Unit 4 Lesson 1

What is Big Data?

Resources

Activity Guide - Big Data Sleuth Card

Directions:

- With a partner, select one of the tools in the list to the right.
- Determine what the tool is showing.
- Find the source of the data it allows you to explore.
- Complete the table below.

Web Sites:

1. Web archive http://www.archive.org
2. Measure of America http://www.measureofamerica.org/maps/
3. Wind Sensor network http://earth.nullschool.net/
4. Twitter sentiment https://www.csc.ncsu.edu/faculty/healey/tweet_viz/tweet_app/
5. Alternative Fuel Locator http://www.afdc.energy.gov/locator/stations/

Website Name	
What is this website potentially useful for? What kinds of problems could the provided information be used to solve?	
Is the provided visualization useful? Does it provide insight into the data? How does it help you look at a lot of information at once? How could it improve?	
Where is the data coming from? Check for “About”, “Download”, or “API”. You may also need to do a web search. ● Is the data from one source or many? ● Is it static or live? ● Is the source reputable? Why or why not? ● Add a link to the raw data if you can find one.	
Do you consider this “big” data? Explain your reasoning.	

Unit 4 Lesson 2

Finding Trends with Visualizations

Resources

Name(s)_____ Period _____ Date _____

Activity Guide - Exploring Trends

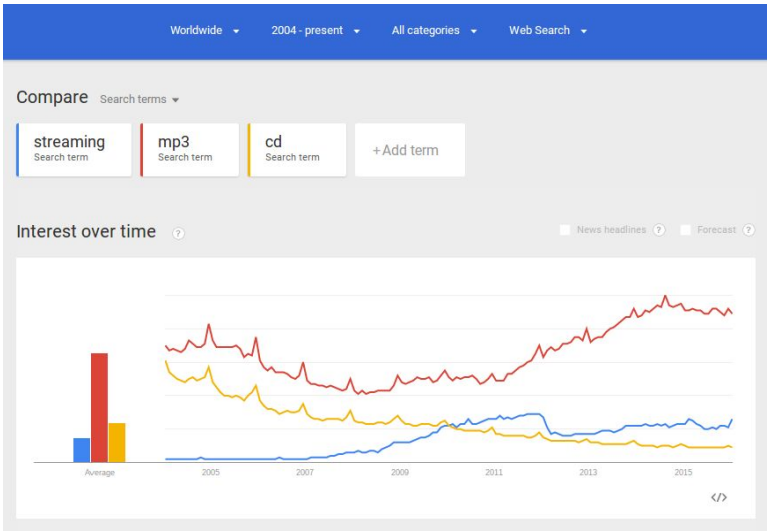
What's a Trend?

When you post information to a social network, watch a video online, or simply search for information on a search engine, some of that data is collected, and you reveal what topics are currently on your mind. When a topic is quickly growing in popularity, it is often said to be **trending**, but there are many different trends or patterns we might find in this data, including historical trends. These patterns might help us to identify, understand, and predict how our world is changing.

Using Google Trends

You will be using Google Trends, which is a tool that allows you to **visualize data** about search history across different times and locations**.** You will be **looking for interesting patterns, trends, or relationships between multiple trends** and **try to tell the story that pattern is showing**.

- Access it here: https://www.google.com/trends/
- To get started, you want to "Explore" a trend of your own.
- Find the Explore text box or go to the Explore section of the site https://www.google.com/trends/explore



Do your first comparison of trends

You can enter any two or three things you like. This example compares **streaming**, **mp3**, and **cd**.

Understand what is being represented: Take a moment to actually understand what the Google Trends data is showing. Respond to the questions below. You can find answers on the Google Trends Help Page.

- Where does Google Trends data come from?

- How is Google Trends data adjusted? What does a value of 100 mean?

Exploring Trends on your own

Continue exploring comparisons of search trends that are interesting to you. You might start by looking up:

- **a current event / social movement / hashtag / meme**
- **your favorite hobby / movie / song / book / celebrity**
- **popular apps / businesses / products / websites**

As you try different terms, here are **some things to look for:**

- **Long-term trends:** Is your topic becoming more popular over time? Less?
- **Short-term trends:** Does your topic suddenly spike or dip in popularity?
- **Patterns:** Does your topic follow some predictable repeated pattern?
- **Relationships:** Does one topic increase or decrease in popularity when another one does?
- **Zoom-in:** You can narrow your search to particular regions, times, and categories

Tell a Story

As you explore, you should eventually **settle on one topic** you find particularly interesting. For your chosen topic, **be prepared to present or share** what you found, responding to the following:

- **Describe** what terms you compared and whether you narrowed your search by using filters.

‘

- **Accurately describe** what the charts or other visualizations are showing.

- Come up with a **possible story or explanation of why** the trend you described might have happened.

Unit 4 Lesson 3

Check Your Assumptions

Resources

Name(s)_____ Period _____ Date _____

Activity Guide - Digital Divide & Checking Assumptions

The "Digital Divide"

Perhaps one of the easiest assumptions to make, when looking at data collected online, is that it actually is a good representation of what the average person is thinking, doing, or cares about. Look through this report from Pew Research, which shows the large differences in access to technology, even in the modern day. Digital Divides 2018 (http://www.pewinternet.org/fact-sheet/internet-broadband/)

1. What is the "digital divide"?

2. What groups are overrepresented or underrepresented online as a result of the digital divide?

3. What was the most surprising piece of information or visualization you found in this report?

Identifying Assumptions in Data Analysis

When you use data to make decisions, you need to be careful to **identify your assumptions** and **reflect on how those assumptions impact your analysis.**

Pick one of the scenarios below. With a partner, respond to the questions you find there about the assumptions made to conduct that analysis.

Scenario	Data	Decision
A city would like to more efficiently locate potholes that need to be filled.	The city builds an app that allows residents to report potholes from their smartphone.	The city will use this app as the primary method of identifying potholes.
A news agency would like to predict the outcome of a coming election.	A social media company (e.g., Twitter) keeps track of how many times each candidate has been mentioned on the platform.	The news agency will use this data to predict the outcome of the election.
A state government is trying to determine which issues are most important for the upcoming year.	The government creates an online survey where citizens can vote for the issue they care most about.	The government will use the results of this survey to help prioritize issues in the coming year.
A chef is deciding where to open a new restaurant in the city.	A restaurant reviews website keeps track of the areas of the city that receive the most restaurant reviews.	The chef will use the data to choose the location for opening a new restaurant.

Respond

Answer each question below for the scenario you chose.

1. What kind of trends or patterns might the decision makers in your scenario be observing in the data when they make their decision? How might they use these patterns to help make their decision?

2. The people making decisions in each instance have not taken into account the “digital divide.” How has the assumption that all people have equal access to technology **possibly biased their analysis?**

3. Are there any other assumptions being made in your scenario? How might they affect the final decision made? What additional data might you want to collect before making the decision for your scenario?

Unit 4 Lesson 4

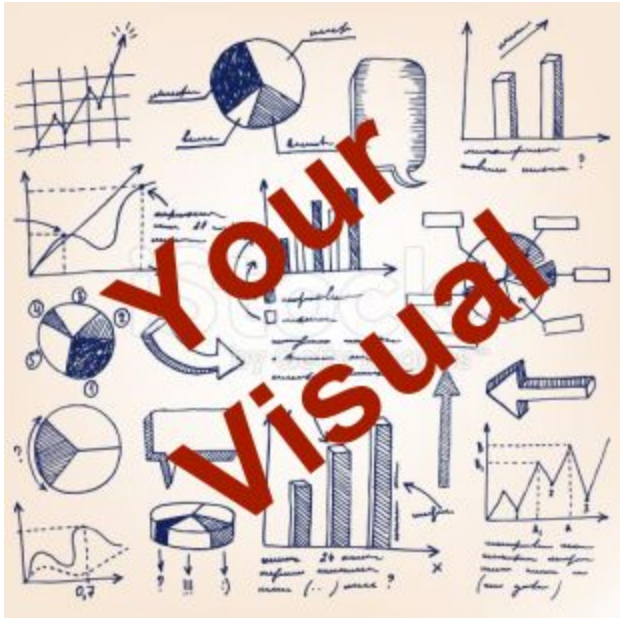
Rapid Research - Data Innovations

Resources

Name(s)_____ Period _____ Date _____

Data Innovation One-Pager Template
<change this to your title>

Note: *All text in Italics, including this text, is intended to be replaced by your responses, and deleted once you've completed your one-pager.*



Innovation Purpose

What is the intended goal or objective of the innovation? Why was it created in the first place? What need or problem led to it being created? Assume that the audience doesn't know much about the topic.

Innovation Function

How does the innovation work? What data does this innovation use, produce, or consume? Where does the data come from? How is it collected? Who is collecting it? If possible, how and where is it stored? Do they provide public access to this data? If so, put a link to it here.

Beneficial Effect on Society, Economy, or Culture

Think big picture. What is a group of people that benefits from this innovation? How specifically do they benefit? State clearly if you think this benefit is on society, economy, or culture.

Sources

List all websites that you used to find any information you wrote here. Include the permanent URL. Identify the author, title, source, the date you retrieved the source, and, if possible, the date the reference was written or posted. You should number your sources, here is a template you can follow:

[1] Author's Last name, First name. "Title of Web Page." Title of Website, Publisher, Date, URL. Date retrieved.
[2] Author's Last name, First name. "Title of Web Page." Title of Website, Publisher, Date, URL. Date retrieved.
[3] ….

Note about the Visual:

The visual should be an image or graphic that you think helps illuminate your topic, the innovation itself, the data it uses, or your specific example. It does not need to be anything you created by hand, but it can be. A good visual would add depth or additional information to your written text. If necessary, one of the sections should explain the purpose of the visual: what it is depicting? Make sure you also cite the source of the image.

Name(s)_____ Period _____ Date _____

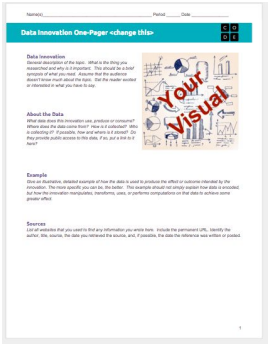
Activity Guide: Rapid Research - Data Innovations

Project Overview: Data Innovations

In this small project, you will quickly research a computing innovation of your choice and present a “one-pager” about it. The computing innovation should be one that produces, uses, consumes or is “driven” by data in some way.

The One-Pager

In the professional world it is common to ask someone to do a bit of online research and then prepare a one-page summary or “one-pager” to show the rest of the team or colleagues about the highlights of what you found. For this project you will prepare a one-pager that explains how some technological innovation uses data.



