**Introduction**

**Bad Graph는 어디에나 있다.**

아무도 Bad Graph를 만들진 않지만 계속 생긴다. 더 잘 알고 싶어할 때 그런 일이 발생한다.

**우리는 원래 데이터와 함께 스토리텔링을 하는 데에 능하지 않다.**

수학을 배울 때 숫자에 대한 느낌을 배우지 이야기를 배우지 않는다.

데이터를 시각화하고 이야기를 할 수 있는 것은 더 나은 의사결정을 하게하는 정보의 열쇠이다.

데이터에는 이야기가 있다. 하지만 우리의 도구들은 그 이야기가 무엇인지 모른다.

이 책이 이야기 만드는 법을 알려줌.

**누구를 위한거 ?**

데이터를 이용하여 효과적으로 커뮤니케이션하는 능력을 개선시키고 싶은 사람이면 아무나.

효과적인 데이터 시각화는 우리의 연구에서 발견한 것을 소통 하는 것, 비영리에서 돈을 버는 것, 우리의 칠판에 표현하는 것, 혹은 간단히 우리의 요점을 관객들에게 전달 하는 것에 대하여 성공과 실패를 의미한다.

작자의 경험상 사람들은 비슷한 어려움에 직면했다. 바로 사람들은 시각화를 하고 싶지만 그 시각화를 하는 데에 있어서의 전문성이 부족하다고 생각한 것이다. 데이터 시각화를 하는 사람들을 찾는 데에는 어려움이 있다. 어려운 부분 중 하나는 데이터 시각화가 분석 과정 중 하나의 단계 라는 것이다. 이러한 분석적인 역할에 고용된 사람들은 대개 다른 과정들에 더 적합한 양적인 배경을 가지고 있지만, 분석(대개 당신의 관중이 보는 분석적인 과정 중 일부분이다.)에 대하여 의사소통을 하는 데에 있어 그들의 이해를 돕게 디자인 하는 것을 배우는 형식적인 과정이 꼭 필요한 건 아니다. 점점 더 데이터 발생이 많아지고 있는 세계에서, 이러한 기술적인 배경지식 없이 분석자의 위치에 간다거나 데이터를 이용한 의사소통을 한다는 것은 물음을 받고 있는 중이다.

이 분야에서의 경험하는 것이 불편한 것은 놀랄 일이 아니다. 데이터를 이용해 효과적으로 의사소통을 가능하게 하는 것이 전통적으로 배워왔던 무언가가 아니다. 뛰어난 사람들은 일반적으로 시행착오를 통해 무엇이 작동하고 안하는지를 배워왔다. 이것은 지루하고 긴 과정 일지도 모르지만, 이 책을 통하여 나는 그것을 더 신속히 처리하는데 도움을 주고 싶다.

**어떻게 데이터와 이야기하는 것을 배웠는가?**

나는 항상 수학과 비즈니스가 교차하는 공간을 그려왔다. 나의 교육적인 배경은 양쪽 모두에게서 효과적으로 의사소통을 가능하게 하고 서로를 이해하는데 돕는 수학과 비즈니스이다. 나는 데이터에 대한 과학을 이용해 더 나은 사업상의 결정을 알리는 사용되는 것을 좋아한다. 시간이 흘러 나는 데이터를 시각적이면서 효과적으로 소통 가능하게하는 성공의 열쇠를 찾아왔다.

나는 원래 첫 직업을 가졌을 때 숙련된 스킬을 가지는 것이 중요하다고 생각했었다. 나는 신용 위험 관리의 분석가로 일하고 있는 중이었다. 나의 일은 비행과 손실을 예측하는 통계적인 모델을 세우고 평가하는 것이었다. 이게 의미하는 것은 복잡한 사물들을 받아서 손실을 무마할 충분한 돈이 있는지, 어떤 시나리오에 우리가 위험에 노출됐는지 등등에 대한 것을 궁극적으로 간단하게 커뮤니케이션 할 수 있도록 바꿔야 하는 것이다. 나는 미적인 작품에 시간을 소비하여 빨리 배웠고, 나의 작업은 나의 상사와 상사의 상상사로부터 많은 관심을 얻었다. 나에게는, 그것이 데이터에 대한 시각적 의사소통에 시간을 소비하여 가치를 보는 것에 대한 시작이었다.

많은 일을 거쳐 나는 은행 업무와 재정 업무 바깥에서 커리어를 계속 하고 싶어졌다. 다른 기술들 다 제쳐두고, 데이터를 사업상의 결정에 영향을 주는 데에 사용하고 하는 것을 하고 싶었다.

