Part A: Theory - Using a **few concise** sentences answer the following questions:

1. Explain carefully what an estimator is, and the three desirable properties of estimators. Explain whether an estimate or an estimator is a random variable.

   While an estimator is a mathematical rule or formula that tells us how to combine data to get an estimate for some unknown population parameter, an estimate is just a numerical value derived from using sample data in an estimator. The estimator is a random variable dependent on the sample, but a number is the outcome of the random variable.

   The three desirable properties of estimators are unbiased, efficient and consistent. An estimator is said to be "unbiased" if the expected value of the estimator is equal to the true population parameter being estimated, that is $E(\hat{\theta}) = \theta$. An estimator is said to be consistent if the difference between the estimator and the true parameter grows smaller as the sample size grows larger. This is usually seen as the variance growing smaller as the sample size gets larger. A relatively efficient estimator is the estimator with the smallest variance when compared to other unbiased estimators of the same unknown population parameter. We do not compare estimators for different population parameters to find the relatively efficient estimator.

2. State the Central Limit theorem. Why is it important in econometrics?

   If Y1, . . . ,YN are independent and identically distributed random variables with mean μ and variance σ^2, and $\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$, then $\bar{Y}$ has a probability distribution that $Z_n = \frac{\bar{Y}-\mu}{\sigma/\sqrt{n}}$ converges to the standard normal $N(0.1)$ as N→∞.
   The theorem says that the sample average of N independent random variables from **any** probability distribution will have an approximate standard normal distribution, after standardizing, if the sample size is sufficiently large.
   This is extremely powerful as it allows us to make confidence intervals and do hypothesis tests using sample statistics.

3. Explain what a Type 1 and Type II error would be, using an example to illustrate.

   A type I error is to reject H_0 when H_0 is true, while a type II error is to not reject H_0 when H_0 is false. For example, in Part B below, question 1, part f, you are asked to test the hypothesis that the mean hourly wage of workers in Alberta and Ontario are equal at a 5% significance level. A type one error would happen if you were to conclude that the mean hourly wages were not equal, and in fact the mean hourly wages were equal. A type II error would happen if you decided that the mean hourly wages were equal, and in fact the mean hourly wages were different.

4. What are the six assumptions in the linear regression model? (Include the optional assumption.)

SR1 $y = \beta_1 + \beta_2 x + e$

SR2. $E(e_t) = 0 \Leftrightarrow E(y_t) = \beta_1 + \beta_2 x_t$

SR3. $\text{var}(e_t) = \sigma^2 = \text{var}(y_t)$

SR4. $\text{cov}(e_i, e_j) = \text{cov}(y_i, y_j) = 0$

SR5. The variable x_t is not random and must take at least two different values.

SR6. (*optional*) The values of *e* are *normally distributed* about their mean

$e_t \sim N(0, \sigma^2) \Leftrightarrow y_t \sim N[(\beta_1 + \beta_2 x_t), \sigma^2]$

5. State the Gauss-Markov theorem. Explain what "BLUE" stands for and how it relates to the Gauss-Markov theorem. Why is the Gauss-Markov theorem important?

Gauss-Markov Theorem: Under the assumptions SR1-SR5 of the linear regression model the estimators b_1 and b_2 have the *smallest variance of all linear and unbiased estimators* of β_1 and β_2. They are the Best Linear Unbiased Estimators (BLUE) of β_1 and β_2.

BEST- <u>smallest variance</u> when compared to similar <u>estimators that are linear and unbiased.</u>
LINEAR- a class of estimators where it is a linear function of the sample observations, that is where we can write it as $Y_i = \sum a_i x_i$, where a_i are constants.
UNBIASED – the <u>expected value</u> of the parameter (or on average) is equal to the <u>true</u> unknown parameter value.
ESTIMATORS- a general mathematical formula for obtaining estimates using a data sample.

The Gauss-Markov theorem is important because it tells us that the least squares estimators have the smallest variance, so they are the best estimators in the class of linear and unbiased estimators.

Part B: Problems

1. The Excel file posted with this assignment contains sample data, from 2004, on the hourly wage rate of a sample of individuals, along with what province the person lives in. **<u>You can use Excel or Stata for this question. Do not print out the data!!</u>**

   a. Calculate the mean hourly wage and variance for workers over all the data, the Canadian average. Now, calculate the mean hourly wage and variance for unionized workers for each province. Using a table summarize the mean hourly wage, the variance, and the sample size for each province and for all the data (the Canadian average).

