

# 11/10 - 11/15 Mini Project

미니 프로젝트 1  
주제 : 웹에서 수집한 데이터 DBMS(MySQL)에 저장하고, 분석하기  
(1) 데이터로부터 얻을 수 있는 질문 만들기 (수정 가능)  
(2) 데이터 수집 : 웹에서 스크래핑  
(3) 데이터 저장 : 테이블 설계, DBMS로 저장 (with pymysql)  
(4) 데이터 분석 : 질문에 답할 수 있는 다양한 분석 쿼리 활용

부트 캠프에서 첫번째로 진행된 미니 프로젝트의 주제로는

무신사 스토어에서 월간 아우터 랭킹 데이터를 웹스크래핑하고, 분석하는 것으로 정해졌다.

## 첫번째. 질문 만들기

1. 월별 전체 아우터 순위
2. 리뷰 수 많은 아우터 브랜드 순위(top 10)
3. 아우터 브랜드 top 10 중 추가 할인 진행 브랜드
4. 브랜드별 아우터 중 쿠폰 할인 브랜드
5. 가격대별 아우터 순위
6. 누적 판매량별 아우터 순위

## 두번째. 데이터 수집

```
import requests
from bs4 import BeautifulSoup

url = "https://www.musinsa.com/ranking/best?period=month&age=ALL&mainCategory=002&subCategory=&leafCategory=&price=&golf=false&kids=fa

user_agent = "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/107.0.0.0 Safari/537.36"

headers = {"User-Agent" : user_agent}
response = requests.get(url, headers)
response.raise_for_status()

soup = BeautifulSoup(response.text, 'lxml')
```

```
# 누적판매량 추가
ii = 0
driver = webdriver.Chrome()
while ii < 90 :

    driver.get("https://www.musinsa.com/ranking/best?period=month&age=ALL&mainCategory=002&subCategory=&leafCategory=&price=&golf=false&kids=fa
    items = driver.find_elements_by_css_selector("img.lazyload.lazy")
    items[ii].click()
    time.sleep(1)
    sum_sale = driver.find_element_by_id("sales_1y_qty")
    sst = sum_sale.text

    if '만' in sst :
        if '.' in sst :
            sst = sst.replace('.', '')
            sst += '000'
        else :
            sst += '0000'

    sst = sst.replace('만 개 이상', '')

    elif '천' in sst :
        if '.' in sst :
            sst = sst.replace('.', '')
            sst += '000'
```

```

        sst += '00'
    else :
        sst += '000'
    sst = sst.replace('천 개 이상','')

else :
    sst = sst.replace('개 이상','')

# 브랜드
brands = soup.select('form#goodsRankForm div.list-box.box div.article_info p.item_title a')

brd = ''+brands[ii].text+''

# 제품명
names = soup.select('form#goodsRankForm div.list-box.box div.article_info p.list_info a')

nm = ''+(names[ii]['title'])+''

# 다 넣은 곳에서 쿠폰 있느거 없느거 찾기
all_items = soup.select('div.right_area.page-contents-best form#goodsRankForm div.list-box.box .li_box')

a = all_items[ii].find('p', { 'class' : 'mu-icon mu-icon-coupon'})
if a == None :
    coup = ''x''
else :
    coup = ''o''

# 리뷰 수
points = soup.select('div.right_area.page-contents-best form#goodsRankForm div.list-box.box div.article_info p.point span.count')

pp = points[ii].text.replace(',','')

# 가격대
prices = soup.select('div.right_area.page-contents-best form#goodsRankForm div.list-box.box div.article_info p.price')

price11 = prices[ii].text.split()

if len(price11) == 2:
    pri1 = price11[0][0:-1].replace(',','')
    pri2 = price11[1][0:-1].replace(',','')
else:
    pri1 = price11[0][0:-1].replace(',','')
    pri2 = 0

print('INSERT INTO musinsa VALUES ('+ii+1, ',', brd, ',', nm, ',', pri1, ',', pri2, ',', pp, ',', coup, ',', sst, ');')

ii += 1

```

순위, 브랜드명, 제품명, 쿠폰의 유무, 리뷰 수, 가격대, 누적 판매량을 수집했다.

