**1.HYPOTHESIS:**

How to increase revenue for the resort by cutting the unwanted cost without undermining the ticket price and utilising the facilities in a proper way? Methods to reduce the operation cost of additional chair lifts should be identified to decrease the cost from $1,540,000 for the next season.

**2.ABSTRACT:**

Big Mountain Resort, is a ski resort located in Montana. They offer spectacular views of Glacier National Park and Flathead National Forest, with access to 105 trails. Every year about 350,000 people ski or snowboard at Big Mountain. This mountain can accommodate skiers and riders of all levels and abilities.They recently installed an additional chair lift to help increase the distribution of visitors across the mountain. This additional chair increases their operating costs by $1,540,000 this season. The business wants some guidance on how to select a better value for their ticket price. They are also considering a number of changes that will either cut costs without undermining the ticket price or will support an even higher ticket price.

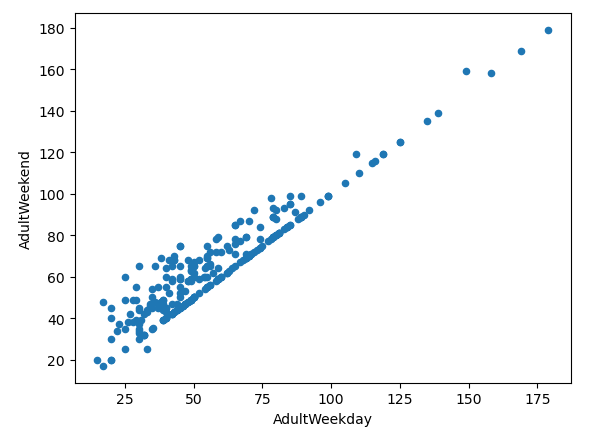
**3.DESIGN & IMPLEMENTATION:**

We used the Ski Resort data for this project. The major four steps of the project are data wrangling, exploratory data analysis, preprocessing & training, and finally modelling the data. In the data wrangling step, we cleaned and transformed the data into analysis ready format. In the second step feature engineered by merging state summary features into the ski resort data and visualized high dimensional data using PCA. In the preprocessing & training step, trained the data with machine learning models, linear regression and random forest and cross validated. Finally, used our model to gain insight into what Big Mountain's ideal ticket price could be, and how that might change under various scenarios.

**3.1.Data Wrangling:**

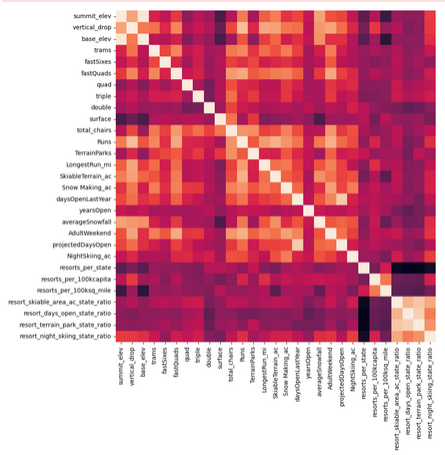
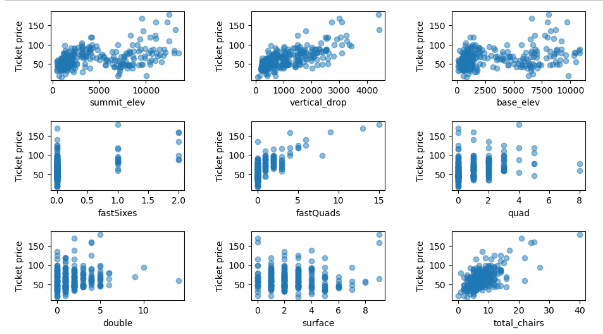
Initially there was 330 rows and 27 columns and after cleaning, transforming the final data frame with 277 rows and 25 columns. Dropped 14% of rows that have no price and 2 columns ('fastEight','AdultWeekday') as weekend prices have the least missing values of the two, so drop the weekday prices and then keep just the rows that have weekend prices. The data got ready for analysis, cleaned, transformed and filtered to the required format. Also found that the distribution for weekday and weekend prices in Montana seemed equal. Big Mountain resort is in Montana, so the relationship between these quantities in this state are particularly relevant.

Figure below shows the relationship between weekday and weekend ticket prices.