General Process & Requirements

- Review the One-Pager Template and the **Rubric** below.
- Choose your innovation using the guide below to help.
- Conduct your research by following the **Research Guide** below.
- Complete the one-pager.

Choosing Your Innovation

You should choose an innovation related to big data that you find personally relevant or interesting. For the purposes of this project, you may need to use a broad definition of an “innovation,” and it doesn’t necessarily mean that it has to be a new invention. In particular, we’re looking for an innovation that uses data.

Brainstorming:

Start by simply going to your favorite search engine and entering the name of a thing you are interested in, followed by the word “data” or “big data.” For example: “soccer big data,” “shopping big data,” “music big data.” It might take some poking around, but you will find a few interesting things.

Make sure you can easily find at least one or two reputable sources for your innovation and how it uses data.

Other Ideas:

- Smart grids, smart buildings, smart homes
- The data behind your favorite app
- Online shopping trends/recommendations
- Crowdsourcing
  - Crowd source inventions and funding (Kickstarter, Quirky, GoFundMe)
  - reCAPTCHA
  - GWAPs - Games With A Purpose
- Assistive technologies - aiding human vision, hearing, movement, etc.
- Machine learning
- Sports analytics
- Any kind of science: physics, biology, chemistry, astronomy, etc.

Name(s)_____ Period _____ Date _____

Research Guide - Data Innovation

Conducting Your Research

You already have some practice finding good resources online. You'll want to find **recently published documents** from **authoritative sources.** There is no need to use overly technical documents, but keep an eye out for familiar terminology and topics.

Key Information to Find

- **Purpose:** the need, goal, or problem that led to the creation of the innovation. This is "what" the innovation was designed to do.
- **Function:** how the innovation actually accomplishes its purpose. This typically means describing the way it consumes, produces, or transforms data. This is "how" the innovation actually works.
- **Beneficial Effect:** try to find a group that has benefitted from the innovation. Describe the specific impact of the innovation on them in terms of society, economy, or culture.

Use the tables below to keep track of your information; you can also add more if you like. **You'll need to include at least 3 sources of information** but you can use more.

My Innovation: _____

Reference Name: **Year Published:**	**URL:**
Key Information	

Reference Name: **Year Published:**	**URL:**
Key Information	

Reference Name: **Year Published:**	**URL:**
Key Information	

Name(s)_____ Period _____ Date _____

Rubric - Data Innovation One-Pager

One-Pager Rubric

Component	1	2	3	Score
Sources & Visual				
Sources	Research Guide includes references to fewer than three sources and the sources listed are not recent and authoritative.	Research Guide includes references to fewer than three sources or the sources are not recent and authoritative.	Research Guide includes references to at least three recent, authoritative sources.	
Visual	Visual does not substantially contribute to expanding the understanding of the innovation provided in the written responses.	Visual may be appropriate but largely repeats information in the written responses, or is only loosely related to the innovation.	Visual is well chosen and enhances or augments the information in the written responses.	
Written Responses				
Purpose	The description does not provide a clear explanation of the purpose. It's possible that this response clearly explains the "function" of the innovation but not the goal or objective that led to its creation.	The description explains what the innovation does, but not the purpose - it's missing the "why", or doesn't directly state why the innovation was created.	The purpose description includes rich details. The specific need or goal that led to the creation of the innovation is noted.	
Function	A brief description of the data used by this innovation is given but only includes superficial detail or generic descriptions.	A description of the data used by this innovation is given but lacks sufficient detail to give a clear picture of exactly how the innovation uses data or what is innovative about it.	A clear description is given of what kind of data the innovation uses, how it's collected, and by whom. It also explains how the innovation uses, produces, or consumes this data.	
Beneficial Effect	Response does not identify a group affected nor an economic or cultural impact. Or the response is overly vague like "it affects society because…"	Describes the group affected OR the economic or cultural impact but not both. Or the response is vague about the group, or impact.	Response clearly identifies BOTH a group affected by the innovation AND an economic or cultural impact.	

Unit 4 Lesson 5

Identifying People With Data

Resources

Name(s)_____ Period _____ Date _____

Activity Guide - Research Yourself

Your Digital Self

You may already be aware of information about you that is freely available online, but you probably haven't thought about it from the standpoint of research. Suppose someone were to research you online. What would they be able to find? What connections could they make from the existing data out there to learn even more about you?

Conducting Your Research

You should look through any publicly available pieces of information online. Start by simply looking up your name in a search engine but then refine your results by adding more specific information, like the place you live. Don't forget social networks, your school website, or any other websites you frequently use.

Record Your Findings

In the space below record the information you find about yourself. If you know something is available online but can't get to it now, record it anyway. If you need more space, you can record your findings on the back of this sheet as well.

Information	Where you found it

Now connect the dots. If someone *really* wanted to find out about you online, given the information above, what would they know about you?

Of the pieces of information you found above, which do you think poses the biggest threat to your security or privacy? Why do you think so?

Unit 4 Lesson 6

The Cost of Free

Resources

Name(s)_____ Period _____ Date _____

Activity Guide - Privacy Policies

Choose a Website and Find the Data Privacy Policy

Choose an app, website, or other online service you are familiar with to research their privacy policy. The easiest way to find a data policy, if it exists, is to search for the company name followed by the terms “data policy” or “privacy policy.”

Your website: _____

What Is Their Data Policy?

Respond to the questions below. Even if you can’t find information, you should record where you looked and the fact that you can’t find it. If there isn’t a policy or it’s hard to find, that can be just as interesting as seeing the policy itself.

What kinds of data are being collected? How many different kinds of data?

What service or feature is enabled by the data they are collecting? Why are they collecting it in the first place?

Who else is given access to that data? How are they using it?

Can you get access to your own data? Can you modify what is collected or used, or delete your data if you wish?

Bottom Line: on a scale of 1-4, rate how comfortable you are with this company’s data policy? Give your rating (no going halfway - no 2.5 or 1.5 - make a choice!) and then justify your choice.

1	2	3	4
very uncomfortable	uncomfortable (but maybe fixable)	comfortable (maybe minor concerns)	very comfortable

THE WALL STREET JOURNAL.

It's Modern Trade: Web Users Get as Much as They Give

Link to Original Article article from WSJ.com

Intro and Background

If you surf the web, congratulations! You are part of the information economy. Data gleaned from your communications and transactions grease the gears of modern commerce. Not everyone is celebrating, of course. Many people are concerned and dismayed—even shocked—when they learn that "their" data are fuel for the World Wide Web.

Who is gathering the information? What are they doing with it? How might this harm me? How do I stop it?

These are all good questions. But rather than indulging the natural reaction to say "stop," people should get smart and learn how to control personal information. There are plenty of options and tools people can use to protect privacy—and a certain obligation to use them. Data about you are not "yours" if you don't do anything to control them. Meanwhile, learning about the information economy can make clear its many benefits.

It's natural to be concerned about online privacy. The Internet is an interactive medium, not a static one like television. Every visit to a website sends information out before it pulls information in. And the information Web surfers send out can be revealing.

Most websites track users, particularly through the use of cookies, little text files placed on Web surfers' computers. Sites use cookies to customize a visitor's experience. And advertising networks use cookies to gather information about users.

A network that has ads on a lot of sites will recognize a browser (and by inference the person using it) when it goes to different websites, enabling the ad network to get a sense of that person's interests. Been on a site dealing with SUVs? You just might see an SUV ad as you continue to surf.

Most websites and ad networks do not "sell" information about their users. In targeted online advertising, the business model is to sell space to advertisers—giving them access to people ("eyeballs") based on their demographics and interests. If an ad network sold personal and contact info, it would undercut its advertising business and its own profitability.

Concerns About Data Tracking

Some people don't like this tracking, for a variety of reasons. For some, it feels like a violation to be treated as a mere object of commerce. Some worry that data about their interests will be used to discriminate wrongly against them, or to exclude them from information and opportunities they should enjoy. Excess customization of the Web experience may stratify society, some believe. If you are poor or from a minority group, for example, the news, entertainment and commentary you see on the Web might differ from others', preventing your participation in the "national" conversation and culture that traditional media may produce. And tied to real identities, Web surfing data could fall into the hands of government and be used wrongly. These are all legitimate concerns that people with different worldviews prioritize to differing degrees.

"Surreptitious" use of cookies is one of the weaker complaints. Cookies have been integral to Web browsing since the beginning, and their privacy consequences have been a subject of public discussion for over a decade. Cookies are a surreptitious threat to privacy the way smoking is a surreptitious threat to health. If you don't know about it, you haven't been paying attention. But before going into your browser settings and canceling cookies, Web users should ask another question about information sharing in the online world. What am I getting in return?

How Companies Make Money

The reason why a company like Google can spend millions and millions of dollars on free services like its search engine, Gmail, mapping tools, Google Groups and more is because of online advertising that trades in personal information.

And it's not just Google. Facebook, Yahoo, MSN and thousands of blogs, news sites, and comment boards use advertising to support what they do. And personalized advertising is more valuable than advertising aimed at just anyone. Marketers will pay more to reach you if you are likely to use their products or services. (Perhaps online tracking makes everyone special!)

If Web users supply less information to the Web, the Web will supply less information to them. Free content won't go away if consumers decline to allow personalization, but there will be less of it. Bloggers and operators of small websites will have a little less reason to produce the stuff that makes our Internet an endlessly fascinating place to visit. As an operator of a small government-transparency website, WashingtonWatch.com, I add new features for my visitors when there is enough money to do it. More money spent on advertising means more tools for American citizens to use across the web. Ten years ago—during an earlier round of cookie concern—the Federal Trade Commission asked Congress for power to regulate the Internet for privacy's sake. If the FTC had gotten authority to impose regulations requiring "notice, choice, access, and security" from websites—all good practices, in varying measure—it is doubtful that Google would have had the same success it has had over the past decade. It might be a decent, struggling search engine today. But, unable to generate the kind of income it does, the quality of search it produces might be lower, and it may not have had the assets to produce and

support all its fascinating and useful products. The rise of Google and all the access it provides was not fated from the beginning. It depended on a particular set of circumstances in which it had access to consumer information and the freedom to use it in ways that some find privacy-dubious.

Some legislators, privacy advocates and technologists want very badly to protect consumers, but much "consumer protection" actually invites consumers to abandon personal responsibility. The caveat emptor rule requires people to stay on their toes, learn about the products they use, and hold businesses' feet to the fire. People rise or fall to meet expectations, and consumer advocates who assume incompetence on the part of the public may have a hand in producing it, making consumers worse off.

