ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE

Paper ID #37323

Andromeda in the Classroom: Collaborative Data Analysis for 8th Grade Engineering Design

Mia Taylor

Mia Taylor is a Master's Student (MS) at Virginia Tech. She holds a Bachelor of Science (BS) degree in Computer Science from the same institute. She has industry experience in applied natural language processing, software engineering, and full-stack web development. Her research interests include applied natural language processing, interactive machine learning, and information visualization. She hopes to continue working in these research areas in industry after her 2022 graduation.

Leanna House

Andromeda in the Classroom: Collaborative Data Analysis for 8th Grade Engineering Design

1 Introduction

Collaborative data analysis enables students to explore and analyze multi-dimensional data together. This case study shows how collaborative data analysis can be successfully integrated into teaching the engineering design process using novel software called Andromeda.

Andromeda is an interactive tool that seeks to make analysis of multi-dimensional data accessible for novices. Multi-dimensional data is often difficult for people to understand because it relates multiple variables, for example all the many factors that might make a racecar design faster or slower. Andromeda visualizes the data in a two-dimensional plot, such that data points (e.g. racecar designs) that are considered similar in the multi-dimensional sense are plotted closer to each other. Analysts can interactively explore relationships and trade-offs in the data by dragging data points and sliders. The original purpose of Andromeda was to enable individual analysts, regardless of skill, to explore multi-dimensional data interactively and in parallel with their personal sensemaking processes. Now, we show that Andromeda has collaborative utility in a classroom setting and as a public, educational resource for teachers and students.

This case study performs an observational, qualitative analysis on the collaborative use of Andromeda in an 8th grade technology education class. Students were given two engineering projects through WhiteBox Learning: Survival Shelter 2.0 and Dragster 2.0. WhiteBox Learning is a web-based STEM education software that allows students to learn STEM concepts, such as introductory physics, and practice the engineering design process. Survival Shelter 2.0 and Dragster 2.0 are two design projects that let students create an emergency survival shelter for hikers and a CO_2 racecar, respectively. In this case, students used WhiteBox Learning to create, analyze, and simulate their project designs. Between design iterations, the class explored their designs in Andromeda with the teacher acting as the facilitator. That is, the teacher uploaded data describing the students' projects to Andromeda; each point in the visualization represented a student's design. With the teacher controlling Andromeda, students used Andromeda to visualize, analyze, and compare their designs.

Data were collected from students, along with impressions from their teacher. We use these data to assess the success of the class collaboration. We discuss the potential implications of Andromeda as a public, educational resource. We also provide an example class activity aligned with Virginia's proposed Standards of Learning in data science. Hopefully, this will inspire educators to introduce collaborative data analysis tools such as Andromeda into their classrooms.

2 Background and Related Work

2.1 Andromeda

Andromeda is an interactive visual analytics tool that was originally designed to enable data analysts of all skill levels to explore multi-dimensional data [1]. Andromeda is available publicly as a web application [2]. The application creates a scatterplot-like 2D visualization of multi-dimensional data using Weighted Multi-dimensional Scaling (WMDS) [3]. WMDS is a dimensionality reduction algorithm that, given high-dimensional data and variable weights as input, can output a corresponding low-dimensional dataset that follows a "near == similar" metaphor. The variable weights represent how WMDS prioritizes the variable's value when outputting the low-dimensional dataset. This means that two points in the projected scatterplot that are near each other are considered more similar in the high-dimensional variables with the most weight than points farther away.

Andromeda offers two forms of interaction: parametric interaction (PI) and observation-level interaction (OLI). Users apply PI by interacting with the slider bars of the weights on the right-hand panel of the application. This directly manipulates the weights assigned to the variables in the dataset for WMDS. Andromeda then updates the projection based on the new weight values. Users apply OLI by interacting with the projection itself. After dragging points to different locations on the projection and clicking the "Update Layout" button, Andromeda solves for the optimal weights that preserve the user-defined projection. Then, Andromeda updates its display with the new weights. With PI and OLI, users can explore the variable weights and the projections to better understand the high-dimensional data. An example screenshot of Andromeda can be found below in Figure 1.

Figure 1. Example screenshot of Andromeda. Each point is assigned a unique number and represents a student's design for the shelter assignment. Design 3 was different from the rest.

