Getting Started with Data

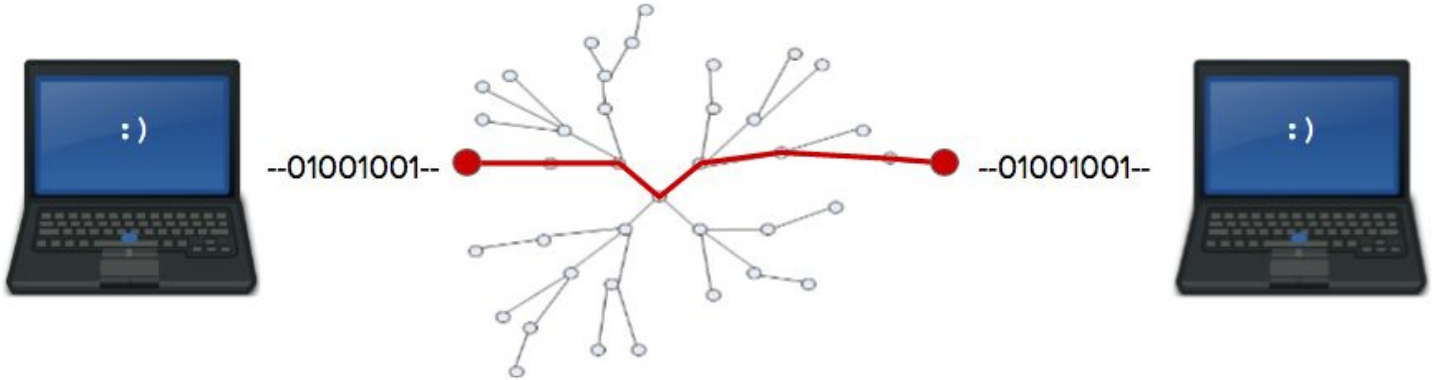
Recap: Digital Information

We learned how basically all information can be represented **digitally** as binary digits, or bits! (0s and 1s)

Sentence	HI	
Character	H	I
Number	72	73
Binary	01001000	01001001
Hardware		

Recap: The Internet

We learned how that binary data can be transferred between different devices across the Internet!



This Data Unit

We'll be learning how computers are used to:

- Collect
- Store
- Manipulate
- Visualize

data, in order to answer questions and gain knowledge about the world!

What is "Data"?

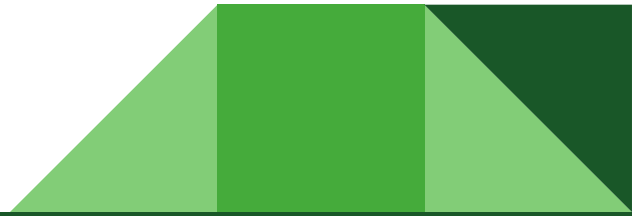
Data is **information** (numbers, words, measurements, observations, etc) that is in a **computer-readable form**.

What is "Data"?

Data is **information** (numbers, words, measurements, observations, etc) that is in a **computer-readable form**.

Human-Readable Information:

- Books
- Notebooks
- Physical Photographs
- Voice



What is "Data"?

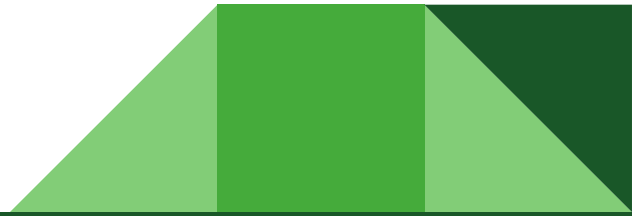
Data is **information** (numbers, words, measurements, observations, etc) that is in a **computer-readable form**.

Human-Readable Information:

- Books
- Notebooks
- Physical Photographs
- Voice

Computer-Readable Information:

- Literally just Binary



The Challenge

The big challenge in working with data is converting from human-readable information to computer-readable information, and vice versa.

