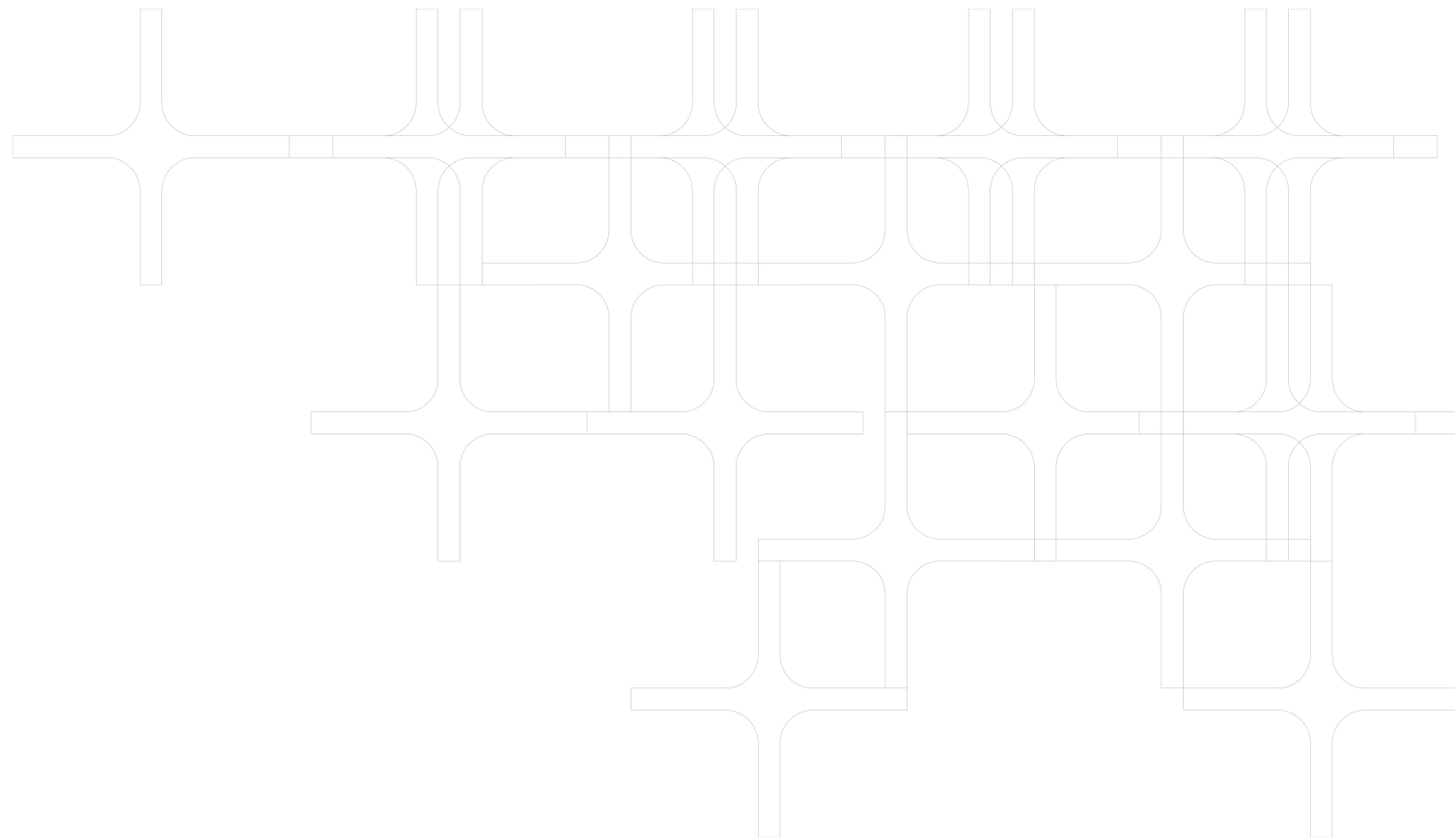


SQL 활용

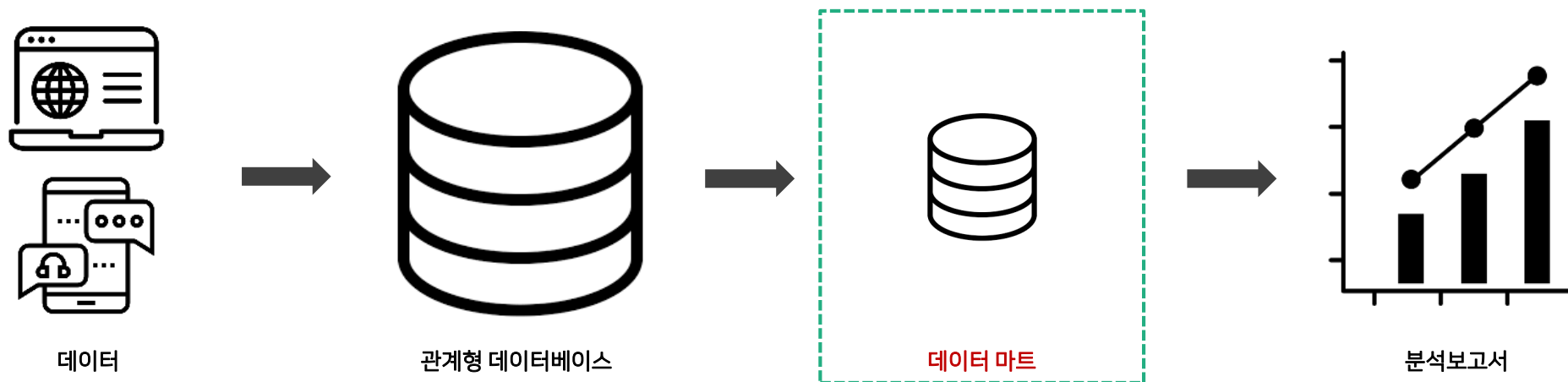
데이터 마트



데이터 마트

