

# Jupyter

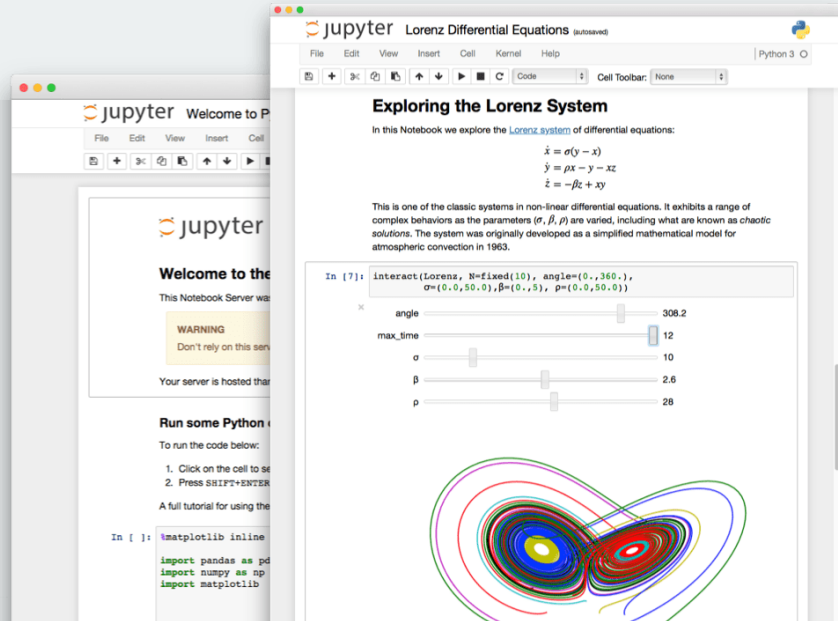
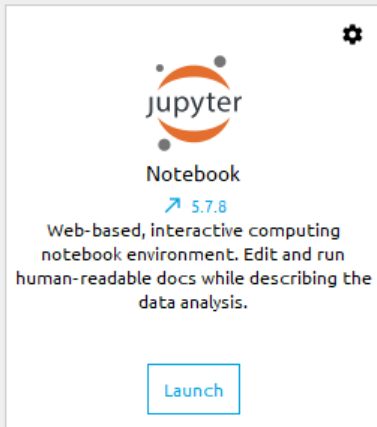
Foundations of AI Academy



