



数理・データサイエンス・AI教育の ための文理を問わないリテラシレベル 教材開発

山梨大学 教育国際化推進機構
大学教育センター

鈴木一克，佐藤友香，森澤正之，日永龍彦，鈴木 裕，埴 雅典

概要

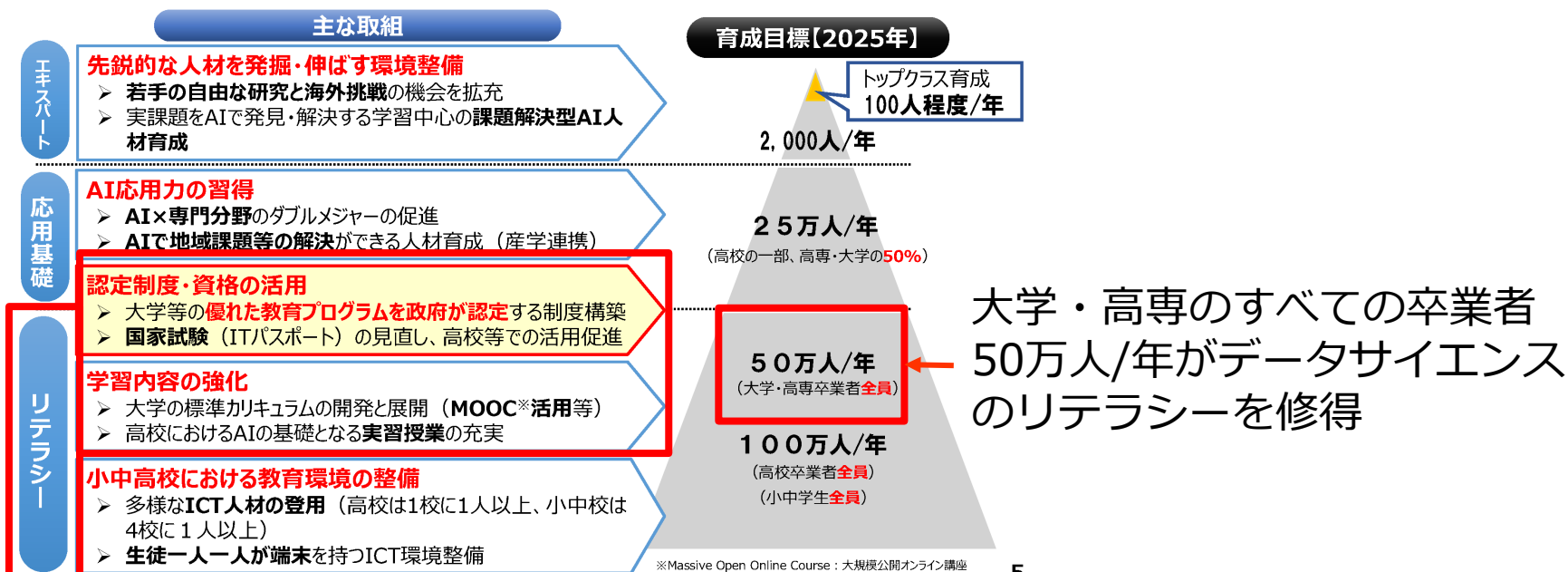
1. 背景
2. 開発した教材
3. 授業の実施方法と実施結果
4. まとめ

数理・データサイエンス・AI教育用 教材開発の背景

AI戦略2019

「数理・データサイエンス・AI」はデジタル社会の「読み・書き・そろばん」

数理・データサイエンス教育強化の要求



大学・高専のすべての卒業生
50万人/年がデータサイエンス
のリテラシーを修得

出典：第1回数理・データサイエンス教育プログラム認定制度検討会議資料（2019年10月29日）

https://www.kantei.go.jp/jp/singi/ai_senryaku/suuri_datascience_ai/dai1/siryou1-2.pdf

