

DATA MINING

Python

Python

- In the last few years there is an increasing community that creates **Data Mining tools in Python**
 - There are also tools in other languages but we will use Python whenever we can for a common point of reference.
- We will use **Iron Python** that interfaces with .NET and we can run Notebooks in a browser.
 - You can also use any editor and compile and run from a terminal

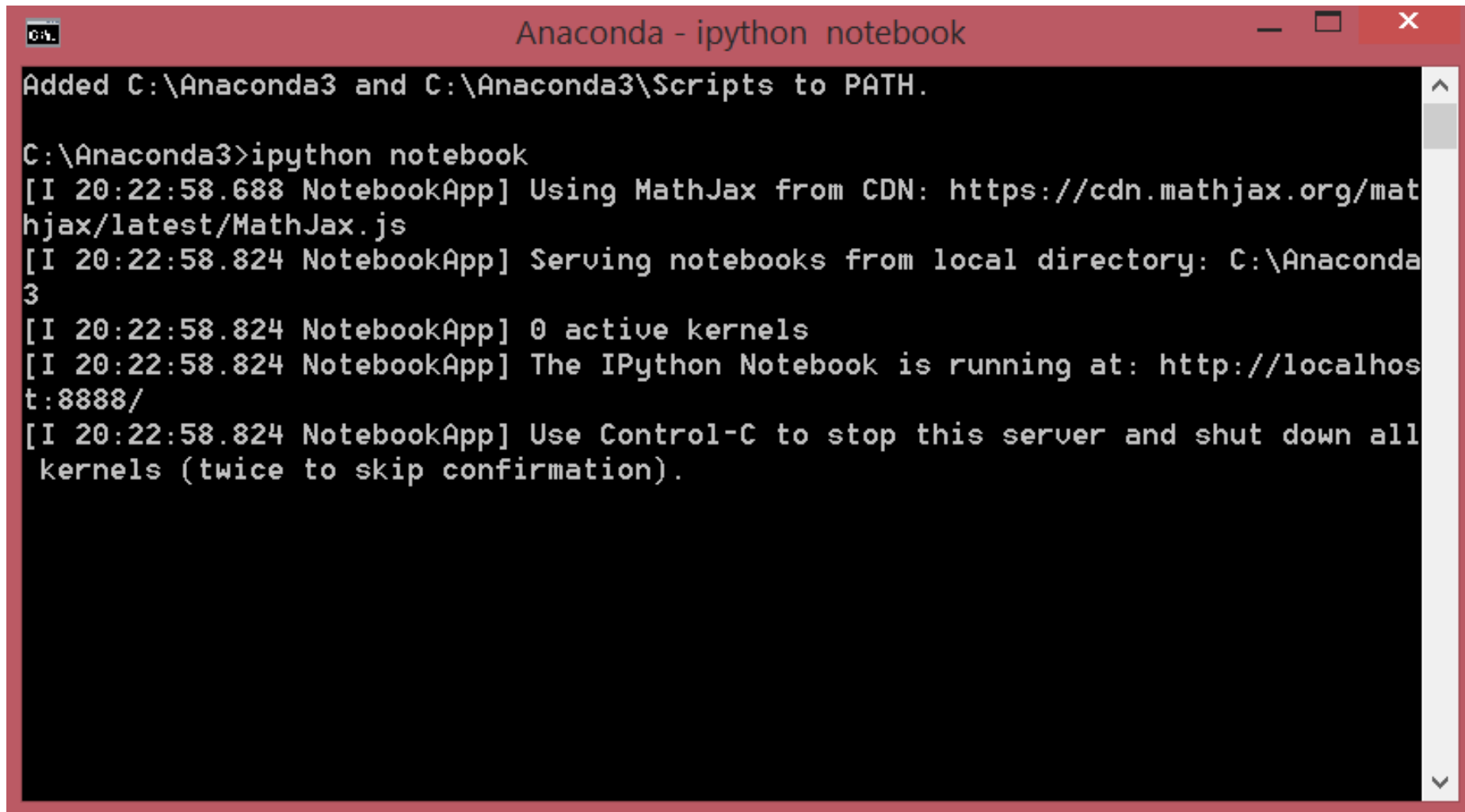
Installing Python

- Installing libraries in Python is complex, so you should download the **Anaconda Scientific Python** distribution which will install most of the libraries that we will use.
 - There are two versions, Python 2.7 and Python 3.0 and they are not compatible. We will use Python 3.0

Resources

- There are tons of resources online for Python.
- For an introduction you can also look at the slides of the [Introduction to Programming](#) course by prof. N. Mamoulis

Starting iPython notebook



```
Added C:\Anaconda3 and C:\Anaconda3\Scripts to PATH.  
C:\Anaconda3>ipython notebook  
[I 20:22:58.688 NotebookApp] Using MathJax from CDN: https://cdn.mathjax.org/mathjax/latest/MathJax.js  
[I 20:22:58.824 NotebookApp] Serving notebooks from local directory: C:\Anaconda3  
[I 20:22:58.824 NotebookApp] 0 active kernels  
[I 20:22:58.824 NotebookApp] The IPython Notebook is running at: http://localhost:8888/  
[I 20:22:58.824 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
```

jupyter

Files Running Clusters

Select items to perform actions on them.

Upload New ↕

- Home
- DLLs
- Doc
- Examples
- Lib
- Library
- Menu
- Scripts
- Tools
- conda-meta
- envs
- include
- info
- libs
- licenses
- node-webkit
- pkgs
- share
- tcl