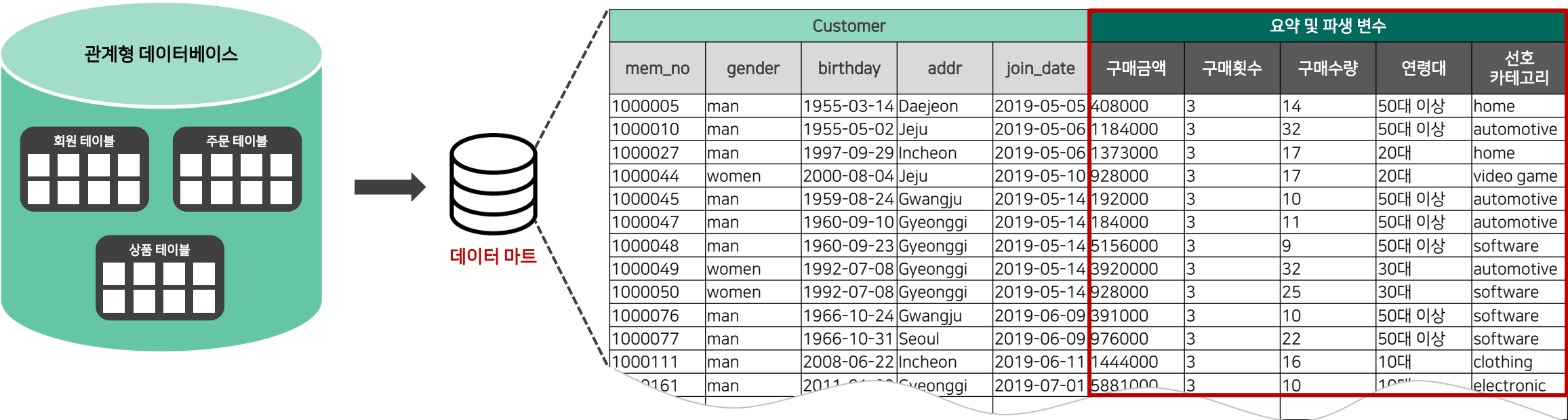
이번 영상에서는
데이터 마트 개념을 학습합니다.

데이터 마트는 **분석에 필요한 데이터**를 가공한 분석용 데이터입니다.

