

# Act-NOW: Data Mapping Project

**Ritik Jain**, [jainrit2@msu.edu](mailto:jainrit2@msu.edu)

**Naamna Modi**, [modinaam@msu.edu](mailto:modinaam@msu.edu)

**Ritin Kalidindi**, [kalidin6@msu.edu](mailto:kalidin6@msu.edu)

**Gabriella Stickney**, [stickn21@msu.edu](mailto:stickn21@msu.edu)

## About QSIDE

The QSIDE (Quantitative Studies of Inclusion, Diversity, and Equity) Institute is dedicated to leveraging data science and mathematical modeling to advance social justice. Inspired by civil rights pioneer Ida B. Wells, who stated that "the way to right wrongs is to turn the light of truth upon them," QSIDE seeks to illuminate systemic inequities and drive meaningful change through data-driven research and action.

QSIDE's work encompasses research-to-action-to-justice projects that address critical issues. It also provides data, evaluation, and research consulting services for non-profits and mission-driven organizations, as well as education and community-building programs to enhance the intersection of data science and social justice, fostering informed and ethical research practices.

Established as a 501(c)(3) nonprofit in 2019, QSIDE operates as a dynamic network of quantitative scholars, justice subject matter experts, activists, and stakeholders. Led by co-founders Chad Topaz, Ph.D., and Jude Higdon, Ed.D., QSIDE is committed to participatory community engagement, ensuring that data science serves as a powerful tool for equity and justice.

At its core, QSIDE upholds fundamental principles that guide its mission, including science as the foundation of knowledge, the recognition of systemic inequities, and the importance of constructive dialogue and activism. QSIDE operates based on scientific consensus and rigorous research, acknowledges the historical exclusion and oppression of marginalized populations and actively works to address these injustices, and fosters a space for respectful discussion and collaboration while firmly supporting established scholarly consensus on social justice issues.

## Background and Motivation

Chronic absenteeism—defined as missing 10% or more of the school year for any reason—is a persistent and complex issue affecting millions of students across the United States (Bauer et al., 2018). Unlike truancy, which refers specifically to unexcused absences, chronic absenteeism encompasses all absences, whether excused or unexcused, making it a more comprehensive indicator of student disengagement or hardship. According to the U.S. Department of Education, students who are chronically absent are at greater risk of falling behind academically, dropping out of school, and facing long-term socioeconomic challenges. (Eklund et al., 2020)

This project stems from the growing need to better understand the complex and interconnected factors contributing to chronic absenteeism in schools. These factors span a wide range of areas, including health challenges, family responsibilities, transportation barriers, school attributes, and broader socioeconomic conditions.

Through our partnership with QSIDE, we took a data-driven approach to explore patterns between school-related factors and chronic absenteeism rates. Our goal was to create accessible tools that empower educators and communities to better understand and address this widespread issue.

## Problem Definition and Goals

The primary goal of this project was to explore the indicators and contributing factors of chronic absenteeism using publicly available data sources and insights from scholarly research. In collaboration with our community partner, QSIDE, we identified key variables linked to absenteeism and developed a synthetic ("toy") dataset that mirrors these variables. This dataset serves as a structured foundation for analysis and helps illustrate potential relationships and trends related to chronic absenteeism.

Another core objective was to design a series of interactive data visualizations using Tableau. These visualizations aim to clearly convey the connections between various indicators—such as socioeconomic status, school resources, and student demographics—and chronic absenteeism rates. By presenting the data in a visual format, we sought to make it easier for users to explore patterns and draw meaningful, data-informed conclusions.

Guided by extensive research and feedback from QSIDE, our team approached the project with a strong grounding in the literature, ensuring our methods and outcomes prioritized clarity, accuracy, and equity.

## Methodology

Our methodology centered around identifying, collecting, organizing, and visualizing publicly available data related to barriers to consistent student attendance. We approached this project in a multi-phase process, designed to ensure both clarity of insights and accessibility for stakeholders with varying levels of data literacy.