### 세번째. 데이터 저장

참고 → [pymysql 기초](#)

```

create table musinsa (
    ranks int not null,
    brand varchar(20) not null,
    title varchar(150) not null,
    ori_price varchar(20),
    dis_price varchar(20),
    review_count int,
    coupon varchar(20),
    sale_amt int(10),
    primary key(ranks)
);

# 데이터 저장 코드 생략

```

### 네번째. 데이터 분석

컬럼순 : ranks / brand / title / ori\_price / dis\_price / review\_count / coupon / sale\_amt

```

# 리뷰 수 많은 아우터와 브랜드 순위(top 10)
select * from musinsa
order by review_count desc
limit 10;

```

6	토피	2WAY 스웨트 후드 집업 (MELANGE GREY)	45000	37000	52900	x	149000
71	무신사 스탠다드	베이식 블레이저 [블랙]	79900	67890	46214	x	21000
37	무신사 스탠다드	릴렉스드 베이식 블레이저 [블랙]	81900	73690	27521	x	44000
82	무신사 스탠다드	오버사이즈 블레 이저 [블랙]	83900	71290	24392	x	20000
15	무신사 스탠다드	캐시미어 블렌드 오버사이즈 싱글 코트 [블랙]	159900	143890	19098	x	12000
7	라퍼지스토어	[리뉴얼 ver.][3M 신솔레이트]오리 지널 M-1965 피 쉬테일 파카 _Original Khaki	138000	89900	17402	o	37000
49	라퍼지스토어	[매트블랙] 버핑레 더 미니멀 포켓 블 루종_Matt Black	94000	65900	12626	o	30000
20	도프제이슨	오버핏 비건레더 싱글 자켓	129000	82550	12423	x	23000
67	파르티멘토	[비건 레더] 필드 자켓 블랙	111700	72600	12414	x	19000
2	스파오	베이직 푸퍼 _SPJPC4TC11	69900	62910	11741	x	41000

→ 아직 한겨울이 오지 않아서(11월 초!) 패딩보다는 후드집업, 블레이저, 자켓등의 제품 리뷰가 많은 것을 확인할 수 있음. 데이터를 수집한 사이트가 무신사 스토어인 만큼, 브랜드도 무신사 스탠다드의 비중이 높은 것을 확인할 수 있다.

```
# 랭킹 중 할인율 가장 많이 하는 브랜드 top10
select brand, count(*) '할인 진행 제품수'
from musinsa
where dis_price != 0
group by brand
order by count(*) desc
limit 10;
```

라퍼지스토어	10
커버넛	7
스파오	4
쿠어	4
파르티멘토	4
코드그래피	4
에일	4
무신사 스탠다드	4
드로우핏	3
도프제이슨	3

→ 가장 할인을 많이 하는 브랜드는 라퍼지스토어로, 90위 랭킹 중 10개의 제품이나 세일 중이다.

```
# 브랜드별 아우터 중 쿠폰 할인 브랜드(1페이지)
select brand
from musinsa
where coupon = 'o'
group by brand;
```

쿠어

라퍼지스토어

예일

노이어

무신사 스탠다드

케이투

파르티멘토

→ 90위 랭킹 중 쿠폰을 증명하는 제품의 브랜드는 총 7개가 나왔다.

