

「KH Coder」を利用したテキストマイニングの授業実践

愛媛県立新居浜東高等学校

本田 知仁

1 主題設定の理由

令和4年度より年次進行で実施される新学習指導要領において、共通教科「情報」では、現在の「社会と情報」「情報の科学」の選択から、必修科目の「情報Ⅰ」と選択科目の「情報Ⅱ」に再編される。「情報Ⅰ」では、「プログラミング」「情報デザイン」「データ分析」などを活用し、情報技術を適切かつ効果的に活用する力を育むことが望まれている。昨年度、データ分析を授業に取り入れる方法を探る第一歩として、立命館大学産業社会学部教授 樋口耕一氏の開発したテキストマイニングソフト「KH Coder」の使用方法について研究を行った。今年度は、「KH Coder」の授業での活用を研究する。

2 実践内容

- (1) 教員が用意したデータを利用して、個々の生徒にテキストマイニングについての説明や「KH Coder」の使用方法を教える。
- (2) 全クラスの生徒から「新居浜市について」のアンケートを取り、その結果をテキストマイニングさせる。活動は複数人のグループで行う。テキストマイニングの結果から「新居浜市の今後」について考えさせる。
- (3) 授業後にテキストマイニングについてのアンケートを行う。
- (4) 考察させた内容をパワーポイントでまとめ、発表を行う。

3 結果・考察

- (1) 「KH Coder」にチュートリアルとして付属している「こころ」を使用し、「KH Coder」の解説を行った。基本的な使用方法をまとめてくれているので、その中から生徒が使いやすい機能に絞って説明を行った。小説とアンケートとでは、解析の方法や考え方が少し違うところもあり、教えにくい部分もあった。
- (2) テキストマイニングは、分析ツールなので、マイニングの目的を明確にしておかなければならない。今回は、生徒が通っている新居浜市について課題と改善点を考えさせることを目的とした。そして、アンケート項目は「新居浜市のイメージ」「新居浜市の良いところ」「新居浜市の悪いところ」の3つに加え「出身市」「性別」を設定した。KHCoderでは、複数の項目を一度にテキストマイニングすることも可能だが、初めて行う生徒には敷居が高いため、個別にマイニングができるよう項目を3つに絞った。考察・結果については、後日のプレゼンテーションスライド作成に使用するため、プリントに記録させた。(図1)

生徒にとっては、初めて行うテキストマイニングであったが、「KH Coder」にある使いやすい機能のみを使用したため、手間取ることなく作業ができていた。しかし、アンケートを1年生約240人のみで行ったため、結果が偏ってしまった。アンケートからテキストマイニングを行うためには、幅広い条件(年齢など)からもっと大量の

データを集めなければならないと感じた。「出身市」についても他市町村から来ている生徒が少なくデータとして意味が無かった。男女別の結果は、男子は「祭り」女子は「ショッピング」というようにその差が顕著に見られて面白い結果となった。生徒が考えやすくなるテーマ選びをする必要があるだろう。

