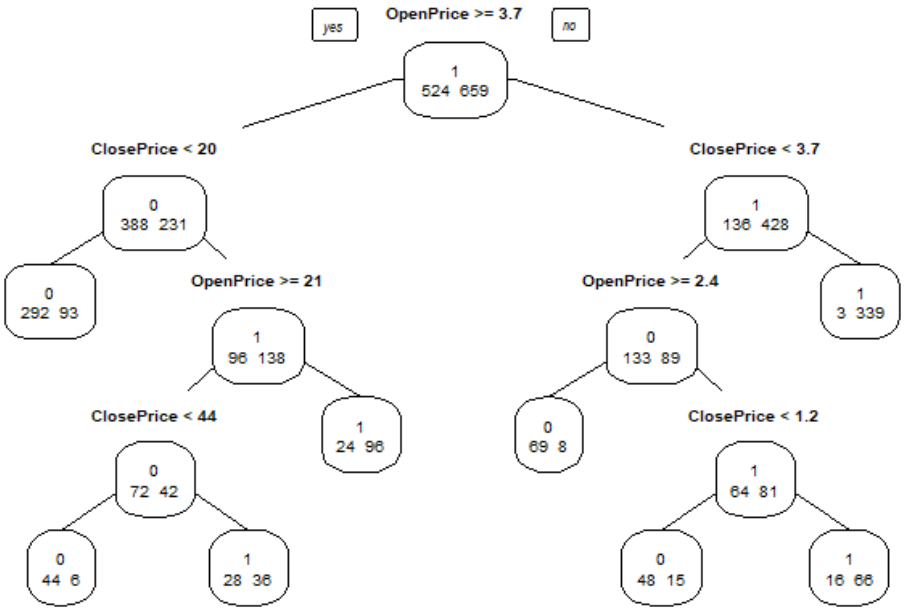
Salwa Niaz Preet
sxp240086

Business Analytics with R

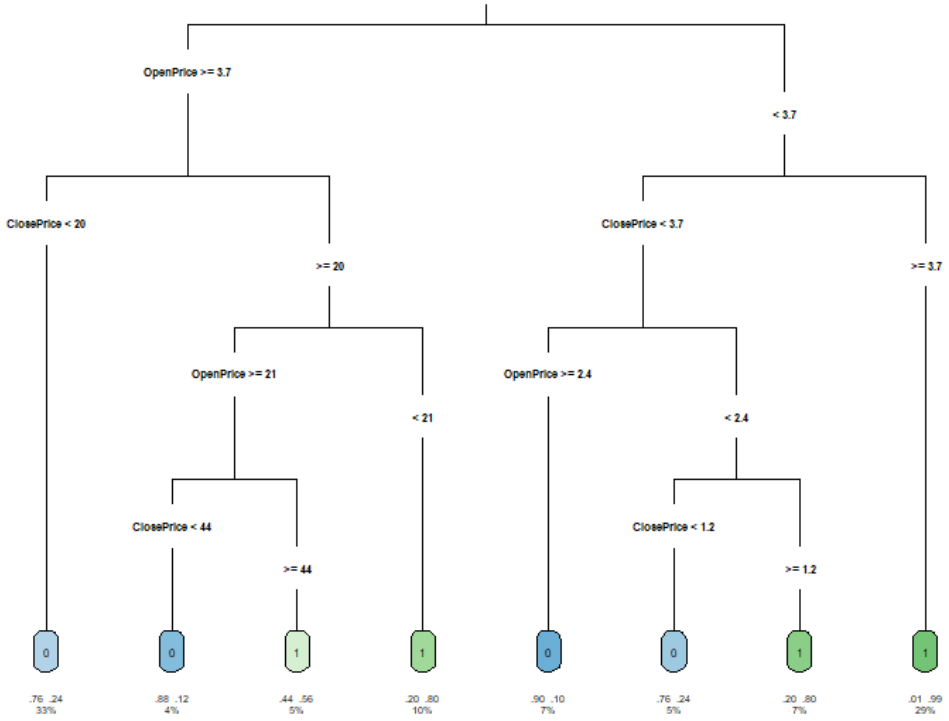
Assignment 3

Data preprocessing done in R

(a) Classification Tree



Pruned Tree:



Salwa Niaz Preet
sxp240086

Rules:

- If Opening Price ≥ 3.685 and Closing Price < 20.195, then Class 0 i.e. noncompetitive auction.
- If Opening Price ≥ 20.5 and Closing Price < 43.64, then Class 0 i.e. noncompetitive auction.
- If Opening Price ≥ 20.5 and Closing Price ≥ 43.64, then Class 1 i.e. competitive auction.
- If Opening Price < 20.5 and Closing Price ≥ 20.195, then Class 1 i.e. competitive auction.
- If Opening Price < 3.685 and Closing Price < 3.685 and Opening Price ≥ 2.445, then Class 0 i.e. noncompetitive auction.
- If Opening Price < 3.685 and Closing Price < 3.685 and Opening Price < 2.445 and Closing Price < 1.24, then Class 0 i.e. noncompetitive auction.
- If Opening Price < 3.685 and Closing Price < 3.685 and Opening Price < 2.445 and Closing Price ≥ 1.24, then Class 1 i.e. competitive auction.
- If Opening Price < 3.685 and Closing Price ≥ 3.685, then Class 1 i.e. competitive auction.

(b) This classification tree only shows opening price and closing price of an auction as predictors of the competitiveness of an auction. While closing price itself, in general, can be a useful indicator of competitiveness, it is not always the most reliable determinant.

Why closing price is a decent predictor:

A higher closing price suggests a more competitive auction, as it reflects strong bidding activity and high demand for the item. Conversely, a low closing price could imply that there was limited interest or bidding, indicating a less competitive auction.

Why closing price is not a good predictor for future data:

However, in the case of a new auction, the closing price is the outcome of an auction rather than a predictor available before the auction starts. Thus, for future auction data, closing price cannot practically be used for predicting competitiveness.

(c) **Interesting Information:**

- The rules give specific thresholds for opening and closing prices of auctions like 3.685 and 20.195 for Opening Price, as well as 1.24, 3.685, and 43.64 for Closing Price. These prices are decisive in determining whether an auction is competitive or not. This indicates that price plays a crucial role in influencing the competitiveness of auctions, which could be tied to perceived value or buyer psychology.
- The rules also show varying price sensitivities. For example, if Opening Price is relatively high (e.g. ≥ 20.5) and Closing Price crosses a certain threshold (≥ 43.64), the probability of a competitive auction increases. On the other hand, when prices are lower, the rules flip, highlighting different dynamics between high and low-price ranges.
- The rules point out that auctions with a higher Closing Price tend to be competitive. This shows that as items close at higher prices, the level of competition among bidders increases. Sellers aiming to optimize pricing strategies can make use of this insight to make their auctions more appealing.

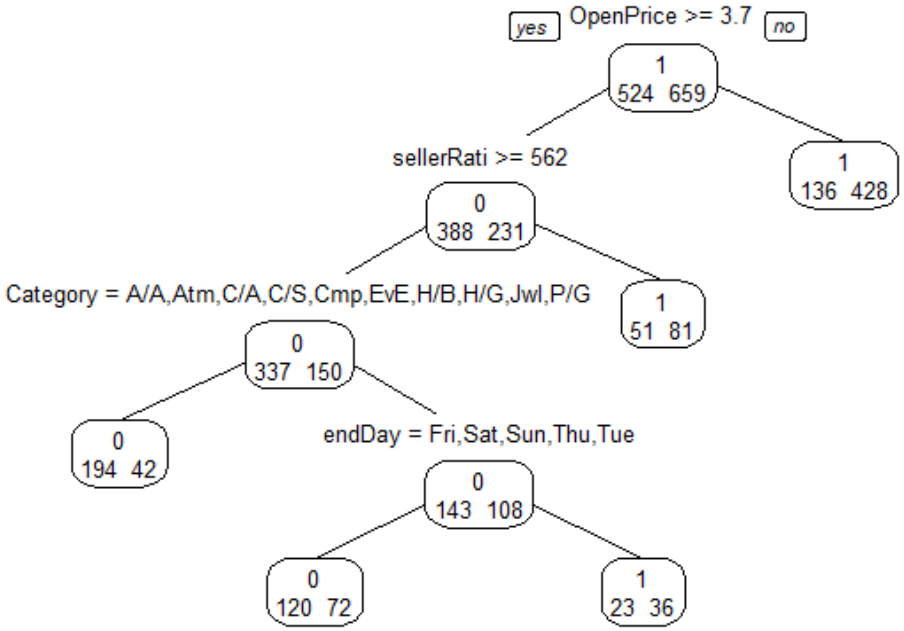
Uninteresting Information:

- The classification tree does not show any other variables like duration of auction, day end of auction, item category, and seller rating. The absence of these possible predictors might make the analysis less comprehensive. Their exclusion from the splits may limit the insights that could be gained about how various auction parameters impact competitiveness.
- The rules do not mention any differences in outcomes based on the currency used (USD, EURO, GBP). This shows that currency does not have a noticeable impact on auction competitiveness

for this dataset. This may be considered uninteresting for sellers looking to optimize auction strategies across different currencies.

