

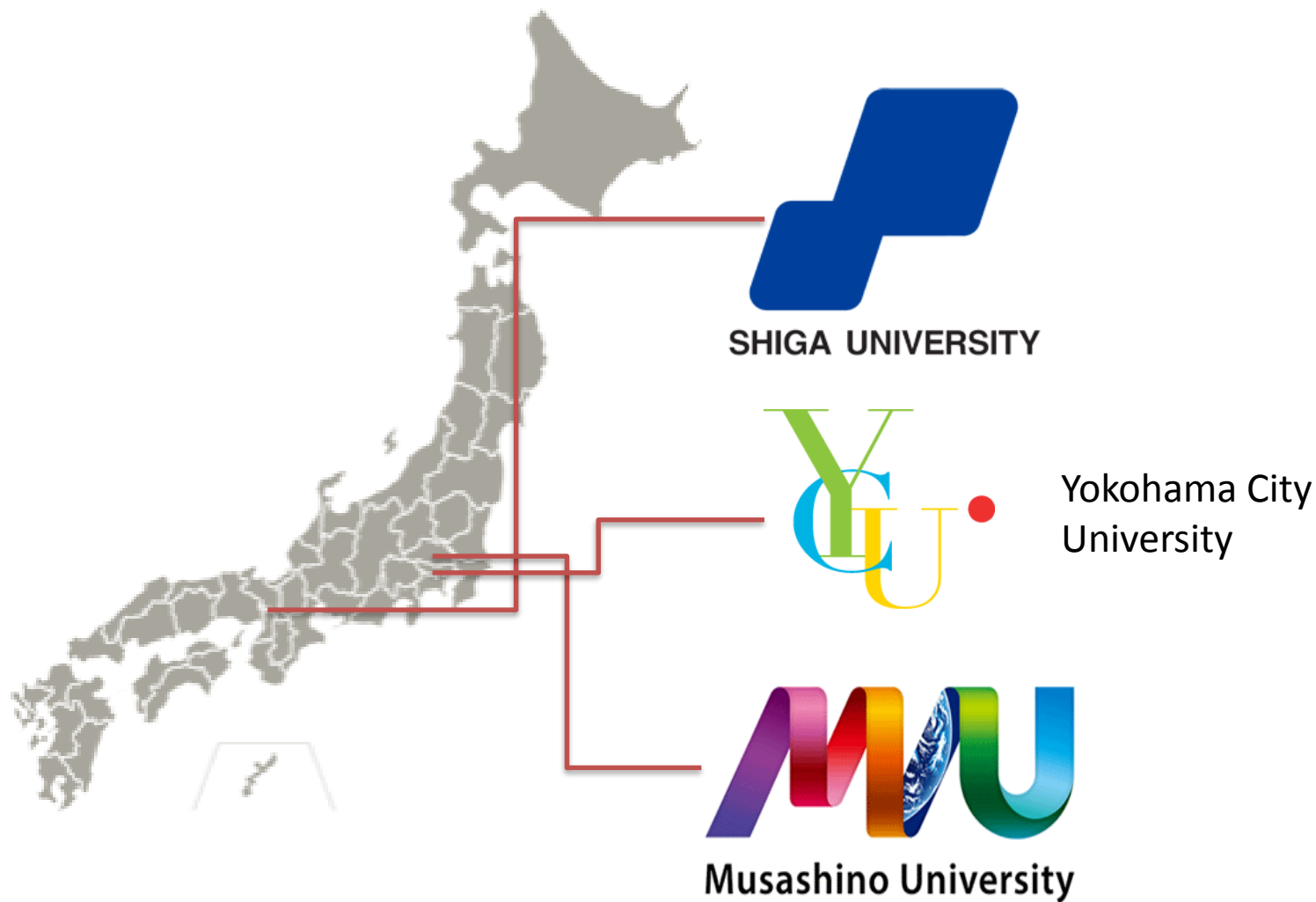
# 横浜市立大学のデータサイエンス 教育---WiDS Tokyo@YCUを通じて

