**Problem statement:**

Overall, the purpose of this project is to assess if there are opportunities to increase Big Mountain Ski Resort’s revenue by at least $1.6M this season through better ticket pricing and reduction of operational costs.

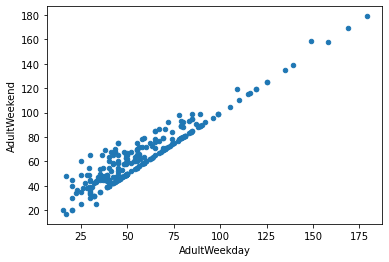
Currently Big Mountain uses average ticket price in its market segment to derive ticket pricing. Big Mountain suspects it may not be maximizing its returns, relative to its position in the market and wants to assess if it could increase its price based on facilities offered relative to those offered by other resorts.

In this project, we used data science principles to assess which facilities offered by ski resorts matter most to visitors and can impact ticket prices. We built a predictive model to provide guidance for Big Mountain's pricing and future facility investment plans which is shared in this report. This report will walk you through each of the steps taken to arrive at the final recommendation

**Step 1: Data Wrangling**

To build a model, we used a dataset provided by the operations team; this dataset included data for 330 resorts in Big Mountain’s market segment as well as a number of facilities provided by each resort e.g., number of ‘chair lifts’, ‘skiable terrain’ etc. Additionally, the state in which the resort functioned, names of resorts and ticket prices for weekday and weekend were also provided. We started by reviewing this dataset for completeness and accuracy, i.e., reviewing if any fields were missing data, if there were any outliers, or if additional data was needed.

Based on the review we dropped some data due to missing values, fixed values where possible, and summarized data by state. We also pulled some external data on states such as population and size of state so we could assess the impact of state related features on ticket price and joined this data with our internal ski resort data.



**Image 1.1: Representing correlation between Weekend and Weekday Ticket Prices**

In our dataset, we had two ticket prices – weekday and weekend prices. Since this is an outcome variable and of interest to us, we spent reviewing both fields of data in detail. We observed that weekend and weekday prices were very close (see image 1.1). Weekend prices could be higher but mostly when weekday ticket prices were less than USD 100. Lastly, weekday and weekend prices are equal for all resorts in Montana where Big Mountain is at. For these reasons we decided to stick with weekend prices and dropped the column with weekday prices.

At the end of data wrangling exercise, we were left with data for 227 resorts. This gave us a good foundation to start analyzing our data.

**Step 2. Exploratory Data Analysis (EDA)**

Exploratory Data Analysis involves reviewing the data for any initial patterns. In this stage of the process, we started by assessing if state characteristics i.e., number of resorts per state, density of people etc., had an impact on ticket prices. We did not see any trend of state on ticket prices and decided to treat all states equally in our model.

We merged the data we had on states with our ski resort data to see how all fields of data correlate with one another and especially with our outcome variable - ‘Ticket Price’. We saw some variables as being strongly correlated with ticket price and this gave us some good early insight and confidence to build the first version of our model.

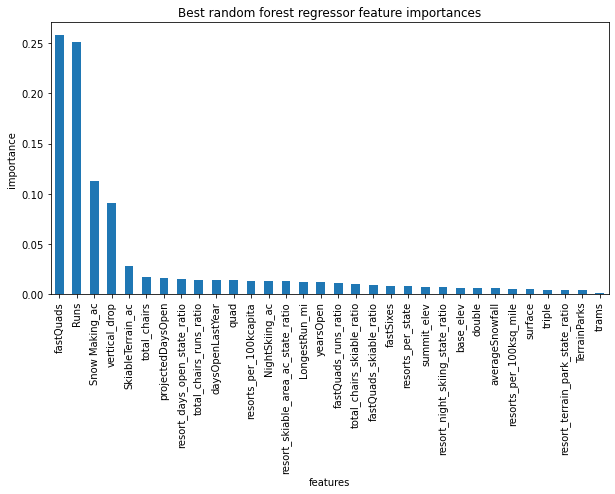
**Step 3. Model Preprocessing**

From data that we derived from the data wrangling, we created a subset data which excluded Big Mountain Resorts data. We divided the subset data into two – train data which consisted 70% of the dataset and test data which consisted of 30% of the dataset. We also used cross validation technique – this technique allowed us to split our train dataset into smaller datasets which further ensures a robust model by checking for overfitting of the model.

To test how well our model was doing in predicting ticket prices, we chose mean absolute error as our assessment metric. Once the above parameters were identified, we worked on testing accuracy of Big Mountain’s current strategy i.e. using average ticket price and saw mean absolute error was 19.13 which means if we use the current strategy, we will be off by USD 19 on ticket price.

We then created two separate models

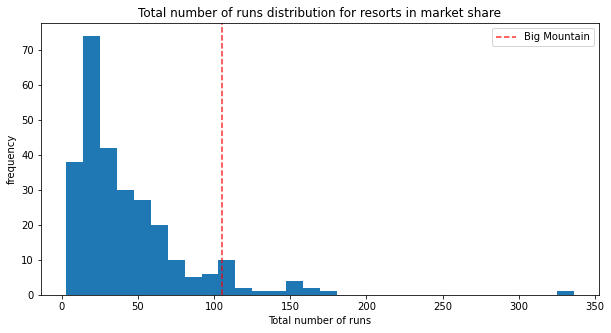
1. **Linear Regression** which explained about 70% of the variance in ticket price and mean absolute error of about 11.7 i.e. the new model was off by USD 11/- compared to USD 19/- when using average - so definitely a better approach. We learned that the following facilities - ‘vertical drop’, ‘snow making ac’, ‘total chairs’ and ‘fastQuads’ have the most impact on ticket price.
2. **Random Forrest Regression** with cross validation also identified facilities that had the most impact on ticket price and the top four facilities were common with the linear regression model – ‘fastQuads’, ‘Runs’, ‘snow making ac’ and ‘vertical drop’ (see image 1.2)



