京都大学
KYOTO UNIVERSITY

CIREDS

全学共通教育

ビッグデータの時代
データ科学を学ぼう

国際高等教育院附属
データ科学イノベーション教育研究センター　山本章博
公式Twitter [@cireds2020]

2023年4月

その裏側には、データサイエンス。

お手元に届いていると思います

今話題のデータサイエンスは、京都大学全学共通教育で学ぶことができます

ともだちの始まりの裏側にも

健康維持の裏側にも

最高の買い物の裏側にも

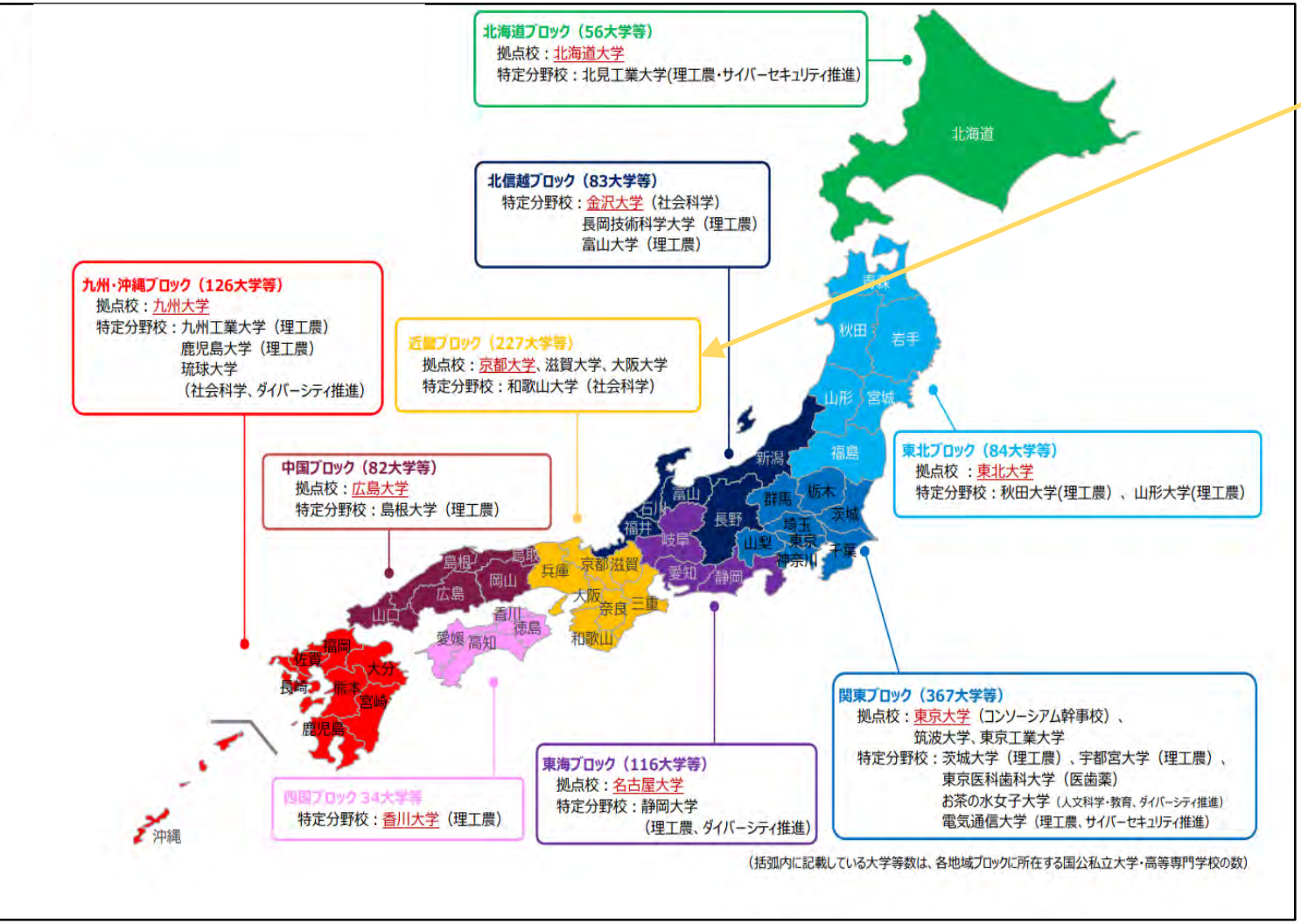
あの宅配の裏側にも

内閣府・文部科学省はデータ科学(データサイエンス)の教育強化を進めています

デジタル社会の「読み・書き・そろばん」である「数理・データサイエンス・AI」の基礎などの必要な力を全ての国民が育み、あらゆる分野で人材が活躍

	主な取組	育成目標【2025年】
エキスパート	**先鋭的な人材を発掘・伸ばす環境整備** ■ **若手の自由な研究と海外挑戦**の機会を拡充 ■ 実課題をAIで発見・解決する学習中心の**課題解決型AI人材育成**	トップクラスの育成 100人程度/年 2,000人/年
応用基礎	**AI応用力の習得** ■ **AI×専門分野**のダブルメジャーの促進 ■ **AIで地域課題等の解決**ができる人材育成（産学連携） **認定制度・資格の活用** ■ 大学等の**優れた教育プログラムを政府が認定**する制度構築 ■ **国家試験**（ITパスポート）の見直し	25万人/年 （高校の一部、高専・大学の50%）