We began by analyzing a collection of datasets provided by QSIDE and supplemented them with data from open-source platforms such as the U.S. Census Community Survey. Our early work focused on identifying which indicators were most relevant to school attendance patterns such as income level, access to transportation, housing instability, public health access, and community safety. To better understand the structure and meaning of these indicators, we created a data map that broke them down into super-classes and sub-classes, allowing for clear organization.

Before working with large-scale datasets, we developed a fake toy dataset to test our approach. This allowed us to experiment with cleaning steps, test visualizations, and build a reproducible

workflow for transforming raw data into insights. We used this smaller dataset to practice operations such as filtering, grouping, and restructuring columns into meaningful categories.

StudentID	StudentFirstName	StudentLastName	Race/Ethnicity	GenderAssignedatBirth	GenderIdentity	Address	City	State	ZipCode	Birthdate	NumofSiblings	SES	StableLivingSituation
1001	John	Doe	Caucasian	Male	Non-Binary Femme	2421 Arrowhead Dr	Springfield	IL	62702	5/14/2008	2	Low	2
1002	Jane	Smith	Asian	Female	Cisgender Female	922 N 4th St	Springfield	IL	62702	9/21/2007	1	Middle	3
1003	Michael	Johnson	African American	Male	Cisgender Male	2213 W Chatsford Ct	Peoria	IL	61615	12/2/2006	3	High	1
1004	Emily	Davis	Hispanic	Female	Pangender Fluid	344 Point Bluff Dr	Decatur	IL	62521	3/11/2009	0	Low	5
1005	William	Brown	Caucasian	Male	Genderqueer Femme	2001 Marland St	Springfield	IL	62702	7/30/2008	2	Middle	3
1006	Olivia	Wilson	Asian	Female	Transfeminine	1921 S Richmond Rd	Decatur	IL	62521	10/15/2007	1	Low	2

Figure 1: Example of our Toy Dataset

We regularly checked in with our community partners to review our findings in relation to their goals and gather feedback to refine our approach. Each step of the process was grounded in a desire to highlight patterns and correlations without oversimplifying the complex, structural causes behind school attendance challenges.

Throughout the project, we documented our process, built a reproducibility guide, and created visual tools that prioritize accessibility and transparency. Our focus was on surfacing meaningful insights through structured analysis and storytelling that can inform future work or intervention design by stakeholders.

## Results

Our team found the following as key indicators of chronic absenteeism: food insecurity, stability of living situation, grade distribution, disability status, attendance rate, literacy rate, and incidence of illness.

After creating the toy dataset, we used Tableau to create various interactive maps to visualize the different attributes: identification numbers, days present, excused absences, tardies, dates absent, reason given, and absence note. We used different mapping techniques, such as color maps to differentiate various attributes. We show various findings below associated with figures.



Figure 2: Multi-Layer Map

In this multi-layer map, we plot the GEOIDs of individual students. The map is color-coded by default to show unexcused absences, but users can interact with the visualization by selecting other metrics—causing the point colors to update accordingly. This allows for a visual comparison between unexcused absences and various other factors across the area. Additionally, the map includes a filter option by school for more focused analysis.



Figure 3: SAT Average Per School

This map currently serves as a placeholder, displaying the exact longitude and latitude of each school, with color coding based on average SAT scores. However, our goal is to update this visualization by replacing SAT average with a school's rank within its district or an overall performance metric, providing a more comprehensive view of school quality.

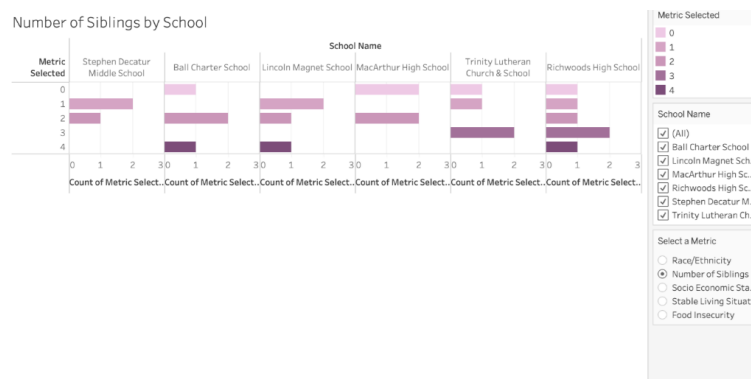


Figure 4: Bar Plot

This bar plot allows users to select from various metrics, dynamically updating the graph to display counts for the selected category. For example, selecting "socioeconomic status" will generate a bar chart showing the number of students in each economic category within a given school. This interactive feature makes it easy to explore and compare distributions across different variables.

