Guided Capstone Project Report

Fei Wang
07-21-2023

Problem Statement

Big Mountain Resort is looking for strategies to optimize revenue for the next ski season by increasing its lift ticket price and cutting operation costs. Leveraging data, the management team wants to find out if their current ticket price undervalues the resort's facilities and to pinpoint potential areas for reducing operational costs.

Data Wrangling

- Count missing values to identify potential data issues.
- Perform feature analysis to check outliers, to drop useless features and to decide how to use categorical features.
- Select *AdultWeekend* price as the modeling target.

Exploratory Data Analysis (EDA)

- State wise, Montana, home to Big Mountain Resort, should draw in ski enthusiasts with its expansive total skiable area and less crowded population.
- From PCA analysis on state-wise statistics, the current average ticket price (indicated by color) in Montana was lower than its neighboring states in Figure 1, indicating potential undercharging.
- The target ticket price was found to be linearly correlated with numerical features, such as *vertical_drop*, *Snow_Making_ac*, *fastQuads*, *Runs* and *total_chairs*.

Model Preprocessing with Feature Engineering

- Missing values were imputed for model building
- More business relevant state-resort-competition features (resort / state total), ratios of existing features were created

Model Building and Evaluation

- A linear regression model and a random forest model were developed through pipeline and refined by hyperparameter tuning, respectively.
- Both models were evaluated by metrics including R^2, mean absolute error and mean square error and through cross-validation on train split and directly on test split.

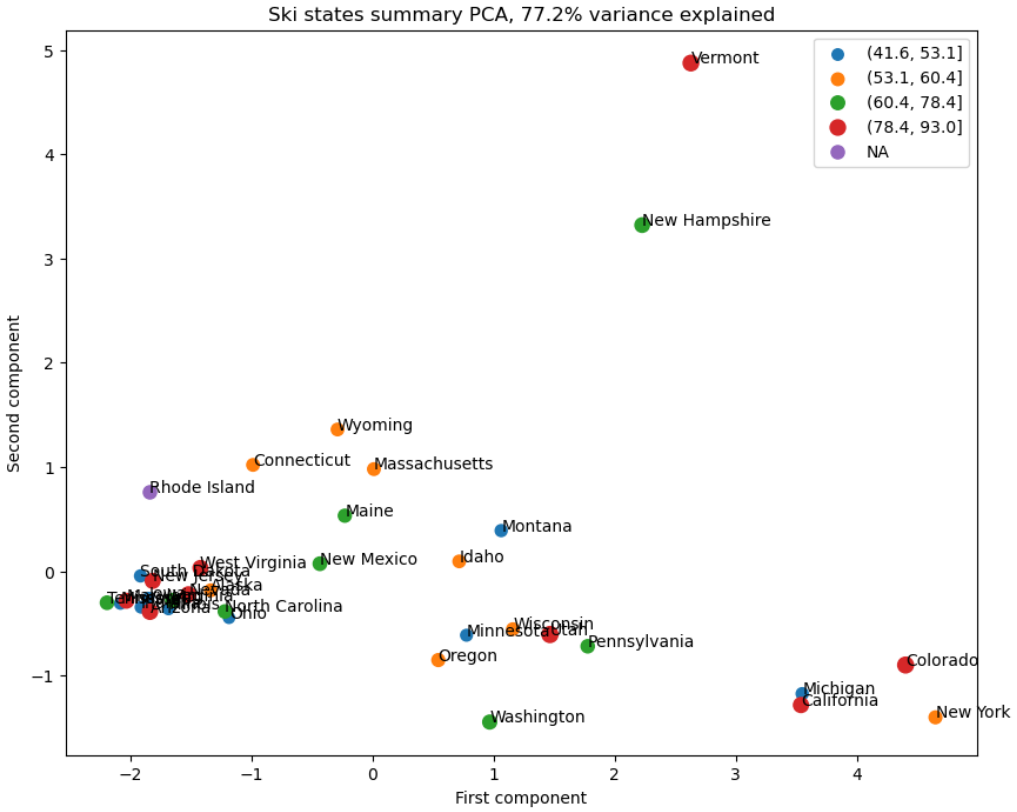

Figure 1. Ski States Summary PCA Analysis

Model Selection

The refined random forest model was selected as the best model due to its lower errors, less oscillations in errors and more consistent performance on train split and test split.

Inferential Modeling

The best model was used to predict the ticket price change in the following scenarios:

- S1: shutting down up to 10 least used runs (lacking operational costs)
- S2: adding 1 run and 1 lift raising vertical drop by 150 feet (recommended by model)
- S3: same as S2 but adding 2 acres of snow making (not recommended by model)

- S4: increasing the longest run by 0.2 mile with adding 4 acres of snow making (not recommended by model)

Pricing Recommendation

- The current ticket price, $81, of Big Mountain Resort is lower than the predicted ticket price, $95.87, from the model.
- Increasing the ticket price to $90 should be reasonable from the model, which will bring an increase in revenue by $15,750,000 considering 350,000 visitors and each buying 5 tickets for the next ski season.

Summary

The model can help solve a variety of business problems, as illustrated in the inferential modeling section. However, for the model predictions to be reliable, it is crucial that the underlying assumptions are valid, which is the free-market assumption, and the ticket price is not correlated with number of visitors.

Future Work

- Include number of visitors per season or number of tickets sold per season to improve the model accuracy on ticket price and to aid revenue estimations.
- Gather operation costs to evaluate cost-reduction strategies. For example, if the operation cost of each run per season is available, then a plot like Figure 2 can be generated to determine the number of runs to shut down to optimize the revenue.
- Clustering all ski resorts to compare the price of Big Mountain Resort and others in the same cluster. Clustering can potentially help with feature engineering to further improve the model accuracy.
- Deploy the model for convenient internal use by business analysts and stakeholders.

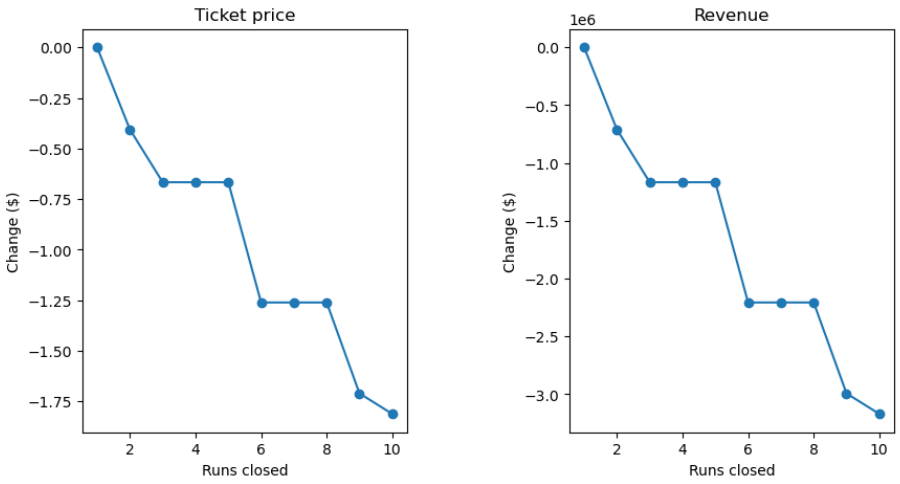

Figure 2. Model Predicted Ticket Price and Revenue upon Number of Runs Closed