Framework of Data Driven Marketing

Problem-driven vs. Data-Driven Marketing

전통적 시장 조사

- 시장 조사 프로젝트는 대부분 특정 문제 해결에 초점 (예: 브랜드 XYZ의 시장점유율 감소. 그 이유는? 대책은?)
- 컨텐츠: 고객 니즈, 선호도, 태도 등
- 방법: 서베이등의 기법을 활용한 stated preference data 수집
- Macro 마케팅에 초점
- "보험"의 일환? ("we know the answer but just to prove we are right we do some MR") - 기존의 인식과 조사 자료간의 불일치 발생시 자료 수집의 오류 가능성 높음

Data Driven Analytics

- 데이터가 비즈니스를 드라이브: 비즈니스 기회 발굴; 신규 가치 창출; 신규 고객 확보 등
- 자료 소스: transaction data (은행, 신용카드, 이동전화, 포털, 항공 등), syndicated services (Nielsen, IRI, IMS 등),
- 컨텐츠: 고객 행동, 구매 정보, 웹브라우징 정보,
- Addressable consumers(개별 고객의 ID 파악, 고객 히스토리 확보) - Macro 및 Micro 마케팅 모두 가능
- 데이터가 가져다 주는 불편한 진실 – 기존의 인식과 실제 자료간의 불일치의 경우 기존의 인식의 오류 가능성 높음

What is “Data-Driven”?

사례: 애틀란타 지역의 Gwinnett 카운티 공립 학교 시스템

- 2002년, 이 공립 학교군의 문제점들
  - 학생들의 성취도가 모든 기준에서 점점 나빠짐
  - 특히 열등 학생들과 불우한 환경에 놓인 학생들의 졸업율이 하락
  - 다양한 교육 정책 측면의 시도가 별 효과를 내지 못함
- **Data Driven Analytics**
  - 풍부한 데이터: Gwinnett은 미국에서 14번째로 큰 공립 학교 시스템 보유 (23,000명 고용), 매일 등교하는 학생/교직원 수가 델타항공이 매일 실어나르는 고객 수보다 많음
  - Analytics로부터 해답을 찾기 시작: “낙제에 근접한 열등생의 경우, 그 학생의 성취도의 어떤 측면이 그 학생의 궁극적 졸업 여부를 가장 잘 예측하는가?”
  - Analytic-driven answer: **Algebra 1** (보통 9학년이나 10학년 학생들이 이수해야 하는 과목)에서의 성공 여부가 졸업 여부를 예측하는 가장 파워풀한 변수로 나타남
  - 문제점: 저학년에서 이미 수학 과목에서 낙제한 학생들이 Algebra 1을 통과하도록 할 수 있는 방법은?
  - 데이터의 대답: 선수 수학 과목에서 낙제한 학생들의 Algebra 1에서의 성공에 대한 best predictor는 8학년 수업인 “create writing”에서의 성공 여부

- 결과
  - Analytics에서 얻은 해답을 기초로 카운티 교육 당국의 투자: 8학년 creative writing 과목에서 학생들이 성공할 수 있도록 관심과 지원 쏟아부음
  - 글쓰기 과목 성공율이 높아짐
  - Algebra 1의 성공율이 높아짐
  - 2010년 가을 Gwinnett은 Broad Prize 수상 (소득과 인종 계층간 차이를 최소화하면서 학생들의 성취도와 개선이 큰 학군에 수여하는 상)
- **Analytics의 불편한 진실**
  - 전통적으로 오랫동안 존중되어졌던 교육적 접근 방법이 별 효과가 없을 뿐만 아니라 심지어 방해 (counterproductive)가 된다는 결과가 Gwinnett의 analytics 에서 나타남
  - 모든 사람이 analytics의 결과를 환영하지는 않음: 오랫동안 존중되어왔던 신념이 틀리다고 증명

Data Driven Business Analytics

Data와 Analytics를 통한 Business value 창출

Big Data

- 자료의 크기 측면: 일반적 소프트웨어로 처리 가능한 범위를 벗어나는 대용량의 자료
- 자료의 이용 측면: 대용량 자료로 부터 가치있는 정보와 actionable recommendation을 추출하는 기법

"high-volume, high-velocity, and/or high-variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization"

Data Mining

- Knowledge Discovery in Database
- 데이터베이스 시스템과 통계 기법 활용

"extract information from a data set and transform it into an understandable structure for further use"

Business Analytics

- 경영상의 과거 경험(자료)을 토대로 가치있는 통찰 획득 및 최적의 경영 계획 수립
- 광범위한 자료의 수집과 분석, 통계 및 계량 분석, 예측 기법 등 이용

Bigger Data

Name(Symbol)	Value	Binary Usage
Kilobyte (kB)	10^3	2^{10}
Megabyte (MB)	10^6	2^{20}
Gigabyte (GB)	10^9	2^{30}
Terabyte (TB)	10^{12}	2^{40}
Petabyte (PB)	10^{15}	2^{50}
Exabyte (EB)	10^{18}	2^{60}
Zettabyte (ZB)	10^{21}	2^{70}
Yottabyte (YB)	10^{24}	2^{80}

