A small introductory example:

```
>>> import numpy as np
>>> a = np.array([0, 1, 2])
>>> a
array([0, 1, 2])
>>> print a
[0 1 2]
>>> b = np.array([[0., 1.], [2., 3.]])
>>> b
array([[ 0.,  1.],
       [ 2.,  3.]])
```

In practice, we rarely enter items one by one...

- Evenly spaced values:

```
>>> import numpy as np
>>> a = np.arange(10) # de 0 a n-1
>>> a
array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9])
>>> b = np.arange(1., 9., 2) # syntax : start, end, step
>>> b
array([ 1.,  3.,  5.,  7.])
```

```
squares = []

for i in range(10):
    squares.append(i**2)
```

squares

```
[0, 1, 4, 9, 16, 25, 36, 49, 64, 81]
```

```
squares = [i**2 for i in range(10)]
```

squares

```
[0, 1, 4, 9, 16, 25, 36, 49, 64, 81]
```

```
squares3 = [i**2 for i in range(30) if i % 3 == 0]
squares3
```

```
[0, 9, 36, 81, 144, 225, 324, 441, 576, 729]
```

```
squares3_dict = {i: i**2 for i in range(30) if i % 3 == 0}
squ
```

squares
squares3
squares3_dict

 Code Cell Toolbar: None

```
totals_both = totals_both[unisex]
totals_both.sort(ascending=False)
totals_both.head(5)
```

```
name
Willie      593888
Jordan      479434
Taylor      416096
Leslie       376587
Jamie       350262
dtype: float64
```

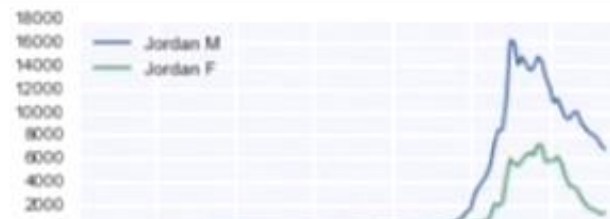
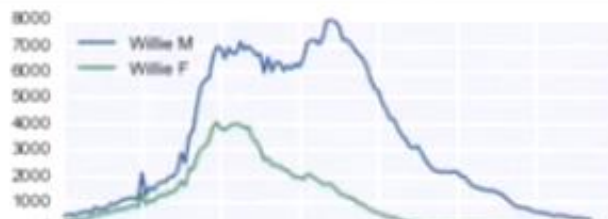
```
names = totals_both.head(10).index.values

pp.figure(figsize=(12,12))

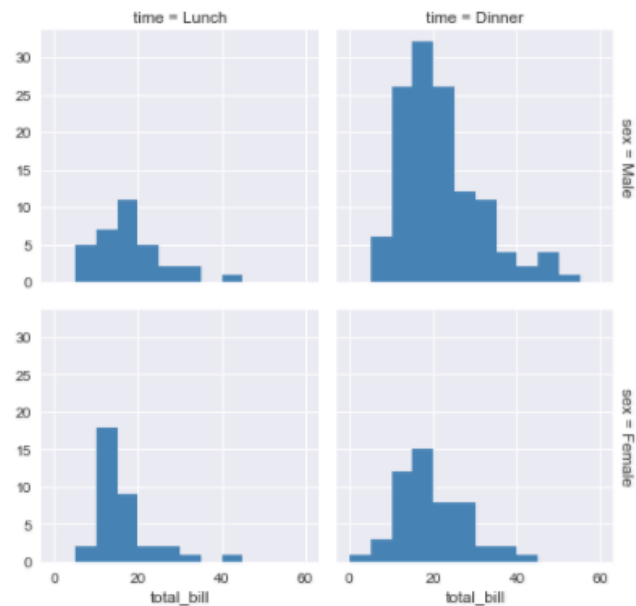
for i,name in enumerate(names):
    pp.subplot(5,2,i+1)

    plotname('M',name)
    plotname('F',name)

    pp.legend([name + ' M',name + ' F'],loc='upper left')
```



Facetting histograms by subsets of data



Python source code: [\[download source: faceted_histogram.py\]](#)

```
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
sns.set(style="darkgrid")

tips = sns.load_dataset("tips")
g = sns.FacetGrid(tips, row="sex", col="time", margin_titles=True)
bins = np.linspace(0, 60, 13)
g.map(plt.hist, "total_bill", color="steelblue", bins=bins, lw=0)
```


¿Cómo usar los paquetes de Python?

► Guías paso a paso:

<http://www.scipy-lectures.org/numpy/numpy.html>

<http://seaborn.pydata.org/examples/index.html#example-gallery>
Guías

¿Cómo importar datos a Python?


Para importar datos desde una URL:

```
import urllib.request
urllib.request.urlretrieve('URL', 'archivo.txt')
open('archivo.txt', 'r').readlines()[:#]
```

¿Cómo importar datos a Python?

Para importar datos desde una URL:

```
import pandas as pd
xls = pd.ExcelFile("yourfilename.xls")
sheetX = xls.parse(2)
                        #2 is the sheet number
var1 = sheetX['ColumnName']
print(var1[1])
                        #1 is the row number...
```



“ Explora sin miedo y sigue
aprendiendo ;)

”

Ángela Devia -
ing.angela.devias@gmail.com

Gracias