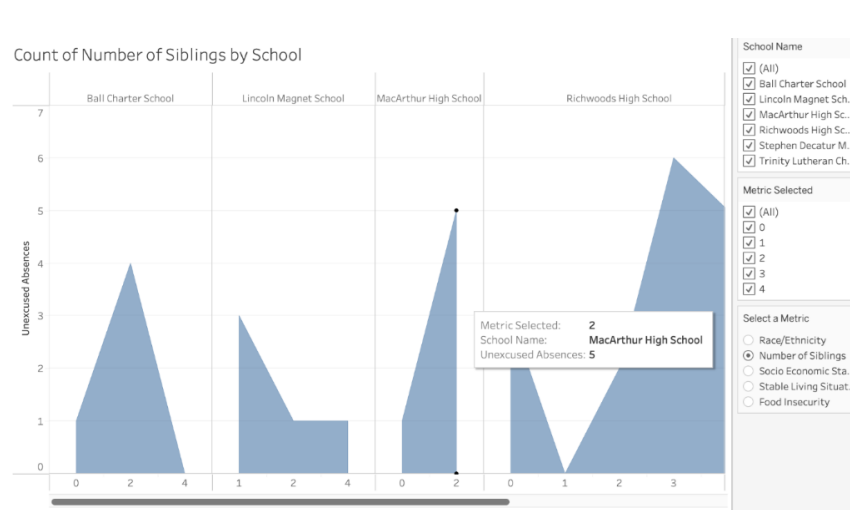


Figure 5: Line Plot

This line plot visualizes unexcused absences based on a selected metric, allowing users to observe trends over different categories. It corresponds with Figure 2, the multi-layer map above, offering a complementary view of the data. Overall, our dashboard includes a variety of visualizations that highlight different dimensions of chronic absenteeism, making the issue easier to explore and understand from multiple angles.

## Conclusions

In conclusion, our project offered a structured and exploratory look into the complex issue of chronic absenteeism using a synthetic dataset and interactive Tableau visualizations. While the data itself was artificial, our analytical framework, variable design, and visualization tools were carefully built to mirror real-world conditions and constraints. The process allowed us to explore how school-level indicators such as socioeconomic status, student demographics, and school resources might relate to absenteeism patterns. Overall, our project includes key research into what it means to be chronically absent, key indicators that go into it as well as the creation and use of our toy dataset (which includes our key indicators of chronic absenteeism) to then use our artificial data for creating visualizations in the form of maps in Tableau. Our team was able to create maps of various schools within a state and include important key information about each school within the map for further interpretations, we have metrics including SAT averages, ACT averages 4 and 5-year graduation rates, rank, school size, etc. Though our current work relied on toy data, it lays a foundation for impactful future work with real datasets, our team would have ideally created our maps based on real data but due to not receiving actual data we had to rely on our toy dataset, and we think there would be great potential in future works for our project using real relevant data.

## Challenges

One of the major challenges our team faced was the lack of real-world data to work with. As a result, we created and utilized a synthetic “toy dataset” using Microsoft Excel, which included variables such as geographic information about schools, academic statistics, and 4-year and 5-year graduation rates disaggregated by race and gender. While this allowed us to move forward with our analysis, the limited size and scope of the dataset made it difficult to draw strong conclusions or identify consistent relationships between variables.

Because the dataset was relatively small, our visualizations—though thoughtfully designed—sometimes appeared sparse or inconclusive at first glance. While they still manage to convey key information, the lack of robust data points made it harder to identify clear patterns or insights without deeper inspection. Chronic absenteeism is a highly nuanced issue, influenced by a wide range of intersecting factors, and representing that complexity accurately within a limited synthetic data set posed a significant challenge for our team.

