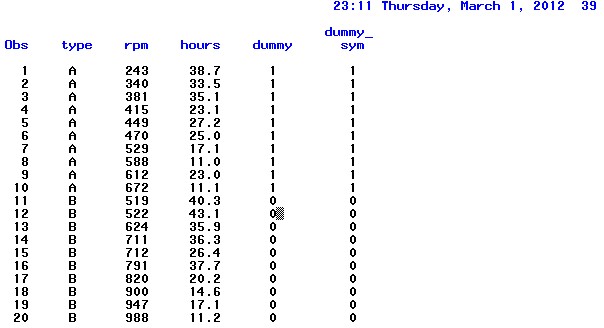
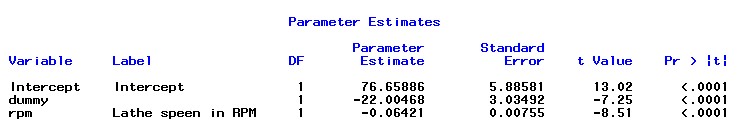
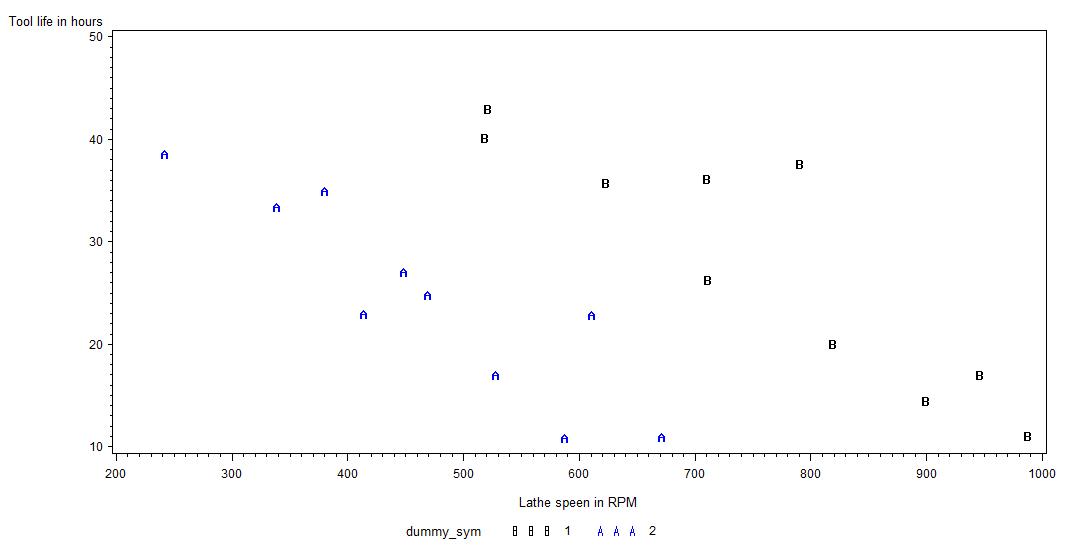
* 1. Create a SAS or R dataset and print it.

