

CASE $Q \rightarrow Q$

Two Variables: Examining Relationships
Unit 1: Exploratory Data Analysis



[Interactive Data Analysis](#)

UF UNIVERSITY of
FLORIDA

Now we look at the last case in examining relationships, two quantitative variables, Case Q-Q and there are a number of new ideas to discuss for this topic.

Let's take a sneak preview using this interactive data analysis applet for two quantitative variables.

{weblink: <http://bolt.mph.ufl.edu/2012/12/24/interactive-applet-case-q-q/>}

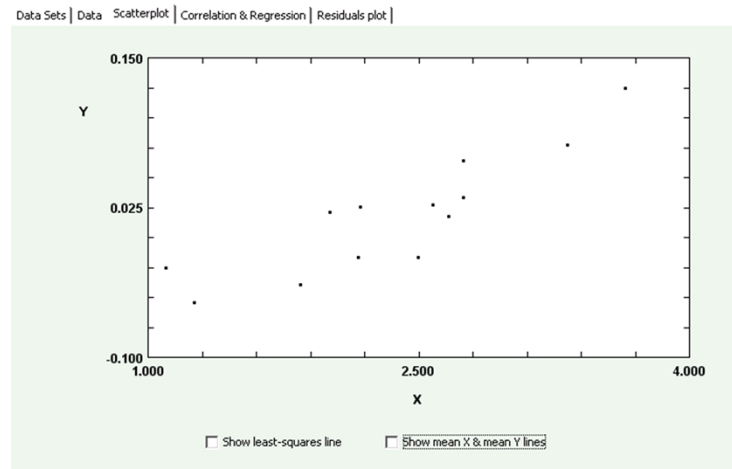
Case Q-Q Applet – Choose Data

Data Sets | Data | Scatterplot | Correlation & Regression | Residuals plot |

- ☒ Social distress and brain activity (Exercise 2.33, page 98)
- ☐ Value and income of NBA teams (Table 2.1, page 99)
- ☐ Nonexercise activity (NEA) and fat gain (Example 2.18, page 108)
- ☐ Fruit and vegetable consumption and smoking by state (Exercise 2.156, page 156)
- ☐ Neuron response to tones and monkey calls (Table 2.7, page 158)
- ☐ Gas mileage versus log of speed for a car (Example 10.10, page 560)
- ☐ In-state tuition and fees in 2000 and 2008 (Table 10.1, page 580)
- ☐ Beer consumption and blood alcohol content (Exercise 10.13, page 581)
- ☐ Watershed area and water quality (Table 10.4, page 584)
- ☐ SAT and ACT scores (Exercise 10.51, page 588)

We will use the first dataset on social distress and brain activity. You can view the data using the data tab.

Case Q-Q Applet – View Scatterplot



We'll skip straight to the scatterplot. For two quantitative variables, our graphical displays and numerical measures are very reminiscent of algebra class. Do you remember $Y = mX + b$? If so, it might help!

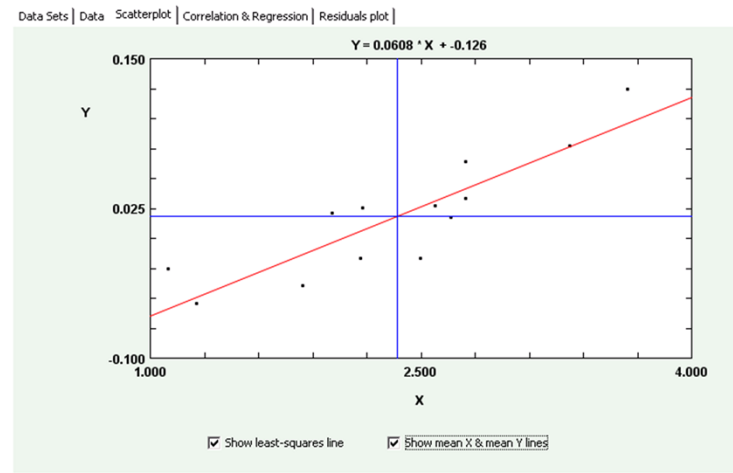
Here in this scatterplot, we plot each observation in two-dimensions by locating the intersection of the x and y coordinate locations for that observation.

