

# テキストマイニングの実習 ー 準備編 ー

2019/6/27

ビジネス科学研究科  
経営システム科学専攻

# スケジュール

- 1日目: 7/4
  - 説明 – データ分析の手順
  - 演習 – データの理解 (Excel)
- 2日目: 7/11
  - 説明 – テキストマイニング ツールの使い方 (KHCoder)
  - 練習 – テキストマイニング ツールの使い方 (KHCoder)
- 3日目: 7/18
  - 演習 – データ分析の実践 (KHCoder)

# 講義で使用するサイト

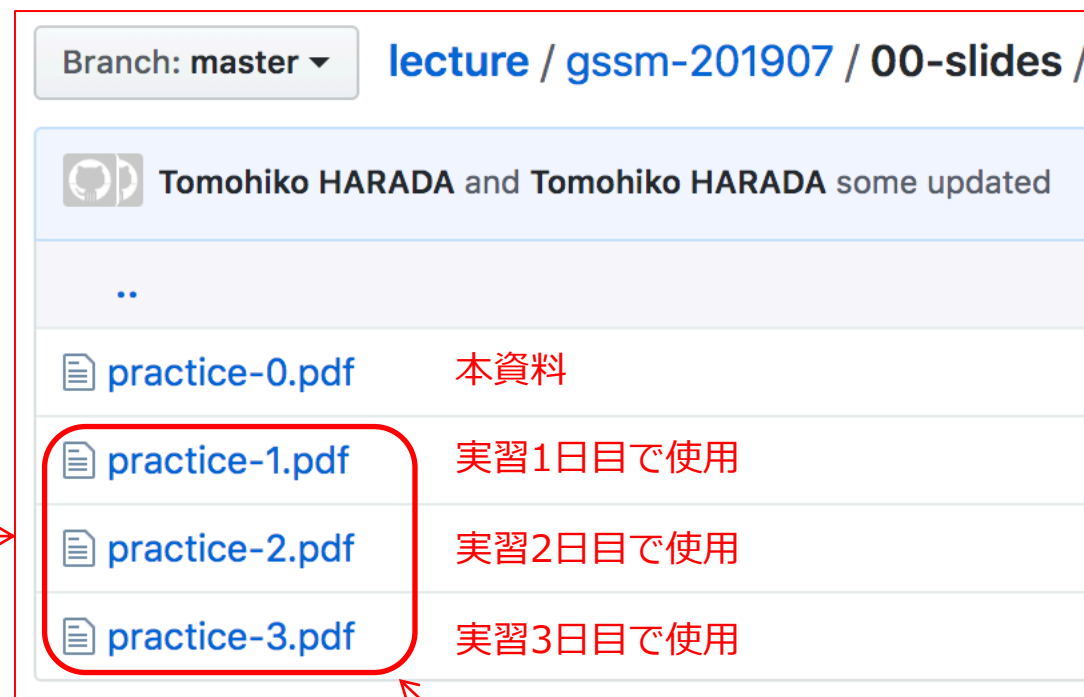
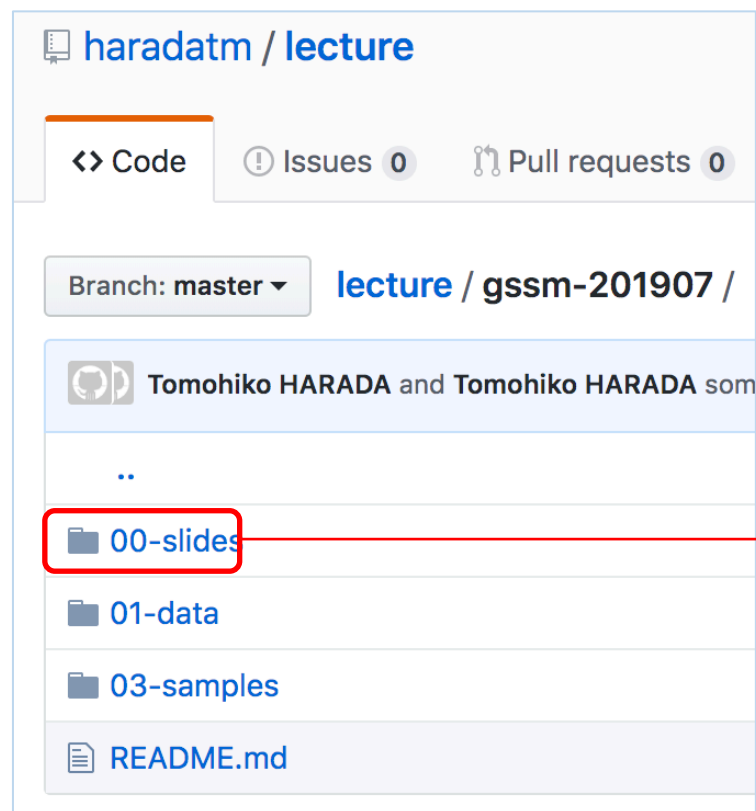
<https://github.com/haradatm/lecture/tree/master/gssm-201907>

