Working with Data

The ABCs of Noteable workbook (CONTINUOUS LEARNING MATERIAL FOR COMPUTING SCIENCE TEACHERS FOLLOWING THE SQA CURRICULUM FOR COMPUTING SCIENCE)

Contents

Welcome to this introductory guide on using computational notebooks and data with Noteable.

Requirements if this is your first time using Noteable

Before making use of the materials described in this workbook, please make sure that you have done the following to gain access to Noteable:

1. Access supporting service Documentation: https://noteable.edina.ac.uk/documentation/
2. Access supporting video playlists for Schools using Noteable:
   https://studio.youtube.com/channel/UCKhcyiuFyq8xTUlg_DpKIlA/playlists
3. Review the template Data Protection Impact Assessment available for teachers, schools and Local Authorities:
   https://blogs.glowscotland.org.uk/glowblogs/digilearn/2021/08/11/noteable-dpia-information/
4. Request access to Noteable from your Local Authority if **access has not already been approved**
5. Once the request is approved, Noteable is available as an Application through the GLOW App Library.

*If you are an independent school, please contact your local GLOW officer or the Scottish Council for Independent Schools.

Definitions and further information:

Authentication: authentication within the Noteable service refers to the ability to provide access to a subset of users for a specified customer. Authentication is pluggable, supporting a number of authentication protocols. For schools in Scotland, authentication takes place through the GLOW system.

Documentation: the documentation made available to schools online via www.noteable.edina.ac.uk/Documentation or such other web address notified by EDINA to the Subscriber on an ad hoc basis, which sets out a description of the Service and the user instructions for the Service.

EDINA: a centre for digital expertise, based at the University of Edinburgh as a division of the Information Services Group providing the Noteable service.

Git: software for tracking changes in any set of files, usually used for coordinating work among programmers collaboratively developing source code during software development. Its goals include speed, data integrity, and support for distributed, non-linear workflows (thousands of parallel branches running on different systems).

GitHub: a provider of Internet hosting for software development and version control using Git. It offers the distributed version control and source code management functionality of Git, plus its own features.

Jupyter notebook: Jupyter notebooks are an open-source web application that facilitates the creation and sharing of documents that contain live code and supporting commentary in the form of an explanatory text. It is a platform that can be used throughout the academic process to organise and articulate elements of a teaching and learning workflow where programming and/or data analysis are involved. The Jupyter notebook web application is open source and supports interactive data analysis in over 40 programming languages.

Learning Management System: the web-based platform used by teachers to access Noteable. For teachers and students in Scottish school accessing Noteable, this will be through the GLOW system App Library.

**Service**: the Noteable service provided by EDINA, the University of Edinburgh

User ID: the unique identifier or username provided by a school for each individual Authorised User.
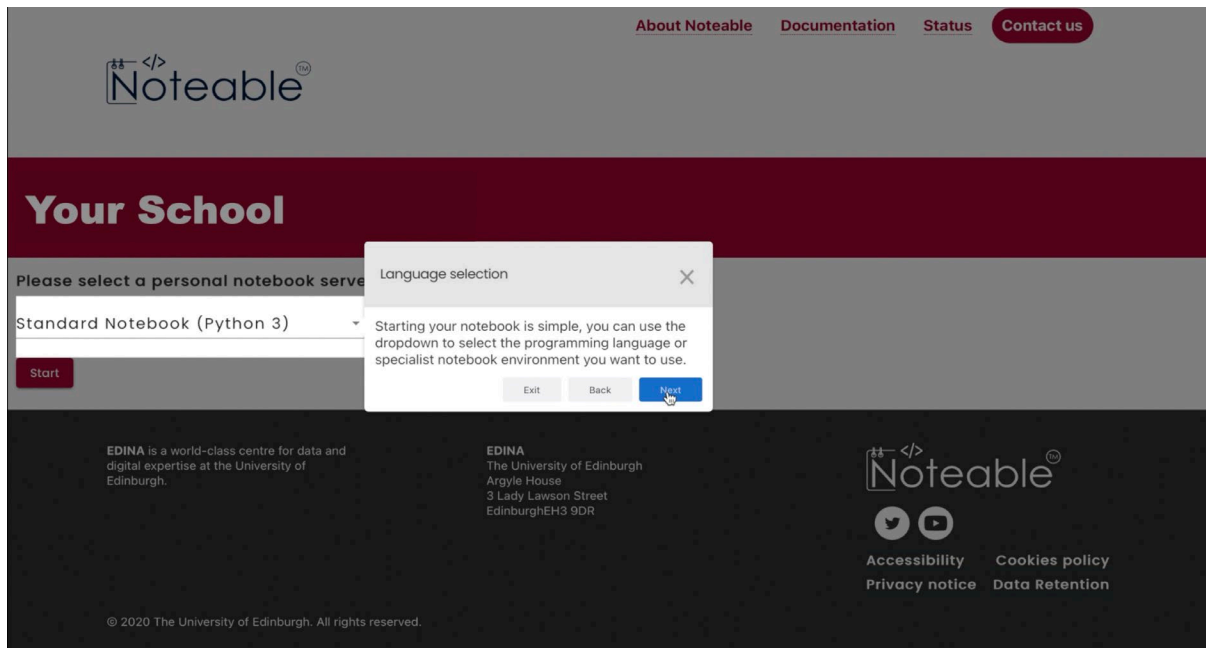
1. Noteable, notebooks and teaching

