

統計データの表現を含むレポート作成指導の支援システム

吉根勝美^{*1}

^{*1} 南山大学

Support System for Teaching Report Writing including Representation of Statistical Data

Katsumi YOSHINE^{*1}

^{*1} Nanzan University

大学の授業では、理解度の確認や成績評価のために、学生にレポートを課すことが多い。大学生に対して、レポートの構成の仕方や倫理的な注意事項について指導されることは多い。統計データを根拠とするレポートの場合、データを適切に要約した文章を書く必要があるが、この指導方法が確立されているとは言いがたい。本研究では、統計データに関わる文章をデータベース化して、これをレポート作成指導に役立てる方法を検討する。

キーワード:大学の授業, レポート作成指導, 統計データ, 統計グラフ

1. はじめに

大学の授業では、理解度の確認や成績評価のために学生にレポートを課すことが多い。このとき、一般的に行われる指導内容は、テーマの見つけ方やレポートの構成の仕方、あるいは倫理的な注意事項であることが多く、関連する図書も数多く刊行されている（例えば文献(1)(2)など）。

筆者が所属する経済学部では、統計データを根拠としたレポートを課すことが多いので、データを適切に要約した文章を書くという指導も必要となる。しかし、この指導方法が確立されているとは言いがたい。そこで、本稿では、実際の授業で筆者が行っている指導例をもとにして、統計データに関わる文章をデータベース化して、これをレポート作成指導に役立てる方法を検討する。

2. 統計データを含むレポートの指導

2.1 レポートを課す目的

筆者が所属する経済学部では、客観的なファクトに基づいたレポートを書く練習の一環として、経済統計データを根拠としたレポートを課すことが多い。この

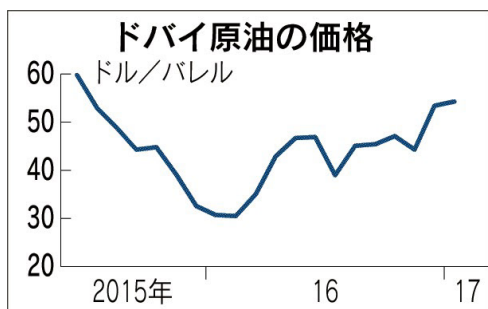
とき、表計算ソフトを使って、統計データに簡単な分析を施して、グラフを作成する指導が主となり、統計データを適切に要約した文章を書くという指導は、どうしても後回しになる。そこで本章では、統計データを文章で表現する事前練習として、実際に行った指導例を2つ紹介する。

2.2 指導例1

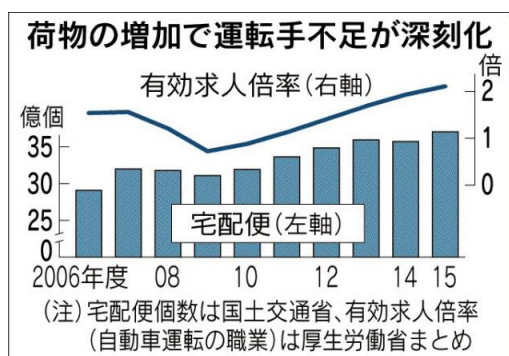
統計グラフの特徴を読み取る練習として、新聞記事に添えられているグラフの伏せられたタイトルを考えさせた。文献(3)によれば、グラフのタイトルの付け方は2通りある。ひとつは *informative title* で、図1 (a)のようにデータを理解するのに必要な情報が入っているタイトルであり、通常はこのタイプのタイトルがつけられることが多い。もうひとつの付け方は *descriptive title* で、グラフに見られるパターンや傾向を強調するタイトルで、図1 (b)のように、グラフから読み取れるストーリーを短い文で示している。

統計グラフの特徴を読み取る練習としては次のように実践した。*descriptive title* がついているグラフを含む新聞記事を提示する。このときグラフのタイトルを伏せておき、記事の内容を踏まえて、グラフに

descriptive title をつけるように指示する。その後、学生ひとりひとりがつけたタイトルをすべて書き出して、学生同士でタイトルの適切さを議論させた。



(a) informative title の実例



(b) descriptive title の実例

図 1 グラフタイトルの 2 通りのつけ方 (2017 年 1 月 31 日付 日本経済新聞朝刊より)

