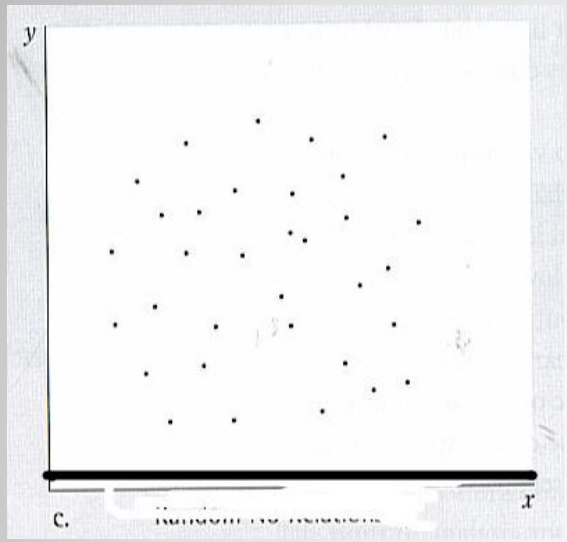
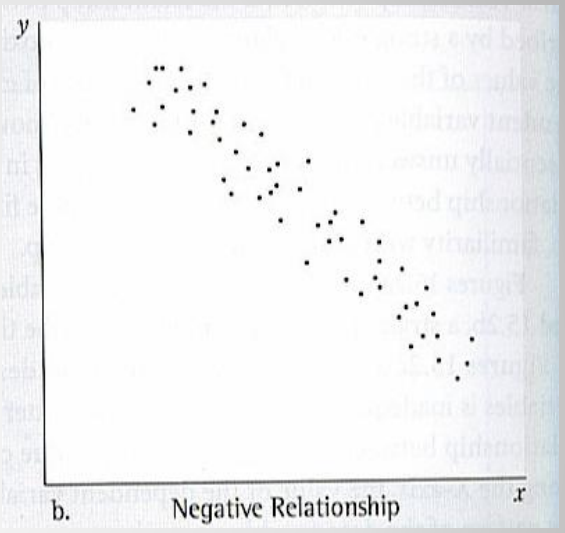
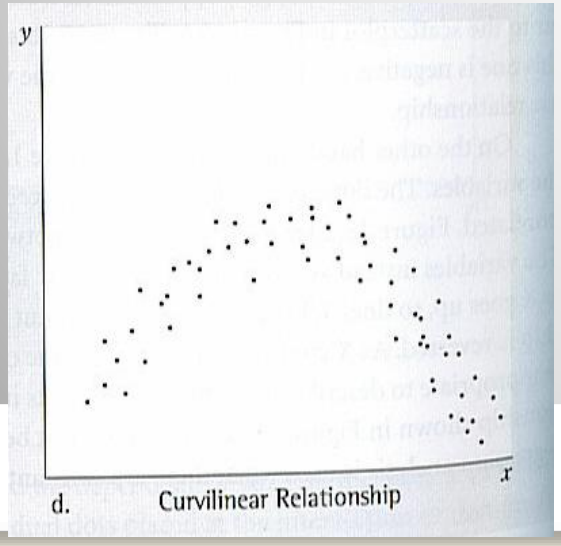
Regression Analysis

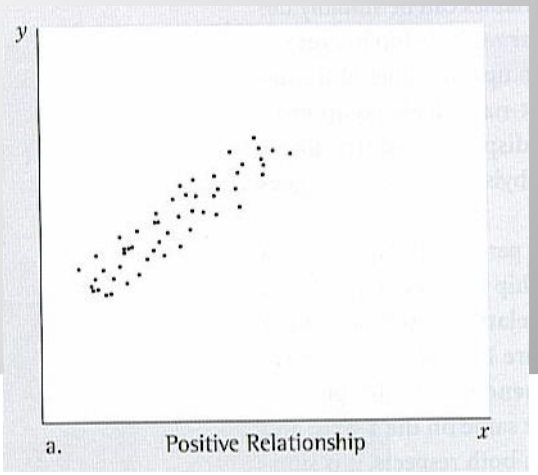
- Regression analysis requires interval and ratio-level data.
- To see if your data fits the models of regression, it is wise to conduct a scatter plot analysis.
- The reason?
  - Regression analysis assumes a linear relationship. If you have a curvilinear relationship or no relationship, regression analysis is of little use.