Noteable is a platform that adopts Jupyter notebooks and expands upon this open-source technology to write code and explanations in that allows for a Jupyter notebook file to be saved as an '.ipynb' file format, with additional features to help you in your teaching of computing science topics and other topics that include the Python and R programming languages.

Once you have accessed Noteable, you will have access to environments for coding activities, explanations, showing input and output, visualisations and more combined in one file and place. With Noteable, your files are saved online, and can be accessed any time by authenticating into GLOW and launching Noteable in the App Library.

If it is the first time you access Noteable, you will see this virtual guided tour screen of the service:
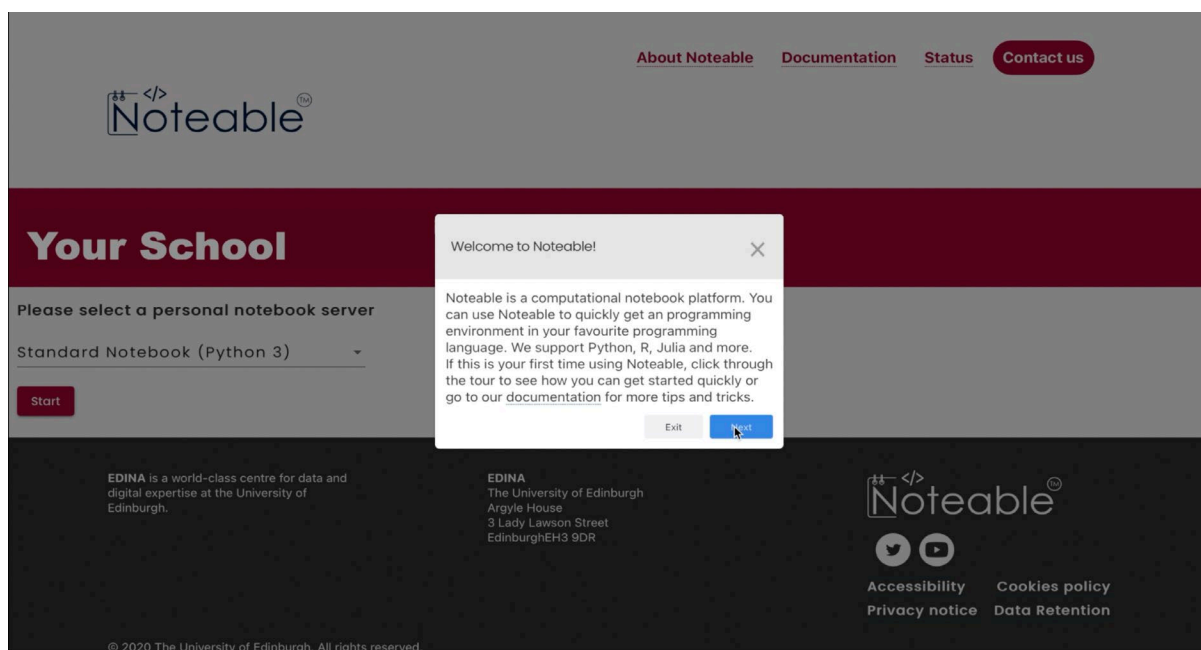
You will then be able to select a notebook server from the drop-down menu. The Standard Notebook on Noteable will be the default selected notebook when you first launch the service. You can go ahead and click 'Start' if you would like to launch to Standard notebook server, alternatively if you would like to choose another server, such as RStudio, you can select a different notebook server from the dropdown menu.