If a central authority such as Congress or the FTC were to decide for consumers how to deal with cookies, it would generalize wrongly about many, if not most, individuals' interests, giving them the wrong mix of privacy and interactivity. If the FTC ruled that third-party cookies required consumers to opt in, for example, most would not, and the wealth of "free" content and services most people take for granted would quietly fade from view. And it would leave consumers unprotected from threats beyond their jurisdiction (as in Web tracking by sites outside the United States). Education is the hard way, and it is the only way, to get consumers' privacy interests balanced with their other interests.

•••

Only one thing is certain here: Nobody knows how this is supposed to come out. Cookies and other tracking technologies will create legitimate concerns that weigh against the benefits they provide. Browser defaults may converge on something more privacy protective. (Apple's Safari browser rejects third-party cookies unless users tell it to do otherwise.) Browser plug-ins will augment consumers' power to control cookies and other tracking technologies. Consumers will get better accustomed to the information economy, and they will choose more articulately how they fit into it. What matters is that the conversation should continue.

Educate Yourself

Unit 4 Lesson 7

Simple Encryption

Resources

Unit 4 Lesson 8

Encryption with Keys and Passwords

Resources

Name(s)_____ Period _____ Date _____

Worksheet - Exploring the Vigenère Cipher Widget

Discover: Try the Vigenère Cipher Widget !

Goals:

- Understand how the Vigenere Cipher Algorithm works
- Understand why simple frequency analysis doesn't work against this cipher
- Figure out what makes for a good v. bad secret key

Instructions:

- You should have a partner for this exploration.
- Go to the interactive Vigenère Cipher Widget
- **Click on buttons and try things out!** Solve the mystery of what this tool is doing and how it's doing it!

You should try each of the following - check off the **DONE** column once you've tried it

Try This	Details	Done
Encrypt a few different messages using different secret keys	• Enter a text message in the box and secret key • Step through the encoding of each character to see what's happening • Try a different secret key	
Decrypt a message	• Copy/paste the ciphertext of an encrypted message into the text message area. • Hit the button to "decrypt" • Now step through and see what happens	
Find a "bad" secret key	• Hint: try "A" or "AAAAA" or "GGGG" or any single character, what about other patterns? • What makes a key bad?	
Find a "good" secret key	• Use what you learned about bad keys and do the opposite • What are the characteristics of a good key?	
Try to decrypt without knowing the key (in other words: try to crack it!)	• Have one partner look away, while the other copy/pastes the ciphertext of an encrypted message into the text area, and deletes the secret key from view • Have the partner who looked away come back and try to crack the message	

Thought Questions:

You might want to play with the widget a little bit more in trying to answer these questions, but they can be answered based only on the properties of the Vigenère cipher.

- Describe in your own words what the Vigenere Cipher Algorithm is doing.

- What makes for a good v. bad secret key using the Vigenere cipher? Give examples of a good key and a bad one and explain why.

- Compare and Contrast the difference between a substitution cipher (Caesar or Random) and Vigenere, using the message "`I think I can I think I can I think I can`" to explain why Vigenère is a stronger form of encryption than a substitution cipher.

- Will frequency analysis work to crack the Vigenere cipher? Why or why not? Keep your answer as simple as possible.

- If I promised you that the message at right was encrypted with the Vigenère cipher widget, would that make it easy to crack (yes or no)? Explain why. Your explanation should include a description of what you would need to know to decrypt this and how you might go about figuring that out.

```
KNEIWNQBKFEOQXCUMSIYBJKEWNLVZ
YBFDBVNSAEZBPADGBAAXFHEHUXSFQ
OFCADPAFAFQGPLZEA
```

- What if I told you that the message above was encrypted with the Vigenère cipher widget *and* the key I used was 10 characters long. Does that make it any easier to crack the message? Again, what would you need to figure out and how would you go about finding it?

Name(s)_____ Period _____ Date _____

Worksheet - Keys and Passwords

Answer these questions

These questions are intended to be answered as part of an activity using http://howsecureismypassword.net. The questions below assume ask you to try things out using that tool.

- Create a few passwords using 8 lowercase ASCII characters (a-z). What's the longest amount of time-to-crack you can generate?

- Using any characters on the keyboard, what's the longest amount of time-to-crack you can generate with an 8-character password?

- As you try passwords, what seems to be the single most significant factor in making a password difficult to crack? Why do you think this is?

- Opinion: Is an 8-character minimum a good password length for websites to require? Give your opinion, yes or no, and explain why you think that.

- The AP CS Principles framework contains the following statement: ***Implementing cybersecurity has software, hardware, and human components.*** Based on what you've learned so far, describe at least one way that cybersecurity involves "human components."

Hopefully you can now appreciate this comic: http://xkcd.com/936/

Unit 4 Lesson 9

Public Key Cryptography

Resources

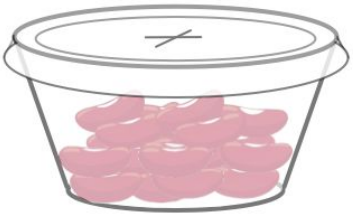
Name(s)_____ Period _____ Date _____

Activity Guide - Public Key Bean Counting

Sending Secret Messages without agreeing a on a secret key ahead of time
In this activity, cups filled with beans will represent information going back and forth between Alice and Bob. We do this activity to show you a simple version of something called **Public Key Encryption** so we can introduce you to the basic process of information exchange and to some of the terminology involved (which we'll get to later).

This activity will show a technique for Alice and Bob to send secret messages to each other, *without agreeing on a secret key ahead of time*, and only by exchanging messages over public, insecure channels.

Background: A metaphor -- cup of beans as one-way function

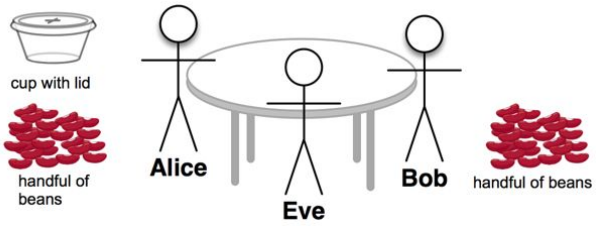


- Imagine that putting some beans into a clear plastic cup and then putting a lid on the cup is an encryption function. Only the person who put the lid on is able to remove it.
- Everyone else can try to count the beans but they can't take the lid off; they just have to stare into the cup (like trying to count the jelly beans in a jar at the carnival). This represents a **computationally hard problem.**
- In this activity, there is a wrinkle: a person *can add beans* to the cup by pushing them through the slot in the top of the lid. The result is that there will be more beans in the cup, but it's still hard to count them by looking in from the outside.

Information Exchange Procedure

Materials

- A clear plastic cup
- A few handfuls of dried beans
- (optional) A lid with a slot in the top that would allow a bean to be pushed through. **If you don't have lids,** you could use plastic wrap, or just use your powers of imagination.



Setup

- Decide who is playing Alice, Bob and Eve.
- Give all of the cups and lids to Alice to start.
- Alice and Bob should each have a handful of beans.

Eve:

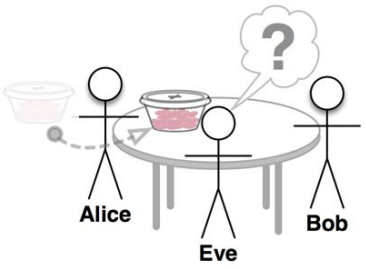
Eve, you will direct all the action. You should read all the instructions of the procedure out loud to everyone, and Alice and Bob should follow along accordingly. (Alice and Bob can follow along on their sheets as well.)
Eve reads….

Alice:

1. Alice, turn your back to Eve and Bob while you do this:

- Put a random number of beans into a cup -- *Remember this number (or write it down in a secret location).*
- Put the lid on the cup.

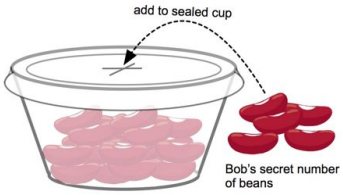
2. Then put the cup onto the table in front of Bob and Eve.
NOTE: Eve and Bob can only guess how many beans are in the cup.



Bob:

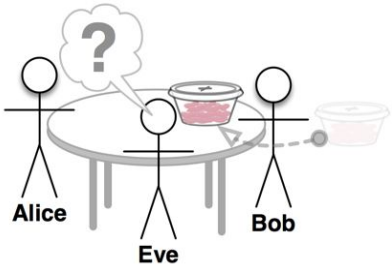
1. Bob, take the cup off the table and turn your back to Alice and Eve while you do this:

- Pick a secret number to send to Alice. *Remember this number*.
- Count out that many beans and add them to the cup.



2. Then put the cup back onto the table.

NOTE: Even though she might be able to see the number of beans is different, Eve can still only guess how many are in the cup.



Eve:

Quick question for Eve: Do you have any idea what secret number Bob is sending to Alice? *Note:* Unless Bob and Alice put so few beans into the cup that you can clearly see from the outside how many there were, your answer should be "No." You might be able to make a guess, but you wouldn't know for sure whether it was right. Okay...move on.

Alice:

Once more, turn your back to Bob and Eve while you do this:

- Take the cup off the table
- Remove the lid and dump out the beans.
- Count off the number of beans originally in the cup.

- What's left is Bob's secret number!

Recap:

- Alice and Bob did not have to agree on anything, or communicate ahead of time.
- Alice and Bob only exchanged information in public, right in front of Eve.
- Eve would have to be able to count the beans in the cup without opening it, both on the way over to Bob and on the way back to Alice, in order to determine what Bob was trying to send Alice.

Try it again?

- Change roles and try the procedure again to see how it works. Try to make it hard for Eve to guess the secret number. And, Eve, do try to guess.

- Here's a fun wrinkle that makes it even more impossible for Eve: Use 3 cups!
  - Alice, put a random number of beans into 3 different cups (you need to remember how many total beans you used, or you could remember the 3 separate numbers).
  - Bob, for the number you wish to send, distribute the beans randomly into the 3 cups; it doesn't matter how many go into each cup as long as the total is the number you want to send.
  - Alice, once you have the 3 cups back, either dump all the beans out and take away the total number of beans you originally put into the cups, or subtract the individual amounts from each cup. Either way, the beans left over are the ones Bob sent you.

Cups and Beans -- What's the point?

What's the point of the cups and beans activity?

Public key cryptography is what makes secure transactions on the Internet possible. Obviously, computers don't exchange information with beans in plastic cups; they use data (numbers mostly) and the methods of encryption use some math, which we will see in a later lesson. Here the **number of beans represented data** and the **cups represented encrypted data.** In order to see how the real thing works, we need to know some terms so we can talk about it accurately.

