# 8-1. 종합실습 1

홍형경 chariehong@gmail.com 2021.06

#### 1. 종합 실습 – 로또 데이터 분석

#### lotto\_master 테이블

컬럼	데이터형	Null	PK	설명
SEQ_NO	NUMBER	N	Υ	로또회차
DRAW_DATE	DATE	Υ		추첨일
NUM1	NUMBER	Υ		당첨번호1
NUM2	NUMBER	Υ		당첨번호2
NUM3	NUMBER	Υ		당첨번호3
NUM4	NUMBER	Υ		당첨번호4
NUM5	NUMBER	Υ		당첨번호5
NUM6	NUMBER	Υ		당첨번호6
BONUS	NUMBER	Υ		보너스번호

#### lotto\_detail 테이블

컬럼	데이터형	Null	PK	설명
SEQ_NO	NUMBER	N	Υ	로또회차
RANK_NO	DATE	N	Υ	등수
WIN_PERSO N_NO	NUMBER	Υ		당첨자수
WIN_MONEY	NUMBER	Υ		1인당 당첨금 액

#### 1. 종합 실습 – 로또 데이터 분석 (955회까지)

· lotto\_master 테이블 생성

```
CREATE TABLE lotto_master (
         NUMBER NOT NULL, -- 로또회차
seq_no
                    -- 추첨일
draw_date DATE,
                    -- 당첨번호1
 num1
         NUMBER,
                    -- 당첨번호2
         NUMBER,
num2
         NUMBER,
                    -- 당첨번호3
num3
                     -- 당첨번호4
         NUMBER,
 num4
                    -- 당첨번호5
         NUMBER,
 num5
                     -- 당첨번호6
         NUMBER,
num6
                     -- 보너스번호
         NUMBER
 bonus
```

**ALTER TABLE lotto\_master** ADD CONSTRAINTS lotto\_master\_pk PRIMARY KEY (seq\_no);

#### 1. 종합 실습 – 로또 데이터 분석 (955회까지)

· lotto\_detail 테이블 생성

```
CREATE TABLE lotto_detail (
seq_no NUMBER NOT NULL, -- 로또회차
rank_no NUMBER NOT NULL, -- 등수
win_person_no NUMBER, -- 당첨자수
win_money NUMBER -- 1인당 당첨금액
);
```

ALTER TABLE lotto\_detail

ADD CONSTRAINTS lotto\_detail\_pk PRIMARY KEY (seq\_no, rank\_no);

# 1. 종합 실습 – 로또 데이터 분석 (955회까지)

· lotto\_master, lotto\_detail 테이블 데이터 생성

02.lotto\_data\_insert.sql 파일 실행

#### 1. 종합 실습 – 로또 데이터 분석

**SELECT** \* FROM lotto\_master ORDER BY 1;

**SELECT** \* FROM lotto\_detail ORDER BY 1, 2;

4	) SEQ_NO  ⊕ DRAW_DATE		⊕ NUM1	⊕ NUM2	⊕ NUM3	⊕ NUM4	⊕ NUM5	⊕ NUM6	∯ BONUS
1	1 2002-12-07	00:00:00	10	23	29	33	37	40	16
2	2 2002-12-14	00:00:00	9	13	21	25	32	42	2
3	3 2002-12-21	00:00:00	11	16	19	21	27	31	30
4	4 2002-12-28	00:00:00	14	27	30	31	40	42	2
5	5 2003-01-04	00:00:00	16	24	29	40	41	42	3
6	6 2003-01-11	00:00:00	14	15	26	27	40	42	34
7	7 2003-01-18	00:00:00	2	9	16	25	26	40	42
8	8 2003-01-25	00:00:00	8	19	25	34	37	39	9
9	9 2003-02-01	00:00:00	2	4	16	17	36	39	14

	∯ SEQ_NO	RANK_NO	⊕ WIN_PERSON_NO	⊕ WIN_MONEY
1	1	1	0	0
2	1	2	1	143934100
3	1	3	28	5140500
4	1	4	2537	50000
5	1	5	40155	5000
6	2	1	1	2002006800
7	2	2	2	94866800
8	2	3	103	1842000
9	2	4	3763	50000

#### (1) 중복 번호 조회

- num1 컬럼에 대해 1회 ~ 916회까지 중복 번호는?

