

데이터분석 환경의 변화

1990년대 데이터 분석(혹은 데이터 사이언스)는 주로 데이터베이스 마케팅에 분야에 활용이 되었습니다. 데이터베이스 마케팅이란 고객과 회사와 접점에서 발생하는 데이터를 데이터베이스화하고 마케팅에 활용하자는 아이디어입니다. 여러번 구매할 수 있는 상품의 경우 데이터베이스 마케팅이 효과가 좋습니다. 주로 트랜잭션(구매이력)을 RFM (Recently, Frequency and Monetary) 등의 인자로 고객을 세분화하고 세분된 고객군별로 별도의 마케팅을 함으로써 더 효과적인 캠페인을 할 수 있었습니다.

2000년 초반에는 CRM(Customer Relationship Management)이 업계의 화두였습니다. 특히 데이터베이스마케팅을 할 수 있는 소프트웨어가 인기가 좋았습니다. 오라클(Oracle), 시벨(Siebel) 등은 CRM 소프트웨어와 컨설팅을 경쟁적으로 시장에 팔았습니다. CRM은 “신규고객 1명 획득에 따른 비용 대비 수익보다, 기존 고객을 유지하는 비용 대비 수익이 훨씬 좋다”라는 기본 철학을 바탕으로 합니다. 하지만 고객 분석과 전략보다는 벤더의 소프트웨어 판매에 치중이 되다보니, 효과를 입증하기 어려웠고 무용론이 대두되었습니다.

또한 데이터분석은 개인 신용평가에도 활용이 많이 되었습니다. 과거 IMF (1997년) 이전에는 대출 신용한도나 신용카드 한도를 대출 담당직원이 심도있는 데이터 분석 없이 결정했습니다. 데이터에 근거하지 않고, 과거 경험으로 결정했습니다. IMF 이후, 많은 신용불량자가 생성되자 은행에서는 연체관리의 중요성을 인식하기 시작했습니다. 데이터에 근거한 통계적인 접근을 시작했습니다. 예를 들어, “이 고객은 통계적으로 연체할 확률이 70% 이므로, 예상되는 손실이 오백만원이 될때까지 신용한도를 낮춰야한다. 이런 식의 접근입니다”

데이터 분석은 회사에서 중요한 업무를 담당했지만, 대우를 잘 받지 못 했습니다. 똑똑한 신입이나 대리급에서 하는 일이라고 생각하는 임원이 대부분이었습니다. 따라서 지원부서의 역할이 강했고 승진과 성과급은 영업부서의 독차지였죠.

하지만, 시대가 변했습니다. 인터넷/모바일의 발달로 온라인 시장이 급격히 커졌습니다. 온라인은 영업부서의 역할을 축소시켰습니다. 은행은 점포보다는 모바일 채널을 통하여 고객에게 더 저렴하고 다양한 서비스를 제공할 수 있게 되었습니다. 딥러닝 기술의 발달로 사람의 판단이 필요없는 단순한 업무는 자동화가 가능해졌습니다. 즉 분석된 결과가 비용절감과 영업의 결과로 곧바로 연결이 되는 시대가 되어 가고 있습니다. 구글은 유튜브 광고를 통하여 많은 수익을 올리고 있습니다. 데이터 분석(추천 알고리즘)이 영업결과가 되는 대표적인 케이스라고 말씀드릴 수 있습니다.

제조업도 센서기술과 빅데이터 저장 기술의 발달로 데이터분석의 혜택을 보기 시작했습니다. 전체 프로세스에서 발생하는 데이터를 연결하여 불량 원인분석, 비용절감 및 생산증대에 데이터를 활용하고 있습니다. 특히, 데이터를 활용하여 최적으로 기계를 제어하는 부분까지 발전하고 있습니다.

분석 툴과 용어도 많은 변화가 있습니다. 과거에는 SAS라는 통계분석툴을 주로 사용했습니다. 개인적으로도 SAS를 이용하여 20년이상 일을 했습니다. 아침에 출근하면 왼쪽 모니터에는 엑셀, 오른쪽 모니터에는 SAS가 있었습니다. 개인적으로 애착이 깊은 소프트웨어입니다. 요즘에는 파이썬이 대세입니다. 두가지 툴 사이에 큰 차이점은 SAS는 분석에 특화되어 있지만, 파이썬은 분석뿐 아니라 자동화와 배포(예를 들면 웹서비스)까지 모든 서비스를 가능하게 합니다. 따라서 분석의 결과를 구현해서 성과로 보여주고 싶다면 파이썬이 좋은 툴입니다.

데이터분석이라고 부르던 업무 영역들이 확장되어 이제는 데이터사이언스라고 불리고 있습니다. 일부는 데이터분석가와 데이터사이언티스트를 구분하고 업무영역을 다르게 보는 시각이 있으나, 제 경험으로는 차이가 크게 없습니다. 요즘 데이터사이언티스트를 하시는 분들은 컴퓨터 혹은 소프트웨어 전공자 분들이 많으셔서 소프트웨어 개발도 잘 하십니다. 분석결과가 실행되어 결과가 되는 속도가 더욱 가속되고 있습니다.