リテラシー	**学習内容の強化** ■ 大学の**標準カリキュラム**の開発と展開 ■ 高校におけるAIの基礎となる**実習授業**の充実 **小中高校における教育環境の整備** ■ 多様な**ICT人材の登用**（高校は1校に1人以上、小中校は4校に１人以上） ■ **生徒一人一人が端末**を持つICT環境整備	**すべての大学卒業生が修得しているべき素養** 50万人/年 （大学・高専卒業者全員） 100万人/年 （高校卒業者全員） （小中学生全員）

数理・データサイエンス教育強化拠点コンソーシアム



（括弧内に記載している大学等数は、各地域ブロックに所在する国公私立大学・高等専門学校の数）

京都大学は拠点校
R4年度より近畿ブロック代表校

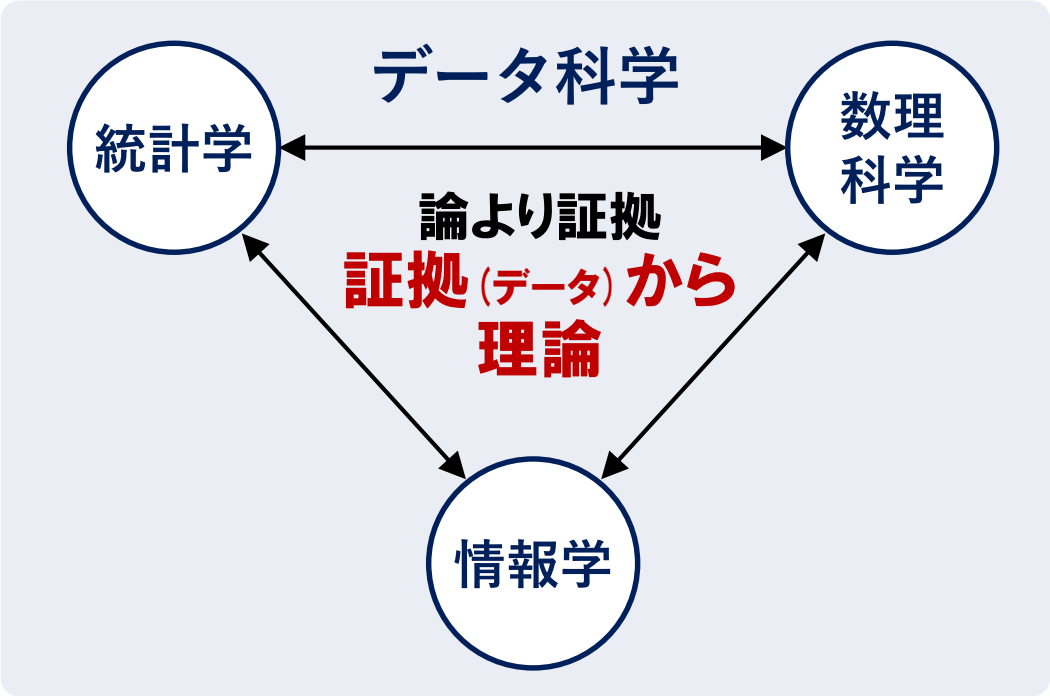
拠点校：滋賀大学，京都大学，大阪大学
特定分野協力校：和歌山大学（社会科学）
連携校：[国立] 三重大学，京都工芸繊維大学，神戸大学
[公立] 京都府立大学，福知山公立大学，大阪公立大学，
兵庫県立大学
[私立] 皇學館大學,鈴鹿医療科学大学，京都光華女子大学，
京都産業大学，京都女子大学，京都精華大学，
京都先端科学大学，京都橘大学，京都ノートルダム女子大学，
同志社大学，同志社女子大学，立命館大学，龍谷大学，
大阪医科薬科大学，大阪大谷大学，大阪経済大学，
大阪工業大学，大阪歯科大学，大阪成蹊大学，
大阪電気通信大学，関西大学，近畿大学，
四天王寺大学・四天王寺大学短期大学部，摂南大学，
阪南大学，桃山学院大学，大和大学，神戸学院大学，
兵庫大学，兵庫医科大学，畿央大学，高野山大学，
大阪キリスト教短期大学

