**Name=Akash Mishra**

**Group=CS1**

**Roll no. = 102197001**

**Probability and Statistics (UCS410)**

**Experiment 2: Descriptive statistics, Sample space, definition of probability**

(a) Suppose there is a chest of coins with 20 gold, 30 silver and 50 bronze coins.

You randomly draw 10 coins from this chest. Write an R code which will give us the

sample space for this experiment. (use of sample(): an in-built function in R)

**CODE**

chest = c(rep("gold", 20), rep("silver", 30), rep("bronze", 50))

sample(x=chest, size=10)

CODE OUTPUT





(b) In a surgical procedure, the chances of success and failure are 90% and 10%

respectively. Generate a sample space for the next 10 surgical procedures performed.

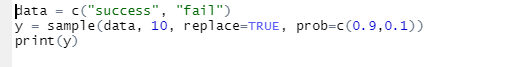
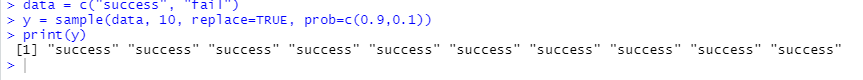
(use of prob(): an in-built function in R)

**CODE**

data = c("success", "fail")

y = sample(data, 10, replace=TRUE, prob=c(0.9,0.1))

print(y)

**2.A room has n people, and each has an equal chance of being born on any of the 365**

**days of the year. (For simplicity, we’ll ignore leap years). What is the probability**

**that two people in the room have the same birthday?**

**(a) Use an R simulation to estimate this for various n.**

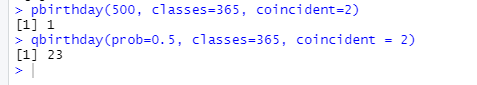
**(b) Find the smallest value of n for which the probability of a match is greater than**

**.5.**

**CODE**



**OUTPUT**



**Write an R function for computing conditional probability. Call this function to do**

**the following problem:**

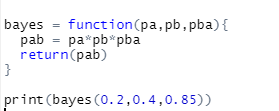
**suppose the probability of the weather being cloudy is 40%. Also suppose the prob-**

**ability of rain on a given day is 20% and that the probability of clouds on a rainy day**

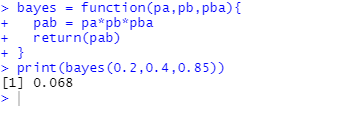
**is 85%. If it’s cloudy outside on a given day, what is the probability that it will rain**

**that day?**

**CODE**



**OUTPUT**



**The iris dataset is a built-in dataset in R that contains measurements on 4 different**

**attributes (in centimeters) for 150 flowers from 3 different species. Load this dataset**

**and do the following:**

**(a) Print first few rows of this dataset.**

**(b) Find the structure of this dataset.**

**(c) Find the range of the data regarding the sepal length of flowers.**

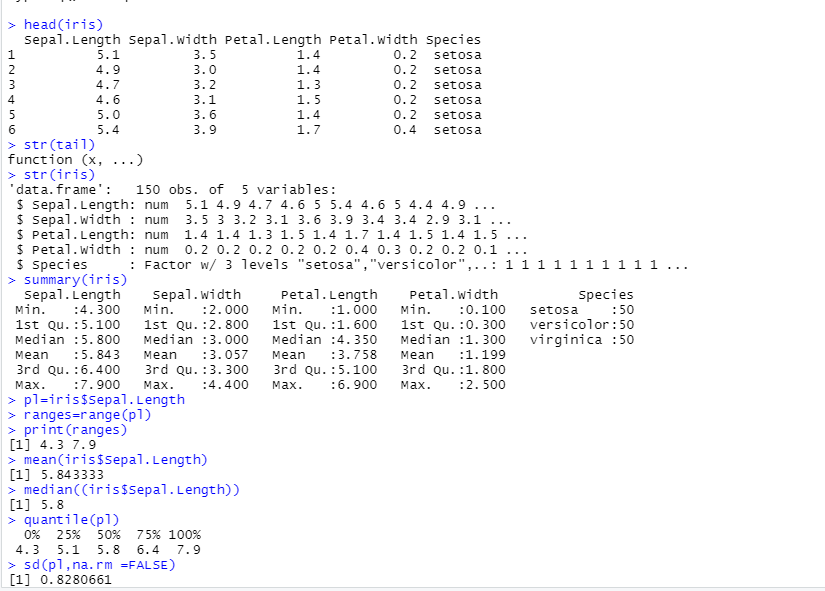
**(d) Find the mean of the sepal length.**

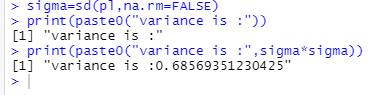
**(e) Find the median of the sepal length.**

**(f) Find the first and the third quartiles and hence the interquartile range.**

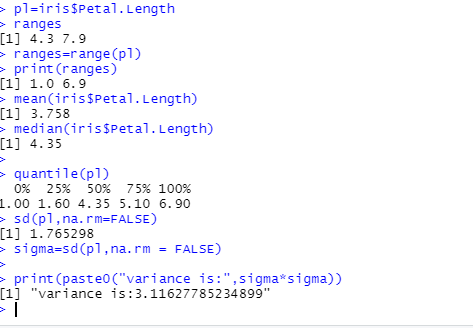
**(g) Find the standard deviation and variance.**

**CODE SNIPPET**





**(h) Try doing the above exercises for sepal.width, petal.length and petal.width.**

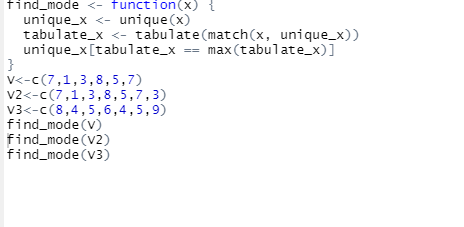


**(5) R does not have a standard in-built function to calculate mode. So we create a user**

**function to calculate mode of a data set in R. This function takes the vector as input**

**and gives the mode value as output.**

**code**



**OUTPUT**

