

ビッグデータの活用を学ぶ中学生対象授業プログラムの開発 —未来の学校における選挙予測を題材として—

藤川 大祐¹⁾ 阿部 学¹⁾ 和田 翔太²⁾ 三倉 知也²⁾

小川 愛³⁾ 塚本 亜紀³⁾ 畠 慎一郎³⁾

千葉大学教育学部¹⁾ NPO 法人企業教育研究会²⁾

日本アイ・ビー・エム株式会社³⁾

情報技術の発展に伴い、大量のデータを研究やビジネスに役立てようとする「ビッグデータの活用」が進み、私たちの生活に大きな影響を与えている。しかし、義務教育における教科等の授業で、ビッグデータの活用が扱われることは非常に少ない。そこで、発表者らは、数万件程度のデータを扱うこと、未来の学校の生徒会選挙を題材とすること、デジタル教材を作成するとともにデータ分析の仕事をする人をゲストとして教室に招くことを方針とし、ビッグデータの活用を扱う授業プログラムを開発することとした。このプログラムを試行的に実践した（中学校3年生選択数学授業、45分1コマ）ところ、架空の物語を描いたデジタル教材を利用することによって、データ分析のあり方に関して一定の学習が可能になることが示唆された。

キーワード：ビッグデータ、デジタル教材、データ・サイエンティスト

1. はじめに

藤川（2012）で論じているように、コンピュータ技術の発展やインターネットの普及は社会のあり方を劇的に変えている。学校教育における教育内容に関しても、こうした社会の変化を前提に検討がなされる必要がある。

ここで言う社会の変化として重要な点の一つが、いわゆる「ビッグデータの活用」である。コンピュータの計算速度の向上と統計的手法の発展により、以前では考えられなかったほどの巨大なデータが、研究やビジネスにおいて活用されるようになってきている。

総務省（2013）によれば、「ビッグデータの定義は依然として曖昧」であるが、多様な捉え方に共通する点として「多量性、多様性、リアルタイム性等」が挙げられる。業務用データや顧客データといった構造化されているデータだけでなく、音声データ、映像データ、ブログやSNSのデータ、各種センサーで検知され送信されるデータ等の構造化されていないデータについても、ICTの進展に伴って分析可能となっており、ビッグデータにはこうした多様なデータが含まれる。

木谷（2014）によれば、CPUの周波数、ストレージ

（記憶装置）の容量、ネットワークの速度といったハードウェアの性能は指数関数的に発展しており、2025年にはスーパーコンピュータが人間の脳をシミュレーションすることが可能となることが予測される。ハードウェアの性能の進化によって複雑なアルゴリズムの実行が可能となり、これまで以上に多量で多様なデータを分析し活用することが容易になると考えられる。

では、「ビッグデータの活用」に関して、初等中等教育の教育課程の状況はどうか。

現行の学習指導要領でも、中学校数学で「資料の活用」が強化される等、統計に関わる内容の充実がはかられている。

しかしながら、高等学校までの数学教育では統計に関して、平均値、相関、確率分布、母集団と標本等の項目を扱うにとどまり、実社会でのビッグデータの活用との間の距離は大きい。社会科や公民科においても、統計資料を扱うことはあるものの、データの扱いは限定的である。

松寄ら（2014）は、「新教育課程編成に向けた系統的な統計指導の提言」の一つとして、「テクノロジー利用を前提とした、ビッグデータや実データを扱う指導」を挙げ、高等学校第1学年において表計算ソフト等のテクノロジーを用いてヒストグラム作成等の学習を行わせることを提案している。扱われるデータは実データが理想だとしつつも、「統計指導のねらいを達成するためには、架空のデータであってもかまわない」とし、「架空のデータを扱う指導」を行うことをも提案している。

Daisuke FUJIKAWA¹⁾ Manabu ABE¹⁾ Shota WADA²⁾
Tomoya MIKURA²⁾ Ai OGAWA³⁾ Aki TSUKAMOTO³⁾
Shinichiro HATA³⁾

