Empowering K-12 Students Through Open Inquiry on Open Government Data: A Data-Driven Approach in CS Education

Sayed Mohsin Reza
Department of Computer Science
Pennsylvania State University Harrisburg
Harrisburg, Pennsylvania, USA
skr6024@psu.edu

Anmol Garg
Department of Computer Science
Pennsylvania State University
Harrisburg, Pennsylvania, USA
abg6200@psu.edu

Michael A. Johnson
University of North Texas
Denton, Texas, USA
michaeljohnson12cbc1@my.unt.edu

Amanda Barany
University of Pennsylvania
Philadelphia, Pennsylvania, USA
amanda.barany@gmail.com

Alex Acquah
ABC Learning Lab
The University of Texas at El Paso
El Paso, Texas, USA
aacquah@miners.utep.edu

Justice T Walker
ABC Learning Lab
The University of Texas at El Paso
El Paso, Texas, USA
jtwalker@utep.edu

***Abstract*—This research-to-practice full paper describes a transformative approach to computer science education that leverages Open Government Data to bridge the gap between theoretical knowledge and practical application, empowering K-12 and higher education students to engage in open inquiry and contribute meaningfully to societal challenges. In the contemporary landscape of computer science (CS) education at K-12 and higher levels, a visible shift towards data science-based teaching has emerged as a pivotal focus across diverse disciplines and age groups. The predominant goal is to equip students with the skills and mindset needed to actively participate as informed citizens, engaging in inquiries grounded in data that intersect with social and civic phenomena. This paradigm shift is catalyzed by the exponential growth of data production and a societal trend towards openness and information sharing, both of which are transformative forces shaping the economy and society. Numerous initiatives have been undertaken to enhance CS skills among K-12 students, with notable programs such as Bootstrap Data Science (BDS), Coding Like a Data Miner (CLDM), and Exploring Computer Science (ECS) leading the charge in integrating data science into the educational framework or curriculum using different perspectives on coding, social media data, and curated dataset. Despite these commendable efforts, a pressing concern appears over the efficacy of current pre-college data science-based CS education strategies. Often, these strategies involve the utilization of datasets and investigations curated by external entities, limiting learners' authentic practice, and constraining their exploration of meaningful lines of inquiry. This replication-centric approach impedes the development of comprehensive knowledge and mastery, hindering the cultivation of literacies with agency that allows learners to pursue personal interests or address pertinent social issues. Simultaneously, governmental bodies at the local (e.g. opendataphilly.org), state (e.g. data.pa.gov), and federal levels (e.g. data.gov) have actively engaged in this educational evolution by opening up their data for access and reuse by public and private entities. The global phenomenon of Open Government Data has gained momentum in recent years, driven by the belief that its utilization possesses the potential to generate both economic and social value. However, there exists a substantial gap in effectively engaging students to harness this valuable resource for the creation of social value and it is due to a limited number of tools and curricula that create the intersection in the learning process. This paper investigates the critical intersection of data science and computer science education, proposing a transformative approach that empowers learners through open inquiry using Open Government Data from Local (e.g. opendataphilly.org) State(data.pa.gov), and Federal(data.gov). By advocating for hands-on experiences that involve students in the generation and analysis of data relevant to real-world and local, state, and federal issues, our proposed framework seeks to bridge the gap between theoretical knowledge and practical application. The framework includes the use of a data science tool that explores open government data, connects students' interests in an area (e.g. economy, public health, transportation), and helps in mining, analyzing, and visualizing based on student socio-cultural perspective through curriculum. Through a comprehensive exploration of Open Government Data, we aim to pave the way for a new era in Data Science-based CS education that not only equips students with technical skills but also inspires them with a sense of social responsibility and the ability to contribute meaningfully to societal challenges.**

***Index Terms*—Data Science Education; K-12 Education; Curriculum Integration; Open Government Data; Culturally relevant computing**

I. Introduction

In today's rapidly evolving landscape of computer science (CS) education, the transformative impact on society is undeniable. It shapes how we interact with technology, drives innovation, and influences public discourse. Government organizations have taken a significant step forward by making their data available as open data resources, such as data.gov (federal), data.pa.gov (state), and opendataphilly.org (local) [1]. This shift aims to measure government performance and actively engage citizens through the power of real data. A critical need arises to enhance educational initiatives and best

practices, starting from high schools, fostering critical thinking and decision-making skills [2].

