

제2부 다차원모델

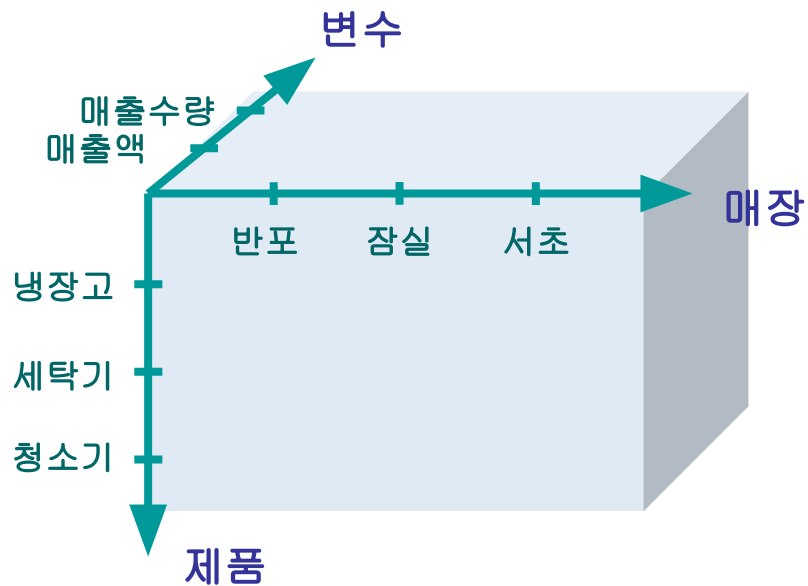
AGENDA

- 다차원모델의 구성요소
- 스타스키마
- 차원 및 큐브의 이해

다차원모델의 구성요소

- 차원과 차원항목
- 큐브와 셀
- 계층구조
- 레벨
- 애트리뷰트
- 관계식

다차원모델의 구성요소 - 차원과 차원항목



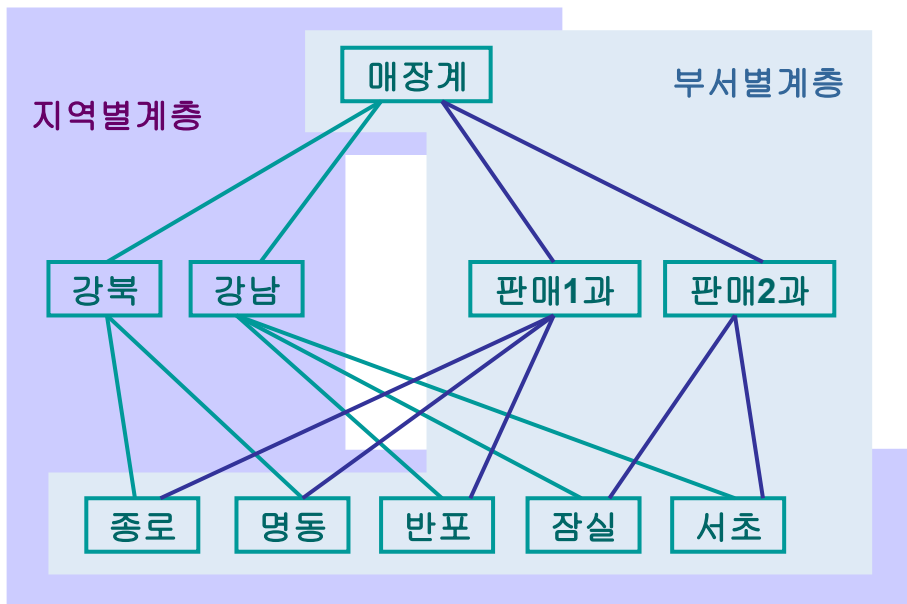
- 차원
 - 큐브를 구성하는 축
 - 사용자가 데이터를 분석하고자하는 관점
- 차원항목
 - 각 축의 좌표

