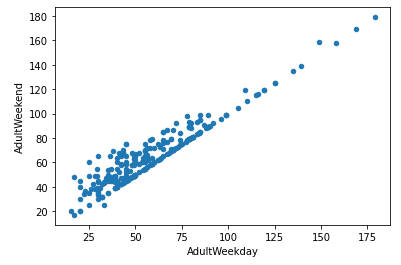
**GUIDED CAPSTONE PROJECT REPORT**

**Introduction**

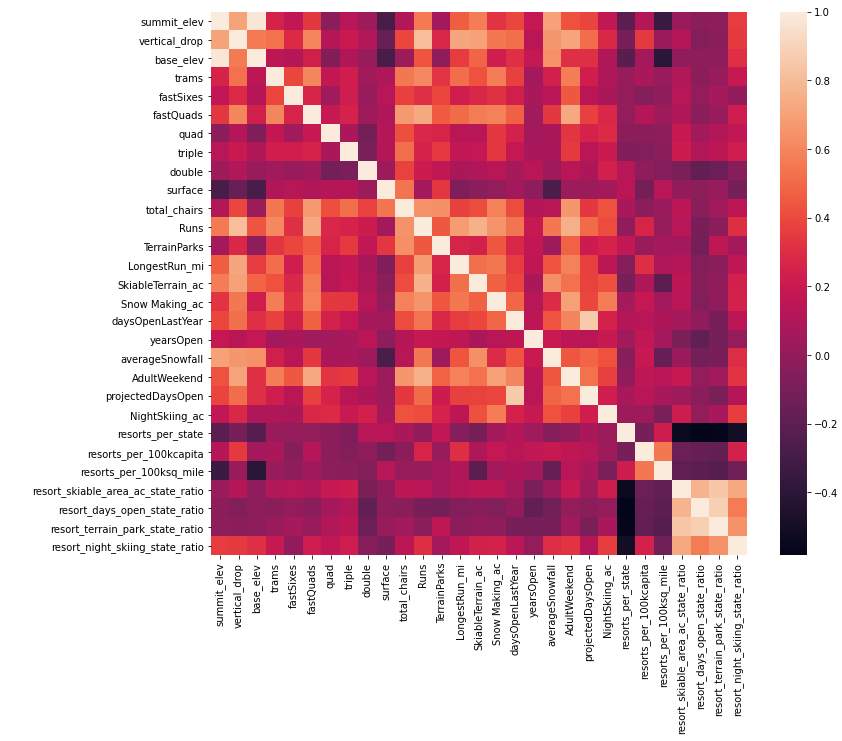
Big Mountain Resort, a ski resort in Montana, has access to 105 trails. This allows them to accommodate skier and riders of all levels and abilities. Big Mountain Resort has recently installed an additional chair lift, costing $1.5M, to help increase distribution of visitors across the mountain. Management is concerned that they aren’t capitalizing on their facilities as much as they could, due to their current pricing methodology. The pricing strategy has been to charge a premium above the average price of resorts in its market segment.

**Data Wrangling**

The first process was to review the data at a high-level by utilizing .head() and .info() functions. In that manner we could see if there were missing values, how the information is presented, the size of the data, etc. We started with 330 rows and 27 columns and ended with 277 rows and 25 columns. We eliminated those that had missing values. We dropped the fastEight column in its entirety. Half the values were missing and all, but the ‘Others’ had the value zero. We also dropped 'AdultWeekday' prices because they appeared to resemble the same pricing as 'AdultWeekend'. When running the analysis, there weren’t any major differences between the two. The reason why 'AdultWeekday' was chosen to be dropped is because it had more empty values compared to 'AdultWeekend'. Below you can see there is a clear line where weekend and weekday prices are equal.

Also 'SkiableTerrain\_ac' was updated to 1819 from 26819 because it was skewing the data and it was confirmed that 1819 was the correct number instead of 26819.

**Exploratory Data Analysis**

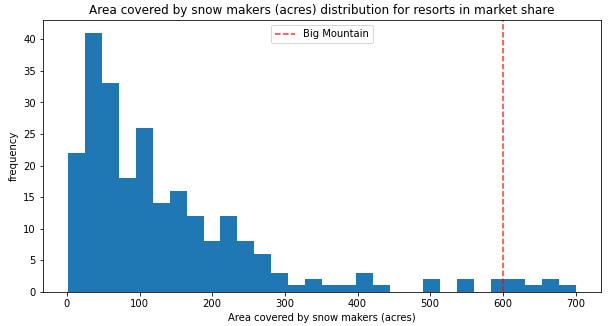
The categorical features within the data were related to resorts and their location (state/region). The remaining fields were numerical, in the form of integers and floats. The numerical data was primarily due to geographical information and information pertaining to the parks (snowfall, days open, runs, chairs, ticket pricing, etc.). There was a strong positive correlation with vertical\_drop and fastQuads seemed to be very useful. Runs and total\_chairs appeared to be quite similar and also appeared to be useful. In this process, the data was scaled and then the Principal Cumulative Analysis was calculated. The PCA was used to reduce the number of features under consideration by correcting for existing features with high correlation. By scaling the data, we were able to create a heat map, which correlated pricing with the following attributes: vertical drop, fastQuads, runs, and snow making area covered.

