

2022年度データサイエンスの基礎を学ぶ フレッシュャーズ・セミナーの実施概要

青山スタンダード教育機構長 稲積 宏誠
aDSec 運営委員会委員長 荒木 万寿夫

1 データサイエンスの基礎を学ぶフレッシュャーズ・セミナーの概要

1.1 開講の背景

2010年代以降、スマートフォンやSNSの普及とビックデータの利活用によるデータ駆動型社会の実現は、やがて第4次産業革命の推進やSociety5.0の実現を担うAI人材の不足を指摘する声へと連なっていった。2016年の第5期科学技術基本計画（2016年1月閣議決定）や文部科学省の人材育成総合イニシアチブ、2017年4月の日本初の滋賀大学データサイエンス学部開設などは、AI人材の育成が急務であるという国家的課題への取り組みという流れの中で捉えることができる。

特に大学におけるデータサイエンス教育という観点からは、統合イノベーション戦略推進会議の『AI戦略2019』（2019年6月11日閣議決定）において、文理を問わずAIリテラシー教育を50万人に展開することが謳われたことがひとつの画期となり、文部科学省を中心とした数理・データサイエンス・AI教育プログラム認定制度が2021年に開始されると、各大学は当該領域における教育プログラムの拡充整備に一様に舵を切った。本学においても、全学展開されるデータサイエンスの基礎教育プログラムの開発が課題となった。

1.2 実施体制

ところで、我が国におけるAI人材育成や、そもそも大学教育におけるデータサイエンス教育の展開自体が諸外国に遅れた背景として、統計学部を持たなかった日本固有の教育事情が指摘されることがある。各学部各学科にデータサイエンス関連分野の教員が分属しているため、学部単独ではデータサイエンス関連の授業体系において、専門性と網羅性とを同時に持たせることが難しい一方、どの学部にも類似した初級クラスが開講されている。

そこで本学では、各学部に分属しているデータサイエンス関連分野の教員が、先ず相互に顔を合わせて情報共有ができる懇談会組織を作るべく2021年度から準備を始め、2022年5月に青山スタンダード教育機構内に青山データサイエンス教育コンファレンス（aDSec）を設置した。2022年度後期からは、aDSecに所属する教員が制作と監修を行うデータサイエンス教育第1弾として、全学教育であるフレッシュャーズ・セミナーとしてデータサイエンスの基礎を学ぶオンデマンド型授業を開講した。

開講準備は2021年度に遡り、本講義の開講趣旨や教育上の位置付け、具体的な講義概要等について、aDSec設立準備委員会（当時）で議論された後、学部長会においても経緯説明をし、また要望や懸念点の聴取を行った。並行して、企業の実務家データサイエンティストに対しても取材や講演を依頼するなど比較的時間をかけて準備を進めた。2023年現在、当該科目は必修とされていないが、オンデマンド型授業とすることで履修者数の上限は設けていないため、希望すれば全

学部全学科の学生が履修することができる。実際、2022年度は、ほぼ全学部全学科から合計640名の正規履修登録が得られた。

なお、aDSecは、運営委員会とDS教育評価委員会から構成されており、運営委員会は、aDSecの運営方針全般を議論すると共に、学部間で当該領域の教育に関わる情報や意見の交換を行う懇談会としての役割を担うほか、フレッシュャーズ・セミナー等のDS関連授業の実施状況を共有し、それら授業内容の評価や改善を行なっている。一方、DS教育評価委員会は、外部有識者から意見や提言を求める外部評価委員会を開催し、以って教育プログラム実施後の評価や改善に関わる助言を行なうものである。

1.3 講義概要

本講義の位置付け（シラバス記載内容）

近年、文理を問わず大学教育において、デジタル社会の「読み・書き・そろばん」として「数理・データサイエンス・AI」の基礎教育が必須とされており、さまざまな領域で活躍できるAI人材の輩出が社会的な要請にもなっている。このセミナーは、どの学部や学科でも必要となるデータサイエンスの入門科目として、数理・データサイエンス・AIに対する関心を高め、それらを理解して活用するために求められる基礎知識と技術を紹介し、各学部で開講されているより上位のデータサイエンス科目への橋渡しを図る。

