**Instructional Guide**

**Title: Data Analysis**

**Suggested Quarter: 4th**

**West Virginia College- and Career-Readiness Standards:**

**M.1HS.31** Represent data with plots on the real number line (dot plots, histograms, and box plots)

**M.1HS.32** Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.

**M.1HS.33** Interpret differences in shape, center and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).

**M.1HS.35** Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

a. Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear and exponential models.

b. Informally assess the fit of a function by plotting and analyzing residuals. (Focus should be on situations for which linear models are appropriate.)

c. Fit a linear function for scatter plots that suggest a linear association.

**M.1HS.36** Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data. Instructional

**Performance Objectives: (Know & Do)**

**Know:**

             How to gather, create, and interpret data from various data displays.

             How to use a linear regression (fitted line) to draw conclusions.

             How to determine and use the measures of central tendency to draw conclusions.

             Vocabulary: linear regression, scatter plots, histogram, normal distribution,

            measures of central tendency, mean, median, mode, and range.

**Do:**

            Gather and represent data in various data displays.

            Use the linear regression equation to make predictions.

            Use technology to calculate and display data.

**Big Idea:** We can analyze and draw inferences from data.

**Enduring Understandings:**

Data displays can be used to make and support conclusions

Measures of central tendency can be used to support data analysis

Linear regressions can be used to make predictions

**Essential Question:**

1.  In what ways can data be gathered and organized so data is clear and concise?

2.  What is the best way to display data?

3. What kinds of inferences can be made about data?

**Learning Plan and Notes to Instructor:**

This unit allows students to gather, display, and analyze data using various types of data displays.  Direct instruction can be utilized to **review**  (most of these concepts have been taught through middle school.) the different types of data displays and measures of central tendency.  The CSO refers to histograms; scatter plots, box plots and normal distribution.   The teacher should also review the line of best fit or the linear regression equation to make a connection between linear equations, slope, rate of change and data prediction.

Since most of these displays should be review, I suggest that the teacher review with the many of the websites that I will list below.  The main new emphasis may be with applying the line of regression to the scatter plot, the difference between bar graphs and histograms, <http://www.amstat.org/publications/jse/v5n2/gnanadesikan.html>,

and how to use normal distributions . <http://www.shodor.org/interactivate1.0/discussions/sd_bell1.html>

The website: <http://www.explorelearning.com/> is an excellent site to review all types of skills.  The website uses a gizmo (applet) for students to actually manipulate data.  There are  printable directions as well as on on-line assessment.  Students can actually print out their graph for review and assessment.  **You must set up an account for this website.  There is a free 30-day trial and you can register as a teacher and put classes in for whole classroom use.  I believe that is well worth the amount to use this site.**

The following website gives you data and activities to use for other activities you may want to use to make data displays:

<http://nces.ed.gov/nceskids/index.asp>

**Lesson Plans for the Academic Prompts:**

For the academic prompts the following student data should be collected: foot length in cm, height in cm., balancing data in sec. (arms over head, one foot, eyes closed- one hop permitted only) and time spent watching TV (includes video games and DVDs/VHS) during the school week.

**Balancing Act:**  Students need to review how to use the measures of central tendency to interpret the data from them.  Use explorelearning.com website and/or <http://nlvm.usu.edu/en/nav/vlibrary.html>

**Foot Size-Height Matters:**For this lesson, review finding the trend or relationship between the independent and dependent variables.  Review correlations and finding the line of best fit (line of regression) using the graphing calculator.  If you have the use of a TI-83 or 84 you may want to discuss the r- value.  Also review using the line to predict future trends.

**TV –Time -** For this lesson,review how to make a histogram. You may want touse the M&M activity, which lends itself to a histogram.  Also review using the displays to give data conclusions.

**Differentiated Instruction Strategies**

Always be aware of students with special needs.  Use foldables, the Frayer model or any other type of graphic organizer to help students with necessary vocabulary.  Tier assignments for student ability and be aware of students who do not have access to technology at home and those who may have to have alternate assignments in case of blindness or deafness.   The use of writing is indispensable for student learning and students should journal often so that the teacher can look at student understanding and learning. Teaching strategies should include a mixture of direct instruction, group and independent learning.  Emphasize the rules of cooperative learning.

**Academic Prompts:**

**Balancing Data**

You work at the circus.  The circus ringmaster, who is planning to add a new tightrope walker, is interested in knowing which one of his employees is the best balancer.  Working in your group of four and using the balancing data that was taken before, create a box and whisker plot for the employee data, and girl vs boy balancing data. Your group will present your conclusions about who is the best balancer to the ringmaster by using chart paper or spreadsheet software.

[Rubric](#Balancing)

**Foot Size or Height- Which Matters?**

You have determined who the best balancer in the circus is, but you need to find out if foot size or height matters in helping to balance.  You and your group want to investigate this question.  Using the employee balancing data and the foot size and heights, create two scatter plots that show the relationships between the data.  Using a fitted line, show the correlation and make a prediction about foot size and balancing and height and balancing.

1.Use graph paper/chart paper to present your scatter plots and to present your conclusions to the ringmaster.  Write an equation for the line of regression.

2. Using graphing calculators, graph and plot the scatter plots to find the real line of regression and print the scatter plots and line of regressions.

[Rubric](#foot)

**TV Time**

You are the national motivational speaker for *TV Turn Around*, and you want to prove to students that they spend more time watching TV or playing video games that they think they do. Using the data gathered from each student, make a histogram. Using the Internet, research the average time that students between the ages of 12 and 18 spend watching TV and playing video games.  Use a media presentation to show the comparison between the national trend and the specific classroom.