Previous Andromeda studies focus on the way data analysts, usually college students, interact with Andromeda [4, 5, 6, 7]. *Be the Data* [8, 9] is the only previous study on Andromeda that analyzes how K-12 students interact with data. Virginia Tech hosted *Be the Data* events for various student groups ranging from 3^{rd} grade to the undergraduate level. In *Be the Data*, students embodied data points and physically moved in a room to interact with WMDS. For example, in a dataset about animals, each student represented an animal and their physical location in the workshop room was projected onto a large screen with Andromeda. Then, if a student moved from one side of the room to the other, their corresponding animal data point would move in the Andromeda projection as well. Collectively, classes would answer questions such as "What characteristics differentiate good pets from bad pets?" Data from these events suggested that the students were engaged and enabled to learn about high-dimensional data and analytics.

While *Be the Data* proved a successful collaborative use of Andromeda, *Be the Data* also required large-scale technological infrastructure such as motion detection systems, cameras, and large projectors. In this paper, we discuss the collaborative educational benefits and potential drawbacks of Andromeda in a typical classroom setting.

2.2 WhiteBox Learning

WhiteBox Learning is a software learning system that offers activity-based STEM assignments for 6^{th} to 12^{th} grade science and pre-engineering classes [10]. Through WhiteBox Learning, students can research, design, analyze, simulate, and compete with classmates. Students can construct 3D renderings and print plans to physically construct their designs. This option lets students complete the engineering design process from research to end product.

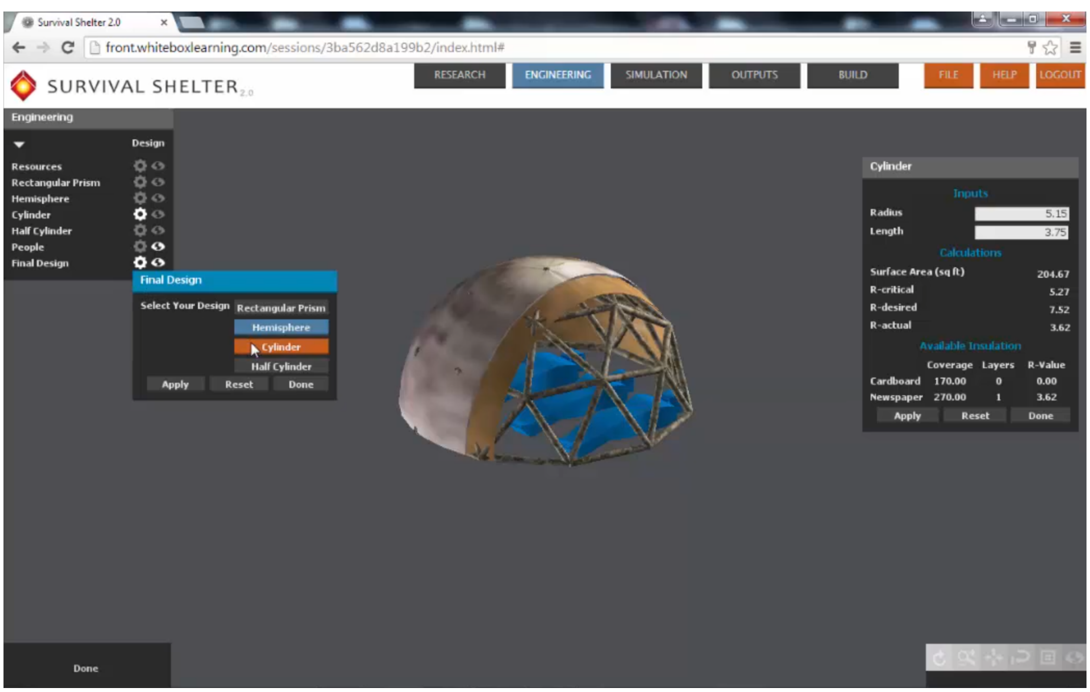


Figure 2. Example screenshot of the WhiteBox learning Shelter 2.0 assignment.

The WhiteBox Learning assignments used in this case study are the Survival Shelter 2.0 (referred to as shelter) [11] and Dragster 2.0 (referred to as dragster) [12]. An example 3D rendering of a shelter can be found above in Figure 2. The shelter assignment enables students to design an emergency shelter for hikers who are trapped in a blizzard. After students perform initial research on topics like energy and heat, they can design, analyze, and compete with their designs within WhiteBox Learning. The most comfortable shelter, in terms of temperature and space, wins the competition. The dragster assignment enables students to design a CO_2 dragster (racecar). Similar to the shelter assignment, students research relevant concepts such as Newton's Second Law of Motion and net force. Then, they design, analyze, and compete with their designs. The fastest car wins the competition. In this case study, the students' designs from these assignments were used as input to Andromeda. Andromeda was then used to enable students to learn about similarities and trade-offs in their designs.

3 Methods

While exploring computer science research for personal reasons, the teacher discovered Andromeda and its potential for their 8^{th} grade technology education class. With no outside involvement from the Andromeda team, the teacher decided to use Andromeda alongside WhiteBox Learning.