Examples from NSF's 10 Big Ideas: "Harnessing the Data Revolution" and "Future of Work at the Human-Technology Frontier" illustrate the significance of data science education [3]. From analyzing social media data to comprehending public perceptions, data-driven analysis has become integral to the success of different industries [4]. Therefore, it is imperative to provide data-driven CS education not only at the higher education level but also at high school education [2], [5]. This educational approach requires a fusion of computational skills and critical thinking, enabling students to mine, analyze, and visualize data effectively.

Challenges intensify when it comes to diversifying the landscape of CS learners, particularly among diverse groups such as women, people of color, and various ethnic backgrounds. Despite the prominence of the field, with 11,281 computing jobs and an average salary of $101,047 in Pennsylvania, CS education in high schools remains inadequate, with only 53% offering CS courses [6]. Furthermore, the existing curriculum often fails to align with the demands of modern data science-oriented CS education. For instance, 93% of parents in Pennsylvania desire to see computer science taught in schools, yet only 53% of high schools offer such programs [6]. A particular concern is that only 20% of female students are interested in computer science education.

To address these challenges, a transformative data science-based CS education framework is proposed to empower students with data analysis and visualization knowledge. It enables them to explore Open Government Data (OGD) and use data science tools and techniques to engage with datasets from federal, state, and local government sources. The project's objective is to utilize federal, state, and local open data initiatives, including U.S. Government's Federal Open Data (data.gov), OpenDataPA (data.pa.gov), and OpenDataPhilly (opendataphilly.org).

The research aims to create and test a data science-based computer science curriculum for pre-college students. The curriculum will focus on analyzing and visualizing data from government efforts related to social and cultural issues. This project involves designing a culturally relevant and data-driven computer science curriculum to address the under representation of certain groups. The interdisciplinary team working on this project has expertise in computer science, data science, and learning sciences. They are implementing a computer science demonstration project that utilizes big data from open government sources such as data.gov, data.pa.gov, and opendataphilly.org. The research question guiding this project is: What learning and instructional resources are necessary to sustain a data-driven computer science curriculum that emphasizes culturally relevant issues and data mining and analytics related to local, state, or federal issues?

II. Related works

Advancements in disseminating computer science knowledge at the K-12 level are not only contributing to the expansion of academic understanding but are also playing a pivotal role in driving job growth and fostering innovation within our economy and society [7], [8]. According to a report from Code.org, computing occupations are the primary source for creating new jobs in the United States [6]. The projections further underscore a 50% increase in future jobs within STEM fields, emphasizing the diverse educational and employment implications. The National Science Foundation (NSF) has laid the groundwork for these transformative initiatives like "Harnessing Data Revolution" [3]. Researchers are actively involved in designing computing education across various educational levels, from Pre-K-12 to undergraduate and graduate studies [9], [10]. These design efforts encompass curriculum interventions, teaching pedagogies, and diverse perspectives on computer science techniques. Notable organizational initiatives, such as Exploring Computer Science, Code.org, and Hour of Code, further contribute to this momentum [11]. While substantial progress has been made, ongoing research endeavors focus on refining curriculum interventions, teaching pedagogies, and incorporating public datasets in data science-based CS education. We discuss existing research on government public dataset programs, data science-based CS education research, and the integration of culturally relevant pedagogy into these educational frameworks.

A. Data science-based CS education

Benita et al. (2021) explored the integration of IoT-based ecosystems into STEM education to enhance data-driven thinking among students. The paper highlights the effectiveness of a smart learning environment that provides immersive, experiential learning opportunities. These are crafted to develop crucial STEM skills like problem-solving, analytical thinking, and cognitive flexibility [12]. K. Miller (2024) argues that as data becomes central to civic participation, data science education is crucial for preparing active citizens. She highlights eleven projects showing how learners from diverse backgrounds engage in civic and social issues through data-driven inquiry, emphasizing the link between data science education and civic engagement [13]. Our research complements this discussion by illustrating how open government data can be utilized within such smart learning ecosystems to foster not only technical skills but also a deeper understanding of civic issues, thereby enriching the CS education narrative with a focus on social good and community engagement.