http://www.mi.u-tokyo.ac.jp/consortium/

データ科学（データサイエンス）って何？

データを収集してコンピュータで管理し、
数理的手法を用いて分析することにより結論を導き、将来の推測を行う学問

情報学・統計学・数学（数理科学）の融合分野



医療情報	気象
創薬	文献解析
経営分析	地球科学
建築設計	生命科学
金融	人工知能
政策決定	画像音声

**データ科学イノベーション教育研究センターは全学共通教育での
統計学を中心としたデータ科学科目を整備・設計・実施する組織です**

Twitter公式アカウント［@cireds2020］

統計学は数学の一部ではありません

日本において令和3年度以前の高校までの学習指導要領では，
統計的手法は数学の一部として組み込まれています

- 高校では数学Ⅰ「データの分析」，数学B「確率分布と統計的推測」
- 歴史的には，著名な数学者が統計的手法の構築に大きく貢献したことは事実です

しかし

統計学は数学を基盤とする別の学問です

- 「日本統計学会」と「日本数学会」は別組織です

さらにITの発達により，情報学との関係が密接になり，

現代は 統計学 × 情報学 × 数学 ＝ データ科学 のどれもが重要です

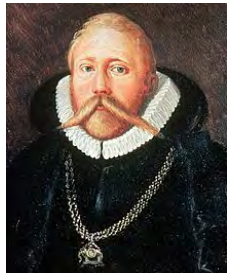
データと物理法則

物理法則の源泉はデータである

16世紀頃まで惑星の運行は謎だった

■恒星は一定の動き：同じ時刻に観測すると位置が少しずつずれていく

■惑星はときどき逆行するなど単純ではない動きをする



Tycho Brahe
(1546-1601)[*1]

Johannes Kepler
(1571-1630) [*2]

ティコ・ブラーエによる
火星の運行に関する**高精度観測データ**

ヨハネス・ケプラーの発想
“惑星の軌道は楕円ではないか?”

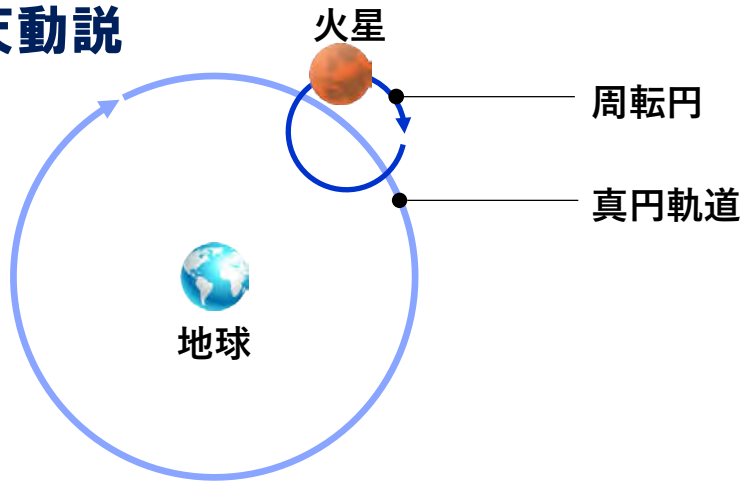
地動説の下で，火星の軌道が楕円だと仮定すると
データが説明できることが分かった！

