



# People Analytics

An introduction

# EXPECTATIONS



## Setting Expectations

1. Basics, Areas of operation and challenges in "People Analytics"
2. **NOT** in-depth People Sciences / Algorithms. **BUT** relatable, tangible things in this field
3. MORE questions than answers
4. Hold your questions for the end

**GET  
INSPIRED  
BY THESE  
EXAMPLES**



# “The Algorithm That Tells the Boss Who Might Quit”

---



Pioneered and championed one of the first examples of the now very popular **Employee Churn Analytics**



**\$70,000,000** annual savings



Engagement is often seen as the holy grail of HR – but its impact is hard to measure

---

*But what if it was possible?*



At Best Buy the value of a **0.1%** increase in employee engagement at a particular store is ...

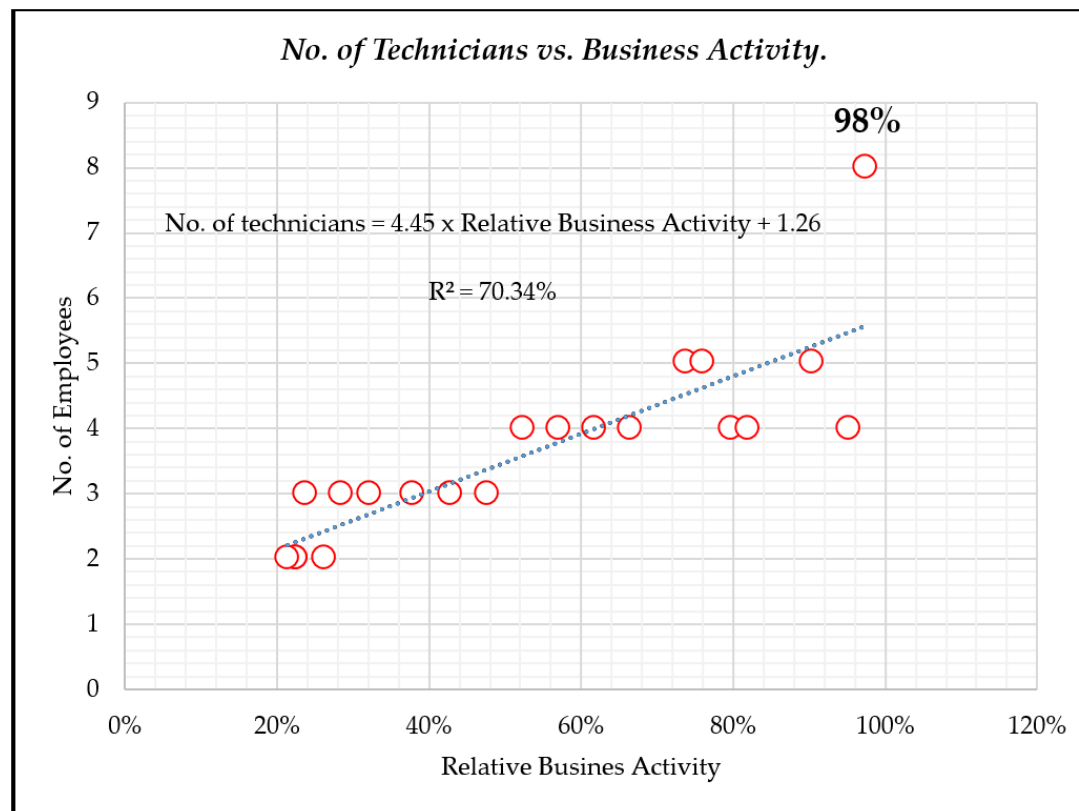
**\$100,000**

The significance of this relationship motivated Best Buy to make **employee engagement surveys quarterly rather than annually**



# A large mining company in Zimbabwe was concerned about losing money because of over or understaffed departments

They took the number of employees of a business unit and compared this to the business activity of this same business unit, measured over 17 quarters



# Content

---



WHAT AND WHY?



PERFORMANCE  
EVALUATION



STAFFING



COLLABORATION



SUMMARIZING  
TALENT ANALYTICS

# Content



WHAT AND WHY?



PERFORMANCE  
EVALUATION



STAFFING




COLLABORATION



SUMMARIZING  
TALENT ANALYTICS





“HR analytics is the systematic  
identification and quantification of the  
people drivers of business outcomes”  
*(Heuvel & Bondarouk, 2016)*

**AIHR**  
BLOG & ACADEMY



---

# What?



A data driven approach to managing people at work

# Why?



Growing availability of data, processing power and analytics tools

HUMAN RESOURCES

MANAGEMENT HR

HR



# Content



WHAT AND WHY?



