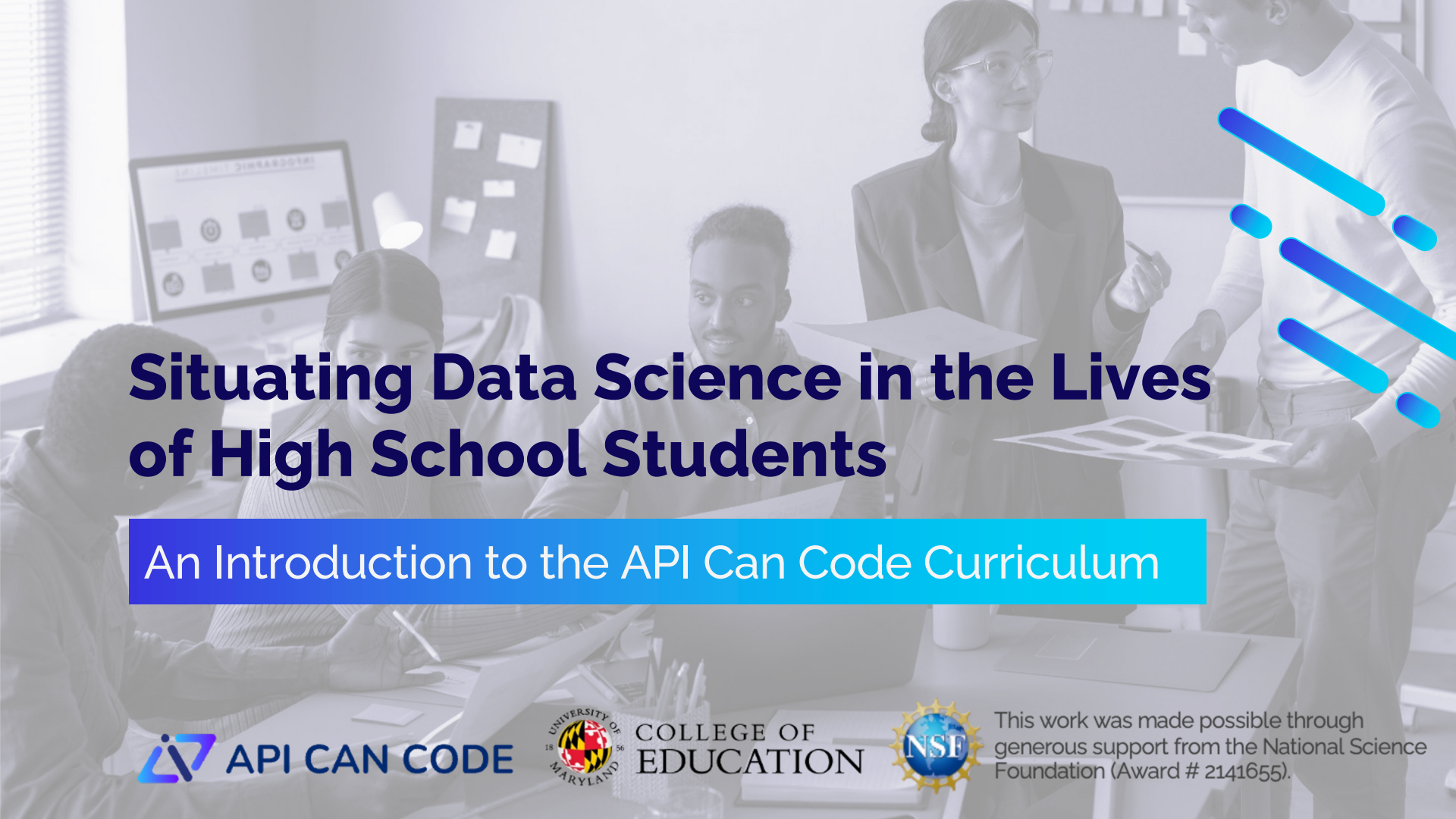


Situating Data Science in the Lives of High School Students

An Introduction to the API Can Code Curriculum

API CAN CODE

UNIVERSITY OF MARYLAND COLLEGE OF EDUCATION

This work was made possible through generous support from the National Science Foundation (Award # 2141655).

The API Can Code Team

David Weintrop

Principal Investigator
weintrop@umd.edu

Rotem Israel-Fishelson

Postdoctoral Researcher
rotemisf@umd.edu

Peter F Moon

Consulting Researcher
pmoon@umd.edu

Yue Xin

Doctoral Student
yxin21@umd.edu

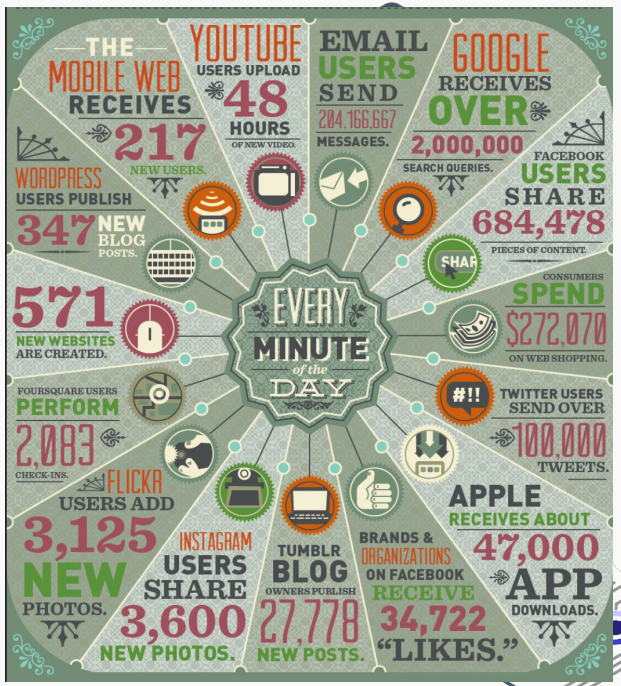
Introduction & Motivation

2013 Edition

(Data Never Sleeps, 2025)

Introduction & Motivation

2013 Edition



(Data Never Sleeps, 2025)

Introduction & Motivation

All students should learn foundational data science concepts and practices as part of their K-12 education.

