

【開発教材の紹介】

リテラシーレベル教材について

東京大学 数理・情報教育研究センター

北川 源四郎

東京大学 MIセンターにおける教材開発

- コンソーシアムでは数理・データサイエンス・AIの教材ポータルを設置し、各大学の教材を集約

<http://www.mi.u-tokyo.ac.jp/consortium/e-learning.html>

- コンソーシアム活動の一環として、東京大学 数理・情報教育研究（MI）センターではモデルカリキュラムに対応した教材を開発・公開

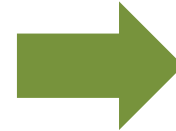
http://www.mi.u-tokyo.ac.jp/6university_consortium.html

- リテラシーレベル教材： 2020年5月
- 応用基礎レベル教材： 2021年6月

この教材の特徴

- モデルカリキュラムに**完全準拠**
 - 全てのキーワード（知識・スキル）を解説
 - キーワードとの対応表がある
- 一部は講義スライドのほかに補助教材も提供
- 全ての大学等で利用できるように、教育目的なら**ページ単位で自由に改変・利用・再配布可能**

CCライセンス CC BY-NC-SA(表示－非営利－継承)



キーワード－教材対応表

キーワード	章-節	ページ
1次データ	1-2	12
1変数関数の微分と積分	4-1	55-90
2軸グラフ	1-4	7
2次データ	1-2	12
AI	1-1	33-34
AIサービスの責任論	3-1	31
AI最新技術の活用例	1-6	5, 6, 7, 8, 9, 10
AI社会原則	3-1	30
AI等を活用した新しいビジネスモデル	1-6	5, 6, 7, 8, 10, 11
AIとビッグデータ	1-4	22
AIの非連続的進化	1-1	35
ELSI	3-1	5-8
EU一般データ保護規則(GDPR)	3-1	29
IoT	1-1	26-29
n-gram	4-5	14
Society 5.0	1-1	42-44
SQL	4-7	12
あーお		
悪意ある情報搾取	3-2	38-41
アクティビティ図	4-2	14
アルゴリズムの表現	4-2	15
アルゴリズムバイアス	3-1	23
アンケート調査	2-1	24
暗号化	3-2	37
異常値	4-7	19
移動平均	4-4	12
今のAIで出来ることと出来ないこと	1-4	23

CCライセンス (Creative Commons Licenses)

<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.ja>



表示 - 非営利 - 継承 4.0 国際
(CC BY-NC-SA 4.0)

あなたは以下の条件に従う限り、自由に：

共有 — どのようなメディアやフォーマットでも資料を複製したり、再配布できます。

翻案 — マテリアルをリミックスしたり、改変したり、別の作品のベースにしたりできます。

あなたがライセンスの条件に従っている限り、許諾者がこれらの自由を取り消すことはできません。

あなたの従うべき条件は以下の通りです。



表示 — あなたは適切なクレジットを表示し、ライセンスへのリンクを提供し、変更があったらその旨を示さなければなりません。これらは合理的であればどのような方法で行っても構いませんが、許諾者があなたやあなたの利用行為を支持していると示唆するような方法は除きます。



非営利 — あなたは営利目的でこの資料を利用してはなりません。



継承 — もしあなたがこの資料をリミックスしたり、改変したり、加工した場合には、あなたはあなたの貢献部分を元の作品と同じライセンスの下に頒布しなければなりません。

追加的な制約は課せません — あなたは、このライセンスが他の者に許諾することを法的に制限するようないかなる法的規定も技術的手段も適用してはなりません。

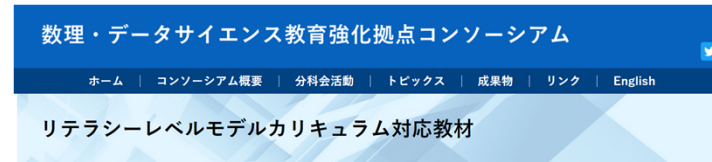
リテラシーレベル教材サイトへのアクセスの方法

● コンソーシアムHPから



リテラシーレベル
教材ポータルサイトへのリンク

教材ポータルサイト



利用条件とアンケート

東京大学と記載のあるスライド教材の利用については[こちら](#)の一部スライドは冒頭の利用条件をご参照ください。
東京大学と記載のある講義動画の利用条件は、各動画の冒頭をご参照ください。
滋賀大学と記載のある教材の利用条件は[CC-BY-NC-SA](#)です。
九州大学と記載のある教材の利用条件は[CC-BY](#)です。
教材のアンケートは[こちら](#)です。

