

데이터가 보인다(8장, 9장)

작성일: 20.03.16

작성자: 도원진

8 장 비즈니스 인텔리전스

8-1. BI 와 PDCA 사이클

BI (Business Intelligence) 란?

데이터를 이용해 기업의 의사결정에 활용하는 방법

- PDCA 사이클의 과정 속에서 정보 발생
- 정보를 유용한 데이터로 분석, 가공
- PDCA 사이클 과정속에서 발생하기도, PDCA을 진행하는데 있어 기본 토대가 되기도 한다.

PDCA 란?

기업 업무프로세스의 한 종류. "계획하고 실행하고 점검하고 보안하고"

- 두 가지 입장에서 프로세스를 진행한다.
 - 경영층 입장(Top-Down)

사측 손익구조, 시장상황, 인사, 설비, 투자계획등을 토대로 예산편성

PDCA	내용	비고
PLAN	경영전략/계획 수립	예산편성 안에서 수행
DO	사업부문의 계획 실행	중요업적지표
CHECK	모니터링, 문제 색출, 평가, 분석 및 검증	
ACT	문제점 수정, 개선, 대처	

- 현장층 입장(Bottom-Up)

각 부문/ 담당자 단위에서 필요한 매출, 경비, 개별계획을 토대로 예산편성

PDCA	내용	비고
PLAN	부문별 계획 수립	예산편성 안에서 수행
DO	부문에서 업무 실행	
CHECK	일별 모니터링	
ACT	경영층으로 보고, 수정	

BI의 역할 (요약)

1. PLAN

- 계획의 근거 획득
- EX)
 - 실적분석, 복수패턴간 비교, 시뮬레이션

2. DO

- 계획을 수행하면서 동향을 관찰하는 것(감시, 모니터링)
- 문제의 조짐을 발견. 이후, check단계로 넘어감

3. CHECK

- 결과 평가, DO단계에서 파악한 문제를 분석, 검증

4. ACT

- 대처방안과 힌트를 얻는다.
- 다음 PLAN단계에 필요한 보완사항을 제공

-----BI역할 (상세)-----

8-2. PLAN 단계

- 계획값 작성
 - DO, CHECK 단계에서 데이터 분석의 기준값으로 활용
 - 예산편성 중에 실시
 - 과거실적 분석, 현재 예산값 작성의 근거로 이용
 - 비용 산출 = 예측매출 - 필요 이익
 - Bottom - Up 데이터를 수집 / 시뮬레이션
 - Top-Down과 데이터 비교
- 예산편성 Process
 - Top-down 실시
 - 목표이익 확정 -> 거시적관점의 매출예측 -> 경비산출
 - 큰 틀의 목표이며 구체적이지 않다. 전망의 의미정도
 - (변동하는 비용 + 고정 비용) 을 구별하여 산출
 - 복수패턴 시나리오 준비(BEST, WORST)
 - Bottom-Up 실시
 - 담당 부문 / 담당자가 필요한 매출, 비용 산출
 - 매출 예측, 수요예측은 직접적으로 연결됨.
 - 수요예측
 - 고객별 / 제품별로 수요예측을 분석

8-3 DO단계

- 문제의 조짐을 발견
 - 기업의 활동이력 축적, 동향을 모니터링

- 데이터 분석 시행
- 목표값에 도달하는지 모니터링
- 빠른 문제점 분석 및 계획 수정을 가능케함
- 리포트 작성
 - 데이터 시각화를 통해 통계데이터 산출
 - 전체를 부감하는 그래프
 - 비정상값, 이슈를 쉽게 보이도록
 - 다양한 관점으로 볼 수 있는 그래프 배치
- 3개의 분야별 분석 프레임워크

유형	특징	비고
판매	1. 시간순 2. 그래프의 개형파악에 중점 3. 예상값, 실적값 비교기능이 있는 그래프제공 4. 시간 뿐만 아니라, 해당 시간에 발생한 품목별 / 고객별 분석 그래프 제공	
회계	이익관리를 위해 비용분석을 해야함. 1. 총이익이 낮다고 판단되면) 낮은 판매단가 높은 원가 영업이익이 적다면) 2. 과다한 경비사용	수익성 평가 segment : 시장, 제품, 고객 비용평가 segment : 영업, 개발, 생산
재고	1. 재고부족) 기회손실, 신용실추로 직결 및 매출저하 2. 재고과잉) 재고비용 증대, 현금흐름 효율저하	'재고 최적화'가 주요 과제

8-4. CHECK 단계

- 문제의 요인을 검증
- 다차원 분석 축을 이용해 과정을 바꿔가며 이슈를 분석
- 문제상황에 대해 DO단계보다 심층적으로 파고들어감.
- 단계
 - 결과평가 --> 요인분석 --> 요인검증
- 1. 결과평가
 - 총체적인 실적을 PLAN의 계획값과 비교/평가 실시
 - 적절한 요소로 분리하여 또 한번 비교 평가
- 2. 요인분석
 - 왜 결과가 그렇게 나왔는가 원인을 찾는일
 - 다차원의 분석축 도입 (다양한 관점)
- 3. 요인검증
 - 분석축을 교체해 가며 분석
 - 담당자 / 고객 / 상품 등등 이 중에서 요인을 검증하기

