STAT 206 Assignment 10 — **Ted Ying (20511999)**

1.

(a) The mean quantity purchased is **3.5 k (3502) licenses**.
(b) The mean price per license is **$28.26 per license**.
(c) The quantity sum of squares S_{xx} is **76.13 license^2**.
(d) The product of the price and quantity relative to their means S_{xy} is **−$384.97 k (−$384,966)**.
(e) The price sum of squares S_{yy} is **^{2}2050.83 per M license^2**.
(f) For the linear regression, the slope is $\boldsymbol{\beta \approx -5.06}$.
(g) For the linear regression, the intercept is $\boldsymbol{\hat{\alpha} \approx 45.97}$.
(h) The residual squared error is $\boldsymbol{s_e^2 \approx 2.17}$.
(i) The residual standard error is $\boldsymbol{s_e \approx 1.47}$.
(j) The 95% confidence interval for the slope is $\boldsymbol{-5.40 \le \beta \le -4.72}$.
(k) The 95% confidence interval for the intercept is $\boldsymbol{44.71 \le \hat{\alpha} \le 47.23}$.
(l) Since the 95% confidence interval for β does not contain 0, we reject the null hypothesis that $\beta = 0$. This implies the unit price is dependent on the order size.
(m) The 95% confidence interval for the unit price for 1.25 k licenses is **$38.78/license ≤ μ(1.25) ≤ $40.52/license**. So, the licenses are overpriced, since the interval doesn't contain $45 and we reject the null hypothesis.

Appendix — R session

The CSV file was saved to the place the R session was started in.

```
> prices <- read.csv('Stat206-A10.csv')
> mean(prices$x)
[1] 3.502
> mean(prices$y)
[1] 28.261
> sum((prices$x-mean(prices$x))**2)
[1] 76.1298
> sum((prices$x-mean(prices$x))*(prices$y-mean(prices$y)))
[1] -384.9661
> sum((prices$y-mean(prices$y))**2)
[1] 2050.826
> model <- lm(x_i..size. ~ y_i..price., prices)
> model <- lm(y_i..price. ~ x_i..size., prices)
> summary(model)

Call:
lm(formula = y_i..price. ~ x_i..size., data = prices)

Residuals:
    Min      1Q  Median      3Q     Max
-2.8789 -0.7867 -0.4245  0.6693  6.6155

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  45.9696     0.6269   73.33   <2e-16 ***
x_i..size.   -5.0567     0.1688  -29.95   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.473 on 48 degrees of freedom
Multiple R-squared:  0.9492,  Adjusted R-squared:  0.9481
F-statistic:   897 on 1 and 48 DF,  p-value: < 2.2e-16

> sum(resid(model)**2)/(length(prices$x)-2)
[1] 2.170117
> confint(model)
                2.5 %   97.5 %
(Intercept) 44.709137 47.23004
x_i..size.  -5.396174 -4.71724
> predict(model, data.frame(x_i..size.=1.25), interval="confidence")
fit      lwr      upr
1 39.6487 38.77699 40.52042
>
```