Using Assessment System Data to Generate Individualized Learning Materials for Students
What We Learned from Developing the iDAP

Alex Brodersen & Ying Cheng
LAMBS lab, Department of Psychology
University of Notre Dame

A presentation at the 19th MARC conference, Nov. 7, 2019
Overview

1 Background
   - AP-CAT
   - iDAP

2 Lessons learned
   - Teacher Access & Teacher Engagement
   - Plan for Heterogeneity
   - Toolchain
   - Quality Control
   - Process Data

3 Future
   - Learning Modules
Background
Full Name: “Cognitive Diagnostic Computerized Adaptive Testing (CD-CAT) for AP Statistics”

A five-year CAREER project funded by the National Science Foundation

Two major components
  - Research
  - Educational outreach
Item Bank

- 850 items
  - 4 sections
  - 16 main topics
  - 158 attributes
### All Questions

<table>
<thead>
<tr>
<th>#</th>
<th>Question</th>
<th>Creator</th>
<th>Type</th>
<th>Status</th>
<th>Attributes</th>
<th>Time Created</th>
<th>Operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Temperature Boxplot Median Qid: 1</td>
<td>Alison_Cheng</td>
<td>Multiple Choice</td>
<td>Updated</td>
<td>Statistics - 1.C.1.</td>
<td>Jul 28, 2014 - 10:48 p.m.</td>
<td>Review</td>
</tr>
<tr>
<td>3</td>
<td>Temperature Summary Outlier 1 Qid: 3</td>
<td>Alison_Cheng</td>
<td>Blank Field</td>
<td>Updated</td>
<td>Statistics - 1.B.5.</td>
<td>Jul 29, 2014 - 10:14 a.m.</td>
<td>Review</td>
</tr>
</tbody>
</table>
The heights and weights of the 2016 USA Olympic cycling team are graphed below. We can predict an Olympian's weight from their height using the least squares regression equation below:

\[ \text{WEIGHT} = -287 + 6.1 \cdot \text{HEIGHT (in)} \]

Interpret the y-intercept in context.

- **A.** This model predicts that a cyclist who is zero inches tall will weigh around -287 pounds.
- **B.** This model predicts that a cyclist who is zero inches tall will weigh around 110 pounds.
- **C.** This model predicts that a cyclist who weighs zero pounds will be -287 inches tall.
- **D.** This model predicts that a cyclist who weighs zero pounds will be around 110 inches tall.
- **E.** This model predicts that a cyclist who weighs zero pounds will be around 62 inches tall.
Scaffolding

- Step-by-step Solutions
  - Visible after assignment completion
In a random sample of 24 people in an aerobics class, their ages are given as follows:

21 38 32 29 27 27 46 18 24 23 30 31
28 20 35 34 31 29 19 48 28 25 22 33

Explain what the interquartile range means for these data.

- Fifty percent of the time, the age of an aerobics participant is between 23.5 and 32.5 years old.
- The interquartile range is the average range of the ages of participants in this aerobics class.
- The middle 25% of ages of the aerobics participants is between 23.5 and 28.5.
- The interquartile range is the participants' average distance away from the mean age.
- The middle 50% of ages of aerobics participants has a range of 9.
1. Find the first quartile and the third quartile:
   a. Split the data into two halves.
   b. Since there are 24 numbers in the dataset and 24 is an even number, each half will have 12 numbers.
   c. Q1 is the median of the lower half and Q3 is the median of the upper half. Each half has an even number of data points, so the median of the lower half is
      \[ Q_1 = \frac{23 + 24}{2} = 23.5 \]
      and the median of the upper half is
      \[ Q_3 = \frac{32 + 33}{2} = 32.5 \]
Feedback

- Granular feedback
  - Areas of opportunity
Test results: 2017 - Mock Assignment 1 - Form A

Below are displayed the different levels of attribute performance for this assignment:

- Green represents attributes that you are proficient in.
- Yellow represents attributes where your skills are still developing.
- Red represents attributes that you need to work on more.

Comparing distributions of univariate data (dotplots, back-to-back stemplots, parallel boxplots) **INTERMEDIATE**

- Comparing and interpreting outliers: 100.0%
- Comparing and interpreting centers: 50.0%
- Comparing and interpreting spreads: 100.0%
- Comparing and interpreting shapes: 50.0%

Exploring bivariate data **INTERMEDIATE**

- Knowing properties of a Correlation coefficient: 50.0%
- Interpreting coefficient of determination: $R^2$ square: 100.0%
- Writing Least Squares Regression Line (LSRL) from computer output: 100.0%
- Obtaining LSRL from dataset: 100.0%
- Writing LSRL from summary statistics (using $b_1=s_{y|x}$ formula) (not likely): 0.0%
- Interpreting a residual: 0.0%

Exploring categorical data **NOVICE**

- Knowing the Difference between categorical and quantitative variables: 50.0%
- Calculating Joint frequencies for two-way tables: 0.0%
Intelligent Diagnostic Assessment Platform (i-DAP) for High School Statistics Education

- GOAL: Develop a holistic, personalized, learning system integrated into the classroom
Leverage attribute system developed in the AP-CAT

- Map to 36 common core state standards on “statistics and probability”
Lessons learned
Lesson 1

Teacher Access & Teacher Engagement
Teacher Access & Teacher Engagement

- Teacher approval
  - delivery engine
  - reporting
  - feature requests
  - rollout of assignments

