

# Big Data in the Classroom

Anne Byford  
Gaston Day School

# Why Big Data?

- Producing data is EASY!
  - InternetLiveStats (per second):
    - 808 Instagram, 7796 Tweets, 71,208 YouTubes watched
- Knowing what to do with data is HARD
- According to Forbes.com:
  - Wide range of fields from health care to retail
  - Median Salary (with experience): 124,000
  - % Demand growth: up to 300% in one year

# Why Big Data?

- Datasets available (and free) for many subjects, not just science
- Ideal for Project-based learning
  - Authentic data, authentic tools
- Adaptable for differentiation
  - Same data set can be used with different questions
  - Curated data provided to some groups

# Skills Used

- Background research
- Question development
- Data organization
  - What data do I need to answer my question?
  - What should my table look like?
  - What information should I compare?
- Data analysis
  - What does this data mean?
  - What statistics should I do and what do they mean?

# Skills Used

- Data visualization
  - Graphs, Tables, other Figures
- Communication
  - Oral Presentation
  - Written Presentation
  - Website Development
  - Poster Session

# Cautions:

- Using “real” data
  - Can be complex
  - Is often messy
  - May not give the expected answer
  - Takes more time than “canned data”
- Students will need help with specific skills
  - Asking good questions
  - Using spreadsheets, both set-up and graphing
  - Relating raw data to their questions
- Students will experience frustration
  - Especially those used to getting right answers immediately

# Zooniverse.org

## Projects

Active

Paused

Finished



ALL DISCIPLINES



ARTS



BIOLOGY



CLIMATE



HISTORY



LANGUAGE



LITERATURE



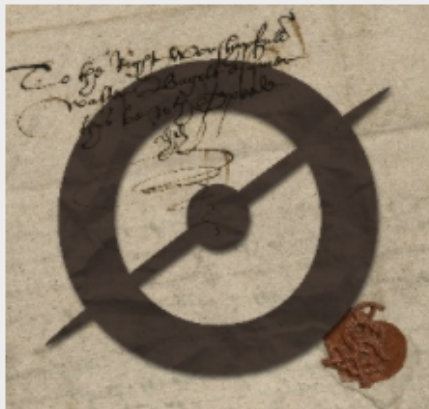
MEDICINE



Most Recently Launched x v

Showing 1-20 of 76 projects found.

Name: x v



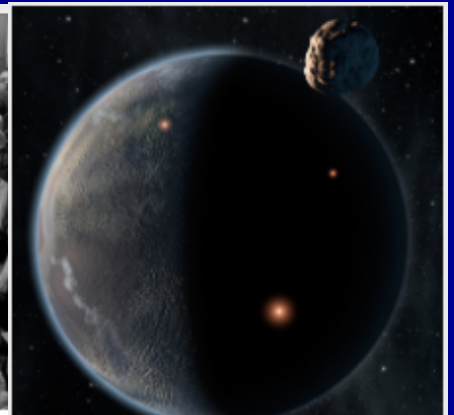
SHAKESPEARE'S WORLD



DECODING THE CIVIL WAR



OPERATION WAR DIARY



PLANET HUNTERS

# HHMI Wildcam Gorongosa

- Lesson plans already developed
- Segments include
  - Making observations
  - Scientific inquiry and data analysis
  - Biodiversity
  - Ecological Pyramids
- Entire sets of raw camera and animal identification data available for download



# Other Sources of Data

- NOAA

- <https://www.ncdc.noaa.gov/>

- CDC and WHO

- <https://data.cdc.gov/>

- <http://www.who.int/gho/database/en/>

- NASA data

- <https://data.nasa.gov/>

# DNA to Proteins Project

- Based on:
  - Salt Lake City Community College's Halobacterium project
  - <http://www.slcc.edu/biotech/halo-project/index.aspx>
- Students
  - analyze 50,000bp DNA to identify ORFs
  - Build a consensus ORF map
  - Analyze an ORF to determine if it is a gene and its potential function
- Files in shared Google folder

# Sea Scallop Project

- Developed during Teacher-At-Sea (NOAA)
- Raw data from yearly Sea Scallop Survey
  - Data from 1979 – 2016
  - Fish Species catch number and mass
  - Scallop catch number, size, and mass
  - Location of all dredge sites (lat/long)
- Full lesson plan in shared folder

# Link to Project Folder

- <https://drive.google.com/drive/folders/0B9tWi0CfAPdZOGVtcmdHckpCS3M?usp=sharing>