다차원모델의 구성요소 - 큐브와 셀



- 셀(Cell)
 - 각 차원을 구성하는 항목들의 조합에 의한 데이터를 저장하는 논리적인 공간

다차원모델의 구성요소 - 계층구조

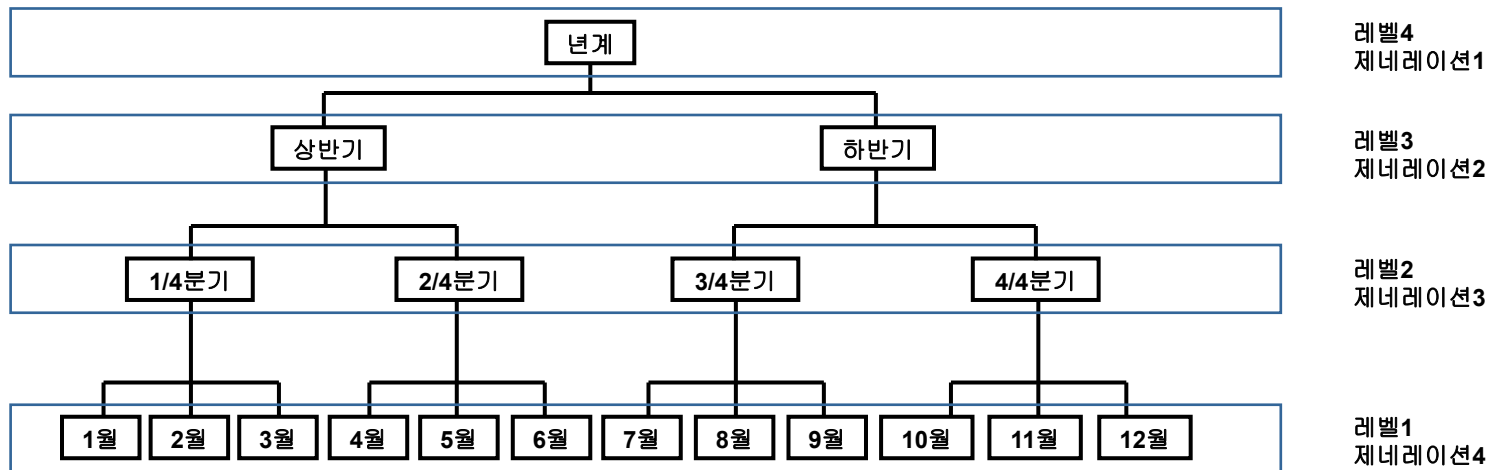


- 계층구조
 - 데이터가 집계되는 기본적인 경로
- 패어런트(Parent)
 - 어떤 항목의 바로 상위항목
- 차일드(Child)
 - 어떤 항목의 바로 하위항목
- 씨블링(Sibling)
 - 동일한 패어런트를 가진 항목
- 루트항목(Root)
 - 패어런트를 갖지않는 항목
- 리프항목(Leaf)
 - 차일드를 갖지않는 항목
- 앤세스터(Ancessor)
 - 어떤 항목의 모든 상위항목
- 디센던트(Descendent)
 - 어떤 항목의 모든 하위항목