## Future Works

Moving forward, we aim to implement our Tableau maps using real data while ensuring student privacy. To maintain confidentiality, we plan to use Census blocks or ZIP codes rather than personally identifiable student data. This approach will allow for meaningful geographic insights without compromising individual student privacy.

Additionally, we plan to develop a front-end interface integrated with Tableau to make the application accessible to site leaders who may not have technical backgrounds. This user-friendly interface will streamline the data input process and allow users to generate visualizations with ease. As part of this effort, we have created a rough wireframe to demonstrate the application's functionality.

The wireframe shows a web interface with the following components:

- Website Header**: A box at the top of the page.
- CSV File Upload**: A text input field for uploading data files.
- Visualization Options**: A section titled "Checkboxes for which types of data visualizations the user wants to view including:" with a list of chart types: Bar Chart, Histograms, Scatter Plot, Ridge Plot, Line Graph, Box Plot, Pie Chart, and Violin Plot.
- Attributes Selection**: A section titled "Checkboxes for which attributes the user wants to visualize:" with a list of attributes: StudentID, SchoolID, DaysPresent, ExcusedAbsences, UnexcusedAbsences, Tardies, DatesAbsent, ReasonGiven, AbsenceNote, and ReasonCategory.
- Map Configuration**: A section titled "Special Section if Tableau Data Map is wanted:" with options for "Filter Color by: (Set of Attributes)", "Type of Data Map (State Level, County Level, National Level)", and "Filter by Specified Threshold".
- Disclaimer**: A box at the bottom of the form.

Figure 6: The input screen of the wire frame



Figure 7: The output screen of the wire frame

The wireframe in both Figure 6 and Figure 7 outlines a process where users can upload a CSV file, select visualization types such as bar charts, scatter plots, and pie charts, and choose specific attributes to visualize, such as Student ID, School ID, and Dates Absent.

A key feature of our design includes a section dedicated to customizing a Tableau map. Users will be able to filter by specific attributes, select the type of data map they want to generate, and define thresholds for analysis. These enhancements will ensure that the platform is both powerful and intuitive, making it easier for stakeholders to interpret and act on data insights effectively.

## Acknowledgments

Our team would like to acknowledge our capstone instructor Dirk Colbry as well as our community partners Jude Higdon, Carlos Alvarez, and Tyrone Bass who have worked with us throughout this semester. They have been invaluable in providing feedback, help and direction to our team as we worked through this project.

## References

- Eklund, K., Burns, M. K., Oyen, K., DeMarchena, S., & McCollom, E. M. (2020). Addressing Chronic Absenteeism in Schools: A Meta-Analysis of Evidence-Based Interventions. *School Psychology Review*, 51(1), 95–111. <https://doi.org/10.1080/2372966X.2020.1789436>
- Gottfried, M. A. (2015). Chronic Absenteeism in the Classroom Context: Effects on Achievement. *Urban Education*, 54(1), 3-34. <https://doi.org/10.1177/0042085915618709> (Original work published 2019)
- Gottfried, M. A. (2014). Chronic Absenteeism and Its Effects on Students' Academic and Socioemotional Outcomes. *Journal of Education for Students Placed at Risk (JESPAR)*, 19(2), 53–75. <https://doi.org/10.1080/10824669.2014.962696>

Bauer, L., Liu, P., Schanzenbach, D. W., & Shambaugh, Jay. Hamilton Project. (2018). Reducing Chronic Absenteeism Under the Every Student Succeeds Act. Retrieved from [https://www.hamiltonproject.org/wp-content/uploads/2023/01/reducing\\_chronic\\_absenteeism\\_under\\_the\\_every\\_student\\_succeeds\\_act.pdf](https://www.hamiltonproject.org/wp-content/uploads/2023/01/reducing_chronic_absenteeism_under_the_every_student_succeeds_act.pdf)

QSIDE Institute. (n.d.). About QSIDE. Retrieved from <https://qsideinstitute.org/#>

Sheldon, S. B., & Epstein, J. L. U.S. Department of Education. (2008). Chronic Absenteeism: An Old Problem in Search of New Answers. ERIC. Retrieved from <https://eric.ed.gov/?id=EJ794822>

Wikipedia contributors. (n.d.). Socioeconomic Status. Wikipedia, The Free Encyclopedia. Retrieved from [https://en.wikipedia.org/wiki/Socioeconomic\\_status#:~:text=Socioeconomic%20status%20is%20typically%20broken,and%20occupation\)%20can%20be%20assessed.](https://en.wikipedia.org/wiki/Socioeconomic_status#:~:text=Socioeconomic%20status%20is%20typically%20broken,and%20occupation)%20can%20be%20assessed.)

