**Teacher’s Instructions**

**Part 1: Goal A:Learn how to interpret data and make a scatterplot**

Assign students to a group. Groups of 3-4 are best. Students will eventually be graphing data from six different countries, so 6 groups are ideal. Alternatively, each group can be asked to graph data from two countries and duplicated graphs are not a problem.

Lead students through the creation of a simple graph using the graph table template and the data in Table 1.

**Additional information that may be useful in explaining the items that should be included in a graph.**

**How to figure out which data set should go on the X-axis vs Y-axis:
 X-axis:** The horizontal (left to right) line which usually holds the independent variable (the variable that stands alone). *Common x-axis variables: age, time, distance, height*
 **Y-axis:** The vertical (up and down) line which usually holds the dependent variable (the variable being measured). *Common y-axis variables: items counted, resulting size/weight, rainfall*

 ***If you are trying to make predictions about data****, put the variable you’re trying to predict on the y-axis and the variable that you’re basing it on on the x-axis. For example, I want to predict the number of tomatoes (y-axis) of different heights of tomato plants (x-axis).* ***If you are trying to show a causal relationship between two observed variables****, put the cause on the x-axis and the result (or effect) on the y-axis.
 For example, I may want to show whether or not parental education level (the cause on the x-axis) relates to the child’s education level (the effect on the y-axis).*

**Explain how to decide the proper scale for each axis.**
 **Proper scale:** Often, each axis will begin at zero and continue until just beyond the last data point. The axis can begin at a number above zero if the smallest data point is a large number. Tick marks designate each unit of measure. *For example, if we are using the tomato example above, we don’t need to start the x-axis at zero if all of the plants are at least 3 feet tall. Likewise, we would want to have the last tick mark point just slightly higher than the last data point.*
**Explain how to write good axis labels, graph descriptions, and keys:
 Axis labels:** Labels should be descriptive and include units of measurement.
 **Graph description:** A good description should summarize the information (data points and axis labels) in the graph. *A complete sentence is appropriate.*
 **Key/Legend**: If multiple data sets are graphed, add a visual picture (such as different colors or shapes) explaining the data sets on the graph.

**Possible followup questions:**After graphing the data, do you have further insight into the data? What story does the data tell? Was this easier to see in graph form or table form? Did your axes both start at zero?

What was the value associated with the highest tick mark on the x-axis (y-axis)?

How would the graph have changed if you increase the y-axis to 200?
 **Part 1: Goal B: Understand Common Trend Curves and their use in interpreting data sets**

Review these common trend curves with students and discuss the answers to the follow up questions. The term ‘trend curve’ is used here to emphasize proper math terminology. Only the linear trend curve uses a line. A line is one of various types of curves.

Depending on the previous math classes of your students, you may choose to include the equation or not. This material is covered in a typical college algebra class.

|  |  |
| --- | --- |
| **Linear trend curve:****Equation**: $y=mx+b$ **Description**: Growth rate is constant and doesn’t change. Graph is a straight line.**Example**: Money earned from a job, without a raise, over a year.*Linear trends show direct correlations. A linear equation is an equation of the form f(x) = mx + b, where m is the slope of the line and b is the y-intercept of the line. For linear equations representing growth, m is the growth rate which is constant and x would be the independent variable representing time.* | **Linear Graph** |
| **Exponential trend curve:****Equation**: $y=Ae^{kx}$**Definition**: Growth rate increases over time.**Example**: Growth rate of bacteria dividing by binary fission.*In biological systems, data of growth rates may be represented with an exponential growth curve. With an exponential growth curve, there is an increase in organismal growth at an exponential rate.* *Most applications of this graph would shift this curve to the right because negative x-axis values would be rare.* | **Exponential Graph** |
| **Logarithmic trend curve:****Equation**: $y=A+Blnx$**Definition**: Growth rate decreases over time.**Example**: Height increase of a person from a baby to an adult.*Most applications of this graph would only apply to the positive function values because negative y-axis values would be unlikely.*  | **Logarithmic Graph** |
| **Logistic trend curve:****Equation**: $y=\frac{B}{1+Ae^{-kx}}$*y=0* and *y=B* are horizontal asymptotes.$$(0, B/(1+A) ) y-intercept$$The domain is all real numbers and the range $0<y<B$**Definition**: Growth rate increases then decreases.**Example**: Number of oak trees in an acre, over time.*In biological systems, data of growth rates is usually represented with a logistic growth curve. This is because rapid organismal growth steadily slows due to restraints of the environment. This type of curve shows an increase in numbers and then a plateau. The plateau represents no net change in the number of living organisms. It should be pointed out that logistic growth combines exponential growth followed by the beginning of a logarithmic growth curve. But note that a logarithmic curve continues to increase, which logistic does plateau.* | **Logistic Graph** |

**Discuss followup questions (provided on student worksheet):**

**Part 2. Goal A. Interpreting complex data sets**
Students will review the data in Table 2 and answer a series of questions about the data. The teacher should lead a discussion about these questions after the groups discuss them.