東大教材サイトへのリンク

モデルカリキュラムと対応する講義動画・スライド

- 1. 社会におけるデータ・AI活用
- 2. データリテラシー
- 3. データ・AI活用における留意事項
- 4. オプション

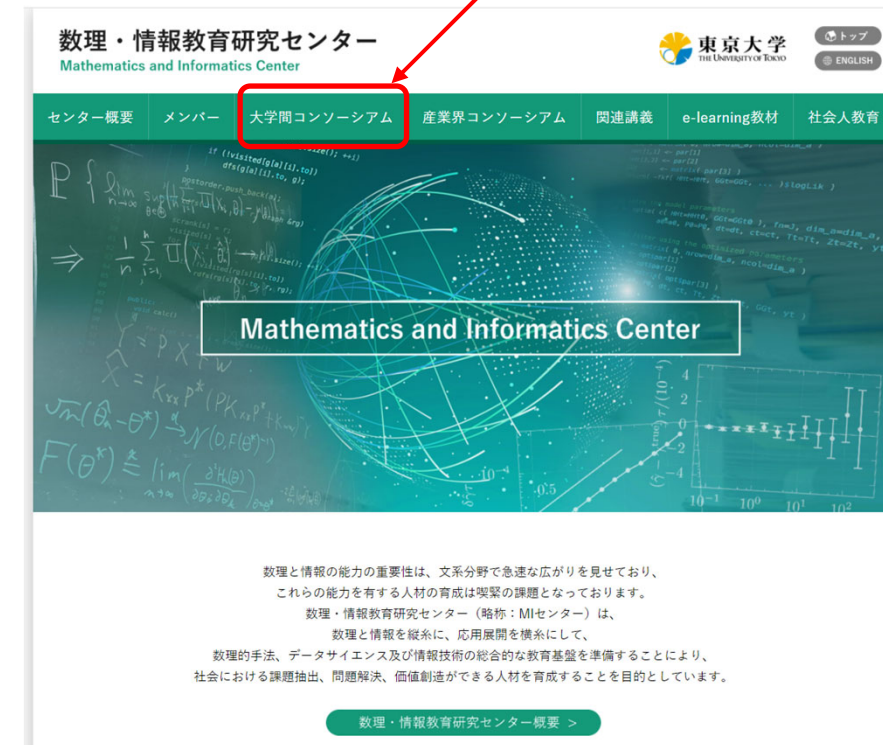
1. 社会におけるデータ・AI活用

1-1. 社会で起きている変化

キーワード	導入	基本	補助教材	教科書シリーズとの対応
	社会で起きている変化(スライド・東京大学)			
ビッグデータ、IoT、AI、ロボット	データサイエンスの役割 (1) (動画・滋賀大学)			数値としてのデータサイエンス (刊行予定)

● 東大 数理・情報教育センターHPから

東大教材サイトへのリンク



数理・情報教育研究センター 教材サイト

数理・情報教育研究センター
Mathematics and Informatics Center



東京大学
THE UNIVERSITY OF TOKYO

[トップ](#)
[ENGLISH](#)

センター概要

メンバー

大学間コンソーシアム

産業界コンソーシアム

関連講義

e-learning教材

社会人教育

数理・データサイエンス教育強化拠点コンソーシアム

M I センターは、2017年度政府予算に盛り込まれた「数理・データサイエンス教育の強化」事業の東京大学における実施主体です。同事業で選定された6大学（北海道大学、東京大学、滋賀大学、京都大学、大阪大学、九州大学）のセンターのコンソーシアムの幹事校として、大学、産業界、研究機関等と幅広くネットワークを形成し、地域や分野における先進的教育モデルの拠点として実践的な教育の普及に努めます。

[\[コンソーシアムホームページ\]](#)

数理・データサイエンス・AIの活用事例動画

本動画集は数理・データサイエンス・AIリテラシーレベル教材の導入となるような活用事例を収集したものです。数理・データサイエンス・AIリテラシーレベル教材との対応は下の教材一覧をご覧ください。