- IDC 추정: 세계 전체 데이터 량이 2012년에 2.7ZB에 이를 것 (2011년 1.8ZB 대비 48% 증가), 2020년까지는 44ZB에 이를 것으로 추정
  (* 1.8ZB: 2시간 분량의 HD영화 2천억편 해당 - Woori 연구소)

Big Data

Big Data의 3V 모델

- Volume: 방대한 양
- Velocity: 데이터 생산, 유통, 분석 속도 증가
- Variety: 다양한 데이터 유형, 데이터 원천 (기존 구조화된 기업 데이터베이스외에도 인터넷 등에서 발생하는 비정형 텍스트, 사진, 비디오, 오디오 등)

Big Data 예 (위키피디아)

- 월마트: 매시간 백만명 이상의 고객 거래 내역 처리, 2.5PB 에 이르는 것으로 추정되는 자료 누적 (미국 국회도서관 소장된 정보 전체의 1000여배)
- Facebook: 500억장에 이르는 사용자 사진
- Google에서의 2012년 8월 한달간 검색건수: 1000억건
- FICO Falcon Credit Card Fraud Detection System: 21억개의 신용카드 계좌에서 비정상적 거래 패턴 탐지하여 계좌 보호
- 인간 게놈 해독에 처음에 10년 걸림, 지금은 하루안에 가능
- 전세계 비즈니스 데이터는 매 1.2년마다 배로 성장하는 것으로 추정

IoT

웹로그

RFID

소셜네트워크

인터넷텍스트

인터넷검색색인

통화내역

게노믹스

군사정찰

의료기록

전자상거래

비디오 아카이브

.....

McKinsey (2011) 추정치

데이터 관련 수치 (McKinsey 2011)

- 전세계 모든 음악 파일을 저장할 수 있는 하드디스크의 값: $600
- 매월 Facebook에서 공유되는 컨텐츠 건수: 300억건
- 년간 생산되는 전세계 데이터 성장율 (40%) vs. 전세계 IT 지출 성장율 (5%)
- 2011년 4월 기준 US Library of Congress 자료 크기: 235TB

빅데이터 활용의 잠재 가치

- 미국 헬스케어 분야: 년간 3000억불 가치 증대 (스페인 년간 헬스케어 전체 지출액의 두배)
- 유럽 공공 행정 부문: 년간 2500억 유로 가치 증대 (그리스 GDP 상회)
- 전세게 개인 위치 정보 활용을 통한 소비자 잉여 증대: 년간 6000억불
- 소매 분야 영업 마진 60% 증가

Data Mining

Knowledge Discovery in Database (KDD)

- 대형 데이터군에서 패턴을 발견하는 프로세스, 통계적 기법 사용
- 컴퓨터 과학과 통계의 결합: 인공 지능, machine learning, 통계, 데이터베이스 관리

Data Mining 활용

- risk scoring
- CRM
- fraud detection (통신, 신용카드)
- Market Basket Analysis: cross-selling
- targeting
- 수요예측, 재고모형
- ……

Data Mining 기법

- Anomaly detection – 비정상적인 outlier 탐지
- Association Rule Learning (Dependency Modeling): 변수간 관련성 탐색, Market Basket Analysis
- Clustering: 그룹화, 군집화
- Classification: CART:(Classification and Regression Tree)
- Variable selection (no substitute for thought there)
- Summarization: 데이터 축약, visualization, 보고서

Data Mining 소프트웨어

- Microsoft Analysis Services
- SAS Enterprise Miner
- Statistica Data Miner
- LIONsolver
- SPSS Clementine

Business Analytics

데이터와 분석적 기법의 결합

- 광범위한 자료의 이용
- 통계, 계량 분석 기법과 predictive modeling 이용
- Fact-based management to drive decision making

Business Analytics의 역사

- Taylor의 과학적 관리법, 시간 관리법
- 헨리 포드의 조립 공정 pacing 측정
- 1960년대의 의사결정 지원체제
- ERP (enterprise resource planning) system, data warehouse 등으로 진화
- 최근 대량의, 양질의 자료 축적이 일반화되면서 business analytics의 정확도가 높아짐에 따라 경영 현장 활용도가 급격하게 증가되는 추세

Business Analytics 주 이용 분야

- 마케팅 분야: pricing, assortment, CRM
- 금융 서비스 분여: 위험/ 신용 분석, Fraud 탐지
- 공급 사슬 분야: 배송, inventory

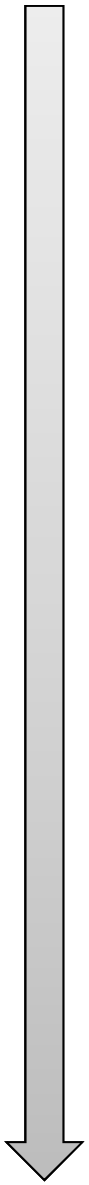
Business Analytics를 성공적으로 활용하는 기업의 특성

- 고위 경영자의 fact-based decision making에 대한 선호가 크다
- 일반 기술적 통계 뿐만 아니라 predictive modeling과 복잡한 최적화 기법의 사용이 보편화
- 경영의 다양한 부문과 기능에 걸쳐 analytics를 실질적으로 활용
- Analytics 사용이 전사적 수준으로 확대