Creative Commons Attribution 4.0 International

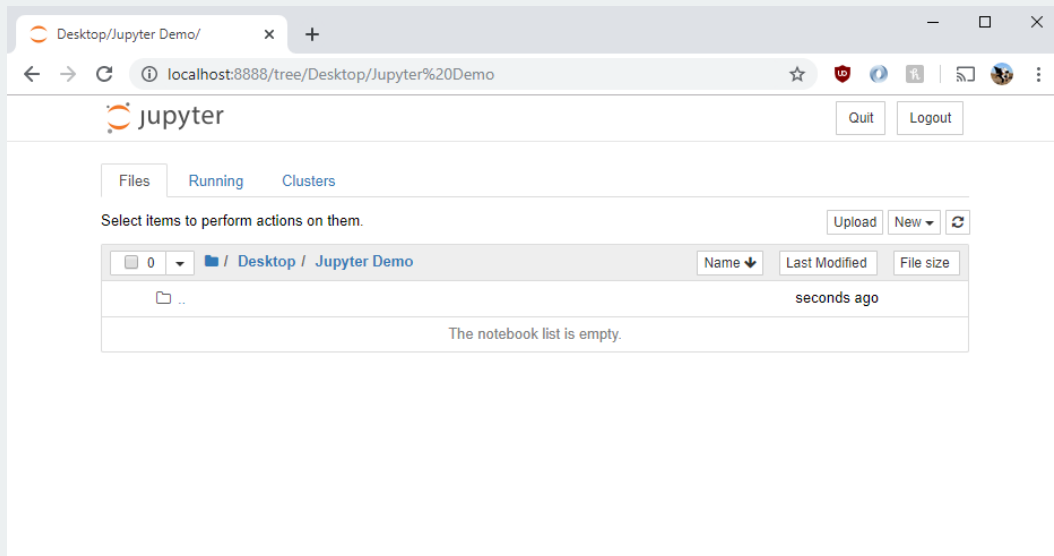
# Jupyter

A programming environment that allows you to run small code snippets at a time



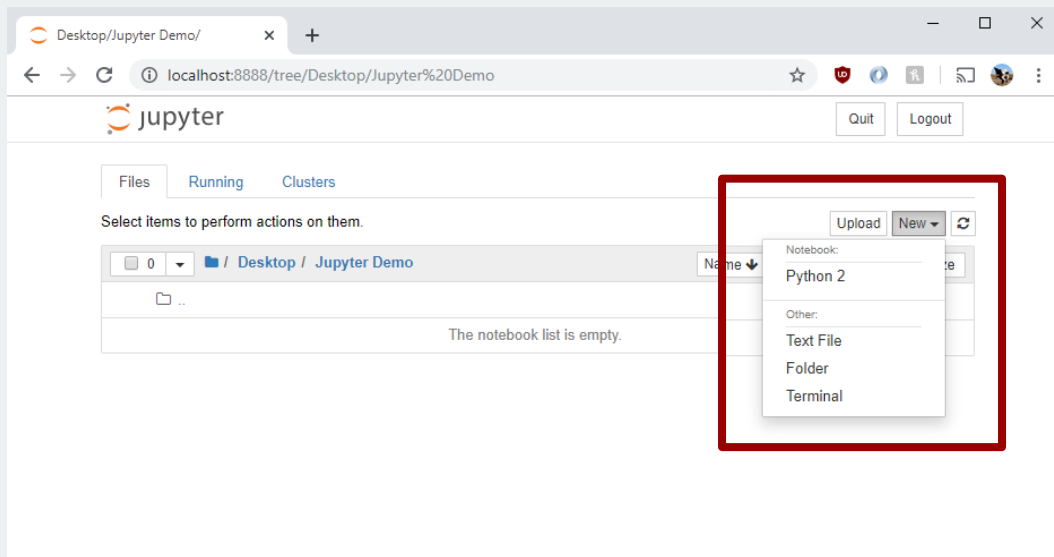
# Jupyter

Upon loading Jupyter, your browser opens a webpage that allow for **notebooks**



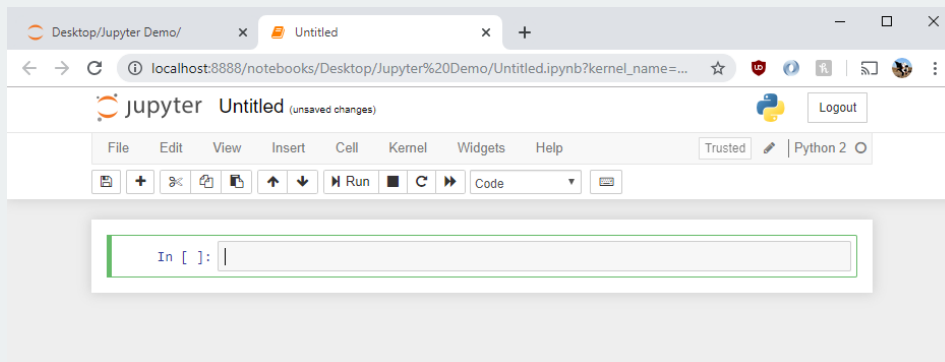
# Jupyter

The **New** option allows you to create a new notebook



# Jupyter Notebooks

You can think of a Jupyter Notebook like the Spyder Console, allowing you to enter in small snippets of code that are run sequentially



```
In [ ]: 1 # IMPORT LIBRARIES
        2 import math as m
        3 import numpy as np
        4 import pandas as pd
        5 import matplotlib.pyplot as plt
```

```
In [ ]: 1
```

```
In [1]: 1 # IMPORT LIBRARIES
        2 import math as m
        3 import numpy as np
        4 import pandas as pd
        5 import matplotlib.pyplot as plt
```

```
In [ ]: 1
```

Now, only this code block is loaded into memory

```
In [1]: 1 # IMPORT LIBRARIES
        2 import math as m
        3 import numpy as np
        4 import pandas as pd
        5 import matplotlib.pyplot as plt
```

```
In [2]: 1 iris = pd.read_csv("../data/iris.csv")
        2
        3 x_axis = "sepal_length"
        4 y_axis = "petal_width"
```

```
In [ ]: 1 |
```

I can load datasets similar to using Spyder, but now can view information line-by-line



```
In [1]: 1 # IMPORT LIBRARIES
        2 import math as m
        3 import numpy as np
        4 import pandas as pd
        5 import matplotlib.pyplot as plt
```

```
In [2]: 1 iris = pd.read_csv("../data/iris.csv")
        2
        3 x_axis = "sepal_length"
        4 y_axis = "petal_width"
```

```
In [3]: 1 iris.head()
```

Out[3]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

Outputs like print() and .head() will be displayed inline with code

```
In [3]: 1 iris.head()
```

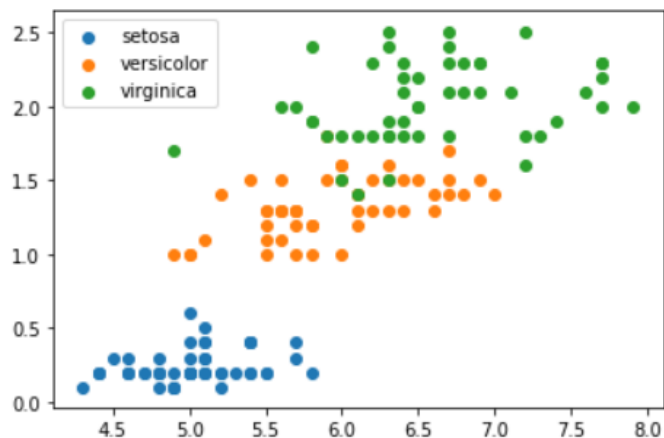
```
Out[3]:
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

Likewise, Matplotlib plots appear in line with the rest of the code.

