Bookmarks

- Module 1: The Basics of R and Introduction to the Course
- Entrance Survey
- Module 2: Fundamentals of Probability, Random Variables, Distributions, and Joint Distributions
- Module 3: Gathering and Collecting Data, Ethics, and Kernel Density Estimates
- Module 4: Joint, Marginal, and Conditional Distributions & Functions of Random Variable

Question 9 - 13

Bookmark this page

Question 9

1/1 point (graded)

On the website www.modelingonlineauctions.com, you will find a number of data sets from actual auctions conducted on eBay.

Download one involving the sale of Cartier watches: *Cartier+3-day+auctions.csv.* There are data on auctions of 18 different watches. For each auction, there is an auction ID, bids, time of each bid, bidder name, bidder rating, minimum bid for the auction, and winning bid for the auction. (Note: the winning bid is not the maximum bid submitted by the highest bidder, but rather the second-highest bid plus an increment.)

Load the data into R. How many auctions are in this data?

18 ✔

18

Submit You have used 1 of 2 attempts

Now clean the data set and create the following variables:

- The id of the auction
- The ratio of the second highest bid to the third highest bid.
- The number of bidders.
- The number of bids.

We can provide you with the following R-code to create these variables, but some information is missing. You will need to either fill the information or create your own code.

```
cartier_data <- read.csv("XXX")
cartier_data$auctionid <- as.character(cartier_data$auctionid)
unique_bids <- unique(cartier_data$XXX)

ratio <- rep(NA, times = XXX)
XXX <- rep(NA, times = length(unique_bids))
number_of_bids <- rep(NA, times = length(unique_bids))

for (i in c(1:length(unique_bids))){
  temp <- subset(XXX, cartier_data$XXX == unique_bids[i])
  bid2 <- XXX[rank(temp$bid, ties.method = 'last') == (length(temp$bid)-1)]
  bid3 <- temp$bid[rank(temp$bid, ties.method = 'last') == (length(XXX)-2)]
  ratio[i] <- bid2 / bid3
  number_of_bidders[XXX] <- length(unique(XXX))
  number_of_bids[i] <- length(temp$bid)
}

data_clean <- data.frame(unique_bids, ratio, number_of_bidders, XXX)
```

Question 10

1/1 point (graded)

What is the mean of the ratio of the second highest bid to the third highest bid?

Please round your answer to the thousandth decimal place, i.e. if it was 2.0086, you would round to 2.009 and if it was 2.0081, you would round to 2.008.

1.083 ✔ **Answer:** 1.083

1.083

Explanation

With the command summary(data_clean) we get that the mean for ratio is 1,083.

Submit You have used 1 of 2 attempts

✔ Correct (1/1 point)

Question 11

1/1 point (graded)

What is the median of the number of bidders?

6 ✔ **Answer:** 6

Explanation

With the command summary(data_clean) we get that the median for the number of bidders is 6.

Submit You have used 2 of 2 attempts

✔ Correct (1/1 point)

Question 12

1/1 point (graded)

What is the maximum value of the number of bids?

33 ✔ **Answer:** 33

33

Explanation

With the command summary(data_clean) we get that the maximum value for the number of bids is 33.

Submit You have used 1 of 2 attempts

✔ Correct (1/1 point)

Question 13

1 point possible (graded)

We can think of ordered bids as being order statistics from some underlying distribution of valuations. Using this perspective, would you expect the number of bidders and the number of bids to inform the ratio between the 2nd and 3rd highest bids?

(Note: We are not looking for a precise, mathematical answer here, just a bit of informed speculation.)

- Yes ✔
- No ✖

Explanation

In general the PDF of any ordered statistic would depend on the number of trials that we have. Therefore, we should expect a relationship between this ratio and the number of bidders and the number of bids.

Submit You have used 1 of 1 attempts

✖ Incorrect (0/1 point)