13.Imagine that you have selected data from the All Electronics data warehouse for analysis. The data set will be huge! The following data are a list of All Electronics prices for commonly sold items (rounded to the nearest dollar). The numbers have been sorted: 1, 1, 5, 5, 5, 5, 5, 8, 8, 10, 10, 10, 10, 12, 14, 14, 14, 15, 15, 15, 15, 15, 15, 18, 18, 18, 18, 18, 18, 18, 18, 20, 20, 20, 20, 20, 20, 20, 21, 21, 21, 21, 25, 25, 25, 25, 25, 28, 28, 30,  30, 30.  (i) Partition the dataset using an equal-frequency partitioning method with bin equal to 3 (ii) apply data smoothing using bin means and bin boundary.(iii) Plot Histogram for the above frequency division .

PROGRAM:

prices <- c(1, 1, 5, 5, 5, 5, 5, 8, 8, 10, 10, 10, 10, 12, 14, 14, 14, 15, 15, 15,

15, 15, 15, 18, 18, 18, 18, 18, 18, 18, 18, 20, 20, 20, 20, 20, 20, 20,

21, 21, 21, 21, 25, 25, 25, 25, 25, 28, 28, 30, 30, 30)

num\_bins <- 3

sorted\_prices <- sort(prices)

bin\_size <- ceiling(length(prices) / num\_bins)

equal\_freq\_bins <- split(sorted\_prices, ceiling(seq\_along(sorted\_prices) / bin\_size))

cat("Equal-Frequency Bins:\n")

print(equal\_freq\_bins)

bin\_means <- lapply(equal\_freq\_bins, mean)

smoothed\_means <- unlist(lapply(1:length(equal\_freq\_bins), function(i) rep(bin\_means[[i]], length(equal\_freq\_bins[[i]]))))

smoothed\_boundaries <- unlist(lapply(equal\_freq\_bins, function(bin) {

lower <- min(bin)

upper <- max(bin)

sapply(bin, function(x) ifelse(abs(x - lower) < abs(x - upper), lower, upper))

}))

cat("\nSmoothed Data (Bin Means):\n")

print(smoothed\_means)

cat("\nSmoothed Data (Bin Boundaries):\n")

print(smoothed\_boundaries)

hist(prices, breaks=10, col="lightblue", main="Histogram of Prices",

xlab="Price", ylab="Frequency", border="black")

OUTPUT:

