Guided Capstone - Step Six - Summary

Big Mountain ski resort in Montana wants to know if its ticket prices can be adjusted to capitalize on its facilities. It also wants to know if changes can be made to its facilities that would improve profitability or support a higher ticket price.

The first question to be answered is whether there are any basic patterns in terms of price, particularly in terms of price per state. The initial dataset included information on all of the ski resorts in the United States. Additional data on state population and area was gained for wikipedia. In examining the data sets, both were cleaned, and a box plot was created of states and weekend and weekday ticket prices (see Figure 1). No discernable pattern was visible from this, and it was decided that all states would be kept for analysis. Adult weekend ticket price was determined to be the target feature because it was missing fewer data than Adult Weekend ticket prices. For Montana resorts there was no difference in price between weekend and weekday tickets.

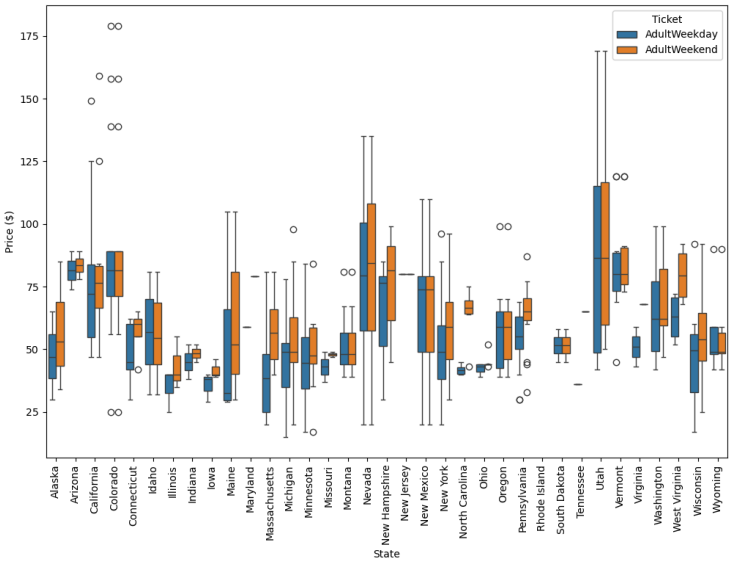

Figure 1

EDA

In the exploratory data analysis phase of the project, Principal Component Analysis was performed and showed that two components account for 77.2% of the variance in the data. The ski data and state summary data sets were merged, and a heatmap was created to study correlations (Figure 2). From this a few features stand out that correlate well with ticket price: fast quads, runs, snow making, the ratio of night skiing area to total state area, total chairs and vertical drop.

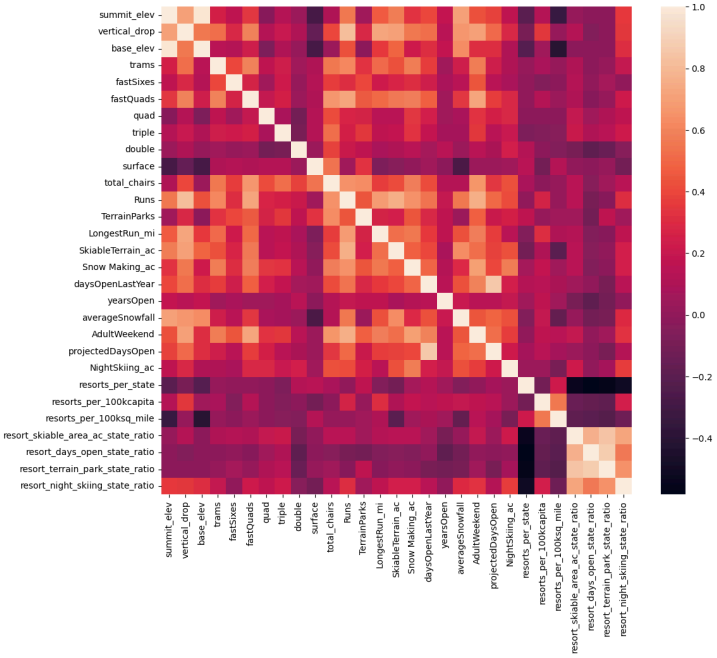

Figure 2

The next step involved using train/test split to divide the data into training and testing sets, and the metrics, r-squared, mean absolute error and mean squared error were calculated for three different models: using the mean, linear regression and random forest. For the latter two, cross-validation was performed on each. Both the linear regression and random forest models found vertical drop, snow-making, fast quads and runs to be the most important features, with linear regression also including total chairs. Other features deemed important were: longest runs, trams, and skiable terrain. This is consistent with what was concluded in the exploratory data analysis step.

The random forest model was selected because it has a lower cross-validation mean absolute error by almost $1 and it exhibits less variability. The random forest model was retrained on all the data, minus Big Mountain and it predicted a price of $95.87, with a mean absolute error of $10.39, meaning that there is opportunity to increase the price. That said, Big Mountain's ticket price is on the higher end of all resorts, and is the highest in Montana. It is also high in vertical drop, snow making area, total number of chairs, fast quads, and several other of the most important features in determining price. See figures below for the ticket price versus the important features (Figures 3-12).