가격대별 아우터 순위 (1페이지)

```
# 10만원 이하
select * # 19개
from musinsa
where ori_price <= 100000
order by ranks;
```

2	스파오	베이직 푸퍼 _SPJPC4TC11	69900	62910	11741	x	41000
6	토피	2WAY 스웨트 후드 집업 (MELANGE GREY)	45000	37000	52900	x	149000
9	스파오	베이직 퍼플리스 집업 _SPFZC4TU01	29900	26910	8273	x	34000
10	라퍼지스토어	버핑레더 오버핏 블루종_Black	94000	65900	6060	o	22000
11	스파오	파스텔 푸퍼 _SPJPC4TG01	69900	62910	5695	x	21000
30	무신사 스탠다드	오버사이즈 MA-1 보머 재킷 [블랙]	89900	0	4664	x	9000
37	무신사 스탠다드	릴렉스드 베이식 블레이저 [블랙]	81900	73690	27521	x	44000
39	예일	(22FW) WARM+ UP QUILTING JACKET BLACK	99000	49500	1090	x	9300
43	예일	(22FW) WARM+ UP QUILTING JACKET IVORY	99000	74250	680	o	4000
44	무신사 스탠다드	리버서블 플리스 숏 패딩 재킷 [블 랙/크림]	89900	0	214	o	2200
47	커버넌트	플리스 집업 자켓 블랙	79000	55300	11611	x	14000
49	라퍼지스토어	[매트블랙] 버핑레 더 미니멀 포켓 블 루종_Matt Black	94000	65900	12626	o	30000
50	제멋	리치 오버 MA-1 자켓 블랙 YHJK2269	66900	0	10497	x	16000
53	커버넌트	숏 키팅 자켓 블랙	89000	62300	3483	x	6400
63	후아유	리버서블 후드웬	99900	69900	568	x	2400

		업 WHJJC4T01U					
64	예일	EMBROIDERY UNIVERSITY DAN FLEECE TRACK JACKET BLACK	79000	59250	409	o	2800
71	무신사 스탠다드	베이식 블레이저 [블랙]	79900	67890	46214	x	21000
75	스파오	리버서블 푸퍼 _SPJDC4VC14	89900	80910	7921	x	24000
82	무신사 스탠다드	오버사이즈 블레 이저 [블랙]	83900	71290	24392	x	20000

```
# 10만원 에서 20만원 사이
select * # 43개
from musinsa
where ori_price between 100000 and 200000
order by ranks;
```

4	내셔널지오그래픽	N224MDW920 헤론 U넥 경량 다 운 점퍼 CARBON BLACK	149000	0	720	x	5200
7	라퍼지스토어	[리뉴얼 ver.][3M 신솔레이트]오리 지널 M-1965 피 쉬테일 파카 _Original Khaki	138000	89900	17402	o	37000
8	파르티멘토	[22F/W 리뉴얼] 오버사이즈 멜톤 더플 롱 코트 네이 비	155000	99900	3990	x	15000
12	무신사 스탠다드	캐시미어 블렌드 오버사이즈 발마 칸 코트 [블랙]	199900	0	1199	x	2600
13	코드그래피	[내피분리형] US- ARMY 밀리터리 피쉬테일 후드 롱 파카_카키	139000	98900	895	x	5300
14	예일	REVERSIBLE WARM UP QUILTING JACKET BLACK / IVORY	129000	96750	591	o	3700
15	무신사 스탠다드	캐시미어 블렌드 오버사이즈 싱글 코트 [블랙]	159900	143890	19098	x	12000
16	코드그래피	[DIVO] RDS 유 털리티 덕다운 푸 파 숏패딩_블랙	124000	99900	3236	x	5100
17	라퍼지스토어	[무봉제]덕다운 심 리스 미니멀 푸퍼 숏패딩_Black	143000	85900	9143	o	16000
19	내셔널지오그래픽	N224MFJ910 남 성 코스토니 플리 스 뽀글이 후드 폴 집업 SUGAR SWIZZLE IVORY	179000	159000	205	x	2200

