

高度統計人材の質・量強化を ALL JAPANで加速

情報・システム研究機構

データサイエンス共同利用基盤施設副施設長

椿 広計

最近の問題意識 1 : 生成AI時代に統計の利活用プ ロセスは変わるのか？

夏目漱石(1901)「データといえば悉く Sense Impression也」

現代は、「Sense Impression (自然言語・プログラム言語・パターン)」
がデータとして大規模蓄積し、入出力データとなる

天文学から社会物理学へ：社会のための統計

L. A. Quetelet (1798~1874)

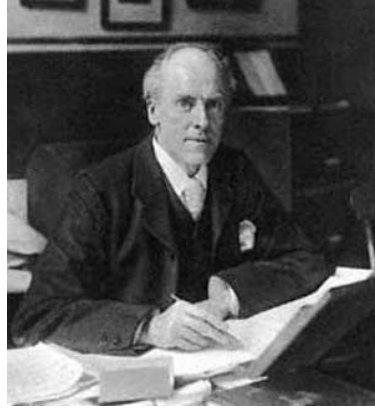
- ブリュッセルに天文台建設活動：1828設立
- 1835: 社会物理学（統計学の社会科学への適用）
 - **社会的要素を変える原因の影響**
 - **犯罪学（犯罪の社会的要因）** ・ **自殺率**など
 - **Gauss分布の平均人、BMI(Body Mass Index)の提唱**
- 1853 第1回国際統計会議議長
- 1855 国際統計協会設立へ
 - 各国官庁統計の整備推進のリーダーシップ
- 応用統計学の勃興
 - Snow：1854, 「この水道の水を飲むとコレラになる」
 - Koch: 1884 コレラ病原体としてのコレラ菌発見

数学は「科学の言語」 統計は「科学の**文法**」

K. Pearson, *The Grammar of Science*, 1892.

何でも科学になる：プロセスに基づく科学の定義

Man Gives a Law to Nature：統計科学の枠組み成立

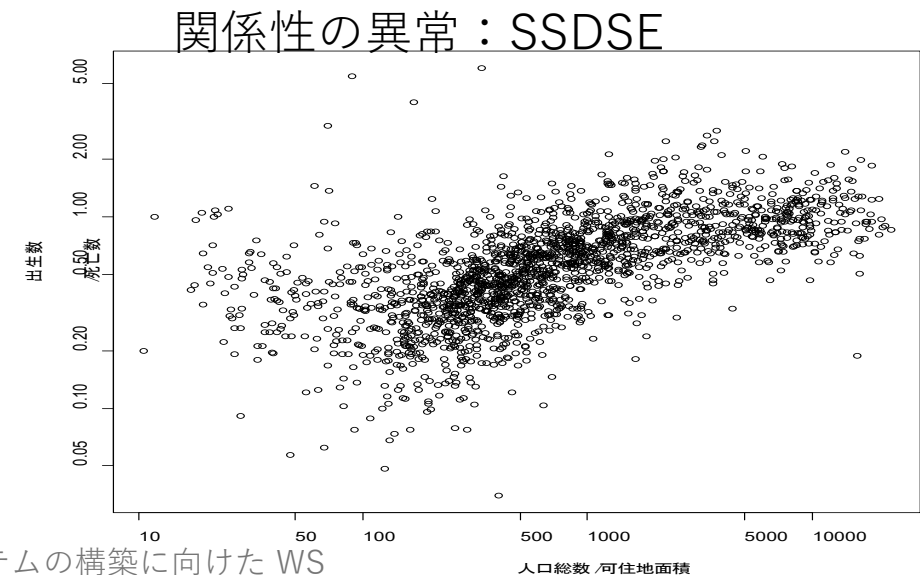
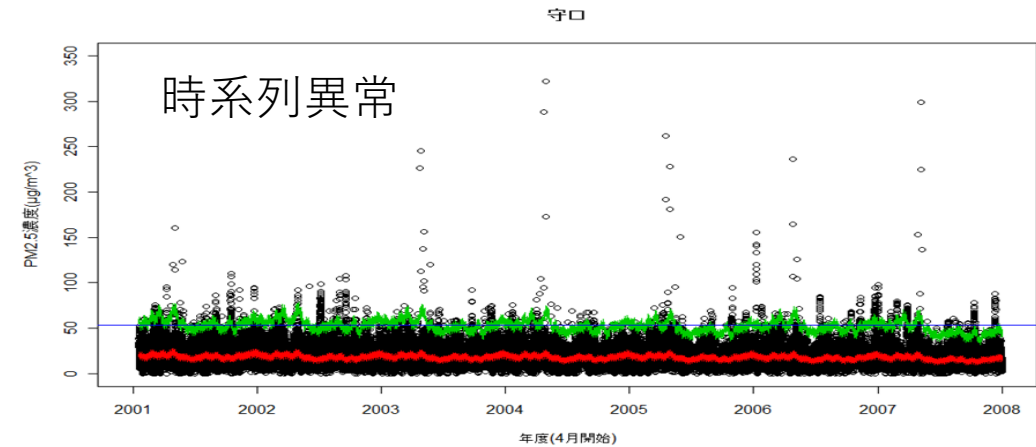