(Dorsey et al., 2025; LaMar & Boaler, 2021)

Introduction & Motivation

"Data science bridges disciplines and thus should be introduced and taught across the curriculum in K-12 schools to help develop informed users of data… All subjects in school should recognize the contribution of data to their discipline and take curricular approaches that integrate data with disciplinary lessons where appropriate."

(NCTM, NSTA, ASA, NCSS, and CSTA joint position, April 2024)

Introduction & Motivation

Harnessing data that resonates with students' experiences and aspirations can deepen engagement and increase the likelihood of knowledge acquisition.

(Brooks et al., 2021; Lee et al., 2021)

Intro to Our Research Project

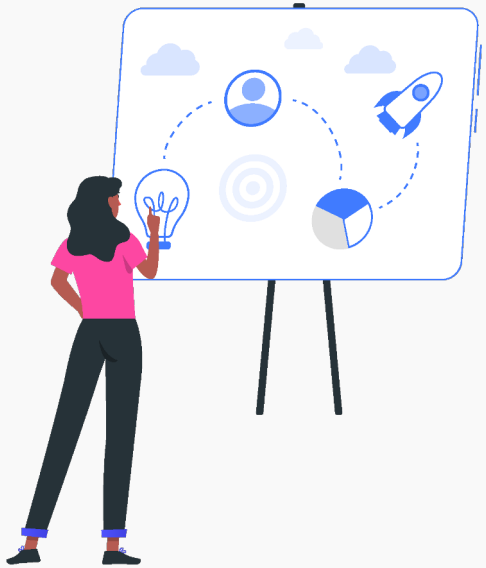
Goal of research:

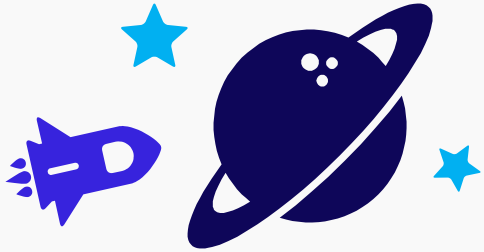
- Understand the state of Data Science in K-12
- Explore how to teach Data Science to high school students in ways that draw on their interests & identities

High Level Research Project Breakdown

Step 1: Study current high school Data Science

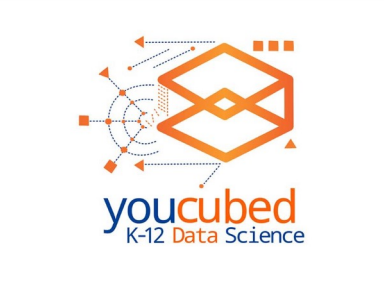
Step 2: Design, teach, and study an interest-driven high school data science curriculum





Studying the Landscape of Data Science

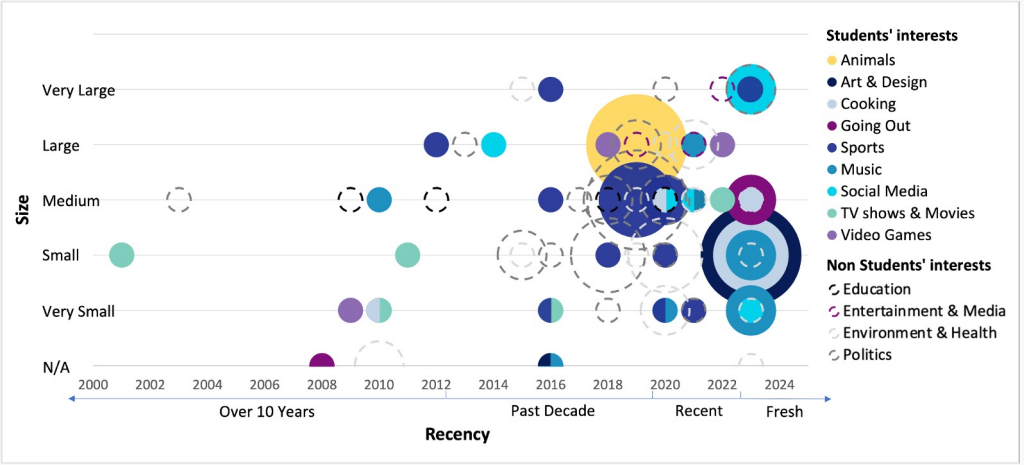
Data Science Curricula



Type Tech Language Topic Size Recency Proximity Interests

Data in K-12 Data Science

Analysis of **296 datasets across 4 curricula** and examination of their **alignment with the students' interests**



Israel-Fishelson, R., Moon, P. F., Tabak, R. & Weintrop, D. (2024). **Understanding the data in K-12 data science**. *Harvard Data Science Review*, 6(2).

K-12 Data Science Tools

Reviews **30 data science tools** used, or designed to be used, in K-12 data science education. Several attributes are used including **accessibility**.

RQ1: What capabilities does the tool possess regarding data science practices, data visualisation, statistical calculations, and extensibility?
RQ2: How do users interact with the tool, and how might the interaction support learning?
RQ3: What accessibility features does the tool provide?
RQ4: If the tool uses a block-based programming approach, what types of blocks does the tool include to support students learning data science?

Israel-Fishelson, R., Moon, P. F., Tabak, R., & Weintrop, D. (2023). **Preparing students to meet their data: An evaluation of K-12 data science tools**. *Behaviour & Information Technology*, Taylor & Francis.