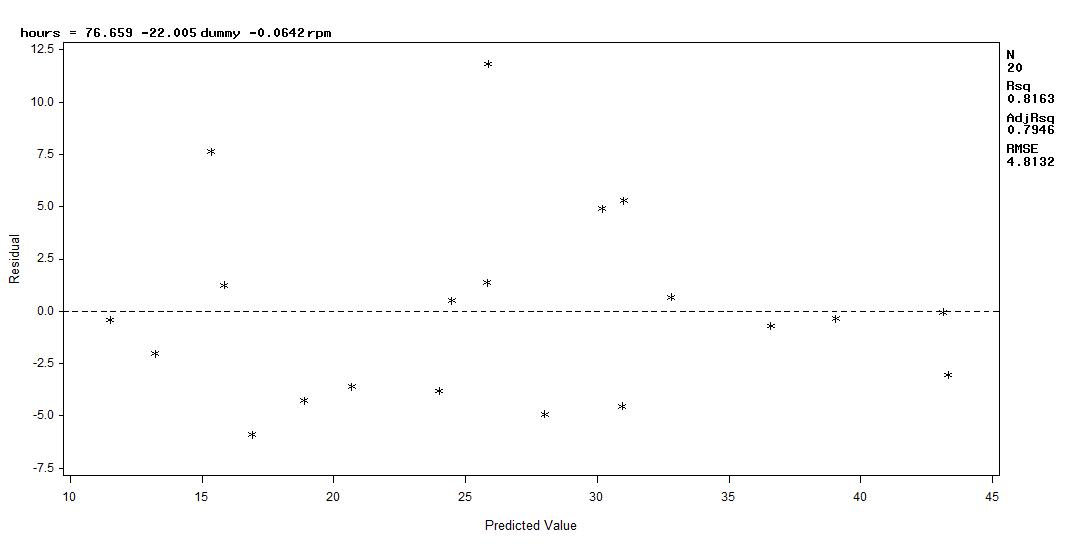
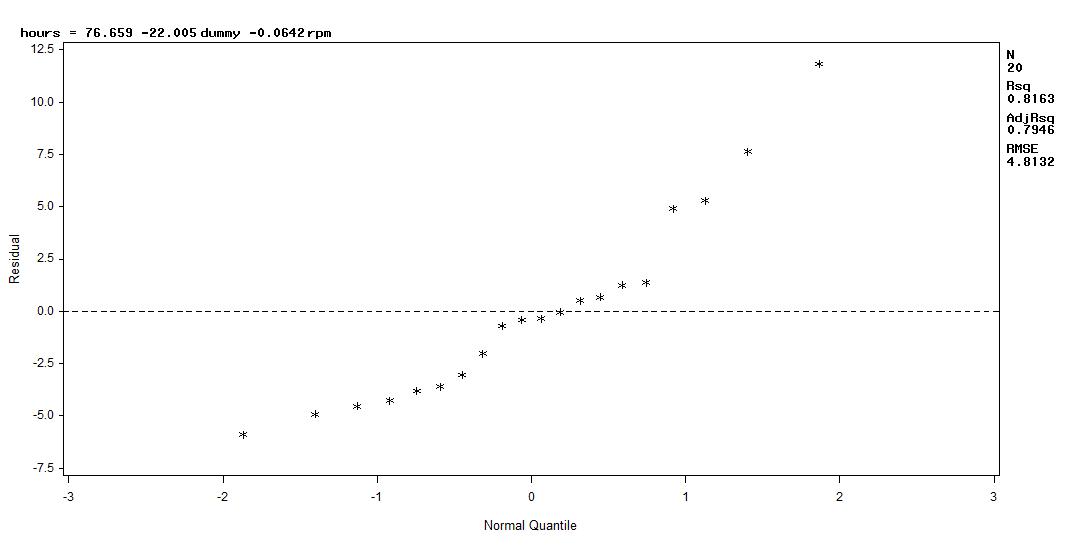


* 1. Create a regression model for predicting hours from type and rpm. Use a dummy variable for type.

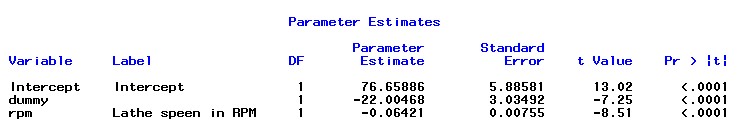


*Hours = 76.65886 - 22.00468dummy - 0.06421rpm*

* 1. Create a scatterplot of hours vs. rpm, using the symbol 'A' or 'B', depending on the type. 
  2. Form the residual plot and the normal plot of the residuals.



* 1. Does tool type have a significant effect on hours?



It`s significant since pr <.0001.

**Part B. SalarySurvey Dataset (65 pts.)**

* 1. Create a SAS or R dataset salaries. Include labels for the variables. Print the dataset.