この講義で身につけることのできる能力（シラバス記載内容）

様々なデータが入手でき、また利用できる機会が飛躍的に増加した現代において、それらデータに基づき、実務や日常生活におけるさまざまな問題の解決に取り組めるようになることは、本学で学ぶ全ての学生にとって必須の目標とすることができる。そのために必要なデータサイエンスに関する知識や技術の基礎を習得することが、まさにこのセミナーの目標となる。このセミナーは、データ分析を通じた問題解決のプロセスと、そこで求められる知識や技術に関する基本的な解説を与えるとともに、今後もいっそう発展し続けるであろうAI等の最新の技術動向についても、関心を持って学び続けることができる素地を身につける。

実施方法と特色

本講義を開講するにあたって、以下の点について配慮ないし工夫をした。

- 数理・データサイエンス・AI教育プログラム認定制度（リテラシーレベル）に応募する場合、多くの大学では複数科目を一体の教育プログラムとして申請しているが、本講義では1科目で申請要件を満たすよう、始めからこれに特化した授業内容を編成した。これは、学部単独では申請要件を満たす授業が開講されていない場合があったり、他学部科目の履修で補完しようとしても、他学部履修科目数の上限制約で履修出来ない場合があるなど、学内諸般の事情を考慮した結果である。
- 履修を希望する全学部全学科の学生が受講可能となるよう、オンデマンド型の授業形態とした。
- 本講義の難易度については、例えば記述統計の範囲でも基本統計量に関する説明から始めるなど、受験で数学を選択していない学生でも無理なく履修できるよう配慮している。

- 履修を検討する際、特に文科系学生にとって授業内容の数理的な難易度が懸念材料になる可能性が見込まれた。そこで、履修登録期間中に全ての授業動画を公開し、難易度を学生が直接確認できるよう配慮した。
- オンデマンド型授業一般に言えることだが、授業時間中、教員が一方的に話し続ける授業形態では、学生の集中力を維持することが難しい。そこで、一回の講義を約 50 分の座学 (Core Lecture) と約 30 分の PC 実習 (Tech Tutorial) とで構成されるよう設計した。50 分と 30 分の授業単位では、次回に持ち越すことなく纏まった単元を取り扱うため、これら授業単位をユニットと捉え、将来的には学部や履修学生の要望に応じて難易度を変えたユニットを複数用意し、授業内容や数理的な難易度のある程度カスタマイズできるよう備える予定である。
- 同様にオンデマンド型授業一般に言えることとして、授業動画の視聴を溜めてしまい、結果として未消化のまま試験を受ける、あるいは途中で履修を放棄する学生が一部に存在しうる。本学では、オンデマンド型授業については実施曜日や時限の指定はないが、毎週水曜日の午前 11 時から午後 1 時までをパブリックビューイングを称して同期視聴時間を設定した。当該時間帯には教員と TA, SA が待機し、リアルタイムで質疑に対応できる体制をとり、併せて履修者の学習支援も実施している。同期視聴時間以外も同様に、教員や TA が本学が採用している LMS (学習管理システム) 上で随時質疑に対応した。
- 授業では Python や R, KH Coder などの使い方も紹介しているが、機械学習の実習に至るまで、基本的に Excel を用いることとした。これは、主として一部からデータサイエンスの入門教育の段階でプログラミングを課すことに懸念が示されたことに基づく。今年度はプログラミングとして簡単な VBA を実習しているが、現在の履修者からの意見も踏まえ、来年度以降は Excel を主とする実習ユニットの他、Python を利用した実習ユニット等の提供も検討する。
- 実務家データサイエンティストの講演を正規授業として組み入れており、次年度以降も新規に講演を収録して差し替えるなど、日進月歩の当該領域における企業の取り組みを随時伝えられるよう準備している。また、授業に直接利用しない講演もアーカイブ化し、補助教材として公開する予定である。

単位認定の要件 (修了要件)

本講義では以下の修了要件を定め、これらを確認するテストやレポートを課し、以って単位の認定を行う。

- 近年におけるビックデータの利活用や AI の活用領域の広がりに関する基礎的な知識を身につけていること。
- AI 倫理や情報セキュリティに関する基礎的な知識を身につけ、当事者意識をもってデータ駆動型社会のリスクについて考えられるようになること。
- Excel や R, Python など利用し、記述統計的なデータの集計を行うことができ、併せて基礎的な機械学習の分析手法を理解していること。

なお、文部科学省の数理・データサイエンス・AI 教育プログラム認定制度 (リテラシーレベル) に対応したモデルカリキュラムに添った学習内容が、このセミナーだけで完結して学べるようシラバスを編成しており、修了すべき単位数は 2 単位である。

