

## **MODULE #1: Introduction to Big Data**

### **Notes & Considerations:**

- Feel free to adapt the timeline and materials to your individual course and student needs.
  - You may edit all of the instruction sheets, activity sheets and project rubrics for your individual course and student needs.
  - Ideas for extensions are listed under applicable activities.
  - Email Lee Cristofano [cristofano.lee@bphawks.org](mailto:cristofano.lee@bphawks.org) or Emily Smoller [smoller.emily@bphawks.org](mailto:smoller.emily@bphawks.org) with any questions.
- 

### **Need to Know Information for Teachers (in order to complete the module):**

- Careers in Big Data, Data Science and Data Analytics
  - Colleges that offer degrees and certificates in Big Data, Data Science and Data Analytics
  - Ethical considerations when collecting, storing and utilizing data
- 

**Module Objective:** To develop a deeper understanding and curiosity about the field of big data.

### **Student Objectives:**

- To identify career opportunities in big data, data analytics and data science.
  - To identify local universities, and other universities of interest, that offer big data, data analytics and data science certificate, Bachelors, Masters and PhD Programs.
  - To identify and analyze how big data applies to our everyday life.
  - To understand ethical issues with data collection and data usage.
- 

### **PA Standards:**

#### **PA Standard Area: 3A.DA Computer Science - Data & Analysis**

- 3A.DA.12 - Evaluate the tradeoffs in how data elements are organized and where data is stored.

#### **PA Standard Area: 3.4 Technology & Engineering Education**

- 3.4.10.B1 - Compare and contrast how the use of technology involves weighing the trade-offs between the positive and negative effects.
- 3.4.10.B4 - Recognize that technological development has been evolutionary, the result of a series of refinements to a basic invention.
- 3.4.12.B1 - Analyze ethical, social, economic, and cultural considerations as related to the development, selection, and use of technologies.
- 3.4.12.B2 - Illustrate how, with the aid of technology, various aspects of the environment can be monitored to provide information for decision making.

#### **PA Standard Area: 13.1 Career Awareness & Preparation**

- 13.1.A - Relate careers to individual interests, abilities, and aptitudes.
- 13.1.B - Analyze career options based on personal interests, abilities, aptitudes, achievements and goals.
- 13.1.C - Analyze how the changing roles of individuals in the workplace relate to new opportunities within career choices.

- 
- 13.1.F - Analyze the relationship between career choices and career preparation opportunities.

**PA Standard Area: 15.4 Computer & Information Technologies**

- 15.4.12.B - Evaluate the impact of social, legal, ethical, and safe behaviors on digital citizenship.

**PA Standard Area: 15.8 Management**

- 15.8.12.C - Critique major social and ethical decisions made by business and organizations throughout history.

**PA Standard Area: English Language Arts**

- CC.1.2.11-12.B - Cite strong and thorough textual evidence to support analysis of what the text says explicitly as well as inferences and conclusions based on and related to an author's implicit and explicit assumptions and beliefs.
- CC.1.2.11-12.G - Integrate and evaluate multiple sources of information presented in different media or formats (e.g. visually, quantitatively) as well as in words in order to address a question or solve a problem.
- CC.1.2.11-12.J - Acquire and use accurately general academic and domain-specific words and phrases, sufficient for reading, writing, speaking, and listening at the college and career readiness level; demonstrate independence in gathering vocabulary knowledge when considering a word or phrase important to comprehension or expression.
- CC.1.2.11-12.L - Read and comprehend literary nonfiction and informational text on grade level, reading independently and proficiently.
- CC.1.5.11-12.A - Initiate and participate effectively in a range of collaborative discussions on grades level topics, texts, and issues, building on others' ideas and expressing their own clearly and persuasively.

**Materials:**

- Hard copies of the Course Curriculum (if a whole course is being taught)
- Course Syllabus (if you wish to create)
- [Pittsburgh DataWorks](#) Website
- Video from the DataJam Mentors [Introducing the Pittsburgh DataJam](#)
- [Big Data Research & Student Info Sheet](#)
- Sticky notes

**Prerequisites (for teachers):**

- Preview Module #1 materials
- Conduct some of your own research in the field

**Activity 1: Course Introduction**

- Review curriculum outline and timeline
- Preview the Pittsburgh Dataworks DataJam Program together as a class:
  - Website: [Pittsburgh DataWorks](#)
  - Video from the DataJam Mentors [Introducing the Pittsburgh DataJam](#)
- Have students work on the following activity individually - [Big Data Research & Student Info](#)

- I typically work through this sheet each year for the most updated information

**Extensions:**

- Pittsburgh Dataworks website: Create a webquest sheet to guide students through the site
- Big Data Research & Student Info Sheet: Add higher level thinking questions

**Activity 2: What is Big Data?**

- **Video Resources** to aid in your discussion. Feel free to use all, or a few.
  - [Pitt Mentor presentation](#) or invite Pitt Mentors to join a class session ○ [World Economic Forum - 2 Minutes: What is Big Data](#)
  - [Big Data in 5 Minutes: the 5 V's of Big Data](#)
  - [Making data mean more through storytelling](#): Ben Wellington - 15 minutes
  - [How Cops are Using Algorithms to Predict Crimes \(Los Angeles\)](#) - 13 minutes Good place to discuss bias in algorithms.
- **Extensions:**
  - Videos: Create a guided note sheet or an Edpuzzle

**Activity 3: Careers in Big Data, Data Analytics & Data Science Programs**

- Discuss responses to [Big Data Research & Student Info](#) throughout the conversation ● Discuss the following topics as a class:
  - Types of careers
  - Salary and demand for these careers
  - Local employers
  - National employers
- **Video:** [Walmart Global Customer Insights and Analytics Team](#) - 4 minutes ● **Extensions:**
  - Videos: Create a guided note sheet or an Edpuzzle

**Activity 4: Colleges offering Big Data, Data Analytics & Data Science Programs ●**

Discuss responses to [Big Data Research & Student Info](#) throughout the conversation

- Discuss the following as a class:
  - Local universities
  - All universities

**Activity 5: Big Data in Our Everyday Life**

- Have students go through a day in their life
  - On one color of sticky notes, write how data is collected
  - On another color of sticky notes, write how big data makes their lives easier
- Have students put their sticky notes on the board

- 
- Do some Affinity Clustering with the students' responses
- First, discuss how data is collected in their everyday lives  
Second, discuss how big data collection and usage makes their lives easier
- Lastly, are there any times they believe data collection infringes on their privacy (leads into Activity #6)

### **Activity 6: Ethical Issues in Big Data**

- Produce a list of data that is collected on a daily basis (look to Activity #5)
  - How is/can this data be used to improve the quality of your life?
  - How is/can this data be used unethically?
- Other examples of how:
  - Data collection and usage benefits society
  - Data collection and usage infringes on our privacy
- **Article:** [One Nation: Tracked](#) - NY Times article about cell phone datasets that allow tracking
- **Video:** [How Data Brokers Sold My Identity](#) - Tedx Talk - 16 minutes
- **Extensions:**
  - Article & Videos: Create a guided note sheet or an Edpuzzle

### **Module Student Evaluation:**

- Completion of [Big Data Research & Student Info](#)
- Participation in activities and discussions

### **Possible Timeline (open to teacher discretion):**

Day 1: Activity 1  
 Day 2: Activity 2  
 Day 3: Activity 3  
 Day 4: Activity 4  
 Day 5: Activity 5 & 6