MEMORANDUM

TO: Mrs. Mowers

FROM: Alex Embry

DATE: March 24, 2010

Mrs. Mowers, I do believe my colleagues and I have found the culprit of last fall’s break in. In this report I have given you the suspect whom my team and I believe is the guilty party. The great news we have done all the work for you, and the suspect was determined using statistics.

The first step we had to do in order to determine a suspect was use the data given to us in the CSI Shoe Data table. This table was a list of people numbered 1 to 39, their shoe length in centimeters, their height in inches, and their gender. After looking over the table, it was determined that we would use shoe length as the explanatory variable (X axis), and height as the response variable (Y axis). From the information shown and common knowledge of the team we know the height of a person and their shoe size typically go hand in hand. We entered the x values, or the shoe sizes, into our calculator in the L1 column. Next, the y values, or the heights, were entered into our calculator in the L2 column. After we graphed the results this is what we came up with...

The graph of the data is described as linear, positive, and moderately strong. After plotting the data, the next step was to find the least squares regression equation. To do this, we used the
same values in L1 and L2, the equation of $y=ax+b$ calculated as [equation deleted]. The least squares regression line was then added to the graph and it appeared to look like this...

The correlation between the length of the shoe and the height, or the “r” value is $r=0.69$, and the coefficient of determination, “r” squared was calculated as 0.47 or 47%. The coefficient of determination tells us that 47% of variation in height is explained by the regression on shoe length.

Now finally we had to measure the footprint. We used a standard ruler and came up with a footprint length of approximately [measurement deleted] cm. Since we have a shoe size, or an “x” value, we can now substitute it into our equation to predict how tall our suspect must be. In the equation [equation deleted], the height comes out to be [height deleted] inches. Since the height of the suspects were shown in front of a scale in 2-inch increments we rounded the height up to [height deleted] to make it easier to follow.

Using our graph and equation we predicted that the person who has a footprint measuring [length deleted] centimeters, should be approximately [deleted] inches tall, or [measurement written as feet and inches deleted] inches tall. Using the predicted height, and looking at the mug shots we determined that using our prediction and the least squares regression equation that suspect number [id number deleted] had to be the suspect.
My team and I are very confident that we have pinpointed the correct suspect. Although it is possible to be some outlying variables we believe that with the information given we have eliminated all the others and have proven just one as guilty. After you question this culprit we are sure you too will agree. If you need any more help in this investigation you know where to find me.