HOMEWORK: CORRELATION, SIMPLE REGRESSION
ELEMENTARY STATISTICS AND INFERENCE (STAT:1020; BOGNAR)

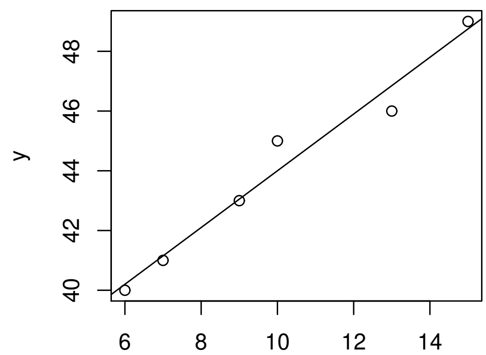
1. At a large company, the salaries (y, in thousands of dollars) and years of experience (x) of six randomly chosen engineers are

$x =$ years:	6	7	9	10	13	15
$y =$ salary:	40	41	43	45	46	49

(a) Find Pearsons sample correlation coefficient r.
$Cov(x, y) = 11.4$, $s_x = 3.464$, $s_y = 3.347$, $r = 0.983$.

(b) Determine the least squares regression line.
$\hat{y} = 34.5 + 0.95x$

(c) Carefully make a scatter-plot of the dataset and draw the regression line (place the explanatory variable x on the horizontal axis, and the response variable y on the vertical axis).



(d) On average, each extra year of experience yields how much extra pay?
$0.95 = \$950$

(e) What is the approximate average starting pay?
$34.5 = \$34500$

(f) Approximate the mean salary for engineers with 6 years of experience, i.e. approximate $\mu_{y|x=6}$.
$\hat{y} = 34.5 + 0.95(6) = 40.2 = \$40200$

(g) Find a 95% confidence interval for the population mean salary for engineers with 6 years of experience, i.e. find a 95% CI for $\mu_{y|x=6}$. Interpret the CI. *Hint: According to Minitab,* $\widehat{se}(\hat{y}) = 0.448$.
$\hat{y} \pm t_{\alpha/2,n-p}\widehat{se}(\hat{y}) = 40.2 \pm 2.776(0.448) = (38.96, 41.44)$

(h) Is there a significant linear relationship between years of experience and salary? *Hint: According to Minitab,* $\widehat{se}(\hat{\beta}_1) = 0.0878$. *You must state* H_0 *and* H_a *(use* $\alpha = 0.05$*), find the test statistic and critical value, plot the rejection region, and state your decision and final conclusion.*
$H_0 : \beta_1 = 0$, $H_a : \beta_1 \neq 0$, $t^* = (\hat{\beta}_1 - \beta_1)/\widehat{se}(\hat{\beta}_1) = 10.82$, $t_{\alpha/2,n-p} = 2.776$, reject H_0, evidence that $\beta_1 \neq 0$, significant linear relationship between years of experience and salary.

(i) Approximate the $p-$value for the test in (1h). Based upon your $p-$value, is there a significant linear relationship between years of experience and salary? Why?
$p - value = 2P(t_{(4)} > 10.82) < 0.001$, significant linear relationship between years of experience and salary since $p - value < \alpha$

(j) Find a 95% confidence interval for β_1. Based upon your CI, is there a significant linear relationship between years of experience and salary? Why? *Hint: According to Minitab,* $\widehat{se}(\hat{\beta}_1) = 0.0878$.
$\hat{\beta}_1 \pm t_{\alpha/2,n-p}\widehat{se}(\hat{\beta}_1) = 0.95 \pm 2.776(0.0878) = (0.706, 1.194)$, significant linear relationship since CI excludes 0

(k) Find a 95% confidence interval for the (population) mean starting salary, i.e. find a 95% CI for $\beta_0 = \mu_{y|x=0}$. *Hint: According to Minitab,* $\widehat{se}(\hat{\beta}_0) = 0.9208$.
$\hat{\beta}_0 \pm t_{\alpha/2,n-p}\widehat{se}(\hat{\beta}_0) = 34.5 \pm 2.776(0.9208) = (31.94, 37.06)$

(l) In reference to question (1k), is the population mean starting salary significantly different than 40 (i.e. $40,000)? Why?
Yes, since the CI excludes 40