K-12 Data Science Tools

API CAN CODE

UNIVERSITY OF MARYLAND

Comparison of Data Science Tools For K-12

- Tools' Capabilities
- Supported Interactions and Educational Features
- Supported Accessibility Features

Table 1: Tools' Capabilities

Tools	Data Manipulations	Statistical Capabilities	Data Visualization			Data Availability				Extensibility
			Tabular Display	Type of Graph	Creation Method	Built-in Data	Import	API	Export	
Blockly	Aggregating	N/A	×	Line	Code	×	✓	×	✓	✓
BlocklySQL	Filtering, Sorting, Aggregating	N/A	✓	N/A	N/A	✓	×	×	✓	✓
BlockPy	Filtering, Sorting	N/A	×	Scatterplot, Bar Chart, Line Chart, Box Plot, Histogram	Blocks / Code	✓	✓	×	✓	×
Bridges CS	Filtering, Deleting, Sorting, Aggregating	Correlation, Linear Regression	×	Line Chart	Code	✓	✓	×	✓	×
CODAP	Filtering, Deleting, Sorting, Aggregating	Correlation, Linear Regression	✓	Scatterplot, Bar Chart, Histogram, Box Plot, Map	GUI	✓	✓	✓	✓	✓

https://go.umd.edu/APICC_DS_Tools

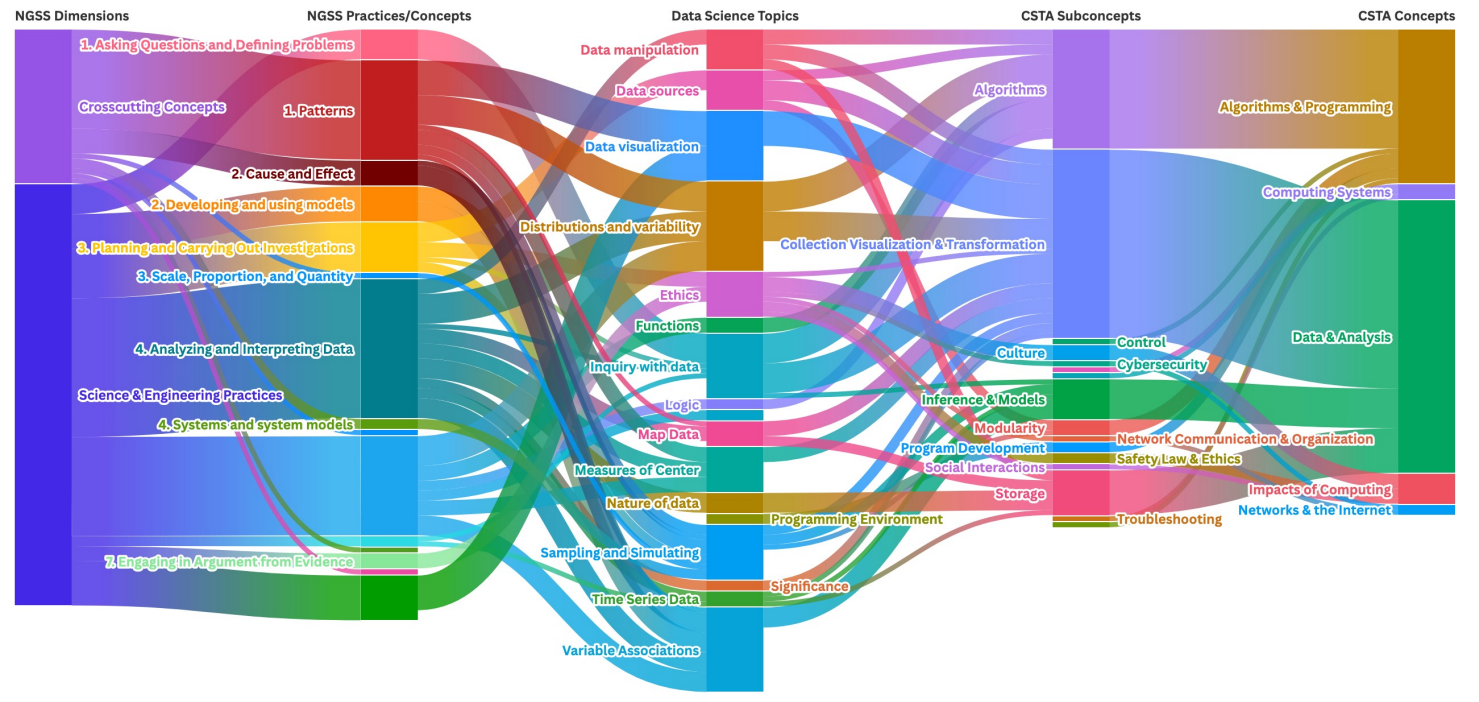
K-12 Data Science Standards

Analyzed alignment between CSTA high school standards and data science concepts and found **16 of the 17 core data science topics are present in CSTA standards.**

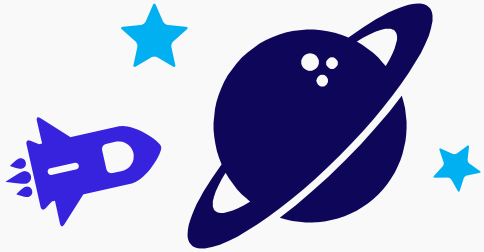
RQ 1. In what ways do data science topics align (and not align) with the CSTA high school standards?

Xin, Y., Moon, P. F., Israel-Fishelson, R., & Weintrop, D. (Under Review). **Data Science in Computer Science Classrooms: Insights from Standards Alignment and Student Perspectives**.

K-12 Data Science Standards



https://apicancode.umd.edu/standards.html



Designing an Interest Driven Curriculum

Create a data science curriculum that

- students find engaging and compelling
- informed by their values and voices

Participatory Design

- A research methodology that involves the end-users in the design process to ensure the outcome reflects their voices, values, and needs.
- PD was found to be effective in improving learning materials and curriculum development, as well as identifying students' interests.

(Coenraad et al., 2022; DiSalvo & DiSalvo, 2014)