20	도프제이슨	오버핏 비컨레더 싱글 자켓	129000	82550	12423	x	23000
21	커버넛	리버시블 웨르파 집업 자켓 아이보 리	129000	90300	5879	x	8900
23	라퍼지스토어	올리치 발마칸 오 버 코트_Black	173000	99900	8270	o	17000
25	내셔널지오그래픽	N224WDW920 여성 헤론 U넥 경 량 덕다운 점퍼 L BEIGE	149000	0	499	x	3000
27	라퍼지스토어	[리뉴얼 ver.][3M 신슬레이트] M- 1965 피쉬테일 숏 파카_Vintage Black	129000	79900	4400	o	17000
28	디스커버리 익스 페디션	픽토르 U 구스다 운 경량패딩 (BLACK)	159000	0	396	x	2000
29	라퍼지스토어	[리뉴얼 ver.][3M 신슬레이트]오리 지널 M-1965 피 쉬테일 파카 _Vintage Black	138000	89900	5089	o	16000
31	무신사 스탠다드	다운 푸퍼 숏 패딩 재킷 [블랙]	119900	0	6492	o	5200
33	내셔널지오그래픽	N224MFJ910 남 성 코스토니 플리 스 뽀글이 후드 폴 집업 CARBON BLACK	179000	159000	216	x	1900
34	도프제이슨	솔리드 무튼 자켓	186000	119000	6993	x	9600
35	라퍼지스토어	덕다운 미니멀 블 루중 패딩_Black	138000	89900	300	o	2500
38	무신사 스탠다드	캐시미어 블렌드 대디 숄더 싱글 롱 코트 [블랙]	189900	0	1983	x	5400
41	커버넛	RDS 커버라이트 카라리스 다운 점 퍼 블랙	119000	83300	2571	x	6600
42	커버넛	리버시블 카라리 스 웨르파 집업 자 켓 아이보리	119000	83300	1914	x	3900
48	코드그래피	[DIVO] RDS 유 틸리티 덕다운 푸 파 숏패딩_콜그레 이	124000	99900	6138	x	8200
51	내셔널지오그래픽	N224WDW920 여성 헤론 U넥 경 량 덕 다운 점퍼 CARBON BLACK	149000	0	290	x	1800
54	내셔널지오그래픽	N224MDW920 헤론 U넥 경량 다 운 점퍼 BEIGE	149000	0	303	x	1600
56	커버넛	RDS 씨빅 다운 숏 푸퍼 블랙	179000	125300	1613	x	3500
57	디스커버리 익스 페디션	픽토르 U 구스다 운 경량패딩 (IVORY)	159000	0	221	x	1200

58	게스	남녀공용 베이직 푸퍼 다운 _MM4W9891	108000	97200	387	x	1800
60	후아유	바시티 자켓 WHJJC4V24U	129000	90300	1656	x	4300
62	코드그래피	헤비 보아 플리스 무튼 자켓_베이지	109000	94900	513	x	2400
65	라퍼지스토어	덕다운 아르틱 후 드 패딩_Black	173000	128000	54	o	1100
67	파르티멘토	[비건 레더] 필드 자켓 블랙	111700	72600	12414	x	19000
68	무신사 스탠다드	오버사이즈 디테 처블 후드 MA-1 보머 재킷 [블랙]	129900	0	866	x	3100
73	커버넛	리버시블 웨르파 후디 집업 자켓 아 이보리	149000	104300	7538	x	4000
79	후아유	리버시블 덤블 숏 푸퍼 [5color] WHJPC4V02U	129000	89900	306	x	1400
80	라퍼지스토어	[리뉴얼 ver.][3M 신슬레이트] M- 1965 피쉬테일 숏 파카_Original Khaki	129000	79900	1620	o	5700
83	파르티멘토	3M 신슬레이트 M-65 피시테일 코 트 패딩 카키	155000	93000	10571	x	5000
87	내셔널지오그래픽	N224WDW920 여성 해론 U넥 경 량 덕 다운 점퍼 BEIGE	149000	0	184	x	1100
88	라퍼지 포 우먼	[Woman]올리치 트렌치 맥코트 _Black	163000	99900	5973	x	11000
89	파르티멘토	[22FW 리뉴얼] 오버사이즈 멜톤 콘보이 더플 코트 네이비	111700	79900	940	o	4400
90	프리즘웍스	VINCENT M65 FISHTAIL PARKA 003 _ OLIVE	198000	148500	2862	x	3800

```
# 20만원 이상
select * # 28개
from musinsa
where ori_price >= 200000
order by ranks;
```

