**Preface**

**What Is Data Science?**

Data science comprises three distinct and overlapping areas: the skills of a statistician who knows how to model and summarize datasets (which are growing ever larger); the skills of a computer scientist who can design and use algorithms to efficiently store, process, and visualize this data; and the domain expertise—what we might think of as “classical” training in a subject—necessary both to formulate the right questions and to put their answers in context.

A diagram of data science

Description automatically generated

**Why Python**

* NumPy for manipulation of homogeneous array-based data,
* Pandas for manipulation of heterogeneous and labeled data,
* SciPy for common scientific computing tasks,
* Matplotlib for publication-quality visualizations,
* IPython for interactive execution and sharing of code, Scikit-Learn for machine learning

**Jupyter: Beyond Normal Python**

**Getting Started in IPython and Jupyter**

* **IPython shell** for trying out short sequences of commands
* **Jupyter Notebook** for longer interactive analysis and for sharing content with others
* **Interactive development environments (IDEs)** like **Emacs** or **VSCode** for creating reusable Python packages.

**The IPython Shell**

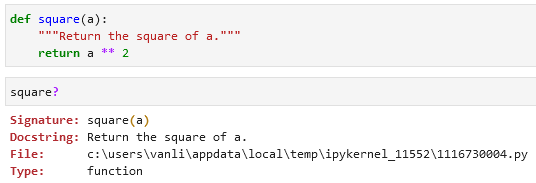
* Start the *IPython Shell* by typing ipython in the *Anaconda Prompt*.
* Launch Jupyter lab (?) $ jupyter lab

**Access documentation** with help() like help(len). The alternative is ? like len?. Get information on objects using ? like in the example below:

A screenshot of a computer

Description automatically generated

You can also get information on functions or other objects you create like in the example below which has a docstring (a description of the function):



?? provides the source code of the object you are curious about:

A screenshot of a computer

Description automatically generated

Use wildcard matching (character \*) to list every object in the namespace whose name ends with *Warning*:

A screenshot of a computer screen

Description automatically generated

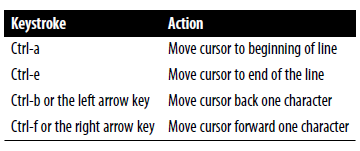
In the example below we area looking for a string method that contains the word *find* somewhere in its name:

A screenshot of a computer

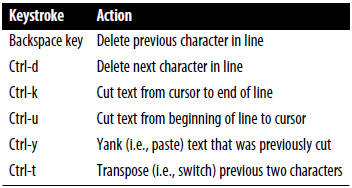
Description automatically generated

**Keyboard Shortcuts in the IPython Shell**

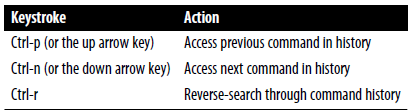
**Navigation Shortcuts**



**Text Entry Shortcuts**



**Command History Shortcuts**

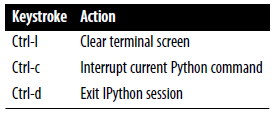


Use *Ctrl-r* to browse the search results. When you’re done press *Enter*.

A close up of text

Description automatically generated

**Miscellaneous Shortcuts**



**Enhanced Interactive Features**

**IPython Magic Commands**

Magic commands are prefixed by the % character.

* **Line magics** are denoted by a single % prefix. They operate on a **single line of input**.
* **Cell magics** are denoted by a double %% prefix and operate on **multiple lines of input**.