https://en.wikipedia.org/wiki/Karl_Pearson

- (統計的認識) 科学の適用範囲 (Scope)
 - 科学を特徴づけるのは対象ではなく、その**プロセス**である。
 - あらゆる知的方法を用いて真実を確かめること：認識科学「あるものの探究」
- (統計的認識) 科学的プロセス：知識マネジメントプロセス
 - 事実認識や事実関係（含む因果的法則）の認識の進化
 - **分解：Analysis**; 事実の周到な分類と事実間の関連性と順序の観察：（現象の可視化と理解）
 - **総合：Synthesis**; 創造的想像に支援された科学的法則の付与（仮説生成）
 - **妥当性検証：Validation**; “自己批判と全ての人が等しく妥当性を認めるか否かを検証（仮説検証）”
- 科学的プロセスを支援する技法：ピアソンの統計的方法開発
 - **視覚に訴える**：ヒストグラム，散布図
 - **指標化**：標準偏差，相関係数
 - **妥当性検証**：モデルの適合度検定
- 統計的認識科学の分析的側面はすでにAIによる自動化が進捗

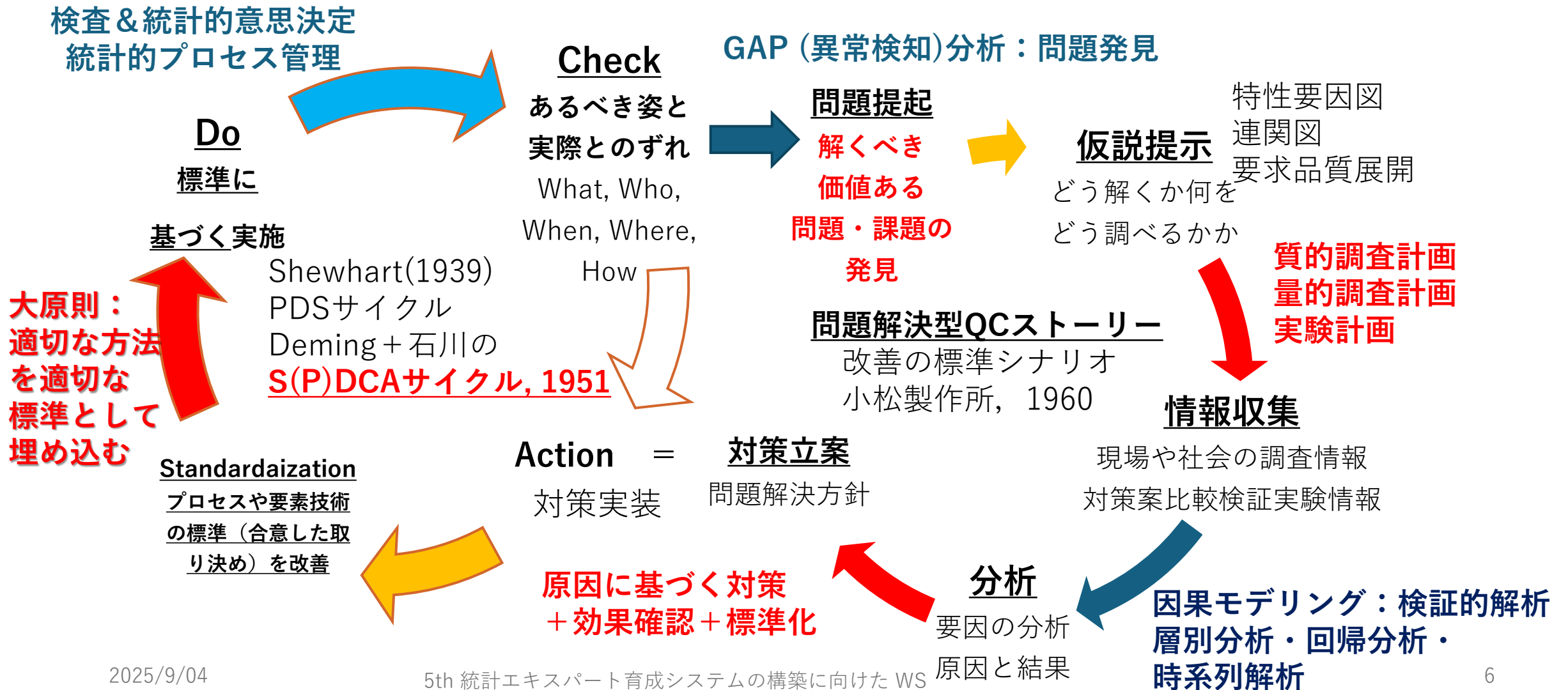
マネジメントを統計科学に: Shewhart(1891~1967)の品質管理と異常検知

- 「科学の文法」から派生した「実証科学のプロセス」
- Shewhart流品質管理：**CAPD₀**：
 - Cは検査ではなく：工程改善のトリガーの発見
 - データに基づく問題（価値）の発見は「外れ値の探索(異常検知)」から
- 問題点（価値）の抽出技法：
- 期待と実際の乖離が改善チャンス
 - Shewhartの工程管理・
 - 改善重視の世界観⇒日本が引き継ぐ
 - 検査は品質管理ではない：Deming
 - 検査を無くすことが品質管理の目的：西堀栄三郎



