Python & VBA for Data Science - Syllabus

Course Summary:

Programming is now a core skill in the financial industry, and while Python is the most widely used language, VBA tools are spread all around. We teach students to use it to handle real cases:

- “Learning by doing” is the best way to process in coding. Students need to concretely write code.
- Real life cases should show a practical usage of programming, and should encourage students to learn to answer problems they can meet in their professional career.

Students are supposed to use their laptops, with Anaconda pre-installed.

Course can be done in French / English or Spanish, but the content is in English.

Course objectives

- Learn the basics of both languages (variables, functions …)
- Code functions to answer specific problems
- Use Python to interact with data (I/O operations, webscrapping, SQL databases)
- Perform data analysis and basics machine learning with Python
- Be familiar with the statistical package commonly used (Numpy, Pandas, Scikit-learn)
- Overview on how coding is applied in the financial industry

Content to be covered within the course:

- VBA within Excel
- Different uses cases of data automation with Excel, where VBA is very convenient
- The Anaconda suite, with a focus on Numpy, Pandas and Scikit-Learn librairies
- Basics of Python
- Functions in Python
- Statistical refresher (mainly errors to avoid in the real life)
- The different steps of data engineering, with Python
- Machine Learning basics
- Example of trading model
- How to interact with data in the real life (databases, Excel files …)

Students assessment

- Coding exercises (number to be determined) to be given along the course.
- A group work for each programming language
- A final MCQ

Course content

Session 1: Introduction and VBA (1)
Global introduction to the course
Introduction to VBA / Programming languages / Lists & Tuples / Loops / …

Session 2: VBA (2)
Exercises, inspired from real cases in financial industry

Session 3: Introduction to Python
How coding is used in the financial industry
Introduction to Python / Programming languages / Lists & Tuples / Loops / …
Objective is to present the basics of programming.

Session 4: VBA (4)
Final VBA project

Session 5: Data structuring (1)
Introduction to the Python project.
Information flow is key in Market finance: everything is about moving data. Session organized around data, with practical cases.
/ Reading writing files (.txt, .xlsx, .csv)
Theory and then practice
/ Web scrapping & data acquisition
Theory and then practice
/ Databases (reading writing)
Theory and then practice: this is very SQL related, but it is also important

Session 6: Data structuring (2)

Session 7: Data Science introduction
Objective is to move forward with more practical things, to give students the tools to apply them during the next two modules
/ Presentation on the importance of interpretability and communication.
/ Numpy and Pandas : libraries for data manipulation
/ Matplotlib: data visualization
/ Sklearn : Machine Learning libraries, go through several algorithms

Session 8: Analysis (1)
Objective is to use the different tools together for analysis purpose.
/ Presentation on a real model currently running
/ Working on modeling futures prices

Session 9: Analysis (2)
Keep working on the modeling exercise
/ Presentation on a second real model currently running

Session 10: Analysis (3)
Keep working on the modeling exercise