

2021

작심두달

- 웹 크롤링과 업무자동화 -

Copyright 2021. STARTCODING All rights Reserved

목차

1. 멘토소개
2. 강의목표
3. 수업방식에 대한 이해
 - flipped learning
 - lecture & assignment
 - slack
4. 웹 크롤링이란 무엇인가?
5. 웹 크롤링 사용예시
6. 웹 크롤링에 필요한 지식
 - 프로그래밍 언어(python)
 - 웹구조(html, css)
7. 커리큘럼

1. 멘토소개



권기준 멘토

(현) 위즈라이브 자바스크립트 온라인 튜터 강사

(현) Reverse Mountain 웹 크롤링 강사

(현) 초중고 4차 산업 진로체험 특강 강사

(현) 탈잉 웹 개발, 파이썬 튜터

(전) NICE I&T 개발실 Web Developer

(전) Aram-In 개발팀 Engineer

(전) Link-B 개발팀 Engineer

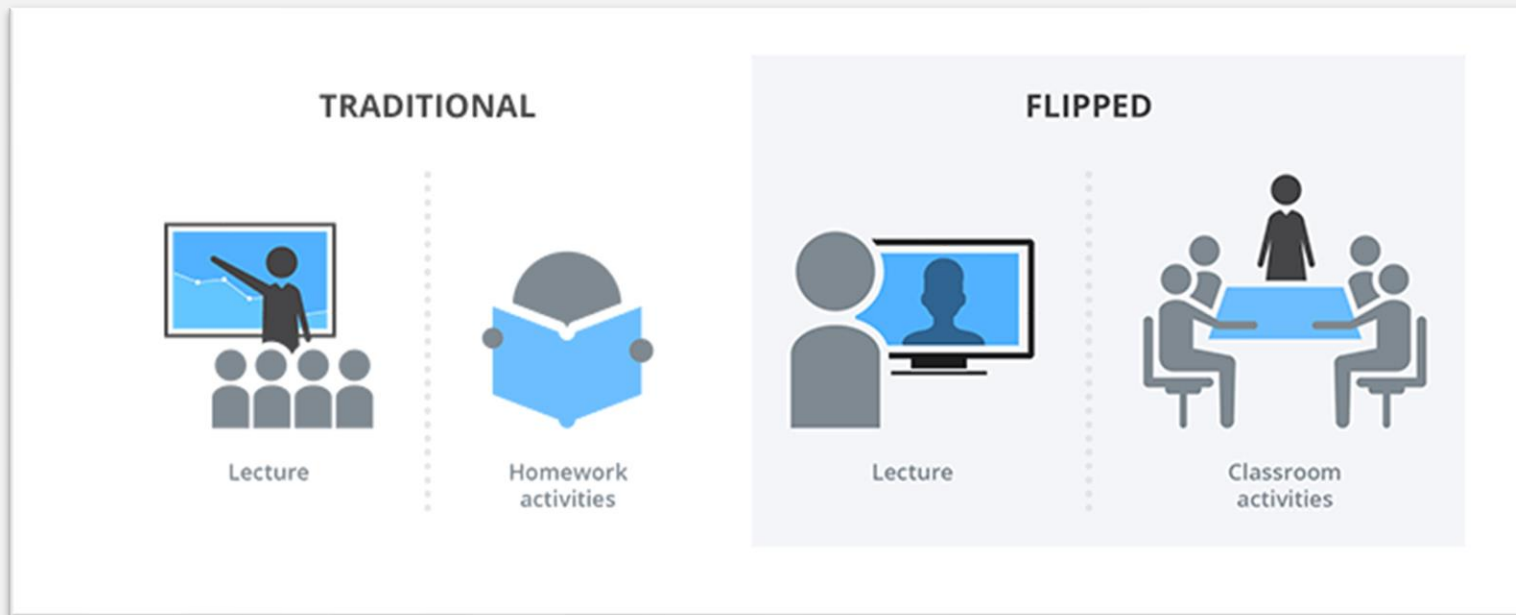
[멘토 운영 블로그 STARTCODING 바로가기](#)

2. 강의목표

웹 크롤링에 필요한 **지식을 습득**하고
이를 활용해
반복적인 업무를 도와주는 프로그램을
"**스스로**" 작성하는 능력을 갖는 것

3. 수업방식에 대한 이해

flipped learning?



