Chapter 10

# Try it Yourself: Calculate Similarity Using Distance

Create the scatterplot:

Open the small\_penguins.csv data in StatCrunch

Click Graph -> Scatter Plot

X variable -> bill\_length\_mm

Y variable -> bill\_depth\_mm

Group by -> species

Click Compute!

Calculate the distances:

Unfortunately, StatCrunch does not offer an easy way to create such a matrix, so you will have to write expression for each column.

Data -> Compute -> Multiple Expressions

Create 15 expression lines with names observation\_1 through observation\_15.

Enter the first expression:  
sqrt((bill\_length\_mm-38.9)2+(bill\_depth\_mm-18.8)2)  
where 38.9 and 18.8 are bill\_length\_mm and bill\_depth\_mm values in the first row.

Copy this expression to other lines and replace these numeric values with those from the corresponding row of the data. For example, the last expression will be  
sqrt((bill\_length\_mm-47)2+(bill\_depth\_mm-17.3)2)

Click Compute!