First, **NOTICE:**

- At no point did Bob or Alice agree on any secret password, number, or key.
- They only exchanged information in public.
- Bob can encrypt a secret message for Alice by using something that Alice puts out in public
- Eve could not tell what was going back forth without simply guessing either Alice or Bob's private number.

Asymmetric Keys
The cups and beans represent **asymmetric** (pronounced "A-symmetric") encryption because the procedure for encrypting a message (which Bob does) is different from the procedure for decrypting the message (which Alice does). *Up to this point, the encryption schemes we've studied have been **symmetric**. This means that the key used to encrypt the message is the same key needed to decrypt the message.*

Private Key
In the case of this activity, Alice's secret number - the number of beans that she put into the cup originally is known as her **private key**. Only she knows it, and she never shares it with anyone.

Public Key
The sealed container sitting on the table represents Alice's **public key**. In the real world a public key is something related to the private key, that can be safely shared in public, that another person can use to encrypt a message. In this case, the cup with the lid on top.

Encrypting (a message)
When Bob adds beans to the sealed cup, he is using a public key to encrypt a message. Since they get mixed in with the other beans (which are related to Alice's private key), no one, not even Bob, knows how many total beans there are.

Decrypting (the message)
When Alice receives the cup back from Bob, she can **decrypt** the message by opening the lid and counting the beans. Since she knows how many beans she put in in the first place, she can subtract that number of beans and arrive at the number that Bob intended to send.

Public Key Cryptography
This entire form of exchange is called **Public Key Cryptography**. In this form of secure communication, every participant has *both* a public and a private key. When sending a message, the sender encrypts his message using the **public key** of the recipient.

The *real math* is actually not that complicated. It essentially uses multiplication and division instead of addition and subtraction. The next lesson shows how it works.

Teacher Guide - Public Key Crypto Widget Activity

Summary

This guide suggests a way to introduce the Public Key Crypto Widget as a series of guided steps. You always have the option of letting students experiment with the tool first *before* explaining how it works. Here is what it looks like:

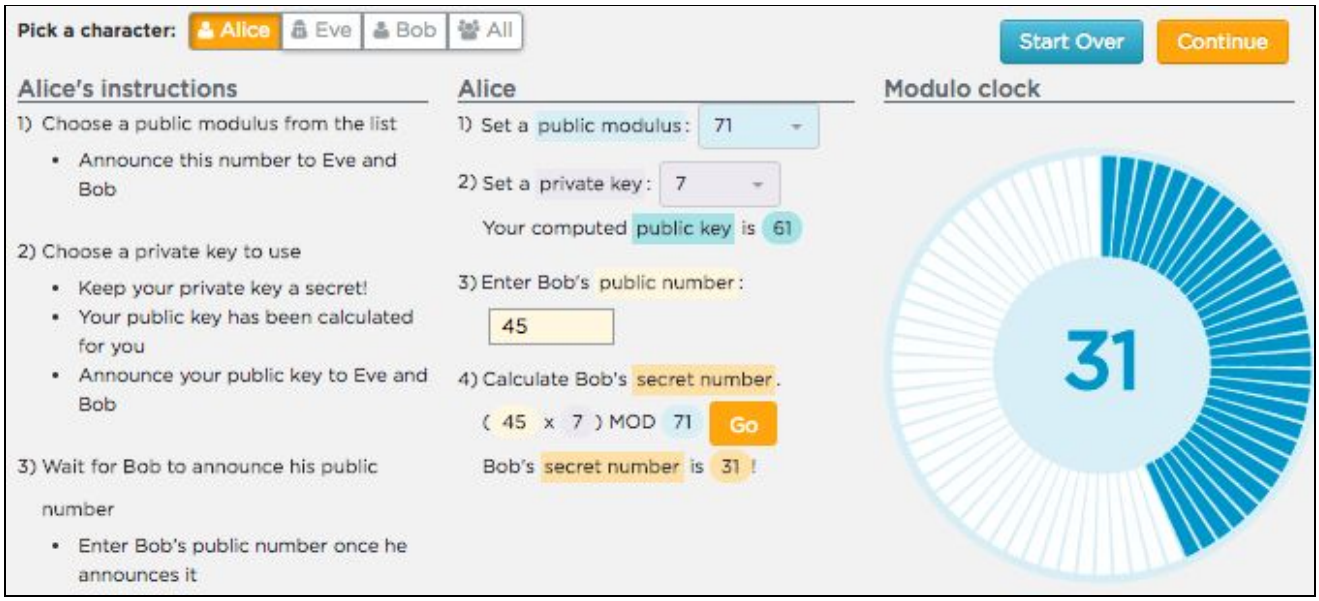


Figure 1: Screenshot of the the Public Key Crypto Widget showing Alice's screen after going through the process of creating a public key and decrypting a message,

Assumption: This guide assumes that you have done the prior activities in the Public Key Cryptography lesson, namely the "Cups and Beans" activity, and the "Multiplication + Modulo" exercises. If you haven't you'll have more explaining to do through this activity and in the wrap up.

Objectives:

- Use the widget to practice the public key encryption process
- Explain how asymmetric encryption works at a high level
- See how multiplication + modulo can be used to create asymmetric keys
- Try to crack messages encrypted with multiplication+modulo

Agenda:

This guide suggests you take the activity in 5 steps:

1. Introduce the widget and give background
2. Just play Alice and Bob
3. Show how Eve works
4. Experiment with cracking bigger numbers
5. (Optional) Use the widget with all 3 characters showing.

Discussion**:** *Details in the lesson plan*

Setup:

Group: Put students into groups of 2 (to play just Alice and Bob initially).

- Each student should be at their own computer, but within speaking distance

Display: the Public Key Crypto Widget Instructions page (in code studio). It's the page just before the widget itself.

- You can ask students to go to that page as well if you want them to read it now, or just have it displayed for you to review the instructions.

Part 1: Introduce the Widget (10 mins)

Introduce the Public Key Crypto widget providing the background and instructions given on the Instructions page in code studio. Make sure to point out the similarities and differences between using this widget and cups and beans.

Key Background Info (from the instructions page):

This widget will use numbers and math to do public key encryption, but it's important to understand that the mechanics of what you're doing are **basically the same as the cups and beans activity**

The Goal just as before, is for Bob to send Alice a secret number. But for that to happen Alice actually has to act first to create a public key for Bob to use.

So, using the widget the process is still:

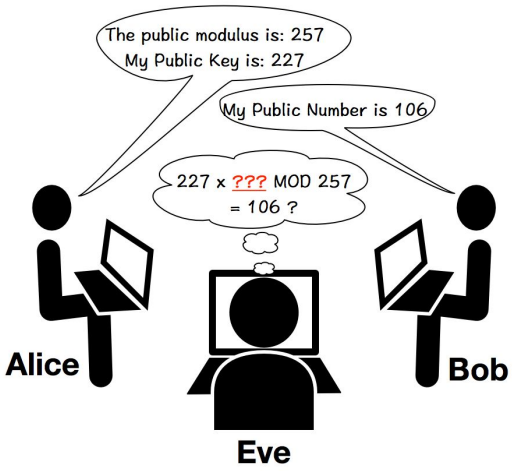
- Alice creates a private and public key
- Bob uses the public key to encrypt a secret number
- Eve can intercept all public information and tries to crack it.

The differences between the public key crypto widget and the cups and beans activity:

- **All data are numbers** - the secret messages are numbers that get encrypted by transforming them into other numbers. (This replaces secret numbers of beans "encrypted" by putting them in a cup).
- **Use your voice** to broadcast encrypted information publicly (this replaces beans in a cup getting passed around)
- **The "public key" is actually a combination of two numbers** - Alice will produce a "public key" number, but there is also a "public clock size" that is used to produce that number. Both are publicly known. Since the clock size could actually be declared by anyone, including Eve, we refer to Alice's public number as her "public key."
- **The Math: Multiplication and Modulo** - rather than simple addition and subtraction of beans, the widget uses multiplication + modulo to compute encrypted values

Demonstrate the first step of using the widget. (Click past the instructions page to get to the widget if necessary)

- Choose a character (Alice)
- Note where that character's instructions are (left side of screen)
- Clarify that the widget **does not send information** you need to **use your voice** to broadcast encrypted data.
- Note the diagram on the instructions page that shows the 3-person setup (at right).
- Note: the computers do not have to be arranged exactly this way - the point is the widget is a standalone tool. Students should be on their own computer and they talk to "send" a message.



Part 2: Just Play Alice and Bob (5 mins)

Prompt: *With a partner, just play Alice and Bob and exchange a few numbers to get the hang of it.*

- Follow the character's instructions on the screen.
- Go through the process a few times:
  - Alice produce a public key
  - Bob encrypt a secret number
  - Alice decrypt
- Communicate by just speaking out loud.
- Exchange roles at least once.
- Verify that you can encrypt and decrypt messages.

Give students some time to work with this.

When Alice reveals the secret number it should be a mini "ooh and ah" moment, or at least a "wait a second, what?" moment.

Encourage students to use both large and small numbers.

Once students have exchanged a few numbers, **regroup** for the next step.

Answers to some FAQs about the widget

Clock size is chosen randomly by Alice but there is a set list of values to choose from. The clock sizes in the list provided are prime numbers between 1 and 10,000. This ensures certain properties of the encryption.

Alice's private key is chosen at "random" but there is also a list to choose from. We've computed pairs of public/private keys behind the scenes so they have the necessary mathematical relationship. Alice simply has to pick one.

Bob is sending a secret number to Alice, not vice-versa. In public key cryptography for Bob to send a secret to Alice, Alice has to act first, producing a public key for Bob to use.

Bob can send any number to Alice - as long as the number is between 0 and (clockSize - 1.)

The clock size limits the range of values - the secret numbers that Bob and Alice use are confined to the output range of the mod clock. For example: if the clock size is 13, then Bob can only send a secret number in the range 0-12. If the clock size is 253 then the secret values can be 0-252.

Part 3: Show how Eve Works (10 mins)

Regroup: After pairs have gotten the hang of playing Bob and Alice, regroup to review how Eve works.

Display: Eve's screen in the widget.

