

# Moodle データベースからの学習分析: Ad-hoc database queries

## プラグイン と R を用いた解析事例

浅田義和<sup>\*1</sup>

<sup>\*1</sup> 自治医科大学

### Learning Analytics from Moodle Database : Using Ad-hoc Database Queries Plugin and R

Yoshikazu Asada<sup>\*1</sup>

<sup>\*1</sup> Jichi Medical University

Moodle は国内外で広く利用されている LMS である。Moodle における教員および学生の活動履歴はログとしてデータベースに保存されているが、Moodle 上から閲覧できるのはそのごく一部である。今回、Moodle のデータベースに保存されたログを直接解析することで、どのような学習分析が可能となるか検証した。Moodle という共通の LMS を利用して教育実践を行っていくにあたっては、再利用可能な解析コードを用い、効率的な学習分析が実施できるよう、環境を整備していくことが必要と思われる。

キーワード: Moodle, 学習分析, e ラーニング, Ad-hoc database queries, R

#### 1. はじめに

EDUCAUSE Center for Analysis and Research (ECAR) の Working Group によるレポートでは「予測的学習分析 (Predictive Learning Analytics)」として「予測」するための学習分析 (例えば成績不良者の事前予測など) を紹介している<sup>(1)</sup>。この学習分析を行うにあたっては、例として LMS 上での「全活動数」と「成績」との相関などを分析対象としている。学習分析に用いるデータの例として、他にもアクセス回数やアクセス日時、フォーラムの投稿・閲覧件数、課題や小テストの成績などが挙げられる<sup>(2)</sup>。学習分析を行うにあたっては、こうした各種データを必要に応じて LMS から抽出し、分析・可視化等を行う必要がある。

Moodle は国内外で多く活用されている LMS であり、Moodle を用いた授業・研修等の実践は数多くなされている。この Moodle での学習結果について、前述のような学習分析を検討する場合、Moodle 上だけでは分析できない、あるいは分析が非常に煩雑になるという課題がある。例えば「成績」や「活動ログ」など、それ単体では Moodle 上からも閲覧することは可能で

あるが、複数の結果を結合して表示することはできない。また、フォーラムでの投稿件数や相互コメントを付けた相手のリストなどは Moodle 上で確認することは不可能である。

こうした解析の一部は、Moodle のプラグインを活用することで対応可能なものもある。例えば前述したフォーラムの相互コメント相手については Forum Graph<sup>(3)</sup> というプラグインを利用することで可視化・解析が可能である。しかし、プラグインを利用して解析を行うにあたっては、そのプラグインが Moodle 本体のバージョンに対応しているかを常に検討する必要がある。実際、Forum Graph は 2016 年 3 月現在で Moodle 2.7 までの対応となっており、本稿執筆時点での最新版である Moodle 3.0 では利用することはできない。

このような事態を回避する方法として、Moodle のデータベースから直接データを集めて解析するという方策が考えられる。Moodle は MySQL や PostgreSQL といったデータベースに様々な情報を蓄積しており、ユーザがアクセスした画面 (トップページや小テスト

などの活動など)に応じて必要なデータを抽出して表示させている。言い換えれば、Moodle 上で何らかの形で閲覧可能なデータであれば、データベースから直接情報を集めることで集計・解析を行うことが可能であるといえる。

本発表では Moodle のデータベースから得られる情報の分析によってどのような結果が得られるか、その実践結果の一部を紹介し、今後の展望について述べる。

## 2. 方法

### 2.1 Moodle からのデータ収集

Moodle のデータベースにアクセスする方法は複数あるが、ここでは Ad-hoc Database queries プラグイン<sup>4)</sup>を利用する。本プラグインは Moodle 上からデータベースに直接アクセスが可能である。また SELECT 文(データ読み出し)のみ利用可能であり、INSERT (追加)や DELETE (削除)などは利用不可能となっているため、安全に利用することができる。

### 2.2 R での解析

得られたデータについて、R を用いて集計・解析・可視化を行った。本研究では R version 3.2.4 および R Studio version 0.99.887 を利用した。また、一部の可視化には Cytoscape 3.3.0 を利用した。

## 3. 結果

Moodle の活動別に整理する。

### 3.1 小テスト

平均点や標準偏差、多肢選択問題における誤答傾向などは Moodle 上からも教師権限で閲覧可能である。そこでそれ以外の切り口として、コース内に存在する複数の小テストについて「学生が合格点を取るまでの回数分布(図 1)」「学生が小テストを実施した日時の分布(図 2)」をそれぞれ可視化した。