Teen Talk. (n.d.). Gender Identity. Retrieved from <https://teentalk.ca/learn-about/gender-identity/#:~:text=There%20are%20many%20different%20gender,or%20a%20combination%20of%20these.>

## Appendix A: Deliverables

Below is a list of all major deliverables completed over the course of the semester, along with short descriptions and direct links to their locations in our GitHub repository.

Deliverable	Description	Link
Project Proposal	Initial project overview outlining goals, data sources, and proposed impact.	<a href="https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/2025-0202-Proposal.pdf">https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/2025-0202-Proposal.pdf</a>
Project Plan Storyboard	Visual storyboard outlining our project plan and research flow.	<a href="https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Qside-CMSE495_Plan_Storyboard.pdf">https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Qside-CMSE495_Plan_Storyboard.pdf</a>
Project Plan Video	Short video explaining our project goals and proposed methodology.	<a href="https://michiganstate.sharepoint.com/sites/Section_SS25-CMSE-495-001-225215054-EL-32-A26-QSIDE/_layouts/15/stream.aspx?id=%2Fsites%2FSection%2FSS25%2DCMSE%2D495%2D001%2D225215054%2DEL%2D32%2DA26%2DQSIDE%2FShared%20Documents%2FQSIDE%2FProject%2Fdeliverables%2FQSIDE%2DCMSE495%5FPan%5FPresentation%5FVideo%2Emp4&amp;referrer=StreamWebApp%2EWeb&amp;referrerScenario=AddressBarCopied%2Evie%2E31fcd6ea%2D462b%2D40ee%2Db335%2D34041fc95fe7">https://michiganstate.sharepoint.com/sites/Section_SS25-CMSE-495-001-225215054-EL-32-A26-QSIDE/_layouts/15/stream.aspx?id=%2Fsites%2FSection%2FSS25%2DCMSE%2D495%2D001%2D225215054%2DEL%2D32%2DA26%2DQSIDE%2FShared%20Documents%2FQSIDE%2FProject%2Fdeliverables%2FQSIDE%2DCMSE495%5FPan%5FPresentation%5FVideo%2Emp4&amp;referrer=StreamWebApp%2EWeb&amp;referrerScenario=AddressBarCopied%2Evie%2E31fcd6ea%2D462b%2D40ee%2Db335%2D34041fc95fe7</a>