```
In [7]: 1 species = iris.groupby('species')  
2 for name, data in species:  
3     plt.scatter(data[x_axis], data[y_axis], label=name)  
4 plt.legend()
```

```
Out[7]: <matplotlib.legend.Legend at 0x20fc6277948>
```



```
In [3]: 1 iris.head()
```

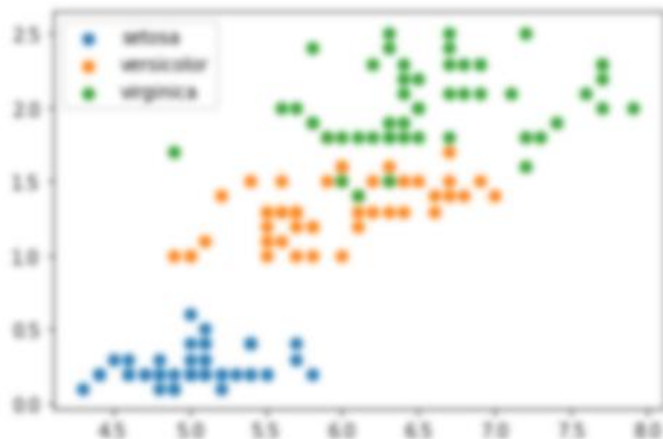
```
Out[3]:
```

```
In [2]: 1 iris = pd.read_csv("../data/iris.csv")
        2
        3 x_axis = "sepal_width"
        4 y_axis = "petal_length"
```

If I want to manipulate the code, I can jump back to a previous "cell", alter it, and rerun the cell's code

```
In [7]: 1 species = iris['species']
        2 for name, data in iris.groupby(species):
        3     plt.scatter(x_axis, y_axis, data=data)
        4 plt.legend()
```

```
Out[7]: <matplotlib.legend.Legend at 0x20fc6277948>
```



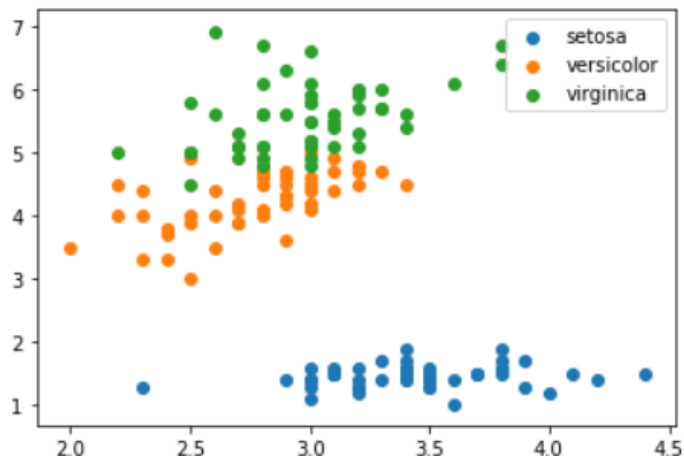
```
In [3]: 1 iris.head()
```

Out[3]:

```
In [2]: 1 iris = pd.read_csv("../data/iris.csv")
2
3 x_axis = "sepal_width"
4 y_axis = "petal_length"
```

```
In [10]: 1 species = iris.groupby('species')
2 for name, data in species:
3     plt.scatter(data[x_axis], data[y_axis], label=name)
4 plt.legend()
```

Out[10]: <matplotlib.legend.Legend at 0x20fc62bbb88>



# Pandas Data Manipulation

Sometimes the data you need doesn't exist, you so you want to calculate it

```
>>> df = pd.DataFrame({'temp_c': [17.0, 25.0]},  
                        index=['Portland', 'Berkeley'])
```

```
>>> df
```

	temp_c
Portland	17.0
Berkeley	25.0

# Pandas Data Manipulation

You can use the **assign** function to assign new values based on other values

```
>>> df.assign(temp_f=df['temp_c'] * 9 / 5 + 32)
```

	temp_c	temp_f
Portland	17.0	62.6
Berkeley	25.0	77.0

# Pandas Data Manipulation

You can use the **assign** function to assign new values based on other values

```
>>> df.assign(temp_f=df['temp_c'] * 9 / 5 + 32)
```

	temp_c	temp_f
Portland	17.0	62
Berkeley	25.0	77

DataFrames will process the new column's data on a row by row basis for each row's temp\_c