データ分析と統計解析基礎

Data Analysis and Foundation of Statistical Analysis

担当教員	粂野文洋（客員）	**単位数**	2単位
開講学期	春学期	**開講曜日・時限**	水曜日・土曜日 1・2時限目
位置づけ	経済・統計・分析 基礎段階		
区分	総合・実践科目		

科目紹介	
科目の重要性・必要性	ビジネスにおいて、様々なデータを分析し、客観的な判断を行うデータサイエンスの重要性が高まっている。本科目では、データサイエンスの実践に必要となる統計学の基礎知識を学ぶ。本科目を受講することにより、①データの「ばらつき」の概念、②基本統計量の理解、③統計手法を用いたデータ分析の仕方、④Excelのデータ分析ツールを用いた統計解析の方法など、データ解析手法に関する基礎知識を修得できることに加え、データ解析結果の読み方についても学ぶことができる。
科目の目的	授業で習得したデータ解析手法を実務で使用できるようになることを目的とする。
到達目標	データサイエンスの重要性を説明でき、その基礎としての数理統計の知識、すなわち、基本統計量、正規分布、検定や推定、多変量解析の基本を説明できる。また、実務データを統計処理し、分析することにより、データに基づく客観的な判断や問題解決に役立てることができる。
受講してもらいたい院生	本科目は、共通基礎科目に位置付けられており、これまでに統計学/統計解析を基礎から学んだことが無い人を対象とした科目である。MOTの他の科目を理解するためには、本科目は必要不可欠な科目の1つであるため、統計学/統計解析の知識を持っていない初学者にとって分かりやすい講義を心掛ける。現在は、ビッグデータ時代とも呼ばれビジネスにおけるデータ分析の重要性が高まっている。データ分析の差、すなわち情報の価値化の違いが、企業の意思決定の差となって現れるため、本科目ではデータの持っている重要な情報を引き出すための手法について学ぶとともに、得られたデータ分析結果をどのように読み解くか、についても演習を通して体験していただきたい。

授業計画		
第1回	予習・復習時間	4時間
	予習・復習内容	予習：シラバスを読み、職場で抱えている問題や今後「統計解析」手法を利用して取り組みたいと考えている問題を洗い出しておくこと。 復習：PPDACサイクルを見直し，職場でどのように適用できるか、または適用してみたいかを考えておくこと。
	授業内容	【データサイエンスと統計学】 本科目のオリエンテーションを述べた後、データサイエンスとは何かと，その問題解決フレームワークであるPPDACサイクルについて、事例を使って説明する。さらにその実践に必要な統計学の役割について学ぶ。 【課題】自分の業務あるいは職場において、データ分析によって解決したい課題は何か、そのためのデータとしてどのようなものが収集されているかを調べること。また、これらのデータとは別に職場で必要なデータにはどのようなものがあるかを考えること。
	授業課題	有
	課題フィードバック方法	解答例示のみ
第2回	予習・復習時間	4時間
	予習・復習内容	予習：事前に、基礎統計量にはどのようなものがあるかを調べておくこと。 復習：データを取る意味、基礎統計量の意味と計算方法、データの種類、データの「ばらつき」について理解すること。現在、職場で収集しているデータの種類について確認すること。
	授業内容	【統計学の基礎】 統計学とはどのような学問か、記述統計学と推測統計学、データに存在する「ばらつき」、基礎統計量など，統計学の基礎について学ぶ。
	授業課題	有
	課題フィードバック方法	解答例示のみ

第3回	**予習・復習時間**	4時間
	予習・復習内容	予習：母集団とサンプルデータ（資料データ）とは何かについて調べておくこと。 復習：サンプルデータから母集団の平均値（母平均）や分散（母分散）の推測の仕方を理解すること。
	授業内容	【確率の考え方】 サンプルデータ（資料データ）による母集団の母平均や母分散の推測、母集団に関する正しい情報獲得のためには、確率の概念が必要であることについて学ぶ。
	授業課題	無
第4回	**予習・復習時間**	4時間
	予習・復習内容	予習：検定と推定の違いを調べておくこと。 復習：検定および推定の仕組みとその進め方について理解しておくこと。
	授業内容	【検定と推定の考え方】 統計量を用いた検定(片側検定、両側検定)と推定(点推定、区間推定)の基本を学ぶ。
	授業課題	有
	課題フィードバック方法	解答例示のみ
第5回	**予習・復習時間**	4時間
	予習・復習内容	予習：品質管理における統計の役割、とりわけ管理図について調べておくこと。 復習：授業で学んだ管理図の内容とその作成手順を理解すること。
	授業内容	【品質管理と統計、管理図】 品質管理における統計の役割り、とりわけ管理図について解説する。さらに管理図の作成を体験する。
	授業課題	無