認識科学の文法から設計科学の文法への発展: STEM教育のE部分の本質

Nair, 2015 ISI会長講演「統計学の産業界への最大の貢献はDeming石川の統計哲学」
小笠原浩安川電機会長2024デミング賞受賞講演「データはビジネスの言語、QCストーリーは文法」

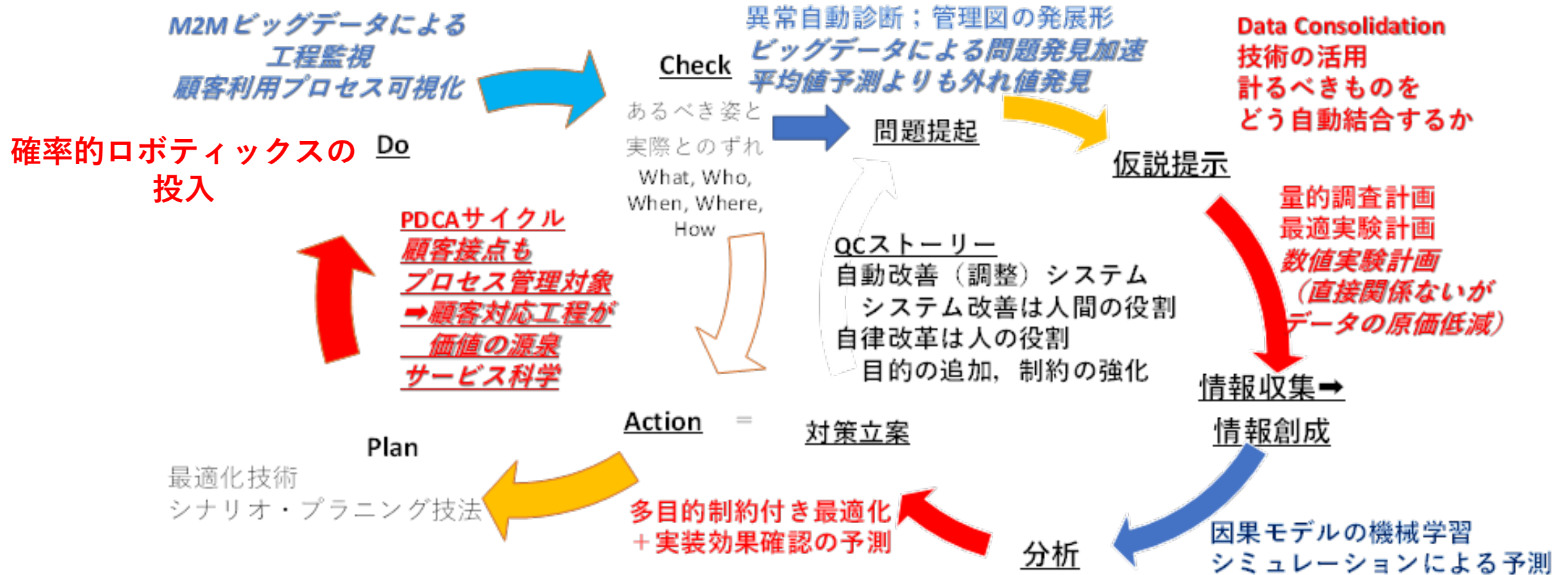


大谷(2024)文理横断によるSTEAM教育の特徴と学校教育改革, 第15回横幹連合コンファレンスB-6-1の示唆

- STEAM
 - Science, Technology, Engineering, Art, Mathematics
- E⇒EDP(Engineering Design Process)
 - ABET(Accreditation Board for Engineering and Technology)
 - 「問題を発見し特定する活動」
 - 「発見した問題を分析し、自分なりのとらえ方で課題を設定する活動」
 - 「設定した課題に基づいて解決策をデザインし、実際に解決する行動」

AI時代でも解決プロセスは維持：その教育が重要 フェイズ毎の適用技術が進化(自動化)するだけ

今、自分はどの立ち位置で人工知能を活用しているかを知る



最近の問題意識 2

統計科学とデータサイエンスに本質的違いはあるのか？

- **データ収集の究極の目的は *Action***

- 科学の経験的知識は予測で表現

- Deming(1942) On a classification of the problems of statistical inference, JASA Vol.37, Issue 218, 173-185.

- ISO/IEC 20546:2019 : Information Technology –Big data-
Overview and vocabulary

- <https://www.iso.org/obp/ui/#iso:std:iso-iec:20546:ed-1:v1:en>

- 「データサイエンス」の定義：国際投票で承認

- **データから発見, 仮説,
その検証のプロセスを通じて *Actionable* 知識を抽出するプロセス**

毎度お馴染み？米国労働統計局職業分類15-2041：数学職 統計家（15-2041） と データサイエンティスト(15-2051)