ケプラーの法則（第1法則）

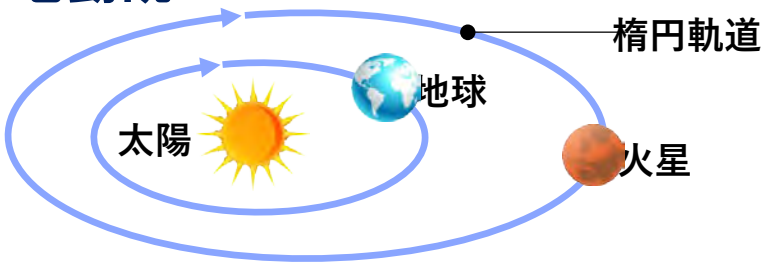
惑星は軌道は楕円である！

さらにニュートンの万有引力の法則へと発展

■ 天動説



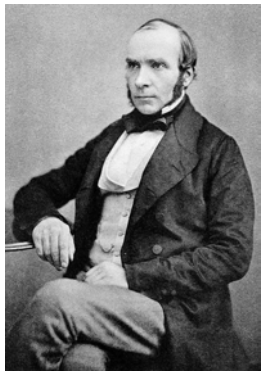
■ 地動説



- 天動説の場合は，真円軌道に小さな周転円を追加するモデル
- 地動説の場合は，真円軌道だと誤差が大きい

データ科学と医療・政策

1854年ロンドンにおいて特定の井戸周辺にコレラ多発

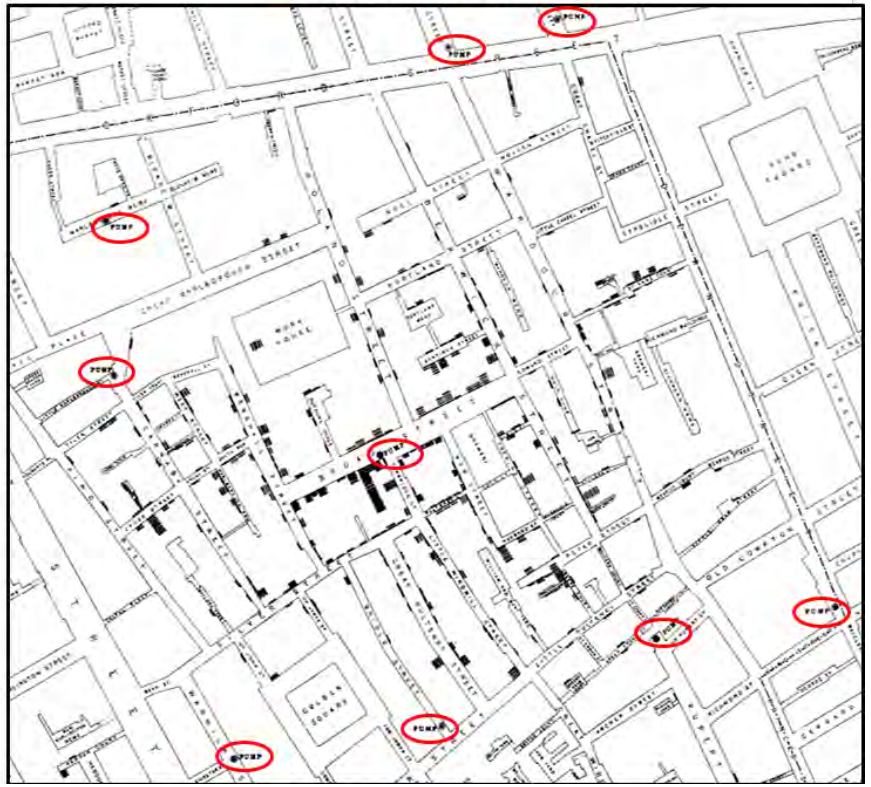


John Snow
(1813-1858) *3

データを集めることにより同一の区域でも、水道供給会社によってコレラ発生率に大きな違いがあることを発見

コレラ対策という政策が進む

赤枠がポンプの場所



*4

流行ピーク、8割がインフルエンザの外来

インフルエンザ

検査	あり	なし	計	
陽性	1440	8 偽陽性	1448	陽性的中率 99.4%
陰性	160 偽陰性	392	552	陰性的中率 71.0%
計	1600	400	2000	
	感度 90%	特異度 98%		