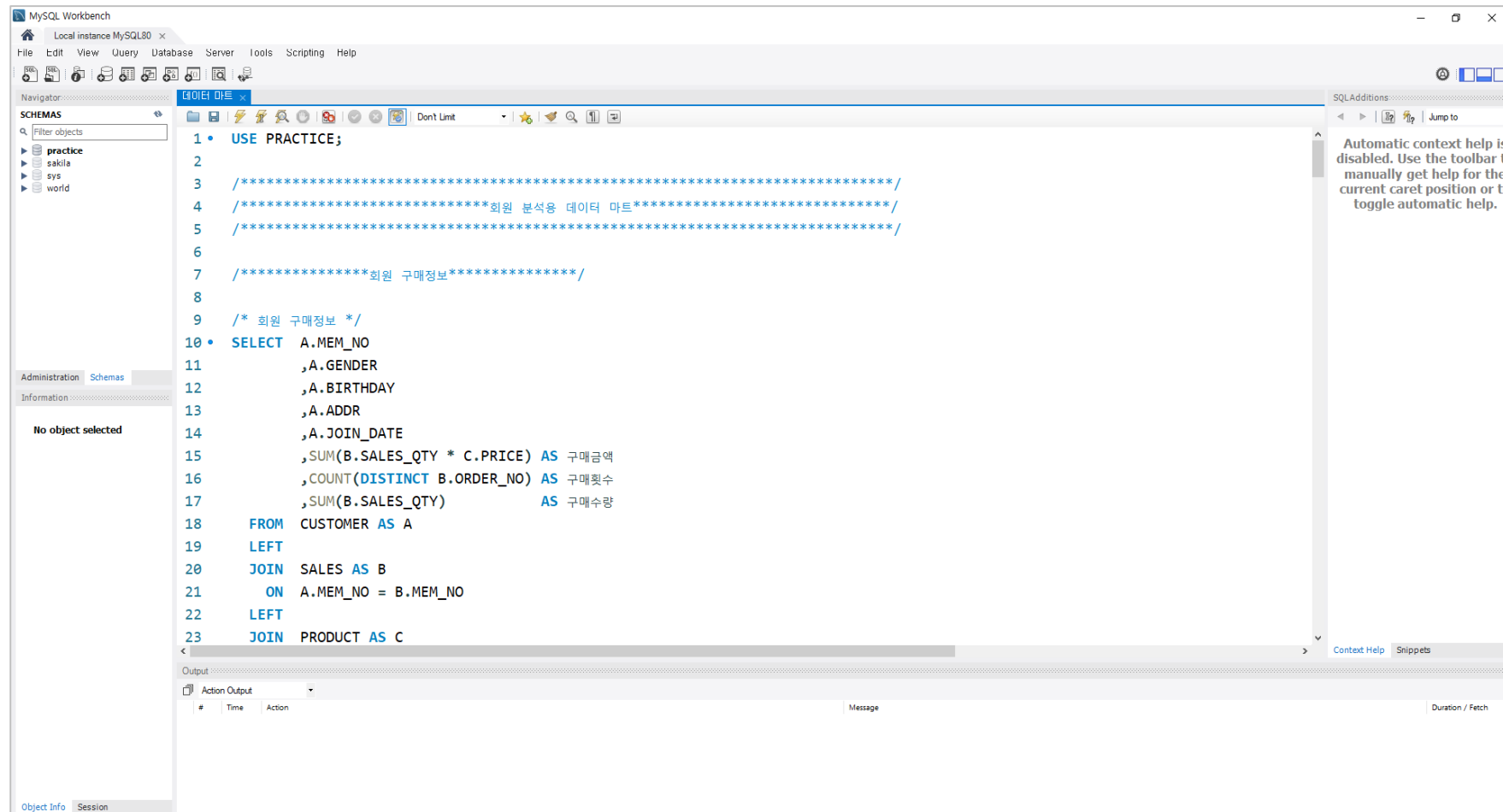


데이터 마트는 **분석에 필요한 데이터**를 가공한 분석용 데이터입니다.

- 요약 변수 : 수집된 데이터를 분석에 맞게 종합한 변수(**기간별 구매 금액, 횟수, 수량 등**)
- 파생 변수 : 사용자가 특정 조건 또는 함수로 의미를 부여한 변수(**연령대, 선호 카테고리 등**)



데이터 마트는 **분석에 필요한 데이터**를 가공한 분석용 데이터입니다.

