

Duty Free Online Shopping

- Lotte Duty Free
- Shilla Duty Free
- Shinsegae Duty Free

Contents

- Make a plan / 계획 수립하기
- Web scraping / 웹 스크래핑하기
- The questions / 질문들
- Create tables / 테이블 생성하기

MAKE A PLAN

MAKE A PLAN



MAKE A PLAN



MAKE A PLAN

LOTTE
DUTY
FREE

메뉴 ▾

홈

스토리

세일

혜택



화제일 베스트

멕

[사중기한 24년 1월] 러스터글
래스 립스틱

₩26
\$7.8 70%
9,852원



화제일 베스트

빠아

래스트 벨벳 립 틴트 21~25

₩7
\$4.9 30%
6,189원

화제일 베스트

라소

라소

₩4
\$4
6,111원



화제일 베스트

조 말론 런던

비타민 E 립 컨디셔너 -

로그인 후 할인을 확인
\$35
44,208원

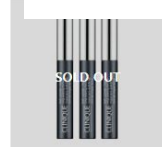


베스트

메이크업 포에버

울트라 HD 프레스트 파우더

로그인 후 할인을 확인
\$40
50,524원



화제일 베스트

크리니크

래쉬 파워 마스크라 프리오
(BLACK ONLY)

로그인 후 할인을 확인
\$67
84,627원



사운들 베스트

헤라

헤라 볼락 :
광택속립 SPF34/PA+

로그인 후 할
\$45
56,839원

SHINSEGAE 세일 혜택 랭킹 SSGDF LIVE 뷰티 패션 공식스토어 Q

트렌드스토리 지원안내 3시간전 | #리디크 #리디크 #리디크 #리디크 #리디크

홈 | 뷰티 | 메이크업

메이크업

전체 | 립 메이크업 | 메이크업 도구 | 메이크업 세트 | 아이 메이크업 | 페이스 메이크업

필터

브랜드

브랜드를 입력하세요

변여 | ABC | 카니온

THE SHILLA
duty free

카테고리

브랜드

혜택

세일

랭킹

3시간전

신라부티크

비밀매인 신규가입 기념 30%사

Q

160여개보기

필터검색

총 3,450건

구매받은 순 · 신상품순 · 낮은 가격순 · 높은 가격순 · 할인율순 · 리뷰많은 순 · 40개 보기

상품비교

카테고리

브랜드

브랜드명을 검색해주세요

카니온 | 일러스트

가

가

가

가

가

가

가

가

가

가

가

가

가

가

가



메이크업 포에버
울트라 HD 프레스트 파우더

로그인 후 할인을 확인
\$40 50,524원
★ 4.8 | 790건
화제일 | 3시간전



바이오더마
아도담 립스틱 1+1

30% 6H
\$7.7 9,725원
★ 5.0 | 1건
화제일 | 3시간전



나스
NATURAL RADIANT LONGWEAR
CUSHION FOUNDATION EMPT...

로그인 후 할인을 확인
\$17 21,472원
★ 4.4 | 5건
3시간전



헤라
UV프로텍터 톤업 SPF50+PA+++ 5
0ml

로그인 후 할인을 확인
\$27 34,105원
★ 5.0 | 1건
시상품



MAKE A PLAN



카테고리 ^	브랜드 v	혜택	세일	랭킹	3시간전
스킨케어		베이스			아이브로우
메이크업 >		아이 >			마스카라
향수/바디/헤어		립 & 치크			아이섀도우
가방/신발/잡화		네일			아이라이너
시계/주얼리		메이크업 도구			
의류/아이웨어		세트/팔레트			
스포츠/레저					
디지털/리빙					
식품					
유아동					
남성관					

WEB SCRAPING

WEB SCRAPING

```
import selenium
import requests
import time
from bs4 import BeautifulSoup
from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
```

웹 스크래핑을 위한 IMPORT 코드

```
import pymysql
```

```
db = pymysql.connect(host='localhost', port=3306, user='root', passwd='1234', db='project1', charset='utf8')
ecommerce = db.cursor()
```

```
sql = """
    create table shilla (
        ranks int,
        dept_name varchar(20) not null,
        brand_name varchar(200) not null,
        item_name varchar(200) not null,
        dollar_price varchar(20),
        won_price varchar(20),
        before_dist varchar(20),
        dist_percentage varchar(20),
        primary key(ranks,dept_name)
    );
"""
```

```
ecommerce.execute(sql)
```

