

# テキストマイニングの実習

## － 3日目 －

2016/7/20

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

- 7/6

- 説明 — データ分析の手順
- 演習 — データの理解 (Excel)

- 7/13

- 説明 — ツール (KHCoder)
- 練習 — ツール (KHCoder)

- 7/20

- 演習 — データの分析 (KHCoder)

# テキスト分析

- ユーザーの声の可視化と生の声(原文)をもとに,数値評価では見えないニーズや課題を発見する
  1. 各人でデータ分析を行う
  2. 周囲の4人前後でグループを作る
  3. グループ内で考察や支持する図,原文を共有し議論
  4. グループごと図と原文を使って考察内容を発表
- 課題は,下記の4つ
  - 個人-1: ユーザーがどの項目に注目しているか (特徴語)
  - 個人-2: ユーザーがどの項目に注目しているか (共起NW)
  - 個人-3: ユーザーが何をどう評価しているか (共起NW)
  - グループ: 分析結果にもとづいて改善案を提案してみる

※ 個人作図は**レジャー**か**ビジネス**のいずれかで構わないが,グループ内には両方あると望ましい

# 課題 一個人1 「特徴語の集計と比較」

- ユーザーがどの項目に注目しているかを確認する
  - 分類=**レジャー**と分類=**ビジネス**について,テキスト中の特徴語を集計し,エリアによって特徴語がどう異なるかを比較し,2つの**注目する項目の違い**を考察する
  - 分類=**レジャー**の5エリアについて,テキスト中の特徴語を集計し,エリアによって特徴語がどう異なるかを比較し,5エリアで**注目する項目の違い**を考察する
  - 分類=**ビジネス**の5エリアについて,テキスト中の特徴語を集計し,エリアによって特徴語がどう異なるかを比較し,5エリアで**注目する項目の違い**を考察する

## 課題 一個人2 「特徴語の共起NW」

- ユーザーがどの項目に注目しているかを確認する
  - 分類=**レジャー**と分類=**ビジネス**について,特徴語の共起ネットワーク図を作成し,エリアによって特徴語がどう異なるかを比較し,2つの注目する項目の違いを考察する
  - 分類=**レジャー**の5エリアについて,特徴語の共起ネットワーク図を作成し,エリアによって特徴語がどう異なるかを比較し,5エリアで注目する項目の違いを考察する
  - 分類=**ビジネス**の5エリアについて,特徴語の共起ネットワーク図を作成し,エリアによって特徴語がどう異なるかを比較し,5エリアで注目する項目の違いを考察する

# 課題 一個人3「ポジ/ネガの共起NW」

- ユーザーが何をどう評価しているかを確認する
  - 分類=**レジャー**の5エリアについて,特徴語とポジティブ意見の共起NW図を作成して比較し,5エリアの何がどう評価されている点を考察する
  - 分類=**レジャー**の5エリアについて,特徴語とネガティブ意見の共起NW図を作成して比較し,5エリアの課題を考察する
  - 分類=**ビジネス**の5エリアについて,特徴語とポジティブ意見の共起NW図を作成して比較し,5エリアの何がどう評価されている点を考察する
  - 分類=**ビジネス**の5エリアについて,特徴語とネガティブ意見の共起NW図を作成して比較し,5エリアの課題を考察する

# 課題 グループ

- あなたは観光地域振興担当者です. 地域の改善案を提案してください
  - レジャーエリアのうち,数値評価の総合点およびポジティブ/ネガティブ」の両方の意見から対照的な2エリアを選択し,担当エリアと比較先エリアとする
  - ビジネスエリアのうち,数値評価の総合点およびポジティブ/ネガティブ」の両方の意見から対照的な2エリアを選択し,担当エリアと比較先エリアとする
  - 担当エリアについて,ポジティブ/ネガティブの両方の意見から,比較先エリアと比較し,改善すべき点を考察する

注意: 主張を支持する ☒ とユーザーの生の声(原文)の両方を使って説明してください

# 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, 京都, 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.

## (Windows環境によるCGM収集の参考に)

- [3] 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)

- [4] 金明哲. "テキストデータの統計科学入門." 岩波書店, 2009.
- [5] 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)

- [6] 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.
- [7] 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)

- [8] 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.