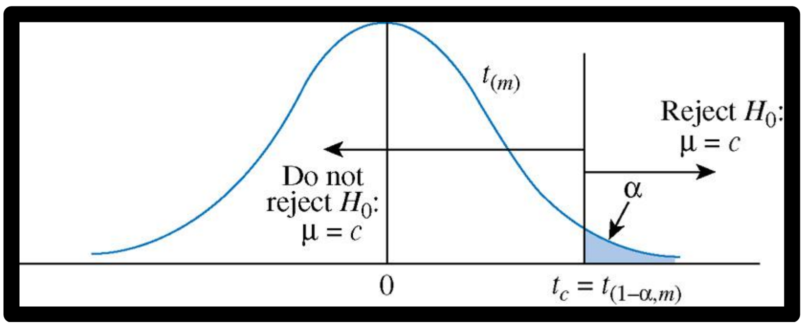
	Data		
Province	**Average of Hourly Wage Rate**	**Variance of Hourly Wage Rate**	**Count of Participants**
Newfoundland 10	$17.98	99.7134	636
P.E.I. 11	$17.56	81.0159	486
Nova Scotia 12	$17.46	99.5078	1109
New Brunswick13	$17.72	99.6493	1031
Quebec 24	$19.70	106.5378	3514
Ontario 35	$21.91	135.9888	5325
Manitoba 46	$18.54	74.9743	1244
Sask. 47	$19.38	108.0025	1211
Alberta 48	$21.68	157.9969	1609
B.C.59	$21.30	130.8565	1414
Canada	$20.20	120.8189	17579

b. Using the data for Saskatchewan, test the hypothesis that the mean hourly wage for workers is greater than $19.85 at the 5% significance level. Be sure to state the null and alternative hypotheses, give the test statistic and its distribution, indicate the rejection region, including a sketch, state your conclusion, and calculate the p-value for the test.

$H_0 : \mu = 19.85$

$H_1 : \mu > 19.85$

α=0.05, since we have $\widehat{\sigma^2} = s^2 = 108.0025$,
this is a Student *t*-distribution with df = n-1 = 1210.



Test statistic: $t = \frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{n}} = \frac{19.38-19.85}{10.39/34.80} = -1.5738$
Rejection region: $t > t_{0.95,1210} = 1.645$ from Table 2, Appendix E,
(Rejection region: $t > t_{0.95,1210} = 1.6461$ from Excel)
Since -1.5738 < 1.645 **Do Not Reject H_0**.

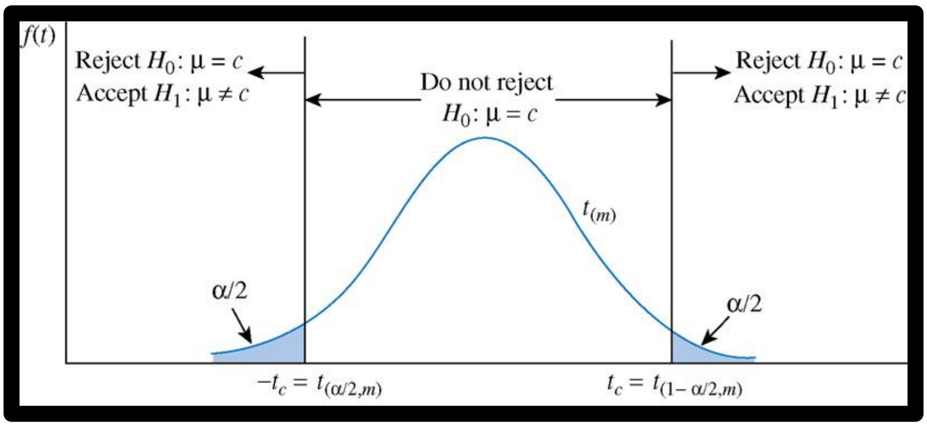
Conclusion: There is not enough evidence in this sample to reject the hypothesis that the mean hourly wage in Saskatchewan is $19.85.

p-value: Using the table at df = ∞, -1.645< - 1.5738 < -1.282,
so 0.1> $P(t_{1210} < -1.5738) > 0.05$. The $P(t_{1210} < -1.5738) = 0.0579$ (from Excel)
The $P(t_{1210} < -1.5738) = 0.0579 > 0.05$, so do not reject the null hypothesis.

c. Using the data for British Columbia, test the hypothesis that the mean hourly wage for workers is equal to \$20.75 at the 5% significance level. Be sure to state the null and alternative hypotheses, give the test statistic and its distribution, indicate the rejection region, including a sketch, state your conclusion, and calculate the p-value for the test.

$H_0 : \mu = 20.75$
$H_1 : \mu \neq 20.75$

$\alpha = 0.05$, since we have $\widehat{\sigma^2} = s^2 = 130.8565$, this is a Student *t*-distribution with df = n-1 = 1413.



