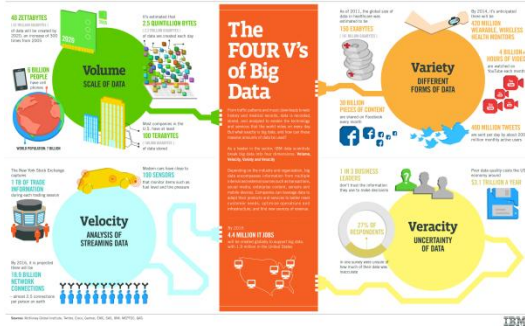


빅데이터란 무엇인가?

What is Big Data?



IBM에서 말하는 4V

Volume: 데이터의 양과 스케일

Velocity: 데이터의 속도 그리고 분석

Variety: 여러 다른 데이터의 다양성

Veracity: 데이터의 불확실성

*한가지 V를 더 추가 하자면 **value** 또한 추가 될 수 있지 않나 싶다.

빅데이터 역시 value 즉 사용할 수 있고 적용 할 수 있는 가치를 창출해낼 수 있어야 한다.



나의 개인적인 생각: 시대의 흐름이 빠르게 변화되고 있다. 2013년에 IBM의 정의 그리고 지금의 정의는 몇가지 추가 된것이 있다. 하지만 변화된것이 아니라 같은 근본에서 추가 된것이기 때문에, 과거의 생각 그리고 오늘날의 생각들을 같이 고려하면서 이책을 읽는 것이 좋다고 생각한다.

시대별 IT 및 비즈니스 변화

책 초반부에는 IT의 트렌드와 디지털이징 비즈니스에 대한 설명이 이루어 지고 있다.

Centralizing Business:	IBM
Downsizing:	Oracle, Cisco, HP, Microsoft, Intel
eBusiness:	Oracle, SAP
Digitizing Business:	Google, Apple, Amazon, Salesforce, Samsung

나의 개인적인 생각: IT기업과 IT 트렌드, 몇가지를 나열해 보면: 클라우드 컴퓨터, 모바일, IOT, 빅데이터, 소셜미디어 등등 길게 나열할 수 있다. 게다가 새로운 혁신을 통해 나날이 새로운 IT 영역이 탄생하고 있다. 시간이 갈수록 경쟁 업체들은 하나 둘씩 경쟁에서 밀리기 시작하며 새로등장하는 기술 또는 업체들이 나타난다. 미래로 갈수록 누가 first mover가 되고 누가 혁신을 꾸준히 하며, 누가 가장 적응을 잘하는지에 따라 한 회사의 생명줄이 결정된다고 생각한다.

다양한 빅데이터 전문가

Different Big Data Specialist

빅데이터 IT 전문가:

빅데이터 분석 전문가:

빅데이터 도메인 전문가:

빅데이터 전략 전문가:

*빅데이터의 프로세스 체인을 산업에 대한 전문 지식, 데이터 처리 분석 모델링 결과 해석 실행 등으로 생각할 수 있다.
기업에서 빅데이터의 방향과 전략을 수립하기 위해서는 산업에 대한 이해와 업무 실행에 대한 이해가 반드시 필요하다.

빅데이터 테마 결정에 필요한 정보

산업 분류: 산업에 따라 구현 방법 및 기술, 분석 방법이 달라짐

업무 분야: 고객, 마케팅, 품질, 위기관리(Risk Management), 고객의 소리, 사기방지 등과 같은 업무 영역

분석 기법:

- 서술 분석(Descriptive Analysis): 과거의 일 분석 SQL 사용
 - AS-IS 분석 : SQL
- 진단 분석(Diagnostic Analysis): 원인을 찾는 분석 (왜 이런 현상이 일어났는가?)
 - 근본 원인 분석 : 통계
- 예측 분석(Predictive Analysis): 앞으로 일어날 일들 예측(Forecasting)
 - 예측 : 통계 + Business Logic
- 처방 분석(Prescriptive Analysis): 앞으로 이런일이 일어나려면 어떻게 해야한다 를 분석. 분야의 관한 전문성 필요
 - 시뮬레이션: 통계 + Business Logic
- 최적화(Optimization): 이런 일이 일어나면 어떻게 업무 조정해서 대처해야 하는가를 알아내고 전략을 세우는 것. 컨설팅 마인드 비즈니스 경험 지식 필요.
 - 통계 + Business Logic + 업무 처리 프로세스

책을 읽은 후 느낀 점...