Data Driven Marketing Analytics

- Data–Driven Marketing : 데이터가 마케팅 활동의 근간
  - 데이터가 마케팅 value 창출의 근원

- 마케팅 데이터의 종류
  - 고객 태도 자료(stated preference data): survey 등을 통해 얻은 고객의 구매의도(intention) 및 인지적 태도 정보
  - 고객 행동 자료(revealed preference data): 실제 시장에서 고객의 구매 행동을 추적한 자료
  - 고객 행동 자료에 focus
- 분석
  - 고객 구매 행동의 계량적 분석
  - 마케팅 프로그램의 평가(How profitable are my promotions?)
  - 최적 마케팅 프로그램 탐색(What price should I charge?)
  - Direct marketing: 개별 고객별 특화된 마케팅 활동을 통해 최적의 가치 창출/획득



마케팅 목표 설정

- Measurable
- 시장/고객반응 기반 (판매량, 고객 선택율, 재구매율, etc)
- 종속변수 측면

마케팅 모형 설계

- customer choice drivers 파악 (개인 변수, 기업 마케팅 투입량 , etc)
- 종속변수-독립변수 관계 모형

모형 추정과 확정

- 모형 추정
- 모형 진단 및 평가, 최종 모형 선택

마케팅 프로그램 개선

- 마케팅 목표 달성을 위한 프로그램의 개선 방향과 정도

데이터 이해

통계 기법

최적화기법

Various Sources of Marketing Data

- Improved aggregate data: 점포 수준에서 수집한 스캐너 자료, 수퍼마켓에서 시작되어 전분야의 소매점으로 확대
  - A.C Nielsen (ScanTrack), IRI (InfoScan)
  - On-line retailing as a source of POS data

- The addressable consumer: 개별 고객에 대한 자세한 정보 - 빌링데이터(통신회사 월별 요금 고지서), 개인 신용카드 사용 자료 등
  - Metromail: 미국 가정의 인구 통계/행동 정보
  - Equifax: 전세계 6억개의 신용카드 계좌의 이용 내역 정보

- Specialized samples of households
  - 미국 인구조사국의 패널: Consumer Expenditure Survey
  - Nielsen/IRI의 household scanner panel
  - 인터넷 웹사이트 등록/브라우징 내역(Media Matrix)

- 개별 기업이 보유한 거래 내역
  - frequent shopper program (마일리지 프로그램) 통한 고객 거래 내역 축적
  - Web/app usage data

More on Scanner Data

- 주요 유형
  - Store POS scanning data: 대부분의 중/대형 소매점
  - Household Scanner Panel
  - 미디어 사용데이터
    - TV 시청률(people meter)
    - 프린트 광고 audit 데이터
    - 인터넷 광고 노출 및 웹 브라우징 자료 (Media Matrix, Nielsen Net Ratings)
- 주요 Data Vendor
  - Nielsen Market Research (NMR)
    - SCANTRACK
    - Pioneer in store panels/TV monitoring
    - Pioneer in Europe
  - Information Resources Inc. (IRI)
    - INFOSCAN
    - Pioneer in household panel experimentation
    - 데이터 공급에서 NMR과 비슷한 수준
    - P&G와 같은 거대 제조업자와 계약

- Scanner Data의 이용자
  - 제조업자
    - 대부분의 CPG 제조업자들이 syndicated service 이용
    - 대부분 실험보다는 모니터링 목적
  - 소매업자
    - 원자료를 NMR/IRI에 제공: cleaned and summarized data를 NMR/IRI로부터 다시 받음
  - 광고 대행사
    - 광고주(제조업자)들이 광고 캠페인의 효과를 모니터하고자 요구
    - 단순히 브랜드 인지도만이 아니라 실제로 판매 증대 여부가 광고 효과 증명의 도구
  - 니치 플레이어
    - EMS(Efficient Marketing System) - POS장비에 직접 연결, 데이터 클리닝/에디팅, "품절" 경고를 위한 예측 모듈 제공
    - Catalina Marketing Inc - 약 60% ACV정도의 점포들에 전자 쿠폰 발행/배부 장치 설치, POS 터미널과 연결하여 쿠폰 출력
    - Valassis - frequent shopper data를 사용하여 target couponing
    - Planet U - 인터넷 쿠폰 발행

- Data Cube
  - Geography x Product x Time x Variable
  - 한 카테고리 내에서도 G x P x T x V가 10,000을 쉽게 넘어감
  - Example: Kraft Frozen Dinner Product Category
    - "Markets"
      - Total US $2 mm +
      - Total US $4 mm +
      - 지역: 서부/중서부/남부/동부
      - 주요시장: LA/시애틀
      - 어카운트: Lucky's/Ralphs/Vons
    - "Products"
      - Total frozen dinners/entrees
      - Total frozen dinners
      - Total beef entrees
      - Budget Gourmet Low Cholesterol Beef Stroganoff Dinner 11 oz
    - "Facts"(variables)
      - Sales in equivalent units
      - Average price
      - Any feature/display units
      - Price reduction $
    - "Periods"
      - Week
      - 4 weeks
      - Quarter
      - year

