



# 한양대 경상대학

## 취업특강

마이티하이프 한명주. June 2021



**HELLO!**

**I AM ISABEL HAN (한명주)**

**MightyHive(S4 Capital), 상무 Director of Analytics**

**연세대학교 글로벌 인재대학 겸임교수**


## 오늘의 주제

- **우리회사 이야기**
- **데이터 리터러시 이야기**
- **데이터와 직업 이야기**

# 우리 회사 이야기



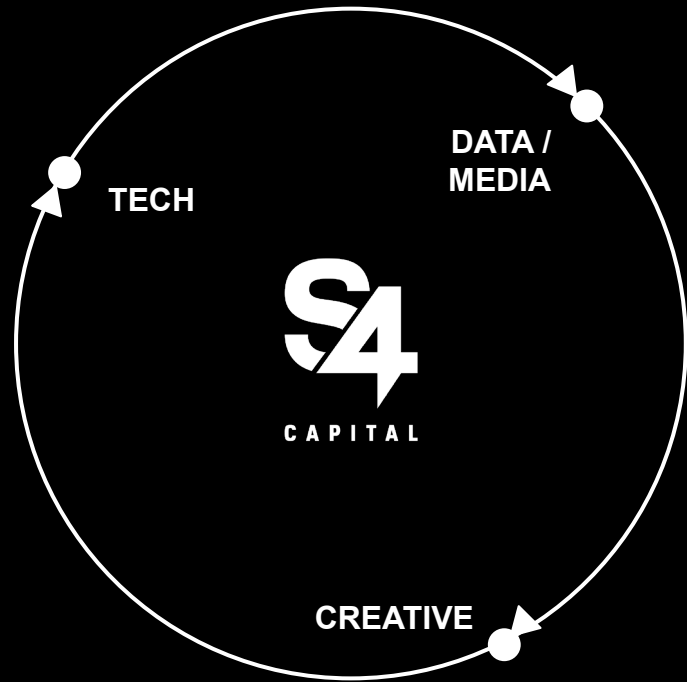




# HELLO 한양대학교, WE'RE S4

© 2020 S4 Capital. All rights reserved. Any copying or use of this confidential information is strictly prohibited without the express written permission of S4 Capital.





" — **Traditional agencies are out of touch with today's marketing needs.**

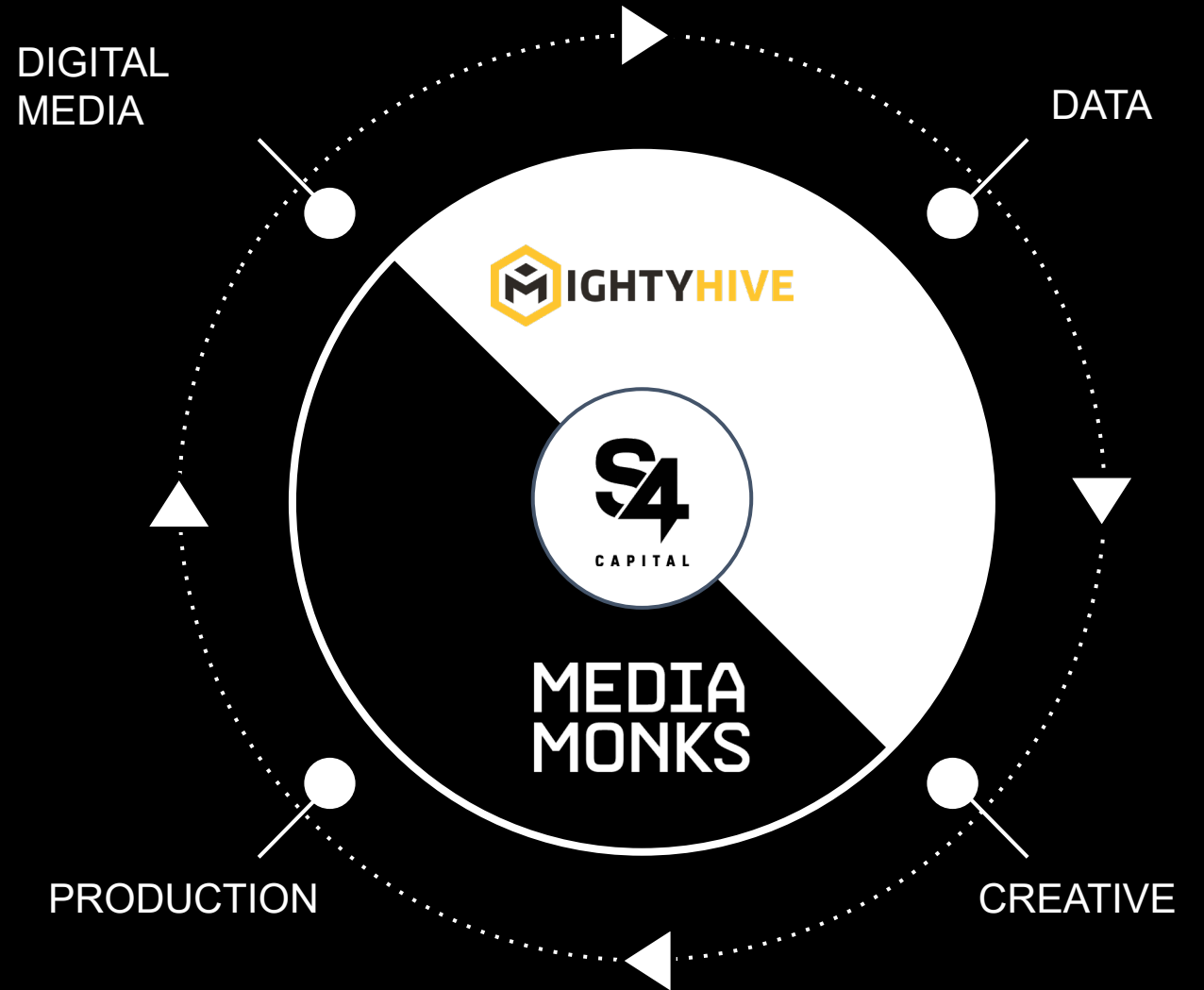
Sir Martin Sorrell

# new era new age new model

S4 is a new era, new marketing services organization embracing **data**, **content**, and **technology** in an **always on** environment for global, multi-national, regional and local clients and for millennial-driven digital brands.