The data is started once the country has its 5th death. This is done so that the growth rates of all countries can be compared, despite the fact that they had their first deaths on different days of the year.

Discuss followup questions (provided on student worksheet).
Additional questions you may wish to add:
Which state saw the fastest increase in the number of deaths? *An easy way to quickly estimate this is to look for a doubling of deaths between 40 and 60 days and then between 60 and 80 days, for example.*
Which state saw the slowest increase in the number of deaths?
Is the population size of each country factored in this data? What problems does that present? *This question is dealt with in the next section, but you may want to address it here as well.*

**Part 2. Goal B. Understand the benefits and drawbacks of various graph types**

Lead a discussion which helps students understand the different types of charts to determine how useful they will be with this data set. If time allows, this section could be expanded to include other types of graphs or to require students to graph other pie/bar/line charts. You could, for example, assign each group a type of chart and ask them to create 2 different charts that tell different stories about the data.

**Aside from the questions on their worksheets, here are additional questions and comments:**

**Pie graph:**
How was this data turned into a pie chart that represents 100% of the deaths across these 6 countries? *You may need to explain how the frequency count was found.*
Can we see a trend using a pie chart?
**Bar graph:**
Bar charts can be useful for looking at the total of some variable over time, but can be confusing if too many data points are added. Also, they encourage the viewer to think about each bar as a discrete category rather than a trend.
**Line chart**:
A line chart can be useful for looking at the total of some variable over time, but the lines between points can be misleading. Viewers may interpret that the data followed the straight lines between points. A grapher who is attempting to accurately portray the data would not choose data points that are too distantly spaced, but a negligent or deceptive grapher could use specific data points, connected by a line, to skew the actual trend.

**Part 2. Goal C. Graph data sets and compare resulting charts to trend lines**

Assign a country (or two) to each group.

Here are examples of ways to assign graphs to different groups:

|  |
| --- |
| **Six groups?**Assign each group a different country. |
| **Six groups, but you want each group to graph two graphs which are good to compare to one another?**Group 1: United States and Brazil (different trends)Group 2: Canada and Italy (different trends)Group 3: Mexico and Australia (different trends)Group 4: United States and Canada (similar trends)Group 5: Brazil and Mexico (similar trends)Group 6: Italy and Australia (similar trends) |
| **Only 3 groups?**Group 1: United States and BrazilGroup 2: Canada and ItalyGroup 3: Mexico and Australia |

When students are done graphing, they should answer the questions on their worksheet. Here are some suggested answers to the first question:

Question: Describe the trend curve that best fits your data set.
*Possible answers:
United States: logistic
Brazil: logistic or exponential
Italy: Logarithmic
Canada: logistic
Mexico: logistic or exponential
Australia: logarithmic*

**The teacher will then lead a discussion comparing the newly created graphs to each other and a prepared graph of all of the countries together.** The Teacher’s Powerpoint contains copies of each of the finished graphs if you want to display them instead of the student’s finished graphs.

Aside from the questions on their worksheets, here are additional questions:

Which countries have similar curves?

How well did the data fit a perfect example of that curve? Why don’t some of them fit perfectly?

Imagine if we had graphed all of these countries on one graph. What are the benefits of looking at the graphs of each country individually instead?

What are the drawbacks of looking at each county’s graph individually?

**Part 3: Goal: Use various methods to make predictions.**

In this section, students will use the data and/or the scatter plots completed in Part 2 to predict the next data point (120 days after the 5th death).

Assign students one or two of the five methods of prediction. If you do assign more than one method, you can have each group do Method 1 or 2 and then Method 3 or 4 or 5.

The linear equation will not be used to make predictions for these data sets because none of the countries exhibited linear growth. Instead of using the linear equation, we will use the average rate of change (method 2) to make a prediction.

