

開講期間	配当年	単位数	科目必選区分
半期	1 年	2	選択
担当教員			
築 雅之			

授業のねらい（概要）	データ処理のためのプログラミング、統計分析のためのプログラミングを学ぶ。 具体的には製品のアンケート等の統計的分析、webサーバ等のログ解析、SNS等のテキストマイニングなど、インターネットビジネスやマーケティングの現場で必要となるデータ分析を想定し、データ構造とアルゴリズム、統計学的手法を実データを使った実習も交えて学ぶ。 上記のため環境として、PC上に仮想マシンを作成し、OSを含むソフトウェア環境をゼロから構築する。 可能であれば、協定締結企業との協働して、実際のマーケティングデータの分析やサーバログ解析も行ないたい。		
授業計画	1, 2, 3回	データ処理環境の整備とawkによるテキスト処理 教科書1に準拠して、awkを使ったテキストデータ処理、データ構造とアルゴリズムを学ぶ。並行してdebianのインストール、文字コード体系の確認と相互変換を行なう。 予習（時間）：教員の指示で以下のいずれか(あるいは複数)の準備が求められる。(180分) 1. PC上の実習環境の構築 2. 教科書あるいは教員の指示によるプログラミング課題 3. コンピュータ科学あるいは統計学に関わる事項の確認 復習（時間）：自らの予習内容と授業時の講義・実習とを振り返り、教員の助言を踏まえて知識面、実習面での不足を補う。(60分)	
	4, 5, 6回	Rのデータ構造、統計処理 教科書2に準拠して、Rの制御構造と、R特有のデータ構造を学び、統計分析の手法を学ぶ。併せてPSPP(SPSS互換のOSS)にて同様の処理を行なう。 予習（時間）：教員の指示で以下のいずれか(あるいは複数)の準備が求められる。(180分) 1. PC上の実習環境の構築 2. 教科書あるいは教員の指示によるプログラミング課題 3. コンピュータ科学あるいは統計学に関わる事項の確認 復習（時間）：自らの予習内容と授業時の講義・実習とを振り返り、教員の助言を踏まえて知識面、実習面での不足を補う。(60分)	
	7, 8, 9回	awkとRを組み合わせたビジネスデータ処理・統計分析 教科書3に準拠して、生データ処理のフロントエンドにawkを使い、Rの分析に適した形式に変換し、統計分析を行なう。よりビジネス寄りの素材を扱う。併せてR-Studio、R-Markdownのインストールを行ない、統合処理環境を構築する。 予習（時間）：教員の指示で以下のいずれか(あるいは複数)の準備が求められる。(180分) 1. データ分析の試行 2. 分析結果から得られる知見の文書化 復習（時間）：自らの予習内容と授業時の講義・実習とを振り返り、教員の助言を踏まえて知識面、実習面での不足を補う。(60分)	
	10, 11, 12回	ビジネス課題解決のための事例 教科書4に準拠して、ビジネス上の課題解決のためのデータ分析の事例を学ぶ。 予習（時間）：教員の指示で以下のいずれか(あるいは複数)の準備が求められる。(180分) 1. 教科書事例データの再分析 2. 分析結果を踏まえた課題解決のための提案の文書化 復習（時間）：自らの予習内容と授業時の講義・実習とを振り返り、教員の助言を踏まえて知識面、実習面での不足を補う。(60分)	
	13, 14回	実データの分析と検討 協定締結企業との協働で、マーケティングデータやサーバログ解析を行なう 予習（時間）：教員の指示で以下のいずれか(あるいは複数)の準備が求められる。(180分) 1. 協定締結企業から提供されたデータの分析 2. 分析結果の文書化 復習（時間）：自らの予習内容と授業時の講義・実習とを振り返り、教員の助言を踏まえて知識面、実習面での不足を補う。(60分)	
	15回	まとめ 予習（時間）：教員の指示で期末課題に必要なものを補う。(180分) 復習（時間）：教員の助言を踏まえて期末課題の準備を行なう。(60分)	
授業を通して身に付けることができる能力（DP）	DP(商学研究科)の以下の項目を意識した科目となっている。 ●調査、研究のために必要となる様々なスキルを身に付け、企業や産業が直面する諸課題に対し高度な判断能力、分析能力を活用することができる。		
到達目標	1. 独力でlinux環境を構築し、必要なソフトウェアパッケージの選択ができる。 2. プログラミング言語awkに習熟し、ある程度構造化されたテキストデータの処理ができる。 3. データ構造とアルゴリズムの基本的な知識の裏付けられたプログラミングができる。 4. 代表的な多変量解析までの統計手法の知識がある。		

	5. プログラミング言語RおよびR-Studioが使える、上記の統計手法を使った分析ができる。 6. 統計学的検定を利用した論文を理解することができる。また検定結果と考察が含まれた研究レポートを書くことができる。
課題や小テスト等のフィードバックの方法	少人数が想定されるため、授業時間内外を通して、想定する成果が得られるようにコメントや修正指示を加えつつ個別指導を行う。
履修上の注意	以下の事項は習得済みとして授業を行なう。履修希望者で未修の者は開講時までに自学すること。自学のための教材や学習法についての質問は前年度中から開講年度履修登録期間まで受け付ける。 1. プログラミングの基礎(手続き型言語の制御構造と基本データ型) 高等学校あるいは本学商学部「プログラミング基礎実習」で学習する水準 2. 統計学の入門(データの散らばり・相関) 高等学校「数学I」「情報の科学」あるいは本学商学部「統計学」で学習する水準 また、ノートPCやタブレットPCを所有する受講生は、持参すること。
成績評価の方法・基準	毎回の課題等 50% 期末の課題 50%
教科書	ISBN番号：978-4904807002 書籍名：プログラミング言語AWK 著者名：エイホ、カーニハン、ワインバーガー 出版社：USP研究所 発行年：2010 価 格：3,500円(税抜き) 備 考：他にも新紀元社、アジソン ウェスレイ・トッパンからも出版されているが同一書籍である。古本含めていずれを購入しても構わない。 ISBN番号：978-4274067570 書籍名：Rによる統計解析 著者名：青木繁信 出版社：オーム社 発行年：2009 価 格：3,800円(税抜き) ISBN番号：978-4798134901 書籍名：Rではじめるビジネス統計分析 著者名：末吉 正成， 里 洋平， 酒巻 隆治， 小林 雄一郎， 大城 信晃 出版社：翔泳社 発行年：2014 価 格：3,200円(税抜き) ISBN番号：978-4797376333 書籍名：ビジネス活用事例で学ぶ データサイエンス入門 著者名：酒巻 隆治， 里 洋平 出版社：SBクリエイティブ 発行年：2014 価 格：1,980円(税抜き)
参考書・教材	
備考	講義科目
教員との連絡方法	大学ドメインのメール、Mellyでの個別連絡を併用。Line、Gmail等は原則として使用しない。