**Answer script**

**Task [E]**

|  |  |
| --- | --- |
| **Input File** | data\_cases\_valid.csv |
| **Output File** | Output\_enhanched.json |
| **Installed Packages** | dplyr |

**Program**

**# Program for cleaning data\_case\_corrupted.csv and Find required Output**

**#STEP1** To Know about working directory

getwd()

**#STEP2** Set working directory in folder Test\_wd

setwd(“C:/Users/HP/OneDrive/Documents/Test\_wd”)

**#STEP3** Install package and load libraries

install.packages("dplyr")

library(“dplyr”)

**#STEP4** Read Data from **data\_cases\_valid.csv**

data <- read.csv(“data\_cases\_valid”)

**#STEP4** Replace Disease name for **data\_cases\_valid.csv according to disease id**

data$disease\_id[data$disease\_id == "0"] <- "Babesiosis"

data$disease\_id[data$disease\_id == "1"] <- "Brucellosis"

data$disease\_id[data$disease\_id == "2"] <- "Ephemeral Fever"

data$disease\_id[data$disease\_id == "3"] <- "Gangrene"

data$disease\_id[data$disease\_id == "4"] <- "Infectious Coryza"

data$disease\_id[data$disease\_id == "5"] <- "Necrotic Enteritis"

data$disease\_id[data$disease\_id == "6"] <- "Worm Infestation"

data$disease\_id[data$disease\_id == "7"] <- "Tetanus"

data$disease\_id[data$disease\_id == "8"] <- "Tape Worm"

data$disease\_id[data$disease\_id == "9"] <- "Mycoplasmosis"

**# Program for output file as given examples**

**#STEP6** Find total number of sick cases to calculate average {inc}

tcases <- sum(data$total\_number\_cases)

sprintf(“Total number of reported cases is: %d“ , tcases)

tcase\_op <- sprintf(“Total number of reported cases is: %d“ , tcases)

**#STEP7** Find total number of death cases at each location

deathbyloc\_op <- "Total number of deaths from each disease: "

**#STEP8** Find Sum for Group by data : disease id/ name

aggregate(data$number\_mortality , list(data$disease\_id), FUN = sum)

groupdata <- aggregate(data$number\_mortality , list(data$disease\_id), FUN = sum)

**#STEP8** install Package for json

install.packages("rjson")

library(rjson)

**#STEP9** convert data frame to Json

myfile = toJSON(groupdata)

**#STEP10** store data in tcase\_op for showing output of Total number of reported death cases

avr\_sick\_cats <- print ("Average number of sick cats reported in reports from villages up to two decimal points: ")

**#STEP11** data bind

Output <- rbind( avr\_sick\_cats, deathbyloc\_op)

**#STEP12** final code for showing output in output.json file

write(output, "Output\_enhanched.json")

write(myfile, "Output.json", append = TRUE)