# Solution: Final Exam

Business Cycles & Forecasting

Economics 392

Fall/2015

Instructor: Dr. I-Ming Chiu

*It seems that most of you had trouble in the second question, so here is the solution for your reference. Remember, your estimates and predictions are different from mine because the sample is randomly drawn from the population.

Q2. **Answer:**

Study Time vs. Exam Score

Score = 56.15 + 3.71*Time - 1.98*D*Time
(D = 0 if Male, D = 1 if Female)
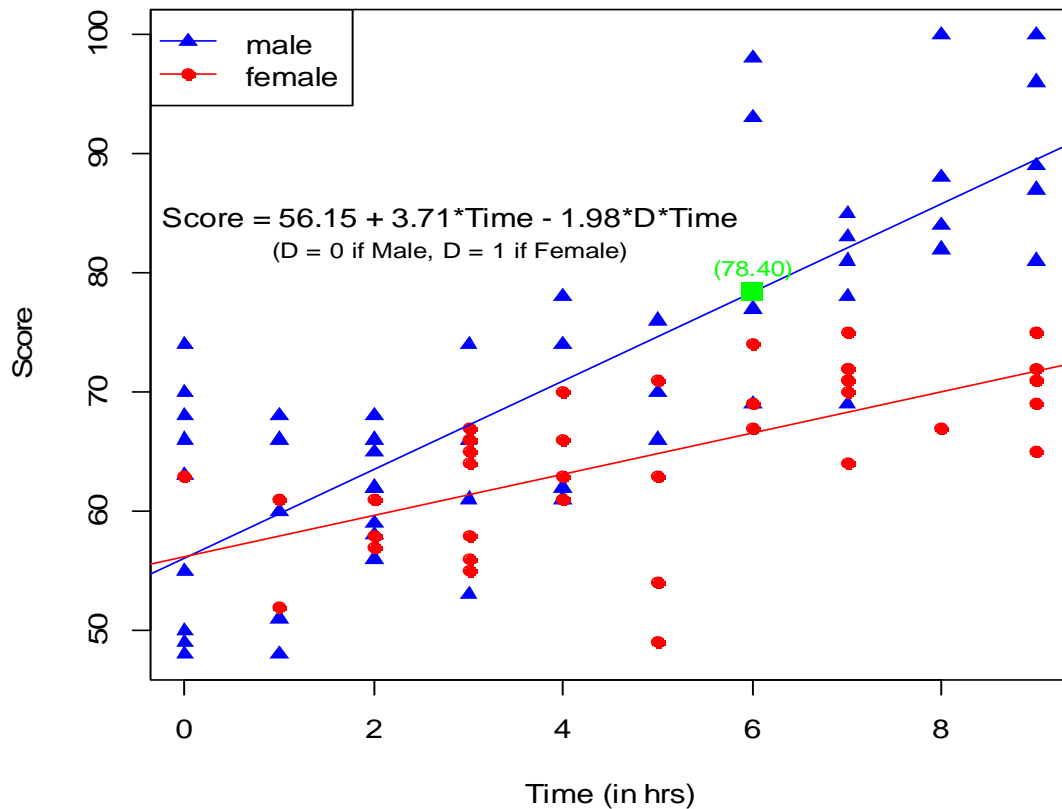
#**Data Analysis using Linear Regression Model w/ Dummy Intercept? Slope? or Both?**

```
rm(list=ls())
data = read.csv("392_finalq2.csv", header = T)
set.seed(2015)
random = sample(1:1000, 100, replace = F)
da = data[random,] #xx: the last two digits of your student id (I don't have an ID)
da.m = da[da$gender == 0,]
```

```r
da.f = da[da$gender == 1,]

plot(da.m$time, da.m$score, pch = 17, col = "blue", main = "Study Time vs. Exam Score",
xlab = "Time (in hrs)", ylab = "Score")
points(da.f$time, da.f$score, pch = 19, col = "red")
#From the plot you should see that the scores for male(triangle) are often above female
(circle) given the same study time
#Based on this information, you know the slopes must be different due to the gender
difference
#Plot the regression line for each gender

model.m = lm(da.m$score~da.m$time)
summary(model.m)
model.f = lm(da.f$score~da.f$time)
summary(model.f)
abline(model.m$coef, col = "blue", lwd = 1.5)
abline(model.f$coef, col = "red", lwd = 1.5)
#What do you find? It seems that both genders have about the same intercept but different
slope
#What is your best model? Introduce dummay to the slope and call it "int" for interaction:

int = da$gender*da$time #same as D*Time in your handout

BestModel = lm(da$score~da$time + int)
summary(BestModel)

#Score = 56.1466 + 3.7089*Time - 1.9778*D*Time
#To make the estimates as close as possible to the maximal decimal points, the prediction:

predict = BestModel$coef[1] + BestModel$coef[2]*6 + BestModel$coef[3]*0*6
predict #a male student who studied 6 hours

#Plot the outcome
plot(da.m$time, da.m$score, pch = 17, col = "blue", main = "Study Time vs. Exam Score",
xlab = "Time (in hrs)", ylab = "Score")
points(da.f$time, da.f$score, pch = 19, col = "red")
abline(BestModel$coef[1], BestModel$coef[2], col = "blue")
abline(BestModel$coef[1], BestModel$coef[2]+BestModel$coef[3], col = "red")

text(2.8, 85, "Score = 56.15 + 3.71*Time - 1.98*D*Time")
text(2.8, 82, "(D = 0 if Male, D = 1 if Female)", cex = 0.8)
legend("topleft", legend = c("male","female"), lty = c(1, 1), col = c("blue","red"), pch = c(17,
19))
points(6, predict, pch = 15, col = "green", cex = 1.5) #add prediction point
text(6, predict+2, "(78.40)", cex = 0.8, col = "green")
```