We are looking for trends in the plot. In particular, in this course, we will be very happy if we see a linear trend - such as the one in this plot - as all of our methods will work well in the case of linear relationships.

When looking at scatterplots, we look at the overall pattern which will include the direction (positive for increasing trends, negative for decreasing trends), form (linear, curvilinear or non-linear), and strength (how closely the data points fit the overall pattern).

We will also look for outliers.

Case Q-Q Applet – Scatterplot

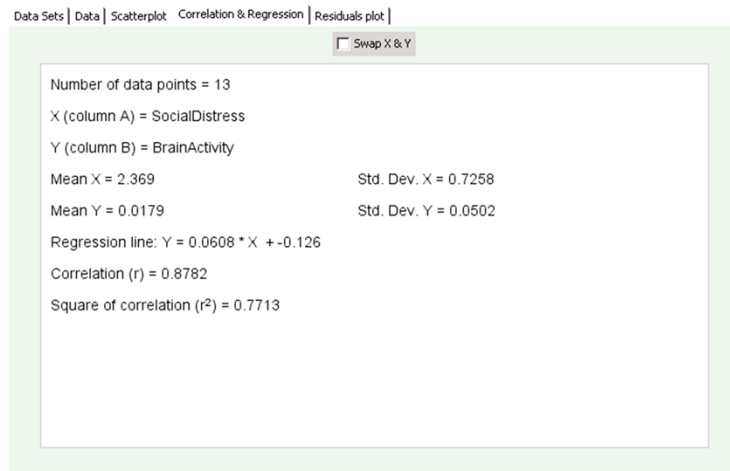


In the applet, we can add the linear regression equation and lines for the mean X and mean Y.

As we can see in the plot, for this dataset, as X = social distress increases, Y = brain activity also tends to increase.

This scatterplot shows a positive linear trend.

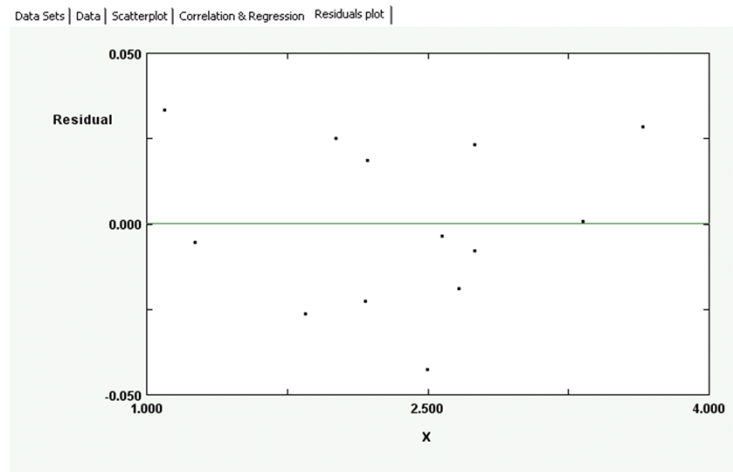
Correlation and Linear Regression



In the correlation and regression tab, we see some information including

- Which variables are defined to be X and Y – you can swap these if you wish using the checkbox at the top
- The mean and standard deviation for both X and Y
- The number of data points
- The simple linear regression equation
- The correlation and it's square

Residual Plot



The final tab shows the residual plot, which is a plot showing how far each data point fell from the regression line. We will learn that we want to see a random scatter in this plot with no clear trend – this will indicate the best case scenario for a linear regression analysis.



CASE $Q \rightarrow Q$

Two Variables: Examining Relationships
Unit 1: Exploratory Data Analysis

You can use this applet to help you understand and apply the concepts in the section on Case Q-Q, examining the relationship between two quantitative variables.