Faculty of Education, Chiba University¹⁾ NPO, the
Association of Corporation and Education²⁾ IBM Japan
Co. LTD³⁾

松寄らは、東京都統計協会・東京都総務局統計部が作成した「どこで何を売る？ 自動販売機戦略」という教材¹を架空のデータを扱う教材の例として取り上げている。この教材はビッグデータの活用をうたうものではないが、多様なデータを扱って提案書を作成することを課すものである。このような教材が多く作成され、広く活用されることが求められる。

私たちの生活への影響の大きさを考えれば、ビッグデータの活用に関わることを学習者に明示した授業を、義務教育段階で学べるようにすることが求められるはずである。筆者らは、ビッグデータの活用について中学生が概観できる授業の開発を試みた。本稿は、筆者らが行った授業プログラムの開発のあり方について報告し、中学校での試行的な授業の実施をふまえて授業プログラムについて考察するものである。

なお、この取り組みは日本アイ・ビー・エム株式会社の社会貢献事業の一環としてなされているものであり、同社の協力を得て主に NPO 法人企業教育研究会（発表者のうち藤川、阿部、和田、三倉が担当）が授業プログラムの開発、実施を行った。²

2. 授業プログラムの開発

2.1. 授業プログラム開発の方針

では、中学生が「ビッグデータの活用」を概観できるようになることを目指す授業プログラムは、どのような方針で開発されるべきであろうか。

第一に、取り上げるデータの規模についてである。ビッグデータと言うと、数千万件、一億件、あるいはそれ以上の規模の数のデータと考えられるが、こうした規模のデータについて中学生たちが具体的に想像することは困難だと考えられる。他方、数百件から数千件程度のデータではビッグデータとの規模の違いが大きすぎる。こうしたことから、数万件程度の規模のデータを扱うことが適切であると考えられる。数万件程度のデータを扱うを通して、より規模の大きいビッグデータについて想像できるようにすることが考えられる。

第二に、題材についてである。ビッグデータの活用例としては、典型的には工業生産や商業活動に関わるものが考えられる。だが、こうした例を取り上げた場合、中学生はそもそもの産業や仕事の状況に親しんでいない場合が多く、そうした状況について学んでもらう負担が大きくなる。これらとは異なり、ビッグデータの活用例があるものの中で中学生が理解しやすいものを扱う必要がある。そこで、社会科でも扱われる選挙を取り上げ、選挙の予測報道を扱うこととした。選挙の予測報道は、過去の選挙の実績、各種世論調査、選挙区の様子等、各

種データを活用して行われる。しかしながら、現実の選挙をそのまま取り上げると中学生が理解しにくい要素が多くなると考えられる上に、実際に利用できるデータが限定されることから、設定にフィクションを取り入れることとした。具体的には、生徒数 35,000 人の未来の学校の生徒会選挙に関する予測報道を取り上げることとした。

第三に、教材や指導方法についてである。未来の学校における生徒会選挙という架空の設定を学習者が理解しやすくするために、報道番組やある種のテレビドラマの雰囲気に近いデジタル教材を作成することとした。また、実際にデータ分析をしている人をゲストとして教室に招いて専門の立場からコメントをしてもらえるようにし、さらにデジタル教材と教室とを円滑に接続するために、デジタル教材の物語の中で授業者やゲストが登場する場面を設けることとした。こうした方法を用いることにより、前述の「どこで何を売る？ 自動販売機戦略」と比較しても、生徒が教材の内容に興味を持ちやすくなり、複雑な作業を扱いやすくなると考えられる。

2.2. デジタル教材の作成

上記の方針に基づいて、デジタル教材の作成を行った。

デジタル教材の作成にあたっては、NPO 法人企業教育研究会側の担当者が日本アイ・ビー・エム側の担当者に実際の選挙予測で使われている手法についてヒアリングし、そうした手法が効果的に適用できる物語を構築することとした。