[Rubric](#tv)

**Culminating Assessment or Product:**

**Grammy Groupie**

You are a music producer and you need to present to your company  the trends in Grammy winners and genres of music so they know where to put the company‘s advertising monies.

Provide the president of your company with at least two different data displays containing your predictions for which genre will win next year based on past trends. Use the tools of your choice to make your presentation. Information on the Grammy Awards can be found at <http://www.grammy.com/>.

[Rubric](#grammy)

**Links and Other Resources:**

<http://wps.aw.com/wps/media/objects/52/54247/activities/histogram/index.html>

<http://www.shodor.org/interactivate1.0/discussions/sd_bell1.htm>

***Student Materials:***

* Graphing Calculators
* White Boards
* Chart Paper
* Graph Paper
* Computers with Word Processing/Presentation Software
* Internet Access
* Pre and Post Assessments
* Timers

**Contact Author:**

[**mholland@k12.wv.us**](mailto:mholland@k12.wv.us)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Employee Box and Whisker  Balancing Data | Girl and Boy Box-whisker Plot | Predictions | Presentation |
| **4** | Thorough, complete and accurate | Thorough, complete and accurate | At least two accurate predictions based on measures of central tendency | Highly effective; Clear and easy to follow |
| **3** | Mostly thorough, complete and accurate | Mostly complete and accurate, may contain minor computational error | At least two accurate (may contain minor inaccuracies that do not effect overall result) predictions based on measures of central tendency | Generally effective; Generally clear and able to be followed |
| **2** | Partially complete and accurate | Partially complete and accurate, may contain more serious errors | One accurate prediction based on measures of central tendency | Somewhat effective; Lacks clarity and /or difficult to follow |
| **1** | Serious misunderstandings; numerous errors | Serious misunderstandings; numerous errors | One accurate (may contain minor inaccuracies) prediction based on measures of central tendency | Ineffective; Unclear and impossible to follow |

**BALANCING ACT**

**Foot Size or Height**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **First**  **Scatter Plot** | **First Line-**  **of-Best Fit** | **Second Scatter Plot** | **Second Line-of-Best Fit** | **Predictions** | **Presentation** | **Comparison of lines of regression** |
| **4** | Thorough, complete and accurate | Thorough, complete and accurate | Thorough, complete and accurate | Thorough, complete and accurate | At least two accurate predictions based on lines-of-best-fit | Highly effective; Clear and easy to follow | Compares the lines of regression and makes an effective reasoning for the comparison |
| **3** | Mostly thorough, complete and accurate | Mostly complete and accurate, may contain minor computational errors. | Mostly thorough, complete and accurate | Mostly complete and accurate, may contain minor computational error | At least two accurate (may contain minor inaccuracies that do not effect overall result) predictions based on lines-of-best-fit | Generally effective; Generally clear and able to be followed | Compares the lines of regression but the conclusion is generally effective |
| **2** | Partially complete and accurate | Partially complete and accurate, may contain more serious errors | Partially complete and accurate | Partially complete and accurate, may contain more serious errors | One accurate prediction based on lines-of-best-fit | Somewhat effective; Lacks clarity and /or difficult to follow | Somewhat effective comparison of the lines of regression |
| **1** | Serious misunderstandings; numerous errors | Serious misunderstandings; numerous errors | Serious misunderstandings; numerous errors | Serious misunderstandings; numerous errors | One accurate (may contain minor inaccuracies) prediction based on lines-of-best-fit | Ineffective; Unclear and impossible to follow | Makes unclear comparisons of the lines of regression |

# TV Time

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Histogram** | **Presentation** | **Conclusions** |
| **4** | Thorough, complete and accurate | Highly effective; Clear and easy to follow | Compares the data and makes an effective conclusion based on the data. |
| **3** | Substantially thorough, complete and accurate | Generally effective; Generally clear and able to be followed | Compares the data and makes a generally effective conclusion based on the data. |
| **2** | Partially complete and accurate | Somewhat effective; Lacks clarity and /or difficult to follow | Somewhat effective conclusion about the data. |
| **1** | Serious misunderstandings; numerous errors | Ineffective; Unclear and impossible to follow | Makes unclear comparisons of the data and the conclusion is reflects little of the data. |

Grammy Groupie

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Data display of Grammy Data -1 | Data display of Grammy Data -2 | Prediction/Trend Data Display-1 | Prediction/Trend Data Display-2 | Presentation |
| **4** | Thorough, complete and accurate | Thorough, complete and accurate | At least two accurate predictions based on measures of central tendency or other data from data display | At least two accurate predictions based on measures of central tendency or other data from data display | Highly effective; Clear and easy to follow |
| **3** | Substantially thorough, complete and accurate | Substantially complete and accurate, may contain minor computational error | At least two accurate (may contain minor inaccuracies that do not effect overall result) predictions based on measures of central tendency or other data from data display | At least two accurate (may contain minor inaccuracies that do not effect overall result) predictions based on measures of central tendency or other data from data display | Generally effective; Generally clear and able to be followed |
| **2** | Partially complete and accurate | Partially complete and accurate, may contain more serious errors | One accurate prediction based on measures of central tendency or other data from data display | One accurate prediction based on measures of central tendency or other data from data display | Somewhat effective; Lacks clarity and /or difficult to follow |
| **1** | Serious misunderstandings; numerous errors | Serious misunderstandings; numerous errors | One accurate (may contain minor inaccuracies) prediction based on measures of central tendency | One accurate (may contain minor inaccuracies) prediction based on measures of central tendency | Ineffective; Unclear and impossible to follow |