

データフレームを対象とした ブロックプログラミングに基づく簡易な可視化環境の設計

Design of Simple Visualization Environment for Data Frame based on Block Programming

三宅 礼華^{*1}, 小島 匡顕^{*2}, 香山 瑞恵^{*1}, 永井 孝^{*1}

Ayaka MIYAKE^{*1}, Masaaki KOJIMA^{*2}, Mizue KAYAMA^{*1}, Takashi NAGAI^{*1}

^{*1} 信州大学工学部

^{*1} Faculty of Engineering, Shinshu University

^{*2} 信州大学大学院理工学系研究科

^{*2} Graduate School of Science & Technology, Shinshu University

Email: 14t5084e@shinshu-u.ac.jp

あらまし：本研究の目的は、データフレームを対象とした簡易な可視化環境の提案である。対象とするデータフレームの種類は限定せず、汎用性の高い環境の提案を目指す。そのために、Web ベースの既存ブロックプログラミングツールを拡張することで、カラム操作や可視化に特化した環境を設計した。本稿では、提案する可視化環境について述べた上で、複数形式のデータフレームに対するカラム操作と可視化の方法を示し、提案環境の汎用性評価の結果を示す。

キーワード：可視化、データフレーム、ブロックプログラミング、Web ツール

1. はじめに

医理工学分野、社会・人文科学分野も含む多くの分野でデータ解析は古くから行われている。近年、計測機器や計算機の高度化、通勤環境の整備に伴い、関連研究が数多く報告されるようになってきている。

データ解析を行う際、データの可視化には表計算ツールのグラフ作成機能が用いられることが多い。しかし、列単位でのデータ計算を含むカラム操作や、グラフ作成などの可視化がユーザの意図に沿わない等の問題がある⁽¹⁾。また、表計算ツールでは計算過程が明示されないことも多い。追試や結果トレースのためには計算過程の明示が必要な場合も多い。

また、Jupyter⁽²⁾等のエディタツールを使用してプログラムを記述し、データ計算や可視化を行う方法もある。この方法はプログラミングに不慣れなユーザには操作が困難であろう。

データ管理およびデータ分析をサポートするツール、PerCon の開発も行われている⁽³⁾。しかし、これは大規模なデータセットへのアクセスを前提としており、簡易的にデータの可視化を支援するものではない。また、特定ソフトウェアが必要であり、使用環境が制限されていた。

本研究に関連する先行研究では、これまでに歩行動作解析を対象とした簡易可視化ツールの開発を行ってきた⁽⁴⁾。このツールでの処理対象データは歩行動作における加速度データに限定されていた。

そこで、本研究では使用環境の制限を軽減し、使用データの種類を制限しないことを目的に、データフレームを対象としたカラム単位での操作が可能な簡易可視化ツールを提案する。カラム操作の記述方法としてブロックプログラミングツールを採用した。

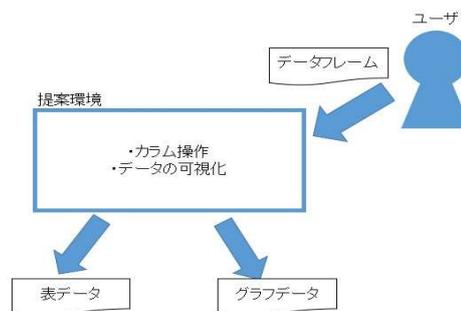


図 1 提案環境の処理フロー概念図

2. 提案する可視化環境の設計と実装

2.1 提案環境の概要

図 1 に、提案環境の処理フロー概念図を示す。まず、ユーザは解析対象となるデータフレームを提案環境にアップロードする。ここでのデータフレームは、名前付けされたカラムの集合と定義される。提案環境においてアップロードされたデータフレームのカラム操作やデータの可視化を行い、必要に応じて表形式データやグラフデータを出力する。

2.2 提案環境の特徴

提案環境の特徴は以下の 2 点である。1 点目は、Web ベースのツールであるため、ブラウザ以外に特別な環境を必要とせずに簡易的なカラム操作や可視化が可能となることである。Web ブラウザがあれば本提案環境を利用することができる。

2 点目は、可視化対象とするデータフレームの種類を限定しないことである。Plain text で表現されているデータフレームであればよい。読み込まれたデータフレームに対しては、可視化対象とするカラムに制約はない。