# 演習環境について

- 演習では,648計算機室の Windows PC を使用します
  - PC は 10 台のため, 複数人に1台の割当てになります
- 個人のPCを使用しても構いません
  - ただし, Windows OS (7, 8.1, 10) を搭載した PC が必要です
  - 講義前に,以下のツールが使用できることを確認しておいてください
- 演習では,以下のツールを使用します
  - 次回: Microsoft EXCEL (用途: データの加工や修正)
  - 次々回以降: KHCoder (用途: テキストマイニング) ※フリーソフト

# 講義スライド

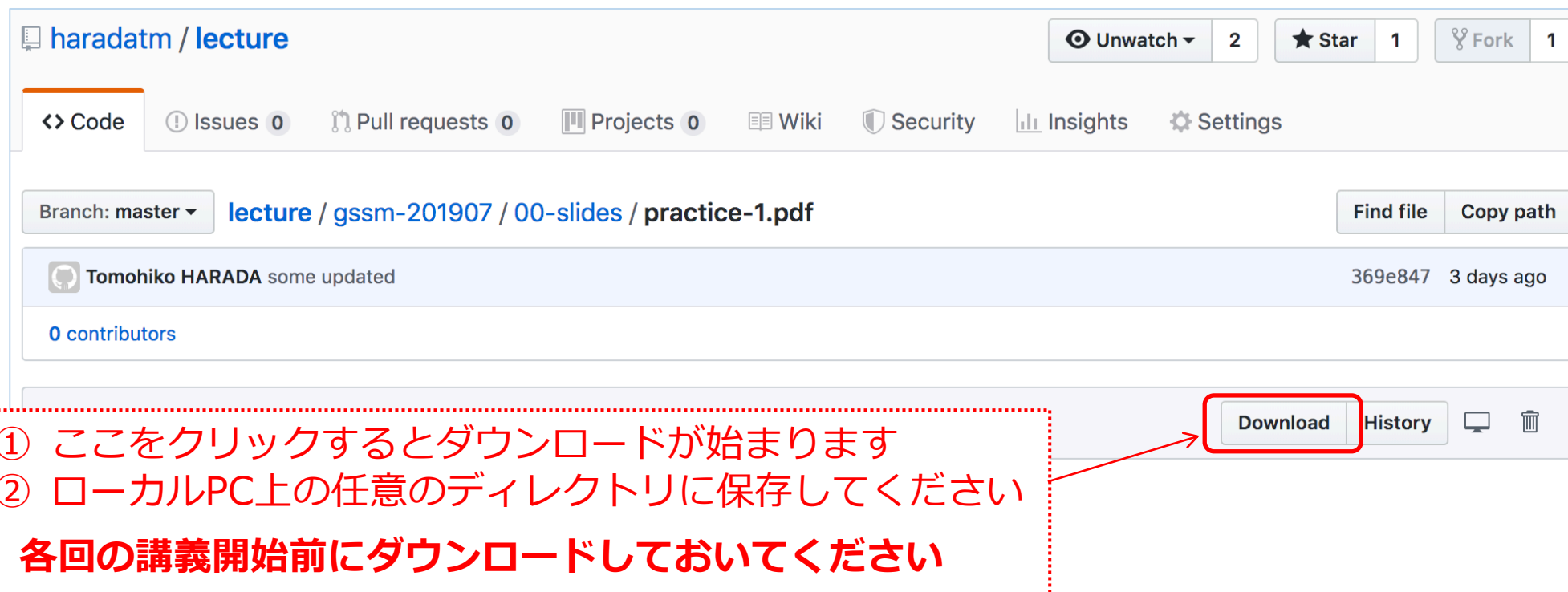
- <https://github.com/haradatm/lecture/tree/master/gssm-201907>