小野陽子

横浜市立大学 DS学部

WiDS Tokyo @ YCU アンバサダー

# 日本のデータサイエンス学部2019



歴史から見るデータサイエンス

# データサイエンスとは？

# データサイエンスとは？

- データサイエンスという言葉には欠陥がある，でも使えるのはどうして？
- データサイエンスは単に統計の名称変更？

データサイエンティストとは，統計学者より計算機科学が理解できて，計算機科学者より統計を使えるひとのこと

Joshua Blumenstock

# DSの歴史を紐解く (--1980)

**1962:**Data Analysis, Tukey :The Future of Data Analysis, Ann. Math. Statist.

**1968:**Data Science, Naur:'Datalogy', the science of data and data processes, and its place in education, Proc. of the IFIP 68

**1974:**Naur published *Concise Survey of Computer*

**1977:** \*Tukey published *Exploratory Data Analysis*

\*The International Association for Statistical Computing (IASC) is established.

# 1990'sは学会

**1989** Piatetsky-Shapiro organizes and chairs the first Knowledge Discovery in Databases (KDD) workshop.

**1994** \*BusinessWeek “Database Marketing”

\*Gibbons et al. “The New Production of Knowledge:  
The Dynamics of Science and Research in Contemporary  
Societies. “

**1995** Knowledge Discovery and Data Mining (KDD).

**1996** \*For the first time, the term “data science” is included in the title of the conference (“Data science, classification, and related methods”). IFCS, Kobe, Japan

\* Fayyad, Piatetsky-Shapiro, and Smyth publish  
“From Data Mining to Knowledge Discovery in Databases.”

**1997** The journal “Data Mining and Knowledge Discovery” is launched

# 2000'sは広がりに

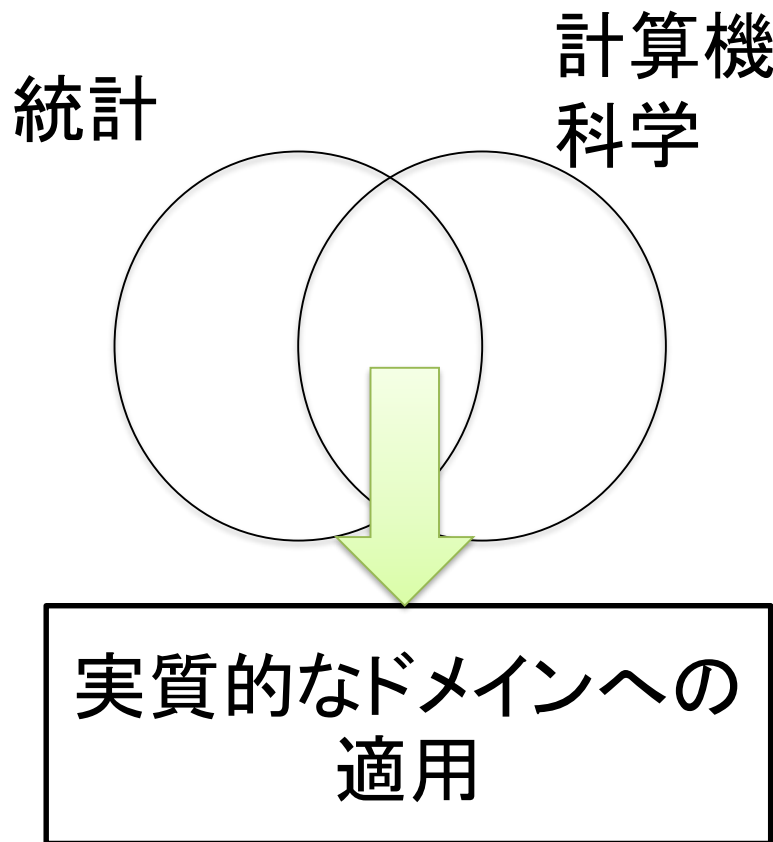
- 2001** Cleveland publishes “Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics.”
- 2003** Launch of “Journal of Data Science”
- 2005** Davenport et al. publish “Competing on Analytics
- 2009** Sadkowsky creates the data scientists group on LinkedIn

