

RとQuartoではじめるデータサイエンス《2026》

#1 イントロダクション

苅谷千尋

Wed, 8, Apr, 2026

目次

1. 本日の目標
2. 自己紹介
3. 学修目標と成績評価
4. データサイエンスの第一歩
5. なぜRなのか / なぜQuartoなのか
6. R環境の整備
7. Rにふれてみよう：palmer penguins
8. Rを使ってデータ分析を深める

0. 本日の目標

1. 授業趣旨と成績評価方法を**しっかりと**理解する
2. 表とデータセットの違いを**なんとなく**理解する
3. Rの便利さ、凄さを**なんとなく**理解する
4. Rを使える環境を各自のPC/macに整備する
5. Rで散布図を作る

1. 自己紹介

1. 名前、学類
2. 受講動機
3. データ分析したい対象（の有無）
4. 高校「情報」科目で習ったこと（該当者のみ）
5. プログラミングの経験の有無

2. 学修目標と成績評価

(1) 学修目標（到達目標）

1. 受講生が、R言語の基本的な操作（データの読み込み、加工、抽出、結合など）ができる
2. 受講生が、ggplot2パッケージを使い、基本的な図（棒グラフ；ヒストグラム；箱ひげ図；散布図；折れ線グラフ）を作成できる
3. 受講生が、自分の関心のあるテーマにかかわるデータから適切な図を作成できる

(2) 成績評価

- 授業別ルーブリック

(3) スケジュール

- 4月30日（木）：4月29日（水・祝）の振替授業日

(4) 問い合わせ先

- メール：kariyach@staff.kanazawa-u.ac.jp
- 研究室：総合教育2号館 670号室 (**Advance Appointment Required**)

3. データサイエンスの第一歩

(1) ツールに依拠した学びが推奨される傾向にあり

- データサイエンス=Pythonを学ぶこと、生成AIを使えるようになること、という誤解
 - **もっとも重要なのはデータの形**についての思想・哲学=データサイエンスの出発点

(2) データの形

- 表とデータセットを区別する（日本はエンジニアを除いてできていない）
 - 表：人間が目で見えて理解できる
 - データセット：コンピュータが理解できる
 - よいデータ構造は、分析を簡単にし、エラーを減らし、再現性を高める
- tidy（整然な）データ

4. なぜRなのか / なぜQuartoなのか

1. コード（スクリプト）
2. R言語の特性
3. 豊富な可視化ツール
4. オープンソースと親切なコミュニティ
5. 簡易的な記述法（Quarto）
6. レポート機能と多様な出力形態（Quarto）
7. 無料

5. R環境の整備

- 視認性と可読性
- Setup Chunk

6. Rにふれてみよう：palmer penguins

- データセット penguins
- view()関数；head()関数；colnames()関数；distinct()関数
- パイプ演算子（'|>'）
- 散布図：ggplot; geom_point

宿題

- 授業の感想：
 - 回答先：Google Forms
 - 締め切り：4月10日（金）23時59分
- 初回授業アンケート：
 - 回答先：Google Forms
 - 締め切り：4月10日（金）23時59分
- ※演習の宿題はありません