- Data Aggregation
  - 분석을 하기 위해 어느 정도의 aggregation 필요
  - 타입
    - Temporal
    - Item Aggregates(UPC/SKU → brand): 보통 brand/size/flavor/packaging이 aggregation 기준
  - 변수
    - 판매량: 쉬움(단순히 더함)
    - 가격: 평균 가격?
    - Display/feature/distribution ? : 주로 % ACV

- Typical Store Level Scanner Variable
  - 판매량 측정치
    - Sales in Equivalent Units(oz/LBS)
    - Sales in dollars
    - Sales Velocity = $ sales of item/ACV
    - ACV(**A**ll **C**ommodity **V**olume): $ sales of all items in a store (점포 규모 지표)
  - Distribution
    - 제품의 유통망 커버리지 측정치
    - % ACV: 제품을 판매하는 점포들의 percentage (단, 점포 크기를 고려)
      - Ex) 64oz Hellman's 마요네즈

점포	ACV	제품 취급 여부
1	200/주	No
2	300/주	Yes
3	400/주	Yes

% ACV = (300+400)/(200+300+400)

  - Baseline Equivalent Units
    - 프로모션 활동이 전혀 없을 경우에 예상되는 "normal" 판매량
  - 가격
    - 평균가격 = $ Sales / Sales in Equivalent Units

- 프로모션 활동
  - Sales dollars, Sales Units, % ACV 등으로 측정
  - Display/Feature
    - Display 유형: any type/front/end-of-aisle/in-aisle
    - Feature 유형: Major ad("A"), Minor ad(B, C), line ads, all types
  - Display only
  - Feature only
  - Display and Feature
- 가격 인하
  - "정규"가격으로부터 인하폭
  - Price decrease of 10% (5.1 ~ 15)
    20% (15.1 ~ 25)
    30% (25.1 ~ 35)
- 쿠폰
  - Retailer 쿠폰
  - 제조업자 쿠폰 - Circulation/average face value

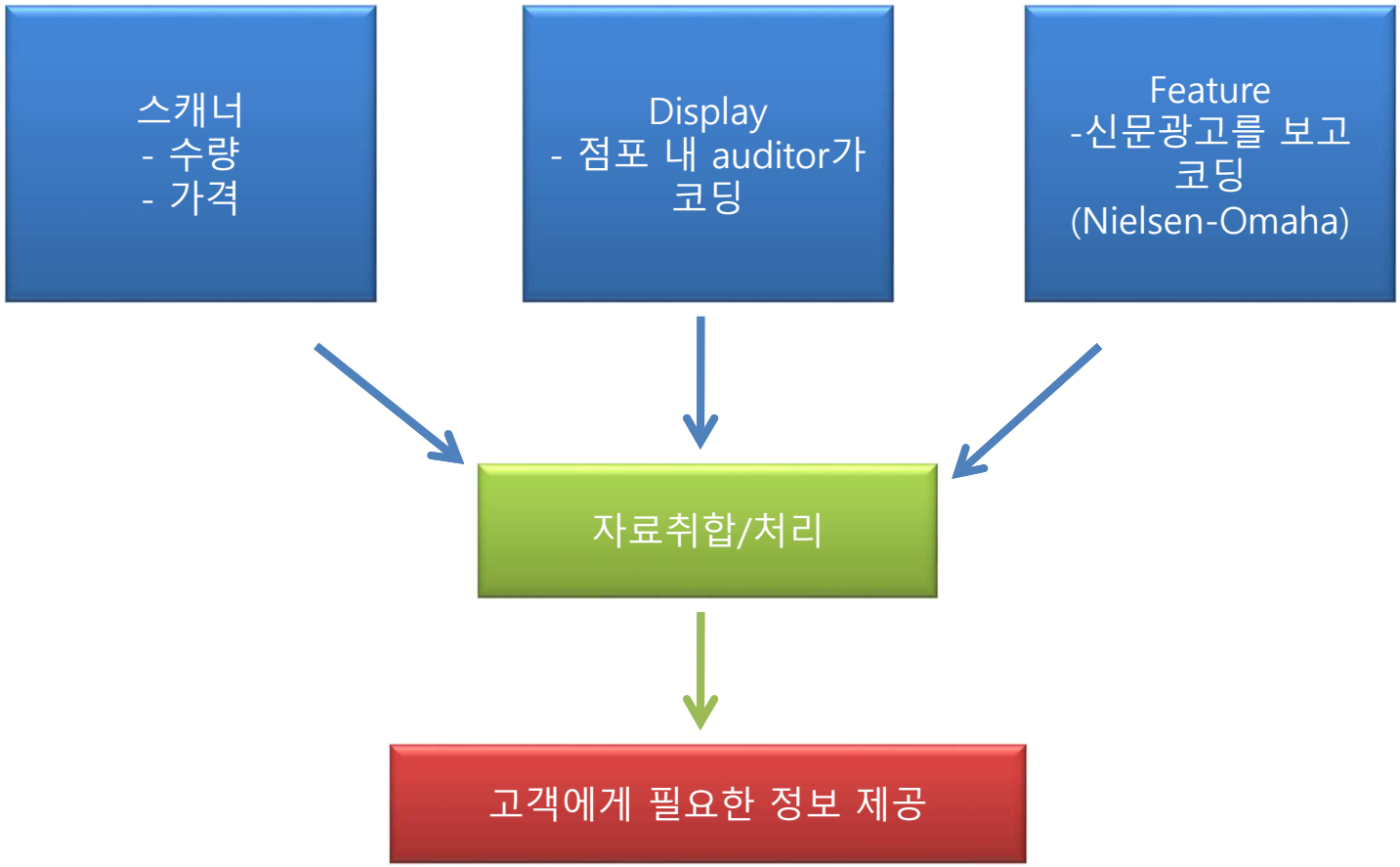
- TV Media 측정
  - GRP(Gross Ratings Points) = Reach x Frequency
    - Reach: 광고에 노출된 가구의 비율
    - Frequency: 광고에 노출된 횟수
    - 예: 80%의 가구가 평균 3회 광고 시청 = 240GRPs
  - GRP 자료는 세부적 자료로도 이용 가능
    - 고객 집단별 (예: 25 – 54세 여성)
    - 시간대별 (아침, 오후, 프라임 타임)
    - 매체 (케이블/지상파)

