1. [9] Eighteen readers took a speed-reading course. The data set contains the number of words that they could read before and after the course. At the 5% significance level test the hypothesis that the reading-speed course was effective.

a. Specify the appropriate test to check the effectiveness of the course and explain your reason. [2]

b. What is the Null and Alternative hypothesis of interest? Explain your reason. [2]

c. Which of the following outputs represent your rational to test the hypothesis of interest? Explain your reason. [2]

1. t = 11.985, df = 17, p-value = 1.025e-09
   95 percent confidence interval:
   7.141056 10.192278

2. z = 11.985, df = 17, p-value = 1.025e-09
   95 percent confidence interval:
   7.141056 10.192278

3. t = 11.985, df = 17, p-value = 5.126e-10
   95 percent confidence interval:
   7.141056 10.192278

4. z = 11.985, df = 17, p-value = 5.126e-10
   95 percent confidence interval:
   7.141056 10.192278

d. According to part c, what is your conclusion? [1]

e. Interpret the 95% confidence interval. Does your conclusion agree with the conclusion reached in part d? [2]

2. [6 Marks] Laughter is often called ‘‘the best medicine,’’ since studies have shown that laughter can reduce muscle tension and increase oxygenation of the blood. In the International Journal of Obesity (January 2007), researchers at Vanderbilt University investigated the physiological changes that accompany laughter.
   Ninety subjects (18–34 years old) watched film clips designed to evoke laughter. During the laughing period, the researchers measured the heart rate (beats per minute) of each subject with the following summary results: $\bar{y}$ = 73.5, s = 6. It is well known that the mean resting heart rate of adults is 71 beats/minute. At α = .05, is there sufficient evidence to indicate that the true mean heart rate during laughter exceeds 71 beats/minute?

a) Include definition of parameter of interest, μ [1]

b) Setup the null and alternative hypothesis of interest.[1]

c) Estimate the appropriate test statistic. [2]

d) Use critical value approach or P-value approach to make your conclusion in the context of question. [2]

3. [14 Marks] In Brain and Behavior Evolution (April 2000), Zoologists conducted a study of the feeding behavior of blackbream fish. The zoologists recorded the number of aggressive strikes of two blackbream fish feeding at the bottom of an aquarium in the 10-minute period following the addition of food. Data set contains information regarding weekly number of strikes and age of the fish (in days). Considering the following outputs:

```
Call:
lm(formula = Numberofstrikes ~ Ageoffish)

Residuals:
    Min      1Q  Median      3Q     Max
-18.781  -9.684  -2.211   7.634  27.164

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 175.7033    38.6284   4.549  0.00264 **
Ageoffish    -0.8195     0.2386  -3.434  0.01092 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.43 on 7 degrees of freedom
Multiple R-squared:  0.6276,    Adjusted R-squared:  0.5743
F-statistic: 11.79 on 1 and 7 DF,  p-value: 0.01092


Analysis of Variance Table

Response: Numberofstrikes
          Df Sum Sq Mean Sq F value  Pr(>F)
Ageoffish  1 2809.9 2809.91  11.795 0.01092 *
Residuals  7 1667.7  238.24
---
```

a) What is the least squares prediction equation. [1]

b) Give a practical interpretation of intercept, if possible. [1]

c) Provide an interpretation of the slope estimate of the Least Square Regression Line in the context of the question. [2]

d) What is the interpretation of estimated standard error of the model in the context of question? [2]

e) Is the overall model significant? Why? [1]

f) The estimated coefficient of variation is 34.89% in this model. Are you recommending this model as a precise prediction model? Why? [1]

g) Is there enough evidence of a negative linear relationship between the response variable and the predictor, at the significant level of 0.05? Use a proper procedure to justify your answer. (Null & Alternative, use the output and specify the exact test statistics and p-value and conclusion) [4]

h) What is your conclusion regarding the impact of the explanatory variable using the following is 95% confidence interval. Does it have significand effect? Why? [2]

```
confint(model1, level=.95)
                2.5 %      97.5 %
(Intercept) 84.361583 267.0450774
Ageoffish   -1.383713  -0.2552483
```

4. Suppose that you have been asked to join the team supporting a young New York City chef who plans to create a new Italian restaurant in Manhattan. The stated aims of the restaurant are to provide the highest quality Italian food utilizing state-of-the- art decor while setting a new standard for high-quality service in Manhattan. The creation and the initial operation of the restaurant will be the basis of a reality TV show for the US and international markets (including Australia). You have been told that the restaurant is going to be located no further south than the Flatiron District and it will be either east or west of Fifth Avenue. You have been asked to determine the pricing of the restaurants dinner menu such that it is competitively positioned with other high-end Italian restaurants in the target area.

In particular, your role in the team is to analyze the pricing data that have been collected in order to produce a regression model to predict the price of dinner. Actual data from surveys of customers of 168 Italian restaurants in the target area are available (Data set: NYCR).

The data are in the form of the average of customer views on:

Price: the price (in $US) of dinner (including one drink & a tip)

Food: customer rating of the food (out of 30)

Décor: customer rating of the decor (out of 30)

Service: customer rating of the service (out of 30)

District: East or West of Fifth Avenue

a) [5] Develop a regression model that directly predicts the price of dinner using all four potential predictor variables listed above. Give the least squares prediction equation. (Model 1)

b) [5] Determine which of the numerical variables has the largest estimated effect on Price? Explain your reason. Is this effect also the most statistically significant?

c) [5] If the aim is to choose the location of the restaurant so that the price achieved for dinner is maximized, should the new restaurant be on the east or west of Fifth Avenue? Explain your reason. (Hint: interpret the estimated β coefficient for the related variable.)

d) [5] Does it seem possible to achieve a price premium for "setting a new standard for high-quality service in Manhattan" for Italian restaurants?

e) [4] Fit the new model to the data without the insignificant variable(s) and give the least squares prediction equation. (Model 2)

f) [7] Using Model 1, There is general consensus amongst the team that restaurants on the east of Fifth Avenue are very different from those on the west side with service and décor thought to be more important on the east of Fifth Avenue. As such you have been asked to consider different models for the East and West. (Hint: Investigate whether the effect of the predictors depends on the location of the restaurant.) (Model 3)

g) [10] Conduct an appropriate test to compare Model 3 with Model 2. (Set the hypotheses, identify the reduced and full models, and include the test statistic, your decision and conclusion.)