Guided Capstone Project Report

Phuc Nguyen

December 2022

1 Problem Statement

How can Big Mountain Resort improve its pricing strategy so that its revenue is increased by at least $ 1,540,000 within one season, by taking into account the importance of each facility or geographical features such as the vertical drop of the resort ?

2 Data Wrangling

During the data wrangling step, we decided that the target feature should be the weekend price because the data for weekend prices is more complete than data for weekday prices. Also, we dropped a few rows and columns, as follows: rows with missing data for the weekend price, one row with wrong number of years the resort is open, one column where half the values are missing and the remaining values are basically zero, and the weekday price column.

3 Exploratory Data Analysis

In this step, we did a PCA analysis to investigate whether there is a relationship between the state and the ticket price, and did not observe any such relationship. Therefore, we decided to treat all states equally. We also studied the correlation between features, and saw that the weekend price is highly correlated with features such as fastQuads, Runs, snowMaking_ac, and resort_night_skiing_state_ratio (this latter is a new feature engineered by us).

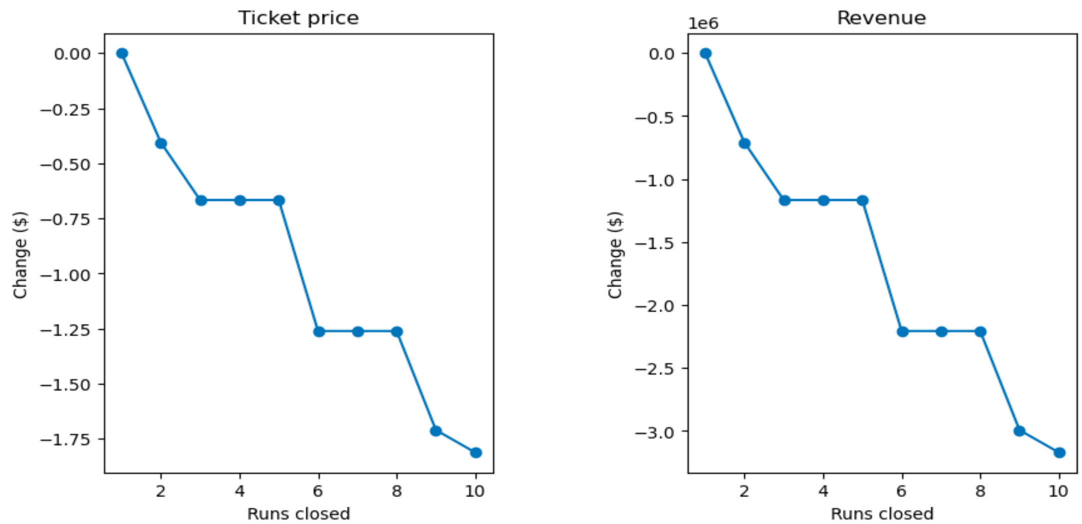
4 Preprocessing and training

We tried three models: a baseline model, a linear regression model, and a random forest model. To handle missing values, we imputed them by using both the mean and the median of the data. We scaled the data using the StandardScaler, so that the scaled data have zero mean and unit variance. To measure the performance of the models, we used 3 metrics: the coefficient of determinant (R squared), the mean absolute error and the mean squared error. We further refined our performance assessment by using cross-validation as well as hyperparameter search using GridSearchCV (to optimize the model with respect to the number of best features in the case of linear regression, or the number of trees to use in the case of random forest).

5 Modelling

The winning model is the random forest one. According to this model, Big Mountain should price its weekend ticket at $95.87 with a MAE of $10.39. This represent a sharp increase from the current price of $81.00. If the model is correct, the increase in revenue will be more than enough to offset the operating cost of installing the recent chair lift.

Among the four scenarios under consideration, I would recommend further consideration of the first two of the four modelled scenarios. The last two scenarios do not seem to make much of a difference. If the business leadership is very interested in bringing down the ticket price as much as possible, then I would recommend Scenario 1 (i.e. closing up to 10 of the least used runs). In the figure below, we plot the decrease in ticket price (as well as in revenue) as a function of the number of runs closed, according to the model.



On the other hand, if the business leadership is willing to invest money into more equipment and installations, then I would recommend scenario 2.

6 Conclusion and future scope of work

In conclusion, we recommend that Big Mountain increases its weekend adult ticket price to \$95.87.
However, one deficiency of our model is that it assumes that other resorts correctly set their prices according to the market. But, it could be that other resorts are also clueless about how to set prices, just like our Big Mountain resort. Another limitation is that we only know the operating cost of the newly-installed lift at Big Mountain, and it would be great to know the operating costs of all the other lifts at Big Mountain as well as at the other resorts.
Thus, in the future, it could be worthwhile to study further whether the free-market assumption is accurate, and to obtain additional data for the operating costs.