数理・データサイエンス教育強化拠点コンソーシアム 標準カリキュラム策定・教育用教材の作成など

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

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

学修項目の4分類

- ▶ モデルカリキュラムの構成を以下のとおり「導入」「基礎」「心得」「選択」に分類し、学修項目を体系的に示した。
- ▶ 「導入」「基礎」「心得」はコア学修項目として位置付ける。「選択」は学生の学習歴や習熟度合い等に応じて、適切に選択頂くことを想定している。
- ▶ 次頁よりそれぞれの分類における「学修目標」「学修内容」「スキルセット（キーワード）」をまとめた。

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

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）

出典：数理・データサイエンス教育強化拠点コンソーシアム「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～」
http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_literacy.pdf

データサイエンス教育のリファレンス(1/2)

モデルカリキュラムに先駆けた独自リファレンスの策定(2019年6月)

0. 高校数学の基礎的素養

数列, 数列の和 (Σ), 指数関数, 対数関数などを活用できる

1. データサイエンスの概要 (**DP教養 : 多様な知識**)

導入

- データとは何か, 現代社会(ビッグデータの時代)においてなぜデータサイエンスを学ぶ必要があるのか, を説明できる
- データサイエンスの基本的な用語(ビッグデータ, データマイニング, テキストマイニング, クラウドコンピューティング, 並列計算など)を説明できる

2. 機械学習とAIの概要と可能性 (**DP教養 : 様々な学問分野の考え方**)

導入

- 機械学習やAIの代表的な手法(ニューラルネットワーク, ディープラーニング, サポートベクターマシンなど)の概要を説明できる
- 教育プログラムに関連するデータサイエンスの応用事例を説明できる

3. データの選定・収集・整理 (**DP汎用能力 : 情報リテラシー**)

基礎

- 様々な手段を用いて多様なデータを合法的・合理的に探索・取得できる

心得

- 多様なデータの中から適切なデータを選択することができる

選択

- データの不整合を修正する(データクレンジング)ことができる

データサイエンス教育のリファレンス(2/2)

4. データ分析の基礎（**DP汎用能力：数量的リテラシー，論理的思考力**）

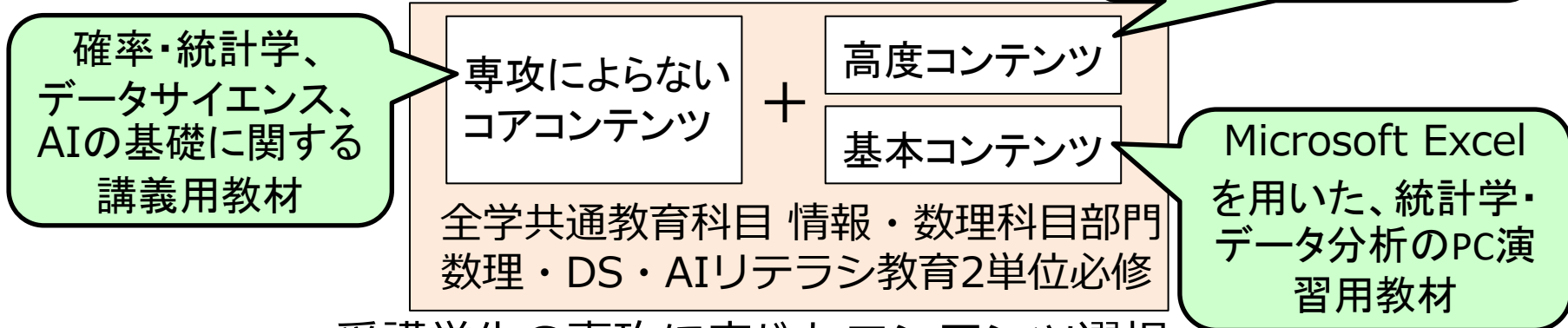
- 選択**
- データの傾向を表す様々な基本統計量（平均値，中央値，標準偏差，散布図，相関係数など）を説明でき，データの理解に活用できる
 - 確率変数，確率密度関数，分布関数の持つ役割と性質を説明でき，正規分布，2項分布などの基本的な分布について数式やグラフを用いて説明できる
 - 統計的推定と仮説検定などの基礎的事項を説明でき，データの理解に活用できる
 - 多変量解析の基本的な手法（重回帰分析，判別分析，主成分分析，因子分析，クラスター分析など）の概要を説明できる