If you are looking to use Python on Noteable, select the default 'Standard Notebook'. The section below outlines some of the characteristics of Python as a programming language in Jupyter notebooks on Noteable.

2. Introduction to Python in Noteable for teaching the Scottish Computing Science curriculum

Python is an increasingly popular programming language.

Python can be used to code for a variety of reasons, including on a server through web applications such as Noteable, or on a server to create web applications!

It was created by Guido van Rossum and released to the public in 1991.

Python can be used for:

- Learning general programming tenets,
- web development (server-side),
- software development,
- mathematics,
- system scripting.

Python works on different platforms (Windows, Mac, Linux, Raspberry Pi, etc) and has a simple syntax similar to the English language.

Python has syntax that allows developers to write programs with fewer lines than some other programming languages.

Python runs on an interpreter system, meaning that code can be executed as soon as it is written. This means that prototyping can be very quick.

Python can be treated in a procedural way, an object-oriented way, or a functional way.

Quick guide to Python Syntax on Noteable

- Python was designed for readability and interactivity between coders, and has some similarities to the English language, with influence from mathematics.

- Python uses new lines to complete a command, as opposed to other programming languages which often use semicolons or parentheses.

- Python relies on indentation, using whitespace, to define scope; such as the scope of loops, functions and classes. Other programming languages often use curly brackets {} for this purpose.

- Noteable provides access to a Python coding environment with the additional Jupyter Notebook interface, which follows the standard Python syntax and also includes **cell type distinctions** so that you can add natural language explanations and text using **Markdown** and **create assignments.**

3. How to launch a blank Jupyter notebook with Noteable for teaching in Python

When you first access a new Python 3 Jupyter notebook, you will see these cell distinctions in the header, including a Notebook name, a Menu Bar including menus to navigate the formatting and management of the notebook file, the Toolbar to work through cells and content in the notebook, and an example of a blank Code cell ready for input from the user.
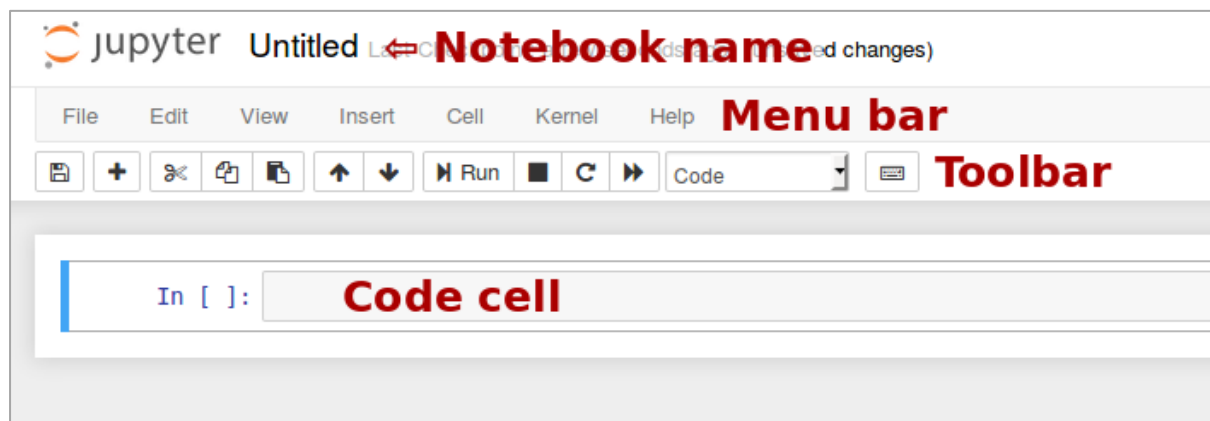
The code cell is where you will be able to code in Python straightaway, and you will see a blinking cursor in the code cell.