- 집계쿼리 사용

SELECT num1, COUNT(\*)
FROM lotto\_master
GROUP BY num1
ORDER BY 1;

	⊕ NUM1	⊕ COUNT(+)
1	1	128
2	2	101
3	3	88
4	4	83
5	5	80
6	6	69
- 7	7	56
8	8	45
9	9	31
10	10	36
11	11	33
12	12	31
13	13	22
14	14	26
15	15	11
16	16	21
17	17	15
18	18	7
19	19	12
20	20	4
21	21	5
22	22	3
23	23	2
24	24	2
25	25	1
26	26	2

#### (1) 중복 번호 조회

- num1 ~ num6 컬럼에 대해 1회 ~ 916회까지 중복 번호는?

SELECT num1 ,num2 ,num3 ,num4 ,num5 ,num6 , COUNT(\*)

FROM lotto master

GROUP BY num1, num2, num3, num4, num5, num6

ORDER BY 1, 2, 3, 4, 5, 6;

	\$ <b>∀</b>	*		∯ NUM4		∯ NUM6	COUNT(⋆)
1	1	2	3	9	12	23	1
2	1	2	3	14	27	42	1
3	1	2	3	15	20	25	1
4	1	2	4	8	19	38	1
5	1	2	4	23	31	34	1
6	1	2	5	11	18	36	1
7	1	2	6	9	25	28	1
8	1	2	6	16	19	42	1
9	1	2	6	16	20	33	1
10	1	2	7	9	10	38	1
11	1	2	8	17	26	37	1
12	1	2	8	18	29	38	1
13	1	2	9	17	19	42	1
14	1	2	10	13	18	19	1
15	1	2	10	25	26	44	1
16	1	2	15	19	24	36	1
17	1	2	15	28	34	45	1
18	1	2	16	22	38	39	1
19	1	2	23	25	38	40	1
20	1	3	4	6	14	41	1
21	1	3	7	8	24	42	1
22	1	3	8	12	42	43	1
23	1	3	8	21	22	31	1
24	1	3	10	27	29	37	1
25	1	3	11	24	30	32	1
26	1	3	12	14	16	43	1

#### (1) 중복 번호 조회

- num1 ~ num6 컬럼에 대해 1회 ~ 916회까지 중복 번호는?

SELECT num1 ,num2 ,num3 ,num4 ,num5 ,num6 , COUNT(\*)

FROM lotto\_master

GROUP BY num1, num2, num3, num4, num5, num6

**HAVING COUNT(\*) > 1** 

ORDER BY 1, 2, 3, 4, 5, 6;



#### (2) 가장 많이 당첨된 번호 조회

- num1 컬럼에 대해 1회 ~ 916회까지 가장 많이 당첨된 번호는?

SELECT num1 lotto\_num, COUNT(\*) CNT

FROM lotto\_master

**GROUP BY num1** 

ORDER BY 2 DESC;

		⊕ CNT
1	1	128
2	2	101
3	3	88
4	4	83
5	5	80
6	6	69
- 7	7	56
8	8	45
9	10	36
10	11	33
11	12	31
12	9	31
13	14	26
14	13	22
15	16	21
16	17	15
17	19	12
18	15	11
19	18	7
20	21	5
21	20	4
22	22	3 2
23	26	2
24	24	2

#### (2) 가장 많이 당첨된 번호 조회

- num1과 num2 컬럼 통틀어 가장 많이 당첨된 번호는?