物語の舞台は、西暦 2200 年の巨大な学園とした。生徒数 35,000 人の学校を現代の日本において想定することは困難であることから、現在の中学生がほぼ寿命を終える程度の未来である西暦 2200 年の設定とした。現状のままでは日本でかなり少子化が進むはずであり、そうした中で大規模な学園を設定することには無理があるため、留学生や移民を積極的に受け入れ、機能を集約した学園都市が作られたという設定とした。また、現実の選挙において、一部の富む者を優遇するか広く富を再分配するかという対立軸が生じやすいことから、これを学園の状況に置き換えて、成績優秀者を優遇する政策を継続するか否かを選挙の主要な争点とすることとした。

成績優秀者を優遇する政策は、「AT システム」という名称とし³、成績によって生徒を 10 のレベルに分け、レベルが高いほど電子機器の自由な利用等の優遇がなされる一方で、レベルが低い生徒には恋愛禁止等の制約が課されることとした。

授業の中では、生徒たちに 3 種類の課題（ミッション）を与えることとし、選挙戦の進展とともに新たな課題が与えられるよう物語を展開することとした。このこ

とを可能にするため、選挙の告示から投票までの期間を日本の国政選挙よりも長く、8週間とした。また、途中で選挙戦の情勢が一変するエピソードを入れ、最終的には接戦となるようにするなど、生徒の興味をひきつけられるようドラマとしての劇的な展開にも留意することとした。

登場人物である3人の立候補者については、その区別が容易になるよう、性別や国籍を多様とし、専門のスタッフによるキャラクターデザインを施し、声に特徴あるスタッフが声優の役割を担って音声で台詞をつけることとした。また、これら以外の登場人物として新聞部の部長が登場するが、これについては3人の候補者ほどには目立たないようにする必要があるため、音声の台詞はつけるものの、デザインはシルエットのみとした。また、教室で授業を進める授業者は新聞部の副部長、ゲスト（データ分析を担当している日本アイ・ビー・エムの社員）は新聞部の顧問という設定とした。