- 数学家（15-2021）と統計家
 - 特定の問題や疑問に答えるために
どのようなデータが必要であるかを判断
 - ビジネス、工学、科学、その他の分野の
実問題解決に数学理論や技法を適用
 - データを収集するために調査、実験等を計画
 - データ分析のための数学的・統計的モデルを
開発
 - データを解釈し、分析結果を専門家や
非専門家に説明
 - 統計ソフトウェアを使用してデータを分析し、
ビジネスにおける
意思決定を支援する可視化を作成
- データサイエンティスト
 - どのようなデータが利用可能で、
プロジェクトに有用かを判断
 - データの収集、分類、分析
 - アルゴリズムとモデルの作成、検証、
検査、および更新
 - データ可視化ソフトウェアを使用して
調査結果を提示
 - データ分析に基づくステークホルダー
へのビジネス提案

米国職業分類：15:計算機・数学職 15-20XX 数学職

米国労働統計局Occupational Outlook Handbook<https://www.bls.gov/ooh/>

2023/09改訂

職名	職業小分類コード	2022就労者数	2032までの増加率(%)	主要な学歴	主要就職先
保険数理士	15-2011	30,200	22 全米12位伸び率	学士	金融保険78.2%
数学家	15-2021	2,500	4	修士	政府51.0%,教育機関18.7%
OR分析家	15-2031	123,300	23 全米10位伸び率	学士	金融保険22.6% 科学技術サービス18.7%政府18.6%
統計家	15-2041	32,400	12	修士	科学技術サービス32.4% 政府24.5%、教育機関13.8%
データサイエンティスト	15-2051	202,900	36 全米4位の伸び率	学士	科学技術サービス30.1% 金融保険18.7% 企業マネジメント8.7%
マーケットリサーチ分析家	13-1061	903,400	8	学士	マネジメント、科学技術コンサル10%
金融・投資分析家	13-2051	380,500	10	学士	セキュリティ、契約、金融投資25%
エコノミスト	19-3011	17,500	5	修士	政府28%,マネジメント、科学技術コンサル16%
調査研究者	19-3022	9,700	0	修士	専門科学技術サービス32%

15-2041.1 生物統計家:博士比率が大きい（統計家：Healthcare and Social Service 7.0%:今後10年で10.5%増）
全職業での10年増加率は4%

労働統計局（トランプリスクの統計への影響）WEBより (<https://www.bls.gov/ooh/math/>)

• 統計家（15－2041：数学職）

- 統計家雇用増加が、ビジネス、**医療、政策の意思決定**に情報を提供するための統計分析のより広範な使用によってもたらされると予想。
- デジタルに保存されたデータの量は、人々や企業がオンラインでビジネスを行い、ソーシャルメディア、スマートフォン、その他のモバイルデバイスを使用し続けるにつれて、10年間で増加、その結果、企業には収集された大量の情報やデータを分析するための統計家をますます必要
- 統計分析は、企業のビジネス プロセスを改善、新製品設計開発、潜在的顧客に製品を宣伝するのに有用
- **統計家は通常、修士号が必要**、一部のエントリーレベルのポジションでは学士号を持つ候補者を受け入れる場合
- **統計学を専攻する学生は、コンピューターサイエンス、ライフサイエンス、物理科学などの別の分野のコースを受講可能、これらのコースは、学生がさまざまな業界で働く準備をするのに有用、生物学、化学、健康科学のコースワークは、医薬品や農産物の試験。物理学は、品質向上に携わる製造業に有用**

• データサイエンティスト（15-2051、数学職）

- データサイエンティスト雇用増加は、**データ主導の意思決定に対する需要の増加**に起因
- 利用可能なデータの量とそのデータの潜在的な用途は、10年間で増加、その結果、組織は、収集された大量の情報とデータマイニングのために、より多くのデータサイエンティストが必要になる可能性
- データサイエンティストの分析は、**組織が情報に基づいた意思決定**を行い、ビジネスプロセスを改善し、新製品を設計および開発し、製品をより適切にマーケティングするのに有益
- **データサイエンティストの職業に就くには、少なくとも数学、統計学、コンピューターサイエンス、または関連分野の学士号が必要**
- ただし、一部の雇用主は、候補者が修士号または博士号を取得していることを要求または優先
- **データサイエンスにはアルゴリズムと統計手法の使用が含まれるため、学生は数学と統計学の広範な学習が必要**
- データサイエンティストになりたい**高校生は、線形代数、微積分、確率と統計などの科目の授業を受講する必要**
- 大学レベルでは、**数学と統計に加えて、コンピューターサイエンスのコースが重要**
- 学生は、データ指向のプログラミング言語、および分析を提示するための統計、データベース、およびその他のソフトウェアを学習する必要