Outreach PowerPoint Slide	Single-slide summary used for community outreach and awareness.	<a href="https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Outreach_PPslide.pdf">https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Outreach_PPslide.pdf</a>
Outreach Blog	Public-facing blog post summarizing our project's purpose and early findings.	<a href="https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Outreach_BlogPost.pdf">https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Outreach_BlogPost.pdf</a>
Install Instructions	Step-by-step guides for setting up all tools (Tableau, Excel, Census API).	<a href="https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Install_Instructions.md">https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Install_Instructions.md</a>
MVP Storyboard PowerPoint	Slide deck showing our Minimum Viable Product (MVP) with early visualizations.	<a href="https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Qside-CMSE495_MVP_Storyboard.pdf">https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Qside-CMSE495_MVP_Storyboard.pdf</a>
MVP Storyboard Video	Brief walkthrough of our MVP, including visuals and explanations.	<a href="https://michiganstate.sharepoint.com/sites/Section_SS25-CMSE-495-001-225215054-EL-32-A26-QSIDE/_layouts/15/stream.aspx?id=%2Fsites%2FSection%5FSS25%2DCMSE%2D495%2D001%2D225215054%2DEL%2D32%2DA26%2DQSIDE%2FShared%20Documents%2FQSIDE%2FProject%5FDeliverables%2FQSIDE%2DCMSE495%5FMVP%5FPresentation%5FVideo%2Emp4&amp;referrer=StreamWebApp%2EWeb&amp;referrerScenario=AddressBarCopied%2Eview%2E896c405b%2Dff98%2D4003%2D83cf%2D18c33ec91370">https://michiganstate.sharepoint.com/sites/Section_SS25-CMSE-495-001-225215054-EL-32-A26-QSIDE/_layouts/15/stream.aspx?id=%2Fsites%2FSection%5FSS25%2DCMSE%2D495%2D001%2D225215054%2DEL%2D32%2DA26%2DQSIDE%2FShared%20Documents%2FQSIDE%2FProject%5FDeliverables%2FQSIDE%2DCMSE495%5FMVP%5FPresentation%5FVideo%2Emp4&amp;referrer=StreamWebApp%2EWeb&amp;referrerScenario=AddressBarCopied%2Eview%2E896c405b%2Dff98%2D4003%2D83cf%2D18c33ec91370</a>
Reproducibility Document	Documentation to help users test and replicate our work and visualizations.	<a href="https://github.com/ritikjain1/QsideCapstone25/tree/main/Reproducibility_Documents">https://github.com/ritikjain1/QsideCapstone25/tree/main/Reproducibility_Documents</a>
Final Report Storyboard	Final visual storyboard summarizing our results and key findings.	<a href="https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Qside-CMSE495_Final_Presentation_Plan.pdf">https://github.com/ritikjain1/QsideCapstone25/blob/main/Deliverables/Qside-CMSE495_Final_Presentation_Plan.pdf</a>
Final Project Storyboard Video	Concluding video presentation detailing our process, insights, and outcomes.	<a href="https://michiganstate-my.sharepoint.com/personal/jainrit2msu_edu/_layouts/15/stream.aspx?id=%2Fpersonal%2Fjainrit2%5Fmsu%5Fedu%2FDocuments%2FAttachments%2FSemester%20Final%20Storyboard%20Video%2Emp4&amp;ct=1745194934361&amp;or=OWA%2DNT%2DMail&amp;cid=4884e4eb%2D03c6%2D6fe2%2Df3e7%2D90359ed2c747&amp;ga=1&amp;referrer=StreamWebApp%2EWeb&amp;referrerScenario=AddressBarCopied%2Eview%2Ed6e74621%2D44a6%2D4811%2D8462%2Dcc4b1dd3da19">https://michiganstate-my.sharepoint.com/personal/jainrit2msu_edu/_layouts/15/stream.aspx?id=%2Fpersonal%2Fjainrit2%5Fmsu%5Fedu%2FDocuments%2FAttachments%2FSemester%20Final%20Storyboard%20Video%2Emp4&amp;ct=1745194934361&amp;or=OWA%2DNT%2DMail&amp;cid=4884e4eb%2D03c6%2D6fe2%2Df3e7%2D90359ed2c747&amp;ga=1&amp;referrer=StreamWebApp%2EWeb&amp;referrerScenario=AddressBarCopied%2Eview%2Ed6e74621%2D44a6%2D4811%2D8462%2Dcc4b1dd3da19</a>

## Appendix B: Installation / Reproducibility instructions

To get started with the project on your local machine, follow the steps below:

1. **Clone the GitHub Repository**

Open your terminal and run the following commands:

```
Bash
git clone https://github.com/ritikjain1/QsideCapstone25.git
cd QsideCapstone25
```

2. **Follow the Full Reproducibility Instructions**

Detailed instructions for setting up Tableau, Excel, and the US Census API are provided in the repository:

[https://github.com/ritikjain1/QsideCapstone25/tree/main/Reproducibility\\_Documents](https://github.com/ritikjain1/QsideCapstone25/tree/main/Reproducibility_Documents)

These steps will guide you through everything needed to replicate our analysis and visualizations.