検査結果が陰性でも感染している確率は，全体の感染率と感度から計算される

数学

$$P_B(A) = \frac{P_A(B)\,P(A)}{P(B)}$$

データ科学と人工知能（AI）

2011年衆議院会議録作成システムに京都大学で開発された音声認識技術が採用

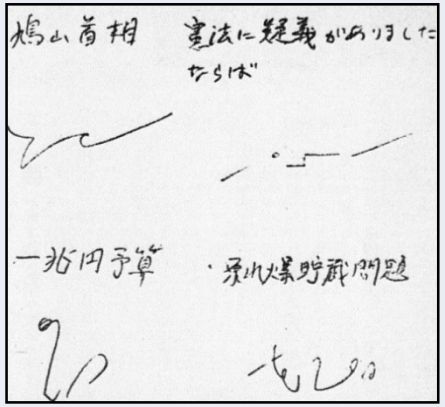
従来

衆議院本会議

*5

新システム

*6



*7

事務事業の仕分けをはじめとした
ご提案をいただきました.

情報学 膨大な量の国会質疑音声記録と国会議事録のデータベース

統計学 音声記録と議事録の違いを数式化

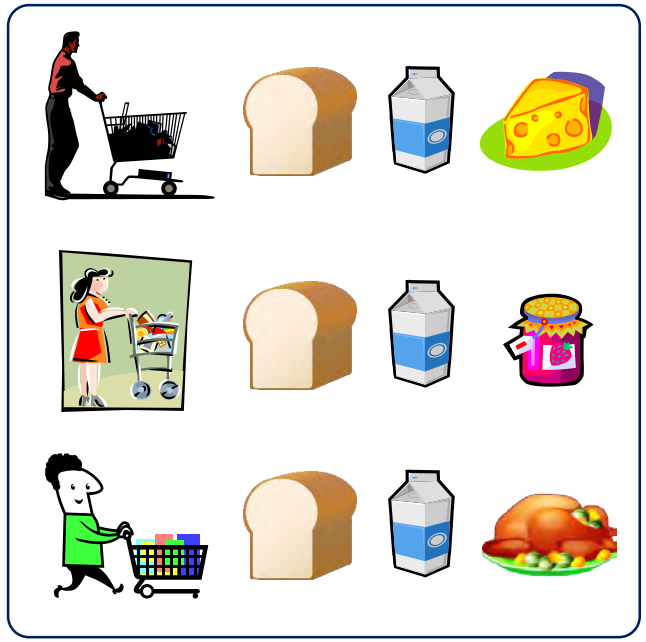
数 学

$$P_B(A) = \frac{P_A(B)\,P(A)}{P(B)}$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt, (n = 0, 1, 2, 3, \cdots)$$

データ科学とビジネス

■ お店の販売データ（POSデータ）分析



販売・仕入れ・商品開発などにデータを活用

- ■ 一緒に売れる商品の組み合わせは？
- ■ 最近よく売れている/売れ残りがちな商品は？
- ■ これから売れそうな商品は？

データ科学分野・情報群科目の例

	データ科学分野 （自然群）	情報群
リテラシーレベル 情報・統計に関する 一般常識・基本知識	■ 統計入門 ■ 数理・データ科学のための数学入門Ⅰ・Ⅱ ■ 数理統計	■ 情報基礎 ■ 情報基礎演習 ■ 情報と社会 ■ 情報基礎実践 ■ プログラミング演習（Pythonなど）
応用基礎レベル データから分布などの 情報を抽出する力	■ 統計と人工知能 ■ データ分析演習Ⅰ・Ⅱ ■ データ分析基礎	■ 情報企業論 ■ 情報ネットワーク

データサイエンススクール（課外）

統計入門

■全学部・全回生向（医学部医学科1回生は必須）

※学部・学科によっては「数理統計」が指定・推奨されているので「履修の手引き」「履修要覧」などをよく確認してください.