第6回	**予習・復習時間**	4時間
	予習・復習内容	予習：管理図が示す内容の意味を理解した上で、管理図の点の動きから工程の変動が掴めることを調べておくこと。 復習：管理図の読み方、使い方をマスターすること。
	授業内容	【管理図の読み方・使い方及び対処の仕方】 管理図の読み方・使い方及び問題が発見されたときの対処の仕方について学ぶ。
	授業課題	有
	課題フィードバック方法	解答例示のみ
第7回	**予習・復習時間**	4時間
	予習・復習内容	予習：散布図、相関分析、回帰分析について調べておくこと。 復習：散布図、相関分析、回帰分析の意味する内容をよく理解した上で、Excelを用いて解析できるようにすること。
	授業内容	【相関分析と回帰分析】 相関分析・回帰分析とは何か、散布図、単相関分析の理論と単相関係数、回帰分析を学び、それらの分析を Excelを用いて実践する方法について学ぶ。
	授業課題	有
	課題フィードバック方法	解答例示のみ

第8回	予習・復習時間	4時間
	予習・復習内容	予習：重回帰分析とは何かを調べておくこと。 復習：重回帰分析の内容とExcelを使った分析手順について理解しておくこと。
	授業内容	【重回帰分析の基礎と応用】 重相関分析の解析手順、重回帰式の推定、回帰式の有意性の検定などExcelの分析ツールを用いて演習を行い、解析結果の読み方についても議論する。
	授業課題	無
第9回	予習・復習時間	4時間
	予習・復習内容	予習：時系列データの具体例としてどのようなものがあり、どのような分析がされているかを調べておくこと。 復習：学んだ時系列データ分析の手法を理解し、Excelを用いて分析ができるようにすること
	授業内容	【時系列分析】 時間の経過に従って観測値が変わる時系列データに対して、過去データの傾向から今後の予測を行う分析手法の基本を学ぶ。
	授業課題	有
	課題フィードバック方法	解答例示のみ
第10回	予習・復習時間	4時間
	予習・復習内容	予習：演習テーマを読んでおくこと 復習：取り組んだ演習テーマを改めて自分で実践できることを確認しておくこと
	授業内容	【時系列分析実践演習】 前回に学んだ手法を活用し，具体的な時系列データに対する分析に取り組み，その分析結果をプレゼンテーションする演習を行う。
	授業課題	有
	課題フィードバック方法	全体へのフィードバック

第11回	**予習・復習時間**	4時間
	予習・復習内容	予習：提示するワークショップ課題説明書を読んでおくこと 復習：議論した結果の仮説が正しいかどうか改めて確認すること
	授業内容	【データ分析実践ワークショップ：テーマ選定と問題分析】 PPDACサイクルに基づいたデータ分析実践を行うワークショップである。提示するテーマの中から一つを選び（あるいは自らの問題で分析を実践したいテーマでもよい）、PPDACの手順で分析を試みる。原則、グループ実習とする予定である。この回はテーマ選定、グループ分け，問題分析（仮説設定）までを行う。
	授業課題	有
	課題フィードバック方法	個別フィードバック
第12回	**予習・復習時間**	4時間
	予習・復習内容	予習：仮説検証のために必要なデータとその入手先を検討しておくこと 復習：データ収集と分析法が現実的なものであるかを確認しておくこと
	授業内容	【データ分析実践ワークショップ：データ収集と分析法の検討】 前回に定めた仮説の検証に必要なデータの特定とその入手方法、そして分析手順について調査や検討を行う。
	授業課題	有
	課題フィードバック方法	個別フィードバック

