

授業科目区分		科目名		単位	科目コード	開講時期	履修方法		
(全課程からの提供) リベラルアーツ系科目 文理横断		ビジネスデータサイエンス(夏期集中講義)		1	G246-01	1期(前学期)	修学規程第4条を参照		
担当教員名		研究室	内線電話番号	電子メールID			オフィスアワー		
授業科目の学習・教育目標									
キーワード		学習・教育目標							
1	統計学の基礎	近年急速に発展しつつあるAI(人工知能)の利用には、膨大なデータが必要である。しかし、AIを活用する以前に、蓄積されたデータを集約・整理・分析するデータ分析を行うことにより、解決できる問題も少なくない。本講義では、データ分析に関する基本的な概念や手法を学ぶとともに、実践的な演習を行う、このような講義と演習を通じて、データ解析の基本的な手法を習得して、実際にデータ活用を行えるようになることを目指す。							
2	データ分析								
3	データに基づく予測								
4	データに基づく分類								
5	データサイエンス								
授業の概要および学習上の助言									
<p>本授業では、データを活用するための手法としてのデータ分析を、講義と演習を通じて学ぶ。最初に、記述統計の基本的な概念を復習するとともに、推測統計の点推定と区間推定および仮説検定の概要を学ぶ。基本統計量を求めること、2群の平均値に「統計的に有意な」差があるかどうかの検定などに関して、演習を通じて実際のデータ処理手順を習得する。次に、統計的手法の代表である回帰分析を主として、多変量解析の理論と実践を学ぶ。さらに、いわゆるデータマイニングの諸手法に関して、理論と実践を学ぶ。いずれの回も、座学による講義を前半に実施して、後半には実際に手を動かして分析を行う演習を実施する。その演習を実施できるように、毎回各自のパソコンを持参することが必須である。なお、データ分析のツールとしては、ビジネスにおいて広く用いられている表計算ソフトであるマイクロソフト・エクセル、フリーウェアである統計ソフトKNIMEおよびテキストマイニング・ツールのKHCoderを使用する。各手法の数理的背景となる数式的説明は、手法の理解と適用限界に関する最小限にとどめるように配慮して、利用方法と適用範囲に時間を割く。また、ビジネスにおけるデータ利用とされているビジネスインテリジェンスの活用事例についても適宜紹介する。予習は特に必要とはしないが、不明な点は授業時間中に解決するとともに、復習には時間を割いて欲しい。本授業を通じて、受講生の所属部門や企業に蓄積された数値データやテキストデータを基に、要因分析や知識発見などのデータ分析に基づくビジネス展開ができる力を身につけることを期待する。なお、本授業は遠隔リアルタイムで実施予定である。教科書は用いないが、講義資料等に関しては本学eシラバスの他に、バックアップ用としてBox等のツールを用いて配布する予定である。</p>									
【教科書および参考書・リザーブドブック】									
教科書：指定なし									
参考書：指定なし									
リザーブドブック：指定なし									
履修に必要な予備知識や技能									
<p>本講義で用いるソフトウェアは 1) マイクロソフト・エクセル、2) 統計ソフト KNIME(https://www.knime.com/)、3) テキストマイニング・ツール KHCoder(http://khcoder.net/d13.html、「AI基礎」でインストール済である人は再インストールは不要。アンインストール等で使えない人はStarting Editionをインストール願います)、4) ウェブGIS jSTAT MAP(https://jsstatmap.e-stat.go.jp/jsstatmap/main/trialstart.html)である。KNIME と KHCode はフリーウェアであるので、各自のPCに事前にインストール、さらにjSTAT MAPは事前登録をしておいてほしい。なお、詳細は第1回目のガイダンスでも説明する。</p>									
No.	学科教育目標 (記号表記)	学生が達成すべき行動目標							
①	A, D	平均値などの代表値が統計的に有意な差があるかどうかの検定ができる							
②	A, D	所与の多変量データに対して、適切な重回帰分析を行うことができる							
③	A, D	データサイエンスにおける、いくつかの統計的手法およびデータマイニング手法を説明できる							
④	A, D, G	統計的手法およびデータマイニング手法を用いて、実際に予測や分類さらに要因分析や知識発見を行うことができる							
⑤	A, D, G	データサイエンスに関する各手法の意義が認識でき、さらにその限界についても述べることができる							
⑥									
達成度評価									
評価方法		試験	クイズ 小テスト	レポート	成果発表 (口頭・実技)	作品	ポートフォリオ	その他	合計
指標と評価割合									
総合評価割合		0	0	70	0	0	0	30	100
総合力指標	知識を取り込む力	0	0	40	0	0	0	0	40
	思考・推論・創造する力	0	0	20	0	0	0	0	20
	コラボレーションとリーダーシップ	0	0	0	0	0	0	0	0
	発表・表現・伝達する力	0	0	10	0	0	0	0	10
	学習に取組む姿勢・意欲	0	0	0	0	0	0	0	30