■数理・データサイエンス・ＡＩ（リテラシーレベル）モデルカリキュラムに準拠

■合格者は修了証を受領可能

■より広範な学生が統計学の基本的な考え方を理解する

■数学的な厳密な理論を理解することよりも，データを扱うエンドユーザーとしてデータの性質に応じた適切な分析方法を選択できるようになることを目指す

■統計解析ソフトを用いたデータ分析を自習形式で習得する

修了証番号：1234567890

CIREDS

修　了　証

生年月日　2022年01月01日

氏　名　mdash事務局

上記の者は、下記の科目を修得し、本学の「数理・データサイエンス・AI 教育プログラム（リテラシーレベル）」を修了したことを認めここに証します。

記

令和３年度　「統計入門」

以上

令和４年３月31日

京都大学国際高等教育院長　宮　川　恒

統計と人工知能，データ分析基礎，データ分析演習Ⅰ,Ⅱ

■数理・データサイエンス・ＡＩ（応用基礎レベル）モデルカリキュラムに準拠

■4科目中1科目以上の合格者は修了証を受領可能

統計と人工知能

■「統計入門」の発展として，より高度な内容を習得

■人工知能・機械学習の基本的な考え方や手法を習得

データ分析基礎

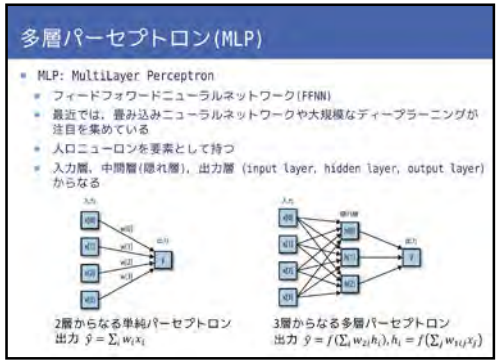
■「統計入門」より先の分析法や推測アルゴリズムを習得

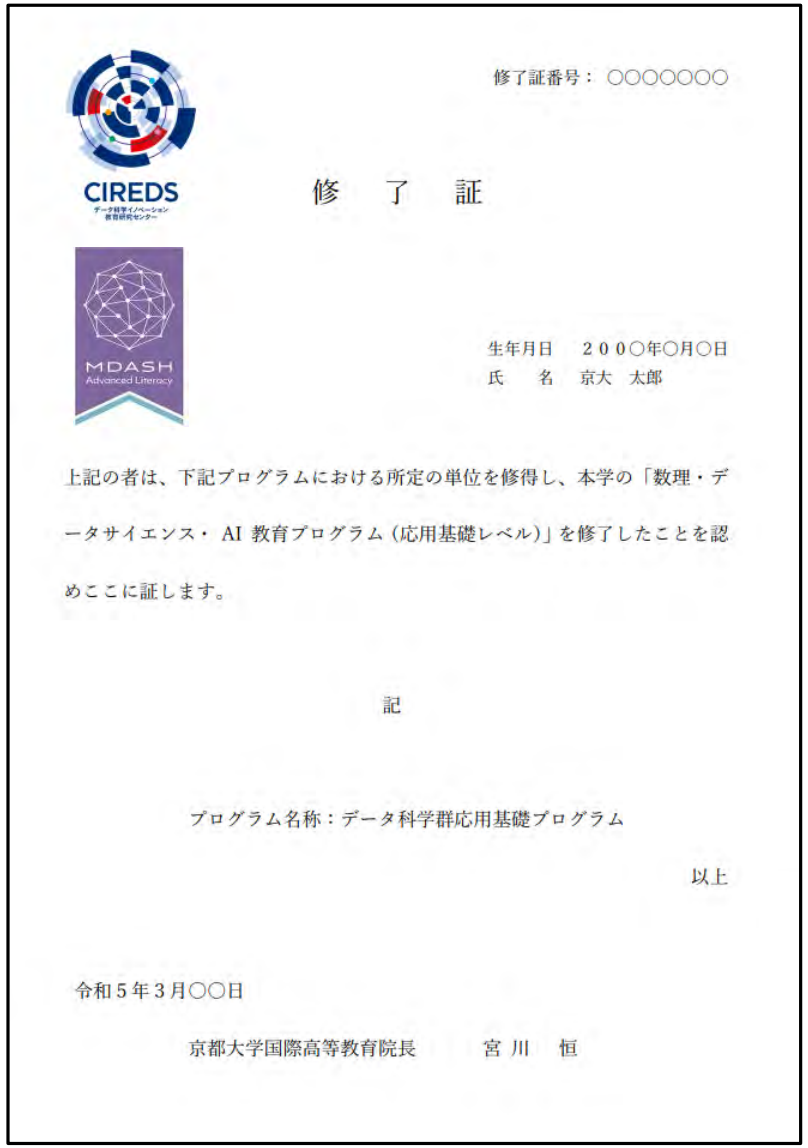
■分析の使いどころを統計解析ソフトを用いて実践的に習得

データ分析演習Ⅰ,Ⅱ

■大規模なデータ分析の演習を通じてデータ分析の基礎の体得を目標

■今後専門とする分野に合わせて幅広い選択ができるように，異なる内容の科目を揃える





修了証番号：○○○○○○○