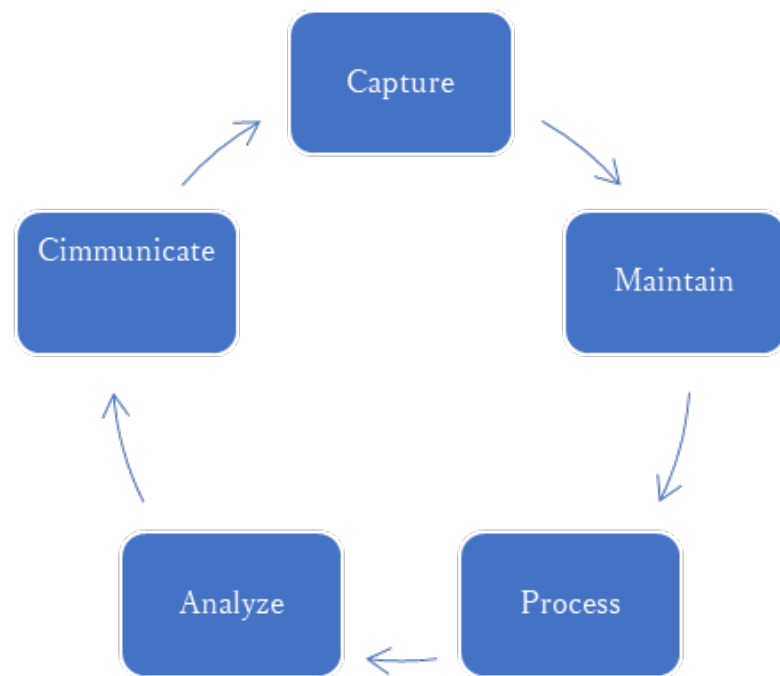
UC Berkeley データサイエンス学部

2025/08時点で全米1位

- <https://cdss.berkeley.edu/dsus/data-science-prospective-students>

- コンピューティング、データサイエンス、社会学部の一部として、私たちのカリキュラムは、コンピューティングと統計の教育と、社会的および倫理的影響を強調する実際の例を組み合わせています。
- 大学全体の学部の教員によって作成および指導される私たちのカリキュラムは、データ分析とデータ主導の意思決定の学際的使用についての批判的理解を促進します。
- この包括的なアプローチにより、卒業生は、データ集約型戦略による産業や分野の進歩、人工知能イノベーションの拡大、社会の進歩への貢献など、緊急の課題に取り組む準備が整います。

2年前の米国データサイエンス高等教育 UC Berkeley: オンライン大学院MIDS



データサイエンスのライフサイクル
これができるエキスパートを育成
プロセスを担える人材教育

- **Capture**

- データ取得、データ入力、信号入力とデータの抽出

- **Maintain**：統計学の教程ではこれまで軽視されていたこと：前処理

- データウェアハウスの作成、分析可能なデータの編成（クレンジング）、データ処理、データアーキテクチャー

- **Process**

- データマイニング、クラスタリング／分類、データモデリング、データの要約

- **Analyze**

- 探索的分析と検証的分析、予測分析、回帰分析、テキストマイニング、質的分析

- **Communicate**

- レポートニング、可視化、ビジネスインテリジェンス、意思決定

UC Berkeley社会人大学院MIDSの生成AIの影響

同様全米1位:この2年で自然言語処理・生成AIを補強
統計の本質、データのデザインは以前基盤となる骨格

- 基礎コース
 - データサイエンスプログラミング入門
 - **データと分析のための研究デザインと応用**
 - **データサイエンスの統計学（研究の計画、統計分析）**
 - データエンジニアリングの基礎
 - 応用機械学習（**実験計画、特徴量・予測と説明**、ネットワーク分析）
- 上級コース
 - **実験と因果推論**
 - データの背景：人間と価値観（倫理と法、政策分析）：すごく重要！
 - **離散応答・時系列・パネルデータの統計的手法**
 - 大規模機械学習
 - 深層学習による自然言語処理
 - **データの可視化（探索的データ分析、プレゼンテーション、人間の認識のデザインなど）**
 - 機械学習システム工学
 - コンピュータビジョン
 - 最新のデータアプリケーション
 - **生成AI：基盤・技術・課題・機会**

All Japan の取り組みとは？

全国民が統計家になれる機会が到来

- 新T型人材：専門知識（縦型知識）と横断型知識
 - 統計学科がどこでもある国：一般統計家から分野専門統計家に
 - 日本：当面分野専門家が統計家に：統計エキスパート人材育成プロジェクト
- 生成AIの影響
 - 丸山宏先生：Software 2.0；プログラムは書く時代から探す時代
 - 急速に実現
 - 本来のプログラミング教育とは？
 - ビジネススクールのには、適切な仕事の流れを論理的に記述できること
 - 人間の役割
 - 形式的なプログラミング教育や統計手法の技術側面教育の破綻
- 椿(2000)コンピュータ vs 統計家、日本統計学会会報No.105
 - ドライバー：問題の表現方法
 - 燃料：問題を解くために必要な情報
 - エンジン：問題の解決方法；コンピュータの得意とするところ：統計家は駆逐
 - **すべての人が自分の知識を必要程度に統計モデル化することが可能になれば統計家は不要**
 - **すべての人が統計家となる時代」⇒25年経ってそれが実現できる環境としての生成AIが整っている**
 - **その教育現場での使い方の標準とそこにSDCAを回すのは必須**
- 日本の産業界のかつての特徴：全社的品質管理（専門家が行う品質管理と差別化）
 - **ビジネスの文法；全社員がQCストーリーに基づく統計的問題解決が可能**
 - **目指すは；全国民が必要に応じて統計的問題解決を可能にすること**
 - 初中等統計教育⇒大学での統計教育（統計家である専門家の輩出）⇒大学院での統計教育（専門統計家の輩出）

