

シラバス

授業科目名	年度	学期	開講曜日・時限	学部・研究科など	担当教員	配当年次	単位数
AI・データサイエンスツールⅢ	2021	前期	他		酒折 文武、武石 智香子、中村 周史	1年次配 当	2

履修条件・関連科目等

履修者数の上限は、100名程度となります。

授業で使用する言語

日本語

授業で使用する言語（その他の言語名）

授業の概要

AI・データサイエンス全学プログラムの応用基礎レベルに位置づけられ、スキルを身につけるツール科目の中でも、今後の応用力につながるツール科目である。BI（ビジネス・インテリジェンス）ツールツールやデータ分析言語を用いて、社会的な実課題に対してデータを活用することができる基本的なスキルを学んでいく。

前半は Exploratory というツールを扱う。大規模なデータを扱って可視化やデータ分析を行うデータサイエンスでは、R や Python などでのプログラミングが必須だと思われるがちである。しかしデータを扱う現場では Tableau や Microsoft Power BI などのBI（ビジネス・インテリジェンス）ツールを用いて、簡単な操作で効率的に可視化や分析を行い、迅速な意思決定につなげている。この科目で扱う Exploratory は「探索的データ分析ツール」と銘打っているが高機能な BI ツールとも考えられ、学生は無料で使用できることも大きな特徴である。この科目の前半では、それぞれの分野でデータ分析や可視化の活用方法を体得したい学生を対象とし、Exploratory を用いてデータの加工、可視化、分析を行い、他人に伝えるための技法について学ぶ。

後半は R というツール（言語）を扱う。R はデータ加工、可視化、分析などを行うための、データサイエンスでは欠かせないプログラミング言語であり、オープンソースで誰でも無料で新たなパッケージを開発・配布することが可能であるという特徴をもつ。同様の用途で用いられる Python 言語と比べて汎用性には劣るが、統計的な分析や可視化ならば R のほうが簡潔に行うことが可能である。さらに、tidyverse と呼ばれるパッケージ群を活用することで、データの加工を自由自在に行うことができる。この科目前半で扱った Exploratory も内部ではこの R が動いている。この科目の後半では、R 言語に触れたことのない学生を対象にして、tidyverse を用いたデータ加工や可視化を含んだ、データ分析の基礎的な技術を学ぶ。

本科目は【動画配信型オンライン授業】（オンデマンド）であり、動画によるレクチャーと実習部分からなる。

実習部分については、動画あるいは静的なウェブページの内容を参考にして、自身の PC で同じことを実行しながら進め、適宜演習問題にチャレンジしながら理解を深める。わからないところは manaba のスレッドなどを活用して講師やTAによるサポートが受けられるようにする。小課題は、課題がクリアできていない場合にはTAのサポートを得ながら再提出することが必要である。

科目目的

データサイエンスを各専門分野やビジネスで活用していくには、ビッグデータをラングリングし、視覚化・分析するツールが不可欠である。本科目の目的は、プログラムの基礎知識を要することなくこれらを行うことができる BI（ビジネス・インテリジェンス）ツールと、データサイエンスにおける基礎言語の一つである R の基礎、そしてデータ加工、可視化、データ分析の基本的な考え方を理解することである。

到達目標

本科目の到達目標は、数万件程度までの比較的小規模のデータについて、データ加工、可視化、基本的なデータ分析を自身で行えるようになることである。

授業計画と内容

【動画配信型オンライン授業】（オンデマンド）

■ 前半：BIツール

第1回 AIとBI、データ・エンジニアリングとツール

標題についての動画レクチャー

Exploratory の概要説明、R についての説明

Exploratory のインストール

○前半レポートの提示

第2回 データの読み込みとサマリー

ローカルデータの読み込み、Exploratory や R におけるデータの「型」

サマリーとデータテーブル、ハイライトと関連の活用

第3回 チャートや表による可視化と分析

棒グラフ・円グラフ・ヒストグラム・折れ線グラフなど基本的なグラフによる可視化

ピボットテーブルの活用、層別の重要性

○小課題1：規定のデータへのグラフ作成と解釈

第4回 データラングリング

データ整理と加工の大変さと重要性、外部データの取り込み、データ型の変換、フィルタ、集計
新規変数の追加、データテーブルの結合

第5回 データ・アナリティクス（統計と機械学習）

教師あり学習（予測・識別）と教師なし学習（クラスタリング・特徴抽出）
重回帰モデルの考え方と予測、クラスタリングの考え方とk-平均法
○小課題2：規定のデータへの重回帰モデルによる予測

