

MET CS 555 Assignment 4 – 20 points

SUBMISSION REQUIREMENTS: **Please submit a single document (word or PDF) for submission.** Your submission should contain a summary of your results (and answers to questions asked on the homework) as well as your R code used to generate your results (please append to the end of your submission). Please use R for the calculations whenever possible. You will lose points if you are not utilizing R. You will also lose 10 points per day for late submissions unless prior arrangements are made with your facilitator.

Question 1/2. The data in this document is from 3 groups of students (math, chemistry, and physics) on an IQ related test. Save the data, and read the data into R. Use this data to address the following questions (14 points)

- (1) How many students are in each group? Summarize the data relating to both test score and age by the student group (separately). Use appropriate numerical and/or graphical summaries. **(3 points)**
- (2) Do the test scores vary by student group? Perform a one way ANOVA using the aov or Anova function in R to assess. Use a significance level of $\alpha=0.05$. Summarize the results **using the 5-step procedure**. If the results of the overall model are significant, perform the appropriate pairwise comparisons using Tukey's procedure to adjust for multiple comparisons and summarize these results. **(7 points)**
- (3) Create an appropriate number of dummy variables for student group and re-run the one-way ANOVA using the lm function with the newly created dummy variables. Set chemistry students as the reference group. Confirm the results are the same (specifically point out test statistics, p-values, etc. that show the results are equivalent). What is the interpretation of the beta estimates from the regression model? **(4 points)**

Data set for Assignment 4 Question 1/2

group	iq	age
Physics student	34	15
Physics student	33	17
Physics student	32	15
Physics student	25	14
Physics student	36	19
Physics student	30	18
Physics student	31	16
Physics student	34	17
Physics student	29	16
Physics student	34	17
Physics student	39	16
Physics student	33	18
Physics student	39	19
Physics student	42	20
Physics student	41	20
Math student	36	20
Math student	38	28
Math student	37	22
Math student	35	18
Math student	41	19
Math student	40	23
Math student	36	19
Math student	38	16
Math student	24	18
Math student	39	20
Math student	29	19
Math student	38	20
Math student	45	23
Math student	44	24
Math student	44	22
Chemistry student	52	46
Chemistry student	46	38
Chemistry student	51	41
Chemistry student	52	39

Chemistry student	45	44
Chemistry student	49	33
Chemistry student	47	41
Chemistry student	46	36
Chemistry student	41	40
Chemistry student	47	44
Chemistry student	46	46
Chemistry student	42	38
Chemistry student	43	32
Chemistry student	47	41
Chemistry student	40	42

Question 2/2. In the United States, there is a strong relationship between education and smoking: well-educated people are less likely to smoke. Does a similar relationship hold in France? To find out, researchers recorded the level of education and smoking status of a random sample of 459 French men aged 20 to 60 years. The two-way table below displays the data. Is there convincing evidence of an association between smoking status and educational level among French men aged 20 to 60 years? (Identify the right type of chi-square test, explain why, and check the conditions. State an appropriate pair of hypotheses to test in this setting and carry out the test by calculating the test statistic, degree of freedom, and P-value. What conclusion would you draw at $\alpha = 0.05$?) (6 points)

Smoking Status	Education		
	Primary School	Secondary School	University
Nonsmoker	56	37	53
Former	54	43	28
Moderate	41	27	36
Heavy	36	32	16

Extra Credit Question for Assignment 4 (4 point)

The endorphins released by the brain act as natural painkillers. For example, a study monitored endorphin activity and pain thresholds in pregnant rats during the days before they gave birth. The data showed an increase in pain threshold as the pregnancy progressed. The change was gradual until 1 or 2 days before birth, at which point there was an abrupt increase in pain threshold. Apparently, a natural painkilling mechanism was preparing the animals for the stress of giving birth. The following data represent pain-threshold scores. Do these data indicate a significant change in pain threshold? Use a repeated-measures ANOVA with $\alpha = .01$.

Subject	Days Before Giving Birth			
	7	5	3	1
A	39	40	49	52
B	38	39	44	55
C	44	46	50	60
D	40	42	46	56
E	34	33	41	52