

ドリトル言語でのデータ処理機能の対応

大阪電気通信大学・兼宗進、小林史弥
 武庫川女子大学・白井詩沙香、イーテキスト研究所・山澤昭彦

教育用プログラミング言語「ドリトル」のデータ分析機能を紹介する。新課程の情報 I、II でデータサイエンスを学習する環境を目的にしている。ドリトルのデータサイエンス機能として、データを加工するデータ処理機能、データを分析する分析機能、結果を表示するグラフ機能を用意した。

1. はじめに

新課程での共通教科情報において、情報 II の「情報とデータサイエンス」では統計的なデータ分析が行われる予定である[1]。表計算ソフトなどを用いた対話的な分析に加え、プログラミングによる分析を可能にするために、教育用言語「ドリトル」[2]にデータ処理機能とデータ分析機能を実装した。

2. ドリトルのデータサイエンス機能

ドリトルは教育用に設計されたプログラミング言語であり、高等学校の「情報の科学」の教科書[3]や教材など [4]で用いられている。

ドリトルには命令を実行した結果に続けて次の命令を送るカスケードと呼ばれる機能が用意されている。これは「かめた！100 歩く 90 左回り。」のような初歩的なプログラムでも使われる機能で、「かめた！100 歩く」の実行結果であるかめた自身に、続けて「90 左回り」という命令が送られる。

今回、ドリトルにデータサイエンス用の機能を検討する際に、ファイルなどから読み込んだ表形式のデータに対して、次々と処理の命令を送る形でデータの処理や分析を行えるようにした。

用意した機能を表 1 に示す。扱うデータはファイルから入力する形にした。現在は表計算ソフト等から出力される CSV 形式に対応している。1 行目は項目名と解釈される。

データ処理としては、sAccess[5]の機能を参考に、整列や関係データベースの選択・射影・結合といった操作を用意した。

データ分析では、代表的な統計関数を用意した。

グラフ描画では、代表的なグラフに対応した。

その他、度数分布表のような表を生成する命令も用意した。

表1 命令の一覧

処理	命令	動作
データ入力	ファイルから作る	データファイルを読み込む
データ処理	選択	行を取り出す
	射影	列を取り出す
	結合	データ同士を結合する
	大きい順	大きい順にソートする
データ分析	小さい順	小さい順にソートする
	合計	合計を求める
	平均	平均を求める
	中央値	中央値を求める
	最大値	最大値を求める
	最小値	最小値を求める
	第1四分位数	第1四分位数を求める
	第3四分位数	第3四分位数を求める
	度数	度数を求める
	標準偏差	標準偏差を求める
グラフ描画	分散	分散を求める
	クロス集計	クロス集計をする
	円グラフ	円グラフを作る
	棒グラフ	棒グラフを作る
	折れ線グラフ	折れ線グラフを作る
	散布図	散布図を作る
表生成	箱ひげグラフ	箱ひげグラフを作る
	描画	グラフを描画する
表生成	度数分布表	度数分布表を生成
出力	表示	結果を表示

3. 使用例

設計した機能の使用例を示す。図 3 は使用したデータの一部である。図 4 にプログラム例を、図 5 に実行結果を示す。

図 4 のプログラムでは、1 行目でファイル

からデータを読み込んでいる。2 行目では使用する項目（横と縦）を指定して箱ひげ図を生成している。この時点ではグラフオブジェクトを作成しただけで、画面の描画は行われない。その後、今回は使用していないが必要に応じてグラフに関する細かい設定を行った後で、3 行目でグラフを画面に表示している。

	A	B	C	D	E
1	Sepal Length	Sepal Width	Petal Length	Petal Width	Species
2		5.1	3.5	1.4	0.2 setosa
3		5.9	3	5.1	1.8 virginica
4		5.6	2.9	3.6	1.3 versicolor
5		4.6	3.1	1.5	0.2 setosa

図 3 読み込んだデータの例(一部)

花=テーブル！"iris.csv"ファイルから作る。
 箱ひげ=花！"Species" "Sepal Length" 箱ひげグラフ。
 箱ひげ！ 描画。

図 4 箱ひげ図を表示するプログラム例

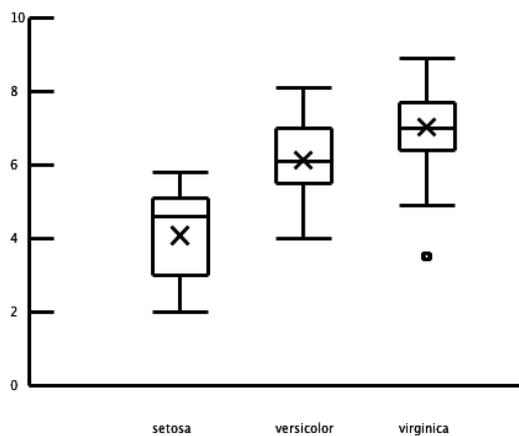


図 5 箱ひげグラフの描画例

4. おわりに

教育用プログラミング言語「ドリトル」でデータサイエンスを扱う機能を紹介した。表計算ソフトと比較して、処理の内容を言葉で明確に記述することができることが特徴である。今後は授業での評価を行いながら、オンラインプログラミング環境 Bit Arrow で動作するドリトルで利用できるように開発を進めたい。

参考文献

- (1) 文部科学省：幼稚園、小学校、中学校、高等学校及び特別支援学校の学習指導要領等の改善及び必要な方策等について（2016）
- (2) 兼宗研究室：プログラミング言語「ドリトル」. <http://dolittle.eplang.jp>
- (3) 東京書籍：情報の科学（2017）
- (4) 日本文教出版：IT・Literacy「Scratch・ドリトル編」（2016）
- (5) sAccess. <http://saccess.eplang.jp/>