Settings & Participants

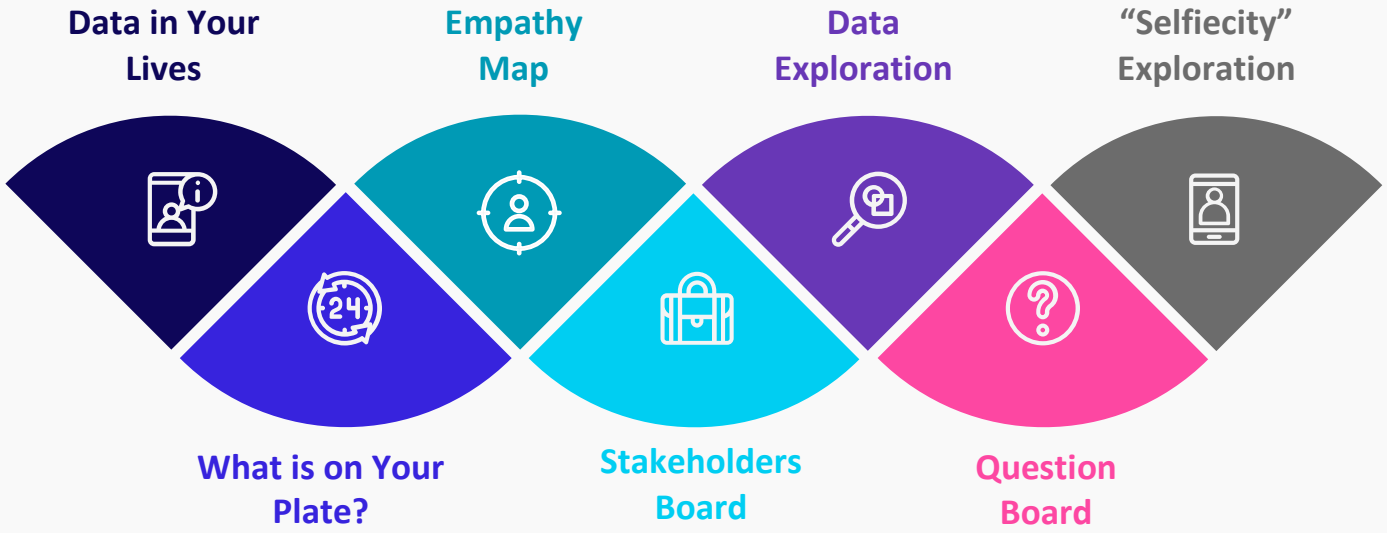
Urban, public charter high school

28 9-12 grade students

Gender	N
Male	17
Female	11
Race	
Black or African American	22
Native American	2
Hispanic	1
White	1

Participatory Design Activities

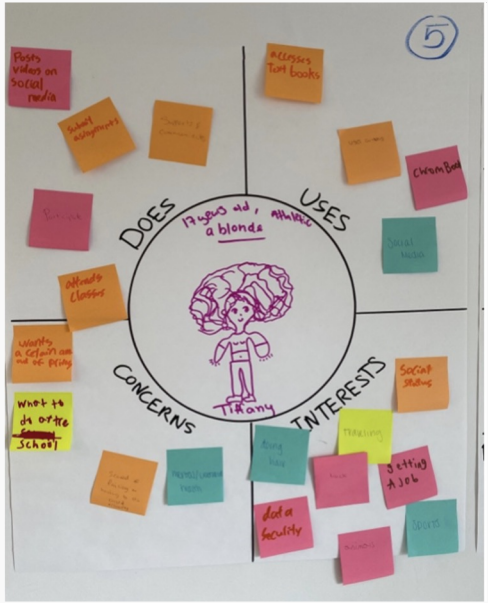
Discussions and 7 hands-on design activities:



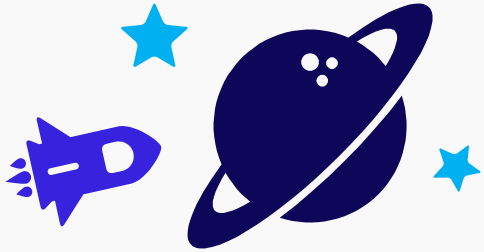
Israel-Fishelson, R., Moon, P. F., Pauw, D., & Weintrop, D. (2024). **Using Participatory Design to Gain Insight into How Students Make Sense of Data in Their Lives**. Proceedings of the Symposium on Learning, Design and Technology, 85–94.

Empathy Map

A User-Centered Design concept of a persona (Miaskiewicz & Kozar, 2011)



Code	Examples
Where	“Social Media (e.g., Instagram, Twitter)”; “Netflix”; “Hulu”; “Canvas”; “Google”; “Online shopping”; “iMessage”
Concerns	“Being tracked by the apps used”; “Being recorded without his consent”; “Someone wants to use your data against you”
Area of Interest	“Sports”; “Social Media”; “Video Games”; “Music”; “Movies and TV Shows”; “Animals”



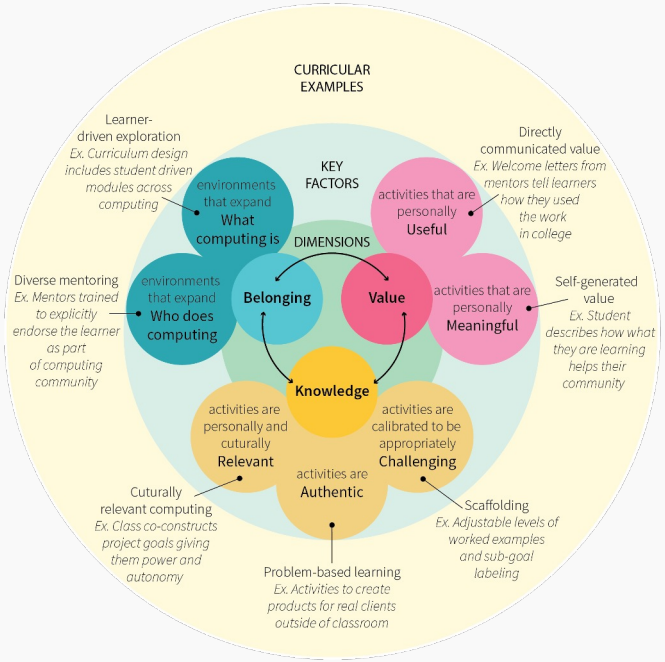
Meet

API CAN CODE

The API Can Code Curriculum

An interest-driven curriculum that introduces high-school students to computational foundations of data science by having them explore meaningful and authentic data that align with their interests using APIs.

