

# On Site 파이썬 중급 실무 교육

2020. 1차:11.24~11.27.

2차:12.8.~12.11.

한국지질자원연구원 국제지질자원인재개발센터에서는 연구실 맞춤형 교육과정인 On Site 파이썬 실무 교육 과정을 개설하고자 합니다. 교육과정은 국제지질자원인재 개발센터에서 총 2회 진행되며, 1회: 2020년 11월 24일 ~ 11월 27일. 2회: 2020년 12월 8일 ~ 12월 11일. 4일간 다음의 주제로 진행될 예정입니다.

■ 강의일정

일정	주제	대표강사
1차:11.24.(화) 2차:12.8(화)	<b>Anaconda 패키지의 설치와 python의 기본 습득</b> <ul style="list-style-type: none"> <li>Anaconda 패키지의 설치와 이용</li> <li>Jupyter notebook 활용과 기본적 구동원리 이해</li> <li>오픈소스 문화에 기반한 코딩법 습득</li> <li>python 기본 문법 및 활용</li> </ul>	박찬희 (KIGAM)
1차:11.25.(수) 2차:12.9.(수)	<b>개발환경설정과 Pandas의 다양한 활용</b> <ul style="list-style-type: none"> <li>Shell에서의 개발환경 설정</li> <li>git의 간단 설명과 github을 통한 자료 습득</li> <li>Pandas의 다양한 활용</li> </ul>	
1차:11.26.(목) 2차:12.10(목)	<b>Time series data 예제와 응용</b> <ul style="list-style-type: none"> <li>Time series 데이터 처리, 분석 및 데이터 시각화</li> <li>Customized python coding</li> </ul>	
1차:11.27.(금) 2차:12.11.(금)	<b>오픈소스와 연계하여 python을 이용한 DIYs</b> <ul style="list-style-type: none"> <li>Linear Regression 외</li> <li>세가지 형태의 Machine Learning</li> <li>특화된 모듈 설치 (다양한 OS에서의 설치)</li> <li>Python in action (Analytical solutions, ParaView, Raspberry Pi, etc) - 여력이 될 경우</li> </ul>	

## ■ 과정안내

### • 교육목적 및 필요성

- ① 연구를 수행하면서 마주치는 다양한 상황과 생산되는 데이터의 가공 및 해석에 특화된 맞춤형 해결에 대한 요구도가 기하급수적으로 증가
- ② 오픈소스로 인해 과거와는 전혀 다른 방식인 단순 코딩의 습득이 아니라 코딩 문화의 체득으로 얻어지는 효율의 학습과 이에 대한 경험
- ③ 수많은 컴퓨터 언어중 하나인 개인에서 시작된 python이 자연스럽게 커뮤니티를 통해서 성장하고 발전한 이유처럼 수강자는 이를 수용하고 이용하여 자신만의 솔루션 구현

### • 내용 및 범위

- ① python을 자신의 컴퓨터에 설치하여 매일 옆에 두고 사용하는 계산기와 같이 항상 접근하여 생활 속에서 활용 가능한 환경 구현
- ② python의 기본 문법을 예제와 실습을 통해서 습득
- ③ pandas, matplotlib를 사용하여 데이터를 수집하고 분석하여 표현
- ④ 간단한 통계적 분석 및 머신러닝의 구현

### • 수강을 위해 필요로 하는 조건

- ① 개인의 wifi가능한 노트북과 누구나 부담 없이 재미로 접근해 보는 관심과 개성
- ② Learning-by-doing-it (일단 하면서 배우는)에 대해 부담이 없는 자세
- ③ 실패해도 실망하지 않는 낙관주의

### • 교육대상

- ① 자신의 직업이나 살아가면서 컴퓨터 언어 하나는 배워보고 싶다고 생각한 사람
- ② 내가 생산하는 데이터가 엑셀로만 표현하려니 좀 부족하다고 생각하는 사람
- ③ 뭔가 나의 연구를 자동화해보고 싶은 사람
- ④ 세상에 있는 python 프로그래머가 과업을 수행하는데 처음부터 끝까지 자기가 혼자서 다하는지 궁금한 사람
- ⑤ 이 모든 것을 수행하는데 영어로 적힌 문서를 읽는데 부담이 없는 사람

## • 강의 주제 및 내용 요약

과정 첫째 날은 오픈소스로 공개되어 있는 python을 개인의 노트북에 다운받아 설치하여 직접 사용할 수 있는 환경을 구현한다. 이때 노트북은 윈도우, 맥, 리눅스 등의 다양한 환경의 컴퓨터 OS를 지원하는 걸 원칙으로 한다. 강의의 처음부터 끝까지 다양성과 개인화는 존중하며, 문제가 생길 시에 직접 개인의 컴퓨터에 접근하여 문제를 해결하고, 이미 단계별 과업을 스스로 수행한 수강자는 그러지 못한 수강자를 도우면서 2차 혹은 3차 학습을 자연스럽게 하도록 유도한다.

과정 둘째 날은 python에서 광범위하게 사용되는 pandas와 matplotlib를 이해하고 실습한다. 이때 실습 데이터는 수강자가 자신의 데이터를 사용해서 학습하고 싶어 할 경우에는 그 데이터를 이용하고 그러지 않을 경우 학습에 이용할 데이터를 제공해서 학습을 진행한다.

과정 셋째 날은 시계열 데이터를 예제로 처리하고 응용한다. 시계열의 데이터가 아닌 자료라도 충분히 범용성이 있어 보이면, 직접 그 데이터의 활용 목적에 맞는 python 코딩을 함께 수행한다.