To prime the students for Andromeda, the teacher introduced Andromeda by exploring a dataset describing animals. This allowed the students to understand the basic interactions available within Andromeda with an approachable multi-dimensional dataset. The class explored Andromeda by answering questions such as "What makes a good pet?" and "What differentiates wild and domestic animals?". The class grouped together animals they considered similar and increased the weight of variables they suspected would answer the questions. Once the students were familiar enough with Andromeda to understand the software's capabilities, they then used Andromeda for their WhiteBox Learning assignments.

The class completed the two assignments, shelter and dragster, through WhiteBox Learning. They created and competed their designs using WhiteBox Learning simulations. Using the engineering design process, the students designed multiple iterations for each assignment. Their intermediate designs were converted to high-dimensional data where each design decision (such as material for the shelter or surface friction for the dragster) represented a single dimension, or variable, of data. The teacher consolidated the students' designs into a single, multi-dimensional dataset to be used in Andromeda.

Using this multi-dimensional dataset, the students who opted-in explored and interacted with Andromeda's visualizations. The teacher controlled the web version of Andromeda and shared their screen with these students. Of the students, about 30% were present on a video-conferencing software. The majority of discussion took place in the classroom environment, yet it was not uncommon for virtual students to add insights via chat. Due to COVID-19 protocols during 2020-2021, the students were not able to physically build their designs.

The teacher gave the students a survey after each assignment that asked students to rate the following statements on a scale from 0 (strongly disagree) to 10 (strongly agree).

1. Analyzing data with Andromeda is fun.
2. I know what is meant by the term high-dimensional data.
3. Understanding math makes using Andromeda easier.
4. I know which interaction (map/observable vs. slider/parametric) to move to better understand the data.
5. Andromeda helped me understand the raw data.
6. Using Andromeda helped me understand how different variables affect my design.

Students rated statements 7a and 7b after completing the shelter and dragster assignments, respectively.

7a. Using Andromeda allowed me to design a better shelter.

7b. Using Andromeda allowed me to design a better dragster.

The survey described above also included free response questions about how the students used the engineering design process, including application of knowledge and skills gained through the use of tools such as Andromeda and WhiteBox Learning. After the teacher collected the survey data throughout the 2020-2021 school year, the Andromeda team became involved in summer 2021 to analyze the data. The team selected not to compare the data to similar, past assignments due to the large number of unique circumstances under which the class was conducted.

4 Results

The following heatmaps represent students' responses to each survey. For all questions, responses ranged from 0 (strongly disagree) to 10 (strongly agree). The surveys were offered to 36 students, but only 30 and 26 students voluntarily completed the shelter and dragster survey, respectively. There were no incomplete responses when the surveys were completed. Most students were enrolled in pre-algebra level math. While they had previous experience in technology education, they had no educational background in multi-dimensional data or math related to visualizing multi-dimensional data.

4.1 Shelter

We see in Figure 3, below, that the statements with the highest and lowest average are "Using Andromeda helped me to design a better shelter." and "Using math makes using Andromeda easier.", with respective means of $\bar{x} = 7.93$ and $\bar{x} = 4.20$. The statement with the most variability in responses is "Using math makes using Andromeda easier." ($s = 2.68$). The statement with the least amount of variability is "Analyzing data with Andromeda fun." ($s = 1.74$). The statement "Using Andromeda allowed me to understand the raw data." received relatively high responses ($\bar{x} = 7.87$), while the statement "I know what is meant by the term high-dimensional data." received the second-lowest average ($\bar{x} = 6.90$).

	$\bar{x}$ (s)	0 (Strongly disagree)	1	2	3	4	5	6	7	8	9	10 (Strongly agree)
Analyzing data with Andromeda is fun.	6.87 (1.74)	0	0	0	0	4	2	7	4	10	0	3
I know what is meant by the term high-dimensional data.	6.90 (2.16)	0	1	0	1	3	2	4	2	13	1	3
Understanding math makes using Andromeda easier.	4.20 (2.68)	0	5	5	4	6	1	2	1	3	3	0
I know which interaction (map/observable vs. slider/parametric) to move to better understand the data.	7.80 (1.77)	0	0	0	0	1	3	4	4	4	9	5
Andromeda helped me understand the raw data.	7.87 (1.80)	0	0	0	0	2	2	3	3	6	9	5
Using Andromeda helped me understand how different variables affect my design.	7.60 (1.83)	0	0	0	0	2	1	7	5	2	8	5
Using Andromeda allowed me to design a better shelter.	7.93 (2.03)	0	0	1	0	0	3	3	4	5	5	9

Figure 3. Survey responses after the shelter assignment. The mean ($\bar{x}$) and standard deviation (s) for each survey question is provided with a visual count of responses.