SELECT num1 lotto\_num, COUNT(\*) CNT

FROM lotto master

**GROUP BY num1** 

#### **UNION ALL**

SELECT num2 lotto\_num, COUNT(\*) CNT

FROM lotto\_master

**GROUP BY num2** 

ORDER BY 1;

	⊕ LOTTO_NUM	⊕ CNT
1	1	128
2	2	101
3	2	19
4	3	28
5	3	88
6	4	37
7	4	83
8	5	80
9	5	41
10	6	69
11	6	33
12	7	55
13	7	56
14	8	59
15	8	45
16	9	44
17	9	31
18	10	36
19	10	59
20	11	50
21	11	33
22	12	31
23	12	53
24	13	2.2

```
(2) 가장 많이 당첨된 번호 조회
  - num1과 num2 컬럼 통틀어 가장 많이 당첨된 번호는?
SELECT lotto_num, SUM(CNT) AS CNT
 FROM ( SELECT num1 lotto_num, COUNT(*) CNT
        FROM lotto_master
       GROUP BY num1
        UNION ALL
      SELECT num2 lotto_num, COUNT(*) CNT
       FROM lotto_master
      GROUP BY num2
GROUP BY lotto num
ORDER BY 2 DESC;
```

	⊕ LOTTO_NUM	⊕ CNT
1	1	128
2	5	121
3	2	120
4	4	120
5	3	116
6	7	111
-7	8	104
8	6	102
9	10	95
10	12	84
11	11	83
12	9	75
13	14	69
14	13	65
15	15	58
16	17	56
17	16	53
18	18	46
19	19	38
20	20	31
21	21	29
22	22	23
23	23	20
24	25	17

#### (2) 가장 많이 당첨된 번호 조회

- num1 ~ num6 컬럼 통틀어 가장 많이 당첨된 번호는?

```
SELECT lotto num, SUM(CNT) AS CNT
 FROM ( SELECT num1 lotto_num, COUNT(*) CNT
     FROM lotto master
    GROUP BY num1
    UNION ALL
    SELECT num2 lotto num, COUNT(*) CNT
     FROM lotto master
    GROUP BY num2
    UNION ALL
    SELECT num3 lotto_num, COUNT(*) CNT
     FROM lotto master
    GROUP BY num3
```

```
UNION ALL
SELECT num4 lotto_num, COUNT(*) CNT
     FROM lotto master
    GROUP BY num4
    UNION ALL
    SELECT num5 lotto_num, COUNT(*) CNT
     FROM lotto master
    GROUP BY num5
    UNION ALL
    SELECT num6 lotto_num, COUNT(*) CNT
     FROM lotto_master
    GROUP BY num6
GROUP BY lotto_num
ORDER BY 2 DESC;
```

#### (2) 가장 많이 당첨된 번호 조회

- num1 ~ num6 컬럼 통틀어 가장 많이 당첨된 번호는?

	LOTTO_NUM	⊕ CNT
1	34	143
2	27	135
3	43	133
4	18	132
5	17	132
6	40	132
- 7	39	131
8	14	130
9	45	130
10	12	130
11	19	128
12	1	128
13	13	128
14	20	128
15	5	127
16	33	127
17	37	127
18	10	125
19	26	124
20	8	124
21	11	124
22	21	124
23	4	124
24	31	123
25	38	122
	4.5	400

- (3) 가장 많은 당첨금이 나온 회차와 번호, 금액 조회
  - 가장 많은 당첨금? → 1등
  - 1등 → lotto\_detail 테이블의 rank\_no = 1
  - 1인당 당첨 금액 → lotto\_detail 테이블의 win\_money
  - 해당 로또 번호 → lotto\_master 테이블
  - lotto\_master와 lotto\_detail 테이블 조인
  - → 두 테이블 조인하고, 1등을 구한 다음 당첨금액 순으로 내림차순 정렬

#### (3) 가장 많은 당첨금이 나온 회차와 번호, 금액 조회

SELECT a.sed_no
,a.draw_date
,b.win_person_no
,b.win_money
,a.num1 ,a.num2 ,a.num3
a.num4 ,a.num5 ,a.num6 ,a.bonus,
FROM lotto_master a
,lotto_detail b
WHERE a.seq_no = b.seq_no
AND b.rank_no = 1
ORDER BY b.win money DESC;