Integrated Interest Development for Computing Education Framework



(Michaelis & Weintrop, 2022)

The Curricular Units

Unit 1: Data in Learners' Lives

- 1.1 Introduction to Data
- 1.2 Data Collection and its Purpose
- 1.3 Using Data: the DIKW Model
- 1.4 Sources of Data
- 1.5 Evaluating Datasets
- 1.6 Data Collection: Impact and Equity

Unit 2: Computational Foundations of Data Science

- 2.1 What is Data Science?
- 2.2 Manual Data Processing
- 2.3 Intro to Programming with EduBlocks
- 2.4 Accessing Data with APIs using RapidAPI
- 2.5 Preparing Data for Analysis
- 2.6 Data Analysis in Practice

Unit 3: Data Science Practices

- 3.1 Introduction to Data Visualization
- 3.2 Exploratory Analysis with CODAP
- 3.3 Graphs and Figures: One Variable
- 3.4 Graphs and Figures: Two Variables
- 3.5 Statistical Testing
- 3.6 Linear Models

Final Project: Doing Data Science

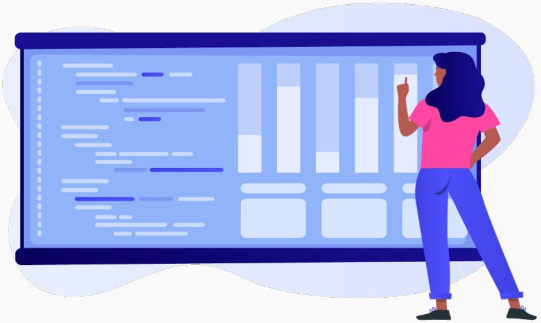
Finding Data
Access and evaluate a relevant dataset

Research Questions
Identify questions about this dataset

Data Refining
Filter, clean and trim the data

Data Visualization
Create data visualizations and interpret

Communicating Results
Share your conclusions and insights



The Curricular Approach

Unit 1:
Data in
Learners' Lives

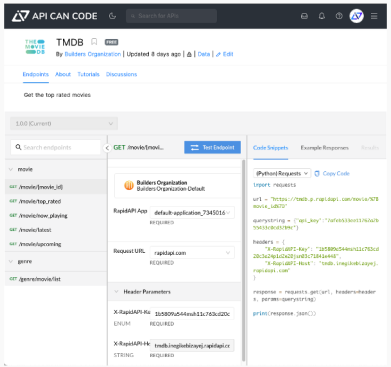
Rapid

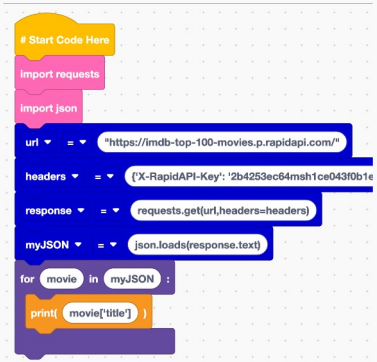
Unit 2:
Computational
Foundations

edublocks

Unit 3:
Data Science
Practices

CODAP





Integrating Students' Area of Interest

Topic	Lessons	Examples of Datasets
Social Media	1.2, 1.4, 2.4, 3.5	TikTok's Privacy Issues, Facebook Report
Music	1.5, 2.5	Billboard Hot 100 API
Sports	1.4, 2.6, 3.1, 3.5	NFL game scores, NBA API
Video Games	2.3, 2.5	Mario Kart Data
Movies & TV Shows	1.2, 1.6, 2.3, 2.4	IMDB Top 100 Movies API
Animals	1.5, 3.2, 3.3	Mammals Data, Four Seals Data
School / Education	1.4, 1.6, 2.2, 3.1, 3.3, 3.4	Schools' locations, Choosing a Collage
Going Out	3.2, 3.4	Roller Coasters Data
Environment & Health	1.1, 1.3, 3.1, 3.6	Plastic Bottle Waste, Earthquakes (USGS)
Food / Cooking	1.4, 1.5, 2.1, 3.2, 3.3	Starbucks Yearly Data, Food Deserts
Money / Jobs	1.1, 3.3, 3.5	Data Science Salaries,
Community	1.3, 1.4, 1.5, 1.6, 2.1, 3.5	DC COVID-19 Data, Zillow API

An API Can Code Example

"How much does a house in my neighborhood cost?"