4.2 Dragster

The results are similar for the dragster assignment as shown in Figure 4 below. The statement "Using Andromeda helped me understand the raw data." has the highest average ($\bar{x} = 8.96$) and the lowest variability ($s = 1.34$). Similar to the shelter assignment, the statement with the lowest average is "Using math makes using Andromeda easier." ($\bar{x} = 4.15$). This statement also has the highest variability with ($s = 2.75$). The statements "Using Andromeda helped me understand how different variables affect my design." and "I know which interaction ... to move to better understand the data." also had high average responses with $\bar{x} = 8.38$ and $\bar{x} = 8.27$, respectively. The impact of the results represented in Figures 3 and 4 are in the next section of this paper.

	$\bar{x}$ (s)	0 (Strongly disagree)	1	2	3	4	5	6	7	8	9	10 (Strongly agree)
Analyzing data with Andromeda is fun.	7.62 (1.65)	0	0	0	0	1	2	4	4	6	6	3
I know what is meant by the term high-dimensional data.	7.42 (2.21)	0	1	0	1	0	3	1	5	6	5	4
Understanding math makes using Andromeda easier.	4.15 (2.75)	0	5	4	3	4	3	2	2	0	1	2
I know which interaction (map/observable vs. slider/parametric) to move to better understand the data.	8.27 (1.59)	0	0	0	0	0	3	1	2	6	8	6
Andromeda helped me understand the raw data.	8.96 (1.34)	0	0	0	0	0	1	1	1	4	7	12
Using Andromeda helped me understand how different variables affect my design.	8.38 (1.86)	0	0	0	0	1	1	4	2	1	7	10
Using Andromeda allowed me to design a better dragster.	7.77 (2.12)	0	0	1	0	0	3	3	4	4	3	8

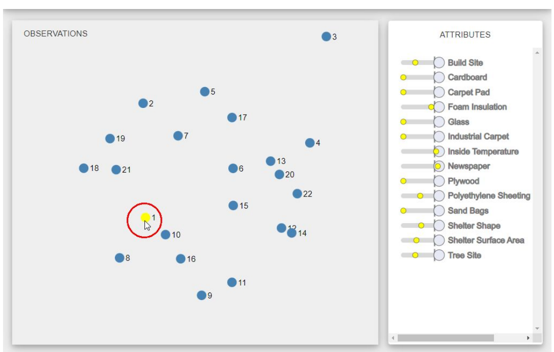
Figure 4. Survey responses after the dragster assignment. The mean ($\bar{x}$) and standard deviation (s) for each survey question is provided with a visual count of responses.

5 Discussion

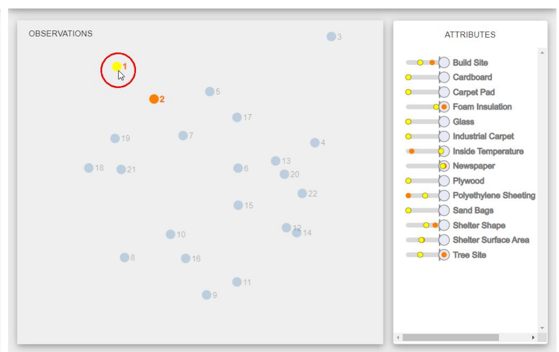
The following discussion puts the results presented in the previous section into context with impressions shared by the teacher. We break down context into four themes: collaboration, data analysis, design influence, and sentiment.

5.1 Collaboration: Andromeda supported student collaborations

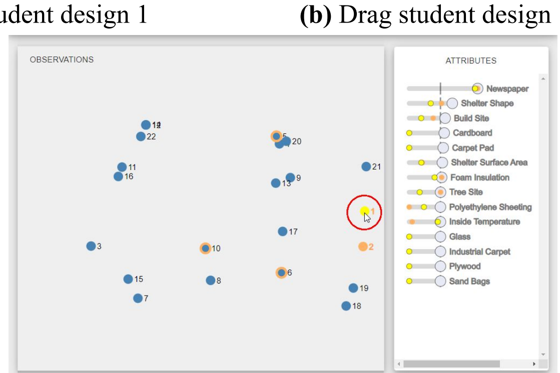
This case study presents findings on the collaborative utility of Andromeda. The class analyzed and discussed their designs visualized within Andromeda together. According to the students' teacher, "students didn't start discussion with other students until Andromeda". As the students became more comfortable exploring their data within Andromeda, they requested more interactions from their teacher on the shared screen. For example, two friends had different internal shelter temperatures and they wanted to analyze the reason why. They requested that the teacher drag their two points together and update the layout. This used Andromeda's observation-level interaction as discussed in subsection 2.1. After seeing the updated layout and variable weights, the students discussed the most prominent difference in their designs. Example images are shown below in Figure 5. Interactions like this demonstrate the potential for Andromeda to serve as a conversation piece as a student group collaboratively explores a dataset.