[人工知能（AI）の活用事例動画一覧](#)

数理・データサイエンス・AI教材

本教材は数理・データサイエンス教育強化コンソーシアムが公開したリテラシーレベル及び応用基礎レベルのモデルカリキュラムの全国展開に向け、東京大学 数理・情報教育センターが開発したものです。モデルカリキュラムに完全準拠したものであり、全ての大学、高専において、学生への教育にご利用いただけます。

[本教材の利用についての詳細](#)

下に移動

章	節	教材	実習用補助教材
1. 社会におけるデータ・AI活用	1-1. 社会で起きている変化	スライド 	
	1-2. 社会で活用されているデータ	スライド 	
	1-3. データ・AIの活用領域	スライド 	人工知能（AI）の活用事例動画
	1-4. データ・AI活用のための技術	スライド 	
	1-5. データ・AI活用の現場	スライド 	
	1-6. データ・AI活用の最新動向	スライド 	
2. データリテラシー	2-1. データを読む	スライド 	
	2-2. データを説明する	スライド 	
	2-3. データを扱う	スライド 	
3. データ・AI活用における留意事項	3-1. データ・AIを扱う上での留意事項	スライド 	
	3-2. データを守る上での留意事項		
4. オプション	4-1. 統計および数理基礎	スライド 	
	4-2. アルゴリズム基礎	スライド 	
	4-3. データ構造とプログラミング基礎		
	4-4. 時系列データの解析	スライド 	データ  , 解析例 (pdf  , excel )
	4-5. テキスト解析	スライド 	
	4-6. 画像解析	スライド 	Pythonサンプルコード (pdf1  , pdf2  , pdf3 )
	4-7. データハンドリング	スライド 	
	4-8. データ活用実践（教師あり学習）	スライド 	データ解析例 
	4-9. データ活用実践（教師なし学習）	スライド 	データ解析例 

応用基礎レベル教材の詳細目次

- ◆ [教材目次詳細版](#) 
- ◆ [キーワードと教材の対応表](#) 

リテラシーレベル教材の詳細目次

- ◆ [教材目次詳細版](#) 
- ◆ [キーワードと教材の対応表](#) 

スライド教材の構成

章	節	教材	実習用補助教材
1. データサイエンス 基礎	1-1. データ駆動型社会とデータサイエンス	スライド PDF	
	1-2. 分析設計	スライド PDF	
	1-3. データ観察	スライド PDF	
	1-4. データ分析	スライド PDF	補助資料
	1-5. データ可視化	スライド PDF	補助教材

1. タイトル
2. 概要（学修目標）
3. 本教材の目次
4. 内容
5. . . .

1-1 データ駆動型社会と データサイエンス

東京大学 数理・情報教育研究センター
2021年4月30日

東京大学 数理・情報教育研究センター 久野遼平 2021 CC BY-NC-SA

1

概要

- 本節では
 - データ駆動型社会においてデータサイエンスを学ぶことの意義を理解する
 - データを活用した一連のプロセスを体験し、データ活用の流れ（進め方）を理解する（仮説検証、知識発見、原因究明、計画策定、判断支援、活動代替など）
 - 来るべき未来社会について理解を深める
- ことを目標とします。

東京大学 数理・情報教育研究センター 久野遼平 2021 CC BY-NC-SA

2

本教材の目次

1. データサイエンス活用事例	4ー7
2. データを活用した新しいビジネスモデル	8ー11
3. 進むAIの社会実装とデータ駆動社会	12ー15
4. Society 5.0とスーパーシティ	16ー24

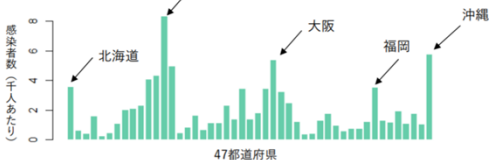
東京大学 数理・情報教育研究センター 久野遼平 2021 CC BY-NC-SA

3

スライドの例

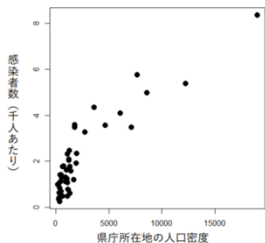
説明変数との関係を見る

下の棒グラフは左から北海道、青森、・・・、沖縄の順に47都道府県の人口千人あたりの新型コロナウイルスの累積感染者数を示しています。都道府県によって感染した人の割合が大きく異なることがわかります。