## Appendix C: Data Bibliography

Our toy dataset consists of six different Excel sheets, covering student demographics, metadata, individual school trajectories, school demographics, and student school presence. It includes 25 fake students with attributes. Additionally, it incorporates data from six real Illinois schools and 80 fake excused absence letters to mimic real-world attendance records. Our fake toy dataset will serve as an outline for site leaders to collect real-world data. It will help for a smoother transition to incorporate the real data into our existing Tableau workbooks.

## Appendix D: Data Dictionary

Below are the 6 different Excel sheets as well as some information on the data included within each sheet. Regarding the locations included in the toy dataset, we focused on 3 cities within Illinois (Springfield, Decatur, and Peoria). Most of the information was generated using ChatGPT (our prompts and results are provided in our [Github](#)). We also used resources such as Zillow and Google to find real addresses and schools.

- **StudentDemographics:** This sheet gives information regarding the student's personal data. There are 25 unique students listed within this Excel sheet.
  - StudentID
    - This is currently a stand in for the unique identifier for each of the students
  - StudentFirstName

- This is the first name of the student
- StudentLastName
  - This is the last name of the student
- Race/Ethnicity
  - This is the race/ethnicity of each student
  - The values within this column consist of African American, Asian, Caucasian, Hispanic
- GenderAssignedatBirth
  - This is the gender assigned at birth of each student
  - The values within this column are Male and Female
- GenderIdentity
  - This is the gender identity that the student most recently identifies as
  - The values within this column consist of Agender, Agender Neutrois, Cisgender Female, Cisgender Male, Gender Neutral, Genderqueer Androgynous, Genderqueer Femme, Non-Binary, Non-Binary Femme, Non-Binary Masc, Pangender, Pangender Fluid, Third Gender, Third Gender Androgynous, Third Gender Fluid, Transfeminine, Transgender Female, Transgender Male, Transgender Non-Binary, Transmasculine, Two-Spirit Feminine
  - \*We did our best to have varying gender identities to best reflect the numerous gender identities. We used this [source](#) as a guide for all the gender identities included in the toy dataset:
- Address
  - This is the student's current address \*All of the addresses are real Illinois homes. These addresses were found through [Zillow](#), inputting the wanted city of each student.
- City
  - These are the cities that each student's address is in (Springfield, Decatur, or Peoria)
- State
  - Because we are only focusing on the state of Illinois, all the states are "IL"
- ZipCode
  - The five-digit postal code corresponding to the student's address
- Birthdate
  - This is the birthdate of each student. This variable will be mainly used to look at the ages of each student for further analysis
- NumofSiblings

- This is the number of siblings that each registered student has
- SES (High/Middle/Low)
  - This is the Socioeconomic status of the student. These are categorized into (High, Middle, or Low) categories. We used [Wikipedia](#) to determine these categories. These categories were also discussed and agreed upon with the community partners.
- StableLivingSituation (Scale 1-5: 1 Not stable - 5 Very Stable)
  - This is a scale of the living situation of the student. 1 is considered not stable while 5 is considered very stable.
- FoodInsecure (Rating 1-5: 1 Not Insecure- 5 Very Insecure)
  - This is a scale of the food security of the student. 1 is considered not insecure while 5 is considered very insecure.
- ID\_MetaData: This sheet contains a preliminary list of potential values for the StudentID column in the StudentDemographics sheet. The StudentID must be a unique, unchanging identifier assigned to each student, ensuring there is no risk of duplication or confusion, regardless of the school they attend or their grade level.
- IndividualSchoolTrajectory: This sheet gives information regarding the educational path that each of the 25 students is currently on.
  - StudentID
    - This is currently a stand in for the unique identifier for each of the students
    - It references the StudentID that is assigned within the StudentDemographics sheet
  - ExpectedGraduationDate
    - This is the expected graduation date for each student
    - All the graduation dates are between 2024 and 2027
  - CourseofStudy
    - This is the course of study that each student is currently on.
    - Consists of Science, Engineering, Business, Arts, Math, Literature, and Technology
  - GPA
    - On a scale of a 4.0
  - GradeDistribution(A-F)
    - This column has how many A/B/C/D grades that each student has in the gradebook
    - Each student has a total of 6 grades distributed amongst the 4 letter grades
  - PhysicalDisabilityStatus