(a) Select student design 1



(b) Drag student design 1 next to student design 2



(c) The weights and visualization are updated with new weights based on the similarities and differences of students 1 and 2. They both heavily used Newspaper in their shelter designs. Students 18, 19, 21 also had similar designs to 1 and 2.

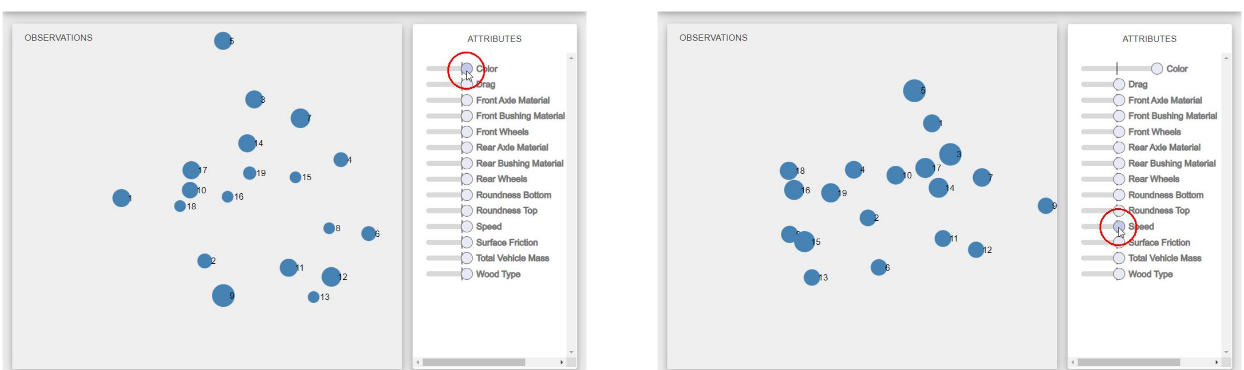
Figure 5. An example of two students comparing and contrasting their shelters. Each point is assigned a unique number and represents a student's design for the shelter assignment. The exact scaling has been altered for purposes of presentation in this paper.

5.2 Data Analysis: Andromeda helped students practice data analysis

Andromeda can be engaging without requiring an extensive math or data analysis background. The students did not fully understand dimensionality reduction before, or after, using Andromeda. Regardless, the class was able to meaningfully interact with Andromeda to improve their designs without being bogged down by Andromeda's underlying math. These results are supported by the student responses to the statement "Understanding math makes using Andromeda easier." ($\bar{x}_{shelter} = 4.20$, $\bar{x}_{dragster} = 4.15$). While not understanding the math, the students were still able to gain insights into their high-dimensional data.

Before Andromeda, "students did not know what high-dimensional data was", according to their teacher. The students began to feel more confident in their understanding of the concept of high-dimensional data once they used Andromeda. This is shown by the positive responses to the statement "I know what is meant by the term high-dimensional data." ($\bar{x}_{shelter} = 6.90$, $\bar{x}_{dragster} = 7.42$). Students also felt confident in their ability to interact with high-dimensional data by the positive responses to the statement "I know which interaction (map/observable vs. slider/parametric) to move to better understand the data." ($\bar{x}_{shelter} = 7.80$, $\bar{x}_{dragster} = 8.27$).

While introducing concepts like high-dimensional data, Andromeda also allowed students to approach their design analytically rather than anecdotally. Andromeda helped the students better understand their designs by giving them "a new way to analyze the designs", according to their teacher. Students tested negative hypotheses by analyzing variables they suspected would not impact their designs' performances. Figure 6 shows an example interaction where the class increased the weight on the color variable for their dragsters. This used Andromeda's parametric interaction as discussed in subsection 2.1. As expected, faster dragsters were not plotted near each other. Students were able to verify their negative hypotheses.

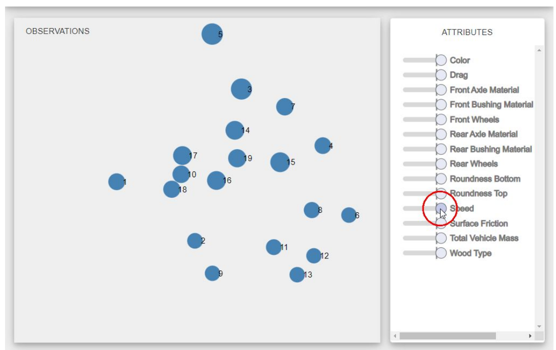


(a) Increase the weight of the color variable (attribute).

