Python Workshop 1 Demonstration

2026 – Wendy Olsen v1.0

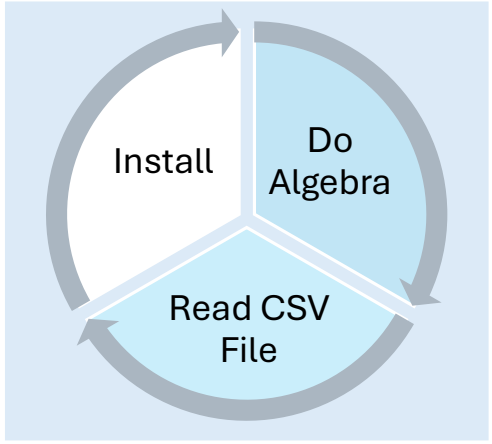
Online location of key documents

12.00 Introduction, 12.10 Demonstration of Python for population (or investment & savings) growth forecasting 12.20 Installing & Setup Q&A; speedqueries. 12.45 Demonstration of setup and 7 basic python skills. These are how to output a CSV file; make a line graph; add a package; do algebra; make a list and transform it; what is the 'pandas dataframe' punctuation vs. the Pandas Series? Subset a data table. 1.15 activity on your own laptop.

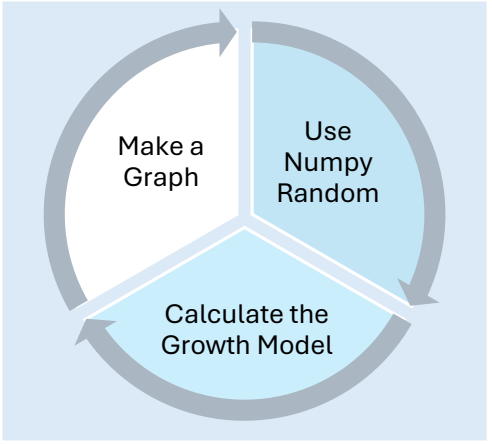
HOW TO USE PYTHON and PYSPARK

v 1.1 2026

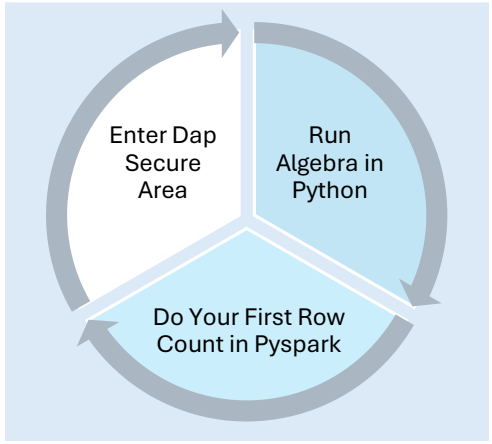
Hints, Workshop 1



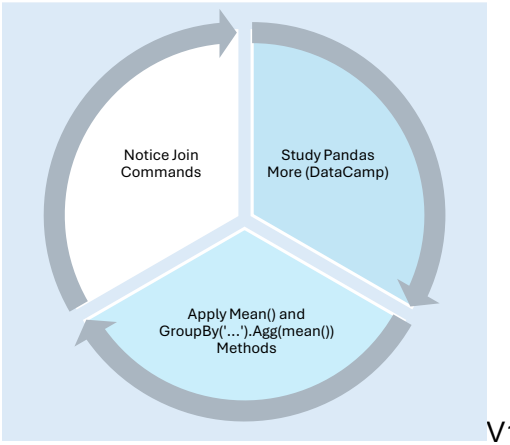
Workshop 2



Workshop 3



Self-Training in Pyspark



V1.0 + minor edits

Basic python programming

Input	Method()	Function	A type of entity	Put it together
Read…	Sort_values()		Tuple; list;	*See program.*
		Can be user-defined	Dataframe;	*The sample code is about*
		Can have default parameters	Array	*Population.*

print	Display	[entity name]	File type csv output	Image .png output
Print(mydata)	Display(mydata)	mydata	Write_csv(mydata.,,,)	***First, make a graph***

Help is at hand! See Zed Shaw's book
(online $ 29)
URL Learn Python the Hard Way
Buy this at a used bookshop URL
https://www.abebooks.co.uk/

Four kinds of loops and where to get help with writing computer code

If loops	For loops	While loops	Iterative use of object	Iterative Function, ie a function embedded in a for loop for example
If condition: indent… is python code format.	For condition or term: Indent … is python code format.	While … action	A. what is an iterable object? Something like a sentence or a word	Def myamends(): activity 1 activity 2 return amendeddf
	You need to avoid the risk of unending looping	You need to avoid the risk of unending looping	B. What is itertools? In python, it can easily discover each digit in a number, each letter in a word, or each word in a sentence.	Notice above that the indentation is crucial to the defined function. Suppose Activity 2 is a "for" loop. It has an output product, the amended dataframe