1	SEQ NO # DRAW DATE		∯ WIN_PERSON_NO	<b>⊕ WIN_MON</b>	IΕΥ	NUM1	∯ NUM2	NUM3	® NUM4	∯ NUM5 {	NUM6 ∯ B	BONUS
1	19 2003-04-12	00:00:00	1	4072295	9400	6	30	38	39	40	43	26
2	25 2003-05-24	00:00:00	2	2422774	15300	2	4	21	26	43	44	16
3	20 2003-04-19	00:00:00	1	1935221	2800	10	14	18	20	23	30	41
4	43 2003-09-27	00:00:00	1	1774963	30800	6	31	35	38	39	44	1
5	15 2003-03-15	00:00:00	1	1701424	15000	3	4	16	30	31	37	13
6	36 2003-08-09	00:00:00	1	1601447	75800	1	10	23	26	28	40	31
7	62 2004-02-07	00:00:00	1	1581728	6400	3	8	15	27	29	35	21
8	33 2003-07-19	00:00:00	1	1490351	7600	4	7	32	33	40	41	9
9	82 2004-06-26	00:00:00	1	1456249	4400	1	2	3	14	27	42	39
10	862004-07-24	00:00:00	1	1425218	6400	2	12	37	39	41	45	33

#### [복권의 세계]④'당첨금 407억'..전설이 된 19회 로또

역대 최고 당첨금 407억원..실수령액만 317억 넘어 역대 최저 당첨금은 4억원..한꺼번에 1등 30명 나와

등록 2015-02-20 오전 6:10:22



traffic\_accident 테이블

컬럼	데이터형	Null	PK	설명
YEAR	NUMBER	N	Υ	년도
TRANS_TYPE	VARCHAR2(30)	N	Υ	교통수단
TOTAL_ACCT_NUM	NUMBER	Υ		사고발생건수
DEATH_PERSON_NUM	NUMBER	Υ		사망자수

- traffic\_accident 테이블 생성

```
CREATE TABLE traffic_accident (
           NUMBER NOT NULL, -- 연도
 vear
 trans_type VARCHAR2(30) NOT NULL, -- 교통수단
                               -- 사고발생건수
 total_acct_num NUMBER,
                                 -- 사망자수
 death_person_num NUMBER
```

**ALTER TABLE traffic\_accident** ADD CONSTRAINTS traffic\_accident\_pk PRIMARY KEY (year, trans\_type);

- traffic\_accident 테이블 데이터 입력

03.traffic\_accident\_insert.sql 파일

**SELECT** \*

FROM traffic\_accident;

	∯ YEAR	♦ TRANS_TYPE		⊕ DEATH_PERSON_NUM
1	1980	자동차	120182	5608
2	1981	자동차	123373	5804
3	1982	자동차	141128	6110
4	1983	자동차	170026	6834
5	1984	자동차	134335	7468
6	1985	자동차	146836	7522
7	1986	자동차	153777	7702
8	1987	자동차	175661	7206
9	1988	자동차	225062	11563
10	1989	자동차	255787	12603
11	1990	자동차	255303	12325
12	1991	자동차	265964	13429
13	1992	자동차	257194	11640
14	1993	자동차	260921	10402
15	1994	자동차	266107	10087
16	1995	자동차	248865	10323
17	1996	자동차	265052	12653

- (1) 연도, 교통수단별 총 사고건수 조회
  - 연도(year), 교통수단(trans\_type), 사고건수(total\_acct\_num)
  - 집계쿼리 → 연도, 교통수단 별로 사고건수 SUM