You will need to provide various data points for the different methods.

|  |  |  |
| --- | --- | --- |
| **Method** | **Before they start** | **After they complete their prediction** |
| **1- Visual estimate** |  | Provide the actual data on day 120 (provided below table) |
| **2 - Rate of change** | May need to explain the y and x variables. For example, to figure out the rate of change from 20-40 days in the United States, the equation would be (18,777-108) / (40-20)  | Review student’s average rates of change (provided below table)Then provide the actual data on day 120. |
| **3- Exponential**  | k values for each country:United States:Brazil:Italy:Canada:Mexico:Australia: | Provide the actual data on day 120 |
| **4 - Logarithmic** |  | Provide the actual data on day 120 |
| **5 - Logistic** |  | Provide the actual data on day 120 |

**Actual data on day 120**:
United States: 127,410
Brazil: 77,851
Italy: 34,675
Canada: 8,810
Mexico: 41,190
Australia: 108

**Rate of change data for method 2:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|   | 20-40 days | 40-60 days | 60-80 days | 80-100 days | Estimated rate of change for 100-120 days  |
| United States | 933.45 | 2314.55 | 1418.55 | 928.35 | 700 - 900 |
| Brazil | 235.90 | 530.80 | 1075.05 | 1024.40 | 800 - 1000 |
| Italy | 686.55 | 483.00 | 317.95 | 16.05 | 0 - 50 |
| Canada | 104.25 | 172.35 | 99.25 | 29.35 | 0 - 50 |
| Mexico | 114.90 | 234.50 | 452.70 | 669.75 | 750 - 1000 |
| Australia | 2.15 | 0.95 | 0.20 | 0.05 | 0 -1 |

**Method 3:**
Teacher will provide A and k for each country or provide the Exponential equations below.
USA: A = 161.98555; k = 0.7747
Brazil: A = 203.40188 ; k = 0.06262
Italy: A = 24861.42593; k = 0.03052
Canada: A = 357.18708; k = 0.03688
Mexico: A = 231.76708; k = 0.052
Australia: A = 38.55723 ; k = 0.01179

Equations for the different models of each data set:

USA Data

Linear $y=1492.29 x - 31657.8$

 Exponential $y=161.98555 e^{0.07747x}$

 Logarithmic $y=72549.10712 ln(x) - 228923.75187$

 Logistic $y=\frac{111333.50376}{1 + 165.36564 e^{-0.08912x}}$

Brazil Data

Linear $y=733.815 x - 20887.9$

 Exponential $y=203.40188 e^{0.06262x}$

 Logarithmic $y=33477.45824 ln(x) - 109203.11305$

 Logistic $y=\frac{70711.55859}{1 + 218.49838 e^{-0.06866x}}$

Italy Data

Linear $y=380.805 x - 1301.7$

 Exponential $y=24861.42593 e^{0.03052x}$

 Logarithmic $y=19828.14277 ln(x) - 56838.6212$

 Logistic $y=\frac{32220.62416}{1 + 48.45825 e^{-0.09209x}}$

Canada Data

Linear $y=108.2 x - 1464.4$

 Exponential $y=357.18708 e^{0.03688x}$

 Logarithmic $y=5407.68031 ln(x) - 16350.20739$

 Logistic $y=\frac{8544.46456}{1 + 88.35674 e^{-0.08835x}}$

Mexico Data

Linear $y=363.09 x - 10426.4$

 Exponential $y= 231.76708 e^{0.052x}$

 Logarithmic $y=16371.83294 ln(x) - 53362.6313$

 Logistic $y=\frac{50063.67314}{1 + 157.80241 e^{-0.05448x}}$

Australia Data

Linear $y=0.785 x + 36.5$

 Exponential $y=38.55723 e^{0.01179x}$

 Logarithmic $y=42.93589 ln(x) - 86.13546$

 Logistic $y=\frac{102.96376}{1 + 11.35237 e^{-0.09049x}}$

After all groups have completed their predictions, received the actual data on day 120, and answered their questions, encourage each group to share their predictions with the entire class (or simply another group that used different methods). Discuss the various pros and cons of each method. Which method was best able to predict the actual data?

Finally, display the actual graphs that show data on day 200. If the students would have used their methods to predict this far into the future, which methods would still have yielded good predictions? If you’d like to expand this lesson, you could provide more data points (day 140, 160, 180) and see if the students, using the best method AND more data points, could better predict the data on day 200.