From Data Consumers to Data Makers: An Agency-Based Contract Model for Teaching Data Analytics to Learners Aged 10–17

Heba Mohammed Fadhil[1]

[1] Department of Data Science, College of Excellence, University of Baghdad – Baghdad, Iraq.
Email: heba@kecbu.uobaghdad.edu.iq

Corresponding Author: Heba Mohammed Fadhil[1]

ABSTRACT

Data literacy is now regarded as a civic ability, but in the Arabic-speaking world, data literacy education for learners before university age is still primarily presented as a consumer-oriented skill, and the short-term courses that are offered are more focused on teaching data literacy software features without connection to professional practice. This paper presents the design and implementation of Data Makers Camp which is a 12-hour long program, held in Baghdad, Iraq, in collaboration with Coding for Kids Iraq and College of Excellence, University of Baghdad. The camp restructured the learning space into a network of data agencies: 75 learners between the ages of 10 and 17 were divided into 25 three-person data agencies, each agency having a data engineer, a dashboard designer, and a story narrator. Each agency signed a different client contract that was based on a matrix of five fictional clients and five stakeholder briefs per client, and each agency received one of five parallel datasets that contained the same defects to ensure that the difficulty of the tasks was similar across agencies. The curriculum covered the difference between data and information, to structured data entry, import and clean data in Microsoft Power BI and Power Query, to select charts based on questions, to recognize inaccurate visualizations and finally to publish the dashboard and have a public demonstration day. Evaluation was based on a six criterion, hundred point rubric where teamwork was given weighting the same as presentation, with a five medal system to back up this rubric. All 25 agencies provided their contracted dashboards, 10 projects made it to the closing panel and awarded agencies averaged above eighty-five. The study reports the feasibility of the model under the conditions described and specifies the evidence that a claim of transferability or of learning gain would require.

KEY WORDS: data literacy, project-based learning, authentic assessment, dashboard design, youth education

1. INTRODUCTION

The skills needed to thrive in a data-driven society are not the exclusive domain of specialists. The ability to explain, analyses, interpret, and manipulate artefacts, like dashboards and indicators and automated health and education recommendations, has become a general, rather than technical, literacy (Wolff et al. 2016; Engel 2017). Faced with concerns about who gathered a data set, how and why, established thinking about statistical literacy needs to be expanded to make room for such questions, as they now hold the power to assess the claims made before them (Gould, 2017). So data literacy, in this reading, is not something that would be a natural extension of school mathematics, but a condition for informed participation.

The literature reviewed for this study records no systematic survey of pre-university data education provision in Arabic-speaking countries, and no region-wide claim is made here. In the setting in which this program was delivered, provision outside the formal curriculum consists predominantly of short courses in software operation, in which learners reproduce a sequence of interface actions on a dataset selected for its tractability. The activity terminates in a chart, the chart is graded for correctness, and no stakeholder ever receives it. The absence of a systematic account of provision across the region is itself a gap worth addressing, and the

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

characterization offered here should be read as describing the context of this implementation rather than the region.

Here are two absences that have consequences. The first is the lack of a client. The first step in any analytical work is asking a question by a person who requires an answer that is for a specific purpose, and the role of a professional analyst is to answer that question, not to impress the elegance of the answer. There is no standard to which a dashboard can be compared or evaluated, so long as it is not a client, and the learner is forced to conclude that more visualization, is indeed, better visualization. The second is that there is no defect. Real data are not delivered complete, not duplicated, not consistently typed, and sometimes incorrect; the act of cleaning is not just a preliminary disruption, but the place where most of the analytical judgment is applied (Kjelvik & Schultheis 2019). Datasets designed to be clean do away with that judgment entirely.

The following paper presents a program which focuses on the conscious restoration of both. The Data Makers Camp was held in Baghdad in cooperation with Coding for Kids Iraq and the College of Excellence at University of Baghdad and reconstructed the learning environment as a network of data agencies that are simulated. The participants were not students who were going through exercises, but rather young analysts with named roles in a three-person firm, who had a signed contract with a client with a clear, answerable question, and who had data that was known to contain defects that they were expected to discover and fix.

The study has three questions. First, is it possible to achieve the agency-and-contract structure of learning with learners as young as 10 in a twelve-hour time frame, and with normal institutional resources? Second, what instrument design is needed for such a structure to provide similar assessment at different agencies engaged in broadly similar assignments. Third, the programme's delivery will shed light on the degree to which the model can be transferred to other schools and child-focused organizations within similar constraints.

The contribution is offered as a design case rather than an experimental trial, and the distinction governs what is claimed throughout. The study reports that an instructional architecture was implemented and that the assigned work was completed. It does not report that participants' data literacy improved, since no instrument capable of detecting such improvement was administered, and completion of an assigned task is not evidence of learning. Nor does it establish transferability: the implementation is bounded by a single cohort, a single city, a twelve-hour duration and one software environment, and transfer beyond those boundaries is advanced here as a hypothesis the design makes testable, not as a finding. Section 2 situates the design within scholarship on data literacy, project-based learning and authentic assessment. Section 3 sets out the methodology and the instruments. Section 4 reports outcomes. Sections 5 and 6 interpret the findings and state the evidentiary limits of a first implementation.

2. LITERATURE REVIEW

Theories of this design come together from three sources: first, data literacy as an educational object; second, evidence for project-based instruction; and third, theory of authentic assessment. Both provide a part of the model and the lack of interaction between them is what the present study tries to remedy.