When you think about programming generally, and about coding specifically with the Python programming language, the first thing to note is that coding languages aren't like most other languages spoken between human beings. There are no vocabularies, alphabets or dialects. Instead, each coding language is unique and uses special commands and abbreviations in order to work properly. The Python and R programming languages are readily available on Noteable for teaching the Scottish Curriculum in Computing Science, Mathematics and other subjects. Additional materials have been developed focussing on each programming language as well. Please refer to the Github page with available resources for importing into Noteable: https://github.com/edina/Python-and-RStudio-school-resources

4. How to launch an RStudio notebook with Noteable for teaching in R

Similarly, to Python, R is not just a programming language, but it is also an interactive ecosystem including a runtime, libraries, development environments, and extensions. All these features help you and your students think about problems as a data scientist, while supporting fluent interaction between the user and the computer.

R has become a popular programming language used for statistical computing and graphical presentation, among other applications in computing, statistics, mathematics and other areas. Its most common use is to analyse and visualise data.

R can be used in the RStudio user interface, for uses including:

- data analysis, data visualization, data science and machine learning
- statistical techniques (such as statistical tests, classification, clustering and data reduction)
- It is easy to draw graphs in R, like pie charts, histograms, box plot, scatter plots
- It works on different platforms (Windows, Mac, Linux)
- It has many packages (libraries of functions) that can be used to solve different problems

When using R, it is up to you as the user whether you want to use the print() function to output code. However, when your code is inside an R expression (inside curly braces {}, for example), use the print() function to output the result.

Noteable provides access to R coding environments both through the RStudio user interface and with the additional Jupyter Notebook interface through the R with Stan notebook, which follows the standard classic Jupyter notebook view**.**

To launch RStudio with Noteable, once you have accessed Noteable through GLOW, you will have access to environments for coding activities, explanations, showing input and output, visualisations and more combined in one file and place. You can select RStudio from the notebook selection dropdown menu.

When you first access an RStudio notebook with Noteable, you will see 4 panels (shown in the image below):

- top-left: scripts and files
- bottom-left: R console
- top-right: objects, history, and environment
- bottom-right: tree of folders, graph window, packages, help window, viewer
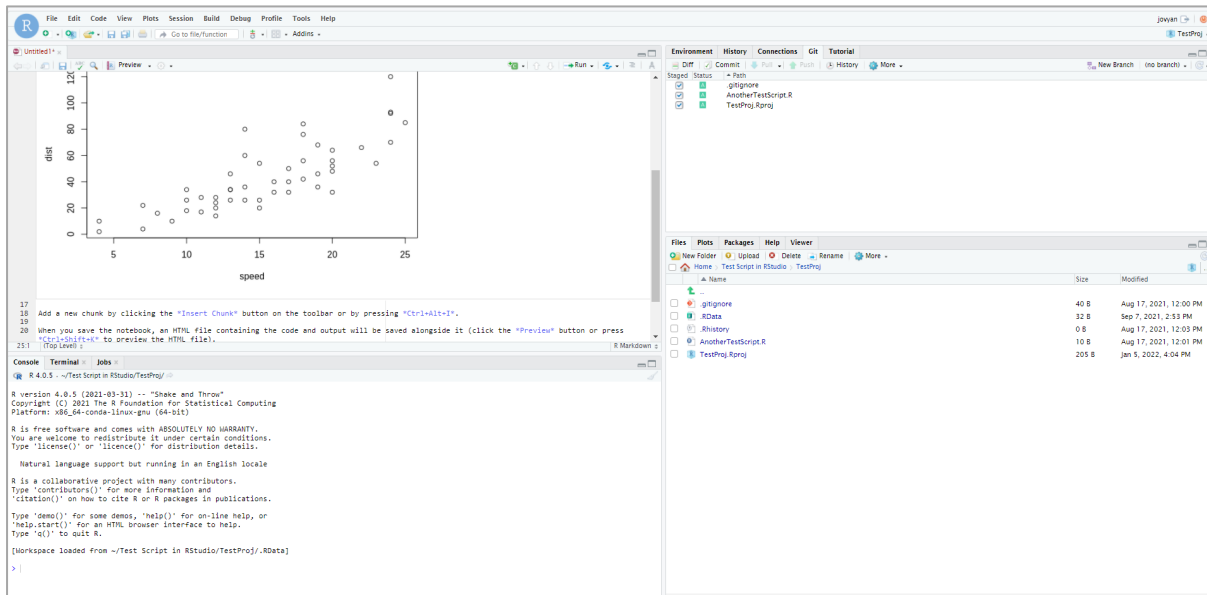
If you are new to RStudio, or would like to refer to this resource for learning how to programme in R, the following link leads to an excellent online book that provides information to learn how to programme in R, from loading data to writing functions: https://rstudio-education.github.io/hopr/preface.html

5. Available resources in Python and R

There are a number of available resources in Python that follow the Scottish Computing Science curriculum for schools and can be used with Noteable.