- This column details the student's physical disability status
  - Listed disabilities include Visual impairment, Hearing impairment, and Physical Disability
- Neurodiverse/Learning Disability
  - This column details the student's Neurodiverse/Learning Disability status
  - Listed disabilities include ADHD, Dyslexia, and Autism
- IEP(Y/N)
  - This value is whether the student is on an Individual Educational Plan
- SchoolDemographics: This sheet gives information about the 6 unique, real Illinois schools. \*While all of the school names and addresses are real, the remaining values are artificially generated and do not reflect the actual status of each of the schools.
  - SchoolID
    - A unique numerical identifier assigned to each school. This ensures that schools can be easily referenced without ambiguity
  - SchoolAddress
    - The full street address of the school, including building numbers and street names
  - City
    - The city where the school is located (Springfield, Decatur, or Peoria)
  - State
    - The state abbreviation (e.g., IL for Illinois) where the school operates
    - Because we are only focusing on the state of Illinois, all the states are "IL"
  - ZipCode
    - The five-digit postal code corresponding to the school's address
  - SchoolDistrict
    - The numeric identifier of the public school district the school belongs to. \*Private or charter schools may not be directly associated with a public district
  - SchoolType
    - The classification of the school, such as Public District, Private, Charter, or Magnet.
    - This indicates whether the school is government-funded, independently run, or has specialized programs
  - EducationLevel
    - The grade levels offered at the school (PreK-8, 6-8, 9-12)

- SchoolName
  - The full name of the school
- TotalSchoolDays
  - The total number of school days in a given academic year
- SchoolSize
  - The total number of enrolled students at the school
- Student:TeacherRatio
  - The ratio of students to teachers in the school, displayed in the format X:01 (e.g., 18:01 for 18 students per teacher)
  - Lower ratios generally indicate smaller class sizes
- YearOpened
  - The year the school was originally established. This helps provide historical context about the institution's longevity
- Four-YearGraduationRate(Overall | By Race/Gender)
  - The percentage of students who graduate within four years, both as an overall percentage and broken down by race/ethnicity and gender
- Five-YearGraduationRate(Overall | By Race/Gender)
  - The percentage of students who graduate within five years, also segmented by race/ethnicity and gender. This can highlight differences in graduation timelines among student demographics
- RankedinDistrict
  - The school's ranking within its respective school district
- RankedinState
  - The school's ranking within the state based on standardized assessment scores or overall performance compared to other schools
- ACTAvgerage
  - The average ACT score of students at the school
- SATAvgerage
  - The average SAT score of students at the school
- StudentSchoolPresence: This sheet gives information regarding the attendance, specifically the absences and tardies of each of the students at their recorded school. It also includes made up absence notes for the number of ExcusedAbsences (totaling 80 fake absence letters).
  - StudentID
    - This is currently a stand in for the unique identifier for each of the students

- It references the StudentID that is assigned within the StudentDemographics sheet
- SchoolID
  - A numeric identifier linking the student to their school
  - It references the SchoolID that is assigned within the SchoolDemographics sheet
- DaysPresent
  - The total number of school days the student attended
- ExcusedAbsences
  - The number of days the student was absent for a valid, documented reason
- UnexcusedAbsences
  - The number of days the student was absent without an acceptable excuse
- Tardies
  - The number of times the student arrived late to school
- DatesAbsent
  - A semicolon-separated list of dates when the student was absent (formatted as MM/DD/YYYY)
- ReasonGiven (Y/N)
  - Indicates whether a reason was provided for absences
- AbsenceNote
  - A detailed text description of why the student was absent
  - There is a different, unique absence note for the number of excused absences
  - These were AI generated
- ReasonCategory
  - A semicolon-separated list categorizing the absence reasons
  - Reasons include Medical, Athletics, Family, Appointment, Mental health, Transportation
- StudentSchoolPresenceUpdated: This sheet is a modified version of the StudentSchoolPresence sheet, with the "ReasonCategory" column split into separate columns. This change was made to simplify future coding tasks. Although we have not yet implemented any code using "ReasonCategory", pre-splitting the values eliminates the need for additional processing to separate them later, as the original format used a ";" as a delimiter.