# 2010'sは世間への認知

- 2011** Warden writes “Why the term ‘data science’ is flawed but useful”.
- 2011** David Smith writes in “‘Data Science’: What’s in a name?”
- 2012** Davenport and Patil publish “Data Scientist: The Sexiest Job of the 21st Century”
- 2013** First Data Science in Aviation Workshop  
“Complex Data Mining”
- 2015** Women in Data Science(WiDS) launched @ Stanford U



# なぜDS? どうして今DS?



Skiena (2017)

- 新しい技術が大量データに適用できるようになった
- 計算機能力が上がった
- GAFAM などの企業が現代的データ解析に対する力を持った

## モード1とモード2の違い

	モード 1	モード 2
問題設定	内在的 専門分野内部で生まれる理論上の欠陥、矛盾の解消 より高度な理論体系の形成	外在的 社会・企業等の実際場面に起こった、あるいは起こることが想定される問題の解決 ミッション・オリエンテッド
最終目標	理論的完成度 ディシプリン内の価値の向上	社会問題の発見・解決 新しい価値の社会への展開
他分野との関係	単一のディシプリン内部での解決 同輩集団内部での自己完結性	Transdisciplinary、異分野間のコミュニケーションが必須
研究開発体制	制度的に安定してる組織体制	大学以外の研究機関、シンクタンク、政府、NGO 分野横断的
研究分野の評価	その専門分野での高度化への貢献	社会からの期待をどれだけ満たすことができたか 社会への付加価値付与
評価基準	同輩集団による評価	社会の期待をどれだけ満たすことができたか
社会状況	投入資源量の拡大、安定 規模の経済 効率性重視	投入資源量の減少 ストックの利活用 安全・安心、満足の重視、Society 5.0
基礎研究	(純粋) 基礎研究	社会状況の理解、アイディアの構造化にかかる基礎研究 ※ モード1とモード2の基礎研究は、本質的に異なる
例	核物理、電気、機械、金属、化学、統計学、数学	環境保全学、安全学、デザイン学、データサイエンス、原子力学

Gibbons(1994)、潮木(2012)を村田・汪が改変

Mode2をいかに教えるか

# データサイエンス教育とは

# データ解析教育 (Tukey, 1963)

Data analysis, and the parts of statistics which adhere to it, must then take on the characteristics of a science rather than those of mathematics, specifically:

(b1) Data analysis must seek for scope and usefulness rather than security.

(b2) Data analysis must be willing to err moderately often in order that inadequate evidence shall more often *suggest* the right answer.

(b3) Data analysis must use mathematical argument and mathematical results as bases for judgment rather than as bases for proof or stamps of validity.