Help is at hand! See Zed Shaw's book (online $ 29) Chapters 29 to 35. URL Learn Python the Hard Way Or buy it at used bookshop URL https://www.abebooks.co.uk/

Sample Program Chunk in Python

- `myagesexdat = pd.read_csv("D:\\pyprogramsAndSims/data/SyntheticAgeSexImmigDataLarge.csv")`
- `mynoneu = myagesexdat[myagesexdat["type"] == "Noneu"]`

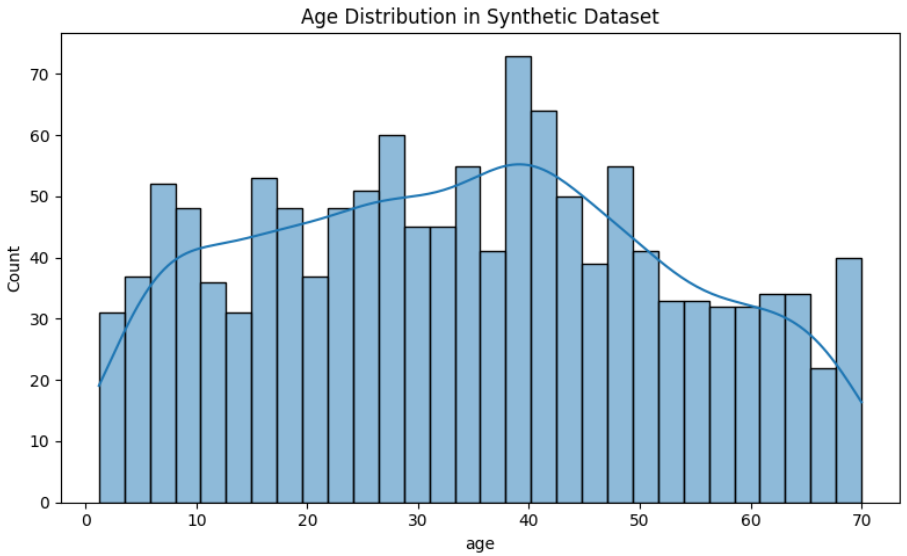
Print out 20 rows of data to the console.	Head() Tail()	What are head() and tail()?
Mydata.head(20)	Mydata.head()	These are 'methods'. Here the default is 5 rows.
	Notice the format is Entity.method()	They were written into Python language.
20 is the parameter	The parameter is used by the method.	

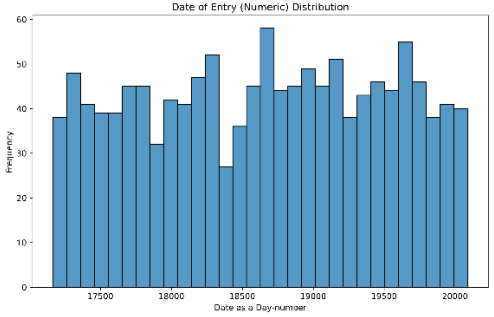
Sample Program Chunk in Python

- myagesexdat = pd.read_csv("D:\\pyprogramsAndSims/data/SyntheticAgeSexImmigDataLarge.csv")
- mynoneu = myagesexdat[myagesexdat["type"] == "Noneu"]
- mynoneu.head()
- myagesexdat = myagesexdat[myagesexdat["type"] == "Noneu"]
- myagesexdat.head()

Notice the input step	= is an assignment operator
[ ] is bracket notation, ideally square bracket notation, indicating rows by default	Dataframe[rownames] is an example of Entity[listnumbers]
To choose by row and column, use the .loc() or .iloc() method	***Can you see the assignment operator again?***
myagesexdat["type"] == "Noneu"] is a condition, which rows must conform to.	Entity[rowname] Is similar to Entity[rowcondition]
Six easy conditionals	==, ~=, >=, <=, <, >
These are operators.	= is the assignment operator. == checks for equality so it's a conditional operator.