Rejection region: $t > t_{0.975,1413} = 1.96$ from Table 2, Appendix E, or if $t < -t_{0.975,1413} = -1.96$
(Rejection region: $t > t_{0.975,1413} = 1.9616$ from Excel or if t< -1.9616)

Test statistic: $t = \frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{n}} = \frac{21.30-20.75}{11.44/37.60} = 1.8080$
Since -1.9616 < 1.8080 <1.9616 **Do Not Reject H₀**.

There is not enough evidence to reject the mean hourly wage for workers is equal to \$20.75 at the 5% significance level.

The P(t_{1413}> 1.8080) = 0.0354 (from Excel), since this is a two-tail test we must double this value, or 0.0708. 0.0708 > 0.05, so do not reject the null hypothesis.

d. Comment on your results from parts b) & c).
In both hypothesis tests we were not able to reject the null hypothesis. The test in part b) was a single-tail test with the rejection region in the right tail, while the test in part c) was a two-tail test with the rejection region in both tails.

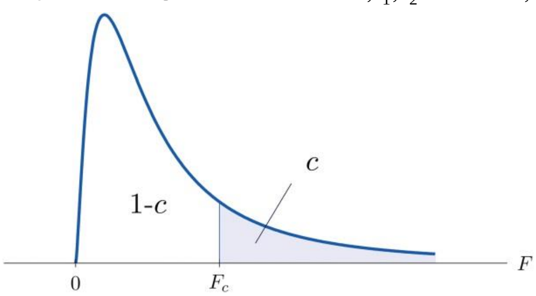
e. Assume that the hourly wage rates for workers in Alberta and Ontario are normally distributed with means μ_A and μ_O, and variances σ_A^2 and σ_O^2 respectively. Assume the hourly wage rates in Alberta and Ontario are independent of each other. Test the hypothesis that the variances σ_A^2 and σ_O^2 are equal, at a 5% significance level. Be sure to state the null and alternative hypotheses, give the test statistic and its distribution, indicate the rejection region, including a sketch, state your conclusion, and calculate the p-value for the test.

Let Alberta be population 1 and Ontario be population 2.

$$H_0 : \sigma_1^2 / \sigma_2^2 = 1$$

$$H_1 : \sigma_1^2 / \sigma_2^2 \neq 1$$

Rejection region: $F > F_{\alpha/2,\nu_1,\nu_2} = F_{.025,1608,5324} = 1.0812 \approx 1$



F = s_1^2 / s_2^2 = 157.9969/135.9888 =1.16,

$F > F_{\alpha/2,\nu_1,\nu_2} = F_{.05,1608,5324} = 1.0812 < 1.16$, **Reject the null hypothesis**

There is enough evidence to reject the hypothesis that the variances in hourly wage rates in 2004 are equal between Alberta and Ontario at the 95% confidence level. The evidence supports the hypothesis that the variance for Alberta is greater than the variance for Quebec.

	Sample 1	Sample 2	F Stat	1.16
Sample variance	157.9969	135.9888	**P(F<=f) one-tail**	**0.0001**
Sample size	1609	5325	**f Critical one-tail**	**1.0677**
Alpha	0.05		**P(F<=f) two-tail**	**0.0002**
			f Critical two-tail	**0.9234**
				1.0812

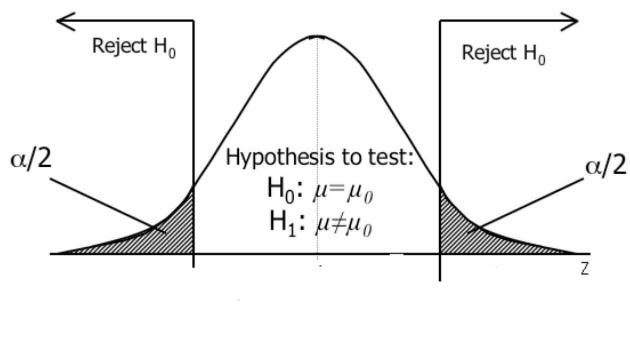
f. Test the hypothesis that the mean hourly wage of workers in Alberta and Ontario are equal at a 5% significance level. Use your conclusion in part e) to do the appropriate version of the test.

$$H_0 : (\mu_1 - \mu_2) = 0$$

$$H_1 : (\mu_1 - \mu_2) \neq 0 \qquad \alpha/2 = 0.025$$

Degrees of freedom: $\nu = \dfrac{(s_1^2/n_1 + s_2^2/n_2)^2}{\dfrac{(s_1^2/n_1)^2}{n_1-1} + \dfrac{(s_2^2/n_2)^2}{n_2-1}} = 2502.03$

Reject H_o if $t < -t_{0.025,2502} = -1.9609$ or if $t > t_{0.025,2502} = 1.9609$



$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}} = \frac{(21.68 - 21.91) - 0}{\sqrt{\frac{157.9969}{1609} + \frac{135.9888}{5325}}} = -0.6539$$

$-1.9609 < -0.6539 < 1.9609$ **Do Not Reject H_o**
There is insufficient evidence to reject the hypothesis that the mean hourly wage in Alberta and Ontario in 2004 are equal at the 5% significance level.