Access the Zillow dataset

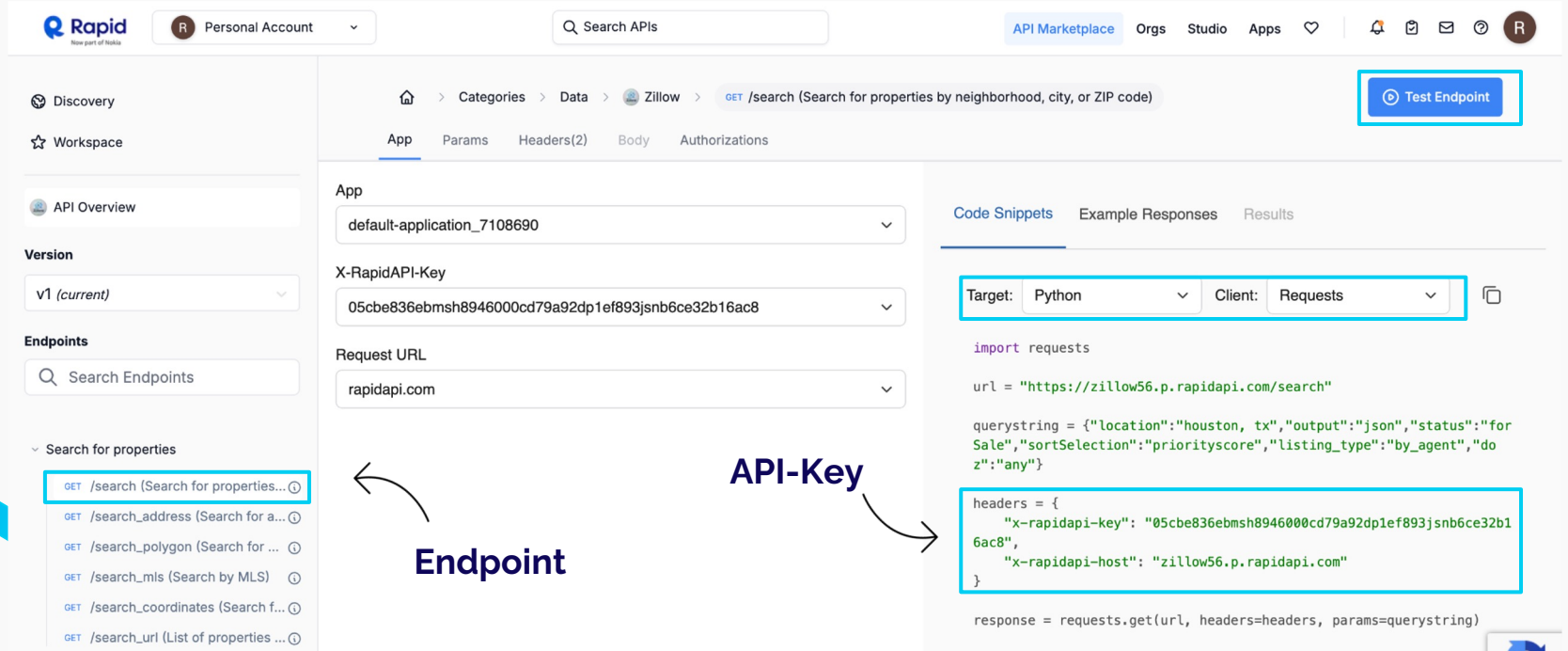
Identify questions about this dataset

Filter data from the API source

Create data visualizations and interpret

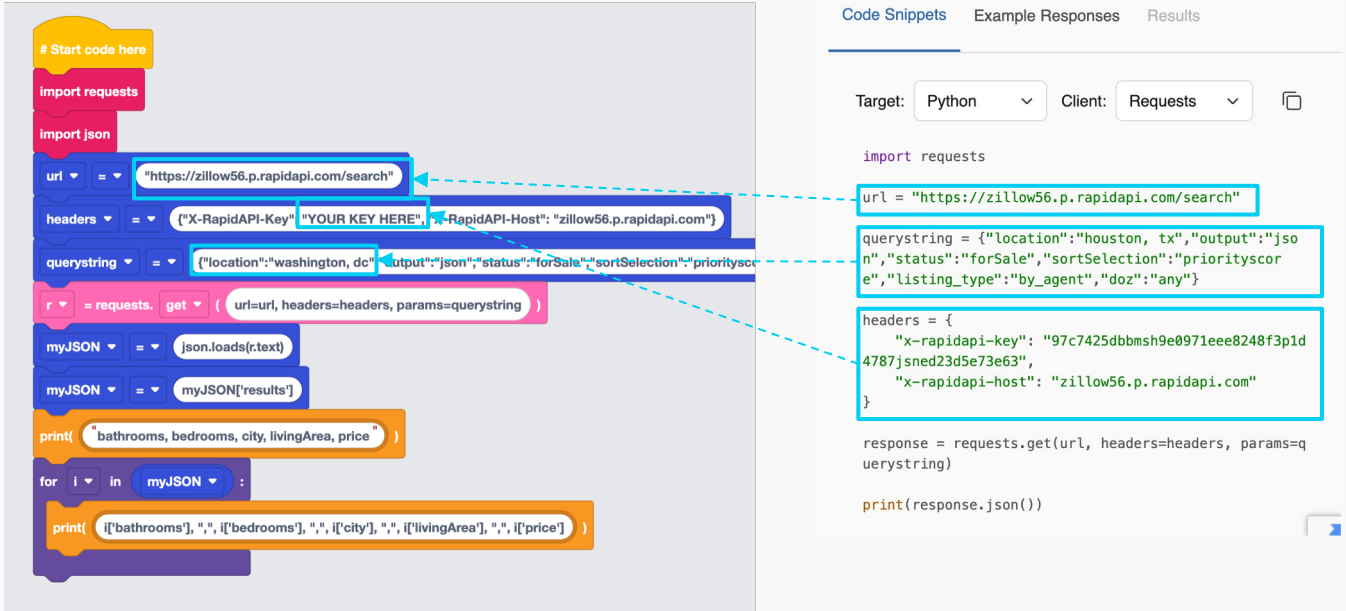
Access the Zillow Dataset

Choose the **"/search properties by neighborhood, city, or ZIP code" endpoint** and **find your API-Key** in the code snippets.



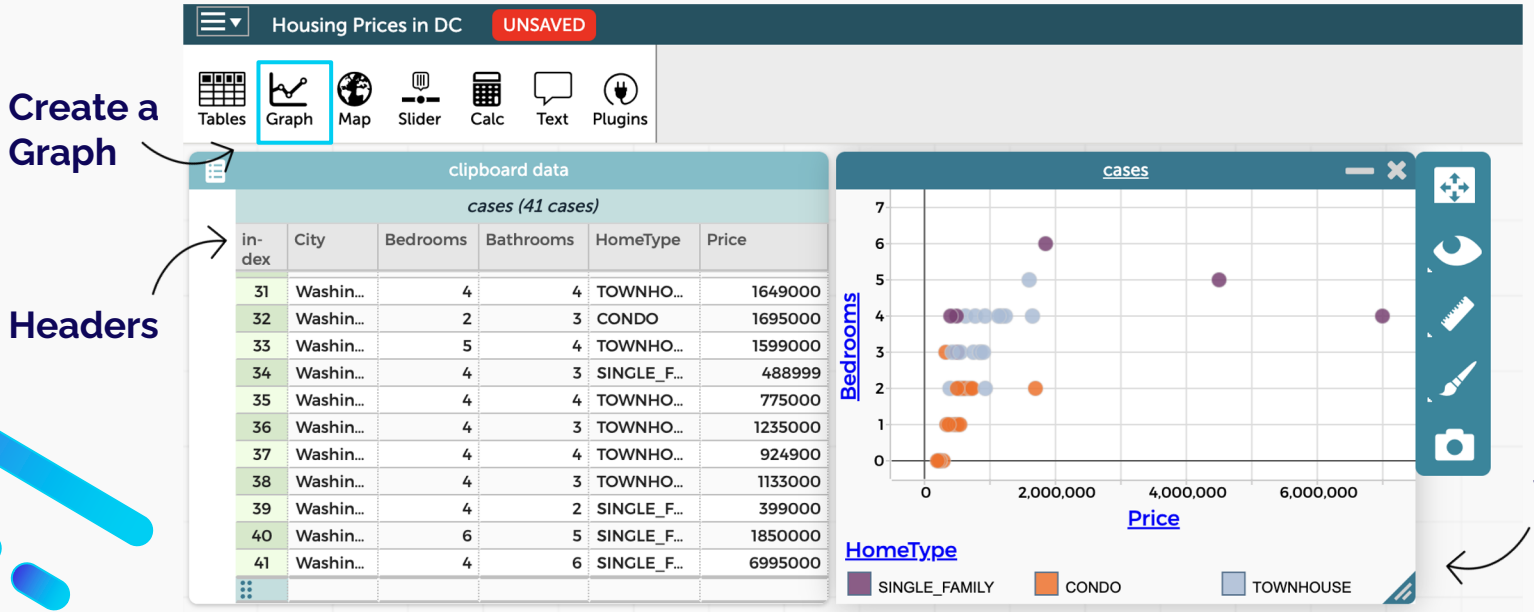
Retrieve & Filter the Data

- Open [this EduBlocks program](#) **(Clone, rename, and save!)**
- **Insert your API-Key** and **change the location**