<https://www.epiphan.com/blog/what-is-a-flipped-classroom/>

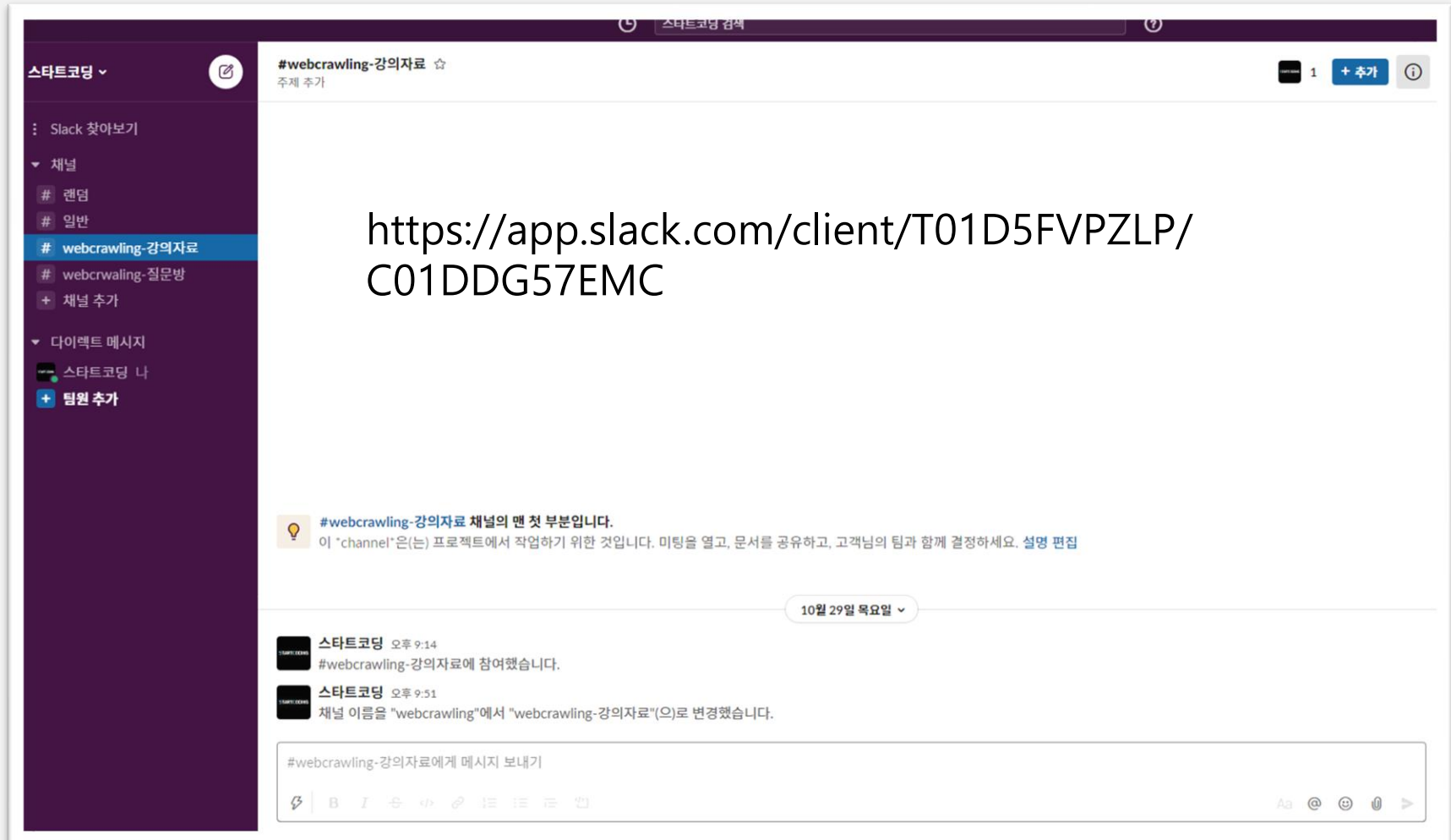
3. 수업방식에 대한 이해

lecture & assignment

멘토가 직접 작성한 자료와 과제
+외부자료 (유튜브 강의, 웹사이트)