PERFORMANCE  
EVALUATION



STAFFING



COLLABORATION



SUMMARIZING  
TALENT ANALYTICS



# Performance Evaluation

---



# Performance Evaluation

## Why?



Feedback



Reward

## Risks

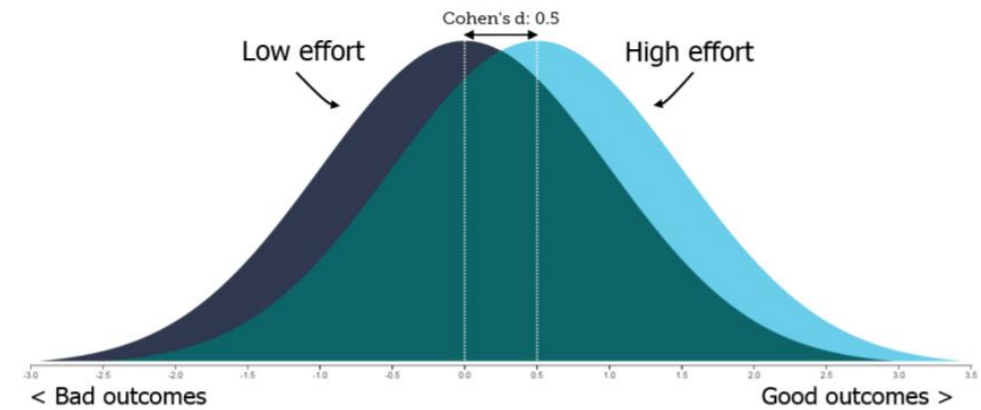
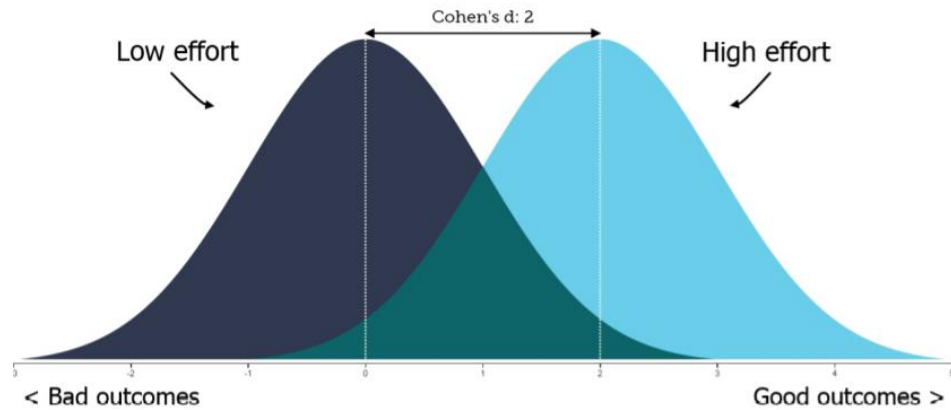


Each employee is unique



Performance measures are noisy

# Process v/s Outcome



- Consider broader set of outcomes
  - What impact does the person have on other areas
  - From 100% result-based rating -> 50% result-based + 50% effort-based
- Focus on process



# Process v/s Outcome

	2	FINAL	1	
Belgium				United States
93'		Kevin De Bruyne		
105'		Romelu Lukaku Assisted By: Kevin De Bruyne		
107'		Julian Green Assisted By: Michael Bradley		



**World Cup: USMNT goalkeeper Tim Howard sets World Cup record with 16 saves vs. Belgium**

# Regression to the mean

Any time you sample based on extreme values of 1 attribute, other attributes that is NOT perfectly related will tend to be closer to mean value

A study was recently conducted examining the performance of the 283 stock mutual funds that existed during the 1990s. The study divided the 1990s into an early period (1990-1994) and a late period (1995-1999). Here are the 10 funds that had the highest rate of return in the early period, ranked from 1 to 10

