

# 데이터 집계하기 : 집계쿼리

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

- 그룹화: 특정 항목을 기준으로 데이터를 집계하는 것
- 집계 값은 집계 함수로 계산
- 집계 쿼리: 그룹화와 집계 함수를 결합해 작성한 쿼리

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (1) 데이터 그룹화하기

- GROUP BY 절

SELECT \*

FROM 테이블

WHERE 조회 조건

GROUP BY 칼럼[표현식, 순번]1, 칼럼[표현식, 순번]2, ...

ORDER BY ...

LIMIT n;

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (1) 데이터 그룹화하기

- GROUP BY 절
  - GROUP BY 절은 WHERE와 ORDER BY 절 사이에 위치
  - 그룹화할 칼럼(표현식)은 GROUP BY 절 다음에 넣음
  - GROUP BY 절에는 칼럼, 표현식, 순번 명시 가능  
(순번 → SELECT 절에 명시한 칼럼(표현식)의 순서)
  - SELECT 절에는 GROUP BY 절에 기술한 칼럼(표현식)만 명시해야 올바른 데이터 조회 가능
  - SELECT 절의 칼럼(표현식)에 별칭 사용 시, GROUP BY 절에도 해당 별칭 사용 가능

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (1) 데이터 그룹화하기

- GROUP BY 절

```
SELECT continent  
FROM country  
GROUP BY continent;
```

	name	continent ▲
	Holy See (Vatican Ci...	Europe
	Yugoslavia	Europe
	Aruba	North America
	Anguilla	North America
	Netherlands Antilles	North America
	Antigua and Barbuda	North America
	Bahamas	North America
	Belize	North America
	Bermuda	North America
	Barbados	North America
	Canada	North America
	Costa Rica	North America

	continent
▶	North America
	Asia
	Africa
	Europe
	South America
	Oceania
	Antarctica

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (1) 데이터 그룹화하기

- GROUP BY 절

```
SELECT continent, region  
FROM country  
GROUP BY continent, region  
ORDER BY 1, 2;
```



```
SELECT continent, region  
FROM country  
GROUP BY 1, 2  
ORDER BY 1, 2;
```

	continent	region
▶	Asia	Eastern Asia
	Asia	Middle East
	Asia	Southeast Asia
	Asia	Southern and Central Asia
	Europe	Baltic Countries
	Europe	British Islands
	Europe	Eastern Europe
	Europe	Western Europe
	Oceania	Melanesia
	Oceania	Micronesia
	Oceania	Micronesia/Caribbean
	Oceania	Polynesia
	Antarctica	Antarctica
	South America	South America

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (1) 데이터 그룹화하기

- GROUP BY 절

```
SELECT SUBSTRING(district, 1, 2) DO  
FROM city  
WHERE countrycode = 'KOR'  
GROUP BY SUBSTRING(district, 1, 2)  
ORDER BY 1;
```

	DO
▶	Ch
	In
	Ka
	Kw
	Ky
	Pu
	Se
	Ta

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (1) 데이터 그룹화하기

- GROUP BY 절
  - 그룹화 대상 칼럼과 조회하는 칼럼이 일치해야 올바른 집계 쿼리

```
SELECT continent  
  FROM country  
GROUP BY region  
ORDER BY 1;
```

	continent
▶	Asia
	Asia
	Asia
	Asia
	Europe
	Europe
	Europe
	Asia
	Oceania
	Oceania
	Antarctica
	South America



# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (1) 데이터 그룹화하기

- DISTINCT
  - SELECT 절에 DISTINCT 키워드를 넣고 그 다음에 칼럼(표현식)을 명시
  - GROUP BY 절 없이도 중복 제거 → 고유한 값 추출
  - DISTINCT는 그룹화는 아니지만, GROUP BY 절을 사용한 것과 같은 효과

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (1) 데이터 그룹화하기

- DISTINCT

```
SELECT DISTINCT continent  
FROM country;
```

	continent
▶	North America
	Asia
	Africa
	Europe
	South America
	Oceania
	Antarctica

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (2) 집계 함수 사용하기

- 테이블의 전체 또는 특정 칼럼 값에 대한 로우 수, 합계, 최댓값, 최솟값, 평균 등의 집계 값을 계산하는 함수
- COUNT([DISTINCT] expr)
  - 테이블 전체 로우 수, ()안에 \* 명시 가능
- MAX([DISTINCT] expr): 최댓값
- MIN([DISTINCT] expr): 최솟값

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (2) 집계 함수 사용하기

- AVG([DISTINCT] expr): 평균
- SUM([DISTINCT] expr): 합계
- VAR\_POP(expr): 분산
- STDDEV\_POP(expr): 표준편차

# 1. 집계 쿼리 다루기: 데이터 그룹화와 집계 함수

## (3) GROUP BY 절과 집계 함수로 집계 쿼리 완성하기

- 연도별 개봉 영화 편수 집계하기
- 2019년 개봉 영화의 유형별 최대, 최소 매출액과 전체 매출액 집계하기
- 2019년 개봉 영화 중 매출액이 1억 원 이상인 영화의 분기별, 배급사별 개봉 영화 수와 매출액 집계하기

## 2. 총계 산출과 HAVING 절

### (1) WITH ROLLUP으로 소계와 총계 구하기

- 항목별 소계와 총계를 구하려면 GROUP BY 절 맨 끝에 WITH ROLLUP 키워드 추가

```
SELECT movie_type 영화유형, SUM(sale_amt) 금액
FROM box_office
WHERE YEAR(release_date) = 2019
AND sale_amt > 100000000
GROUP BY movie_type
ORDER BY 1 DESC;
```



```
SELECT movie_type 영화유형, SUM(sale_amt) 금액
FROM box_office
WHERE YEAR(release_date) = 2019
AND sale_amt > 100000000
GROUP BY movie_type WITH ROLLUP;
```

	영화유형	금액
▶	단편	29142700
	올니버스	75428940
	장편	1870000814478
	HULL	1870105386118

→ GROUP BY movie\_type

→ WITH ROLLUP

## 2. 총계 산출과 HAVING 절

### (2) GROUPING( ) 함수

- GROUP BY 절과 WITH ROLLUP을 사용하는 집계 쿼리에 추가 기능 제공
- 반환값: 1 → 총계, 나머지 → 0

```
SELECT movie_type 영화유형, SUM(sale_amt) 금액, GROUPING(movie_type)
FROM box_office
WHERE YEAR(release_date) = 2019
GROUP BY movie_type WITH ROLLUP;
```

	영화유형	금액	GROUPING(movie_type)
▶	NULL	4517000	0
	단편	29395700	0
	올니버스	92895690	0
	장편	1870657354798	0
	NULL	1870784163188	1

## 2. 집계 산출과 HAVING 절

### (3) HAVING 절

- 집계 함수 결과값에 대한 조건 처리 → WHERE 절처럼 조회되는 로우를 걸러내는 필터 역할
- 집계 쿼리에서만 사용
- HAVING 다음에 집계 함수나 GROUPING() 함수만 사용 가능