The SAS System 00:49 Friday, March 2, 2012 3

Obs exper educ mgt salary

1 1 1 1 58418

2 1 3 0 48870

3 1 3 1 78731

4 1 2 0 47501

5 1 3 0 49539

6 2 2 1 87871

7 2 2 0 49560

8 2 1 0 44352

9 2 3 0 51341

10 3 2 0 51838

11 3 1 1 63045

12 3 2 1 89972

13 3 3 1 83358

14 4 1 0 48066

15 4 3 1 85307

16 4 3 0 55703

17 4 2 0 54242

18 5 2 0 55761

19 5 3 0 57580

20 5 1 1 67213

21 6 1 0 51935

22 6 3 1 89892

23 6 2 0 58262

24 6 2 1 96342

25 7 1 1 71477

26 8 2 0 62321

27 8 1 1 73271

28 8 3 1 93395

29 8 1 0 57037

30 10 1 0 60906

31 10 2 0 67116

32 10 3 1 97563

33 10 2 1 100114

34 11 2 1 106976

35 11 1 0 62565

36 12 2 0 71073

37 12 3 1 101764

38 13 1 0 67318

39 13 2 1 110849

40 14 2 0 75565

41 15 3 1 108134

42 16 2 1 117194

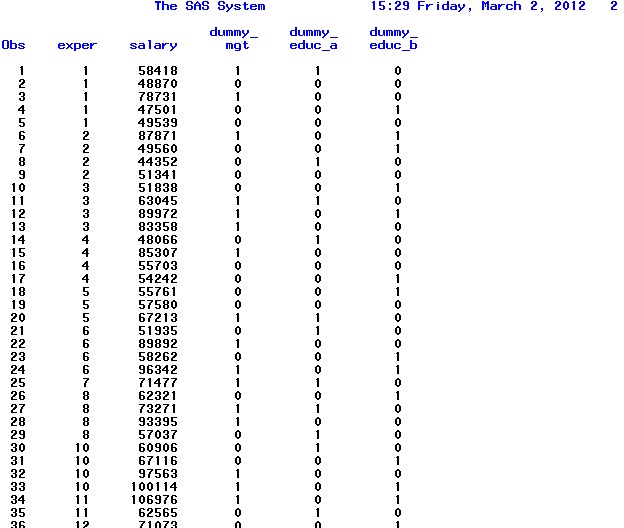
43 16 2 0 79308

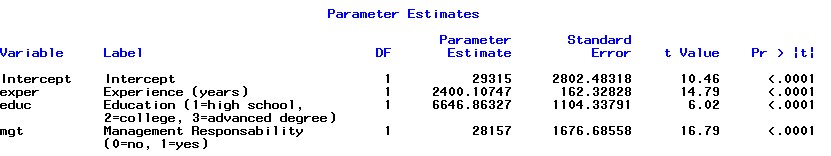
44 16 1 0 73603

45 17 2 0 80861

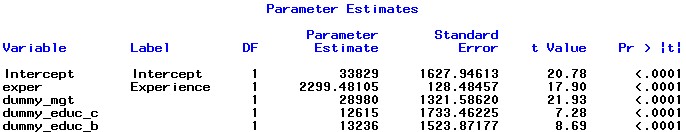
46 20 1 0 81447

* 1. Create and print a new dataset that uses dummy variables to represent educ and mgt. You should have two dummy variables for educ and one for mgt.



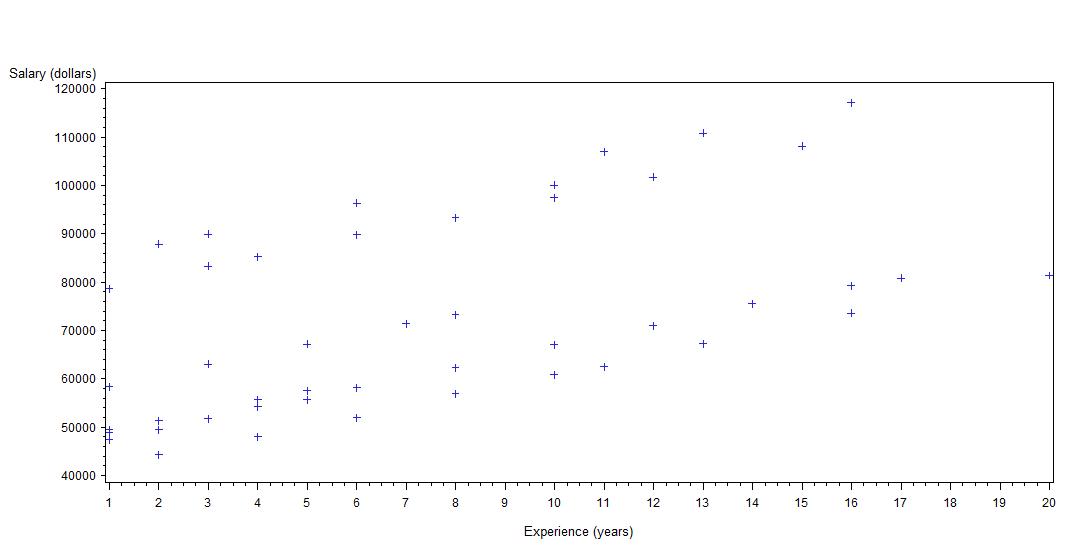
* 1. Create a regression model that predicts salary from exper, educ, and mgt.

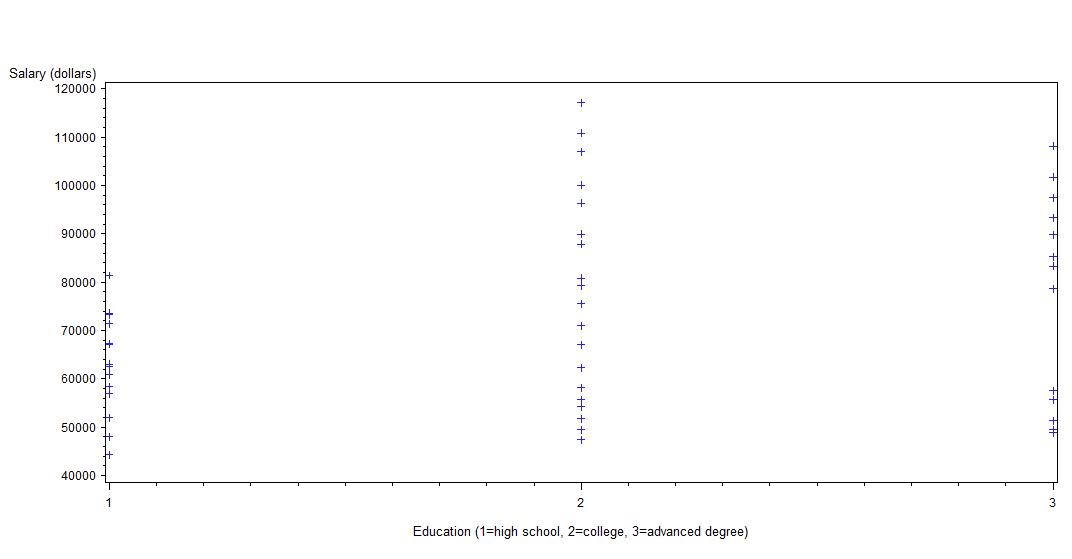
*Salary = 29315 + 2400.10747exprience + 6646.86327 education + 28157 management*

