

# EllipSeeIt: Visualizing Strength and Direction of Correlation

Miguel Olvera

Rafael E. Diaz

Cosumnes River College

California State University Sacramen  
rdiaz@csus.edu

Marco Molinaro, Matthew Steinwachs,  
Amy Smith, Tobias Spencer, Alberto  
Guzman, Alyssa Okita

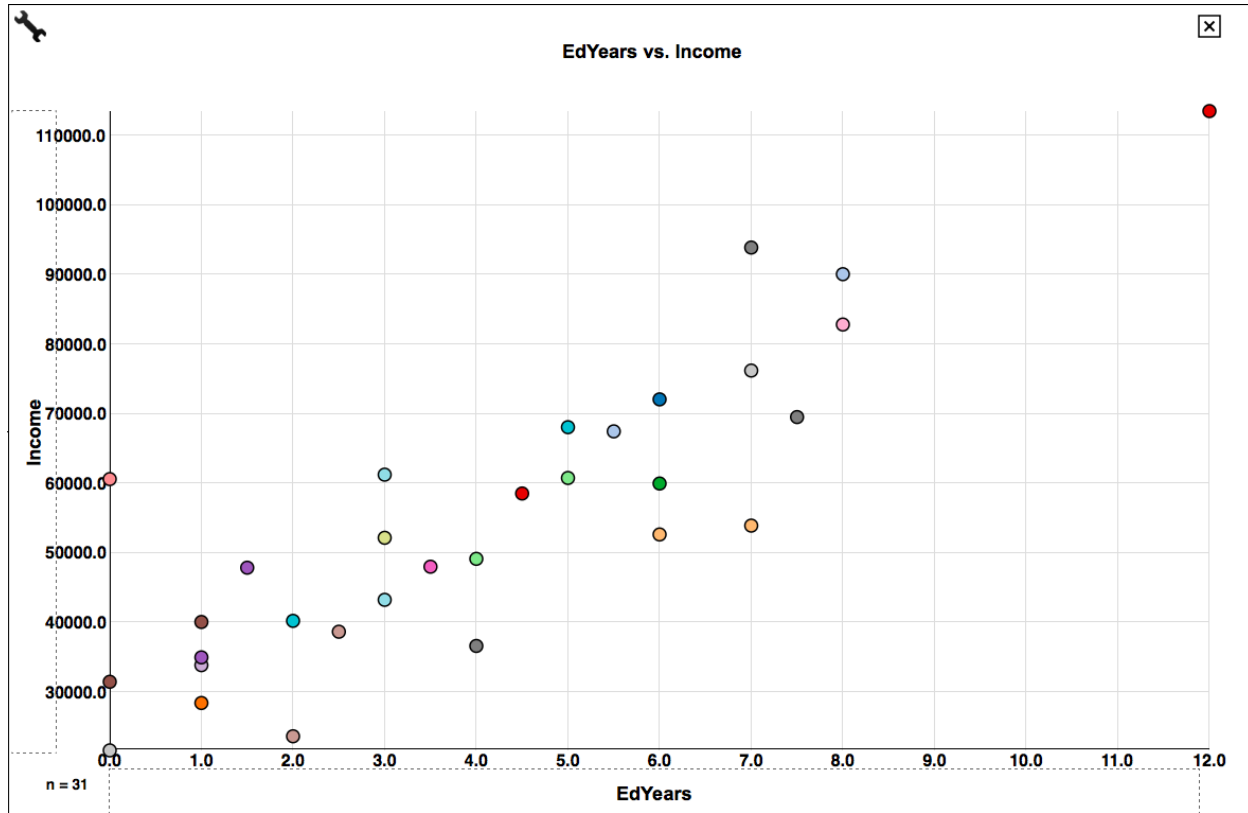
University of California Davis, iAMSTEM  
Hub