During the Digital Information unit, we learned about how we can represent all kinds of things using Binary, though!

The Challenge

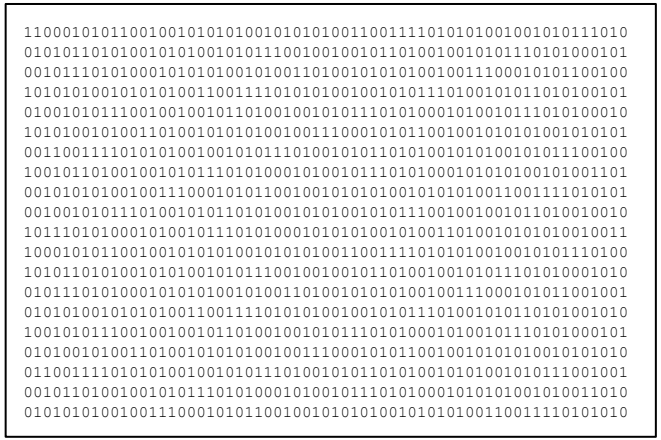
The big challenge in working with data is converting from human-readable information to computer-readable information, and vice versa.

During the Digital Information unit, we learned about how we can represent all kinds of things using Binary, though!

The Challenge

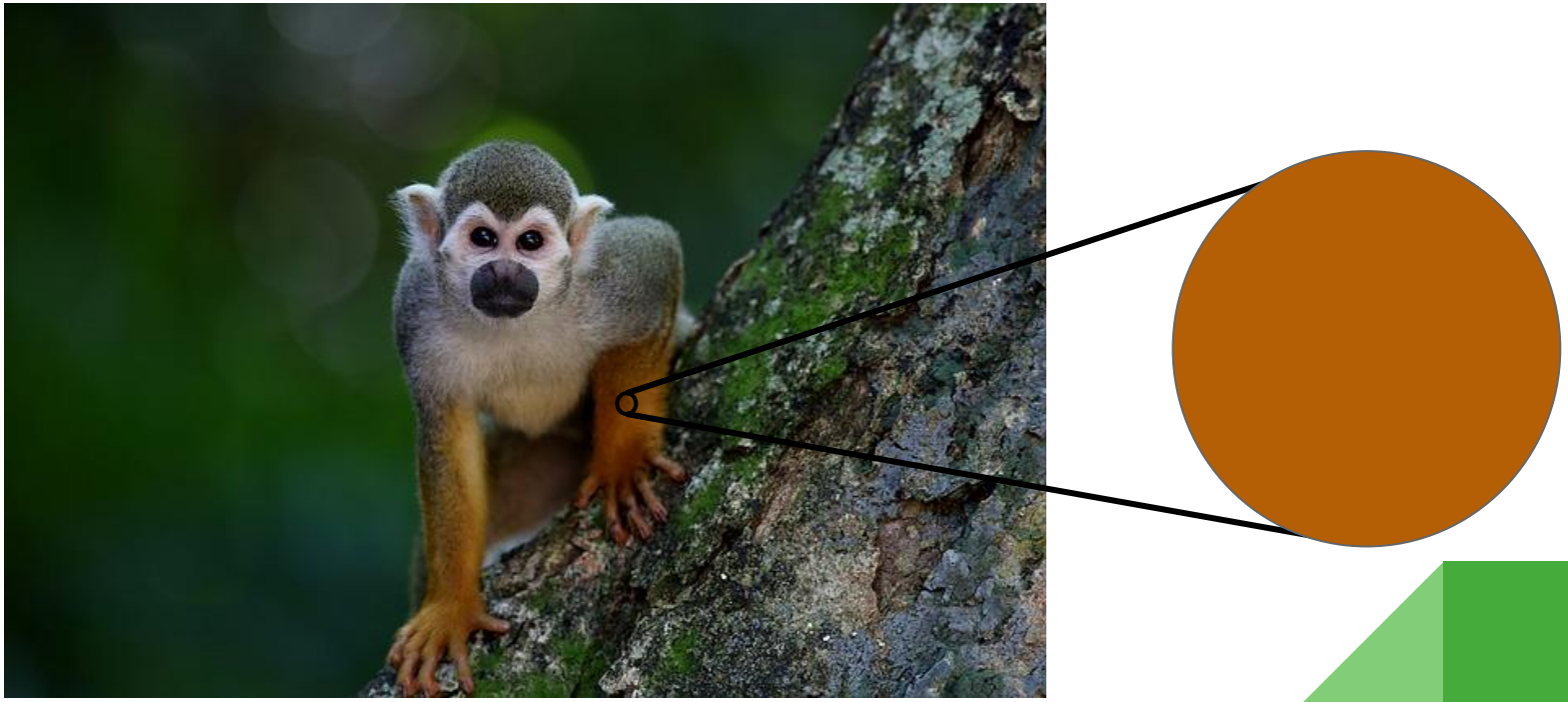
The big challenge in working with data is converting from human-readable information to computer-readable information, and vice versa.