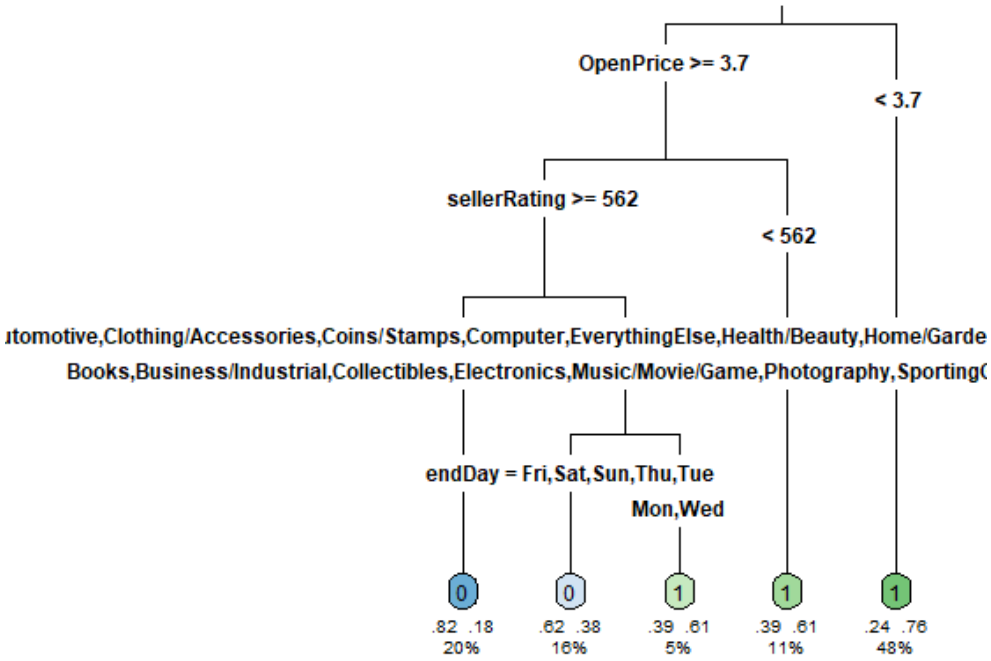
- Variables like duration of auction and day end of auction do not appear in the rules, suggesting they probably do not significantly influence the auction's competitiveness. This might be surprising, as one might expect the timing and duration to influence bidder behavior. The exclusion makes these variables seem unimportant, which could be seen as uninteresting or even disappointing from a strategic standpoint.

d) Classification Tree:



Salwa Niaz Preet
sxp240086

Pruned Tree:

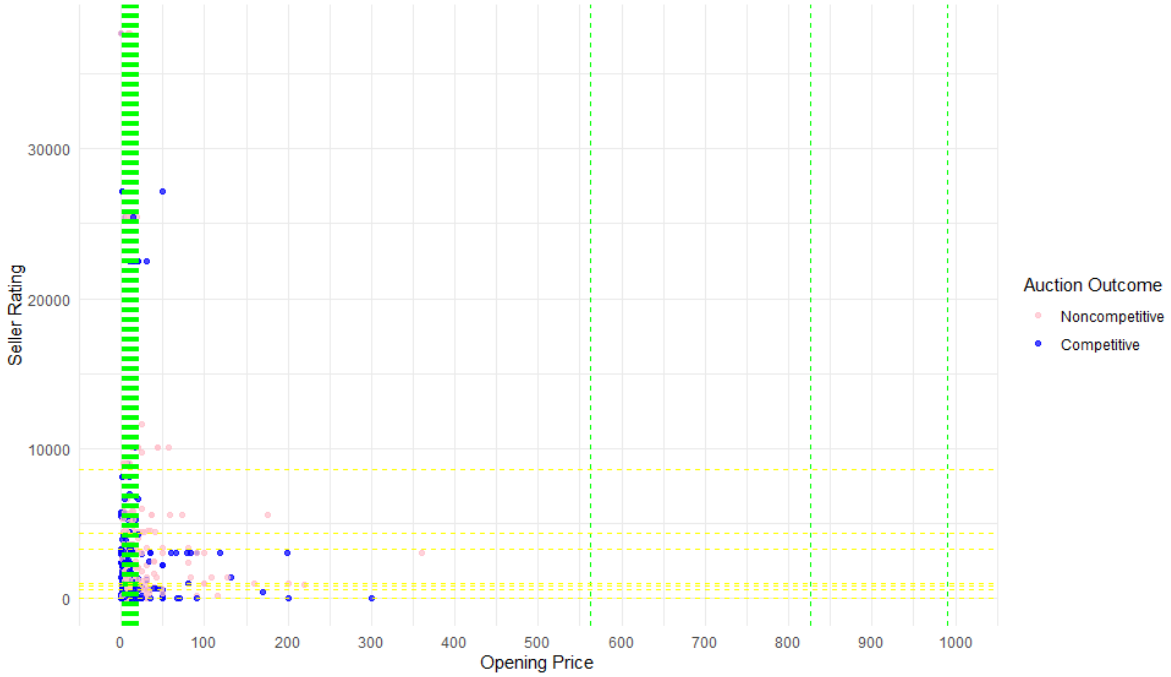


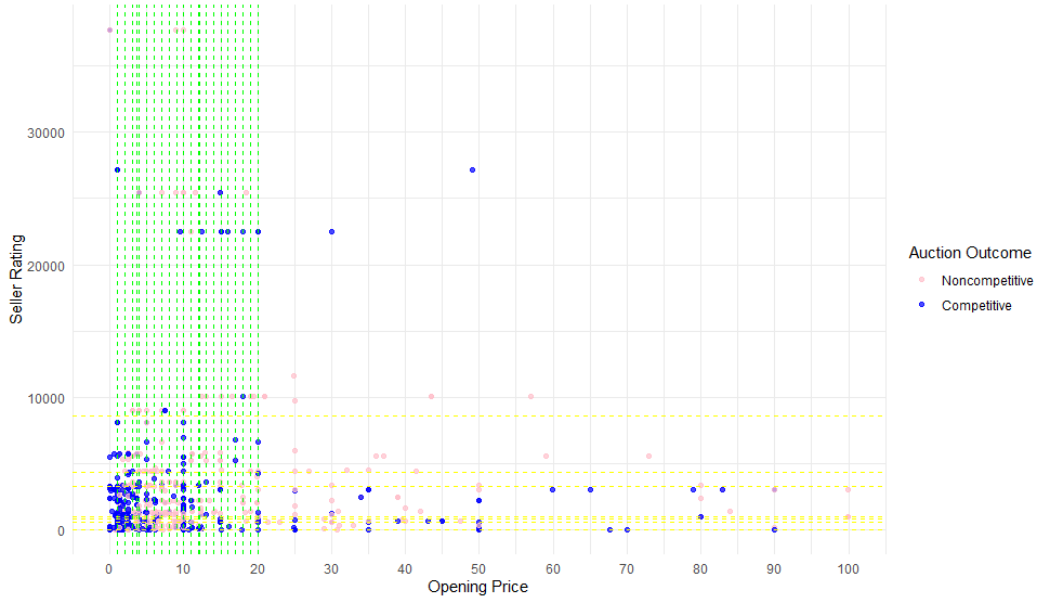
Rules:

- If Opening Price ≥ 3.685 and Seller Rating ≥ 562 and Category is one of Antique/Art/Craft, Automotive, Clothing/Accessories, Coins/Stamps, Computer, EverythingElse, Health/Beauty, Home/Garden, Jewelry, Pottery/Glass, then the auction is not competitive.
- If Opening Price ≥ 3.685, Seller Rating ≥ 562, Category is one of Books, Business/Industrial, Collectibles, Electronics, Music/Movie/Game, Photography, SportingGoods, Toys/Hobbies, and end day is Friday, Saturday, Sunday, Thursday, or Tuesday, then the auction is not competitive.
- If Opening Price ≥ 3.685, Seller Rating ≥ 562, Category is one of Books, Business/Industrial, Collectibles, Electronics, Music/Movie/Game, Photography, SportingGoods, Toys/Hobbies, and end day is Monday or Wednesday, then the auction is competitive.
- If Opening Price ≥ 3.685 and Seller Rating < 562, then the auction is competitive.
- If Opening Price < 3.685, then the auction is competitive.

e) The first scatter plot is with the whole range of opening price and the second one shows the zoomed in version, limiting the opening price axis to 100 so that we can see the separation in splits. The lines don't perfectly separate the two classes, and there is overlap, especially as seller ratings are becoming high. Overall, while the splits capture general trends and differentiate the classes to some extent, there is still room for improvement. The overlap suggests that a more complex model or additional predictors might be needed to achieve a clearer separation.