第6回 ダッシュボードやノートによるレポート作成

ダッシュボードとは、Exploratory におけるダッシュボードとノートの作成
これまで学んだチャートの知識などを活用して、
自分なりのダッシュボードやノートを作成する演習中心回

第7回 BIツールのさらなる活用

スクレイピング、twitterとの連動とテキストマイニングなどの紹介
○前半レポート：与えられるテーマ（複数個ある中から自身で選択）に従って
データを収集・加工・可視化・分析し、結果を「ノート」にまとめてレポートとして提出

■ 後半：R

第8回 RとRStudio

R、RStudioのインストール（もしくは RStudio.cloud の活用）
BIツールに加えてR言語を学ぶ意味
コマンドの実行とスクリプト、四則演算、関数、代入（<-）、ベクトル・行列演算
○後半レポートの提示

第9回 データの読み込みと操作

パッケージのダウンロードと読み込み、データの読み込み（read.csv）
data.frameとtibble、read_csv、tidyverseの活用によるデータハンドリングの容易化
パイプ処理 %>%、簡単なデータ加工、日本語の文字コードとR（windowsの問題点）

第10回 グラフ作成とカスタマイズ

plot（基本的で汎用的だが汚い）、ggplotによる美しいグラフの作成、グラフへの日本語表示
○小課題3：既定のデータの加工とggplotによるグラフ作成

第11回 データラングリングの詳細

データ型の変換、フィルタ、集計、新規変数の追加
ワイド型データとロング型データの変換（pivot_wider, pivot_longer）、データテーブルの結合

第12回 データ・アナリティクス（統計と機械学習）

重回帰モデルによる予測、ロジスティック回帰モデルによる識別
k-平均法・階層型方法によるクラスタリング
○小課題4：規定のデータへのロジスティック回帰モデルによる識別

第13回 Rmarkdown によるレポート作成

markdown と Rmarkdown、チャック、数式の挿入（LaTeX形式）
これまでの知識を活用して自分なりのドキュメントを作成する演習中心回

第14回 Rによるプログラミングの基礎

分岐（if）、ループ（for）、自作関数の定義（function）
○後半レポート：与えられたデータに対する問題を解く（データ加工、分析まで）コードを書き、
Rmarkdown でのドキュメント作成して提出

授業時間外の学修の内容

授業終了後の課題提出

授業時間外の学修の内容（その他の内容等）

授業時間外の学修に必要な時間数/週

- ・毎週1回の授業が半期（前期または後期）または通年で完結するもの。1週間あたり4時間の学修を基本とします。
- ・毎週2回の授業が半期（前期または後期）で完結するもの。1週間あたり8時間の学修を基本とします。

成績評価の方法・基準

種別	割合 (%)	評価基準
レポート	80	●前半（Exploratory）レポート：40% 与えられたテーマに合致したデータを収集し、Exploratory を用いたデータ加工、可視化、分析を適切に行って考察をまとめることができるかを問う。 ●後半（R）レポート：40% 与えられたデータについて、R言語を用いて効率的にかつ適切にデータ加工、分析を行い、考察ができるかどうかを問う。
その他	20	小課題 20%（第3回、第5回、第10回、第12回の4回、各5点）。ただし、小課題は解けるまで再提出が必要であり、未提出の場合には単位認定しない。 小課題では、データ加工、可視化、分析などの与えられたタスクを実行できるかどうかを問う。

成績評価の方法・基準（備考）

課題や試験のフィードバック方法

授業時間に限らず、manabaでフィードバックを行う

課題や試験のフィードバック方法（その他の内容等）

アクティブ・ラーニングの実施内容

実習、フィールドワーク

アクティブ・ラーニングの実施内容（その他の内容等）

授業におけるICTの活用方法

その他

授業におけるICTの活用方法（その他の内容等）

各自のPCでExploratoryおよびRの実習を行う。

実務経験のある教員による授業

はい

【実務経験有の場合】実務経験の内容

（武石）シンクタンク勤務経験有り。

【実務経験有の場合】実務経験に関連する授業内容

（武石）時代は大きく変わり、ツールも進化したが、実務時代に行っていた現場のニーズを起点として最新ツールの活用に挑戦できるための基礎力を養う授業を行う。

テキスト・参考文献等

特定のテキストは使用しない。毎回資料を配布する。

その他特記事項

動画配信型オンライン授業を受講できる環境と、実習を行うことのできるPCが必要です。

参考URL

Exploratory: <https://exploratory.io>

R: <https://www.r-project.org>

RStudio: <https://rstudio.com>

コメント1

コメント2

コメント3

コメント4
