I

SPRING SEMESTER 2023
IST 3015 A: BUSSINESS DATA ANALYTICS
INSTRUCTOR: JAPHETH MURSI
DATE: 13TH March 2023

ASSIGNMENT TWO

Total marks (40)

Question 1 (8mks)

a) Differentiate between a package and a library in R and show example (2mks)
b) How many atomic classes does R programming language have? Discuss them (3mks
c) Write a R program to convert given dataframe column(s) to a vector (3mks)

Color	Numbers
Red	10
Yellow	22
Green	11
Blue	5
Violet	2

Question 2

Given the shared Excel dataset (Ngara_district dataset); (16 Mks)

a) Write a R program to create a simple histogram for second and third column (2mks)
b) Rename the three dataframe columns after "ward" to "A", "B", "C" (2mks)
c) Write a R program and display a line plot for column B and C as X &Y axis respectively (3mks)
d) Label the Line plot as "Total Rural population vs Population with_ Clean Water" (2mks)
e) Exclude the variable 'B' and display the top 6 rows of the data frame. (2mks)
f) Add another column "D" to the dataframe where column"A-B" =D (2mks)
g) Write a R program to drop ward "Kabanga" and "Bukiriro" from "ward" column and show the new total number of wards (3mks)

Question 3 (16Mks)

a) Write a R program to call the (built-in) dataset '*iris'.* Check whether it is a data frame or not? Display the last rows of the dataset (3mks)

b) Write a R program to create a matrix from a list of given vectors (3mks)

Vec1_Age<- c(34,40,30,23,21,24,44,60,90,94)

Vec2_Height<-c(1.7,1.65,1.2,1.69,1.51,1.36,2.37,2.08,1.5,1.6)

c) Write a R program to find Sum, Mean and Product the given vectors (4mks)

d) Convert the above vectors in section D to a dataframe and write a R program to create; (6mks)

i) Line chart of Age and Height

ii) Boxplot of Height

iii) Histogram of Age