主要なキャラクターのデザインは、図1から図4の通りである。

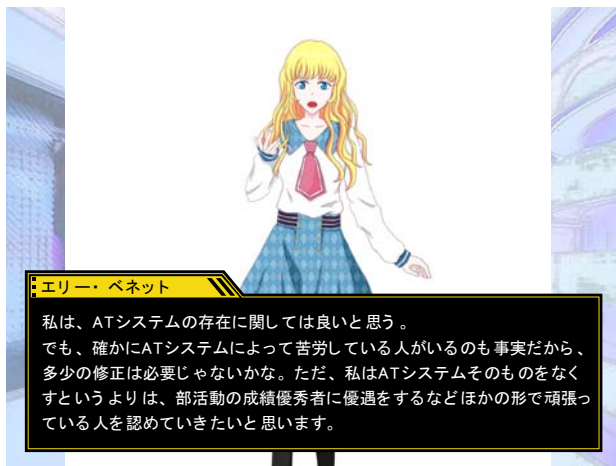


図1 エリー・ベネット候補

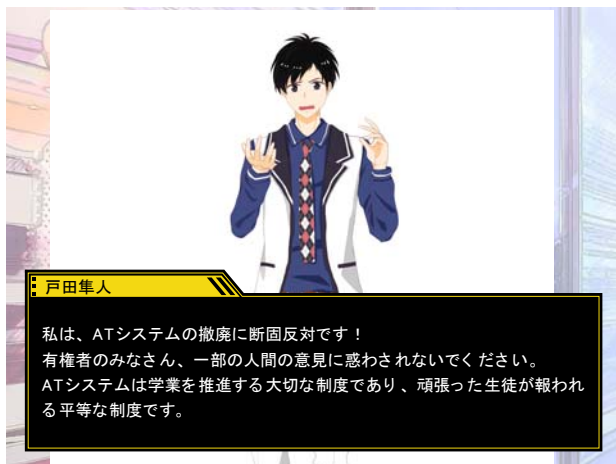


図2 戸田隼人候補

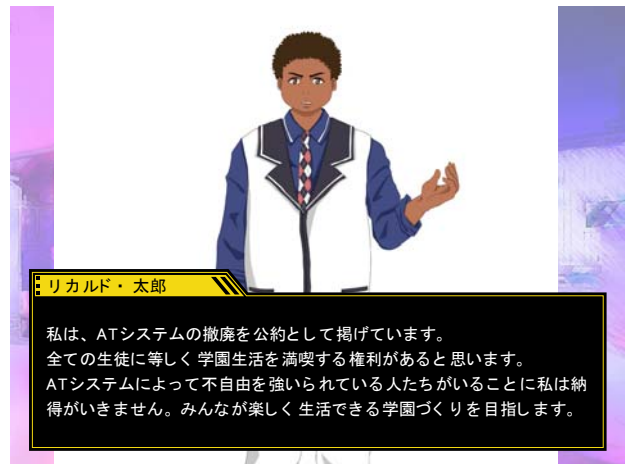


図3 リカルド・太郎候補

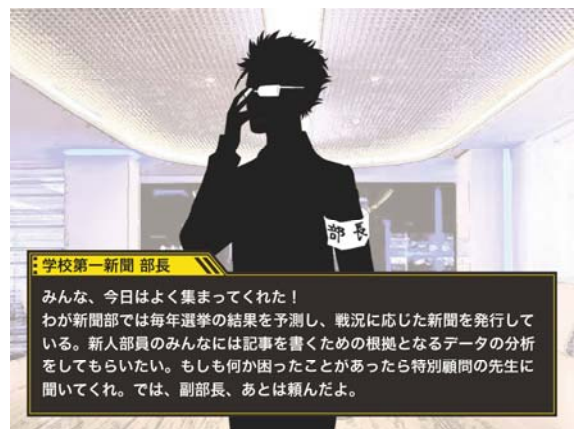


図4 新聞部顧問

物語の設定の説明に多くの時間をとることは、授業時間の使い方として不適切であると考えられるため、約5分の導入映像のみで、生徒たちに設定を説明し、物語の世界に入り込むことが可能となるようにした。このために、映像ではテンポのよい音楽とナレーションをつけ、一部にアニメーションや効果音も入れ、生徒をひきつけられるようにした。他方で、映像を繰り返し見ることが難しいため、必要な情報は配付資料にも掲載することとした。導入の映像のデザインについては図5、図6を参照。



図5 タイトル画面

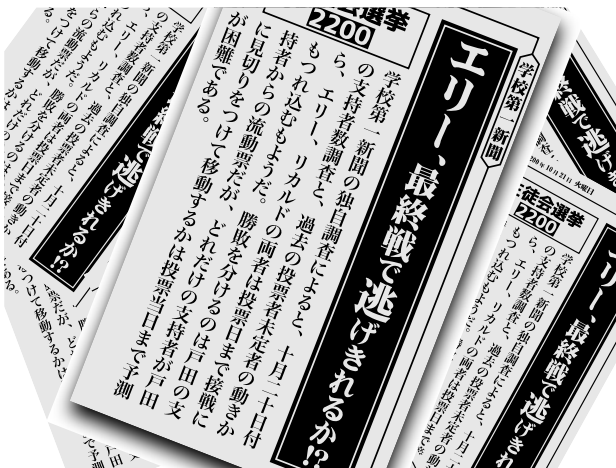


図6 予測報道の紙面

また、未来の学校の話が一通り終わった後に、補足として、ビッグデータの活用がどのようになされているかを、日本アイ・ビー・エムが関係している事例を中心に紹介することとした。具体的には、名古屋鉄道におけるICカード「manaca」のデータを活用する事例⁴を中心として紹介することとした。

