

This is a data set that is put together that looks at student interactions at the action level for all of the students who completed a particular problem set.

There are 4 files in the data set; `priors`, `exp_slogs`, `exp_plogs`, and `exp_alogs`.

- The **priors** data gives information about who students have done prior to completing given problem set.
- **Action Logs**: looks at students actions while taking the assignment, as in when they clicked next question, entered an answer, asked for a hint etc. This is the biggest of the files (compared to ProblemLog and Student-Logs) in terms of number of rows, as for every action they take they get a new row.
- **Problem Logs**: file share information about the problems that were completed in the problem set, like score and total time on task. This file has one row per student per problems. So if a student does 5 problems, and took 25 actions across those that student would contribute 25 action to the ActionLog but only 5 rows to the problem-Log file.
- **Student Log**: file gives a summary about how students completed the entire problem set. The file has the smallest number, compared to the other files, because it one row per student. We summarize in the files a few key variables, like condition but we loose lots of low level details that are recorded in the files.
- Below is a detailed breakdown of all the columns that are in each of these files

Table of Contents

Priors

Experiment_id: the code for the problem set that is being studied

Student_id: a unique and non identifiable id for students completing work. Student ids will remain the same across files

Class_id: the class in which the student completed these assignments in the past

Class_most_assigned_grade: if the teacher is assigning work from curriculum folders they will belong to some grade. We looked at the problem sets assigned to the class and provided the grade(s) that the teacher assigned most. If the column is null this would mean that the teacher is assigning their own work, or a lot of subsets of problem sets.

Student_prior_median_time_on_task: the median time it takes for a student to complete a problem completely

Student_prior_median_first_response_time_task: the median time it takes for a student to first respond to a problem

Student_prior_average_problem_score: the students average score on problems that have been graded prior to completing the problem set (only problems that the student has completed)

Student_prior_completed_problem_count: out of the problems this student started, this is the total number that they completed

Student_prior_problem_set_started: how many other assignments has the student started before completing this problem set

student_prior_problem_sets_completed: how many other assignments has the student completed prior to this experiment

Student_prior_started_skill_builder_count: how many skill builders has the student started prior to this experiment

Student_prior_average_attempt_count: the average number of times a student attempts a problem before completing it

student_prior_started_skill_builders: how many skill builders has the student started prior to this experiment

student_prior_completed_skill_builders: how many skill builders has the student completed prior to this experiment

Student_prior_viewed_assignment_report_percent: the percentage of reports that the teacher viewed after the student completed some work.

Class_creation_date: the date that the class was originally created

Class_student_count: the number of students that are in the class

Class_prior_median_time_on_task: the median amount of time spent on a problem for a given classroom

Class_prior_median_first_response_time: the median time it takes for a student to enter a response for a problem

class_prior_average_problem_score: the classes average on problems prior to the experiment

class_prior_completed_problem_count: the total number of problems completed by the class prior to the experiment

class_prior_problem_sets_started: the average percentage of the class that started a given assignment that was not a skill builder

class_prior_problem_sets_completed: the average completion rate of students who began an assignment assigned to them

Class_prior_average_attempt_count: the average number of attempts for a class to complete a problem

Class_prior_skill_builder_count: the number of skill builders assigned to this class prior to this experiment

Class_prior_problem_set_count: the number of non skill builder assignments assigned to this class prior to the experiment

class_prior_started_skill_builders: The number of skill builders someone in the class started

class_prior_completed_skill_builders: the number of skill builders some student in the class completed

class_prior_viewed_assignment_report_percent: the percentage of reports viewed by the teacher

Teacher_id: a unique non identifiable id for teachers that is shared across files.

Teacher_account_creation_date: the date the teacher made their account

Teachers_grades_taught: if the teacher filled out their settings, these are the grades that they selected

Teachers_curriculum: if the teacher filled out their settings, these are the curriculums they selected

School_free_lunch_count: if the teacher filled out their settings, this is the number of students in the school who have free lunch

School_reduced_lunch_count: if the teacher filled out their settings, this is the number of students in the school who have a reduced price for lunch

Action Logs:

Experiment_id: the code for the problem set that is being studied

Student_id: a unique and non identifiable id for students completing work. Student ids will remain the same across files

Problem_id: the PR code of the problem the student was working on

Problem_part: the part of the problem the student is working on. Some questions have multiple parts

Scaffold_id: ?

Experiment_tag_path: this tells the researcher the name of the problem set being completed, the condition it is in (pre-test, post-test, treatment-1, control-1, etc) and the version of the experiment

Action: the action the student took at the time (assignment_started, problem_started, correct_response, hint_requested, wrong_response, etc...)