Types of Lines

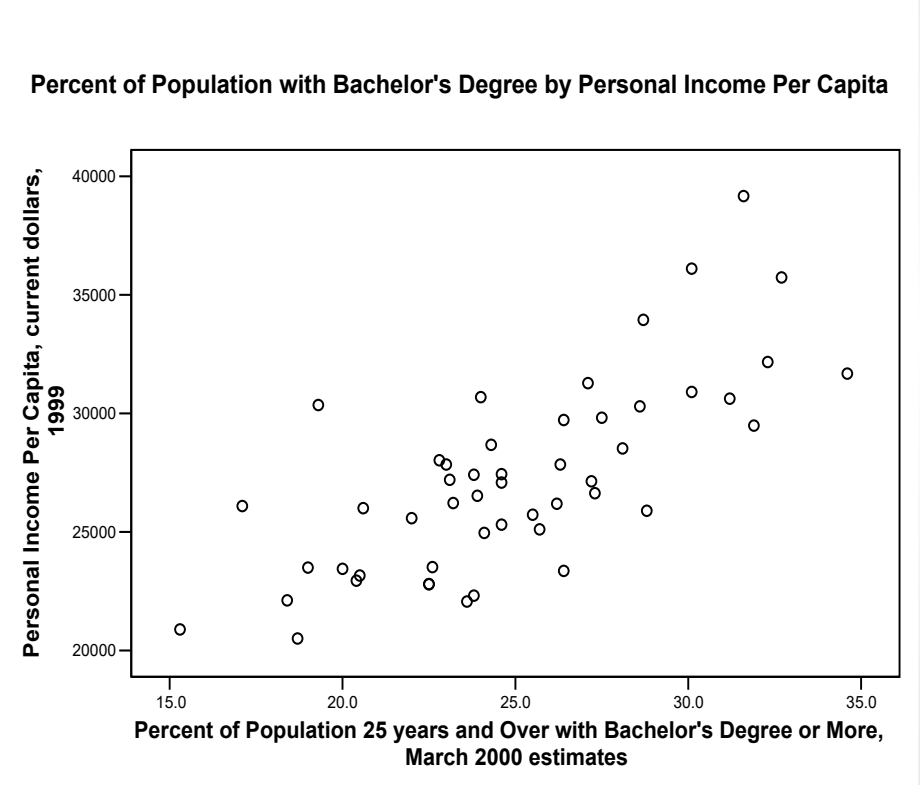






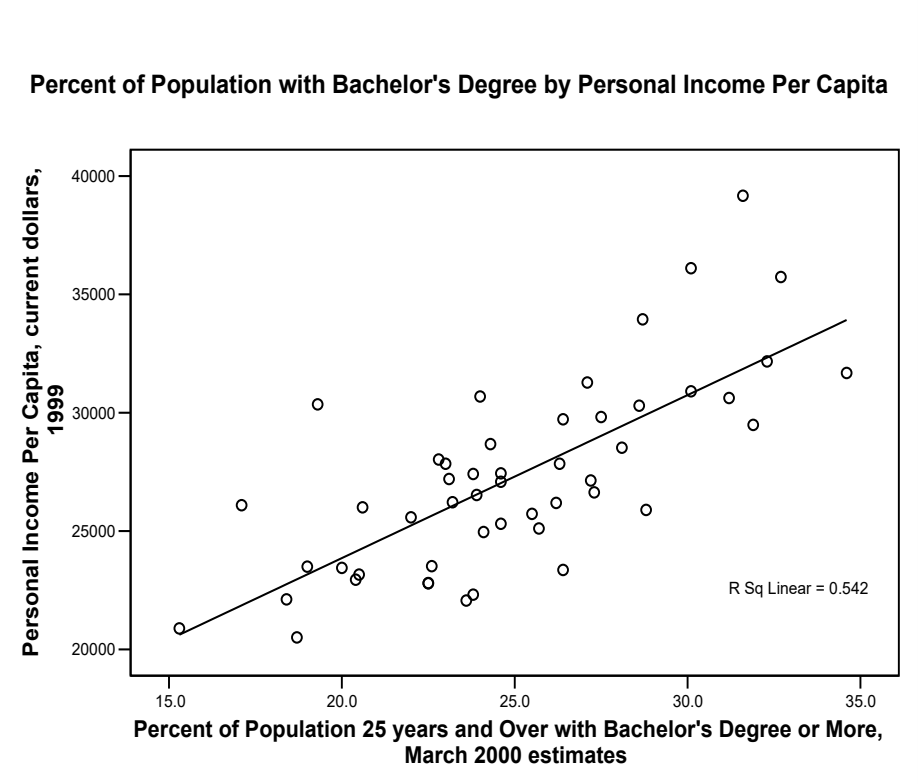


Scatter plot



- This is a linear relationship
- It is a positive relationship.
- As population with BA's increases so does the personal income per capita.

Regression Line

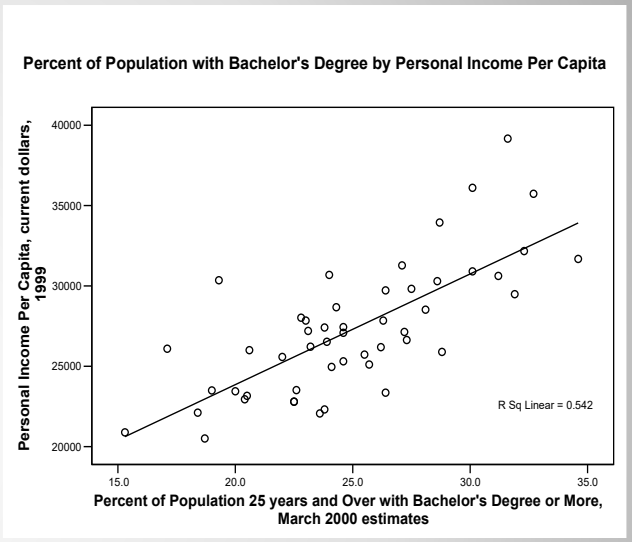


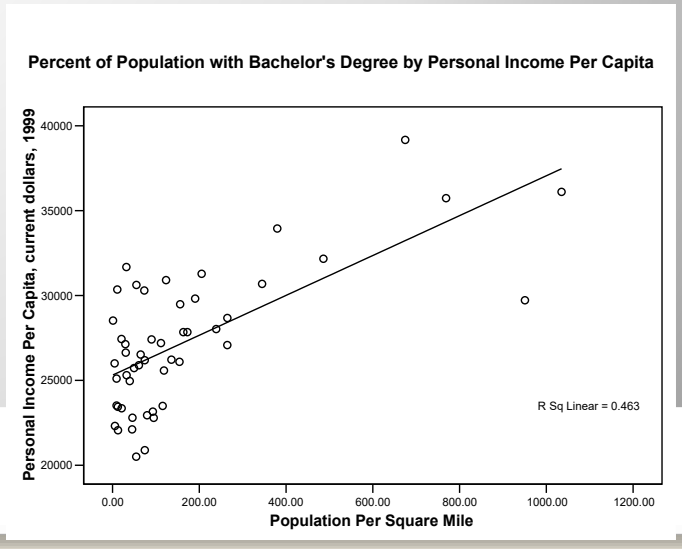
- Regression line is the best straight line description of the plotted points and use can use it to describe the association between the variables.
- If all the lines fall exactly on the line then the line is 0 and you have a perfect relationship.

- Regressions are still focuses on association, not causation.
- Association is a necessary prerequisite for inferring causation, but also:
  1. The independent variable must preceded the dependent variable in time.
  2. The two variables must be plausibly lined by a theory,
  3. Competing independent variables must be eliminated.

Regression Table

- The regression coefficient is not a good indicator for the strength of the relationship.
- Two scatter plots with very different dispersions could produce the same regression line.





- The regression coefficient is the slope of the regression line and tells you what the nature of the relationship between the variables is.
- How much change in the independent variables is associated with how much change in the dependent variable.
- The larger the regression coefficient the more change.

Regression coefficient

- To determine strength you look at how closely the dots are clustered around the line. The more tightly the cases are clustered, the stronger the relationship, while the more distant, the weaker.
- Pearson's r is given a range of -1 to + 1 with 0 being no linear relationship at all.