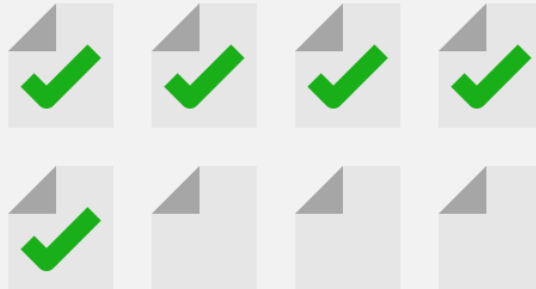
3. 수업방식에 대한 이해

slack (업무용 협업툴)



4. 웹 크롤링이란 무엇인가?

웹 크롤링?



4. 웹 크롤링이란 무엇인가?

웹(Web)상에 존재하는 수많은 데이터를
선별적으로 수집하는 것

5. 웹 크롤링 사용예시

시연

1. 프로그래밍언어



* 각자의 실력이 다르지만
같은 출발선에 있다 생각하기

* 기초가 탄탄해야 응용을 할 수 있음

6. 웹 크롤링에 필요한 사전지식

2. 웹구조(HTML, CSS선택자)

```
<!DOCTYPE html>
<html lang="ko" data-dark="false" data-useragent="Mozilla/5.0 (Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/86.0.4240.183 Safari/537.36">
  <head>...</head>
  <body cz-shortcut-listen="true"> == $0
    <div id="u_skip">
      <a href="#newsstand">...</a>
      <a href="#themecast">...</a>
      <a href="#timesquare">...</a>
      <a href="#shopcast">...</a>
      <a href="#account">...</a>
    </div>
    <div id="wrap">
      <style type="text/css">...</style>
      <div id="NM_TOP_BANNER" data-clk-prefix="top" class="_1hiMwemA" style="background-color: #3c3f48;">...</div>
      <div id="header" role="banner">
        <div class="special_bg">...</div>
        <!--EMPTY-->
        <div id="gnb">...</div>
      </div>
      <div id="container" role="main">
        <div style="position:relative;width:1130px;margin:0 auto;z-index:11">...</div>
        <div id="NM_INT_LEFT" class="column_left">...</div>
        <div id="NM_INT_RIGHT" class="column_right">...</div>
        <a id="NM_scroll_top_btn" href="#wrap" class="content_top">...</a>
        <button id="NM_darkmode_btn" type="button" role="button" class="btn_theme" aria-pressed="false">...</button>
        ::after
      </div>
      <div id="footer" role="contentinfo">...</div>
    </div>
    <div id="adscript" style="display:none">
      <script src="https://siape.veta.naver.com/fxshow?su=SU10578&da_dom_id=p_main_business_2&tb=BUSINESS_1&calp=-
&rui=1604563220342&main_svt=20201105170015" charset="utf-8" data-da="true"></script>
      <script src="https://siape.veta.naver.com/fxshow?su=SU10577&da_dom_id=p_main_business_1&tb=BUSINESS_1&calp=-
&rui=1604563220343&main_svt=20201105170015" charset="utf-8" data-da="true"></script>
    </div>
    <div style="display: none;">...</div>
  </body>
</html>
```

6. 커리큘럼

커리큘럼

1st week

Part1. 오리엔테이션 |
프로그래밍이란 무엇일까요?