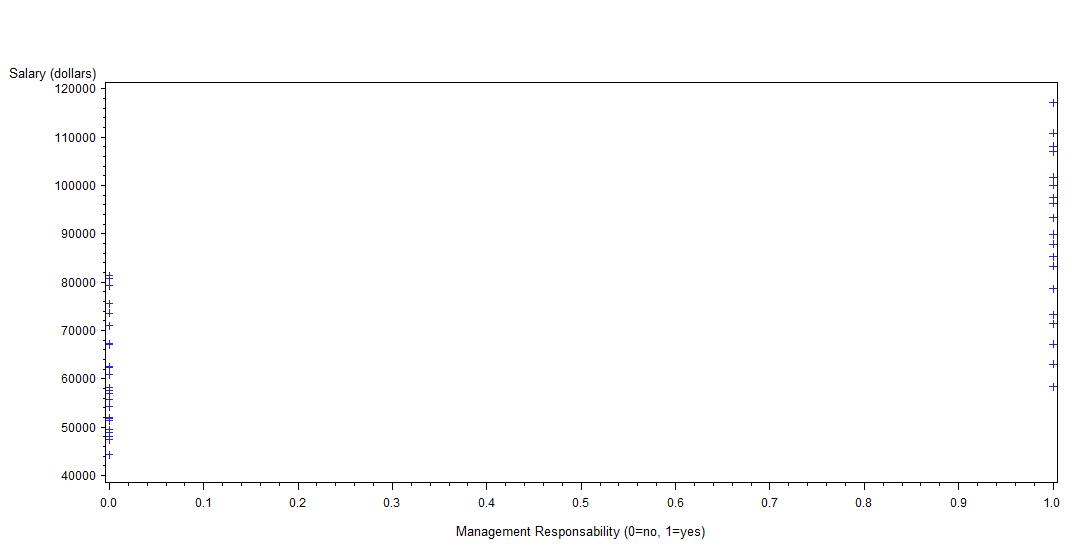
**

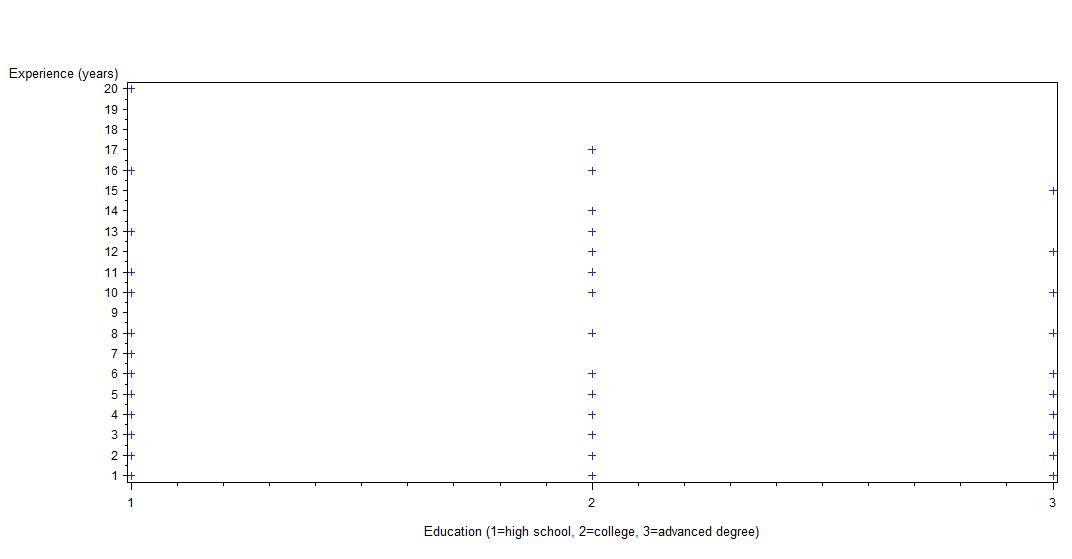
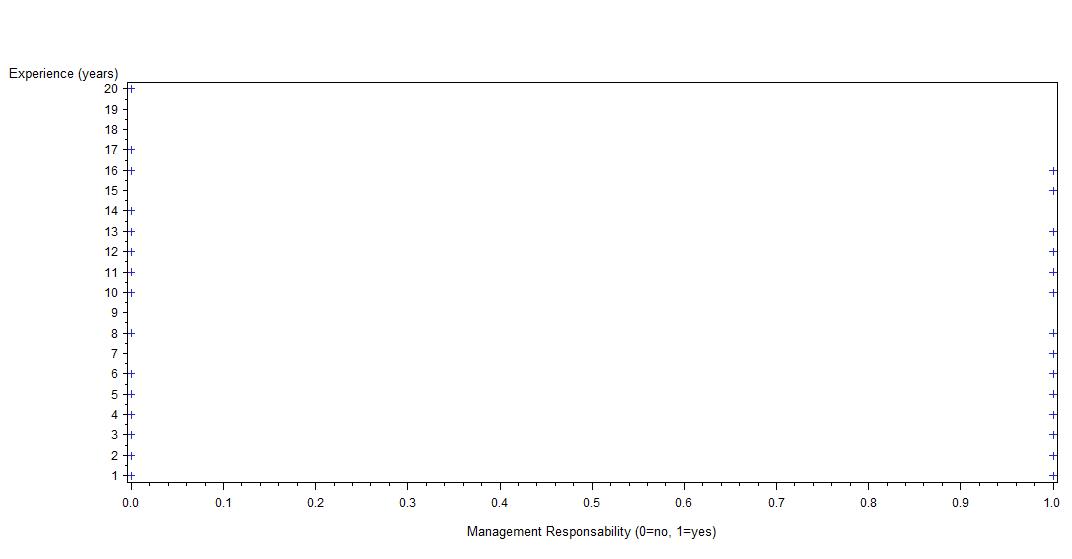
*Salary =33829 + 2299.5 exper + 13236 dummy\_edu\_b +12615 dummy\_edu\_c +28980 dummy\_mgt*