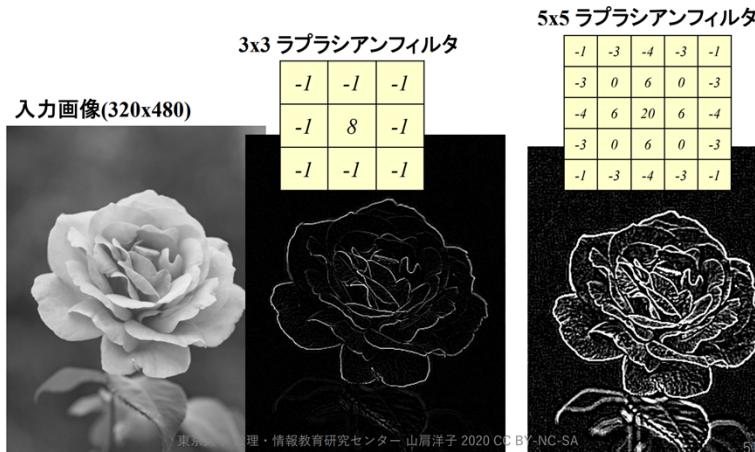
このような都道府県（以下では単に県と書きます）による感染者の違いを各県の県庁所在地の人口密度と関連づけて説明することを考えてみましょう。

実際に、各県の県庁所在地の人口密度を横軸に、千人当たりの感染者数を縦軸にとって**散布図**を描いてみると、右のように人口密度と感染者数には明確な関係がみえます。



エッジ抽出：ラプラシアンフィルタ

その画素の値が周辺の画素と違うほど大きな値とすることによって、周辺との輝度が大きく変化する輪郭線を取り出す

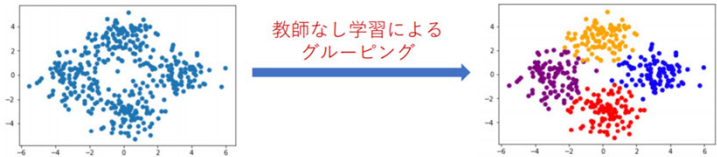


教師なし学習

訓練データは特徴ベクトルのあつまり $x_1, \dots, x_n$ です。

教師値が与えられない点が教師あり学習と異なります。

教師なし学習によるグルーピング：データ同士の類似性に基づくグルーピング（クラスタリング）は教師なし学習の代表例です。



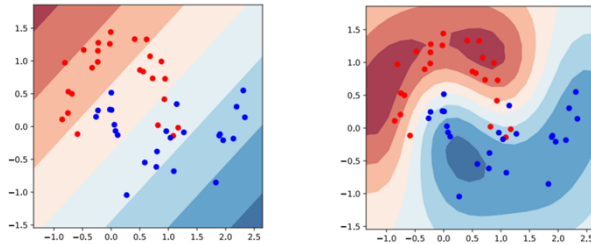
クラスタリングの例

顧客セグメンテーション→ 顧客情報や利用状況によるクラスタリング

店舗クラスタリング→ 店舗や周辺情報からクラスタリング

カーネルSVM

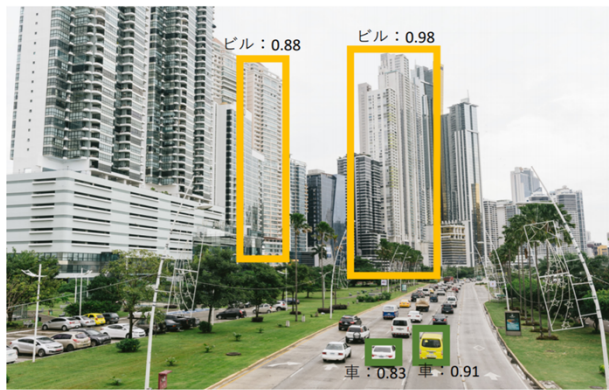
- 線形SVMは分離境界が直線になるため、データによっては必ずしもうまく予測ができません。
- カーネルSVMは、データ点同士の類似度を表すカーネル関数を用いる手法で、複雑な識別境界を実現可能です。