다차원모델의 구성요소 - 레벨

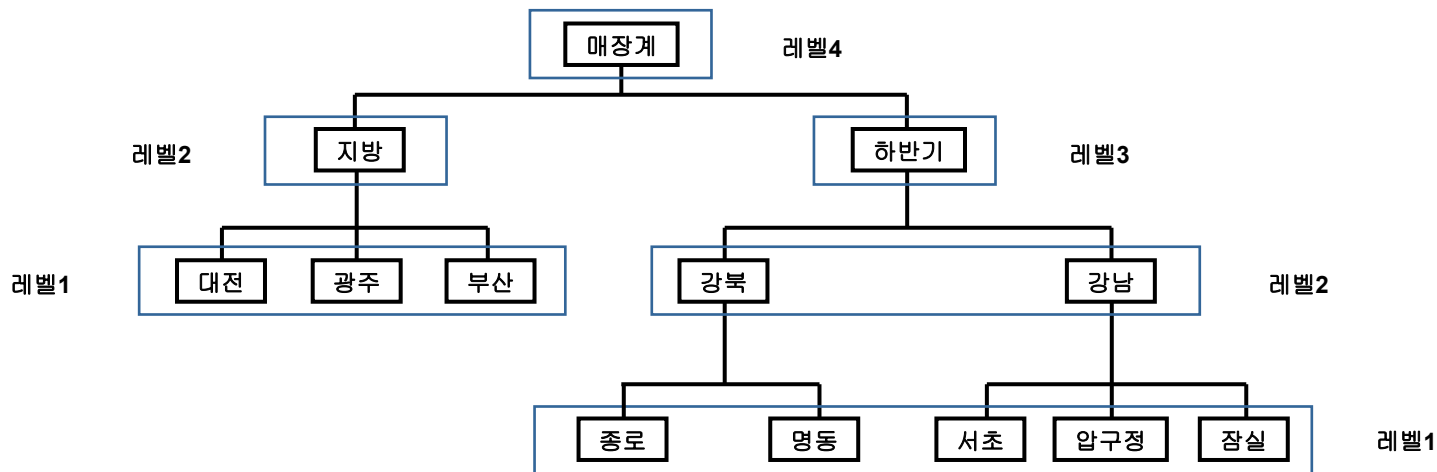
- 레벨

- 계층구조상의 거리
- 차원항목들의 부분집합



<대칭 계층구조>

다차원모델의 구성요소 - 레벨



<비대칭 계층구조>

- 년 : 1998년, 1999년 등의 항목으로 구성
- 반기: 상반기, 하반기 항목으로 구성
- 분기: 1/4분기, 2/4분기, 3/4분기, 4/4분기
- 월 : 1월, 2월, 3월 등의 항목으로 구성

다차원모델의 구성요소 – 애트리뷰트/관계식

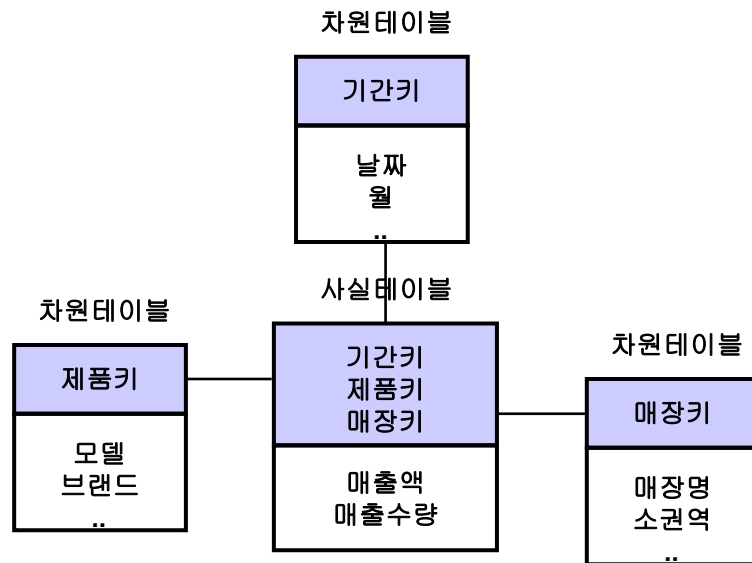
- 애트리뷰트(Attribute)
 - 차원항목들의 특성을 나타내는 텍스트형태의 정보
 - 매장 : 매장의 주소, 전화번호, 담당자, 매장크기, 직영여부, 개점일 등
 - 제품 : 제품의 색상, 크기, 신제품여부 등
- 관계식
 - 항목들간의 관계를 식으로 정의
 - 평균매출가격 : 매출액/매출수량
 - 세금 : 매출액 * 0.2

- 스타스키마
- 사실테이블
- 차원테이블
- **E-R**모델과 스타스키마
- 스노우플레이크 스키마

스타스키마 - 스타스키마

- 스타스키마

- 다차원데이터를 효과적으로 저장하고 조회하기위한 관계형 데이터베이스의 설계기법
- 사실(Fact)테이블과 차원(Dimension)테이블로 구성
- 사실 : 분석하고자 하는 대상항목
- 차원 : 사실을 보는 관점



