

One-Day Project. 스타벅스 이디야 데이터 분석 - SQL 사용하기

여러분은 이전 과제에서 스타벅스 이디야 데이터 분석 프로젝트를 진행하였습니다.

이번 시간에는 분석할 데이터를 Database 로 저장하고, 분석에 필요한 데이터를 저장된 Database 에서 쿼리하여 가져와 보도록 하겠습니다.

문제 1. AWS RDS (MySQL) 에 프로젝트 관련 Database 를 생성하고, 접근 가능한 사용자 계정을 생성하세요.

- Database Name : oneday
- User Name / Password : oneday / 1234

제출 1.

- Database 생성문 조회 결과 : *SHOW CREATE DATABASE oneday;*
- 사용자 권한 확인 결과 : *SHOW GRANT FOR 'oneday'@'%'*

문제 2. 스타벅스 이디야 데이터를 저장할 테이블을 다음의 구조로 생성하세요.

Schema

COFFEE_BRAND		
Name	Type	Description
id	int	PRIMARY KEY 로 지정, auto_increment 설정
name	String	커피브랜드 이름

COFFEE_STORE		
Name	Type	Description
id	int	PRIMARY KEY 로 지정, auto_increment 설정
brand	int	FOREIGN_KEY (COFFEE_BRAND 테이블의 id) 로 지정
name	String	커피매장 이름 (최대 32자)
gu_name	String	구 이름 (최대 5자)
address	String	매장 전체 주소 (최대 128자)
lat	decimal(16,14)	위도정보
lng	decimal(17,14)	경도정보

문제 3. Python 코드로 COFFEE_BRAND 데이터를 다음과 같이 입력하고 확인하세요.

Sample Data

COFFEE_BRAND	
id	Name
1	STARBUCKS
2	EDIYA

제출 2.

- Table 생성 결과 : *Desc COFFEE_BRAND; Desc COFFEE_STORE;*

제출 3.

- COFFEE_BRAND 조회 결과 : *SELECT * FROM COFFEE_BRAND;*

문제 4. 스타벅스 페이지에 접근하는 코드에서 팝업창이 없는 경우, 팝업창을 닫는 코드에서 에러가 발생합니다. 예외처리 해서 에러 메시지를 출력하고 실행이 중단되지 않도록 수정해주세요.

```
# 팝업창 닫기
driver.find_element_by_css_selector('.holiday_notice_close a').click()

⊗ 0.3s Python

-----
NoSuchElementException Traceback (most recent call last)
<ipython-input-3-8ce710d70159> in <module>
      1 # 팝업창 닫기
----> 2 driver.find_element_by_css_selector('.holiday_notice_close a').click()

~/miniconda3/envs/ds_study/lib/python3.8/site-packages/selenium/webdriver/remote/webdriver.py in find_element_by_css_selector(self, css_selector)
    596     element = driver.find_element_by_css_selector('#foo')
    597     """
--> 598     return self.find_element(by=By.CSS_SELECTOR, value=css_selector)
    599
    600     def find_elements_by_css_selector(self, css_selector):

~/miniconda3/envs/ds_study/lib/python3.8/site-packages/selenium/webdriver/remote/webdriver.py in find_element(self, by, value)
    974     by = By.CSS_SELECTOR
    975     value = '[name="%s"]' % value
--> 976     return self.execute(Command.FIND_ELEMENT, {
    977         'using': by,
    978         'value': value})['value']

~/miniconda3/envs/ds_study/lib/python3.8/site-packages/selenium/webdriver/remote/webdriver.py in execute(self, driver_command, params)
    319     response = self.command_executor.execute(driver_command, params)
    320     if response:
--> 321         self.error_handler.check_response(response)
    322         response['value'] = self._unwrap_value(
    323             response.get('value', None))

~/miniconda3/envs/ds_study/lib/python3.8/site-packages/selenium/webdriver/remote/errorhandler.py in check_response(self, response)
    240     alert_text = value['alert'].get('text')
    241     raise exception_class(message, screen, stacktrace, alert_text)
--> 242     raise exception_class(message, screen, stacktrace)
    243
    244     def _value_or_default(self, obj, key, default):

NoSuchElementException: Message: no such element: Unable to locate element: {"method":"css selector","selector":".holiday_notice_close a"}
(Session info: chrome=98.0.4758.109)
```

문제 5. Python 코드로 스타벅스 페이지에서 데이터를 가져올때, COFFEE_STORE 테이블에 바로 입력하도록 수정하세요.

- 데이터 세트: 매장 이름, 매장이 위치한 구 이름, 매장 주소, 위도, 경도
- 필요한 데이터를 한세트씩 가져와서 COFFEE_STORE 테이블에 각각INSERT 하도록 합니다.
- 입력된 데이터의 총 갯수를 쿼리하여 결과를 확인합니다.
- 입력된 데이터 상위 10개를 쿼리하여 결과를 확인합니다.

제출 4.

- 팝업 예외처리 코드 & 실행 결과 (ipynb)

제출 5.

- 스타벅스 데이터 관련 코드 & 실행 결과 (ipynb)

문제 6. Python 코드로 이디야 페이지에서 데이터를 가져올때, COFFEE_STORE 테이블에 바로 입력하도록 수정하세요.