石川薫が目指したコト： 人間性尊重の経営

- 全員参加の経営：QCサークル（小集団改善活動）というボトムアップ環境の意義
 - 全ての従業員が自主的に物事を考え、
 - 実行して成果も確認できる筋道を与えたことで、
 - 喜びや達成感を味わい成長し、
 - メンバーの自己実現が達成するようなQCサークルという環境を整備したことが大きなコトづくり
- 1960年第1回QCサークル大会：現場の女性職長：材料を変えて大きな効果
 - 質問：部課長はなかなか私たちのいうとおりにやってくれない。
材料を変えようと思ってもかわらない。それをあなたはどのようにしてできたのか？
 - 回答：だから私は急がしくてしょうがないんだ。
私の仕事の半分はデータでもって上司の尻をひっぱたくことです。
- 石川(1980)品質管理学会誌
 - Q,C,Dは経営の第二の目的であり、第一の目的は人間性尊重の経営である。
 - 人間性尊重の経営とは、**人から言われたからではなく、自主性、自分の意志により、自発的に行い、そして、頭を使ってよく考え、人間の無限の能力を発揮させる経営**である

日本の絶頂期の欧米の反攻：米国労働省SCANS Report (1992) The Secretary's Commission on Achieving Necessary Skills (SCANS) **日独の労働者の70%が持つ力量!：教育の全面改革**

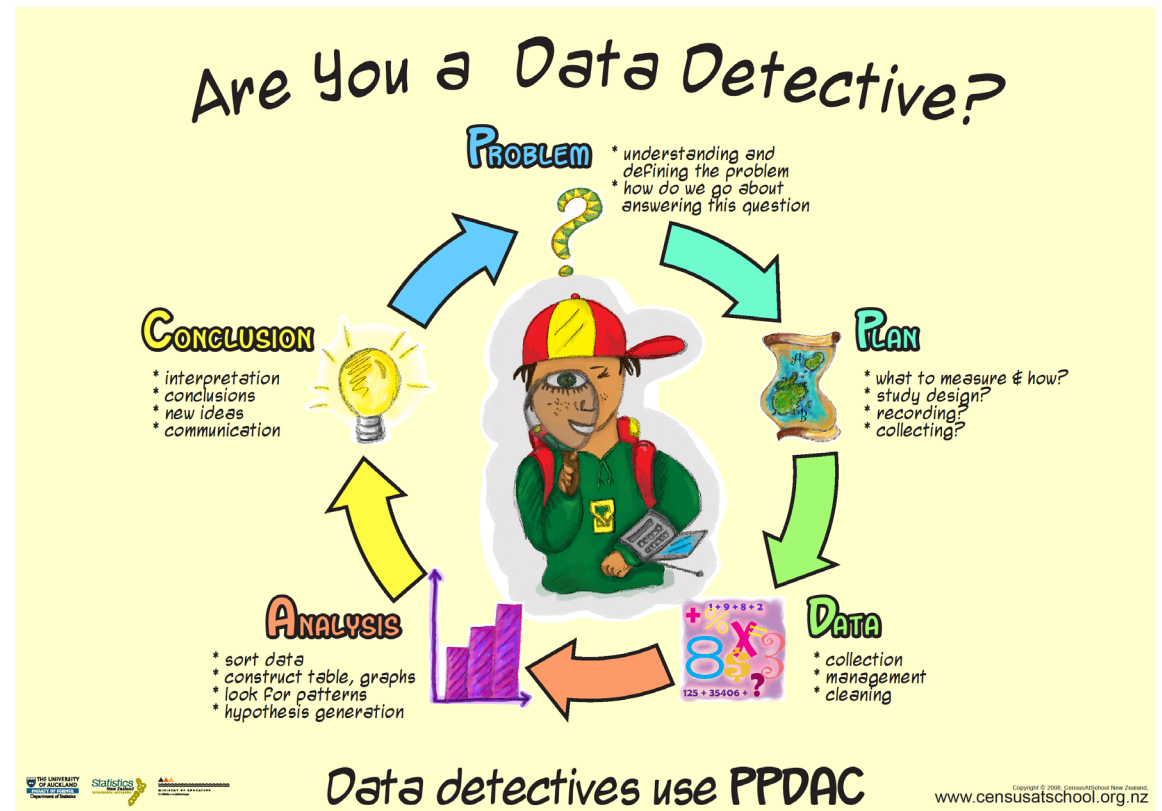
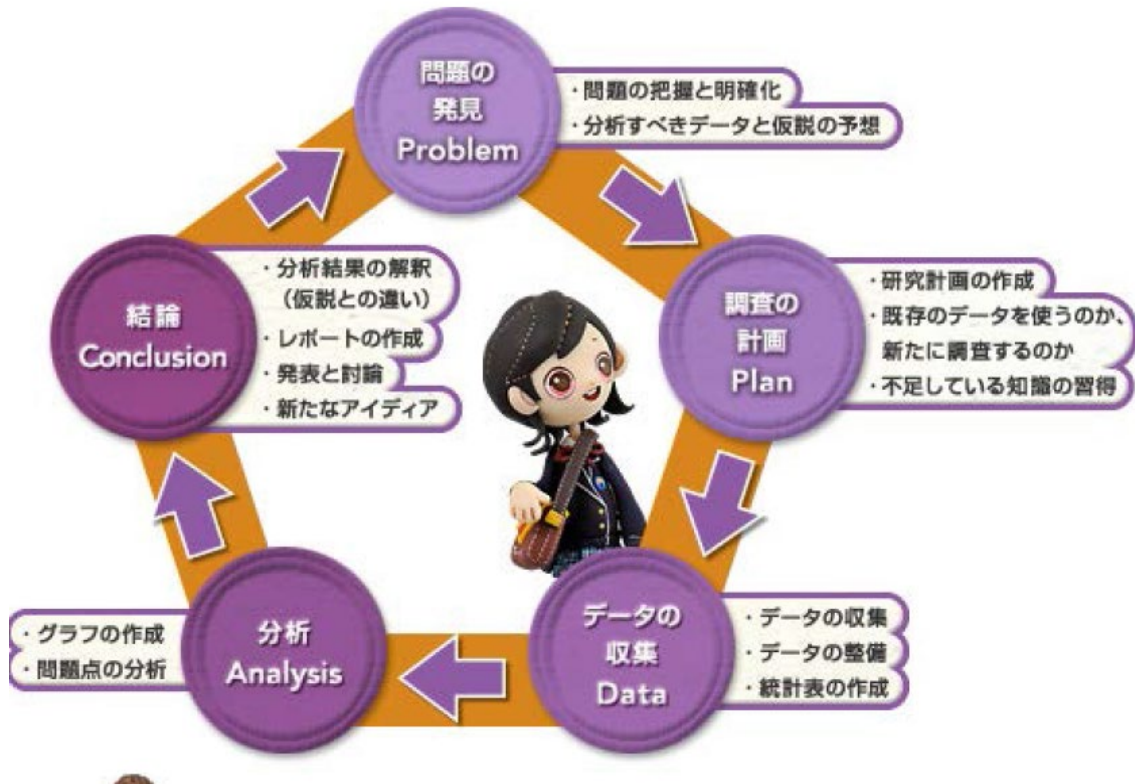
“Learning a Living: A Blueprint for High Performance — a SCANS Report for America”. <http://wdr.doleta.gov/SCANS/>

