Manipulating Large Data Where Can Big Data Be Used?

- Students will explain how analyzing Big Data is different from the way ordinary data is analyzed.

- Students will describe how computers can make predictions and answer questions through the use of Big Data, storage of data, and processing data.

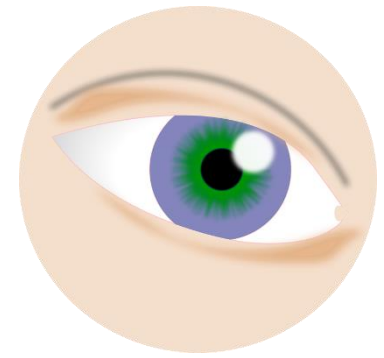- Students will synthesize the relationship(s) between causation and correlation.

Think about your daily and weekly activities. What types of data are being stored about you?
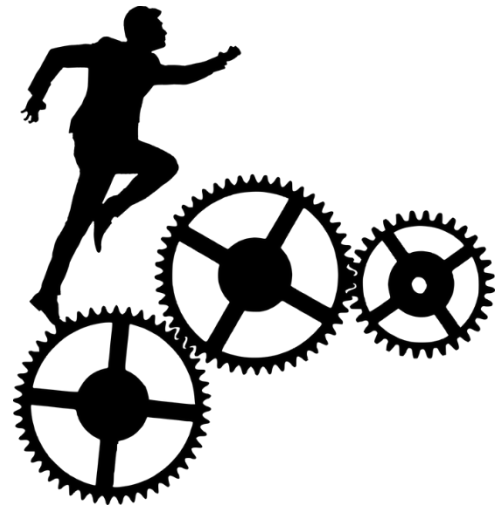
Hans Rolling

- A very cool data visualization

Another Tool for Visualization

- gapminder

Processing Big Data

1. Identify appropriate data source and form questions
2. Extract data source into format supported by underlying tools
3. Normalize data (remove redundancies, irrelevant details)
4. Import data into tool
5. Perform analysis
6. Visualize results

Where do we find data sets?

- http://catalog.data.gov/dataset
- http://www.census.gov/data.html
- https://github.com/awesomedata/awesome-public-datasets
- https://www.kaggle.com/datasets
- Search for data sets

Homework

- Google classroom

Name: _____ Class: _____ Date: _____

Big Data Sets Worksheet

Approved Topic: _____

Directions: Select a topic that includes the use of big data. After your teacher approves of your topic, search for reputable big data sets that are related to your topic. Record the information about each big data set below.

Big Data Set Name	How is the data collected?	How is the data usually analyzed?	What problems are addressed or solved by analyzing the data?	Citation

Unit 5: Data Manipulation

5-1-B