**3.2.Exploratory Data Analysis:**

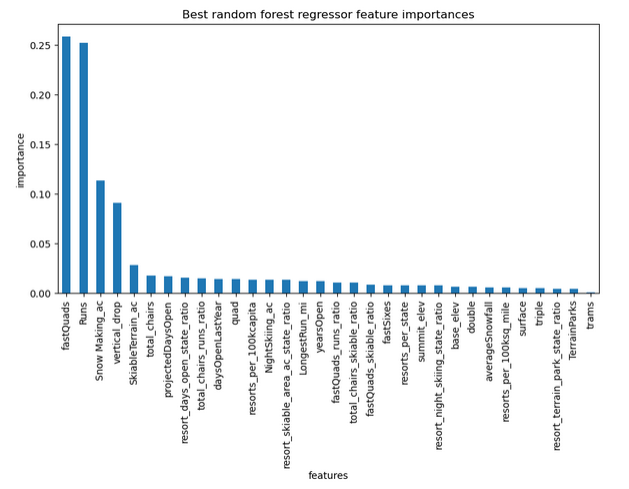
Used principle components analysis to explore how the original features contribute to the derived features. This done by scaling the data, fit the PCA transformation to learn the transformation from the data, applying the transformation to the data to create the derived features. Also produced feature engineering heatmap and scatterplots of numeric features against ticket price.



The figures shows the scatterplots of numeric features against ticket price and Feature correlation heatmap.

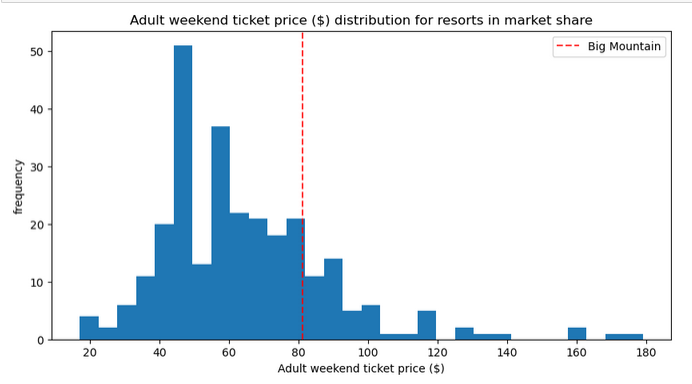
**3.3.Preprocessing & Training:**

In this step preprocessed the data and trained using machine learning models, like linear regression model and random forest model. The random forest regression model was chosen because the random forest model has a lower cross-validation mean absolute error by almost $1. It also exhibits less variability. Verifying performance on the test set produces performance consistent with the cross-validation results. Figure below illustrate best random forest regressor feature importances.



**3.4.Modeling:**

Modelled the features and data we got after training using four different scenarios. In scenario1, the model says closing one run makes no difference. Closing 2 and 3 successively reduces support for ticket price and so revenue. If Big Mountain closes down 3 runs, it seems they may as well close down 4 or 5 as there's no further loss in ticket price. Increasing the closures down to 6 or more leads to a large drop. In scenario 2 , Big Mountain is adding a run, increasing the vertical drop by 150 feet, and installing an additional chair lift, this increases support for ticket price by $8.61and over the season, this could be expected to amount to $15065471. By scenario 3, increases support for ticket price by $9.90 and over the season, this could be expected to amount to $17322717. According to scenario 4, it calls for increasing the longest run by .2 miles and guaranteeing its snow coverage by adding 4 acres of snow making capability but found no difference whatsoever. Although the longest run feature was used in the linear model, the random forest model only has longest run way down in the feature importance list. Big Mountain Resort modelled price is $95.87, actual price is $81.00. Even with the expected mean absolute error of $10.39, this suggests there is room for an increase. Figure below shows where Big Mountain sits overall amongst all resorts for price and for just other resorts in Montana.



**CONCLUSION & FUTURE WORK:**

We analysed that Big Mountain Resort modelled price is $95.87, actual price is $81.00. Even with the expected mean absolute error of $10.39, this suggests there is room for an increase.We took our model for ski resort ticket price and leveraged it to gain some insights into what price Big Mountain's facilities might actually support as well as explored the sensitivity of changes to various resort parameters. So this model can be used by analysts to explore the changes by varying the parameters and trying out various scenarios for better results.