クリックするとページ遷移します

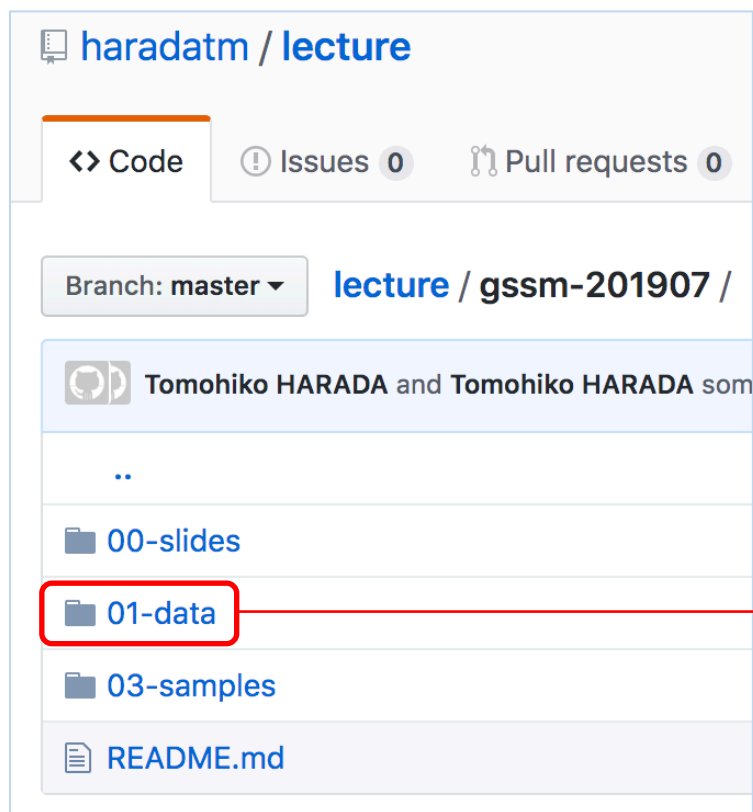
# ダウンロード方法

- Download ボタンをクリックするとダウンロードを開始



# 使用するデータ (次回の演習で使用)

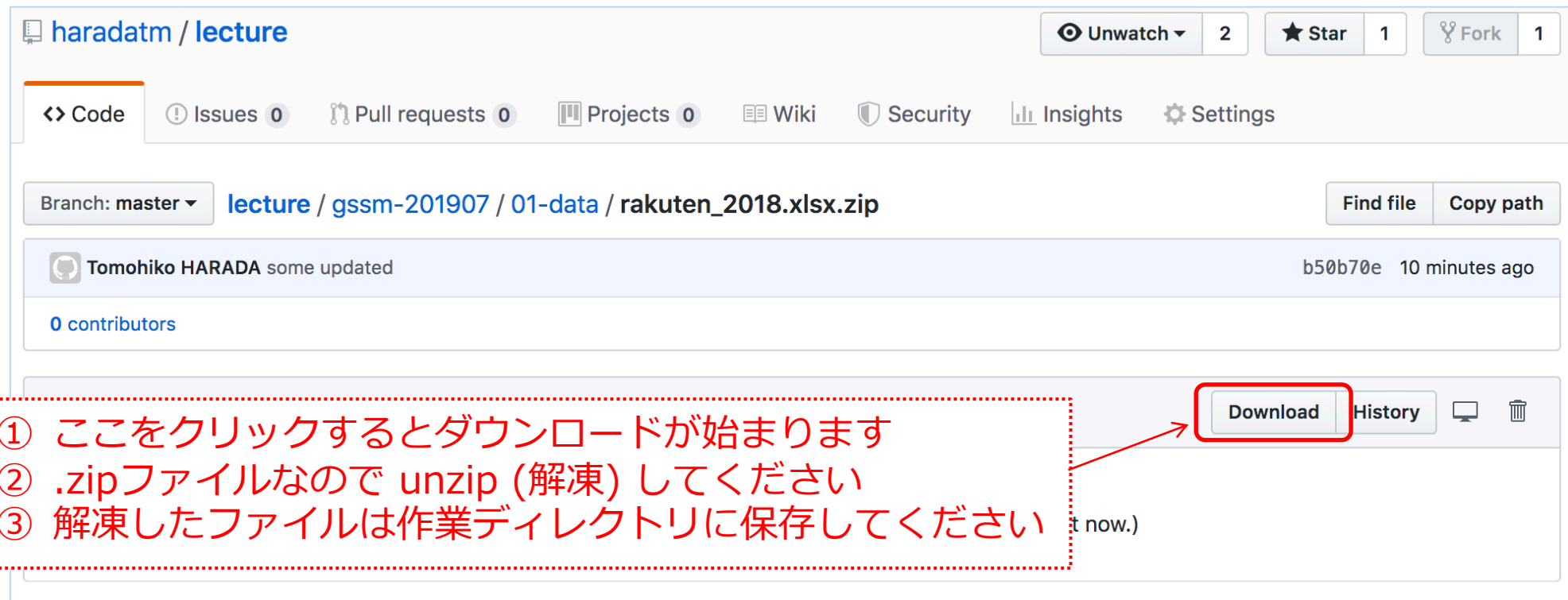
- <https://github.com/haradatm/lecture/tree/master/gssm-201907>



