Stat 407 Quiz 4 Name _____

1. *(2 pts)* True or false

   (a) The pooled variance-covariance matrix is equal to the variance-covariance matrix of all of the data, ignoring group labels. **T** or **F**

   (b) Quadratic discriminant analysis assumes that the samples are take from a multivariate normal population with possibly different means, and different variance-covariance matrices. **F**

2. *(4 pts)* The following problem was conducted on data collected from the 2012 Australian Open womens tennis tournament. It contains standardized data, of statistics measuring the performance of players in the first three rounds of play. We want to predict if a player will make it to the quarterfinals.

```
lda(quarters ~ Aces + Double.Faults + Unforced.Errors + Winners +
    Receiving.Points.Won, data = tennis.sub.std, prior = c(0.5,
    0.5))

Group means:
        Aces Double.Faults Unforced.Errors Winners Receiving.Points.Won
FALSE 0.0079       -0.0095          -0.032  -0.054                -0.07
TRUE  0.0434       -0.5434           0.213   0.589                 1.34

Coefficients of linear discriminants:
                         LD1
Aces                   0.041
Double.Faults         -0.395
Unforced.Errors       -0.146
Winners                0.323
Receiving.Points.Won   0.910
```

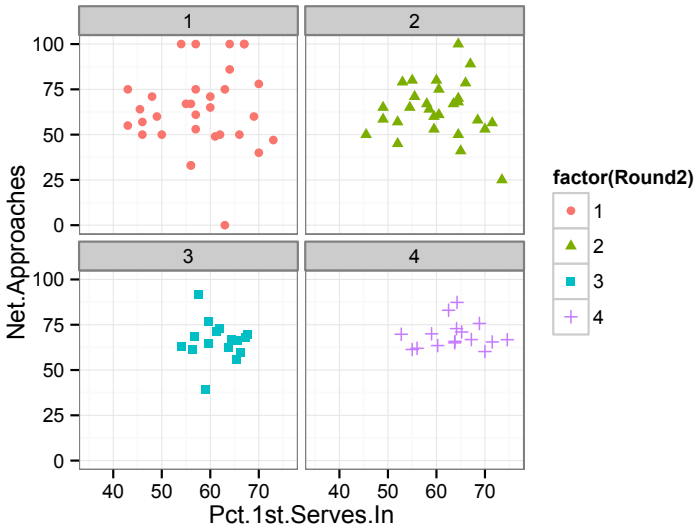
	Predicted	
Actual	FALSE	TRUE
FALSE	88	24
TRUE	0	8

   (a) *(2)* Suppose that the constant for the LDA rule is 0.76, write down the classification rule.

   (b) *(1)* Calculate the error rate for the rule. _____

   (c) *(1)* Which is the most important statistic? _____

3. *(4 pts)* This question looks at the same 2012 Australian Open tennis data, and explores differences between the means for players who got knocked out in the 1st, 2nd, 3rd or 4th and later rounds, on two statistics.




```
> summary(manova(cbind(Pct.1st.Serves.In,Net.Approaches)~Round2, data=tennis),
  test="Wilks")
          Df Wilks approx F num Df den Df Pr(>F)
Round2     1  0.91     4.43      2     89  0.015 *
```

(a) *(1)* What is the value of Wilks Λ? _____

(b) *(1)* What null hypothesis is being tested by the MANOVA calculations?

(c) *(1)* Would you conclude "Reject the null hypothesis?" _____

(d) *(1)* Name one assumption that is required in order for the MANOVA to be conducted. (An extra credit point is awarded if you also say whether it is likely satisfied by the data or not.)