Continuing Trends in Marketing Data

- Increased Coverage of scanner system
  - 지역적: 유럽/아시아
  - 카테고리: 자동차(J.D Power), 의약품(IMS)
- 인터넷, 모바일 사용 자료와 연계
  - IRI와 Media Matrix의 파트너쉽: e-SCAN
  - Nielsen Korea
    - Web: web URLs of 12,000+ PC users
    - Mobile: application installation records and application usage records of 6,000+ mobile users
  - 다양한 모바일 자료, Location Based Services 자료
- 소비자 행동 자료 추가
  - eye-ball movement 분석
  - 쇼핑카트 위치 추적: 고객의 점포 내 이동경로 유형 파악
  - RFID: 구매 후 소비시점/장소 파악 가능
- 고객 마일리지 제도(frequent shopper program) 확대 지속

마케팅 정보회사의 자료 구조

• 자료 수집 Flow



- 스캐너 자료 – UPC
  - Universal Product Code

제조업자가 부여한 제품식별 번호

Uniform Code Council에서 부여한 제조업자 ID

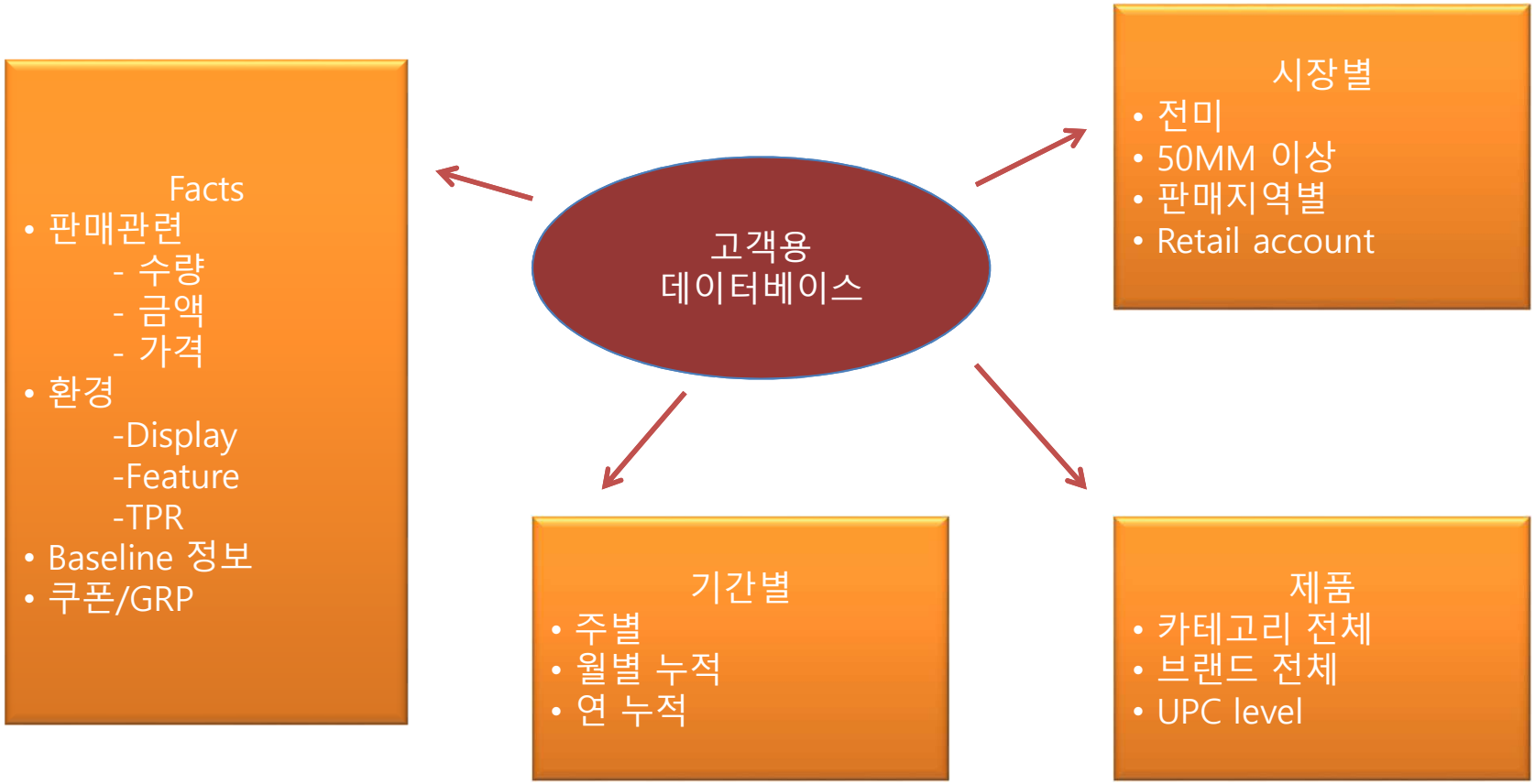
제품유형(축산품, 공산품 등)