# SPEED. QUALITY. VALUE.

End-to-end solutions for efficient,  
always-on, hyper-relevant and  
personalized strategy that turns  
insights into results







# Marketers are in the Middle of the **Big Bang of Data**

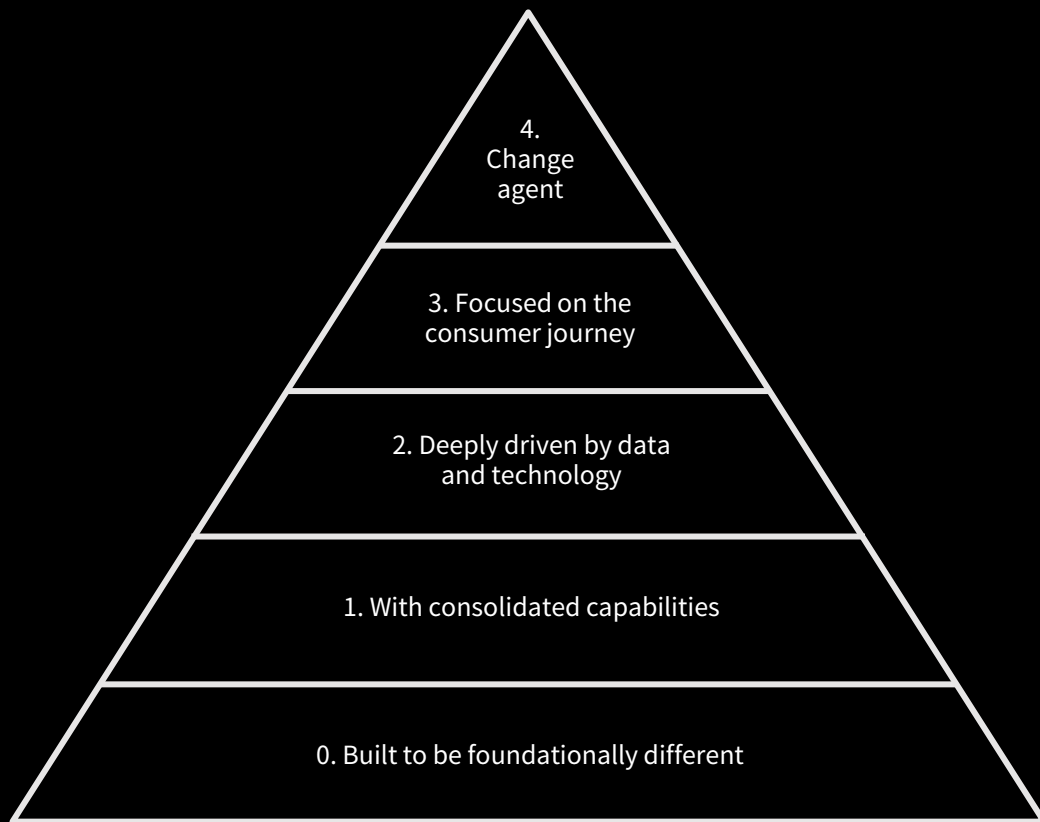
Marketers spent over \$15B on marketing data in 2019, but only 12% think they have “fully-functioning and effective data marketing strategies”







CAPITAL



4.  
Change  
agent

3. Focused on the  
consumer journey

2. Deeply driven by data  
and technology

1. With consolidated capabilities

0. Built to be foundationally different

# Built digital & different

**A marketing services partner that's built to be foundationally different in the way we help brands innovative their marketing and media and digitally transform their way of working.**

**MEDIA  
MONKS**

# **Most Awarded Digital Production Company in the World**

In a fast-changing business environment brands need to move faster than ever. They need smarter, integrated solutions. Delivered with speed, scale & quality. Our solution is simple and singular. One company creating work above the line and below the fold, across all media.



Genesis.com  
Brand.com and content



Victoria  
Integrated film and VFX production



RedBull  
Snapchat digital activation



## The Data & Digital Media Consultancy that Helps Marketers Take Control

MightyHive helps brands bring together their data, find the signal in the noise, and activate on those insights with speed, efficiency, and results.