また、Moodle 3.0 より利用可能となった Drag & Drop 形式の小テストについて、学習者全員の実践結果を集約して表示することを試みた(図 3)。Drag & Drop 形式のテストは画像やテキストを正しい位置に Drag & Drop して答える形式の小テストである。Moodle のデフォルト機能では学習者個人が自分の解

答結果を閲覧することはできたが、複数回の解答結果を 1 つに集約して表示することはできなかった。

### 3.2 フォーラム

前述のようにフォーラムの投稿件数や相互コメントの相手などを集計する機能は Moodle の標準機能に含まれていない。そこで単一のフォーラムについて、「投稿の文字数(図 4)」「投稿の時間帯(図 5)」「投稿の曜日(図 6)」を例として可視化した。また、フォーラム上でのコメントについて、「誰が誰にコメントを投稿したか(図 7)」を可視化した。

### 3.3 レッスン

レッスンモジュールでは分岐をもったストーリー型の教材を作成することが可能である。ここでは学習者がどのような分岐でストーリーをたどったかを可視化した(図 8 および図 9)。

## 4. 考察

Moodle での学習分析を行うにあたって、データの切り口としては以下のようなカテゴリが考えられる。

1. 粒度：学年全体、グループ単位、個人単位など
2. ユーザの属性：教師か学生か、性別、学年、グループなど
3. 日時情報：日付・曜日・時刻など
4. 活動の種類：小テスト、フォーラムなど
5. 学習行動の種類：閲覧、投稿など
6. 学習行動の内容：投稿内容、解答結果など
7. 学習行動の結果：正誤判定など
8. 関わった相手：採点者、コメント相手など

これらのデータは Moodle のデータベース上にそれぞれ保存されているため、目的に合わせて必要に応じて組み合わせることで結果を得ることができる。例えば今回はコース全体の小テストに関する受講回数や日時を可視化した。ある特定の学生の受講日時、あるいはある特定のテストの受講日時といった切り口でのデータ収集・可視化も可能である。また、「学習を支援するにあたって、どの活動が多く使われているか」といった視点での解析につながることもできる(図 10)。

学習分析を行うにあたり、個々の施設の特徴(学生数や授業形態)によって、同一の解析方法であっても

異なる結果が得られるため、施設ごとにその背景を考慮した分析を行うことは不可欠である。一方で解析手順については車輪の再発明をする必要はない。特に Moodle という共通の LMS を利用しているのであれば、再利用可能な解析コードを利用し、効率的に学習分析を進めていけるよう環境を整えることも必要であろう。

今回は参考までに、付録として図 1 に示した「小テスト合格までの回数分布」を可視化するために用いた SQL のコード (Ad-hoc database queries プラグイン) および R のコードを添付する。なお、コードは可視化の検証を目的としたものであり、リファクタリングについては検討の余地がある。今後、学習分析を進めるにあたって、利用した SQL や R のコードについては分析ツールとして公開・共有していくことを検討している。

## 5. 今後の課題

今回利用した Ad-hoc database queries プラグインは 5000 行までの結果しか返すことができないという制限がある。このため、大規模なデータ処理を行うには件数が不十分である。また、解析するにあたって一度データを CSV でダウンロードし、R で読み込んだうえで解析を行うという手間が発生している。この作業を簡素化するため、データベース操作の SQL 文および解析用の R コードを入力データとし、得られる解析結果を出力として返すプログラムあるいはプラグインの開発を検討する必要がある。

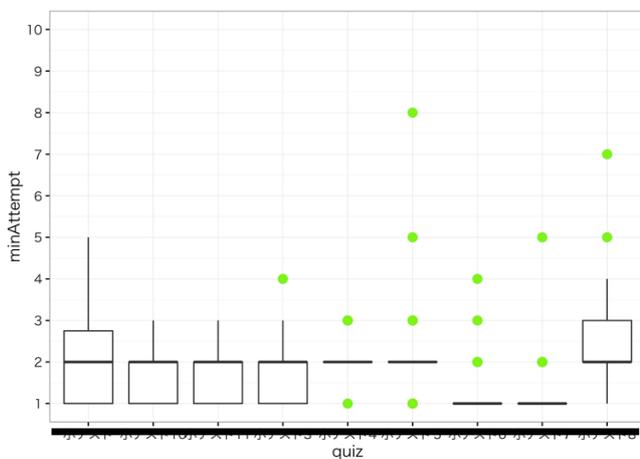


図 1 学生が合格点を取るまでの回数 (小テスト別)