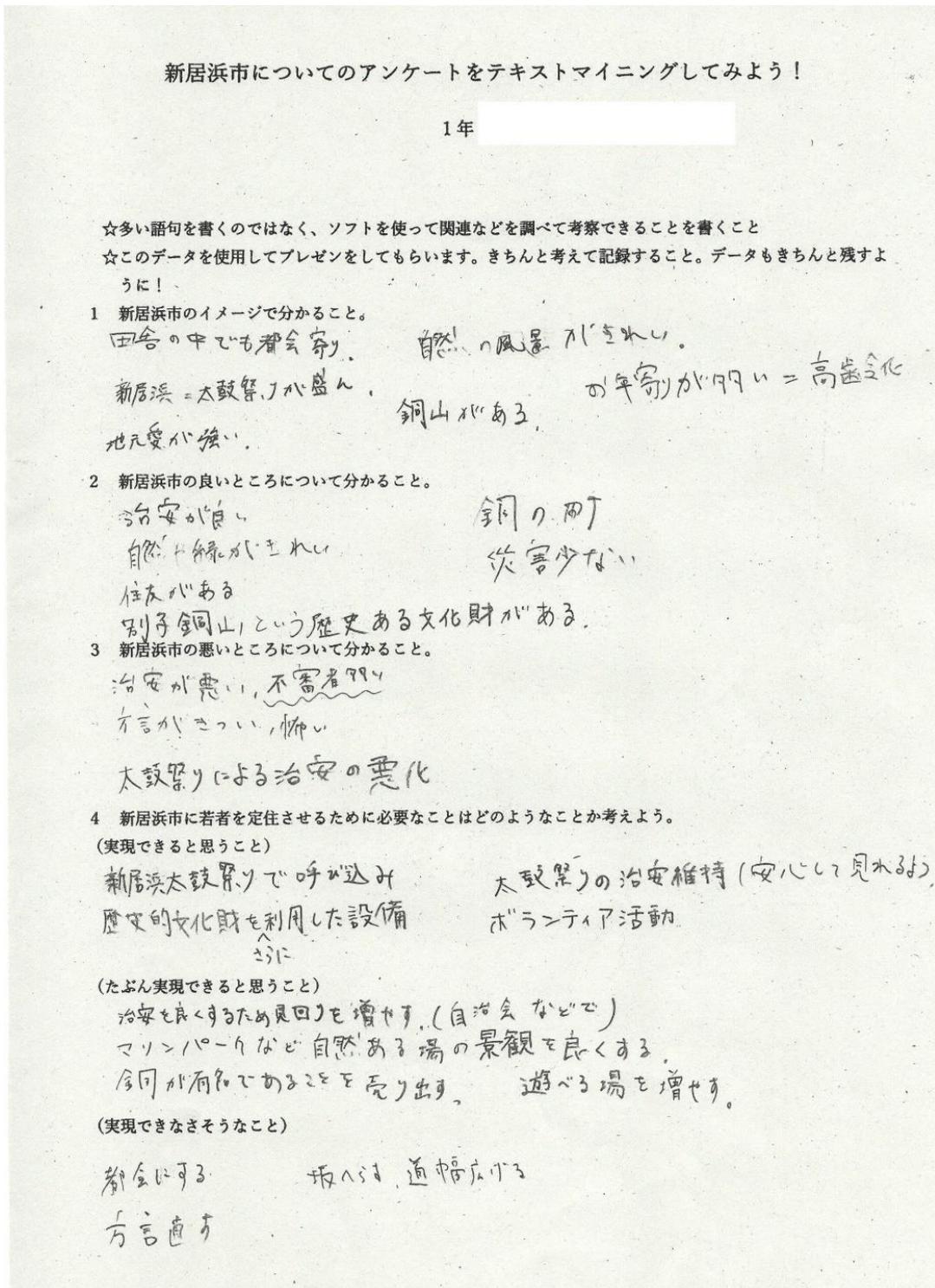


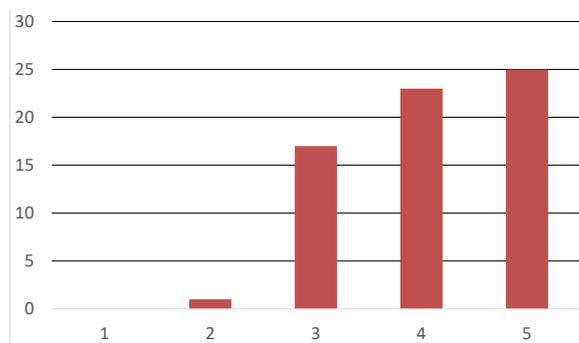
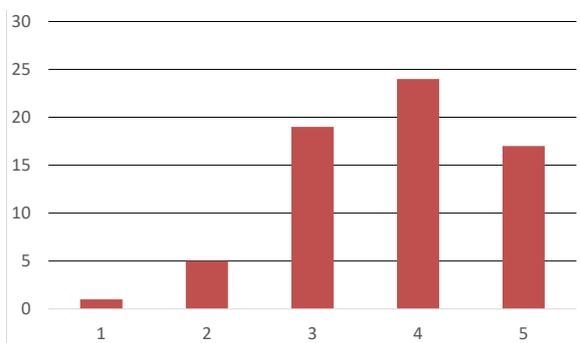
図1 テキストマイニングまとめプリント

(3) 授業後アンケートをグラフ1～3に示す。それぞれ、1「思わない」～5「思う」の5段階で回答させた。

グラフ1「テキストマイニングの意義（有効性）が理解できましたか。」については、3や4と答えている生徒が多く、「理解できた気がする。」という生徒が多いのだと考えられる。授業開始時に、WEB上にあるヒットソングの歌詞のテキストマイニング結果を見せて説明した際、比較的生徒の反応が良かった。生徒の興味関心の高いテーマで行うことでもう少し理解が深まるかもしれない。さらに理解を深めるためにどうすればよいのか、今後の課題である。

