Exercise Manual for Course 1264

Introduction to Data Science, Machine Learning and AI with Python

1264/MA/A.2/003/

by Mary Flynn

Exercise Manual Contents

Legend for Course Icons..... i
Hands-On Exercise 1.1: Exploring Jupyter and Loading Data In Python..... 1
Hands-On Exercise 2.1: Working With Data in Python..... Jupyter
Hands-On Exercise 2.2: Exploring and Visualizing Data in Python..... Jupyter
Hands-On Exercise 2.3: Transforming Data in Python in Preparation for Analysis... Jupyter
Hands-On Exercise 3.1: Text Mining..... Jupyter
Hands-On Exercise 4.1: Performing Regression Analysis With Python..... Jupyter
Hands-On Exercise 5.1: Decision Tree Classifiers..... Jupyter
Hands-On Exercise 6.1: Working With a Neural Network Using Python..... Jupyter
Hands-On Exercise 6.2: Working With Logistic Regression Using Python..... Jupyter
Hands-On Exercise 6.3: Working with a Naive Bayes Model in Python..... Jupyter
Hands-On Exercise 7.1: Cluster Analysis on Structured Data With Python..... Jupyter
Hands-On Exercise 8.1: Generating Association Rules From Transaction Data..... Jupyter
Hands-On Exercise 8.2: Generating Recommendations From Transaction Data. ..... Jupyter
Hands-On Exercise 9.1: Creating and Visualizing Network Graphs..... Jupyter
Hands-On Exercise 9.2: Analyzing Networks..... Jupyter
Hands-On Exercise 10.1: Data Visualization ..... Jupyter

Standard icons are used in the hands-on exercises to illustrate various phases of each exercise.

Major step

Warning

1. O **Action**

Hint

Checkpoint

Stop

Question

Congratulations

Information

Bonus

Solution/Answer

Hands-On Exercise 1.1:

Exploring Jupyter and Loading Data In Python

Objectives

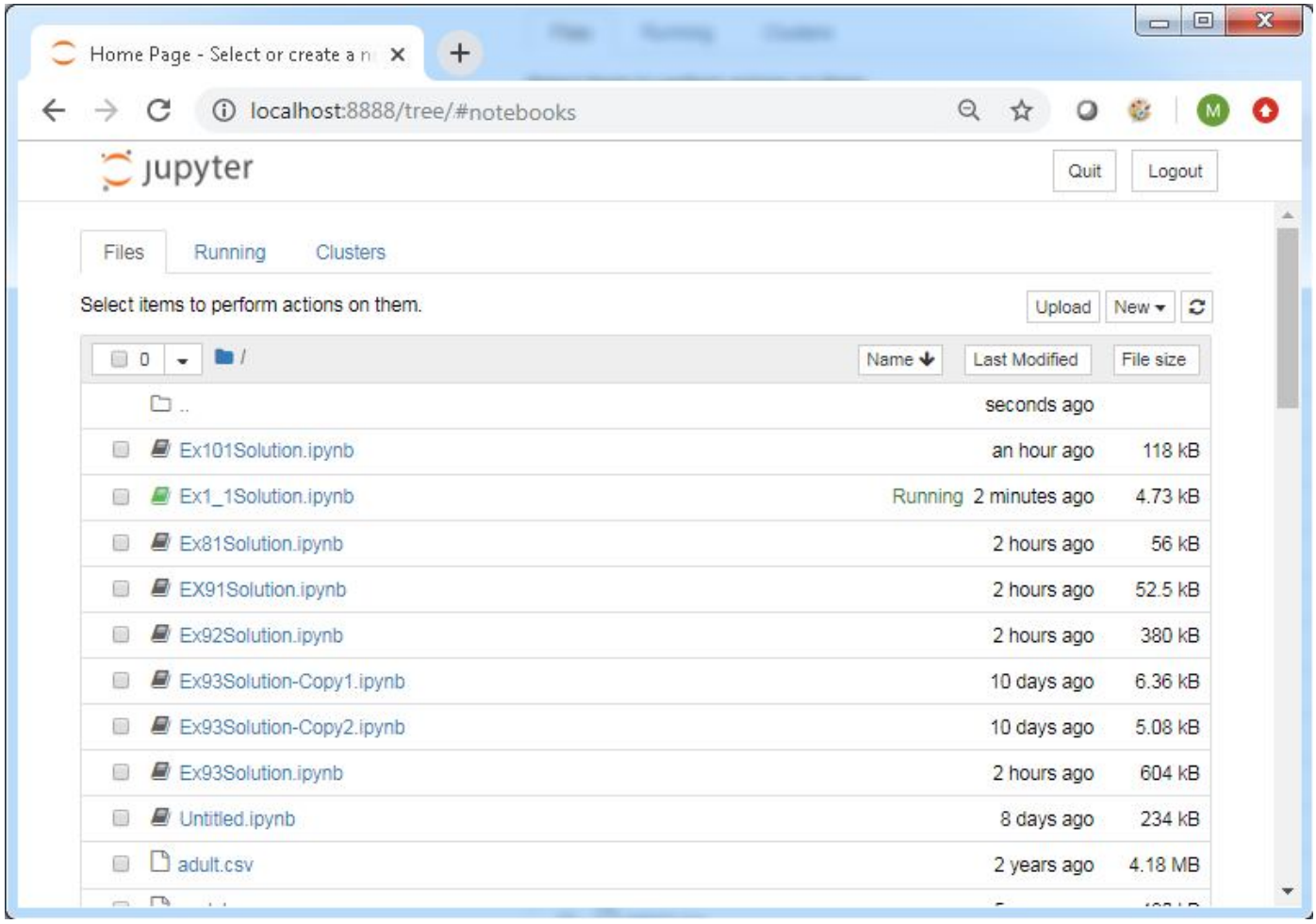
In this exercise, you will explore the Jupyter environment and the different types of data that can be loaded into Python for processing.