授業内容

この講義は、文部科学大臣が認定・選定する数理・データサイエンス・AI教育プログラム認定制度(リテラシーレベル)に応募できるよう、講義内容についても当該の認定要件を満たすよう特化した編成となっている。具体的には、授業内容は数理・データサイエンス教育強化拠点コンソーシアムがまとめたモデルカリキュラムに添っている必要があるが、以下に本講義のシラバスとモデルカリキュラム¹との対応関係を図1として示す。なお、図1を拡大した表も併せて付す。

授業日	区分	Core Lecture	対応モデルカリキュラム	講義内容	Tech Tutorial: Excel	講義内容とTech
第1回	導入	ガイダンス	-	科目概要、教員紹介、履修準備、成績評価の方法、その他	準備と数値表	準備メソッドでMS Officeのインストールを無印- https://www.aizaiyama.ac.jp/software/ees-office-install/
第2回		データの基礎とその活用	1.2.1 データの種類、1.2.2 データの所有者、1.2.3 機械化データと非機械化データ、2.1.3 SDDSデータを取り、2.1.8 図表と棒グラフ	*** 表形式データの構造 表頭、表体、ヘッダ、(データ)行 *** データの種類 - EDI、オープンデータ (datago.jp)、教育用データ提供システム、一般用クラウドデータ、his データ、クラウドデータ 1.2次データ、2次データ、3次データ、メタデータ、付帯表、属性表 (調査表) *** データの整理 調査データ、顧客データ、実験データ、ログデータ、アンケートデータ 加味表、有意味表 - 顧客表、契約表、経済センサス (装置) 表、簿記表、標準表、会計表、請求、家計簿、企業活動簿表 - 機械化データ、非機械化データ	Excelの基礎	*** Excelとは 代表的なワークシート関数 (sumaverage)、レンジ (maxmin) オートフィル、フィルハンドル、フィルターオプション グラフ作成 必須のキーボードショートカット
第3回		ビジュアルデータとデータ駆動型社会	1.1.1 ビジュアルデータ、1.1.2 機械化データとSNS、1.1.3 集約化データ、1.1.4 AIの倫理的側面、1.1.5 人間の知的知覚とAI、1.2.4 自動運転	*** ビジュアルデータとは ビジュアルデータという概念 (3つのV、認知の過程と背景)、ムアの原則 ビジュアルデータの活用EC *** AIとは - DX、AI、インダストリー4.0、IoT、スマートファクトリー *** AIとは - 強AI、汎用人工知能、弱AI、特化人工知能 機械学習、教師あり学習 (回帰分析)、強化学習、深層学習 (ニューラルネットワーク)	ワークシートの関数の活用 (1)	*** 基本統計 (分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準)) *** 対応関数 (分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準)) *** 対応関数 (分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準))
第4回		データサイエンスとは	1.4.1 数値的現象にデータを解析して生きている、1.4.2 さまざまなデータ解析の目的、1.4.3 さまざまなデータ解析のグループ、1.4.4 さまざまなデータ解析の発展、1.4.5 データ解析の発展、1.5.1 データ分析による意思決定、1.5.2 意思決定による意思決定、1.5.3 データ分析、意思決定の連携、1.5.4 総論の考え方、1.6.1 AI 等を適用した新しいビジネスモデル、1.6.2 AI活用後の活用	*** データ分析とは データ分析の種類、予測、分類、検出、推定 データ分析の目的、記述的データ分析、予測的データ分析、推定的データ分析 データ分析のプロセス データ分析のワークフロー、シミュレーション、データ利用、最適化	ワークシートの関数の活用 (2)	*** 基本統計 (分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準)) *** 対応関数 (分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準))
第5回		AI:その活用 (1)	1.3.1 事業活動におけるデータ-AI活用の広がり	*** 企業におけるデータとAIの活用領域 - 顧客関係 (顧客体験)、生産 (工場の自動化)、マーケティング、サプライチェーン最適化、リスク管理 - 顧客関係 (顧客体験)、生産 (工場の自動化)、マーケティング、サプライチェーン最適化、リスク管理 - 顧客関係 (顧客体験)、生産 (工場の自動化)、マーケティング、サプライチェーン最適化、リスク管理	ワークシートの関数の活用 (3)	*** 分布とヒストグラム 分散、標準偏差の性質と活用 ヒストグラムの作成、散布グラフの作成 (ヒストグラムによるヒストグラムの作成、石田正広 (石田 正広) の分析、所得分布と標準偏差の分散、度数分布表の作成、count
第6回		AI:その活用 (2)	1.3.2 活用目的ごとのデータ-AI活用の広がり	*** 企業活動におけるデータとAI活用の広がり - 顧客関係 (顧客体験)、生産 (工場の自動化)、マーケティング、サプライチェーン最適化、リスク管理 - 顧客関係 (顧客体験)、生産 (工場の自動化)、マーケティング、サプライチェーン最適化、リスク管理	Excelのデータベース的な使い方	*** Excelのデータ分析機能 - フィールド、レコード、SQL 文字列検索の関数 (left, mid, right) - 検索の関数 (find, findlast) - 検索の関数 (find, findlast) - 検索の関数 (find, findlast) - 検索の関数 (find, findlast)