Timestamp: the exact time the action took place

Assistments_reference_action_log_id: the actual action log id if needed to possibly request more data to specific actions later on

Problem Log:

Experiment_id: the code for the problem set that is being studied

Student_id: a unique and non identifiable id for students completing work. Student ids will remain the same across files

Problem_id: the PR code of the problem the student was working on

Problem_part: the part of the problem the student is working on. Some questions have multiple parts

Scaffold_id: ?

Problem_condition: the condition the problem is in (pre-test, post-test, treatment-1, control-1, etc)

Start_time: the exact time the student started the problem

End_time: the exact time the student completed the problem

Session_count: the number of different sessions the student took to complete the problem. ie started the problem, left the problem to complete a different one, or signed out while completing a problem and signed back in to complete it later.

Time_on_task: the difference in time in seconds between the start and end time of the problem

First_response_or_request_time: how long it took the student to make some action (submit a response, ask for a hint, etc) in the problem

First_answer: the first answer given by the student, right or wrong

Correct: whether the first answer the student gave was correct or not

Reported_score: the score the student received for completing the problem

Answer_before_tutoring: whether the student entered a response to the question before asking for a hint, explanation, or scaffold

Attempt_count: the number of attempts the student took to complete the problem

Hints_available: the total number of hints the student could have asked for

Hints_given: the total number hints the student requested while completing the problem

Scaffold_problems_available: the total number of scaffold problems a student could receive

Scaffold_problems_given: the total number of scaffold problems the student was given while completing the problem

Explanation_available: if there was an explanation available for the student

Explanation_given: if the student received an explanation while completing the problem

Answer_given: if the student was shown the answer before submitting the correct response.

Assustnebts_reference_problem_log_id: the unique problem_log_id if the researcher needs more information from later on

Student Log

Experiment_id: the code for the problem set that is being studied

Student_id: a unique and non identifiable id for students completing work. Student ids will remain the same across files

Release_date: the exact time the assignment became available for students to complete

Due_date: the date the assignment is due (sometimes there is no due date)

Start_time: the exact time the student started the assignment

End_time: the exact time the student completed the assignment

Assignment_session-count: how many sessions it took for the student to complete the assignment (reloaded the page while completing)

Pretest_problem_count: the number of pretest problems in the experiment

Pretest_correct: the number of pretest problems answered correctly

Pretest_time_on_task: the total time spent on completing pretest problems

Pretest_average_first_response_time: the average for the first answer for pretest problems

Pretest_session_count: the number of sessions it took to complete complete pretest problems in the assignment

Assigned_condition: the condition the student was assigned assigned to (treatment-1, condition-1, treatment_2, etc..)

Condition_time_on_task: the total time taken to complete problems that are part of the experiment

Condition_average_first_response_or_request_time: the average time for the first hint or answer when completing problems that are part of the experiment

Condition_problem_count: how many problems the student completed that are part of the experiment

Condition_total_correct: the total number of problems that are part of the experiment that were answered correctly

Condition_total_correct_after_wrong_response: how many problems did the student answer correctly after answering the previous question incorrectly (may just be as part of the same problem) for problems that are part of the experiment

Condition_total_correct_after_tutoring: how many problems were answered correctly after the student received help on the previous question (again could be help for this question i'm not sure) for problems that are part of the experiment

Condition_total_answers_before_tutoring: how many problems were answered before asking for hints, explanations, or scaffolding for problems that are part of the experiment

Condition_total_attempt_count: the total number of responses from the student for problems that are part of the experiment

Condition_total_hints_available: the total number of hints available for problems that are part of the experiment

Condition_total_hints_given: the total number of hints the student asked for while completing problems that are part of the experiment

Condition_total_scaffold_problems_available: the total number of scaffold problems that the student could have completed as part of problems that are in the experiment

Condition_total_scaffold_problems_given: the total number of scaffold problems the student received while completing problems in the experiment

Condition_total_explanations_available: the total number of explanations the student asked for while completing problems that are part of the experiment

Condition_total_explanations_given: the total number of explanations the student asked for while completing problems that are part of the experiment

Condition_total_answers_given: the total number of answers the student requested for problems that are part of the experiment

Condition_session_count: the total number of sessions that it took to complete the problems that are part of the experiment

Posttest_problem_count: the total number of problems that are in the posttest

Posttest_correct: the total number of problems answered correctly in the posttest

Posttest_time_on_task: the total amount of time taken to complete the problems in the posttest

Posttest_average_first_response_time: the average amount of time it took a student to enter their first answer to a question in the posttest

Posttest_session_count: the total number of sessions that it took to complete the problems that are part of the posttest

Assistments_reference_assignment_log_id: the assignment log id that researchers can use to possibly ask for more information pertaining to the assignment for a given student.

Additional Notes

- How to read timestamps. Time stamps in the file will look like 2022-05-13 13:27:33.573000+00:00. The first part is the date, the second part is the time on a 24 hour clock. The last +00:00 refers to which time zone it is in. So this time is at GMT. any minus are time zones west of GMT and pluses are time zones east of GMT