Early 1990s  
Fund & Rank

A 1  
B 2  
C 3  
D 4  
E 5  
F 6  
G 7  
H 8  
I 9  
J 10

Total # funds = 283

# Regression to the mean

Early 1990s Fund & Rank	
A	1
B	2
C	3
D	4
E	5
F	6
G	7
H	8
I	9
J	10

Total # funds = 283

Late 1990s Estimated Rank (median)	
	10
	20
	20
	28
	44
	37
	42
	31
	31
	25

Avg. = 25,  $r = .51$


Late 1990s Actual Rank	
	129
	134
	261
	21
	210
	53
	183
	105
	275
	54

Avg. = 142.5,  $r = -.03$

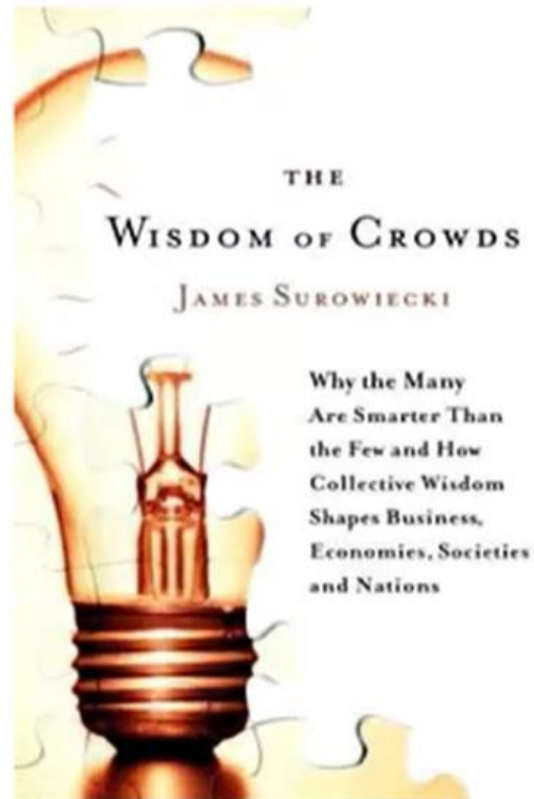


# Extrapolation from small numbers

•Your firm has two plants, one large and one small, which mass produce a standard computer chip. Other than the amount they produce, the two plants are identical in all essential regards. Both use the same technology to produce the same product. When properly functioning, this particular technology produces one percent (1%) defective items. Whenever the number of defective items from one day's production exceeds two percent (2%), a special note is made in the quality control log to "flag" the problem. At the end of the quarter, which plant would you expect to have more "flagged" days in its quality control log? Please mark one.

- |                               |  |     |
|-------------------------------|--|-----|
| A) The small plant            |  | 22% |
| B) The large plant            |  | 30% |
| C) The same number on average |  | 48% |

# Wisdom of crowds



The average of a large number of forecasts reliably outperforms the average individual forecast.

Our idiosyncratic guesses offset each other  
Eg: Cow's weight in a fair

**VALUE OF THE CROWD CRITICALLY DEPENDS ON THE INDEPENDENCE OF OPINIONS**

# Keep in mind

- Are the **differences persistent or random?** I.e., how do we know this isn't just good/bad luck?
  - The **more fundamental (skill-related) a performance measure is, the more it will persist over time**
  - The more chance-related a performance measure is, the more it will regress to the mean over time
- What else do we care about? Are we measuring enough? **What can we measure that's more fundamental?**
- **Is the sample large enough to draw strong conclusions?** How can we make it larger?
- How many **different signals** are really tapping into here? **How can we make them as independent as possible?**

# Content



WHAT AND WHY?



