

# Python 데이터 분석과 이미지 처리

나동빈

# 네이버 영화 리뷰 데이터 분석 1. 영화 정보 크롤링

## 리뷰 정보 클래스

```
import urllib.request
from bs4 import BeautifulSoup

class Review:
    def __init__(self, comment, date, star, good, bad):
        self.comment = comment
        self.date = date
        self.star = star
        self.good = good
        self.bad = bad

    def show(self):
        print("내용: " + self.comment +
              "\n날짜: " + self.date +
              "\n별점: " + self.star +
              "\n좋아요: " + self.good +
              "\n싫어요: " + self.bad)
```

# 네이버 영화 리뷰 데이터 분석 1. 영화 정보 크롤링

## 리뷰 정보 크롤링 함수

```
def crawl(url):
    soup = BeautifulSoup(urllib.request.urlopen(url).read(), "html.parser")
    review_list = []
    title = soup.find('h3', class_='h_movie').find('a').text
    div = soup.find("div", class_="score_result")
    data_list = div.select("ul > li")

    for review in data_list:
        star = review.find("div", class_="star_score").text.strip()
        reply = review.find("div", class_="score_reple")
        comment = reply.find("p").text
        date = reply.select("dt > em")[1].text.strip()
        button = review.find("div", class_="btn_area")
        sympathy = button.select("strong > span")
        good = sympathy[0].text
        bad = sympathy[1].text
        review_list.append(Review(comment, date, star, good, bad))

    return title, review_list
```

# 네이버 영화 리뷰 데이터 분석 1. 영화 정보 크롤링

## 리뷰 정보 크롤링 실습

```
title, review_list = crawl("https://movie.naver.com/movie/bi/mi/basic.nhn?code=36944")
print('제목: ' + title)
for review in review_list:
    review.show()
```