ID: 00019117

Task 1:

a. To find the answer, we use R's built-in pnorm() function, specifying the mean, upper and lower bounds, and standard deviation. The probability is then calculated by subtracting the two results: pnorm(46, 31, 17) - pnorm(31, 31, 17).

b. Using the normal distribution, the pnorm() function helps determine the probability. Since the question asks for "more than," we subtract the result from 1 and multiply by 100 to get the desired percentage: (1 - pnorm(29, 31, 17)) \* 100.

Task 2:

a- Using built-in r function to get desired answer like at least we have to use pbinom() and give varibales, after that subtracting form 1. 1-pbinom(2,21,0.7)

b- to get answer (80 – e), e = 7 persentile for x and I used binom() that will work 1-pbinom(2,21,0.7)

Task 3:

a-So in order to get quartiles and giving from a=39, b=21, x=56, here I uniform distribution used. With punif(a,b,x)

b- after that we same logic punif(a,b,x)\*252 and round it to get nearest integer. round(punif(41,21,56)\*252)

Task 4:

a- Poisson distribution we can use ppois(a, b/c) , a=2, b=5, c=12

b-so we have interval so we need to subtract both values of it ppois(2, 5) - ppois(6,5)

Task 5:

a- First of all, in my case it must start from 11 but I need to get even number so I will add 1 and until 34 like this, and I will add 2 to get even number sequence with r built-in seq()

b- My task was to go from 1 and until length of the task\_5a length, similarly seq() using from 1 to length of the sequence which above I created

c- After that task\_5a and task\_5b just adding in total, and creating file I need to use write() function to write a file I will write commands each of them to save it and after that to open I will use source() function give my file name

d- I use built-in function sum() and it will sum the whole sequences of numbers

e -I use built-in function median() and it will find the median form sequence

f- I use built-in function mean() and it will find the mean form sequence

g- I use built-in function sd() and it will find the standard deviation form sequence

h- I use built-in function quantile() and it will find the quantile form sequence, which of them you need in my case was 0.41

i- I use built-in function IQR() and it will find the IQR form sequence meaning that subtracting from q3-q1 and you will get answer

j- Task it says to get absolute value I will use built-in abs() function to get answers and subtracting given quantiles like form 0.81 and 0.21 and using task\_5c sequance

Task 6:

a- bag contains different keys we have chance of getting key is ½, and one bag has 1/3 and 1/19 to get probability. So we use multiply ½ each number of keys and subtract form it.

Task 7:

a- we have table which represents some value and use sqrt() r built-in (a/b)+(c/d) and dividing (f/a) after we use pnorm() of the result.(14/14.5)/sqrt((16/36)+(25/47))

Task 8:

Before starting the work we need to get rid of some rows form given data, using readr library we need to get data from link and save to to the variable and using -c(6,16,114,264,406,456,601,701,1008) function we will give numbers which we wanted to remove rows

a- Once we have all the data, we need to filter the city, control, and tuition fields from the dataset. We can achieve this using the %>% operator along with the filter() function from the dplyr package. After filtering the data, we calculate the average tuition using the mean() function, ensuring we include na.rm = TRUE to exclude any missing values from the calculation.

b-so using same technique filtering data like give properties and after that we need to get length if the data, we will use nrow() to get answer.

c- in that case we will use %>% sign and we have to group them by state by using group\_by and also summarise function will be helpful for us to get average tuition fee and arrange them increasing order and after that we get first value suing slice(1) and selecting from state column and getting answer using pull(), and select() fucntion.

e- ggplot(my\_data, aes(x = region, y = tuition, fill = region)) sets up the plot, using the region column for the x-axis and tuition for the y-axis. Stating region will be represented by a different color. The box\_colors variable, allowing you to set specific colors for each region. theme\_minimal() will help to background style to the plot. Using labs adds a title to the plot and labels for the x-axis and y-axis. theme(plot.title = element\_text(hjust = 0.5)) centers the plot title horizontally.

f- ggplot(my\_data, aes(x = tuition, fill = control)) will create the plot, where the tuition column represents the x-axis (tuition rates) and the control column (indicating whether the university is private or public) determines the fill color. geom\_histogram(alpha = 0.5, position = "identity", bins = 30) generates overlapping histograms for private and public universities.