**Guided Capstone Project Report**

**1. Problem setup**

Provide some guidance to the management of Big Mountain ski resort in Montana on how to adjust their pricing and investment into existing and new facilities to increase revenue. The steps are as follows.

* Build a model that predicts daily ticket prices for US ski resorts.
* Identify the most important features affecting the price.
* Propose a balanced price based on model prediction for existing facilities.
* Evaluate four development scenarios with different upgrades.
* Determine which important data are missing and which are worth collecting.

**2. Key factors determining ticket price and its proposed adjustment**

After testing multiple linear regression and random forest regression methods for data from 330 US ski resorts, random forest regression was chosen as a more consistent predictor of daily ticket prices. The quality of its predictions for test sets during cross-validations did not reduce much compared to the training set, in contrast to linear regression. The random forest regression model allowed us to identify eight main features that affect the ticket price (Figure 1).

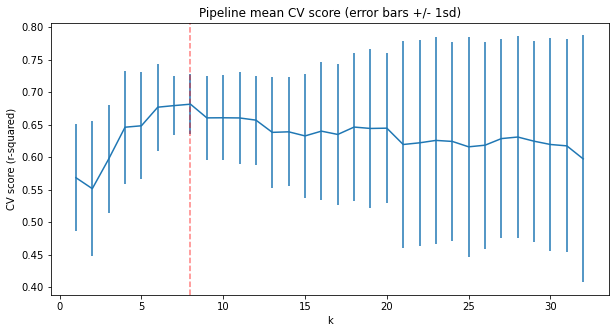


Figure 1. Cross-validation score for a varying number of core features included in a random forest regression model for ticket price. Inclusion of eight features (indicated by a red line) yields the highest score with the lowest standard deviation.

Table 1 shows the ranking of Big Mountain Resort among the US ski resorts for each of the eight most influential features. For instance, only 2.2% of resorts have larger skiable terrain area. All features have a percentile higher than 80.4% for the price, indicating that our tickets may be **underpriced**. The market-based **predicted** daily ticket price for our resort is $95.87 with a confidence interval of $[85.48, 106.26].

**Table 1.** Ranking of Big Mountain Resort among 276 US ski resorts.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Features** | Price, $ | Vertical Drop, ft | Snow Making Area, acres | Total Chairs | Fast Quads | Runs | Longest Run, mi | Trams | Skiable Area, acres |
| **Current Value** | 81 | 2353 | 600 | 14 | 3 | 105 | 3.3 | 0 | 3000 |
| **Percentile** | 80.4 | 89.9 | 84.4 | 95.3 | 95.3 | 92.4 | 94.2 | 92.0 | 97.8 |

The ranking of the feature importance for the predicted price is as follows: the number of fast quads, runs, snow making area, vertical drop, ski area, total chairs, length of the longest run, trams.

**3. Evaluation of business scenarios**

Four development scenarios were proposed by the management with different upgrades, characterized by the parameters shown in Table 2. Scenario 1 proposes to reduce operating costs by closing up to 10 least used runs, which is highly likely to result in partial loss of profit. **Scenario 2** of increasing the vertical drop by adding one run, served by an already added chair lift, is the most promising. Additional seasonal profit is forecast at a little under $2M. Scenario 3 differs from Scenario 2 only in the additional operating costs of 2 acres of artificial snow. Scenario 4 does not offer a sufficient increase in the length of the longest run and area with guaranteed snow, so the balanced price is predicted to remain the same, only to add the cost of making 4 acres of artificial snow. Ranking of the scenarios from best to worst is 2, 3, 4, and 1.

**Table 2.** Business development scenarios proposed by the management of Big Mountain Resort. The cells display the proposed increments in the resort’s feature values, together with forecasts of changes in the justified price, revenue, and seasonal profit.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Features** | Vertical drop (ft) | Snow making (acres) | Total chairs | Runs | Longest run (mi) | Price ($) | Revenue ($) | **Profit** ($) |
| **Scenario 1** | 0\* | 0\* | 0\* | [-10, -1] | 0\* | [-1.81, 0] | [-1.8M, 0] | [-10, 0] × RC\*\* |
| **2** | 150 | 0\* | 1 | 1 | 0\* | 1.99 | 3474638 | **1934638** |
| **3** | 150 | 2 | 1 | 1 | 0\* | 1.99 | 3474638 | 1934638 - 2×SMOC\*\*\* |
| **4** | 0 | 4 | 0 | 0 | 0.2 | 0 | 0 | -4×SMOC |

**\***Note that many features are interrelated, so changing some of them as a result of development could potentially change other features. For instance, the increase in the length of the longest run in Scenario 4 may be due to the addition of a skiable area rather than merging with an existing area. This will require a recalculation of price, revenue, and profit.

\*\*RC – unknown Run preparation Cost; this is a fixed single-time cost and is thus expected to have much less impact than variable costs.

\*\*\*SMOC – unknown Snow Making Operating Cost per acre per season.

**4. Future improvements of price model**

To improve the accuracy of price predictions, we must first collect missing price data for some resorts, build a complementary model for daily ticket prices on weekdays, and collect customer count data for all resorts over the past season. In addition, to accurately estimate operating costs, collect resource usage data for all types of equipment and resource cost (electricity, gas, water, and labor) for each resort location. All of the above will help determine which resorts are underpriced, justified, or overpriced. It is important to add a risk analysis based on the standard deviation of the predictions. Scenarios involving features not yet tested should be evaluated.

**5. Recommendations**

Proceed with Scenario 2, predicted to be the most profitable. Start collecting visitor and cost data.