PERFORMANCE  
EVALUATION



STAFFING



COLLABORATION



SUMMARIZING  
TALENT ANALYTICS

Staffing





# Staffing Cycle



# Staffing Cycle



Most fundamental mantra

**HIRE RIGHT PERSONNEL!**



# Getting selection right

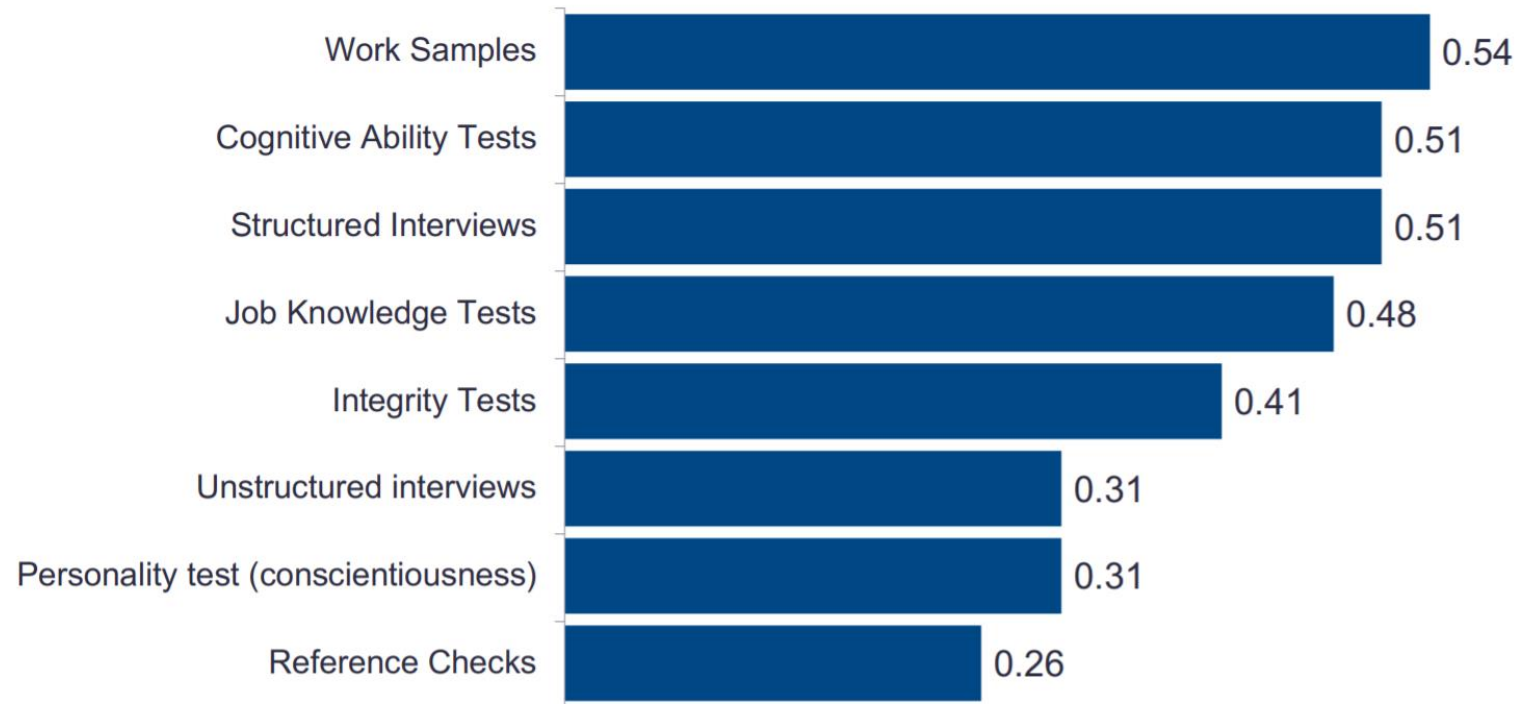
---

- Job knowledge tests
- Cognitive ability tests
- Personality tests
- Reference checks
- Structured interviews
- Unstructured interviews
- Work samples
- Integrity tests



# Getting selection right

Correlation with subsequent performance (0-1)



# Using Data Analysis to Predict Performance



# Staffing Cycle



Peter principle:

“In time, every post is occupied by an employee who is incompetent to carry out its duties”

- Peter and Hull, 1969



## Analyzing Promotability

---

- How well does success in the current job predict performance in a higher-level job?

OR

- Which dimensions of lower level performance best predict performance in the higher-level job?

**PROMOTES v/s HIRES**

# Does it Matter How People Enter Jobs?

## Performance

- Hires performed substantially worse than similar promotes
  - **75% less likely to get top** rating 270% more likely to get lowest rating
- Takes **3 years to acquire similar performance** to those promoted into the job
- Manager posting job internally & inviting interested candidates to apply, sees better performers than when identifying people from personal networks

## Pay

- New hires receive 18% more compensation than promotes
  - Pay gap only closes very slowly (up to 7 years)



# Importance of Causality and why do we care about it?

- People who enter jobs through formal posting perform worse  Should we avoid posting?
- People who have been in the job longest have lower performance  Should we move people around more?
- People who have taken a training program perform better  Should we send more people to training?
- People who have taken a training program show greater performance improvements  Should we send more people to training?

# Problems with causality



Omitted variable bias

Eg: Job posting -> Lower performance



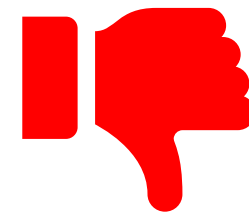
Reverse causality

Eg: Training vs Job performance

# Staffing cycle



**Do millennials switch jobs more frequently than the previous generation?**



# Attrition



## Problems

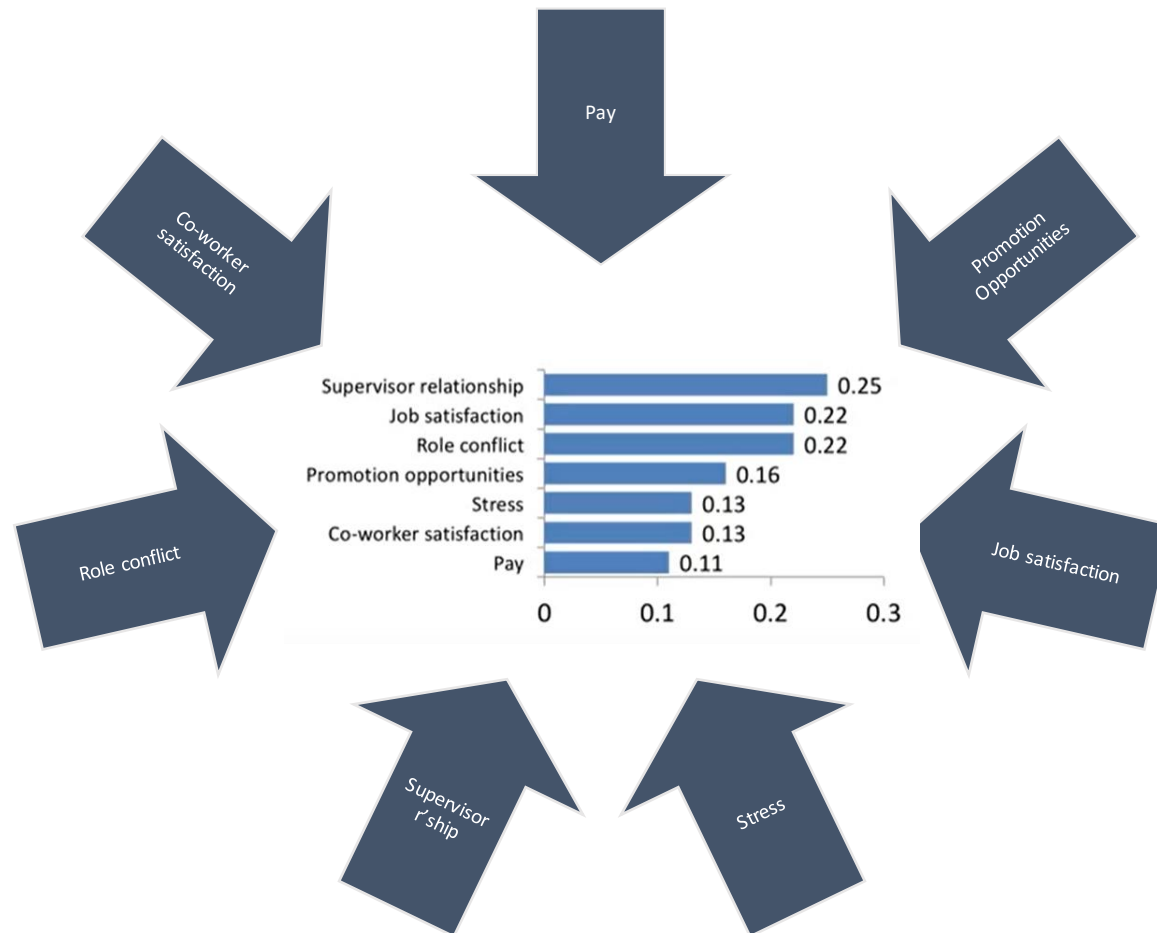
- Increased hiring and training cost
- Loss of unique info
- Hampered client relationships