Create Data Visualization

- **Create a new graph** and **drag** the relevant **headers** to the x and y axis



Final Project: Doing Data Science

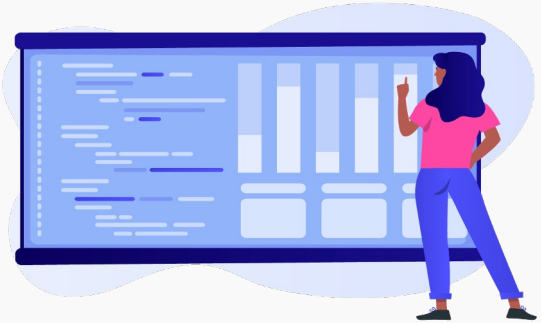
Finding Data
Access and evaluate a relevant dataset

Research Questions
Identify questions about this dataset

Data Refining
Filter, clean and trim the data

Data Visualization
Create data visualizations and interpret

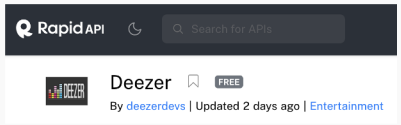
Communicating Results
Share your conclusions and insights



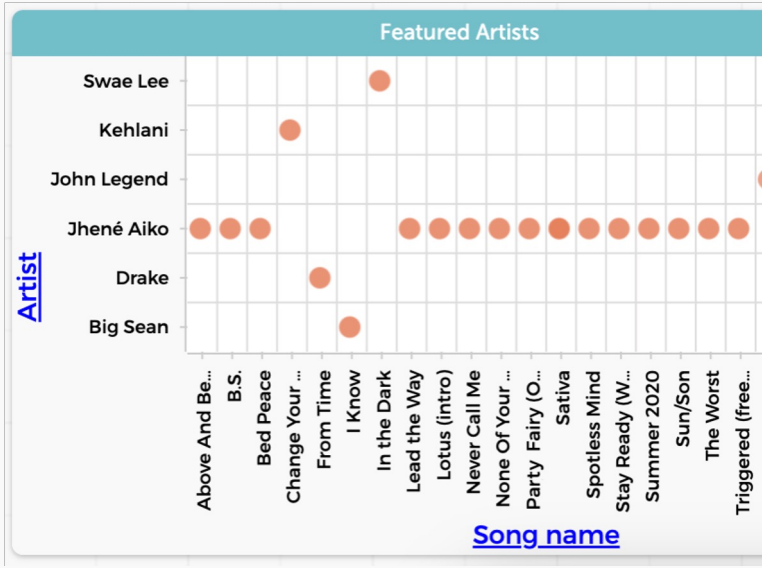
Final Project: Doing Data Science

"Is Jhené Aiko a star?"

DEEZER

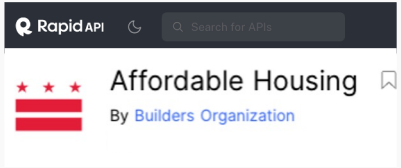


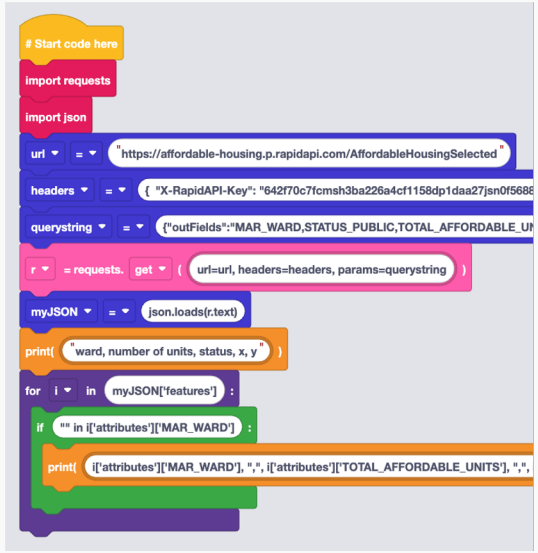

```
# Start code here
import requests
import json
url = "https://deezerdevs-deezer.p.rapidapi.com/search"
headers = {"X-RapidAPI-Key": "e2f29b2b80mshd2c4cf79aa84cacp119f2ejsn80180e7f58
querystring = {"q":"Jhene Aiko"}
r = requests.get(url=url, headers=headers, params=querystring)
myJSON = json.loads(r.text)
myJSON = myJSON['data']
print("Song name, Album, Artist, Duration, Rank")
for i in myJSON:
    print(i["title"], ", ",i["album"]["title"], ",",i["artist"]["name"], ",", i['duration'], ",", i['rank'])
```