구글에 있는 the People Analytics Team에 들어갔다. 구글에 있는 모든 서비스에서 롤 아웃 된 데이터의 시각화 과정을 개발했다.

이러한 코스의 개설을 구글 내외적으로 잡음이 많았다. 그러다가 구글을 나왔다.

지난 몇 년 동안 나는 미국과 유럽에서 100개가 넘는 기관에서 강연을 했다. 그것 보기 흥미로워

**The importance of context**

직관에 어긋날 수 있지만, 데이터 시각화의 성공은 데이터 시각화로 시작하지 않는다. 어떤 경로로 만들지 정하기 전에, 의사소통에 필요한 **문맥**을 이해하는데 시간과 관심을 기울여야 한다.

**Exploratory vs. explanatory analysis**

문맥에 대해 구체적으로 들어가기 전에 exploratory와 explanatory 분석의 대한 중요한 차이를 짚고 넘어가야 한다.   
 Exploratory analysis는 당신이 데이터를 이해하기 위해 하는 것이고, 다른 것들을 강조할 주목할만한 혹은 흥미로운 것이 무언지 이해하는 것이다. 이것은 마치 굴 속의 진주를 찾는 것과 같다. 우리는 아마 2개의 진주를 찾기 위해 100개의 굴(100가지 다른 가설에 대한 시험이나 100가지 다른 방법에 의해 데이터를 보는 것)을 열어볼 것이다. 우리의 분석을 청중에게 알려줘야 하는 시점이 오면, 우리는 explanatory 단계로 가기를 원할 것이다. 이 explanatory 단계 라는 것은 내가 설명할 것에 대한 구체적인 것을 가지고 있어야 한다는 것과 이 2개의 진주에 대해 내가 말하고자 하는 구체적인 이야기를 의미한다.

너무나 자주, 사람들은 explanatory를 보여줘야할 때 exploratory 분석을 보여줘도 괜찮다 라는 실수를 범하고 생각한다 이것은 이해할 수 있는 실수이다. 전체적인 분석을 착수하고 나서, 분석결과에 대해 당신의 잘 해냈다 라는 것에 대한 증거로 모든 것을 청중들에게 보여주고 싶어하는 유혹을 받을 수도 있다. 이런 유혹을 뿌리치기를 권장한다. 왜냐하면 당신의 청중들은 다시 100개의 굴을 열어봐야 할 수도 있기 때문이다. 청중들이 알고 싶어하는 것은 진주지 굴이 아니다.

여기에서, 우리는 explanatory 분석과 의사소통에 대해 초점을 맞춘다.

**Who, what and how**

Explanatory 분석에 대해서, 무언가에 대해 생각하고, 컨텐츠를 만드는 것이나 데이터를 시각화 하기 전에 최종적으로 분명하게 해야하는 몇가지가 있다. 첫번째로, ***누구와 소통하는가?*** 이는 당신의 청중이 누구이고 어떻게 그들이 당신을 인식할 것인지에 대한 좋은 이해를 위해 중요하다. 이것은 당신의 청중이 당신의 메시지를 확실히 받아들였는지 확인하는 공통점에 대해 도움을 줄 것이다. 두번째로, ***당신은 당신의 청중이 무엇을 알았으면 좋겠고 무엇을 하고 싶어하는지 알아야한다.*** 당신은 당신의 청중이 어떻게 행동하기를 바라는지 명확하게 하고, 어떻게 그들과 의사소통 할 것이고 그들과 의사소통 할 때 쓰길 원하는 전반적인 톤을 계산에 넣어야 한다.

세번째 질문으로 넘어가기 위해서 앞선 두 질문에 대한 간결한 답변이 필요하다: ***어떻게 당신의 요점을 만드는데 데이터를 사용할 것인가?***

Who, What, and how의 문맥에서 더 자세하게 보자.

# Who

청중에 대한 폭을 좁혀야 한다. 아무나 하고는 안된다.  
당신을 어떻게 받아들이게 할 것 인지에 대한 것도 생각해봐야 한다.

# What

간단히 데이터에 대한 설명을 묻는 질문이 아니라 청중의 액션을 이끌어내는 질문을 해야한다. 더 생산적인 반응이 더 생산적인 대화로 이어지기 때문이다.

3분 이야기

만약에 이야기 할 수 있는 시간이 3분밖에 없다면 어떻게 말할 것인가? 이 방법은 어떤 것을 말할 것인지 명확하게 생각할 수 있게 해준다. 이렇게 생각이 명확하면 당신이 준비되지 않았더라도 시간 내에 잘 마무리 지을 수 있다.