스타스키마 – 사실테이블/차원테이블

- 사실테이블
 - 사용자가 요구하는 정보의 내용이 포함됨
 - 경영의사결정지원을 위한 소스 테이블
- 차원테이블
 - 사실테이블에서 정보를 추출하여 사용자에게 제공하기 위한 일종의 경로(Path)가 되는 테이블
- 차원테이블의 공유
- 애트리뷰트
 - 사용자의 다양한 의사결정 지원
 - 일반적으로 텍스트 형태의 값
 - 수치 데이터도 가능
- 차원과 사실이 분리되어 저장되고
검색은 차원테이블에 대해 먼저 수행됨

```
SELECT s.매장명, SUM(f.매출액)  
FROM “사실테이블” f,  
      “매장차원테이블” s  
WHERE f.매장키 = s.매장키  
      AND s.소권역 = ‘강남권’  
GROUP BY s.매장명  
ORDER BY s.매장명
```

스타스키마 – 사실테이블/차원테이블

- 차원테이블의 변화
 - 차원항목들의 애트리뷰트가 시간이 흐름에 따라 변화
- 처리방법
 - 종로 매장의 유형이 '위탁' 에서 '직영' 으로 1999년 4월 1일 변경시

a) 새로운 값으로 덮어쓰기

매장키	매장명	소권역	대권역	담당자	유형
10010	반포	강남	수도권	홍길동	위탁
..



매장키	매장명	소권역	대권역	담당자	유형
10010	반포	강남	수도권	홍길동	직영
..

- 구현이 쉬움
- 변경과정을 파악할수 없음
- 과거 데이터가 중요하지 않은경우 사용

스타스키마 – 사실테이블/차원테이블

b) 레코드 새로 추가

매장키	매장명	소권역	대권역	담당자	유형
10010	반포	강남	수도권	홍길동	위탁
..



매장id	매장키	매장명	소권역	대권역	담당자	유형
100100	10010	반포	강남	수도권	홍길동	위탁
100101	10010	반포	강남	수도권	홍길동	직영

c) 새로운 필드 추가

매장키	매장명	소권역	대권역	담당자	유형
10010	반포	강남	수도권	홍길동	위탁
..

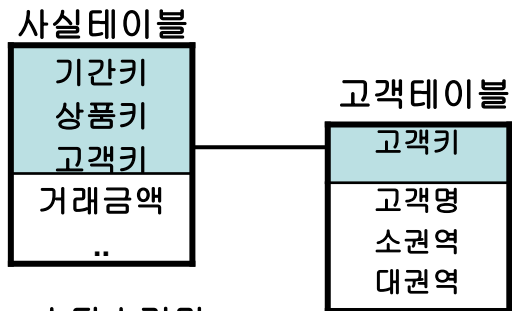


매장키	매장명	소권역	대권역	담당자	시작 유형	현재 유형	최종 변경일
10010	반포	강남	수도권	홍길동	위탁	직영	19990401
..

