**Instructional Days:** 7-10

**Topic Description** : This lesson will introduce Data Analyzer as a data analysis tool. In class, you will explore and use its’ basic features in order to better understand the data that you will be taking a look at as well as create different types of graphs.

**Objectives :**

* Understanding the uses of latitude and longitude versus an address
* Understand files of data
* Create subsets from original data
* Create different types of graphs based on data
* Understand Sonification of data

**Outline of the lesson**

* Journal Entry (5 min)
* Describe Sonification (30 min)
* Explore data and Data Analyzer (30 minutes)
* Sonification Activity (45 minutes)
* Local Cuisine Activity (45 minutes)
* Participate in discussion of scatter plots vs other graph types (20 minutes)
* Good Nights Rest Activity (45 minutes)

**Student Activities**

* Complete journal entry
* Participate in discussion of location, Billboards 100 Data, and Data Analyzer
* Complete Sonification Activity
* Complete Local Cuisine Activity
* Participate in discussion of scatter plots
* Complete Good Nights Rest Activity

**Teaching/Learning Strategies**

* Journal Entry: What are other ways to interpret the data you have been collecting? Are there ways to organize it to make it easier to understand.
* Open up Data Analyzer in the web browser of your choice.
* Learning to describe location: Which way is best? Why use latitude and longitude sometimes?
* Open Pointz in the web browser to demonstrate finding locations of local restaurants.
* Ask questions about these locations: Have you been to these places? How would you tell someone to get to them? When would you use latitude and longitude?
* Go to **getlatlon**.yohman.com/ (may need more accessible option) and explain how there are multiple tools available online that are able to translate an address into a set of coordinates. The Pointz application will also return latitude and longitude in its csv form.
  + Compare some of the coordinates from Pointz to ones from GetLatLon. Do the coordinates given match the coordinates in already in the data set? If they don’t, why might that be?
  + Have students give familiar locations to them to find coordinates for
  + Discuss how longitude and latitude can be used to create spatial objects such as points on a map of where these locations are.
* Exploring data in the Data Analyzer
  + Load sushi1.csv into the Data Analyzer
  + Point out the following features of the application
    - You can either load your own data set or create a new one
    - The table within the app is completely editable. Explain how you can change values and which parts of the table affect the rest of the outputs for the application. Also explain how you would add or remove a row or column from the table.
    - Explain how the Graph Data section works. You can change line colors or toggle what data is displayed on the graph. This is also where the max, min, and average of each row of data can be found.
    - Show how the Audio Controls under the graph work. You can select a line to play and at what speed.
  + Point out the following parts of the data set
    - Header, number of rows, and categories
    - The first line is the header and works as the labels for the graphs in the application. In the case of the sushi example, the variables are the price level and rating of the different restaurants.
    - Each row in the table represents a different sushi location. Each column represents the various data of those restaurants in reference to the variables.
    - Briefly discuss frequency and how one could find the mode of this data for each variable.
    - Briefly demonstrate how one could make a table from scratch.
* Sonification Activity
  + Students complete on own, answer questions.
* Local Cuisine Activity
  + Students complete on own, answer questions
* Scatter Plots
  + Discuss what you would use a scatter plot for
  + What relationships could they show in their final project with a scatter plot?
* Good Nights Sleep Activity
  + Students complete on own, answer questions

**Resources**

* Data Analyzer web app
* Pointz web app
* Activities 1, 2, and 3
* topArtists.csv
* peakArtists.csv
* sleepStudy.csv

**Analyzing Data**

\*\* **Data Description**

For part of this lesson, data will be provided for the students that was gathered from the Music 100 Billboards. It provides information about various artists and their standings in the charts. For the other part, the data that will be collected by students from the web app Pointz, which returns data on 20 places based on search results. This data includes the name of these places, their latitude and longitude, their price level, and their rating as an establishment. Students should use their hometowns as a term of search as this will help with the next section of the activity.

\*\* **Describing location**

When someone asks you to describe your location, you will usually respond by simply stating that you are in class or at school. Those who know the area will understand, but if for example, a relative comes from out of town, they may have no clue where your school is. At this point, you become more specific, giving them the address for the school, or a nearby intersection they could easily find. These descriptions are excellent for looking up locations on a map for navigating to somewhere new. The roads and addresses create a network to help us navigate areas. However, this network can change with time as new roads are added and old ones are replaced or sometimes there are locations that simply do not have addresses. For this, it is necessary to have a more consistent way to specify positions, which is why there is a need to associate places with a fixed coordinate system for the earth.

One of the most popular coordinate systems for the planet involves specifying a point’s latitude and longitude, two numbers that represent angles from the center of the earth to a point on the surface. Latitudes are the angles that go from north to south, where the North Pole is assigned a value of 180 degrees, the equator is 0, and the South Pole is -180 degrees. Longitudes are the angles from east to west with 0 being the Prime Meridian, a line that runs from the North to the South Pole.

These two ways of describing a location (address versus lat-long) can be compared, as is showcased in the data you retrieve from the Pointz web app.

