



Curriculum Overview

March 2024



COLLEGE OF
EDUCATION



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

API Can Code

About the Project

API Can Code (APICC) is a research project aimed at [situating data science in the lived experiences of today's students](#). The project explores ways to introduce high-school students to the computational foundations of data science by having them explore meaningful and authentic data that align with their interests. In particular, the learning experiences will center on application programming interfaces (APIs) for popular platforms and services, allowing students to better understand the digital worlds surrounding them.

APICC is part of [a research-practice partnership between the University of Maryland and Washington Leadership Academy](#). As part of this partnership, the APICC team will work with students and teachers to [design an interest- and data-driven data science curriculum and iteratively refine and study the curriculum in classrooms](#) over three years.

This document will provide an overview of the curricular units to be developed and tested in the 2023-2024 school year.

For additional information related to API Can Code, visit: apicancode.umd.edu

APICC Curriculum

Overview

The APICC team will work collaboratively with teachers and students from Washington Leadership Academy in the coming year to test and refine the curricular materials. The curriculum will include three units that will introduce students to the computational foundations of data science. Below is a description of these units.

Unit 1 - Data in Learners' Lives

The "Data in Learners' Lives" unit aims to empower students with a comprehensive understanding of data, its significance, and its impact on various aspects of their lives. The unit will cover the following goals and learning activities:

- **Introduction to Data:** Students will gain a fundamental understanding of what data is, including its definition, types (qualitative and quantitative), and significance in today's information-driven world. Students will explore and discuss how much data is being created every minute and which data they consume and create.
- **Data Collection and its Purpose:** Students will explore the diverse entities that collect data, such as governments, corporates, educational/research institutions, and non-profit organizations. They will investigate the motivations for collecting data, including its applications in decision-making, research, and problem-solving. Additionally, students will be introduced to the Data-Information-Knowledge-Wisdom (DIKW) model and learn about the transformation of data into valuable insights and inform decisions.
- **Sources of Data:** Students will differentiate between primary and secondary data sources, understanding how information is gathered firsthand or from existing datasets and how these distinctions impact data reliability and accuracy.
- **Evaluating Datasets:** Students will be introduced to the 5V's model of evaluating a dataset's quality and will be equipped with critical thinking skills to assess datasets, ensuring they can discern biased, incomplete, or misleading information. They will be taught how to verify the data's reliability and trustworthiness.

-
- **Data Impact and Equity:** Students will explore how data can influence social equity and inclusivity. By analyzing real-world situations, students will understand how data can perpetuate biases or contribute to creating a more equitable and just society when used responsibly.

Through this curricular unit, students will become informed and responsible consumers of data, enabling them to make more informed decisions and fostering a deeper awareness of data's role in shaping the world around them. Moreover, they will be encouraged to consider the ethical implications of data use, promoting a sense of responsibility as future citizens and potential data collectors themselves.

Unit 2 - Computational Foundations of Data Science

The "Computational Foundations of Data Science" unit will introduce students to the field of data science and help students gain essential computational skills required for it. Using a block-based platform, they will learn how to programmatically retrieve and manipulate data from diverse Application Programming Interfaces (APIs). This unit will cover the following learning goals and activities:

- **What is Data Science:** Students will discuss the different fields that contribute to data science and what kind of work the field entails.
- **Accessing Data:** Students will learn to use APIs to access and save data from government databases and private companies for analysis.
- **Data Processing:** Students will learn to clean and preprocess data using programming, specifically in Python, to prepare a dataset for analysis.
- **Forms of Data:** Students will learn to recognize and use data in different forms, including JSON and CSV formats.

Through this curricular unit, students will acquire the skills of working with data and manipulating it to meet their analysis requirements. They will also recognize the value of programming techniques in automating efficient workflows for this process.

Unit 3 - Data Science Practices

The "Data Science Practices" unit will focus on the practical aspects of data science, where learners will analyze and visualize data to extract meaningful insights. They will be introduced to various data analysis techniques and statistical methods.

- **Data Visualization with CODAP:** Students will learn to use the CODAP platform to communicate findings effectively to a chosen audience.
- **Data Analysis:** Students will learn to use informal and formal techniques to describe patterns and findings in a dataset.
- **Graphs and Figures:** Students will learn to create, customize, and interpret a variety of appropriate summary plots for different kinds of data, including histograms, bar charts, and scatter plots, as well as map-based and time-series visualizations.
- **Statistical Testing:** Students will learn to create and use appropriate statistical models, including linear models and confidence intervals.

The unit will emphasize the importance of storytelling with data and conveying insights clearly and compellingly. In addition to the skills of creating data visualizations, students will be asked to think critically about what kinds of data visualizations will best serve the needs of the questions they seek to answer and the stories they seek to tell.

Final Projects

Students complete a final project on a topic of their choice, drawing data from an API source, processing it with EduBlocks, visualizing the data with CODAP, and presenting their findings to the class. The final project is intended to include all three phases covered in the curriculum applied to a subject of interest for the students.

Overall, this data science curriculum is designed to provide learners with a well-rounded understanding of data science, starting from its foundational concepts and ethics, advancing to essential computational skills, and culminating in applying data analysis and visualization techniques to draw meaningful conclusions and effectively communicate their insights. By the end of the curriculum, learners will be equipped with the knowledge and practical skills necessary to tackle real-world data science challenges.