  - 국가별로 번호 체계 다름

- UPC 정보
  - 제품의 standard characteristics: 모든 UPC에 대해 다음의 정보들이 자동적으로 입력
    - UPC 번호
    - 브랜드 이름
    - 제조업자 번호
    - Size and weight code
    - 멀티팩
  - Nonstandard characteristics들은 UPC 번호가 데이터베이스에 추가되기 전에 담당자가 직접 입력
    - 맛(flavor), form(액체, 가루), 타입

- Nielson에서의 UPC 코딩
  - 아이템 데이터베이스에 150만개 이상의 UPC와 각 UPC의 해당 정보 수록
  - 매주 3천개의 신규 아이템 등록

- 데이터베이스 내용


Facts
• 판매관련
- 수량
- 금액
- 가격
• 환경
-Display
-Feature
-TPR
• Baseline 정보
• 쿠폰/GRP
고객용
데이터베이스
시장별
• 전미
• 50MM 이상
• 판매지역별
• Retail account
기간별
• 주별
• 월별 누적
• 연 누적
제품
• 카테고리 전체
• 브랜드 전체
• UPC level

Need for Data-Driven Marketing

Data-Driven Changes in Standard Consumer Research

- Audit/share report: 자세한 시장 점유율 정보가 며칠 내로 획득
- 테스트 마케팅: 필드에서 테스트 마케팅 실험이 용이해짐(standard test marketing)
- 시장 세분화: 인구통계학적 세분화에서 행위 중심 세분화로 무게 중심 이동(ex. Kraft 치즈를 사는 사람은 어떤 사람이며 또한 그들은 어떤 제품을 주로 구매하는가?)
- 광고 테스트: "rented" 테스트 마켓이나 split cable을 통한 직접 실험 가능
- Demand for accountability: "pay for performance" 비효율적인 마케팅 비용 지출 감액

Mass vs. Micro-Marketing

- 자세한 개별 고객의 구매 행동 정보가 이용 가능해 짐에 따라 mass marketing에서 micro-marketing으로의 이동 가속화
- 실제 구매행동을 바탕으로 고객 파악 및 잠재 고객 발굴
- 개별 고객 특성에 맞춘 제품 차별화(보험 등의 금융 상품), 가격 차별화, 제품 구색 차별화
- 광고 메시지 차별화 (split cable), targeted banner advertising (Doubleclick)

Customer Data as the Source of Business Opportunities

- 고객 데이터베이스가 신규 사업의 기회 창출
- Bloomingdale's 백화점은 고객 정보를 기반으로 해서 카탈로그 사업 진출
- AT&T - 신용카드 사업 진출(billing database)
- 다양한 플랫폼 비즈니스: 넷플릭스, 카카오톡, 우버, 에어비앤비, …
- 개인 정보 보호 이슈

Data-driven Market Analytics 분야

- Pricing
- Promotion
- Micromarketing
- Targeting
- Direct Marketing
- …….

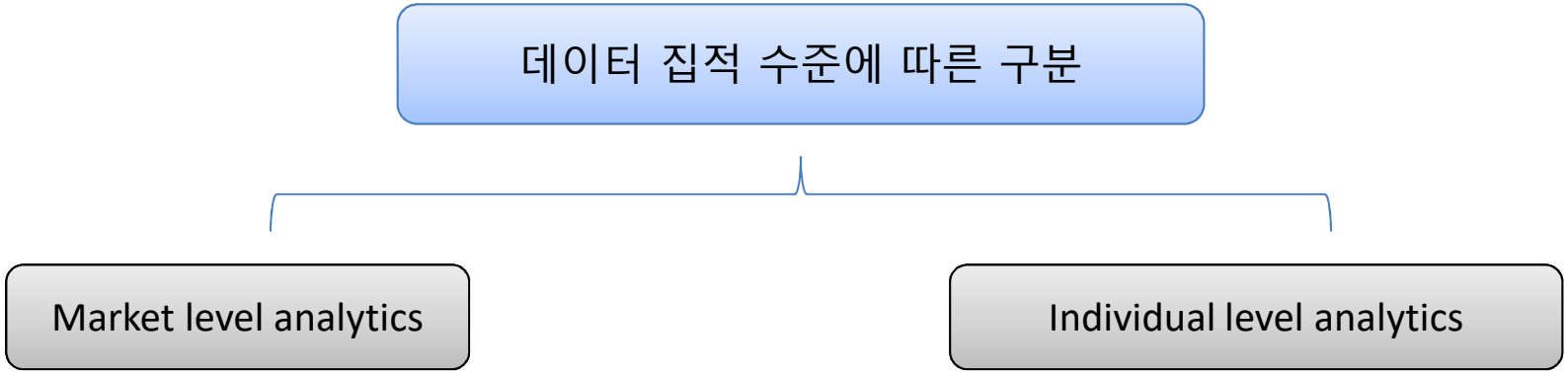
Why data-driven marketing analytics?