## Levers

- Informed hiring strategy
- Target interventions
- Improve work conditions
- Address unmet needs
- Train managers

# Possible reasons



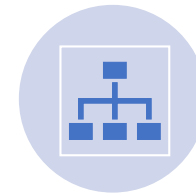
# Predicting attrition



Managers



Pre-hire  
background



Type of work /  
Project / Function



Performance  
evaluations



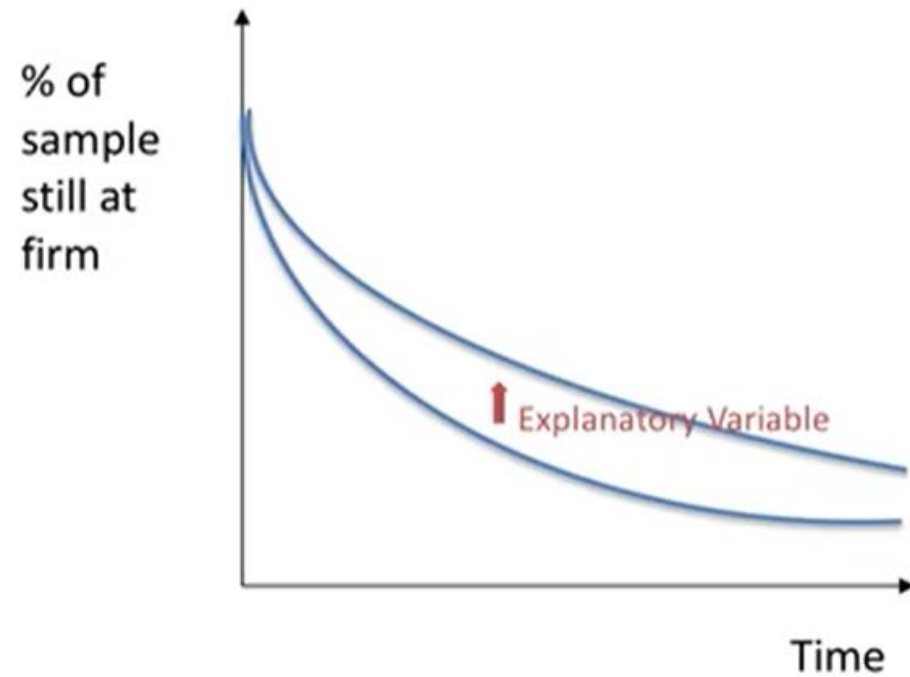
Geography



Social Network  
behaviour

# Predicting attrition

## The Survival Model



# Content



WHAT AND WHY?