**DIGIDAY**

Inside MightyHive,  
the agency helping  
brands create in-  
house agencies

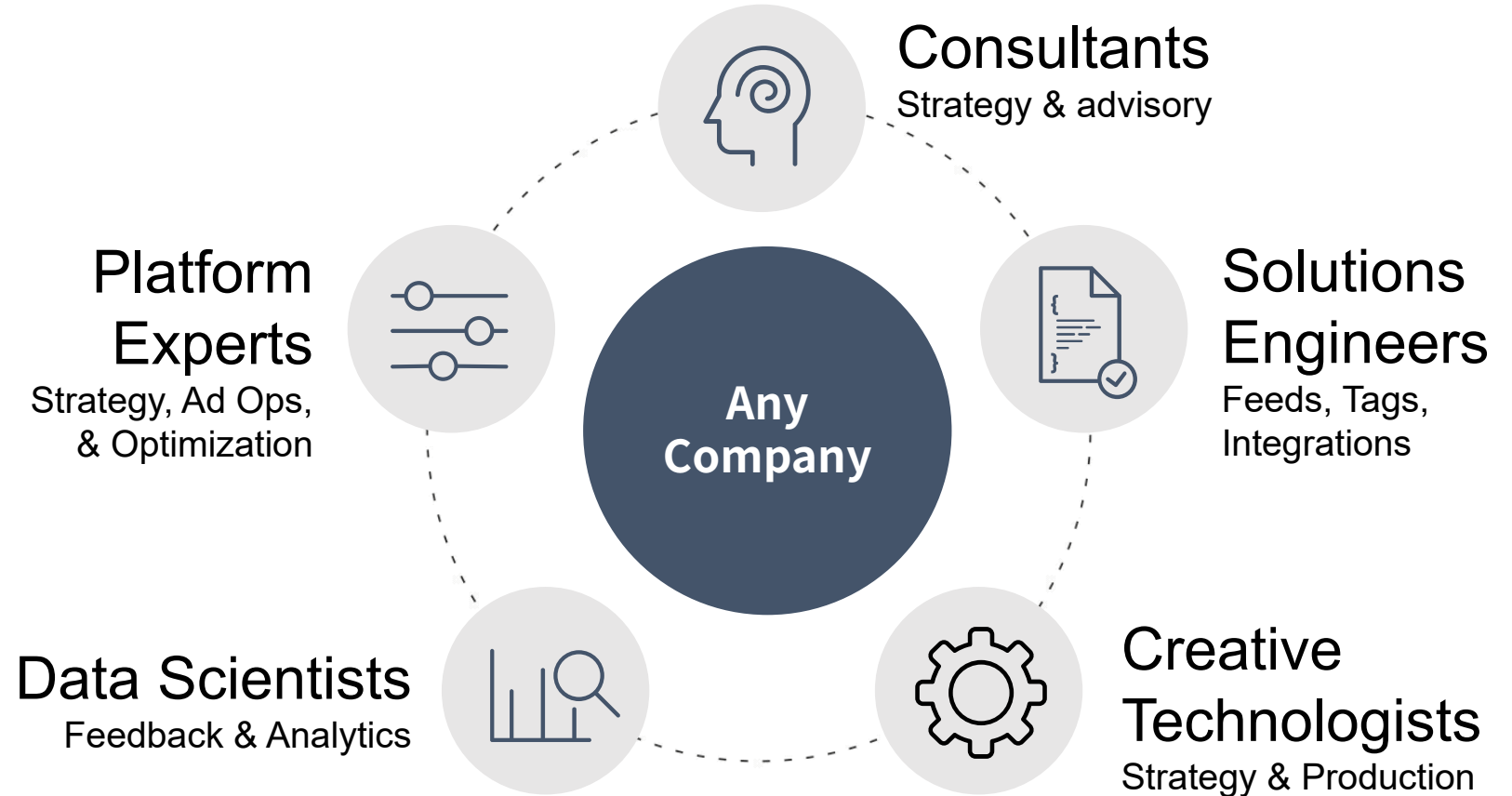


2019 Best  
Programmatic  
Consultancy - Finalist



Why We Are Different

# Cross-Domain Talent Integrated into One Team





Why We Are Different



CAPITAL



MEDIA  
MONKS

# Global Expertise & Partnerships

S4Capital Stats



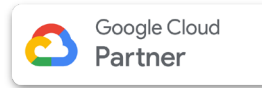
30

Countries



2,700+

Employees



## NORTH AMERICA

Canada  
United States

## LATIN AMERICA

Argentina  
Brazil  
Colombia  
Costa Rica  
Chile  
Mexico

## EUROPE & AFRICA

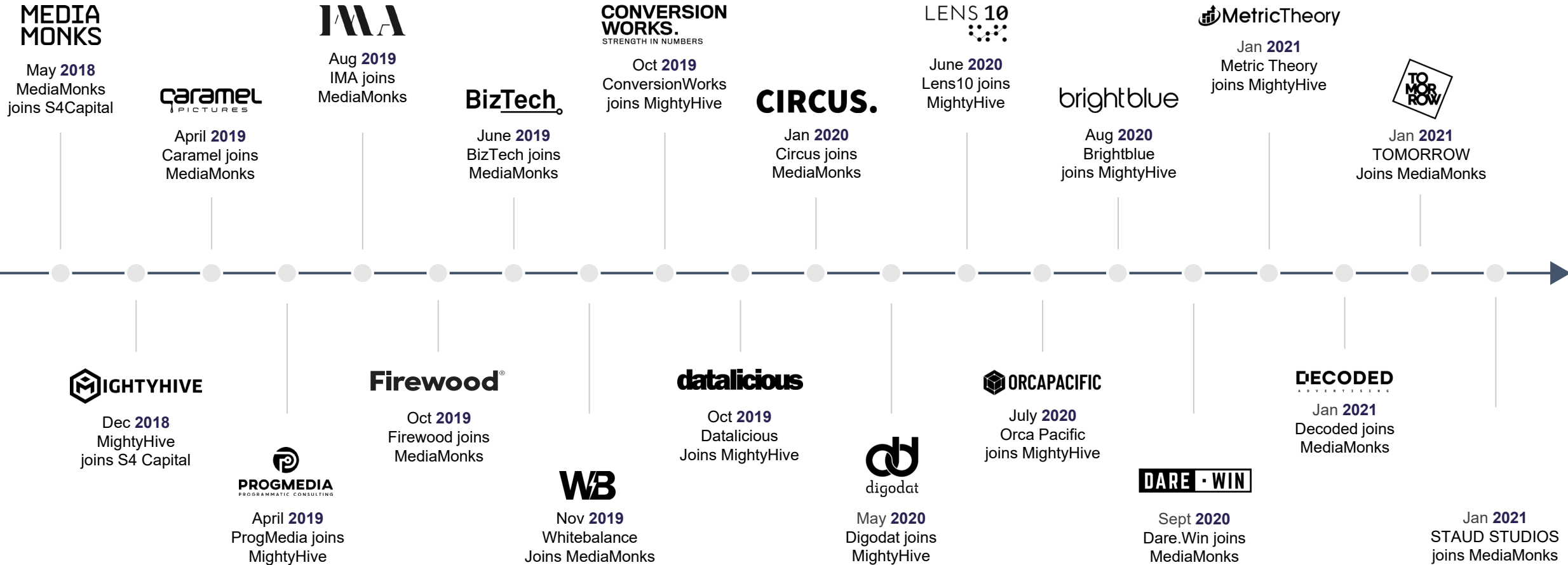
France  
Germany  
Italy  
Ireland  
Kazakhstan  
The Netherlands  
UK  
South Africa  
Russia  
Spain  
Sweden

## ASIA-PACIFIC

Australia  
China  
Hong Kong  
India  
Indonesia  
Japan  
New Zealand  
South Korea  
UAE



# Milestones



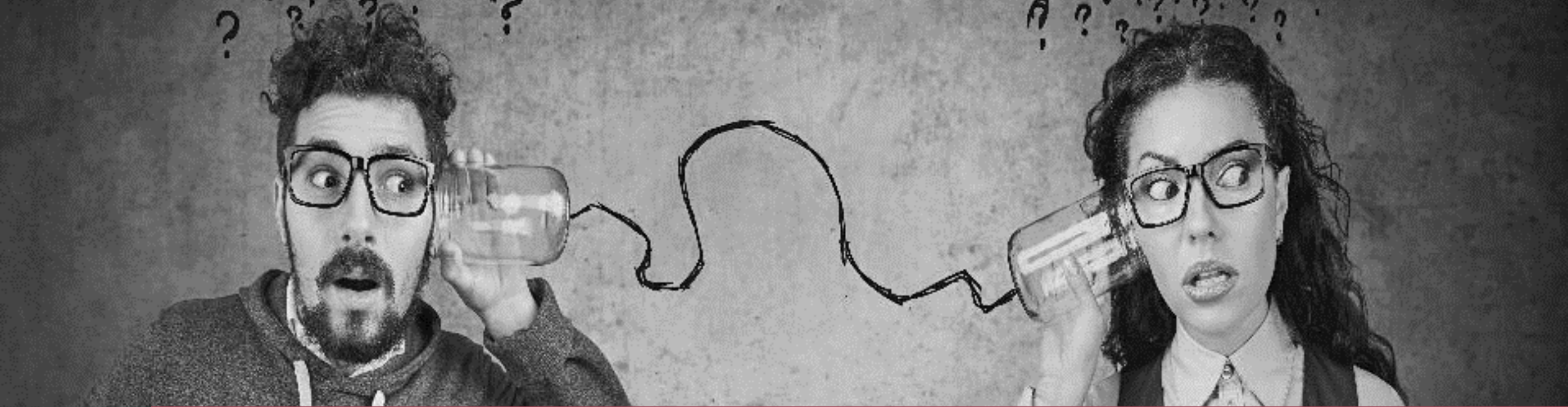
# 데이터 리터러시 이야기



**“Do you speak data?”**

Data is everywhere. BUT, 대부분의 사람들은

- Data를 어디에 사용할지 모름
- Data를 읽고 쓸 수 있는 문법을 모름



**Data is the New Language of Business..**  
However, not everyone is comfortable





# WHAT CAN **DATA** DO TO YOU?

1. Overwhelm You
2. Scare You
3. Confuse You



The background of the slide is split diagonally from the top-left to the bottom-right. The upper-left portion is black with white CSS code, including properties like 'position', 'display', 'top', 'left', 'line-height', 'padding-right', and 'z-index'. The lower-right portion is white.

## WHAT DOES DATA DO FOR YOU?

1. **Improve** capability of solving problem.
  2. **Improve** decision making.
- Moreover,
3. **Improve** your salary and career opportunity.