2.1 The application of data literacy in the classroom

It is hard to "settle" on what data literacy actually means. It was discovered by Koltay (2015) from the circulation of this term in library science, information literacy and statistics education and as observed the term is often used to describe competences in which the only common denominator is the shared object. Wolff et al. (2016) then challenge this by arguing that the construct should be built on inquiry, and that data literacy is the ability to interpret phenomena in the world through the use of data, and not as a tool for the execution of technical procedures. What they form have consequences for curriculum design since the real question is the criterion for a data-literate activity. If an exercise yields a correct output chart when no query is entered, then it is not an exercise in data literacy, on this definition.

Ridsdale et al. (2015) provide the other perspective through a synthesis of literature into a competency framework covering data collection, management, evaluation and ethical use. They

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

emphasize the phases that precede and follow analysis and view the presentation of results as a skill as opposed to a postscript. Gould (2017) and Engel (2017) build on the idea of citizenship, suggesting that data literacy is closely related to the concept of civic capability and that it must be developed early on, and before tertiary specialization.

Empirical research in early instruction underpins the possibility of this stance. According to Frischemeier (2020), statistical project work can be performed with and by primary-age students, but only when the statistical data is relevant to something with which they have an interest. Erwin (2015) gets similar results when secondary students are engaged in real datasets, and Konold and Higgins (2003) show how children think about distribution and variation that conventional curricula do not typically elicit. The consistent finding across this body of work is that learners demonstrate the relevant capabilities earlier than conventional curricula assume, and that what limits performance is typically the design of the task rather than the developmental readiness of the learner.

2.2A Project-based learning

The pedagogical apparatus is well established with such task design. Blumenfeld et al (1991) suggest that there are two structural aspects that help maintain motivation during extended projects; both the driving question and the tangible artefact must be authentic: If the learner does not recognize the question as having been contrived, it will not drive anything. Krajcik and Blumenfeld (2006) continue that projects are more likely to be successful as a function of the fidelity of the practices from which they are recruited; Thomas (2000), who has reviewed the literature, warns that the approach is extremely sensitive to implementation quality. Projects where neither the division of labour is specified nor the outcome is defined, consistently underperform.

Lave and Wenger (1991) present the theoretical explanation of why role significance is important. In their formulation, learning is characterized by becoming a member of a community of practice, and their pathway is from periphery to full participation, as they gradually take on the full responsibility. The pedagogical implication is that an apprenticeship structure provides what a classroom structure cannot: an identity in a practice. Both Papert (1980), from constructionist premises, and Wing (2006), within the context of the larger story of computational thinking, agree on this related conclusion: that understanding is durable when it is created as an artefact that is public.

This optimism is qualified in a specific and relevant manner in the cooperative literature. Likewise, Slavin (1996) and Johnson and Johnson (2009) have proved that gains in learning are only likely to result from group work when there is positive interdependence between the group members and when each member is individually responsible. This trend is also observed in inquiry-based environments, according to Barron and Darling-Hammond (2008). This is the main risk that needs to be engineered against for a design that has three learners in one agency.

2.3 Authentic assessment

The idea to let the learner produce a complete performance judged against public criteria and not to answer decontextualized items is advocated by Wiggins (1998) who states that assessment should mimic the conditions in which competence is exercised. Hattie and Timperley (2007) set out conditions under which the feedback that ensues is effective; feedback given on the task and process helps improve performance, whereas feedback given to the person does not. Both arguments are linked to data specifically, with Kjelvik and Schultheis (2019) claiming that authentic data – variable, incomplete, resistant – produce the best learning just by requiring the application of judgement to them that clean data do not.

From this literature, two tensions in designing emerge, which do not get resolved within it. The first one is the tension between authenticity and comparability: the assignments are distinguished enough to feel authentic but not enough to be judged on a common scale; the assignments are standardized enough to be judged on a common scale but not enough to feel authentic. The second is the conflict of collaboration versus accountability discussed above. Both these tensions are poorly considered in the literature on pre-university data education, and the design reported here was developed to investigate a resolution of both..

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

3. METHODOLOGY

The study is a design case study. It documents an instructional architecture ; its component structures, the instruments built to operate it, and the sequence through which it was delivered , in sufficient detail for another practitioner to reconstruct and critique it. What the case examines is whether the architecture can be operated under the stated constraints of age range, duration and resource, and whether its assessment instruments function as designed when applied across a differentiated cohort. What the case can support is a judgement about feasibility and about the internal coherence of the instruments. What it cannot support is any claim about effect: no control condition was constituted, no counterfactual is available, and no measure of change in participant capability was administered. Claims throughout the paper are confined accordingly.

3.1 Setting and participants

The camp was hosted with the College of Excellence at the University of Baghdad in Baghdad, Iraq, by Coding for Kids Iraq. A total of 75 persons (ages 10-17) signed up. There were 6 sessions of two hours each, with a total of 12 hours of contact time. The software interface and the broader professional literature is in English – for this reason, instruction was done in Arabic, with some technical terminology in both Arabic and English. The main analytical environment chosen was Microsoft Power BI Desktop as it is free of license costs, can be run on standard hardware and can be shared with others who are not part of the programme.