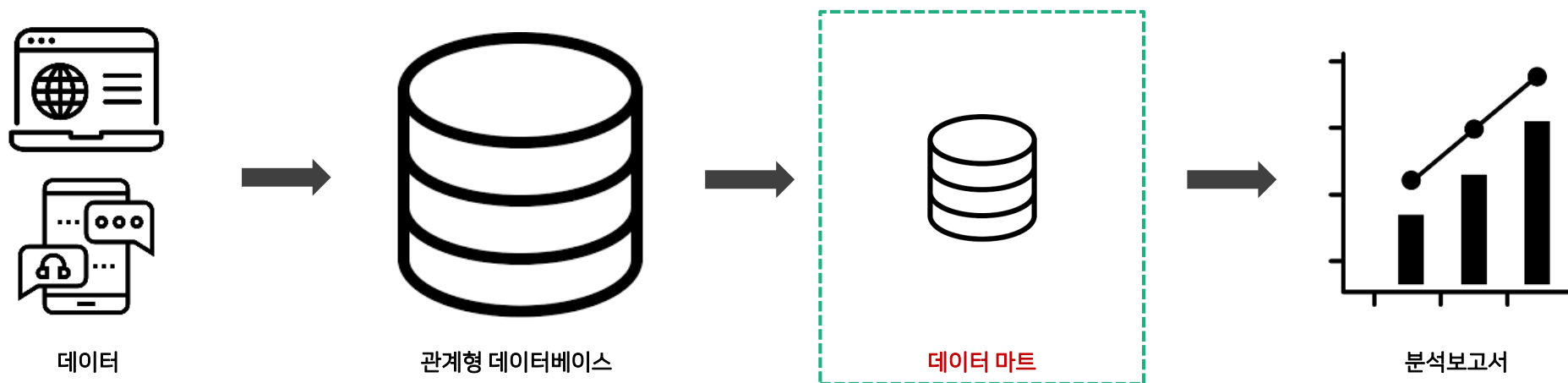


데이터 정합성은 데이터가 서로 모순 없이 **일관되게 일치함**을 나타낼 때 사용됩니다.

1. 데이터 마트의 회원 수의 **중복**은 없는가?
2. 데이터 마트의 **요약 및 파생변수**의 오류는 없는가?
3. 데이터 마트의 **구매자 비중(%)**의 오류는 없는가?

Customer					요약 및 파생 변수				
mem_no	gender	birthday	addr	join_date	구매금액	구매횟수	구매수량	연령대	선호 카테고리
1000005	man	1955-03-14	Daejeon	2019-05-05	408000	3	14	50대 이상	home
1000010	man	1955-05-02	Jeju	2019-05-06	1184000	3	32	50대 이상	automotive
1000027	man	1997-09-29	Incheon	2019-05-06	1373000	3	17	20대	home
1000044	women	2000-08-04	Jeju	2019-05-10	928000	3	17	20대	video game
1000045	man	1959-08-24	Gwangju	2019-05-14	192000	3	10	50대 이상	automotive
1000047	man	1960-09-10	Gyeonggi	2019-05-14	184000	3	11	50대 이상	automotive
1000048	man	1960-09-23	Gyeonggi	2019-05-14	5156000	3	9	50대 이상	software
1000049	women	1992-07-08	Gyeonggi	2019-05-14	3920000	3	32	30대	automotive
1000050	women	1992-07-08	Gyeonggi	2019-05-14	928000	3	25	30대	software
1000076	man	1966-10-24	Gwangju	2019-06-09	391000	3	10	50대 이상	software
1000077	man	1966-10-31	Seoul	2019-06-09	976000	3	22	50대 이상	software
1000111	man	2008-06-22	Incheon	2019-06-11	1444000	3	16	10대	clothing
1000161	man	2011-01-02	Gyeonggi	2019-07-01	5881000	3	10	10대	electronic

- 데이터 마트: **분석에 필요한 데이터**를 가공한 분석용 데이터
- 데이터 정합성: 데이터가 서로 모순 없이 **일관되게 일치함**을 나타낼 때 사용



1. 데이터 마트의 회원 수의 **중복**은 없는가?
2. 데이터 마트의 **요약 및 파생변수**의 **오류**는 없는가?
3. 데이터 마트의 **구매자 비중(%)**의 **오류**는 없는가?

End of Document

-

Thank you.