LET'S DIVE IN



# CONTENTS

- Data Literacy
- What is Data
- Why Now
- Data related Jobs

# **DATA LITERACY**

# Data Literacy in general

문해력 (Literacy) 이란,

- 문자를 읽고, 쓰고, 이해하는 능력
- 해석의 문제이지, 문자의 해독을 의미하지는 않음

데이터 문해력 (Data Literacy) 이란,

- 데이터를 읽고, 쓰고, 이해하는 능력
- 데이터의 맥락(context)을 이해하는 능력



# Quiz #1 :

질병 X를 치료할 수 있는 두 가지 약이 있다.

## Question 1.

작년 한 해 동안, 첫 번째 약은 1,000만명의 사람들을 치료한 반면, 두 번째 약은 100만명만 치료했다. 첫 번째 약이 더 효과적이라고 가정하는 것이 옳은가?

## Question 2.

그렇다면, 어떤 약이 더 효과적인지 이해하는 데 도움이 될 수 있는 지표는 무엇이 되어야 할까?

- 1) 일정 기간 동안 약을 복용한 사람
- 2) 약을 먹은 후 치료된 사람의 비율
- 3) 의약품 판매 수
- 4) 전혀 모르겠다



# Quiz #2 :

평균값의 함정  
4 명의 평균 연봉은 얼마입니까?



5천만원



5천만원



8천만원



2억원

평균 연봉이 9,500만원이라고?

# Quiz #2 :

중앙값...연봉 6.5천만원

4명의 연봉의 중앙값(Median)은 6,500만원

fx =MEDIAN(B3:E3)						
A	B	C	D	E	F	G
					Average	Median
Salary	5,000	5,000	8,000	20,000	9,500	6,500

# Data Literacy in business



## Read with data

데이터가 무엇이고,  
데이터가 나타내고자  
하는 관점을 이해하는  
능력



## Work with data

데이터를 생성, 획득,  
정제하고 관리하는 능력



## Analyze data

데이터를 필터링, 정렬,  
집계, 비교하고 분석하는  
작업을 수행할 수 있는  
능력



## Argue with data

특정의 청중에게 어떤 메시지나  
이야기를 전달하기 위해  
데이터를 사용할 수 있는 능력  
(데이터 스토리텔링)

# Data Literacy in business

비즈니스에서 기대하고 있는 Data Literacy 란,  
데이터를 알고 이해하는 단계를 넘어서 데이터를 통해 문제를 해결할 수 있는 능력



## What

- 어떤 데이터가 있나
- 분석의 목적은 무엇인가
- 무엇에 기여 할 것 인가 (비즈니스의 성장 혹은 비용의 감소)
- 누구에게 어떤 이야기를 통해 전달할 것인가

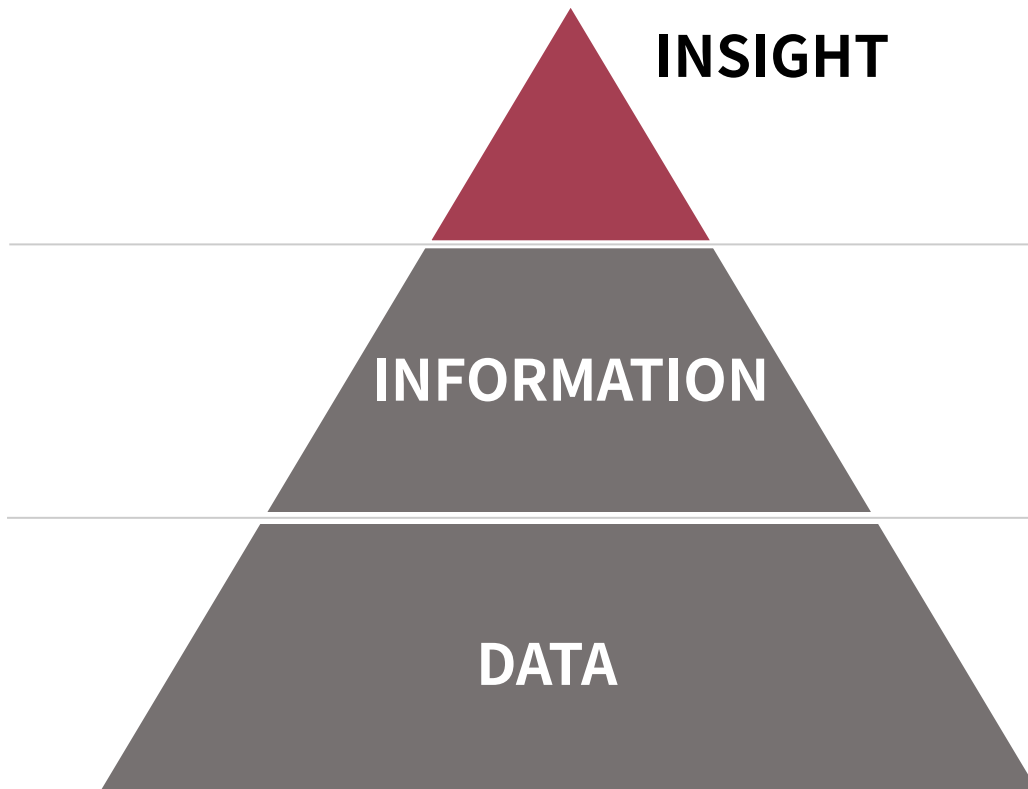
## How

- 무엇을 봐야 하나
- 데이터 구조화
- 서비스와 제품 이해
- 비즈니스 도메인 이해
- 비즈니스 성공지표와 분석의 측정지표
- 어떻게 시각화 할까
- 분석 후 인사이트와 제안은 무엇인가
- 의사결정자가 무엇을 의사 결정하게 지원할 것인가
- 어떻게 데이터를 통해 소통을 최적화 할까

