Create code for these instructions using the Excel file.

Save it in:

1. Word Format
2. Knitted HTML
3. Knitted PDF

# square three times four

# create a new variable called x that is blank

# create a numeric vector called y that contains the numbers 2 thru 2931.

# create a numeric vector called z that is a sequence starting from 1 thru 1000 that increments by 3.

# create a character vector called 'myBuds' that contains names 'Cavin', 'Judy', and 'Kai Wei'. 'Cavin should be repeated 5 fives. Also have the output print when we do the assignment.

# how many factor levels does the myBuds character vector have?

#None, because its not a factor vector

# coerce the myBuds vector to a factor vector by overwritting it.

# write a statement that would tell you how many total characters the myBuds vector has. Hint: work from the inside out.

# create a Date vector called 'today' that has today's date. If this date was converted to a number in R, what would it be?

# create a logical vector called 'a' that has 34 TRUE values and 300 FALSE values.

# write a statement that tests if the length of vector 'a' is the same as vector 'z'.

# What is the sum of vector a multiplied by vector z?

# What would happen if you added vector y and vector z? Would it work, why or why not?

# use the any() function to see if vector z contains the value 100.

# set variable z to NULL. Did this delete variable z?

# what statement would delete variable z from your environment

# what data structure type is 'd'?

# how could you find the names of all the columns in your dataset?

# if 'd' was a matrix type, would names(d) provide you the correct column names?

# what statement would give you the dimensionality of 'd'?

# what statement would give you the number of columns in 'd'?

# what statement would provide you the data structure type, column vector types, and dimensionality all in one?

# print the invoice\_number column.

# print the first three rows of your dataset 'd'.

# select rows 1 thru 10 for colums vendor\_id and invoice\_number.

# what 'is' function asks if the 'terms\_id' column is an integer vector?

# coerce the terms\_id column to a numeric vector type.

# what 'is' function asks if the 'terms\_id' column is a numeric vector?

# create a new data.frame called 'd2' that has columns 'vendor\_id', 'invoice\_total',

# and payment\_total' columns in it from 'd'.

# coerce the 'vendor\_id' column in the d2 data.frame to a factor vector.

# In the d2 data.frame, change the payment\_total and invoice\_total values in the second row to 400.20 instead of 40.20.

# create a vector called 'dummies' using the model.matrix function where the columns are dummy variables for 'vendor\_id' in 'd2'.

# create a new data.frame called 'd3' that contains 'd2' and 'dummies'

# create a list called 'recs' that contains vendor\_id vector from 'd3' in the first element and vector 'a' in the second element

# assign names to the two elements in your 'recs' list

# access the last value in your first element of your 'recs' list.

# create an empty data.frame called 'results' that has 100 rows and 4 columns

# set the column names for 'results' to the following: model, vars, accuracy, runtime.

# load the library data.table (first install it if its not installed)

# coerce data.frame 'd' to a data.table called 'd'

# use the str() to confirm 'd' is a data.table

# set a data.table key on the column invoice\_id in 'd'. Then show that there is indeed a key on that field.

# remove all the objects in your environment