2.3. 課題の作成

先述のように、今回の授業プログラムにおいては、3つの課題を「ミッション」として課すこととした。

第一のミッションは、序盤戦における選挙結果の予測である。当該新聞部では生徒会選挙の情勢について1週間に1度、全校生徒に対する調査を行っているという設定を設け、過去3年の候補者別支持者数推移のデータと、2200年度における告示日から2週目までの序盤の支持者数調査のデータをグラフで示し、この状況からどのような予測報道が可能かを検討させることとした(図7)。

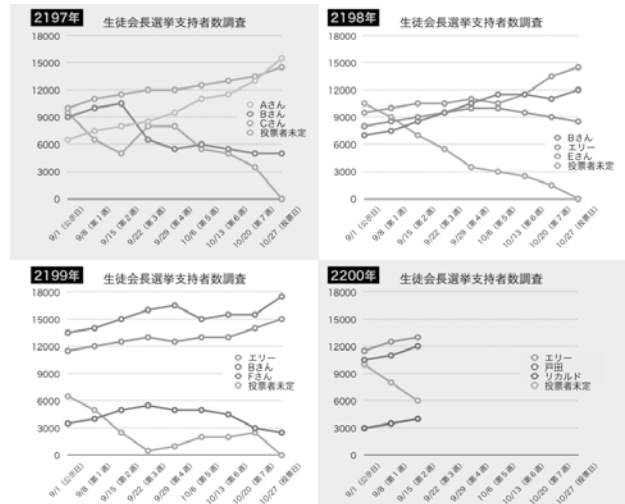


図7 序盤戦における選挙結果予測のためのデータ

この段階で言えることは限定されており、エリー候補と戸田の2候補が争っていてリカルド候補が離されていること、複数年続けて立候補している候補は票を伸ばす例が見られること等に着目がなされ、エリー候補と戸田候補の争いを示唆する報道が考えられるということが生徒から出されることを想定した。

第二のミッションは、中盤に選挙情勢が大きく変わり、生徒の成績階層別の支持傾向のデータにもとづいて予測を検討するというものとした。選挙戦の中盤で、最も支持者が少なかったリカルド候補が、ATシステムがもともと説明されていた以上に差別的なシステムであり、成績中位の者には実質的な優遇はなく、上位の者のみが不当に優遇されるシステムであることを暴く動画を配信したことがきっかけで、ATシステム推進論者の戸田候補の支持数が落ち込んだことを示している(図8、図9、図10)。



図8 ATシステムの問題を暴くりカルド候補の動画

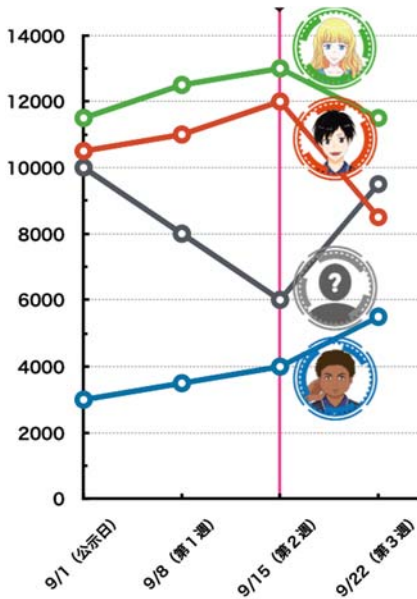


図9 情勢が大きく変わり始めた選挙戦

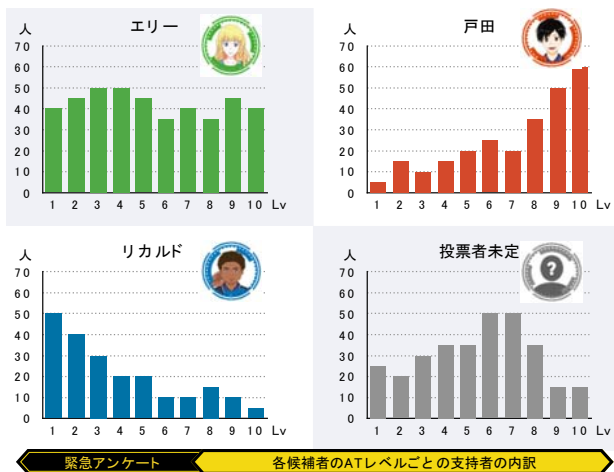


図10 レベル別の支持傾向データ

このミッションにおいては、全校生徒を一様の集団として見るのではなく、レベル別に傾向を分けてとらえ、分析することを求めている。リカルド候補の演説はレベル注意者に大きな影響を与える可能性が高いことを前提に、各候補の支持傾向がこのあとどのように変動するかを予想させるものである。