第7回		心構え	2.1.1 分散、2.1.2 分散、2.1.3 分散、2.1.4 分散、2.1.5 分散、2.1.6 分散、2.1.7 分散、2.1.8 分散、2.1.9 分散、2.1.10 分散	*** AI活用 - シンクロナイズドとAI活用 - AI活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用	VBの活用 (1)	*** プログラミング入門 VBの活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用
第8回		データサイエンスとソフトウェアの活用	2.2.1 数値的現象にデータを解析して生きている、2.2.2 さまざまなデータ解析の目的、2.2.3 さまざまなデータ解析の発展、2.2.4 データ解析の発展、2.2.5 データ分析による意思決定、2.2.6 意思決定による意思決定、2.2.7 データ分析、意思決定の連携、2.2.8 総論の考え方、2.2.9 AI 等を適用した新しいビジネスモデル、2.2.10 AI活用後の活用	*** データサイエンスとソフトウェアの活用 - データサイエンスとソフトウェアの活用 - データサイエンスとソフトウェアの活用 - データサイエンスとソフトウェアの活用 - データサイエンスとソフトウェアの活用 - データサイエンスとソフトウェアの活用	VBの活用 (2)	*** 条件付き VBの活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用 - AI活用によるAI活用
第9回		基礎	2.3.1 データの種類、2.3.2 データの種類、2.3.3 データの種類、2.3.4 データの種類、2.3.5 データの種類、2.3.6 データの種類、2.3.7 データの種類、2.3.8 データの種類、2.3.9 データの種類、2.3.10 データの種類	*** データの種類 - データの種類 - データの種類 - データの種類 - データの種類 - データの種類 - データの種類	データサイエンスの基礎と活用 (1) データの整理に応じた可視化	*** データの種類 - データの種類 - データの種類 - データの種類 - データの種類 - データの種類
第10回		データサイエンスの基礎と活用 (2) データの分布と代表値	2.1.2 データの分布と代表値、2.1.3 代表値の性質の違い、2.1.4 データの分布、2.1.5 データの分布、2.1.6 データの分布、2.1.7 データの分布、2.1.8 データの分布、2.1.9 データの分布、2.1.10 データの分布	*** 記述統計 - 分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準)) - 分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準)) - 分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準)) - 分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準))	データサイエンスの基礎と活用 (2) データの分布と代表値	*** 対応関数 (分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準)) *** 対応関数 (分布の中心と分散 (度数)、平均とメジアン (性質の違い) (標準))
第11回	データサイエンスの基礎と活用 (3) 相関と回帰	2.1.7 相関と回帰、2.1.8 相関と回帰、2.1.9 相関と回帰、2.1.10 相関と回帰	*** 相関と回帰 - 相関と回帰 - 相関と回帰 - 相関と回帰 - 相関と回帰 - 相関と回帰	データサイエンスの基礎と活用 (3) 相関と回帰	*** 相関と回帰 - 相関と回帰 - 相関と回帰 - 相関と回帰 - 相関と回帰 - 相関と回帰	
第12回	機械学習 (1) 教師あり学習	2.1.5 教師あり学習に求められる前提条件、4.8 データ活用後の活用	*** 機械学習とは - 機械学習とは - 機械学習とは - 機械学習とは - 機械学習とは - 機械学習とは	機械学習 (1) 教師あり学習	*** 機械学習とは - 機械学習とは - 機械学習とは - 機械学習とは - 機械学習とは - 機械学習とは	
第13回	機械学習 (2) 教師なし学習	4.9 データ活用後の活用 (教師なし学習)	*** 教師なし学習 - 教師なし学習 - 教師なし学習 - 教師なし学習 - 教師なし学習 - 教師なし学習	機械学習 (2) 教師なし学習	*** Excelを用いたk-means法の活用 k-means法とは、ユークリッド距離 k-メータスを用いたk-means法 (k-メータス)	
第14回	深層学習 画像認識	4.8 画像認識、1.4.8 画像認識技術、1.4.9 人工知能	*** 画像認識 - 画像認識 - 画像認識 - 画像認識 - 画像認識 - 画像認識	深層学習 画像認識	*** 画像認識 - 画像認識 - 画像認識 - 画像認識 - 画像認識 - 画像認識	
第15回	総論と総括	-	講義内容のまとめと振り返り、試験に関する留意事項	総論と総括	-	

