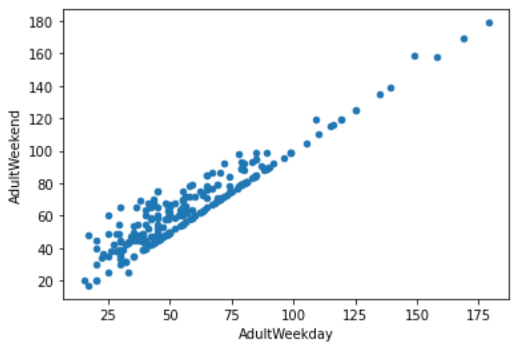
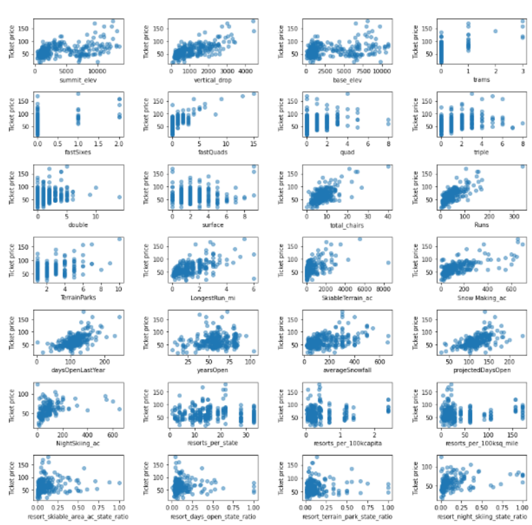
Guided Capstone Project Report

Within the target feature and looking within the graph below we can make a few assumptions. Firstly, there is a clear line where weekend and weekday prices are equal. Weekend prices being higher than weekday prices seem restricted to sub hundred-dollar resorts. Big Mountain resort is in Montana, so the relationship between these quantities in this state are particularly relevant.



In the scatterplots you see what some of the high correlations. There's a strong positive correlation with vertical_drop. fastQuads. Furthermore, Runs and total_chairs appear quite similar. However, resorts_per_100kcapita shows something interesting that you don't see from just a headline correlation figure. When the value is low, there is quite a variability in ticket price, although it's capable of going quite high. Ticket prices may drop a little before then climbing upwards as the number of resorts per capita increases. Ticket price could climb with the number of resorts serving a population because it indicates a popular area for skiing with plenty of demand. The lower ticket price when fewer resorts serve a population may similarly be because it's a less popular state for skiing. The high price for some resorts when resorts are rare (relative to the population size) may indicate areas where a small number of resorts can benefit from a monopoly effect. It's not a clear.



Within the training and test set data produced a zero, because it tells us how much of the variance. Here we see that our simple linear regression model explains over 80% of the variance on the train set and over 70% on the test set. Using this model, then, on average you'd expect to estimate a ticket price within $9 or so of the real price. This is much, much better than the $19 from just guessing using the average. These results suggest that vertical drop is your biggest positive feature. This makes intuitive sense and is consistent with what you saw during the EDA work. Also, you see the area covered by snow making equipment is a strong positive as well. People like guaranteed skiing! The skiable terrain area is negatively associated with ticket price! There could be all manner of reasons for this. It could be an effect whereby larger resorts can host more visitors at any one time and so can charge less per ticket. As has been mentioned previously, the data are missing information about visitor numbers. Bear in mind, the coefficient for skiable terrain is negative *for this model*. For example, if you kept the total number of chairs and fast Quads constant, but increased the skiable terrain extent, you might imagine the resort is worse off because the chairlift capacity is stretched thinner.

Big Mountain Resort modelled price is $108.30, actual price is $81.00. Even with the expected mean absolute error of $10.24, this suggests there is room for an increase. This result should be looked at optimistically and doubtfully! The validity of our model lies in the assumption that other resorts accurately set their prices according to what the market (the ticket-buying public) supports. The fact that our resort seems to be charging that much less that what's predicted suggests our resort might be undercharging. It's reasonable to expect that some resorts will be "overpriced" and some "underpriced."