第三のミッションは、選挙終盤において、投票先未定の人に対する調査をふまえた予測をすることである。投票先未定の人々の投票行動を予測するために、投票時に重視する項目と、各項目に関する各候補の評価を問い、この結果から投票先未定の人々がどの候補に投票するかを予測させるものである(図11、図12)。

緊急アンケート② 各候補者に対する意識調査

質問項目	エリー	戸田	リカルド	計
最も勢いがあると感じるのは	60	10	130	200
信頼できるのは	100	10	90	200
熱意を感じるのは	100	10	90	200

※1200人にアンケートを行い、投票者未定と答えた200人の結果を表示

図11 投票者未定の人による各候補への評価

今回調査した200人が、どの候補者に投票しようか計算しよう。

	エリー	戸田	リカルド
「勢いがあると感じる人」に投票するのは	$60 \times 0.5 =$ <input type="text"/>	$10 \times 0.5 =$ <input type="text"/>	$130 \times 0.5 =$ <input type="text"/>
「信頼できると感じる人」に投票するのは	$100 \times 0.3 =$ <input type="text"/>	$10 \times 0.3 =$ <input type="text"/>	$90 \times 0.3 =$ <input type="text"/>
「熱意を感じる人」に投票するのは	$100 \times 0.2 =$ <input type="text"/>	$10 \times 0.2 =$ <input type="text"/>	$90 \times 0.2 =$ <input type="text"/>
合計	<input type="text"/>	<input type="text"/>	<input type="text"/>

図12 投票者未定の人々の投票を予測するワークシート

このミッションでは重みづけ計算を行うことによって投票者未定の人々の投票行動を予測するものであり、実際の選挙予測でも用いられている手法を反映したものであるが、予測の精度はかなり限定される。このことを印象づけるために、教材ではこの後に選挙結果を示すこととし、ミッション3での計算とは異なる結果となるという展開とした。すなわち、ミッション3ではエリー候補がリカルド候補を僅差で破って当選することが示唆されるのであるが、結果としてはリカルド候補が僅差で当選することとした。

3. 授業の実施と考察

3.1. 授業の実施

本授業プログラムを用いた試行的な授業を、2014年7月9日に千葉大学教育学部附属中学校3年生選択数学授業にて実施した(45分、1コマ、受講者15名)。

この選択授業は同校が独自に設けているものであり、筆者のうち藤川が研究室の学生たちとともに、選択数学のクラスを担当している。クラス自体は「ゲームで学ぶ数学」というテーマで設けられており、全13回の授業が行われる。その中の1回を特別授業として、本稿で述べている授業にあてた。主な授業担当者は、筆者のうち、千葉大学の学生でありNPO法人企業教育研究会のスタッフでもある三倉が担当した。

導入で動画が流されると生徒たちには興味を持って視聴する様子が見られた。そして、ミッション1の作業開始を指示されると、生徒たちはしばらく頭を抱えたりしながらデータを眺めてから、徐々に考えをワークシートに書いていった。なかなか書けない生徒もいた。ミッション1について拱手した生徒に発表させたところ、エリーが3位、2位と前年まで順位を上げているので今回は1位になる可能性がある、過去の例から第3位の候補が上昇することは難しい等の意見が出され、ゲスト講師からは「非常に鋭い分析」「トレンドを見て予測している」などのコメントがなされた。