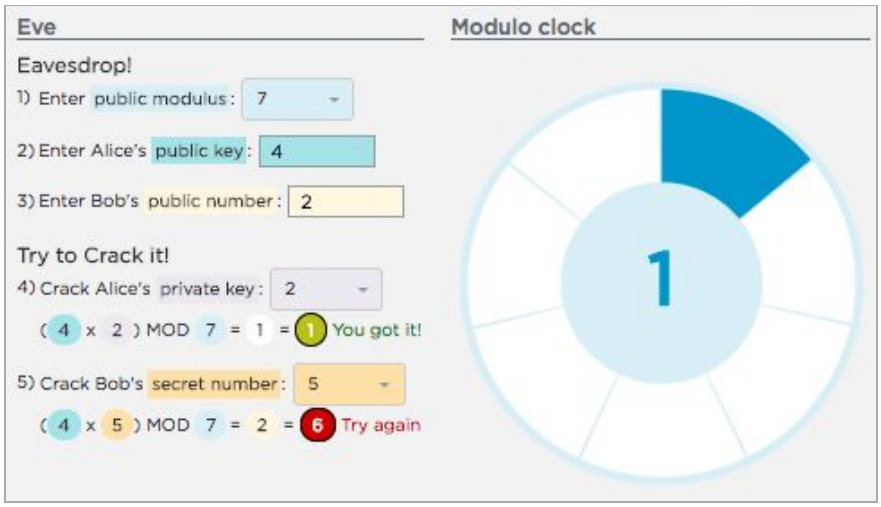


Figure 2. Screenshot of Eve's screen shown having successfully cracked Alice's private key (very easy) when the public clock size is 7.

Remark: *Okay, now that you have the hang of playing Alice and Bob let's look to see whether or not this is encryption is hard to crack. In theory, you're broadcasting encrypted values by saying them out loud, and an eavesdropper should have a hard time figuring out the secret. Let's see how hard it is.*

Pick 2 students on opposite sides of the room to play Alice and Bob.

Explain: You (teacher, and rest of the class observing) are going to play Eve.

- The widget lets Eve record the numbers being spoken in public.
- Eve also knows what computations were performed to produce those numbers - it's an "open standard" just not the exact values.
- She has to guess either Alice's private key, OR Bob's secret number. She'll know she's right if the math works out.

Prompt: For the first round let's pick a small clock size, let's say 7. Alice and Bob set your public clock size to 7 and go through the process.

- Students playing Alice and Bob should call out numbers (they will be small)
- Record them in the appropriate boxes on the screen

After Bob has announced his public number, demonstrate how to try out values as Eve to crack the code. It shouldn't take long since the clock size is so small. In fact, because of the way we create the public/private key pairs, Alice only had 2 choices!.

Prompt: Well, that wasn't too hard. Is this really secure? What could make it harder to crack?

- The answer of making the public clock size bigger should come out quickly
- That's what we'll try next

Note: There is an optional **Student Handout** that explains Eve's part.

Part 4: Experiment with Cracking bigger numbers (5-10 mins)

There are two suggested options for this part. Pick one:

Option 1: Crowd-source cracking

- Continue as a whole class, with 2 students playing Bob and Alice, and everyone else playing Eve. See how long it takes to crack.

Option 2: Small group experimentation

- Have previous Alice-and-Bob pairs get together in groups of 4.
- One pair plays Bob and Alice, the other pair plays Eve as a team of 2 (on one computer or two)

Prompt + Thinking Challenge: Exchange numbers a few more times, trying to make it hard for Eve to crack. See how long it takes and what makes it hard. At what point would you feel "safe" as Alice or Bob that your messages were basically secure?

Thinking Challenge: As you play with the widget can you figure out why it works? Why can Alice decrypt the message but Eve can't? The widget is using multiplication + modulo to encrypt but it's not very obvious why Alice can decrypt using her private key. See if you can get to the bottom of it.

Encourage students to start with small clock sizes to get the hang of cracking the message, and then increase the clock size a few times.

Notes:

- It's actually quite tricky to understand why it works, and students are not expected to figure it out completely. We're going for an intuitive sense.
- **Hint**: Alice's public key is not random. It is computed carefully based on the public clock size and her choice of private key.
- **Short explanation:** Alice's public/private key pair is chosen so that when Alice does `pvt * pub MOD clock` the result is `1.` This means when Alice multiplies Bob's encrypted message by her private key, it effectively cancels out the public key, because it leaves Bob's secret number * 1...which is just Bob's number.
- See lesson plan for more details.

(Optional) Part 5 - Use the “show all 3” version of the widget

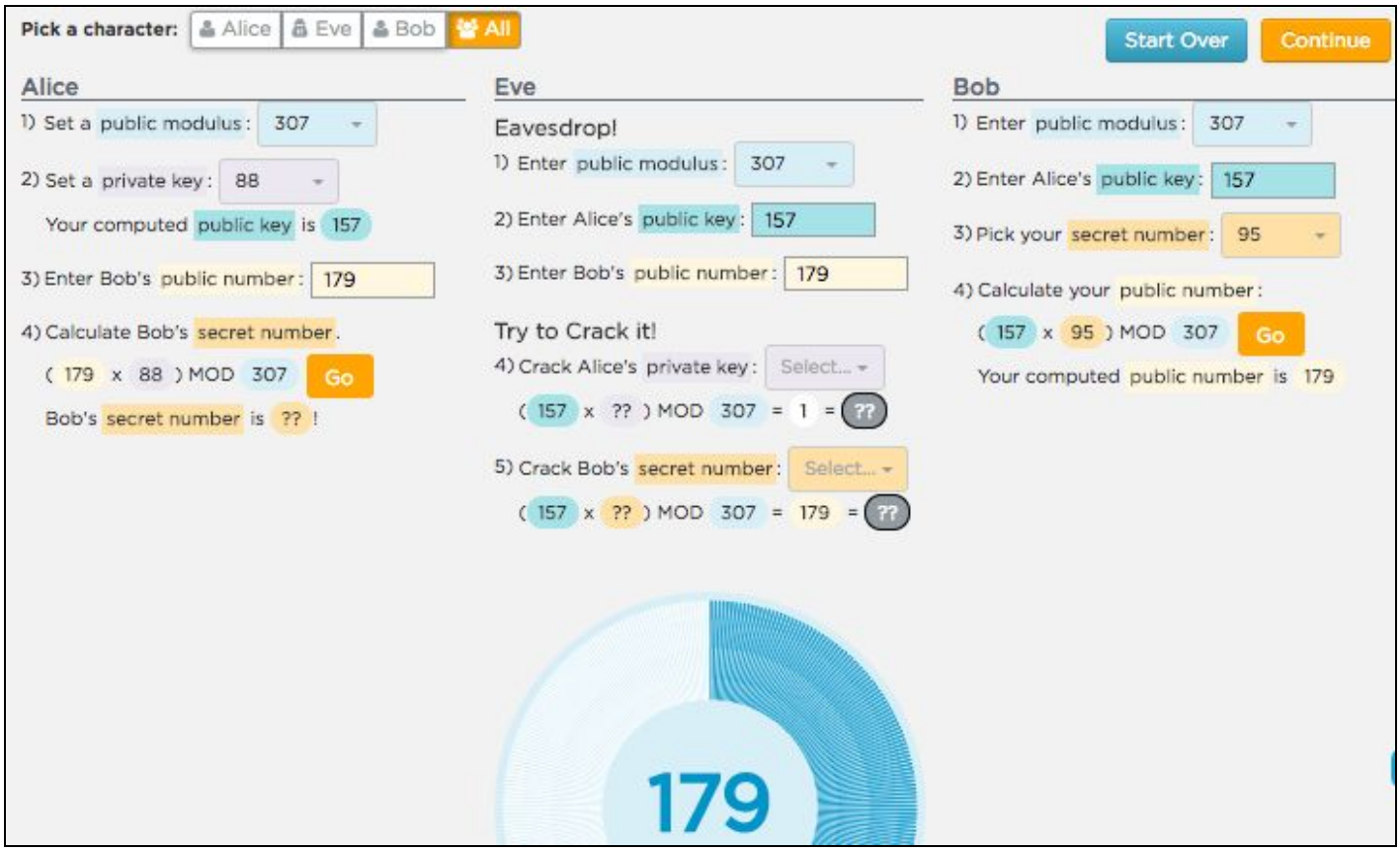
Prompt: Look at the “all” tab in the widget, which lets you act out and see all 3 characters at the same time by yourself. Try this out for a few rounds and see if you get a sense for why it works.

Encourage students here to play with small values so the can get a sense of the relationships between the numbers.

Note: Eve’s inputs are in the middle of the screen to visually represent her “man-in-the-middle” status. When using the “all 3” version you should still do the process in order:

1. Alice create public key
2. Bob encrypt secret number
3. Alice decrypt

The “all 3” version automatically fills in values for Eve.



Name(s)_____ Period _____ Date _____

Activity Guide - Multiplication + Modulo

Summary

In this activity you'll multiply numbers as input into the modulo operation and explore some interesting properties that relate to cryptography.

Goal: Understand how multiplication + modulo can be used to make computationally-hard-to-crack encryption.

Tools:

- **Calculator:** you probably want a calculator handy for multiplying big numbers
- **The "Mod Clock" widget** in code studio (pictured at right)

Assumption: You have been introduced to the modulo operation and the "clock" analogy for it.

Step 1: Experiment with the Mod Clock

Goal: familiarize yourself with properties of the Modulo operation

Get your feet wet - play

- Try inputting different values into the mod clock for *both* the "number" and the "clock size".
- Try big numbers and small numbers for both

Questions:

1. Using a clock size of 50, write a list of 5 numbers that produce a result of 0.

2. With clock size of 50 how many total numbers are there that produce a result of 0? (If the list is short, write it out. If the list is long, describe a pattern of what the numbers are).

3. Using a clock size of 13, can you find a number to input that produces a *result* of 13? (If so, what is it? If not, why not?)

4. Using a clock size of 13, find the answers to the following:

1 MOD 13	
10 MOD 13	
100 MOD 13	

1,000 MOD 13	
10,000 MOD 13	
100,000 MOD 13	

Are these results surprising or interesting? Why or why not?

Step 2: Toward encryption - Use multiplication to produce inputs

Experiment - Small changes to inputs, big changes to outputs.
Using a clock size of 37, let's multiply two numbers (we'll call them A and B) to use as input, then make small changes to each while holding the other constant. We'll always use the formula **`A * B MOD M.`** We'll start with A=20 and B=50 and M=37. So here is the first result...

`20 * 50` MOD 37 = **`1`**

Now find in the rest of these values making small adjustments to A and B individually.
*Use a calculator, if necessary, to compute A * B. Use the Mod Clock to compute the modulus of the result.*

Increment A	
21*50 MOD 37	
22*50 MOD 37	
23*50 MOD 37	

Increment B	
20***51** MOD 37	
20***52** MOD 37	
20***53** MOD 37	

Result: What do you notice about these results? Is there a pattern? Could you predict the result of **`25 * 50 MOD 37`**?