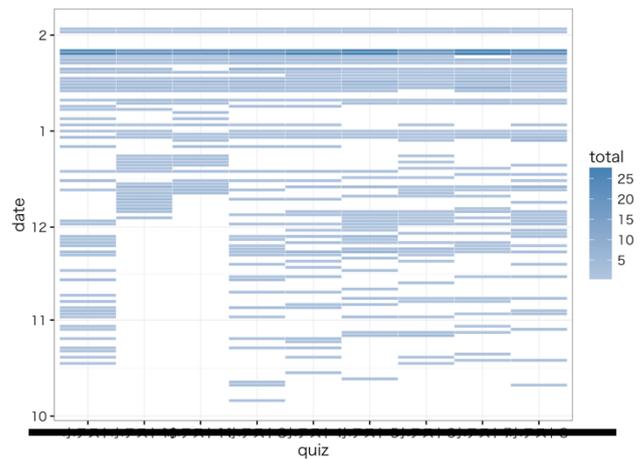


図 2 学生が小テストを実施した日時の分布

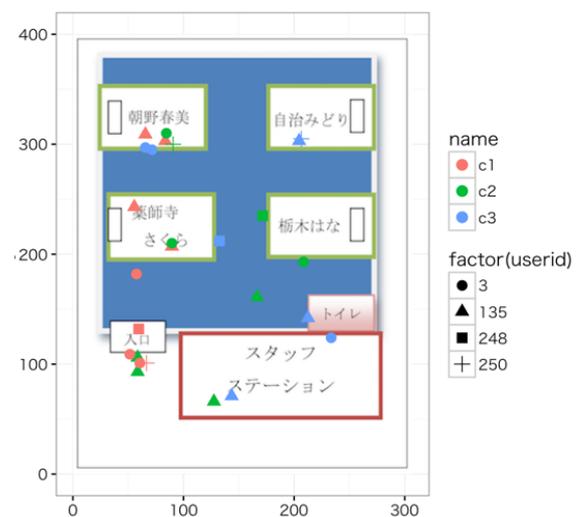


図 3 Drag & Drop テストの結果集約例

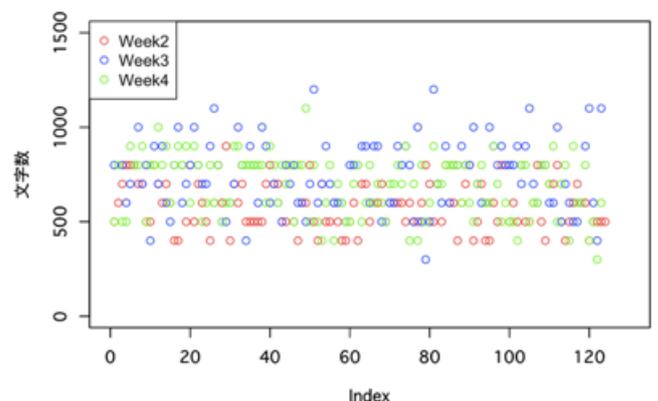


図 4 フォーラムの投稿文字数

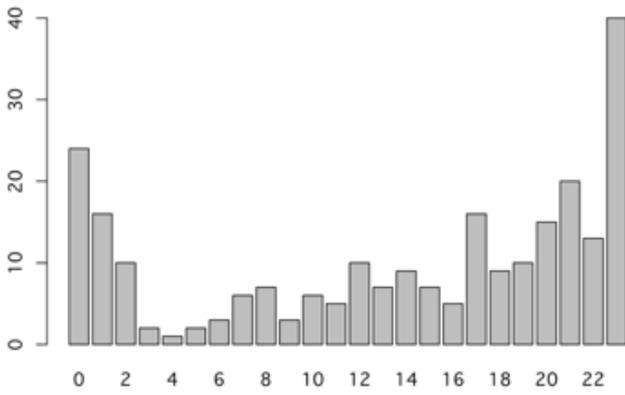


図 5 フォーラムの投稿時間帯



図 8 レッスの画面遷移（全体像）

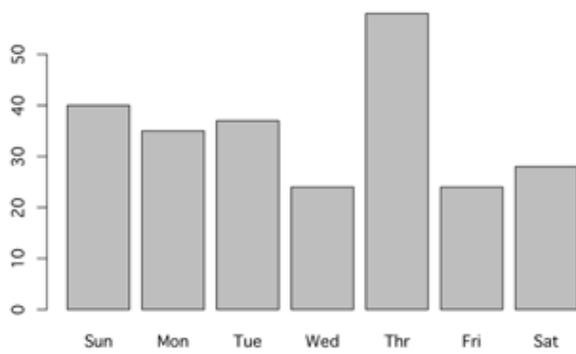


図 6 フォーラムの投稿曜日

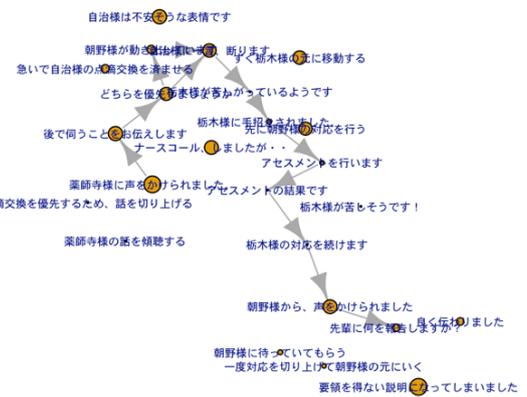


図 9 レッスの画面遷移（学習者の解答結果）

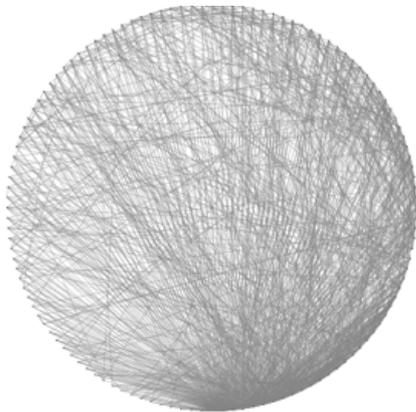


図 7 フォーラムのコメント返信先

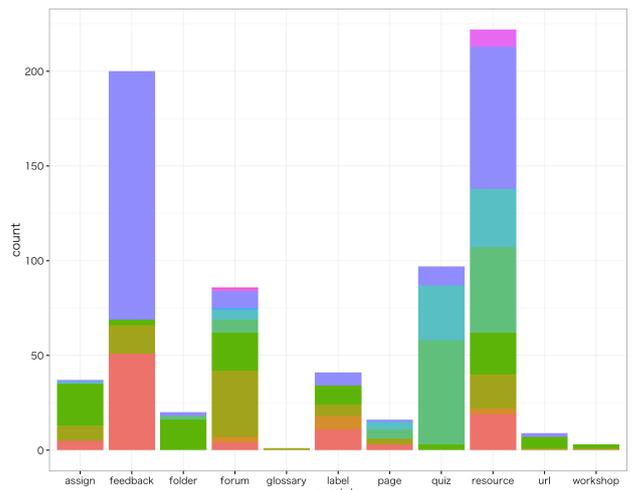


図 10 モジュール別使用状況の可視化

参考文献

(1) “The Predictive Learning Analytics Revolution: Leveraging Learning Data for Student Success”  
<http://net.educause.edu/ir/library/pdf/ewg1510.pdf>

(2016年4月1日確認)

- (2) Beth Dietz-Uhler, Janet E. Hurn: “Using Learning Analytics to Predict (and Improve) Student Success: A Faculty Perspective”. *Journal of Interactive Online Learning*, 12(1), pp.17-26, 2013
- (3) [https://moodle.org/plugins/report\\_forumgraph](https://moodle.org/plugins/report_forumgraph) (2016年4月1日確認)