# WHAT IS DATA

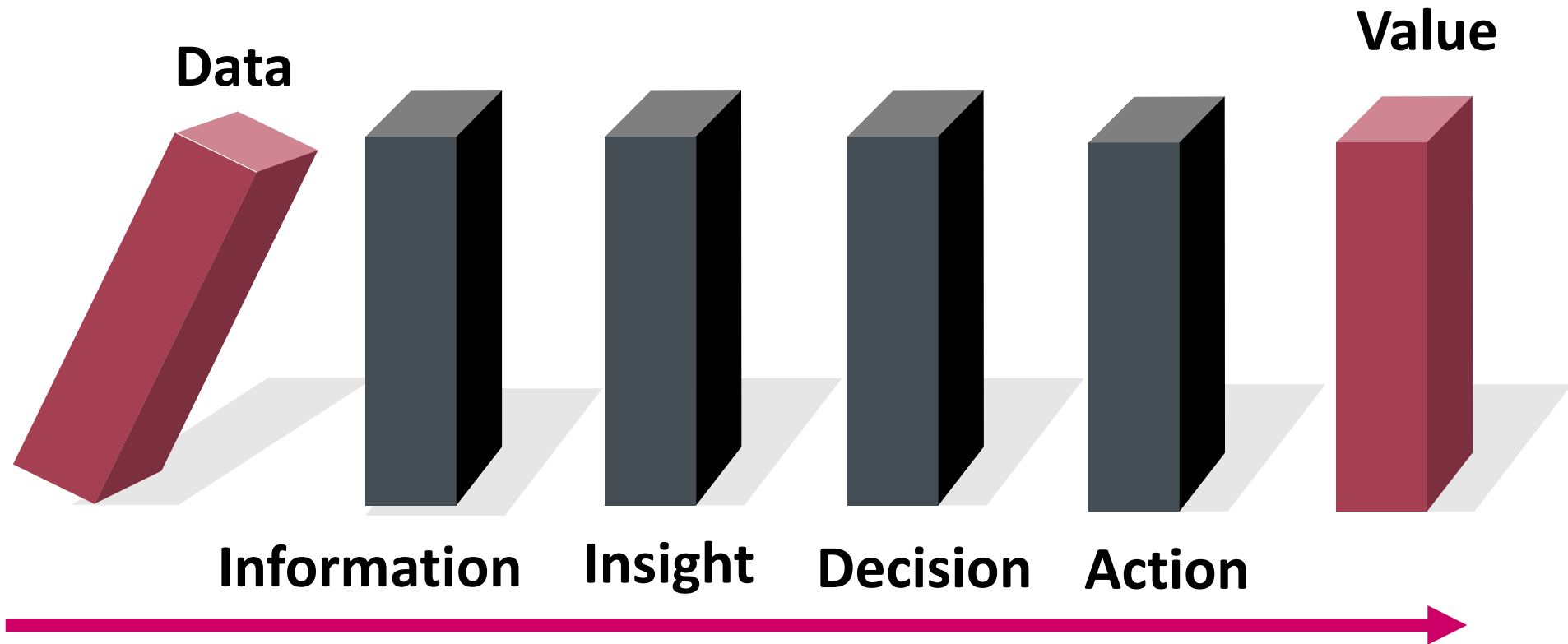


# Data != Insight

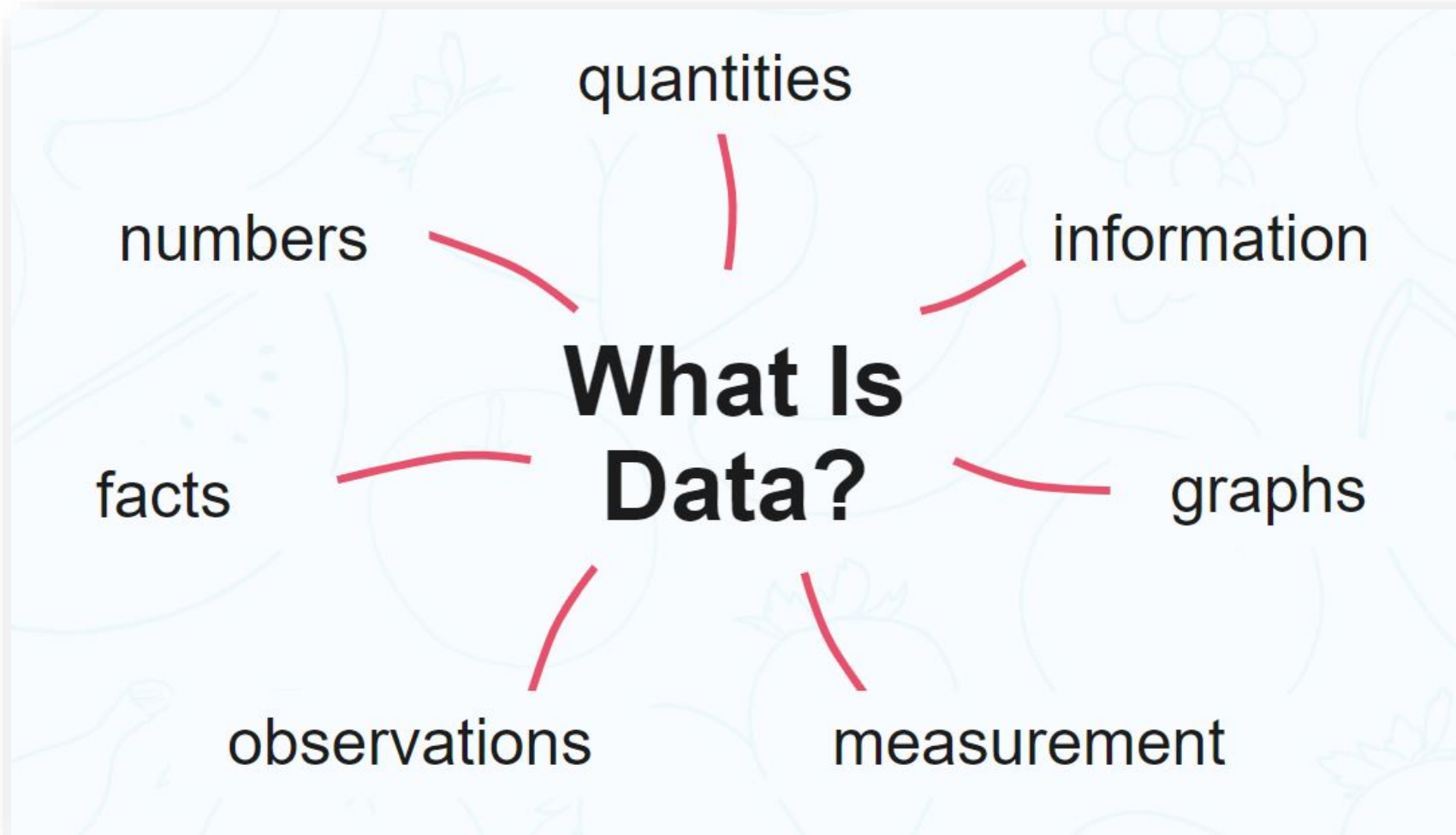


- Generated by analyzing information and drawing conclusions
- Prepared data : processed, aggregated and organized
- Human-friendly format
- Data visualizations, reports and dashboards
- Raw and unprocessed facts : numbers or text
- Quantitative (measured) or qualitative (observed)
- Computer-friendly formats: in databases and spreadsheets

