Unit 9 Lesson 1

Learning from Data

Resources

Unit 9 Lesson 2

Exploring One Column

Resources

Name(s)_____ Period _____ Date _____

Activity Guide - Exploring One Column

Bar Charts

Do This

- Make a bar chart for each column in the "Dogs" dataset and categorize it as either "useful" or "not useful" by placing that column name in the table below.

Useful	Not Useful

Chart

Paste one of the charts that you believe is "useful" in the space below.

Questions

1. What is one piece of information someone could learn from the chart you pasted?

2. Do you notice any patterns in which charts are or are not useful?

Histograms

Do This:

- Make a histogram for the "Max Weight" column of the "Dogs" data set and choose the "Bucket Size" that you think is most helpful for interpreting the data.

Chart

Questions

3. What bucket size did you choose? Why do you think this is the most helpful bucket size for this chart?

4. What is the most common range of maximum weights for dog breeds?

5. What is the least common range of maximum weights for dog breeds?

Do This:

- Make a histogram for one other column in the Dogs dataset and choose the "Bucket Size" that you think is most helpful for interpreting the data.

Chart

Questions

1. What bucket size did you choose and why?

2. What is one question the chart you created could help answer?

Unit 9 Lesson 3

Filtering and Cleaning Data

Resources

Name(s)_____ Period _____ Date _____

Activity Guide - Filtering Data

Data Set: Female State Legislators

Do This:

- Filter the Female State Legislators dataset for Year by the year you were born.
- Create a histogram for Percentage of Females in Legislature. Choose a bucket size of 5.
- Copy/paste the chart below.

Chart 1

- Filter the Female State Legislators dataset for Year by the year 2019.
- Copy/paste the chart below.
- Answer the questions.

Chart 2

Questions

1. In the year you were born, how many states had between 25 and 30 percent of their legislatures made up of female legislators?

2. In 2019, how many states had between 25 and 30 percent of their legislatures made up of female legislators?

3. What does this data show?

4. Why might that be the case?

Data Set: US Women Running for Elected Office in 2020

Do This:

- Filter the US Women Running for Elected Office in 2020 dataset to answer the following questions. The chart type is up to you. Paste the chart below each answer.

Question: Which states had at least one woman run for Governor?

Answer:

Chart:

Question: Which states had more than 10 women win their primary race?

Answer:

Chart:

Unit 9 Lesson 4

Exploring Two Columns

Resources

Name(s)_____ Period _____ Date _____

Activity Guide - Exploring Two Columns

Cross Tab

1. **Crosstab - Use the “Words” dataset**
   Use: Length & Part of Speech
   a. Chart - Copy and paste the chart you created in the space below

   b. Which part of speech shows up most often in this data set?

   c. Which part of speech seems to be longest, on average?

   d. Which part of speech seems to be shortest, on average?

   e. About how long is the "typical" noun? How can you tell?

2. **Crosstab - Use the “Favorite Classes” dataset**
   Use: Grade & Favorite Class
   a. Chart - Copy and paste the chart you created in the space below

   b. Which class do Seniors like the most?

   c. Which grade likes History the most?

   d. What is one other interesting pattern you can see in this diagram?

Scatter

1. **Scatter - Use the “US States” dataset**
   Use: Median Household Income & Percent Adult with Bachelor's Degree
   a. Chart - Copy and paste the chart you created in the space below

   e. What is the range of incomes on this chart?

   f. What is the range of percentages of adults with a bachelor's degree?

   g. Do you see a relationship between income and percent of adults with a bachelors degree? What does the relationship look like?

3. **Scatter - Use the “US States” dataset**
   Use: You pick the columns to compare
   In the States table find a combination of columns in a Scatter Plot that you think show an interesting pattern

   a. Chart - Copy and paste the chart you created in the space below

   b. What is the pattern that you observe in this data?

   c. Describe what this pattern might mean or indicate.

   d. Usually we can't be sure whether the pattern we observe in a scatter chart (a correlation) means that one of the columns is causing the other. Give one way you might do more research to try to investigate the cause of the pattern you found.

Unit 9 Lesson 5

Big, Open, and Crowdsourced Data

Resources

Name(s)_____ Period _____ Date _____

Activity Guide - Big, Open, and Crowdsourced Data

Background

Today you will research one of three topics at the intersection of data analysis, computing, and society. These topics all use the data analysis process in new and interesting ways to address problems.

Topic	Links	Key Vocabulary to Look For
Topic 1 Big Data	• AI and Parallel Systems https://www.youtube.com/watch?v=1XGo8K1boH4 • Big Data and Medicine Code.org https://www.youtube.com/watch?v=bMrDHtGHFR4	Scalability Parallel systems
Topic 2 Crowdsourced Data	• How Pokemon Inspired A Citizen Science Project... https://www.npr.org/2018/04/20/597972310 • What is Citizen Science? https://youtu.be/81hhecl0p5k	Citizen science Crowdsource
Topic 3 Open Data	• What is Open Data? https://www.youtube.com/watch?v=qSD9ob8rGcs • The Case for Open Data https://www.youtube.com/watch?v=iOrPK7p2AwI	Open data Open access

What topic did you choose? _____

As you watch the videos keep notes on examples of how they complete each step of the data analysis process.

Collect or Choose Data	Clean and/or Filter Data	Visualize and Find Patterns	New Information

What is the core idea of your topic? What is it about?

_____

_____

Give two examples of the problems / questions your topic is being used to solve / answer.