**Published: February, 2017**

**Student Handouts and Technology Directions**

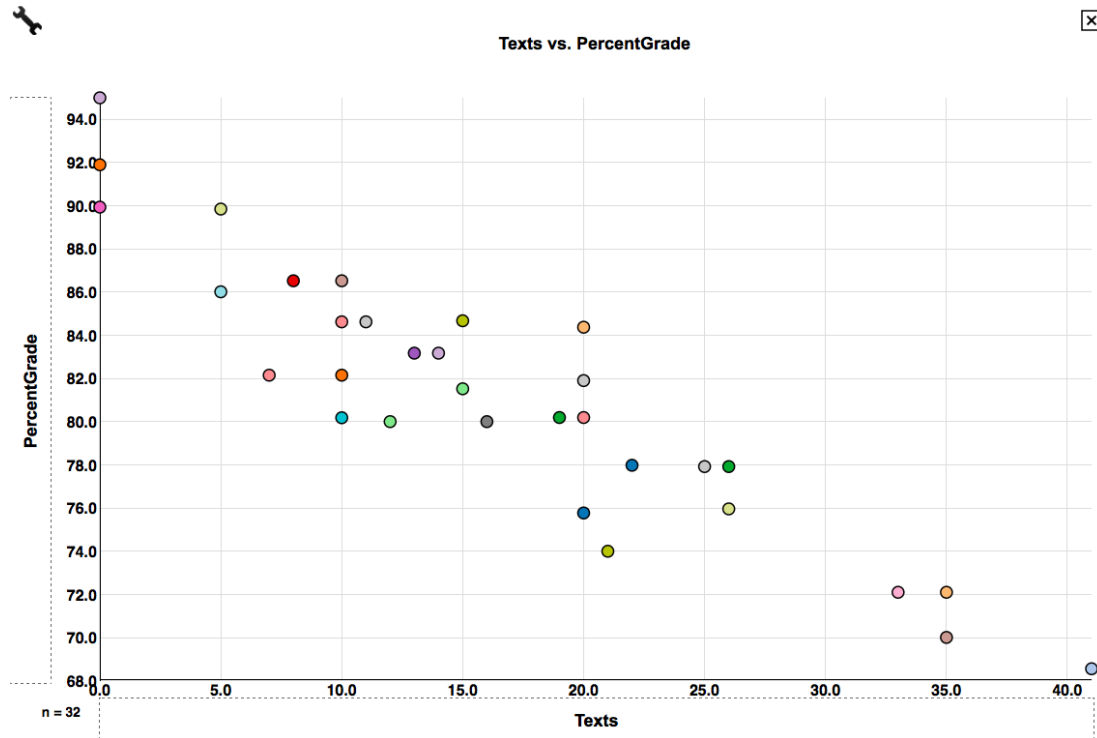
## EllipSeeIt: Visualizing Strength and Direction of Correlation Student Handout Day 1

1. Answer the following questions by analyzing the following scatterplot .



- For people in this dataset who have 1 year of education after high school, annual income ranges from \_\_\_\_\_ to \_\_\_\_\_.
- For people in this dataset who have 6 years of education after high school, annual income ranges from \_\_\_\_\_ to \_\_\_\_\_.
- Repeat the instructions of parts a and b above scanning the whole range of the years of education. What happens to the ranges of annual income in general as the number of years of education after high school increases? Do these ranges tend to shift upward (as in a line with a positive slope) through the entire range of years of education? What can you say about the general trend between the annual income of a person and the number of years of education after high school?

2. Answer the following questions by analyzing the following scatterplot.



a. For students in this dataset who never sent a single text message during class, their percentage grade on the exam ranges from \_\_\_\_\_ to \_\_\_\_\_.

b. For students in this dataset who sent an average of 10 text messages during class, their percentage grade on the exam ranges from \_\_\_\_\_ to \_\_\_\_\_.

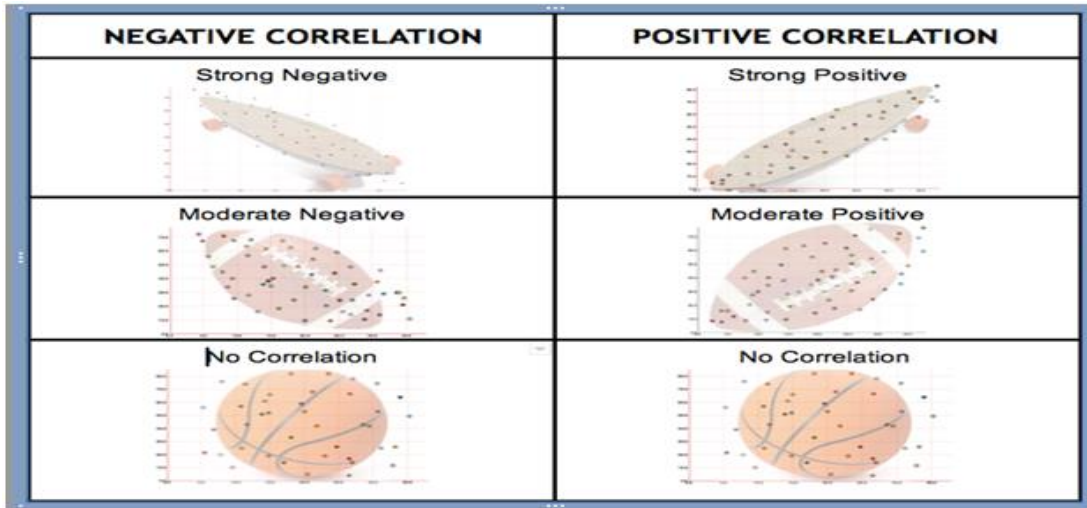
c. Repeat the instructions of parts a and b above scanning the whole range of average text messages sent during class. What happens to the ranges of grades in general as the number of text messages sent during class increases? Do these ranges tend to shift downward (as in a line with a negative slope) through the entire range of number of text messages? What can you say about the general trend between the number text messages sent in class and exam grades?