# Data + Action = Value



# DATA DEFINITION



# DATA TYPE



**Data**

**Numeric  
(number)**

**Quantitative**

- The number of jellybeans in a bag

**Non-numeric  
(categorical)**

**Qualitative**

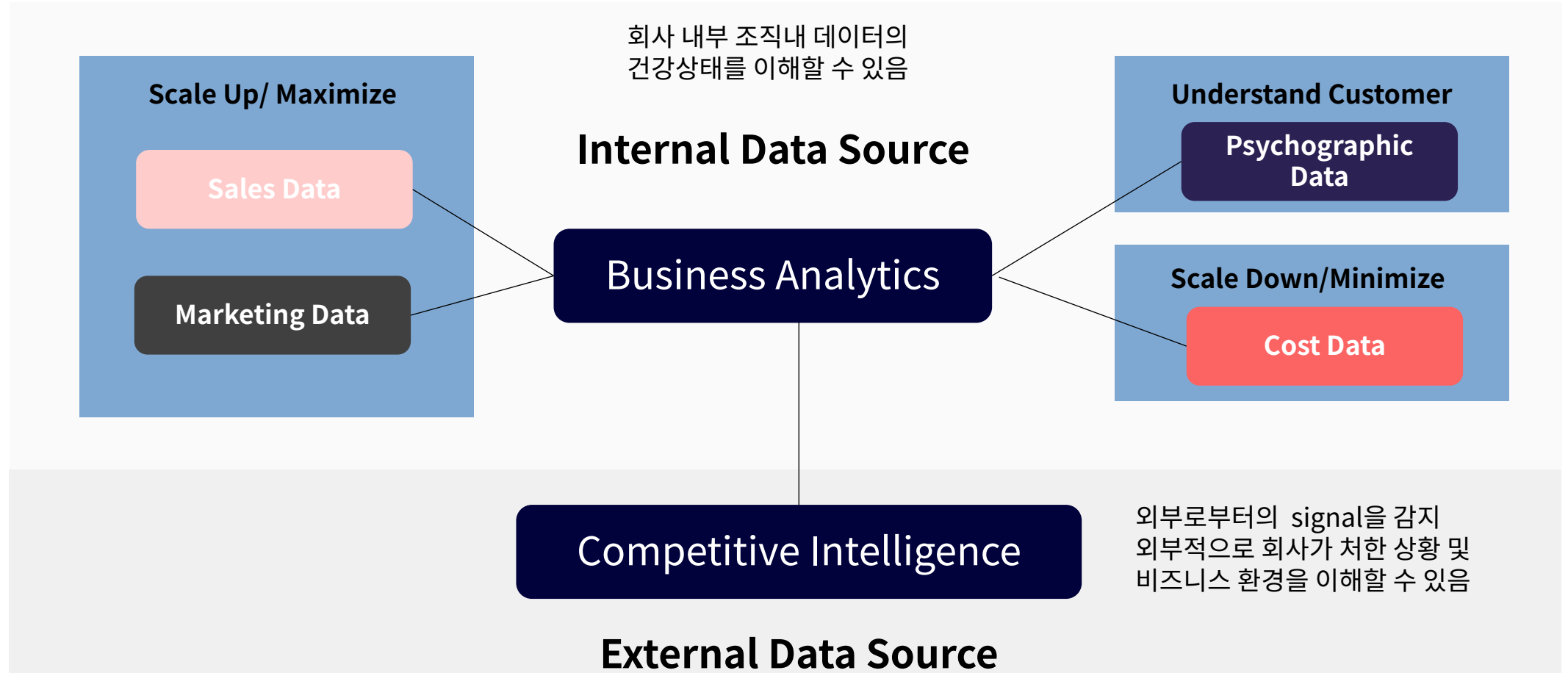
- The colors of jellybeans in a bag

# DATA TYPE EXAMPLE

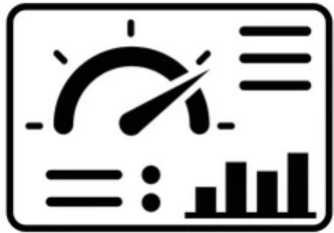
Numeric	Quantitative Data	Categorical	Qualitative Data
젤리빈(34g)의 가격	11,800원	어디에서 젤리빈을 샀나	스타필드 코엑스몰
지난 달 팔린 젤리빈 갯수	500개	가장 많이 팔리는 맛은?	버터팝콘맛
지난달 주문당 평균 젤리빈 매출	30,000 원	젤리빈의 색깔은?	red, black, white, green, yellow, brown, orange, pink, ,purple.



# DATA SOURCE MAP



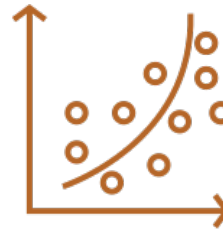
# 데이터를 가지고 할 수 있는 것



조직이나  
프로세스의  
현황 파악



이상 현상이  
일어나는 지점을  
탐지



어떤 이벤트나  
행동이 발생한  
원인을 진단하고,  
이들 간의  
상관관계를 이해



향후 일어날  
이벤트에 대한  
예측

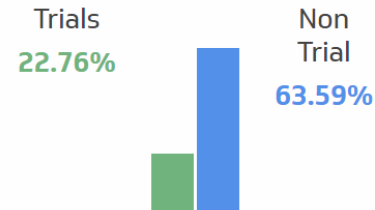
# CASE 1. DIGITAL MARKETING

- Web/App 의 고객 행동을 수집하여 비즈니스 의사결정에 활용
- 고객의 유입 및 전환을 측정
- 마케팅이 효율적/효과적으로 수행되고 있는지 검증

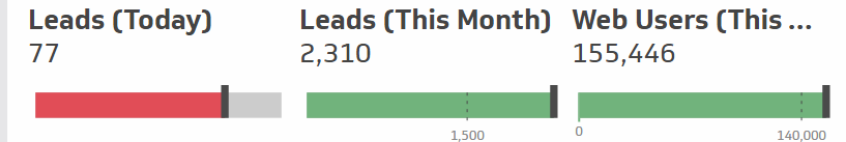
## Digital Marketing Dashboard

Klipfolio

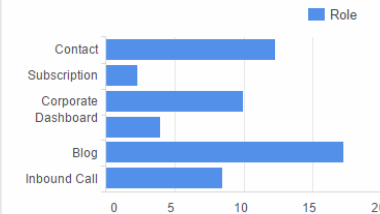
### Lead Breakdown (Today)



### Leads This Month



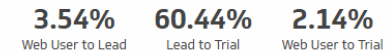
### Today Lead by Program



### Google Analytics Web Traffic Targets (This Month)

Channel	Progress	Sessions	Conversion Rate	Avg Time On Site	Bounce Rate	Events:Sessions Ratio
Organic Search	<div></div>	243,737	5.63%	5m:48s	41.66%	1.00 : 1
Display	<div></div>	112,378	1.98%	0m:37s	81.12%	1.01 : 1
Direct	<div></div>	47,324	8.53%	5m:12s	35.13%	1.01 : 1
Paid Search	<div></div>	13,790	13.09%	1m:47s	66.15%	1.02 : 1
Referral	<div></div>	12,023	26.95%	6m:52s	34.11%	1.01 : 1
Social	<div></div>	3,864	5.61%	4m:6s	55.27%	1.03 : 1
Email	<div></div>	3,549	13.01%	8m:3s	31.77%	1.03 : 1
(Other)	<div></div>	1,943	11.89%	7m:31s	28.21%	1.00 : 1
		438,606	5.91%	4m:19s	51.61%	1.00 : 1

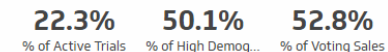
### Key Conversion Metrics



### Quality of Trial Leads (Today)



### Quality of This Month's Trial Leads



### Top Social Media Channels



### Trial Leads By Region (Today)



Source:

<https://www.klipfolio.com/resources/dashboard-examples/marketing/digital-marketing-dashboard>

Connected

Copyright © 2016 Klipfolio Inc. Trust & Terms of Use

BUILD YOUR OWN DASHBOARDS

Powered by Klipfolio

# CASE 2. FRAUD DETECTION

- 신용카드의 도난이나 오용 방지를 위해 활용
- 새로운 거래가 발생할 때마다 금액 및 날짜, 지역 등 알고리즘에 필요한 정보를 보내, 유효한 거래인지 사기인지를 알려주도록 학습
- 전통적인 방법의 Machine Learning (ML) 의 사례



## 전제 조건

### 1. 잘 정의된 문제

- 이 거래가 사기일 확률은 얼마나 될까?