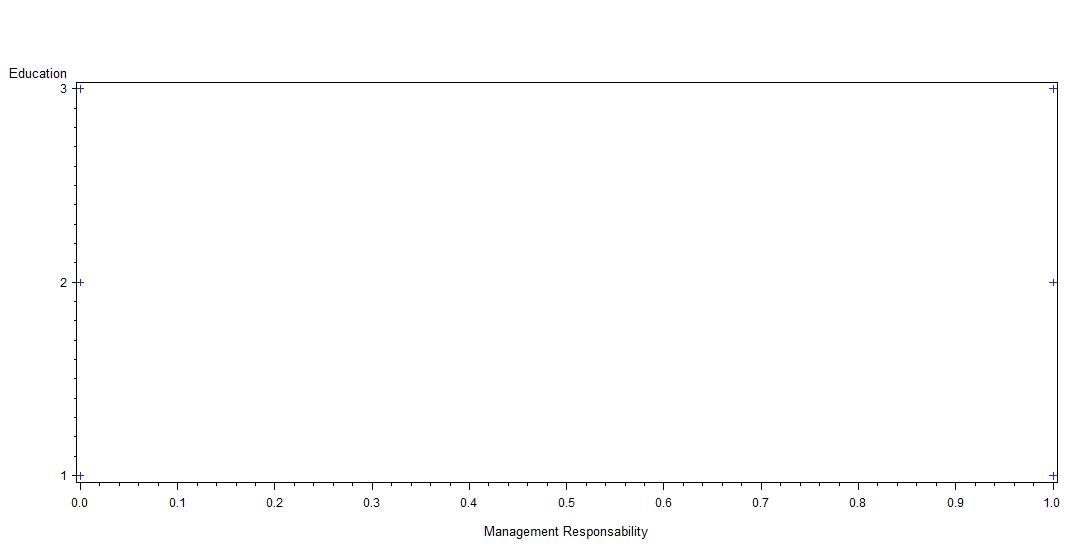
* 1. Create the six pairwise scatterplots?