5. ソフトウェアを活用した統計学的手法による分析・表現・活用（**DP汎用能力：問題解決力**）

- 選択**
- ソフトウェア（注）を活用して様々なデータを統計学的手法で分析できる
 - 分析結果を適切な図表で表現し，問題解決に活用できる

データサイエンス入門授業の当初計画

- 2020年度入学者から、1年次**全学必修**科目
- **リテラシーレベル**教育
- **2段階モジュール構成**の教材



受講学生の専攻に応じたコンテンツ選択

- 「高度コンテンツ」：医学部医学科
- 「基本コンテンツ」：医学部看護学科、教育学部

- **反転授業**の導入
 - 講義動画 + PC演習による知識と技能の同時習得
- AI学習に**プログラミングレスAI基盤**を活用

表1: リテラシレベルシラバス (i)コアコンテンツ+基本コンテンツ

コア:コアコンテンツ、基本:基本コンテンツ

回	内容	回	内容
1	データサイエンス概論 (コア) 導入	9	回帰分析 (コア、基礎) 選択
2	オープンデータの取得と整理 (コア)、 Excelの基礎1 (基本) 導入 基礎 心得	10	確率と場合の数 (コア) 選択
3	Excelの基礎2 (基本) 基礎 選択	11	確率分布 (コア) 選択
4	ベクトルと行列の基礎 (コア) 選択	12	二項分布 (コア) 選択
5	統計のグラフ (コア、基本) 基礎 選択	13	正規分布 (コア) 選択
6	ヒストグラムとデータの代表値1 (コア、 基本) 基礎 選択	14	推測統計 (コア) 選択
7	ヒストグラムとデータの代表値2 (コア、 基本) 基礎 選択	15	機械学習・AI概論 (コア) 導入 選択
8	相関 (コア) 基礎 選択		

表1: リテラシレベルシラバス (ii)コアコンテンツ+高度コンテンツ

コア:コアコンテンツ、高度:高度コンテンツ
赤字部分はAI応用基礎

回	内容	回	内容
1	データサイエンス概論 (コア) 導入	9	確率分布、二項分布 (コア) 選択
2	オープンデータの取得と整理 (コア)、Pythonの基礎1 (高度 導入 基礎 心得)	10	正規分布 (コア) 選択
3	Pythonの基礎2 (高度)、ベクトルと行列の基礎 (コア) 基礎 選択	11	仮説検定 (コア) 選択
4	統計のグラフ (コア、高度) 選択	12	機械学習・AI概論、ニューラルネットワーク(1) - NNの仕組みと数学基礎- (コア) 導入 選択
5	ヒストグラムとデータの代表値 (コア、高度) 基礎 選択	13	ニューラルネットワーク(2) - Neural Network Consoleを用いた実装 1- 選択
6	相関 (コア) 基礎 選択	14	ニューラルネットワーク(3) - Neural Network Consoleを用いた実装 2- 選択
7	回帰分析 (コア、高度) 基礎 選択	15	総括
8	確率と場合の数 (コア) 選択		

2020年度講義の授業内容の例（教育学部）

後期開講

コア:コアコンテンツ、基本:基本コンテンツ

回	内容	回	内容
1	データサイエンス概論(コア)	9	相関(コア)
2	オープンデータの取得と整理(コア)、 e-Statからのデータ取得(基本)	10	回帰直線と決定係数、相関関係と 因果関係(コア)
3	データサイエンスとプログラミング(コ ア)、Excelの基礎(基本)	11	Excelで回帰分析を行う(基本)
4	ベクトルと行列の基礎(コア)	12	機械学習・AI概論(コア)
5	統計グラフ(コア)、Excelを用いた統 計グラフの作成(基本)	13	確率と場合の数(コア)
6	データの代表値(コア)、データの代 表値をExcelで計算する(基本)	14	二項分布(コア)
7	箱ひげ図(コア)	15	正規分布(コア)
8	Excelで箱ひげ図を描く(基本)		