_____

_____

Unit 9 Lesson 6

Machine Learning and Bias

Resources

Unit 9 Lesson 7

Project - Tell a Data Story Part 1

Resources

Name(s)_____ Period _____ Date _____

Project Guide - Tell a Data Story

Background

Computing tools allow people from all different fields to learn from data. By transforming and visualizing data we can find patterns that help us learn about our world.

For this project you will

- Select a dataset from the Data Library
- Use the Data Visualizer to make a visualization that helps us learn something interesting about the data
- Write a response that explains:
  - what your visualizations show
  - new insights or decisions that can be made based on this visualization
  - possible problems with your analysis including potential bias in the data or your interpretation

Tell a Data Story

Data Analysis Process	Written Response
 Collect or Choose Data	Select a dataset from the Data Library. Read the metadata to understand what information is available in the table **Dataset Name:** **Short Description:**
 Clean and/or Filter	**Did you filter or clean the data? Why or why not?**
 Visualize and Find Patterns	**Create a visualization. Copy/paste the chart below.**

New Information

What information does your visualization show:

What new insights or decisions can be made based on this visualization:

What are the possible problems with your analysis including potential bias in the dataset or your interpretation:

Rubric

Category	Extensive Evidence	Convincing Evidence	Limited Evidence	No Evidence
Collect or Choose Data	Dataset is correctly identified and description is complete.	Dataset is correctly identified and description is mostly complete.	Dataset is correctly identified and description is somewhat complete.	Dataset is not identified or description is missing.
Clean/Filter Data	Explanation for cleaning and/or filtering is complete and reasonable.	Explanation for cleaning and/or filtering is complete and mostly reasonable.	Explanation for cleaning and/or filtering is somewhat complete or somewhat reasonable.	Explanation for cleaning and/or filtering is incomplete.
Visualize and Find Patterns	Visualization is readable and includes a title, and citation.	Visualization is mostly readable and includes a title, and citation.	Visualization is somewhat readable and/or is missing a title or citation. .	Visualization is unreadable or missing.
New Information: Interpreting the visualization	The visualization is thoroughly explained.	The visualization is mostly explained.	The visualization is somewhat explained.	The visualization is not explained.
New Information: Insights or Decisions	Insights or decisions are reasonable and effectively linked to information displayed in the visualization	Insights or decisions are mostly reasonable and effectively linked to information displayed in the visualization	Insights or decisions are mostly reasonable and somewhat effectively linked to information displayed in the visualization	Insights or decisions are missing.
New Information: Bias	Possible problems with analysis or potential bias are reasonable and thoughtfully addressed.	Possible problems with analysis or potential bias are mostly reasonable and thoughtfully addressed.	Possible problems with analysis or potential bias are somewhat reasonable and addressed.	Possible problems with analysis or potential bias are not addressed.

Unit 9 Lesson 8

Project - Tell a Data Story Part 2

Resources

Name(s)_____ Period _____ Date _____

Project Guide - Tell a Data Story

Background

Computing tools allow people from all different fields to learn from data. By transforming and visualizing data we can find patterns that help us learn about our world.

For this project you will

- Select a dataset from the Data Library
- Use the Data Visualizer to make a visualization that helps us learn something interesting about the data
- Write a response that explains:
  - what your visualizations show
  - new insights or decisions that can be made based on this visualization
  - possible problems with your analysis including potential bias in the data or your interpretation

Tell a Data Story

Data Analysis Process	Written Response
 Collect or Choose Data	Select a dataset from the Data Library. Read the metadata to understand what information is available in the table **Dataset Name:** **Short Description:**
 Clean and/or Filter	**Did you filter or clean the data? Why or why not?**
 Visualize and Find Patterns	**Create a visualization. Copy/paste the chart below.**

New Information

What information does your visualization show:

What new insights or decisions can be made based on this visualization:

What are the possible problems with your analysis including potential bias in the dataset or your interpretation:

Rubric

Category	Extensive Evidence	Convincing Evidence	Limited Evidence	No Evidence
Collect or Choose Data	Dataset is correctly identified and description is complete.	Dataset is correctly identified and description is mostly complete.	Dataset is correctly identified and description is somewhat complete.	Dataset is not identified or description is missing.
Clean/Filter Data	Explanation for cleaning and/or filtering is complete and reasonable.	Explanation for cleaning and/or filtering is complete and mostly reasonable.	Explanation for cleaning and/or filtering is somewhat complete or somewhat reasonable.	Explanation for cleaning and/or filtering is incomplete.
Visualize and Find Patterns	Visualization is readable and includes a title, and citation.	Visualization is mostly readable and includes a title, and citation.	Visualization is somewhat readable and/or is missing a title or citation. .	Visualization is unreadable or missing.
New Information: Interpreting the visualization	The visualization is thoroughly explained.	The visualization is mostly explained.	The visualization is somewhat explained.	The visualization is not explained.
New Information: Insights or Decisions	Insights or decisions are reasonable and effectively linked to information displayed in the visualization	Insights or decisions are mostly reasonable and effectively linked to information displayed in the visualization	Insights or decisions are mostly reasonable and somewhat effectively linked to information displayed in the visualization	Insights or decisions are missing.
New Information: Bias	Possible problems with analysis or potential bias are reasonable and thoughtfully addressed.	Possible problems with analysis or potential bias are mostly reasonable and thoughtfully addressed.	Possible problems with analysis or potential bias are somewhat reasonable and addressed.	Possible problems with analysis or potential bias are not addressed.

Unit 9 Lesson 9

Assessment Day

Resources