Mode 1として

# YCU での教育プログラム

Mode 2 のデータサイエンスには不可欠

ドメインでの学び

柱となる学問の融合領域

アルゴリズム

統計


計算機科学

数学を含めた基礎教育

ドメインへの興味喚起

# YCUデータサイエンスのプロセス

課題  
発見



データ  
定義



データ  
収集




データ  
加工



データ  
分析



解釈・  
提案




提案完了



# 大学の外で学ぶこと

課題  
発見

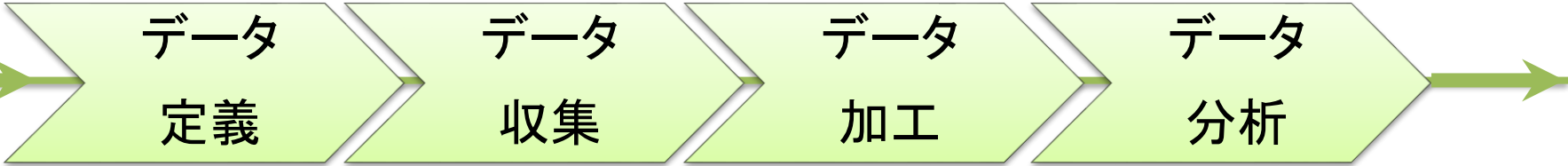


データ  
定義

データ  
収集

データ  
加工

データ  
分析

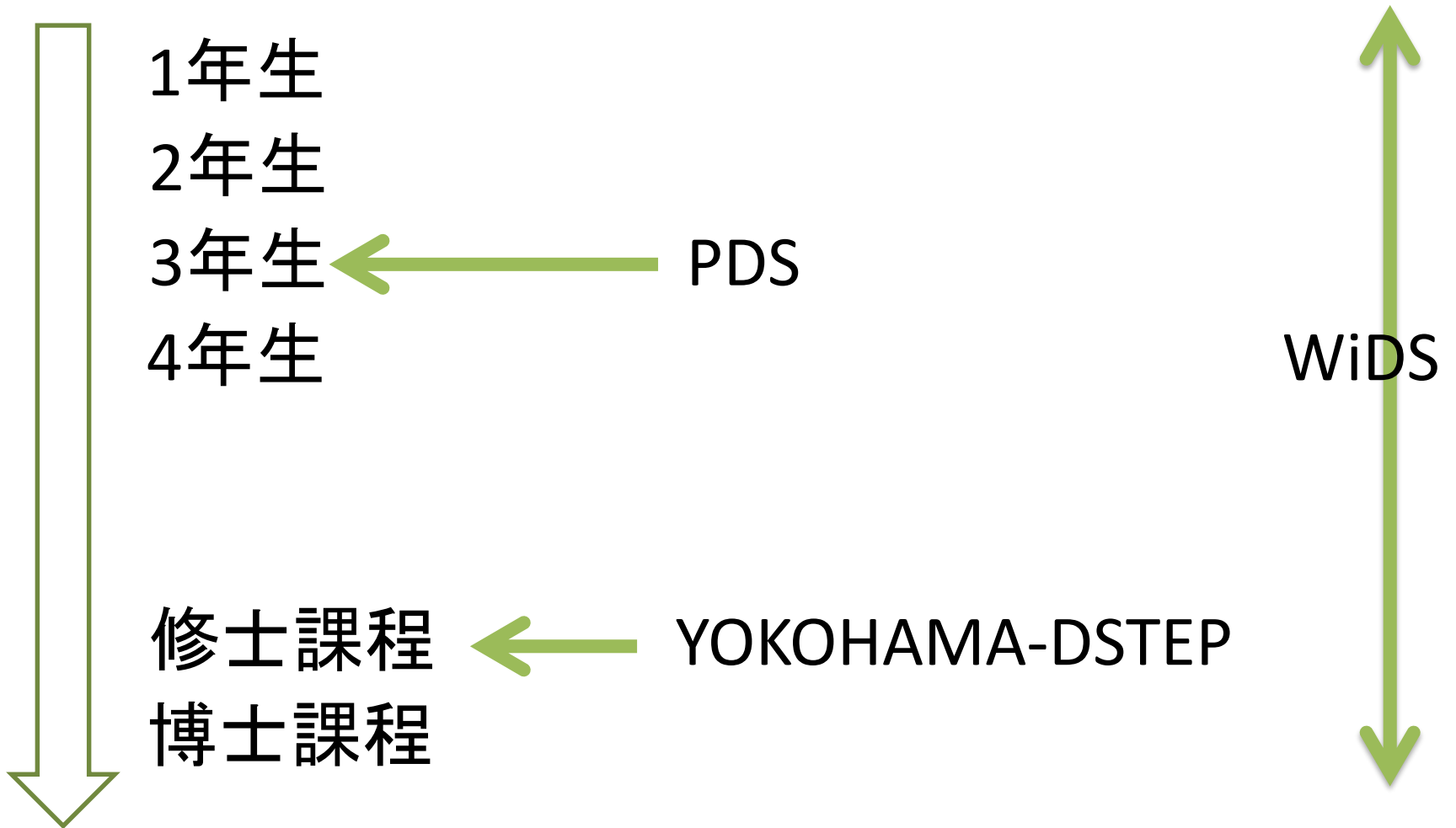


解釈・  
提案

提案完了



# 外での学び





# PDS: Practical Data Science

- 学部3年生対象 PBL = “Problem” Based Learning
- 産官学連携
- 事前にデータマッピング・クレンジングを学ぶ
- DS プロセスを現地で2週間程度 体験する。