테이블 생성하는 코드

WEB SCRAPING

shilla online shopping code

```
options = webdriver.ChromeOptions()
options.add_argument('user-agent=Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/109.0.0.0 Safari/537.36')
```

```
url = 'https://www.shilladfs.com/estore/kr/ko/c/2'
driver = webdriver.Chrome(options=options)
driver.get(url)
```

```
elem = driver.find_element_by_xpath('//*[@id="facetListCount"]')
elem.click()
elems = driver.find_element_by_xpath('//*[@id="facetListCount"]/option[4]')
elems.click()
```

사이트 URL 입력 후
웹 스크래핑 페이지 오픈

```
dept_name = driver.find_element_by_xpath('//*[@id="footer"]/div[2]/div/div[1]/strong').text
totals = driver.find_elements_by_css_selector('div.pro_detail')
ranks = 1
```

```
for total in totals:
```

```
    if total.find_element_by_css_selector('strong.info_brand').text == '':
        continue
```

```
    brand_name = total.find_element_by_css_selector('strong.info_brand').text
    item_name = total.find_element_by_css_selector('p.info_name').text
    dollar_price = total.find_element_by_css_selector('p.p_sum>strong').text
    won_price = total.find_element_by_css_selector('p.p_sum>i').text
```

```
    try:
```

```
        before_dist = total.find_element_by_css_selector('i.strike').text
        dist_percentage = total.find_element_by_css_selector('p.p_rate>strong').text
```

```
    except:
```

```
        before_dist = None
        dist_percentage = None
```

```
    sql = """INSERT INTO shilla VALUES(
```

```
    {0}, "{1}", "{2}", "{3}", "{4}", "{5}", "{6}", "{7}")""".format(ranks, dept_name, brand_name, item_name, dollar_price, won_price, before_dist, dist_percentage)
```

```
    ranks += 1
```

```
    print (sql)
```

```
    ecommerce.execute(sql)
```

1. 만들어 높은 테이블 안에
필요한 정보 삽입 하는 코드

WEB SCRAPING

```
driver = webdriver.Chrome()

driver.get('https://kor.lottedfs.com/kr/display/category/first?dispShopNo1=10031766&reDpth=1');

elem = driver.find_element_by_css_selector('#cntPerPage-button')
elem.click()

elem = driver.find_element_by_css_selector('#cntPerPage-menu > li:nth-child(4) ')
elem.click()

soup = BeautifulSoup(driver.page_source, 'lxml')
prices = soup.select('strong.price02')
brand = soup.select('a > div.unit_info > span.brand > i')
name = soup.select('a > div.unit_info > span.name')
or_pri = soup.select('a > div.unit_price > span.price01')
won = soup.select('a > div.unit_price > span.price03')
dept = soup.select('#lotteFooter > div > div.areal > dl > dd:nth-child(2) > address > span:nth-child(1)')

ranks = 1
for i in range(240):
    brand_name = brand[i].text
    dept_name = name[i].text
    before_dist = or_pri[i].text
    dollar_price = prices[i].text
    won_price = won[i].text
    dept_text = dept[0].text
    # for i in range(240):
    sql = """INSERT INTO lotte_duty_free VALUES("{0}", "{1}", "{2}", "{3}", "{4}", "{5}", "{6}");""".format(dept_text, product_code, brand[i].text, won_price, before_dist, dollar_price, brand_name)
    ranks += 1
    print (sql)
    cursor.execute(sql)
db.commit()
```

사이트 URL 입력 후
웹 스크래핑 페이지 오픈

2. 만들어 높은 테이블 안에
필요한 정보 삽입 하는 코드

QUESTIONS

QUESTIONS

1. 브랜드

1.1 각 면세점 별 100위안에 가장 많이 포함되어 있는 브랜드

1.2 면세점별 1~15위에 모두 포함되어 있는 브랜드

2. 가격

2.1 각 면세점별 100위 안 제품들 평균 가격

2.2 제품이 10만원 이상인 아이템

3. 이름

3.1 제품 이름이 가장 긴 아이템 / 짧은 아이템

CREATE TABLES

CREATE TABLES - 각 면세점 별 100위안에 가장 많이 포함되어 있는 브랜드

```
create or replace view table_1 as (  
select rank() over (PARTITION BY dept_name order by count(*) desc) brand_rank, brand_name, dept_name,  
count(*) as total  
from duty_free  
where ranks <= 100  
group by brand_name, dept_name  
order by count(*) desc  
);  
select * from table_1  
where brand_rank = 1;
```

brand_rank	brand_name	dept_name	total
1	입생로랑	(주)신세계면세점 본점	14
1	맥	(주)호텔롯데 롯데면세점	10
1	나스	㈜호텔신라	9
1	베네프트	㈜호텔신라	9

