

科目名	データサイエンス演習 a (国)				授業コード	D001030a			
科目ナンバリング	授業形態	履修形態	単位数	年次	開講期				
AA130bJ	演習	必修	2単位	1年	後期				
担当者名	阿部 真育								
学修目標との対応									
DP1			DP2			DP3			
1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3	
		60	20	20					
授業概要									
<p>本演習では、データを取り扱うための基礎的な知識・技能を習得することを目的とする。医学・保健科学だけでなく人文・社会科学においてもデータが持つ性質を理解し、適切な方法でデータの分析を行う視点は重要であるため、データ分析を行なうための基礎的な数理・データサイエンスの技能を身に付けることを目指す。実際にさまざまな領域でデータ分析が行なわれていることを知るために、人文・社会科学や保健科学、公衆衛生、環境科学などの具体的な統計データを使用し、各領域に共通するデータの処理方法、検定方法、分析方法を学習する。また、領域横断的な課題に対して、効果的なデータ分析手法を判断し応用する力を身に付けるために、多様かつ新たなデータの取り扱いについて実技指導を含めた演習を行う。</p>									
到達目標				成績評価の方法と基準					
データの種類に応じた分析手法を選択することができ、誤った分析を適切に批判する判断力を持ち、同時に自らの分析手法が適切であることを証明し、社会の諸課題を解決する構想力がある。				授業ごとに学習した内容の振り返りとなる問題を課す。課題の提出回数および理解度に応じて評価する。					
学習目標		評価項目と割合							
具体的学習目標		配点比率	授業態度	小テスト	自主学習態度	レポート	プレゼンテーション	学期末筆記試験	その他
データの集計を行なうための計算機を操作することができる。		25				25			
データの集計・分析を実践できる。		25				25			
データの正当性を示すための検定ができる。		25				25			
新たなデータを分析することができる。		25				25			
授業の項目と内容・方法				事前事後学習の内容と目安時間					
第1回	オープンデータの調査と取り扱い								
具体的なOpenDataの取り扱いと、ファイルの形式、利用可能なアプリケーションについて演習する。				授業で演習した課題を完成させる。概ね4時間以内に終了させること。					
第2回	基本統計量の取り扱い方の学習								
中央値などの四分位数、最頻値、尖度、歪度、標準誤差などの概念を用いてデータを整理する。				授業で演習した課題を完成させる。概ね4時間以内に終了させること。					
第3回	データの性質に応じたグラフの作成								
与えられたデータを可視化するために適切かつ効率的なグラフを作成する。				授業で演習した課題を完成させる。概ね4時間以内に終了させること。					
第4回	ヒストグラムや代表的な確率分布の理解								
確率分布について理解し、データを確率的に解釈する。				授業で演習した課題を完成させる。概ね4時間以内に終了させること。					
第5回	データ間の関係の理解								
データが持つ関係について、相関関係・因果関係・疑似相関などを区別しデータの取舍選択を行なう。				授業で演習した課題を完成させる。概ね4時間以内に終了させること。					
第6回	クロス集計の考え方の理解								
複数のデータ系列を関連させて集計する。				授業で演習した課題を完成させる。概ね4時間以内に終了させること。					
第7回	検定の学習(1): χ^2 乗検定								
統計的検定について理解し、与えられたデータの独立性の検定を行なう。				授業で演習した課題を完成させる。概ね4時間以内に終了させること。					
第8回	検定の学習(2): t検定								
与えられたデータの2群の平均値に差があるかを検定する。				授業で演習した課題を完成させる。概ね4時間以内に終了させること。					
第9回	検定の学習(3): 分散分析と多重比較								
与えられたデータの3群以上の平均値に差があるかを検定する。				授業で演習した課題を完成させる。概ね4時間以内に終了させること。					

第10回	回帰分析の考え方 (1) : 単回帰	
独立変数と説明変数の関係を定式化する方法について演習する。		授業で演習した課題を完成させる。概ね4時間以内に終了させること。
第11回	回帰分析の考え方 (2) : 重回帰	
独立変数と複数の説明変数の関係を定式化する方法について演習する。		授業で演習した課題を完成させる。概ね4時間以内に終了させること。
第12回	テキスト分析の考え方の理解	
与えられたデータについて、KH-coderを使用したテキスト分析を行ない、形態素解析を行なう。		授業で演習した課題を完成させる。概ね4時間以内に終了させること。
第13回	多次元データの取り扱い方の理解	
与えられたテキストデータに距離尺度を導入してクラスター分析を行なう。		授業で演習した課題を完成させる。概ね4時間以内に終了させること。
第14回	線形代数の基礎	
ニューラルネットワークの理解のため、表計算ソフトを用いたベクトル演算を行なう。		授業で演習した課題を完成させる。概ね4時間以内に終了させること。
第15回	機械学習の基礎	
表計算ソフトを用いた教師あり学習の演算を演習する。		授業で演習した課題を完成させる。概ね4時間以内に終了させること。
テキスト、副読本、教材		
受講生へのメッセージ		
履修条件及び備考 (レポート評価基準・その他の具体的評価内容基準等)		YFL科目