Participants were distributed into twenty-five agencies of three members by the organizing team rather than by self-selection, on the principle that self-selected groups reproduce existing friendships and concentrate competence. Composition deliberately mixed ages, so that the span within an agency covered several years wherever enrolment permitted. Clients were then allocated to agencies by a randomized draw conducted on screen, which distributed the five client datasets evenly across the twenty-five agencies and removed any perception that assignments had been chosen for particular groups. Within each agency the three roles were allocated by the members themselves during the first session, with instructor guidance where an agency could not agree; roles were fixed once assigned and were not rotated, since the twelve-hour duration did not permit a participant to develop competence in more than one. In each case a member had a named role: a role that was described in Table 1, and that determined the main responsibility of the member, but did not exclude the other work of the agency.

Table 1. Agency Roles and Primary Responsibilities

Role	Primary responsibility
Data engineer	Importing the dataset, diagnosing defects, and performing cleaning operations in Power Query
Dashboard designer	Selecting chart types appropriate to the contract question, laying out the dashboard, and configuring filters and interactivity
Story narrator	Formulating the answer to the client's question, drafting the recommendation, and delivering the presentation to the panel

Note. Roles defined primary responsibility only; all three members were accountable for the delivered dashboard.

3.1.1 Age range and design accommodations

The ten-to-seventeen range was determined by the partner organization's enrolment rather than selected for research purposes, and it spans substantial differences in reading fluency, abstraction, keyboard competence and tolerance for unstructured problems. Three features of the design responded to this spread. First, agencies were composed to mix ages rather than to group participants of similar age, so that the older members of an agency carried the parts of the work requiring sustained abstraction while younger members were not isolated in a group operating beyond them. Second, the three roles differ in their cognitive demand as well as their content: the data engineer role proceeds through a defined diagnostic sequence, whereas the story narrator role requires formulating a recommendation from evidence, and the allocation of roles within an agency therefore permitted entry at different levels of demand. Third, every session was accompanied by a printed step-by-step booklet in Arabic, so that a participant who lost the thread of a live demonstration could recover it without waiting for an instructor.

Prior experience with analytics software was not assessed. A pre-camp readiness form was collected from families, but it addressed device availability and operating system rather than the participant's prior competence, and the study therefore cannot

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

report the distribution of entering ability or control for it. This is recorded among the limitations in Section 5.3.

3.1.2 Tool selection and its trade-offs

Power BI Desktop is a professional instrument, not an educational one, and Bhargava and D'Ignazio (2015) argue persuasively that tools built for users and tools built for learners embody different design priorities; the former optimise for rapid output, the latter for the visibility of the underlying process. Their DataBasic suite (D'Ignazio & Bhargava 2016) represents the alternative we did not take. The decision to use professional software was therefore a deliberate trade and is stated as such.

Three considerations governed it. The first is continuity of artefact: a dashboard published from Power BI is the same kind of object a working analyst produces, which is what permits the contract to function as a genuine deliverable rather than a classroom fiction. A purpose-built learner tool would have weakened precisely the authenticity the design depends upon. The second is cost and availability: Power BI Desktop carries no license fee and runs on the hardware families already own, which was a binding constraint in this setting. The third is continuation: participants retain access to the environment after the camp ends, which a hosted teaching tool with institutional provisioning would not have provided.

The cost of this choice is interface complexity, and the design absorbed it through sequencing rather than by assuming it away. The tool was not introduced until the second session, and the first dataset each participant imported was one they had generated themselves in Session 1, so that the unfamiliarity of the interface was not compounded by unfamiliarity of the data. Power Query operations were rehearsed on a simplified simulator before being performed in the live application. Printed Arabic booklets carried each operation as a numbered sequence. We nonetheless make no claim that the interface is developmentally appropriate for the full range of participants, and whether younger participants required disproportionate instructor support was not measured. Section 5.3 records this as a limitation and Section 5.4 identifies it as a question for replication.

3.2 The contract matrix

The client contract was the main design tool used. Five fictional clients were built — a school library, a sports club, a juice café, a toy shop and a veterinary clinic — each of which was selected because it was familiar to the age group and because it was possible to think of a real question that could be posed about it based on the operations of the business. Five different stakeholder briefs were authored for each client on a common business question. The outcome was 25 contracts (one for each agency), none of which were duplicates.

The following were included in each contract: The client, the stakeholder, the question to be answered, the data provided, and the deliverable needed. Contracts were made and signed by the three members of the agency and kept by the agency for the entire camp period. The signature wasn't just an afterthought: The contract was the basis on which the agency's final dashboard was measured, and an agency that made an elegant dashboard that answered a question besides its own was rated on that contract.

Table 2. Structure of the Contract Matrix

Client	Basis of the five stakeholder briefs
School library	Borrowing patterns viewed from the perspectives of the librarian, the school administration, the teaching staff, the student council and the supplier
Sports club	Attendance, membership and scheduling questions posed by the coach, the manager, the membership officer, the equipment officer and the treasurer
Juice café	Product mix, peak hours and stock questions posed by the owner, the shift supervisor, the purchaser, the marketer and the accountant
Toy shop	Seasonal demand, category performance and pricing questions posed by five distinct commercial stakeholders
Veterinary clinic	Visit frequency, case category and capacity questions posed by the veterinarian, the receptionist, the pharmacy officer, the owner and the auditor

Note. Five clients × five stakeholder briefs = 25 contracts, one per agency, none duplicated.

3.3 Parallel datasets and planted defects

Suppose that twenty-five different agencies are all given different questions: one agency given an easy question, another an impossible one; then there is an

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

obvious issue of comparability: the score obtained is a measure of allocation, not of performance. This was overcome by creating five datasets, all within the same system, but each representing a different client, and each differing in domain, but each having the same number and types of columns, and — most importantly — the same defects intentionally seeded into them, but different in content.