책을 통해 제목 그대로 빅데이터로 일하는 기술에 관한 인사이트(insight)을 많이 얻게 되었다. 사실 십 년 전부터 빅데이터라는 소리는 떠돌아 다녔고 지금도 많은 사람들이 들으면서 살고 있다. 그러나 그것이 무엇인지 정확히 알고 있는 사람이 많지 않았고 IT관련 전공을 했던 나도 정확히 무엇이다 정의하기 어려웠다. 하지만 이 책을 통해 이해 감이 생기고 활용방법에 대한 생각을 많이 하게 된 계기가 된 것 같다. 현업에서도 빅데이터를 정확히 활용하고 사용하는 사람들은 많지 않다 또한 지금 현재 100%활용 한다고 생각하는 사람들조차 몇 년 후면 다시 공부해야 하는 것이 오고 또한 innovate 혁명 해야 하는 것이 온다. 이를 대비해서 구축해 놓 데이터를 통해 끊임없는 연구를 하는 것이 옳다고 본다.

한가지 아쉬운 점은 책의 내용의 2014년 데이터로 한정적이라는 것이다. 미래 전망을 예측한 부분이 맞아 떨어진 부분도 있고 빗나간 부분도 있다. 책을 읽으면서 내내 지금 데이터로 2015년-2019년 데이터도 들어 갔으면 좋았을 텐데 라는 생각을 많이 하였다. 모든 사람들이 아마 느낄 것이다 지금 세상은 빠르게 돌아가고 있다. 매일 같이 새로운 기술들이 도입이 되고 있고 혁신을 하려고 노력한다. 이 책을 읽으면서 insight이 많이 생긴 것은 사실이다. 빅데이터의 실무 경험이 전혀 없는 나로서는 일반 회사의 시각에서 바라보는 빅데이터는 어떤 것인지, 또한 어떤 부분이 문제점이고 그것의 해결방안은 무엇인지 또한 R의 한계점등 전에는 몰랐던 지식을 습득하고 배웠다. 기업이 빅데이터를 도입하는 목표는 데이터에 의한 의사 결정을 하기 위해서다. 한마디로 problem solving method으로 바라보고 있다.

초반에는 많은 사람들이 wearable technology에 관심을 가졌다. 지금의 테크 회사들은 끊임없는 연구를 하는 것도 사실이다. 하지만 Google Glasses, Snapchat Spectacles, Apple/Samsung watch 등 웨어러블 테크는 사실상 크게 성공을 거두진 못했다. 이는 여러가지 이유가 있는데, 물론 아직 미완성 단계인 기술로 볼수 있지만 사람들의 심리라는 것이 큰 역할을 한다고 본다. 이러듯이 미래를 예측하는 것이 여러가지 missing factors 생각하지 못한 변수들로 크게 빗나갈수 있다는 것이다. 그러므로 데이터를 모을수록 창의성도 큰 역할을 한다고 생각한다.

데이터 분석 기법에 따를 비즈니스 가치 및 필요 인력에서는 IT 전문가 통계 전문가 그리고 비즈니스 전문가들이 각각 필요하다. 기업의 입장에서 이러한 기술들을 모두 가지고 있는 인재들을 찾을 것이고 (사실 오늘날 컴퓨터 못 다루는 사람은 없다) 전문성의 깊이 그리고 경험에서 판가름이 날것이다. 이러한 매력적인 인재로 성장을 하기 위해서는 끊임 없는 배움, 사람들을 상대하는 기술 (발표 기술, 말하기 기술, 컨설팅 마인드) 많은 노력이 필요할 것 이다.