During the Digital Information unit, we learned about how we can represent all kinds of things using Binary, though!

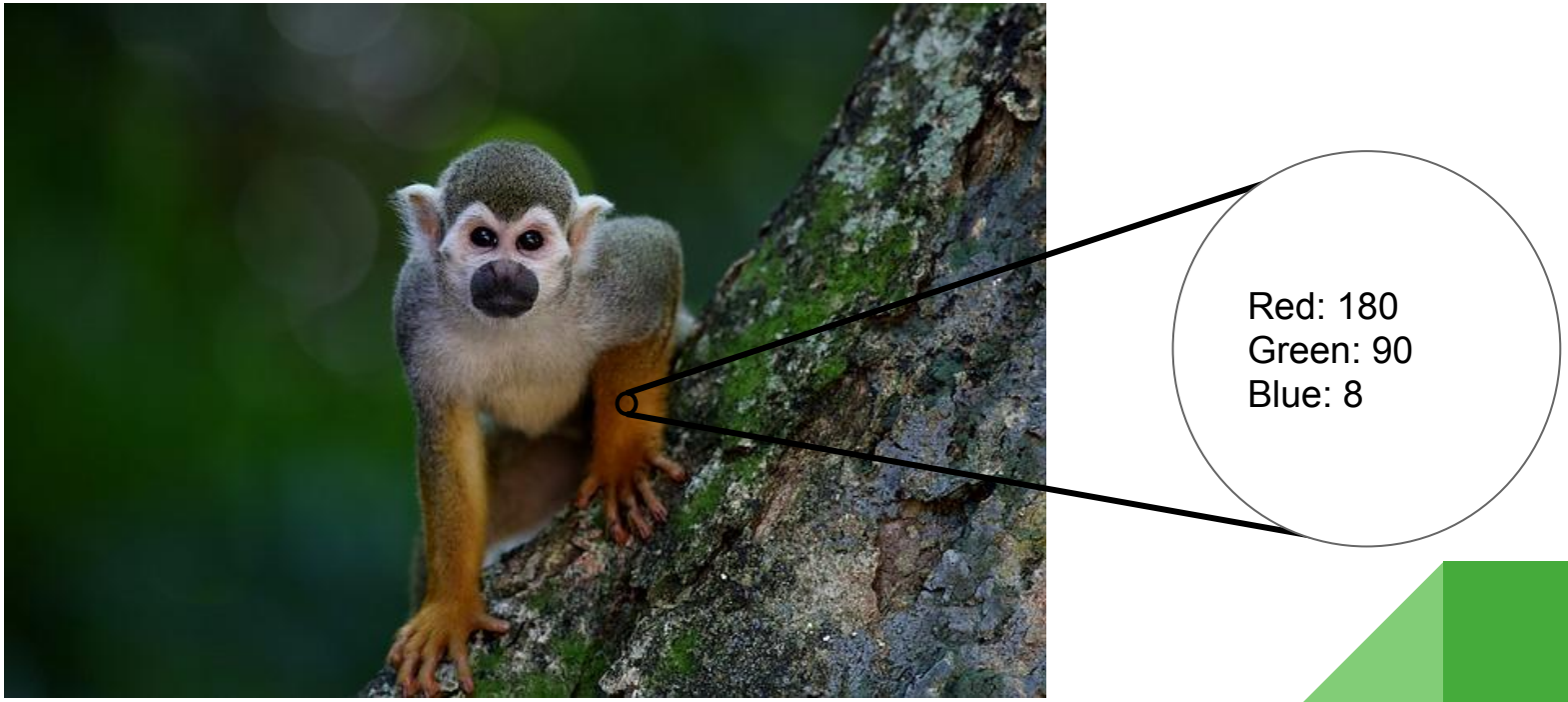


Making Information Digital

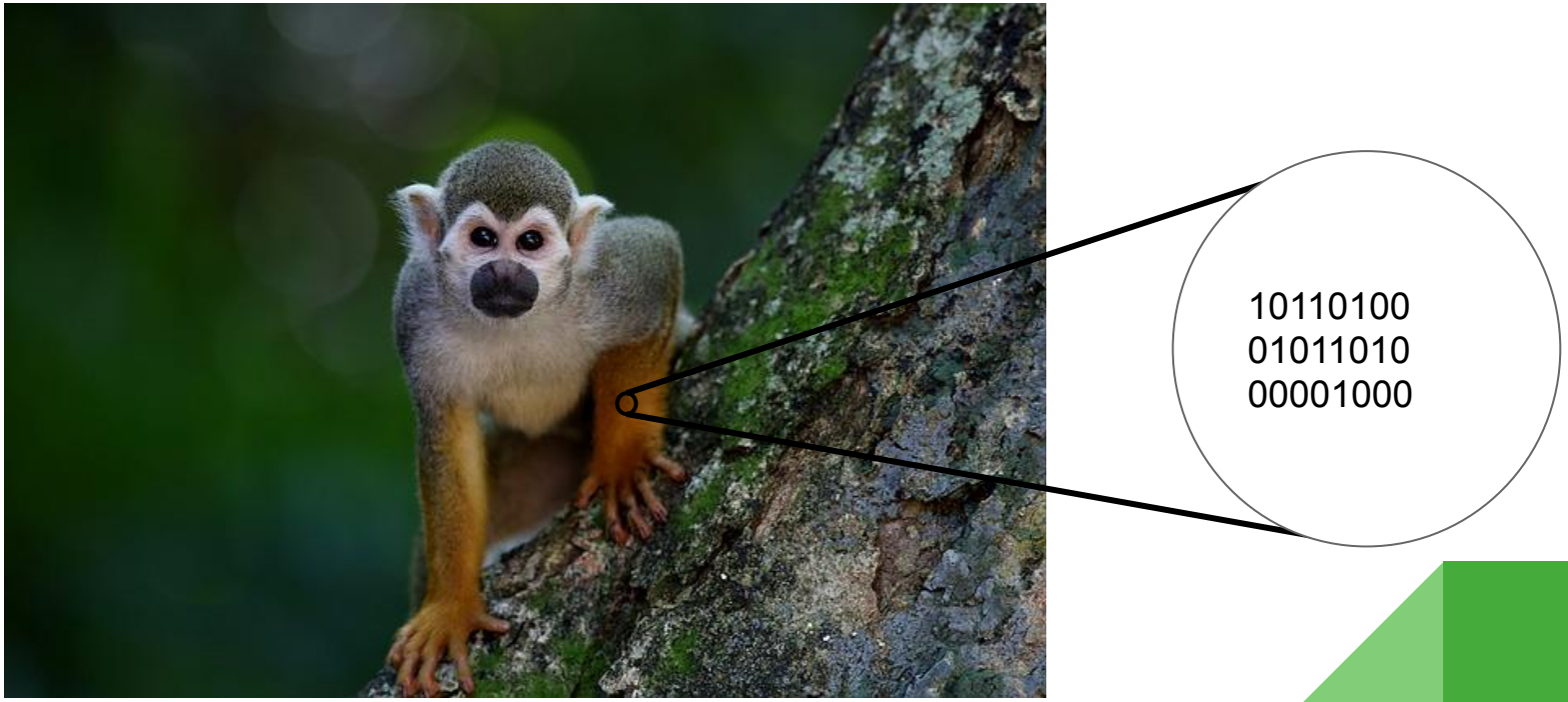
Making Information Digital



Making Information Digital



Making Information Digital



Making Information Digital

We can represent an entire image using only 0's and 1's using this process!

Making Information Digital

```
1100010101100100101010100101010100110011110101010010010101110100101011010100
1010100101011100100100101101001001010111010100010100101110101000101010100101
0011010010101010010011100010101100100101010100101010100110011110101010010010
1011101001010110101001010100101011100100100101101001001010111010100010100101
1101010001010101001010011010010101010010011100010101100100101010100101010100
1100111101010100100101011101001010110101001010100101011100100100101101001001
0101110101000101001011101010001010101001010011010010101010010011100010101100
1001010101001010101001100111101010100100101011101001010110101001010100101011