- 데이터가 비즈니스의 중심이 되는 시대 도래
  - 데이터베이스 그 자체가 경영 자산
  - 효과의 존재 (e.g. 가격이 오르면 수요가 감소) 를 넘어서는 효과의 크기 (e.g., 가격을 1% 올리면 수요가 x% 만큼 감소) 측정의 정확성 중요
  - Pay for performance

- 경영 직관이 도전받을 수 있음
  - 막연히 믿어왔던 신념들이 부정될 가능성
  - Let the data speak
  - 새로운 통찰의 기회로 이용

- Data-driven analytics에 대한 능동적 수용 필요

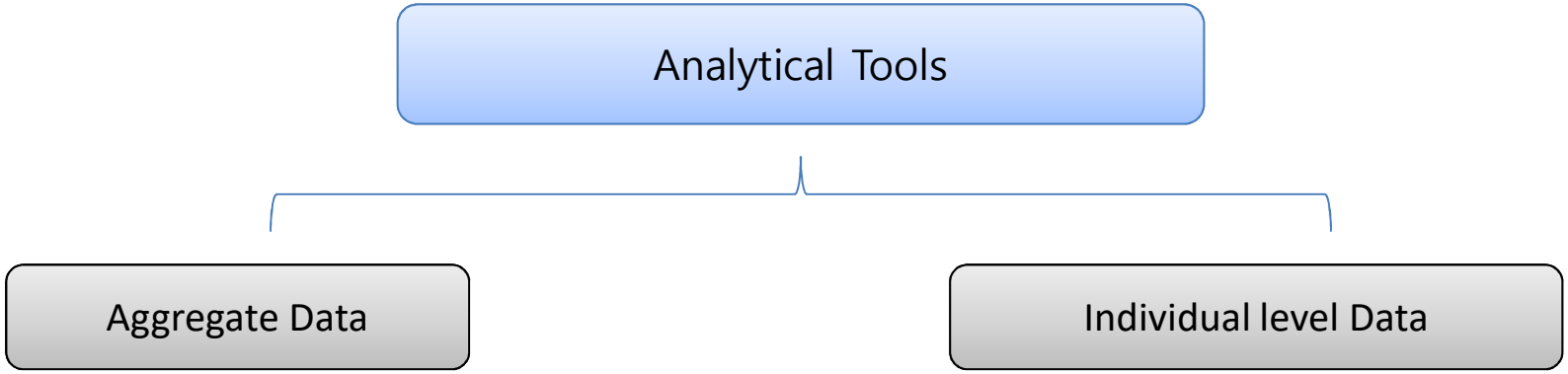
Roadmap, again



- Aggregate data
- 시장 수준의 마케팅 활동 효과 측정과 최적화
- Continuous variables: regression based model
- Pricing, promotion, advertising, positioning, demand forecasting, etc.

- Customer level data
- 고객 이질성 (customer heterogeneity) 초점, 고객별 상이한 반응 행태 파악 및 타겟팅
- Discrete variables: classification based model
- Customer life time value, customer, choice, targeting, etc.

Tools used in Marketing Analytics



- Regression / Model Selection
- Demand forecasting
- Bayesian hierarchical models
- Lift Modeling
- Clustering for segmentation
- Perceptual mapping for positioning
- …

- Conjoint analysis
- Choice modeling: Logit/Probit
- Classification: logistic, discriminant, tree, neural nets, SVM, etc
- Recommendation: collaborative filtering / market basket analysis
- Social Media and text: sentiment / topic modeling
- …

Objectives of Marketing Analytics

Model: $y = f(x, \theta)$

Descriptive analytics

- Interested in how X affect Y in terms of the functional forms and parameters
- Functional forms matter. We want to figure out the data generating process (relationship between X and Y) and sometimes want to test theories.
- Focus: $\theta, f(\cdot)$

Predictive analytics

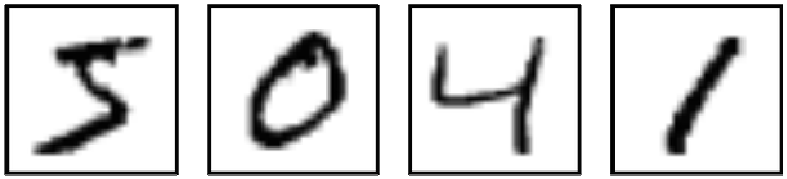
- Interested in predicting Y for a given value of X
- Not much interest in the nature of the relationship between X and Y
- Focus: $\hat{y}(x_0)$ for a given value x_0
- Typical machine learning applications are interested in prediction.

