Review Exercises : Part1

1

a. create a vector containing following values : 20,17,99,-100,35,60,70,“apple”,19
b. find out class of this vector. what is the reason for this class to be not integer/numeric?
c. convert the vector to be numeric type
d. use R function to find number of elements in the vector
e. find out mean of the values in the vector
f. create a subset of the vector which contains all value greater than 20, calculate its mean
g. replace all values which are less than 20 or missing with the mean calculates above

2

a. Create a vector containing all values between 1-100 which are divisible by 3
b. Create a vector containing all values between 1-100 which are divisible by 5
c. Using the vectors created above, find out the values between 1-100 which are divisible by both 3 and 5
d. Combine vectors created in a and b , remove repeating values
e. Find out values between 1-100 which are either divisible by 3 or 5 but not by both

3

a. write a for loop to extract names of the character columns in the data ‘product_train.csv’
b. write a for loop over the names obtained above to print for each of those columns ; frequency of categories in them

4

a. read data ‘housing_train.csv’ in R , use stringsAsFactors=F
b. How many unique values variable postcode takes
c. Find out average price across all CouncilAreas
d. What is the difference in average price between house type h and t

5

a. read data ‘store_train.csv’ in R , use stringsAsFactors=F
b. create a subset of the data which belongs to Supermarket Type1 in area Kennebec County, ME
c. what is the total sales (sum of all sales) of Supermarket Type1 in area Kennebec County, ME
d. Find out frequency of 0/1 (variable store) across Grocery Store.
e. Convert the result obtained above to %, round off to two decimal digits