1001001001011010010010101110101000101001011101010001010101001010011010010101
0100100111000101011001001010101001010101001100111101010100100101011101001010
1101010010101001010111001001001011010010010101110101000101001011101010001010
1010010100110100101010100100111000101011001001010101001010101001100111101010
1001001010111010010101101010010101001010111001001001011010010010101110101000
1010010111010100010101010010100110100101010100100111000101011001001010101001
0101010011001111010101001001010111010010101101010010101001010111001001001011
0100100101011101010001010010111010100010101010010100110100101010100100111000
1010110010010101010010101010011001111010101001001010111010010101101010010101
0010101110010010010110100100101011101010001010010111010100010101010010100110
1001010101001001110001010110010010101010010101010011001111010101001001010111
0100101011010100101010010101110010010010110100100101011101010001010010111010
1000101010100101001101001010101001001110001010110010010101010010101010011001
1110101010010010101110100101011010100101010010101110010010010110100100101011
1010100010100101110101000101010100101001101001010101001001110001010110010010
1010100101010100110011110101010010010101110100101011010100101010010101110010
0100101101001001010111010100010100101110101000101010100101001101001010101001
0011100010101100100101010100101010100110011110101010010010101110100101011010
10010101001010111001001001011010010010101110101000010
```

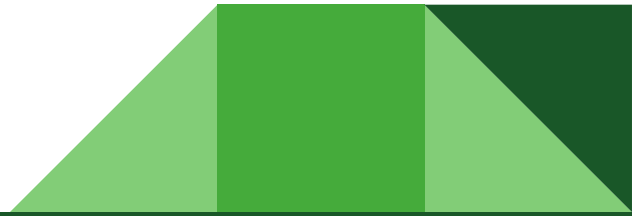
We can represent an entire image using only 0's and 1's using this process!

Making Information Digital

As our world grows ever more digital, we are constantly finding ways to turn human readable, unstructured data into computer readable, structured data.

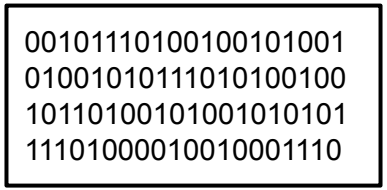
Song files, video files, image files, text files, spreadsheets with measurements and statistics are all examples of mechanisms we've found to store human-readable information in a digital format!

Once data is structured in a uniform way, we can write programs to store, process, manipulate, and visualize the data!



Making Information Digital

Just about every single program is going to interact with data in some way, shape, or form.