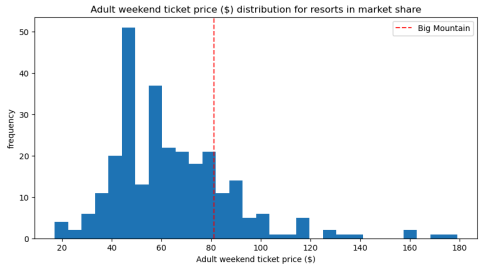

Figure 3

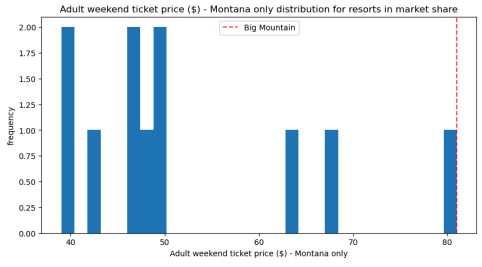

Figure 4

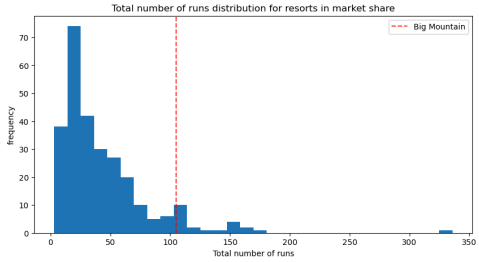

Figure 5

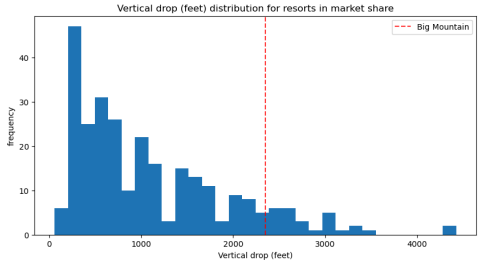

Figure 6

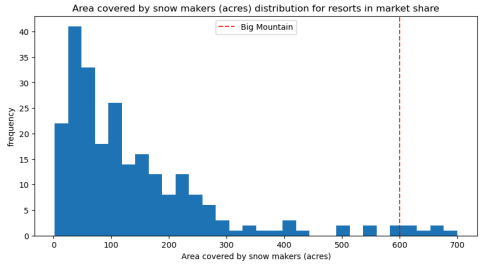

Figure 7

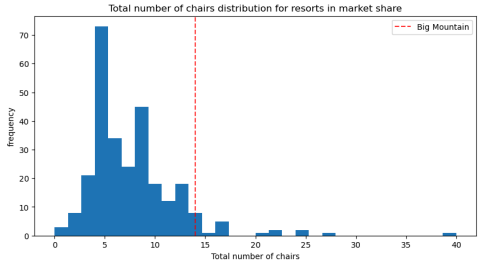

Figure 8

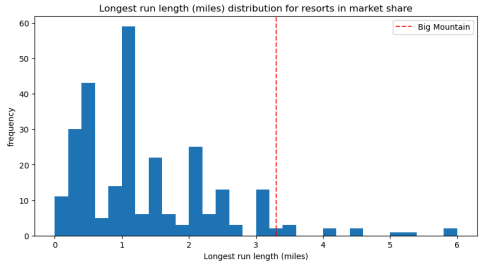

Figure 9

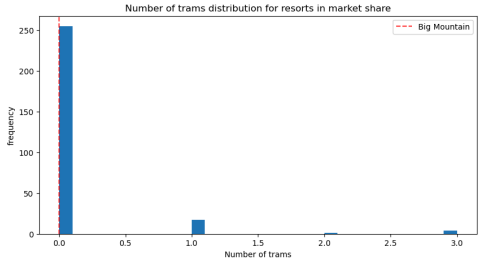

Figure 10

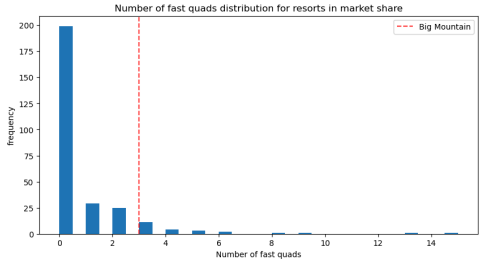

Figure 11

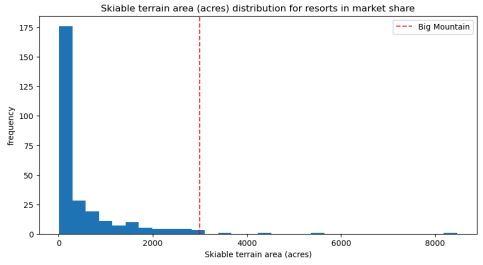

Figure 12

Next, information about the expected visitors per year was added and four scenarios were modeled. Of these, the second scenario (increasing vertical drop and installing a new chairlift) has the most potential. Taking the recent installation of the new chair as an estimate for additional annual operating cost, the new revenue would be $1.11 per ticket. Scenario 1 showed a decrease in projected price with the increase in number of runs closed, but the model does not account for savings in operational costs. This should be

estimated before moving forward. The resort could temporarily close runs to determine the effect it has on attendance before determining to permanently close the runs.

The model predicts that the cost of the ticket is lower than it should be based on the current amenities at Big Mountain. The prediction is for an increase of $10 or more. At a minimum, prices should be increased by $0.88 per ticket to cover the operational cost of the new chair lift. In terms of additional amenities to add beyond the proposed scenarios, it should be noted that Montana already ranks high in some of the most important features for determining ticket price.

Although outside the scope of the project, it might be worthwhile understanding the capital cost of the newest chair in order to justify part of the price increase. Other future efforts could include assessing the impact of other amenities such as distance to major population hubs, size and condition of other amenities such as lodging and food, as well as evaluating two subsets of resorts: one that includes resorts that are geographically close to Big Mountain, but may be in other states, and one that examines other resorts with similar profiles of characteristics to Big Mountain (e.g similar vertical drop, total chairs, fast quads, skiable area). Additional work could include creating a simple app to allow the resort leadership to test out their own scenarios and see the resulting predicted price changes.