

Hanya Elhashemy^{1, 2}, Robert Parks¹, David Y J Kim¹, Evan Patton¹, and Hal Abelson¹
¹Massachusetts Institute of Technology, ²Technical University of Munich,
hanya.elhashemy@tum.de, rparks@mit.edu



Introduction

Barriers, such as cost, access, and complexity, can prevent students from building working prototypes for performing authentic data science practices across **the data science life cycle** — **collecting data, data cleanup, data visualization, and prediction** (Mike, 2022). Few products enable students to build projects that visualize data on mobile platforms — which, after all, is where many scientists and citizen scientists collect and consume data visualizations today (Silvertown, 2009).

The toolkit is a free suite of new, block-based-programming data science features that allow students to solve real-world problems that affect them or their community.

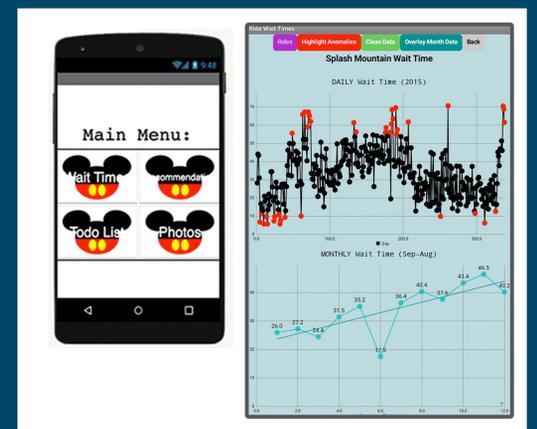
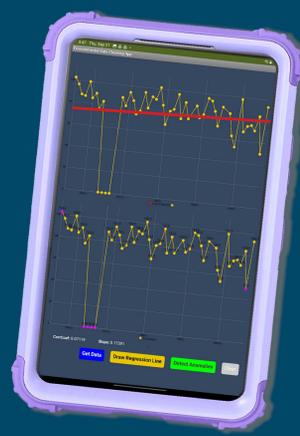
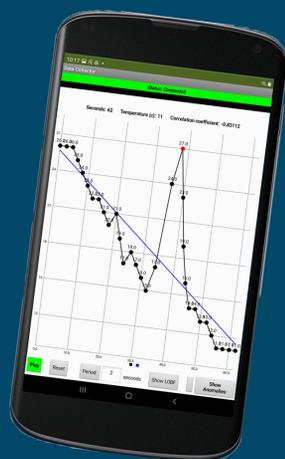
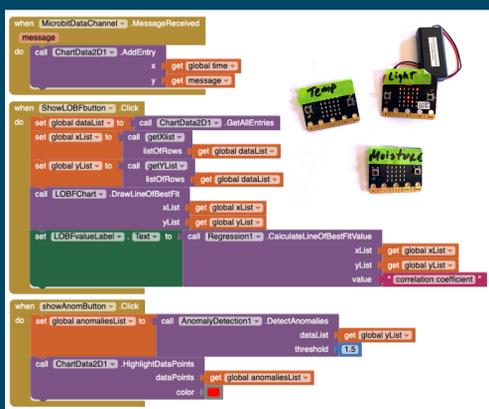
The demo and discussion are **relevant for middle and high school teachers and facilitators of out-of-school programs**. Participants can build simple mobile apps using MIT App Inventor and test these systems for themselves during the conference presentation.

Toolkit

- **Data Collection:** IOT connection and data management features for mobile phone apps
- **Data Visualization:** mobile charting for visualizing data in mobile applications

- **Data Cleanup:** anomaly detection and data cleaning techniques for mobile phone data analysis
- **Prediction:** linear regression for predictive insights in mobile applications

Examples



a.) **App Inventor programming blocks** show the App Inventor IoT component for receiving serial messages from a low-cost Bluetooth sensor.

b.) The **charting feature** shows a steady decrease in temperature briefly interrupted by an increase (a warm coffee cup on the sensor for illustration).

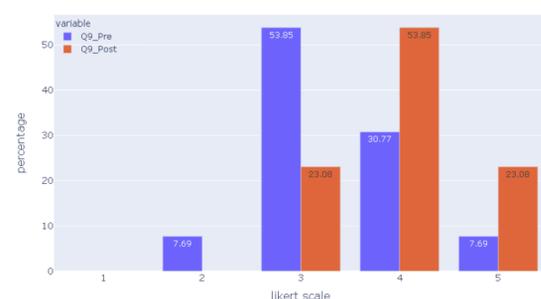
c.) The **linear regression line** in red shows a steady increase for the data. Data points in pink, at bottom, are identified as anomalies based on a z-score greater than 1.5.

d.) The **student-built app uses the data cleaning and predictive components** to analyze wait times for various rides at Disney World to determine the average time a child might have to wait in line.

June 2023 Pilot Study

Highlights from Elhashemy, H. (2023). From curious kids to data wizards: an empowerment approach to data science education (Unpublished master's thesis). Technical University Of Munich, Munich, Germany.

- A pilot study with 14 high school students measuring the efficacy of the data action framework on students' ability to understand the influence of data and use data science to impact their communities.
- "I thought data science was something you only observe rather than change and interact with"
- "Data science means to have a challenge you can solve by analyzing and exploring data"



(a) Results of Q9: I feel confident in my ability to use data to solve a real-life challenge.

Acknowledgements

Jacky Chen, Tommy Heng, Gisella K Kakoti, Susan Lane, Evaldas Latoškinas, Ava G Muffoletto, Matthew R Quispe, Arianna C Scott, Isabela Naty Sanchez Taipe, and Jennet Zamanova