Experiment 2 - Guessing inputs is hard?: Using a clock size of 101, we'll give you the value of A and even hold the result of the modulo operation constant. **Your task:** find a value for B (the blank) that makes the math work out.

1) **`2 * ____ MOD 101 = 1`**

2) **`3 * ____ MOD 101 = 1`**

3) **`4 * ____ MOD 101 = 1`**

Takeaways:
Solving modulus equations like **`2 * ___ MOD 101 = 1`** is "hard" because you can't solve it like a typical equation. There are no easy patterns or shortcuts like other equations you might see in a math class. As you learned in step 1 (hopefully) there is an infinite list of single values for which **`___ MOD 101 = 1`**. (The list is 1, 102, 203, 304, 405...etc). With multiplication, to solve **`2 * ___ MOD 101 = 1`** you end up randomly guessing to find some number to multiply by 2 that gives you a result in that list.

Things get especially "hard" when you use a prime number as the clock size. Thanks to some special properties of prime numbers **with a prime clock size there's only one solution to each modulus equation.** You are guaranteed that there is the number less than the clock size itself, but there are still 100 different values you have to try. With a **brute force search** you could go through them all in a couple minutes. But what if the clock size were a 50-digit prime number?

Encryption!
Whenever you have a problem for which the only way to solve it is by random guessing or brute force search over a large range of values, you have a candidate for a encryption. Next you'll get to try it!

Teacher Guide - Modulo Clock Thought Experiment

Thought Experiment - Clock as a one-way function

Any kind of encryption requires transforming information in a way that is hard to reverse without a key.

A “one-way function” is a math operation that is impossible to reverse or solve even if you know some of the inputs that went into it. But it’s not random. GIven the same inputs, it will produce the same result. There is just no way to reverse the process.

As an example, let’s do a thought experiment:

> *Imagine that you are a person who loses complete track of time when you close your eyes. When you open your eyes, a minute could have passed or an hour…or a day…or a week…or a year…you don’t know.*
>
> *So, now imagine a clock that reads 4:00.*

Show a clock of some kind that shows 4:00; here is an interactive one.

> *Now close your eyes and I’m going to add some time to the clock - I’m going to simulate that some amount of time is passing. Remember, with your eyes closed, any amount of time could be going by.*

Set the clock to show 3:00.

> *Now open your eyes, look at the clock and, without saying anything to anyone, write down how much time has passed.*

Wait a few seconds for students to write.

Prompt: So, how much time passed? What are the possibilities?

Let students share answers:

- There are an infinite number of possibilities, including: 11, 23, 35, 47 hours, etc. Or 1 day and 11 hours, and so on.
- If students want to know what you were thinking, make up something that no one has said yet, something like, “Oh, I was actually imagining that I was adding 13 years, 47 days and 11 hours.”

BEFORE: clock is showing 4:00.

AFTER: clock is showing 3:00.

Takeaway: Clock is a one-way function
There is no way to know the original input just from looking at the face of the clock. No matter what number you put into it, only numbers 1-12 can show afterward. Even if the number is 2,023,789 hours, if you wind the clock around, it will still come out as a number 1-12. We cannot know what the original number was that went into the clock.

Clock is a metaphor for modulo

Real cryptography uses this “clock” technique to obscure information, but with clocks that can have a wide range of possible values on their faces. **The operation is called modulo.** Modulo is a math operation that returns the *remainder* from dividing two integers. It is important for cryptography because it can act as a one-way function - the output obscures the input.

More points about the modulo operation can be found in the lesson plan.

Teacher Demonstration Guide - Public Key Bean Counting

Synopsis:

This guide shows you how to run the “Cups & Beans” Demonstration as a teacher. *Detailed instructions are below.*

1. Introduce the characters "Alice, Bob and Eve"
2. Explain the "Cups & Beans" metaphor - how beans in a cup represents encrypted information
3. Demonstrate the Information Exchange process
    - Alice make a public key
    - Bob use public key to encrypt message
    - Alice decrypt
    - (Eve sees all information going back and forth but cannot decrypt)
4. Recap the mechanics of how public key cryptography works
5. Cups and Beans - What’s the point?

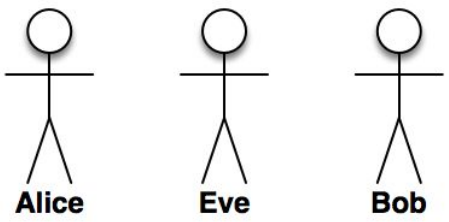
Detailed instructions:

1. Introduce the characters "Alice, Bob and Eve"

- Explain by saying:

> *In cryptography scenarios computer scientists use stock characters:*
>
> - ***Alice** and **Bob** (“A” and “B”) who are trying to send messages to each other*
>
> *(This is because many diagrams show messages going from point “A” to point “B”)*
> - ***Eve** the ‘eavesdropper’ is listening in.*
>
> *You should always assume that Eve can see everything that Alice sends to Bob and vice versa. Eve is a smart adversary who knows what Alice and Bob are trying to do. The goal is to make it hard for Eve to crack the encryption even if she knows how it’s done.*

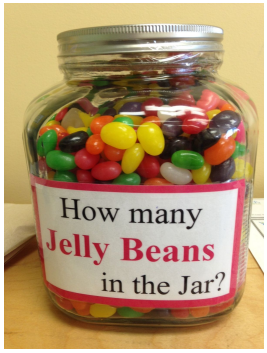
- Then select 3 student volunteers to act as Alice, Bob and Eve
- Have them stand in the front of the room, an arm's width apart, with Eve standing between Alice and Bob.



2. Explain the "Cups & Beans" metaphor - how beans in a cup represents encrypted information

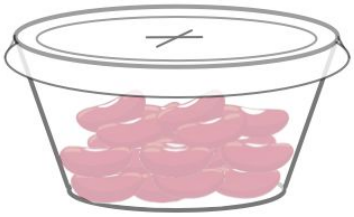
- Explain the metaphor by saying

  - Have you ever been to a carnival or fair where there's a big glass jar of candy (usually jelly beans) and you're supposed to guess how much candy is in the jar?
  - ***We're going to use that idea as a metaphor for an encryption function***
  - It's easy to count out some candies and put them into a jar, but really hard, if not impossible, to guess how many there are once the jar is closed and locked.



- Explain how the demonstration will work, saying:

For to today:



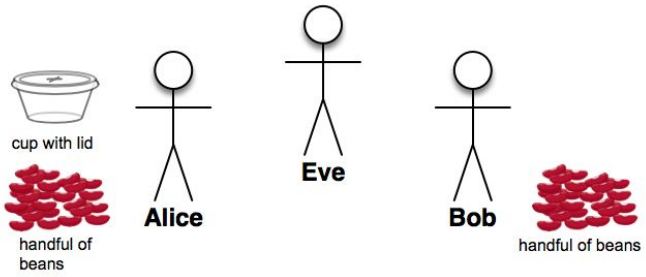
- Our secret information will be some number of beans
- We'll "**encrypt**" that information by putting some beans into a clear plastic cup and then putting a lid on the cup.
- We'll simulate information traveling across **the Internet** by passing this cup of beans from one person to another.
- We'll have to imagine that whoever puts the lid on the cup also locks it - so o**nly the person who put the lid on, can take it off**
- Anyone can try to count the beans in the cup but they can't take the lid off; they just have to stare into the cup (like trying to count the jelly beans in a jar at the carnival).
- This represents a **computationally hard problem,** since it reduces your ability to count the beans to essentially random guessing
- One wrinkle: **We'll allow a person to add beans to the cup** after the lid has been put on by pushing them through the slot in the top of the lid.
  (*Note: if you don't have lids it's okay, imagine that too)*

3. Demonstrate the Information Exchange process

Now we're ready for the demonstration.

- Give Alice a cup and a handful of beans
- Give Bob a handful of beans.

Read aloud these instructions and aid in counting out beans and passing the cup.

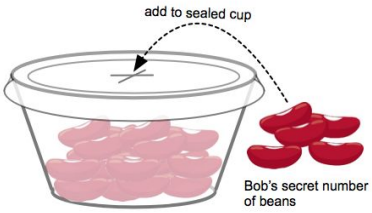


The Setup

1. **Alice Bob and Eve** none of you can move. You can only pass a cup of beans back and forth.
2. Our goal is for **Bob to send a secret message to Alice.** For this to work though, Alice must produce a public key for Bob to use. So...

3. **Alice**: secretly count out some number of beans place them in the cup. Don't let anyone see (except me)
4. **Alice:** put a lid on the cup and pass the cup to Bob out in the open where everyone, including Eve, can see the cup. We'll call this the **public key cup.**

   (*Remark that in our metaphor Eve cannot tell how many beans are in the cup, she can only guess.Typically for this the teacher might carry the cup across the room, or ask Alice to pass it to Bob right in front of Eve, or even have Eve hand it to Bob, or even have every student in the class pass it around until it ends up at Bob)*



5. **Bob**: pick a secret number you wish to send to Alice and add that many beans to the cup. So the cup now contains Alice's Beans and Bob's Beans.
6. **Bob**: pass the cup back to Alice in plain view of everyone.
   (*Again, the metaphor here is that we've just added beans to the cup, it's still just takes random guesses to know what's in there)*
7. **Alice**: upon receipt of the cup, dump out the beans, and subtract the number of beans you secretly placed in the cup in the first place.

   The remainder is the secret number Bob sent you!

Verify that this worked with Alice and Bob - Did Alice correctly receive Bob's "secret message"?

(Applause Applause Applause)

4. Recap the mechanics of how public key cryptography works

Review with students the following points about the demonstration. Did you notice...?

- At no point did Bob or Alice agree on any secret password or key.

- They only exchanged information in public.
- Crucial: If Bob wants to send a message to Alice, Alice has to act first by producing a cup of beans that Bob could use (her public key)
- Bob can encrypt a secret message for Alice by using something that Alice puts out in public -- If Alice wanted to send a message to Bob, then Bob would have to produce his own cup to put out in public
- Eve could not tell what was going back forth without simply guessing either Alice or Bob's private number.