1	드로우핏	프리미어 퓨어 캐 시미어 히든 싱글 코트 [BLACK]	278000	208500	4805	x	8600
3	노스페이스	NJ1DN75A 남성 1996 예코 낚시 자켓	339000	0	277	x	2200
5	쿠어	MTR 발마칸 코트 (블랙)	329000	296100	2732	o	4700
18	알파 인더스트리	MA-1 레귤러 Black	279000	0	9660	x	8000

22	비슬로우	22FW OVERSIZED CASHMERE BALMACAN COAT BLACK	299000	209300	411	x	1800
24	노이어	캐시미어 더블 브 레스티드 코트 (블 랙)	425000	297000	55	o	1000
26	쿠어	MTR 발마칸 코트 (멜란지토프)	329000	296100	2235	o	3100
32	무신사 스탠다드	우먼즈 캐시미어 블렌드 발마칸 로 브 코트 [블랙]	209900	0	370	x	1600
36	드로우핏	프리미어 퓨어 캐 시미어 발마칸 코 트 [BLACK]	308000	231000	1113	x	2800
40	노스페이스	NJ1DN50J 화이 트라벨 노벨티 늑 시 다운 자켓	348000	0	529	x	2000
45	쿠어	MTR 오버핏 더블 롱 코트 (블랙)	334000	300600	6580	o	2900
46	쿠어	퍼프 플라이트 MA-1 (블랙)	235000	211500	507	x	1300
52	메종미네드	BLACK DOUBLE LONG GOOSE DOWN	389000	289000	1845	x	2700
55	드로우핏	프리미어 퓨어 캐 시미어 로브 더블 코트 [BLACK]	308000	231000	1527	x	2100
59	케이투	씬에어 (THINAIR)ECO LITE_Black	350000	252000	1260	o	1500
61	도프제이슨	오버핏 램스킨 싱 글 자켓	339000	198000	7321	x	8800
66	노스페이스	NC3NN50J 화이 트라벨 알칸 EX 온볼 코트	258000	232200	88	x	700
69	엠엘비	베이직 스몰로고 숏패딩 NY (Black)	289000	0	58	x	500
70	키뮤어	서플러스 울 (WOOL) 65 세미 오버 발마칸 코트 _블랙	209000	169000	651	x	1300
72	프리즘웍스	VINCENT M1965 FISHTAIL PARKA 004 _ OLIVE	219000	186150	171	x	800
74	캠빈클라인 진	여 아이보리 숏 령 스 스탠드 카라 푸 퍼 다운 재킷 J221163 ACF	279000	251100	97	x	600
76	키뮤어	서플러스 울 (WOOL) 65 세미 오버 더플 코트_ 네이비	209000	169000	217	x	900
77	인사일런스	캐시미어 더블 롱 코트 BLACK	299000	164450	10769	x	6000
78	디스이즈네버멧	PERTEX® T	219000	153300	66	x	900



		Down Jacket Black					
81	캠빈클라인 진	여 블랙 숏 렁스 스탠드 카라 푸퍼 다운 재킷 J221163 BEH	279000	251100	71	x	500
84	노스페이스	NJ1DN75B 남성 1996 에코 늑시 자켓	339000	0	85	x	500
85	노스페이스	NC1DN00A 고 프리 다운 코트	329000	312550	127	x	600
86	비슬로우	22FW OVERSIZED M65 HOOD SHORT DOWN PARKA BLACK	269000	188300	965	x	2000

```
# 30만원 이상 높은 가격순 top10
select *
  from musinsa
 where ori_price >= 300000
 order by ori_price desc
 limit 10;
```