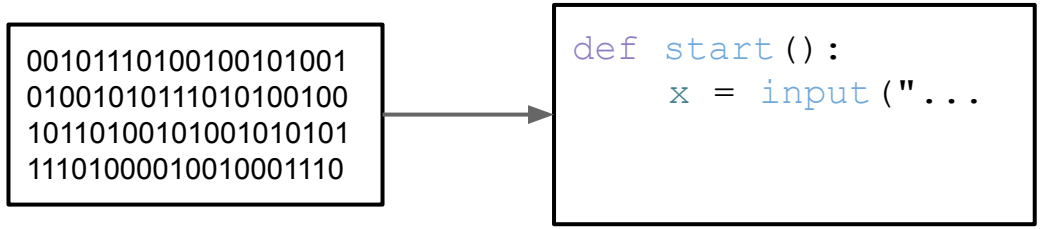


Input data

- Spreadsheets
- Images
- Mouse movements
- Keyboard inputs
- HTTP Request
- etc

Making Information Digital

Just about every single program is going to interact with data in some way, shape, or form.



Input data

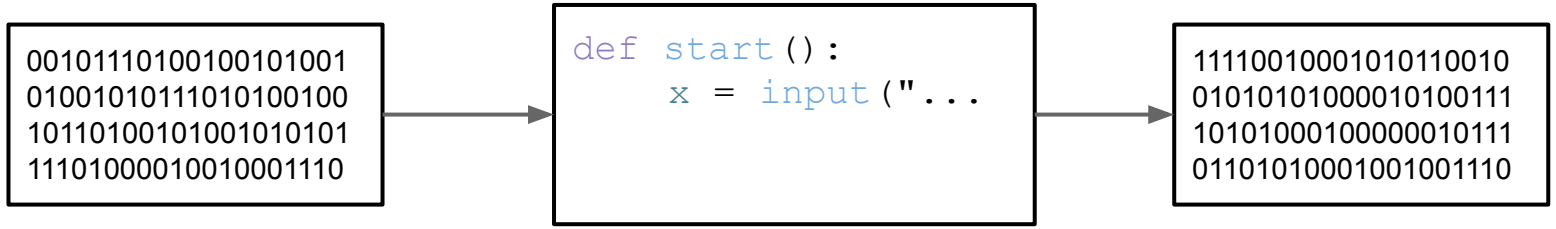
- Spreadsheets
- Images
- Mouse movements
- Keyboard inputs
- HTTP Request
- etc

Program

- Takes in input data
- Analyzes and manipulates data
- Produces output data

Making Information Digital

Just about every single program is going to interact with data in some way, shape, or form.



Input data

- Spreadsheets
- Images
- Mouse movements
- Keyboard inputs
- HTTP Request
- etc

Program

- Takes in input data
- Analyzes and manipulates data
- Produces output data

Output data

- Spreadsheets
- Images
- Sounds
- HTTP Response
- etc

Processing Digital Data

Once we have our data in a digital form, we can manipulate it in any number of different ways using a program!

Programs can manipulate the numbers to modify the image!

Processing Digital Data

Once we have our data in a digital form, we can manipulate it in any number of different ways using a program!

Programs can examine the numbers to locate edges, shadows, and even faces with enough training!

Why is data relevant?

Data (and information in general) are both used in the creation of **knowledge**.

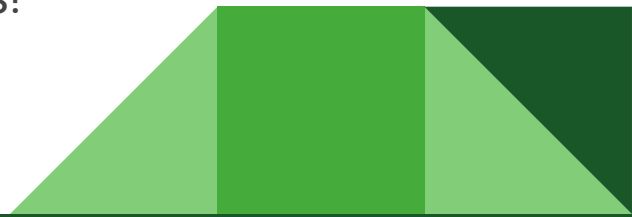
Raw information doesn't tell us anything by itself - we need to do some processing, some evaluation, or some computation to understand what that data actually represents.

We can run computations on the data we have to recognize patterns, gain new insights and knowledge, predict future events based on past ones, realize problems, and find solutions to those problems!