PERFORMANCE  
EVALUATION



STAFFING



COLLABORATION



SUMMARIZING  
TALENT ANALYTICS



Collaboration





# How can we improve collaboration inside organizations?

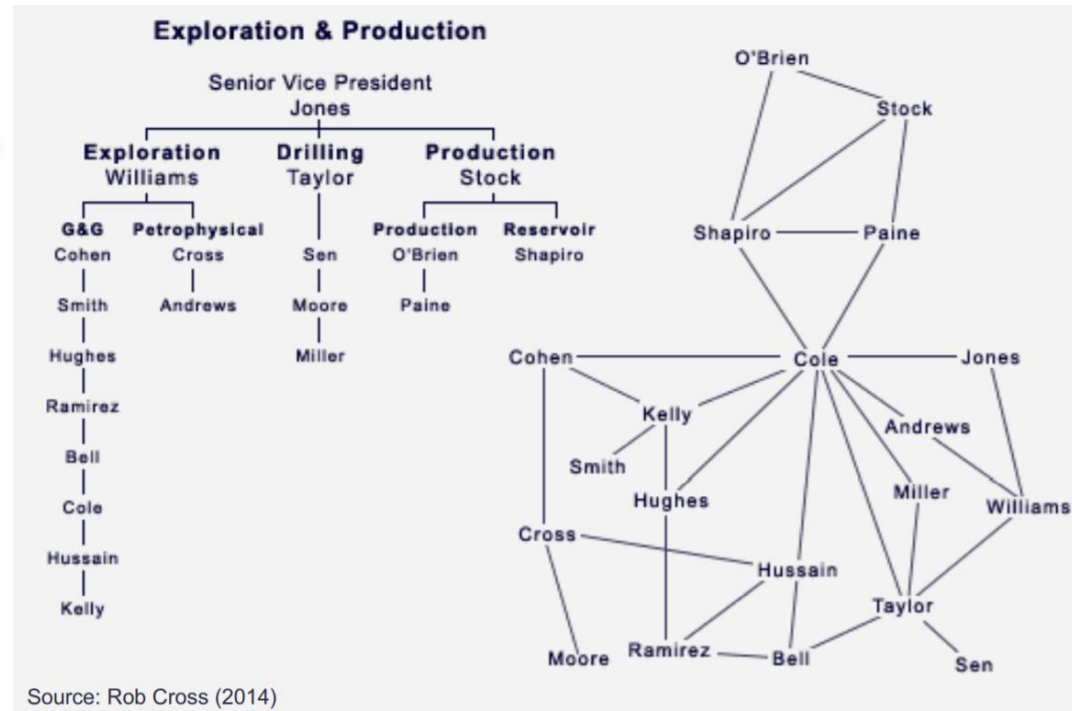
---

- How can we **describe** collaboration patterns between employees?
- How can we **measure** these collaboration patterns?
- How can we **evaluate** these collaboration patterns?

Ans: By the tools and techniques of  
**Organizational  
Network Analysis**

# What are Organizational Networks?

Formal structure  
(org. chart)

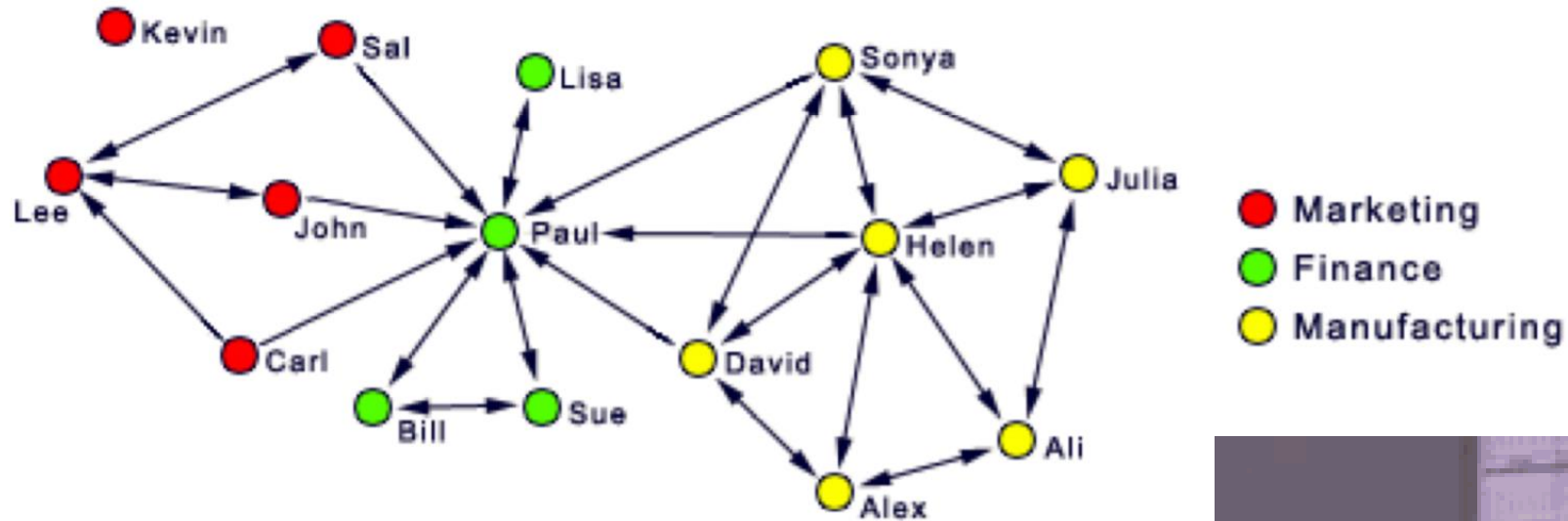


Informal structure  
(network map)

Types:

- Collaboration networks (information flows, knowledge sharing)
- Communication networks
- Friendship networks
- Advice networks
- Trust networks

# Collaboration Networks: An Example



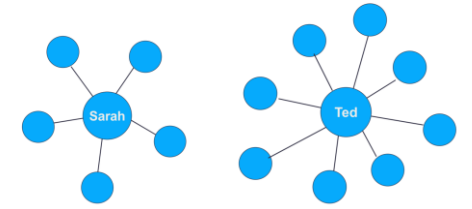
- Marketing
- Finance
- Manufacturing

A → B A seeks information from B  
A ↔ B A and B seek information from each other