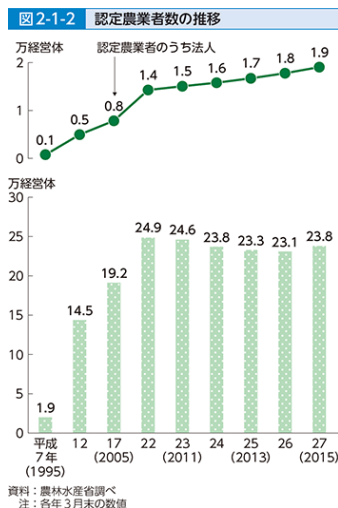
2.3 指導例 2

統計グラフから読み取ったことを文章として表現する練習として、統計データを要約して説明している実際の文章から作成した穴埋め問題を出題した。各省庁が刊行する白書、年次報告等には、統計グラフが多く掲載され、データを要約した文章が数多く見られる。そこで、穴埋め問題の題材として、白書類に掲載されている統計グラフと文章を使用した。

例えば、農林水産省刊行『平成 27 年度 食料・農業・農村白書』の 99 ページには、認定農業者数の動向を説明する統計とグラフが記載されている (図 2)。これを題材として、図 3 の穴埋め問題を作成し、WebClass のアンケート機能を使って出題した。

(4)と(6)は、グラフから数値を読み取り、そのまま回答すればよい。(4)は、23.8 万あるいは 238,000 を正解とし、(6)は、1.9 万あるいは 19,000 を正解とした。

(3)は、23.1 万経営体から 23.8 万経営体への増加率を問われているので、 $(23.8 - 23.1) / 23.1 * 100 (=3.03\cdots)$ の計算をしてもらい、3 あるいは 3.0 と正解とした。



認定農業者数は平成 22 (2010) 年までは一貫して増加してきましたが、高齢等を理由に 5 年間の計画期間終了後に再認定申請を行わない者がいること等から、平成 23 (2011) 年より減少に転じていました (図 2-1-2)。しかしながら、平成 27 (2015) 年においては、前年に比べ 3%増加し、23 万 8 千経営体となりました。この要因としては、平成 27 (2015) 年産からは、経営所得安定対策が、全ての販売農家を一律に対象とするのではなく、認定農業者等の担い手を対象に規模要件を課さずに実施されることとなったこと等が考えられます。また、認定農業者のうち法人の数は一貫して増加しており、平成 27 (2015) 年は、1 万 9 千経営体となっています。

図 2 統計グラフとその説明文の実例 (平成 27 年度 食料・農業・農村白書より)

認定農業者数は平成 22 (2010) 年までは【 (1) 】してきたが、高齢等を理由に 5 年間の計画期間終了後に再認定申請を行わない者がいること等から、平成 23 (2011) 年より【 (2) 】。しかしながら、平成 27 (2015) 年においては、前年に比べ【 (3) 】%増加し、【 (4) 】経営体となった。この要因としては、平成 27 (2015) 年産からは、経営所得安定対策が、全ての販売農家を一律に対象とするのではなく、認定農業者等の担い手を対象に規模要件を課さずに実施されることとなったこと等が考えられる。また、認定農業者のうち法人の数は【 (5) 】しており、平成 27 (2015) 年は、【 (6) 】経営体となっている。

図 3 統計グラフの要約文の穴埋め問題の例

(1),(2),(5)には、文章として意味が通るように語句を当てはめることになる。白書の原文に従うなら、(1)と(5)にはどちらも「一貫して増加」が、(2)には「減少に転じた」が正解となる。ただし、(1)と(5)の場合には、単に「増加」と書いても、(2)の場合には、単に「減少した」と書いても、統計グラフの説明として間違いではない。

本研究の目的は、統計データを根拠としたレポートを書くにあたり、データを適切に要約する指導方法を確立することである。この目的に沿うならば、単に「増加」「減少」と書かせるだけでは不十分である。前述の(1)(5)の場合は、指定された期間においては、一度も減少することなく増加し続けたことを明記するように指導すべきである。(4)の場合は、それまで増加が続いた

のに、指定された時期を境にして減少に転じたことを指導すべきである。このように、単なる増加・減少ではなく、データの動向をより適切に表現するような語彙を使いこなせるような指導が必要である。

