Stat 322 (F17) - Assignment #1

(Due Wed. Oct. 4 at 4:00 pm in appropriate STAT 322 slot in assignment box #15 outside the Math Tutorial Centre (MC 4066/4067). Electronic submissions or in-class submissions will not be accepted under any circumstances.)

1) In response to a complaint that a particular property tax assessor (A) was biased, a study was conducted in which eight properties were randomly selected and each was assessed by assessor A and by an impartial assessor (B). The assessments (10^5 dollars) are shown in the table below.

Property	A	B
1	36.3	35.1
2	48.4	46.8
3	40.2	37.3
4	54.7	50.6
5	28.7	29.1
6	42.8	41.0
7	36.1	35.3
8	39.0	39.1

a) Describe how replication, randomization and blocking (if relevant) are employed in this study.

b) Give the response model that describes the experimental design of this study.

c) Is there a difference in assessed values between assessors A and B? **Answer this question** by calculating manually (i.e. without using dedicated functions in R, such as *anova, t.test, confint,* etc.) a 95% confidence interval for the difference in treatment effects, $\tau_1 - \tau_2$ (you may use R to perform arithmetic operations and to obtain standard deviations and means, using the *sd* and *mean* functions, respectively).

d) Verify the confidence interval obtained in c) by using the *t.test* function in R.

e) Note the small sample sizes. For confidence intervals and hypothesis tests for $\tau_1 - \tau_2$, what assumptions are required about the differences in assessments between the two assessors for each property?

f) Suppose the data were (incorrectly) analyzed as having come from a sampling protocol in which 16 properties were randomly selected, and each assessor was randomly assigned to assess eight of these properties each. Answer the same question posed in c) by (manually) performing a (two-sided) hypothesis test. As with all hypothesis tests, be sure to include the null and alternative hypothesis, value of test statistic (with degrees of freedom), p-value (using the *pt* function in R), and conclusion in the context of the study.

g) Verify the hypothesis test results obtained in f) by using the *t.test* function in R.

h) Was your conclusion in incorrectly analyzing the data in this way different from that in c)? Briefly explain why.

2) A major greeting card company has negotiated a deal with a credit card company to include a coupon for a greeting card set in every cardholder's monthly statement. The greeting card company has come up with 4 coupon designs that it wishes to test to determine which coupon is most effective in generating greeting card set sales. After sending out the 4 different coupons to randomly selected credit card customers in a trial run, a random sample of 8 customers who redeemed their coupons is selected from each coupon design (8 customers for each design), and the dollar value of the card set ordered for each customer is recorded. The data are provided in the following table.

	Coupon Design			
Customer	**#1**	**#2**	**#3**	**#4**
1	86	123	75	33
2	119	204	243	88
3	77	145	98	65
4	127	143	123	109
5	107	98	34	88
6	123	202	76	68
7	111	209	69	88
8	44	156	89	88
mean	**99**	**160**	**101**	**78**
std. dev.	**28**	**41**	**63**	**23**

a) Define the factor, factor levels, treatment, and response variate for this study.

b) Describe how replication, randomization and blocking (if relevant) are employed in this study.

c) Give the response model that describes the experimental design of this study.

d) Use the means and standard deviations of the sales for the different coupon designs to generate manually an ANOVA table for this data. Use this table to test the hypothesis that there is no difference in mean sales between the 4 coupon designs. Be sure to include the appropriate hypotheses, value of the test statistic, p-value (using the *pf* function), and conclusion in the context of the study

e) Use R to verify that the values in the ANOVA table and subsequent *F*-test that you calculated by hand in d) are correct (Note: there may be some discrepancy due to round-off error)