

RとQuartoではじめるデータサイエンス：データを可視化する《2024》

Home

荻谷 千尋

12, Jun, 2024

Note

毎回、PCと教科書は持参してください

Warning

このウェブサイトは2024年9月30日に削除します

I. シラバス

- シラバス

II. 講義スケジュール

Note

授業日当日までに、各回のレジメにアクセスできるようにします

Important

欠席者はご自身で授業内容、授業課題をフォローしてください（欠席連絡は不要です）

Week 1 (6/12) : イントロダクション

- キーワード :
 - 図とデータ
 - Rのインストールとプロジェクトの作成
- 『データ分析のためのデータ可視化入門』 :
 - まえがき
 - 第1章 データを見る

Week 2 (6/19) : Rの基本的な操作方法 (1)

- キーワード :
 - 整然データ (tidy data) と雑然データ (messy data)
 - データの型
 - 代理演算子
 - YAML Part I: 基本
- 『データ分析のためのデータ可視化入門』 :
 - 第2章 さあ、始めよう !
 - 特に2.3、2.4
 - 付録A.3.2 プロジェクトを組織化する
 - 付録A.1.3 tidyデータ

Week 3 (6/26) : Visualization (1)

Important

受講生の理解度を考慮して、Week3とWeek4を入れ替えて授業をおこないます。教科書どおりの順序（簡単にggplotにさわりの、次いでデータの加工、その後、ggplotに本格的に触れる）になります

- キーワード :
 - データの読み込み
 - ggplot2

- マッピングとレイヤー
- 『データ分析のためのデータ可視化入門』 :
 - 第2章 さあ、始めよう！
 - 2.5、2.6
 - 第3章 プロットを作る

Week 4 (7/3) : Rの基本的な操作方法 (2)

- キーワード :
 - データの加工
 - データの抽出
- 『データ分析のためのデータ可視化入門』 :
 - 第4章 正しい数値の示し方
 - 第5章 データの整形・ラベル・メモの追加
 - 特に5.1と5.2
 - 付録A.2 データ読み込み時の共通問題

Week 5 (7/10) : Visualization (2)

- キーワード :
 - データの結合
 - 欠損値
 - 5 Named Graphs(5NG) :
 - 棒グラフ；ヒストグラム；箱ひげ図；散布図；折れ線グラフ
 - YAML Part II: 多様な出力

Warning

7月17日は月曜授業日のため授業はありません

Week 6 (7/24) : Visualization (3)

- キーワード :
 - プロットを整える
 - 色とラベル
 - 横持ちデータと縦持ちデータ
 - チャンクオブション
- 『データ分析のためのデータ可視化入門』 :
 - 第5章 データの整形・ラベル・メモの追加
 - 第8章 プロットを整える

Week 7 (7/31) : Visualization (4)

- キーワード :
 - 地図；自作関数
- 『データ分析のためのデータ可視化入門』 :
 - 第7章 地図を描画する
 - A.2.3 コードの重複を減らすために関数を書く
 - A.2.1 データ形式

Week 8 (8/7) : プレゼンテーション

III. 提出物（提出先）

Important

提出物（アンケート類を含む）は、すべて評価対象物です

Important

「授業の感想」を除き、欠席を理由とする、授業課題の未提出は認めません（遅延提出は認めます）

- 授業の感想：Google Forms
- 演習：Google Forms

IV. 講評

- 成績評価後に掲載します（8月末を予定）

V. 役立つサイト

- 役立つサイト

VI. QRコード

- QRコード

VII. 追加連絡

- 実習データについて

VIII. 最終成果物サンプルページ

- 金沢広域急病センター利用者数
 - データ入手先：URL
- 野生鳥獣の放射線モニタリング調査結果
 - データ入手先：URL
- 初回アンケート集計