Modern Data Wealth

We live in a time of ever-growing data sets. Information about all kinds of things from all over the world is collected, stored, and processed! Things like financial transactions, all kinds of personal and medical data, information about web browsing history, and natural measurements about the world around us are all being tracked constantly!

These huge data sets provide us with opportunities to identify trends, make predictions about the future, make connections that might not be readily apparent, and recognize various problems with the world around us!

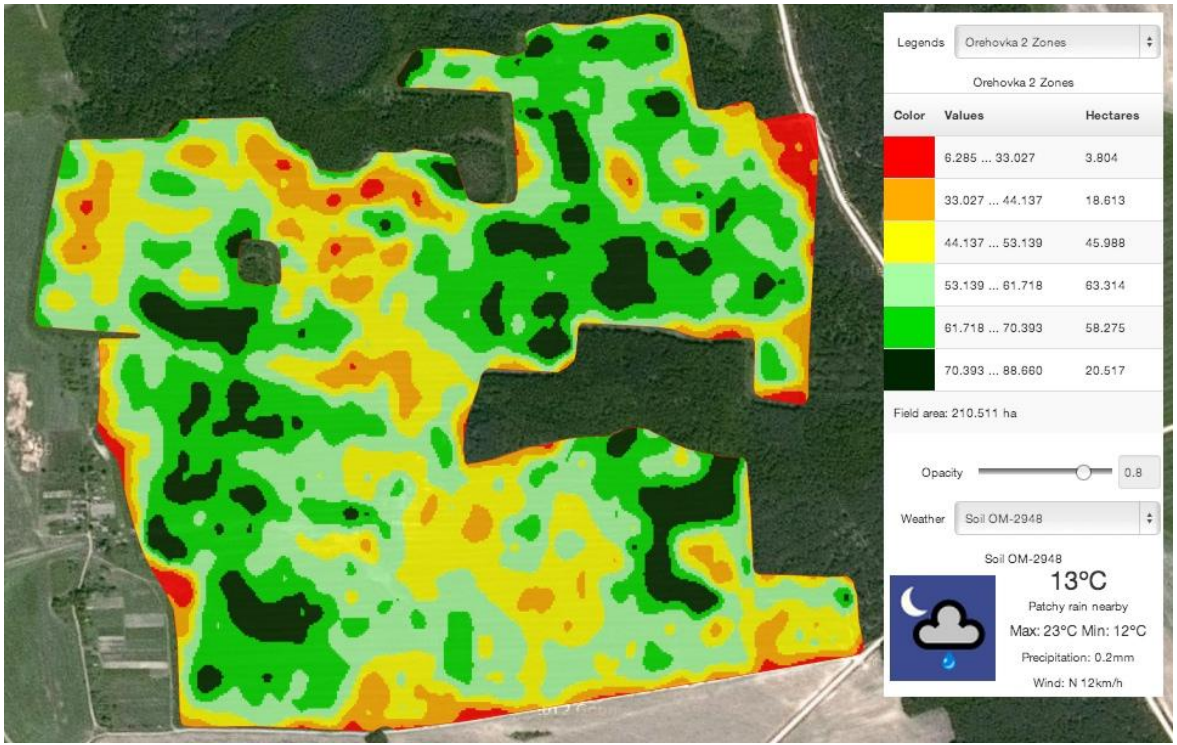


Computing Tools

Computing tools are essential when working with large data sets in order to extract useful information and gain knowledge, because computers are capable of processing information **SO** much faster than humans.

Computing tools provide ways to search, filter, store, and visualize data!