2. Consider the following five observations. You are to do all the parts of this exercise using only a calculator, or in an Excel spreadsheet.

y	x	$x-\bar{x}$	$(x-\bar{x})^2$	$y-\bar{y}$	$(x-\bar{x})(y-\bar{y})$
13.5	4	2	4	5.5	11
11.5	3	1	1	3.5	3.5
8	2	0	0	0	0
5.5	1	-1	1	-2.5	2.5
1.5	0	-2	4	-6.5	13
$\sum y_i =$	$\sum x_i =$	$\sum(x-\bar{x}) =$	$\sum(x-\bar{x})^2 =$	$\sum(y-\bar{y}) =$	$\sum(x-\bar{x})(y-\bar{y}) =$
40	10	0	10	0	30

a. Fill in the table. Put the sums in the last row. What are the sample means $\bar{x}$ and $\bar{y}$?

$\bar{x} = {\sum x_i}/{n} = \frac{10}{5} = 2$, $\bar{y} = {\sum y_i}/{n} = \frac{40}{5} = 8$

b. Calculate b_1 and b_2 and state their interpretation.

$b_2 = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2} = \frac{30}{10} = 3,\quad b_1 = \bar{Y} - b_2\bar{X} = 8 - (3*2) = 2$

$$y_i = 2 + 3x_i$$

The intercept is 2. The average value of *y*, when *x* is zero is 2. This interpretation is valid as we have data for value both larger, smaller and equal to when *x* is zero. The slope is 3. On average, a one unit increase in *x* leads to a three unit increase in *y*.

c. Compute $\sum_{i=1}^{5} x_i^2$, $\sum_{i=1}^{5} x_i y_i$. Use the values to show that $\sum(x-\bar{x})^2 = \sum x_i^2 - n\bar{x}^2$ and $\sum(x-\bar{x})(y-\bar{y}) = \sum x_i y_i - n\bar{x}\bar{y}$.

$$\sum_{i=1}^{5} x_i^2 = 0^2 + 1^2 + 2^2 + 3^2 + 4^2 = 30, \sum_{i=1}^{5} x_i y_i = 54 + 34.5 + 16 + 5.5 + 0 = 110$$

$$\sum (x-\bar{x})^2 = 10, \sum x_i^2 - n\bar{x}^2 = 30 - (5 * 2^2) = 10,$$

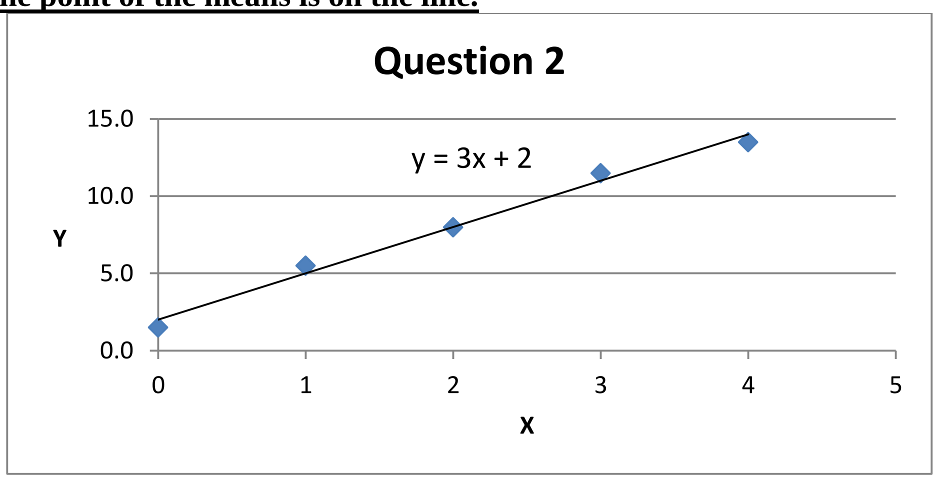
$$\sum(x-\bar{x})(y-\bar{y}) = 30, \quad \sum x_i y_i - n\bar{x}\bar{y} = 110 - (5 * 2 * 8) = 30.$$

d. Use the least squares estimates from part b to compute the fitted values of *y* and complete the remainder of the table below. Put the sums in the last row.

y	x	$\hat{y}_i$	$\hat{e}_i$	$\hat{e}_i^2$	$x_i\hat{e}_i$
13.5	4	14	0.5	0.25	2
11.5	3	11	-0.5	0.25	-1.5
8	2	8	0	0	0
5.5	1	5	-0.5	0.25	-0.5
1.5	0	2	0.5	0.25	0
$\sum y_i =$	$\sum x_i =$	$\sum \hat{y}_i =$	$\sum \hat{e}_i =$	$\sum \hat{e}_i^2 =$	$\sum x_i\hat{e}_i =$
40	10	40	0	1	0

e. On graph paper (or a grid you draw), plot the points and sketch the fitted regression line $\hat{y}_i = b_1 + b_2 x_i$. Locate and plot the point of the means, $(\bar{x}, \bar{y})$. Does your fitted line pass through that point?