ミッション2についても生徒たちは詳細に分析を行い、両極端の意見をもった人がお互いにつぶしあうのではないかという意見が出された。ゲスト講師からは、常識から考えることとデータから言えることを区別し、投票未定者のデータから言えることを考えるようにという話がなされた。

ミッション3については残り時間が厳しくなったことから、早めに生徒を指名し、ホワイトボードに数値を記入させ、予測の難しさを確認した。

全体として、生徒は動画教材や予測作業に強い興味を示して取り組んだものの、45分の授業の中で三つのミッションに取り組むことは困難であり、時間が足りない様子が見られた。

3.2. 事後アンケートの結果

事後アンケートの結果は、表1の通りである。

表1 事後アンケートの結果

(1)今日の授業は楽しかったですか？

1.とても楽しかった	15
2.まあまあ楽しかった	0
3.あまり楽しくなかった	0
4.つまらなかった	0

(2)今日の授業の内容はわかりましたか？

1.とてもよくわかった	11
2.まあまあわかった	4
3.あまりわからなかった	0
4.わかりにくかった	0

(3)今日の授業を受けて数学への関心が高まりましたか？

1.とても高まった	11
2.まあまあ高まった	4
3.あまり高まらなかった	0
4.まったく高まらなかった	0

(4)今日の授業で、はじめて知ったことはありましたか？

1.たくさんあった	9
2.いくつかあった	6
3.あまりなかった	0
4.まったくなかった	0

(5)演習問題によって、学習の理解がふかまりましたか？

1.とてもふかまった	13
2.まあまあふかまった	2
3.あまりふかまらなかった	0
4.まったくふかまらなかった	0

(6)ビックデータの話はわかりましたか？

1.とてもよくわかった	9
2.まあまあわかった	6
3.あまりわからなかった	0
4.わからなかった	0

また、自由記述で書かれた授業の感想は、次の通りであった（全員分、全文、原文のまま）。

- ・ 選挙という身近なもので、今後の変化予想するのは、おもしろかった。データの分析ははじめてで、難しかったけど、ある程度変化を考えることはできた。
- ・ 選挙が展開していくにつれ、自分でもドキドキしました。データを色々な角度から分析し、予測をしていくことはとても楽しかったです。
- ・ MISSION③の投票者未定の動きの予想のところで数値を出してその後の動きを考えたりして楽しかったです。
- ・ 「数学」と言うと、答えが一つに定まるが、今回は「予想」という一つに定まらない答だったので、とても不思議で、とてもおもしろかった。また、データから読み取るのも、新しい発見があっておもしろかった。

- ・ 予測というのは、変化し続けるので、未来を完璧に予知するのは、魔法じゃない限り、無理ですが、未来に近づける考えを出すということは、今回の特別授業で理解することができました。
- ・ 数学はいろいろなところで使えるので、大切だと思った。これからも便利な数学を頑張っていきたいと思う。
- ・ 今日はすごく楽しい授業だったのでこのことを糧にしてもっと数学を頑張りたいと思います。
- ・ 今まで習っていた数学とは違い、流れるように頭のなかに入り、分かりやすく楽しかった。
- ・ 選挙の支持者数の移り変わりは、データを見ることによって、ある程度の予測を付けることが可能だと分かりました。データだけでなく状況を実際に見ることも大切だと思いました。
- ・ 具体的なストーリーがあって楽しかったです。戸田の声とリカルドの声のイメージが違った…
- ・ みんなが賛成するような公約を出して、逆転でリカルドさんが当選したことに、とても驚きました。他の物事についてもデータをもとに分析してみたいと思いました。
- ・ またデータの分析をしてみたいと思いました！ストーリーが面白かったです。ありがとうございました。
- ・ データの分析と聞いて、とても難しいことをするのかと、初めは思いましたが、ゲーム形式の演習で、とても楽しく数学を学ぶことが出来ました。ありがとうございました。
- ・ アニメーションがとてもおもしろかったです！データについて色々学べたのでよかったです。
- ・ とても分かりやすいアニメーションで、演習問題をすることでわかったことがたくさんあった。