B. Social and Culturally Relevant Pedagogy

Culturally relevant pedagogy, developed by Gloria Ladson-Billings, positions diverse learners and their cultural backgrounds at the center of instructional practices. This framework has shown promise in STEM education by prioritizing students' social and cultural assets [14]. Practical implementations of this pedagogy include empowering students to create computational projects like games, websites, and wearable textiles [15], [16]. This approach fosters an inclusive learning environment, validating students' lived experiences and

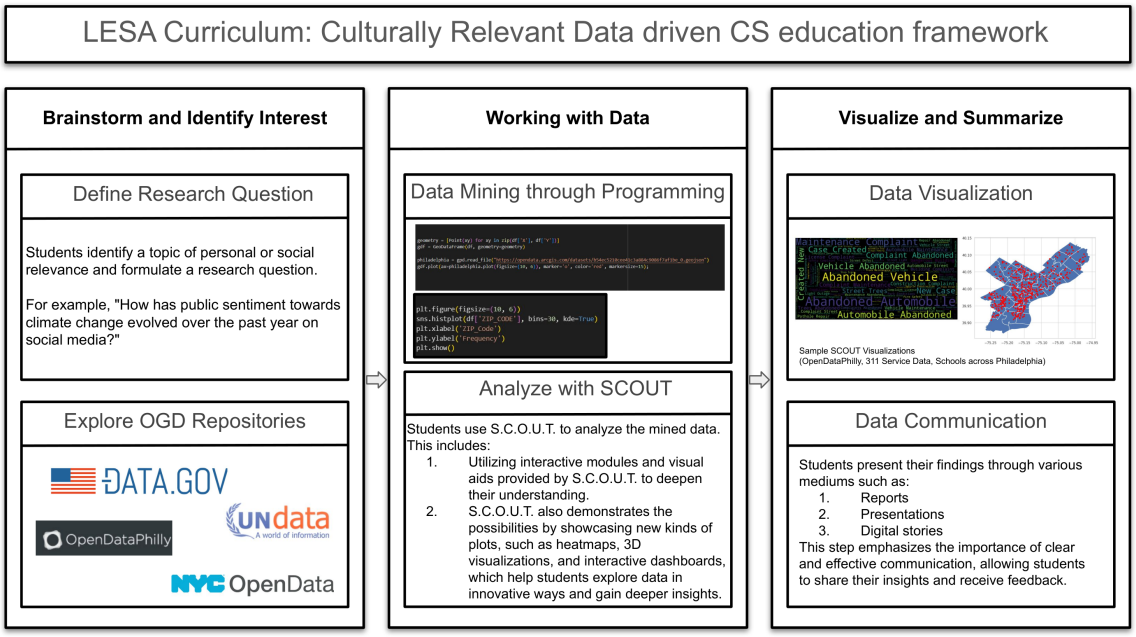

Fig. 1. Data driven CS education framework

ensuring better engagement and learning outcomes. Integrating culturally relevant pedagogy with data science education allows students to connect their personal and community interests with data mining and analysis, thereby enhancing their educational experience [17], [18]. Applying these principles, we extend this approach into the realm of CS education by promoting data literacy through the lens of societal and civic engagement, aiming to make CS education both culturally pertinent and practically impactful [19]. Walker et al. (2024) focus on embedding social and cultural relevance into the design of CS education. The paper argues for curricula that are inherently aligned with students' social contexts and cultural identities, using these elements to enrich the learning experience and outcome. It emphasizes the role of social media data in making learning relevant and engaging [20], [21].