<u>Yes, the point of the means is on the line.</u>



f. Show that for the point of the means, $(\bar{x}, \bar{y})$, $\bar{y} = b_1 + b_2\bar{x}$.

$\bar{y} = b_1 + b_2\bar{x}, \qquad 8 = 2 + (3 * 2) = 8$

g. Show that the mean of the predicted $\hat{y}$ is equal to $\bar{y}$, the mean of the y.

$$\bar{y} = {\sum y_i}/{n} = \frac{40}{5} = 8, \qquad \bar{\hat{y}} = {\sum \hat{y}_i}/{n} = \frac{40}{5} = 8$$

h. Compute $\hat{\sigma}^2$.

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{5} e_i^2}{n-2} = \frac{1}{3}$$

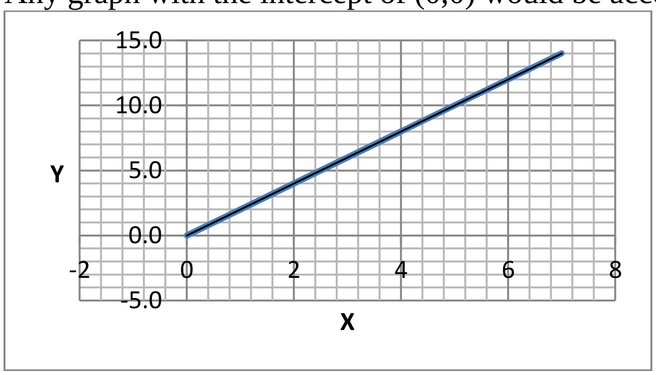
i. Compute the $\widehat{var(b_2)}$.

$$\widehat{var(b_2)} = \frac{\hat{\sigma}^2}{\sum(x-\bar{x})^2} = \frac{1/3}{10} = \frac{1}{30}$$

3. The simple linear regression model is defined as $y = \beta_1 + \beta_2 x + e$.

a. Suppose that we know $\beta_1 = 0$. What is the algebraic model now?
<u>$\boldsymbol{y = \beta_2 x + e}$</u>

b. Suppose that we know $\beta_1 = 0$. What does the model look like if we draw the graph?
Any graph with the intercept of (0,0) would be acceptable.



c. If $\beta_1 = 0$, the least squares "sum of squares" function would become

$$S(\beta_2) = \sum_{i=1}^{n} (y_i - \beta_2 x_i)^2.$$

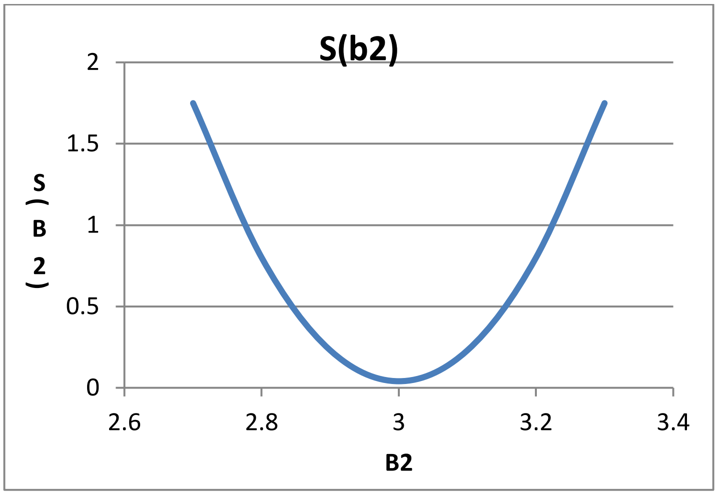
Using the data points below,

x	-2	-1	0	1	2	3
y	-5.9	-3.1	0	3	5.9	9.1

plot the value of the sum of squares function, $S(\beta_2)$, against enough values of β_2 (guess some values for β_2) for you to locate the approximate minimum of $S(\beta_2)$. What is the significance of the value of β_2 that minimizes $S(\beta_2)$?