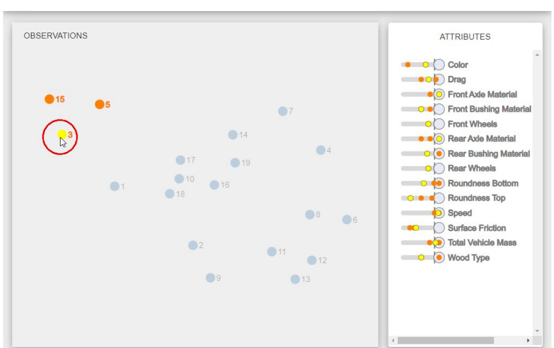
(b) Hover over speed so each circle size is proportional to the design's speed. There is no clear pattern of circle size, so dragster color does not correlate with speed, as expected.

Figure 6. An example of the class analyzing if color correlates with dragster speed. Each point is assigned a unique number and represents a student's design for the dragster assignment. The exact scaling has been altered for purposes of presentation in this paper.

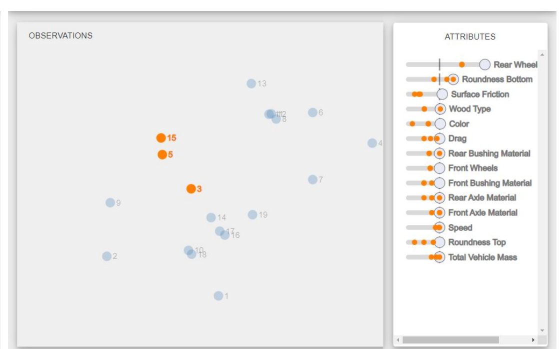
Another common interaction the class used was to cluster the top-performing designs together via Andromeda's observation-level interaction and look at the most heavily weighted variables. These variables represented the most common values the top-performing designs had in common compared to the worse-performing designs. In the example images from Figure 7, the top-performing dragsters have similar values for total vehicle mass and surface friction. Interactions like these helped students understand their data and what variables affected their performance. This is supported by the high averages on the following statements: "Using Andromeda helped me understand the raw data." ($\bar{x}_{shelter} = 7.87$, $\bar{x}_{dragster} = 8.96$) and "Using Andromeda helped me understand how different variables affect my design." ($\bar{x}_{shelter} = 7.60$, $\bar{x}_{dragster} = 8.38$). Through using Andromeda, students were able to understand and interact with their high-dimensional data in a new and pedagogically impactful way.



(a) Hover over speed so each circle size is proportional to the design's speed. The student designs numbered 5, 3, and 15 are fastest.



(b) Drag student designs 5, 3, and 15 together



(c) The weights and visualization are updated based on the similarities and differences of the fastest dragsters. Large rear wheels and round bottom were important variables.

Figure 7. An example of the class analyzing the fastest dragsters. Each point is assigned a unique number and represents a student's design for the dragster assignment. The exact scaling has been altered for purposes of presentation in this paper.

5.3 Design Influence: Andromeda influenced the engineering design process

The class explored students' overall design and individual components within their designs in Andromeda. The class did this together by identifying the top-performing designs, such as the fastest dragster based on WhiteBox Learning simulations, and comparing the designs' projections in Andromeda. Once the class identified which variables were of importance, students focused their next design iteration on tuning those variables in WhiteBox Learning. Andromeda helped students prioritize their design work and served as an impactful tool in the analysis stage of the engineering design process.

According to the teacher, "students produced higher quality products than in previous years- a direct result of the analysis and rapid prototyping provided by the use of Andromeda and WhiteBox, respectively." This statement is supported by the student responses to the following statements: "Using Andromeda allowed me to design a better shelter." ($\bar{x}_{shelter} = 7.93$) and "Using Andromeda allowed me to design a better dragster." ($\bar{x}_{dragster} = 7.77$). Students were able to meaningfully analyze designs in Andromeda and use this analysis to help create better end products.

5.4 Sentiment: The class had a positive experience using Andromeda

A class' attitude toward a new concept, such as data analysis and software, is important in maintaining active engagement. The teacher found that Andromeda has a "low intimidation factor" which is important for its adoption by non-expert data analysts. As a whole, the class enjoyed using Andromeda. Students generally found Andromeda fun and useful. When asked to rate the statement "Analyzing data with Andromeda is fun.", students responded with averages $\bar{x}_{shelter} = 6.87$ and $\bar{x}_{dragster} = 7.62$. Students also responded to the post-survey saying Andromeda "is going to be good for upcoming projects" and is "kind of cool". The class was able to use Andromeda meaningfully while learning new concepts and enjoying the learning experience.