- (4) [https://moodle.org/plugins/report\\_customsql](https://moodle.org/plugins/report_customsql) (2016年4月1日確認)

## 付録 1: 用いた SQL の例

```
SELECT u.username, qa.userid as id, qa.attempt, q.id
as Qid, qa.sumgrades as QAsumgrades, q.sumgrades
as Qsumgrades, q.grade, q.name as Quiz
FROM prefix_quiz_attempts qa
LEFT JOIN prefix_quiz q ON qa.quiz = q.id
LEFT JOIN prefix_course c ON q.course = c.id
LEFT JOIN prefix_course_categories cc ON c.category
= cc.id
LEFT JOIN prefix_course_modules cm ON ( cm.course
=c.id AND cm.instance = q.id )
LEFT JOIN prefix_user u ON qa.userid = u.id
LEFT JOIN prefix_modules m ON cm.module = m.id
WHERE m.name = 'quiz' AND c.id = :course_id
```

## 付録 2: 用いた R のコードの例

```
library(dplyr)
library(ggplot2)
datatemp <- read.csv("a.csv")
data <- datatemp %>% dplyr::arrange(qid) %>% filter
(qasumgrades == qsumgrades) %>% dplyr::group_by(qi
d, quiz, username) %>% summarize( minAttempt = m
in(attempt))
grp <- ggplot(data, aes(x=factor(qid), y=minAttempt) )
grp <- grp + geom_boxplot(outlier.colour = "green", ou
tlier.size = 3)
grp <- grp + theme_bw(base_family = "HiraKakuPro
N-W3")
plot(grp)
```