Table of some values,

b2	2.7	2.8	2.9	3	3.1	3.2
S(b2)	1.75	0.8	0.23	0.04	0.23	0.8



The value of β_2 that minimizes $S(\beta_2)$ would be our estimate for b_2, the slope of the regression line that has the smallest sum of errors squared.

d. Using the derivation found in the Chapter 2 slides, number 62 to 65, (calculus) show that the formula for the least squares estimator of β_2 in this model is $b_2 = \sum x_i y_i / \sum x_i^2$. Use this formula to compute b_2 and compare this result to the value you obtained in part c).

$$S(\beta_2) = \sum (y_i - \beta_2 x_i)^2,$$

$$\frac{\partial S}{\partial \beta_2} = -2\sum x_i(y_i - \beta_2 x_i) = -2\sum x_i y_i - (-2\sum \beta_2 x_i^2$$

$$= 2\beta_2 \sum x_i^2 - 2\sum x_i y_i,$$

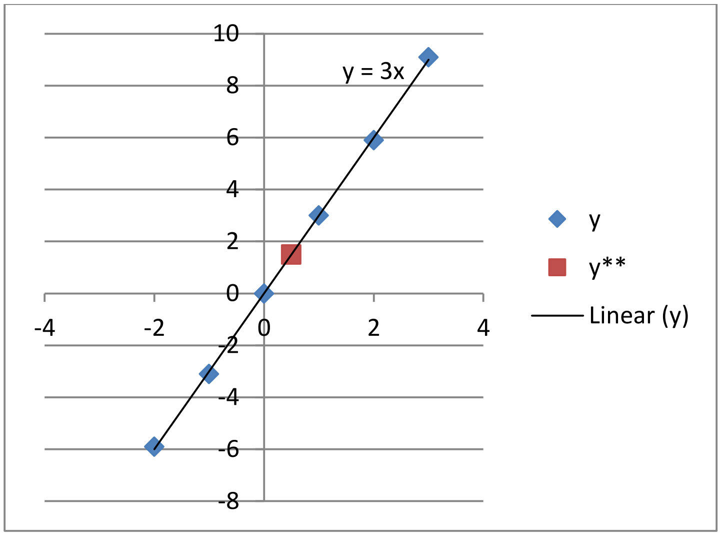
$$0 = 2\beta_2 \sum x_i^2 - 2\sum x_i y_i,$$

$$2\beta_2 \sum x_i^2 = 2\sum x_i y_i,$$

$$b_2 = \frac{\sum x_i y_i}{\sum x_i^2} \qquad b_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{57}{19} = 3$$

The numerical value agrees with the value in the graph that minimizes $S(\beta_2)$.

e. Now, plot the data and the estimated regression line using the value for b_2 from part d). Calculate and graph the point of the means, $(\bar{x}, \bar{y})$. What do you observe?
In this case the point of the means, (0.5, 1.5), is on the line.



f. Using the estimates obtained with the formula from part d), find the least squares residuals, $\hat{e}_i = y_i - b_2 x_i$. Now, what does the sum of the errors equal ($\sum \hat{e}_i =$)?

x	-2	-1	0	1	2	3
y	-5.9	-3.1	0	3	5.9	9.1
$\hat{e}_i = y_i - b_2 x_i$	0.1	-0.1	0	0	-0.1	0.1
$x_i \hat{e}_i$	-0.2	0.1	0	0	-0.2	0.3

$$\sum \hat{e}_i = 0.1 + (-0.1) + (-0.1) + 0.1 = 0$$

g. Calculate $\sum x_i \hat{e}_i$. $\sum x_i \hat{e}_i = -0.2 + 0.1 + (-0.2) + 0.3 = 0$

4. How much does experience affect wage rates? The data file *LandruAB2007.dat* contains 1719 observations on hourly wage rates, years of experience, and other variables from 2007. *The data definitions are on the next page.* **Submit your Stata log file in your assignment.**

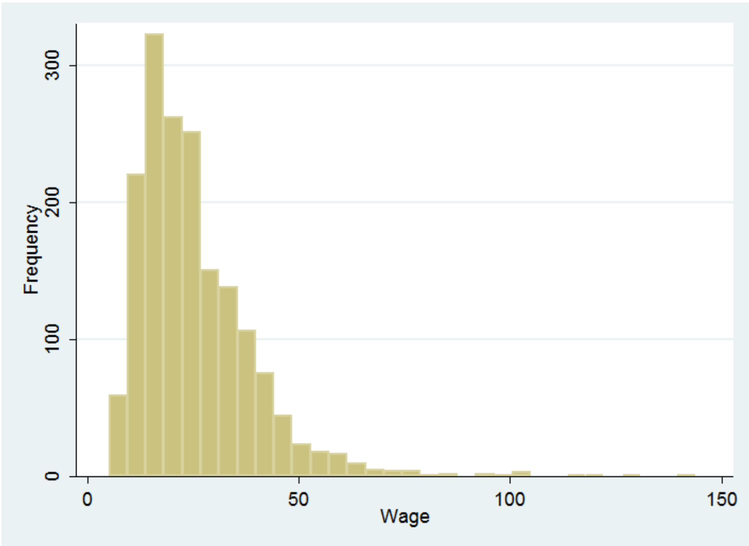
a. Obtain the summary statistics and histograms for the variables *HourlyWageRate (Wage)* and *YrsExperience (Exper)*. Comment on the data characteristics.