#### **SELECT YEAR**

,trans\_type

,SUM(total\_acct\_num) AS 사고건수

,SUM(death\_person\_num) AS 사망자수

FROM traffic\_accident

**WHERE 1=1** 

**GROUP BY YEAR, trans\_type** 

**ORDER BY 1, 2;** 

	♦ YEAR	♦ TRANS_TYPE	∯ 사고건수	∜ 사망자수
1	1980	선박	255	147
2	1980	자동차	120182	5608
3	1980	지하철	7	4
4	1980	철도	2136	778
5	1980	항공기	4	17
6	1981	선박	336	123
7	1981	자동차	123373	5804
8	1981	지하철	15	7
9	1981	철도	1894	858
10	1981	항공기	2	0
11	1982	선박	301	42
12	1982	자동차	141128	6110
13	1982	지하철	8	4
14	1982	철도	1849	755
15	1982	항공기	0	0
16	1983	선박	300	74
17	1983	자동차	170026	6834
		-1 -1 -1	-	-

#### (1) 연도, 교통수단별 총 사고건수 조회

		♦ TRANS_TYPE	∯ 사고건수	∜ 사망자수
1	1980	선박	255	147
2	1980	자동차	120182	5608
3	1980	지하철	7	4
4	1980	철도	2136	778
5	1980	항공기	4	17
6	1981	선박	336	123
- 7	1981	자동차	123373	5804
8	1981	지하철	15	7
9	1981	철도	1894	858
10	1981	항공기	2	0
11	1982	선박	301	42
12	1982	자동차	141128	6110
13	1982	지하철	8	4
14	1982	철도	1849	755
15	1982	항공기	0	0
16	1983	선박	300	74
17	1983	자동차	170026	6834
		TI TI EI	-	_

- · 원하는 결과는 얻었으나 결과 로우가 너무 많아 한 눈에 파악하기 힘듬
- 개별 년도 별이 아닌 년대(예, 1980년대) 별로 다시 집계해보자.

#### (1) 연도, 교통수단별 총 사고건수 조회

SELECT CASE WHEN year BETWEEN 1980 AND 1989 THEN '1980년대'
WHEN year BETWEEN 1990 AND 1999 THEN '1990년대'
WHEN year BETWEEN 2000 AND 2009 THEN '2000년대'
WHEN year BETWEEN 2010 AND 2019 THEN '2010년대'
END AS YEARS

,trans\_type

,SUM(total\_acct\_num) AS 사고건수

FROM traffic\_accident

WHERE 1=1

GROUP BY CASE WHEN year BETWEEN 1980 AND 1989 THEN '1980년대'

WHEN year BETWEEN 1990 AND 1999 THEN '1990년대'

WHEN year BETWEEN 2000 AND 2009 THEN '2000년대'

WHEN year BETWEEN 2010 AND 2019 THEN '2010년대'

END, trans\_type

**ORDER BY 1, 2;** 

	♦ TRANS_TYPE	∜ 사고건수
1 1980년대	선박	3965
2 1980년대	자동차	1646167
3 1980년대	지하철	208
4 1980년대	철도	18681
5 1980년대	항공기	20
6 1990년대	선박	6605
7 1990년대	자동차	2581517
8 1990년대	지하철	311
9 1990년대	철도	13415
10 1990년대	항공기	24
11 2000년대	선박	7780
12 2000년대	자동차	2331063
13 2000년대	지하철	732
14 2000년대	철도	5083
15 2000년대	항공기	63
16 2010년대	선박	11840
17 2010년대	자동차	1564103
18 2010년대	지하철	17
19 2010년대	철도	1546
20 2010년대	항공기	77

- (2) 연대별 교통사고 추이 분석
  - 교통사고 유형을 로우로, 연대를 컬럼으로 두고 교통사고 건수 조회
  - 1980년대에 발생한 교통사고 총 건수를 컬럼으로 만들려면?