2020年度前期の授業内容の例（医学部医学科）

コア:コアコンテンツ、基本:基本コンテンツ
赤文字はAI応用基礎教材

前期開講

回	内容	回	内容
1	データサイエンス概論(コア)	9	データの代表値をExcelで計算する(基本)
2	統計グラフ(コア)	10	Excelで回帰分析を行う(基本)
3	データの代表値(コア)	11	Excelで箱ひげ図を描く(基本)
4	相関(コア)	12	機械学習・AI概論、ニューラルネットワーク (1) — NNの仕組みと数学基礎—(コア)
5	回帰直線と決定係数、相関関係と因果関係(コア)	13	ニューラルネットワーク(2) — Neural Network Consoleを用いた実装 1—
6	ベクトルと行列の基礎(コア)	14	ニューラルネットワーク(3) — Neural Network Consoleを用いた実装 2—
7	オープンデータの取得と整理(コア)、e-Statからのデータ取得(コア)	15	総括
8	データサイエンスとプログラミング(コア)、Excelの基礎(基本)		高度コンテンツから基本コンテンツに変更

教材（動画+スライド+小テスト）の形式

講義用教材

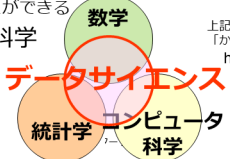
動画

山梨大学 統計学からデータサイエンスへ

- コンピュータを用いたデータの収集・整理・分析
 - 大量のデータを収集・整理・分析し、データの全体像を把握することは、もはや人手ではほぼ不可能
 - 高性能なコンピュータと高速なインターネットが身近な道具になっている
 - 便利なソフトウェア（アプリ）があり、しかも日々進化している
 - 大量のデータを短時間で計算・処理できる
 - 計算結果や図を保存でき、他者と共有・改善できる
 - きれいなグラフを簡単に描くことができる
- 統計学・数学・コンピュータ科学
→ データサイエンス

上記イラストの出典：
「かわいいフリー素材集 いらすとや」
<https://www.irasutoya.com/>

17



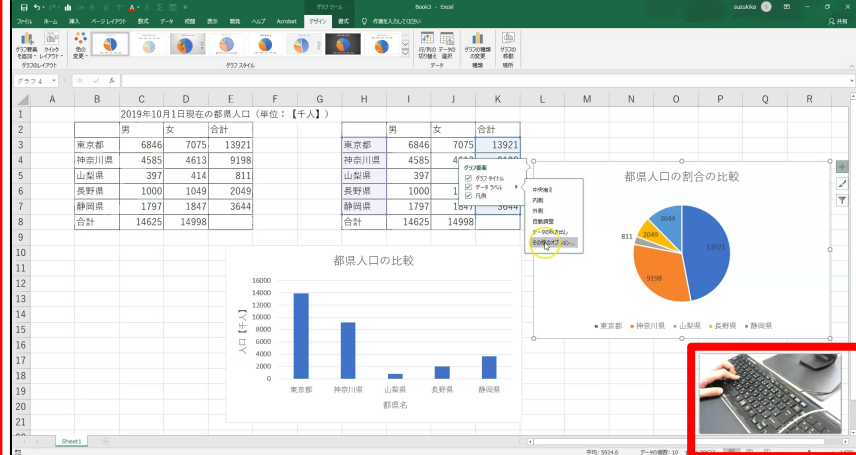
スライドとナレーション
(ナレーションにはプロ
のナレーターを起用)