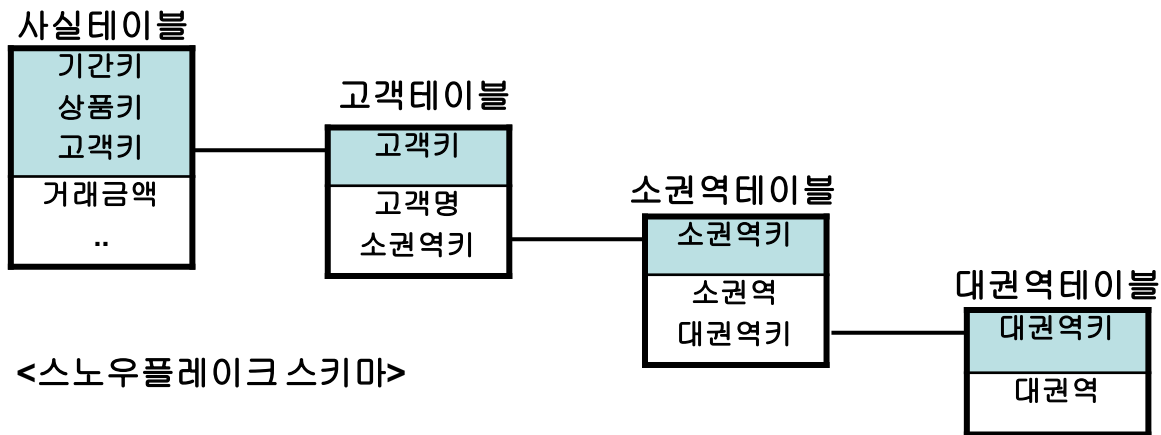
스타스키마 – E-R모델과 스타스키마

- E-R모델
 - 갱신 성능을 최적화하기 위해 정규화됨
 - OLTP
 - 다수의 테이블로 구성
 - 소수의 열, 테이블간 여러 조인경로
- 스타스키마
 - 최적의 질의 응답성능을 위한 설계기법
 - 비정규화됨
 - BI
 - 소수의 테이블로 구성
 - 다수의 컬럼