**Preprocessing and Training**

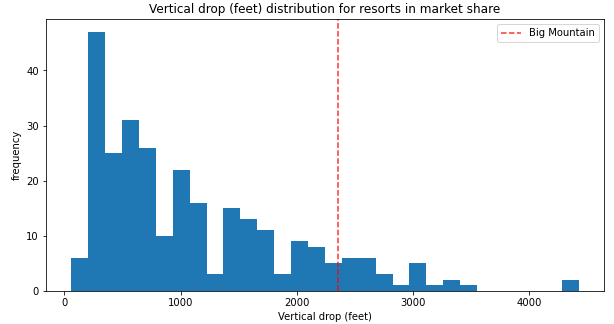
In summary, we worked with and tested a few models to gain confidence in the data set and models. The models were stripped of Big Mountain Resort. We started with a machine learning model using the mean value as the predictor, then assessed if this approach was useful and determined it was not. We wanted to make sure we didn't overfit the models and did so by partionining the data (which included a 70/30 split). With our models we were able to identify four categories that had a correlation with pricing. The top two were vertical drop and area covered by snow making equipment. Vertical drop was the biggest positive feature. Area covered by snow making equipment was a strong positive as well.

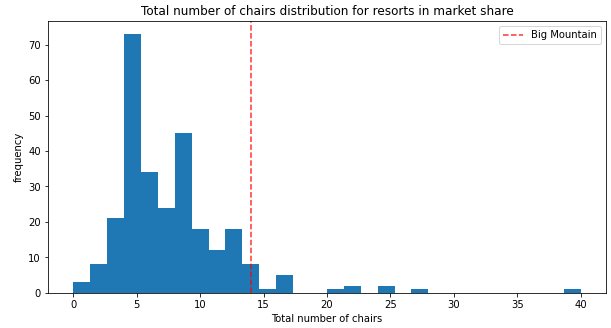
**Modeling**

Current pricing for Big Mountain Resort is $81.00. Big Mountain has a few advantageous characteristics that other resorts don't have that could set them apart. For example, Big Mountain offers some of the best area covered by snow making (Exhibit A), competitive verticals drops (Exhibit B), total chairs (Exhibit C), fast quads (Exhibit D), runs (Exhibit E), run length (Exhibit F) and some of the largest skiable terrain (Exhibit G). That being said, Big Mountain could leverage their pricing strategy based on those advantages and comparability to resorts of similar caliber. After analysis and modeling, the Big Mountain Resort modelled price is $95.87. Therefore, there is room for ticket price increase in comparison to today's pricing, while improving profitability (cutting costs by eliminating some runs without having a negative impact on revenue).

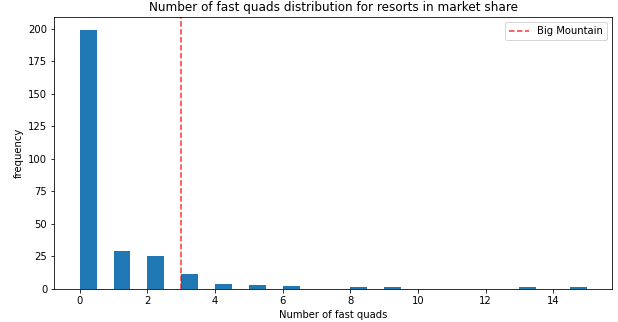
**Exhibit A**

**Exhibit B**

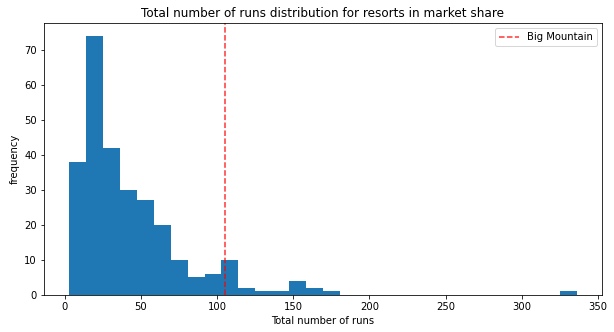
****

**Exhibit C** ****

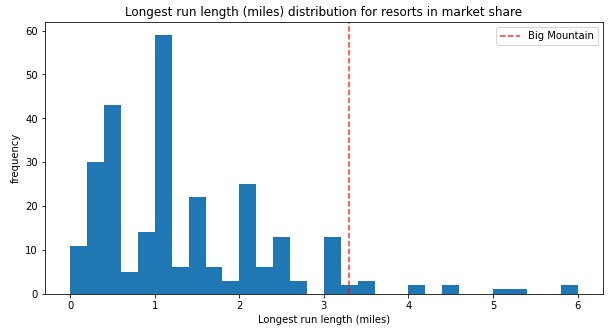
**Exhibit D**



**Exhibit E**



**Exhibit F**



**Exhibit G**