スライド

各動画の時間は30～40分

PC演習用教材

実演動画



2019年10月1日現在の都県人口 (単位: 【千人】)

	男	女	合計
東京都	6846	7075	13921
神奈川県	4585	4613	9198
山梨県	397	414	811
長野県	1000	1049	2049
静岡県	1797	1847	3644
合計	14625	14998	

都県人口の割合の比較

都県人口の比較

Excel操作のデスクトップ画面、
キーボード操作の映像、
および教員解説ナレーション

内容要約スライド

PC演習実演動画作成のきっかけ

PC演習は対面授業でも難しい。遠隔で実施することは可能か？

心配だったこと

- 1年生の、PCおよびExcel操作への不慣れさ
- リアルタイム配信演習の場合の受講者側機器環境の負担
- スライドと文章のみの資料や音声のみによる情報伝達の限界
- 学生の進捗状況確認の困難さから生ずる、各学生の理解度や達成度の差

それなら教員が演習課題を実演する動画を作ろう！

オンデマンド動画配信の利点

- 学生が各自のペースで動画を観ながら演習を実施可能
- 繰り返し視聴可能

小テストと演習課題レポートによる 理解度確認 (1/2)

毎回、LMS上での小テスト実施による理解度確認

問題 1

未解答

最大評点 1.00

問題にフラグを付ける

問題を編集する

1TB (1テラバイト) は1MB (1メガバイト) の何倍でしょうか。選択肢の中から1つ選んでください。

1つ選択してください:

- a. 1,000倍
- b. 100倍
- c. 1,000,000倍
- d. 10,000倍

選択式問題

問題 2

未解答

最大評点 1.00

問題にフラグを付ける

問題を編集する

次の一文の「???'の箇所に入る語句を解答欄に記入してください。

ある問題を解決するための意思決定を行うには、専門家の「経験」と「勘」だけでなく、「???'の分析とその結果の検証が必要である。

解答:

