

科目番号	教室	登録人数	履修登録方法
B学共119	[火6]文講215	60	抽選対象
開講年度	期間	曜日時限	開講学部等
2025	後学期	火6	国際地域創造学部国際地域創造学科
講義コード	科目名[英文名]	単位数	
g00584002	データサイエンス初級	2	
担当教員[ローマ字表記]			
山田 健太 [Kenta Yamada]			

#### 授業の形態

講義、演習又は実験

#### アクティブラーニング

学生が自身の考えを発表する、学生が文献や資料を調べる

#### 授業内容と方法

##### 【授業内容】

コンビニの売り上げ、電車の乗降記録、インターネット上の書き込みなど身の回りの様々なものやサービスが電子化され、コンピュータには日々膨大なデータが蓄積される時代になりました。そして、多くの企業がこれらのデータを活用した事業に取り組もうと始めていますが、膨大なデータの解析は、まだ歴史も浅くExcelなど表計算ソフトでは解析できない場合もあります。このような背景から、膨大なデータを解析でき、そこから有用な知見を引き出せる人材(データサイエンティスト)が社会から強く求められていますが、まだ少ないので現状です。そこで、Excelを用いた基本的なデータ分析から始め、大規模なデータも処理可能であり多くのデータサイエンティストにも利用されているプログラミング言語「Python」を用いて、できる限り身の回りのデータの解析を行い、社会のニーズに応えられる技術力を身につけられるよう授業を行います。

##### 【授業の方法】

最初に統計学の知識などを説明した後、実データを「Excel」「Python」を用いて自らの手で解析、グラフ化を行い簡単なレポートにまとめる。データ解析プロジェクトでは、自ら解析するデータを選び、これまでの知識をフル活用したデータ解析を行い、得られた結果の発表を行う。

#### URGCC学習教育目標

情報リテラシー、問題解決力

#### 達成目標

- ・「Excel」や「Python」を用いたデータ解析ができるようになる。[情報リテラシー]
- ・自分の興味のある現象や課題に対して、データサイエンスの基礎的な知識を用いることで、現象の理解や課題の解決を図れるようになる。[問題解決力]
- ・得られた知識を効果的に他人に説明できるようになる。

#### 評価基準と評価方法

毎回授業内で作成する小レポート 50%

データ分析プロジェクト発表会の内容 25%

データ分析プロジェクトを基にした期末レポート 25%

#### 履修条件

ノートパソコンを持参すること

#### 授業計画

1. ガイダンス：データサイエンスとは何か？なぜデータサイエンスを学ぶのか？
2. データ分析のための環境構築：「Excel」「Python」「MeCab」などのインストール, Pythonの基本的操

3. みんなの好きな数を調べてみよう 1 (Excelによる単変量解析) : 平均値, 標準偏差, 中央値, 四分位, 標準化(z値), 箱ひげ図
4. みんなの好きな数を調べてみよう 2 (Pythonによる単変量解析) : 平均値, 標準偏差, 中央値, 四分位, 標準化(z値), 箱ひげ図
5. みんなの好きな数字のヒストグラムを描いてみよう 1 (Excelによる分布の描画) : 確率関数, 確率密度関数, 累積分布関数, 期待値
6. みんなの好きな数字のヒストグラムを描いてみよう 2 (Pythonによる分布の描画) : 確率関数, 確率密度関数, 累積分布関数, 期待値
7. サッカーの得点分布を描いてみよう (Excelによる 2 項分布とポアソン分布の描画) : 二項分布, ポアソン分布
8. サッカーの得点分布を描いてみよう (Pythonによる 2 項分布とポアソン分布の描画) : 二項分布, ポアソン分布
9. 身長の分布を描いてみよう (ExcelとPythonによる正規分布の描画) : 正規分布
10. 小説(走れメロス)を解析してみよう (Pythonによるテキストマイニング) : 形態素解析, ベキ分布, ジップの法則
11. テストの点数や野球のスコアの関係を調べよう (Excel, Pythonによる相関解析) : ピアソンの積率相関係数, スピアマンの順位相関係数
12. 築年数などの各種条件を用いて家賃モデルを作成しよう (Excel, Pythonによる回帰分析)
13. まとめ, データ解析プロジェクト概要説明
14. データ解析プロジェクト準備
15. データ解析プロジェクト発表会/確認問題

#### 事前学習

事前に配布資料に目を通しどのような内容を学習するのかを確認する。

#### 事後学習

授業を行った内容を確認し, 「python」のコマンドをもう一度実行する。

#### 教科書にかかわる情報

#### 教科書全体備考

#### 参考書にかかわる情報

参考書	書名	Pythonによるデータ分析入門 : NumPy、pandasを使ったデータ処理			ISBN	487311845X	備考			
	著者名	Wes McKinney著 ; 濑戸山雅人, 小林儀匡, 滝口開 資訳								
	出版社	オライリー・ジャパン	出版年	2018	NCID					
参考書	書名	ソーシャルメディアの経済物理学 : ウェブから読み解く人間行動			ISBN	4535556784	備考			
	著者名	高安美佐子編著								
	出版社	日本評論社	出版年	2012	NCID					

#### 参考書全体備考

#### 使用言語

日本語

#### メッセージ

本講義は対面とオンデマンドのハイブリッドで行う予定です。

#### オフィスアワー

木曜10:30～11:30 文系研究棟 521

事前に下記メールアドレスへ連絡しアポイントメントをとるようにしてください

 メールアドレス

この項目は教務情報システムにログイン後、表示されます。

 URL

データサイエンティスト養成履修カリキュラム : <https://mds.skr.u-ryukyu.ac.jp/mds/program/>