

教育大數據微學程教學用開放資料的公開說明

一、 資料背景簡介

本次教育大數據微學程教學用開放資料包含了 397 位使用者 (共 287,137 筆) 在某平臺 (data platform K, dp00k) 上學習時各類操作的紀錄檔 (log)。每一筆資料包含一個匿名處理後的使用者 ID (PseudoID) 與 38 個描述其操作行為的欄位。由於此資料檔經匿名處理，並依照國際評量資料庫資料處理原則產生，基於此資料檔的分析結果與原始資料檔結果可視為統計與測量上相等。唯此資料檔的結果無法用於實務上的現象詮釋，僅供教育大數據微學程課程進行教學與練習使用。

本次教育大數據微學程開放資料來自 dp001 及 dp002 兩個平臺，此兩個平臺的資料各佔三分之一，dp001+dp002 平臺都有的資料佔三分之一。使用者可依據欄位的名稱判斷資料來自哪個平臺，dp001 平臺上的使用者是透過影片學習時的各類行為的操作紀錄，例如：每一次的影片瀏覽活動 (review session) 都包含一個獨特的影片瀏覽序號(欄位名稱：dp001_review_sn)；而 dp002 平臺針對教學教材的操作行為，例如：每一次的操作行為 (object_definition_type) 都包含一個獨特的操作行為 ID(欄位名稱：dp002_object_definition_type)。本次開放資料檔案格式為 Big5 編碼的 CSV 檔案。

二、 資料格式與欄位說明

1. 使用者基本資料

- PseudoID: 使用者流水號

2. dp001:影片基本資料

- dp001_indicator: 能力指標 (格式: X-X-XX-YYY 其中 X 為教育部編列指標，Y 為平臺編列指標)
- dp001_video_item_sn: 影片 ID

3. dp001: 使用者瀏覽影片時的基本資料

- dp001_review_sn: 影片瀏覽序號
- dp001_review_start_timestamp: 影片瀏覽開始的秒數
- dp001_review_end_timestamp: 影片瀏覽結束的秒數
- dp001_review_start_time: 影片瀏覽開始的時間戳記
- dp001_review_end_time: 影片瀏覽結束的時間戳記
- dp001_review_finish_rate: 影片瀏覽的進度
- dp001_review_seme_year: 影片瀏覽學年度

4. dp001: 使用者於瀏覽影片時影片操作的行為資料紀錄

- dp001_record_plus_view_time: 執行影片操作的時間戳記
- dp001_record_plus_view_action: 影片操作的行為名稱(包含：play-播放、paused-暫停、end-結束、normal-恢復正常撥放速度、slowdown-降低播放速度、speedup-提升播放

速度、checkpoint-遇到檢核點、note-停下做筆記、question-學生提問、backward-快轉、forward-倒轉)

- dp001_record_plus_entertime: 影片操作開始秒數
- dp001_record_plus_timestamp: 影片操作結束秒數

5. dp001: 使用者於瀏覽影片時進行檢核點作答時的行為資料紀錄

- dp001_question_sn: 檢核點試題 ID
- dp001_item_plus_item_num: 檢核點試題的題目編號
- dp001_item_plus_question_timestamp: 影片中跳出試題的時間點 (秒)

6. dp001: 使用者於瀏覽影片時進行檢核點作答時的作答資料紀錄

- dp001_exam_video_exam_sn: 檢核點測驗記錄 ID
- dp001_exam_question_timestamp: 跳出試題的影片秒數
- dp001_exam_ans_time: 跳出試題後，學生回答問題花費的毫秒數
- dp001_exam_binary_res: 作答結果 (1 答對，0 答錯)
- dp001_exam_timestamp: 檢核點測驗記錄時間

7. dp001: 使用者於 dp001 平臺上進行練習題作答的行為資料紀錄

- dp001_prac_date: 練習時間戳記
- dp001_prac_during_time: 練習花費時間
- dp001_prac_score_rate: 練習正確率
- dp001_prac_start_time: 練習作答開始時間 (Unix Timestamp)
- dp001_prac_stop_time: 練習作答結束時間 (Unix Timestamp)
- dp001_prac_binary_res: 作答結果 (1 答對，0 答錯)

8. dp002: 使用者於 dp002 平臺上的學習資料行為歷程資料操作紀錄

- dp002_timestamp: 執行操作的時間戳記
- dp002_verb_id: 操作行為的 ID (URL)
- dp002_verb_display_zh_TW: 操作行為名稱
- dp002_object_id: 操作行為的目標對象 ID (URL)
- dp002_object_definition_name_zh-TW: 操作行為目標的名稱
- dp002_object_definition_type: 操作行為目標的類型
- dp002_extensions_alignment: 操作資源的知識架構分類
- dp002_context_platform: 操作平臺
- dp002_classification: 操作行為的簡述
- dp002_result_response: 搜尋相關行為的關鍵字
- dp002_result_duration: 影音相關行為的執行時間