図1: 本講義とモデルカリキュラムの対応

¹ より正確には、モデルカリキュラムに準拠した教科書である北川源四郎、竹村彰通 編『教養としてのデータサイエンス (データサイエンス入門シリーズ)』(講談社, 2021) の章立てとの対応を示している。

授業回	区分	対応モジュール/カリキュラム	講義内容	Tech Tutorial : Excel	講義内容メモ
第1回	コア講義	—	科目概要、教員紹介、履修準備、成績評価の方法、その他 ## 実証方式データの構造 - 表頭、表側、表体、ヘッダ、(データ) 行列 ## データの具体例 -ID-POS、オンラインデータ (data.gov.jp)、教育用データ提供システム、一般用ミッドプロデータ、ins データ、タイムシリーズデータ 1次データ、2次データ、3次データ、メタデータ、符号表、質問紙 (調査票) ## データの識別 - 調査データ、観測データ、実験データ、ログデータ、アンケートデータ - 無作為抽出、有為抽出 - 志望調査、国勢調査、経済センサス (基礎) 調査、名簿情報 - 標本調査、母集団、標本、家計調査、企業活動基本調査 - 構造化データ、非構造化データ	準備と注意点	・情報システムでCEMS Officeのインストールを指示 https://www.aimaoyama.ac.jp/software/cees-office-install/
第2回	データの類型とその活用	1.2.1 データの種類、1.2.2 データの所有者、1.2.3 構造化データと非構造化データ、2.3 SSDSEデータを扱う、2.1.8 母集団と標本抽出	## データとは - データの種類 (分布の中心を測る尺度) - 平均、中央値、最頻値 (average, median, mode)、平均とメディアンの性質の違い (傾性) - 五数要約 (quartile, inc, quartile, exc), - 箱ひげ図、箱ひげ図の読み方と作成、外れ値 - 絶対参照と相対参照	ExcelのBriefing	## Excelとは - 代表的なワークシート関数 (sum, average), レンジ (max, min) - オートフィル、フィルハンドル、フィルオプション - グラフの作成 - 必須のキーボードショートカット
第3回	ビッグデータとデータ駆動型社会	1.1.1 ビッグデータ、1.1.2 検索エンジンとSNS、1.1.3 第4次産業革命、1.1.4 AIの驚異的発達、1.1.5 人間の知的活動とAI 1.2.4 自動翻訳	## ビッグデータとは - ビッグデータという概念 (3つのV、認知の過剰と豊穡)、ムーアの法則 - ビッグデータの活用とEC ## 第4次産業革命 - DX、AI、インダストリー4.0、IoT、スマートファクトリー データ駆動型社会とその社会像 ## AIとは - 狭いAI、汎用人工知能、弱いAI、特化型人工知能 - 機械学習、教師あり学習 (回帰問題)、教師なし学習 (k-means法)、強化学習、深層学習 (ニューラルネットワーク)	ワークシート関数の利用 (1)	## 基本統計量 (分布の中心を測る尺度) - 平均、中央値、最頻値 (average, median, mode)、平均とメディアンの性質の違い (傾性) - 五数要約 (quartile, inc, quartile, exc), - 箱ひげ図、箱ひげ図の読み方と作成、外れ値 - 絶対参照と相対参照