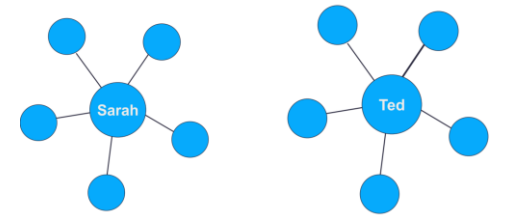


# How Can We Describe Collaboration Patterns?

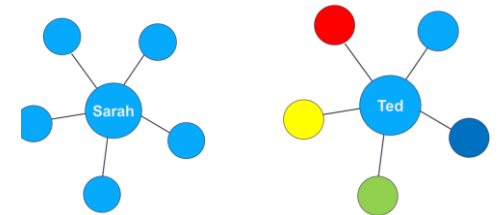
Network size



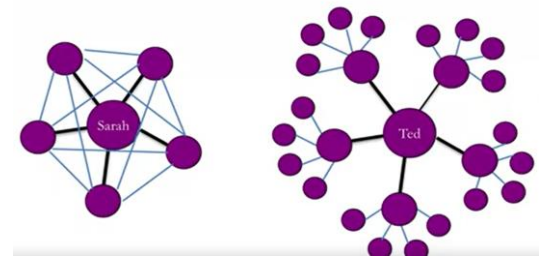
Network strength



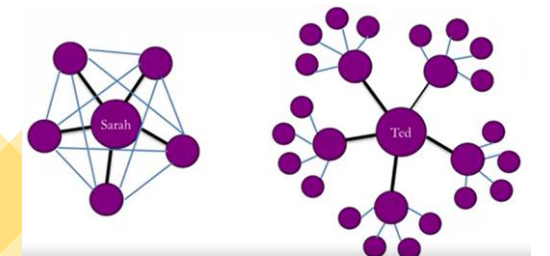
Network range



Network density



Network centrality





# How can we measure Collaboration patterns?

## 2. Big Data:

- interactions via email, computer conferencing, intranet etc.

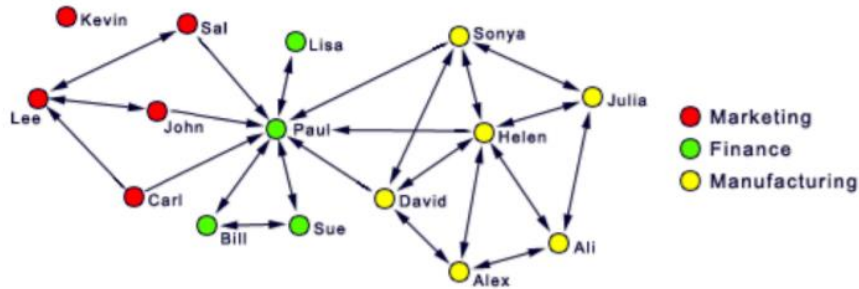
## 3. Archival Records:

- corporate databases - e.g. info on shared project assignments, work histories, event attendance
- public databases - e.g. info on co-patenting, co-authorship, co-citations

## 4. Fieldwork

- observations, diaries, electronic tags, etc.

# How can we evaluate these collaboration patterns?



How do collaboration patterns vary?

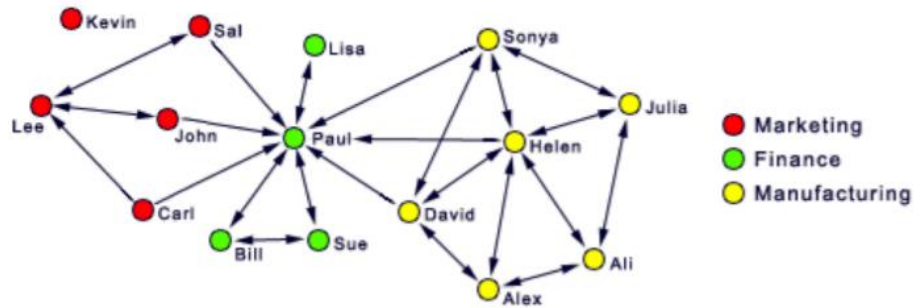
How do collaboration patterns matter for important outcomes? (individual, group, or organizational)

- Network size
- Network strength
- Network range
- Network density
- Network centrality

- Performance
- Satisfaction
- Commitment
- Burnout
- Turnover



# How can we evaluate these collaboration patterns?



	Lee	John	Paul	Helen	Julia
Network size (inbound ties: number of people who seek information from X)	3	1	9	5	3
Network size (outbound ties: number of people from whom X seeks information)	2	2	3	6	3

	Outcome variable:
	Performance
<b>Network variables:</b>	
Network size (inbound ties: number of people who seek information from X)	+
Network size (outbound ties: number of people from whom X seeks information)	-

# Content



WHAT AND WHY?