C. Open Data Programs for Governments

L. Hagen, et al. (2019) explore the transformative potential of open government data (OGD) through visual analytics and topic modeling, focusing on the "We the People" petition platform data [1]. They demonstrate how large volumes of structured and unstructured data can enhance policy-making processes by applying topic modeling techniques to identify latent themes and using tools like LDAvis for visualization. The study assesses the usability of these tools through interviews with policy makers and data analysts, revealing significant adoption barriers such as insufficient resources, integration difficulties, and the need for training. Despite these challenges, the authors propose a policy informatics framework that suggests practical measures to lower these barriers, emphasizing the potential of visual analytics to make policy-making more data-driven and informed. Their work underscores the importance of making analytics tools accessible to non-technical users and providing the necessary training to ensure effective use in government workflows.

We build upon these insights by suggesting that open government data when leveraged effectively within CS education, can provide a rich, contextually relevant resource that not only enhances learning but also encourages students to participate actively in their communities.

III. Data-driven Framework in CS education

The proposed framework is designed for learning through data-driven thinking using open government data. It is built upon project-oriented, problem-based learning and collaborative learning. The journey of K-12 students through data-driven thinking is illustrated in Figure 1, and the main stages of the learning process can be summarized as follows:

- Brainstorm about the topic or the domain.
- Identify the subtopic of a domain of Context, Interests, or Trends.
- Identify the open government website the student will work with.
- Define the research question and formulate a hypothesis to develop cognitive skills such as logical and analytical thinking and gain a comprehensive insight into the usefulness of data to draw effective problem solutions.
- Dataset selection and collection through directly from an open government data website or our SCOUT tool, proposed in later sections and is illustrated with various features in Figure 2 and 3.

- Data analysis and processing can be conducted using various quantitative, qualitative methods with Python or our SCOUT tool that automates the process. These tools are utilized following the curriculum structure available at https://sites.psu.edu/sayedreza/d3m/ddcs/
- Data visualization enables the creation of 2D or 3D charts, maps, graphs, and networks to highlight patterns, trends, and correlations using effective techniques.
- Develop a Summary report Where K-12 students can elaborate on important discovered insights and data-driven results for government data analysis.

The framework includes a S.C.O.U.T. tool, a curriculum using Python language, open inquiry-based learning approach to enable a Data-Driven Approach in CS education.

A. SCOUT tool

S.C.O.U.T. (Statistics and Data Computing Observation and Understanding Tool) is an educational tool designed to empower K-12 students with the skills to explore, analyze, and understand data using real-world open government datasets. SCOUT aims to create an interactive, user-friendly environment where students can engage with data science concepts through hands-on experimentation. By leveraging open government data repositories, the tool provides a rich resource for students to explore diverse datasets, thereby enhancing their understanding of data analysis, statistical methods, and the importance of data-driven decision-making. Features:

- User-Friendly Interface: SCOUT offers a web-based platform accessible from any device with an internet connection, eliminating the need for software installation or account registration. The interface is intuitive, allowing students to easily navigate through various features and tools.
- Dynamic Data Exploration: Students can manipulate and visualize data using dynamic graphs, tables, and maps. The platform supports various graph types, including bar graphs, dot plots, box plots, segmented bar graphs, and scatter plots, to help students explore relationships between different data attributes.
- Hierarchical Data Organization: SCOUT allows the creation of hierarchical tables, enabling students to organize data in meaningful ways and uncover subtle relationships between attributes. This feature encourages a deeper understanding of data structures and enhances analytical skills.
- Statistical Analysis Tools: The platform includes robust statistical tools, such as mean, median, regression lines, and movable values, which can be applied to graphs for in-depth analysis. These tools help students learn key statistical concepts and methods in an interactive environment.
- Integration with Open Government Data: SCOUT seamlessly communicates with open government data repositories, providing students with access to a vast array of real-world datasets. This helps students have access to numerous datasets spanning various fields, such as public health, transportation, environmental studies, etc.
- Collaborative Learning: Teachers can design and share data activities and projects with students using the platform's sharing features. Collaborative features enable students to work together on data analysis projects, fostering teamwork and peer learning.