The same "defect profile" was present in each dataset: missing data on a mandatory analysis field, duplicate records, at least one column with data type stored incorrectly, and inconsistent data labelling for categories, caused by non-standardized manual data entry. So the defect profile was the same in each data set and the amount of cleaning regardless of client drawn was the same so that the differences in outcome were due to the agency but not to the allocation. No announcement was made of the defects. In line with the argument put forward by Kjelvik and Schultheis (2019), their discovery was included within their assessed work.

3.4 Curriculum sequence

The six sessions were ordered in such a way that the capability needed for the next level of contract work was introduced in each session as described in Table 3. The sequence deliberately chooses to introduce the analytical tool as a participant has produced data of their own so that the first data set the participant imports in to Power BI is one that they have produced and therefore know about.

Table 3. Curriculum Sequence Across the Six Sessions

Session	Focus	Principal activity
1	Data and information	Distinguishing raw data from interpreted information; collecting data from peers in the room; structured entry into a spreadsheet under professional table conventions
2	Import and first visual	Introduction to Power BI Desktop; importing the participant's own file; constructing a first chart from data the participant generated
3	Cleaning	Power Query: diagnosing and repairing missing values, duplicates, incorrect types and inconsistent categories in the client dataset
4	Chart selection and deception	Matching chart type to question family; a deception laboratory in which participants manipulate axis, scale and selection to produce misleading visuals, then reverse the manipulation
5	Dashboard construction	Distinguishing a collection of charts from a dashboard; layout, filters, cross-highlighting, and publication of the contracted deliverable
6	Demonstration day	Gallery round in which all agencies exhibit; panel presentations by finalist agencies; awards and closing

Note. Each session ran for two hours. Total contact time: twelve hours.

The deception lab in Session 4 is an interesting case of the norm of instruction being reversed. Instead, participants weren't given a list of tricks to avoid, but asked to create their own misleading charts, for example by truncating a vertical axis, by choosing a favorable time window, by changing the aspect ratio, etc., and then say what they found the chart to be telling their viewer. The activity is based on the principles of treating the critical reading of visual evidence as an active rather than a receptive capacity found in Tufte (2001) and Cairo.(2019)

3.5 Assessment instruments

Two instruments governed assessment. The first was a six-criterion rubric scored out of one hundred points, presented in Table 4. Its allocation encodes the design's priorities: depth of discovery carries twenty-five points and fidelity to the contract question twenty, so that analytical substance accounts for forty-five of the hundred, while design quality carries twenty, decision and recommendation fifteen, and presentation and teamwork ten each. Fidelity to the contract is thus weighted equally with design quality and below depth of discovery, on the reasoning that a dashboard which is visually accomplished but answers a question other than the client's constitutes a professional failure rather than a partial success.

Table 4. Six-Criterion Judging Rubric (100 Points)

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

Criterion	Points	What it captures
Depth of discovery	25	Defects identified and repaired; non-obvious patterns surfaced from the data
Answer to the contract	20	Whether the dashboard answers the specific question the stakeholder asked
Design quality	20	Chart selection, layout, legibility, absence of misleading representation
Decision and recommendation	15	A stated recommendation the client could act upon, grounded in the data
Presentation	10	Clarity and structure of the oral delivery to the panel
Agency teamwork	10	Evidence of distributed contribution across all three roles

Note. Teamwork carries weight equal to presentation, by design.

The second instrument was a five-medal system awarded independently of the rubric total, covering discovery, design, narration, technical cleaning and collaboration (see Table 5). The collaboration medal was weighted equal to the design medal in its public standing. This was a direct response to the risk identified by Slavin (1996) and Johnson and Johnson (2009): where visible reward attaches only to the most conspicuous output, a single capable member will absorb the work of the group. Attaching equivalent recognition to the act of supporting a The struggling agency, that is: an agency other than the one's self, was to be used to make distributed contribution rational, not just good.

Table 5. Medal Categories and Award Criteria

Medal	Criterion
Discovery	The most consequential finding surfaced from the data that the client had not anticipated in the contract brief
Technical cleaning	The most complete diagnosis and repair of the planted defects, judged on the Power Query steps retained in the file
Design	The clearest dashboard, judged on chart selection, legibility and absence of misleading representation
Narration	The clearest delivery of the answer and recommendation to the panel, judged on structure and grounding in the displayed evidence
Collaboration	Sustained support given to other members or to another agency, judged on instructor observation across the six sessions; equal in standing to the design medal

3.5.1 Monitoring of individual contribution

Three provisions were intended to counter the stratification that cooperative learning research predicts in small groups (Slavin 1996; Johnson & Johnson 2009): named roles carrying distinct primary responsibilities, a rubric criterion allocating ten points to evidence of distributed contribution, and a collaboration medal equal in public standing to the design medal.

Their operation was monitored by instructor observation during sessions and by the panel's questioning at the demonstration day, where questions were directed to specific role-holders rather than to the agency as a whole, so that a member who had not participated in the work could not be concealed behind a spokesperson. No structured observation protocol, contribution log or individual artefact was used, and the ten teamwork points were therefore awarded on judges' holistic assessment rather than on recorded evidence. It follows that the study cannot distinguish agencies in which contribution was genuinely distributed from agencies in which a capable member carried the work while satisfying the panel that they had not. This is the principal evidentiary gap in the present implementation, and Section 5.4 specifies the instrumentation a replication would require to close it.