クリックするとページ遷移します

# ダウンロード方法 (次回の演習で使用)

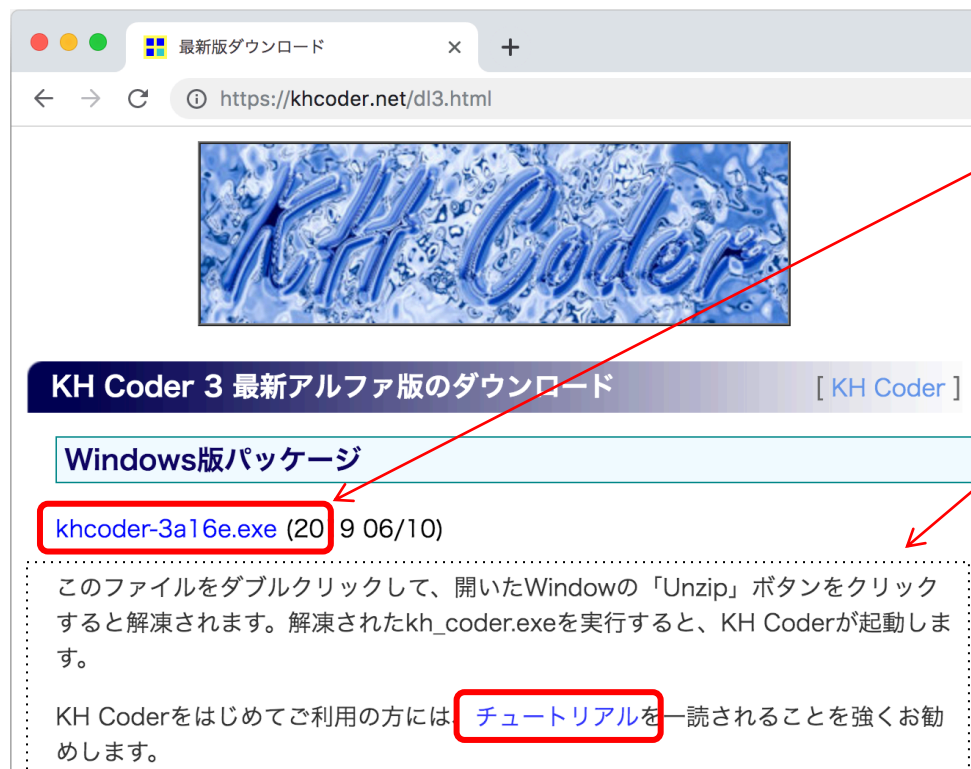
- Download ボタンをクリックするとダウンロードを開始





# KH Coder のインストール (次々回の演習で使用)

- ダウンロードとインストール <https://khcoder.net/dl3.html>

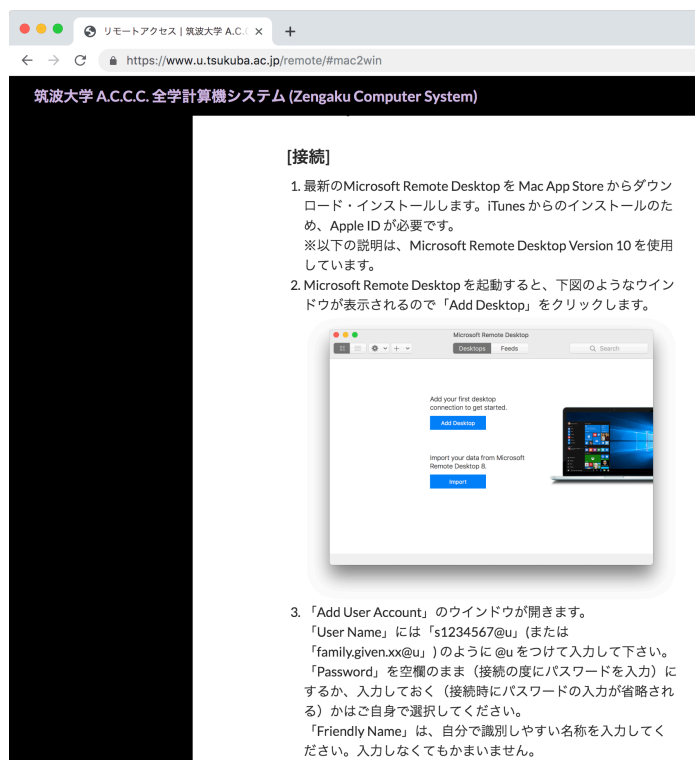


- ① ここをクリックすると遷移先のページからダウンロードが始まります
- ② 指示に従いインストール

自己解凍ファイルです。このファイルを実行（ダブルクリック）し、開いたWindowの「Unzip」ボタンをクリックすると、（特に変更しなければ）「C:¥khcoder3」というフォルダにすべてのファイルが解凍されます。解凍された kh\_coder.exe を実行すると、KH Coder が起動します。

# 参考 – Windows OSがない場合 (同時実行未検証)

- 全学計算機システムに Windows デスクトップがあります
  - <https://www.u.tsukuba.ac.jp/remote/#mac2win>



左記のページにある説明に従って,

- ① ツール (Microsoft Remote Desktop) のインストール
  - ② 全学計算機システム (Windows) へのログイン
- ができることを確認してください

## KH Coder インストール時の注意:

全学の Windows の場合は,ログイン後の**デスクトップ上**  
**に「khcoder」というフォルダを作成**して,その中に解凍  
してください

# 参考書

(KH Coder)