2次元入力変数を2クラス(赤or青)で2クラス分類した例。赤い領域は赤のクラスに、青の領域は青のクラスに分類されます。

事例：物体検出

画像認識より一段むずかしい問題設定。
どこに何が写っているかを検出します。



「どこに」+「なにが」

補助教材

4. オプション

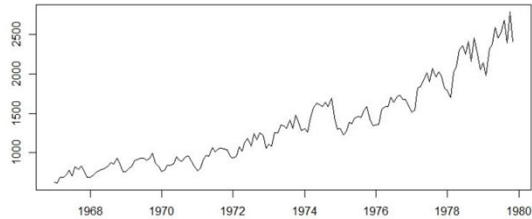
4-2. アルゴリズム基礎	スライド	
4-3. データ構造とプログラミング基礎		
4-4. 時系列データの解析	スライド	データ 解析例 (pdf , excel)
4-5. テキスト解析	スライド	
4-6. 画像解析	スライド	Pythonサンプルコード (pdf1 , pdf2 , pdf3)

1. データ (csvファイル)
2. 計算コード (Excel, R, Python)
3. 解析例詳細

卸売り高データ

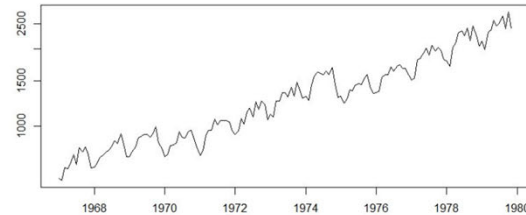
データ名	WHARD
内容	あるハードウェアの毎月の卸売り高
観測間隔	一月
データ数	155
出展	合衆国Bureau of Labor Statistics

データを読み込む
`data( WHARD )`
データをプロットする
`plot(WHARD)`

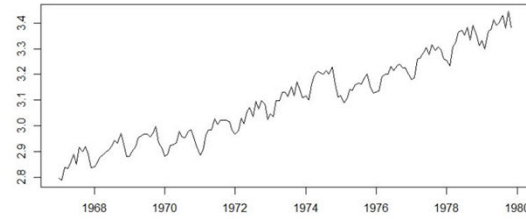


金額、人数、個数などの計数データは値が大きくなると変動も大きくなる場合が多い。そのような場合、対数変換すると変動がレベルによって変化しなくなることが多い。

Y軸を対数目盛に変えて表示
`plot( WHARD, log="y" )`

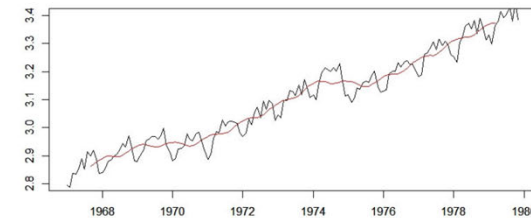


データを対数変換してプロット
`plot( log10( WHARD ) )`



卸売り高データ(2)

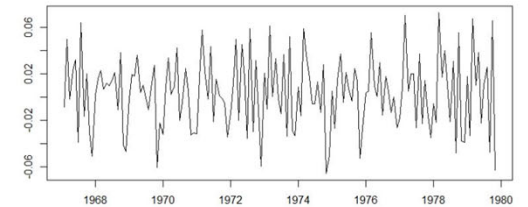
移動平均を計算してプロットする
`data( WHARD )`
`x <- log10( WHARD )`
`plot( x,type="l",ylim=c(2.8,3.4) )`
`y <- x`
`ndata <- length( x )`
`y[1:ndata] <- NA`
`kfilter <- 8` ← この値を大きくすると滑らかになる
`n0 <- kfilter+1`
`n1 <- ndata-kfilter`
`for(i in n0:n1){`
`i0 <- i-kfilter`
`i1 <- i+kfilter`
`y[i] <- mean( x[i0:i1] )`
`}`
`par( new=T )`
`plot( y,ylim=c(2.8,3.4),type="l" )`
`lines(y,col=2,ylim=c(2.8,3.4))`