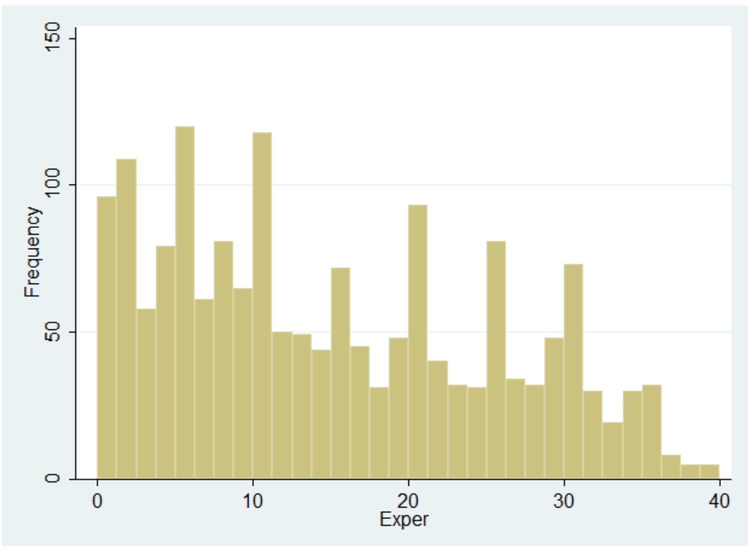
From log file
summarize HourlyWageRate YrsExperience

```
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
HourlyWage~e |      1719    25.48925    14.03004       5.12        144
YrsExperie~e |      1719    14.95637    10.40903          0         40
```

The Hourly Wage Rate has a mean of $25.49 per hour with a standard deviation of $14.03 per hour. As can be seen from the histogram below the data is skewed with a long right tail, up to a value of \$144 per hour, while the lowed value of \$5.12 is much closer to the mean. Most of the data is between \$5.12 and \$50 per hour.

In contrast, the data for the number of years of experience is more spread out with an almost triangular shape. The smallest value is 0 years of experience and the largest value is 40 years, with a mean of 14.96, or approximately 15 years of experience, and a standard deviation of 10.4 years. There is no single mode, but instead looks like there are several modes.





b. Estimate the linear regression

$$Wage = \beta_1 + \beta_2 Exper + e,$$

and interpret the estimates for the intercept and the slope.

<u>*From log file*</u>

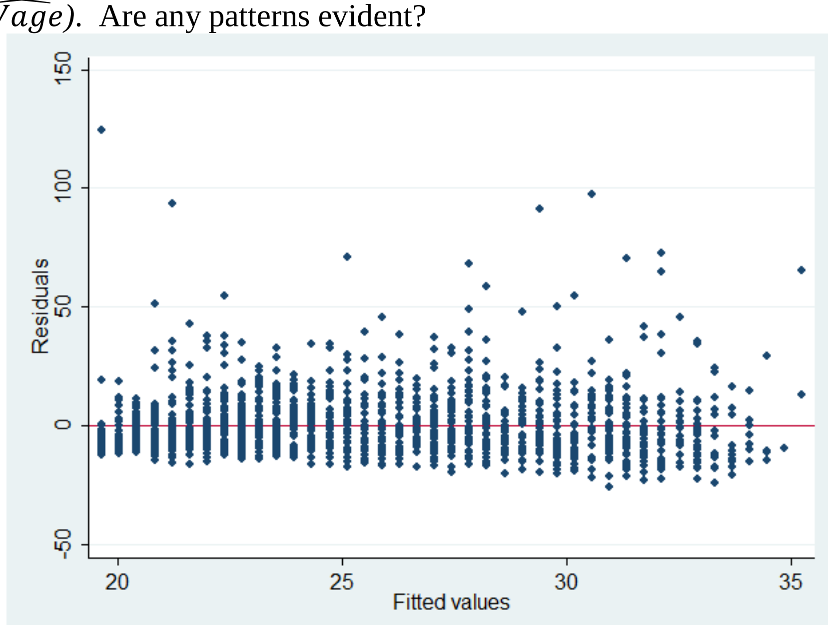
```
. regress HourlyWageRate YrsExperience

      Source |       SS       df       MS              Number of obs =    1719
-------------+------------------------------           F(  1,  1717) =  156.93
       Model |  28319.7626     1  28319.7626           Prob > F      =  0.0000
    Residual |  309854.615  1717  180.462793           R-squared     =  0.0837
-------------+------------------------------           Adj R-squared =  0.0832
       Total |  338174.377  1718  196.841896           Root MSE      =  13.434

------------------------------------------------------------------------------
HourlyWageR~e |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
YrsExperience |   .3900524   .0311367    12.53   0.000     .3289826    .4511222
        _cons |   19.65548    .567318    34.65   0.000     18.54277    20.76819
------------------------------------------------------------------------------
```