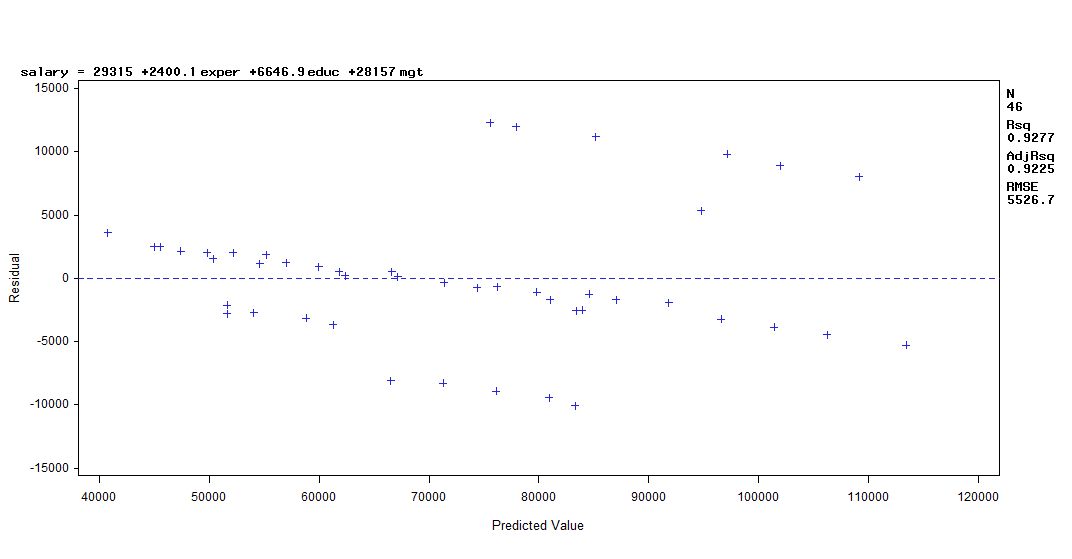
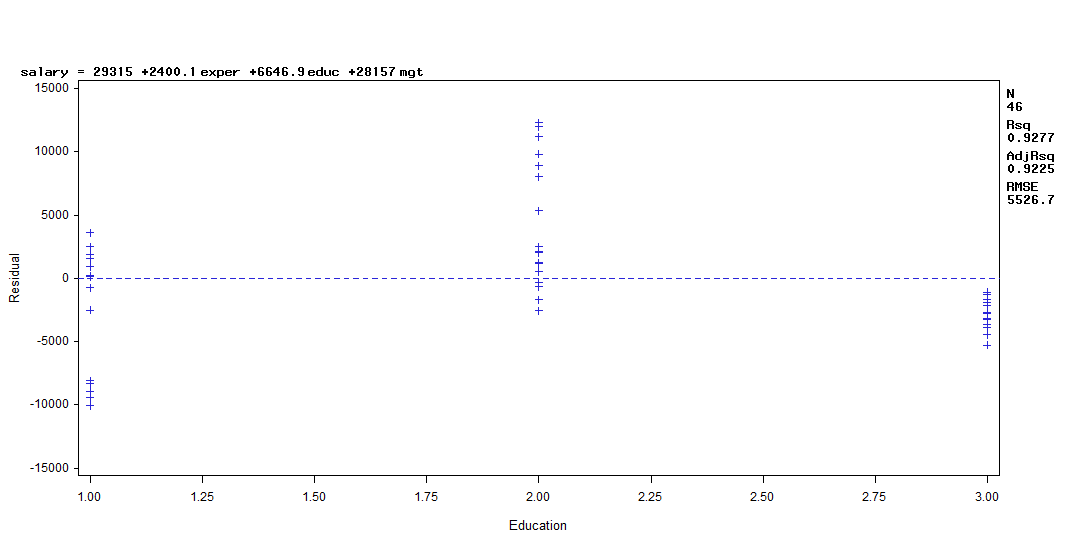