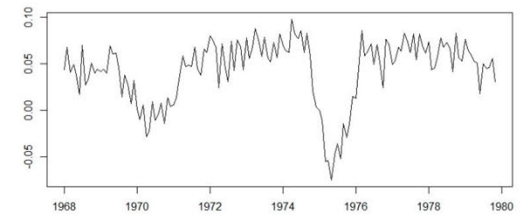
— WHARDの対数
— 移動平均

- ・ 移動平均で滑らかなトレンドが得られる
- ・ 階差によってトレンドを除去できる
- ・ 季節階差で周期を除去できる

`log10( WHARD )`の階差を計算し表示する
`plot( diff( log10( WHARD ) ) )`



季節階差を計算し表示する
`plot( diff( log10( WHARD ), lag=12 ) )`



その他の教材

http://www.mi.u-tokyo.ac.jp/teaching_material.html

数理・データサイエンス関連教材

データサイエンスに関する公開講義をご覧ください。各講義からはさらに詳細な内容一覧をご覧ください。

[数理・情報教育研究センターUTeLFページ](#)

講義内容・講義動画

Open All

複素解析学I 平地 健吾 教授 **New** [講義動画](#)

▼ [講義内容](#)

文科系のための線形代数・解析II 藤堂 眞治 教授・松尾 泰 教授・藤原 毅夫 教授 **New** [講義動画](#)

▼ [講義内容](#)

確率過程論（数理手法VI）荻原 哲平 准教授

▼ [講義内容](#)・[資料](#)・[動画](#)

文科系のための線形代数・解析I 藤堂 眞治 教授・松尾 泰 教授・藤原 毅夫 教授

▼ [講義内容](#)・[資料](#)・[動画](#)

メディアプログラミング入門 山肩 洋子 准教授

▼ [講義内容](#)・[資料](#)・[動画](#)

計算機実験I 藤堂 眞治 教授

▼ [講義内容](#)・[資料](#)・[動画](#)

最適化手法（数理手法III）寒野 善博 教授 [講義動画](#)

▼ [講義内容](#)

確率論（数理手法IV）楠岡 成雄 特任教授 [講義動画（OCW）](#) [講義動画（OCWx）](#)

▼ [講義内容](#)

工学のための現代数学入門（数理手法V）藤原 毅夫 特任教授 [講義動画](#)

▼ [講義内容](#)

確率過程論（数理手法VI）楠岡 成雄 特任教授 [講義動画](#)

▼ [講義内容](#)

時系列解析（数理手法VII）北川 源四郎 特任教授 [講義動画](#)

▼ [講義内容](#)

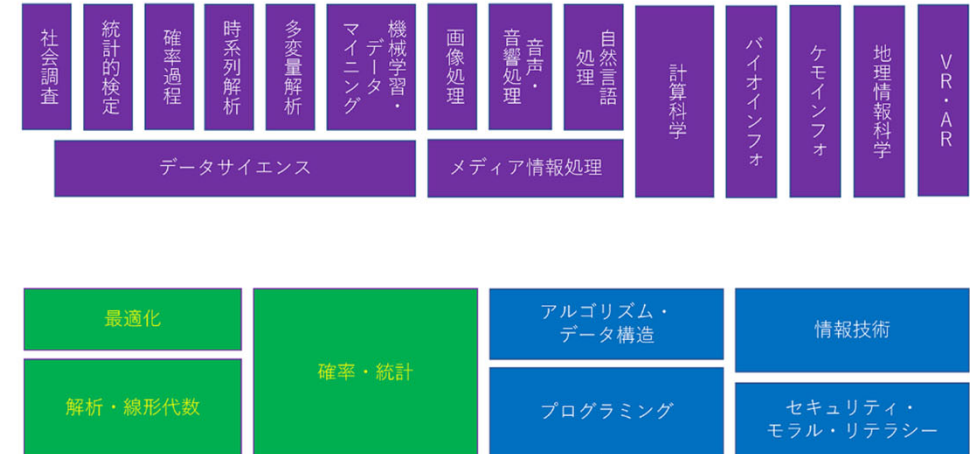
<http://www.mi.u-tokyo.ac.jp/infoportal.html>

情報教育コンテンツリンク集

東京大学で公開している情報分野のe-learning教材や授業ページを紹介しています。東京大学の教員が関わったコンテンツを中心にしていますが、一部、海外の著名なコンテンツも含めました。