Overview

Jupyter Notebooks allow the creation and sharing of documents that contain live code, equations, visualizations and narrative text. This exercise will allow you to explore that environment as well as load .csv files and web pages (webscraping) as data sources.

Major Step 1: Reviewing the Jupyter environment

1. ❑ Start jupyter_notebook from the desktop. A jupyter notebook window will open and you will see some log output. A browser will then open with the following URL: http://localhost:8888/tree/#notebooks (http://localhost:8888/tree/#notebooks) showing a folder of files similar to the following:



2. ❑ To start a new notebook, click on **New** towards the top right of the dashboard. On the dropdown that appears, select **Python 3** as your kernel. A new tab will appear in the browser.

3. ❑ Rename the new Notebook from *Untitled* to **MyFirstNotebook** by selecting File->Rename or double-clicking the name, to the right of the Jupyter icon at the top

4. ❑ On the first cell, try some simple computations such as the following:

```
print(3+2)
```

Hint: Python is case sensitive

It should look like the cell below

In [1]:

```
print(3+2)
```

```
5
```

5. ❑ Execute the cell by pressing Ctrl+Enter and notice a number appearing within the [ ] to the left of the cell once it has executed. Try re-running the cell and notice the number change

6. ❑ Try printing some text such as the following in a new cell:

```
print('Hello' + ' 1 ' + 'World')
```

It should look like the cell below

In [2]:

```
print('Hello' + ' 1 ' + 'World')
```

```
Hello 1 World
```

7. ❑ Experiment with inserting cells above and below previous cells containing similar Python commands (i.e. calculations or printed text), and note how cells can be executed independently or how all cells in the notebook can be executed, depending on your requirements

Hint: Menu->Insert->Insert Cell Above

8. ❑ Experiment with clearing the output of cells and re-executing them

Hint:
Menu->Cell->Current Outputs->Clear or
Menu->Cell->All Output->Clear

9. ❑ Insert a cell and change its type to *Markdown*. Enter some plain text in the cell and then execute the cell. Note the difference in indentation of the output

10. ❑ Double-click on the Markdown cell to edit it. Change the font of the text by placing a # symbol at the beginning of the line, and execute the cell

11. ❑ Repeat step 10 with ## and ### symbols sequentially and note the difference in font size each time. This is a method of placing headers in your Jupyter Notebooks for presentation purposes

12. ❑ Experiment with typing several commands (for example, type a few print commands of text and computations, similar to steps 4 and 6) into one cell and then splitting the cell in order to run the commands separately from each other

    Hint: Menu->Edit->Split Cell

13. ❑ Experiment with deleting cells

    Hint: Menu->Edit->Delete Cells

14. ❑ View the Python Documentation under the Help menu

15. ❑ Click on the menu Help -> User Interface Tour for an overview of the Jupyter Notebook App user interface

16. ❑ Go to Notebook Ex1_1 and continue with step 17 for the remainder of the exercise