**CASE WHEN year** 

**BETWEEN 1980 AND 1989** 

THEN total\_acct\_num ELSE 0 END

(2) 연대별 교통사고 추이 분석

#### **SELECT trans\_type**

,CASE WHEN year BETWEEN 1980 AND 1989 THEN total\_acct\_num ELSE 0 END "1980년대"

,CASE WHEN year BETWEEN 1990 AND 1999 THEN total\_acct\_num ELSE 0 END "1990년대"

,CASE WHEN year BETWEEN 2000 AND 2009 THEN total\_acct\_num ELSE 0 END "2000년대"

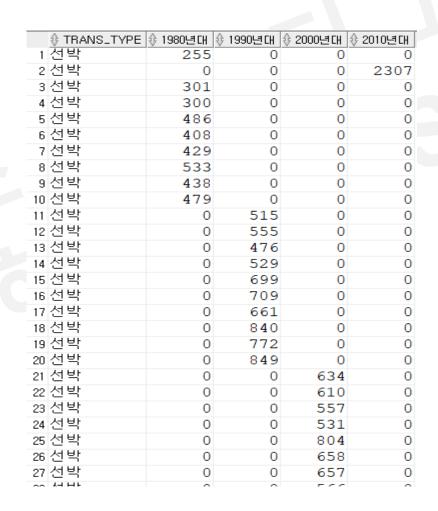
,CASE WHEN year BETWEEN 2010 AND 2019 THEN total\_acct\_num ELSE 0 END "2010년대"

FROM traffic\_accident

**WHERE 1=1** 

ORDER BY trans\_type;

(2) 연대별 교통사고 추이 분석



(2) 연대별 교통사고 추이 분석

```
SELECT trans_type
```

```
,SUM(CASE WHEN year BETWEEN 1980 AND 1989 THEN total_acct_num ELSE 0 END) "1980년대" ,SUM(CASE WHEN year BETWEEN 1990 AND 1999 THEN total_acct_num ELSE 0 END) "1990년대" ,SUM(CASE WHEN year BETWEEN 2000 AND 2009 THEN total_acct_num ELSE 0 END) "2000년대" ,SUM(CASE WHEN year BETWEEN 2010 AND 2019 THEN total_acct_num ELSE 0 END) "2010년대"
```

FROM traffic\_accident

**WHERE 1=1** 

**GROUP BY trans\_type** 

ORDER BY trans\_type;

#### (2) 연대별 교통사고 추이 분석

⊕ TRANS_TYPE	♦ 1980년대	🕸 1990년대	♦ 2000년대	♦ 2010년대
1 선박	3965	6605	7780	11840
2 자동차	1646167	2581517	2331063	1564103
3 지하철	208	311	732	17
4 철도	18681	13415	5083	1546
5 항공기	20	24	63	77

- (3) 교통수단별 가장 많은 사망자 수가 발생한 연도 구하기
  - 사망자 수 → death\_person\_num 컬럼
  - 교통수단 별 가장 많은 사망자 수
    - → trans\_type을 group by, MAX(사망자수)

**SELECT trans\_type** 

,MAX(death\_person\_num) death\_per

FROM traffic\_accident

**GROUP BY trans\_type**;

⊕ TRANS_TYPE	⊕ DEATH_PER
1 지하철	256
2 철도	858
3 항공기	276
4 선박	477
5 자동차	13429

- (3) 교통수단별 가장 많은 사망자 수가 발생한 연도 구하기
  - 교통수단별 가장 많은 사망자 수까지 구했으니, 연도만 구하면 됨
  - 교통수단 별 가장 많은 사망자 수를 교통수단별 사망자 수가 같은 건을 구하면?
  - → 가장 많은 사망자 수가 발생한 연도를 구할 수 있음
  - 앞에서 구한 쿼리 결과와 traffic\_accident 테이블을 조인해 연도를 구함

(3) 교통수단별 가장 많은 사망자 수가 발생한 연도 구하기

```
SELECT a.*
FROM traffic accident a
    ,( SELECT trans_type
              ,MAX(death_person_num) death_per
        FROM traffic accident
       GROUP BY trans_type
     ) b
WHERE a.trans_type
                       = b.trans_type
 AND a.death_person_num = b.death_per;
```

			↑ TOTAL_ACCT_NUM	⊕ DEATH_PERSON_NUM
1	1991	자동차	265964	13429
	1981	_	1894	858
3	2003	지하철	103	256
4	1987	선박	533	477
5	1983	항공기	5	276