Salwa Niaz Preet
sxp240086





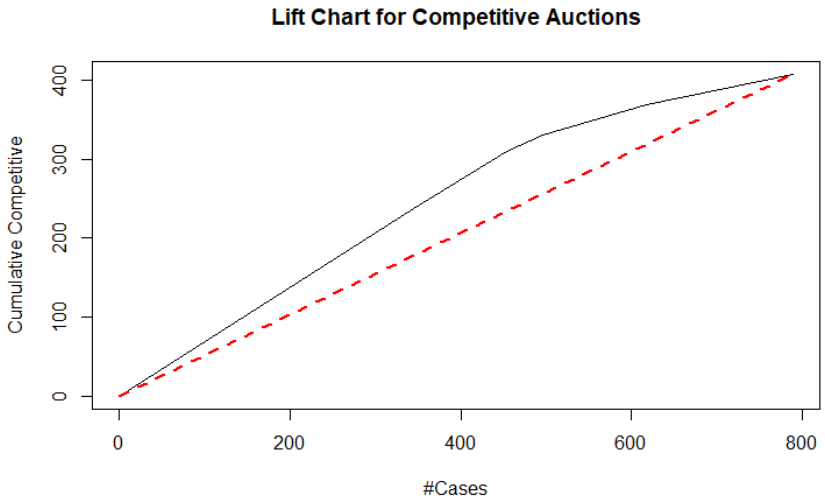
f)

```
Confusion Matrix and Statistics

          Reference
Prediction   0   1
         0 216  77
         1 166 330

               Accuracy : 0.692
                 95% CI : (0.6585, 0.7241)
    No Information Rate : 0.5158
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.379

 Mcnemar's Test P-Value : 1.65e-08

            Sensitivity : 0.5654
            Specificity : 0.8108
         Pos Pred Value : 0.7372
         Neg Pred Value : 0.6653
             Prevalence : 0.4842
         Detection Rate : 0.2738
   Detection Prevalence : 0.3714
      Balanced Accuracy : 0.6881

       'Positive' Class : 0
```

Analysis:

216 non-competitive and 330 competitive auctions have been correctly predicted. 166 non-competitive auctions have wrongly been predicted as competitive and 77 competitive have wrongly been predicted as non-competitive auctions.

- Accuracy: The model has an accuracy of 69.2%, which is okay and not exceptional. This implies that around 69.2% of the predictions made by the model are correct.
- Sensitivity: The sensitivity is 56.5%, indicating that the model is only able to correctly identify 56.5% of the non-competitive auctions. This is relatively low and suggests that the model might be missing some true non-competitive auctions.
- Specificity: The specificity is 81.1%, which means that the model does a good enough job at correctly identifying competitive auctions, with an 81.1% success rate.



Analysis:

The lift chart shows the cumulative number of competitive auctions captured by the model in (d) as more cases are evaluated. The black line represents the performance of the model and the red dashed

line represents a random model. Because the model's line lies above the random model line, we can say that the model performs better than random guessing. The lift is also not very high or significant. The lift appears to flatten out as it approaches the end, suggesting that while the model initially identifies competitive auctions well, its effectiveness decreases as more cases are included.

g) Recommendations:

To maximize the chances of having a competitive auction, a seller should consider the following strategy:

- Keeping the opening price below $3.685 can help attract more bidders and make the auction more competitive.
- Maintaining a high seller rating, above 562, by providing excellent service and positive buyer experiences can significantly increase competitiveness.
- Sellers can list items in categories that have higher competitiveness, such as Books, Collectibles, Electronics, and Sporting Goods.
- Sellers can schedule the auction to end on days with higher bidder activity, such as Monday or Wednesday to increase engagement.
- Currency is not very informative in determining auction competitiveness.