3.6 Demonstration day protocol

The final session adhered to a standardized format: 20 minutes of set-up, 45 minutes of the gallery round (with all agencies on display and questions from their peers and judges), 30 minutes of the agencies' presentations in front of the panel (finalists) and 25 minutes of awards and closing. The gallery round was also significant as it ensured that each agency would show to an audience, regardless of whether they got to the finalist round. No agency ever had its projects end without any impact.

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

3.7 Ethical considerations

All participants were minors, and participation was arranged on a dual basis. Written consent was obtained from a parent or guardian at enrolment, following written information supplied by the partner organization describing the programme's content, duration and requirements. Separately, each participant's own assent was sought at the first session in terms appropriate to their age, and participants were informed that they could withdraw from any activity, or from the program, without consequence to themselves and without a reason being required. Enrolment was voluntary throughout; no participant was placed in the program as a course requirement or an institutional assignment.

All datasets were synthetic and all clients fictional, so no personal data belonging to third parties were processed at any stage. Data that participants collected from one another in the first session remained on the participant's own device, were used only within that session's activity, and were not transferred to or retained by the organizers. No participant is identified by name, image or identifiable detail in any analysis, presentation or publication arising from the program, and no individual-level record was retained after the event..

4. RESULTS / ANALYSIS

Results are presented at the stage where there was systematic evidence retention: delivery against contract, progression to final round, and rubric performance in awarded band. The boundaries of this evidence are clearly defined in Section 5.

4.1 Delivery against contract

Each of the twenty-five agencies finished their delegated work, providing a dashboard that contained the answer to the question that had been given by the agency's stakeholder. None of the agencies dropped out, changed the question or were unable to deliver within the twelve hour envelope. The completion rate was thus 100 per cent across agencies and had no drop-outs in the seventy-five participants that were enrolled through the final session. Table 6 summarizes delivery.

Table 6. Delivery Outcomes

Indicator	Value	Basis
Agencies formed	25	Three members per agency
Participants enrolled	75	Ages 10–17
Contracts issued	25	No duplication across agencies
Dashboards delivered	25	100% of contracts
Agencies advancing to panel	10	40% of agencies
Rubric score, awarded agencies	> 85	Out of 100 points

Note. Figures refer to the single delivered cohort. No comparison cohort was constituted.

Completion is reported here as an indicator of the feasibility of the task design under the stated conditions. It is not evidence that participants acquired capability they did not possess beforehand, and no inference about learning should be drawn from it.

4.2 Progression and Rubric.

Assessment proceeded in two stages. During the gallery round, every agency exhibited its dashboard and was questioned by the circulating panel, which recorded a holistic judgement against the six rubric criteria for each of the twenty-five agencies. Ten agencies were advanced to the panel session on the basis of these judgements. In the second stage the ten finalist agencies presented formally and were scored in full on the hundred-point rubric; awards were determined from these scores, and all awarded agencies scored above eighty-five.

Complete scores are therefore available only for the ten finalist agencies. For the remaining fifteen, the gallery-round judgements were recorded as selection decisions on the judges' working sheets and were not transcribed into a retained scoring record before those sheets were discarded at the close of the event. The study can consequently report the full rubric distribution for the finalist group and the fact of selection for the cohort, but it cannot report a distribution across all twenty-five agencies. This was a failure of data retention in the event design rather than a decision to withhold results, and the protocol for any replication now specifies that gallery-round

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

scores be entered on a retained instrument before the finalist round begins.

Within the finalist group, scores above eighty-five are not attainable through visual polish alone. Because depth of discovery and fidelity to the contract question together carry forty-five of the hundred points, a total in that band requires both that the agency located and repaired the planted defects and that its dashboard answered the question its own stakeholder had posed.

4.3 Observations on the instruments

Two items of design were observed as doing their designed function and are not measured as such but reported here. The first was that the parallel-dataset construction was maintained, since each agency had a similar defect profile, the panel could use a single rubric for the twenty-five nominally different assignments without adjusting for difficulty of the assignments. Second, the contract served as an assessment anchor. The sufficiency of a dashboard was determined by reference to a document signed by the agency, which eliminated the ambiguity that is inherent to open-ended project work and made the criterion visible to the participants at the first session, but not at the time of judgement.

Note that there was no pre-instruction measure and that the scores of non-awarded agencies were not saved in a way that allowed a distributional analysis to be performed. The programme therefore provides evidence of delivery and evidence of a level of performance in the upper band and does not provide evidence of learning gain. This constraint will be addressed in the next section.

5. DISCUSSION

This research outcome is that an agency and contract structure is feasible with as young as ten years old learners, within twelve hours of contact time, without any special infrastructure and generates overall complete delivery in all the participating groups. This is important because the learner is typically thought to be the barrier to authentic data instruction at this age. The delivery in this report places the barrier somewhere else: in the design of the task and the instruments that evaluate it.

5.1 Resolving the authenticity–comparability tension

The literature in Section 2 does not leave us with a clear resolution of the dilemma between assignments which are authentic enough to motivate and assignments which are standardized enough to assess. This contract matrix and parallel defect matched data sets provide a viable solution. The differentiation is represented on the surface of the task, that is, on the client or stakeholder, domain or question, and the analytical load under the surface is fixed by construction. The rubric allocation reinforces this: fidelity to the contract question carries twenty of the hundred points, equal to design quality, and together with the twenty-five points for depth of discovery it places analytical substance above presentation in the determination of success.