5. Cups and Beans - What's the point?

Main Takeaways and Terminology:

- Obviously on the Internet information is not exchanged as beans in cups.
- Our demonstration **DOES NOT show or explain** how the math or encryption works (we'll get to that next)
- What it **DOES show are the mechanics of public key communication**: How public and private keys are used to encrypt information.
- Here are the terms you should know:

Term	Description	Cups and Beans Metaphor
Private Key	A secret piece of information, like a password.	Alice's secret number of beans
Public Key	Information produced using the private key, but transformed in such a way that it's difficult to determine the private key. This can be safely shared in public, and used to encrypt other information.	Alice's secret beans sealed inside a plastic cup
Encrypted Message	Information encrypted using the public key. Because the *private key is subtly mixed into the public key*, this transforms a secret message in such a way that only the person who knows the private key can decrypt it	Bob adding beans to Alice's public cup. The "encrypted" cup of beans contains Bob's secret message and Alice's private key, but only Alice knows how many beans were in there in the first place. So only she can decrypt the message.
Asymmetric Encryption	Encryption that uses *different* keys for encrypting and decrypting. It allows for sender and receiver to communicate without having to agree on a shared encryption key ahead of time.	Bob used the public key to encrypt his message, but Alice used her private key to decrypt.
	Point of confusion: there is *some* amount of encryption involved to produce the public/private pair of keys in the first place. But Alice isn't encrypting a message to send, rather she is producing a key that others can use to send her a message).	

Next we'll learn about the mathematical principles that make public/private keys possible.

Activity Guide - Recap: The Public Key Cryptography Widget

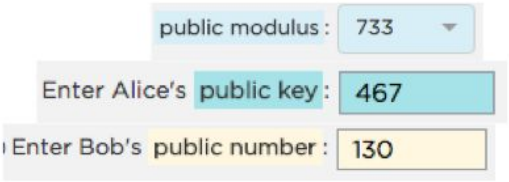
This document assumes you have used the Public Key Crypto Widget.

What does Eve have to do to crack the message?

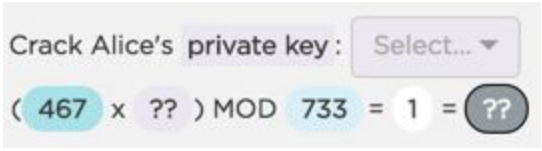
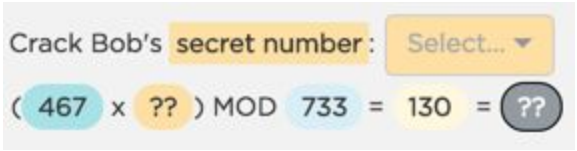
It turns out that figuring out what numbers to plug in to crack Alice or Bob's secret numbers so that the equations work out is essentially random guessing. If you want to read more about why check out the "How and why does it work?" document which gives a deeper explanation of the math behind the widget.

What does Eve know?

Eve knows only public information announced by Alice or Bob



To crack the message she must use this info to guess Alice's Private Key or Bob's Private number.

To Crack Alice's Private Key...	...or to Crack Bob's Private Number
 To crack Alice's private key Eve must figure out what number to multiply by 467 then MOD by 733 to come out = 1 (why 1? It's explained in the *How and Why does it work?* document)	 To crack Bob's secret number Eve must figure out what number to multiply by 467 then MOD by 733 to come out = 130.

Recap - Properties of Public Key Encryption:

- Alice and Bob did not have to agree on anything, or communicate ahead of time.
- Alice and Bob only exchanged information in public, right in front of Eve. Eve could even chose the public modulus if you wanted to!
- Eve would have to guess either Alice's private key or the secret number Bob is trying to send.
- *Note:* The encrypted messages only go one way. If Alice wanted to send a message to Bob, Bob would have to generate his own public/private key pair.
- However, because the message can only go one way, the sender can feel safe that ONLY the intended recipient can decrypt the message.
- The *real* public key crypto does not quite work this way. What we have setup here emulates many important properties, but it is in reality not hard for a computer to crack - just hard for you.
- The *real* version is called RSA encryption and rather than simple multiplication it uses exponentiation in combination with properties of modulo to create even larger numbers that are even harder to guess.

What do you actually need to know?

- **Cryptography has a mathematical foundation.**
- It relies on **asymmetric** keys, which you can make using numbers and math.
- The **modulo** operation acts as a one-way function.
- When you multiply big numbers and mod them by other big numbers, it's really hard to figure out what the original numbers were; the technique is essentially reduced to random guessing.
- The security of publicly known encryption protocols is based on the fact that cracking a message by brute force would take an unreasonable amount of time.
- With a sufficiently large modulus (say, 256 bits, which would be roughly a 77-digit number), random guessing would take an unreasonable amount of time. Even if you had millions of computers working on it constantly, it would take trillions of years.
- Because the method of encryption is public, it actually *increases* the security, because good guys and bad guys know how hard it is to crack.

Name(s)_____ Period _____ Date _____

Resource - How and Why Does the Public Key Crypto Work?

How does the *Public Key Crypto Widget* actually work?

Short Version:

Alice's public/private key pair is chosen so that when Alice does `pvt * pub MOD clock` the result is `1.` This means when Alice multiplies Bob's encrypted message by her private key, it effectively cancels out the public key, because it leaves Bob's secret number * 1...which is just Bob's number.

Medium Version:

1. Alice picks a clock size (`clock`) and a private key (`pvt`). Her public key (`pub`) is *specifically* chosen so that `pvt * pub MOD clock = 1`
2. Bob picks a secret value (`secret`) and his public value is computed as: `pub * secret MOD clock` let's call the result `bobPub` . **Important note:** Bob's choice is limited to values strictly less than `clock` (the range: 0 to `clock-1)`
3. When Alice gets `bobPub` she computes: `bobPub * pvt MOD clock`

Now let's look at Alice's final equation, but substitute in the expression for `bobPub`. This gives us:

`(pub * secret MOD clock) * pvt MOD clock`

If you read more below (about how modulo distributes) you know we can refactor this equation to be:

`(pub * pvt MOD clock) * secret MOD clock`

The new underlined portion in the expression is Alice's equation from step 1 -- it works out to `1` because of how we chose Alice's public/private key pair. Thus to finish, we can substitute 1 for that expression, giving us:

`(1) * secret MOD clock` which is just `secret MOD clock`

Because Bob's secret number must be less than clock, we know that `secret MOD clock = secret`

Longer Version with more Details:

Reference: Public Key Crypto Widget Activity

0. **Modulo distributes**
First, a fact about modulo that's important to realize: If you MOD a number once, the result is less than the modulus. And any number that is less than the modulus, when MODed, results in itself. Therefore, if you MOD a number *and then* MOD the result of *that*, you get the same number. For example:

23 MOD 17 = 6
→ 6 MOD 17 = 6
→ 6 MOD 17 = 6…..and so on.

Also, modulo is distributive, which means that if you multiply some numbers and MOD the result, that has the same effect as MODing all the numbers in the first place, then multiply them, then MODing the result of that.

(**27** * **49**) MOD 17 == 14 is equivalent to: [(**27** MOD 17) * (**49** MOD 17)] MOD 17 == 14

The effect of this, as you'll see later on, is that it means as long as MOD is being applied to some string of multiplied numbers, you can MOD any of the individual terms as much as you like. As long as you MOD again at the end, the result will come out the same as if you just did one single MOD operation.

1. **Prime numbers**
The numbers we let you use for modulus (clock size) are all prime numbers. Prime numbers are useful because no numbers divide them, except for 1 and themselves. Modulo is a form of division. So if we only divide values by prime numbers, the results will be unique and distributed evenly over the range of possible values.

2. **Modular multiplicative inverse**
Alice's public key is not random; it is generated based on the public modulus (clock size) and her choice of private key. Alice's public key is calculated in a special way. Here is the equation:

```
(private * public) MOD clock = 1
```

Thinking about Alice's public/private key pair, for example: let's say the public modulus (clock) is 17, and Alice's public key is 5. To figure out her private key X then you have solve this:

```
( X * 5 ) MOD 17 = 1
```

What's interesting about this equation is that you cannot just "solve for x" in a traditional way. The only way to find X is by brute force trial and error - trying out different values for X and plugging them in. You know that X must be in the range 0-16 but beyond that, not much. Furthermore, there is no way to approximate, "get close", or narrow down the answer in any way. You simply have to try out all the values. After some experimentation you'd find that X must be 7, satisfying the equation: (7 * 5) MOD 17 = 1

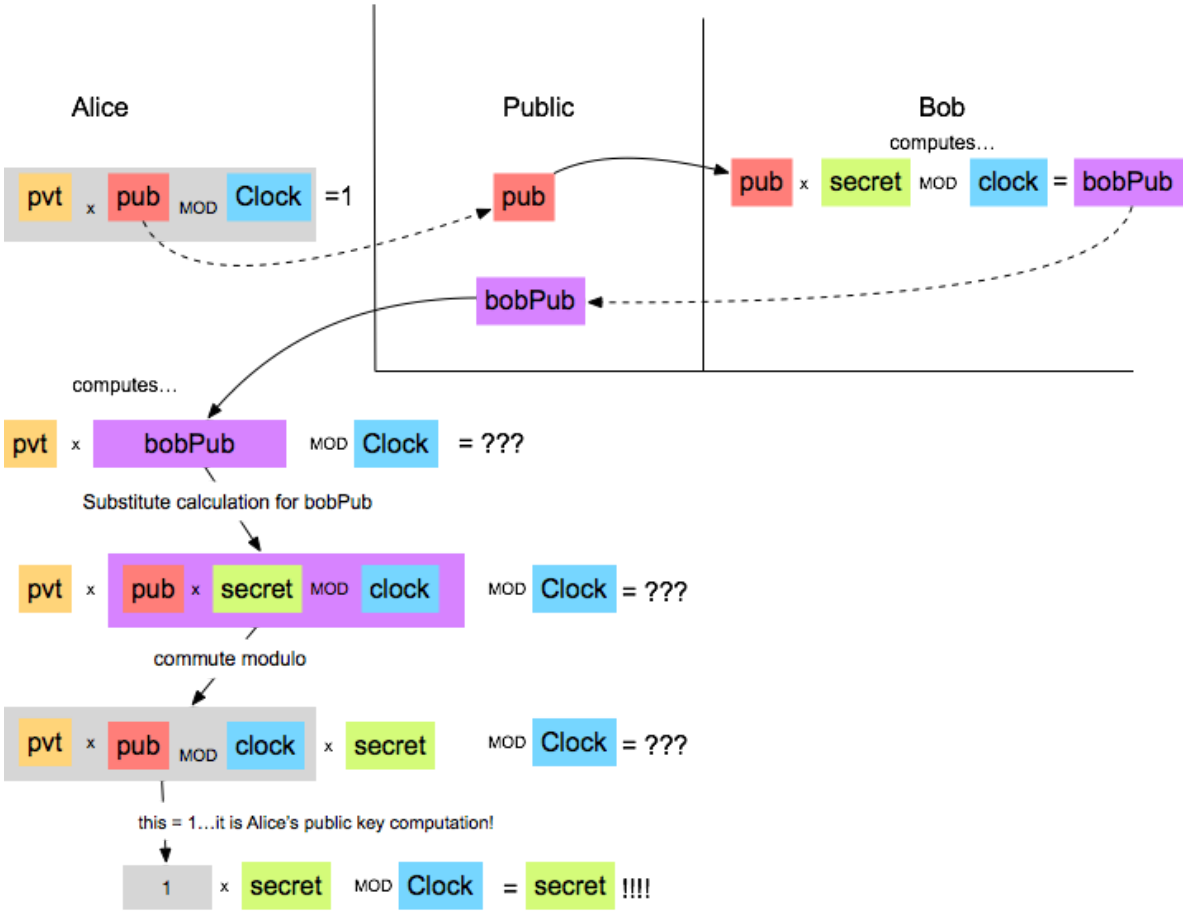
Thus, 5 and 7 form a public/private key pair in our widget (which one is public doesn't matter, they are a reciprocal pair). In the widget we simply calculate ahead of time all of these possible pairs based on whatever is chosen as the public modulus. Bob's secret number, and the resulting public value that he can share, works the same way. In both cases, **the secret value is obscured by the modulo operation, which makes cracking it a more computationally intensive process.**

The math-y term for what we're calculating is the "modular multiplicative inverse".