Alongside this resource, you can access activities in Python and in R programming languages.

These resources are available as Jupyter notebook and R files and can be accessed with Noteable in a few easy steps, described below.

Resources are available through the following link, and where suitable, PDFs can also be downloaded directly from GitHub.

Link to Python and RStudio school resources: https://github.com/edina/Python-and-RStudio-school-resources

6. Accessing Available Coding Resources for use with Noteable

Noteable includes a button that allows you to collect data from an external source on the internet and import it into Noteable. This feature works with Github, a provider of Internet hosting for software development and version control, including the ability to organise files such as Jupyter notebooks and R files that include teaching and learning materials.

To import Python and R files into Noteable using Github, you can use the '+GitClone' button available on Noteable when you access a Jupyter notebook environment. When you have uploaded data into Noteable using this method, the data is stored on your Noteable user profile throughout the duration of access to Noteable, unless you delete data. Guidance on how to do this is provided below.

Alternatively, you can also work on materials locally on your device and upload them to Noteable using the 'Upload' buttons available on the Jupyter notebook and RStudio user interfaces.

How to import data from Github on Noteable:

1. Launch Noteable through GLOW
2. Access a 'Standard Notebook' in Python 3 from the dropdown menu
3. Click on the '+GitClone' button from the Noteable service dashboard
4. Insert the link to Github for the desired data to import in the blank field for 'Git Repository URL' in the pop-up tab, remove any pre-existing text in remaining fields.
5. Click on 'Clone', this will import all data from the specified Github repository

Enter the details of the Git Repository to clone:

Git Repository URL:	insert link here
Branch*:	<default>
Username*:	
Password*:	

* Optional   Clone   Cancel

To import teaching and learning content for Python and RStudio, please use this URL in the pop-up tab: https://github.com/edina/Python-and-RStudio-school-resources

Available Resources for Python, R and Noteable on GitHub

This repository serves to host the teaching and learning content: **https://github.com/edina/Python-and-RStudio-school-resources**

The repository is for coding activities developed in Jupyter notebooks and RStudio using the Noteable platform (www.noteable.edina.ac.uk) for Scottish teachers and learners. It includes the following materials:

Jupyter notebooks for Python in the Scottish Computing Science Curriculum:

1. About the Python Lab
2. Computer Programmes
3. Algorithms
4. Flowcharts
5. Sequence, Selection, Iteration
6. Printing Messages
7. Inputs and Variables
8. Maths
9. If Statements
10. Loops
11. Lists
12. Functions
13. Activities and Solutions (from across all the notebooks above)

This repository also includes a sample of activities that can be carried out in RStudio for Statistics Higher and Advanced Higher. These activities can be referred to as classroom materials or independently. The topics these RStudio activities cover include:

1. Using comparative boxplots
2. Using standard deviation
3. Correlation and regression, choosing an appropriate model
4. Testing Hypotheses

Available External Resources for Python, R and Noteable

References for Python

1. Callysto Lesson Plans, including lesson plans to solve a variety of problems (including TED-Ed Riddles) using Python code in Jupyter notebooks:

https://www.callysto.ca/lesson-plans/

For each problem, there are three resources:

- Teacher lesson plan (how to introduce and teach the problem)
- Teacher Jupyter notebook (including instructions)
- Student Jupyter notebook

2. An online starter's guide to Jupyter Notebooks:
   https://www.analyticsvidhya.com/blog/2018/05/starters-guide-jupyter-notebook/

3. Markdown for Jupyter notebooks cheat sheet, including information on how to format Markdown cells in Jupyter notebooks: https://www.ibm.com/docs/en/watson-studio-local/1.2.3?topic=notebooks-markdown-jupyter-cheatsheet

4. IBM guidance on using Jupyter notebooks: https://www.ibm.com/docs/en/watson-studio-local/1.2.3?topic=data-notebooks
5. Python W3 Tutorials: https://www.w3schools.com/python/

References for R and RStudio

1. Hands-on programming with R: https://rstudio-education.github.io/hopr/basics.html#summary

2. R W3 Tutorials: https://www.w3schools.com/r/

Licensing information