※総合力指標で示す数値内訳は、授業運営上のおおよその目安を示したものです。

評価の要点

評価方法	行動目標	評価の実施方法と注意点
試験	①	
	②	
	③	
	④	
	⑤	
	⑥	
クイズ 小テスト	①	
	②	
	③	
	④	
	⑤	
	⑥	
レポート	①	レ
	②	レ
	③	レ
	④	レ
	⑤	レ
	⑥	
成果発表 (口頭・実技)	①	
	②	
	③	
	④	
	⑤	
	⑥	
作品	①	
	②	
	③	
	④	
	⑤	
	⑥	
ポートフォリオ	①	
	②	
	③	
	④	
	⑤	
	⑥	
その他	①	
	②	
	③	レ
	④	レ
	⑤	レ
	⑥	

具体的な達成の目安

理想的な達成レベルの目安	標準的な達成レベルの目安
<p>講義で紹介した各種のデータ分析手法に関して統計分析ソフトKNIMEを利用して適切に実施できるだけでなく、各手法の意義を理解しており、さらにその適用限界について説明することができる。</p> <p>さらに、自ら保有しているデータに関して、適切な解析手法を適用して要因分析や知識発見などを行うことができる。</p>	<p>所与の数値データに対して、基本統計量を求めて、2群の平均値に有意な差があるかどうかを検定できる。</p> <p>また、所与の数値データに対して、重回帰分析、決定木分析、クラスター分析などの主なデータ分析手法を、統計分析ソフトKNIMEを利用して適切に実施できる。</p>

CLIP学習プロセスについて

一般に、授業あるいは課外での学習では：「知識などを取り込む」→「知識などをいろいろな角度から、場合によってはチーム活動として、考え、推論し、創造する」→「修得した内容を表現、発表、伝達する」→「総合的に評価を受ける、Good Work!」：のようなプロセス（一部あるいは全体）を繰り返し行いながら、応用力のある知識やスキルを身につけていくことが重要です。このような学習プロセスを大事に行動ください。※学習課題の時間欄には、指定された学習課題に要する標準的な時間を記載してあります。日々の自学自習時間全体としては、各授業に応じた時間（例えば2単位科目の場合、予習2時間・復習2時間/週）を取るよう努めてください。詳しくは教員の指導に従ってください。

回数 日付	学習内容	授業の運営方法	学習課題(予習・復習)	時間(分)※
1 /	統計の基礎 ・集計 ・記述統計と推測統計 ・基本統計量：平均，分散，標準偏差 ・パラメトリック検定：F検定，t検定	講義 Excelを用いる演習 討議	演習の復習 レポート作成	120分 180分
2 /	多変量解析Ⅰ ・線形重回帰分析 ・数量化Ⅰ類 ・予測	講義 Excelを用いる演習	演習の復習 レポート作成	120分 180分
3 /	多変量解析Ⅱ ・ロジスティック回帰分析 ・主成分分析 ・因子分析	講義 統計解析ソフトKNIMEを用いる演習	演習の復習 レポート作成	120分 180分
4 /	クラスター分析 / 決定木 ・類似度と距離 ・凝集法とデンドログラム（樹形図） ・K-Means法 ・交差妥当性 ・CART法 ・C5.0法	講義 統計解析ソフトKNIMEを用いる演習	演習の復習 レポート作成	120分 180分
5 /	ニューラルネットワーク ・学習パラメータ ・予測と分類 ・AIとディープラーニング	講義 統計解析ソフトKNIMEを用いる演習	演習の復習 レポート作成	120分 180分
6 /	テキストマイニングの基礎 ・形態素解析 ・潜在意味解析 ・共起ネットワーク	講義 テキストマイニングソフトKH Coderを用いる演習	演習の復習 レポート作成	120分 180分
7 /	GIS(地理情報システム)の基礎 ・統計情報の地図上への展開 ・地図上での情報の編集・検索・分析・管理 自己点検	講義 ウェブGIS jSTAT MAPを用いる演習 討議	演習の復習 振り返り	120分 80分