課題発見

解釈・提案

提案完了

Orchestrating a brighter world

NEC



全日空商事株式会社  
ALL NIPPON AIRWAYS TRADING CO.,LTD.



OPEN  
YOKOHAMA

# YOKOHAMA D-STEP

- [文部科学省] 超スマート社会の実現に向けたデータサイエンティスト育成事業
- 文理融合・実課題解決型データサイエンティスト育成
- 修士レベル相当
- 産官学連携
- 実課題解決型PBL←DEJIMAにご協力をいただいております

# WiDS: Women in Data Science

- Stanford, ICME(2015年より)を中心に始まる
- Goals
  - **Inspire** 女性をDS分野に誘う
  - **Educate** ジェンダーを問わず参画
  - **Support** DS分野の女性を支援
- アンバサダー制度の導入←この仕組みでアメリカ最大の計算機科学奨学補助組織から表彰
- データソンとPodcastを用いた教育
- 2019 Global WiDS Events:150+ 都市, 50カ国+  
←WiDS Tokyo @ YCUもこのひとつ

どうしてWomen inなの？

# WOMEN IN DATA SCIENE

# WiDS2019: 開催国50+, イベント数150+

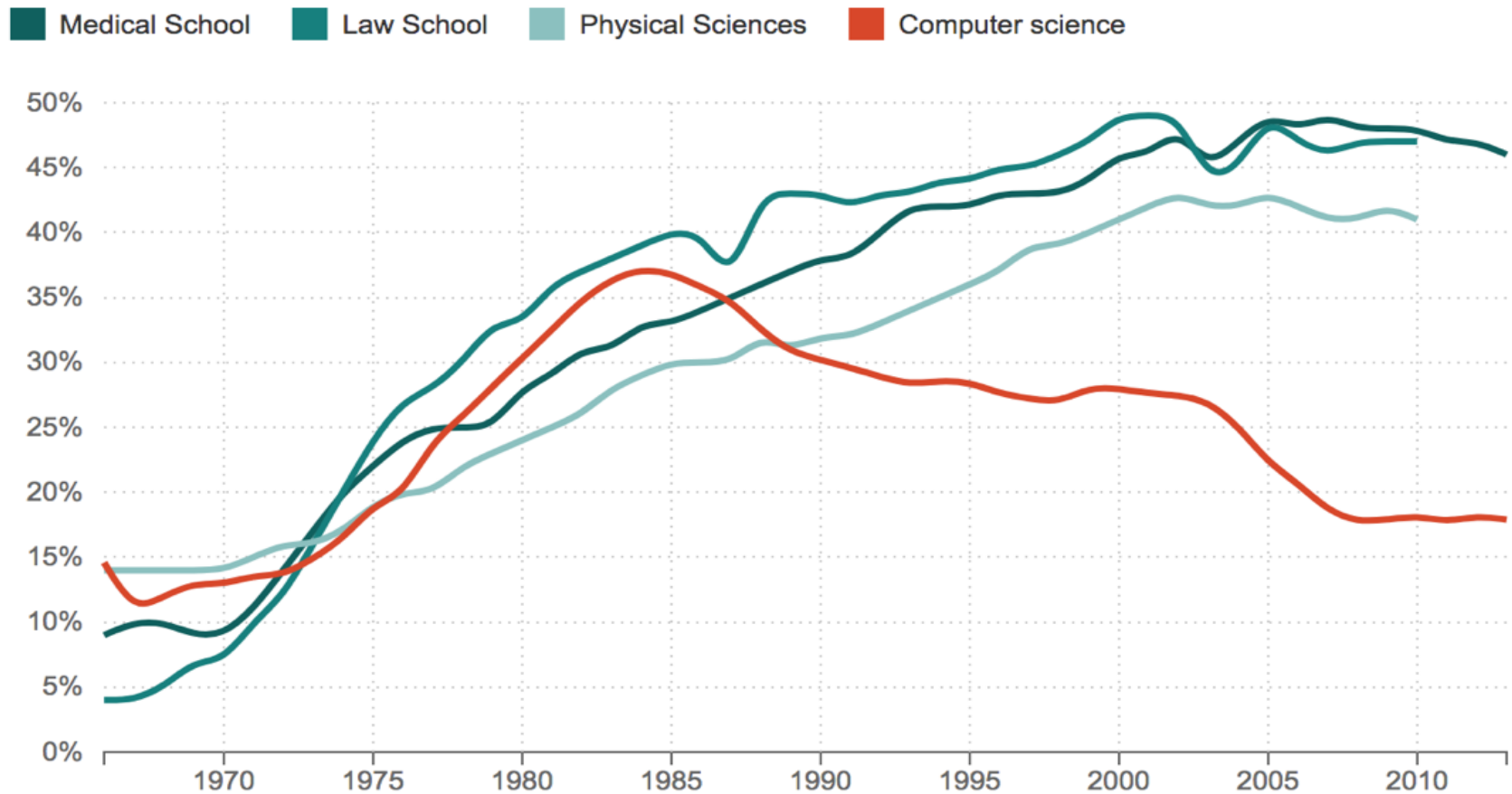


# Women in ...の趣旨と類似点

- 前提: 女性限定のイベントではない
- 所属割合の少ない女性を分野へ誘うことで、分野の発展と継続を検討している
- シンポジウム内で行っていること
  - 女性が企業内でのキャリアを語る
  - 女性研究者が自分の研究を語る
- 女性を管理職にしようキャンペーンを無理矢理展開している訳ではない
- “ユニコーン”を発見し、大事に育てる(大学入学前から)

# STEM Education Challenge

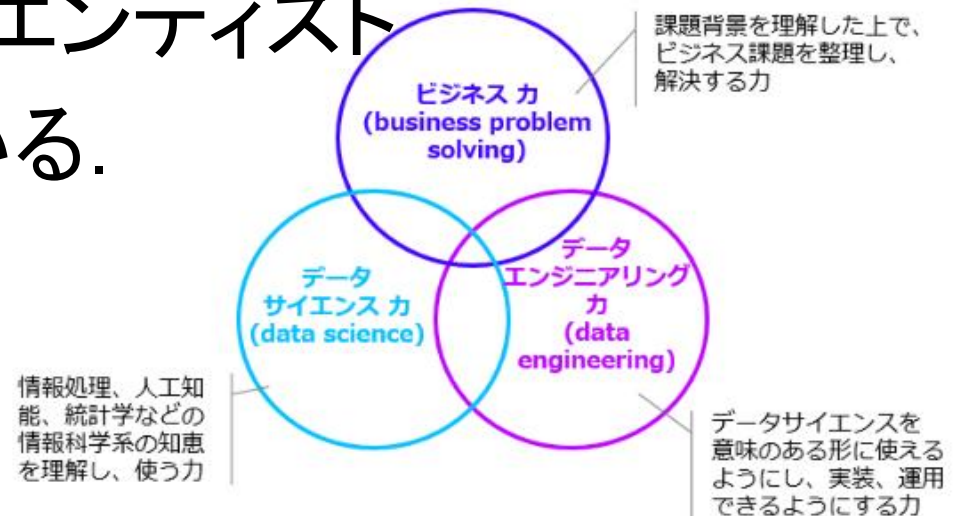
What Happened to Women in CS?