This builds on Blumenfeld et al. (1991's) stance on driving question. While their account uses the question as a motivational tool, the present design uses the question as an assessment tool in addition. The criterion for adequacy is pre-determined in the contract, and recorded in writing, so the criterion is fixed before work starts and is available to the learner all the time. The rubric weighting validates this: Fifteen points are given for fidelity to the contract question, the same as are given for design quality, so that answering the wrong question well is not a 50% success.

5.2.The free-rider condition

The cooperative learning literature suggests that, in a group of 3 people working to a common artefact that is visible, the most talented one will absorb the work (Slavin 1996; Johnson & Johnson 2009). The design's answer was threefold: named roles with clearly defined primary roles, a rubric criterion that gives ten points to evidence of distributed contribution, and a collaboration medal that is as public as the design medal. The mix is designed to create distributed contribution instrumentally rational, not based on the good will of the strongest member. The theoretical basis is Lave and Wenger (1991) who state that the “vehicle of participation” is role, not a “slot in a group”.

The evidence presented here does not demonstrate the effectiveness of the provision. Twenty-five agencies for delivery is distributed

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

contribution and it also is capable people carrying agencies. This is the most critical of the study's evidentiary gaps, as individual-level observations of these are necessary to discriminate between them.

5.3 Limitations

There were three restrictions on the statements made here. First, and most importantly, no pre-instruction measure was used, and thus no learning outcome can be reported for the study. Rubric performance in completion is an indicator of what the cohort produced, rather than that participants couldn't have produced it beforehand. Second, the scores of non-awarded agencies were not preserved in an analyzable form, and so quantitative results cannot be made for the entire cohort; only an insider could make a statement about the distribution of performance across the cohort since only awarded agencies scores were retained in a form that allowed such analysis. Third, the program was implemented in a single city, to a single cohort that self-selected via a partner, and delivered once. The findings should not be interpreted as a generalization to compulsory settings as participants are not representative of the age group and they have volunteered to participate in a voluntary analytics camp.

These limitations are properties of a first implementation rather than of the model, and the instruments set out in Section 3 require no modification to support a study capable of addressing them. Such a replication would administer a matched pre- and post-instruction instrument assessing chart interpretation, defect identification and the reading of misleading visualizations; would retain the full rubric distribution across every participating agency rather than the finalist group alone; would record contribution at the individual level through a structured observation protocol; and would include a delayed follow-up some months after the programme, since retention rather than immediate performance is the outcome of interest for a civic competence. Until such a study is conducted, the present findings bear on the feasibility of the design and not on its effect.

5.4 Implications

In the case of schools and child-oriented organizations in Iraq and similar environments, the impact of limitation of authentic data instruction is not equipment. It was set up on open source software and standard equipment, it simply needed some contracts in writing, some data sets that were defect matched and a rubric given a tilt towards discovery rather than appearance. These are one-off authoring costs, and can be used across cohorts. What the present implementation establishes is that these components can be authored and operated within the stated constraints. Whether the model transfers to compulsory classroom settings, to larger cohorts, to instructors who did not design the materials, or to curricula with different time allocations is not established here and would require implementation under those conditions. The design is offered as a candidate for such testing rather than as a validated solution.

6. CONCLUSION

The purpose of this study was to answer the questions: What kind of a contract model can provide authentic data analytics instruction to learners ages ten to seventeen; What instrument design is needed for delivery in an agency-based contract model; and What it means to deliver in such a contract model to transfer to similar settings. In 12 hours of contact time the 75 participants in 25 different agencies created 25 unique client contracts, and 10 of those agencies were able to move on to a public panel, and awarded agencies that reached scores above eighty-five on a rubric skewed toward analytical discovery over visual presentation.

The contribution is of an architectural nature. A differential contract matrix combined with structurally aligned, defect seeded datasets bridges a tension that is left unaddressed by the current literature and allows for assignments that are truly different to be evaluated by one standard that can be defended. The contract is also used as a motivator and as an assessment anchor and the emphasis on discovery over fidelity in the rubric makes it clear that answering the client's question is the measure of success. The infrastructure cost is low, the cost of the authoring is paid only once and it can be transferred

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

to any domain where plausible client scenarios can be written.

The study does not provide evidence of learning gain. No pre-instruction measure and no retained score distribution limits evidence to delivery and upper-band performance; and will not distinguish between contribution distributed and capable individuals carrying their agencies. These are the questions that any second implementation should answer and the instruments set out here are adequate to support a second implementation without change. Despite that evidence, the overall take-home message is one that is independent: The region's pre-university students are trapped in analytical software programs, not by their performance, but by the way the work is presented.

REFERENCES

Barron, B. & Darling-Hammond, L. 2008, 'Teaching for meaningful learning: a review of research on inquiry-based and cooperative learning', in Powerful Learning: What We Know About Teaching for Understanding, Jossey-Bass, San Francisco, pp. 11–70.

Bhargava, R. & D'Ignazio, C. 2015, 'Designing tools and activities for data literacy learners', in Workshop on Data Literacy, Web Science 2015, Oxford.

Blumenfeld, P.C., Soloway, E., Marx, R.W., Krajcik, J.S., Guzdial, M. & Palincsar, A. 1991, 'Motivating project-based learning: sustaining the doing, supporting the learning', Educational Psychologist, vol. 26, no. 3–4, pp. 369–398.

Cairo, A. 2019, How Charts Lie: Getting Smarter about Visual Information, W.W. Norton, New York.

D'Ignazio, C. & Bhargava, R. 2016, 'DataBasic: design principles, tools and activities for data literacy learners', The Journal of Community Informatics, vol. 12, no. 3, pp. 83–107.

