

# Python과 OpenAPI를 활용한 데이터

## 수집과 분석 대시보드 구축

### 강의목표

- Data Science와 Engineering의 강력한 도구인 Python과 주요 라이브러리를 활용하여 필요한 데이터를 발굴/수집/분석하는 방법 학습을 목표로함
- 다양한 분야의 OpenAPI를 찾고, 문서를 읽고, 데이터를 수집하는 방법을 학습하고 웹사이트에서 데이터를 Scraping(또는 Crawling)하는 방법을 학습
- 수집한 데이터를 목적에 맞게 가공하고 정제하는 방법, 저장하는 방법 학습(OpenRefine 활용)
- 수집한 데이터를 기반으로 데이터 대시보드를 구성하고분석하는 방법을 학습 (Google Data Studio등을 활용)

### 강의계획 및 학습내용

날짜	내용
1주차	데이터 수집/가공/처리를 위한 기본개념과 Python 기초 학습 - 1  학습목표: Python을 개괄적으로 다루는 것이 아닌, 데이터 수집/가공/처리와 관련한 핵심 내용을 중심으로 학습 <ul style="list-style-type: none"><li>- 강의개요</li><li>- Google Colab 활용법 소개<ul style="list-style-type: none"><li>- Google Colab 이용방법 소개</li></ul></li><li>- 환경설정<ul style="list-style-type: none"><li>- Python 개발환경 및 Google Refine 설치 등</li></ul></li><li>- Python 기본 학습 (1)<ul style="list-style-type: none"><li>- 데이터 처리와 관련한 Python 주요 문법 학습</li></ul></li></ul>
2주차	데이터 수집/가공/처리를 위한 기본개념 및 Python 기초 학습 - 2  학습목표: 1주차에 이어 데이터의 처리를 위한 Python의 기본기 학습 <ul style="list-style-type: none"><li>- Python 기본 학습 (2)</li></ul>

	<ul style="list-style-type: none"> <li>- 데이터 처리와 관련한 <b>Python</b> 주요 문법 학습 계속</li> <li>- 데이터의 형식과 구조 학습</li> <li>- 데이터 수집/가공/처리를 위한 기본개념 등 학습</li> </ul>
3주차	<p>데이터 유형별 수집 방법 (1)</p> <p> 학습목표: <b>OpenAPI</b>를 찾고, 개발 문서를 읽는 방법 그리고 원하는 데이터를 가져오는 방법 학습</p> <ul style="list-style-type: none"> <li>- <b>OpenAPI</b>를 통한 데이터 수집 (네이버 검색, 기업정보, 공공데이터 등 <b>OpenAPI</b> 활용방법 소개)</li> <li>- <b>Python</b>을 활용한 <b>OpenAPI</b> 데이터의 저장 (파일)</li> </ul>
4주차	<p>데이터 유형별 수집 방법 (2)</p> <p> 학습목표: <b>OpenAPI</b>가 제공되지 않거나 웹사이트를 통해서만 확인할 수 있는 데이터를 수집할 수 있는 방법 학습</p> <ul style="list-style-type: none"> <li>- 웹사이트 <b>Scraping(Crawing)</b>을 위한 기초</li> <li>- <b>Python</b>을 활용한 웹사이트 데이터 수집과 저장</li> </ul>
5주차	<p>데이터 저장방법</p> <p> 학습목표: 수만건 단위는 데이터 수집과 저장에 유용한 <b>Google Spreadsheet</b>의 <b>API</b> 활용 방법 학습</p> <ul style="list-style-type: none"> <li>- 파일 또는 <b>Google Spreadsheet</b>를 활용한 데이터의 저장</li> <li>- <b>Google Spreadsheet API</b> 활용 방법</li> <li>- <b>Database</b>를 활용한 저장(<b>SQLite</b> 또는 <b>MongoDB</b>)</li> </ul>
6주차	<p><b>Google Spreadsheet, OpenRefine</b> 및 정규식을 활용한 데이터 정제</p> <p> 학습목표: 강력한 데이터 정제 도구인 <b>OpenRefine</b>을 활용하여 지저분한 데이터를 정리하고 품질을 높일 수 있는 방법을 학습</p> <ul style="list-style-type: none"> <li>- <b>OpenRefine</b> 기능을 활용한 다량의 데이터 정제</li> <li>- 정규식 기본 학습</li> <li>- <b>OpenRefine</b> 소개 및 기본 학습</li> </ul>
7주차	<p> 학습목표: 실제 비즈니스 환경에서의 데이터 수집, 정제, 자동화 구축 프로세스를 사례 중심으로 학습하고 학습한 내용을 기반으로 조별 프로젝트 수행</p> <ul style="list-style-type: none"> <li>- 데이터 정제 및 수집자동화 사례공유</li> </ul>

	<ul style="list-style-type: none"> <li>- 조별 프로젝트 수행 (데이터 가공/정제와 저장)</li> </ul>
8주차	<p>데이터 대시보드 구축</p> <p> 학습목표: 수집하고 정제한 데이터를 활용하여 분석 관점을 도출할 수 있도록 데이터를 시각화하거나 대시보드로 구축하는 방법을 학습</p> <ul style="list-style-type: none"> <li>- <b>Google Data Studio</b> 기본 학습</li> <li>- 수집한 데이터 기반의 대시보드 구축 학습</li> </ul>
9주차	<p> 학습목표: 지금까지 학습한 기술이 어떤 형태로 구현되어 데이터 제품으로 시장에 제공되는지를 실제 솔루션 등의 사례를 중심으로 학습</p> <ul style="list-style-type: none"> <li>- 데이터 대시보드 구축 및 사례공유</li> <li>- 조별 프로젝트 수행 (대시보드 구축)</li> </ul>
10주차	<p> 학습목표: 지금까지 학습한 것을 종합하여 대시보드를 구축하는 팀 프로젝트 수행결과를 발표</p> <p>프로젝트 발표</p>