- 데이터 세트 : 매장 이름, 매장이 위치한 구 이름, 매장 주소, 위도, 경도
- 이디야 페이지에서 검색에 사용할 구 이름은 COFFEE_STORE 에서 중복을 제거하는 쿼리를 사용하여 가져와서 {‘서울 ‘ + 구이름} 형식으로 변환하여 사용하도록 합니다.
- 필요한 데이터를 한 세트씩 가져와서 COFFEE_STORE 테이블에 각각 INSERT 하도록 합니다.
- (주의) COFFEE_STORE 테이블에 입력할 구 이름은 {‘서울 ‘} 이 제거된 구 이름입니다.
- 입력된 데이터의 총 갯수를 쿼리하여 결과를 확인합니다.
- 입력된 데이터 상위 10개를 쿼리하여 결과를 확인합니다.

제출 6.

- 이디야 데이터 관련 코드 & 실행 결과 (ipynb)

문제 7. Python 코드에서 다음의 데이터를 쿼리를 사용하여 조회하세요.

- 스타벅스 매장 주요 분포 지역 (매장수가 많은 상위 5개 구이름, 매장 개수 출력)
- 이디야 매장 주요 분포 지역 (매장수가 많은 상위 5개 구이름, 매장 개수 출력)
- 구별 브랜드 각각의 매장 개수 조회 (구이름, 브랜드이름, 매장 개수 출력)
- 구별 브랜드 각각의 매장 개수 조회 (구이름, 스타벅스 매장 개수, 이디야 매장 개수 출력)

제출 7.

- 관련 코드 & 실행 결과 (ipynb)

- 문제 8. 시각화 프로젝트를 위하여 다음의 규칙으로 쿼리하여 CSV 파일로 저장합니다. (Python 코드로 작업)
- 전체 데이터를 가져오는데, 각 스타벅스 매장별로 이디야 전체 매장정보가 매칭되어 있어야 합니다. (정렬 : s_id, e_id 순)
 - 다음의 형식으로 저장되어야 합니다. (브랜드 이름, 칼럼 명 주의)

s_id	s_brand	s_name	s_gu	s_address	s_lat	s_lng	e_id	e_brand	e_name	e_gu	e_address	e_lat	e_lng
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	567	EDIYA	강남YMCA점	강남구	서울 강남구 논현동	37.51353040000000	127.03153410000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	568	EDIYA	강남구청역아이티웨딩점	강남구	서울 강남구 학동로 338 (논현동, 강남파라곤)	37.51657850000000	127.04034180000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	569	EDIYA	강남논현학동점	강남구	서울 강남구 논현로131길 28 (논현동)	37.51422720000000	127.02866620000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	570	EDIYA	강남대치점	강남구	서울 강남구 역삼로 415 (대치동, 성진빌딩)	37.50429570000000	127.06305270000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	571	EDIYA	강남도산점	강남구	서울 강남구 도산대로37길 20 (신사동)	37.52224130000000	127.03148150000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	572	EDIYA	강남세곡점	강남구	서울 강남구 현릉로569길 9 (세곡동, 강남지웰파인즈)	37.46697310000000	127.10052620000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	573	EDIYA	강남율현점	강남구	서울 강남구 방고개로21길 8 (율현동, 세곡프라자)	37.47350070000000	127.11438010000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	574	EDIYA	강남자곡점	강남구	서울 강남구 자곡로 180 (자곡동, 강남유타입블레스)	37.47348460000000	127.10309790000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	575	EDIYA	개포동역점	강남구	서울 강남구 개포로82길 11 (개포동, 삼우빌딩)	37.48852140000000	127.06743250000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	576	EDIYA	건설회관점	강남구	서울 강남구 언주로 711 (논현동)	37.51630030000000	127.02991230000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	577	EDIYA	대청역점	강남구	서울 강남구 개포로109길 34 (개포동, 대청프라자)	37.49412880000000	127.07768410000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	578	EDIYA	대치역점	강남구	서울 강남구 남부순환로 2937 (대치동, 까치빌딩)	37.49410070000000	127.06151900000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	579	EDIYA	대치포스코점	강남구	서울 강남구 테헤란로78길 16 (대치동, 노벨빌딩)	37.50506260000000	127.05338270000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	580	EDIYA	도곡2동점	강남구	서울 강남구 논현로 142 (도곡동, 동현빌딩)	37.48243510000000	127.04299330000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	581	EDIYA	도산사거리점	강남구	서울 강남구 논현동	37.51353040000000	127.03153410000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	582	EDIYA	삼성동점	강남구	서울 강남구 테헤란로88길 10 (대치동)	37.50680620000000	127.05944950000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	583	EDIYA	삼성중앙점	강남구	서울 강남구 봉은사로 442 (삼성동)	37.51127280000000	127.04821230000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	584	EDIYA	삼성한전점	강남구	서울 강남구 테헤란로107길 11 (삼성동)	37.51039080000000	127.06504420000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	585	EDIYA	선릉사거리점	강남구	서울 강남구 선릉로 519 (역삼동, 삼능빌딩)	37.50590500000000	127.04793700000000
1	STARBUCKS	역삼아레나빌딩	강남구	서울특별시 강남구 언주로 425 (역삼동)	37.50108700000000	127.04306900000000	586	EDIYA	선릉역점	강남구	서울 강남구 테헤란로 332 (역삼동, H타워)	37.50383620000000	127.04817030000000

제출 8.

- 시각화 프로젝트 관련 코드 (ipynb), 결과 파일 (csv)