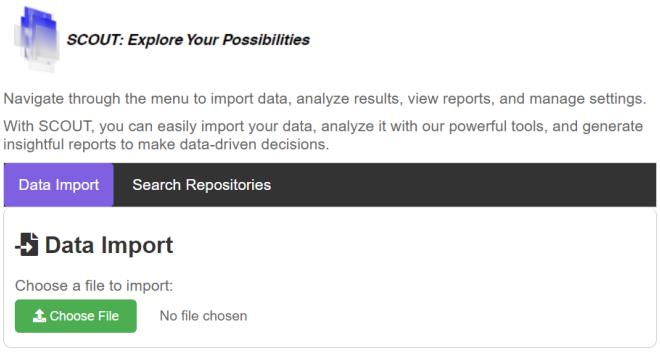

Fig. 2. SCOUT Research Version Preview - Data Import

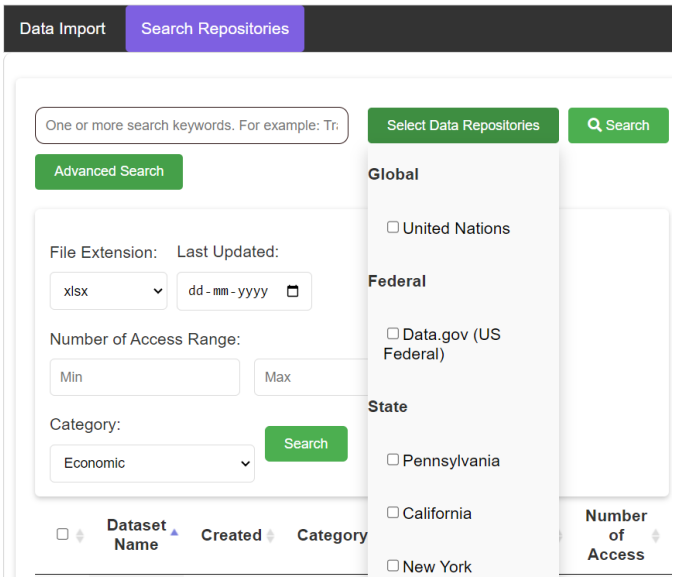

Fig. 3. SCOUT - Repository Search Options

B. Data Driven CS Curriculum

A data science-based CS curriculum is integrated into our framework, focusing on mining, analyzing, and visualizing open data repositories from federal, state, or local government efforts. The curriculum can then be used to inform national CS interventions and practice. The curriculum will be developed using a culturally relevant pedagogical approach to situate learners and their diverse interests, competencies, and cultural contexts at the center of instructional practice [14]. We will accomplish it by cultural relevance framing in instruction materials and a co-design approach. This Curriculum will be

TABLE I
CONSECUTIVE MODULE STRUCTURES FOR DATA SCIENCE-BASED CS CURRICULUM

Modules	Content Ideas
Learn	In this module, Students will learn data, different types of structures, and basic computer science techniques to mine, analyze, and visualize data.
Explore	In this module, Students will explore different open data science efforts from • Federal Open Data (data.gov), • State efforts like OpenDataPA (data.pa.gov), and • Local efforts like OpenDataPhilly (opendataphilly.org).
Select	After exploration, students will identify their interests that intersect with social and cultural aspects and select a topic to analyze. Details of such social and cultural relevance identification are discussed in Table II.
Analyze	Students will analyze the data mined from their selection of data sources and connect with students' interest to rationale.

explored in four consecutive modules mentioned in Table I and found in the following link https://sites.psu.edu/sayedreza/ddcs/.

Cultural Relevance Framing in Instructional Materials: Instructional materials will be prioritized along three themes (consistent with Ladson-Billings's (1995) culturally relevant pedagogies framing: (1) the curriculum will leverage next-generation data science practices using Government Open Data efforts (data.gov), state efforts like OpenDataPA (data.pa.gov), and local efforts like OpenDataPhilly (opendataphilly.org) (2) curricular design will enable learners to make personal connections between data mining techniques on governments open data platforms and the local Susquehanna Township, Middletown school districts region needs, priorities, and interests, (3) learning and assessment strategies will include multiple opportunities for learners to engage in discourse (e.g., about data science interests, communities, and CS project outcomes). Also, the arc of activities will shift from guided engagement as learners leverage scaffolds to develop mastery toward more open engagement where learners are free to decide and carry out their pursuits.