Marketing analytics use both descriptive and predictive analytics

Marketing Analytics vs. Machine Learning

Typical marketing analytics projects suffer from data problems, unlike machine learning projects.

Example: MNIST project (recognizing handwritten digits)

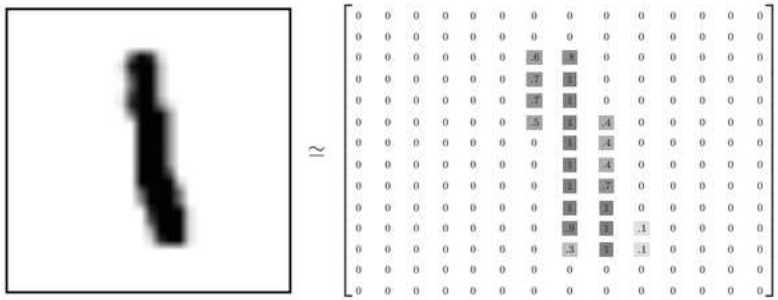


Each image (handwritten digit) is represent by 784(=28*28) pixels. So we have 784 feature variables $x_1, x_2, \cdots, x_{784}$.

Each pixel is 8-bit gray scale (0,1,…,255) but normalized to a number between 0 and 1 (or 0/255, 1/255, … ,255/255).

Each image has label in the training data (i.e, $y = 5,0,4,1, \cdots$). So each y takes a value out of 10 labels (0,1,…,9).

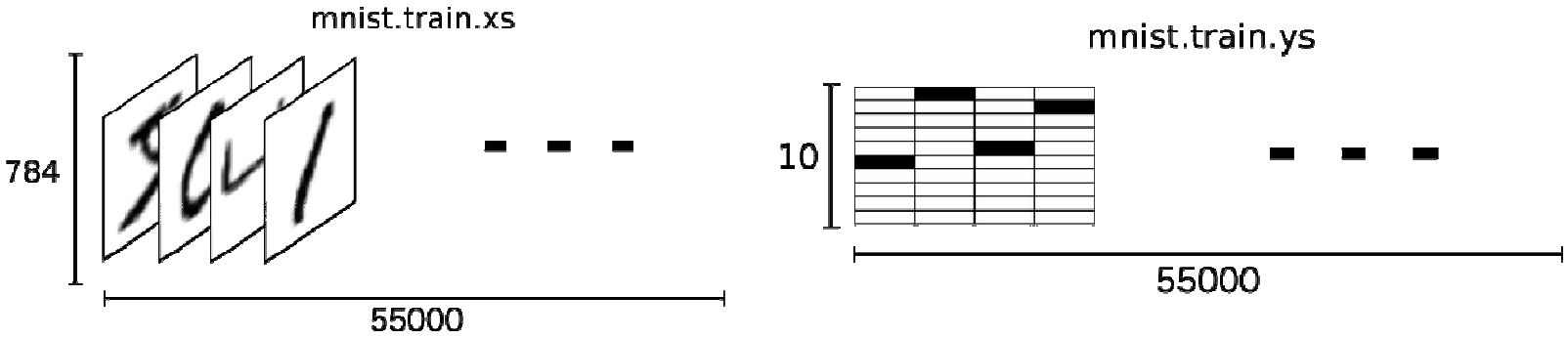
An image looks like

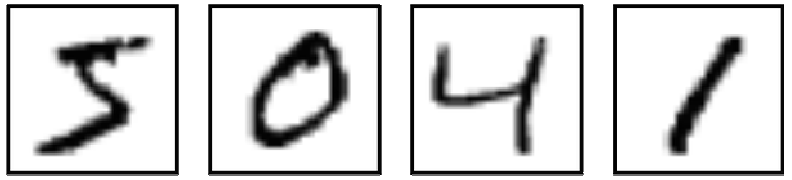


If we have 55000 images, the data set look like

$$\begin{matrix} y_1 & x_{1,1}, \cdots x_{784,1} \\ \vdots & \vdots \\ y_n & x_{n,1}, \cdots, x_{784,n} \end{matrix}$$

When n=55000, the feature (X) and label (Y) look like





Machine learning

- You fit you model using training data and apply the fitting result to predict the value of Y (label) for a new image. So in this case, the relationship between Y and a particular X variable is not a ,major interest.
- The point here that you have all pixels in your data. That is, the 784 pixels are sufficient to predict the label of each image.

Marketing analytics

- Consider a marketing analytics project where you want to predict a customer churn based on customers characteristics such as age, income, gender, prior purchase amount, etc.
- The fundamental problem here is that the features (age, income, etc) may not be complete. The worse thing here is that we do not know whether our feature data are complete. Marketing deals with customers (people) whose behaviors are probably affected by many factors. Much of them are not recorded in our database. Marketing analytics would utilize those (incomplete) data, unlike in machine learning.
- To overcome the incomplete data problem, marketers need to rely on some external information, “marketing theory”.