- 日本の奇跡に学べ
→ **自立的に問題を解決することができる力量**
- 20年間に欧米が打った3つの方策
 - **初中等教育統計的問題解決プロセス教育**
→PPDACサイクル
 - Problem, Plan, Data, Analysis, Conclusion
 - 2010年日本の初中等教育に逆輸入
 - **大学における統計教育強化**
 - **企業等への国家品質賞制定** (MB賞)
- SCANS Skills and Competenciesの定義
 - 基本スキル、思考スキル、資質
 - <https://programs.honolulu.hawaii.edu/intranet/sites/programs.honolulu.hawaii.edu.intranet/files/general-college-mtg-2009-08-20-scans.pdf>
- 日本はここを巻き返すためにT型統計教員を分野ごとに育成：彼らは機会があればΠ型、剣山型人材として他分野で専門統計家育成
- 産業界：シックスシグマ活動
- Six-Sigma: DMAICサイクル
 - Define, Measure, Analyze, Improve, Control
 - 最近が開発系サイクル：Design For Six Sigma
 - ブラックベルト国際認証規格；英国統計学会
 - 試験による専門家認定の中止
 - ISO 18404:2015 **シックスシグマ及びリーン実施に関する主要専任者の能力と組織の適格性**
 - 2025改訂開始：英国MS規格あるいはShallを含む認証用規格とすることを画策
- 産業界の統計のProfessionalの認証
- 日本の理工系・農学系大学カリキュラムを認定
 - JABBE (ISO型)：日本技術者教育認定機構への期待
 - 専門職大学院の認定 (ABEST 21：アジア太平洋地域)
- 統計学・データサイエンスのカリキュラムの認定機関と国際認定

国際初等・中等教育への展開と逆輸入：PPDACサイクル：データを使った問題の発見から解決まで
www.censusatschool.org.nz: オリジナル

総務省統計局「なるほど統計学園高等部」：逆輸入された問題解決プロセス
<http://www.stat.go.jp/koukou/howto/process/index.htm>

現在：日本の小学校5年の算数教科書に収録



AI時代の統計家型国民の役割

- ◎**価値命題（技術を含む施策の目的と制約）の明確化とAIへの伝達：マネジメントの役割**
 - 倫理・法や標準の枠組みによる制約は人間が制御
 - **標準等の妥当性についてS(≒P)DCAサイクル発動の判断：統計家であるマネジメントの役割**
- ○**新たな「データを創り」それに基づくActionで変革：専門統計家の役割**
 - 実験あるいは能動的計測（存在しない技術情報などを新たに創る）
 - 実実験でなくても良い：モデリングや分析はAIの支援を受けられる
- ○**Actionの経済価値評価と経済価値を創出する知識価値・感情価値の評価を行う意思：統計家である分野専門家の役割**
- ○**Actionの価値を生み出すシステム（階層的因果関係）の設計：専門統計家と分野専門家の協業**
 - システムを何で制御するか、何が制御できない環境要因なのかを明確化
- △**AIの品質マネジメント：日常的マネジメントはAIの支援を受け効率化可能**
 - 設計・**入力情報（内容の妥当性とデータの代表性・偏り）**・**AI出力（生もの）の品質マネジメント**
 - AI利用のプロセス標準
 - 事実判断（真偽）の場合、**価値判断**の場合・順問題の場合、**逆問題**（結果から原因を予測）の場合
 - 解釈のプロセス標準：Blackbox説明責任