第4回	データサイエンスとは	1.4.1 誰もが無意識にデータを解析して生きている、1.4.2 さまざまなデータ解析一歩測、1.4.3 さまざまなデータ解析一歩測とビッグデータサイエンス、1.4.4 さまざまなデータ解析一歩測、1.4.5 データ解析の関連用語、1.5.1 データ分析による意思決定、1.5.2 情報技術による自動化、1.5.3 データ分析・自動化の実例、1.5.4 組織的学習と1.6.1 AI等を活用した新しいビジネスモデル、1.6.2 AI最新技術の活用例	## データ分析とは - データ分析の種類、予測、分類、発見 - データ分析の目的、記述的データ分析、予測的データ分析、指示的データ分析 - データ分析のプロセス - データ分析のアプリケーション、データ同化、最適化 ## データサイエンスとは - 保科先生の講演	ワークシート関数の利用 (2)	## 基本統計量 (分布の中心を測る尺度) - 平均、中央値、最頻値、レンジ、五数要約の活用と実習 - オートフィル補足 - 分散と標準偏差
第5回	AIとその活用 (1)	1.3.1 事業活動におけるデータ・AI活用の広がり	## 企業におけるデータとAIの活用領域 - 山口雄大氏 (経営者 当時) の講演 - 毛澤レター：竹田豊 経営学部教授 第4次産業革命、Society5.0時代におけるビジネス環境の変化 - AI活用の取り組みとビジネスにおいてAIが可能にしたこと - データドリブン経営、AIを用いた需要予測、企業が求めるこれからのAI人材	ワークシート関数の利用 (3)	## 分布とヒストグラム - 分散、標準偏差の活用と実習 - ヒストグラムの作成、棒グラフの作成と構構によるヒストグラムの作成 - 右に歪んだ (右裾が長い) 分布、所得や資産資産階級の分布 - 度数分布表の作成、countif
第6回	AIとその活用 (2)	1.3.2 活用目的ごとのデータ・AI活用の広がり	## 企業活動におけるデータとAI活用の広がり - 高橋裕美氏 (NEC) の講演 - 毛澤レター：竹田豊 経営学部教授 - Society5.0とSDGs - 社会課題とAIの活用、スマートシティ、スマートロブ、労働力不足	Excelのデータへの使い方	## Excelのデータ分析的な使い方 - フィールド、レコード、SQL - 文字列操作の関数 (left, right, mid) - 関数の使い方とラベルの付与 (格付け) - vlookup関数の使い方 - レコード検索とオートフィル - ビットデータとクロス集計、スライサー
第7回	心得	3.1.1 ELSI、3.1.2 一般データ保護規則(GDPR)、3.1.3 十分に限定、3.1.4 AI倫理、3.1.5 AI倫理論、3.1.6 プラウドプロセス化、3.1.7 説明可能性、3.1.8 アカウンタビリティ、透明性、トランスパレンシー、3.1.9 公平性、3.1.10 データ・AI活用における負の事例紹介 データの活用・目的外利用、3.1.11 データ・AI活用における負の事例紹介 データサイエンスとプライバシー、3.2.2 データサイエンスと情報セキュリティ、3.2.3 データサイエンスとプライバシー	## AI倫理 - ELSI研究 - プレキキュラティブとAI倫理論 - AI倫理のガイドライン、安全性、アカウンタビリティ、説明可能性、トランスパレンシー - ## データと情報セキュリティ - GDPR (一般データ保護規則)、データオーナーシップ権、十分に限定 - 情報セキュリティの3要素 (Confidentiality, Integrity, Availability) - OECDの原則、プライバシー保護、個人情報保護法 - データサイエンスと個人データ、統計的差別、データバイアス、アルゴリズムバイアス	VBAの利用 (1)	## マクロプログラミング入門 - VBA開発環境の準備 - マクロミニマムとは、変数、数値や文字列の代入 (cell(0), Range(0), Row(0)とマクロ実行のシミュレーション - 繰り返す処理 (Forループ)