3. レポート指導支援のシステム化

3.1 システム化の必要性

2章で示したレポート指導の実践例は、紙のプリントや WebClass で学習者に提示した課題である。しかし、課題の題材となる新聞記事や白書は、教員が実際に目を通して探している。指導例1では、descriptive title がついているグラフを含む新聞記事を見つけ出し、グラフのタイトルを隠した新聞記事を紙のプリントにして配布している。指導例2では、さまざまな白書から、統計グラフと説明文がうまく対応している部分を見つけ出し、空欄にすべき箇所を手作業で見つけ、WebClass のアンケート形式として穴埋め問題を作成している。

こうした現状のままでは、作成できる問題量が圧倒的に不足するため、統計データに関わる文章をデータベース化することを検討した。以下では、指導例2のシステム化について検討する。

3.2 ウェブで公開されている白書

本研究では、統計データを要約して説明する文章の穴埋め問題の素材を、各省庁が刊行する白書、年次報告等に求めている。これらの白書類のほとんどはウェブで公開されているので、これを利用したシステムを考える。

総務省が運営するポータルサイト e-Gov (電子政府の総合窓口) によれば、現在刊行されている白書は少なくとも 38 タイトルあり、その大半は、全文がウェブで公開されている⁽⁴⁾。ただし、その公開方法は、PDF 形式であったり HTML 形式であったり、それぞれを併用したりまちまちである。PDF の場合は、全文をひとつの PDF ファイルで公開しているところもあれば、いくつかの PDF ファイルに分割して公開しているところもある。HTML の場合は、図4のように、左右に画面を分割して、左側に目次、右側に本文を掲載しているものが多い。

3.3 白書本文とグラフのデータベース化

データベース化にあたっては、新しい白書が公開されるのが原則として年1回であることを踏まえて、ある白書を閲覧するために初めて省庁のウェブサイトアクセスするときに、データベースを構築する PC に白書のコピーを作成し、2回目以降のアクセスでは、その PC に保存されているコピーを利用するものとする。

以下では、2016年版中小企業白書を例として、データベース化に必要な要件を検討する。同白書は、その全文が複数の PDF ファイルで提供されているとともに、図4に示すように HTML 形式でも公開されている。本研究で必要なのは、統計グラフとその説明文を容易に引き出せるようにすることである。



図4 ウェブで公開されている白書の例 (2016年版中小企業白書より)

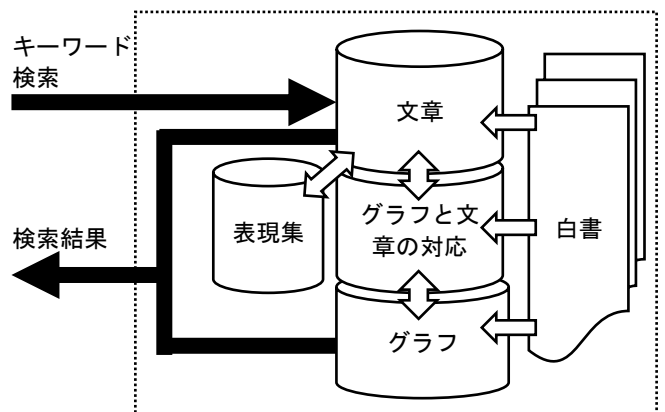


図5 データベースの概要

データベースの概要を図5に示す。白書に含まれる文章は、文章データベースに、図表はグラフデータベースにそれぞれ格納する。また、グラフと文章の対応表を別途作成する。さらに、統計データの説明文に特