스타스키마 – 스노우플레이크 스키마

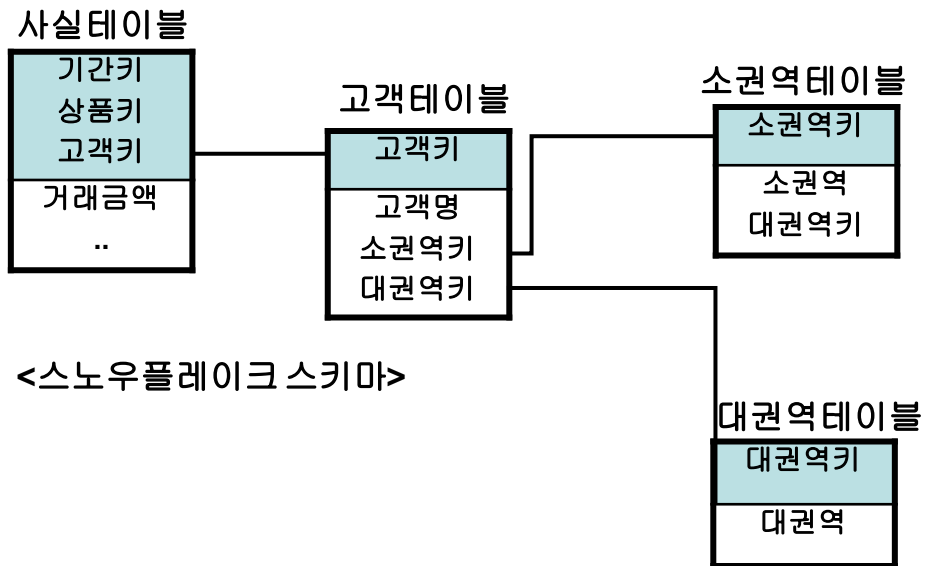


<스타스키마>



<스노우플레이크 스키마>

스타스키마 – 스노우플레이크 스키마

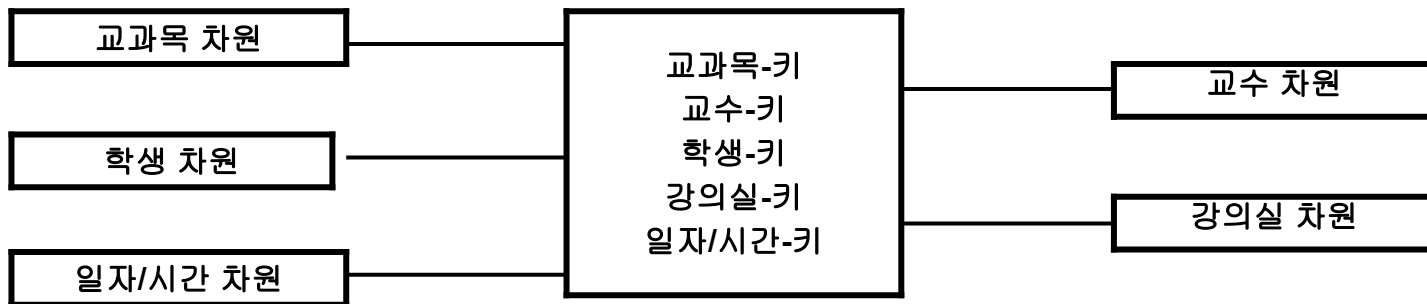


차원 및 큐브의 이해

- 변수차원
- 하이퍼큐브 와 멀티큐브
- 기간차원
- 유형차원

차원 및 큐브의 이해 - 변수차원

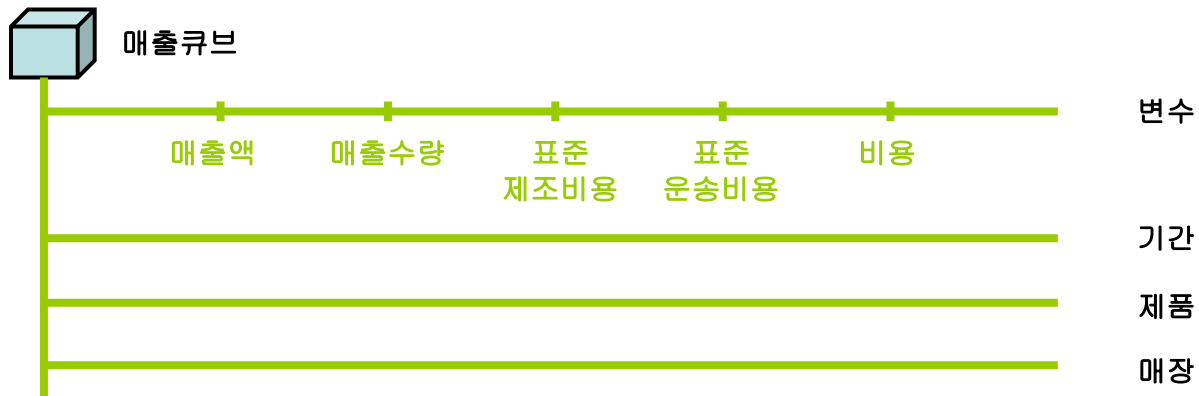
- 변수차원
 - 자산, 부채, 수익, 비용, 매출액 등과 같이 비즈니스 성과를 측정하기 위한 항목
- 특성
 - 다른 차원들의 존재기반을 제공
 - 나머지 차원들의 상세정도를 결정



<사건기록모델>

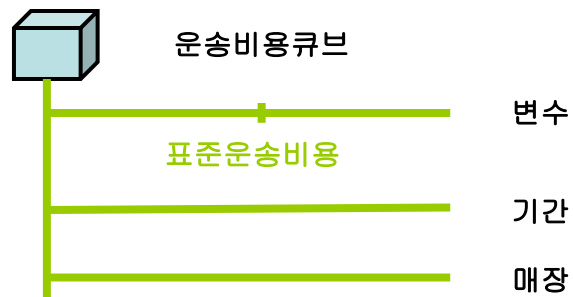
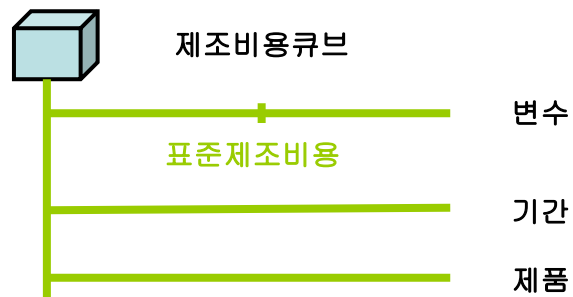
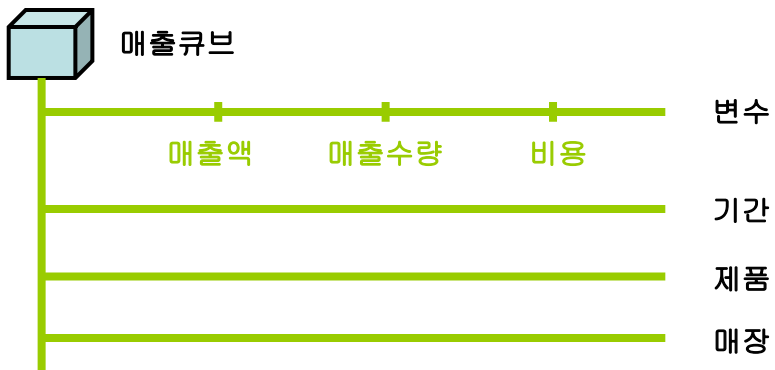
차원 및 큐브의 이해 – 하이퍼큐브와 멀티큐브