一般教育分野動向：統計的探究の機会増加

- 初中等教育
 - 学習指導要領改訂
 - 機会：探究型学習（総合的学習）の実質化
 - その中での算数・数学の統計学充実
 - 記述統計
 - 統計的推測
 - 高校情報Ⅱ：高度データ解析体験実践
 - レベル的には大学理工系専門科目
 - 教員用テキストの文科省開発
 - 機会：教科書の電子化・一人一台端末
 - **課題：探究型学習を指導できる教員不足**
- 大学教育：専門課程は探究型だが?!
 - 内閣府・文科省・経産省
 - 数理データサイエンスAIプログラム認定
 - リテラシーレベル
 - 応用基礎レベル
 - 機会：大学標準カリキュラム策定事業
 - 仮題：教養と手法の教育
 - **問題解決プロセス体験の不足**
 - 卒論に入って統計の力を知る

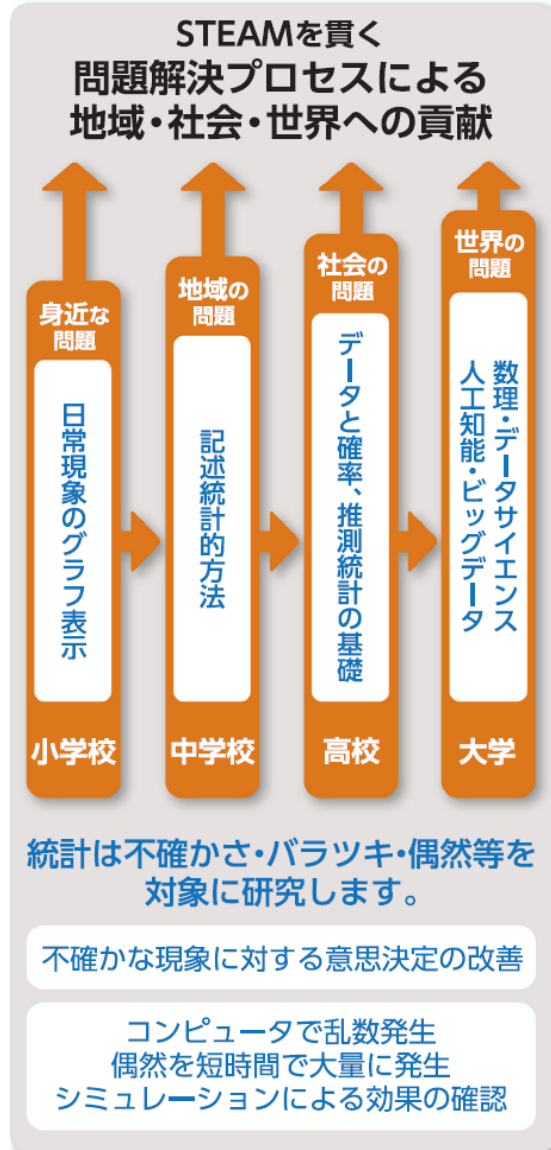
高度探究活動が前提の高専教育に学べないか

- 学生自らの研究における問題解決
 - **専門分野の新しい標準の確立と改善**
 - SDCAサイクルのプロセス教育
 - データに基づく問題発見
 - 理論的探究：技術仮説の明確化
 - データ（実験）に基づく検証的探究
 - 少なくともここには標準的プロセス
 - 技術系問題解決に特化した体系的教育
 - 異常検知による問題発見技法
 - 解決仮説網羅的整理の技法
 - 特性要因図・品質機能展開・連関図等々
 - 問題の原因の探求技法
 - 因子の分類：制御因子×標示因子
 - 予測モデル当てはめ：回帰モデルから深層学習モデルまで
 - 対策の最適化技法
- **目指すのは：T型人材の体系的輩出**

社会に活かす統計の考え方

デジタル社会に求められる
データへのリテラシー(素養)

データに基づき
社会問題を解決するプロセス



おわりに：

日本沈没を意識した人材の海外流出抑止

- 科学技術立国の再生・産業競争力の再生
- **日本に資する人材育成のための人材の育成**が急務
 - 初等中等教育も高等教育も研究者育成も皆同じ問題を抱えている
 - 欧米流横型専門職人材育成拠点が滋賀大DS学部設立前皆無だった
 - 日本人ファーストでは必ずしもないのですが
- そのために、自身の問題を統計的探究マインドを存分に発揮して解決しようとする（Aspiration Ethics, Virtue, **Tugend**を有する）学生・生徒の向上心をコーチングできる教員・教授の育成が初中等教育でも高等教育でも求められている
 - 統計エキスパート人材育成プロジェクトはその重要なピースの一つ

ご清聴有難うございました

すべての子供たちにTugendを勧める

emphelt euren Kindern **Tugend**

Beethoven(1802) Heiligenstadtの遺書より