Homework #1  
Economics 113  
Fall 2011  
Professor Spearot  
Due Thursday, November 17th in class.  

ANSWERS  

Please include all code and computer printouts with your answers!  

1. Using the Wage data from the course website, and the computer program R, summarize the variables wage, hours, and iq. Report the mean, median, max, min, and standard deviation.

```r
> summary(x$wage)
Min. 1st Qu. Median  Mean 3rd Qu.  Max.
115.0 669.0  905.0  957.9  1160.0 3078.0
> summary(x$hours)
Min. 1st Qu. Median  Mean 3rd Qu.  Max.
20.00  40.00   40.00  43.93  48.00  80.00
> summary(x$iq)
Min. 1st Qu. Median  Mean 3rd Qu.  Max.
50.0  92.0   102.0 101.3  112.0  145.0

> sqrt(var(x$wage,na.rm=TRUE))
[1] 404.3608
> sqrt(var(x$hours,na.rm=TRUE))
[1] 7.224256
> sqrt(var(x$iq,na.rm=TRUE))
[1] 15.05264
```

2. Please regress wage on hours. Interpret the coefficient on hours. Is it statistically different from zero? Please use a two-sided test, and a 95% level of confidence.

```r
> summary(lm(wage~hours,x))
Call:
  lm(formula = wage ~ hours, data = x)

Residuals:
     Min      1Q  Median      3Q     Max
-839.72 -287.21  -52.38  200.46 2131.26

Coefficients:  
             Estimate Std. Error t value  Pr(>|t|)    
(Intercept)  981.315     81.575  12.03   <2e-16 ***
hours       -0.532      1.832  -0.29    0.772
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 404.6 on 933 degrees of freedom  
Multiple R-squared: 9.033e-05,  Adjusted R-squared: -0.0009814  
F-statistic: 0.08429 on 1 and 933 DF,  p-value: 0.7716
```
A one hour increase in average hours worked per week yields a 0.532 decrease in the average monthly wage. The t-stat is extremely close to 0 (-0.29), and we cannot reject that the coefficient on hours is equal to zero.

3. Suppose that iq is an omitted variable. Do you think the coefficient on hours is biased? If so, in which direction, and why? Please run a regression to evaluate the bias.

*Probably biased upward, since high iq people work more and earn a higher wage.*

```r
> summary(lm(wage~hours+iq,x))
```

Call:  
`lm(formula = wage ~ hours + iq, data = x)`

Residuals:

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-888.1</td>
<td>-257.0</td>
<td>-46.4</td>
<td>204.3</td>
<td>2046.9</td>
</tr>
</tbody>
</table>

Coefficients: 

|                  | Estimate | Std. Error | t value | Pr(>|t|) |
|------------------|----------|------------|---------|----------|
| (Intercept)      | 190.3847 | 110.9182   | 1.716   | 0.0864   |
| hours            | -1.8194  | 1.7474     | -1.041  | 0.2981   |
| iq               | 8.3675   | 0.8386     | 9.977   | <2e-16 *** |

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 384.7 on 932 degrees of freedom  
Multiple R-squared: 0.09659, Adjusted R-squared: 0.09465  
F-statistic: 49.82 on 2 and 932 DF,  p-value: < 2.2e-16

*The suspicion is warranted. There was an upward bias in the coefficient on hours.  
(Note: Multiple answers are fine here. Just note how your conjecture regarding bias is related to the actual bias)*

4. Please run a regression of log(wage) on log(hours). Please interpret the coefficient on log(hours), and construct a 99% confidence interval for this coefficient. What does this say about the relationship between hours, and the wage rate, wage/hours?

```r
> summary(lm(log(wage)~log(hours),x))
```

Call:  
`lm(formula = log(wage) ~ log(hours), data = x)`

Residuals:

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2.0191</td>
<td>-0.2759</td>
<td>0.0291</td>
<td>0.2777</td>
<td>1.2958</td>
</tr>
</tbody>
</table>

Coefficients:


| Estimate | Std. Error | t value | Pr(>|t|) |
|-----------|------------|---------|----------|
| (Intercept) | 7.17786 | 0.34071 | 21.067 | <2e-16 *** |
| log(hours) | -0.10578 | 0.09029 | -1.172 | 0.242 |

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4211 on 933 degrees of freedom
Multiple R-squared: 0.001469, Adjusted R-squared: 0.0003988
F-statistic: 1.373 on 1 and 933 DF,  p-value: 0.2417

This says that the as hours goes up, the wage rate (wage/hours) is falling. In terms of the coefficient, a one percent increase in hours worked yields a -0.11 % decrease in the average monthly wage. The confidence interval on log(hours) is:

\[-0.10578-2.55*0.09029 < B < -0.10578+2.55*0.09029\]
\[-0.3360195 < B < 0.1244595\]