큰 생각

하나의 문장이다. Nancy Duarte 라는 사람이 말했는데 빅 아이디어는 3가지로 나뉜다고 한다.   
1. 당신만의 독특한 관점을 분명하게 표현하라.  
2. 지금 중요한 것을 잘 전달하라.  
3. 완벽한 문장 이어야 한다.

스토리보딩

의사소통에 대한 구조를 확립하게함.  
소프트웨어를 이용한 프레젠테이션으로 시작하지마라. 생각을 하지 않게 하고, 많은 양의 프레젠테이션은 집중도를 떨어트린다. 따라서, 시작할 때는 화이트보드와 포스트잇과 순수한종이를 사용해라.

Chapter2

효과적인 비주얼 선택

엄청나게 다양한 그래프 종류가 있는데 자기가 사용하는건 12개뿐. 이거 집중할게.

12개 -> Simple text, Scatterplot, Table, Line, Heatmap, Slopegraph, Vertical bar, Horizontal bar, Stacked vertical bar, Stacked horizontal bar, Waterfall, Square area.

1. Simple Text: 한 개나 두가지의 정보를 전달해야할 때 좋다.
2. Tables: 테이블은 verbal system(언어구조)에 적합하다. 그들이 관심있는 분야가 다 다르고 섞여있는 청중들에게 효과적이다. 어두운 계열의 색 보다는 밝은 계열의 색이 좋다. 보더는 당신의 테이블에 대한 가독성을 높이는데 사용되어야 한다.
3. Heatmap: 테이블의 형식에서 수의 규모를 표현할 때 색의 정도로 표현할 수 있다. 여러 수치가 분포되어있을 때 우리는 일일이 다 확인해서 어느것이 높고낮은지 직접 봐야하는데 그 과정을 줄여준다. Color saturation 즉, 색깔 포화도로 그 과정을 스킵할 수 있다.
4. Graphs: 그래프는 시각 구조에 적합하다. 정보를 빠르게 전달해야할 때. 시각적으로 빠르게 전달하려면, 잘 정리된 테이블이보다는 잘 디자인된 그래프가 낫다. 많은 그래프 종류가 있지만, 나는 points, lines, bars, and areas를 쓴다.
5. Points:   
   -Scatterplot: 두 가지의 관계를 보여줄 때 유용하다.
6. Lines:
7. Bar 차트에서 정확성을 보여주기 위해선 제로베이스라인이 있어야한다. 왜냐하면 우리는 바를 볼 때 끝선을 보기때문이다. 바와 바 사이의 간격은 내가봤을 때 바 자체의 두께의 2분의 1정도면되고 배경은 흰색

파이차트, 도넛차트, 3D, secondary y-axes는 피해야한다.

1. 파이차트가 안 좋은 이유는 한 번에 보기 힘들다. Horizontal bar chart로 대체하는게 좋다
2. 도넛도 마찬가지
3. 3D차트는 절대 쓰지 않는게 좋다. 왜냐하면 대략적인 값을 측정하기 어렵기때문이다
4. y축을 두 번 쓰는 것은 추천하고 싶지 않은 방법

Clutter: 뭔가 이미지나 설명은 엄청 많지만 이해에는 도움이 안되는 것들

-> Cognitive Load: 인지 부하.

게슈탈트 요인: 6가지가 있는데.

Proximity(가까움), Similarity(유사성), enclosure, closure, continuity, connection

Proximity: 우리는 물리적으로 가까운 것끼리 하나의 그룹으로 여기는 경향이 있다.

Similarity: 같은 모양, 색깔, 크기에 따라 하나의 그룹으로 여기는 경향이 있다.

Enclosure: 우리는 물리적으로 enclose하면 하나의 그룹으로 생각한다.

Closure: 사람들은 일반적으로 자신이 알고 있던 생각에 맞춰서 생각하거나, 간단하게 하려한다. 처음에는 크게 간단하게 인지하려하고 나중에는 더 자세하게 보이게된다.

Continuity: 사람들은 가장 부드러운 경로를 찾고, 자신이 보는 것이 존재하지 않더라도 논리적 연속성을 만들어 낸다.

Connection: 주변에 모양이 같고 색깔이 같고 크기가 같은 것이 있다고 하더라도, 선으로 연결되어 있으면 그것이 한 그룹인 줄 안다.

게슈탈트 인지론이 사람들이 어떻게 무엇을 보는지 파악할 수 있게 하고, 필요 없는 요소들의 identify를 피하고, 비쥬얼 커뮤니케이션을 쉽게 하게 하기 위해 하는 것을 돕는다.

Lack of Visual Order

대각선 글씨 읽는데 오래걸려, 세로로 적히면 더 오래 걸려

흰색 공간을 무서워해.