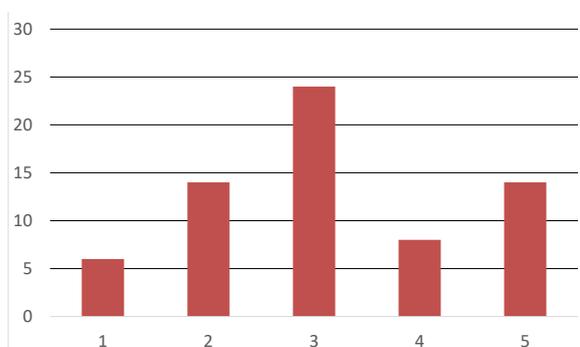
グラフ2「今後の新居浜市の課題・改善点が発見できましたか。」については、4や5と答えている生徒が非常に多く、テキストマイニングが問題解決の手段として利用できたと言える。しかし、マイニング用のデータを見て考え出そうとする班もあり、生徒の様子を小まめに観察し、助言をする必要を感じた。このことに関しては、マイニング前のデータから考察することができないようなテーマを設定する工夫を行えば解決できるのではないかと考える。

グラフ3「またテキストマイニングに挑戦したいと思いましたか。」については、答えが分散されている。これは、興味を持ち主体的に考えることができた生徒、興味がわかず他のメンバーに任せてしまった生徒、やってみたら興味・関心を得ることができなかった生徒、に分かれてしまったのではないかと考えられる。今後様々な改善を行い、4、5と答える生徒を増やしていかなければならない。



グラフ1 テキストマイニングの意義（有効性）が理解できましたか。

グラフ2 今後の新居浜市の課題・改善点が発見できましたか。



グラフ3 またテキストマイニングに挑戦したいと思いましたか。

(4) テキストマイニングにより得られた結果をパワーポイントでまとめさせ、プレゼンテーションを行った。分かりやすいスライドを作ろうとする様子は見られたが、テキストマイニング結果を使わずにスライドを作成しようとする生徒が多く見られた。何のためのテキストマイニングであったのか、マイニング後に整理する時間を設ける必要を感じた。また、性別による違いを使用する班が1つしかなく、どのデータを活用すべきなのか選択する力も伸ばすことも今後の課題である。発表に使用したスライドの一部を図2に示す。

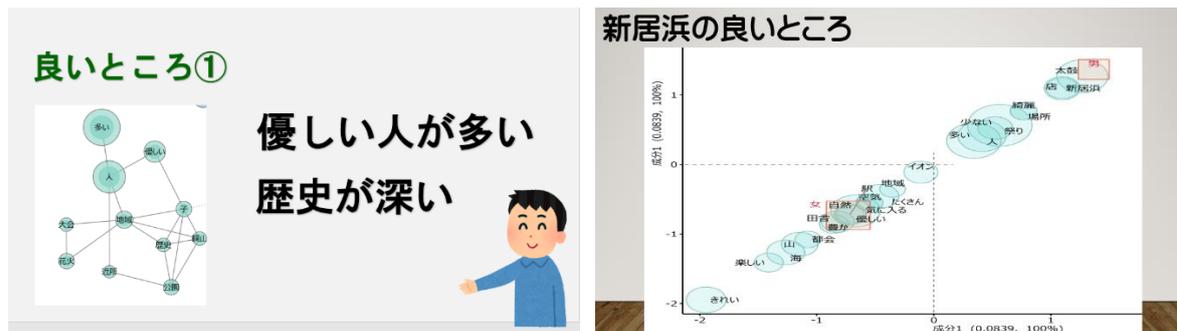


図2 発表に使用されたプレゼンテーションスライド

4 まとめ

「KH Coder」の操作方法はそれほど難しくないので、初めてテキストマイニングを行う生徒たちも戸惑うことなく操作できていた。しかし、生徒にはテキストマイニング後の考察が難しく、適時助言を与えながら授業を行わなければ、ただの作業になってしまう可能性がある。また、生徒に興味・関心を持たせるためには、テーマ設定が重要となる。アンケート結果のマイニングでは、結果をまとめるだけの作業になりがちで、語句の関連性まで考えることが難しい。生徒にとって身近で、テキストデータのみでは想定しにくく、データ数が多いテーマを考えていくことが今後の課題である。

今回の授業では、生徒がとても積極的に活動していた。初めて使うソフトで、出力されたデータを用いて自分たちで答えを導き出すことに面白みがあったのではないかと考えている。新学習指導要領においては、Word、Excel等の生徒が使い慣れたソフトだけではなく、新しいソフトを積極的に取り入れ、生徒の思考力を伸ばす実習を行う必要を感じた。

参考文献

- ・樋口耕一 (2014) 『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
- ・末吉美喜 (2019) 『テキストマイニング入門－Excel と KH Coder でわかるデータ分析』 オーム社
- ・牛澤賢二 (2018) 『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