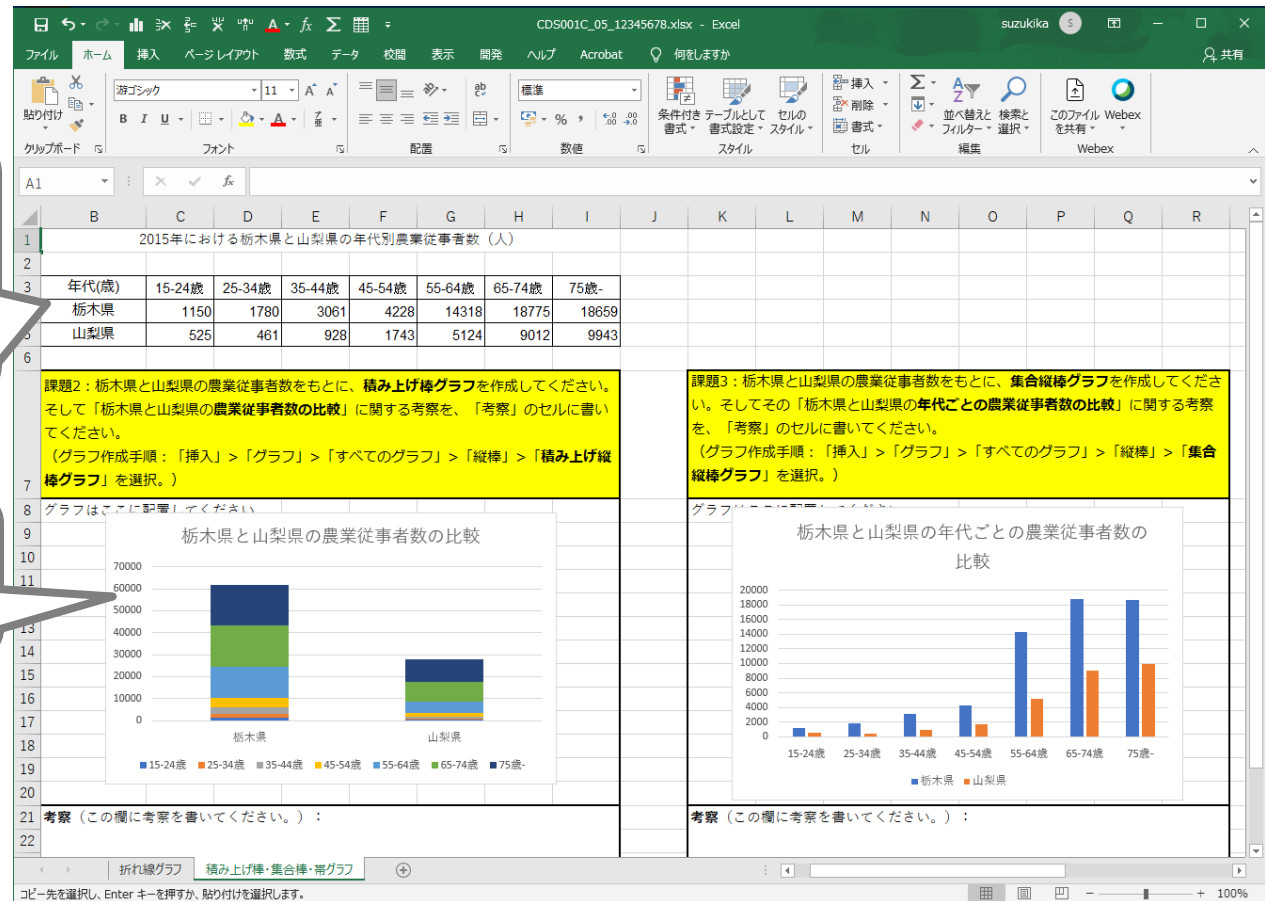
記述式問題

小テストと演習課題レポートによる 理解度確認 (2/2)

演習課題を行ったExcelファイルをレポートとして提出

データは教員が
提供または公
的データを各自
で取得






データからの
グラフ作成



課題提出用ファイルのテンプレートを教員が作成する場合もあり

教材の提供方法

実際のMoodleの画面例

<p>管理</p> <ul style="list-style-type: none">▼ コース管理<ul style="list-style-type: none">設定を編集する編集モードの開始ユーザ▼ フィルタ<ul style="list-style-type: none">レポート設定表セットアップ<ul style="list-style-type: none">バッジバックアップ↑ リストア↑ インポート← リセット問題バンク	<h3>第2回授業 (2020年5月14日)</h3> <p>二つの動画を観て (スライド資料を見て) から小テストに解答してください。</p> <p>今回の小テストの解答提出期限は5月20日 (水) 23:59です。</p> <div data-bbox="560 514 956 664"><p> 第1回授業補足説明と小テスト解説の動画</p><p>利用制限 2020年 05月 14日 08:45 より利用可</p><p>第1回授業で説明しなかった、カタカナ語について補足します。第1回小テストの解説も行います。</p><p>動画の最初で教員が自分の顔を映して挨拶します。動画の時間は、約6分30秒です。</p></div> <div data-bbox="1159 514 1439 585"><h2>講義動画</h2></div> <div data-bbox="560 678 956 792"><p> 第1回授業補足説明と小テスト解説のスライド</p><p>利用制限 2020年 05月 14日 08:45 より利用可</p><p>動画で使っているスライドです。</p></div> <div data-bbox="1159 692 1516 763"><h2>スライド資料</h2></div> <div data-bbox="560 806 956 1006"><p> データサイエンス入門第2回授業動画</p><p>利用制限 2020年 05月 14日 08:45 より利用可</p><p>今回は、授業内容の順序を変更して、統計グラフについて学びます。</p><p>動画及びスライド資料の中では「第4回」となっていますが、それは気にしないでください。また、動画中の「授業内容」で触れられている、「Excelで統計グラフを作成する」は、後日演習で行います。</p><p>動画の時間は、約21分です。</p></div> <div data-bbox="1159 806 1439 878"><h2>講義動画</h2></div> <div data-bbox="560 1021 956 1135"><p> 第2回授業動画のスライド</p><p>利用制限 2020年 05月 14日 08:45 より利用可</p><p>第2回授業動画のスライドです。</p></div> <div data-bbox="1159 1013 1516 1085"><h2>スライド資料</h2></div> <div data-bbox="560 1149 956 1263"><p> 第2回小テスト</p><p>利用制限 2020年 05月 14日 08:45 より利用可</p><p>動画を観た (スライドを読んだ) 後に小テストに解答してください。今回の解答期限は5月20日 (水) 23:59です。</p></div> <div data-bbox="1159 1142 1400 1213"><h2>小テスト</h2></div>
--	--