第8回	データサイエンスとソフトウェアの利用	2.3.1 垂形式のデータ、2.3.2 データ解析ツール、1.4.6 非構造化データ処理、4.5 デキスト解析	## データサイエンスとソフトウェアの利用 ~ Python ~ - 構造化データの分析 - Kaggle と タイタニック データのダウンロード - Python と Google Colab の利用の仕方 - Python とライブラリ、pandas - 垂形式データとデータフレーム - ソフトの概念と使い方 (info, describe, head, value_count, isnull, sum, mean, filna, groupby, pd.cut, pd.value_counts) ## データサイエンスとソフトウェア ~ KH Coder ~ - 非構造化 (テキスト) データの分析 - 共起ネットワーク分析, ショウカード集数, サブグラフ, 共起ネットワーク図	VBAの利用 (2)	## 条件分岐 - VBAからのワークシート関数の利用, Rept() - 条件分岐 (If文), フロッキング, mod, int - 繰り返し処理と条件分岐の組み合せ処理 (各歳から年齢階級区分への格付け処理) - リンクダイナミックとピボットテーブル
第9回	基礎	1.4.7 データ可視化、2.1.1 データの種類、2.2.1 データの表現、2.2.2 データの図解表現、2.2.3 データの比較、2.2.4 不適切なグラフ表現、2.2.5 優れた可視化の例	## データの可視化と工夫されたグラフ - 2軸複合グラフ, ABC分析とパレート図、パウルチャート - 特異表示	データサイエンスの基礎と実習 (1) データの種類の応じた可視化	## データの可視化と工夫されたグラフ - 2軸複合グラフ, ABC分析とパレート図、パウルチャート - 特異表示
第10回	データサイエンスの基礎と実習 (2) データの分布と代表値	2.1.2 データの分布と代表値、2.1.3 代表値の性質の違い、2.1.4 データのばらつき、2.1.6 打ち切りや脱落を含むデータ、層別の必要なデータ	## 記述統計 - 分布の中心を測る尺度 (復習), リリ及び平均 - 分布のバラバラを測る尺度 (復習) - 総和の演算子と - 標準化 (標準化), standardize(), 偏差値 - 異なる属性が混在したデータや測定単位異なるデータの取り扱い - 変動係数	データサイエンスの基礎と実習 (2) データの分布と代表値	## 採点結果表のグラフ別集計 (実習) - 繰り返し処理, 条件分岐の復習, 偏差値の算出 - ワークシート関数の利用 (復習) - 条件分岐と計算の方法 (カンツダの利用) - ワークシートの追加と名前の定義 - 配列の利用 (array)
第11回	データサイエンスの基礎と実習 (3) 相関と因果	2.1.7 相関と因果性、2.1.9 クロス集計表、相関係数行列、散布図行列、2.1.10 統計情報の正しい理解	## 相関係数の実習 - Pythonを利用した相関分析の例 - [japanize, mathlibとseasonomの利用, pairplot() - 相関係数行列と散布図行列 - 2重ループの利用 - VBAを利用した相関係数行列の算出	データサイエンスの基礎と実習 (3) 相関と因果	## 相関分析の実習 - Pythonを利用した相関分析の例 - [japanize, mathlibとseasonomの利用, pairplot() - 相関係数行列と散布図行列 - 2重ループの利用 - VBAを利用した相関係数行列の算出
第12回	機械学習(1) 教師あり学習	2.1.5 教師データに含まれる誤差の扱い、4.8 データ活用実践 (教師あり学習)	## 機械学習その1 - 機械学習とは、その類型 - 教師あり学習 (回帰問題, 分類問題) と教師なし学習 (特徴抽出), 特徴量決定木 (分類木) とは、ジニ不純度 - Pythonを利用した決定木分析, ライブラリ - sklearn, 訓練データとテストデータ - Iris データを用いた分析例 - タイタニック データを用いた分析例 (欠測値補完は第8講)	機械学習(1) 教師あり学習	## 回帰問題 - 回帰分析とは (単回帰分析) - 最小二乗基準とソルバーを利用した最適化 - 解析的に解くということ - 回帰係数の解釈 - 重回帰分析とは - 重回帰分析の分析例, Excelの分析ツールの利用
第13回	選択 機械学習(2) 教師なし学習	4.9 データ活用実践 (教師なし学習)	## 機械学習その2 - 数ベクトルと幾何ベクトル, スカラー倍とノルム - ベクトルの内積, 射影 - 主成分分析と次元の削減 - Excelのソルバーを用いた主成分分析の求め方 - Iris データを用いた主成分分析の分析例 - Excelを用いた識別散布図の描き方 - 主成分分析に関する備忘と諸注意	機械学習(2) 教師なし学習	## Excelを用いたk-means法の実習 - k-means法とは、アルゴリズムの説明 - Iris データを用いたk-means法の分析例 - average()
第14回	深層学習 画像認識	4.6 画像解析、1.4.8 パターン認識技術、1.4.9 人工知能	## 深層学習 - 非構造化データとしての画像データ、画像データの記録の仕組み - MNIST データセットとは - VBAを用いた画像表示 (RGB画数の利用) - 画像認識に用いるニューラルネットワークの解説 - keras と tensorflow - 過学習とは	深層学習 画像認識	## パーセプトロンと論理回路 - 単純パーセプトロンとは - 単純パーセプトロンと隠形性、多層パーセプトロンと非線形性 - Excelを用いたニューロンの学習
第15回	総括と補足	-	講義全体のまとめと振り返り、試験に関する留意事項等	総括と補足	

