**AP Computer Science I**

Performance Task

# **Create — Data Lab**

## **Overview**

Students are more motivated if they feel connected to the material they are studying.

One way to make computer science content relevant, meaningful, relatable, and exciting

is to incorporate real-world data sets into hands-on programming assignments.

Data are values. Values by themselves are not knowledge, knowledge is what you gain

from studying data. This lab is designed to encourage students to derive knowledge

about a topic of interest to them by examining a set of freely available data from the

Web. In order to gain knowledge from data in this lab, students will pose a question,

search for a data set on that topic, and then process that data to determine the answer

to their question.

Being able to access and process data is an important part of many industries. Whether

it is determining the best allocation of resources or deciding what new product to

create, analyzing data is a necessary component of those conversations. Luckily there

are tools available to help read and process data, and this lab will introduce one such

tool to students and ask them to utilize existing real-world data to answer a question

of their own interest.

## **Assessment**

You will be provided with 12 hours of class time to complete and submit the following:

* A video of your program running
* Written responses about your program and development process
* Program Code

Your teacher will share submission guidelines that include suggestions for creating video and PDF files.

## **General Requirements**

You are required to:

* Iteratively design, implement, and test your program.
* Independently create at least one significant part of your program.
* Create a video that displays the running of your program and demonstrates its functionality.
* Write responses to questions about your program.
* Include your entire program code.

## **Program Requirements**

Your program must demonstrate:

* Students are able to create a new class and attempt to create a constructor with the same name as the class, but incorrectly include a return type for the constructor
* Students are able to write program code to declare instance variables and write headers for constructors with or without parameters, but fail to assign value to the instance variables in the constructor. Often they unknowingly create local variables in the constructor with the same name as the instance variables instead.
* Students are able to use variables that store a single value, but can’t understand how related values can be stored and accessed together using a single variable name.
* Students are able to represent related data using an array or ArrayList, but can’t write program code that requires them to manipulate the elements while traversing the array or ArrayList.
* Students are able to write program code using some of the data available, but students often are unaware of what variables (instance variables/ parameters) they have available to solve a problem and the associated data that can be accessed by calling the methods of reference data.
* Students are able to traverse and manipulate elements in a single String, 1D array, 2D array or ArrayList object, but can’t write program code to perform parallel/simultaneous traversals of two or more structures.

## **Submission Requirements**

### 1. **Video**

Submit one video in .mp4, .wmv, .avi, or .mov format that demonstrates the running of at least one significant feature of your program. Your video must not exceed 1 minute in length and must not exceed 30MB in size.

### 2**. Written Responses**

Submit one PDF document in which you respond directly to each prompt. Clearly label your responses 2a – 2e in order. Your response to all prompts combined must not exceed 750 words, exclusive of the Program Code.

## **Program Purpose and Development**

1. Provide a written response or audio narration in your video that:

Identifies the programming language.

* Identifies the purpose of your program.
* Explains what the video illustrates.

(Approximately 150 words)

1. Data science is a field that utilizes programming to bring together data engineering, mathematics, and statistics. How does this lab, and computer science in general, relate to data science? What are some other uses of data science that have impacted your everyday life? What are some specific challenges that you imagine data scientists encounter?

(Approximately 200 words)

1. Data structures such as arrays and ArrayLists are used to organize data and simplify data processing. Discuss one data structure that you used in the completion of this lab. How did it help in the processing of data?

(Approximately 200 words)

1. Would you have been able to complete this lab without the use of a data structure? Why or why not?

(Approximately 200 words)

1. Capture and paste your entire program code into the PDF.

* Include comments or citations for program code that has been written by someone else.

## Tasks

### **Activity 1 - Explore**

**Description**

Although it is not assessed, students should spend time discussing the ethical

implications of computing and this lab provides an appropriate scenario to have this discussion. The ease with which data can be stored, cataloged, and searched in an ever more-connected society is an important point for students to discuss.

Spend a few minutes discussing privacy policies. Regardless of whether a service is offered for free or at a cost, the companies that students interact with on a regular basis collect data about them. Privacy policies are commonplace everywhere from social media to your local doctor’s office. By agreeing to use a service, and often just by visiting a website, you are agreeing to the privacy policy. If students have not already been provided with the handout for this activity, they should be given it now.