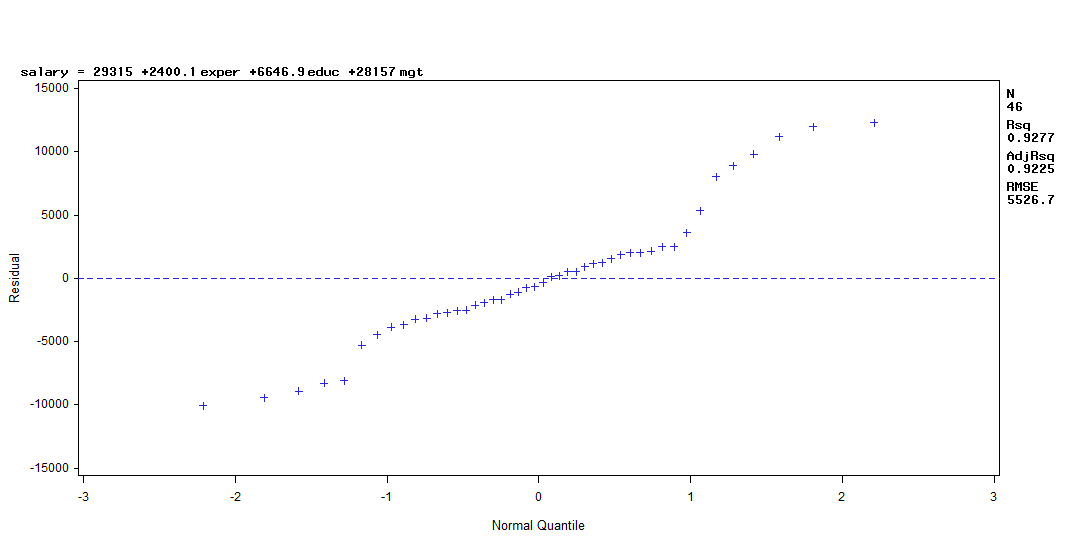


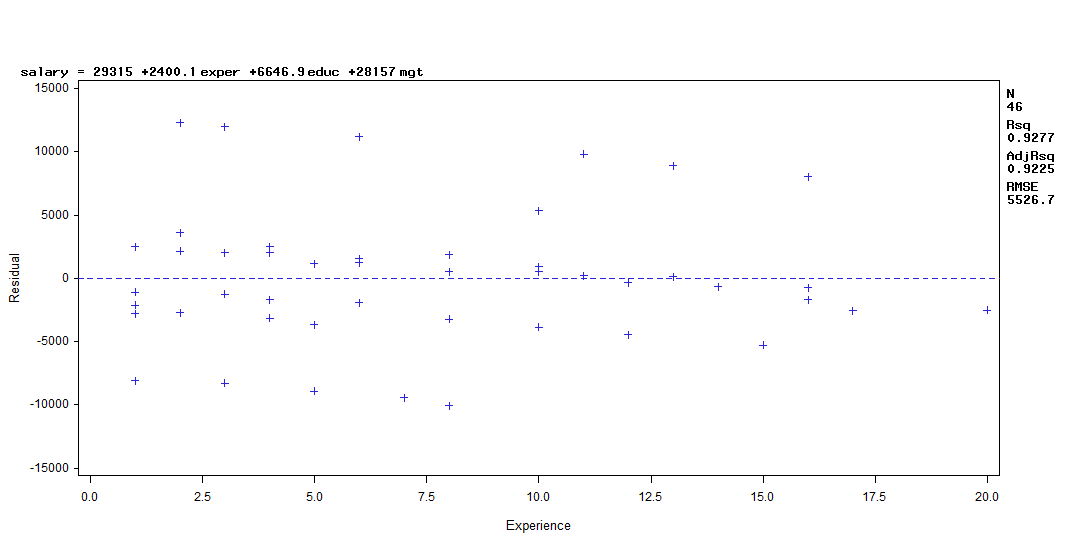


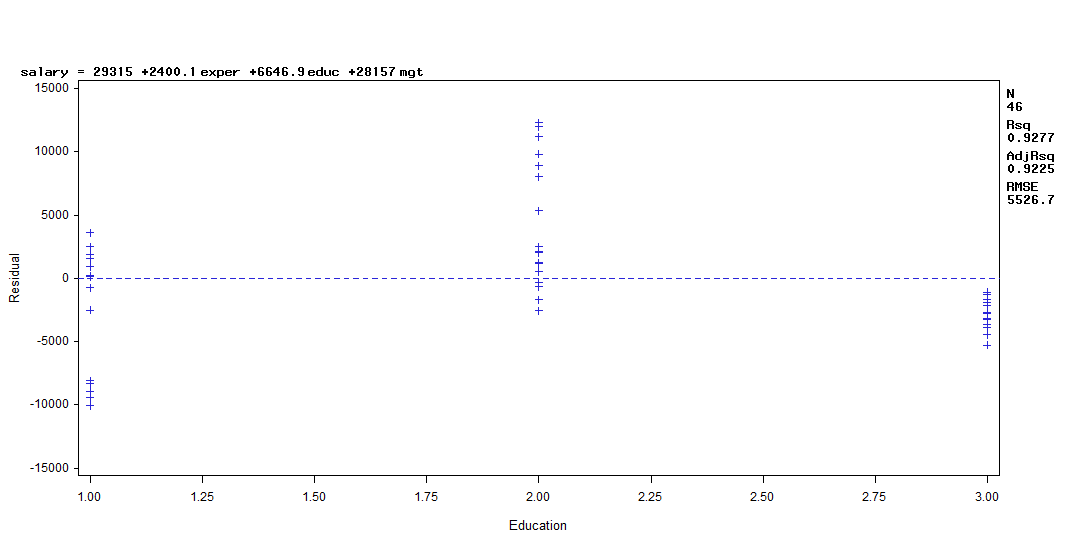


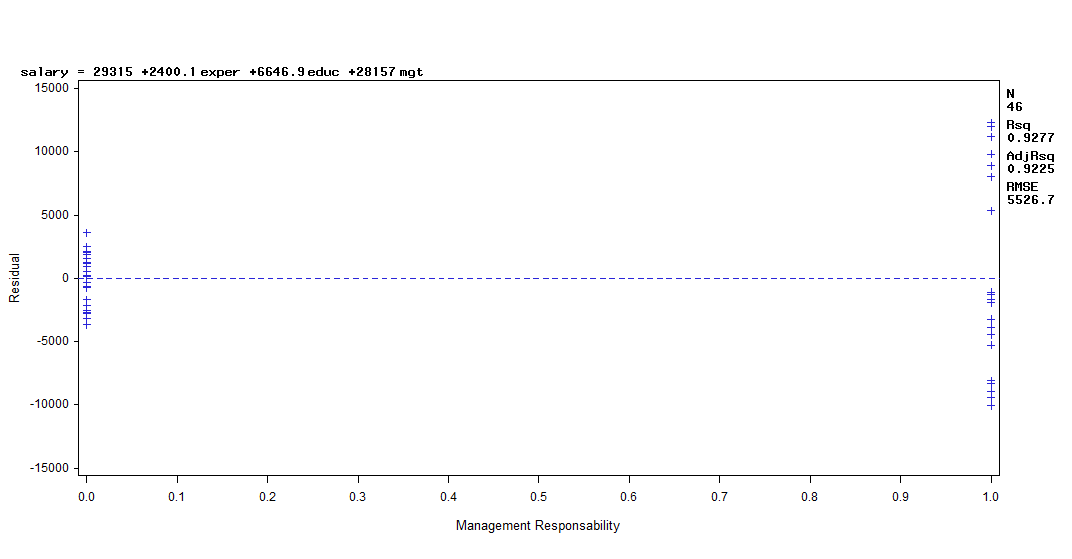


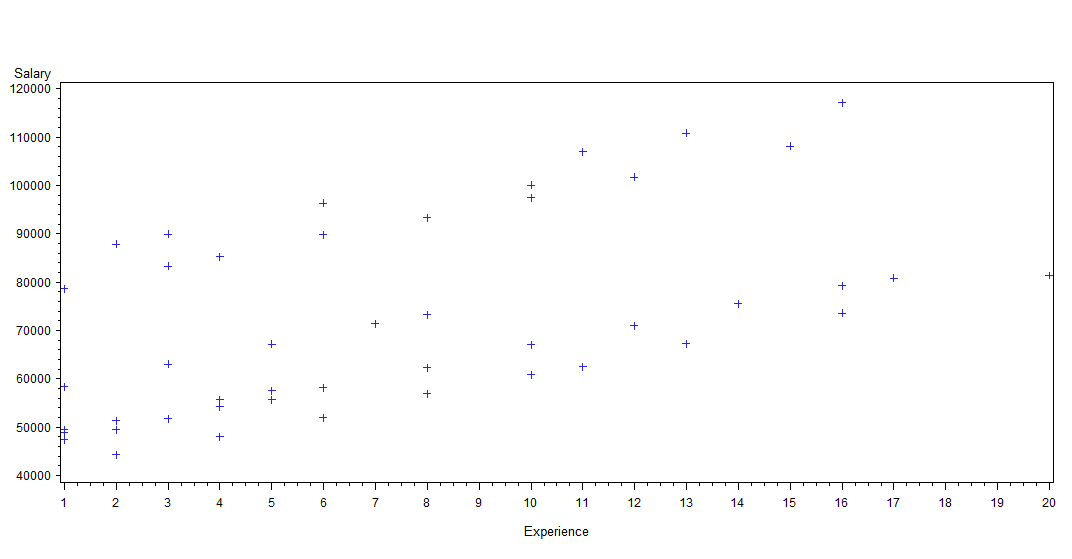
* 1. Create the residual and normal residual plots plot the residuals vs. predicted plots of the residuals. For values and the residuals vs. each independent variable. 











* 1. How much of an increase in salary is one additional year of experience likely to produce?

*Salary = 29315 + 2400.10747exprience + 6646.86327 education + 28157 management.*

*Almost 2400 dollars will be gained if you have one additional year experience.*

*In another model:*

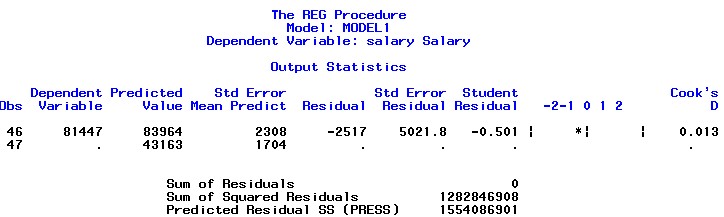
Salary=33829 + 2299.5 exper + 13236 dummy\_edu\_b +12615 dummy\_edu\_c +28980 dummy\_mgt

The figure is 2299.5.

* 1. How much higher is the predicted salary of a college graduate than the salary of a person with only a highschool degree?

*That will be the coefficient of dummy\_edu\_b, say, 13236*

* 1. For a high school graduate with 3 years of experience with no management responsabilities, what is the predicted salary?



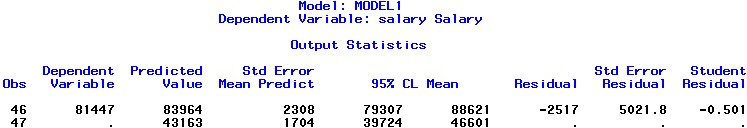
*As shown above by the 47th obs, predicted salary is 43163.*

* 1. Does a person with an advanced degree have a higher predicted salary than a college graduate for this dataset?

Salary=33829 + 2299.5 exper + 13236 dummy\_edu\_b +12615 dummy\_edu\_c +28980 dummy\_mgt

It`s not the case, since dummy\_edu\_c, which denotes AD education, has a lower coefficient than dummy\_edu\_b does.

10. Find 95% prediction interval for the person in Problem 8.

**

*It`s [39724, 46601]*