CIREDS

修　了　証

MDASH Advanced Literacy

生年月日　２００○年○月○日
氏　　名　京大　太郎

上記の者は、下記プログラムにおける所定の単位を修得し、本学の「数理・データサイエンス・ AI 教育プログラム（応用基礎レベル）」を修了したことを認めここに証します。

記

プログラム名称：データ科学群応用基礎プログラム

以上

令和５年３月○○日

京都大学国際高等教育院長　　宮 川　恒

データ科学の学び方

現代の「読み・書き・算盤」

▶ **誰でも将来必要となるときが来ます.**

各学部・学科の履修方針に従って情報学・統計学・数学（数理科学）をバランスよく学んでください

■全学共通教育で学習しなければならない内容は，学部・学科によって異なります.

全学共通教育での統計学関連科目についての詳細は

データ科学イノベーション教育センターが提供する科目一覧をご覧ください

データサイエンスを学ぶためのガイダンス

https://ds.k.kyoto-u.ac.jp/guidance/

Twitter公式アカウント [@cireds2020]

科目を履修できない場合のために

履修計画は必ず各学部・学科の履修方針に従って立ててください．

- 単位数や時間割などの制約によって，データ科学に関する科目を希望通りに履修できない可能性もあります．

データ科学イノベーション教育研究センター（CIREDS）では
データ科学に関する自習用の教材や企画を用意しています．

協力：京大オリジナル(株)

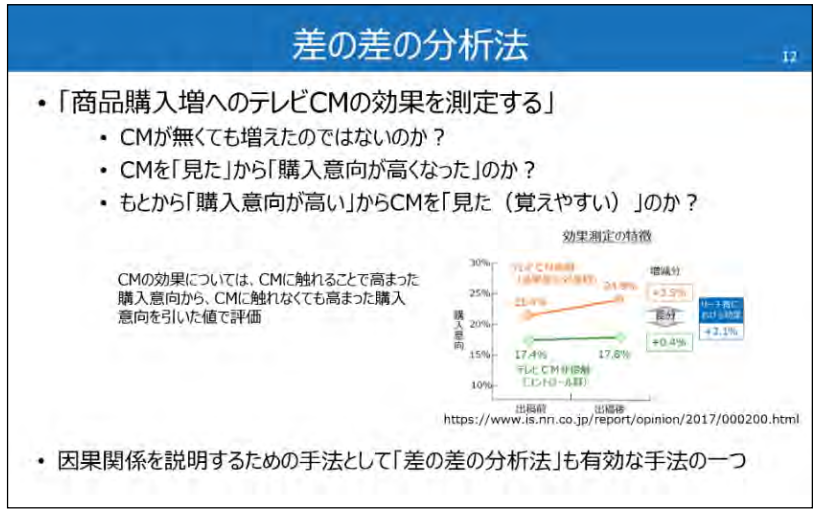
- MOOC（e-learning）
- プログラミング自習キット
- データサイエンス・スクール
- データサイエンス入門講座