Taster Figures



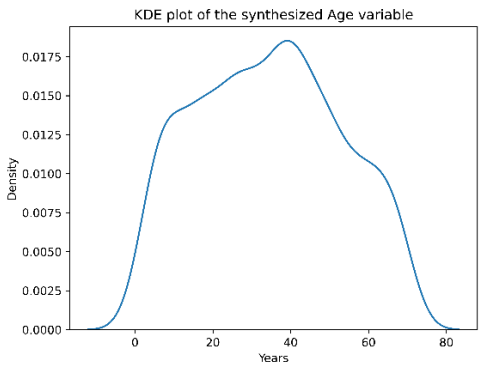


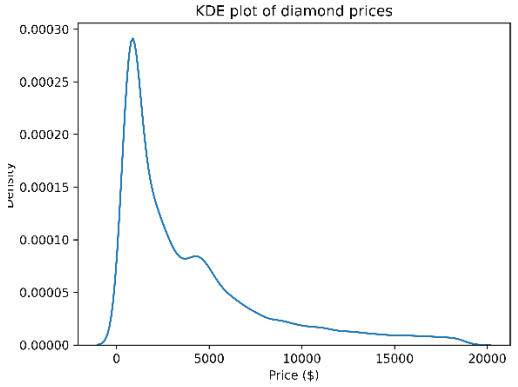
Data-science style

Histograms

Kernel density diagrams

Raise your ambitions





Resource documents

- ***Basic Python*** – installation and coding
- PSD_Capability_Building – uses basic Python.
- ***Basic resources:*** Session Resources - Folder containing materials for some coding upskilling sessions (Defines what “DAP” “CATS” Wiki is; shows how one runs Python in DAP by demonstration, step by step)
- Introduction to Python course; one of these is at URL https://learninghub.ons.gov.uk/course/view.php?id=1168
- also webpage file:///C:/Users/OlseNW/Downloads/ITP_course_intro%20(1).html.

Frequently asked questions

Questions & Answers

Learning Task 1

- Learning task 1: You are a Winner if you can calculate an answer to this problem using Python: 😊 Suppose the world population is 8 billion. Suppose the net pop growth rate is +0.85% per year. Calculate what year the world population would reach 14 billion. Later on, using a loop, substitute more accurate starter populations such as 8.2 billion, 8.45b.

Your activities

- Learning task 1- Extra challenge - You could gather the sum from this kaggle file of country populations, https://www.kaggle.com/datasets/asadullahcreative/world-population-by-country-2025 , accessed Dec. 2025. Print a list of the starting population and year of reaching the target.
-

Tasks 3 and 4

- Learning task 2: can you write a bit of code, which sets the path for output files out of Python in your Python interpreter? You will have a chance to ask questions about key terms like interpreter, path, and output files so it is a workshop open to all, including beginners. You may want to try xlsxwriter package (you must use pip to install this, then import it) and write a csv or xlsx file. I myself use VS Code instead of Spyder but it's up to you. Spyder, Jupiter and VS Code are known as 'interfaces' or 'VDI's for Python. Choose one.
- Learning task to jump into this as a Newbie: RunMS Copilot by using https://m365.cloud.microsoft/chat/?home=1&auth=2 [You will need to log in to MS 365 beforehand. Do not worry, it is fine to login to MS 365 but do it in your laptop. Don't do this in your secure area (DAP).] Ask copilot to write you some code to set the path for output files using your Python interpreter version 3.11 or which version you use. A beginner book: McKinney, Wes (2012), *Python for Data Analysis. 3rd edition available https://wesmckinney.com/book , 3rd ed. Pub. 2022.(Open Access)*

Answers

- See two types of document. The html or word document containing a full rundown of Introductory Exercises. Or the Python code files:
  - The file .ipynb is a Jupyter format document including some results. It's like an R markdown file.
  - The file .py is a Python base code document. You can run single lines, chunks or all of it, by SELECT ALL, Run, in your interpreter.

After you get experienced, you get stuck. This is normal.

- Climb the mountain of solving problems by trial-and-error.
- Make notes of your Debug solutions.
- Study the notes, study your 'operating system' and become more expert.

Legitimate sources on the internet

- A pandas tutorial for beginners
  - https://sparkbyexamples.com/python-pandas-tutorial-for-beginners/
- A high-schools website for python users, really accurate too:
  - https://www.w3schools.com/python/python_for_loops.asp
- How you make Spreadsheet outputs as tables,
  - See https://xlsxwriter.readthedocs.io/tutorial01.html, note tutorial 1:

- Facilitator email wendy.olsen@ons.gov.uk
- Facilitator personal github https://github.com/WendyOlsen

References / Where to go for help

- Chambers and Zaharia, *Spark: The Definitive Guide,* https://www.oreilly.com/library/view/spark-the-definitive/9781491912201/ helps you study chunks of the book. (Use the ***Read Now button***)
- McKinney, Wes (2012), *Python for Data Analysis: Data Wrangling with Pandas, Numpy, and Ipython*. Singapore: O'Reilly. *(3rd edition available https://wesmckinney.com/book , 3rd ed. Pub. 2022.(Open Access))*
- Shaw, Zed (2018) *Learn Python the Hard Way*, various editions, PDF is not free, available at https://www.informit.com/store/learn-python-the-hard-way-9780138270575 , see also https://learnpythonthehardway.org/) 5th ed., ISBN 0-13-827057-0, London: Addison-Wesley. Don't forget to scroll down….
- Practice skills: https://www.datacamp.com/courses/intro-to-python-for-data-science.