- [1] 樋口耕一. 社会調査のための計量テキスト分析 —内容分析の継承と発展を目指して—. ナカニシヤ出版, 2014.
- [2] 樋口耕一. テキスト型データの計量的分析 —2つのアプローチの峻別と統合—. 理論と方法, 数理社会学会, 2004, 19(1): 101-115.
- New** [3] **牛澤賢二. やってみよう テキストマイニング —自由回答アンケートの分析に挑戦!. 朝倉書店, 2019**

(Windows環境によるCGM収集の参考に)

- [4] テキストマイニングソフトを利用した新未来洞察手法の研究. 第10分科会, (財)市場創造研究会. [http://www.shijo-sozo.org/news/%E7%AC%AC10%E5%88%86%E7%A7%91%E4%BC%9A\\_1.pdf](http://www.shijo-sozo.org/news/%E7%AC%AC10%E5%88%86%E7%A7%91%E4%BC%9A_1.pdf)

# 参考書

## (R を使った参考書)

- [5] 金明哲. "テキストデータの統計科学入門." 岩波書店, 2009.
- [6] 石田基広. "RMeCab によるテキスト解析. R によるテキストマイニング入門." 森北出版, 2008, 51-82.

## (他のツールを使った参考書)

- [7] 那須川哲哉. "テキストマイニングを使う技術/作る技術: 基礎技術と適用事例から導く本質と活用法." 東京電機大学出版局, 2006.
- [8] 上田隆穂, 黒岩祥太, 戸谷圭子. "テキストマイニングによるマーケティング調査." 講談社, 2005.

## (統計解析を中心とした参考書)

- [9] 前田忠彦; 山崎誠. 言語研究のための統計入門. くろしお出版株式会社, 東京, 2013.