1.4 履修者から寄せられた授業の感想

ここでは、期間中に随時募集していた授業アンケートに寄せられたコメント（自由記述抜粋，原文ママ）を紹介する。本フレッシュャーズ・セミナーの履修を検討する際の参考として活用されたい。

- 大人数クラスのオンデマンド授業という他のフレッシュャーズセミナーとは大きく異なる授業形態で、履修登録時は授業形態が全く想像できず心配だったが、いざ始まってみるとデータサイエンスという今の時代に合った内容を自分のペースで学習することができ、非常に有意義であった。しかしながら、授業内容のボリュームが多く、多くの人が実習までこなすことができなかつたために、掲示板もあまり有効に活用できなかつたと思う。このボリュームであれば、（できるかできないか分からないが）通年開講の形をとってみても面白いかもしれない。半年間ありがとうございました。
- 新一年生にお勧めできる授業であると思います。昨今、理系の学生だけでなく、文系の学生にも AI 教育（+自主的な学習）が必要な時代になってきました。この授業は、AI に関する知識をつけたり、プログラミングに関する学習をスタートする最初のステップとして最適であると思いました。個人的には、教場テストよりも小テスト+レポートの方が学習しやすかつた気がします。私はエクセルの知識など非常に浅い人間だったので、詳しくエクセルの実用の仕方やその他知識に関して学べてとてもよかつたです。
- 授業内容は今の時代にぴったりの構成になっており、非常に興味深いものであつた。又、オンデマンド形式で 14 回分の講義が予め配信されていることも自分のペースで学習できるという点において、非常にやりやすかつた。しかしながら、私の場合、水曜日に授業が入っており、その分、他の曜日の空きコマで受講したのだが、空きコマで受講すると動画を見るだけで 90 分経ってしまい（動画と並行して練習するのは個人的に難しかつた）、Tech Tutorial を実際に自分で練習する時間を別日に取らなくてはいけない→他のオンデマンド授業（期限あり）も見なくてはならない→次の本授業講義の設定日近辺→先に動画（Core/Tech）を見る→…（繰り返し）…→Tech Tutorial の演習を十分消化できていないまま、次の週の内容の学習を行ってしまい、知識と技術の習得がアンバランスになってしまった。動画がもう少し短いと、一コマの空きコマの中で動画の視聴から演習まで一通り行えると思う。又、時間があるときに次週の内容を受講しておいたり、逆に時間が取れなかつた週は受講を行わなかつたりと、自分なりのペースで学習を行つたのだが、事前に資料や動画を開いてしまったのもあつて、どの講義を受講し、どの講義を受講していないのか分からなくなつてしまった時があつた。出席申請を各講義の受講想定日をパスワードとして設定するのも良いかもしれない。半年間ありがとうございました。
- 大学の場合、テストが返却されないため、知識を詰め込んで、吐き出すだけの生徒が多く見られます。今回のような時間制限が長く、調べながら学習を行える（学習の振り返りができる）小テストや出した解を形式や文章を正しながらまとめ直す（頭の中でも学んだことを整理できる）というステップが加わるレポートの製作の方が、一発の教場試験よりもデータサイエンスという大きな分野の入門編としては体感的にやりやすかつた気がします。ありがとうございました。
- 私はこの授業をするまではパソコンに関する仕組みやエクセルの使い方もままならないような状態だったので、基礎からやっただけで良かつたです。データサイエンスという言葉も聞いたことはありましたがどのようなものかは知らなかつたし、かなり広義なのだともわかりました。私にとっては全体的に難しかつたのですが、特に難しかつたのはベクトルが出てき

たところですが。改めて数学の重要性を知ることができたので勉強しなおさなければならないと思いました。この分野に興味を持つことができたのでこれからも学んでいきたいと思えます。

- オンデマンド授業なので、自分のペースで講義が進められ、また、わからないところをいつでも見直すことができるため、非常に良かった。そして、小テストも復習として自分の得意な分野と苦手な分野の区別がしっかり確認できたので良かった。後期の間ありがとうございました。
- 経済学部に所属しているので、このような内容は学習できないかと思っていましたが、今注目されている AI のことなどを学べてうれしく思います。ありがとうございました。
- まずは半年間ありがとうございました。私はこれまであまりデータ分析に触れてくることがなかったため、初めはよく分からないことばかりで理解するのに講義を3回見直したこともありました。しかし、特に Excel に関して、1度活用できるようになると、とても楽しくて、すごく達成感がありました。期末課題では分からない問題が多く、自力での解決が難しいものもありました。今後、今回できなかった部分を含めて、Excel を利用したデータ分析を個人的にも学習していきたいです。また、AI が持ち合わせる将来の可能性に関するお話もすごく興味深かったです。マーケティング学科の学生として、AI とデータ分析・マーケティングについてより詳しく学習していきたいと考えます。ありがとうございました。