PERFORMANCE  
EVALUATION



STAFFING



COLLABORATION



SUMMARIZING  
TALENT ANALYTICS

Final thoughts I  
want you all to take  
away from this  
session

# Importance of context

- Situation v/s Personal Disposition – Fundamental Attribution Error
- **“Are we comparing apples to apples”:**
  - The need to compare on a level playing field

# Interdependence

- **“How interdependent are each other”**: The need to parse individual’s contribution from the team’s performance
- Performance evaluation is often done at a team level
  - Identify newer methods to understand individual contribution. Eg: Network Analysis

# Self fulfilling prophecies

- **“How have expectations colored evaluations”**
- Greater your expectations, greater the performance

Are the factors  
truly causal?

- Are charismatic leaders successful or leaders who enjoy success are hence charismatic?

People Analytics is more of an  
Organizational Challenge than an Analytics  
Challenge



# Prescriptions

- **NO BLACK BOXES**
  - **Be transparent**
  - **Embed yourself**
  - **Share control**
- **Ask the critical questions**
  - **Are we comparing apples to apples**
  - **How interdependent are each other**
  - **How have expectations colored evaluations**
  - **Are the factors truly causal**

- PAFOW – People Analytics and Future of Work

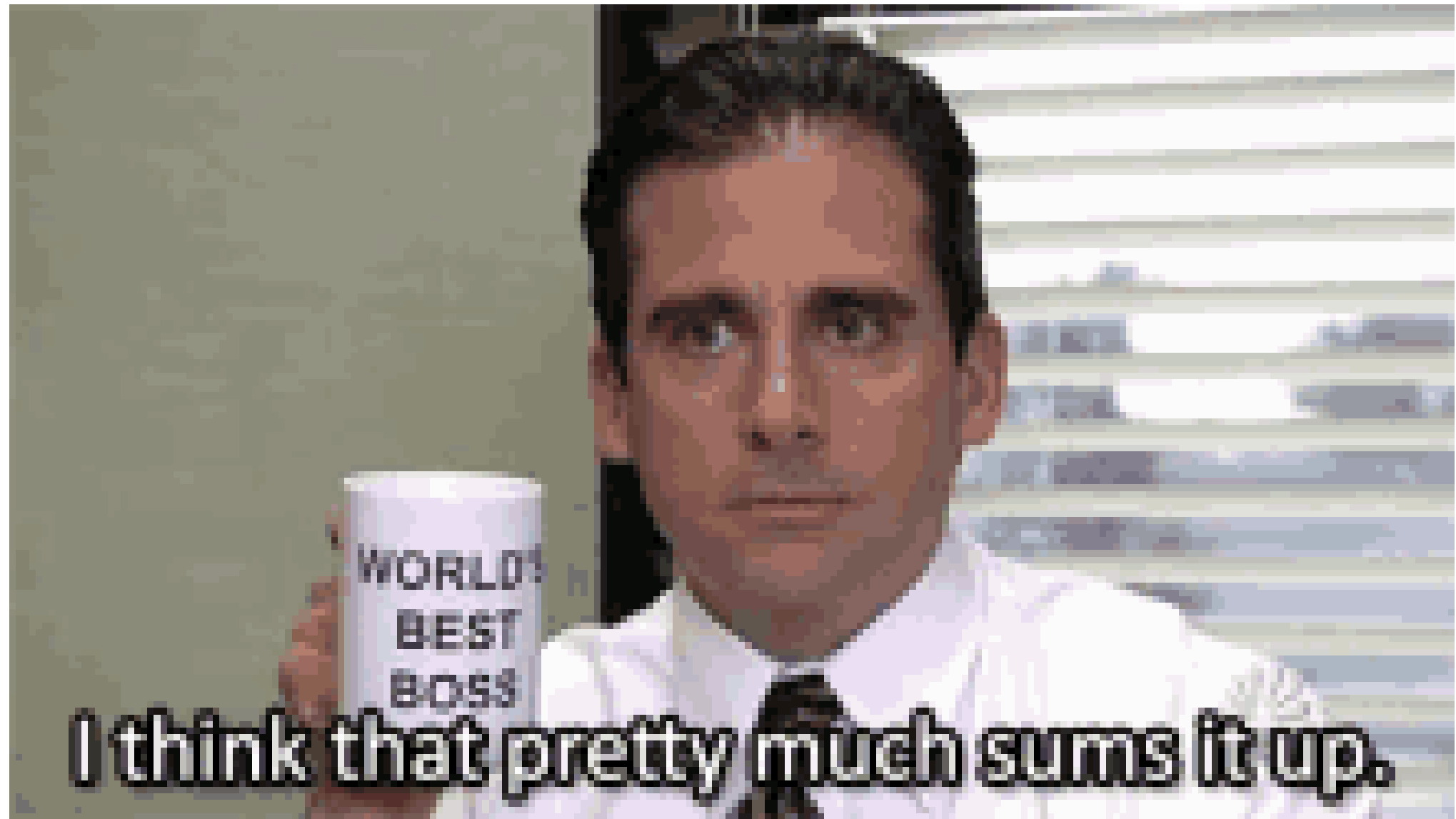
- David Green



- Josh Bersin



- Digital HR leaders podcast



**I think that pretty much sums it up.**