6 Limitations

The results show that the class benefited from using Andromeda, however, this case study can not definitively show that a class without Andromeda would have performed worse on the assignments. Though the teacher reflected on class projects from previous years without Andromeda, there is no formal control group for this study. While the class agreed that using Andromeda helped them design better products, there is no data to support causation. A large, controlled experiment is necessary to further explore the causal influence of Andromeda on K-12 engineering design projects.

Though students were able to explore high-dimensional data and improve their designs, the projections visualized by Andromeda cannot be considered truly accurate. Web-based Andromeda in its current form requires all variables to be continuous. However, in this case study, the definition of continuous was conflated with categorical. Thus, nominal categorical variables, such as the dragster colors, were assigned numeric values to import data into Andromeda. This means that Andromeda considered, for example, red to have a lesser distance to blue when compared to the distance to green. In reality, these colors should be considered equidistant. This conflation may have affected the class' visualizations, however, it does not take away from the fact that students were able to explore and understand their high-dimensional data for their engineering design projects. This error highlights potential feature improvements and the responsibility of the creators of Andromeda to communicate the software's assumptions about its input data.

7 Next Steps

This case study highlights Andromeda as an educational tool within engineering education. At the time of this paper, Andromeda does not support the idea of a dependent variable, such as the final speed of the students' dragsters. For this class, it would have been useful to conveniently display each dragster's speed since that was what the students were seeking to improve. This could be supported in Andromeda by designating a specific variable to display on each data point. While Andromeda could still take the variable as input to its visualization algorithm, the option to explicitly display a specified variable's value on each point may help users who are strongly interested in optimizing a single aspect of the data. This would benefit the class in this case study and Andromeda users in general.

Another future improvement of Andromeda could be to support categorical variables through using a different distance metric that supports numerical (continuous) and categorical variables. This would address the incorrect assumption in this study, as discussed in the previous section, and improve Andromeda's flexibility as a data analysis tool. Regardless, the public Andromeda web page should be updated to specifically define what types of data it supports.

While not used in this case study, Andromeda also supports collaborative analysis that enables multiple users to interact with the same, synchronized visualization. The class would then be able to manipulate the visualization together rather than having the teacher interact with Andromeda based on the class discussion. Exploring the use of this capability in the classroom would help to further illuminate the effects of collaborative data analysis in engineering education.

This case study shows Andromeda's potential to be used as a collaborative data analysis tool in the classroom. Though we discussed its use in K-12 classrooms, we fully expect Andromeda to foster collaborations at the college level as well. We recommend Andromeda to teachers and professors who want to encourage both group and whole-class collaborations. As shown, Andromeda can be used in combination with services such as WhiteBox Learning to teach students about the engineering design process. It can also be taught as a standalone lesson, using the provided sample dataset about animals, to teach introductory data science principles. To encourage educators to introduce Andromeda into their classrooms, we provide an example class activity, with and without solutions, in the appendix. This activity, like our case study, aligns with three guidelines within the proposed data science Standards of Learning for the state of Virginia [13]. These guidelines cover topics such as interpreting data in visualizations, formulating hypotheses from data, and utilizing appropriate tools for data analysis. We hope that educators will explore collaborative data analysis tools like Andromeda to enhance the K-12 data science education experience.

References

[1] J. Self, M. Dowling, J. Wenskovitch, I. Crandell, M. Wang, L. House, S. Leman, and C. North, "Observation-level and parametric interaction for high-dimensional data analysis," *ACM Transactions on Interactive Intelligent Systems*, vol. 8, pp. 1–36, 06 2018.

[2] Andromeda. [Online]. Available: http://nebula.cs.vt.edu/cosmos/andromeda.html

[3] J. B. Kruskal and M. Wish, *Multidimensional Scaling*, ser. Quantitative Applications in the Social Sciences. Thousand Oaks, CA: SAGE Publications, Aug. 1978.

[4] M. Wang, J. Wenskovitch, L. House, N. Polys, and C. North, "Bridging cognitive gaps between user and model in interactive dimension reduction," *Visual Informatics*, vol. 5, no. 2, pp. 13–25, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2468502X21000085

[5] J. Zeitz, N. Self, L. House, J. R. Evia, S. Leman, and C. North, "Bringing interactive visual analytics to the classroom for developing eda skills," *J. Comput. Sci. Coll.*, vol. 33, no. 3, p. 115–125, jan 2018.