Using latitude and longitude, people can create spatial objects such as points or lines on a map, where longitude is considered the x-direction and latitude is the y-direction to help plot the coordinates given.

**Exploring Types of Data**

Open up the Data Analyzer web app. Initially you should have two options to choose from: Load CSV or Create New Table. For now, select Load CSV and open topArtists.csv. Select the “Line Graph” option in the pop up screen. The file you open should resemble this:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Week 1 | Week 2 | Week 3 | Week 4 | Week 5 |
| Taylor Swift | 1 | 1 | 1 | 1 | 1 |
| Maroon 5 | 3 | 3 | 3 | 3 | 2 |
| Ed Sheeran | 5 | 6 | 5 | 5 | 4 |

The first line of the file is the header and shows the different weeks that data was recorded for. Each row refers to a different artist that, at the time of the data collection, had been on the Billboard Top 100 Artists chart for more than 20 weeks. Each piece of data is their standing on the chart at that week. This type of data is **continuous data.** This means that the values can change continuously and you can measure or count it. Other examples include weight, price, and profit. This data is usually best represented in a line graph.

Now go back to Load CSV and open up peakArtists.csv. Select the “Bar Graph” Option in the pop up screen. You should now have a file that looks like this:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Taylor Swift | The Weeknd | Maroon 5 | Ed Sheeran |
| Weeks on Chart | 46 | 35 | 50 | 50 |
| Peak Position | 1 | 2 | 1 | 2 |
| #1 hits | 4 | 0 | 3 | 0 |

With this file, the artists are now in the header row instead of a measurement of weeks. Now in the rows you should have different categories that all pertain to the musicians. You have the number of continuous weeks they were on the chart, the highest position they ever held, the total #1 hits they’ve had, and the total number of songs they’ve ever had on the chart. While you could show these rows as lines, it wouldn’t make that much sense as one doesn’t really lead to the next. These rows represent **categorical data** where you can make a distinction between the groups and list the categories. This is data best represent in a bar graph.

**Activity 1 - Sonification**

The main feature of the Data Analyzer is its’ ability to sonify the data it is given. This means that the application uses different tones to help convey the table information.

1. Open up the topArtists.csv again.
   1. Go to the Audio Controls and play through the different lines to hear how they are sonified. You can change the speed of the play through as well.
   2. Go to Graph Data and record the Max, Min, and Average of each artists. Who has the lowest average? Who has the highest average?
   3. What makes this kind of data continuous? Can you think of other examples of continuous data sets?
2. Open up peakArtists.csv
   1. Again, go through and play through the different rows. Does playing through these rows make as much sense as with the last data set? What may be another way that this data could be sonified?
   2. Go to Graph data and record Max, Min, and Average of each category. Are these maxes, mins, and averages as comparable to each other?
   3. What makes this data categorical? Can you think of other examples of categorical data sets?

**Activity 2 - Local Cuisine**

For this activity, you will be using the Data Analyzer again as well as the web application Pointz.

1. Open the Pointz app. In the first field, fill out a type of cuisine you would go out to eat, such as pizza or sushi. In the second field, fill in your hometown. Once you have results, use the download.csv button to get the results.
2. Open up your downloaded file. What are the variables in the first row? What is the variable used in the first column?
3. How many total restaurants did your search return?
4. What kind of data would you call your restaurant data?
5. Go back to Data Analyzer and load your data in. What happened? Why might be causing it to mess up?
6. Using the Data Analyzer, create a new data table to make a data set that works with it’s sonification tools. How many rows does your table have? How many rows?
7. For your new table, does it have all the same variables? Are there any variables that you left out? Why?
8. What kind of graph did you choose to represent your data? Why?
9. Is the data that you submitted continuous or categorical? What makes it one versus the other? Could you make another data table with the other type?
10. For the variables that you left out of your new data set, what would they be better used for?
11. How could you use the Data Analyzer to represent data that you have gathered from your phone?

**Scatter Plots**

There is one more chart type available to us in Data Analyzer. A scatter plot is similar to the line and bar graphs as it also uses a vertical and horizontal axis, but just like the line and bar graphs, there are differences in its’ usage. Scatter plots are used to show how often one variable is affected by another. Each row’s marker position depends on the values in the columns set in the x and y.

The relation between the variables is also known as their correlation. If all the markers come close to making a straight line, those variables have a high correlation. If they are all evenly distributed around the plot, the correlation will be low or even zero.

Some scatter plots can also include a third variable that corresponds to a color of size of the markers, but for now, don’t worry about that.

**Activity 3 - Good Nights Rest**

For this activity, you will be observing the correlation between the amount of sleep a student gets and their exam grade.

1. Go to Data Analyzer and load sleepStudy.csv
2. What are the variables here? Does the format of the data make sense for this graph type?
3. What are the maxes, mins, and averages for this study?
4. Before you play the graph, based on the data, do you think that there will be a high or low correlation? Why?
5. What do you think a high correlation would sound like? A low or zero correlation?
6. Go ahead and play the scatter plot’s line. What do you hear? What kind of correlation do you think the data has?
7. Can you think of other data relationships that could be represented as a scatter plot?