Read all questions carefully. Submit both R command (R code file with extension R- Name this file as

Ex1)) and accompanied results where necessary. All possible points for each problem are enclosed

within the [ ], partial results are shown for the most of the problems.

Q.1. [Total Points- 28]

(a) Create a folder HW 1 in your computer. What is the path of this folder? (just copy and paste

from the computer) [1]

(b) Create a vector vec by using the following values (just copy and paste the values in the vector

constructor, don’t modify the values) [1]

2100,1,1,30,165, 2101,2,1,39,880, 2102,1,3,39,902, 2103,1,3,33,674, 2104,2,3,40,260,

2105,1,3,40,100, 2106,1,1,30,229, 2107,2,3,25,764, 2108,1,3,33,901,

2109,1,1,54,868, 2110,1,3,39,472, 2111,2,1,49,254, 2112,2,3,44,739,

2113,1,1,32,845, 2114,2,3,32,694, 2115,2,3,35,555, 2116,1,3,66,287,

2117,1,3,58,872, 2118,1,3,60,981, 2119,1,3,42,395, 2120,2,5,48,344

(c) Create a matrix cost\_mat of 5 columns (filled by row) by using the previous vector vec. The cost\_mat should look like this: [2]

(d) Find the dimension of this matrix by using one of the R functions you learned. [1]

(e) Change the column names of cost\_mat to SubjectID, Gender, Race, Age and Price (fromcolumn 1 to column 5 in order). A part of result should looks like: [2]

(f) Create a vector price which contains the elements of Price column of cost\_mat. The price vector should look like: [1]

(g) Use conditional statement (ifelse or if else) to create another vector Tax which calculates 6% of tax if the price is less than 600, otherwise it calculates 8% tax on price. (Hint: if you are calculating 5% of x then you can calculate as x\*0.05). Tax vector should look like: [3]

(h) Create a new matrix cost\_tax by adding Tax vector as a column to the previous cost\_mat .

The cost\_tax matrix which should look like: [2]

(i) Create a matrix total\_cost by adding a column Total\_price in cost\_tax matrix . Total\_price column is created by adding price and tax (Use any method you have learned). Total\_cost matrix should look like [2]

(j) Create a data frame cost\_data by using matrix total\_cost. (Notice that total\_cost is a matrix and we are converting this into data frame and giving this data frame a new name cost\_data).

The cost\_data should look like [1]

(k) Create a vector Sex by using the values of Gender column of cost\_data which assigns labels 1- as

Female and 2- as Male. The vector Sex should look like [2]

(l) Create vector Ethnicity by using the Race column of cost\_data which assigns labels as 1- Whites, 2- AAs, and 5 others. The vector Ethnicity should look like: [2]

(m) Create a vector Price\_range which groups the Total\_price of cost\_data as follows: [2]

< 300 –Low

300 to 599 –Affordable

600 and up –High

Price\_range should look like:

(n) EXTRA CREDIT

Create a vector Pr\_range by using ifelse statement. This vector groups the Total\_price of total\_cost

matrix as follows: (This is a bit challenging) [3]

< 300 –Low

300 to 599 –Affordable

600 and up –High

Pr\_range should look like:

(o) Create Final\_cost data by adding Sex, Ethnicity and Price\_range vectors as columns to the cost\_data.

The final\_cost should look like: [2]

(p) Remove Gender and Race columns of final\_cost. After removing these variables, final\_cost should look like: [2]

(q) EXTRA CREDIT Create two vectors count\_male and count\_whites which counts the number of males and number of whites in final\_cost data obtained in part (n). The output should look like this: [3]

(r) Export the final\_cost data as expt\_cost.csv file in the Exam1 folder that you’ve already created.

[2]

Q.2. [Total Points=15]

(a) Save Cars.txt data from blackboard in folder. Create R data cars by importing Cars.txt. The

cars data has 77 rows and it should look like:

[2]

(b) Create a data num\_cars using cars data which only contains MSRP, Horsepower, MPG\_City and MPG\_Highway . The num\_cars should look like: [1]

(c) Use one of the functions to obtain the averages of MSRP, Horsepower, MPG\_City and MPG\_Highway of num\_cars all at once. Use only one statement (you are not asked to get them separately). Your result should look like: [1]

(d) Use cars data and calculate the average MSRP of each car type (Make). Result should look like: [2]

(e) Create a vector Acura by using cars data, this vector shows the MSRP of Acura vehicle. Acura vector should look like [2]

(f) Make a scatter plot between MPG\_Highway and MSRP of cars data. Label the x and y axes as

Highway Mileage and MSRP ($). Your graph should look like: [3]

(h) Create a vector mileage\_line by using the MPG\_Highway of cars data such that:

Mileage\_line=95000 - 2300\*MPG\_Highway

Then add the line plotted between MPG\_Highway and Mileage\_line in the previous graph. Also provide

the legend. The final graph should look like: [4]

(i) EXTRA CREDIT

Create a custom function sumf that calculates the sum of first 3 values of any numeric vector. If the vector is non-numeric (character) or has length less than 3, it will return the message “Non-numeric or length is less than 3”. Test your function using the following three vectors- a1, a2, and a3. a1 contains

elements from 1 to 4, a2 is the vector of A, B, C and 2, and a3 contains 1 and 2 as elements. When you test the function sumf, for vector a1, the function should return value 6, and for other two vectors, a2 and a3, it should return “Non-numeric or length is less than 3”. [4]