8-5. ACT단계

- CHECK단계의 차트분석 보다 더 면밀한 데이터를 기반으로 해결방안 모색
- 분석자의 기준이 중요한 판단요소가 됨.
- 따라서, 분석자의 역량이 중요시된다.
- DataMining을 통해 미지의 관계, 경향 도출

마케팅에서 ACT활용

1. 마케팅 전략/목표설정
2. 소비집단구분(Segmentation) / 타겟팅
 - 전략적 판매대상 선정
3. 포지셔닝 진행
 - 비교우위 선점

-----BI역할 (끝)-----

8-6. '포지셔닝' 을 위한 구체적인 방법

- 무엇을
 - 어떤 가격에
 - 뭐를 통해
 - 어떻게 알릴까
1. 고객은 무엇을 필요로 하는가?
 - 연령별 / 성별 / 취미 별 속성으로 앙케이트 조사
 - 그러나, 더 다양한 방법으로 조사해야함.
 - 소비자의 요구사항은 점점 복잡화되고 있음.
 2. 얼마에 팔아야 할까?
 - 구매자의 적정가격 감각을 바탕
 - 판매 데이터를 기반으로 **판매단가 - 매출수** 관계를 찾기(회귀분석 이용)
 3. 유통망 선택
 - 제조사 채널 - 대리점, 직영점, 인터넷
 - 도매업 채널 - 약국, 편의점, 슈퍼,백화점, 양판점
 - 점포판매 채널 - 도심, 교외판매
 - 유통망선택은 제고효율성 증대에 기여할 수 있다.
 4. 판매촉진
 - 비용대비 효과가 가장 중요함.
 - 광고매체별 수익 - 비용 데이터 분석
 - 캠페인의 경우 광고매출에 얼마나 기여를 하는지 파악하기 어려움(규모의 한계)

9. 데이터 웨어하우스

업무 애플리케이션에서 발생하는 데이터를 저장하는 저장매체

데이터 웨어하우스 저장 및 이용 순서

[업무어플리케이션 ----> DSA ----> 센트럴웨어하우스 ----> 데이터 마트 ----> 비즈니스 인텔리전스]

- 용어 설명
 - DSA
 - ETL(Extract, Transform, Load)하는 곳.
 - 센트럴웨어하우스에 저장할 데이터를 Staging하는 곳
 - 데이터 마트
 - 센트럴웨어하우스에 접근을 최소화하는 방법
 - BI툴에 맞는 최적의 DB구성
 - 클라이언트 어플리케이션인 BI툴에서 접근

9-1. DSA 정리

과정	특징	비고
추출처리	수신파일로 저장 기간제) 추출프로그램을 통해 파일전송 오픈계열DB) 반대로 웨어하우스가 SQL로 row정보를 가져옴.	
1차처리	비정상데이터 처리 예) 문자코드 변환, 취소레코드 삭제, 트랜잭션 취소	
변환처리	웨어하우스에 로드가능한 파일로 변환	아래 설명 참고
로드처리	파일로 변환된 정보를 DB에 저장. 1레코드씩이 아닌 일괄적으로 INSERT	

- 변환처리
 - 단순변환
 - 날짜형식 ----> 문자열
 - Null, 공백 ----> 0
 - 클렌징
 - 부적합 데이터 삭제
 - 통합화
 - 복수의 테이블에 저장된 데이터를 하나의 테이블에 통합
 - 테이블 갯수 절약
 - 집계
 - 데이터 입도(집계단위)에 따라 집계한다.

9-2. 스타스키마

데이터베이스의 구조를 뜻함. ★ 별모양

- 별의 중심 (사실테이블)
 - PK키가 되는 키값 예) 시간, 고객, 부서, 매출수량, 매출금액

- 별모양 끝 (차원테이블)
 - 각 키값들의 속성
 - 예)
 - 시간(일 / 월 / 사분기 / 반기)
 - 고객테이블(고객코드/고객명/주소)
 - 부서테이블(부서코드/ 부서명)
- 장점
 - 빠른 검색속도
 - 중복저장 최적화

써로게이트 키

업무시스템의 코드만을 검색키로 사용했을 때, 문제점 발생 예) 부서 분할

따라서, dummy number 도입.

- 고유키가 있음에도 '아무의미 없는 값'을 가진 써로게이트 키 도입
- 부서가 분할되어 2개 부서가 되도 다른 테이블에서 FK 참조가 가능하다.

9-3. 다차원 데이터베이스

OLAP 툴의 기반이 되는 데이터베이스 (데이터마트 Layer)

- 특징
 - 빠르고 간단한 크로스 집계
- 크로스집계란?
 - 2가지 항목에 대해 특정수치(SUM, MUL, SUB, DIV)를 집계해서 그 결과 보여줄 때, 한 항목은 row에, 다른 한 항목은 column에 배치하여 표를 보여주는 처리
 - 관계형데이터 베이스의 취약점을 개선
- 방대한 양의 SQL문 작성을 요구한다.
- 방대한 수의 크로스 집계 표를 3차원 주사위같은 이미지로 테이블 저장
- SQL없이 OLAP 툴, 플래닝 툴을 통해 사용
 - 검색기준 3가지
 - 다이스(행, 열)
 - 슬라이스(데이터 범위)
 - 드릴(집계단위)