CREATE TABLES - 면세점별 1~15위에 모두 포함되어 있는 브랜드

```
select brand_name, item_name
from lotte_duty_free
where ranks <= 15
and brand_name in (
    select brand_name
    from shilla
    where ranks <= 15)
and brand_name in (
    select brand_name
    from shinsegae
    where ranks <= 15);
```

	brand_name	item_name
▶	메이크업포에버	울트라HD 프레스드 파우더

CREATE TABLES - 각 면세점별 100위 안 제품들 평균 가격

```
select a.dept_name, round(avg(a.won_price)) '평균값(원)', b.dept_name, round(avg(b.won_price)) '평균값(원)', c.dept_name, round(avg(c.won_price)) '평균값(원)'
from shilla a
inner join shinsegae b
on a.ranks = b.ranks
inner join lotte_duty_free c
on a.ranks = c.ranks
where a.ranks <= 100
and b.ranks <= 100
and c.ranks <= 100
group by 1,3,5;
```

```
SELECT a.dept_name , round(avg(a.won_price)) '평균가격(원)', b.dept_name ,round(avg(b.won_price)) '평균가격(원)',c.dept_name, round(avg(c.won_price)) '평균가격(원)'
FROM shilla a, shinsegae b, lotte_stock c
WHERE a.ranks <=100 AND b.ranks <=100 AND c.ranks <=100 AND (a.ranks = b.ranks AND b.ranks = c.ranks);
```

dept_name	평균값(원)	dept_name	평균값(원)	dept_name	평균값(원)
(주)호텔신라	28257	(주)신세계면세점 본점	34221	(주)호텔롯데 롯데면세점	41249

CREATE TABLES - 제품이 10만원 이상인 아이템

```
update duty_free
set won_price = replace(won_price, ',', '')
where won_price like '%원';

update duty_free
set won_price = replace(won_price, '원', '')
where won_price like '%원';

select brand_name, item_name, cast(won_price as SIGNED INTEGER) won_price
from duty_free
where won_price >= 100000
order by 3 desc;
```

brand_name	item_name	won_price
라 메르	루미너스 리프팅 쿠션 파운데이션 SPF20	170518
에스티로더	Futurist Aqua Brilliance 메이크업 SPF20/PA+++ 듀오	152835
에스티로더	Double Wear 스테이-인-플레이스 메이크업 듀오 COOL VANILA 2C0	132625
에스티로더	Double Wear 스테이-인-플레이스 메이크업 듀오 SAND 1W2	132625
바비브라운	디럭스 사이즈 비타민 인리치드 페이스 베이스	132625
랑콤	랑콤 그랑디오즈 스머지프루프 마스크라 트리오 01 10ml*3 볼륨 검색	121257
라 메르	립 볼류마이저	121257
통 포드 뷰티	[69][유통기한 23년 6월] EYE COLOR QUAD N04 허니문	117468
통 포드 뷰티	EYE COLOR QUAD	117468
통 포드 뷰티	EYE COLOR QUAD	117468
통 포드 뷰티	EYE COLOR QUAD 03 BODY HEAT	117468
에스티 로더	면세전용 더블웨어 랭쓰닝 마스크라 트리오	116205
통 포드 뷰티	EYE COLOR QUAD 04 SUSPICION	116205
에스티로더	Double Wear 면세전용 더블웨어 랭쓰닝 마스크라 트리오	116205
헤라	헤라 UV 미스트쿠션 커버 C21호 트리오 세트 바닐라	114942
헤라	UV 미스트쿠션 커버 C21호 트리오 세트 바닐라	114942
에스케이투	에트모스피어 CC크림 30g	108626
에스케이투	에트모스피어 CC크림 30g	108626

CREATE TABLES - 제품 이름이 가장 긴 아이템 / 짧은 아이템

```
CREATE OR REPLACE VIEW len_test as
(
SELECT ranks,dept_name, brand_name, item_name,char_length(item_name) item_len,length(item_name) byte_len
FROM duty_free
ORDER BY item_len DESC
);
SELECT dept_name, item_name,char_length(item_name) item_len
FROM len_test
WHERE byte_len = (select max(byte_len) from len_test) OR byte_len = (select min(byte_len) from len_test);
```

dept_name	item_name	item_len
썬호텔신라	트래블 익스클루시브: 라이트풀 C+ 코랄 그라...	133
(주)호텔롯데 롯데면세점	립 밤	3

THANK YOU