Engel, J. 2017, 'Statistical literacy for active citizenship: a call for data science education', Statistics Education Research Journal, vol. 16, no. 1, pp. 44–49.

Erwin, R.W. 2015, 'Data literacy: real-world learning through problem-solving with data sets', American Secondary Education, vol. 43, no. 2, pp. 18–26.

Frischemeier, D. 2020, 'Building statisticians at an early age: statistical projects exploring meaningful data in primary school', Statistics Education Research Journal, vol. 19, no. 1, pp. 39–56.

Gould, R. 2017, 'Data literacy is statistical literacy', Statistics Education Research Journal, vol. 16, no. 1, pp. 22–25.

Hattie, J. & Timperley, H. 2007, 'The power of feedback', Review of Educational Research, vol. 77, no. 1, pp. 81–112.

Johnson, D.W. & Johnson, R.T. 2009, 'An educational psychology success story: social interdependence theory and cooperative learning', Educational Researcher, vol. 38, no. 5, pp. 365–379.

Kjelvik, M.K. & Schultheis, E.H. 2019, 'Getting messy with authentic data: exploring the potential of using data from scientific research to support student data literacy', CBE—Life Sciences Education, vol. 18, no. 2, es2.

Koltay, T. 2015, 'Data literacy: in search of a name and identity', Journal of Documentation, vol. 71, no. 2, pp. 401–415.

Konold, C. & Higgins, T.L. 2003, 'Reasoning about data', in J. Kilpatrick, W.G. Martin & D. Schifter (eds), A Research Companion to Principles and Standards for School Mathematics, National Council of Teachers of Mathematics, Reston, VA, pp. 193–215.

Krajcik, J.S. & Blumenfeld, P.C. 2006, 'Project-based learning', in R.K. Sawyer (ed.), The Cambridge Handbook of the Learning Sciences, Cambridge University Press, New York, pp. 317–334.

Lave, J. & Wenger, E. 1991, Situated Learning: Legitimate Peripheral Participation, Cambridge University Press, Cambridge.

Papert, S. 1980, Mindstorms: Children, Computers, and Powerful Ideas, Basic Books, New York.

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin, S. & Wuetherick, B. 2015, Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report, Dalhousie University, Halifax.

Slavin, R.E. 1996, 'Research on cooperative learning and achievement: what we know, what we need to know', Contemporary Educational Psychology, vol. 21, no. 1, pp. 43–69.

Thomas, J.W. 2000, A Review of Research on Project-Based Learning, Autodesk Foundation, San Rafael, CA.

Tufte, E.R. 2001, The Visual Display of Quantitative Information, 2nd edn, Graphics Press, Cheshire, CT.

Wiggins, G. 1998, Educative Assessment: Designing Assessments to Inform and Improve Student Performance, Jossey-Bass, San Francisco.

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]

پوختە

ئێستا خوێندەواری داتا وەک توانایەکی مەدەنی سەیر دەکرێت، بەڵام لە جیهانی عەرەبی زماندا، پەروەردەی خوێندەواری داتا بۆ فێرخوازان پێش تەمەنی زانکۆ هێشتا بە پلەی یەکەم وەک کارامەییەکی بەکاربەر-ئاڕاستەکراو پێشکەش دەکرێت، و ئەو خولە کورتخایەنانەی کە پێشکەش دەکرێن زیاتر سەرنجیان لەسەر فێرکردنی تایبەتمەندیەکانی نەرمەکاڵای خوێندەواری داتا بەبێ پەیوەندی بە پراکتیکی پیشەییەوەیە. ئەم توێژینەوەیە دیزاین و جێبەجێکردنی کەمپی داتا مەیکەرز دەخاتە ڕوو کە بەرنامەیەکی ١٢ کاتژمێرییە، کە لە شاری بەغدای عێراق بە هاوکاری کۆدینگ بۆ منداڵانی عێراق و کۆلێژی نایابی زانکۆی بەغدا بەڕێوەدەچێت. کەمپەکە فەزای فێربوونەکەی داڕشتەوە بۆ تۆڕێک لە دەزگاکانی داتا: ٧٥ فێرخواز لە نێوان تەمەنی ١٠ بۆ ١٧ ساڵدا دابەشکرابوون بەسەر ٢٥ دەزگای داتا سێ کەسیدا، هەر دەزگایەک ئەندازیاری داتا و دیزاینەری داشبۆرد و گێڕەرەوەی چیرۆکی هەبوو. هەر دەزگایەک گرێبەستێکی جیاوازی موکلیفی واژۆ کرد کە لەسەر بنەمای ماتریکسێکی پێنج موکلیفی خەیاڵی و پێنج کورتەنامەی لایەنە پەیوەندیدارەکان بوو بۆ هەر موکلێک، هەر دەزگایەک یەکێک لە پێنج کۆمەڵە داتا هاوتەریبەکانی وەرگرت کە هەمان کەموکوڕییان تێدابوو بۆ دڵنیابوون لەوەی کە سەختی ئەرکەکان هاوشێوەیە لە سەرانسەری دەزگاکاندا. مەنهەجەکە جیاوازی نێوان داتا و زانیاریی گرتەوە، تا داخڵکردنی داتا بە ستراکتۆر، هاوردەکردن و پاککردنەوەی داتا لە مایکرۆسۆفت پاوەر بی ئای و پاوەر کوێری، هەڵبژاردنی چارتەکان لەسەر بنەمای پرسیارەکان، بۆ ناسینەوەی بینینی نادروست و لە کۆتاییدا بۆ بڵاوکردنەوەی داشبۆرد و هەبوونی ڕۆژی نمایشی گشتی. هەڵسەنگاندن لەسەر بنەمای ڕوبریکێکی شەش پێوەر و سەد خاڵ بوو کە کاری تیمی بە هەمان قورسایی پێشکەشکردن درا، لەگەڵ سیستەمی پێنج میدالیا بۆ پشتگیریکردنی ئەم ڕوبریکە. هەر ٢٥ دەزگاکە داشبۆردە گرێبەستکراوەکانیان دابین کرد، ١٠ پڕۆژە گەیشتنە پانێڵی داخستنی و خەڵاتیان بەو دەزگایانە بەخشی کە تێکڕایان لە سەرووی هەشتا و پێنجەوە بوو. توێژینەوەکە ڕاپۆرتی ئیمکانی مۆدێلەکە دەکات لە ژێر ئەو مەرجانەی کە باسکراوە و ئەو بەڵگانە دیاری دەکات کە ئیدیعای گواستنەوە یان قازانجی فێربوون پێویستی پێیەتی.