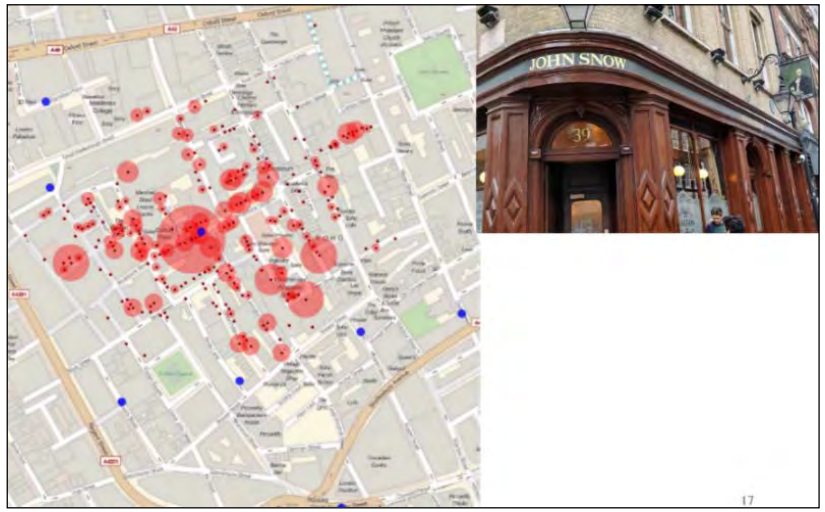
詳しくは下のHP または Twitter公式アカウント［@cireds2020］でご案内
https://ds.k.kyoto-u.ac.jp/category/event/

MOOC「統計の入門」

全国向けe－learningサイトJMOOCで毎年度公開

https://www.jmooc.jp/





■全学共通科目「統計入門」の要点を習得できます
■「統計入門」の予習に利用することもできます
■最後まで履修すると修了証が発行されます

Pythonプログラミング自習キット

https://kubar.rd.iimc.kyoto-u.ac.jp/intro_python/

データ科学で注目されている
プログラミング言語Python
の初歩を習得するためのキットです

Internet環境とWebブラウザだけを用いて学習できます.

■Google社が提供するプログラミング環境Google Collaboratoryを利用しますので,Googleの個人アカウントが必要です

■京都大学の学生だけが利用できます

データサイエンススクール（随時開催）

詳細は
データ科学イノベーション教育研究センターホームページ
http://ds.k.kyoto-u.ac.jp/ で確認してください.

開講科目だけではカバーできない内容を補います

■ 実践的（演習的）内容
■ 先端的内容（機械学習・AIなど）

春期休業，夏期休業，週末などを利用します

■ 1日～4日間の集中講義

学内の研究科・研究所・附属病院所属，他大学，および企業から講師を招聘して開講します

■ 先端研究で用いられる手法を計算機演習を通じて体得

データサイエンス入門講座（有料）

主催：京大オリジナル（株）
共催：京都大学データ科学イノベーション教育研究センター

講義動画(15分程度×18回)
＋1回のライブ配信
（質疑応答と「統計検定」の模擬試験）

「統計検定」を実施する
一般財団法人統計質保証推進協会との
連携セミナー

- 「統計検定」の過去問を利用します

レベルに合わせて学習可能

- 文系学生向3級，2級，統計ソフトR

詳細は京大オリジナル（株）ホームページ，
データ科学イノベーション教育研究センターの
ホームページで確認してください．

おわりに

01 データ科学は現代の「読み・書き・算盤」です

■誰でも必要となるときが来ます

02 データ科学を学ぶには，各学部・学科の履修方針に従って統計学・情報学・数学（数理科学）をバランスよく学んでください

■各学部・学科で学ばなければならない基礎科目や専門科目をおろそかにしないでください．

03 データ科学イノベーション教育研究センターは，皆さんのデータ科学の学習をサポートします

■全学共通教育の科目を担当しています

■MOOC，自習キット，データサイエンススクール，データサイエンス入門講座を提供しています

・データ科学科目を履修できない場合に活用してください

Twitter公式アカウント［@cireds2020］で情報発信中

図版等出典（最終確認日：2023年2月14日）

*1 https://commons.wikimedia.org/wiki/File:Tycho_Brahe.JPG

*2 https://commons.wikimedia.org/wiki/File:Kepler.png

*3 https://commons.wikimedia.org/wiki/File:John_Snow.jpg

*4 https://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak#/media/File:Snow-cholera-map-1.jpg

*5 http://www.shugiin.go.jp/internet/index.nsf/html/images/gijidou002.jpg/$File/gijidou002.jpg

*6 猿谷 豊, 衆議院における音声認識を利用した会議録作成業務,情報管理

Journal of Information Processing and Management : Vol. 55 (2012) No.6 p.392-399
（https://doi.org/10.1241/johokanri.55.E1）

*7 衆議院式速記の一例
https://commons.wikimedia.org/wiki/File:Shugiin-system_Shorthand_01.jpg