Have students list sites that they typically visit and then compile a class list. In groups, have students find the privacy policy for a specific site. Ask them to identify two or three pieces of information that are collected. This information could include information about the device used to access the site, such as type or IP address, as well as information about specific content viewed.

* Data, Data, Data everywhere

Time To Complete: 1 Hours

### **Activity 2 - Research**

**Description**

In this lesson, students will explore and learn the anatomy of classes. They will take a deeper dive into what the access specifier does, and how it can be used within programs to make data public or private. Students will learn about encapsulation and the responsibility programmers have to choose whether data should be accessible, modifiable, both or neither. Students will learn about and create arrays. The use of array objects allows multiple related items to be represented using a single variable. Students will learn about and use ArrayLists. ArrayLists are similar to arrays, except that they are a mutable list of object references. ArrayLists provide a convenient way to create adjustable arrays.

* Unit 5 - Writing Classes
* Unit 6 - Arrays

Time To Complete: 14 Hours

### **Activity 3 - Ideate**

**Description**

Display the given table found at the beginning of the student handout for the class, or have students review it on their own. The table is also provided in the file Cereal.csv.

If you choose to display the file, limit the visibility to the first 21 lines to match the image provided to students. Discuss the following questions:

* What do you think the table is describing?
* What do the letters ‘C’ and ‘H’ represent in the Type column?

In groups or partners, have students answer a specific question about the cereals on the list, for example: Which is the cereal with the most sugar on the list? Which cereal is the highest in fiber on the list?

Highlight the fact that this table includes only the first 21 rows of a much larger table. While it was relatively easy to answer the questions while only looking through 21 rows of data, briefly discuss the time it would take to determine the answers given 100 rows, then 1,000 rows, and 10,000 rows.

* Unit 7
* Designing Classes

Time To Complete: 8 Hours

### **Activity 4 - Evaluate**

**Description**

Display the given table found at the beginning of the student handout for the class, or have students review it on their own. The table is also provided in the file Cereal.csv.

If you choose to display the file, limit the visibility to the first 21 lines to match the image provided to students. Discuss the following questions:

* What do you think the table is describing?
* What do the letters ‘C’ and ‘H’ represent in the Type column?

In groups or partners, have students evaluate the questions about the cereals on the list.

Time To Complete: 1 Hours

### **Activity 5 - Construct a Prototype**

**Description**

Before beginning this activity, make sure that students have imported the appropriate libraries to their IDE. Information on how to complete this process can be found at the beginning of this lab in the “Setup” section and at the following link: http://berry-cs.github.io/sinbad/install-java.

This activity will utilize the tutorials for the library that can be found at https://github.com/berry-cs/sinbad/blob/master/tutorials/java/. Specifically, students will complete the welcome01.md, welcome02-obj.md and welcome03-objs.md tutorials.

It is recommended to give students the Welcome01.java file to ensure that the library is set up properly and it will compile, although the first tutorial could be given as a homework assignment before beginning this activity. If you plan on having students work through the tutorial, you are encouraged to work through it yourself ahead of time. The tutorial also contains possible extensions, which could be assigned to the entire class or individual students if desired.

While using the library, you or your students might notice a pop-up box asking for permission to collect data on the usage of the Sinbad library. Whether you allow data to be collected or not, this provides an opportunity to discuss the ethics of data collection as a class, including potential benefits of such data collection.

* Putting it all together

Time To Complete: 3-5 Hours

### **Activity 6 - Improve the Design**

**Description**

Students should seek feedback to improve their program designs and retest their solutions with appropriate test cases.

As a class, spend a few minutes reviewing the requirements of the open-ended activity.

The goal of the activity is to allow students to demonstrate their knowledge of class

design and data structures in a way that is interesting and engaging to them. While

it is possible to create additional constraints or requirements, it is best to provide

students with as much freedom as possible. The use of any existing classes is intentionally not included in the list of requirements.

* Open Ended Activity

Time To Complete: 3-5 Hours

### **Activity 7 - Share Solutions**

**Description**

Students share their solutions and how they differentiate from the original design.

Time To Complete: 1-3 Hours

### **Activity 8 - Reflect**

**Description**

Students answer the reflections questions and submit all evidence.

Time To Complete: 1 Hour