### 2. Old data set

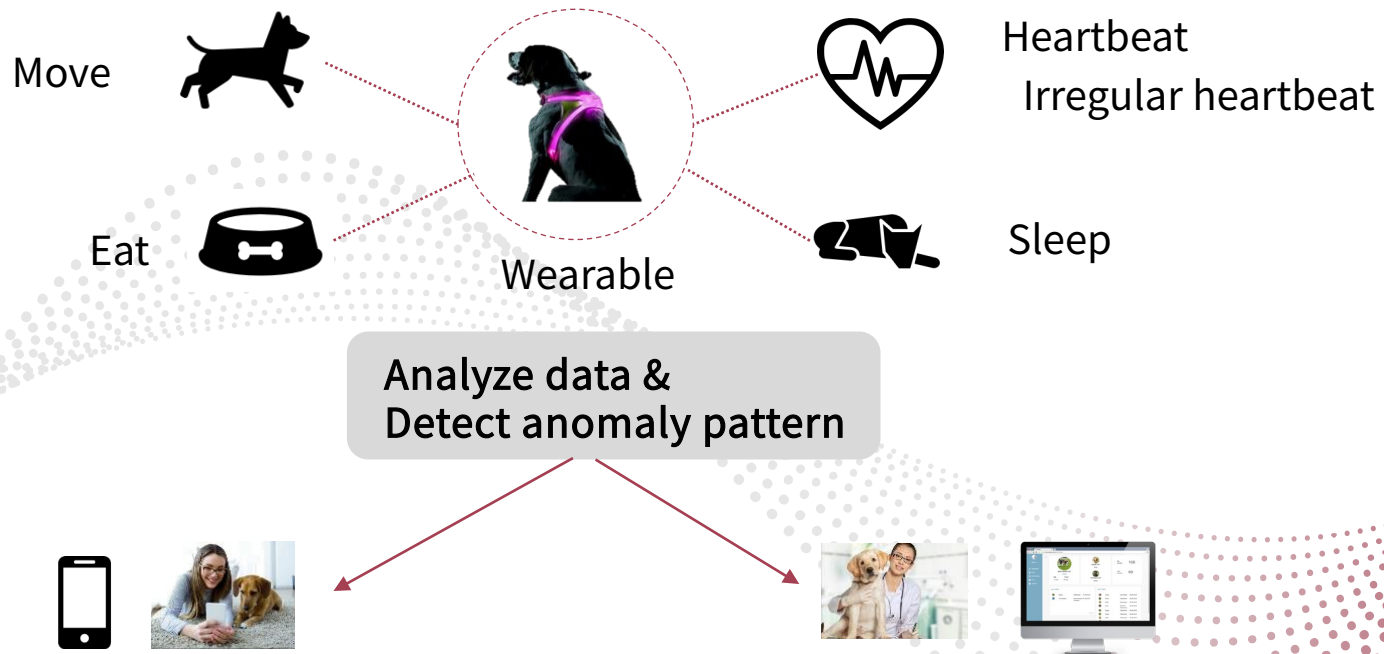
- 지난 몇 달/몇 년 신용카드 거래와 관련되어 ‘사기’와 ‘정상적인 거래’를 구분한 데이터가 있어야 함

### 3. New data set

- 새로운 구매에 대해 알고리즘에서 사용할 수 있게 정의가 되어 있어야 함
- 새 구매 정보에 동일하게 ‘사기’와 ‘정상적인 거래’를 구분할 수 있는 데이터 유형이 정의되어 있어야 함

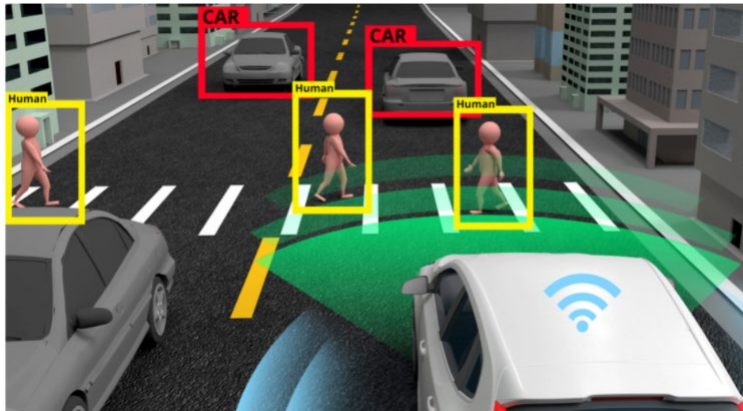
# CASE 3. PET WEARABLE

- 개의 움직임과 여러 신체 신호를 통해서 병원에 가야할 경우, 주인과 병원에 신호를 전달
- 정상적인 개와 아픈 개의 신호를 데이터로 전송 받아서 패턴을 학습함
- 전통적인 방법의 Machine Learning (ML) + Internet of Things (IoT) 의 사례



# CASE 4. IMAGE RECOGNITION

- 자율주행 자동차의 핵심 기술은 이미지에 사람이 포함되어 있는지 식별하는 것
- 각 픽셀의 숫자 행렬로 그림을 표현, 기존의 전통적인 ML 방식으로는 문제해결이 어려움
- 인간의 신경세포 뉴런과 같은 여러 계층의 미니 알고리즘을 활용하여 복잡한 문제 해결
- Deep Learning의 대표적 사례



157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	165	15	56	180
194	68	137	251	237	239	228	227	87	71	201	
172	126	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	95	90	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	165	15	56	180
194	68	137	251	237	239	228	227	87	71	201	
172	126	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	95	90	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Deep Learning 활용되는 문제의 예

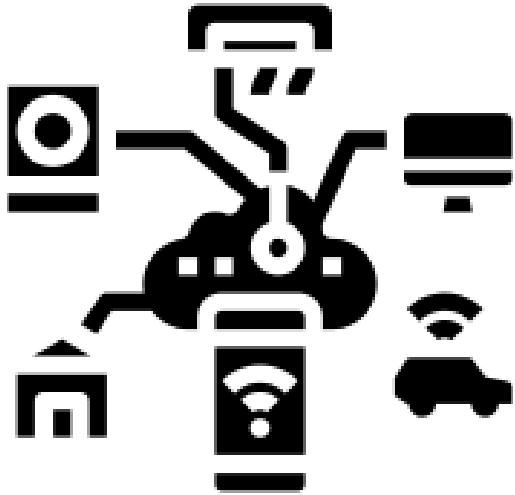
1. Image classification 또는
2. Language learning /understanding 등 데이터 집약적인 문제를 활용하는데 사용됨