마지막으로 과정 넷째 날은 강사가 직접 국가 R&D를 수행하는데 이용했던 자료를 수강자가 선호할 경우에 예제로 이용을 하거나, 특별한 데이터에 대한 지원이 없을 경우에는 학습 자료로 제공되는 데이터를 이용하여 간단한 통계처리 분석이나 초보 단계의 Machine Learning 예제를 수행하며 마무리 한다. 그 외에 여유가 되어 한주 동안 배웠던 부분을 더 강화하거나 놓쳐서 단계학습에 필요한 부분에 질문이 있을 경우 답변하고, 동의된 주제를 더 깊게 탐구하는 방법을 수행한다.

### Day 1. Anaconda 패키지의 설치와 python의 기본 습득

과정 첫째 날은 오픈소스로 공개되어 있는 python을 개인의 노트북에 다운받아 설치하여 직접 사용할 수 있는 환경을 구현한다. 이때 노트북은 윈도우, 맥, 리눅스 등의 다양한 환경의 컴퓨터 OS를 지원하는 걸 원칙으로 한다. 강의의 처음부터 끝까지 다양성과 개인화는 존중하며, 문제가 생길 시에 직접 개인의 컴퓨터에 접근하여 문제를 해결하고, 이미 단계별 과업을 스스로 수행한 수강자는 그러지 못한 수강자를 도우면서 2차 혹은 3차 학습을 자연스럽게 하도록 유도한다. Anaconda 패키지로 설치되는 항목 중에서 Jupyter notebook을 활용하고 기본적 구동원리를 OS 플랫폼 별로 설명하여 이해시킨다.

그리고 나서 Jupyter notebook에서 직접 python의 기본 문법을 예제와 설명을 통해서 익히고 다양한 코딩을 실제로 수행한다. 코딩 시 주입식이 아닌 오픈소스 문화에 기반을 둔 코딩 법을 강조하여 단순 암기식이 아니라 필요에 따른 코드를 즉시 탐구해서 구성해 가는 방법을 소개한다. 이러한 식의 코딩을 위해서 각 수강자의 인터넷 연결은 필수적이다.

### Day 2. 오픈소스를 활용한 다양한 라이브러리의 개인화

과정 둘째 날은 서로 공유해야할 데이터나 학습하고 난 후의 코드는 github을 통해서 공유할 계획이다. 따라서 github을 이용할 수 있는 기본적 git에 대해서 간단한 강의를 한다. 이는 그다음에 배울 python의 예제들에서 필수 보완적인 단계학습의 지식과 경험이 된다. 이 과정속에서 shell을 이용하여 개발환경을 설정하고, 개발자의 중요한 도구중 하나인 vscode를 소개하고 이용한다. 이런 툴과 플랫폼 하에서, python에서 광범위하게 사용되는 pandas와 matplotlib를 이해하고 실습한다.

### Day 3. 오픈소스와 연계하여 python을 이용한 DIYs (Do It Yourself)

과정 셋째 날은 pandas 자체만으로도 광범위한 자료가 있기에 article 별로 소개하면서 자연스럽게 익숙

해지고, 데이터가 처리된 경우엔 데이터의 이해를 위해서 자연스럽게 시각화를 시도할 계획이다. 주로 Time series 데이터 처리와 분석 그리고 전공적 영역에서 어떻게 활용될 수 있는지에 대한 예제들을 소개할 계획이다. 시계열의 데이터가 아닌 자료라도 충분히 범용성이 있어 보이면, 직접 그 데이터의 활용 목적에 맞는 python 코딩을 함께 수행한다.

## Day 4. 오픈소스와 연계하여 python을 이용한 DIYs (Do It Yourself)

마지막으로 과정 넷째 날은 간단한 통계처리 분석이나 초보 단계의 Machine Learning 예제를 수행한다. Machine Learning의 경우 간단하게 python에서 구현할 수 있는 세 가지 형식들을 소개한다. 통계 처리와 Machine Learning에서 특화된 모듈을 이용할 경우, 이러한 모듈들의 설치를 하는 방법을 소개하여 다양한 모듈의 검색과 자신의 OS에 맞는 설치를 경험한다. 단순 데이터 처리 뿐만 아니라, python을 이용하여 해석 해를 구하는 프로그램이라든지, 3차원 고단위 graphics를 외부 전문 software, 가령 ParaView와 연동해서 사용하는 방법이라든지, 혹은 IoT 장비인 Raspberry Pi에서 사용된 예들을 간단히 소개한다. 그 외에 여유가 되어 한주간 배웠던 부분을 더 강화해서 배우고 싶어 할 경우, 동의된 주제를 더 깊게 탐구하는 방법을 수행한다.

## ■ 강사소개

- 현재 KIGAM 소속 지질환경연구본부, 심지층연구센터에서 근무하는 책임연구원으로 30년 이상의 다양한 코딩 경험을 가지고 있으며, KIGAM내 오픈소스플랫폼 동아리에서 활동중
- THMC 모델 중 하나인 OpenGeoSys(UFZ, Germany)의 개발자들 중 한명으로 다양한 국가 R&D에 소스코드(Kernel 중심)를 특화 수정하여 활용하고 있음
- 1996년부터 Linux를 메인 OS로 사용하여 석박사 과정(토목환경공학) 모두 Linux 및 윈도우기반으로 연구 모델(FEM, FDM, FVM)을 개발하여 학위 취득 (석사: University of Waterloo, Canada, 박사: GeorgiaTech, USA)
- C/C++, Python, Java, zsh, bash, 등을 활용한 다양한 환경의 플랫폼을 개발하여 IoT등 다중학제를 융합하여 지질자원 분야에 적용하는 연구(OGSFLAC, DECOVALEX 2023, OpenIoT 지중열교환기 및 열저장, etc)를 수행하고 있음