  Example: In-Class Assignment Length
Lesson 2

- Plan for Heterogeneity
  - in learning environment
  - in students
Sample Heterogeneity

- Not surprising
  - learning outcomes depend on school
- Surprising
  - Student self-predictions of their final exam scores depend on school (Ober et al., In Prep)
  - Patterns of incorrect responses
Lesson 3

Choose your toolchain wisely
```
con <- DBI::dbConnect(RMySQL::MySQL(), ## Connect to DB
  host = "domain.name.edu",
  port = 3306,
  user = "db_user",
  password = "db_pw",
  dbname = "db_name")

tbl(con, "activity") %>%
  filter(role == "student") %>% ## Only student data
  select(\-firstname, \-lastname, ## de-identify
         \-role, \-username,
         \-id, \-meta)
```
Lesson 4

Have Quality Control and Contingency Plans
Lesson 4

- Not all data are worth being collected
- Not all data that have been collected are worth being analyzed
# Source:  lazy query [?? x 4]
# Database:  mysql 5.7.27-0ubuntu0.18.04.1
#  [user@domain.name:/dname]

<table>
<thead>
<tr>
<th>url</th>
<th>tabname</th>
<th>t</th>
<th>info</th>
</tr>
</thead>
<tbody>
<tr>
<td>/results/461</td>
<td>results(student)</td>
<td>2019-09-25 11:30:28</td>
<td>&lt;NA&gt;</td>
</tr>
<tr>
<td>/results/461</td>
<td>results(student)</td>
<td>2019-09-25 11:30:30</td>
<td>&lt;NA&gt;</td>
</tr>
<tr>
<td>/results/461</td>
<td>student_info_btn</td>
<td>2019-09-25 11:30:34</td>
<td>&quot;btn&quot;:&quot;info&quot;,&quot;qid&quot;:</td>
</tr>
<tr>
<td>/results/461</td>
<td>student_info_btn</td>
<td>2019-09-25 11:30:36</td>
<td>&quot;btn&quot;:&quot;info&quot;,&quot;qid&quot;:</td>
</tr>
<tr>
<td>/results/461</td>
<td>student_info_btn</td>
<td>2019-09-25 11:31:30</td>
<td>&quot;btn&quot;:&quot;info&quot;,&quot;qid&quot;:</td>
</tr>
<tr>
<td>/results/461</td>
<td>student_info_btn</td>
<td>2019-09-25 11:31:38</td>
<td>&quot;btn&quot;:&quot;info&quot;,&quot;qid&quot;:</td>
</tr>
<tr>
<td>/results/461</td>
<td>Attributes(student)</td>
<td>2019-09-25 11:33:20</td>
<td>&lt;NA&gt;</td>
</tr>
<tr>
<td>/results/461</td>
<td>results(student)</td>
<td>2019-09-25 11:33:29</td>
<td>&lt;NA&gt;</td>
</tr>
<tr>
<td>/results/787</td>
<td>results(student)</td>
<td>2019-09-25 11:43:31</td>
<td>&lt;NA&gt;</td>
</tr>
<tr>
<td>/results/787</td>
<td>results(student)</td>
<td>2019-09-25 11:46:55</td>
<td>&lt;NA&gt;</td>
</tr>
</tbody>
</table>

# ... with many, many, many more rows
Lesson 4

- Quality Control
  - Dummy items
  - Response time data (Qualtrics hack)
  - Open text responses to surveys

- Contingency Plan
  - End users are used to having "delete" mean "put in trash"
  - Choosing not to take the exam
System paradata (Process data) can provide a wealth of information
Example: Procrastination

- Active Procrastination (Chu & Choi, 2005)
  - Criticisms (e.g. Delay?) (Krause & Freund, 2014)
- Attempting to validate a survey scale via process data
Data Source

- Data
  - \( \text{lag} = \text{Assignment Deadline} - \text{Assignment Submission Time} \)
  - across 5 Assignments
  - Active Procrastination Scale
- Modeling lag as a behavioral indicator of procrastination
Histogram of Time to Due Date

Lessons learned
Mean vs Standard Deviation

Lessons learned

34
Lessons learned

Diagram showing relationships between Outcome Satisfaction, Preference for Pressure, Intentional Decision, Ability to Meet Deadlines, and Behavioral LV, with arrows indicating directions of influence and correlation coefficients.
## Model Fit

<table>
<thead>
<tr>
<th>Model</th>
<th>cfi</th>
<th>tli</th>
<th>rmsea (5%)</th>
<th>rmsea (95%)</th>
<th>srmr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Procrastination</td>
<td>0.964</td>
<td>0.957</td>
<td>0.074</td>
<td>0.082</td>
<td>0.071</td>
</tr>
</tbody>
</table>
Summary

- Summary
- Limitations
- Future Studies
  - Estimate Latent Procrastination to provide reminders
  - Other applications
To calculate the sample mean, first find the sum of all the values:

$$\sum x_i$$

1. What is the sum?
   - 42
   - 37
   - 64
   + Other
A holistic student view

- Multiple Data Sources
  - Assessment Data
  - Big 5
  - Statistics Anxiety
  - Free Response
  - Help Forum
Thank You
Acknowledgments

https://lambslab.nd.edu/
Questions?

abroders@nd.edu