授業実施方法

2020年度

- 前期・後期とも全15回オンデマンド開講
 - ✓ 原則として、指定された授業開講日時に受講
 - ✓ 小テストは授業開講日から6日後まで受験可能
 - ✓ 動画は年度末までいつでも閲覧可能
- 質問対応
 - ✓ 前期:Moodleのフィードバック機能を用いた質問箱設置
 - ✓ 後期:Moodleの質問箱 + ビデオ会議システムによる授業時間中のリアルタイム質疑応答
- 成績評価方法
 - ✓ 筆記試験は行わなかった
 - ✓ 毎回の小テストの結果と最終レポート(前期・医学部、後期・教育学部で実施)
 - ✓ PC演習のレポート(後期・教育学部)
 - ✓ 最終レポート

最終レポート課題

記述式レポート

次の2点について、それぞれ200字以上で回答せよ:

1. すべての回の内容の中で、最も面白いと思った話題とその理由
2. すべての回の内容の中で、今後最も自分の生活や学業等に役に立つと思うものとその理由

出題の意図

学生に、次の点を意識してもらいたい

- データサイエンスの知識や技能が日常生活で役に立つ場面がある
- データサイエンスは自分の生活や人生とどう関係しているだろうか

授業実施結果

- 授業の難易度
 - 授業難易度を授業初期と中期に学生に質問
 - 5段階で真ん中のちょうどよいという学生が最多
- 授業実施方法
 - 高度教材とPythonを使うはずだったクラスでもMicrosoft Excelで実施
 - 小テストはほとんどの学生が期限内に受験
 - PC演習
 - ✓ PC演習を遠隔で実施するために予定外の実演動画作成・配信
 - ✓ 実演動画配信に対する学生からの好意的な意見（学生が自分のペースで動画を見ながら演習を実施可能）
 - ✓ 対面授業再開後もPC演習動画は独習・復習用として利用可能
 - ✓ Mac利用者に十分な指導ができたかったのが悔やまれる
 - 質問対応
 - ✓ LMS上に質問箱を設置、後日スライドまたは動画で回答を公開
 - ✓ Zoom によるリアルタイム質問の利用者はほとんどなし



まとめと今後の課題

まとめ

- 学部1年生にリテラシーレベルの数理・データサイエンス・AI教育を行うことを目的として、文理を問わない数理・データサイエンス・AI教育リテラシーレベル教材を開発した
- 開発した教材は専攻に応じた2段階モジュール構成
 - コアコンテンツ：専攻に依らない共通教材
 - 高度／基本コンテンツ：専攻に応じて選択可能なPC演習用教材
- 教材を用いた授業の実施結果
 - 授業難易度は、ちょうどよいという学生が最多
 - PC演習の実演動画に対する学生からの意見は概ね好意的

今後の課題

- 様々な専門用語（統計学、コンピュータ、インターネット関連）の解説
- 数学やコンピュータが得意な学生に対する、少し高度な内容や課題の提供

本発表でご紹介した教材について詳しいことを知りたい方は下記までご連絡ください。

suzukika@yamanashi.ac.jp