有な表現集を随時蓄積していく。例えば、前章の指導例2で示したように、「一貫して増加」、「減少に転じた」のような表現を、表現集データベースに格納しておく。

例えば、「減少」をキーワードとして検索すると、図6のように検索結果が得られる。この検索結果を利用した指導例を次に示す。

- ・単純にグラフと文章を学習者に提示して、統計データを要約する文章の実例を数多く読ませる。
- ・統計数字を空欄にした穴埋め問題を作成し、グラフの数値を読み取る練習問題とする。例：規模別に見ると、2012年から2014年の2年間で、小規模企業の数【 】万者減少しているものの、中規模企業の数逆に【 】万者増加しており、合計で約【 】万者の減少となった。
- ・簡単な計算を含む穴埋め問題を作成する。例：2009年7月から2012年2月にかけて、年平均で▲【 】万者であったのに対し、2012年2月から2014年7月にかけては年平均で▲【 】万者と、減少のペースは緩やかとなった。
- ・データの動向をより適切に表現する練習としての穴埋め問題を作成する。例：2009年7月から2012年2月にかけて、年平均で▲13.5万者であったのに対し、2012年2月から2014年7月にかけては年平均で▲1.8万者と、減少のペースは【 】となった。
- ・統計データの説明文に特有な表現集に蓄積する語彙を抽出する。例：「長期にわたり減少傾向にある」「減少のペースは緩やかとなった」

中小企業数の推移を見てみると、長期にわたり減少傾向にある。

足下の2012年から2014年にかけての推移についても、2年間で4.4万者の減少となったが、2009年7月から2012年2月にかけて、年平均で▲13.5万者であったのに対し、2012年2月から2014年7月にかけては年平均で▲1.8万者と、減少のペースは緩やかとなった。

規模別に見ると、2012年から2014年の2年間で、小規模企業数は9.1万者減少しているものの、中規模企業数は逆に4.7万者増加しており、合計で約4.4万者の減少となった。

次に、小規模企業数の推移の内訳を確認すると、2012年から2014年の2年間で9.1万者の減少であったが、そのうち、開業については、開業が28.6万者、廃業が45.7万者であり、開業から廃業を引いた数が▲17.1万者と廃業が大きく上回った。

中規模企業と小規模企業間での移動が▲0.5万者であることを考えると、小規模企業の減少のほとんどの要因が廃業であり、廃業数が開業数を大きく上回ったことが影響しているといえる。

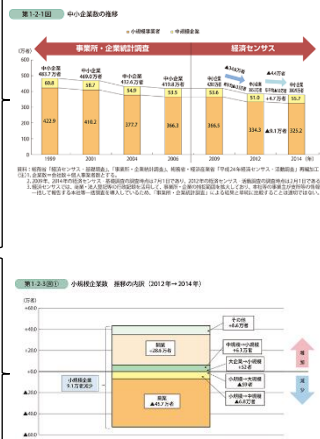


図 6 検索結果の例

4. おわりに

本稿では、大学生に対して、統計データを根拠としたレポートを課すとき、データを適切に要約する指導方法が確立することを目的として、統計データに関する文章を多く含む白書をデータベース化して、これをレポート作成指導に役立てる方法を検討した。

昨今は、人工知能が作成した新聞記事が見られるようになった。日本経済新聞社が日経電子版等で2017年1月25日に配信開始した『決算サマリー』は、元データの上場企業の開示資料や決算データから人工知能が文章を作成し、配信するまで完全自動化されている⁶⁾。

しかし、統計データをよく観察して、特徴を見つけ出し、レポートの文章にまとめる能力自体は、人工知能の進展に関わらず、大学生・社会人に必要であろう。本研究では、大量の統計データとその文章表現をデータベース化し、大学生にその能力を育成するための指導を支援することを目標としている。

今後は、このデータベースを活用して、穴埋め問題作成の完全自動化や、学生が書いた文章の評価への利用方法を検討したい。なお、本研究の一部は科研費26282052の助成を受けたものです。

参考文献

- (1) 木下是雄: “レポートの組み立て方(ちくま学芸文庫)”, 筑摩書房, 東京 (1994)
- (2) 戸田山和久: “新版 論文の教室—レポートから卒論まで (NHK ブックス)”, NHK 出版, 東京 (2012)
- (3) United Nations Economic Commission for Europe: “What makes an effective chart”, Making Data Meaningful Part 2: A guide to presenting statistics”, p.23, Geneva (2009), <http://www.unece.org/stats/documents/writing/> (2017年2月6日確認)
- (4) 総務省行政管理局: “白書、年次報告書等”, e-Gov 電子政府の総合窓口, http://www.e-gov.go.jp/link/white_papers.html (2017年2月7日確認)
- (5) 日本経済新聞社, 言語理解研究所, 東京大学: “完全自動「決算サマリー」ベータ版(試用版)”, <http://pr.nikkei.com/qreports-ai/> (2017年2月6日確認)