Nick Foster

Excel Lab

A)

1) There is not a true linear relation between the x and y data. Although the -.70 correlation factor would seem to suggest a strong correlation, this is only because the graph of the plots is more of a curve with a long right tail. This long right tail gives the appearance of a line. By looking at the residuals you can tell that it is actually a curve, as there is a pattern in the residuals of starting at positive, then all the residuals become negative for a long time, and then they go back to positive. The residuals show very little randomness.

2) There is a stronger linear relation between the ln(y) and x than between just normal y and x. The correlation between these two is -.95, indicating a very strong linear relationship. While the residuals show a pattern of going from positive to negative back to positive, indicating a curve, the residuals are so close to zero that the curve is very minimal, almost to the point of non-existence.

3) There is an even stronger linear relationship between ln(x) and ln(y) than the previous relationships. The correlation factor is -.98, meaning it’s almost a straight positive line. The residuals are in a range of about -.4 to about .3, meaning they are very small and very close to the actual line. Also they are randomly scattered, indicating no pattern and therefore a strong linear relationship.

4) For step 1 omitting 1-8, the relationship becomes slightly more linear, but it’s still not a straight line. The correlation went up from -.7 to -.89, but the residuals still show a positive to negative to positive pattern, indicating a curve still exists. For step 2, the relationship went up from -.95 to -.98. For step 3, the relationship is almost completely linear, as the correlation is now up from -.98 to -.99, and the residuals are almost completely zero.

5) The functional relation is exponential curve. As the rank increases the population decreases exponentially, creating a curve.

B)

1) Standard Deviation should be used to measure uniformity because it measures how close all the wafer thicknesses are to each other and how uniform the thickness is over the entire wafer. Larger SD’s will show that the wafer has greatly varying thickness over the wafer where as smaller SD’s show that the thickness is very uniform.

2) I think that location 13 is a representative outlier because it is always significantly less than the other 12 locations on the wafer; approximately 200 angstroms less. This may be a result of the manufacturing of the semi-conductor or just an anomaly, but the location should still be disregard and considered and unrepresentative outlier.

3) After doing the data analysis, the y is the SD, and the two coefficients are m1=-.072 and m2=1.12. The intercept is 115.86. Therefore the equation is y=-.072x1+1.12x2+115.86

4) Yes, relationship is a good fight for the data because the graphs have a football shape and the residuals are randomly scattered around zero, showing that it’s a good linear relationship.

5) Oxide thickness has very little effect on the measure of the polysilicon thickness because the coefficient is -.07. This is a very miniscule amount and therefore shows very little relation, whereas the coefficient for deposition time is a more significant amount of 1.12, showing a more considerable relationship.