こうしたアンケートの結果から、受講した生徒たちがこの授業を肯定的に受け止めていることがうかがえ、アニメーションを中心としたデジタル教材やデータ分析のミッションといった本授業プログラムの中心部分が生徒たちに十分に受け入れられたと考えられる。特に、授業の楽しさについての評価が非常に高く、データを分析することへの関心の喚起に成功していると考えられる。他方、「ビッグデータの話はわかりましたか」という問いには「まあまあわかった」という答えが多く、授業中の状況とあわせて考えても、全体としてもっと丁寧に内容を扱う必要があったと考えられる。

3.3. 成果と課題

未来の学校の生徒会選挙を題材に、フィクションでデータ分析を行うことを中心とした本授業プログラムは、

中学生対象のプログラムとして十分に成立しうることが示唆された。この点は本研究の最大の成果である。

他方、授業を行ったクラスが、国立大学附属中学校の選択クラスという特殊なクラスであったことについては、留意が必要である。公立学校の一般クラスで授業を行い、状況を確認することが求められる⁵。

また、本授業プログラムがそもそも数学という教科の中に位置づけられるべきなのかという点について、検討が必要である。本授業プログラムでは、数学的な知識を学んだり数学的な問題解決を行ったりすることが中心とはなっておらず、むしろ社会の中でのデータ分析を疑似体験することが中心となっている。このことから、本授業プログラムは社会科で扱われるべきであると考えることが可能である。現状では、社会科の授業においてデータの分析が扱われることはほとんどないと考えられるが、社会的な問題の解決にデータ分析が活用されることは多く、今後は社会科においてデータ分析を扱うことが求められるはずである。こうしたことを踏まえ、教科での扱いを含む本授業プログラムの教育課程上の位置づけについては再検討が必要であると考えられる。

さらに、本授業プログラムでは、ビッグデータそのものを扱うまでには至っていないことについても、注意が必要である。本授業プログラムがビッグデータ活用について学ぶ入口とはなりえても、その後には中学校もしくは高等学校の段階でどのようにビッグデータの活用が扱われるべきかについては、さらに検討が必要である。

1 『中学生のための統計学習 まなぼう統計』内の教材。
<http://www.toukei.metro.tokyo.jp/manabou/tyuu/mission2/ma1205t301.htm> (2014年2月3日最終閲覧)

2 本稿は、2014年9月19日、日本教育工学会第30回全国大会(岐阜大学)における著者らによる同名発表をもとにしている。

3 「AT」の名称は「到達度及び処遇」の意で「Achievement & Treatment」を省略したものとして考案されたものである。

4 以下参照。
<http://www-06.ibm.com/jp/solutions/casestudies/20121002meitetsu.html> (2012年10月17日掲載、2015年2月3日最終閲覧)

5 その後、実際に公立中学校で授業を実施しており、生徒たちが附属中学校の場合と同様に興味をもちミッションに取り組む様子が見られている。

引用文献

- 藤川大祐 (2012) 「「社会とつながる教員養成」がなぜ求められるのか—情報革命以降の教師を育てるために—」、藤川大祐編 千葉大学大学院人文社会科学部 研究プロジェクト報告書 第249集 『社会とつながる教員養成に関する実践的研究』、pp.1-6
- 木谷強 (2014) 「イノベーション創出に向けて」、総務省第3回「ICT新事業創出推進会議」配付資料、
http://www.soumu.go.jp/main_content/000277812.pdf (2015年2月2日最終閲覧)
- 松崎昭雄・金本良通・大根田裕・青山和裕他5名(日本数学教育学会「資料の活用」検討WG)(2014)「新教育課程編

成に向けた系統的な統計指導の提言：義務教育段階から
高等学校第1学年までを対象として」、社団法人日本数学
教育学会『日本数学教育学会誌』96(1)、2-12
総務省（2013）『平成25年版情報通信白書 ICT白書』