[수업 내용]
ZOOM 활용방법 알기
강의 소개 및 Slack 사용법, 구글드라이브 이용방법 알기
웹 크롤링 예제 소개
파이썬 프로그래밍 기초 외부 강의자료 소개 및
이론과제 진행방법 소개
[이론 과제]
한입에 쏙 파이썬: 3시간 만에 끝내는 파이썬 기초 강의 완강
Palette를 통해 수강생들끼리 하루 공부한 양을 공유

2nd week

Part2. 준비운동 |
파이썬 걸음마 떼기

[수업 내용]
실시간 파이썬 문제풀이를 통해 학습한 내용을 문제에 적용
자료형, 변수, 출력문, 입력문, 조건문, 반복문, 함수 등 프로그래밍을 하는데 필수인 개념들을 반복 학습
파이썬 프로그래밍 기초 외부 강의자료(2) 소개 및 이론과제 진행방법
[이론 과제]
초단기 속성 코딩 / 프로그래밍 강좌(파이썬) 완강
Palette를 통해 수강생들끼리 하루 공부한 양을 공유

3rd week

Part2. 준비운동 |
파이썬 활용 + 웹 구조 학습

[수업 내용]
파이썬 중급 문제풀이를 통해 프로그래밍 실력을 조금 더 향상
PyautoGUI 라이브러리 사용법을 익히고, 간단한 매크로 프로그램 제작
웹 구조(HTML, CSS) 학습방법 소개
[이론 과제]
멘토 블로그를 활용해서 간단한 웹페이지 제작
Palette를 통해 수강생들끼리 학습 진행상황을 피드백
[실습 과제]
PyautoGUI를 활용한 매크로 프로그램 개발

4th week

Part3. 실전! 웹 크롤링 |
(1) Request, BeautifulSoup

[수업 내용]
Request, BeautifulSoup 모듈을 이용해 웹사이트 정보를 가져오는 방법을 학습
다양한 웹구조에 적용해 봄으로써, 크롤링 기초를 탄탄하게 만들 수 있다.
네이버 지식인 검색결과 크롤링, 네이버 금융 실시간 주가 크롤링
[실습 과제]
실시간 환율 정보 크롤링
지식인 검색결과 크롤링 프로그램 업그레이드(페이지 여러 개 가져오기)
Palette를 통해 수강생들끼리 과제 진행상황을 공유

6. 커리큘럼

커리큘럼

5th week

Part3. 실전! 웹 크롤링 I
(2) BeautifulSoup 심화과정
(페이징과 엑셀 저장)

[수업 내용]
페이징(Paging) 처리 방법 피드백
크롤링 결과 엑셀 저장 방법 학습 (openpyxl 모듈 사용법)
실시간 주가 크롤링 및 엑셀 저장
[실습 과제]
네이버 뉴스 크롤링
Palette를 통해 수강생들끼리 과제 진행상황을 공유

6th week

Part3. 실전! 웹 크롤링 I
(3) 이미지 크롤링 방법,
Selenium 사용법 기초

[수업 내용]
네이버 뉴스 크롤링 프로그램 피드백
네이버 이미지 크롤링 프로그램 제작
Selenium 사용법 기초 학습
네이버 자동 로그인 프로그램, 다음 자동 로그인 프로그램 제작
[실습 과제]
유튜브 댓글 수집 프로그램
Palette를 통해 수강생들끼리 과제 진행상황을 공유

7th week

Part3. 실전! 웹 크롤링 I
(4) Selenium 심화과정

[수업 내용]
유튜브 댓글 수집 프로그램 피드백
인스타그램 자동 좋아요, 댓글작성 프로그램 제작
Selenium 심화 크롤링 기법 정리 (팝업창 제어, iframe 제어)
[실습 과제]
이제까지 배운 내용을 토대로 내가 원하는 크롤링 프로그램 제작
Palette를 통해 수강생들끼리 제작과정을 공유

8th week

Part4.
나만의 크롤링 프로그램 제작하기

[수업 내용]
크롤링 프로그램 제작 피드백
데이터분석 및 업무자동화 소개
수업 마무리 정리

감사합니다

design icon from flaticon.com