C. Data Driven Based Learning

The curriculum is designed to position students as authentic data scientists, engaging them in real-world data science practices through a structured, yet flexible approach. Drawing from Luna-Reyes et al. (2019) and Barany et al. (2023), the curriculum emphasizes four key stages: (1) Data Gathering, (2) Data Pre-processing, (3) Data Analysis, and (4) Data Visualization. (1) Students begin by selecting topics of personal or social relevance. (2) Next, they pre-process this data to ensure cleanliness and readiness for analysis, involving tasks like removing duplicates and correcting errors. (3) The analysis stage allows students to employ various statistical techniques to identify patterns, trends, and outliers in their data, adapting the complexity of these techniques to their learning context. (4) Finally, students visualize their findings using charts, graphs, and other visual tools to communicate their insights effectively.

TABLE II
DIVERSE PERSPECTIVES OF SOCIAL & CULTURAL RELEVANCE IN OPEN GOVERNMENT DATA

Personal Interests	Interest in Planning Data: Mine, Analyze, and Visualize AMS's Latest Air Quality Sensor Readings
Socio-political Issues	Interest in Politics Data: Mine, Analyze, and Visualize Election Results Database in Philadelphia.
Cultural Aspects	Interest in Cultural Data: Mine, Analyze, and Visualize Inventory of African American Historic Sites

D. Open Inquiry-Based Learning

Open inquiry-based learning is a central component of the curriculum, designed to shift from guided instruction to more autonomous student-led inquiry. This method encourages students to explore data-driven questions that are personally or socially significant. The curriculum is structured into four modules that gradually increase in complexity and student autonomy. Initially, students receive guided introductions to key concepts and inquiries that help them formulate questions and conduct preliminary analyses. As they progress, the support is gradually reduced, allowing students more freedom to pursue their own investigative paths. They engage in hands-on projects that require them to gather, analyze, and interpret open government data, fostering critical thinking and problem-solving skills. This approach not only makes the learning process more engaging by connecting it to real-world issues but also helps students develop a deeper understanding of data science and its applications in various contexts.

IV. DATA DRIVEN ACTIVITY DESIGN

The LESA (Learn, Explore, Select. Analyze) Curriculum design consists of five key phases: data gathering, pre-processing, analysis, visualization, and communication, all aimed at positioning students as authentic data scientists. This approach not only enhances technical skills but also fosters critical thinking and problem-solving abilities, ensuring a robust understanding of data science practices.

A. Data Gathering

The first phase involves students selecting topics of personal or social relevance. For example, they might use various openly available APIs to collect data from Open Government Data repositories using specific keywords or explore various civic agency sites related to their chosen topics. The S.C.O.U.T. tool facilitates this process by providing an intuitive interface for setting parameters and retrieving data, thereby empowering students to define their inquiries and gather meaningful, real-world data readily.

B. Data Pre-processing

The pre-processing phase involves cleaning and organizing the collected data to ensure its readiness for analysis. This involves tasks such as removing duplicates, correcting errors, and formatting the data. The pre-processing step is crucial as it helps students understand the nature of raw data and the importance of data quality in drawing accurate conclusions.

C. Data Analysis

During the analysis phase, students perform descriptive statistics, identify patterns, detect outliers, and explore relationships within the data. This iterative process encourages critical thinking and helps students develop their analytical skills. S.C.O.U.T. provides robust analytical tools and visual aids to support students in understanding complex data analysis concepts.

D. Data Visualization

In this phase, students use various Python libraries to create visual representations of their data, such as charts, graphs, and word clouds. Visualization helps students communicate their findings effectively, making complex data more accessible and understandable. S.C.O.U.T. also demonstrates the possibilities by showcasing new kinds of plots, such as heatmaps, 3D visualizations, and interactive dashboards, which help students explore data in innovative ways and gain deeper insights. By creating visual narratives, students can highlight key insights and trends identified during the analysis phase.

E. Communication

The final phase involves students presenting their findings through various mediums, such as reports, presentations, or digital stories. This step not only allows students to demonstrate their understanding of data science concepts but also emphasizes the importance of sharing insights with others. S.C.O.U.T. supports this by offering templates and tools for creating professional-quality presentations and reports. By sharing their work, students can engage with their peers and receive feedback, further enhancing their learning experience.