$$Wage = 19.66 + .39Exper$$

The average wage rate would be $19.66 for someone with no experience. This is usable, as we have a substantial amount of data for individuals with no experience.

The model predicts that the average wage rate increases by $39 for each extra year of experience.

c. Calculate the least squares residuals and plot them against *Predicted HourlyWageRate* ($\widehat{Wage}$). Are any patterns evident?



There is a pattern evident, as the number of years of experience increases the residuals spread out further from the expected value of zero, creating a fan shape. The values for the residuals start above the "-50" value, but the positive values go up above a value of 150. This regression has some relatively larger positive errors, with many small negative errors.

d. Using the variable *UnionorCollA*, estimate separate regressions for members of a union and covered by a collective agreement (code 1), workers covered by a collective agreement but not in a union (code 2), and workers not a member of a union nor covered by a collective agreement (code 3). Compare the results. Does it appear that there is a difference between the effects of experience on wage rates for each of the three groups of workers?

regress HourlyWageRate YrsExperience if UnionorCollAgree==1

```
      Source |       SS       df       MS              Number of obs =     438
-------------+------------------------------           F(  1,   436) =   12.02
       Model |  1247.69828     1  1247.69828           Prob > F      =  0.0006
    Residual |  45245.5021   436  103.774087           R-squared     =  0.0268
-------------+------------------------------           Adj R-squared =  0.0246
       Total |  46493.2003   437  106.391763           Root MSE      =  10.187

------------------------------------------------------------------------------
HourlyWageR~e |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
YrsExperience |   .1695023   .0488838     3.47   0.001     .0734251    .2655796
        _cons |   24.21304   .9472394    25.56   0.000     22.35132    26.07477
------------------------------------------------------------------------------
```

$Wage = 24.2 + .17 Exper$ for Unionized workers

. regress HourlyWageRate YrsExperience if UnionorCollAgree==2

```
      Source |       SS       df       MS              Number of obs =      34
-------------+------------------------------           F(  1,    32) =    0.18
       Model |  100.376121     1  100.376121           Prob > F      =  0.6747
    Residual |  17907.0105    32  559.594077           R-squared     =  0.0056
-------------+------------------------------           Adj R-squared = -0.0255
       Total |  18007.3866    33  545.678381           Root MSE      =  23.656

------------------------------------------------------------------------------
HourlyWageR~e |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
YrsExperience |  -.1930885   .4559085    -0.42   0.675    -1.121744    .7355667
```

```
       _cons |   32.12252   6.864948     4.68   0.000     18.13907    46.10596
------------------------------------------------------------------------------
```

$Wage = 32.12 - .19Exper$ for Collective Agreement workers

```
. regress HourlyWageRate YrsExperience if UnionorCollAgree==3

      Source |       SS       df       MS              Number of obs =    1247
-------------+------------------------------           F(  1,  1245) =  155.97
       Model |  30221.6791     1  30221.6791           Prob > F      =  0.0000
    Residual |  241245.781  1245  193.771712           R-squared     =  0.1113
-------------+------------------------------           Adj R-squared =  0.1106
       Total |  271467.461  1246  217.871156           Root MSE      =   13.92

------------------------------------------------------------------------------
HourlyWageR~e |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
--------------+----------------------------------------------------------------
YrsExperience |   .4676089   .0374428    12.49   0.000     .3941509    .5410669
       _cons |   18.07514   .6693444    27.00   0.000     16.76197    19.38831
------------------------------------------------------------------------------
```

$Wage = 18.08 + .47Exper$ for Workers not unionized nor in a collective agreement.
$Wage = 24.2 + .17Exper$ for Unionized workers
$Wage = 32.12 - .19Exper$ for Collective Agreement workers

Yes, it appears that the regression results are different between the three groups. The intercept value is the smallest for workers not unionized nor in a collective agreement and largest for workers not unionized but in a collective agreement. The slope value is negative for the workers not unionized but in a collective agreement, while positive for the other two groups, the unionized workers and workers not unionized nor in a collective agreement. The negative slope would not be expected for any group. Experience does appear to have a different effect on wage rates between the three groups. This would need to be verified with hypothesis testing.