3. Answer the following question by analyzing the scatterplots of the pair of variables you chose to analyze.

What can you say about the ranges of the response variable as the values of explanatory variable increase? Is the relationship between these variables approximately linear?

## EllipSeeIt: Visualizing Strength and Direction of Correlation Student Handout Day 2

1. Using the shapes shown below, how would you describe the strength and direction of the correlation between the variables you chose to analyze?



2. Answer the following questions after creating scatterplots for the datasets TVHoursVsGPA ( $x$ -axis: HoursTV,  $y$ -axis: GPA) and TextingVsPercentGrade ( $x$ -axis: Texts,  $y$ -axis: PercentGrade) in *SeeIt*.

- Is there an approximate linear relationship between GPA and the number daily hours of watched TV increases for this group of students? Explain.
- Using the shapes shown in, how would you describe the strength and direction of the correlation between TVHours and GPA after enclosing most of the points in an ellipse?
- Is there an approximate linear relationship between text messages sent during class and percentage grades on the exam for this group of students? Explain.
- Using the shapes, how would you describe the strength and direction of the correlation between Texts and PercentGrade after enclosing most of the points in an ellipse?
- Is the strength of the correlation of TVHours versus GPA stronger, weaker or equal to that of Texts versus PercentGrade?

## EllipSeeIt: Visualizing Strength and Direction of Correlation Technology or Simulation Directions

### Creating a Data Set in *SeeIt*

<http://centerforbiophotonics.github.io/SeeIt3/correlations-stew.html>

Once students have loaded the *SeeIt* tool using the link above, ask them to click on “Add a Worksheet” at the bottom of the dataset panel located on the left of the screen in order to display a *SeeIt* edit worksheet. Ask students to add a title to their dataset in the rectangle that says \*\*\*Enter Worksheet Title\*\*\*. Students must follow the following format when entering the dataset into the worksheet:

- The first row must have the labels (each one a single word) of the columns separated by a comma followed by a space, e.g., Name, Height(cm), Armspan(cm), RightFoot(cm)
- All rows after the first one need to have the name of a student (one single word), and the measurements of the variables they have chosen to analyze separated by a comma followed by a space. Note that students may choose to measure more than two variables to be analyzed pairwise. Figure below illustrates the format needed for a data set from a group of students collecting measurements of their height, armspan and right foot length.

Note: if the teacher has already collected the students’ data on a spreadsheet (Excel or Google Sheet), he/she can email it to the students so that they can simply select and copy the data from the worksheet, and paste it directly onto the *SeeIt* edit worksheet. Note that when pasting the data from an Excel, or a Google Sheet, the columns of data are separated by tabs, which is another format of datasets accepted by *SeeIt*.

**Values MUST be separated by tabs or commas but not both.**

Title:

First Column is Label:

Data:

Name	Height(cm)	Armspan(cm)	RightFoot(cm)
Tony	172	125	40
Gaby	68	115	32
Dylan	148	120	52
Jennifer	140	102	38
Danielle	162	130	35

Format required to enter data onto an edit worksheet in *SeeIt*.

Once students have finished collecting their dataset, ask them to select the entire dataset in the *SeeIt* edit worksheet, copy it, and paste it into an Excel or Google Sheet. Then students can save this spreadsheet by emailing it to themselves as an attachment or storing it in a portable USB drive to use the second day of the lesson plan. Students will need to paste back their dataset into *SeeIt* the second day of the lesson plan since their data will be lost once they close *SeeIt* (datasets entered into *SeeIt* will appear again in the dataset panel of another *SeeIt* session only if the browser's cookies are not deleted in between *SeeIt* sessions).

After students are done saving their data, ask them to click on the “Load Worksheet From Form” button located at bottom of the *SeeIt* edit worksheet. This will refresh the data set panel (left side of *SeeIt* window) where students will be able to see the title of their data set). Note that the dataset panel can be shown or not by clicking on the button that says “Show Datasets” or “Hide Datasets” located at the upper left corner of the graphing window of *SeeIt*.