وشە سەرەکییەکان: خوێندەواری داتا، فێربوون لەسەر بنەمای پڕۆژە، هەڵسەنگاندنی ڕاستەقینە، دیزاینی داشبۆرد، پەروەردەی گەنجان

Wing, J.M. 2006, 'Computational thinking', Communications of the ACM, vol. 49, no. 3, pp. 33–35.

Wolff, A., Gooch, D., Cavero Montaner, J.J., Rashid, U. & Kortuem, G. 2016, 'Creating an understanding of data literacy for a data-driven society', The Journal of Community Informatics, vol. 12, no. 3, pp. 9–26.

ملخص

اليوم مهارةً مدنيةً أساسية. ومع ذلك، (Data Literacy) تُعدّ الثقافة البيانية لا يزال تعليم هذه المهارة للمتعلمين دون سن الجامعة في العالم العربي يُقدَّم في الغالب كمهارة موجهة للمستهلك؛ إذ تركز الدورات التدريبية قصيرة المدى المتاة على تعليم خصائص البرمجيات المتعلقة بالبيانات، دون ربطها بالممارسات المهنية الفعلية. تستعرض هذه الورقة تصميم ، وهو (Data Makers Camp) "وتنفيذ "مخيم صنّاع البيانات برنامج مدته 12 ساعة أُقيم في بغداد بالعراق، بالتعاون مع مبادرة وكلية (Coding for Kids Iraq) ""البرمجة للأطفال في العراق التميز بجامعة بغداد. أعاد المخيم هيكلة بيئة التعلم لتتحول إلى شبكة من "وكالات البيانات"؛ حيث تم تقسيم 75 متعلماً، تتراوح أعمارهم بين 10 و17 عاماً، إلى 25 وكالة، ضمت كل منها ثلاثة أفراد: مهندس بيانات، ، وراوٍ للقصص(Dashboard Designer) ومصمم لوحات معلومات وقّعت كل وكالة عقداً مع عميل افتراضي -من .(Story Narrator) (Briefs) بين خمسة عملاء محتملين- بناءً على مجموعة من المتطلبات الخاصة بكل عميل. كما تلقت كل وكالة واحدة من خمس مجموعات بيانات متوازية، تضمنت العيوب ذاتها لضمان تكافؤ مستوى الصعوبة بين الوكالات. غطى المنهج الدراسي مفاهيم متنوعة، بدءاً من التمييز بين البيانات والمعلومات، وإدخال البيانات المهيكلة، واستيراد البيانات "وتنظيفها باستخدام برنامجي "مايكروسوفت باور بي آي ، (Power Query) "و"باور كويري (Microsoft Power BI) وصولاً إلى اختيار المخططات البيانية المناسبة للإجابة على الأسئلة، وتحديد التمثيلات البصرية غير الدقيقة، وانتهاءً بنشر لوحة المعلومات وعرضها في يوم ختامي مفتوح. اعتمد التقييم على نموذج يتكون من ستة معايير بمجموع 100 نقطة، حيث مُنح العمل الجماعي وزناً مساوياً لوزن العرض التقديمي، مع استخدام نظام من خمس ميداليات لتعزيز هذا النموذج التقييمي. وقد نجحت جميع الوكالات الـ 25 في تسليم لوحات المعلومات المطلوبة، وتأهلت 10 مشاريع للمرحلة النهائية، حيث حصلت الوكالات الفائزة على متوسط درجات تجاوز 85 نقطة. وتخلص الدراسة إلى إثبات جدوى هذا النموذج في ظل الظروف المذكورة، كما تحدد الأدلة المطلوبة لدعم أي ادعاءات تتعلق بإمكانية .نقل التجربة إلى سياقات أخرى أو تحقيق مكاسب تعليمية ملموسة

الكلمات المفتاحية: الثقافة البيانية، التعلم القائم على المشاريع، التقييم الواقعي(Authentic Assessment) ، تصميم لوحات المعلومات، تعليم الشباب.

- **Publisher:** University of Sulaimani, **ISSN (Print):** 1813-0852, **ISSN (Online):** 2617-3034
- **Journal Websitehttps://sjh.univsul.edu.iq/, Volume:** [s], **Issue:** [3], **Year:** [2026], **DOI:** [10.17656/jzsb.12627]
- **Submission Date:** [01/08/2026], **Revised Date:** [01/08/2026], **Accepted Date:** [25/10/2026], **Published Date:** [04/11/2026]