V. Assessment Framework

The outcome of the student project or report will be going through mixed-methodological approaches to assess student learning outcomes and teacher-student perspectives. Learning outcomes related to data science perspectives, and study instruments will be grounded in theoretical or open data perspectives. The learning experiences and outcomes are referred to as (O) and will be analyzed from cognitive, affective, and behavioral perspectives in our future research. The details of the measurement are shown in Table III.

TABLE III
Measurements of Learning Experiences & Outcomes

(O1) Cognitive Outcome	Knowledge Acquisition and Problem-Solving Skills
(O2) Affective Outcome	Self-Efficacy
(O3) Behavioral Outcomes	Engagement and Collaboration

Surveys will be conducted as part of the evaluation and assessment process. The surveys will be carried out before and after the data-driven activities and will include questions based on established measures of learning outcomes adapted from Guo (2020) and Vogler (2018) [22]. The surveys will be mostly analyzed using quantitative methods such as correlation analysis, regression analysis, T-tests, and Analysis of Variance. Qualitative data will be analyzed through the NVivo tool, which provides efficient and systematic coding, organization, and analysis of qualitative data

VI. Conclusion

The objective of this research and design project is to develop a comprehensive plan for creating a curriculum that is culturally relevant for pre-college students. This curriculum will provide a step-by-step guide for students to engage in computer science processes using Open Government Data and an open inquiry-based, constructionist approach, to model authentic data-driven computer science education practices. The detailed design and development of the curriculum are intended to serve as a potential model for future practice, focusing on two primary areas: (1) expanding student engagement with data to include producing their own interpretations and meanings, and (2) establishing connections between computer science activities and culturally and personally relevant real-world issues related to local, state, and federal governments from a humanistic perspective. Moving forward, the curriculum will be tested with underrepresented high school student groups in Central Pennsylvania to investigate the learning experiences and outcomes resulting from implementing a data-driven computer science education program that emphasizes culturally relevant and open inquiry data mining and analytics using open government data.

References

[1] L. Hagen, T. E. Keller, X. Yerden, and L. F. Luna-Reyes, "Open data visualizations and analytics as tools for policy-making," *Government Information Quarterly*, vol. 36, no. 4, p. 101387, 2019.

[2] V. R. Lee, M. H. Wilkerson, and K. Lanouette, "A call for a humanistic stance toward k–12 data science education," *Educational Researcher*, vol. 50, no. 9, pp. 664–672, 2021.

[3] "NSF 10 Big Ideas - Special Report — NSF - National Science Foundation — nsf.gov," https://www.nsf.gov/news/special_reports/big_ideas/harnessing.jsp, [Accessed 21-05-2024].

[4] A. C. Ikegwu, H. F. Nweke, C. V. Anikwe, U. R. Alo, and O. R. Okonkwo, "Big data analytics for data-driven industry: a review of data sources, tools, challenges, solutions, and research directions," *Cluster Computing*, vol. 25, no. 5, pp. 3343–3387, 2022.

[5] J. T. Walker, A. Barany, A. Acquah, S. M. Reza, A. Barrera, K. D. R. Guzman, and M. A. Johnson, "Coding like a data miner: A sandbox approach to computing-based data science for high school student learning," in *2023 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2023, pp. 1–5.

[6] code.org, "Support k-12 computer science education in pennsylvania," https://code.org/advocacy/state-facts/PA.pdf, 2021, [Accessed 21-05-2024].

[7] C. Proctor, M. Bigman, and P. Blikstein, "Defining and designing computer science education in a k12 public school district," in *Proceedings of the 50th ACM technical symposium on computer science education*, 2019, pp. 314–320.

[8] A. K. Chatterji, "Innovation and american k–12 education," *Innovation Policy and the Economy*, vol. 18, no. 1, pp. 27–51, 2018.

[9] A. Oleson, B. Wortzman, and A. J. Ko, "On the role of design in k-12 computing education," *ACM Transactions on Computing Education (TOCE)*, vol. 21, no. 1, pp. 1–34, 2020.