Sources: NSF, ABA, American Assn. of Medical Colleges

# WiDSと他のデータサイエンス関連の 学会・シンポジウムとの違い

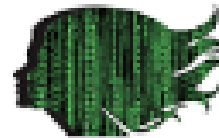
アメリカ型のオールラウンド型データサイエンティスト育成を目指すのではなく、「研究に情熱を持ち、周囲の人々が暮らしていく上でのより良い環境を構築するため」に広範囲に渡って活躍できるデータサイエンティストの育成を目指している。





# WiDS Tokyo @ YCU

- 2019年3月22日新宿ミライナタワー
- 主催:横浜市立大学データサイエンス推進センター  
<https://wids-tokyo.jp/>
- 来場者数230名程(男性36% 女性64%)
- 第1部 データサイエンス領域で活躍する女性たちによる発表とパネルディスカッション
  - データサイエンスとは何か?
  - データサイエンス分野に多くの人材を誘う際の課題
  - SDGsとデータサイエンティスト
- 第2部 アイディア・チャレンジ2019 WiDS Tokyo @ YCU25



**WOMEN IN DATA SCIENCE**  
TOKYO @  
YOKOHAMA CITY UNIVERSITY

# アイデア・チャレンジ「新しい働き方」

- データを収集・分析し、新たな価値(アイデア)を社会に提示し、安寧で豊かな、そして持続可能な『超スマート社会』を築くこと」を目指す
- SDGsを基調としたアイデアをデータから語る
- 学生の部:横浜市立大学データサイエンス学部チーム「37,420,000パターン～新しい時代の働き方～」
- 一般の部:全日空商事(株)チーム「トイレと頭の回転率を上げよう！ートイレの利用データから考えた新しいオフィスでの息の抜き方」

# WiDS Tokyo @ YCU 今後の展望

- データサイエンスの基本は「データからストーリーを語る」こととした上でデータサイエンスを定義し、広く伝えていきたい
- データを扱う研究がデータサイエンス研究ではない。従来の統計や情報の研究との違いを明らかにしていきたい
- SDGsを取り込んだ今回のWiDS Tokyo @ YCUの方向性を継続

# データサイエンスと女性を考える・・・

- データサイエンスは在宅勤務形態に向いているのではとの声
- 女性をターゲットにした市場を男性だけで解析する事に限界があるので女性の感性がデータサイエンスに欲しいとの現場からの声
- 女性ならではの細やかさはデータサイエンスに向いているとの声(幻想?)

女性データサイエンティストの行く末は？

# 本年度ワークショップの開催

第1回: SDGsとデータサイエンス

第2回: データサイエンスへの誘い (YCUオープンキャンパス特別企画)

第3回: データサイエンス教育の未来とWiDS

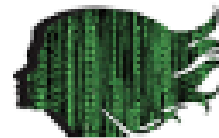
第4回: データサイエンスの未来とWiDS

← 3, 4: Co-DirectorのJudy Logan氏が来日登壇

第5回以降はアイデアチャレンジについて開催予定

# 第2回WiDS Tokyo @YCU

- 2020.3.18
- テーマ:すこやかな働き
- アイディアチャレンジ開催
- SDGsとDSの関わりを引き続き検討



**WOMEN IN DATA SCIENCE**  
TOKYO @  
YOKOHAMA CITY UNIVERSITY

# WiDS Tokyo@YCUの展開: SDGs x DS



Japan, 2019

DSはSDGs だけでなくすべての取組みへのイノベーションに繋がる！

# DS for Good



- データからストーリーを語る
- データから予兆を発見する
- データから感性へ切り込む
- DS は多くの分野を繋げて混ぜる
- DS x SDGs
- Data Ethics

“Data Science” という言葉は次の10年間に耐えうるだろうか？

Pieter Lastman, Jonah and the Whale



# 参考文献一部

Blumenstock(2013)Introduction to Data Science,  
University of Washington School of Information  
Skiena(2017)The Data Science Design Manual