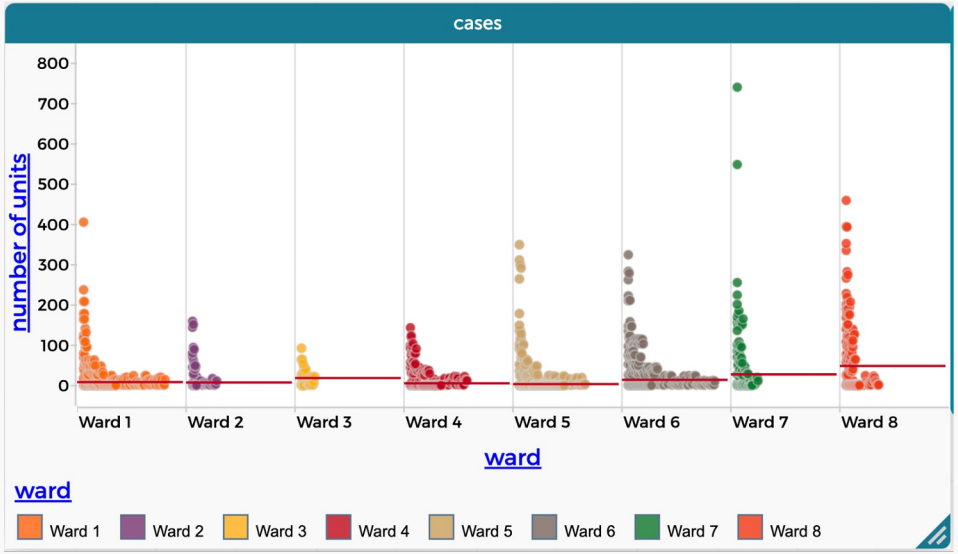


Final Project: Doing Data Science

"Which Ward has the most affordable houses?"







API Can Code at a Glance

Supportive Teacher Resources
Lesson plans, Ready-to-use slides, Videos, Discussion prompts, Assessment rubric

Student Learning Materials
Example programs, Exit tickets, FAQs, Glossary

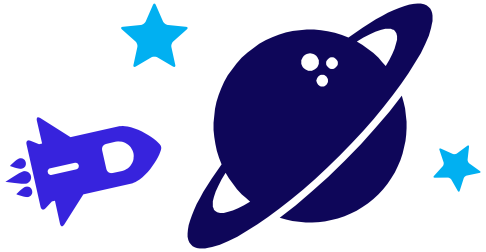
Scaffolded Activities
Guided coding activities following the Use > Modify > Create structure

Meaningful Datasets
Datasets and APIs informed by student interests

Available online

****Free****

apicancode.umd.edu/curriculum.html



Emerging Findings from a Classroom Implementation

Class Implementation Details

- 1 Teacher, 2 Classes
- Public charter high school
- Grades 9-12
- Tier 1 school
- PMF Score: 75.7% (⬇)
- At-risk percentage: 59%

Gender	**N**
Female	18
Male	16
Non-binary	1
Race	
Black or African American	24
Native American	1
Hispanic	4
Latin	2
Prefer not to specify	4

Published

Student Approach to Questioning

RQ: How do students' data science questions evolve in type, scope, and complexity over the course of the final project?

A total of **213 student-generated questions** were collected from **4 student artifacts** over the course of the final project:

Questions were categorized by:

Type (6) Scope (2) Complexity (2)

Student Approach to Questioning

Published

Shifting from broad single variables descriptive inquiries to more complex, evaluative, and exploratory ones:

"What is the average age of players in the NBA?"

↓

"Is there a correlation between a team's payroll and their number of wins in a season?"

Zhou, X., Israel-Fishelson, R., Weintrop, D., Moon, P. F., & Xin, Y. (2025). **Investigating the Evolution of Interest-Driven Data Science Questions Posed by High-School Students**. *Data Science Education K-12: Research to Practice Conference*, San Antonio, TX.

Under Review

Authenticity in Data Science

RQ: How do students perceive the different forms of authenticity integrated into a DS curriculum?

The paper examines how the APICC curriculum aligns with the four dimensions of authenticity proposed by Shaffer and Resnick (1999):

- **Personal Authenticity**
- **Real-world Authenticity**
- **Disciplinary Authenticity**
- **Authentic Assessment**

Two-phase Coding Analysis:

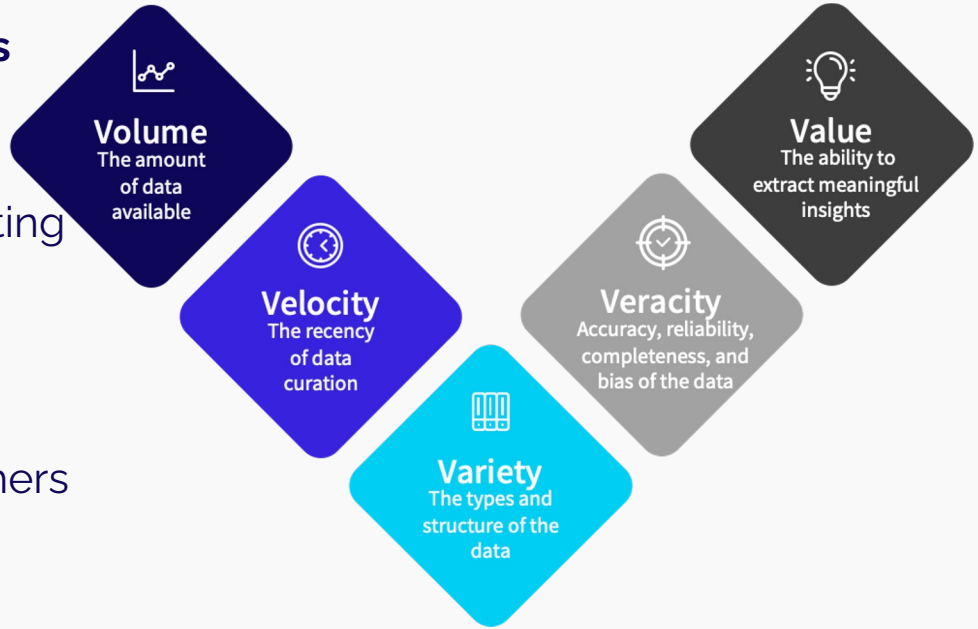
1. **Relevancy Coding:**
   350 interview segments
2. **Dimension coding**

Under Review

5Vs for K-12: Data Evaluation Framework

RQ: How can a structured data evaluation framework support K-12 students in evaluating real-world datasets?

The 5Vs for K-12 framework r**eframes the original 5Vs dimensions**, **Volume, Velocity, Variety, Veracity,** and **Value**, for the purpose of evaluating data in K-12 educational settings.



The framework includes:

- Pedagogical guidance for teachers
- Guiding questions for students

Thank You

Questions?

http://apicancode.umd.edu

Scan me

API CAN CODE

UNIVERSITY OF MARYLAND COLLEGE OF EDUCATION

This work was made possible through generous support from the National Science Foundation (Award # 2141655).