[10] Y. Kafai, C. Proctor, and D. Lui, "From theory bias to theory dialogue: embracing cognitive, situated, and critical framings of computational thinking in k-12 cs education," *acm Inroads*, vol. 11, no. 1, pp. 44–53, 2020.

[11] E. W. Huff Jr, K. Boateng, M. Moster, P. Rodeghero, and J. Brinkley, "Exploring the perspectives of teachers of the visually impaired regarding accessible k12 computing education," in *Proceedings of the 52nd acm technical symposium on computer science education*, 2021, pp. 156–162.

[12] F. Benita, D. Virupaksha, E. Wilhelm, and B. Tunçer, "A smart learning ecosystem design for delivering data-driven thinking in stem education," *Smart Learning Environments*, vol. 8, pp. 1–20, 2021.

[13] K. M. Miller, J. L. Polman, S. A. Yoon, J. Shim, V. Y. Leung, Y. Nguyen, A. Rubin, A. Rubin, T. Higgins, J. M. Karch, J. K. Hammerman, C. Matuk, K. DesPortes, A. Amato, S. Dikker, X. Ochoa, E. Romero, S. Podworny, Y. Fleischer, R. Biehler, J. T. Walker, A. Barany, A. Acquah, A. Scarola, S. Reza, T. C. Tran, R. Vacca, M. Silander, P. J. Woods, C. Fernandez, A. Eloy, P. Blikstein, R. de Deus Lopes, J. Radinsky, I. Tabak, V. R. Lee, D. Demszky, S. Levine, and J. Louie, "Data and social worlds: How data science education supports civic participation and social discourse," in *Proceedings of the 18th International Conference of the Learning Sciences - ICLS 2024*, ser. ICLS 2024. International Society of the Learning Sciences, Jun. 2024. [Online]. Available: http://dx.doi.org/10.22318/icls2024.702524

[14] G. Ladson-Billings, "Toward a theory of culturally relevant pedagogy," *American educational research journal*, vol. 32, no. 3, pp. 465–491, 1995.

[15] G. T. Richard and S. Giri, "Digital and physical fabrication as multimodal learning: Understanding youth computational thinking when making integrated systems through bidirectionally responsive design," *ACM Transactions on Computing Education (TOCE)*, vol. 19, no. 3, pp. 1–35, 2019.

[16] J. Morales-Chicas, M. Castillo, I. Bernal, P. Ramos, and B. L. Guzman, "Computing with relevance and purpose: A review of culturally relevant education in computing." *International Journal of Multicultural Education*, vol. 21, no. 1, pp. 125–155, 2019.

[17] T. Smith, L. Avraamidou, and J. D. Adams, "Culturally relevant/responsive and sustaining pedagogies in science education: theoretical perspectives and curriculum implications," *Cultural Studies of Science Education*, vol. 17, no. 3, pp. 637–660, 2022.

[18] J. Gao, "Culturally relevant pedagogy in teaching proficiency through reading and storytelling," *Chinese Language Teaching Methodology and Technology*, vol. 3, no. 2, p. 63, 2020.

[19] A. Barany, S. Reza, M. Johnson, A. Barrera, O. Badreddin, C. Fuentes, and J. T. Walker, "Towards the design of a culturally relevant curriculum for equitable, data mining-based cs education," in *Proceedings of the 17th International Conference of the Learning Sciences-ICLS 2023, pp. 1498-1501*. International Society of the Learning Sciences, 2023.

[20] J. T. Walker, A. Barany, A. Barrera, M. A. Johnson, and S. M. Reza, "Perspectives on research and practice in and around cultural relevance for pre-college data science in computing," in *Improving Equity in Data Science*. Routledge, 2024, pp. 12–36.

[21] J. T. Walker, A. Barany, A. Acquah, S. M. Reza, K. Guzman, M. Johnson, O. Badreldin, and A. Barrera, *Sandbox Data Science: Culturally Relevant K-12 Computing*, 2024.

[22] P. Guo, N. Saab, L. S. Post, and W. Admiraal, "A review of project-based learning in higher education: Student outcomes and measures," *International journal of educational research*, vol. 102, p. 101586, 2020.