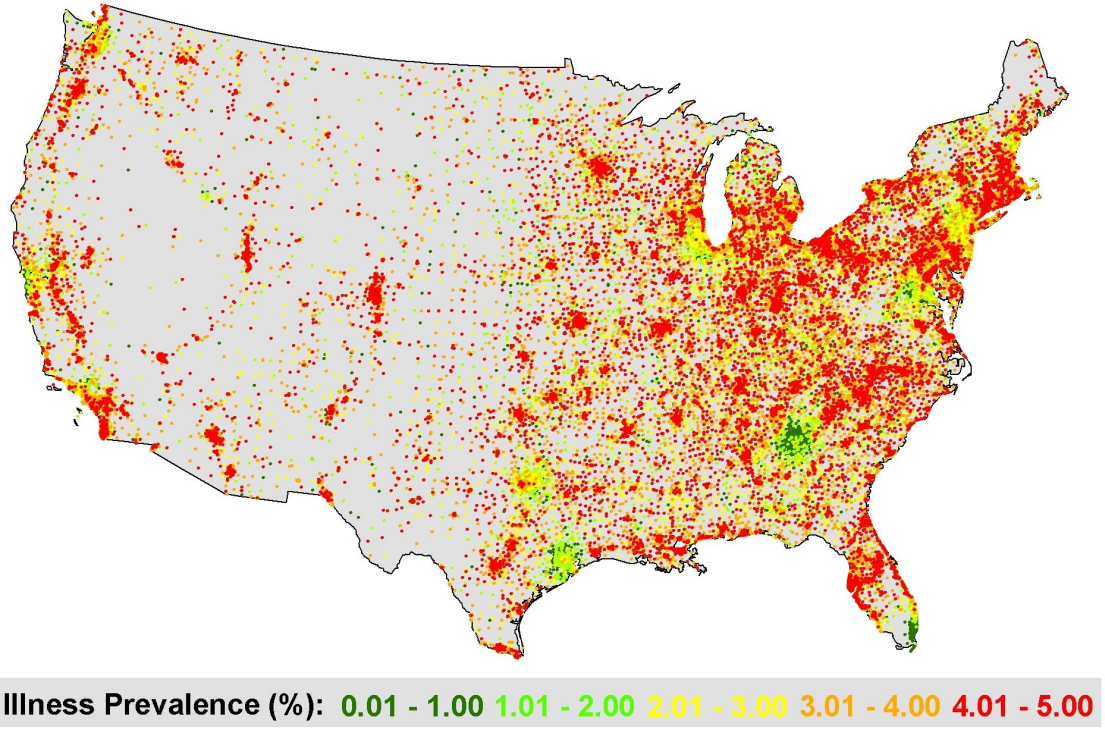
Computing Tools - Agriculture



A farmer could use a tool that uses sensors to capture information about the state of their land, then visualizes that information for them to make decisions on how to plant their crops.

source: https://zoner.bayer.com/static/H4zStdSxLEe27Xs2lDW1VZTEQCE4azjUd4T3RTmvjHm.jpg

Computing Tools - Health and Medicine



We can use computing tools to track how a disease spreads, then make predictions on how it will grow and potential ways to prevent that spread!

source: https://www.fredhutch.org/content/dam/public/diseases/disease-reskin/influenza-high-res.png

Concerns

There are definitely some concerns that are raised as a result of all of the data swirling about in the world around us!

- Storing and processing massive amounts of data efficiently is challenging
- Keeping data private is challenging
- Not everyone knows how much of their personal data is being tracked

Concerns

There are definitely some concerns that are raised as a result of all of the data swirling about in the world around us!

- Storing and processing massive amounts of data efficiently is challenging
  - Databases and Database Languages like SQL
- Keeping data private is challenging
- Not everyone knows how much of their personal data is being tracked

Concerns

There are definitely some concerns that are raised as a result of all of the data swirling about in the world around us!

- Storing and processing massive amounts of data efficiently is challenging
  - Databases and Database Languages like SQL
- Keeping data private is challenging
  - Encryption and anonymized data
- Not everyone knows how much of their personal data is being tracked

Concerns

There are definitely some concerns that are raised as a result of all of the data swirling about in the world around us!

- Storing and processing massive amounts of data efficiently is challenging
  - Databases and Database Languages like SQL
- Keeping data private is challenging
  - Encryption and anonymized data
- Not everyone knows how much of their personal data is being tracked
  - Actually read the privacy policy for things you're signing up for!