24	노이어	캐시미어 더블 브 레스티드 코트 (블 랙)	425000	297000	55	o	1000
52	메종미네드	BLACK DOUBLE LONG GOOSE DOWN	389000	289000	1845	x	2700
59	케이투	씬에어 (THINAIR)ECO LITE_Black	350000	252000	1260	o	1500
40	노스페이스	NJ1DN50J 화이 트라벨 노벨티 늑 시 다운 자켓	348000	0	529	x	2000
3	노스페이스	NJ1DN75A 남성 1996 에코 늑시 자켓	339000	0	277	x	2200
84	노스페이스	NJ1DN75B 남성 1996 에코 늑시 자켓	339000	0	85	x	500
61	도프제이슨	오버핏 램스킨 싱 글 자켓	339000	198000	7321	x	8800
45	쿠어	MTR 오버핏 더블 롱 코트 (블랙)	334000	300600	6580	o	2900
5	쿠어	MTR 발마칸 코트 (블랙)	329000	296100	2732	o	4700
26	쿠어	MTR 발마칸 코트 (멜란지토프)	329000	296100	2235	o	3100

```
# 누적 판매량 top 10
select *
  from musinsa
 order by sale_amt desc
 limit 10;
```

6	토피	2WAY 스웨트 후드 집업 (MELANGE GREY)	45000	37000	52900	x	149000
37	무신사 스탠다드	릴렉스드 베이식	81900	73690	27521	x	44000

		블레이저 [블랙]					
2	스파오	베이직 푸퍼 _SPJPC4TC11	69900	62910	11741	x	41000
7	라퍼지스토어	[리뉴얼 ver.][3M 신솔레이트]오리 지널 M-1965 피 쉬테일 파카 _Original Khaki	138000	89900	17402	o	37000
9	스파오	베이직 퍼플리스 집업 _SPFZC4TU01	29900	26910	8273	x	34000
49	라퍼지스토어	[매트블랙] 버핑레 더 미니멀 포켓 블 루종_Matt Black	94000	65900	12626	o	30000
75	스파오	리버서블 푸퍼 _SPJDC4VC14	89900	80910	7921	x	24000
20	도프제이슨	오버핏 비건레더 싱글 자켓	129000	82550	12423	x	23000
10	라퍼지스토어	버핑레더 오버핏 블루종_Black	94000	65900	6060	o	22000
11	스파오	파스텔 푸퍼 _SPJPC4TG01	69900	62910	5695	x	21000

```
# 5만원 ~ 20만원 누적판매량 top10
select *
from musinsa
where ori_price between 50000 and 200000
order by sale_amt desc
limit 10;
```

37	무신사 스탠다드	릴렉스드 베이식 블레이저 [블랙]	81900	73690	27521	x	44000
2	스파오	베이직 푸퍼 _SPJPC4TC11	69900	62910	11741	x	41000
7	라퍼지스토어	[리뉴얼 ver.][3M 신솔레이트]오리 지널 M-1965 피 쉬테일 파카 _Original Khaki	138000	89900	17402	o	37000
49	라퍼지스토어	[매트블랙] 버핑레 더 미니멀 포켓 블 루종_Matt Black	94000	65900	12626	o	30000
75	스파오	리버서블 푸퍼 _SPJDC4VC14	89900	80910	7921	x	24000
20	도프제이슨	오버핏 비건레더 싱글 자켓	129000	82550	12423	x	23000
10	라퍼지스토어	버핑레더 오버핏 블루종_Black	94000	65900	6060	o	22000
11	스파오	파스텔 푸퍼 _SPJPC4TG01	69900	62910	5695	x	21000
71	무신사 스탠다드	베이식 블레이저 [블랙]	79900	67890	46214	x	21000
82	무신사 스탠다드	오버사이즈 블레 이저 [블랙]	83900	71290	24392	x	20000