- 변수항목들과 나머지 차원들이 연계되는 방식 – 하이퍼큐브, 멀티큐브
- 매출분석모델
 - 비용 = 매출수량 * (표준제조비용 + 표준운송비용)
 - 매출액, 매출수량, 비용: 기간별, 제품별, 매장별 데이터
 - 표준제조비용: 기간별, 제품별 데이터
 - 표준운송비용: 기간별, 매장별 데이터
- 하이퍼큐브



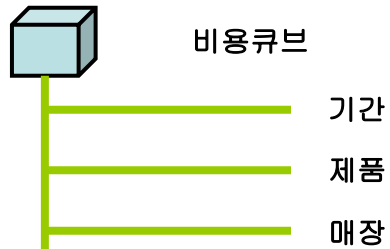
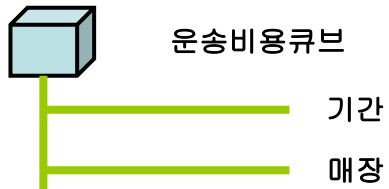
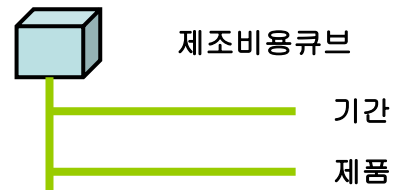
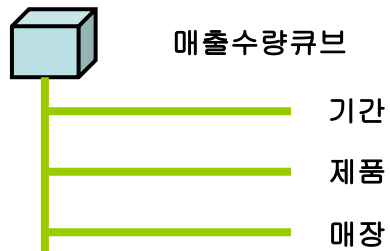
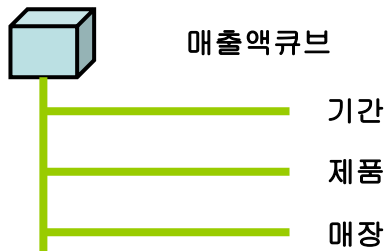
차원 및 큐브의 이해 – 하이퍼큐브와 멀티큐브

- 블록 멀티큐브



차원 및 큐브의 이해 – 하이퍼큐브와 멀티큐브

- 시리즈 멀티큐브



차원 및 큐브의 이해 – 기간차원

- 기간차원
 - 의사결정을 지원하기위한 정보의 가장 기본적인 특성
 - 연속성, 순차성
 - 변수항목의 성격에 따라 데이터 통합시 세심한 주의요함

	1월	2월	3월	1/4분기
매출액	100	200	300	600
기초재고량	100	200	300	100
기말재고량	100	200	300	300
종업원수	100	200	300	<u>300 혹은 200</u>
최고값	100	300	200	300
최소값	100	300	200	100

<변수항목과 기간차원의 데이터집계>

차원 및 큐브의 이해 – 유형차원

- 유형차원
 - 항목들 간의 차이를 측정
 - 실적, 예산, 계획, 추정 등과 같은 항목으로 구성
 - 예산대비 실적의 차이, 계획대비 실적 달성율