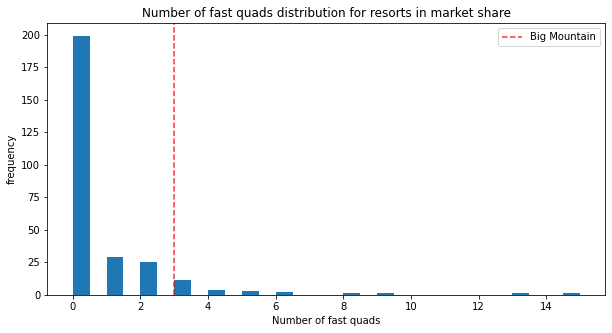
The random forest model had a lower cross-validation mean absolute error by almost $1(9.5 vs. 11.7). It also exhibited less variability. For these reasons, we chose to keep the random forest regression model as our final model.

Additionally, we did a data quantity assessment to see if we had sufficient data for our analysis. This assessment revealed that a sample size of 40 - 50 is good enough and that we have more than enough data. Additional data collection was not needed for our model.

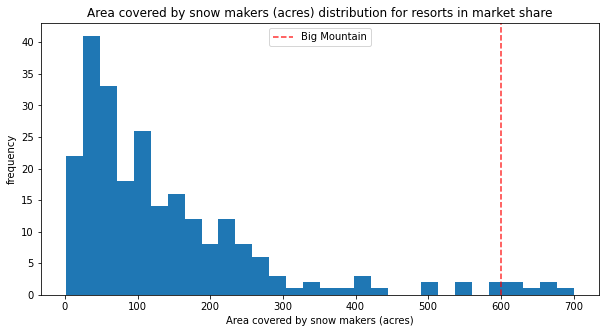
Based on our model, clearly there are facilities that impact ticket prices positively. Close assessment shows Big Mountain is on the higher end of most of these facilities offered (see images 1.3 – 1.6) validating a higher ticket price.



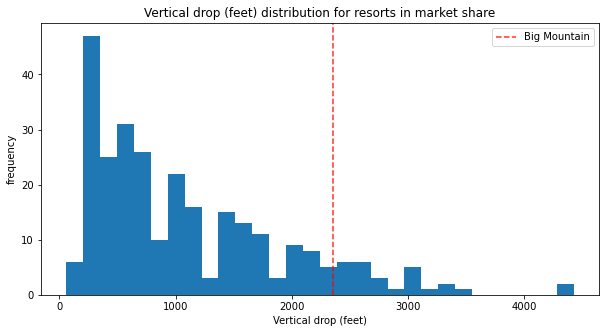
**Image 1.4**



**Image 1.3**



**Image 1.5**



**Image 1.6**

**Pricing recommendation based on model:**

Based on the model built and excluding any mean error, the model recommends that Big Mountain should charge USD 85/- (current price = USD 82/-) which would lead to an increased revenue of USD 7M. This increased revenue would help cover the increased operational costs of USD 1.5M incurred from adding an additional chair lift and provide USD 5.5M additional revenue.

**Winning model and scenario planning:**

Using our model, we also assessed various scenarios for what other steps could Big Mountain take to increase revenue and decrease operational costs. These scenarios and outcomes are reviewed below:

1. **Scenario 1:** Permanently closing up to 10 of the least used runs without impacting any other resort statistics.  
   **Outcome:** Model shows that closing one ‘Run’ does not impact revenue but could reduce operational cost. Also, decreasing two or three ‘Runs’ decreases revenue but removing three runs has the same impact as removing five runs on revenue and ticket price.
2. **Scenario 2:** Increase the ‘vertical drop’ by adding a run to a point 150 feet lower down but requiring the installation of an additional chair lift to bring skiers back up, without additional snow making coverage  
   **Outcome**: This scenario could validate ticket price increase of USD 2/- along with an increased revenue of USD 3.5M
3. **Scenario 3:** Same as number 2, but adding 2 acres of snow making cover
4. **Scenario 4:** Increase the longest run by 0.2 mile to boast 3.5 miles length, requiring an additional snow making coverage of 4 acres  
   **Outcomes for scenarios 3 & 4:** Both don’t make any other impact on ticket price or revenue

**Conclusion**

Our model shows that facilities offered by ski resorts impact ticket pricing. It also recommends, based on the facilities offered, Big Mountain could charge up to USD 85/- per ticket. However, we recommend taking these inputs with caution due to the following reasons:

* The model assumes that all other resorts are largely setting prices based on how much people value certain facilities
* We could be missing key facilities that customers value more.
* It is likely that some resorts in our dataset are over or under priced using a strategy similar to Big Mountain (i.e., average ticket price in market segment)
* Apart from the operational cost of adding a new chair lift, this model does not consider other operational costs that could be driving ticket prices
* We also have limited information on visitors. Note: we used state population data to derive some of this information, but seasonal visitor data was not available
* Data on why visitors prefer one resort over another and how much more are they willing to pay for certain facilities is not available – having this information could further increase confidence in making ticket prices higher without losing foot traffic
* Lastly, our model only takes into account weekend ticket prices due to limited data availability on weekday ticket prices.

**Future scope of work**

We need to review the accuracy of the model over time with new and additional data. Also, in the future, Big Mountain could consider two of the scenarios reviewed to further increase revenue and decrease operational costs i.e., closing one less used ‘Run’ and adding a new ‘Run’ to increase the vertical drop by 150 feet and adding a chair lift. As mentioned earlier, doing so could further validate a ticket price increase of USD 2/- leading to an increased USD 3.5M revenue for Big Mountain.