3. Why Alice can reveal Bob's secret number and Eve can't

Let's now look at the rest of the exchange since it's still not totally obvious *why* Alice can reveal Bob's secret number at the end but Eve cannot. It hinges on the fact that for Alice (public * private) MOD clock **is equal to 1.**

The diagram below traces the exchange of information and tries to reveal what's happening "underneath the hood". Notice the last few steps where we rearrange the values to show how the fact that Alice's public/private key equation is equal to 1 is what allows the secret to be exposed for her. Eve does not have those numbers.



Is this *actually* computationally hard for Eve to crack?

The answer is, no, not really. It's hard for a *human* to figure out, because a feature of modular multiplicative inverses is that they are a unique pair for every modulus you could choose AND they are randomly distributed across the number line. So that means there's no heuristic for "getting close"; if you discover a number that gets you one away from your target, you are no closer than if you were 1000 away. So you can't narrow down the field; the only thing you can do is try every possibility. Since the widget maxes out at only 4-digit numbers, you could brute force attack the problem in seconds with a computer program. (In fact, that's what our widget does.) Even for large numbers, it wouldn't be hard to find the multiplicative inverse.

How close is this to the *real* thing?

0. Our scheme here most closely resembles the RSA public key encryption system. (RSA doesn't stand for anything related to cryptography; it's just the initials of the last names of the 3 people who invented it: Rivest, Shamir, and Adleman.)

1. RSA does rely on multiplication and modulo, but instead of just multiplying numbers plainly, the private keys and secret numbers are used as *exponents* when multiplying large numbers, which makes even bigger numbers that are even harder to reverse after modulo.

2. Rather than finding the multiplicative inverse, to crack RSA you need to find the prime factors of a very large number (i.e., given a large number, find two prime numbers that, when multiplied together, yield that number). This is a much harder problem (computationally) than finding the multiplicative inverse (which is what our widget does).

3. The real thing uses VERY large numbers. In this widget, the biggest numbers only had 4 digits (~13-14 bits). The actual numbers used are over 75 digits long (256 bits)! If you think you can fathom how big a number with 75 digits is, you can't. The distance to the edge of our solar system in *inches* is only a 14-digit number.

Unit 4 Lesson 10

Rapid Research - Cybercrime

Resources

Name(s)_____ Period _____ Date _____

Rapid Research - Cybersecurity and Crime

Project Overview: Cybersecurity and Crime

In this small project, you will research a recent cybercrime event and present a "one-pager" about it. In particular you will focus on the data privacy and security concerns raised by the event.

The One-Pager

You will do a bit of online research and then prepare a one-page summary or "one-pager" to show the rest of the team or colleagues about the highlights of what you found. For this project you will prepare a one-pager that explains:

- The details of a specific recent cybercrime event
- The specific data security or privacy concerns raised by the event.

General Process & Requirements

- Review the One-Pager Template (provided by your teacher) and the **Rubric** below.
- Choose a cybercrime event using the guide below to help.
- Conduct your research by following the **Research Guide** below.
- Complete the one-pager.

Choose Your Cybercrime Event

You should choose a recent cybercrime event that you find personally relevant or interesting. For the purposes of this project, we'll define a cybercrime event as any instance where digitally stored data falls into the hands of someone not originally intended to have access to it.

Read *How Not To Get Hacked*

Get some ideas of different types of cybercrimes or risks by reading the

- (link in Code Studio) https://code.org/curriculum/csp/docs/hownottogethacked
- Each of the 9 tips listed is related to a particular type of cybercrime

Choose an Industry / Product / Company of Interest

Choose an industry, product, or company of interest and try to find instances of it having been hacked or leaked in some way. Aim to find stories where a piece of technology actually was hacked or failed, leading to the release of data. Avoid stories where someone just posts private information online.

Potential search terms include:

- "_____ leak"
- "_____ hack"
- "_____ breach" or "_____ data breach"

Check the news

The rate of cybercrime seems only to be increasing, and it's likely that recent news includes some instance of cybercrime. Search through recent news stories and see if you can quickly identify a cybercrime event as your topic. Just make sure your cybercrime actually involves data falling into the wrong hands.

Research Guide - Cybersecurity and Crime

Conduct Your Research

You already have some practice finding good resources online. You'll want to find **recently published documents** from **authoritative sources.** There is no need to use overly technical documents, but keep an eye out for familiar terminology and topics.

Key Information to Find

- **Overview:** Whose data was stolen? When did this happen? Briefly explain the context of the event.
- **Data Specifics:** What specific data fell into the wrong hands?
- **How was it stolen / How to Prevent:** How specifically was the data stolen? Is this a flaw in the technology? Were there any cybersecurity measures in place? How might this type of attack be prevented in the future?
- **Data Privacy / Security Concerns:** What specific concerns arise from this data being stolen? Is there already evidence of the data being used in concerning ways? Try to find how the privacy or security of some people were compromised.

Use the tables below to keep track of your information; you can also add more if you like. **You'll need to include at least 3 sources of information** but you can use more.

My Cybercrime Event: _____

Reference Name: **Year Published:**	**URL:**
Key Information	

Reference Name: **Year Published:**	**URL:**
Key Information	

Reference Name: **Year Published:**	**URL:**
Key Information	

Rubric - Cybersecurity and Crime One-Pager

Component	1	2	3	Score
Research Guide				
Sources	Research Guide includes references to fewer than three sources and the source listed is not recent and authoritative.	Research Guide includes references to fewer than three sources or the sources are not recent and authoritative. Sources may be cited in one-pager.	Research Guide includes references to at least three recent, authoritative sources. Sources are cited throughout the one-pager.	
Written Responses				
Overview	The response does not clearly describe the event. Does not describe whose data was stolen or accessed, or when it happened.	The response describes a generic or class of cybercrime event, but is not specific or vague or unclear about who was affected or when.	Response clearly describes a specific event, when it happened and who or what was affected. The context of the event is clear.	
How and How to Prevent	The response provides a vague description of how the data was lost / stolen. The response does not describe either the type of cybercrime used to access the data or the cybersecurity methods that could be used to defend it.	The response provides some details about how the data was lost / stolen. The response may describe the type of cybercrime used to access the data or the cybersecurity measures that could defend it in the future.	The response provides clear details about how the data was lost / stolen. The response correctly and clearly identifies both the type of cybercrime used to access the data and the cybersecurity measures that could defend it in the future.	
Data Specifics	The description of the data that was lost / stolen lacks any specific details. The description may discuss the device that was compromised rather than the data it was capturing.	The response identifies the category of data lost / stolen but may not provide specific details. The response does refer to data specifically, rather than the device used to store or capture it.	The response specifically identifies the types of data that were stolen / lost in the event. If the response describes both the device capturing the data and the data itself it clearly distinguishes between the two.	
Data Concern	The concerns described are not directly related to the data that was lost / stolen.	The response describes general data security or privacy concerns without specifically tying them to the data released in the event.	The response describes a data security or privacy concern directly related to the specific data that was leaked. It may be reinforced with a citation to a news story about the aftermath of the leak.	

Name(s)_____ Period _____ Date _____

Cybersecurity and Crime One-Pager Template
<your title goes here>

Note: *All text in Italics, including this text, is intended to be replaced by your responses, and deleted once you've completed your one-pager.*

Overview

When did the event happen?
Whose data was lost / stolen / leaked? How many people / organizations were affected?
Provide any other context necessary to understand the "big picture" of the event.

How and How to Prevent

What specific type of attack / mistake led to the data falling into the wrong hands? Reference terms in "How Not to Get Hacked" where applicable. What types of cybersecurity techniques might be used to help prevent this from happening again?

Data Specifics

What specific data was stolen? Try to avoid vague terms like "financial data" and instead find the specific pieces of information like "credit card numbers". Specific answers here will strengthen your explanation in the next section.

Data Privacy / Security / Storage Concern

What specific concerns arise from this data being used in unintended ways or by unintended people? Is there already evidence of the data being used in these ways? Cite sources if you can find specific news stories.

Sources

List all websites that you used to find any information you wrote here. Include the permanent URL. Identify the author, title, source, the date you retrieved the source, and, if possible, the date the reference was written or posted. You should number your sources, here is a template you can follow:

[1] Author's Last name, First name. "Title of Web Page." Title of Website, Publisher, Date, URL. Date retrieved.
[2] Author's Last name, First name. "Title of Web Page." Title of Website, Publisher, Date, URL. Date retrieved.
[3] ….

Name(s)_____ Period _____ Date _____

Worksheet - Video Guide for “Cybersecurity and Crime”

Overview

Cybercrime causes huge problems for society - personally, financially, and even in matters of national security. In this video, Jenny Martin from Symantec and Parisa Tabriz from Google explain what cybercrime is, how the same advantages in the Internet’s structure can be exploited as disadvantages, and how to defend against attacks with cybersecurity.

Directions

1. Watch the video, “The Internet - Cybersecurity and Crime”.
2. Answer the questions below.

Questions

1. Name 3 specific examples of cybercrime.

2. What is a computer virus?

3. What is a Distributed Denial of Service attack?

4. What is a phishing scam?

5. Pick a type of cybercrime and explain how to defend against it using cybersecurity techniques mentioned in the video.