**WHY NOW**

# 왜 지금인가?

데이터를 어디서든 수집이 가능해짐



데이터의 양이 폭발적으로 증가



=



개인이 하루에  
만들어내는 데이터의 양  
2.5 Quintillion Bytes

100 million (1억)개  
Blu-ray disc

# The world's most valuable resource is no longer oil, but data



Source :

The economist <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>

# 데이터와 직업 이야기





Line of Business

Focused on  
Business  
questions

Easily applies  
advanced analytics  
to answer question

Not a member of  
IT/Analytics teams

## **Citizen Data Scientist :**

A new generation of business users who can bring a business-side perspective to data science

**Market demand :  
Citizen Data Scientist**

# ‘데이터 분석’은 무엇을 하는 것인가



- 데이터 안에서 노다지를 발견하는 과정
- 끊임없이 What, when, who, how, where, why, why, why...라고 질문하면서 원인을 규명하는 과정



# DATA 관련 직무의 주요 업무



## 데이터 수집 및 보관

다양한 데이터 소스에서 데이터를 모아  
안전하고 접근이 가능한 방법으로 데이터를 보관

## 데이터 준비

Raw 형태의 데이터를 정제(cleaning) : 없는 데이터, 중복된 데이터를 찾아내고, 구조화된 데이터 형태로 변경

## 데이터 분석 및 시각화

여러 분석 방법을 통해서 인사이트를 도출하고 Dashboard나 report, 여러 시각화 도구를 활용하여 시각화

## 실험 및 예측

도출된 인사이트를 바탕으로 여러 실험을 통해 예측 모델을 만들고 더 많은 비즈니스 Return을 만들어 냄

# DATA 관련 주요 직무



# DATA ENGINEER



- 정보의 설계
- 데이터 파이프라인을 만들고, 저장 (Storage) 솔루션을 기획하고 설계하는 역할
- 데이터의 수집 및 저장의 책임
- 데이터 권한 설정 및 유지보수

## 사용 Tool / 기술

1. **SQL**
  - 데이터를 저장, 구조화
2. **Java, Scala or Python**
  - 데이터 프로세싱을 위한 Language
3. **Shell**
  - Task를 수행하기 위한 Commend line
4. **Cloud Computing**
  - AWS, Azure, Google Cloud Platform

# DATA ANALYST



- 데이터 탐색하고 분석
- 데이터 시각화 및 Dashboard 생성
- 데이터의 정제작업
- 다른 직무보다 덜 Technical함
- 데이터의 준비와 데이터 분석 및 시각화에 중점

## 사용 Tool / 기술

1. **SQL**
  - 기존 DB를 사용, 데이터를 추출하고 집계
2. **Spreadsheets (Excel or Google Sheets)**
  - 간단한 분석
3. **Web Analytics tools (Google Analytics, Adobe Analytics etc.)**
  - 웹/앱의 사용자 행동 데이터 분석
4. **BI Tools (Tableau, Power BI, Data Studio, Looker etc)**
  - Dashboard 와 시각화
5. **Python or R**
  - 데이터 정제 또는 분석

# DATA SCIENTIST



- 통계적인 지식을 기반으로 업무 수행
- 인사이트를 도출하기 위해 실험과 분석 업무 수행
- 전통적인 Machine learning 을 활용
- 데이터의 준비, 분석, 시각화 및 실험 예측의 과정에 모두 참여함

## 사용 Tool / 기술

1. **SQL**
  - 기존 DB를 사용, 데이터를 추출하고 집계
2. **Python or R**
  - Pandas(Python) 과 Tidyverse(R) 과 같은 Data Science 라이브러리 활용

# MACHINE LEARNING SCIENTIST



## 사용 Tool / 기술

### 1. Python or R

- Deep learning 알고리즘을 실행하기 위해 TensorFlow나 Spark과 같은 Machine learning 라이브러리 활용

- 통계적인 지식이 필요한 직무
- 이미 알고 있는 사실을 바탕으로 예측과 추정하는 업무
- Training 데이터를 사용하여 더 크고, 비정형의 정의되지 않은 데이터를 분류
- 전통적인 ML을 뛰어 넘어 deep learning을 활용하여 imaging processing 이나 Natural language processing을 수행
- 특히 실험 및 예측 부분에 집중











# DATA 관련 직무 요약



Data Engineer	Data Analyst	Data Scientist	Machine Learning Scientist
Store and maintain data	Visualize and describe data	Gain insights from data	Predict with data
SQL + Java/Scala/Python	SQL + BI Tools + Spreadsheets	Python/R	Python/R
	Data Citizen		

# CAREER로서 데이터 분석가의 장점과 단점

PROS	CONS
<ul style="list-style-type: none"><li> High salaries</li><li> Excellent career prospects</li><li> Satisfying work</li><li> Mix of technical and non-technical work</li></ul>	<ul style="list-style-type: none"><li> Programming can be tough</li><li> Companies may not know how to best use your skills</li><li> Lack of consistency across companies</li><li> Dealing with messy data</li></ul>

Source :  
<https://www.dataquest.io/blog/10-data-analytics-jobs/>

# 성공적인 DATA ANALYST가 되기 위한 역량

Category	Attribute	Type
Business Acumen	Big picture thinking	Augmentable
	Customer-centric	Learnable
	Marketing acumen	Learnable
Analytical Skills	Intelligent	Core
	Curious	Core
	Detail-oriented	Augmentable
	Open-minded	Augmentable
	Objective	Augmentable
	Problem-solving skills	Augmentable
People Skills	Tool mastery	Learnable
	Communication skills	Augmentable
	Interpersonal skills	Augmentable
Web Savviness	Online marketing	Learnable
	Online business models	Learnable
	Web design	Learnable
	Online technology	Learnable
	Online trends	Learnable

- **Core** : 필수역량
- **Augmentable** : 타고난 기본 소양을 기본으로 확장될 수 있는 역량
- **Learnable** : 배울 수 있는 역량

## Hero Factor

+

Passion **열정**  
 Proactive **주도적**  
 Confidence **자신감**  
 Relentlessness **주저함이 없음**  
 Resourceful **문제해결이 빠르고 명확**  
 Innovative **혁신적**

**강의를 마무리 하면서.**

**우리 회사에서 일 잘하는  
데이터 전문가들에게 물어보았습니다.**

**성공하는 데이터 분석가에게 가장 중요한 것은  
무엇이라고 생각하세요?**



호기심



무거운 엉덩이

# THANK YOU



ISABEL HAN

한명주



[https://www.linkedin.com/  
in/isabelhan/](https://www.linkedin.com/in/isabelhan/)



[myeongju.han@gmail.com](mailto:myeongju.han@gmail.com)



데이터 읽어주는 여자

<https://www.facebook.com/Isabel.datastoryteller>

<http://thetales.co>