[6] J. Z. Self, R. K. Vinayagam, J. T. Fry, and C. North, "Bridging the gap between user intention and model parameters for human-in-the-loop data analytics," in *Proceedings of the Workshop on Human-In-the-Loop Data Analytics - HILDA '16*. ACM Press, 2016. [Online]. Available: https://doi.org/10.1145%2F2939502.2939505

[7] J. Z. Self, N. Self, L. House, S. Leman, and C. North, "Improving students' cognitive dimensionality through education with object-level interaction," 2014.

[8] X. Chen, J. Z. Self, L. House, J. Wenskovitch, M. Sun, N. Wycoff, J. R. Evia, S. Leman, and C. North, "Be the data: Embodied visual analytics," *IEEE Transactions on Learning Technologies*, vol. 11, no. 1, pp. 81–95, 2018.

[9] X. Chen, L. House, J. Z. Self, S. Leman, J. R. Evia, J. T. Fry, and C. North, "Be the data: An embodied experience for data analytics," in *2016 Annual Meeting of the American Educational Research Association (AERA)*, 04/2016 2016, p. 20.

[10] A flinn scientific company. [Online]. Available: https://www.whiteboxlearning.com/

[11] Survival shelter 2.0. [Online]. Available: https://www.whiteboxlearning.com/applications/shelter

[12] Dragster 2.0. [Online]. Available: https://www.whiteboxlearning.com/applications/dragster

[13] V. D. o. Education. Data science. [Online]. Available: https://www.doe.virginia.gov/testing/sol/standards_docs/data-science/index.shtml

Appendix A

Andromeda Activity Worksheet without Solutions

To complete the activity, go to `http://nebula.cs.vt.edu/cosmos/andromeda.html`. On the top right, click the following: "Start Here", "Load Data", "highD/Animal_Data_square.csv", and "Use Selected Data". This loads a dataset describing animals where the minimum value is 0 and the maximum value is 100. Answer the following questions using Andromeda.

1. 0 Approximately how fierce is a Tiger? 100
2. 0 Approximately how timid is a Wolf? 100
3. 0 Approximately how big is a Deer? 100
4. If you increase the weight on Swims, will the Dolphin and Otter move closer together or farther apart? Circle one.
   a. Closer Together
   b. Farther Apart
5. If you increase the weight on Size, will the Squirrel and Giant Panda move closer together or farther apart? Circle one.
   a. Closer Together
   b. Farther Apart
6. What differentiates a Mouse, Raccoon, and Squirrel from a Blue Whale and a Dolphin? Fill in the 3 blanks.

   ______________ ______________ ______________

7. What differentiates a Tiger, Squirrel, and Elephant from a deer? Fill in the 3 blanks.

   ______________ ______________ ______________

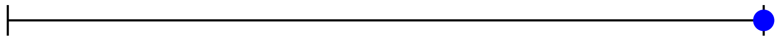
8. What differentiates a good pet from a bed pet? Fill in the 3 blanks.

   ______________ ______________ ______________

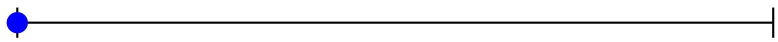
Appendix B

Andromeda Activity Worksheet with Solutions

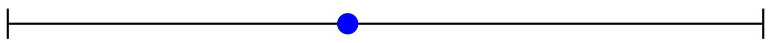
To complete the activity, go to `http://nebula.cs.vt.edu/cosmos/andromeda.html`. On the top right, click the following: "Start Here", "Load Data", "highD/Animal_Data_square.csv", and "Use Selected Data". This loads a dataset describing animals where the minimum value is 0 and the maximum value is 100. Answer the following questions using Andromeda. Solutions in blue. For questions 6-8, a few possible solutions are listed.

1. 

   0 Approximately how fierce is a Tiger? 100

2. 

   0 Approximately how timid is a Wolf? 100

3. 

   0 Approximately how big is a Deer? 100

4. If you increase the weight on Swims, will the Dolphin and Otter move closer together or farther apart? Circle one.

   a. Closer Together

   b. Farther Apart

5. If you increase the weight on Size, will the Squirrel and Giant Panda move closer together or farther apart? Circle one.

   a. Closer Together

   b. Farther Apart

6. What differentiates a Mouse, Raccoon, and Squirrel from a Blue Whale and a Dolphin? Fill in the 3 blanks.

   swims, furry, claws

7. What differentiates a Tiger, Squirrel, and Elephant from a Deer? Fill in the 3 blanks.

   timid, grazer, size

8. What differentiates a good pet from a bed pet? Fill in the 3 blanks.

   good pets are highly domestic, good pets are agile, bad pets are grazers