第13回	**予習・復習時間**	4時間
	予習・復習内容	予習：データの分析手順や分担を検討しておくこと 復習：分析結果や結論が妥当なものであるかを確認しておくこと
	授業内容	【データ分析実践ワークショップ：データ分析】 収集したデータに対して集計・分析作業を行う。その分析結果の解釈に関する見直しや議論を行い、現時点で分かったこと，そして不明な点を整理する。さらに追加のデータ収集や分析が必要ないかの検討を行う。
	授業課題	有
	課題フィードバック方法	個別フィードバック
第14回	**予習・復習時間**	4時間
	予習・復習内容	予習：データの分析手順や分担を検討しておくこと 復習：分析結果や結論が妥当なものであるかを確認しておくこと
	授業内容	【データ分析実践ワークショップ：データ分析と結論】 前回に続いて収集したデータに対して集計・分析作業を行う。その分析結果の解釈に関する議論を行い、結論をまとめる。PPDACの一連の流れをプレゼンテーション資料としてまとめることに着手する。
	授業課題	無

第15回	**予習・復習時間**	4時間
	予習・復習内容	予習：問題分析から結論までの一連の結果をスライドでまとめ、プレゼンテーションの準備をすること 復習：ワークショップの振り返りを行い，自らの業務あるいは職場でどのように役立つかを検討し，考えをまとめておくこと
	授業内容	【データ分析実践ワークショップ：プレゼンテーション】 PPDACの一連の流れをプレゼンテーション資料としてまとめ、プレゼンテーションを行う。その内容に関し、質疑応答および議論を行う。 【最終課題】 この科目で学修したことやワークショップの実践内容が自身の業務あるいは職場の課題の解決にどのように役立つかを検討し，レポートとしてまとめること
	授業課題	有
	課題フィードバック方法	個別フィードバック

授業情報	
授業方法	本科目は担当教員による講義が中心となるが、理解を深めるための演習（Excelを使った分析演習）を各回で適宜行う。11回以降はグループワークによるワークショップを行う予定である。ただし，10回までの進み具合によって一部変更する場合もある。
テキスト	特に指定しない。毎回配布する資料を中心に講義・演習を行う。参考図書も記載するが、ここにあげられているものに限らず，必要に応じて、適切な書籍を紹介する予定。

参考図書	1. 工藤卓哉,保科学世 ,「データサイエンス超入門 ビジネスで役立つ「統計学」の本当の活かし方」,日経BP(2013) ISBN-13 : 978-4822249786 2. 大村 平,「改訂版　統計のはなし ―基礎・応用・娯楽―」,日科技連出版社(2002). ISBN-13: 978-4817180100 3. 羽山 博,「やさしく学ぶ データ分析に必要な統計の教科書」,インプレス(2018). ISBN-13 : 978-4295003656 4. 小島 寛之,「完全独習 統計学入門」,ダイヤモンド社 (2006). ISBN-13 : 978-4478820094 5. 東京大学教養学部統計学教室編,「基礎統計学Ⅰ 統計学入門」,東京大学出版会(1991). ISBN-13: 978-4130420655 6. 宮川 公男, 「基本統計学 第4版」,有斐閣 (2015). ISBN-13 : 978-4641164550 7. 鐵 健司,「新版　品質管理のための統計的方法入門」,日科技連出版社(2000). ISBN-13: 978-4817103420 8. 栗原 伸一,「入門統計学ー検定から多変量解析・実験計画法までー」.オーム社 (2011). ISBN-13 : 978-4274068553 9. 上田太一郎 他, 「Excelで学ぶ時系列分析 理論と事例による予測 Excel2016/2013対応版」, オーム社(2016).ISBN-13 : 978-4274219177

評価方法		
評価の視点	**評価 ウェイト**	**備考**
課題の取り組み	80%	演習課題での解答とワークショップの取り組み、最終課題レポートを評価対象とする。
授業への貢献度	20%	授業中における積極的な質問や私からの問い掛けに答えて戴いた内容を評価する。
合計	100%	

受講生へ （授業科目のアピールポイント、必要な基礎となる科目の履修や知識・スキル）
職場で継続的に取り続けているデータには、どのようなものがあるか、またそれらのデータは何のために使用しているのかを整理しておくこと。Excelのデータ分析ツールが使えることが望ましいので、未経験者は体験しておくことを勧める。データ解析に関して職場で困っていることがあれば、積極的に提示して戴き、受講生の皆さんと討論できる機会を持ちたいと考えている。受講生の皆さんの積極的な授業参加を期待する。

その他
【受講生へのお願い】 疑問点があれば遠慮せず質問して欲しい。授業およびワークショップでの積極的な取り組みに期待する。なお、本科目の受講生の理解度を考慮し、予定している授業計画を変更することもある。