図 2 提案環境画面

2.3 ブロックプログラミングツールの拡張

カラム操作を行う際、ユーザは操作対象カラムと処理方法を指定する必要がある。本提案環境では、これらの指定を表現するためにブロックプログラミングツール Blockly (Google 社)⁽⁶⁾を採用した。このツールを採用することで、プログラミングに不慣れなユーザであっても希望する処理を平易な方法で容易に記述できる環境の実現を意識した。

2.4 提案環境の機能

図 2 に、本研究で開発した開発環境画面を示す。開発画面は大きく 3 つの領域に分けられる。図 2 では、それぞれの領域を①、②、③と表記した。

①はグラフ描画を行う領域である。ユーザが Blockly で指定したグラフが①領域に描画される。

②は、Blockly を用いてカラム操作や計算、グラフ描画やテキスト表示の処理を記述する領域である。ユーザはこの領域上で操作対象カラムを含むブロックを組み合わせることでプログラミングを行う。②領域の上部にある“execute code”リンクをクリックすることで、記述した処理が実行される。グラフ描画には、本研究独自に作成した getGraphplot 内のブロック（この例では“plot line graph”ブロック）と、描画データを選択するブロックを用いる。描画データを選択するブロックでは、グラフの描画範囲を指定する。テキスト表示には、Blockly の print ブロックを拡張したブロックを用いる。なお、Getgraphplot ブロック群と getDataset ブロック群は Blockly のカスタムブロックを拡張することで実現した。

③は解析対象となるデータフレームの確認とカラムブロックの作成、およびプログラムで指定されたデータ表示を行う領域である。③領域の上部にある“ファイルを選択”をクリックして対象データフレームを選択すると、この領域に対象データフレームが表形式で表示される。表示された表のカラム名をクリックすると、カラム情報を持ったブロックが②領域の getDataset ブロック群に生成される。

3. 汎用性評価

本章では、提案環境の特徴の 1 つである対象データフレームの汎用性に関する評価結果を示す。

3.1 実験方法

複数種類のデータフレームを評価データとして用い、提案環境におけるグラフ表示の有無を検証した。

使用したデータフレームは以下の 4 種類である。

- ・慣性計測装置計測データ (dat 形式)
- ・歩数計データ (xlsx 形式)
- ・脳波データ (txt 形式)
- ・心内心電図データ (CSV 形式)

3.2 実験結果

4 種類のデータフレームを用いて実験を行った結果、どのデータフレームもカラムデータの表示は可能であった。しかし、事前に 1 行ヘッダを有する CSV 形式に変換する処理を要した。具体的には、データフレームの 1 行目にカラム名を示すヘッダ行を追加し、データの区切り文字をコンマに統一するという処理である。また、マルチバイト文字の文字コードを UTF-8 に統一する必要があった。

3.3 考察

実験結果より、本提案環境は、データフレームの種類の制限は無いが、形式の制限があることが示唆された。現在は手動で行っている下処理を行うことなく、本提案環境のみで多種類のデータの取り扱いを可能とするためには、ユーザが読み込むファイル形式を指定する機能の追加が求められる。また、読み込まれたデータフレームに対する行と列の追加・削除機能の追加なども求められる。

4. おわりに

本稿では、plain text で表現されたデータフレームを対象に簡易な可視化環境を提案した。提案環境は、web ツールで開発されているため、制限の少ない環境で使用できる。

本提案環境では簡易的なカラム操作や可視化の方法を表現する方法としてブロックプログラミングツールを採用した。対象とするデータフレームの種類は問わないが、取り扱えるデータフレームの形式が限られることが汎用性評価実験から示唆された。

今後は、汎用性を高めるため、取り扱えるデータフレームの種類を拡張することが必要である。また、本研究では汎用性評価のみを行ったため、プログラミングに不慣れなユーザを対象とした、ツールの扱いやすさを評価する評価実験も行ってゆきたい。

参考文献

- (1) 古賀明彦, 谷口洋司, 広瀬正: “表計算用ビジュアルプログラミングシステムのプロトタイプ開発”, 情報処理学会第 50 回全国大会講演論文集, pp333-334(1995)
- (2) Project Jupyter: “Project Jupyter | Home”, <http://jupyter.org/index.html>(2018/02/10 参照)
- (3) Su Inn Park, Frank Shipman.:” PerCon : A personal digital library for heterogeneous data”, 2014 IEEE/ACM Joint Conference on Digital Libraries, pp.97-106 (2014)
- (4) 松澤拓也.:”歩行特徴量をフィードバックする簡易可視化ツールの開発と評価”, 信州大学大学院総合理工学系研究科修士論文 (2016)
- (5) Google Inc: “Blockly | Google Developers”, <https://developers.google.com/blockly/> (2018/02/10 参照)