情報分野の自習に際に活用してください。

以下は情報教育の俯瞰図です。下の方には数学・統計を含む基礎分野、上の方には応用分野が示されています。分野を示す長方形をクリックしてください。（教材を参照している分野はクリックできます。）



プログラミング

[アルゴリズム入門](#)

前期課程1年冬学期のアルゴリズム入門の授業のページで、過去の試験問題を公開しています。

[Pythonプログラミング入門](#)

理学部のPythonプログラミング入門の授業のページで、ノートブック、ビデオなどを公開しています。アクセスするにはECCSクラウドメールアカウントでログインする必要があります。

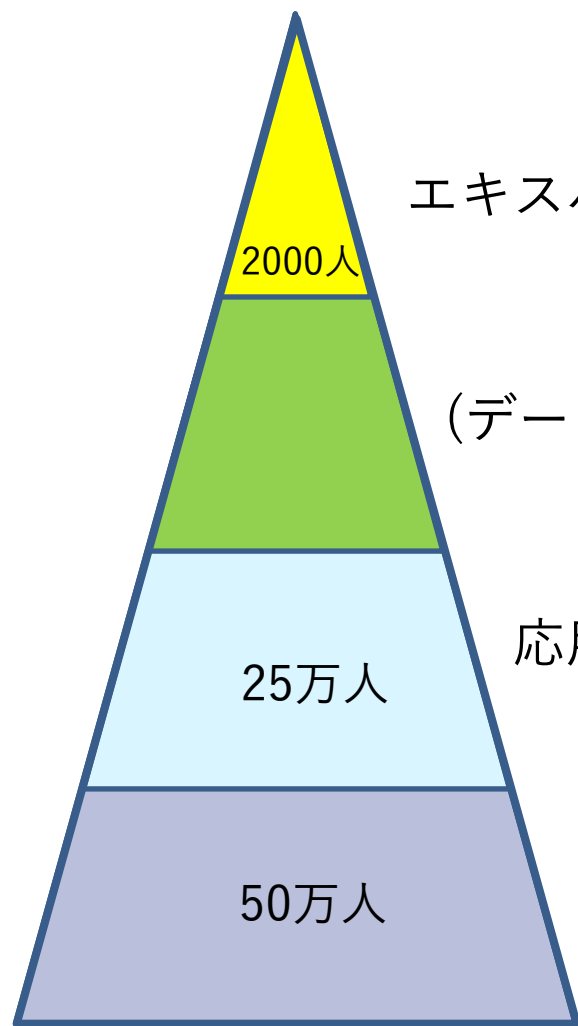
[統計データ解析I](#)

前期課程の統計データ解析Iの授業のビデオおよび講義スライドをOCWxから公開しています。この授業はRのプログラミングを含んでいます。

参考資料1：育成レベルとコンソーシアムが公開したモデルカリキュラム

AI戦略2019

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



エキスパート

2000人

(データサイエンティスト)

大学におけるデータサイエンス教育に関するスキルセット及び学修目標 第2次報告 (2021年6月)

応用基礎レベル

25万人

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 ～ AI×データ活用の実践 ～ (2021年3月)

リテラシーレベル

50万人

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 ～ データ思考の涵養 ～ (2020年4月)

参考資料2： リテラシーレベル・モデルカリキュラムの構成

コア学習科目	導入	1. 社会におけるデータ・AI利活用	
		1-1. 社会で起きている変化 1-3. データ・AIの活用領域 1-5. データ・AI利活用の現場	1-2. 社会で活用されているデータ 1-4. データ・AI利活用のための技術 1-6. データ・AI利活用の最新動向
	基礎	2. データリテラシー	
		2-1. データを読む 2-3. データを扱う	2-2. データを説明する
	心得	3. データ・AI利活用における留意事項	
		3-1. データ・AIを扱う上での留意事項	3-2. データを守る上での留意事項
オプション	選択	4. オプション	
		4-1. 統計および数理基礎 4-3. データ構造とプログラミング基礎 4-5. テキスト解析 4-7. データハンドリング 4-9. データ活用実践（教師なし学習）	4-2. アルゴリズム基礎 4-4. 時系列データ解析 4-6. 画像解析 4-8. データ活用実践（教師あり学習）